Immunogenetics (2019) 71:363–372
https://doi.org/10.1007/s00251-019-01114-z

ORIGINAL ARTICLE

Convergent inactivation of the skin-specific C-C motif chemokine ligand 27 in mammalian evolution

Mónica Lopes-Marques[1] · Luís Q. Alves[1,2] · Miguel M. Fonseca[1] · Giulia Secci-Petretto[1,2] · André M. Machado[1,2] · Raquel Ruivo[1] · L. Filipe C. Castro[1,2]

Received: 28 January 2019 / Accepted: 5 April 2019 / Published online: 2 May 2019

Abstract

The appearance of mammalian-specific skin features was a key evolutionary event contributing for the elaboration of physiological processes such as thermoregulation, adequate hydration, locomotion, and inflammation. Skin inflammatory and autoimmune processes engage a population of skin-infiltrating T cells expressing a specific C-C chemokine receptor (CCR10) which interacts with an epidermal CC chemokine, the skin-specific C-C motif chemokine ligand 27 (CCL27). CCL27 is selectively produced in the skin by keratinocytes, particularly upon inflammation, mediating the adhesion and homing of skin-infiltrating T cells. Here, we examined the evolution and coding condition of *Ccl27* in 112 placental mammalian species. Our findings reveal that a number of open reading frame inactivation events such as insertions, deletions, and start and stop codon mutations independently occurred in Cetacea, Pholidota, Sirenia, Chiroptera, and Rodentia, totalizing 18 species. The diverse habitat settings and lifestyles of *Ccl27*-eroded lineages probably implied distinct evolutionary triggers rendering this gene unessential. For example, in Cetacea, the rapid renewal of skin layers minimizes the need for an elaborate inflammatory mechanism, mirrored by the absence of epidermal scabs. Our findings suggest that the convergent and independent loss of *Ccl27* in mammalian evolution concurred with unique adaptive roads for skin physiology.

Keywords Chemokines · Gene loss · Skin · Inflammation

Introduction

The mammalian skin performs a plethora of biological functions, including that of acting as a protective barrier from external harmful insults, such as invading pathogens and noxious stimuli. In this context, the role of the immune system is fundamental, involving a coherent and highly coordinated network of innate and adaptive components to ensure an adequate response to ensure homeostasis (Di Meglio et al. 2011; Pasparakis et al. 2014). Chemokines, a superfamily of polypeptides, are central in the unfolding of immune and inflammatory responses, serving as chemoattractant signals that drive the movements of immune cells in response to stimuli. In the skin, a tissue-specific T cell–homing chemokine has been described. Initially named as the cutaneous T cell–attracting chemokine, C-C motif chemokine ligand 27 (CCL27, also known as ESkine, ALP, ILC, or ILRα locus chemokine (Baird et al. 1999; Ishikawa-Mochizuki et al. 1999; Pasparakis et al. 2014)) plays a central role in the skin-homing process (Morales et al. 1999). *Ccl27* maps to human chromosome 9 in a tandem gene arrangement with two other chemokines, *Ccl19* and *Ccl21*, and presents two alternative transcripts, yielding secreted and intracellular forms (Fig. 1) (Baird et al. 1999; Gortz et al. 2002). The latter, designated PESKY, includes a different exon 1 and acts as an

Mónica Lopes-Marques and Luís Q. Alves contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00251-019-01114-z) contains supplementary material, which is available to authorized users.

✉ Mónica Lopes-Marques
monicaslm@hotmail.com

✉ Raquel Ruivo
rruivo@ciimar.up.pt

✉ L. Filipe C. Castro
filipe.castro@ciimar.up.pt

[1] CIIMAR-UP, Av. General Norton de Matos s/n, 4450-208 Matosinhos, Portugal

[2] Department of Biology, Faculty of Sciences, Rua do Campo Alegre 1021/1055, 4169-007 Porto, Portugal

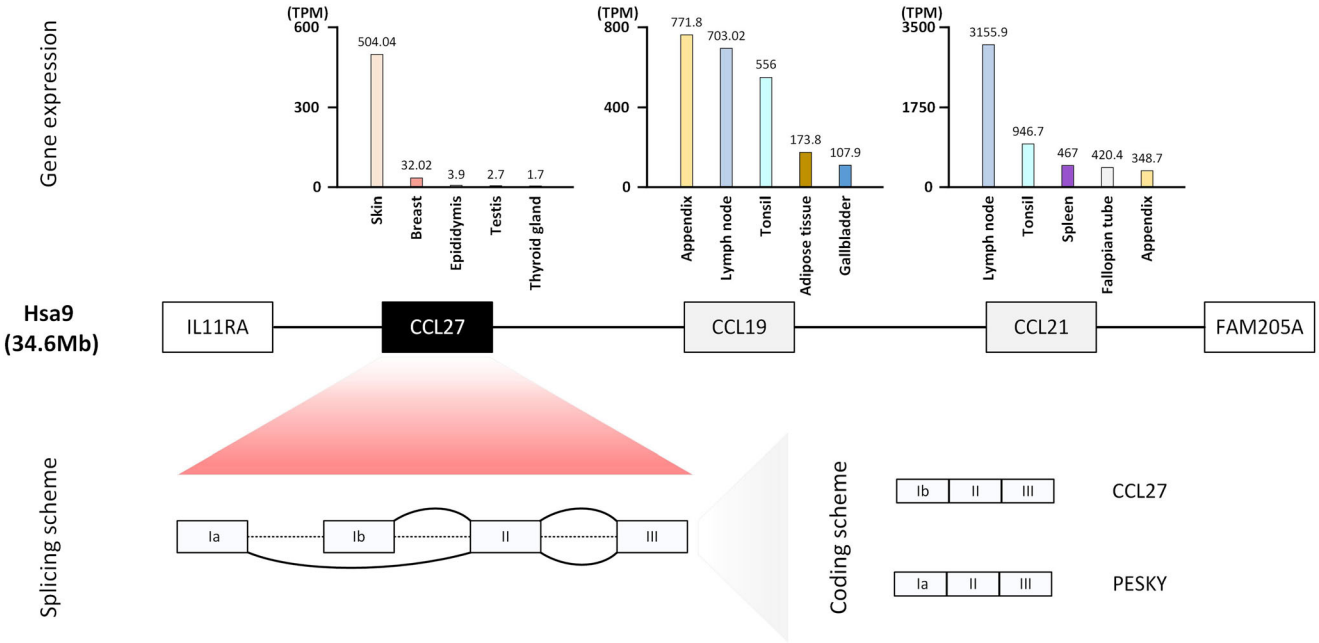


Fig. 1 The human orthologue *Ccl27* gene expression, genomic locus, and structure. In the center, the genomic region of *Homo sapiens* at chromosome 9, containing *Ccl27* gene (in black box) and tandem gene duplicates, *Ccl19* and *Ccl21* (gray boxes). The corresponding *Ccl* gene expression data was retrieved directly from the Human Protein Atlas (https://www.proteinatlas.org/). Only five tissues with the highest values of transcripts per million (TPM) are represented. Bottom figure represents *Ccl27* gene structure and alternative splicing producing two transcripts: CCL27 and PESKY

intracellular chemokine (Fig. 1) (Gortz et al. 2002). PESKY transcripts may be found in various mucosal tissues (Ledee et al. 2004; Nibbs and Graham 2003), but CCL27 secretion is mostly restricted to skin keratinocytes, having a critical role in skin homeostasis (Homey et al. 2002; Morales et al. 1999). To provide a skin-specific cue to attract memory T cells in normal or inflamed skin, CCL27 specifically binds to the CCR10 receptor in vivo (Homey et al. 2002; Morales et al. 1999). While CCL27 is exclusive towards the CCR10 receptor, other chemokines such as CCL8 also bind CCR10 (Homey et al. 2002; Morales et al. 1999; Nibbs and Graham 2003). Although a number of key morpho-functional skin components have been conserved throughout mammalian evolution, specific lineages experienced secondary episodes of phenotypic simplification or/and elaboration (e.g., Lopes-Marques et al. 2019, 2018; Sharma et al. 2018). In this context, Cetacea offer an illustrative example, with the exclusive aquatic dependence underscoring unique anatomical signatures (e.g., Lopes-Marques et al. 2019; McGowen et al. 2014; Sharma et al. 2018). For example, their skin is smooth with no pelage, presenting a thick stratum corneum, while the upper layers of the epidermis are not fully cornified (Mouton and Botha 2012; Sokolov 1982; Spearman 1972). Moreover, to improve smoothness and reduce drag, Cetacea skin is rapidly renewed (Hicks et al. 1985; Mouton and Botha 2012; Zabka and Romano 2003). These intensive cellular replacement and epidermal thickness reduce scab formation and the risk of pathogen invasion (Sokolov 1982; Spearman 1972; Zabka and Romano 2003). Accordingly, skin inflammation is apparently reduced in Cetacea (Zabka and Romano 2003). The underlying genomic events connected with skin repair mechanisms and whether other mammalian lineages display similar traits are presently unknown. A growing number of full genome sequences currently available have provided valuable insights into the role of gene loss as the foundation for phenotypic alterations and consequently on the perception of adaptive landscapes (Albalat and Canestro 2016; Hecker et al. 2019; Lachner et al. 2017; Lopes-Marques et al. 2019, 2018; Sadier et al. 2018; Sharma et al. 2018; Strasser et al. 2015). Given the key role of CCL27 in the process of skin inflammation, we hypothesized that the *Ccl27* coding sequence might be compromised in Cetacea as suggested by the overall skin inflammatory physiology observed in this lineage (Mouton and Botha 2012; Zabka and Romano 2003).

Methods

Sequence retrieval

Ccl27 coding nucleotide sequences were searched and collected from NCBI for a set of mammalian species representative of the major mammalian lineages (see Supplementary table 1). Searches were performed through tblastn and blastn queries using the human *Ccl27* sequence as reference. Full coding sequences and corresponding genomic sequences were

collected, for phylogenetic analysis and gene annotation respectively. Coding sequences were next uploaded into Geneious R7.1.9 curated by removing 5′ and 3′ UTR (untranslated regions) and aligned using the translation align option. Sequence alignment was inspected and exported for phylogenetic analysis. Maximum likelihood phylogenetic analysis was performed in PhyML3.0 server (Guindon et al. 2010), with best sequence evolutionary model determined using smart model selection (Lefort et al. 2017), and branch support with the aBayes algorithm (Anisimova et al. 2011). The resulting phylogenetic tree was then visualized in Figtree (Supplementary material 1).

Gene annotation

For gene annotation, the genomic sequence of *Ccl27* annotations tagged as LQ was collected from NCBI. For species with no *Ccl27* annotation (*Balaenoptera acutorostrata*, *Orcinus orca*, and *Rhinolophus sinicus*), the genomic sequence ranging from the upstream to the downstream flanking genes was collected. Finally, for species with no annotated genome (*Balaenoptera bonaerensis*, *Eschrichtius robustus*, *Balaena mysticetus*, *Hippopotamus amphibius*, and *Manis pentadactyla*), genomic sequences were recovered through tblastn searches in the whole genome assembly and scaffolds corresponding to the highest identity hits were taken. Collected genomic sequences were next loaded to Geneious R7.1.9 for manual annotation as previously described (Lopes-Marques et al. 2018, 2017). Briefly, using human and *Bos taurus Ccl27* coding sequence (CDS) as reference, each individualized exon was mapped on the corresponding genomic sequences using the built-in map to reference tool in Geneious R7.1.9. Aligned regions were manually inspected to verify coding status and identify ORF-disrupting mutations (frameshifts, premature stop codon, loss of canonical splice sites). The identified mutations were next validated in at least two independent SRA projects (when available) (see Supplementary material 2).

Transcriptomic analysis

RNA-Seq analysis was performed to assess the functional condition of *Ccl27* in 6 cetacean species and *H. amphibius*. For each of the 6 cetaceans, using the discontiguous megablast task from Blastn, the *B. taurus Ccl27* CDS was used as a query sequence to recover reads from the totality of the available transcriptomic Sequence Read Archive (SRA) projects available at NCBI. Supplementary table 2 provides an in-depth description of the explored NCBI SRA projects per species. In the case of *H. amphibius*, through megablast from Blastn, the CDS of the annotated gene in the same species was used as the query sequence and reads were recovered from the available *H. amphibius* skin transcriptome (accession number PRJNA507170). The collected mRNA reads were mapped against the corresponding annotated gene using the map to reference tool from Geneious R7.1.9. The aligned regions were manually curated, and poorly aligning reads manually removed. Next, reads were then classified as spliced reads (reads spanning over two different exons) and exon-intron reads (reads containing intronic sequence). Reads fully overlapping a single exon, exonic reads, were considered inconclusive for this analysis, given that it is infeasible to infer the nature of the corresponding transcript (spliced or unspliced).

Comparative homology modeling

Comparative homology modeling was performed for *O. orca* representative of Odontoceti and *B. mysticetus* representative of Mysticeti, and for *Aotus nancymaae* (Platyrrhini) representing a coding CDS with a short C-terminal. Predicted CDSs of *O. orca* and *B. mysticetus* were determined using the annotated exons, and premature stop codons identified in exon 2 were reverted to the residue observed in *B. taurus*, while mutations in exon 3 were left as observed. Corresponding protein sequences were then next submitted to the SWISS-MODEL (Benkert et al. 2011; Waterhouse et al. 2018) for homology modeling using the human CCL27 crystal structure as reference (2KUM) (Jansma et al. 2010). Resulting models were downloaded and analyzed in PyMOL V1.74 (Schrodinger 2010).

Results and discussion

To investigate the distribution and annotation tags of the *Ccl27* gene in mammals, we scrutinized a total of 114 selected mammalian genomes available at NCBI and Ensembl genome browsers (Supplementary table 1). This search retrieved 14 *Ccl27* annotations tagged as "low-quality" (LQ) and uncovered 9 species with no *Ccl27* gene annotation (Supplementary table 1). Next, we investigated the genomic sequences corresponding to the *Ccl27* LQ annotations to determine the CDS through manual annotation. This step revealed coding *Ccl27* genes, tagged as LQ for the following species: *Saimiri boliviensis* (black-capped squirrel monkey), *Galeopterus variegatus* (Sunda flying lemur), *Peromyscus maniculatus bairdii* (North American deer mouse), *Loxodonta Africana* (African bush elephant), and *Chrysochloris asiatica* (Cape golden mole). Also, the analysis of the genomic sequence corresponding to the *Ccl27* locus in *Ochotona princeps* (American pika) showed that the missing annotation in this species is most probably due to poor genome coverage in this locus (not shown). Importantly, all cetacean species analyzed presented sequences tagged as LQ or no *Ccl27* annotation. This impelled us to further explore other cetacean species with unannotated genomes: *Balaenoptera bonaerensis* (Antarctic

minke whale), *Eschrichtius robustus* (gray whale), *Balaena mysticetus* (bowhead whale), and *Sousa chinensis* (Indo-Pacific humpback dolphin).

Ccl27 gene sequence contains inactivating mutations in Cetacea

Annotation of collected cetacean genomic sequences revealed *Ccl27* gene erosion across all analyzed species (Fig. 2a). In detail, gene sequence examination in Odontoceti showed a non-disruptive insertion of a codon in exon 2 in all species with the exception of *Lipotes vexillifer* (Yangtze River dolphin). In addition, in exon 2, a frameshift mutation (deletion of 1 nucleotide) was identified and validated by Sequence Read Archive (SRA) analysis in *Physeter catodon* (sperm whale; Supplementary material 2). A conserved premature stop codon was found in *Orcinus orca* (orca) and *Lagenorhynchus obliquidens* (Pacific white-sided dolphin), as well as a non-conserved premature stop codon in *L. vexillifer*. These observations were confirmed in *O. orca* and *L. obliquidens* through SRA analysis (Supplementary material 2). In *L. vexillifer*, exon 2 also presented a frameshift mutation (4 nucleotide deletions) and the loss of the canonical splice site (GT>CC). Next, in exon 3, a conserved premature stop codon was identified in all Odontoceti (Fig. 2b, Supplementary material 2). A frameshift mutation by deletion was identified in all species apart from *P. catodon*, which in turn shows a frameshift mutation before the identified stop codon (Fig. 2b).

Regarding the Mysticeti, all identified mutations were conserved across all 4 analyzed species (Fig. 2a and c). Non-disruptive mutations consisting in the deletion and insertion of 1 codon were identified in exon 1 and exon 2, respectively. Also, two conserved premature stop codons were identified in exon 2 and exon 3 (the former was validated by SRA; Supplementary material 2), which were followed by a 1 nucleotide deletion identified in all analyzed species (Fig. 2c, black arrow). Interestingly, this 1 nucleotide deletion is conserved among all cetacean species (Supplementary material 3), suggesting that *Ccl27* pseudogenization preceded the divergence of Odonticeti and Mysticeti.

Transcriptomic analysis supports *Ccl27* gene erosion in Cetacea

To further scrutinize the functional condition of *Ccl27*, we next analyzed multi-tissue RNA-Seq projects available at NCBI for 6 cetacean species: *Tursiops truncatus* (common bottlenose dolphin), *Delphinapterus leucas* (beluga whale), *Neophocaena asiaeorientalis* (finless porpoise), *P. catodon* (sperm whale), *Balaenoptera acutorostrata* (common minke whale), and *B. mysticetus* (Supplementary table 2).

Overall, RNA-Seq analysis revealed a considerably low number of *Ccl27* mRNA reads across all the 6 species, especially in *N. asiaeorientalis* (Fig. 3). Moreover, for the remaining species, we observed a substantially high proportion of reads spanning adjacent exonic and intronic regions, exon-intron reads, versus spliced reads, connecting contiguous exons and containing no intronic remnants, especially in the case of *T. truncatus* (121 exon-intron reads against 3 spliced reads). In the latter, the higher number of skin-specific sequencing runs available for this species, compared with the remaining ones (25 skin-specific sequencing runs in *T. truncatus* vs. an average of 6.6 skin-specific sequencing runs per species), probably explains the variation in the number of exon-intron reads (see Supplementary table 2). As we observed a specific case with a considerably distinct ratio of exon-intron reads/spliced reads among the remaining species, namely *B. mysticetus* (49 spliced reads vs. 52 exon-intron reads), we decided to further verify the presence of ORF disruptive mutations in the produced transcripts of *Ccl27* in each of the referred species. We were able to detect at least one premature stop codon in the transcripts of the analyzed 6 cetacean species (see Supplementary material 4), revealing that *Ccl27* transcripts contained the genome predicted ORF mutations (Fig. 3).

The conserved mutational pattern observed between Odontoceti and Mysticeti suggests that *Ccl27* inactivation occurred in the Cetacea ancestor. To further survey and estimate the approximate timing of *Ccl27* loss in Cetacea, we next investigated the genome and the skin transcriptome of the extant sister clade of the Cetacea, the Hippopotamidae. The current version of the *H. amphibius* genome, available at NCBI, is fragmented and unannotated (GCA_002995585.1). However, we were able to deduce the full coding ORF of the *Ccl27* gene orthologue in *H. amphibius*, and without any intervening inactivating mutation (Fig. 3). Furthermore, by examining a skin-specific transcriptome, we identified a very high proportion of spliced/exon-intron mRNA reads (1995 spliced reads against 379 exon-intron reads), a clear indication that the gene is functional in this species (Fig. 3).

Ccl27 is eroded in other non-cetacean mammals

We next investigated the uniqueness of *Ccl27* inactivation in other mammalian lineages with absent or LQ annotations. Our initial analysis revealed that several genes annotated as LQ were in fact coding. For example, the analysis of the retrieved genomic sequence for *L. africana* revealed poor genome coverage in the *Ccl27* locus. However, blastn search of the whole genome sequence recovered a genomic scaffold (NW_003573426.1:65683000–65687000) that contained an intact *Ccl27* gene sequence. Yet, in the case of LQ tagged *Ccl27* from *Hipposideros armiger* (great roundleaf bat), *Trichechus manatus* (West Indian manatee), *Heterocephalus*

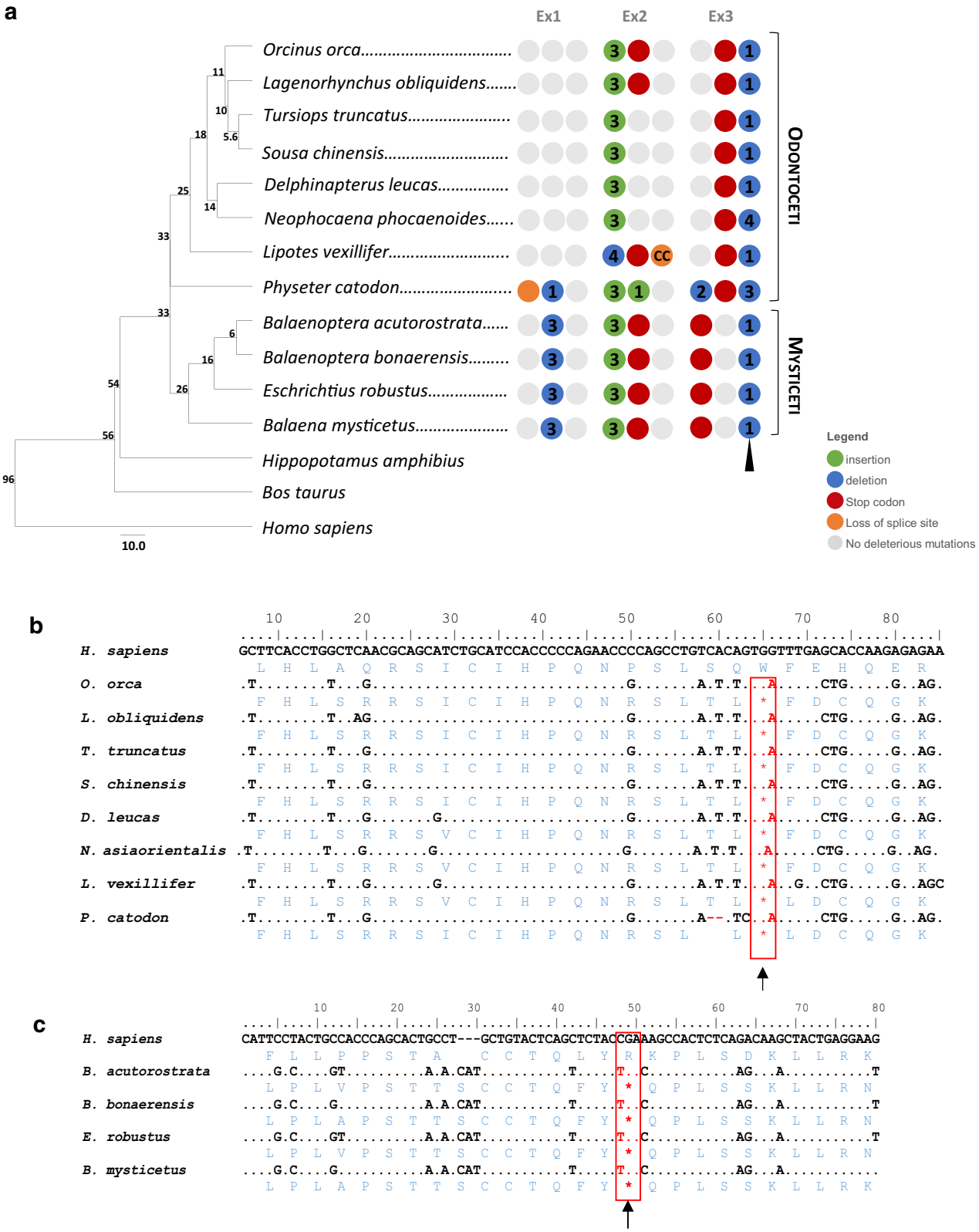


Fig. 2 Pseudogene annotation in Cetacea. **a** Schematic representation of the *Ccl27* gene and identified mutations in Cetacea, each group of 3 circles represents one exon, red represents stop codon; orange, non-AG-GT splice site; blue, deletion; and green, nucleotide insertion; numbers at tree nodes indicate million years. Numbers in the circles indicate number of nucleotides inserted or deleted and dark gray circles represent regions or exon not found. **b** Sequence alignment of the identified premature stop codon in exon 3 of Odontoceti. **c** Sequence alignment of the identified premature stop codon in exon 2 of Mysticeti

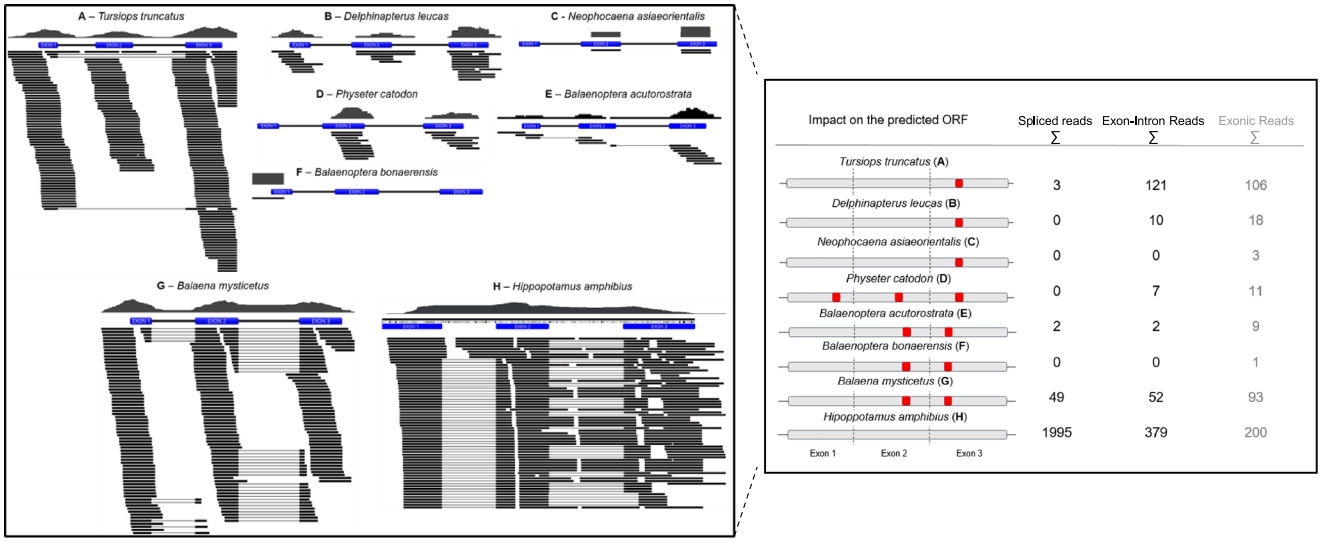


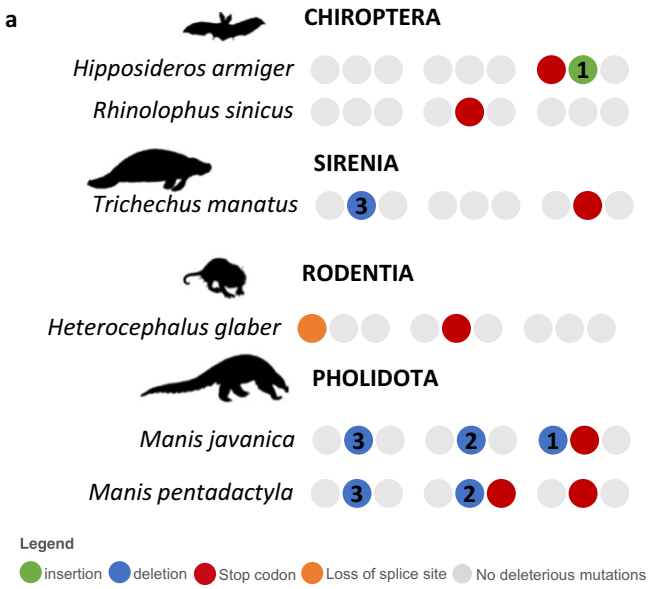
Fig. 3 Gene expression of *Ccl27* across Cetacea species. In the left box: mapping of the NCBI Sequence Read Archive (SRA) recovered multi-tissue RNA-Seq reads (black) for each of the 7 represented species against the corresponding *Ccl27* annotated gene (blue). Right box: impact of the annotated mutations in the open reading frame (ORF) of the *Ccl*27 gene. Premature stop codons are represented with a red squared marker at the corresponding exon. Overall count of RNA-Seq mapped reads for each species. Reads are classified into spliced reads (reads spanning over two different exons), exon-intron reads (reads containing exonic and intronic sequence), and exonic reads (reads fully overlapping exonic regions)

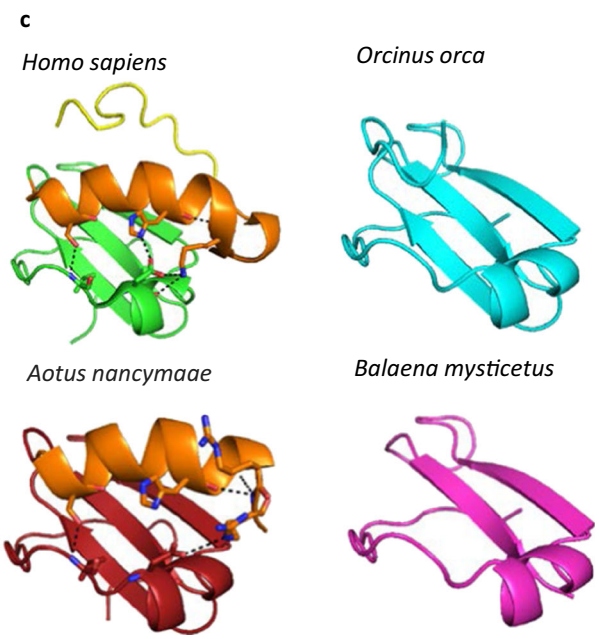
glaber (naked mole rat), and *Manis javanica* (Sunda pangolin), the analysis and manual annotation of the corresponding genomic sequence revealed a number of ORF-disrupting mutations (Fig. 4a). Our findings were further supported by searching the available unannotated genomes of *Manis pentadactyla* (Chinese pangolin) and *Rhinolophus sinicus* (Chinese rufous horseshoe bat), which after annotation also presented a non-coding *Ccl27* ORF (Fig. 4a). Briefly, *Ccl27* annotation in *H. armiger* revealed a premature stop codon in exon 3 followed by 1 nucleotide insertion, and in *R. sinicus*, a single premature stop codon was identified in exon 2 (all confirmed by SRA search; Supplementary material 5). In the Pholidota *M. javanica* and *M. pentadactyla*, a shared frameshift mutation in exon 2 was identified (validated by SRA in *M. javanica*; Supplementary material 5). Additionally, *M. pentadactyla* presents a premature stop codon in exon 2 while *M. javanica* presents a premature stop codon in exon 3 preceded by a 1 nucleotide frameshift mutation. In Rodentia, the *Ccl27*gene annotation in *H. glaber* revealed a missing start codon in exon 1 combined with a premature stop codon in exon 2, and fichrodingernally in *T. manatus*, *Ccl27* gene annotation uncovered a premature stop codon in exon 3 (stop codons validated by SRA; Supplementary material 5).

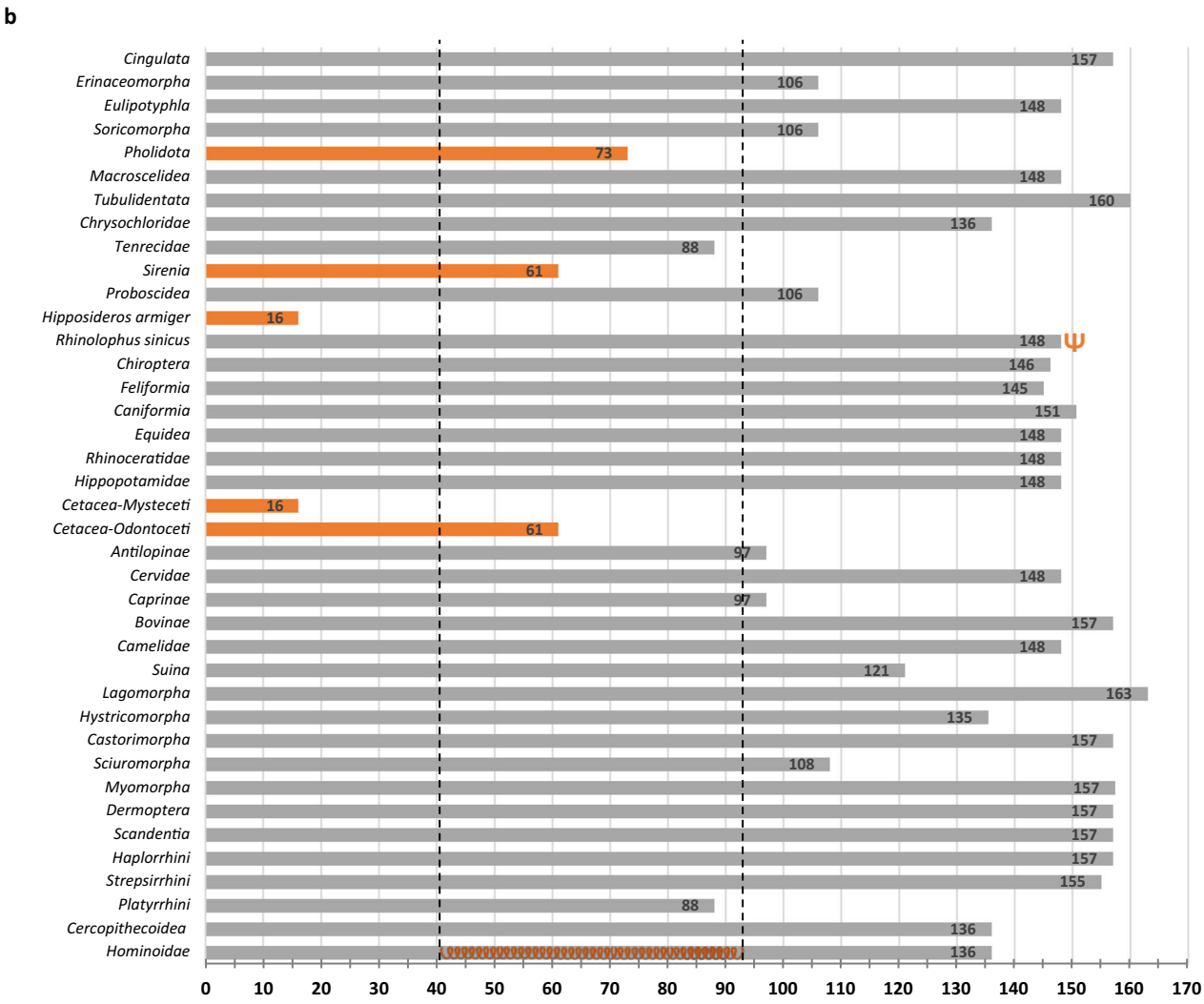
Exon 3 length reduction parallels gene inactivation

The survey of 112 placental mammalian *Ccl27* CDSs exposed a variable C-terminal length in different mammalian species. Thus, we compared the predicted length of exon 3 in the annotated pseudogenes regardless of prior ORF-disrupting mutations (Fig. 4b). This analysis showed that all annotated pseudogenes were severely truncated in exon 3 with the exception of *R. sinicus*. Also, the analysis of the observed truncations in the overall structure of CCL27 using homology modeling for *O. orca* and *B. mysticetus* showed that premature stop codons occur early in the C-terminal α-helix. CC chemokines present a highly conserved quaternary structure characterized by disordered N-terminal region followed by a 3_{10}-helix, 3 antiparallel β-strands, followed a C-terminal α-helix and ending with a disordered stretch of positive residues (Jansma et al. 2010). Interestingly, the C-terminal region, specifically the disordered region, is a feature that differentiates CCL27 from the majority of CC chemokines, and has been shown to be involved in nuclear import (Nibbs and Graham 2003). In agreement, both *Ccl27* transcript variants, the intracellular chemokine PESKY and the internalized CCR10-bound CCL27, target the cell nucleus, modulating morphology and motility via transcriptional modification (Hromas et al. 1999; Ledee et al. 2004). Moreover, the remaining mammals including Tenricidae, Antilopinae, Caprinae, and Platyrrhini exhibit a sequence deletion pattern at the end or shortly after the α-helix (Fig. 4b), which implies the loss of the final C-terminal disordered region involved in nuclear targeting. Yet,

Fig. 4 **a** Gene annotation of *Ccl7* in non-cetacean mammals. **b** Analysis of exon 3 length in nucleotides: orange bars highlights species with severe exon 3 truncation, orange helix in Hominoidae bar corresponds to extension C-terminal α-helix in human crystal structure (2KUM). **c** Comparative analysis of the human crystal structure 2KUM (green) and calculated homology models in red *Aotus nancymaae*, blue *Orcinus orca*, and magenta *Balaena mysticetus*. In humans, structural features highlighted in orange terminal α-helix, in yellow disordered terminal region


a
CHIROPTERA
Hipposideros armiger
Rhinolophus sinicus
SIRENIA
Trichechus manatus
RODENTIA
Heterocephalus glaber
PHOLIDOTA
Manis javanica
Manis pentadactyla
Legend
insertion
deletion
Stop codon
Loss of splice site
No deleterious mutations


c
Homo sapiens
Orcinus orca
Aotus nancymaae
Balaena mysticetus


b
Cingulata
Erinaceomorpha
Eulipotyphla
Soricomorpha
Pholidota
Macroscelidea
Tubulidentata
Chrysochloridae
Tenrecidae
Sirenia
Proboscidea
Hipposideros armiger
Rhinolophus sinicus
Chiroptera
Feliformia
Caniformia
Equidea
Rhinoceratidae
Hippopotamidae
Cetacea-Mysteceti
Cetacea-Odontoceti
Antilopinae
Cervidae
Caprinae
Bovinae
Camelidae
Suina
Lagomorpha
Hystricomorpha
Castorimorpha
Sciuromorpha
Myomorpha
Dermoptera
Scandentia
Haplorrhini
Strepsirrhini
Platyrrhini
Cercopithecoidea
Hominoidae
157
106
148
106
73
148
160
136
88
61
106
16
148
Ψ
146
145
151
148
148
148
16
61
97
148
97
157
148
121
163
135
157
108
157
157
157
157
155
88
136
136
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170

contrarily to the annotated pseudogenes, the coding *Ccl27* *Aotus nancymaae*, which presents the shortest exon 3, still conserves the full α-helix which has been reported to stabilize the overall fold (Jansma et al. 2010) (Fig. 4c). The biological significance of this plasticity remains to be studied.

Ccr10 coding status in species with *Ccl27* loss

Since convergent evolution of chemokines and respective receptor was previously detected (van der Loo et al. 2012), we next investigated the coding status *Ccl27* receptor *Ccr10* in all species presenting an eroded *Ccl27* gene. Gene annotation was conducted similarly to that previously described for *Ccl27*. Initial search at NCBI revealed poor annotation of CCR10 exon 1 in several species such as *L. obliquidens*, *D. leucas*, *B. acutorostrata*, and *H. glaber* (Supplementary material 6A), while other species with annotated genomes presented no annotation of CCR10 in current genome release, namely *T. truncatus* (Supplementary material 6A). Yet, manual annotation of *Ccr10* revealed that all species with *Ccl27* gene disruption presented no ORF-disrupting mutations in the corresponding receptor gene *Ccr10* (see Supplementary material 6B and 6C). Thus, the loss of *Ccl27* in the analyzed species is unrelated to the coding status of the corresponding receptor *Ccr10*.

Ccl27 gene loss correlates with alternative protection and healing programs

Previous reports have shown considerable differences in chemokine gene repertoire in different mammalian lineages (Nomiyama et al. 2010; Shibata et al. 2013). While some chemokine gene families expanded due to duplication events, others contracted as a consequence of gene loss (Nomiyama et al. 2010; Shibata et al. 2013). For example, chemokine *Ccl16* pseudogenization in leporids was attributed to a random effect (Neves et al. 2015; Neves et al. 2019; Shibata et al. 2013). Here, our analysis indicates that *Ccl27* is most likely non-functional in all of the examined cetacean species. Inactivating mutations are also present in species of Pholidota, Sirenia, Chiroptera, and Rodentia. Even if a full phenotypic description of mouse knockout (KO) for this gene is presently unavailable, the initial data suggests a decrease of the T cell population in intact skin (Davila et al. 2016). On the other hand, constitutive production of keratinocyte CCL27 enhanced the inflammatory response in mice (Kagami et al. 2008). In agreement, chronic inflammatory skin diseases, such as atopic dermatitis and psoriasis, are characterized by increased serum levels of the T cell–attracting chemokine CCL27 (Homey et al. 2002), while CCL27-neutralizing antibody treatment reduced skin inflammation in a transgenic animal model (Chen et al. 2006). Thus, upon insult or infection, *Ccl27* KO would likely show an attenuated inflammatory response in the skin. This hypothesis remains to be verified.

Nevertheless, it could be argued that the premature stop codon in exon 3 of *Ccl27* in *T. truncatus*, *D. leucas*, *S. chinensis*, and *N. asiaeorientalis* could still encode a functional shorter isoform. Yet, RNA-Seq transcriptome and structural analysis supports a different interpretation. Since *Ccl27* prime expression site is the skin, we analyzed the available skin and multi-tissue RNA transcriptomes from Cetacea and found two distinct scenarios. First, in the majority of the species, RNA-Seq searches recovered reads covering exon-intron. Second, in *B. mysticetus*, the recovered RNA-Seq reads presented a higher number of spliced reads. However, in both cases, a detailed analysis of the collected reads confirmed the presence of the previously identified ORF-disrupting mutations. Thus, we suggest that these mRNA mature sequences do not translate into a functional protein. In addition, previous studies, addressing the loss of visual opsins in Chiroptera, highlighted possible discrepancies between gene integrity and protein production, further suggesting post-transcriptional mechanisms as regulators of evolutionary gene silencing (Sadier et al. 2018).

Our analysis strongly supports that *Ccl27* gene pseudogenization compromises both canonical CCL27 and PESKY transcripts. This is in accordance with previous findings reporting a distinct inflammatory and wound healing program in cetacean skin (Zabka and Romano 2003; Zasloff 2011). Interestingly, scarless and low inflammation wound repair have also been reported in several mammalian fetuses, including human, as well as in human adult oral mucosa (Iglesias-Bartolome et al. 2018; Moore et al. 2018). Both observations might correlate with decreased or null CCL27 secretion. In fact, embryonic keratinocytes are more proliferative and less immunogenic than adult cells, inhibiting T cell proliferation (Tan et al. 2014). Similarly, oral mucosa exhibits rapid wound healing due to accelerated re-epithelialization (Iglesias-Bartolome et al. 2018). In wounded oral mucosa, the overall expression of *Ccl27* is also downregulated when compared with wounded skin (Iglesias-Bartolome et al. 2018). Thus, in scarless and low inflammation wound repair, increased epithelial renewal seems to parallel the downregulation or absence of CCL27 secretion. Yet, in oral mucosa, the possible maintenance of PESKY could participate in the healing circuitry by stimulating cell migration and proliferation.

Additionally, we found convergent inactivation of *Ccl27* in other non-cetacean mammalian species, namely in pangolins (*M. javanica* and *M. pentadactyla*), in the naked mole rat (*H. glaber*), in the sirenian *T. manatus*, and in two Chiroptera (*H. armiger* and *R. sinicus*). Curiously, with the exception of Chiroptera, these species share some of the distinctive features of Cetacea skin: for example, the hairless phenotype is observed in Pholidota, Sirenia, and naked mole

rat; increased epidermal thickness is observed in Sirenia and naked mole rat; and Sirenia skin is also smooth (Daly and Buffenstein 1998; McGowen et al. 2014). The diversity of skin phenotypes along with the scarce information regarding species-specific inflammatory and wound healing programs hampers the anticipation of the possible outcomes of *Ccl27* pseudogenization. Nonetheless, the available information suggests that *Ccl27* erosion occurred in species exhibiting singular epidermal renewal, or even protective mechanisms or structures, reducing the need for CCL27-dependent inflammatory processes. For instance, pangolins present a protective armor with keratin-derived scales, which was suggested to reduce epithelial immune requirements (Choo et al. 2016; Meyer et al. 2013). In agreement, pseudogenization of interferon epsilon, which confers protection against viral and bacterial infections, was also reported in these species (Choo et al. 2016). On the other hand, the naked mole rat abundantly produces high molecular weight hyaluronic acid, suggested to underscore their peculiar longevity and cancer resistance and also contributing to cell motility, rapid wound healing, and immunity (Daly and Buffenstein 1998; Fisher 2015). Regarding Chiroptera, although their skin is generally similar to most mammalian species, interdigital skin membranes are thinner, and thus more susceptible to damage; yet, interdigital membranes have an enhanced healing capacity (Ceballos-Vasquez et al. 2015). Nonetheless, the inflammatory circuitry of this healing process is still poorly studied. Also, *Ccl27* pseudogenization was only detected in two Chiroptera species. Again, post-translational events could promote CCL27 loss in additional species (Sadier et al. 2018). In conclusion, our findings reinforce gene loss mechanisms as evolutionary drivers of skin phenotypes in mammals and correlate *Ccl27* loss with species-specific scarless and/or low inflammation wound repair.

Funding information This work was supported by Project No. 031342 co-financed by COMPETE 2020, Portugal 2020, and the European Union through the ERDF, and by FCT through national funds.

References

Albalat R, Canestro C (2016) Evolution by gene loss. Nat Rev Genet 17: 379–391. https://doi.org/10.1038/nrg.2016.39

Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol 60:685–699. https://doi.org/10.1093/sysbio/syr041

Baird JW, Nibbs RJB, Komai-Koma M, Connolly JA, Ottersbach K, Clark-Lewis I, Liew FY, Graham GJ (1999) ESkine, a novel β-chemokine, is differentially spliced to produce secretable and nuclear targeted isoforms. J Biol Chem 274:33496–33503. https://doi.org/10.1074/jbc.274.47.33496

Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343–350. https://doi.org/10.1093/bioinformatics/btq662

Ceballos-Vasquez A, Caldwell JR, Faure PA (2015) Seasonal and reproductive effects on wound healing in the flight membranes of captive big brown bats. Biol Open 4:95–103. https://doi.org/10.1242/bio.201410264

Chen L, Lin SX, Agha-Majzoub R, Overbergh L, Mathieu C, Chan LS (2006) CCL27 is a critical factor for the development of atopic dermatitis in the keratin-14 IL-4 transgenic mouse model. Int Immunol 18:1233–1242. https://doi.org/10.1093/intimm/dxl054

Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S, Tamazian G, Antunes A, Wilson RK, Warren WC, Koepfli KP, Minx P, Krasheninnikova K, Kotze A, Dalton DL, Vermaak E, Paterson IC, Dobrynin P, Sitam FT, Rovie-Ryan JJ, Johnson WE, Yusoff AM, Luo SJ, Karuppannan KV, Fang G, Zheng D, Gerstein MB, Lipovich L, O'Brien SJ, Wong GJ (2016) Pangolin genomes and the evolution of mammalian scales and immunity. Genome Res 26:1312–1322. https://doi.org/10.1101/gr.203521.115

Daly TJ, Buffenstein R (1998) Skin morphology and its role in thermoregulation in mole-rats, Heterocephalus glaber and Cryptomys hottentotus. J Anat 193(Pt 4):495–502

Davila ML, Fu Y, Yang J, Xiong N (2016) Role of CCR10 and CCL27 in skin resident T cell development and homeostasis. J Immunol 196: 137.137

Di Meglio P, Perera GK, Nestle FO (2011) The multitasking organ: recent insights into skin immune function. Immunity 35:857–869. https://doi.org/10.1016/j.immuni.2011.12.003

Fisher GJ (2015) Cancer resistance, high molecular weight hyaluronic acid, and longevity. J Cell Commun Signal 9:91–92. https://doi.org/10.1007/s12079-015-0278-6

Gortz A, Nibbs RJB, McLean P, Jarmin D, Lambie W, Baird JW, Graham GJ (2002) The chemokine ESkine/CCL27 displays novel modes of intracrine and paracrine function. J Immunol 169:1387–1394. https://doi.org/10.4049/jimmunol.169.3.1387

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010

Hecker N, Sharma V, Hiller M (2019) Convergent gene losses illuminate metabolic and physiological changes in herbivores and carnivores. Proc Natl Acad Sci 116:3036–3041. https://doi.org/10.1073/pnas.1818504116

Hicks BD, St Aubin DJ, Geraci JR, Brown WR (1985) Epidermal growth in the bottlenose dolphin, Tursiops truncatus. J Invest Dermatol 85: 60–63

Homey B, Alenius H, Müller A, Soto H, Bowman EP, Yuan W, McEvoy L, Lauerma AI, Assmann T, Bünemann E, Lehto M, Wolff H, Yen D, Marxhausen H, To W, Sedgwick J, Ruzicka T, Lehmann P, Zlotnik A (2002) CCL27-CCR10 interactions regulate T cell-mediated skin inflammation. Nat Med 8:157–165. https://doi.org/10.1038/nm0202-157

Hromas R, Broxmeyer HE, Kim C, Christopherson K 2nd, Hou YH (1999) Isolation of ALP, a novel divergent murine CC chemokine with a unique carboxy terminal extension. Biochem Biophys Res Commun 258:737–740. https://doi.org/10.1006/bbrc.1999.0507

Iglesias-Bartolome R, Uchiyama A, Molinolo AA, Abusleme L, Brooks SR, Callejas-Valera JL, Edwards D, Doci C, Asselin-Labat ML, Onaitis MW, Moutsopoulos NM, Silvio Gutkind J, Morasso MI (2018) Transcriptional signature primes human oral mucosa for rapid wound healing. Sci Transl Med 10:eaap8798. https://doi.org/10.1126/scitranslmed.aap8798

Ishikawa-Mochizuki I, Kitaura M, Baba M, Nakayama T, Izawa D, Imai T, Yamada H, Hieshima K, Suzuki R, Nomiyama H, Yoshie O (1999) Molecular cloning of a novel CC chemokine, interleukin-11 receptor alpha-locus chemokine (ILC), which is located on

chromosome 9p13 and a potential homologue of a CC chemokine encoded by molluscum contagiosum virus. FEBS Lett 460:544–548

Jansma AL, Kirkpatrick JP, Hsu AR, Handel TM, Nietlispach D (2010) NMR analysis of the structure, dynamics, and unique oligomerization properties of the chemokine CCL27. J Biol Chem 285:14424–14437. https://doi.org/10.1074/jbc.M109.091108

Kagami S, Saeki H, Tsunemi Y, Nakamura K, Kuwano Y, Komine M, Nakayama T, Yoshie O, Tamaki K (2008) CCL27-transgenic mice show enhanced contact hypersensitivity to Th2, but not Th1 stimuli. Eur J Immunol 38:647–657. https://doi.org/10.1002/eji.200737685

Lachner J, Mlitz V, Tschachler E, Eckhart L (2017) Epidermal cornification is preceded by the expression of a keratinocyte-specific set of pyroptosis-related genes. Sci Rep 7:17446. https://doi.org/10.1038/s41598-017-17782-4

Ledee DR, Chen J, Tonelli LH, Takase H, Gery I, Zelenka PS (2004) Differential expression of splice variants of chemokine CCL27 mRNA in lens, cornea, and retina of the normal mouse eye. Mol Vis 10:663–667

Lefort V, Longueville J-E, Gascuel O (2017) SMS: Smart Model Selection in PhyML. Mol Biol Evol 34:2422–2424. https://doi.org/10.1093/molbev/msx149

Lopes-Marques M, Ruivo R, Fonseca E, Teixeira A, Castro LFC (2017) Unusual loss of chymosin in mammalian lineages parallels neo-natal immune transfer strategies. Mol Phylogenet Evol 116:78–86. https://doi.org/10.1016/j.ympev.2017.08.014

Lopes-Marques M, Machado AM, Barbosa S, Fonseca MM, Ruivo R, Castro LFC (2018) Cetacea are natural knockouts for IL20. Immunogenetics 70:681–687. https://doi.org/10.1007/s00251-018-1071-5

Lopes-Marques M, Machado AM, Alves LQ, Fonseca MM, Barbosa S, Mikkel-Holger S Sinding, Rasmussen HM, Iversen MR, Bertelsen MF, Campos PF, da Fonseca R, Ruivo R, Castro LF (2019) Complete Inactivation of Sebum-Producing Genes Parallels the Loss of Sebaceous Glands in Cetacea. Mol Biol Evol msz068. https://doi.org/10.1093/molbev/msz068

McGowen MR, Gatesy J, Wildman DE (2014) Molecular evolution tracks macroevolutionary transitions in Cetacea. Trends Ecol Evol 29:336–346. https://doi.org/10.1016/j.tree.2014.04.001

Meyer W, Liumsiricharoen M, Suprasert A, Fleischer LG, Hewicker-Trautwein M (2013) Immunohistochemical demonstration of keratins in the epidermal layers of the Malayan pangolin (Manis javanica), with remarks on the evolution of the integumental scale armour. Eur J Histochem 57:e27–e27. https://doi.org/10.4081/ejh.2013.e27

Moore AL, Marshall CD, Barnes LA, Murphy MP, Ransom RC, Longaker MT (2018) Scarless wound healing: transitioning from fetal research to regenerative healing. Wiley Interdiscip Rev Dev Biol 7. https://doi.org/10.1002/wdev.309

Morales J, Homey B, Vicari AP, Hudak S, Oldham E, Hedrick J, Orozco R, Copeland NG, Jenkins NA, McEvoy LM, Zlotnik A (1999) CTACK, a skin-associated chemokine that preferentially attracts skin-homing memory T cells. Proc Natl Acad Sci 96:14470–14475. https://doi.org/10.1073/pnas.96.25.14470

Mouton M, Botha A (2012) Cutaneous lesions in cetaceans: an indicator of ecosystem status? https://doi.org/10.5772/54432

Neves F, Abrantes J, Lissovsky AA, Esteves PJ (2015) Pseudogenization of CCL14 in the Ochotonidae (pika) family. Innate Immun 21:647–654. https://doi.org/10.1177/1753425915577455

Neves F, Abrantes J, Lopes AM, Fusinatto LA, Magalhães MJ, van der Loo W, Esteves PJ (2019) Evolution of CCL16 in Glires (Rodentia and Lagomorpha) shows an unusual random pseudogenization pattern. BMC Evol Biol 19:59–59. https://doi.org/10.1186/s12862-019-1390-7

Nibbs RJ, Graham GJ (2003) CCL27/PESKY: a novel paradigm for chemokine function. Expert Opin Biol Ther 3:15–22. https://doi.org/10.1517/14712598.3.1.15

Nomiyama H, Osada N, Yoshie O (2010) The evolution of mammalian chemokine genes. Cytokine Growth Factor Rev 21:253–262. https://doi.org/10.1016/j.cytogfr.2010.03.004

Pasparakis M, Haase I, Nestle FO (2014) Mechanisms regulating skin immunity and inflammation. Nat Rev Immunol 14:289–301. https://doi.org/10.1038/nri3646

Sadier A et al (2018) Multifactorial processes underlie parallel opsin loss in neotropical bats. eLife 7:e37412. https://doi.org/10.7554/eLife.37412

Schrodinger L (2010) The PyMOL Molecular Graphics System, version 1.7.4 Schrödinger, LLC

Sharma V, Hecker N, Roscito JG, Foerster L, Langer BE, Hiller M (2018) A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. Nat Commun 9:1215. https://doi.org/10.1038/s41467-018-03667-1

Shibata K, Nomiyama H, Yoshie O, Tanase S (2013) Genome diversification mechanism of rodent and Lagomorpha chemokine genes. Biomed Res Int 2013:9. https://doi.org/10.1155/2013/856265

Sokolov V (1982) Comparative morphology of skin of different orders: Ordo Cetacea. In: In Mammal skin. University of California Press Ltd, Berkeley, pp 284–324

Spearman RI (1972) The epidermal stratum corneum of the whale. J Anat 113:373–381

Strasser B, Mlitz V, Fischer H, Tschachler E, Eckhart L (2015) Comparative genomics reveals conservation of filaggrin and loss of caspase-14 in dolphins. Exp Dermatol 24:365–369. https://doi.org/10.1111/exd.12681

Tan KKB, Salgado G, Connolly John E, Chan Jerry KY, Lane EB (2014) Characterization of fetal keratinocytes, showing enhanced stem cell-like properties: a potential source of cells for skin reconstruction. Stem Cell Rep 3:324–338. https://doi.org/10.1016/j.stemcr.2014.06.005

van der Loo W, Afonso S, de Matos AL, Abrantes J, Esteves PJ (2012) Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus Oryctolagus), but not in species of cottontail rabbit (Sylvilagus) and hare (Lepus). BMC Genet 13:72. https://doi.org/10.1186/1471-2156-13-72

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–w303. https://doi.org/10.1093/nar/gky427

Zabka TS, Romano TA (2003) Distribution of MHC II (+) cells in skin of the Atlantic bottlenose dolphin (Tursiops truncatus): an initial investigation of dolphin dendritic cells. Anat Rec A Discov Mol Cell Evol Biol 273A:636–647. https://doi.org/10.1002/ar.a.10077

Zasloff M (2011) Observations on the remarkable (and mysterious) wound-healing process of the bottlenose dolphin. J Investig Dermatol 131:2503–2505. https://doi.org/10.1038/jid.2011.220