

# SCIENTIFIC REPORTS



OPEN

## Identification and comparative analysis of the epidermal differentiation complex in snakes

Karin Brigit Holthaus<sup>1,2</sup>, Veronika Mlitz<sup>1</sup>, Bettina Strasser<sup>1</sup>, Erwin Tschachler<sup>1</sup>, Lorenzo Alibardi<sup>2</sup> & Leopold Eckhart<sup>1</sup>

Received: 14 November 2016

Accepted: 22 February 2017

Published: 27 March 2017

The epidermis of snakes efficiently protects against dehydration and mechanical stress. However, only few proteins of the epidermal barrier to the environment have so far been identified in snakes. Here, we determined the organization of the Epidermal Differentiation Complex (EDC), a cluster of genes encoding protein constituents of cornified epidermal structures, in snakes and compared it to the EDCs of other squamates and non-squamate reptiles. The EDC of snakes displays shared synteny with that of the green anole lizard, including the presence of a cluster of corneous beta-protein (CBP)/beta-keratin genes. We found that a unique CBP comprising 4 putative beta-sheets and multiple cysteine-rich EDC proteins are conserved in all snakes and other squamates investigated. Comparative genomics of squamates suggests that the evolution of snakes was associated with a gene duplication generating two isoforms of the S100 fused-type protein, scaffoldin, the origin of distinct snake-specific EDC genes, and the loss of other genes that were present in the EDC of the last common ancestor of snakes and lizards. Taken together, our results provide new insights into the evolution of the skin in squamates and a basis for the characterization of the molecular composition of the epidermis in snakes.

Snakes are reptiles that have lost their limbs during evolution and developed a unique predatory lifestyle that involves the ability to swallow prey of a diameter larger than that of their own body<sup>1–3</sup>. The skin of snakes, and more specifically the epidermis, consists of rigid scales and soft inter-scale regions, which together provide both mechanical resistance and flexibility<sup>4,5</sup>. Snakes belong to the squamate reptiles which are characterized by the regular shedding of the outer layers of the epidermis, also known as ecdysis<sup>6,7</sup>. In snakes, the superficial layers of the epidermis are detached as a single, coherent sheet whereas other squamates (lizards and geckos) shed multiple smaller flakes. While the dynamic regulation and composition of snake skin has been partially revealed over the past fifty years<sup>7–9</sup>, the recent availability of whole genome sequences of snakes and other reptiles allows, for the first time, to identify genes that encode epidermal proteins and, by comparative genomics, to establish a basis for building hypotheses on the molecular evolution of the epidermis in snakes<sup>10–12</sup>.

In all amniotes, keratinocytes proliferate in the basal layer and differentiate in the suprabasal layers of the epidermis<sup>13,14</sup>. Signaling between the epidermis and the underlying dermis controls the patterning of the epidermis and the formation of skin appendages<sup>15–18</sup>. While differentiation of keratinocytes in the mammalian epidermis involves continuous alterations of cell structures and movement of cells towards the body surface, keratinocyte differentiation in squamates results in the formation of distinct non-interconvertible layers that remain stable for several weeks before they are shed together<sup>8,9</sup>. In its final differentiation stage, the outer generation of the epidermis comprises a clear, lacunar, alpha, mesos, beta, and oberhautchen layer<sup>7–9</sup>. Prior to shedding, the outer generation of the epidermis protects the resting and newly forming inner generation of the epidermis. At the beginning of skin development, the cell layers of the embryo-specific periderm cover the epidermis<sup>19</sup>. The histological and ultrastructural features of squamate skin have been reported<sup>7–9,20</sup>. However, only few aspects of the molecular architecture of alpha and beta-layers of squamate epidermis have been determined so far<sup>21–24</sup>.

Corneous beta-proteins, traditionally called beta-keratins, have been identified as components of the epidermis in snakes, like in other reptiles<sup>22,25–27</sup>. Keratin intermediate filament proteins, previously referred to as alpha-keratins, are the main cytoskeletal proteins in the epidermis and skin appendages of vertebrates. A

<sup>1</sup>Research Division of Biology and Pathobiology of the Skin, Department of Dermatology, Medical University of Vienna, Vienna, Austria. <sup>2</sup>Dipartimento di Scienze Biologiche, Geologiche ed Ambientali (BiGeA), University of Bologna, Bologna, Italy. Correspondence and requests for materials should be addressed to L.E. (email: leopold.eckhart@meduniwien.ac.at)

cysteine-rich keratin component of reptilian claws has been lost due to gene inactivation during the evolution of snakes<sup>28</sup>. The presence of various ultrastructurally, but not biochemically, defined epidermal components such as fibers in the beta-layer<sup>29</sup> and different granules in the oberhautchen and the clear layer of reptiles<sup>21</sup> indicate that many structural proteins of snake epidermis remain to be identified.

The recent availability of genome and transcriptome sequences from multiple vertebrates has allowed the determination of genes implicated in epidermal structure and function. Based on the dermatologically relevant characterization of human epidermal barrier genes, we have screened non-mammalian tetrapods for homologs of a gene cluster known as the Epidermal Differentiation Complex (EDC)<sup>30–34</sup>. Genes encoding S100 fused-type proteins (SFTPs), which are homologous to a subgroup of mammalian EDC genes<sup>35</sup>, were found in amphibians<sup>34</sup>, and more complex gene clusters homologous to the mammalian EDC were identified in the chicken<sup>30,32</sup>, in the green anole lizard (*Anolis carolinensis*)<sup>30</sup> and in turtles<sup>33</sup>.

Here we extend the comparative analysis of the EDC in sauropsids and determine the gene complement of the EDC in snakes. We characterize the amino acid sequences of EDC-encoded proteins, suggest hypotheses about their contributions to the molecular architecture of the epidermis, and identify cases of gain and loss of specific EDC genes during the evolution of stem lepidosaurs and snakes.

## Results

**Identification of epidermal differentiation complex (EDC) genes in snake genomes.** The EDCs of the Burmese python (*Python bivittatus*) and the king cobra (*Ophiophagus hannah*) were defined as the genomic regions flanked by S100A genes, like in other amniotes<sup>30,33</sup>. The gene complement of the EDC of these snakes was identified by tBLASTn searches using EDC-encoded proteins of *A. carolinensis*, chicken and humans as queries and by *de novo* prediction of genes in an iterative process, as described previously<sup>33</sup>. The predicted amino acid sequences of snake EDC proteins were used as queries in tBLASTn searches in the published transcriptomes of snakes to test for the expression of the predicted genes.

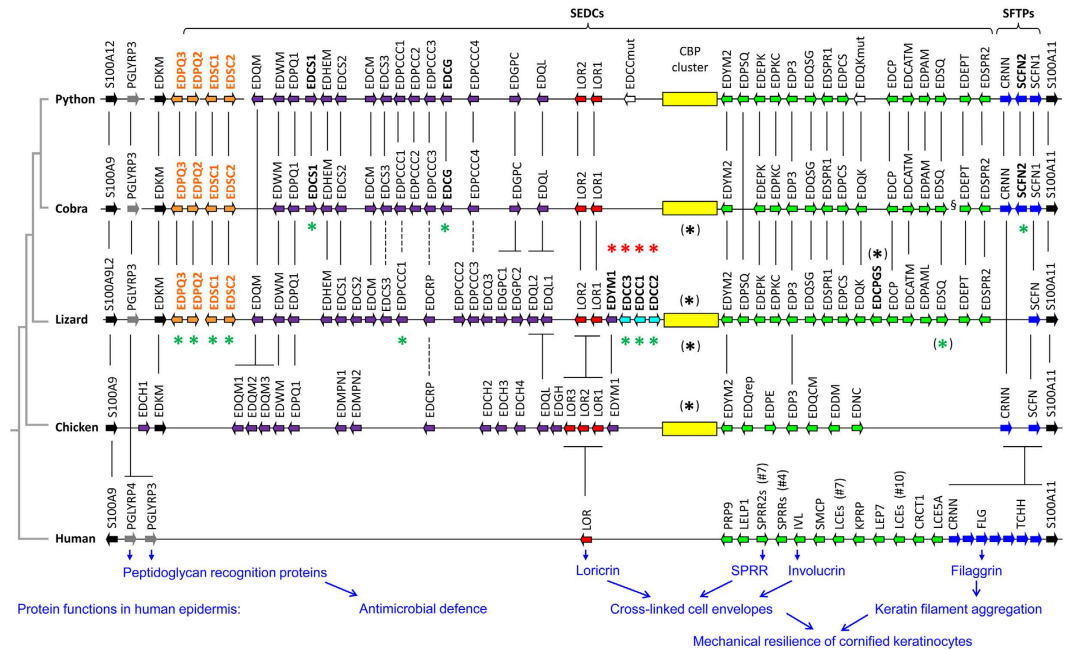
The nomenclature for EDC genes follows the preliminary system defined in previous studies<sup>30</sup>. In brief, gene names consist of the term Epidermal Differentiation (ED) followed by a term that describes the amino acid composition or the presence of particular amino acid sequence motifs in the encoded protein. For easier readability, only the abbreviations are used in the text whereas the full names of genes are summarized in Supplementary Table S1. Exceptions to this naming convention were made to indicate orthologs of human *loricrin* and *cornulin* and chicken *scaffoldin*.

**The EDC of snakes is largely syntenic with that of the green anole lizard.** The structure of the EDC is very similar in the Burmese python, the king cobra, and the green anole lizard (Fig. 1). Like in other amniotes<sup>30,33</sup>, the EDC of snakes is bordered by S100A genes and comprises a peptidoglycan recognition protein 3 (*PGLYRP3*) gene, simple (single coding exon) EDC (SEDC) genes, and SFTP genes (Suppl. Tables S2–6; Suppl. Figures S1 and S2). A gene homologous to *EDKM* of lizards, turtles and birds is located between the *PGLYRP3* and the SEDC genes (Fig. 1). In the draft genome of the king cobra, the genes *EDSQ* and *EDEPT*, the orthologs of which are neighbors in the python genome, were separated by a series of genes not related to the classical EDC genes and a sequence gap. This pattern indicates that a gene rearrangement event might have disrupted the canonical organization of the EDC in this species (Fig. 1, §).

To test whether the predicted EDC genes of snakes are expressed, tissue transcriptomes of snakes were screened by tBLASTn searches. Indeed, transcript reads corresponding to most EDC genes (Suppl. Tables S2 and S3), were detected in skin transcriptomes of the ball python (*Python regius*)<sup>36</sup> (Suppl. Fig. S3) and of the painted saw-scaled viper (*Echis coloratus*)<sup>37</sup>, whereas the transcriptomes of internal organs of snakes included no or only very small numbers of EDC gene transcripts (Suppl. Fig. S4), suggesting a skin-specific expression of most EDC genes.

**A unique corneous beta-protein (CBP) comprising 4 beta-sheets is conserved in squamates.** SEDC genes form a continuous cluster in snakes and include a sub-cluster of genes that encode corneous beta-proteins (CBPs), also known as beta-keratins<sup>14</sup>. These proteins are characterized by a conserved core domain that is predicted to form a beta-sheet<sup>38</sup>. The CBP cluster gene is located between the *loricrin* and *EDYM2* genes of snakes (Fig. 1, Fig. 2). It is syntenic with the CBP locus of the green anole lizard and with the main CBP loci of birds and turtles<sup>30,33</sup>. Within the CBP gene cluster, 35 and 36 CBP genes, here termed *Beta1* through *Beta36* in order of the arrangement of the genes, were identified in the python and cobra, respectively, which is comparable to the 40 CBP genes present in the green anole lizard<sup>26</sup>, but lower than the 71 CBP genes reported for the Japanese gecko<sup>39</sup>.

Both the python and cobra have an ortholog of the lizard gene *Beta1*, previously termed *Li-Ac40*<sup>26</sup> (Fig. 2). This gene encodes a protein that contains 4 CBP core sequence motifs and therefore is predicted to form 4 beta-sheets (Fig. 3; Suppl. Fig. S5) whereas all other CBPs identified so far comprise only a single beta-sheet domain. Remarkably, classical CBPs undergo dimerization via face-to-face interactions between their beta-sheets and subsequently they assemble into a beta-fibril (beta-filament) in which, according to the classical model<sup>38,40,41</sup>, 4 dimers form one turn of the helical structure. We put forward a hypothetical model in which 2 *Beta1* proteins dimerize via their 4 beta-sheets and thereby form one complete turn of the helical structure of a beta-fibril (Fig. 3). The integration of *Beta1* dimers into fibrils likely occurs via edge-to-edge interactions in a manner equivalent to that proposed for dimers of CBPs with 1 beta-sheet domain<sup>41,42</sup>. As there are currently no experimental data on sauropsidian CBPs, that could resolve the structure of beta-fibrils at atomic resolution<sup>41</sup>, the integration of the complete *Beta1* sequences of squamates into ongoing computer modelling attempts<sup>42</sup> will extend the scope of these studies beyond the investigation of interactions between isolated beta-sheets. As *Beta1* homologs are



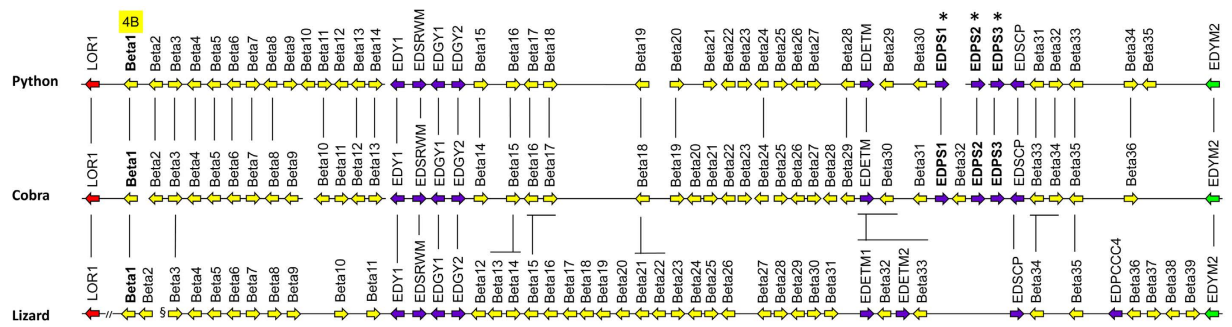
**Figure 1. Organization of the epidermal differentiation complex (EDC) in the Burmese python and the king cobra.** Genes of the EDC in snakes *Python bivittatus* and *Ophiophagus hannah*, in comparison to those of the lizard (*Anolis carolinensis*), the chicken, and human, are schematically depicted. Arrows indicate the orientation of the genes. Simple EDC (SEDC) genes with 2 exons are represented by colored arrows with a black frame whereas other genes are shown as filled arrows. Corneous beta-protein (CBP) gene clusters are shown as boxes in this diagram while detailed information about the genes in these clusters are depicted in Fig. 2. Members of gene families are numbered according to the positions of genes without indicating 1:1 orthology to specific members of the same gene family in other species. The depiction of the human EDC is simplified by representing gene family clusters with arrows and indicating the total number (#) of genes within each cluster. Black vertical lines connect orthologous genes or gene families. Green and red asterisks indicate putative gene gain and loss events whereas black asterisks indicate gene differences that could not be unambiguously assigned to an evolutionary event in a particular lineage. Note that the diagram is not drawn to scale. The symbol § marks a locus in which genes unrelated to classical EDC genes are present in the current genome sequence assembly of the cobra. Because of improved delineation of orthology relationships, the following gene names have been newly assigned to replace previous names<sup>30</sup>: lizard EDQL instead of EDCQ, chicken EDQM3 instead of EDSC, and EDPQ1 instead of EDCH5.

present in snakes, lizards, and geckos (Fig. 3), but not in other saurospids, these proteins represent an evolutionary innovation of squamates.

Interestingly, several genes within the CBP locus encode proteins that lack a beta-sheet-forming domain, i.e. the defining feature of CBPs, but share the exon-intron organization with CBP and other SEDC genes (Fig. 2, Suppl. Fig. S1 and S2). The positions of these non-CBP genes relative to specific CBPs are largely conserved between snakes and the green anole lizard (Fig. 2). BLAST searches with the sequences of newly identified EDC genes allowed us to identify previously uncharacterized homologs of these genes in the green anole lizard (Suppl. Table S6; Suppl. Fig. S6).

**The EDC of squamates contains multiple genes that encode proteins with extremely high cysteine contents.** The amino acid sequences of snake EDC proteins were analyzed for the presence of conserved sequence motifs. Snake S100A genes and SFTPs contain an amino-terminal S100 domain<sup>43</sup> while snake PGLYRP3 is predicted to acquire the characteristic structural fold also found in other PGLYRPs<sup>43</sup>. Among SEDC proteins, only CBPs contain a structural motif, i.e. the beta-sheet forming region, whereas other SEDCs of snakes do not contain sequences associated with the propensity to fold into a known protein domain. However, many snake SEDC proteins have amino-terminal and carboxy-terminal sequence motifs that are conserved in SEDCs of a diverse range of amniotes, including humans (Suppl. Fig. S7).

SEDC proteins of amniotes are generally characterized by amino acid sequence repeats and high abundance of one or several of the following amino acid residues: glycine (G), serine (S), proline (P), glutamine (Q), and cysteine (C)<sup>30,33</sup>. Likewise, the EDC of the python comprises genes that encode proteins with high contents of G (e.g., loricrins 1 and 2, > 34% G), S (EDCS3, 36% S; EDPS1, 33% S), P (EDPQ2, 34% P; EDPCK, 38% P) and Q (EDPQ2, 24% Q) (Fig. 4 and Suppl. Fig. S1). Cysteine-rich amino acid sequences are encoded by many SEDC genes of the snakes, which is surprising in the context of the hypothesis that cysteine-rich proteins function mainly in hard and resistant skin appendages such as claws, hair and feathers, and tend to be lost in species, such as snakes, that lack these skin appendages<sup>28</sup>.



**Figure 2. Organization of the corneous beta-protein (CBP), also known as beta-keratin, gene cluster in snakes.** The CBP gene clusters in python (*P. bivittatus*), cobra (*O. hannah*) and lizard (*A. carolinensis*) are schematically depicted. The CBP genes were tentatively named “Beta” followed by a number that indicates the position in the cluster. The CBP genes of the lizard correspond to the “Li-Ac” genes reported previously<sup>26</sup>, whereby Beta1 and 2 are identical to Li-Ac40 and 39, respectively, and Beta3 through 39 are identical to Li-Ac37 through 1. A gap in the lizard genome sequence assembly (\$) is the likely locus of Beta40 (Li-Ac38). Arrows indicate the orientation of the genes. CBP genes are represented by yellow arrows whereas non-CBP genes inside the cluster are shown by violet arrows. The label “4B” marks the presence of 4 beta-sheets in the encoded protein (Fig. 3) and asterisks indicate snake-specific genes (Suppl. Fig. S8). *LOR1* and *EDYM2* are conserved genes flanking the CBP gene cluster. Vertical lines indicate orthologs. Note that the schemes are not drawn to scale.

Fifteen EDC genes of the python encode proteins containing at least 20% cysteine, and among them 5 proteins have a cysteine content above 35% (Fig. 4B). Five python EDC genes encoding proteins with high cysteine content were clustered between *EDPQ1* and *EDQL*, i.e. two genes evolutionarily conserved in lizard and chicken (Fig. 1). This region of the python EDC is syntenic with the locus of a gene encoding the cysteine-rich feather protein (EDCRP) in the chicken<sup>32</sup>, while the homologous locus of the green anole lizard contains a gene encoding an EDCRP-like protein<sup>32</sup> and 9 other cysteine-rich proteins (EDCS1, EDCS3, EDPCCC1-4, EDGPC1-2 and EDCQ3). Among the latter, 8 were identified in the present study (Fig. 1; Suppl. Fig. S6). The highest cysteine contents so far detected among squamate EDC proteins are present in *A. carolinensis* EDPCCC4 (45% cysteine) and EDPCCC1 of the king cobra (46.6% cysteine). EDPCCC1 is conserved among snakes, and by the analysis of tissue transcriptomes of *E. coloratus* we could confirm expression in the skin whereas internal organs lacked EDPCCC1 (Suppl. Fig. S4C). The cysteine-rich proteins of snakes and other squamates are characterized by repetitive amino acid sequences with clusters of two or more cysteine residues (Fig. 4B). These proteins are candidates to become cross-linked components of the hard scales of squamates but this hypothesis remains to be tested in future studies.

**Identification of snake-specific EDC gene innovations and losses.** Differences in the EDCs of snakes and the anole lizard (Fig. 1) suggest that gene innovations or gene losses have occurred in either one of the evolutionary lineages leading to snakes and iguanians (represented here by *A. carolinensis*). To determine the ancestral condition for each of the clade-specific genes, we searched for orthologous genes in the Japanese gecko (*G. japonicus*), representing Gekkota, a basal clade of squamates and the closest outgroup to Toxicofera<sup>44–46</sup> with a sequenced genome<sup>39</sup>. These comparisons suggested that some EDC genes have been lost in snakes whereas others have originated in the snake lineage.

Genes of the EDCC family (Fig. 5A), which encode proteins rich in cysteine-cysteine motifs, are present in the Japanese gecko ( $n = 6$ ) and in the green anole lizard ( $n = 3$ ) but not in snakes (a sister group of the green anole lizard). This species distribution of EDCCs suggests that EDCC gene(s) were present in the last common ancestor of snakes and lizards and later underwent inactivation in snakes. In agreement with this hypothesis, we identified a mutated remnant of an EDCC gene located between *loricrin 1* and the CBP gene cluster of the python (Fig. 5B). Orthologs of EDCC are absent from the EDCs of birds and turtles<sup>30,33</sup>, indicating that EDCCs represent a squamate-specific gene innovation. The gene *EDYM1*, which is conserved in turtles and birds<sup>33</sup>, is located between *EDCC* and *loricrin* genes in the anole lizard and the gecko whereas it is absent in snakes (Fig. 1). These data suggest that *EDYM1* and the *EDCC* genes of the last common ancestor of Toxicofera (snakes and iguanid lizards) were lost in snakes.

A group of apparently snake-specific proteins are encoded by the genes *EDPS1* through *EDPS3* (*epidermal differentiation proteins rich in proline and serine 1-3*) (Suppl. Fig. S8), which are located within the CBP gene cluster of snakes but not in the CBP cluster of other squamates (Fig. 2). The investigation of tissue transcriptomes of *E. coloratus* suggested that EDPS homologs are expressed in the skin but not in internal organs (Suppl. Fig. S4E).

**Duplication of scaffoldin in snakes.** Snakes have 3 SFTF genes (*Crnn*, *Scfn1*, *Scfn2*) (Fig. 6) whereas the green anole lizard has only one (*Scfn*) while the bearded dragon, the Japanese gecko, the American alligator and the chicken have two (*Crnn*, *Scfn*)<sup>30,31</sup>. Both isoforms of snake scaffoldin proteins are rich in glutamic acid (E) and arginine (R) residues (Fig. 6A), which are also highly abundant in human trichohyalin but not in cornulin<sup>31</sup>. Expression of both *Scfn* genes was confirmed by intron-spanning RNA-seq reads in the painted saw-scaled viper (*Echis coloratus*) (Suppl. Fig. S9).

**A**

```

Core seq.: *****
Consensus: PCVT P KVVI PPP ITLPGP L P EC ET C
Python MSGSGVKCVTTPCITS-CPDAKVVVHPPPLIITLPLGLSLRTSPNQCLLETQTSCLTNGSEVGCNEGSTAIIVTKSSGVCALSGG-----
Cobra MSGSGVKCVTTPCVTS-CPDAKVVVHPPPLIITLPLGLSLRTSPNQCLLVEVTSCLTNGSEVGCNEGSTAIIVTKSSGVCALSGG-----
Lizard MSGTFVKCLTEPCVTK-CPDAKVIHPPPLAITLPLGLLRTNPGKCLVETQTSCLTNGSEVGCNEGSTAIIVTKSSGVCALSGG-----
Gecko MSKTGVKCVTTPCVTS-IPDAKVIHPPPLIITLPLGLLRTNPGKCLVETQTSCLTNGSEVGCNEGSTAIIVTKSSGVCALSGG-----

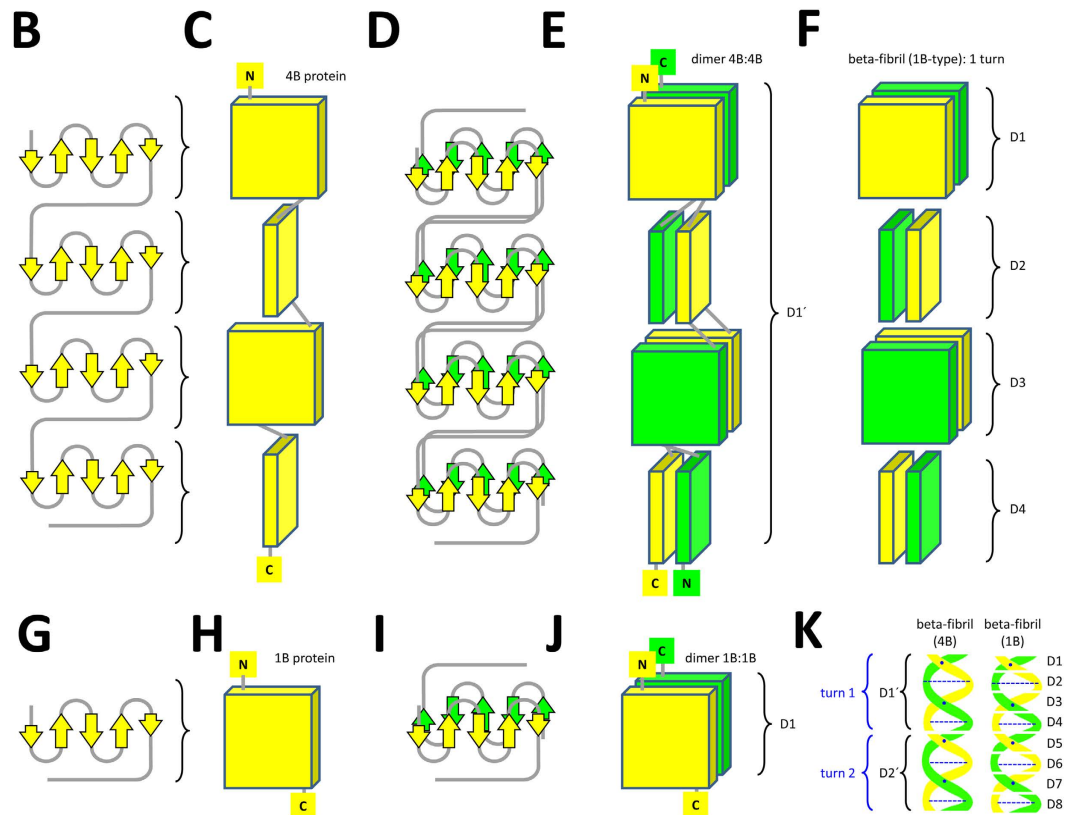
Python -----DTPCCITTCFDSQVVIQPPVCIITIPGAVLTSYPNECLIKTSTPCVPSGQPPALTRSTSVSDCSLTPRRLRSRA-----
Cobra -----DTPCCITTCFDSQVVIQPPVCIITIPGAVLTSYPNECLISSSTNCITSGVQRALPRSTSVSDCSLTPRRLRSRA-----
Lizard -----DTPCSTTACSDSQLVIYPPVCIITIPGAVLTSYPNECLIESSIPCVPSETHPLALTRSSISDSCSVIPQRTTRST-----
Gecko -----TPCCITTCFDSQVVIQPPVCIITIPGAVLTSYPNECLIIATSQPCIPAGSQPPAITRSASVSDGPMVSTPLRSR-----

Python -SVP SGLDLSRCVTRG-PTTKVVIHPPPIEVTIPGPVLEIAAQE CAVEVYNPC---SSNNALTSSDQKAITSSDE--GEVKALTQAKS
Cobra -SVP SGLNSTRCLTQG-PSNKVVIYPPPIEITIPGPVLEIAAEE CTVEVYNPC---TDNNALTSSDQKAITSSDE--EVKTLIPRAKS
Lizard -SVP CGLPSDNPCVTQG-PGSKVVIHPPPIEITIPGPVLEIAAEE CAVEVYNPC---E--AITGEEQCAITSGDEDTEGSKALATRVKD
Gecko -SVP CGLDTASACVAQG-PSSKVIYPPPIEITIPGPVLEIAAEE CAVEVHNPC---E--AITGEEQCAITSGDEDTEGSKALATRVKD

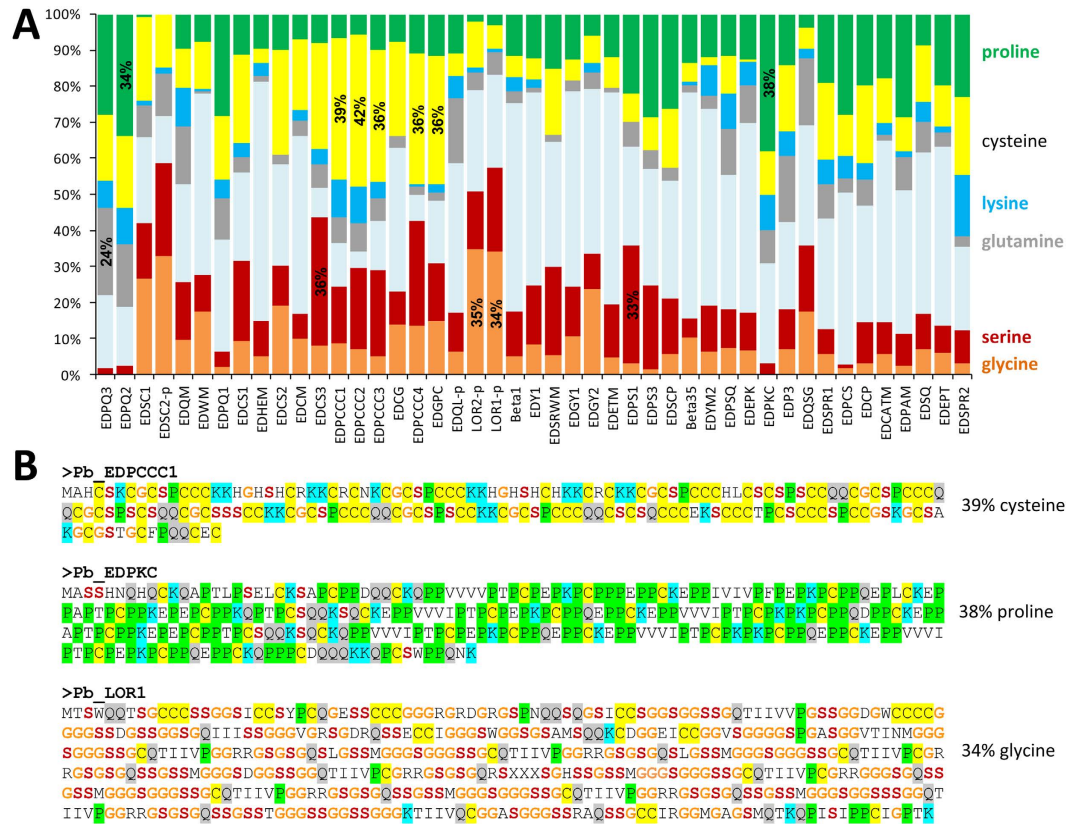
Python CTTVCGLMDTSSCISQG-PEMKIITIQPPPIEVELPGPILEVFPPEACKIETLNPSF---LEPEAITGSETKALCSSK-M-T-STALAT----
Cobra CTTVSGLMDTSCISQG-PEMKIITIQPPPIEVELPGPILEVFPPEACKVETLTPCF---PKPEAITGNETKALCSSK-I-P-STALAT----
Lizard SATICALLGASSCVSEG-PELRIVVKPPPIEVDMPGPIILQVFPETCKVEVLNMSG---SHTPALDSSKFEYALCSSNE-MTP-SRALAT----
Gecko CTEVCGSLGSSPCVSG-PEMKIITIQPPPIEVDLPGPIILRVFPPECKVETLTPCFAPENLALGESSR--P-ALTNGD----SS-LST----

Python TS-N---VKRPLENLRRAFRPFAEMYSRSMTPRTIALSQQTRLAKYRNTLHS-----MNFQPSF-----
Cobra TTS---TKRPLDVRPRPFAEMYSRSMTPRSLALSQQSRLAKYRSALYS-----MHSQSSY-----
Lizard RRSSLATKHLPLDLRRPFRPFAEMYSRSMTPRTIALSQQTRLAKYRNTLHS-----MNFQPSF-----
Gecko -----GKRPLEDIRRFRPFAEMYSRSMTPRSLTAGQEMI-KHRNSFSS-----ASCY-----

```



**Figure 3. A corneous beta-protein with 4 beta-core sequence motifs is predicted to facilitate the formation of unique beta-fibrils in squamates.** (A) Amino acid sequence alignment of Beta1 proteins of python (*P. bivittatus*), cobra (*O. hannah*), lizard (*A. carolinensis*) and gecko (*G. japonicus*). Internal sequence repeats were aligned to illustrate the conservation of segments corresponding to the CBP core sequences (indicated by \* above the alignment) (see also Suppl. Fig. S5). Key residues of the repeat consensus sequence are indicated. Putative beta-strand-forming residues are indicated by yellow shading. Proline and cysteine residues are shown in green and red fonts, respectively. (B–K) Modeling of protein structures. The model for the folding (B,C) and dimerization (D,E) of CBPs with 4 beta-sheets (B4 proteins) was adapted from the model for CBPs with 1 beta-sheet (B1 proteins) (G–J) which also proposes a helical arrangement of 4 CBP dimers in 1 turn of the helical structure of a beta-fibril<sup>41</sup> (F). The model for B4 proteins suggests that a dimer (D1') (E) is equivalent to 1 turn of a beta-fibril (F) and 2 B4-protein dimers (D1' and D2') substitute for 8 B1-protein dimers to form 2 turns of a beta-fibril helix (K). The orientation of the dimerization interfaces is indicated by dashed lines and dots in (K). Note that the schematic depiction of beta-sheets is simplified and does not show a twist typically present in beta-sheets. C, carboxy-terminus; N, amino-terminus.



**Figure 4. SEDC genes encode proteins with extremely biased amino acid composition.** (A) The diagram shows the amino acid compositions of SEDC proteins of *Python bivittatus*. The protein data are shown in the order of the corresponding genes in the EDC (Fig. 1). Translation products of genes within the corneous beta-protein (CBP) gene cluster (Fig. 2) are not included here, with the exception of proteins encoded by the first and the last CBP. (B) Amino acid sequences of exemplary SEDC proteins of the *P. bivittatus*.

Phylogenetic profiling and gene locus comparison suggested that the *Scfn2* gene originated by duplication of the primordial *Scfn* gene specifically in snakes (Fig. 6B). The sequences of the proximal promoters of *Scfn1* and *Scfn2* genes were partly conserved, and they contained homologous TATA boxes. Remarkably, the putative binding sites for 2 transcription factors (KLF4 and AP-1) in promoters of SFTP genes<sup>31</sup> were differentially conserved in snake *Scfn1* and *Scfn2* (Fig. 6B, Suppl. Fig. S10). The predicted binding site for AP-1 was present in the promoters of *Scfn1* but not of *Scfn2* genes whereas the KLF4 binding site was present in the promoters of *Scfn2* but not of *Scfn1* genes of both python and cobra. Taken together, these data suggest a scenario in which a single ancestral *Scfn* gene was present in the last common ancestor of snakes and the green anole lizard, this gene was duplicated in primitive snakes, and the derived genes underwent divergent evolution of their promoter sequences.

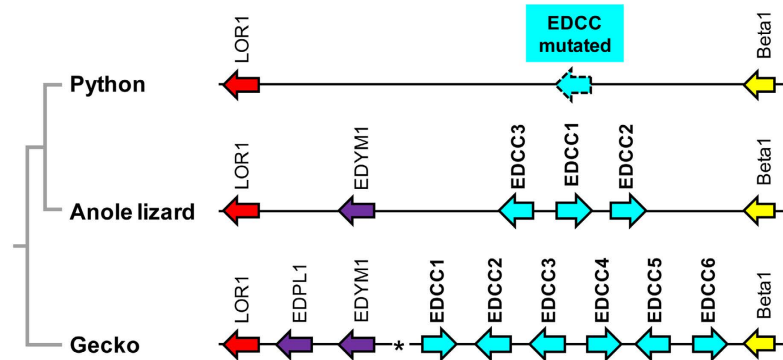
Moreover, the SCFN2 proteins lack a carboxy-terminal sequence motif (CTM) that is present in SCFN1 of snakes and in most other SFTPs of amniotes<sup>31,47</sup> (Fig. 6). As this motif has been implicated in keratin filament binding of SFTPs<sup>47</sup>, divergent amino acid sequence evolution appears to have caused also differences in the functions of scaffoldins 1 and 2 of snakes.

### Discussion

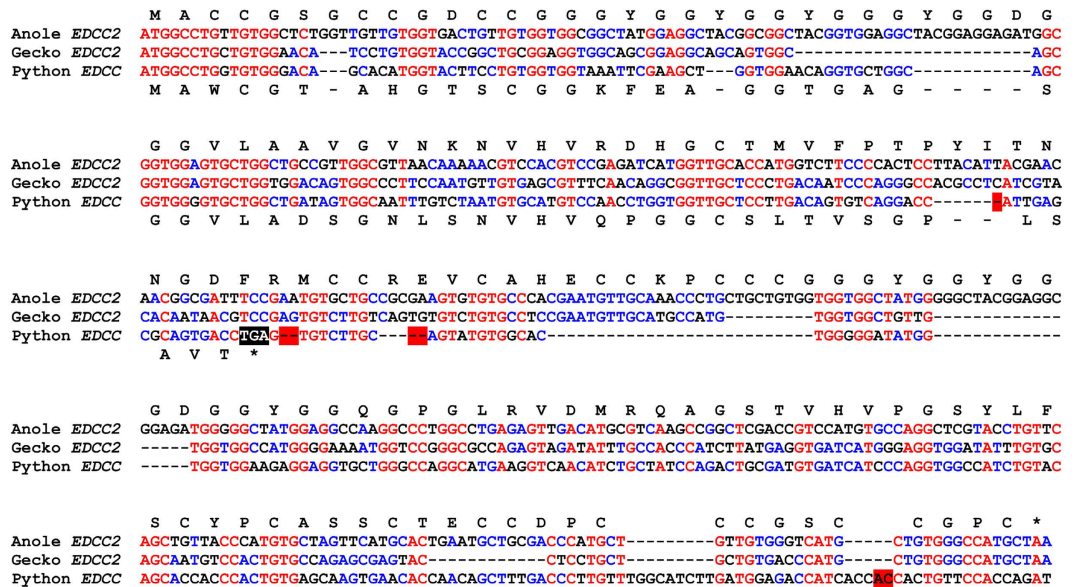
The results of the present study shed new light on the molecular composition and evolution of the epidermis in snakes and other squamates. The comparative analysis of snake EDCs suggests that the epidermis of snakes contains many more proteins than the small set of CBPs (beta-keratins) identified in previous studies<sup>25,26</sup>. Both the number of CBPs and the number of other EDC genes of snakes is similar to those of the green anole lizard, and the total number and sequence diversification of epidermal differentiation genes (including CBPs) in snakes exceeds that present in mammals (Fig. 1). The new data therefore indicate that the process of keratinocyte cornification in the epidermis of snakes requires the participation of numerous proteins aside CBPs, like in other sauropsids. While crucial roles of mammalian EDC genes such as loricrin, LCEs, trichohyalin, and filaggrin in the skin barrier of mammals have been defined by a long series of studies<sup>48–52</sup>, the identification of the EDC gene complement of snakes is the pivotal starting point for a comprehensive investigation of epidermal differentiation in this important subgroup of reptiles.

Important limitations of this study were the quality of the genome sequences that were available for analysis and the focus of our study on genes of the EDC. The current genome sequences of squamates are not of the same quality as those of mammalian model species and, therefore, some gaps are present in our model of the EDC in snakes (Fig. 1). The aim of this study was the characterization of the EDC in snakes; and other genome loci that control

A



B

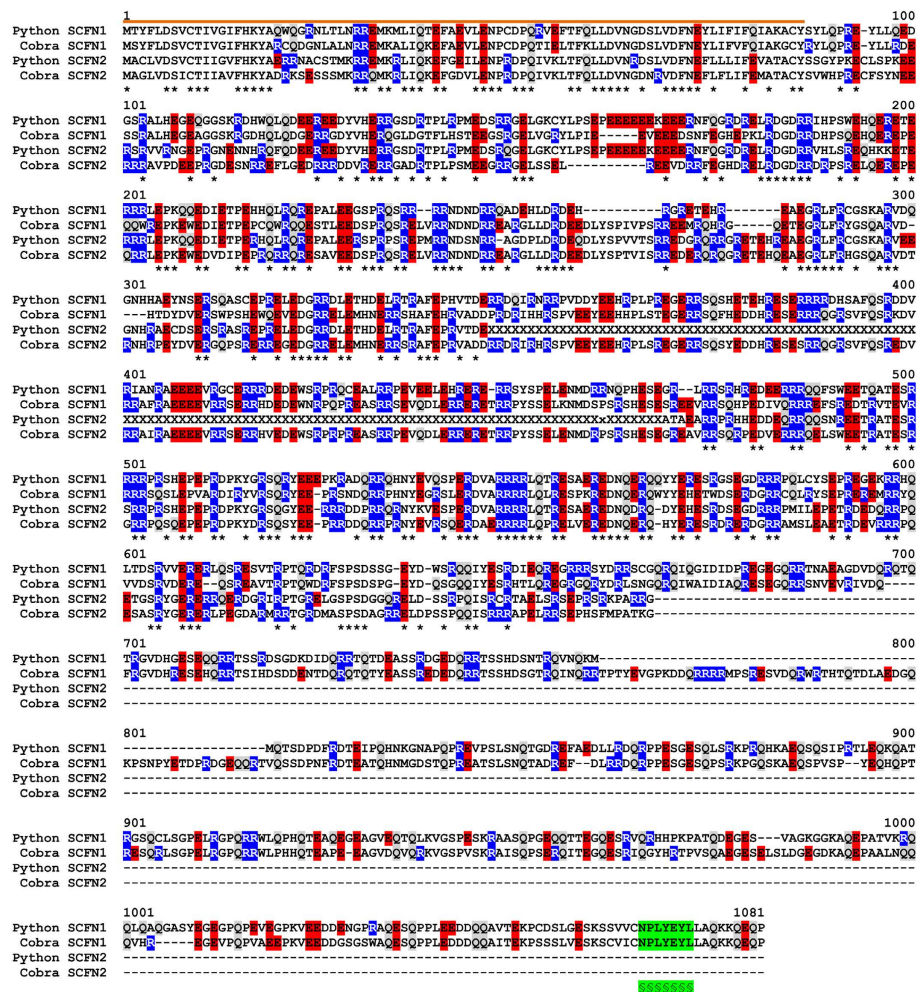


**Figure 5. EDCC genes have accumulated inactivating mutations in snakes. (A)** The gene loci of EDCC and EDYM1 genes in the green anole lizard (*Ac*, *Anolis carolinensis*) and the gecko (*Gj*, *Gekko japonicus*) were compared to that of the python (*Pb*, *Python bivittatus*). **(B)** Nucleotide sequence alignment of the mutated EDCC sequence of the python and EDCC genes of the anole lizard and the gecko. Red fonts indicate conservation in all 3 sequences and blue fonts indicate conservation in 2 of 3 sequences. Red shading indicates frame-shift mutations and a premature stop codon in the python genes is shown with white fonts on black background. Amino acid sequences obtained by *in silico* translation of anole and python genes are shown above and below the nucleotide sequence alignment, respectively.

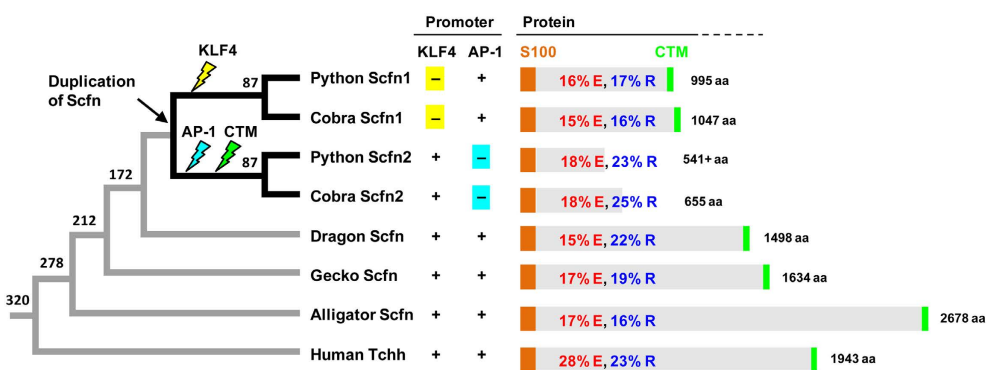
distinct aspects of epidermal differentiation, such as the enzymatic control of protein cross-linking during keratinocyte cornification and the disruption of intercellular junctions during ecdysis have not been investigated here.

The EDC of snakes is largely syntenic with that of the green anole lizard, and only few genes are not orthologous between the two taxa. As snakes lack the specialized epidermal differentiation pathways that lead to the formation of claws and toe pad lamellae, the large degree of EDC gene conservation between snakes and the anole lizard suggests that the great majority of EDC genes play essential roles in skin structures unrelated to limb-specific appendages. Nevertheless, the present identification of several snake-specific changes in the EDC gene complement (loss of EDCCs and EDYM1, origin of EDPS, duplication of scaffoldin) points to unique characteristics of the epidermal differentiation in snakes, likely evolved during their specific adaptation to their environment. EDCCs contain multiple CC dipeptides and stretches of proline residues, respectively, and these proteins show limited similarity with other SEDC proteins, indicating that they have non-redundant roles. Likewise, the amino acid sequences of scaffoldins 1 and 2, as well as their promoters, differ substantially. Recently, we have shown that scaffoldin of the chicken is expressed in the embryonic periderm and in epithelial cells that support the morphogenesis of claws and feathers by providing a transient epithelial scaffold which degenerates after maturation of these skin appendages<sup>31</sup>. These data contributed to the evolutionary-developmental model that connects the embryonic archosaur scale and feathers<sup>16–18,53</sup>. The identification of the genes encoding

**A**



**B**



**Figure 6. Duplication and sequence diversification of scaffoldin in snakes.** (A) Amino acid sequence alignment of scaffoldin (SCFN) 1 and 2 proteins of python and cobra. The S100 domain is indicated by an orange-colored line above the alignment. A carboxy-terminal motif (CTM) (\$) implicated in keratin binding is highlighted by green shading. The predominant amino acid residues, i.e. glutamic acid (E) and arginine (R), are highlighted by red and blue shading. Positions with identical amino acid residues in all 4 sequences are indicated by asterisks below the alignment. (B) Schematic phylogenetic tree for SFTPs in amniotes. Numbers at the branching points of the species tree indicate the divergence times (million years ago) of phylogenetic lineages. The presence of putative binding sites for the transcription factors KLF4 and AP-1 in the promoters (Suppl. Fig. S10) are indicated. The organization of the proteins is depicted schematically with orange boxes indicating the S100 domain and green boxes indicating the CTM. The contents of glutamate (E) and arginine (R) of each protein are shown. aa, amino acids; Tchh, trichohyalin.



scaffoldins 1 and 2 will facilitate the determination of their expression during embryonic development and during the shedding cycle of snakes in future studies. Likewise, the future investigation of other EDC genes, that have either been lost or acquired specifically in the snake lineages, will help to elucidate differences between the epidermis of snakes and other squamates.

The presence of multiple genes encoding cysteine-rich EDC proteins in snakes suggests that a high cysteine content of proteins is not only required for hard skin appendages such as claws, hair and feathers. These skin appendages consist of entirely cornified proteinaceous structures and are absent in snakes. Hard skin appendages utilize cysteine-dependent disulfide protein cross-linking to acquire high mechanical resilience. In mammals, cysteine-rich keratins are components of hair and nails whereas keratins with low cysteine content are components of the soft epidermis<sup>54</sup>. Cysteine-rich keratin-associated proteins (Krtaps) are further components of mammalian hair and nails, and a cysteine-rich EDC protein (EDCRP) is a component of avian feathers<sup>30,32</sup>. Our detection of multiple EDC genes for high-cysteine proteins in snakes suggests that these epidermal proteins can have functions unrelated to claws, hair and feathers. Pythons and boas have spurs (rudimentary claws) that are located next to their cloaca, but other snakes do not have homologs of claws. Thus, disulfide bond-mediated cross-linking of cysteine-rich proteins may contribute to the maturation of hard scales in snakes and probably also in other squamates. In this regard, it will be interesting to compare the expression pattern of cysteine-rich EDC proteins in different types of snake scales, i.e. flat and tough scales on the head, keeled and perhaps softer scales on the dorsum and the sides of the body, and large, mechanically resistant ventral scales (gastrosteges) that are utilized for movement. However, the presence of high cysteine contents in the absence of hard skin appendages may also point to a role of cysteine residues that is unrelated to disulfide bond formation. Cysteine residues have been identified as attachment sites for palmitic acid which allows the anchoring of proteins into membranes<sup>55</sup>. The process of cysteine palmitoylation has been demonstrated in mammalian skin proteins, but whether a similar process also occurs in snakes requires further investigations.

Our finding that a unique CBP (beta-keratin) comprising 4 beta-sheets is conserved among and specific for squamates indicates that this protein contributes to unique properties of epidermal keratinocytes in squamates. Previous immunolabeling studies in the green anole lizard have suggested expression of Beta1, also referred to as Li-Ac40, in the beta-layer of scales on different body sites investigated<sup>56</sup>. The immunolabeling for this large beta-protein was associated with filaments of 3 nm thickness, i.e. the characteristic diameter of beta-fibrils. The classical model of the beta-fibril structure was developed more than 30 years ago on the basis of X-ray diffraction studies<sup>38,40</sup>. According to this model, 4 CBP (then called beta-keratin) dimers form a turn of a left-handed helix with four repeating units per turn. It is now striking that the only CBP comprising more than 1 beta-sheet-forming segment contains 4 such elements, indicating that it may span exactly 1 turn of the fibril. Of note, an alternative model in which 4 CBP dimers would be arranged to form half of a helix turn has also been reported<sup>42</sup>. Many aspects of beta-fibril formation in squamates and sauropsids in general are still open and comparative studies on CBPs with 1 and 4 beta-sheet regions may yield valuable insights in future studies.

In conclusion, the results of this study establish a comprehensive catalog of EDC genes of snakes and, thereby, provide the basis for further studies on the molecular organization and evolution of the epidermis in snakes and other squamates.

## Methods

**Genome sequences and gene identification.** Genome sequences from the following squamate species were used for gene predictions: Burmese python (*Python bivittatus*)<sup>10</sup>, king cobra (*O. hannah*)<sup>11</sup>, painted saw-scaled viper (*Echis coloratus*)<sup>36</sup>, bearded dragon (*Pogona vitticeps*)<sup>57</sup> and the Japanese gecko (*Gekko japonicus*)<sup>39</sup>. The accession numbers of genome sequence scaffolds corresponding to the EDC are listed in Supplementary Tables S2 through S6. Coding sequences of EDC genes were predicted using a combination of the following approaches. Amino acid sequences of EDC proteins of *A. carolinensis*<sup>30</sup> and humans<sup>35</sup> were used as queries in tBLASTn searches against the nucleotide sequence between *S100A12* and *S100A11* genes of the target genome. Information about exon coverage by RNA-seq reads, available in the NCBI browser for “genomic regions, transcripts, and products”, was used to identify transcribed regions in the EDC of *P. bivittatus*. The transcribed regions were translated and the resulting amino acid sequences were compared to those of known EDC proteins. The nucleotide sequence of EDC regions of apparently low gene density was translated *in silico*, and additional open reading frames of candidate EDC genes were identified according to published criteria<sup>33</sup>. Predictions of snake EDC genes were validated by BLAST searches in the transcriptomes of *P. regius* and *E. coloratus*<sup>36,37</sup>.

**Bioinformatic analysis of gene promoters and amino acid sequences encoded by EDC genes.** For the assessment of transcription factor binding scores in the promoter sequences of SFTP genes, the JASPAR 2016 server (<http://jaspar.genereg.net>) was used<sup>58</sup>. Primary and secondary structure analyses of the proteins were performed on the PSIPRED protein structure prediction server<sup>59</sup> and using the software tools at the ExpASY SIB Bioinformatics Resource Portal<sup>60</sup>.

## References

- Shine, R. Reptiles. *Curr. Biol.* **23**, R227–31 (2013).
- Pincheira-Donoso, D., Bauer, A. M., Meiri, S. & Uetz, P. Global taxonomic diversity of living reptiles. *PLoS One.* **8**:e59741 (2013).
- Close, M. & Cundall, D. Snake lower jaw skin: extension and recovery of a hyperextensible keratinized integument. *J. Exp. Zool. A. Ecol. Genet. Physiol.* **321**, 78–97 (2014).
- Klein, M. C. & Gorb, S. N. Epidermis architecture and material properties of the skin of four snake species. *J. R. Soc. Interface.* **9**, 3140–55 (2012).
- Filippov, A. E. & Gorb, S. N. Modelling of the frictional behaviour of the snake skin covered by anisotropic surface nanostructures. *Sci. Rep.* **6**, 23539 (2016).

6. Landmann, L. The skin of reptiles. Epidermis and dermis. In: *Biology of the Integument. 2 Vertebrates* (ed. Bereiter-Hahn, J., Matoltsy, G., Richards, K. S.) 150–187 (Springer, 1986).
7. Maderson, P. F. A., Rabinowitz, T., Tandler, B. & Alibardi, L. Ultrastructural contributions to an understanding of the cellular mechanisms involved in lizard skin shedding with comments on the function and evolution of a unique Lepidosaurian phenomenon. *J. Morphol.* **236**, 1–24 (1998).
8. Maderson, P. F. A. Histological changes in the epidermis of snakes during the sloughing cycle. *J. Zool.* **146**, 98–113 (1965).
9. Landmann, L. Keratin formation and barrier mechanisms in the epidermis of Natrix natrix (Reptilia, Serpentes): an ultrastructural study. *J. Morphol.* **162**, 93–126 (1979).
10. Castoe, T. A. *et al.* The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc. Natl. Acad. Sci. USA* **110**, 20645–50 (2013).
11. Vonk, F. J. *et al.* The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. USA* **110**, 20651–6 (2013).
12. Tollis, M., Hutchins, E. D. & Kusumi, K. Reptile genomes open the frontier for comparative analysis of amniote development and regeneration. *Int. J. Dev. Biol.* **58**, 863–871 (2014).
13. Eckhart, L., Lippens, S., Tschachler, E. & Declercq, W. Cell death by cornification. *Biochim. Biophys. Acta* **1833**, 3471–3480 (2013).
14. Alibardi, L. Sauropsids cornification is based on corneous beta-proteins, a special type of keratin-associated corneous proteins of the epidermis. *J. Exp. Zool. B. Mol. Dev. Evol.* **326**, 338–351 (2016).
15. Chang, C. *et al.* Reptile scale paradigm: Evo-Devo, pattern formation and regeneration. *Int. J. Dev. Biol.* **53**, 813–826 (2009).
16. Dhouailly, D. A new scenario for the evolutionary origin of hair, feather, and avian scales. *J. Anat.* **214**, 587–606 (2009).
17. Musser, J. M., Wagner, G. P. & Prum, R. O. Nuclear  $\beta$ -catenin localization supports homology of feathers, avian scutate scales, and alligator scales in early development. *Evol. Dev.* **17**, 185–194 (2015).
18. Di-Poi, N. & Milinkovitch, M. C. The anatomical placode in reptile scale morphogenesis indicates shared ancestry among skin appendages in amniotes. *Sci. Adv.* **2**, e1600708 (2016).
19. Alibardi, L. Transition from embryonic to adult epidermis in reptiles occurs by the production of corneous beta-proteins. *Int. J. Dev. Biol.* **58**, 829–839 (2014).
20. Alibardi, L. Ultrastructure of the embryonic snake skin and putative role of histidine in the differentiation of the shedding complex. *J. Morphol.* **251**, 149–68 (2002).
21. Alibardi, L. & Toni, M. Cytochemical, biochemical and molecular aspects of the process of keratinization in the epidermis of reptilian scales. *Prog. Histochem. Cytochem.* **40**, 73–134 (2006).
22. Alibardi, L. Review: mapping epidermal beta-protein distribution in the lizard *Anolis carolinensis* shows a specific localization for the formation of scales, pads, and claws. *Protoplasma.* **253**, 1405–1420 (2016).
23. Roberts, J. B. & Lillywhite, H. B. Lipid barrier to water exchange in reptile epidermis. *Science* **207**, 1077–9 (1980).
24. Tu, M. C., Lillywhite, H. B., Menon, J. G. & Menon, G. K. Postnatal ecdysis establishes the permeability barrier in snake skin: new insights into barrier lipid structures. *J. Exp. Biol.* **205**, 3019–30 (2002).
25. Dalla Valle, L., Nardi, A., Belvedere, P., Toni, M. & Alibardi, L. Beta-keratins of differentiating epidermis of snake comprise glycine-proline-serine-rich proteins with an avian-like gene organization. *Dev. Dyn.* **236**, 1939–53 (2007).
26. Dalla Valle, L., Nardi, A., Bonazza, G., Zucal, C., Emera, D. & Alibardi, L. Forty keratin-associated beta-proteins (beta-keratins) form the hard layers of scales, claws, and adhesive pads in the green anole lizard, *Anolis carolinensis*. *J. Exp. Zool. B. Mol. Dev. Evol.* **314**, 11–32 (2010).
27. Dalla Valle, L., Nardi, A. & Alibardi, L. Isolation of a new class of cysteine-glycine-proline-rich beta-proteins (beta-keratins) and their expression in snake epidermis. *J. Anat.* **216**, 356–67 (2010).
28. Dalla Valle, L., Benato, F., Rossi, C., Alibardi, L., Tschachler, E. & Eckhart, L. Deleterious mutations of a claw keratin in multiple taxa of reptiles. *J. Mol. Evol.* **72**, 265–73 (2011).
29. Klein, M. C. & Gorb, S. N. Ultrastructure and wear patterns of the ventral epidermis of four snake species (Squamata, Serpentes). *Zoology (Jena)* **117**, 295–314 (2014).
30. Strasser, B. *et al.* Evolutionary origin and diversification of epidermal barrier proteins in amniotes. *Mol. Biol. Evol.* **31**, 3194–3205 (2014).
31. Mlitz, V. *et al.* Trichohyalin-like proteins have evolutionarily conserved roles in the morphogenesis of skin appendages. *J. Invest. Dermatol.* **134**, 2685–2692 (2014).
32. Strasser, B., Mlitz, V., Hermann, H., Tschachler, E. & Eckhart, L. Convergent evolution of cysteine-rich proteins in feathers and hair. *BMC Evol. Biol.* **15**, 82 (2015).
33. Holthaus, K. B. *et al.* Comparative genomics identifies epidermal proteins associated with the evolution of the turtle shell. *Mol. Biol. Evol.* **33**, 726–37 (2016).
34. Mlitz, V., Tajamul, H., Tschachler, E. & Eckhart, L. Filaggrin has evolved from an “S100 fused-type protein” (SFTP) gene present in a common ancestor of amphibians and mammals. *Exp. Dermatol.* Electronic publication ahead of print. doi: 10.1111/exd.13317.
35. Henry, J. *et al.* Update on the epidermal differentiation complex. *Front. Biosci.* **17**, 1517–32 (2012).
36. Hargreaves, A. D., Swain, M. T., Logan, D. W. & Mulley, J. F. Testing the Toxicofera: comparative transcriptomics casts doubt on the single, early evolution of the reptile venom system. *Toxicon* **92**, 140–56 (2014).
37. Hargreaves, A. D., Swain, M. T., Hegarty, M. J., Logan, D. W. & Mulley, J. F. Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. *Genome Biol. Evol.* **6**, 2088–2095 (2014).
38. Fraser, R. D. & Parry, D. A. The molecular structure of reptilian keratin. *Int. J. Biol. Macromol.* **19**, 207–11 (1996).
39. Liu, Y. *et al.* Gekko japonicus genome reveals evolution of adhesive toe pads and tail regeneration. *Nat. Commun.* **6**, 10033 (2015).
40. Gregg, K. & Rogers, G. Feather keratins: composition, structure and biogenesis. In: *Biology of the Integument* (ed. Bereither-Hahn, J., Matoltsy, G., Sylvia-Richards, K.) 666–694 (Springer, 1986).
41. Fraser, R. D. B. & Parry, D. A. D. Filamentous structure of hard  $\beta$ -keratins in the epidermal appendages of birds and reptiles. *Subcell. Biochem.* **82**, 231–252 (2017).
42. Calvaresi, M., Eckhart, L. & Alibardi, L. The molecular organization of the beta-sheet region in Corneous beta-proteins (beta-keratins) of sauropsids explains its stability and polymerization into filaments. *J. Struct. Biol.* **194**, 282–91 (2016).
43. Marchler-Bauer, A. *et al.* CDD: NCBI’s conserved domain database. *Nucleic Acids Res.* **43**, D222–6 (2015).
44. Wiens, J. J. *et al.* Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biol. Lett.* **8**, 1043–1046 (2012).
45. Vidal, N. & Hedges, S. B. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. *C. R. Biol.* **328**, 1000–8 (2005).
46. Reeder, T. W. *et al.* Integrated analyses resolve conflicts over squamate reptile phylogeny and reveal unexpected placements for fossil taxa. *PLoS One* **10**, e0118199 (2015).
47. Strasser, B., Mlitz, V., Fischer, H., Tschachler, E. & Eckhart, L. Comparative genomics reveals conservation of filaggrin and loss of caspase-14 in dolphins. *Exp. Dermatol.* **24**, 365–369 (2015).
48. Smith, F. J. *et al.* Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat. Genet.* **38**, 337–42 (2006).
49. Palmer, C. N. *et al.* Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat. Genet.* **38**, 441–6 (2006).

50. de Cid, R. *et al.* Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat. Genet.* **41**, 211–5 (2009).
51. Medland, S. E. *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am. J. Hum. Genet.* **85**, 750–5 (2009).
52. Ishitsuka, Y. *et al.* Lce1 family members are Nrf2-target genes that are induced to compensate for the loss of loricrin. *J. Invest. Dermatol.* **136**, 1656–63 (2016).
53. Wagner, G. P. *Homology, Genes and Evolutionary Innovation* (Princeton University Press, 2014).
54. Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochem. Cell. Biol.* **129**, 705–733 (2008).
55. Liu, K. M. *et al.* Cyclic alopecia and abnormal epidermal cornification in Zdhhc13-deficient mice reveal the importance of palmitoylation in hair and skin differentiation. *J. Invest. Dermatol.* **135**, 2603–10 (2015).
56. Alibardi, L. Immunolocalization of large corneous beta-proteins in the green anole lizard (*Anolis carolinensis*) suggests that they form filaments that associate to the smaller beta-proteins in the beta-layer of the epidermis. *J. Morphol.* **276**, 1244–1257 (2015).
57. Georges, A. *et al.* High-coverage sequencing and annotated assembly of the genome of the Australian dragon lizard *Pogona vitticeps*. *Gigascience* **4**, 45 (2015).
58. Mathelier, A. *et al.* JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **44**, D110–5 (2016).
59. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics.* **16**, 404–405 (2000).
60. Artimo, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**, W597–603 (2012).

## Acknowledgements

We thank Wolfgang Sipos for helpful discussions. This work was supported by the Austrian Science Fund (FWF): P23801 and P28004.

## Author Contributions

L.E. conceived the study, K.B.H., B.S., and V.M. performed genome and transcriptome sequence studies, K.B.H., B.S., V.M., and L.E. analyzed the results, K.B.H., L.A., E.T., and L.E. wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Holthaus, K. B. *et al.* Identification and comparative analysis of the epidermal differentiation complex in snakes. *Sci. Rep.* **7**, 45338; doi: 10.1038/srep45338 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017