

The organization of the keratin I and II gene clusters in placental mammals and marsupials show a striking similarity

Alexander Zimek, Klaus Weber*

Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

Received 29 August 2005; accepted 4 October 2005

Abstract

The genomic database for a marsupial, the opossum *Monodelphis domestica*, is highly advanced. This allowed a complete analysis of the keratin I and keratin II gene cluster with some 30 genes in each cluster as well as a comparison with the human keratin clusters. Human and marsupial keratin gene clusters have an astonishingly similar organization. As placental mammals and marsupials are sister groups a corresponding organization is also expected for the archetype mammal. Since hair is a mammalian acquisition the following features of the cluster refer to its origin. In both clusters hair keratin genes arose at an interior position. While we do not know from which epithelial keratin genes the first hair keratins type-I and -II genes evolved, subsequent gene duplications gave rise to a subdomain of the clusters with many neighboring hair keratin genes. A second subdomain accounts in both clusters for 4 neighboring genes encoding the keratins of the inner root sheath (irs) keratins. Finally the hair keratin gene subdomain in the type-I gene cluster is interrupted after the second gene by a region encoding numerous genes for the high/ultrahigh sulfur hair keratin-associated proteins (KAPs). We also propose a tentative synteny relation of opossum and human genes based on maximal sequence conservation of the encoded keratins. The keratin gene clusters of the opossum seem to lack pseudogenes and display a slightly increased number of genes. Opossum keratin genes are usually longer than their human counterparts and also show longer intergenic distances.

© 2005 Elsevier GmbH. All rights reserved.

Keywords: Cytoplasmic keratins; Hair keratins; Keratin gene clusters; Marsupials; *Monodelphis domestica*; Opossum; Placental mammals

Introduction

In mammals such as man, mouse and rat the family of genes encoding the structural proteins of the intermediate filaments (IFs) has nearly 70 members and reflects one of the 100 largest multigene families (Hesse et al., 2001, 2004). Sequence identity levels of IF proteins, their

biochemical properties, the organization of the corresponding genes and their expression patterns define several IF subtypes (Coulombe et al., 2001; Fuchs and Weber, 1994; Herrmann and Aebi, 2000). Type-I and -II keratins are by far the largest subfamilies. They give rise to the epithelial keratin filaments that are based on obligate heteromeric double-stranded coiled coils formed by a type-I and a type-II keratin. Mutations in at least 14 human epidermal keratin genes cause various skin fragility syndromes demonstrating that IFs protect cells against mechanical stress (Irvine and McLean, 1999).

*Corresponding author. Tel.: +49 551 201 1486;
fax: +49 551 201 1578.

E-mail address: office.weber@mpibpc.gwdg.de (K. Weber).

Keratins I and II are restricted to the chordate branch of metazoa (Erber et al., 1998; Zimek and Weber, 2002). While the urochordate *Ciona* seems to have only 1 type-I and 1 type-II gene (Karabinos et al., 2004; Wang et al., 2002) in the human genome this number increases to 26 and 27 (Hesse et al., 2001, 2004; Rogers et al., 2004, 2005). All human type-II genes locate to a tight cluster on chromosome 12q13, that ends with the type-I gene K18. Similarly all human type-I genes, except for K18, form a cluster on chromosome 17q21. While all keratin genes in the type-I cluster have the same orientation a few genes in the type-II cluster have opposite orientation. Highly related keratin gene clusters with nearly perfect synteny are also established in the rodents mouse and rat (Hesse et al., 2004). Two keratin gene clusters are a feature of all terrestrial animals, since they were recently established in the draft genomes of a bird (chicken) and an amphibian (*Xenopus tropicalis*) (Zimek and Weber, 2005). Because hair, with its multiple hair keratins and inner root sheath (irs) keratins, is a mammalian specialization the number of genes in the keratin clusters of chicken and *Xenopus* is reduced to around 50% compared to mammals. In contrast to the well-defined 2 keratin gene clusters 3 teleost fish indicate an excess of type-I over type-II genes, the lack of a type-II cluster and a dispersal of type-I genes. Nevertheless also in teleost fish the genes for keratins 8 and 18, which are the first expression pair in embryogenesis, are adjacent (Zimek and Weber, 2005). Within the current evolutionary framework the lungfish are the closest extant relatives of the terrestrial vertebrates (Schaffeld et al., 2005). Thus keratin gene clusters may also exist in lungfish and even in the lamprey, the most primitive vertebrate.

Mammalian taxons are usually grouped into 3 subclasses (Arnason et al., 2002; De Leo et al., 2005). The largest subclass is formed by the placentalia (placental mammals) (eutheria), which have a sister group relation with the marsupials (metatheria). A clearly older lineage is given by the egg-laying monotremata (protheria), which today are essentially restricted to Australia. Currently keratin gene clusters are available not only for man, mouse and rat (for references see above) but also for the chimpanzee and the dog (<http://genome.ucsc.edu>). In contrast to this wealth of information for placental mammals, nothing was published on the sister group, the marsupials. Thus we were surprised how much genomic information is available in the database for the opossum *Monodelphis domestica* (<http://www.broad.mit.edu/mammals/opossum>). Since large scaffolds were available we analyzed the complete keratin type-I and -II gene clusters of the opossum and compared them with their human counterparts. Because the clusters have the same subdomains a similar organization is also expected for the archetypic mammal. Hair keratin genes arose by multiple duplications and sequence drift as neighboring genes within the

frame of genes encoding cytoplasmic keratins. Interestingly the opossum seems to lack keratin pseudogenes and its keratin genes are generally larger than the corresponding human genes and are also separated by larger intergenic distances.

Materials and methods

We analyzed the draft genome assembly of the opossum (*M. domestica*) and compared it with the human keratin gene clusters on chromosome 17q21.2 and 12q13.13. The opossum genome assembly (MonDom1) was produced in October 2004 by the Broad Institute (Cambridge, MA). It has a 6.5-fold coverage across 3.56 Bbps assembled in 19 358 scaffolds. Since May 2005 a second release of the genome is available (http://www.broad.mit.edu/ftp/pub/assemblies/mammals/monDom/V3.0_prel.agp.chromosome.fasta.gz). In this release the number of scaffolds is reduced to 5677. Automated gene predictions have not yet been performed on the second assembly which only contains the raw sequences of the scaffolds. Chromosomal anchorage is not yet available for the opossum. Therefore we did our initial analysis with the MonDom1 assembly and used the graphical BLAT front-end (genome.ucsc.edu/cgi-bin/hgBlat) for the identification of keratin genes.

The keratin sequences predicted for the opossum were compared with sequences of human type-I and -II keratins present in a local database. A web-based front-end (ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/wwwblast-2.2.11.ia32-linux.tar.gz) allowed easy access to the local BLAST database and was used for the identification of opossum keratin homologs. DNA sequences from the opossum were checked for minimal deletions and insertions to keep the open reading frame (see below). Pairwise BLAST was used to calculate the percent identity value of human and opossum keratins on the amino-acid level. In addition, identity values were also calculated for the sequence of the rod domain of the keratins. In case of pseudogenes percent identity of the DNA sequences is given.

We also screened the genomes of the chimpanzee (*Pan troglodytes*) and the dog (*Canis familiaris*). These databases are also available at <http://genome.ucsc.edu>. The keratin type-I gene cluster of the chimp is on chromosome 19, the type-II cluster is located on chromosome 10. In the dog genome assembly (July 2004) the keratin I and II clusters are on chromosomes 9 and 27, respectively.

Results

We analyzed the initial MonDom1 release (October 2004) of the opossum (*M. domestica*), assembled by the

Broad Institute (Cambridge, MA). Gene identification is based on automated gene prediction and the occasional manual correction (Fig. 1). We identified 66 opossum keratin genes. Following the terminology of mammalian keratins (Hesse et al., 2004), they are abbreviated as Ka and Kb for type-I and -II keratins, respectively, followed by a number.

Our analysis is still limited for several cases. Thus 3 genes (Ka87, Kb87 and Kb94) still contain sequence gaps in the genomic information, and for 7 opossum genes (Ka13, Ka32, Ka36, Ka99, Kb71, Kb91 and Kb96) we did not identify the region encoding the carboxyterminal tail domain which usually is difficult to resolve without the help of a corresponding cDNA. Finally minor manipulations of the genomic sequences were necessary to keep the reading frame. This involved a 1-nt deletion for 7 genes (Ka16, Ka33, Ka36, Kb7, Kb25, Kb26, Kb90) and a 1-nt insertion for 2 genes (Ka28, Ka72). For 4 genes (Ka93, Kb71, Kb98, Kb99) 2 deletions of 1 nt each were necessary, and 1 gene (Kb97) required 3 times a 1-nt deletion. Most of these cases will be resolved once results from a large-scale cDNA sequence project and EST-sequence project on the opossum become available. Currently, there is only a partial cDNA sequence for Ka18 in the database (AY236876).

We built a local database of human and mice keratin protein sequences. Comparison of the predicted opossum sequences with this database resulted in pairs of proteins. Keratin pairs were considered to be homologs when the rod sequences showed an identity of more than 70% and the corresponding genes had the same orientation and relative position in the cluster. In these cases the human keratin names were used for the opossum genes. Otherwise we used our file names (numbers between 71 and 99). Examples for unresolved synteny are described below.

Keratin type-I cluster of the opossum

The opossum keratin type-I cluster is located on scaffolds 12767 and 14960, covering 1.1 and 0.55 Mbps, respectively (Fig. 1). The 2 parts of the keratin cluster are, as already known from other mammals, separated by a region encoding multiple high- and ultrahigh sulfur keratin-associated protein (KAPs) genes (Hesse et al., 2004; Rogers et al., 2001, 2004). We found about 20 KAP genes at one end of the opossum scaffold 12767. The remaining KAP genes are not present in the MonDom1 release. However, analysis of a new release (MonDom2) of the opossum genome (May 2005) showed the complete type-I keratin cluster on the new scaffold 18, which has a size of 4.36 Mbps.

Fig. 1 presents the summary of the opossum keratin type-I gene cluster and illustrates the proposed synteny

(black lines) for the marsupial and placental mammalian keratin genes. Unresolved is the synteny of the opossum genes Ka92, Ka93 and Ka94 which all resemble the Ka10 gene of man, mouse and rat. The opossum genes Ka77, Ka78 and Ka98 may be new genes without an obvious counterpart in the genes of the placental mammals. Finally we note that Ka22, the last gene of the opossum cluster, is clearly well related with the corresponding gene of the mouse and rat, while in the human genome it has drifted to a pseudogene (Hesse et al., 2004; Rogers et al., 2004).

As in placental mammalian genomes the hair keratin genes are inserted as a group into the genes encoding cytoplasmic keratin I proteins. We have verified this assignment by alignment of the corresponding protein sequences, which show the high cysteine content of the terminal head and tail domains also in the opossum hair keratins. While we tentatively identify 11 hair keratin type-I genes in the opossum the human genome shows 11 genes and 1 pseudogene (Fig. 1, Table 1). Similarly clearly defined are the 4 neighboring type-I genes of the *irs* which occur in the same position in the type-I cluster of man and opossum.

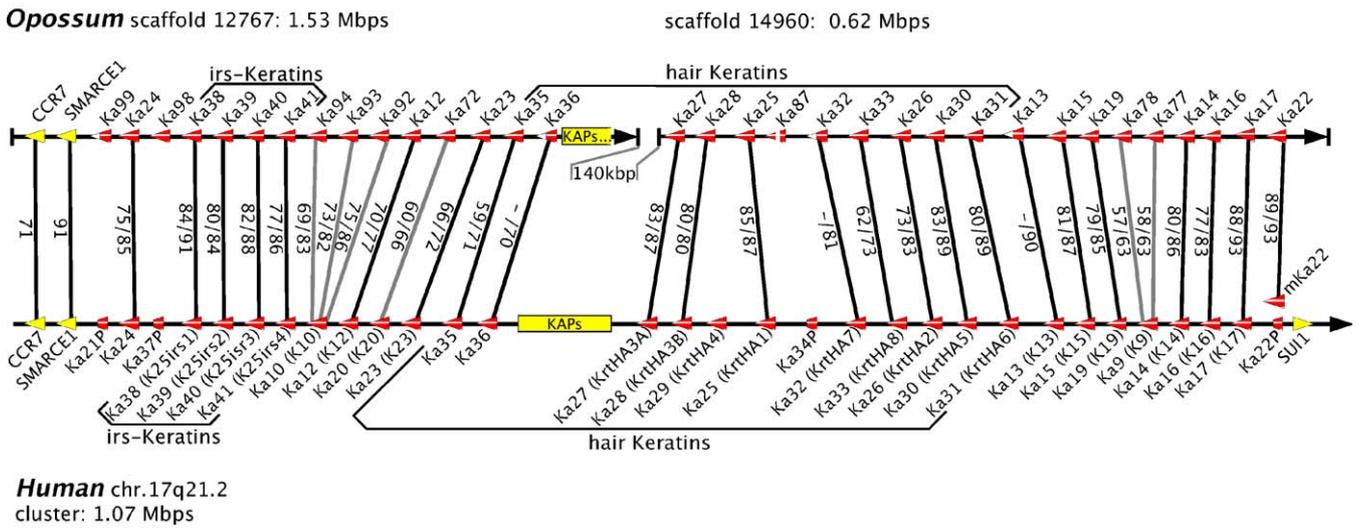
Keratin type-II cluster of the opossum

All type-II keratin genes of the opossum are located on scaffold 18721 which has a size of 1.9 Mbps. The cluster itself spans 1.3 Mbps and contains 32 type-II keratin genes (Fig. 1). In perfect agreement with all vertebrate genomes described so far, the type-I keratin Ka18 is located next to Kb8 in opposite orientation (Zimek and Weber, 2005). The cluster consists of 32 type-II keratin genes. Two subdomains of 8 hair keratin genes and 4 *irs*-keratin genes are embedded in 20 genes for cytoplasmic keratins (Fig. 1, Table 1).

Fig. 1 illustrates the opossum keratin type-II cluster and the proposed synteny for the marsupials and placental mammalian genes. Unresolved is the true synteny between the 4 neighboring opossum genes Kb90, Kb91, Kb92 and Kb94 and the 2 human genes Kb2 and Kb1. Half of the not yet completed gene Kb94 is nearly identical with Kb92. The first 3800 nt show 80% identity including the introns while the last 3100 nt are 99% identical including the introns. While this may indicate a very recent gene duplication it also seems possible that Kb94 was misassembled in the genome draft. Similarly unclear is the relation between opossum genes Kb97, Kb98 and Kb99 and the single human gene Kb38. The opossum genes Kb71 and Kb96 may be new genes without an obvious counterpart in the human genome (Hesse et al., 2004; Rogers et al., 2005).

As in the placental mammalian genomes the hair keratin II genes are inserted as a group into the range of

Keratin I gene cluster



Keratin II gene cluster

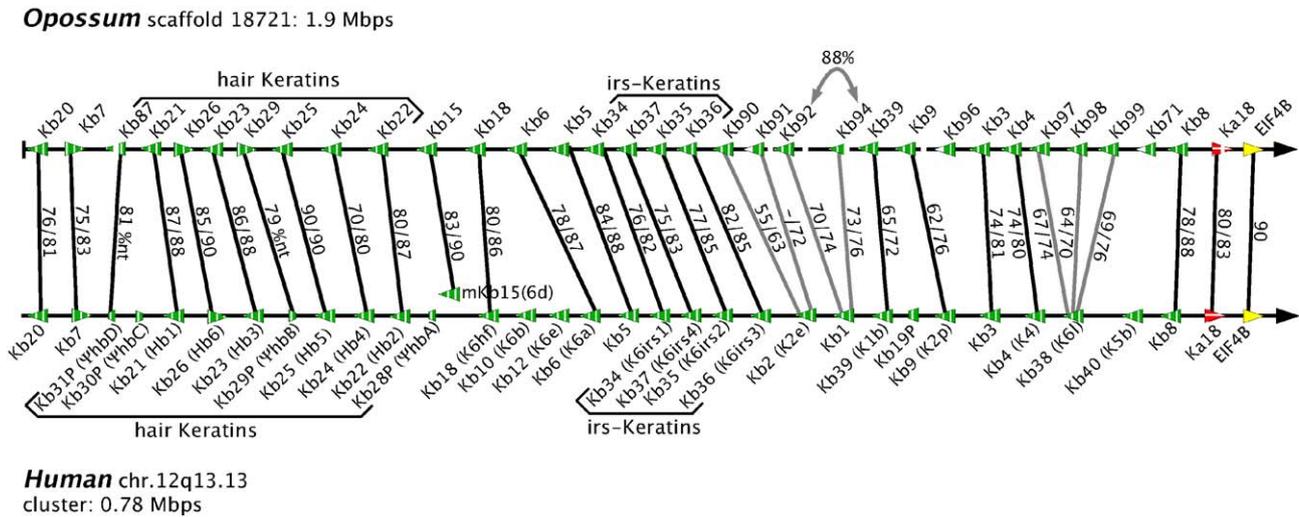


Fig. 1. Pairwise comparison of keratin gene clusters in opossum and man. Type-I keratin genes are marked as red triangles, type-II keratin genes as green triangles and other genes as yellow triangles. The tips of the triangles correspond to the 3'-end of the genes. Pseudogenes are drawn as blunted triangles. Keratins with unidentifiable tail domain have a colorless tip. The horizontal arrowlines display the 5'-3' orientation of the DNA strands. The keratin type-I gene cluster is located on scaffolds 12 767 and 14 960 (MonDom1 assembly). The distance between the scaffolds is about 140 kbp. This is documented in the latest release of the opossum genome (MonDom2, May 2005). Gaps in the assembly are indicated by white spaces (e.g. Ka87). Hair and inner-root-sheath (irs) keratins are marked by bracket lines. For reasons of clarity, the region containing some 25 keratin-associated proteins (KAPs) in the type-I cluster is replaced by a yellow box. Names for the human keratins follow the proposed nomenclature (Hesse et al., 2004), with traditional names added in brackets. Black vertical lines indicate real synteny of the connected genes following the criteria defined in the text. The gray lines connect the remaining opossum keratins with similar human proteins. Their identity values are below our criteria of synteny. Percent identities of the whole protein (first number) and rod sequence (second number) are given along with the connecting lines unless the identity is given on the nucleotide level (pseudogenes). For opossum keratins with synteny the same name is used as for the human homologs. This is the case for identities of the rod greater than 70% and identical orientation and position inside the cluster. For the remaining opossum genes, our sequence file numbers (numbers between 71 and 99) are used to avoid confusion with the accepted keratin names. In case of Ka22 and Kb15 we used the mouse keratin genes instead of the human pseudogenes. Neighboring non-keratin genes (yellow) are identical for human and opossum: SMARCE1 (NM_003079), CCR7 (BC035343), SUI1 (NM_005801), EIF4B (NM_001417).

Table 1. Number of genes for hair, irs and cytoplasmic keratins in various mammals and the chromosomal location (chr) of the 2 keratin gene clusters

| | Chr. | Hair keratins | irs k. | Cytoplasmic k. |
|----------------|----------|---------------|--------|----------------|
| <i>Type I</i> | | | | |
| Man | 17q21.2 | 11+1 P | 4 | 13+3 P |
| Chimpanzee | 19 | 12 | 4 | 13+3 P |
| Dog | 9 | 12 | 4 | 13+1 P |
| Mouse | 11qD | 9 | 4 | 14+2 P |
| Rat | 10q32.1 | 9 | 4 | 15+1 P |
| Opossum | ? | 11 | 4 | 19 |
| <i>Type II</i> | | | | |
| Man | 12q13.13 | 6+4 P | 4 | 16+1 P |
| Chimpanzee | 10 | 6+4 P | 4 | 16+1 P |
| Dog | 27 | 8+1 P | 4 | 18+1 P |
| Mouse | 15qF2 | 6+2 P | 4 | 17+2 P |
| Rat | 7q36 | 6+ >2 P | >2 | >13+1 P |
| Opossum | ? | 8 | 4 | 19 |

P indicates pseudogenes; > because of gaps in the genomic DNA sequence only a minimal value can be deduced. Genomes of man, mouse and rat were analyzed before (Hesse et al., 2004); databases for other species are given in 'Materials and methods'.

genes for cytoplasmic keratins. The assignment of the group of hair keratin genes is made because of the high cysteine content of the terminal domains of the predicted proteins. We tentatively identify 8 hair keratin type-II genes in the opossum while man has 6 type-II hair keratin genes plus 4 pseudogenes. Similarly well defined are in both genomes 4 neighboring genes for irs type-II keratins. Hair keratin genes and irs keratin II genes occur in the same relative position in the type-II clusters of man and opossum.

Differences between the opossum and human keratin clusters

The human keratin gene clusters are 1.07 Mbp (type I) and 0.78 Mbp (type II) in size (Hesse et al., 2004; Rogers et al., 2004, 2005). In the opossum the clusters increase in size to 1.7 Mbps (type I) and 1.3 Mbps (type II). This increase is due to larger intergenic distances and a larger average size of the genes. Opossum keratin genes are on average 49% larger than their human counterparts. Fig. 2 shows the size ranges. There are only 3 cases where the opossum gene is slightly smaller but 5 cases where the opossum gene is more than 100% larger than the human gene.

Discussion

Since the genomic database for the opossum *M. domestica* is highly advanced we could readily identify

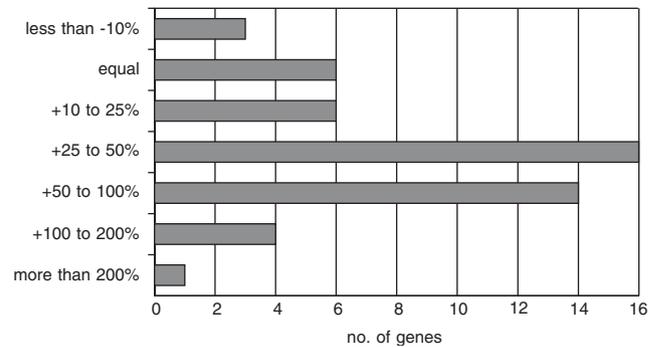


Fig. 2. Ratio of the size of the keratin genes in opossum and man. For the comparison we used only those opossum genes which have no gaps in the sequence and predict complete tail domains. Comparison was made for the genes connected by black and gray lines in Fig. 1. We calculated the gene length from the start to stop codon for a total collection of 50 opossum keratin genes. The figure shows the distribution of the gene size variations, when compared to their human counterparts. We selected 7 categories and assigned the values calculated for the opossum genes. The number of genes in each category is given by the length of the bars. Genes with variations of the length between +10% and -10% were regarded as equal, the limits for the other categories are given in the figure. On average, opossum keratin genes are 49% larger (10 490 bp) than human genes (6860 bp) with a standard deviation of 57% (data not shown). Most opossum keratins are 25–100% larger than their human homologs. Only 3 genes are slightly smaller. One gene, Kb97 (56 kbp), is 356% larger than the most similar human gene Kb38 (12 kbp).

the 2 keratin gene clusters and compare them with their human counterparts. Unfortunately there is lack of large-scale cDNA sequences and EST sequences. This limited our analysis in 2 respects. *First*, in those cases where small manual corrections of the genomic sequences were necessary to keep the open reading frame one cannot yet rigorously exclude a very closely related pseudogene. However, the occurrence of pseudogenes is very unlikely, since the manual changes to be introduced are minimal (see 'Results') and once they are made no unusually positioned stop codons are encountered. Given this situation a large-scale cDNA project will remove also the problem of a few opossum keratin genes where we have not found the regions encoding the terminal tail domain, which often poses problems without corresponding cDNA and/or EST sequence information. The current database favors the view that the 2 keratin gene clusters of the opossum have a few more genes than the human genome (33 versus 28 for type I; 32 versus 26 for type II) and seem to lack pseudogenes (Fig. 1). In contrast the human type-II gene cluster has 4 pseudogenes and the human type-I gene cluster shows 1 pseudogene (Hesse et al., 2004). *Second*, the synteny of opossum and human genes proposed in

Fig. 1 rests solely on a high sequence conservation of the predicted keratin proteins (more than 70% for the rod domain) and conservation of the relative position of the genes within the cluster. Once cDNA and EST sequences become available the tissue-specific expression patterns could also be compared in opossum and man. This would be particularly useful for those few opossum keratin genes where we are unsure of the synteny.

Our results show that the 2 keratin gene clusters of a marsupial and several placental mammals have very similar organizations and subdomains. Since marsupials and placental mammals are sister groups (Arnason et al., 2002; De Leo et al., 2005) we expect a corresponding relation also for the keratin I and keratin II gene clusters of the archetype mammal. This assumption is important as hair keratin genes are a mammalian acquisition. In both keratin clusters the genes encoding hair keratins are not located at one end of the clusters, but are found in an interior position (Fig. 1). Although we do not know from which genes the first hair keratin genes were derived subsequent events involved gene duplications plus sequence drifts to create multiple neighboring genes encoding the closely related hair keratins. A similar conclusion applies for the type-I and -II genes encoding keratins of the IRS. They formed as 4 neighboring genes in an interior position in both the type-I and -II gene clusters that does not overlap with the subcompartment occupied by the genes for the hair keratins (see also Hesse et al., 2004; Rogers et al., 2004, 2005). Finally we note that the acquisition of hair by the archetype mammal involved a subdomain of multiple genes encoding cysteine-rich and -ultrarich KAPs (see Rogers et al., 2001) which evolved at the same position in the type-I gene cluster (see also Hesse et al., 2004). Table 1 summarizes the number of hair keratin genes in the 2 clusters for the various placental mammals and the opossum.

An interesting aspect concerns the type-II keratin Kb3 (K3), which is expressed in the human cornea and which we also found in the genomic database for the chimpanzee and the dog. In contrast, rodents such as mouse and rat, lack a normal Kb3 (K3) gene (Hesse et al., 2004) and we find instead a very remote pseudogene of which only a few exons can be identified (our unpublished results together with T. Magin and M. Hesse). This loss of an active gene seems restricted to the rodent line of placental mammals and our proposed synteny pattern raises the possibility that the opossum has a true Kb3 gene (Fig. 1).

We also note an interesting evolutionary fate of the last keratin I gene in the type-I cluster (Ka22). While this is an active gene in mouse and rat (Hesse et al., 2004) that recently has been cloned from murine nail tissue ((Tong and Coulombe, 2004); Acc. No: AY459292) it is converted to a pseudogene in man

(Hesse et al., 2004) and the chimpanzee. It is also an active gene in the dog and our proposed synteny patterns (Fig. 1) predict that this gene is active in the marsupial *M. domestica*. Thus Ka22 seems to be an old mammalian keratin I gene which was specifically inactivated on the primate line of the placental mammals.

References

- Arnason, U., Adegoke, J.A., Bodin, K., Born, E.W., Esa, Y.B., Gullberg, A., Nilsson, M., Short, R.V., Xu, X., Janke, A., 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. *Proc. Natl. Acad. Sci. USA* 99, 8151–8156.
- Coulombe, P.A., Ma, L., Yamada, S., Wawersik, M., 2001. Intermediate filaments at a glance. *J. Cell Sci.* 114, 4345–4347.
- De Leo, A.A., Wheeler, D., Lefevre, C., Cheng, J.F., Hope, R., Kuliwaba, J., Nicholas, K.R., Westerman, M., Graves, J.A., 2005. Sequencing and mapping hemoglobin gene clusters in the Australian model dasyurid marsupial *Sminthopsis macroura*. *Cytogenet. Genome Res.* 108, 333–341.
- Erber, A., Riemer, D., Bovenschulte, M., Weber, K., 1998. Molecular phylogeny of metazoan intermediate filament proteins. *J. Mol. Evol.* 47, 751–762.
- Fuchs, E., Weber, K., 1994. Intermediate filaments: structure, dynamics, function, and disease. *Annu. Rev. Biochem.* 63, 345–382.
- Herrmann, H., Aebi, U., 2000. Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. *Curr. Opin. Cell Biol.* 12, 79–90.
- Hesse, M., Magin, T.M., Weber, K., 2001. Genes for intermediate filament proteins and the draft sequence of the human genome: novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *J. Cell Sci.* 114, 2569–2575.
- Hesse, M., Zimek, A., Weber, K., Magin, T.M., 2004. Comprehensive analysis of keratin gene clusters in humans and rodents. *Eur. J. Cell Biol.* 83, 19–26.
- Irvine, A.D., McLean, W.H., 1999. Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype-genotype correlation. *Br. J. Dermatol.* 140, 815–828.
- Karabinos, A., Zimek, A., Weber, K., 2004. The genome of the early chordate *Ciona intestinalis* encodes only five cytoplasmic intermediate filament proteins including a single type I and type II keratin and a unique IF-annexin fusion protein. *Gene* 326, 123–129.
- Rogers, M.A., Langbein, L., Winter, H., Ehmann, C., Praetzel, S., Korn, B., Schweizer, J., 2001. Characterization of a cluster of human high/ultra-high sulfur keratin-associated protein genes embedded in the type I keratin gene domain on chromosome 17q12–21. *J. Biol. Chem.* 276, 19440–19451.
- Rogers, M.A., Winter, H., Langbein, L., Bleiler, R., Schweizer, J., 2004. The human type I keratin gene family:

- characterization of new hair follicle specific members and evaluation of the chromosome 17q21.2 gene domain. *Differentiation* 72, 527–540.
- Rogers, M.A., Edler, L., Winter, H., Langbein, L., Beckmann, I., Schweizer, J., 2005. Characterization of new members of the human type II keratin gene family and a general evaluation of the keratin gene domain on chromosome 12q13.13. *J. Invest. Dermatol.* 124, 536–544.
- Schaffeld, M., Bremer, M., Hunzinger, C., Markl, J., 2005. Evolution of tissue-specific keratins as deduced from novel cDNA sequences of the lungfish *Protopterus aethiopicus*. *Eur. J. Cell Biol.* 84, 363–377.
- Tong, X., Coulombe, P.A., 2004. A novel mouse type I intermediate filament gene, keratin 17n (K17n), exhibits preferred expression in nail tissue. *J. Invest. Dermatol.* 122, 965–970.
- Wang, J., Karabinos, A., Zimek, A., Meyer, M., Riemer, D., Hudson, C., Lemaire, P., Weber, K., 2002. Cytoplasmic intermediate filament protein expression in tunicate development: a specific marker for the test cells. *Eur. J. Cell Biol.* 81, 302–311.
- Zimek, A., Weber, K., 2002. The gene for a cytoplasmic intermediate filament (IF) protein of the hemichordate *Saccoglossus kowalevskii*; definition of the unique features of chordate IF proteins. *Gene* 288, 187–193.
- Zimek, A., Weber, K., 2005. Terrestrial vertebrates have two keratin gene clusters; striking differences in teleost fish. *Eur. J. Cell Biol.* 84, 623–635.