HOKKAIDO UNIVERSITY

Title	The SKINT1-like Gene Is Inactivated in Hominoids But Not in All Primate Species : Implications for the Origin of Dendritic Epidermal T Cells
Author(s)	Hassan, Rania Hassan Mohamed
Description	配架番号：2174
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第11692号
Issue Date	2015-03-25
DOI	https://doi.org/10.14943/doctoral.k11692
Doc URL	https://hdl.handle.net/2115/59479
Type	doctoral thesis
File Information	Rania_Hassan_Mohamed_Hassan.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学　位　論　文

The *SKINT1*-like Gene Is Inactivated in Hominoids But Not in All Primate Species: Implications for the Origin of Dendritic Epidermal T Cells

（ヒト及び類人猿で不活化している SKINT1 様遺伝子の旧世界ザルでの機能残存と樹状表皮 T 細胞の由来に関する検討）

2015 年 3 月

北 海 道 大 学

ラニア ハッサン モハメド ハッサン

Rania Hassan Mohamed Hassan

Table of Contents

List of Publications and Presentations at Academic Seminars 1

Introduction 2

Abbreviations List 7

Materials and Methods 9

Results 13

Discussion 33

Summary and Conclusion 37

Acknowledgement 38

References 39

List of Publications and Presentations at Academic Seminars

1. Publications:

1) Rania Hassan Mohamed, Yoichi Sutoh, Yasushi Itoh, Noriyuki Otsuka, Yukiko Miyatake, Kazumasa Ogasawara and Masanori Kasahara. The SKINT1-like gene is inactivated in hominoids but not in all primate species: Implications for the origin of dendritic epidermal T cells. *PLoS One*, Under revision.

2) Shigeru Yoshida, Rania Hassan Mohamed, Mizuho Kajikawa, Jun Koizumi, Minami Tanaka, Kazunori Fugo, Noriyuki Otsuka, Katsumi Maenaka, Hideo Yagita, Hitoshi Chiba and Masanori Kasahara. Involvement of an NKG2D ligand H60c in epidermal dendritic T cell-mediated wound repair. *J. Immunol* **188**, 3972-3979 (2012).

2. Presentations at academic seminars

1) Rania Hassan Mohamed, Shigeru Yoshida, Jun Koizumi, Kazunori Fugo, Noriyuki Otsuka, Masanori Kasahara.. Blockade of interactions between NKG2D and an NKG2D ligand H60c delays wound healing. The 91st Hokkaido Medical Congress of Pathology Association. September 10, 2011, Asahikawa, Japan.

2) Rania Hassan Mohamed, Shigeru Yoshida, Jun Koizumi, Kazunori Fugo, Noriyuki Otsuka, Masanori Kasahara. Blockade of interactions between NKG2D and an NKG2D ligand H60c delays wound healing. The 2012 Annual Meeting of the Japanese Society of Immunology. December 5-7, 2012, Kobe, Japan.

3) Rania Hassan Mohamed, Yoichi Sutoh, Noriyuki Otsuka, Yukiko Miyatake and Masanori Kasahara. Inactivation of *Skint1* in hominoids but not in Old World monkeys. The 32nd Annual Meeting of the Japanese association for research on the Thymus. February 9, 2013, Sapporo, Japan.

4) Rania Hassan Mohamed, Yoichi Sutoh, Yasushi itoh, Noriyuki Otsuka, Yukiko Miyatake, Kazumasa Ogasawara and Masanori Kasahara. Inactivation of SKINT1L in hominoids but not in Old World monkeys. The 2013 Annual Meeting of the Japanese Society of Immunology. December 11-13, 2013, Chiba, Japan.

5) Rania Hassan Mohamed, Yoichi Sutoh, Yasushi itoh, Noriyuki Otsuka,Yukiko Miyatake, Kazumasa Ogasawara and Masanori Kasahara. The SKINT1-like gene is inactivated in hominoids but not in Old World monkeys. The 23rd Annual Meeting of the Japanese Society for Histocompitability and Immunogenetics. September 13-15, 2014, Nagasaki, Japan.

Introduction

γδ T lymphocytes along with αβ T lymphocytes act co-operatively in all vertebrate species to induce the protective immune reactions against various antigens from the inside and outside the body[1]. γδ T cells constitute a small proportion in blood, but they are spreading more widely within the epithelial tissue such as skin, intestine and reproductive organs. γδ T cells develop in the thymus ahead of αβ T cells in most species, suggesting their unique contribution in the immune system. Their development and selection mechanisms are not yet well defined as αβ T cells, although they originate from a common thymic progenitor[2]. γδ T cells are characterized by the expression of heterodimeric γδ T cell receptor (TCR) in a tissue restricted pattern[1,2], that is formed by recombination of the sequence elements; variable (V), diversity (D; for δ chains) and junctional (J) forming γ and δ loci responsible for the TCR diversity. This expression pattern is extensively defined in mice and humans, but in contrary cows lack this highly restricted TCR expression pattern as a result of the presence of more V(D)J cassettes for the γδ TCR rearrangement[1]. The nomenclature of the γδ TCR is designed according to the number of the Vγ and the Vδ segment usage in the TCR rearrangement[3]. The restricted expression of γδ TCR reflects the specificity of the antigen recognition repertoires in different tissues[4,5]. γδ T cells express both monoclonal and oligoclonal γδ TCR, which is characterized by greater potential diversity because of their capacity to use multiple copies of D elements and the junctional diversity generated by the V(D)J recombination[1,2].

γδ T cells antigen recognition manner differs from αβ T cells, where αβ T cells are limited to recognize the peptide antigens via the major histocompitability complexes (MHC), γδ T cells depend on the conformational shape of the intact protein and non-protein antigens for their recognition. γδ T cells carry out various functions including the secretion of IFN-γ[6], induction of cytotoxicity against tumor cells[7,8], production of multiple growth factors contributing to the process of wound healing[9] and antigen presentation[10]. Moreover, the localization of γδ T cells in the epithelial tissue presented as interepithelial lymphocytes (IELs) supports the barrier function. The γδ IELs maintain homeostasis of epithelial tissues through direct association with self antigens expressed by epithelial cells in an “activated-yet-resting” phenotype[11]. γδT cells are "innate like lymphocytes" representing a cross-talk between the adaptive and innate immunity. With the usage of monoclonal and oligoclonal receptors which act as antigen recognition pattern

receptors, they recognize several antigens such as phosphoantigens, nonclassical MHC-molecules and unprocessed proteins. They also perform their immunological functions in reaction with the other adaptive immune cells[12,13]. γδ T cells respond to tissue stress like injuries, cancerous transformation or inflammation within hours, not only through the binding of γδ TCRs and their yet unknown ligands, but also through other co-stimulators such as the engagement between the activating receptor; natural killer group 2 member D (NKG2D) and its ligand NKG2DL[14,15]. γδ T IEL function locally by binding to the self-antigens expressed in the residing epithelial tissue and systemically by their strong interaction with the lymphocyte network surrounding them[16]. These reactions are induced by chemokine production to attract B cells, direct interaction with the dendritic cells (DC) and antigen presentation to the αβ T cells[13] (Figure 1).

Many properties of γδ T cells were elucidated by investigating the γδ IELs residing in mouse skin known as dendritic epidermal T cells (DETCs)[17]. These cells were found to play a key role in epidermal homeostasis, skin tumor surveillance and wound repair mainly in the regulation by the epidermal stress ligands binding to their invariant γδ TCR[18,19].

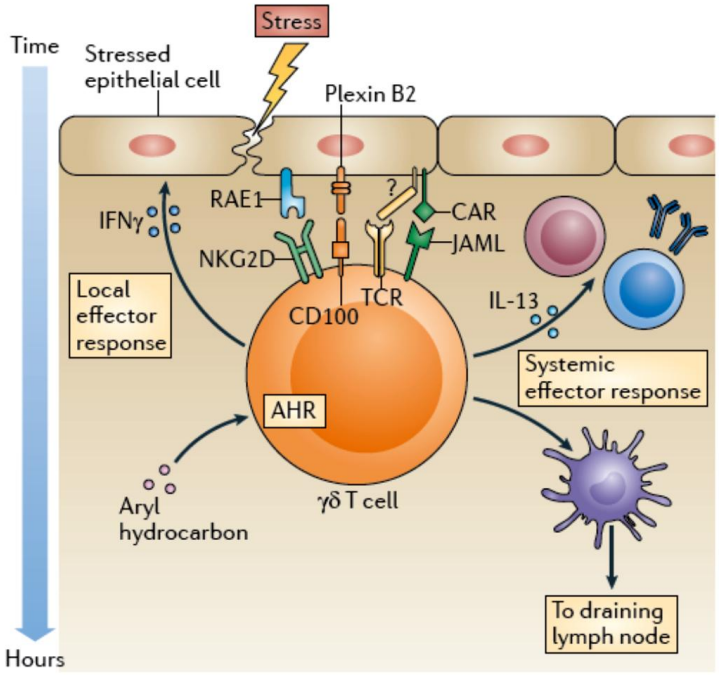


Figure 1: Mouse γδT cells lymphoid stress-surveillance: AHR, aryl hydrocarbon receptor; CAR, coxsackievirus and adenovirus receptor; IFNγ, interferon-γ; IL, interleukin; JAML, junctional adhesion molecule-like; TCR, T cell receptor. This figure is taken from Vantourout et al., 2013. Nature.

Denderetic epidermal T cells (DETC):

DETCs are the specialized epidermal murine γδ IEL expressing monomorphic Vγ5Vδ1 TCR and forming about 95% of epidermal T cell repertoire. The monomorphic DETC TCR binds to an unknown steady state ligand expressed by keratinocytes[20], in addition it recognizes a limited set of "stress antigens" induced on the surface of dysregulated skin epithelial cells. After DETCs respond to the epithelial cell changes, they come in a direct contact with Langerhans cells to contribute to the systemic immune response[21-23].

γδ TCR–deficient ($Tcrd^{-/-}$) mice are susceptible to carcinogenesis, impaired wound healing and skin diseases[24,25], despite the presence of intraepithelial T lymphocytes expressing αβ TCRs. Likewise, the human epidermal resident γδ T cells contribute in the process of acute wound healing, although human epidermis is not identified to have equivalent skin γδ IEL population of limited diversity[26]. DETC is the first wave of γδ T cells to be positively selected in the fetal thymus from embryonic day 14.5 to day 18.5 of gestation, then migrate to the suprabasal epidermis[27,28]. It is believed that DETCs proliferate *in situ* and selectively expand after homing to the skin in response to the keratinocytes, which propose existing of a keratinocyte ligand recognizes the invariant DETC TCR and enhance their expansion[29].

Selection and development of DETC:

The development and the acquisition of the invariant TCR repertoire of DETC are mainly supported by thymic selection. Many studies addressed normal DETC development in mice expressing transgenic γδ TCR suggesting the independence of DETC development on expression of Vγ5 or Vδ1[28]. On the other hand, DETC maturation was found to be retained by specific antibody mediating Vγ5Vδ1 TCR ligation using fetal thymocyte cell suspension of the FVB mice strains (FVB/N-Tac) which is known to be deficient in the canonical mature DETC, this phenomenon is attributed to the lack of the thymic maturation as a result of the inactivation of selection and upkeep of intraepithelial T-cells protein 1 (Skint1). This strongly suggests the essential role of Vγ5Vδ1TCR-Skint1 interaction for selection and development of the invariant DETC in mouse skin. Up to date there is no evidence available whether Skint1 directly or indirectly engages with DETC TCR.

***Skint1* gene**

Skint1 gene is the first identified natural component of the selection machinery for a specific IEL compartment. Skint1 is expressed exclusively in the thymic epithelium and keratinocytes at embryonic day 15 and continue to adulthood[30,31]. Skint1 is a membrane immunoglobulin protein made up of eight modular exons coding for seven domains; a signal peptide (SP), immunoglobulin variable (IgV) domain, an Ig constant (IgC) domain, three transmembrane domains (TMD) and short cytoplasmic domain, each exon expresses one domain (Figure 2)[31]. *Skint1* gene duplicated in mice and gave multiple paralogs forming *Skint* gene family which is composed of 3 subfamilies (*Skint*1-6), (*Skint*7,8) and (*Skint*9-11)[32,33].

It is found that in FVB/N-Tac mice, epidermis harbors T cells expressing diverse γδ TCR in a comparable number to the authentic DETC in the other mouse strains. Moreover, the γδ and αβ T cell repertoires maturation in the other organs appear to be normal, which indicates the exclusive effect of *Skint1* gene to DETC without any impact on the rest of lymphocytic repertoires[34]. *Skint1* gene is necessary for the functional programming of DETC to be either IFN-γ or IL-17 secreting cells, this is determined in the DETC ontogeny depending on the strength of the Skint1 signal which enhances the downstream pathway to produce those cytokines[35].

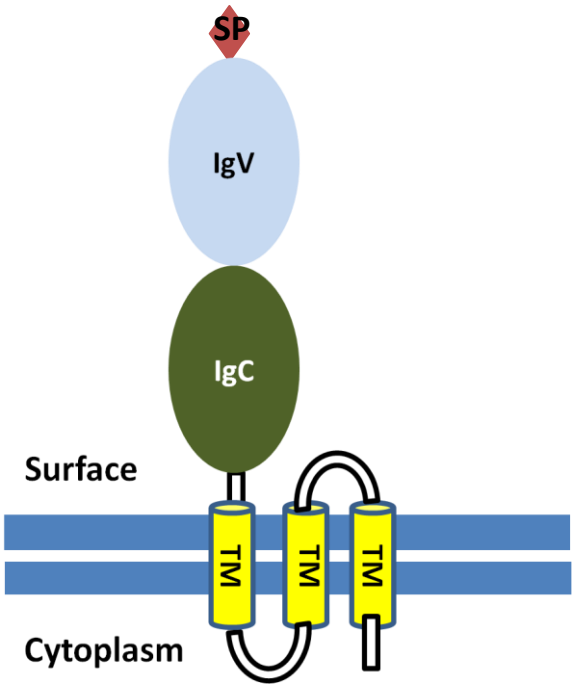


Figure 2: **Mouse Skint1**

***Skint1* gene evolution:**

Skint1 gene has been known to be highly evolved from rodents to humans. Rat and cow Skint1 have identical organization and protein topology to mice Skint1. Previous work revealed that humans and chimpanzee have inactive *Skint1* gene with multiple in-frame premature termination codons[31]. Consistent with this, human epidermis lacks the comparable number of the conservative population of γδ T cells found in mice[36-38]. Previous studies reported that *Skint1* gene duplicated in mice, but this duplication was lost within the primate lineage[31,32]. All these previous observations we mentioned above suggested that *Skint1* underwent multiple evolutionary events, though the information available about the evolution of *Skint1* in mammals remained limited. This led us to determine at which stage in the mammalian phylogeny *Skint1* lost its function and concomitantly the invariant DETCs were lost.

In the present study, for the sake of consistency in nomenclature of the *Skint* family genes, we called *Skint1* gene of all mammalian species other than mice *SKINT1L*, we took an evolutionary approach to analyze *SKINT1L* sequences in the representative mammalian species. The bioinformatics analysis of primate *SKINT1L* sequences revealed that the inactivation of *SKINT1L* took place in the hominoid lineage and that Old World monkeys (OWM) such as olive baboons, green monkeys, cynomolgus macaques and rhesus macaques retain intact *SKINT1L* genes. As a representative of OWM, the cynomolgus macaques were chosen and we found that their epidermis contained skin-resident γδ T cells expressing an invariant TCR. Comprehensive analysis of available mammalian genome sequences indicated that SKINTL family members emerged in an ancestor of placental mammals, but *SKINT1L* was inactivated or lost multiple times in mammalian evolution, suggesting that DETCs expressing an invariant TCR were lost in many mammalian orders.

Abbreviations List

Ab	Antibody
Cγ	TCR γ chain constant segment
Der	Dermis
Epi	Epidermis
IgC	Immunoglobulin constant domain
IL	Interleukin
IgV	Immunoglobulin variable domain
Jγ	TCR γ chain joining segment
mg, ml	milligram, milliliter
MHC	Major histocompatibility complex
NKG2D(L)	Natural killer group 2 member D (ligand)
NWM	New World Monkey
OMW	Old World Monkeys
PCR	Polymerase chain reaction
RACE	Rapid amplification of cDNA ends
RT	Reverse transcriptase
RPMI	Endothelial cell Basal Medium
SKINT	Selection and upkeep of intraepithelial T-cells protein
Tac	Taconic
TCR	T cell receptor
Th22	T helper 22

TMD	Transmemberane domain
Vγ	TCR γ chain variable segment
Vδ	TCR δ chain variable segment
γδ T cells	Gamma delta T cells

Materials and Methods

Animal samples

Tissue samples were obtained from cynomolgus macaques also known as crab-eating macaques (*Macaca fascicularis*) maintained at Shiga University of Medical Science. Skin tissues were excised from the back or palm of adult macaques after euthanasia. All organ samples were obtained from adult female macaques except for the thymus that was biopsied from a 100-day-old male embryo. All animal experiments were conducted according to the Guidelines for the Care and Use of Laboratory Animals at Shiga University of Medical Science and Hokkaido University Graduate School of Medicine.

Database analyses

Ensembl and NCBI databases were searched using human SKINT1L nucleotide (NR_026749.2) and mouse SKINT1 protein (NP_001096132.1) sequences as queries. Non-human primate species subjected to analysis were bushbaby (*Otolemur garnettii*), chimpanzee (*Pan troglodytes*), common baboon (*Papio hamadryas*), cynomolgus macaque, gibbon (*Nomascus leucogenys*), gorilla (*Gorilla gorilla gorilla*), green monkey (*Chlorocebus sabaeus*), marmoset (*Callithrix jacchus*), mouse lemur (*Microcebus murinus*), olive baboon (*Papio anubis*), orangutan (*Pongo abelii*), rhesus macaque (*Macaca mulatta*), squirrel monkey (*Saimiri boliviensis*), tarsier (*Tarsius syrichta*). Protein domains were predicted using the SMART server (http://smart.embl-heidelberg.de/smart/set_mode.cgi). Signal peptides were predicted using the SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/).

Phylogenetic analysis

Amino acid sequences were aligned using the Clustal Omega program[39] with default parameters. The tree was constructed using the neighbor-joining algorithm implemented in the MEGA version 6.0 software[40]. The distance matrix was obtained by calculating p-distances for all pairs of sequences. Gaps were excluded using the pairwise-deletion option. The reliability of branching patterns was assessed by bootstrap analysis (1,000 replications).

Isolation of cDNAs coding for cynomolgus macaque SKINT1L and expression analysis of cynomolgus macaque *SKINT1L* and TCR Vγ/Vδ gene segments

After RNA extraction from cynomolgus macaque organs including thymus, kidney, heart, liver, lung, bladder, uterus, skin and epidermal lymphocytes using the RNeasy mini Kit (Qiagen GmbH, Hilden, Germany), following the instructions of the manufacturer, the purified RNA was then treated with RNase free-DNase (Invitrogen, Camarillo, CA) and converted to cDNA using SuperScript III RT polymerase (Invitrogen), For isolation of *SKINT1L*, cDNA was amplified from skin by 3'-rapid amplification of cDNA ends (RACE) using high fidelity polymerase chain reaction (PCR) polymerase KOD -Plus- (Toyobo, Osaka, Japan). The sequences of gene-specific primers, which were designed based on the genomic sequences retrieved from the NCBI database, were 5'-TTTGGTGTCACCTGGCTCAA-3' for the first round of PCR and 5'-TGGGACCATCTAGTTGCAGGAA-3' for nested PCR. PCR products were cloned into the pGEM-T easy vector (Promega, Madison, WI), and multiple clones were sequenced using an automated sequencer. The cDNA from epidermal lymphocytes was used as a template to amplify cynomolgus macaque TCR Vγ/Vδ gene segments. Specific primers for each segment were used for amplification. Primers are listed in Table 1.

Table 1: Primers used for analyzing Vγ/Vδ gene usage*

Primer	**Sequence**
Cγ reverse	5'-TGACTTTTCTGGCACCGTTA-3'
Vγ1 forward	5'-AGCACAAGGAAGAGCTGGAAT-3'
Vγ2 forward	5'- ATACACGCAAGGTGGAGT -3'
Vγ3 forward	5'-TTGTTGGGTGGCACCTGAGT-3'
Vγ2/4 forward	5'-GATAGTTATGGAAGCACAAGGAACA-3'
Vγ9 forward	5'-CCTGCAGACATGCTGTCACT-3'
Vγ10 forward	5'-ATTCGCCTTCTCGTCCTTCT-3'
Vδ1 forward	5'-GTGGCCCAGAAGGTTACTCA-3'
Vδ1 reverse	5'-GGAGCTAGCTGCTTTCTGGA-3'
Vδ2 forward	5'-CTGTTGAGTTGGTGCCTGAA-3'
Vδ2 reverse	5'-CACCTTGGAAATTGTCCTTGA-3'
Vδ3 forward	5'-TTTCTACAGGGCAATGCTGT-3'
Vδ3 reverse	5'-GCTTCACAGAAAACCGTCCT-3'
Forward primer around the intronic sequence	5'-CAGATACTTTGTTGTTATGAATCTCCT-3'
Reverse primer around the intronic sequence	5'ACCCAGTGCAGCCAAGACT-3'

* Vγ gene usage was evaluated using Vγ/Cγ primer pairs whereas Vδ gene usage was evaluated using pairs of primers designed within the Vδ gene segments.

Immunohistochemistry

To detect crab eating macaque skin gamma delta T cell population, skin tissues were harvested and either embedded in paraffin after formalin fixation. 4 μm-thick paraffin sections were reacted with mouse anti-human TCR CγM1 monoclonal antibody (Ab) (γ3.20; Thermo Scientific, Waltham, MA) or (5A6.E91; Thermo Scientific, Waltham, MA) using an automatic stainer and visualized with 3,3'-diaminobenzidine. The sections were then counterstained with hematoxylin. For detecting CD3+ cells, rabbit anti-human CD3 polyclonal Ab (A0452; Dako, Glostrup, Denmark) was added for 1 h followed by incubation with Alexa Fluor 594-conjugated goat secondary Ab for 30 min at room temperature. The tissue sections were washed in PBS/T20 (0.05 %) and then mounted with VECTASHIELD mounting medium (Vector Laboratories, Burlingame, CA). For double staining, the sections were stained first with mouse anti-human TCR CγM1 monoclonal Ab. After counterstaining and blocking with 10^{-2} goat serum for 1 h, they were stained with rabbit anti-human CD3 polyclonal Ab, and then mounted with VECTASHIELD mounting medium. Cynomolgus macaque skin serial sections were analyzed and Epidermal $CD3^{+}$ $\gamma\delta^{+}$ T cells were counted and their number was calculated per mm basement membrane using the ImageJ software (version 1.46r)[41]. The percentage of γδ TCR+ cells in CD3+ cells was scored in triplicate for each animal, and the data obtained from three animals were combined for statistical analysis.

Isolation of epidermal lymphocytes from cynomolgus macaque skin

To examine the gamma delta TCR expression in crab eating monkey epidermis, the epidermal lymphocytes were isolated as described[42] with some modifications. Briefly, subcutaneous fat was removed from macaque skin tissues with a razor blade. They were then cut into strips, and the epidermis was separated and digested using 10 ml of RPMI 1640 containing 1 mg/ml of collagenase and 1mg/ml of dispase II overnight at 37°C. Epidermal cell suspensions were filtered through a 70-μm pore nylon mesh (BD Biosciences) and centrifuged (300 *g*) for 20 min at 4°C and then enriched for lymphocytes using Ficoll-Paque gradient medium. After washing several times with PBS, cells were immediately frozen using liquid nitrogen for subsequent RNA extraction.

Results

SKINT1L is inactivated in hominoids but is apparently functional in OWMs

Although it's been detected and highly duplicated within rodents, it has been reported that SKINT1L ortholog is pseudogenized by multiple stop codons within the human lineage[31,32]. To verify the existence and the status of SKINT1L within the primate lineages, we undertook an extensive search using the human SKINT1L protein as a query to recognize its orthologs in primate species ranging from prosimians to hominoids (Figure 3). This analysis, which involved human and 15 non-human primate species for which genome information is available, showed that despite the inactivation of SKINT1L within the human linage, great apes (chimpanzees, gorillas and orangutans) and lesser apes (gibbons), it is structurally active in Old World Monkeys (OWM) such as olive baboons, green monkeys, cynomolgus macaques and rhesus macaques which have apparently intact SKINT1L sequences lacking the common stop codon located at the ninth residue of the IgV domain of apes. This stop codon was shared by all the hominoid sequences, suggesting that this mutation was responsible for the initial inactivation of the hominoid *SKINT1L* gene. Two mutations, one located between TMD1 and TMD2, and another located downstream of TMD2 leading to the elimination of TMD3, were shared by humans and great apes, but not by gibbons, suggesting that these mutations occurred in a common ancestor of humans and great apes. Prosimians such as bushbabies also have apparently intact SKINT1L sequences.

Furthermore, in contrast to humans and great apes, the SKINT1L molecules of OWMs and prosimians have three TMDs similar to rodent SKINT1 molecules (Figure 4). Therefore, we conclude that the stop codon at the ninth residue of the IgV domain shared by all the hominoid species took place in a common hominoid ancestor after its separation from OWMs. Interestingly, we could not identify SKINT1L sequences in the genome sequences of common baboons, marmosets, squirrel monkeys or mouse lemurs. It is possible that we failed to identify SKINT1L sequences in squirrel monkeys or mouse lemurs because their genomes have been sequenced only to low coverage.

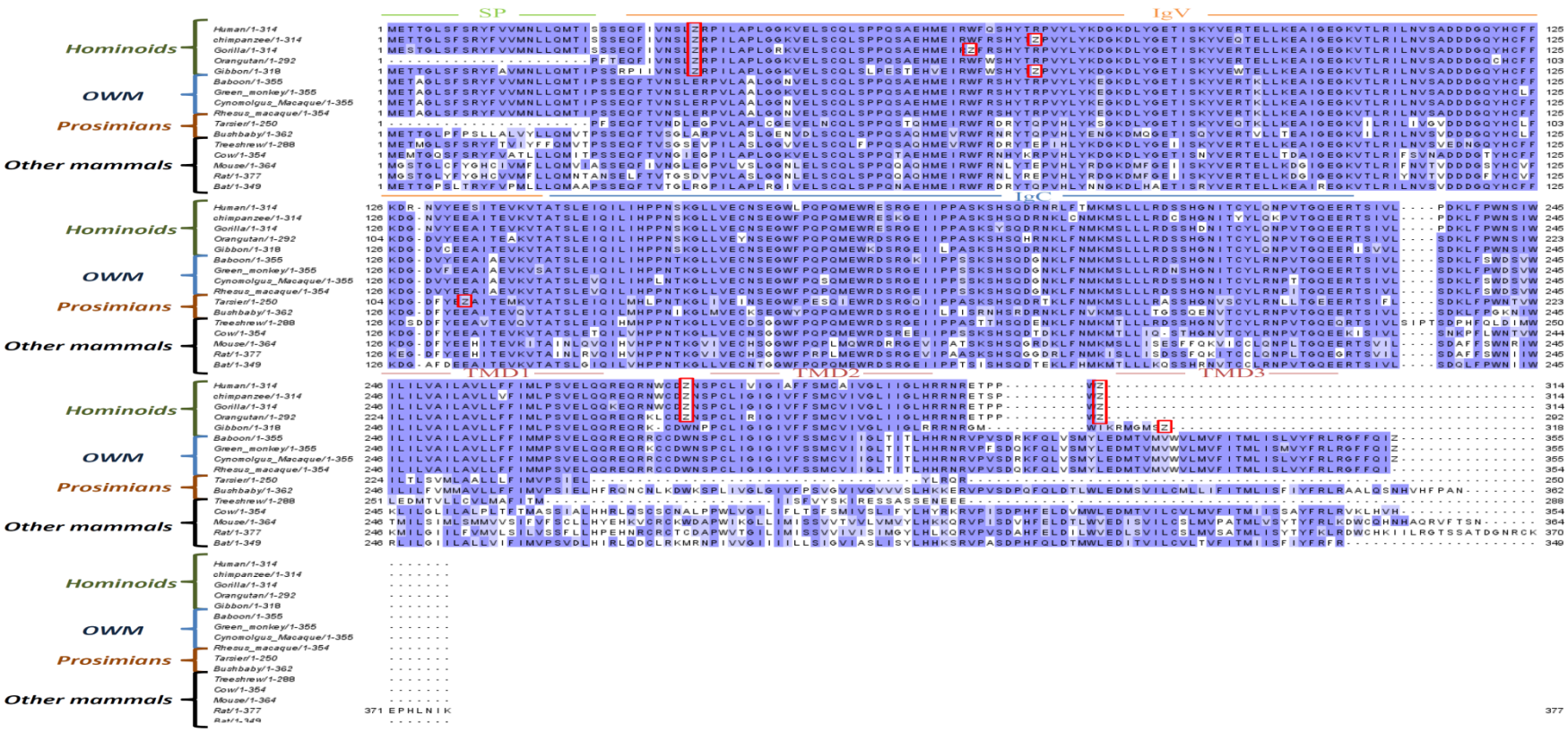


Figure 3: Amino acid sequence alignment of SKINT1 and SKINT1L molecules: Deduced SKINT1 and SKINT1L protein sequences of indicated species were aligned using the Clustal Omega program. Strictly and highly conserved residues are indicated in dark blue and light blue, respectively. Hominoid sequences contain multiple stop codons (indicated by red frams). The location of predicted domains is indicated on top of the sequences. The brackets in green, blue and orange represent hominoids, OWMs and prosimians, respectively. SP stands for signal peptides. Accession numbers of sequences and other relevant information are given in Table 2.

Table 2: Accession numbers of primate SKINT1L sequences

			Regions		
	Species	Accession number	From	To	Databases
Hominoids	Human	NC_000001	48,101,715	48,182,428	NCBI genome
	Chimpanzee	NC_006468.3	48,499,761	48,548,233	NCBI genome
	Gorilla	NC_018424.1	49,735,301	49,819,734	NCBI genome
	Orangutan	NC_012591.1?	181,680,692	181,734,447	NCBI genome
	Gibbon	NC_019827.1	17,090,514	17,112,139	NCBI genome
OWM	Olive baboon	NC_018152.1	50,394,763	50,437,695	NCBI genome
	Common baboon	No hit			Pre-Ensembl
	Cynomolgus macaque	AEHL01229074.1	11,817	27,285	NCBI genome
		AQIA01004292.1	5,930	35,479	NCBI genome
	Rhesus macaque	AEHK01151475.1	1,172	3,795	NCBI genome
		AEHK01151471.1	378	728	NCBI genome
		AEHK01151469.1	2,098	2,382	NCBI genome
		AEHK01151468.1	1,705	4,053	NCBI genome
		AANU01266778.1	5,306	9,035	NCBI genome
	Green monkey	NC_023661.1	84,640,177	84,686,975	NCBI genome
NWM	Squirrel monkey	No hit			Pre-Ensembl
	Marmoset	No hit			Ensembl
Prosimians	Mouse Lemur	No hit			Ensembl
	Bushbaby	GL873729.1	2,367,236	2,386,348	Ensembl
	Tarsier	ABRT02125005.1	11,139	25,919	NCBI genome
		ABRT02314780.1	143,050	127,400	NCBI genome

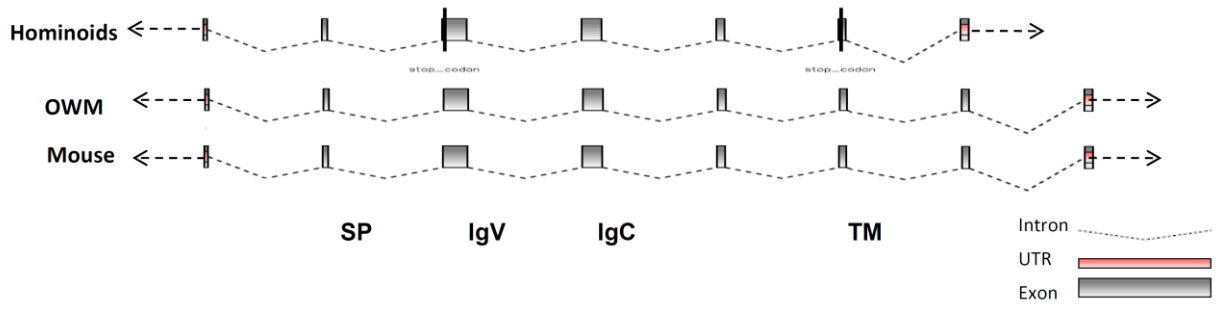


Figure 4: The exon-intron organization of hominoid *SKINT1L* was compared to that of mouse *Skint1* and OWM *SKINT1L*. Exon-intron boundaries were predicted based on the consensus splice junction sequences and similarity of the deduced amino acid sequences to the mouse SKINT1 protein sequence.

Cynomolgus macaques *SKINT1L* expression and it's splicing variants:

To confirm the structure of OWM *SKINT1L*, we chose cynomolgus macaques as a representative of OWMs and two specific oligonucleotides complementary to 5' UTR of the predicted sequence of crab eating macaque (Macaca fascicularis) *SKINT1L* were used for its isolation from the skin tissue, then the whole sequence of *SKINT1L* was obtained by 3' RACE using skin cDNA as a template (accession number: AB974689). Cloning and sequence analysis revealed that, consistent with the predicted sequence by bioinformatics analysis, the crab eating macaque *SKINT1L* is structurally intact. Using the TMHMM program, we confirmed that it has exon-intron organization essentially identical to that of mouse and bovine *SKINT1L* genes (Figure 4). Macaque SKINT1L has seven exons encoding protein out of eight exons for signal peptide (SP), Ig-like domains (IgV and IgC), 3 transmembrane domains (TMD) with an extracellular N terminus and a short cytoplasmic C terminus. The alignment of Macaque *SKINT1L* mRNA shows a 77% (57% amino acids) and 93% (86% amino acids) match with mouse and human *SKINT1L* genes respectively over their entire length, with a comparable exon size between the three species. Regions with the greatest similarity are present within the predicted open reading frame of Ig-like domains and those of less similarity lie between the predicted UTR's and TMD The cynomolgus macaque genome contains only a single copy of the

SKINT1L gene. Furthermore, *SKINT1L* is showed a robust expression only in the thymus and skin, suggesting that it is the functional counterpart of mouse Skint1 (Figure 5).

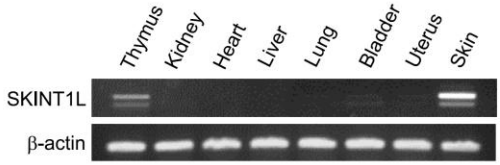


Figure 5: *SKINT1L* expression in cynomolgus macaque tissues. *SKINT1L* expression was examined by reverse transcription PCR. β-actin was used as a control. Thymic tissues were obtained from a macaque fetus at embryonic day 100; the remaining tissues were procured from adult animals. A faint, lower band in the thymus and skin represents alternatively spliced transcripts encoding SKINT1L molecules with only two TMDs.

Sequence analysis of macaque SKINT1L transcripts revealed that the *SKINT1L* gene undergoes alternative splicing; besides the major transcript (corresponding to the upper strong band in Figure 5), we detected two splicing variants (Figure 6): one variant encodes SKINT1L molecules with only two TMRs (corresponding to a lower faint band in Figure 5), expressed at a very low level compared with the major transcript, and another variant produces transcripts that cannot encode functional SKINT1L protein because of insertion mutation. Sequences of both splicing variants are shown in (Figure 7A).

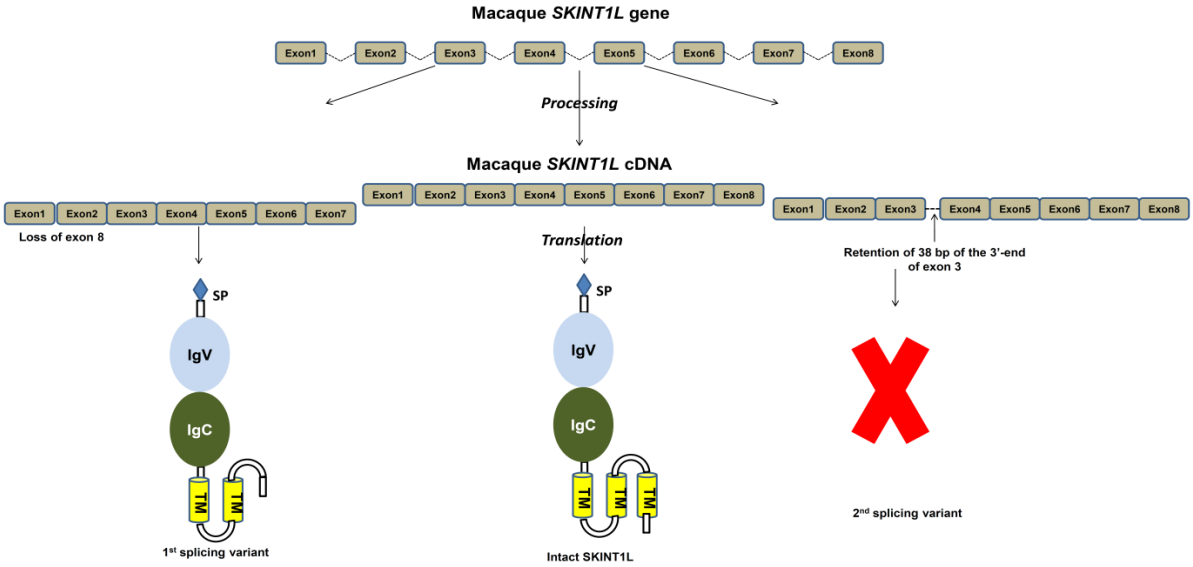


Figure 6: Schematic diagram for Crab eating macaque *SKINT1L* gene splicing variants

The second splicing variant is unfunctional due to the formation of cryptic splicing site at 38 base pairs immediately downstream of the original slice donor site in intron 3 which prevents this intronic sequence to be spliced out between exon 3 and exon 4. This insertion mutation makes a frame shift, which inactivates SKINT1L by omitting the SP (Figure 7A, B).

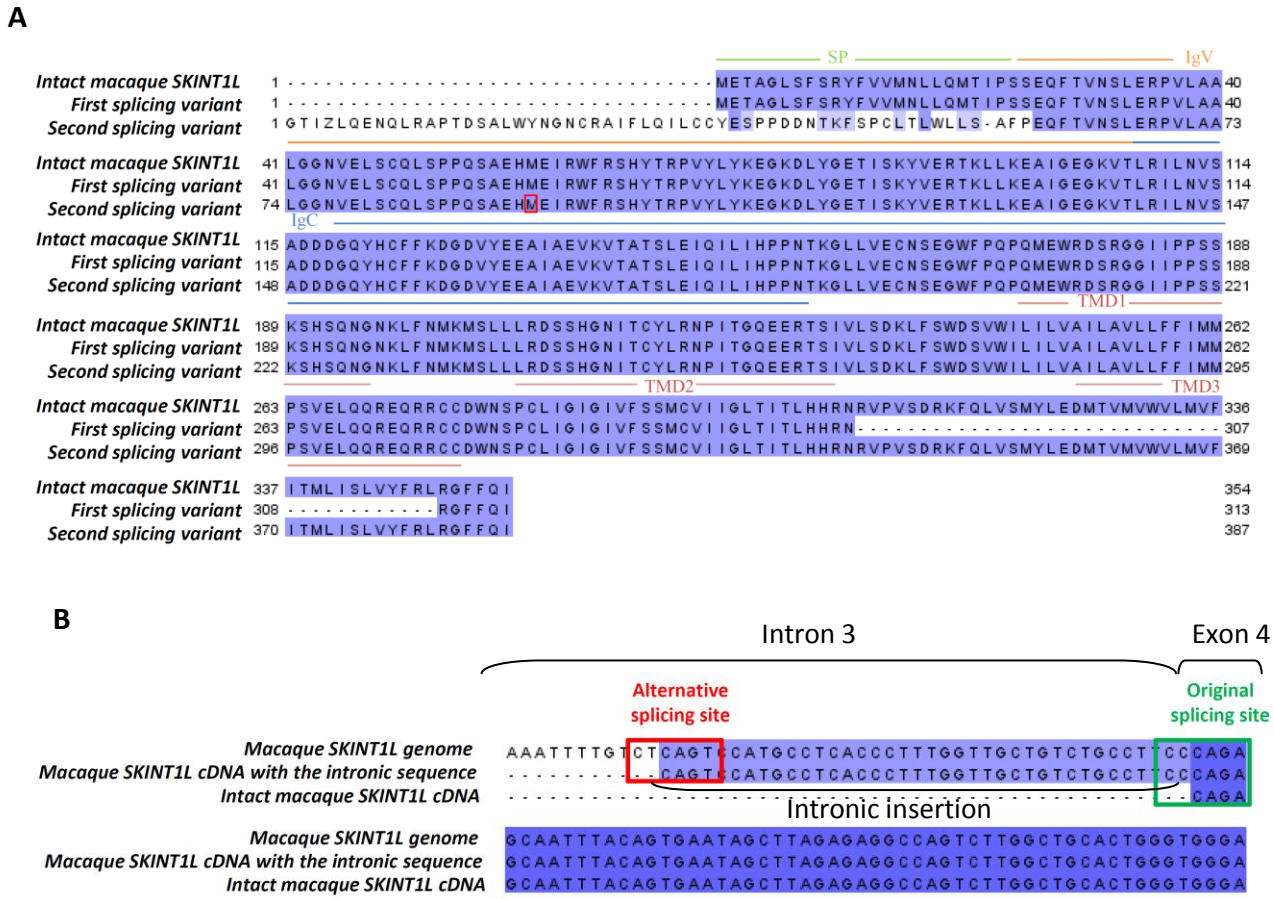


Figure 7: **Sequence alignment of Cynomolgus macaque SKINT1L and its splicing variants.** **(A)** Amino acid alignment of SKINT1L protein sequences of all detected splicing forms were aligned using the Clustal Omega program. Strictly and highly conserved residues are indicated in dark blue and light blue, respectively. The location of predicted domains is indicated on top of the sequences. **(B)** The alternative splicing site of second splicing variant. The brackets show the intron-exon boundaries and the inserted intronic sequence lengh is 38 bp.

Because the inserted intronic sequence is very small, the size of this transcript using primers for the entire SKINT1L sequence is indistinguishable from that of the major transcript. To examine the expression level of the second splicing variant, we just amplified the inserted sequence using primers flanking it. We found that this variant is expressed in an individually different level (corresponding to a higher band in Figure 8).

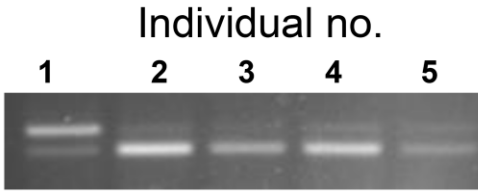


Figure 8: The *SKINT1L* second splicing variant expression: The *SKINT1L* second splicing variant expression in cynomolgus macaque skin was examined by reverse transcription PCR. Functional *SKINT1L* which doesn't have the intronic sequence is corresponding to the lower band and the unfunctional *SKINT1L*, containing the intronic sequence is represented in the upper band.

DETC-like cells in the cynomolgus macaque skin

It has been verified that *Skint1* is an essential component for development and homing of the canonical γδ TCR denderitic epidermal T cells (DETC) to the mouse epidermis[31,43]. To verify the functionality of cynomolgus macaque *SKINT1L*, we examined whether the resident T cell population of macaque skin contains invariant dendritic-shaped γδ T cells. Paraffin sections of the macaque skin were immunohistochemically stained with an Ab for human CD3, and this revealed that macaque skin harbors a high population of $CD3^+$ cells with a highly dendritic morphology (Figure 9A, C). On average, 4.2 ± 0.85 $CD3^+$ cells were detected per mm of the basement membrane. When judged by double staining, we examined this population of $CD3^+$ T cells using two antibody clones for human γδ TCR; clone 5A6.E91 which specifies for the human TCR-δ chain constant region and clone γ3.20 which specifies for the human TCR γ-chain constant region. Only clone γ3.20 could cross react with the crab eating monkey γδ T cells

(Figure 10 B, C) and stained approximately 41% of these $CD3^+$ cells (1.7 ± 0.35 TCR γ-chain+ cells per mm of the basement membrane) (Figure 8, panels B and C).

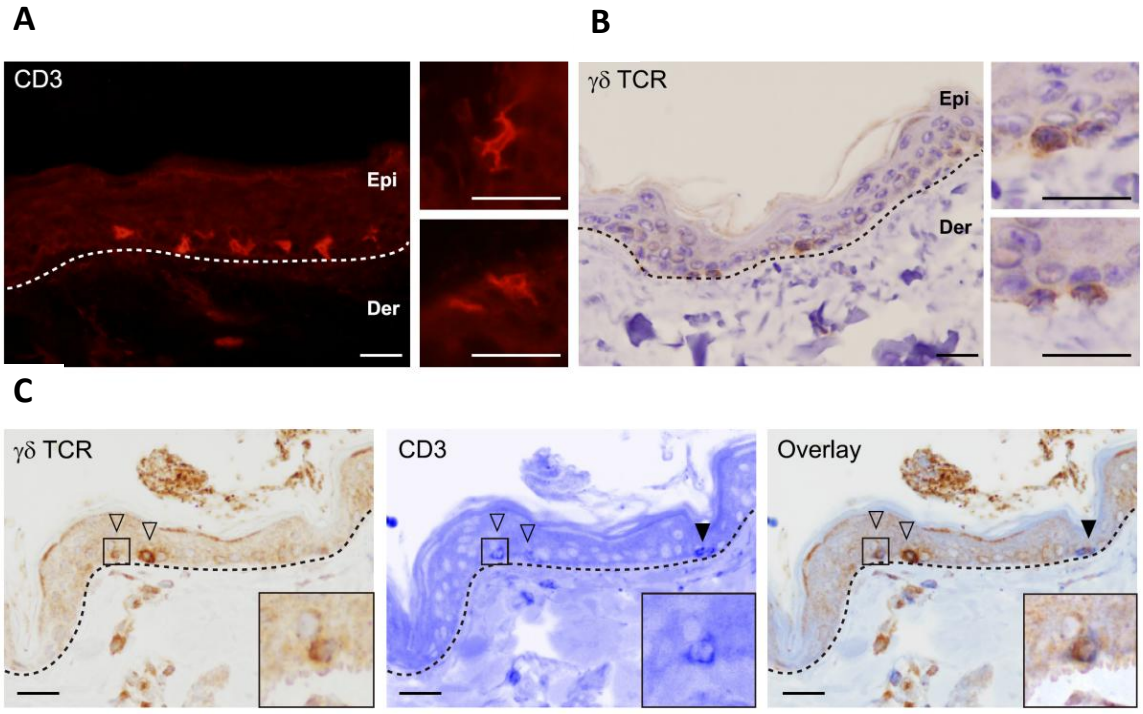


Figure 9: Localization of DETC-like cells in the Cynomolgus macaque: (**A-C**) The epidermis of the cynomolgus macaque contain dendritic-shaped CD3+, γδ TCR+ cells. Paraffin sections of adult macaque skin were stained with cross-reacting antibody for human CD3 (**A**), cross-reacting antibody for human γδ TCR (clone: γ3.20) (**B**) or with both antibodies (C). Insets show higher magnification images of CD3+ dendritic shaped cells (**A**), γδ TCR+ cells (**B**) and CD3/γδ TCR double-positive cells (**C**). Open and filled arrowheads indicate $CD3^+$ γδ TCR^+ cells and $CD3^+$ γδ TCR^- cells, respectively. Scale bar, 10 μm. Epi, epidermis; Der, dermis.

The dendritic shape has been also detected with other $CD3^+\gamma\delta^-$ T cells (Figure 9); this probably because the skin resident T lymphocytes put themselves into the cellular and extracellular matrix of the skin. We found also that, consistent with mouse DETC, the cynomolgus macaque DETC-like cells reside in the basal and suprabasal layer of the epidermis (Figure 10A, B).

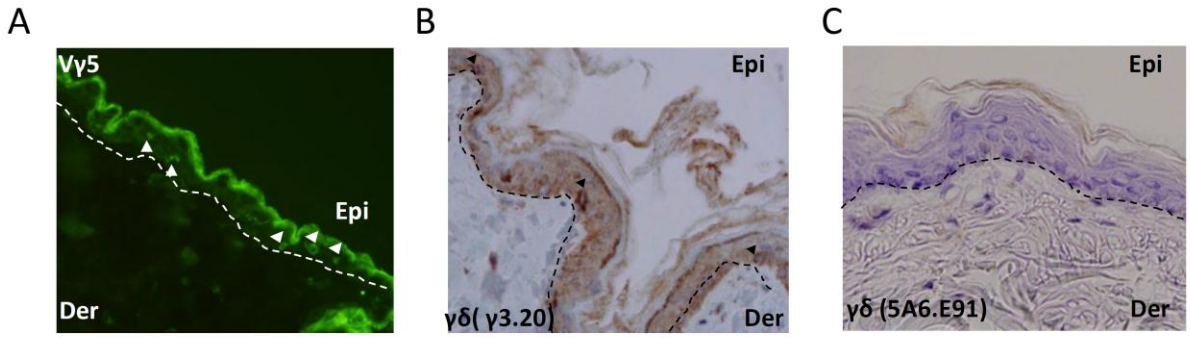


Figure 10: DETC-like cells in cynomolgus macaque populate the basal and suprabasal layer of the epidermis: Consistent with the mouse DETC (**A**), DETC-like cells in cynomolgus macaque populate the basal and suprabasal layer of the epidermis (**B**). Paraffin sections of mouse skin were stained with antibody for Vγ5 TCR (A), and adult macaque skin sections were stained with cross-reacting antibodys for human γδ TCR (clone: γ3.20, (**B**), clone: 5A6.E91, (**C**)). Arrowheads indicate γδ TCR^+ cells in macaque skin or and Vγ5+ T cells mouse. Epi, epidermis; Der, dermis.

It's noteworthy that the presence of *Skint1* is correlated with the presence of a restricted cutaneous T cell population[30,31]. To examine whether the $CD3^+$ cells described above express an invariant TCR, we first analyzed the genomic organization of cynomolgus macaque TCR γ- and δ-chain loci (Figure 11). The organization of the cynomolgus macaque TCR γ- chain locus was very similar to that of the rhesus macaque TCR γ-chain locus[44], and contained six functional Vγ gene segments located on chromosome 3: Vγ1, Vγ2, Vγ3, Vγ2/4, Vγ9 and Vγ10 (genes named according to ref[44]). On the other hand, the TCR δ-chain locus located on chromosome 7 contained three Vδ gene segments: Vδ1, Vδ2 and Vδ3 (Figure 11). The sequences of the functional Vγ and Vδ segments are shown in Figure 12 (A,B).

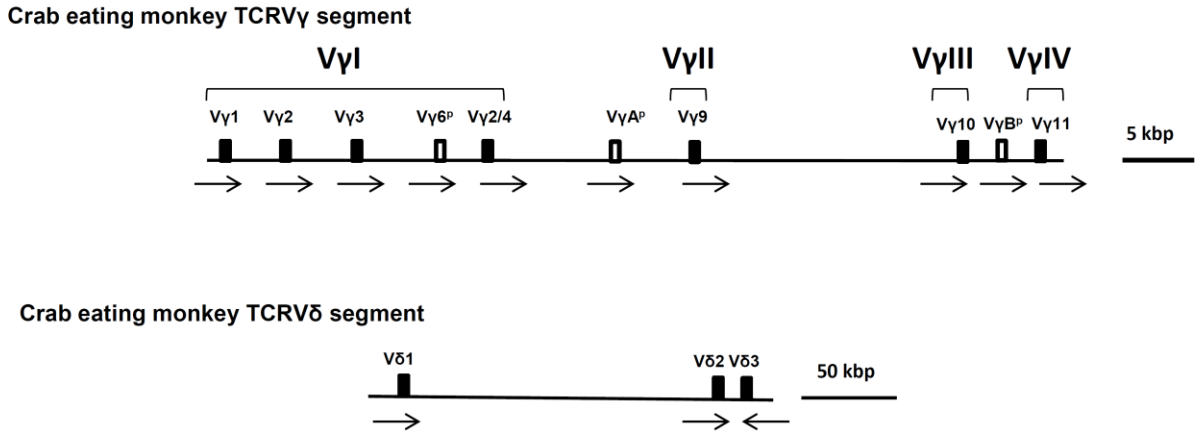


Figure 11: Genomic organization of cynomolgus macaque TCR Vγ and Vδ loci: filled and open boxes indicate functional and non-functional gene segments, respectively. Numbers given in roman numerals over the blankets indicate the subgroups of Vγ gene segments defined by Huck *et al*. 1988.

Vδ1

```
  1  CAGAGAGCTA CATGCCACAT GCTGTTCTCC AGCTTGCTGT GCATATTTGT GGCCTTCAGC
 61  TACTCTGGGG ATTCAGTGTG GCCCAGAAGG TTACTCAAGC CCAGTCATCA GTATCCATGC
121  CAGTGGAGAA AGCAGTCACC CTGAACTGCC AGTATGAAAC AAGTTCGTGG TCATATGACC
181  TTTTTTGGTA CAAGCAACTT CCCGGCAAAG AGATGATTTT CCTTATTCAC CAGGGTTCTT
241  CTCAACAGAA TGCAAGAAAT GGTCGCTATT CTGTCAACTT CCAGAAAGCA GCTAGCTCCA
301  TCACCTTAAC CATTTCAGCC TTACAGCTAG AAGACTCAGC AACATACTTC TGTGCTCTCC
361  GGGAACT
```

Vδ2

```
  1  ATTAACAGGC AGCAGGCCCT GTGAGCCTCC TGTGATCAGC TAGGCTTCTC TGAGGAATGA
 61  AACAGCTGCT GATAGGGACT GACTTCCTCA GTCACAAGTG TTTCACAAGT GTGTTGGTTC
121  AGAGGAGGGA CCAGGCAGAA GGTGGTTCAC AGGCAGAGCT GCCCCTGAGT GAGCCATGCA
181  GAGGATCTCC TCCCTCATCC ATCTCTCCCT CTTCTGGGCA GGAGTCATGT CAGCTGTTGA
241  GTTGGTGCCT GAACACCAAA CAGTGATTGT GTCAGTGGGG GACCCTGCCA CCCTCAAGTG
301  CTCCATGAAA GGAGAAGCAA TCAGTAACTA CTATATCAAC TGGTACAGGA AGACCCAAGG
361  TAACACAATG ACTTTCATAT ACCGAGAAAA GGGCATCTAT GGCCCTGGTT TCAAGGACAA
421  TTTCCAAGGT GACATTGATA CTGAAGAGAA CCAGGCTGTA CTTAAGATCC TTGCACCATC
481  AGAGAGAGAT GAAGGGTCTT ACTACTGTGC CAGTGACA
```

Vδ3

```
  1  TCTTTTGTTA GAACCGTCAG ATAGTCTCTG AGATCTTAGG AACCCTGTTT TGTTGTATAT
 61  CCTCTGCCCT GGTATGAAGA AAAGATGATT CTTACTGTGG GCTTTAGCTT TTTGTTTTTC
121  TACAGGGGCA TGCTGTGTGA CAAAGTAACT CAGAGTTCCC CGGACCAGAT GGTGGCGAGT
181  GGCAGTGAGG TGGTACTGCT CTGCACTTAC GACAGTACAT ATTCAGATCC AGGTTTATTC
241  TGGTACCGCA TAAGGCCAGA TTATAATTTT CAGTTTGTCT TCTATGGGGA TAAAAGCAGA
301  TCTCACGGTT CATATTTTGC TAAAGGACGG TTTTCTGTGA AGCACCTTCT GACCCAGAAA
361  GCCTTTCACT TGGTAATCTC TCCAGTGAGT ACTGAAGACA GTGCCACTTA CTACTGTGCC
```

Figure 12 (A):cDNA sequences of the functional Vδ segments of Cynomolgus macaque.



Figure 12 (B): cDNA sequences of the functional Vγ segments of Cynomolgus macaque.

To determine Vγ/Vδ gene usage in cynomolgus macaque epidermal T cells, we prepared primers specific for each Vγ or Vδ gene segment. Reverse transcription PCR analysis using these primer pairs showed that macaque epidermal T cells predominantly express Vγ10Vδ1 TCR (Figure 13). Because Vδ gene segments other than Vδ1-3 are used as Vα gene segments[45], we amplified only Vδ1-3 gene segments. Interestingly, we found that as in human and mouse epidermis[34,46], Vδ1 is preferentially expressed in the macaque epidermis. These results indicate that cynomolgus macaque epidermal γδ T cells express an invariant TCR, consistent with the existence of a structurally intact *SKINT1L* gene. The examination of the Vγ/Vδ expression in the whole skin revealed that there is expression of other Vγ and Vδ gene segments detected only at low levels. This suggested the different usage of Vγ/Vδ gene usage in cynomolgus macaque dermis.

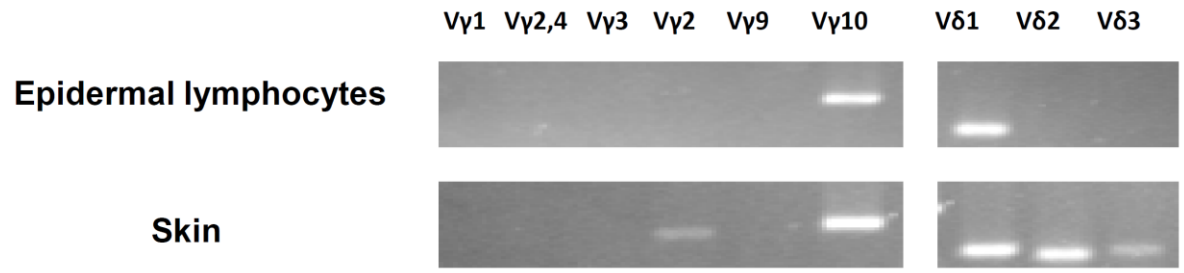


Figure 13: Expression of cynomolgus macaque TCR Vγ and Vδ gene segments in the whole skin and the epidermal T cells. Reverse transcription PCR was conducted using primers specific for each TCR Vγ or Vδ gene segment.

Evolution of the *SKINT* gene family in mammals

Consistent with what has been observed previously[31,32], *SKINT1L* went through different evolutionary events such as duplication within rodents and pseudogenization within primates. For a better understanding of the *Skint1/SKINT1L* gene family evolution, and more generally the entire *SKINT* gene family, an extensive screening using the mouse SKINT protein sequences as query was performed to identify the *SKINT* family members

of representative mammals. A total of 47 species representing 22 orders, for which genome sequences are available, were subjected to analysis. As shown previously[32], phylogenetic analysis using mouse *SKINT* members protein sequences as query and based on the available amino acid sequences including IgV and IgC domains indicates that the SKINT protein family falls into three major subfamilies: SKINT1, SKINT7 and SKINT9 (Figure 14). SKINT2, 3, 4, 5 and 6 proteins are the members of the SKINT1 subfamily; they all appear to have emerged by rodent-specific gene duplication from the *Skint1* gene. Gene duplication is more extensive in mice than in rats, with mice having six and rats having three SKINT1 subfamily members. SKINT8 protein is closely related to SKINT7 protein and seems to have emerged by duplication from the *Skint7* gene in the mouse lineage. SKINT10 and 11 proteins are the members of the SKINT9 subfamily and appear to have emerged by rodent-specific gene duplication.

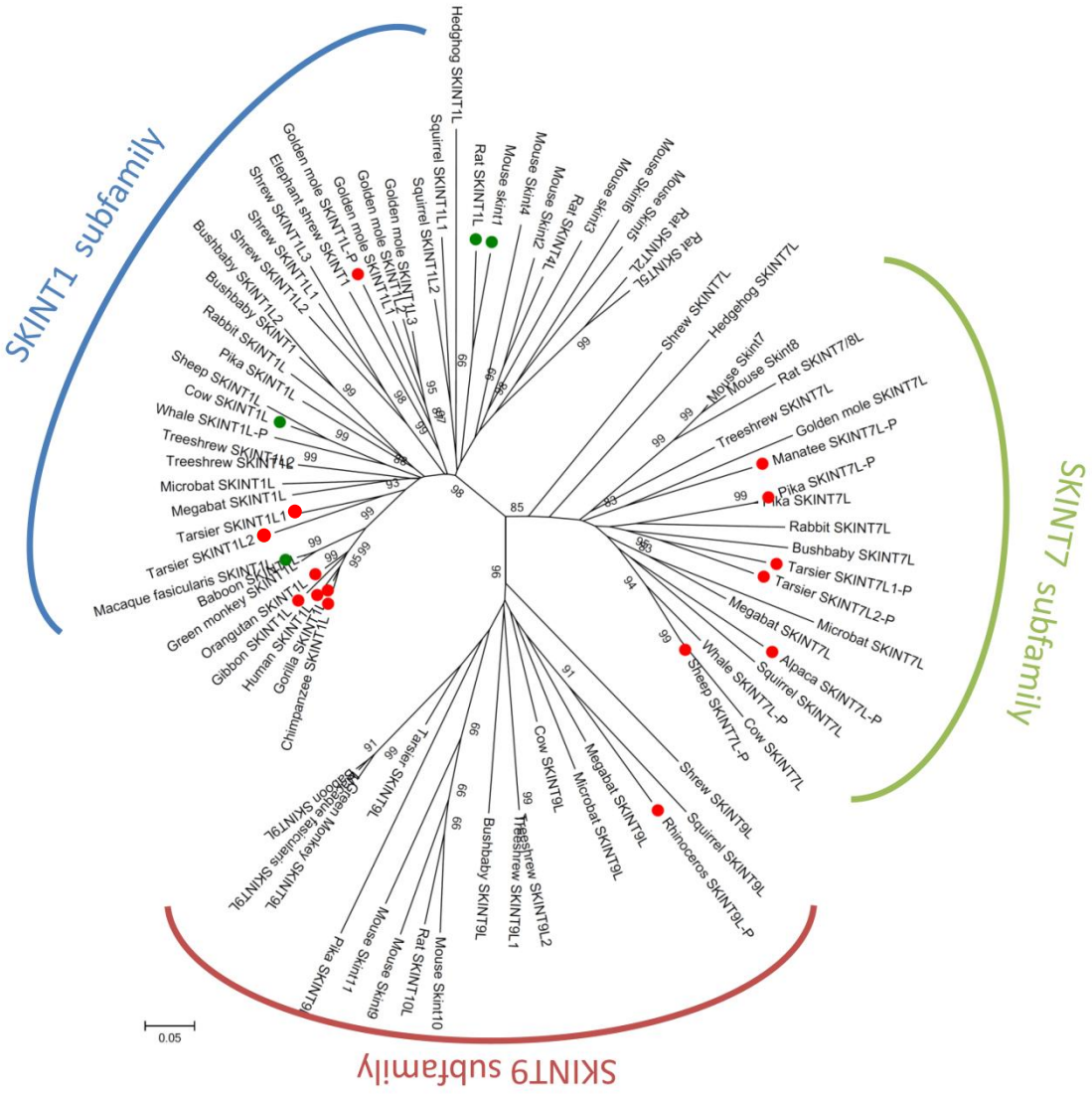


Figure 14. Phylogenetic tree of the SKINT family: The neighbor-joining tree was constructed based on deduced amino acid sequences as described in Materials and methods. The bootstrap confidence values over 80 only are shown (1,000 replications). Genes judged to be non-functional are indicated in red dots. Genes proven to be functional or likely to be functional based on the presence of DETCs expressing an invariant (or restricted) TCRs are labeled by green dots. Genes with no obvious inactivating mutations are not labeled. Accession numbers of sequences and other relevant information are given in Table 3.

Mammals phylogenetic tree of SKINT family indicated that SKINT sequences were detected only in eutherian (placental mammals) and neither marsupials nor monotremes had SKINT-like sequences (Figure 15). Furthermore, we were unable to identify SKINT-like sequences in non-mammalian species (data not shown). Therefore, the SKINT family has presumably emerged in a common ancestor of placental mammals. The distribution of SKINT subfamily proteins across mammalian orders indicates that an ancestor of placental mammals had at least SKINT1 and SKINT7 subfamily proteins and that a boreoeutherian ancestor had all three subfamily proteins. A striking feature of the SKINT family is that its members went through different evolutionary events in boreoeutheria. SKINT members were highly duplicated in rodents and also in small number of species including prosimians, golden moles and tree shrews with less copy number. Interestingly, SKINT family was lost or rendered nonfunctional in many mammalian orders. Indeed, some mammals such as carnivorans and NWMs appear to have lost the SKINT family in its entirety.

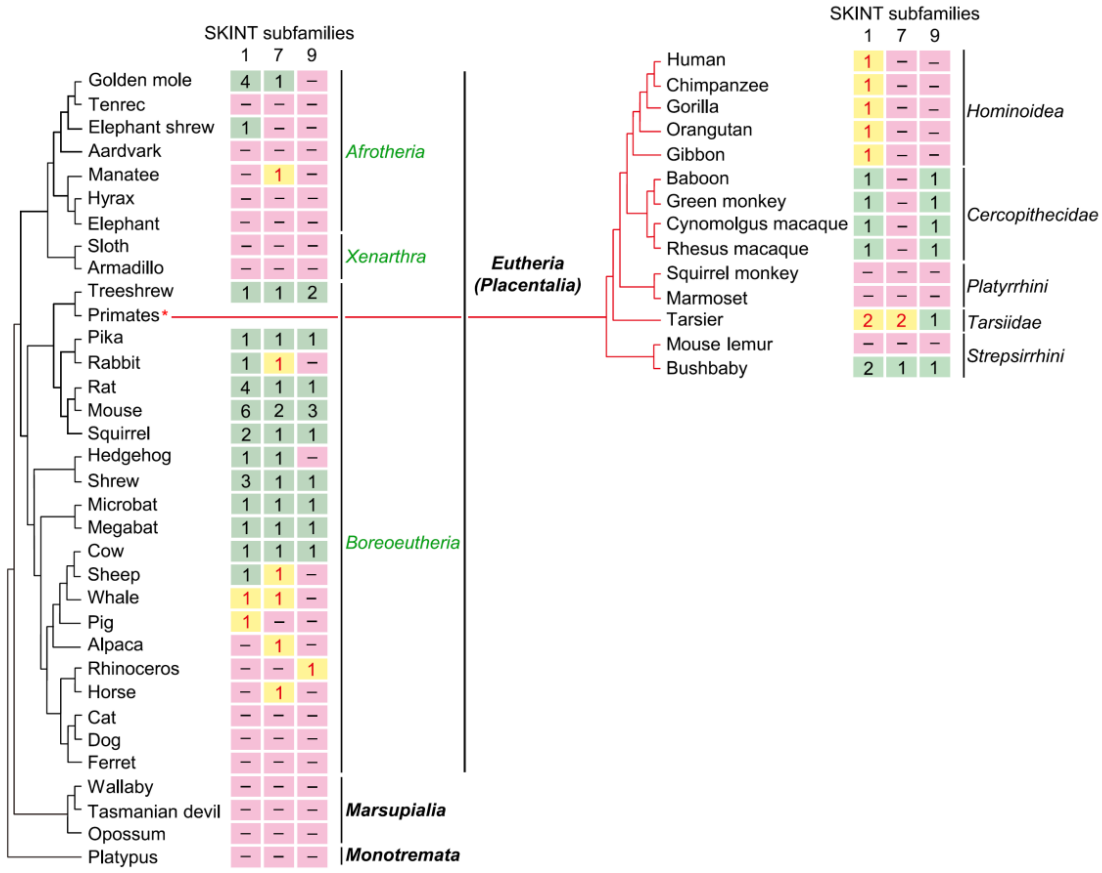


Figure 15. Distribution of SKINT subfamily proteins across mammals. The phylogenetic relationship of mammals shown here is based on Murphy *et al*. and Song *et al*.[47,48]. The copy number of *SKINT* genes belonging to each subfamily is shown for 47 mammalian species representing 22 orders. The copy number of functional and non-functional genes is indicated in black and red, respectively. The distribution of SKINT subfamily proteins in primates is shown separately on the right.

Table 3: Accession numbers of mammalian SKINT family genes

*Species are arranged in the order as they appear in Figure 15.

				Region		
Species	Genes		Accession number	From	To	Database
Golden mole	*SKINT1L*	(Pseudogene)	NW_006408540.1	27,366,667	27,367,369	NCBI genome
	SKINT1L1		XP_006839800.1			NCBI protein
	SKINT1L2		AMDV01072344.1	8,227	3,405	NCBI genome
	SKINT1L3		AMDV01072359.1	10,114	4,656	NCBI genome
	SKINT1L4		XP_006839919.1			NCBI protein
	SKINT7L		XP_006839918.1			NCBI protein
Elephant shrew	*SKINT1L*		XP_006879863.1			NCBI protein
Manatee	*SKINT8L*	(Pseudogene)	NW_004443951.1	7,824,141	7,831,380	NCBI genome
Treeshrew	*SKINT1L*		XP_006147332.1			NCBI protein
	SKINT7L		XP_006147331.1			NCBI protein
	SKINT9L1		ELW67061.1			NCBI protein
	SKINT9L2		XP_006147332.1			NCBI protein
Human	*SKINT1L*	(Pseudogene)	NC_000001	48,101,715	48,182,428	NCBI genome
Chimpanzee	*SKINT1L*	(Pseudogene)	NC_006468.3	48,499,761	48,548,233	NCBI genome
Gorilla	*SKINT1L*	(Pseudogene)	NC_018424.1	49,735,301	49,819,734	NCBI genome
Orangutan	*SKINT1L*	(Pseudogene)	NC_012591.1	181,680,692	181,734,447	NCBI genome
Gibbon	*SKINT1L*	(Pseudogene)	NC_019827.1	17,090,514	17,112,139	NCBI genome
Olive baboon	*SKINT1L*		NC_018152.1	50,394,763	50,437,695	NCBI genome
	SKINT9L		NC_018152.1	50,319,438	50,349,456	NCBI genome
Green monkey	*SKINT1L*		NC_023661.1	84,640,177	84,686,975	NCBI genome
	SKINT9L		XP_007977024.1			NCBI protein
Cynomolgus macaque	*SKINT1L*		AEHL01229074.1	11,817	27,285	NCBI genome
			AQIA01004292.1	5,930	35,479	NCBI genome
	SKINT9L		XP_005543503.1			NCBI protein

Rh. macaque	*SKINT1L*		AEHK01151475.1	1,172	3,795	NCBI genome
			AEHK01151471.1	378	728	NCBI genome
			AEHK01151469.1	2,098	2,382	NCBI genome
			AEHK01151468.1	1,705	4,053	NCBI genome
			AANU01266778.1	5306	9,035	NCBI genome
	SKINT9L		AANU01266770.1	8,733	1,670	NCBI genome
Tarsier	*SKINT1L1*	(Pseudogene)	ABRT02125005.1	11,139	25,919	NCBI genome
	SKINT1L2	(Pseudogene)	ABRT02314780.1	143,050	127,400	NCBI genome
	SKINT7L1	(Pseudogene)	NW_007248563.1	23,884	24,875	NCBI genome
	SKINT7L2	(Pseudogene)	NW_007259162.1	30,102	31,097	NCBI genome
	SKINT9L		Scaffold 8522	11,872	11,591	Ensembl
Bushbaby	*SKINT1L1*		GL873729.1	2,367,236	2,386,348	NCBI genome
	SKINT1L2		XP_003803821.1			NCBI protein
	SKINT7L		XP_003803839.1			NCBI protein
	SKINT9L		XP_003801490.1			NCBI protein
Pika	*SKINT1L*		ALIT01108927.1	2,900	5,400	NCBI genome
	SKINT7L	(Pseudogene)	NW_004535555.1	1,418,868	1,419,855	NCBI genome
	SKINT7L		AAYZ01523420.1|	970	3	NCBI genome
	SKINT9L		XP_004598552.1			NCBI protein
Rabbit	*SKINT1L*		XP_008263849.1			NCBI protein
	SKINT7L	(Pseudogene)	GBCH01073656.1			NCBI TSA
Rat	*SKINT1*		NP_001129388.1			NCBI protein
	SKINT2		XP_002726638.1			NCBI protein
	SKINT5L		XP_006225576.1			NCBI protein
			XP_002729563.3			NCBI protein
	SKINT5/6		XP_006225577.1			NCBI protein
	SKINT7		XP_002726637.2			NCBI protein
	SKINT9		EDL90334.1			NCBI protein
Mouse	*SKINT1*		ABS30712.1			NCBI protein

	SKINT2		NP_001272892.1			NCBI protein
	SKINT3		NP_001095944.1			NCBI protein
	SKINT4		NP_848901.2			NCBI protein
	SKINT5		NP_001096669.1			NCBI protein
	SKINT6		NP_001161348.1			NCBI protein
	SKINT7		XP 006503236.1			NCBI protein
	SKINT8		ABS30725.1			NCBI protein
	SKINT9		NP_808532.1			NCBI protein
	SKINT10		ABS30727.1			NCBI protein
	SKINT11		XP_006543938.1			NCBI protein
Squirrel	*SKINT1L1*		XP_005318191.1			NCBI protein
	SKINT1L2		XP_005318193.1			NCBI protein
	SKINT7L		XP_005318192.1			NCBI protein
	SKINT9L		XP_005318190.1			NCBI protein
Hedgehog	*SKINT1L*		XP_007521449.1			NCBI protein
	SKINT7L		XP_007521448.1			NCBI protein
Shrew	*SKINT1L1*		XP_004607401.1			NCBI protein
	SKINT1L2		XP_004607402.1			NCBI protein
	SKINT1L3		XP_004607400.1			NCBI protein
	SKINT7L		XP_004622390.1			NCBI protein
	SKINT9L		XP_004620605.1			NCBI protein
Microbat	*SKINT1L*		XP_006107913.1			NCBI protein
	SKINT7L		XP_006087439.1			NCBI protein
	SKINT9L		XP_006087438.1			NCBI protein
Megabat	*SKINT1L*		Scaffold 17784	8,267	3,354	Ensembl
	SKINT7L		Scaffold 11943	13,946	14,909	Ensembl
	SKINT9L		Scaffold 19546	7,735	3,366	Ensembl
Cow	*SKINT1L*		NP_001120788.1			NCBI protein
	SKINT7L		DAA31132.1			NCBI protein
			XP 002704034.3			NCBI protein

	SKINT9L		XP_005197953.1			NCBI protein
Sheep	*SKINT1L*		XP 004003585.1			NCBI protein
	SKINT7L	(Pseudogene)	AMGL01001340.1			NCBI protein
Whale	*SKINT1L*	(Pseudogene)	AWZP01078114.1	9,847	19,328	NCBI genome
	SKINT7L	(Pseudogene)	NW_006717733.1	5,059	4,084	NCBI genome
Pig	*SKINT1L*	(Pseudogene)	NW_003610557.1	74,623	76,830	NCBI genome
Alpaca	*SKINT7L*	(Pseudogene)	NW_005882724.1	3,630,483	3,630,215	NCBI genome
Rhinoceros	*SKINT9L*	(Pseudogene)	NW_004454229.1	1,037,412	1,042,202	NCBI genome
Horse	*SKINT7L*	(Pseudogene)	NW_001867402.1	10,495,575	10,495,306	NCBI genome

Discussion

Up to date, no extensive study cleared *SKINT1L* gene evolution within mammals, although *Skint1/SKINT1L* gene is appeared to be highly evolved from rodents to human[31,32]. We provided here a direct evidence that all of the hominoids we have studied including humans, great apes and lesser apes have a single copy of *SKINT1L* gene which is inactivated by a common termination codon and inconsistent with that, *SKINT1L* gene in the Old World Monkeys (OWM) such as olive baboons, green monkeys, cynomolgus macaques and rhesus macaques is structurally active (Figure 1). We have chosen the cynomolgus macaque as a representative for the OWMs and found that, the mRNA of *SKINT1L* was expressed exclusively in the thymus and skin in a similar pattern to its mouse counterpart (Figure 5). We also found a high population of DETC-like cells residing the basal and suprabasal layers of the macaque epidermis and predominantly expressing Vγ10Vδ1 TCR (Figure 13). These observations indicate that the macaque epidermis contains authentic DETCs expressing γδ TCR of limited diversity and suggest that *SKINT1L* presumably has a role in homing of these cells in the OWM epidermis. Therefore, we conclude that *SKINT1L* gene has been inactivated in the hominoid ancestor after the radiation of the OWM and seemingly DETCs expressing an invariant TCR were lost in consequence.

Interestingly, we showed here that *SKINT1L* is represented variably in the majority of primates from great apes to prosimians as structurally active in OWMs, inactivated in hominoids and lost completely in NWMs. Here, we noticed that NWMs genome of marmosets and squirrel monkeys does not have a *SKINT1L* gene sequence. Although the squirrel monkey genome has been sequenced only to low coverage, the marmoset genome is sequenced to high coverage. Thus, it seems likely that the *SKINT1L* gene was lost in marmosets. Our results demonstrated that *SKINT1L* gene is highly evolved even in the same group of species within primates, as in prosimian group, *SKINT1L* gene in bushbabies is structurally intact but is inactivated in tarsiers. Notably, The phylogeny of primates *SKINT1L* formed a clear cluster, however, some primates like the tarsier and the bushbaby

branched out of the primates with a low bootstrap confidence value (Figure 15). Taken together, we conclude that *SKINT1L*was lost at least twice in primate evolution, once by the gene elimination in some NWMs, and then by gene inactivation in some prosimians and hominoid ancestor.

In mammals, by analyzing the phylogeny of SKINT1L protein (Figure 14), we showed by evidence that *SKINT1L* gene emerged in the common ancestor of placental mammals (Figure 15). Also, we presented that, phylogenetically all *Skint1* family members in mice formed one cluster where they are closely related to each other more than to their orthologs which is most probably because the speciation of mice predates the duplication of their *Skint1* gene what makes the number of *skint1* gene paralogs in mice more than in rats, that made the other *Skint1* paralogs act as a co-orthologs to a single-copy gene from the other species and do not represent the true ortholog (Figure 14). However, we confirmed here that the structural similarities and the cellular context in which the protein works can be used as a predictor to the functional orthology[49-53]. We observed that *SKINT1L* gene has the structural and the cellular context equivalent to mouse *Skint1* gene and might have the same impact on DETC homing to the epidermis (Figure 1, 5 and 9), suggesting that the cynomolgus macaque *SKINT1L* gene is the mouse *Skint1* ortholog and this orthology will be most probably true for the *SKINT1L* genes in the other orders of mammals. We demonstrated also that despite the emergence of *SKINT1L* in the placental mammal ancestor, it is lost in some mammalian orders such as xenarthra and carnivora (Figure 15). Because we suggest the functional orthology of the other mammalian *SKINT1L* to mouse *Skint1*, we propose that the DETCs expressing invariant TCR emanated in the ancestor of mammals and lost in some mammalian orders concomitant to the loss of *SKINT1L*. Besides, in common with the presence of 5 splicing variants of mouse *skint1*, we showed that macaque *SKINT1L* gene has 2 splicing variants, one of them has only 2 TMD and the other one is inactive as a result of partial retention of intron 3 (Figure 5, 8); both of these variants mRNA are expressed in all macaque individuals committed in our study. As reported before, the mouse Skint1/Skint2 domain swap chimeras and Skint1 constructs including only one TMD could not retain the Skint1 function in the thymus, indicating that

each Skint1 domain is not redundant[54]. Thus, we assume that both of cynomolgus macaque splicing variants are non-functional because of their domain loss.

Previous studies concluded that the cow epidermis expresses structurally intact *SKINT1L* gene with high level of γδ T cells (44% of epidermal $CD3^+$ T cells) predominantly expressing Vγ3, Vγ7, Jγ5, Cγ5 and Vδ sequences belonging to the Vδ1 family[55-58]. Likewise, in our study we presented that the macaque epidermis has a high number of predominant Vγ10Vδ1 DETC comparable to the number of the epidermal $\gamma\delta^+$T cells found in cow, although the number of $CD3^+\gamma\delta^+$ T cells detected in our experiments retained as a minimal estimate due to lack of specific antibodies for the macaque $CD3^+\gamma\delta^+$ T cells. This result suggests that, while γδ T cells which form 95% of the mouse epidermal T cell population play a unique role in its skin immunosurveillance[11,59], they might contribute with the other T-cells in the macaque and cow for skin homeostasis. Therefore, what remained as unresolved puzzle is the immunological solution in the orders which lost or had inactivated *SKINT1* gene. It is possible that species that lost functional *SKINT1L* have developed compensatory mechanisms for skin immunosurveillance. In humans which are all have non-functional SKINT1L molecules, CD1a molecules expressed by epidermal Langerhans cells play a crucial role for activation of Th22 cells for enhancing wound repair[60]. Because rodents with functional SKINT1L molecules lack CD1a molecules and humans without functional SKINT1L molecules have CD1a molecules, it was suggested that humans and mice have evolved different strategies for immune defense in the skin[61]. However, the presence of SKINT1L and that of CD1a are not always mutually exclusive; for example, cattle have both CD1a[62] and SKINT1L (Figure 1). Therefore, some species may employ both CD1a-mediated and DETC-mediated defense strategies. Furthermore, animals without DETCs expressing an invariant TCR may have other compensatory mechanisms independent of CD1a. In this case, we shall not be able to observe a clear correlation between the absence of functional SKINT1L and the presence of CD1a.

Butyrophilin (*BTN)* gene family belongs to the immunoglobulin subfamily of the transmembrane proteins and shares structural homology with *SKINTL* family members and

is considered as the most closely related family to *Skint* gene family, although the mammalian *BTN* gene emerged in the genome of the therian ancestor, then it underwent multiple duplication before the separation of the marsupials and eutherians[32,63], we found that *SKINTL* appears to be emerged only in the eutherians, then, it went through different evolutionary events inside the boreoeutherians. According to that, *SKINTL* gene seems to be lost more than 2 times in the mammalian phylogeny (Figure 15). Indeed, only a limited number of species retain all of the three *SKINTL* gene subfamilies and some mammals such as carnivorans completely lost *SKINTL* gene family. These observations indicate that the members of the *SKINT* gene family are generally dispensable and that they evolve in a highly order-specific or species-specific manner. Functional characterization of SKINT7L and SKINT9L subfamilies might help us understand why they are lost in some mammalian orders and whether animals without SKINT7L or SKINT9L have any compensatory mechanisms.

Summary and Conclusion

Skint1 gene is expressed specifically in mouse keratinocytes and thymic epithelial cells, suggesting an indispensable role for *Skint1* in the selection machinery for Vγ5Vδ1 DETC, which has an essential role in skin immunosurveillance. Phylogenetically, rodents have functional SKINT1 molecules, but humans and chimpanzees have a *SKINT1-like* (*SKINT1L*) gene with multiple inactivating mutations, urging to determine at which stage in the mammalian phylogeny *SKINT1L* lost its function.

In this study, we conclude that although *SKINT1L* is psuedogenized in the hominoid lineage, it is still structurally intact in the OWM and expressed exclusively in skin and thymus. Moreover, we demonstrated here that the epidermis of crab eating monkey contains a population of dendritic-shaped γδ T cells (DETC- like cells) expressing invariant Vγ10Vδ1 TCR. These observations indicate that the cynomolgus macaque *SKINT1L* is the functional orthology to the mouse *skint1* gene.

By the extensive bioinformatic analysis of several mammalian species, we showed that *SKINT* family has been emerged in an ancestor of eutheria and lost or inactivated multiple times in the mammalian phylogeny which suggest a concomitant loss of the skin-resident γδ T cells in the orders lacking *SKINT1L*.

Acknowledgement

Now at the end of my thesis and my current journey in Japan which was full of good and hard times, I would like to thank everyone was eager to support me approaching my aims in this four year road.

First of all, it is with immense gratitude that I acknowledge the support and help of my Prof. Masanori Kasahara who despite his busy time, always found time for discussion and opening the way for me to learn of his valuable knowledge, he gave me the guidance to be a good researcher and taught me what is the meaning of dedication to work.

Dr. Noriyuki Otsuka who guided me to solve many research problems eagerly and was always here for me whenever I need a support in my scientific research. I owe my deepest appreciation to Dr. Yukiko Miyatake who helped me honestly since my foot stepped inside Japan, not only in my research, but also in every personal daily issue. Dr. Yoichi Sutoh who supported me strongly from the beginning till the end of my research. This research wouldn't appear in this shape without his help.

My parents and my sister, I really can't find the words to express my gratitude to their love, support and encouragement. Despite the long distance separating us, I didn't feel in a moment they are away, they were always with me. They embrace me and my dreams all my life long and believe in me able to make my dreams come true.

And very special thanks to Dr. Utano Tumaro and Dr. Yushida Shigaro who I really feel grateful to their help and guidance through my work life here. I am also indebted to all my lab colleagues who helped me whenever I need their help without any hesitation.

Finally, I need to give special gratitude to all my friends in Egypt and every friend I have met here in Japan. I want to tell them every one of you left an unforgettable print which will keep with me forever. I need to mention thankfully in Japan, Aiman Zidan and Jumana Al-Mallahy and in Egypt, Manal Mohsen and Mira Refaat.

I want to give special thanks for the Egyptian Ministry of Higher Education and Ain Shams University which gave me the chance of my life to come here to Japan and make a fortune of knowledge which will bring to my home new technologies and fresh experience in rapidly developing area.

References

1 Holderness, J., Hedges, J. F., Ramstead, A. & Jutila, M. A. Comparative biology of γδ T cell function in humans, mice, and domestic animals. *Annu. Rev. Anim. Biosci.* **1**, 99-124 (2013).

2 Girardi, M. Immunosurveillance and immunoregulation by γδ T cells. *J. Invest. dermatol.* **126**, 25-31 (2006).

3 Heilig, J. S. & Tonegawa, S. Diversity of murine gamma genes and expression in fetal and adult T lymphocytes. *Nature* **322**, 836-840 (1986).

4 Bonneville, M., O'Brien, R. L. & Born, W. K. Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. *Nat. Rev. Immunol.* **10**, 467-478 (2010).

5 Itohara, S. *et al.* Homing of a gamma delta thymocyte subset with homogeneous T-cell receptors to mucosal epithelia. *Nature* **343**, 754-757 (1990).

6 Carding, S. R. & Egan, P. J. γδ T cells: functional plasticity and heterogeneity. *Nat. Rev. Immunol.* **2**, 336-345 (2002).

7 Gao, Y. *et al.* Gamma delta T cells provide an early source of interferon gamma in tumor immunity. *J. Exp. Med.* **198**, 433-442 (2003).

8 von Lilienfeld-Toal, M. *et al.* Activated gammadelta T cells express the natural cytotoxicity receptor natural killer p 44 and show cytotoxic activity against myeloma cells. *Clin. Exp. Immunol.* **144**, 528-533 (2006).

9 Jameson, J. *et al.* A role for skin gammadelta T cells in wound repair. *Science* **296**, 747-749 (2002).

10 Brandes, M., Willimann, K. & Moser, B. Professional antigen-presentation function by human gammadelta T cells. *Science* **309**, 264-268 (2005).

11 Hayday, A. C. Gammadelta T cells and the lymphoid stress-surveillance response. *Immunity* **31**, 184-196 (2009).

12 Holtmeier, W. & Kabelitz, D. gammadelta T cells link innate and adaptive immune responses. *Chem. Immunol. Allergy* **86**, 151-183 (2005).

13 Vantourout, P. & Hayday, A. Six-of-the-best: unique contributions of γδ T cells to immunology. *Nat. Rev. Immunol.* **13**, 88-100 (2013).

14 Yoshida, S. *et al.* Involvement of an NKG2D ligand H60c in epidermal dendritic T cell-mediated wound repair. *J. Immunol.* **188**, 3972-3979 (2012).

15 Nedellec, S., Sabourin, C., Bonneville, M. & Scotet, E. NKG2D costimulates human V gamma 9V delta 2 T cell antitumor cytotoxicity through protein kinase C theta-dependent modulation of early TCR-induced calcium and transduction signals. *J. Immunol.* **185**, 55-63 (2010).

16 Strid, J., Sobolev, O., Zafirova, B., Polic, B. & Hayday, A. The intraepithelial T cell response to NKG2D-ligands links lymphoid stress surveillance to atopy. *Science* **334**, 1293-1297 (2011).

17 Havran, W. L. & Allison, J. P. Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. *Nature* **335**, 443-445 (1988).

18 Havran, W. L. & Jameson, J. M. Epidermal T cells and wound healing. *J. Immunol.* **184**, 5423-5428 (2010).

19 Asarnow, D. M. *et al.* Limited diversity of γδ antigen receptor genes of thy-1^+ dendritic epidermal cells. *Cell* **55**, 837-847 (1988).

20 Chodaczek, G., Papanna, V., Zal, M. A. & Zal, T. Body-barrier surveillance by epidermal γδ TCRs. *Nat. Immunol.* **13**, 272-282 (2012).

21 Hayday, A. C. γδ T cells and the lymphoid stress-surveillance response. *Immunity* **31**, 184-196 (2009).

22 Witherden, D. A. & Havran, W. L. Molecular aspects of epithelial γδ T cell regulation. *Trends Immunol.* **32**, 265-271 (2011).

23 Komori, H. K. *et al.* Cutting edge: dendritic epidermal γδ T cell ligands are rapidly and locally expressed by keratinocytes following cutaneous wounding. *J. Immunol* **188**, 2972-2976 (2012).

24 Hayday, A. C. γδ cells: a right time and a right place for a conserved third way of protection. *Annu. Rev. Immunol.* **18**, 975-1026 (2000).

25 Girardi, M. *et al.* Regulation of cutaneous malignancy by gammadelta T cells. *Science* **294**, 605-609 (2001).

26 Toulon, A. *et al.* A role for human skin–resident T cells in wound healing. *J. Exp. Med.* **206**, 743-750 (2009).

27 Xiong, N. & Raulet, D. H. Development and selection of γδ T cells. *Immunol. Rev.* **215**, 15-31 (2007).

28 Xiong, N., Kang, C. & Raulet, D. H. Positive selection of dendritic epidermal gammadelta T cell precursors in the fetal thymus determines expression of skin-homing receptors. *Immunity* **21**, 121-131 (2004).

29 Aono, A. *et al.* Forced expression of terminal deoxynucleotidyl transferase in fetal thymus resulted in a decrease in γδ T cells and random dissemination of Vγ3Vδ1 T cells in skin of newborn but not adult mice. *Immunology* **99**, 489-497 (2000).

30 Lewis, J. M. *et al.* Selection of the cutaneous intraepithelial γδ+ T cell repertoire by a thymic stromal determinant. *Nat. Immunol.* **7**, 843-850 (2006).

31 Boyden, L. M. *et al.* Skint1, the prototype of a newly identified immunoglobulin superfamily gene cluster, positively selects epidermal γδ T cells. *Nat. Genet.* **40**, 656-662 (2008).

32 Afrache, H., Gouret, P., Ainouche, S., Pontarotti, P. & Olive, D. The butyrophilin (BTN) gene family: from milk fat to the regulation of the immune response. *Immunogenetics* **64**, 781-794 (2012).

33 Abeler-Dorner, L., Swamy, M., Williams, G., Hayday, A. C. & Bas, A. Butyrophilins: an emerging family of immune regulators. *Trends Immunol.* **33**, 34-41 (2012).

34 Lewis, J. M. *et al.* Selection of the cutaneous intraepithelial γδ+ T cell repertoire by a thymic stromal determinant. *Nat. Immunol.* **7**, 843-850 (2006).

35 Turchinovich, G. & Hayday, A. C. Skint-1 identifies a common molecular mechanism for the development of interferon-γ-secreting versus interleukin-17-secreting γδ T cells. *Immunity* **35**, 59-68 (2011).

36 Groh, V. *et al.* Human lymphocytes bearing T cell receptor γ/δ are phenotypically diverse and evenly distributed throughout the lymphoid system. *J. Exp. Med.* **169**, 1277-1294 (1989).

37 Spetz, A. L., Strominger, J. & Groh-Spies, V. T cell subsets in normal human epidermis. *Am. J. Pathol.* **149**, 665-674 (1996).

38 Foster, C. A. *et al.* Human epidermal T cells predominantly belong to the lineage expressing alpha/beta T cell receptor. *J. Exp. Med.* **171**, 997-1013 (1990).

39 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

40 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725-2729 (2013).

41 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671-675 (2012).

42 Schaerli, P. *et al.* A skin-selective homing mechanism for human immune surveillance T cells. *J. Exp. Med.* **199**, 1265-1275 (2004).

43 Barbee, S. D. *et al.* Skint-1 is a highly specific, unique selecting component for epidermal T cells. *Proc. Natl. Acad. Sci. USA* **108**, 3330-3335 (2011).

44 Kazen, A. R. & Adams, E. J. Evolution of the V, D, and J gene segments used in the primate γδ T-cell receptor reveals a dichotomy of conservation and diversity. *Proc. Natl. Acad. Sci. USA* **108**, E332-340 (2011).

45 Klein, M. H. *et al.* Diversity and structure of human T-cell receptor α-chain variable region genes. *Proc. Natl. Acad. Sci. USA* **84**, 6884-6888 (1987).

46 Ebert, L. M., Meuter, S. & Moser, B. Homing and function of human skin gammadelta T cells and NK cells: relevance for tumor surveillance. *J. Immunol.* **176**, 4331-4336 (2006).

47 Murphy, W. J., Pevzner, P. A. & O'Brien, S. J. Mammalian phylogenomics comes of age. *Trends Genet.* **20**, 631-639 (2004).

48 Song, S., Liu, L., Edwards, S. V. & Wu, S. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. *Proc. Natl. Acad. Sci. USA* **109**, 14942-14947 (2012).

49 Bandyopadhyay, S., Sharan, R. & Ideker, T. Systematic identification of functional orthologs based on protein network comparison. *Genome Res.* **16**, 428-435 (2006).

50 Nehrt, N. L., Clark, W. T., Radivojac, P. & Hahn, M. W. Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Comput. Biol.* **7**, e1002073 (2011).

51 Altenhoff, A. M., Studer, R. A., Robinson-Rechavi, M. & Dessimoz, C. Resolving the ortholog conjecture: orthologs tend to be weakly, but significantly, more similar in function than paralogs. *PLoS Comput. Biol.* **8**, e1002514 (2012).

52 Mohan, A., Uversky, V. N. & Radivojac, P. Influence of sequence changes and environment on intrinsically disordered proteins. *PLoS Comput. Biol.* **5**, e1000497 (2009).

53 Liao, B.-Y. & Zhang, J. Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proc. Natl. Acad. Sci. USA* **105**, 6987-6992 (2008).

54 Barbee, S. D. *et al.* Skint-1 is a highly specific, unique selecting component for epidermal T cells. *Proc. Natl. Acad. Sci. USA* **108**, 3330-3335 (2011).

55 Hein, W. R. & Dudler, L. TCR γδ+ cells are prominent in normal bovine skin and express a diverse repertoire of antigen receptors. *Immunology* **91**, 58-64 (1997).

56 Van Rhijn, I. *et al.* Massive, sustained γδ T cell migration from the bovine skin in vivo. *J. Leukoc. Biol.* **81**, 968-973 (2007).

57 Hein, W. & Dudler, L. TCR γδ+ cells are prominent in normal bovine skin and express a diverse repertoire of antigen receptors. *Immunology* **91**, 58-64 (1997).

58 Mackay, C. R. & Hein, W. R. A large proportion of bovine T cells express the gamma delta T cell receptor and show a distinct tissue distribution and surface phenotype. *Int. Immunol.* **1**, 540-545 (1989).

59 Jameson, J. & Havran, W. L. *Immunol. Rev.* **215**, 114-122 (2007).

60 de Jong, A. *et al.* CD1a-autoreactive T cells are a normal component of the human αβ T cell repertoire. *Nat. Immunol.* **11**, 1102-1109 (2010).

61 Colonna, M. Skin function for human CD1a-reactive T cells. *Nat. Immunol.* **11**, 1079-1080 (2010).

62 Kasmar, A., Van Rhijn, I. & Moody, D. B. The evolved functions of CD1 during infection. *Curr. Opin. Immunol.* **21**, 397-403 (2009).

63 Abeler-Dörner, L., Swamy, M., Williams, G., Hayday, A. C. & Bas, A. Butyrophilins: an emerging family of immune regulators. *Trends Immunol.* **33**, 34-41 (2012).