Available online at www.sciencedirect.com

ScienceDirect

Genomics 89 (2007) 618–629

GENOMICS

www.elsevier.com/locate/ygeno

Members of a novel gene family, *Gsdm,* are expressed exclusively in the epithelium of the skin and gastrointestinal tract in a highly tissue-specific manner☆

Masaru Tamura [a,b,1], Shigekazu Tanaka [a,b,1], Tomoaki Fujii [a,b,1], Aya Aoki [c], Hiromitu Komiyama [a,d], Kiyoshi Ezawa [e], Kenta Sumiyama [e], Tomoko Sagai [a], Toshihiko Shiroishi [a,b,*]

[a] *Mammalian Genetics Laboratory, Genetic Strains Research Center, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-0845, Japan*
[b] *Department of Genetics, School of Life Science, The Graduate University for Advanced Studies, 1111 Yata, Mishima, Shizuoka 411-0845, Japan*
[c] *Department of Neurosurgery, Juntendo University School of Medicine, Shizuoka Hospital, 1129 Nagaoka, Izunagaoka-chou, Izunokuni, Shizuoka 410-2295, Japan*
[d] *Department of Coloproctological Surgery, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan*
[e] *Division of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-0845, Japan*

Received 5 September 2006; accepted 8 January 2007
Available online 12 March 2007

Abstract

Gasdermin (*Gsdm)* was originally identified as a candidate causative gene for several mouse skin mutants. Several *Gsdm*-related genes sharing a protein domain with *DFNA5,* the causative gene of human nonsyndromic hearing loss, have been found in the mouse and human genomes, and this group is referred to as the DFNA5–Gasdermin domain family. However, our current comparative genomic analysis identified several novel motifs distinct from the previously reported domain in the *Gsdm*-related genes. We also identified three new *Gsdm* genes clustered on mouse chromosome 15. We named these genes collectively the *Gsdm* family. Extensive expression analysis revealed exclusive expression of *Gsdm* family genes in the epithelium of the skin and gastrointestinal tract in a highly tissue-specific manner. Further database searching revealed the presence of other related genes with a similar N-terminal motif. These results suggest that the *Gsdm* family and related genes have evolved divergent epithelial expression profiles.

Keywords: *Gsdm* family; *Dfna5*; *Dfnb59* (*pejvakin*); Gastrointestinal tract; Epithelium; Gene duplication; Evolution

Gene duplication events are important for the expansion of gene families and the functional divergence of genes [1–5]. For example, *Hox* genes are arranged in four clusters that were generated by two rounds of duplication of a single ancestral gene cluster [6,7]. Gene duplications and lineage-specific gene loss have greatly contributed to variable functions and expression profiles of individual members of different gene families over the course of evolution.

We originally identified Gasdermin (*Gsdm*) as a candidate gene responsible for the phenotype of a mouse skin mutant, Recombination-induced mutation 3 (*Rim3*) [8], which exhibits hyperproliferation and misdifferentiation of the epidermis and hair follicles [9]. Recently, *Gsdm* was identified as one of three members of a structurally related gene cluster located on mouse chromosome 11 [10,11], and several studies have identified a *Gsdm* family member as the gene responsible for several mutant mouse skin strains [11,12]. Deafness autosomal dominant nonsyndromic sensorineural 5 (*DFNA5*) was identified as the causative gene for nonsyndromic hearing loss in human [13], but

☆ The nomenclature used in this article has not been approved by the human and mouse gene nomenclature committees. However, it is hoped that this article will start a discussion within the community that could lead to a new nomenclature.

* Corresponding author. Mammalian Genetics Laboratory, Genetic Strains Research Center, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-0845, Japan. Fax: +81 55 981 6817.
E-mail address: tshirois@lab.nig.ac.jp (T. Shiroishi).

[1] These authors contributed equally to this work.

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.01.003

its molecular function is largely unknown. The *Gsdm* human orthologue GSDM1 and human DFNA5 proteins have some conserved amino acid residues throughout the coding region, excluding 2 amino acids at the N–terminus and the 25 C-terminal amino acids. This conserved region is now referred to as the DFNA5–Gasdermin domain (Sanger Center, Pfam/getacc PF04598; URL: http://www.sanger.ac.uk/cgi-bin/Pfam/getacc%3FPF04598), and several other genes harboring this domain have been identified. Melanoma-derived leucine zipper extra-nuclear factor (*MLZE*) was originally isolated from a human melanoma cell line as a marker of melanoma progression [14]. It was mapped to chromosome 8q24.1–q24.2, and its mouse homologue, *Mlze,* was mapped to the syntenic region in Band D1 of chromosome 15 (15D1).

Gsdm-related genes have also been identified by searching the available public genome databases. The human gene *PRO2521* [15], synonymous name Gasdermin-like (approved symbol *GSDML*) [10], is chromosomally located near *GSDM1* in the human syntenic region corresponding to mouse chromosome 11. Additionally, human Gasdermin domain containing-1 (*GSDMDC1*), also known as *DFNA5*–like gene (*DFNA5L*), and its mouse homologue *Gsdmdc1,* as well as multiple other *Gsdm*-related genes, are located on human chromosome 8 (8q24.3) and mouse chromosome 15 (15D3), respectively [16].

The majority of *Gsdm*-related genes were independently isolated and assigned to a single gene family based on their amino acid identity to the DFNA5 protein. However, a comprehensive analysis of the genomic characteristics and expression of these gene family members has not been undertaken, and there is confusion surrounding the names and identities of different family members. It remains unclear precisely how many *Gsdm*-related genes exist in the mouse and human genomes in total and whether *DFNA5* and its mouse homologue *Dfna5h* are phylogenetically included in this gene family. Finally, is the previously reported DFNA5-Gasdermin domain shared by all other *Gsdm*-related proteins?

In this study, we first carried out a systematic genomic comparison and analysis of the phylogenic relationship of reported *Gsdm*-related genes. This clearly showed that *Dfna5h* and *DFNA5,* in each species, are phylogenetically distant from *Gsdm* and other *Gsdm*-related genes. We then conducted an extensive genomic database search and found a cluster of three additional *Gsdm*-related mouse genes near the 5′ region of the *Mlze* gene on mouse chromosome 15 (15D1). Among genomes analyzed this duplication is seen only in the mouse. A comprehensive comparison of all *Gsdm*-related genes revealed several conserved amino acid sequences in the N- and C-terminal regions of the proteins and the presence of a conserved leucine-rich motif. In particular, the C-terminus of each protein contains the sequence $PX_{12-13}ALYX_5LX_2L$ outside of the DFNA5–Gasdermin domain. Based on the sequence similarity of *Gsdm*-related genes, we generated a comprehensive list of member genes as well as their phylogenic relationships. Based on this analysis, we feel it is appropriate to exclude *Dfna5h* and *DFNA5* from the *Gsdm* family. In this report, we also propose a new naming scheme to reflect the sequence similarities and refined phylogenetic relationships of this protein family. In addition to the *Gsdm* family genes, our database search also revealed the presence of additional mouse proteins highly similar to the N-terminal half of Gsdm and Dfna5h, and their genes mapped to mouse chromosome 2 in a region linked to the mouse *Hoxd* cluster. These data suggested that the *Gsdm* family genes were derived from an ancestral gene that had been physically linked to an ancestral *Hox* cluster. Following two genome duplication events, they became dispersed over four chromosomes. Further gene duplications occurred only on two mouse chromosomes, Chr. 11 and Chr. 15, and generated gene clusters at these locations.

Information regarding the expression patterns of *Gsdm* family members is very limited. The three genes clustered on mouse chromosome 11 are expressed in the skin and upper gastrointestinal tract [8,11,12], but it is unclear which of these genes are expressed in each location. Therefore, we conducted a systematic expression analysis of *Gsdm* family members in the mouse, and, like the *Gsdm* family members located on chromosome 11, the four genes located on mouse chromosome 15 are also expressed by epithelial cells in the lower gastrointestinal tract.

Results

Refinement of Gsdm-related genes by phylogenetic analysis

We conducted a phylogenetic analysis of the human *DFNA5* and *GSDM1*-related genes and their mouse homologues based

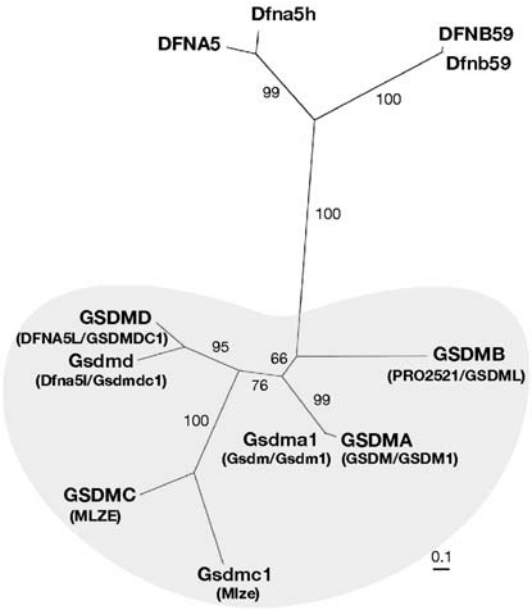


Fig. 1. Phylogeny of the *Gsdm* family genes. Unrooted gene tree inferred from amino acid sequence data using the neighbor-joining method is shown [35]. The bootstrap values (percentage) obtained from 1000 replications are shown at the branches determining each clade. The scale bar shows the estimated evolutionary distance. The tree shows the previously reported DFNA5–Gasdermin domain family genes together with a newly identified *DFNA5*-related gene (in a recent report, the gene was named *DFNB59* [17]). The shaded region indicates the *Gsdm* family, which is classified as a distinct clade separate from another clade including the *DFNA5* and *DFNB59* genes. The distance between the two clades is supported by a 100% bootstrap value. The gene symbols in parentheses are the former gene names.

on their amino acid sequences. Phylogenetic relationships and bootstrap analysis revealed that these genes can be subdivided into two groups: a *Gsdm* group that includes human *GSDM1, PRO2521, MLZE,* and *GSDMDC1* and their mouse orthologues and a *DFNA5* group that includes human *DFNA5,* another *DFNA5*-like gene, and their mouse orthologues (Fig. 1). These two groups are segregated with 100% bootstrap support, indicating that the distance between these two distinct groups is significant. We propose to refer to the genes included in the former group as the *Gsdm* family.

Based on these results, we reclassified and renamed all genes previously classified as belonging to the DFNA5–Gasdermin family (Table 1). The gene we originally isolated as Gasdermin (*Gsdm*) on mouse chromosome 11 is renamed Gasdermin A1 (*Gsdma1*), and the two neighboring paralogous genes, Gasdermin-like 2 (*Gsdml2*) (Gasdermin 2) and Gasdermin-like 1 (*Gsdml1*) (Gasdermin 3), are renamed Gasdermin A2 (*Gsdma2*) and Gasdermin A3 (*Gsdma3*), respectively. Human *GSDM1,* which is located at the human syntenic region 17q21, is renamed Gasdermin A (*GSDMA*). The human gene *PRO2521* (*GSDML*) is renamed Gasdermin B (*GSDMB*); a mouse *GSDMB* homologue has not yet been found. *Mlze,* located on mouse chromosome 15, is renamed Gasdermin C1 (*Gsdmc1*), and its human homologue located centromeric to *BM-009* (human homologue of mouse 0910001A06Rik) on chromosome 8q24.1 is renamed *GSDMC*. Finally, human *GSDMDC1* (synonymous name *DFNA5L*) and its mouse counterpart, *Gsdmdc1,* are renamed human Gasdermin D (*GSDMD*) and mouse Gasdermin D (*Gsdmd*), respectively. In the phylogenetic tree in Fig. 1, the former name of each gene is shown in parentheses.

To determine whether additional *Gsdm*-related gene(s) exist in the mouse genome, we undertook an extensive BLAST search of the public genome databases using *Gsdma1, Gsdmc1,* and *Dfna5h* as query. As a result, we discovered that *Gsdmc1* is not an isolated, solitary gene, but is clustered with three other related genes within a 150-kb stretch on chromosome 15 (band 15D1) (Figs. 2 and 3A). The three newly identified genes, AI987692, 9930109F21Rik, and 9030605I04Rik, were named Gasdermin C2 (*Gsdmc2*), Gasdermin C3 (*Gsdmc3*), and Gasdermin C4 (*Gsdmc4*), respectively. The BLAST search also found a novel *Dfna5h*-related gene located on mouse chromosome 2 (band 2C2) (Fig. 3A). During preparation of this article, *DFNB59* was reported to be the causative gene for nonsyndromic human deafness with a neuronal defect [17]. This gene is the human orthologue of the mouse *Dfna5h*-related gene.

Among the other available species databases, we have also found *GSDM* family members in the following organisms: chicken (*cGsdma,* AJ721093), cow (*bGsdma,* XM_870913; *bGsdmb,* XM_882737; *bGsdmc,* XM_585278; *bGsdmd,* BC113262), dog (*dGsdma,* XM_845245; *dGsdmc,* XM_847168; *dGsdmd,* XM_846639), rat (*rGsdma,* XM_001081404; *rGsdmc,* XM_216925; *rGsdmd,* XM_235434), and chimpanzee (*pGsdma,* XM_523624; *pGsdmb,* hmm23367; *pGsdmc,* DQ031736; *pGsdmd,* hmm39204). Interestingly, we found no *Gsdm* family genes in amphibians or teleosts, having searched the fugu (NCBI), zebrafish (NCBI), medaka (NIG), and *Xenopus* (NCBI) genome databases.

Table 1
Gsdm gene family in mouse and human

Species	Gene name (gene symbol)	Previous gene name (gene symbol)	Reference
Mouse	Gasdermin A1 (*Gsdma1*)	Gasdermin (*Gsdm*), Gasdermin 1 (*Gsdm1*)	[8,11]
Mouse	Gasdermin A2 (*Gsdma2*)	Gasdermin-like 2 (*Gsdml2*), Gasdermin 2 (*Gsdm2*)	[10,11]
Mouse	Gasdermin A3 (*Gsdma3*)	Gasdermin-like 1 (*Gsdml1*), Gasdermin 3 (*Gsdm3*)	[10,11]
Mouse	Gasdermin C1 (*Gsdmc1*)	Melanoma-derived leucine zipper, extranuclear factor (*Mlze*)	[14]
Mouse	Gasdermin C2 (*Gsdmc2*)	–	This study
Mouse	Gasdermin C3 (*Gsdmc3*)	–	This study
Mouse	Gasdermin C4 (*Gsdmc4*)	–	This study
Mouse	Gasdermin D (*Gsdmd*)	Gasdermin domain-containing 1 (*Gsdmdc1*), Dfna5-like (*Dfna5l*)	[16]
Human	Gasdermin A (*GSDMA*)	Gasdermin (*GSDM*), Gasdermin 1 (*GSDM1*)	[8,11]
Human	Gasdermin B (*GSDMB*)	PRO2521 (*PRO2521*), Gasdermin-like (*GSDML*)	[10,15]
Human	Gasdermin C (*GSDMC*)	Melanoma-derived leucine zipper, extranuclear factor (*MLZE*)	[14]
Human	Gasdermin D (*GSDMD*)	Gasdermin domain-containing 1 (*GSDMDC1*), Dfna5-like (*DFNA5L*)	[16]

Gsdm family proteins commonly share amino acid sequences at the N- and C-termini

We aligned the amino acid sequences of all Gsdm family proteins, excluding DFNA5 and Dfna5h, to explore novel domains and motifs. As shown in Figs. 2A and 2B, in addition to the DFNA5–Gasdermin domain, Gsdm family proteins share highly conserved leucine-rich regions throughout the entire protein. They share common N-terminal sequences characterized by a conserved LEP amino acid triplet (except LDT in GSDMB) followed by N/DL/VY/FV/LVM/TE and several more conserved amino acid residues (Fig. 2B). These N-terminal sequences are also shared with DFNA5 and DFNB59, although the extent of conservation for these proteins is lower than for the Gsdm family members. The amino acid sequences of the middle part of the proteins are less conserved, and each family member has a relatively unique primary sequence. At the C-terminus of each protein, five leucine-rich regions are well conserved in all family members (Figs. 2A and 2B). These N- and C-terminus conserved regions were characterized by the sequences F/LE/DX$_3$R/KXL/VXR/K/QQ/E/KL/V/MX$_{3-5}$D/EL/F/MX$_2$L/VX$_2$LX$_5$F/LX$_2$F/YXL/IV/L, T/SL/IX$_2$V/I/LLE/DP(/T), N/DL/VY/FV/LVXEX$_2$E/QX$_2$Q/NK/EE/N/DX$_2$L/V/I, V/IXI/LX$_5$L/MA/SY/FR/KX$_3$L, LX$_{15}$LX$_{8-10}$L, PX$_3$L/I/VL, AX$_3$LX$_3$QX$_{11-12}$LX$_6$V, LX$_3$LLX$_5$EX$_{6-7}$LX$_4$GL, and PX$_{12-13}$ALYX$_5$LX$_2$L (Figs. 2A and 2B). Notably, PX$_{12-13}$ALYX$_5$LX$_2$L at the C-terminus end, not the DFNA5-Gasdermin domain, is the most conserved region of the Gsdm family proteins.

Gsdma1 (formerly Gsdm) contains a putative leucine zipper motif within its region V (Figs. 2A and 2B) [8]. Likewise, mouse Gsdmc1 (formerly Mlze) and human GSDMC (formerly MLZE) are also thought to contain a leucine zipper motif [14]. When considered together, this suggests that a leucine zipper motif could be a conserved feature of all *Gsdm* family members. However, when the amino acid sequences of all *Gsdm* family members are aligned, it becomes clear that, while all Gsdm family members have leucine-rich domains, only Gsdma and Gsdmc possess a leucine zipper. Other leucine residues, as well as other amino acids, are more conserved (Fig. 2A). For example, alanine-348 in GsdmA3, which corresponds to the alanine located in the leucine zipper motif of Gsdmc1, is conserved among all members of the Gsdm family proteins.

Based on our database searching and revised nomenclature, the *Gsdm* family comprises eight mouse and four human genes. *Gsdmd* is the sole member of this subfamily in the mouse genome, but *Gsdma* and *Gsdmc* consist of clusters of three and four genes, respectively. The absence of a *Gsdmd* cluster was confirmed by genomic Southern blot analysis (data not shown). In this study, we isolated cDNAs for all *Gsdm* family genes from mouse skin, stomach, small intestine, and colon cDNA libraries or by RT-PCR, and we verified their sequences.

Genome structures and chromosomal locations of the Gsdm family genes

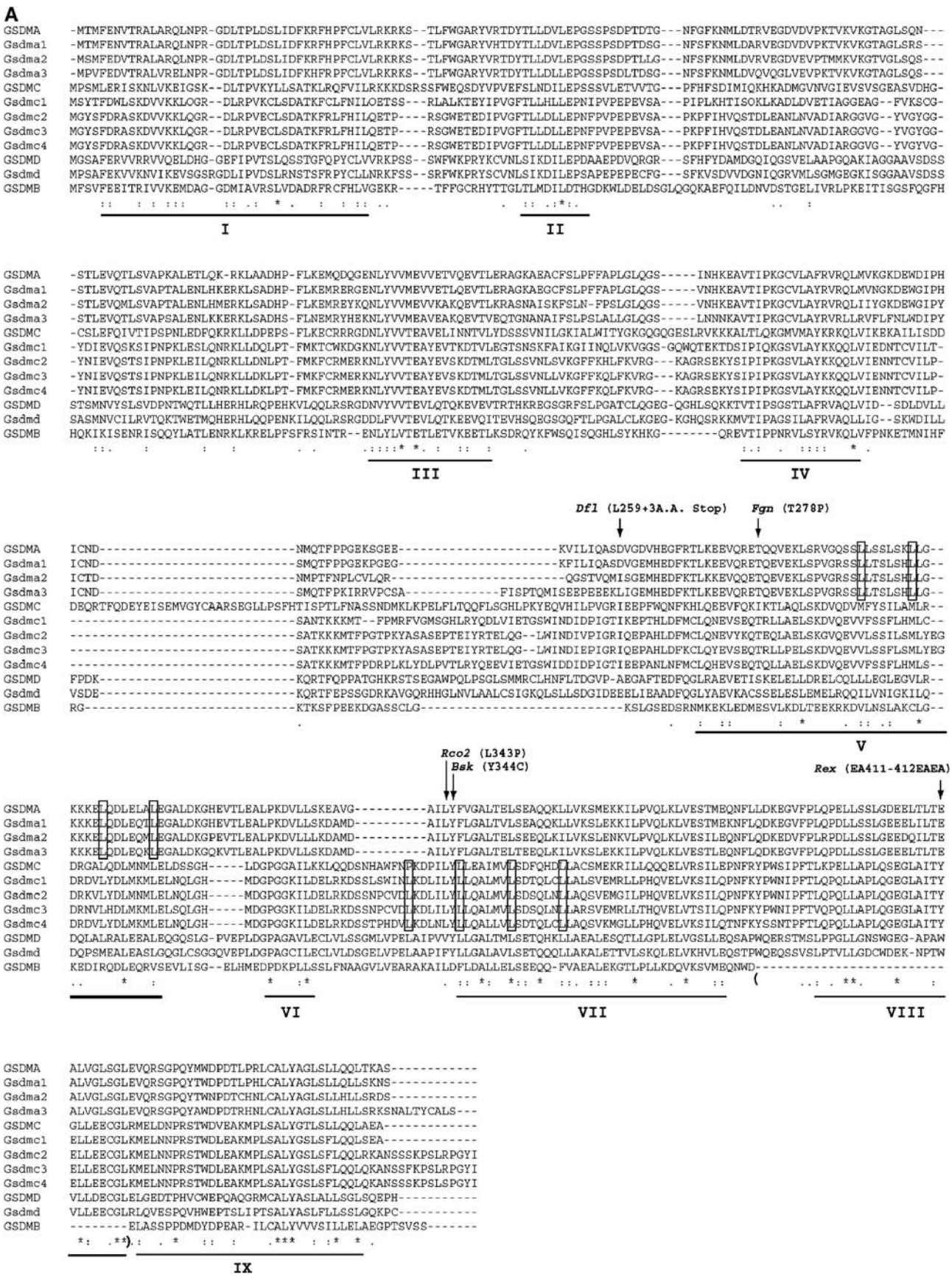
The chromosomal locations of all *Gsdm* family genes, *Dfna5h*, and *Dfnb59* in the mouse genome are shown in Fig. 3A. The *Gsdma* cluster is closely linked to the retinoic acid receptor α (*Rara*) and keratin-1 (*Krt-1*) genes on mouse chromosome 11. Likewise, the *Gsdmc* cluster and *Gsdmd* are closely linked to the retinoic acid receptor γ (*Rarg*) and keratin-2 (Krt-2) genes on mouse chromosome 15 (Fig. 3A). In addition, the *Gsdm* family genes, *Dfna5h*, and *Dfnb59* are all closely linked to *Hox* gene clusters in the mouse genome. For example, the *Gsdma* cluster is linked to *Hoxb,* the *Gsdmc* cluster and *Gsdmd* to *Hoxc, Dfna5h* to *Hoxa,* and *Dfnb59* gene to *Hoxd* (Fig. 3A).

The genome structures of the individual members of the *Gsdm* family are shown in Fig. 3B, based on the available mouse (*Mus musculus* build 35.1) and human genome sequences (*Homo sapiens* build 35.1). The genes surrounding the *Gsdm* family members are highly conserved between the human and the mouse genomes. Human *GSDMA* and *GSDMB* are located in the interval between Zona pellucida binding protein 2 (*ZPBP2*) and *CSF3.* Likewise, the mouse *Gsdma* cluster is located in the interval between *Zpbp2* and *Csf3,* consistent with a previous report [10,11]. A similar conservation of the genes neighboring *Gsdmc* and *GSDMC* also occurs (Fig. 3B), but only a single *GSDMC* subfamily member was found in humans. We confirmed the existence or absence of paralogues of each *Gsdm* (*GSDM*) family gene by Southern blot analysis with genomic DNAs from both species (data not shown). Southern blot analysis of other rodent genomes showed that both *Gsdma* and *Gsdmc* are solitary genes in *Rattus norvegicus* and *Rattus rattus,* respectively, both of which belong to the subfamily *Murinae* (Figs. 4A and 4B). These results suggest that the *Gsdma* and *Gsdmc* clusters are specific to the mouse genome, and the responsible duplication events occurred after the divergence of mice and rats.

Gsdm family genes show tissue-specific and linear expression in the gastrointestinal tract

The genes of the mouse *GsdmA* cluster are expressed in the upper gastrointestinal (GI) tract and skin, but the upper GI tract encompasses a variety of different tissue types. We carried out RT-PCR to define more clearly the expression of the *Gsdma* genes in mouse stomach. *Gsdma1* is specifically expressed in the cardiac region (fore stomach), while *Gsdma2* is expressed in the fundus and pylorus (glandular stomach) (Fig. 5A). This expression pattern suggests that *Gsdma1* is specific for squamous epithelium, and *Gsdma2* is expressed in glandular epithelium. We next examined the kinetics of expression of *Gsdma1* and *Gsdma2.* In the embryonic stomach, the earliest expression of *Gsdma1* was detected at embryonic day (E) 12.5, and *Gsdma2* was seen at E 16.5 to E 17.5 (Fig. 5B). The timing of the appearance of *Gsdma1* and *Gsdma2* mRNA correlates with the histologic transition from squamous to glandular epithelium in the embryonic stomach [18,19].

We next conducted a Northern blot analysis to examine the expression of *Gsdmc* cluster genes and *Gsdmd* in various mouse tissues (Fig. 5C). In addition, we further analyzed the expression of the *Gsdma* cluster genes. Consistent with previous reports, *Gsdma1* and *Gsdma2* were expressed


A
GSDMA
Gsdma1
Gsdma2
Gsdma3
GSDMC
Gsdmc1
Gsdmc2
Gsdmc3
Gsdmc4
GSDMD
Gsdmd
GSDMB
I
II
III
IV
Dfl (L259+3A.A. Stop)
Fgn (T278P)
V
Rco2 (L343P)
Bsk (Y344C)
Rex (EA411-412EAEA)
VI
VII
VIII
IX

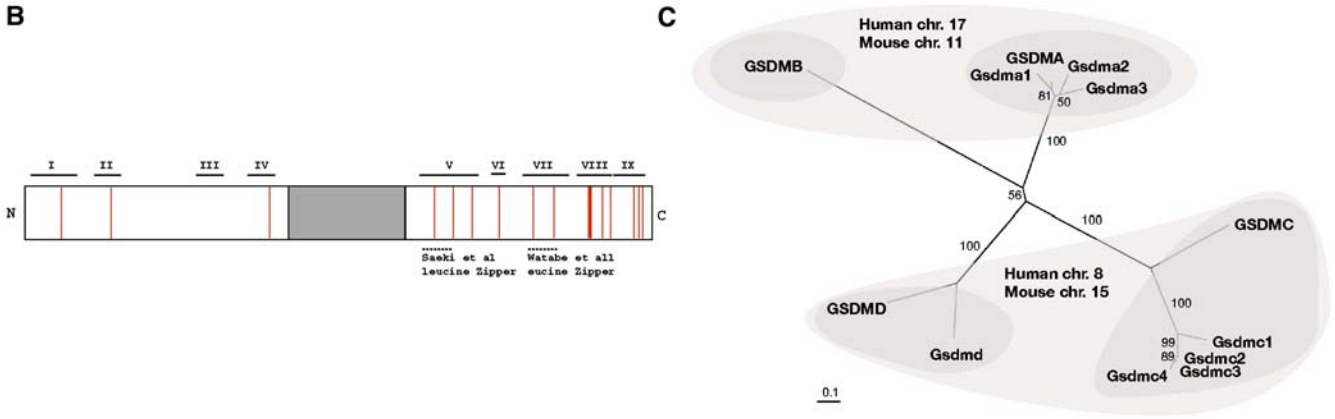


Fig. 2. Gsdm family proteins. (A) Multiple amino acid sequences were aligned using ClustalW [34]. Identical amino acid residues through all the family members are indicated by asterisks below the alignment. Substitutions with highly conserved and less conserved residues are indicated by a colon and a period, respectively. Asterisk, colon, and period in parentheses at the C-terminus region indicate conservation excluding GSDMB. Previously reported leucine zipper motifs are shown in boxes. Arrows indicate positions of the mouse mutations. Amino acid substitutions and type of mutation are shown in parentheses. Underlines indicate conserved sequences, (I) F/LE/DX_3R/KXL/VXR/K/QQ/E/KL/V/MX_{3-5}D/EL/F/MX_2L/VX_2LX_5F/LX_2F/YXL/IV/L, (II) T/SL/IX_2V/I/LLE/DP(/T), (III) N/DL/VY/FV/LVXEX_2E/QX_2Q/NK/EE/N/DX_2L/V/I, (IV) V/IXI/LX_5L/MA/SY/FR/KX_3L, (V) LX_{15}LX_{8-10}L, (VI) PX_3L/I/VL, (VII) AX_3LX_3QX_{11-12}LX_6V, (VIII) LX_3LLX_5EX_{6-7}LX_4GL, and (IX) PX_{12-13}ALYX_5LX_2L, in the Gsdm family. (B) A schematic representation of the mouse and human Gsdm family proteins. They are composed of three parts, the N-terminus, middle, and C-terminus. Red lines show leucines conserved in all Gsdm family members. The gray box representing the middle part is unique for each Gsdm family member. The most conserved sequence regions are indicated on the top. Positions of previously reported leucine zipper motif are showed by broken lines [8,15]. (C) Unrooted NJ tree constructed for eight mouse *Gsdm* family genes and four human *GSDM* family genes. *Gsdm* family genes were divided into four subfamilies, *Gsdma, GSDMB, Gsdmc,* and *Gsdmd.* The mouse counterpart of *GSDMB* was not found in the mouse genome. The darker shaded background depicts each of the four subfamilies and the lighter shaded background two chromosomes on which the *Gsdm* family genes reside.

predominantly in the stomach [8,11], and no expression was detected in the small intestine, colon, brain, liver, kidney, heart, lung, spleen, or skeletal muscle (Fig. 5C). *Gsdmd* was seen predominantly in the small intestine, but weak expression was seen in the colon, stomach, liver, lung, heart, and spleen (Fig. 5C). The *Gsdmc* cluster genes are strongly expressed in the colon and weakly in the small intestine, stomach, and kidney (Fig. 5C). The expression of individual *GsdmC* genes could not be distinguished from one another due to cross-reactivity of the probe used for all genes of the *Gsdmc* cluster. These results clearly show the differential expression of the *Gsdma, Gsdmc,* and *Gsdmd* genes (Fig. 5C).

The expression of the *Gsdmc* cluster genes and *Gsdmd* appeared to depend upon the relative proximal–distal axis along the gastrointestinal tract, and we analyzed the expression of these genes in the intestinal tract in greater detail. We observed increased expression of *Gsdmd* in the proximal region of the small intestine, and its expression gradually decreased toward the distal small intestine. Additionally, *Gsdmd* expression was weaker in the caecum and appendix compared to the distal small intestine, and this low level of *Gsdmd* expression was maintained in the proximal and distal colon (Fig. 5D). In contrast, *Gsdmc* expression was much lower in the proximal and distal small intestine compared with the appendix and caecum, the regions where *Gsdmc* expression were highest. We also observed weak *Gsdmc* expression in the proximal and distal colon (Fig. 5D). Thus, the expression of the *Gsdmd* and *Gsdmc* cluster genes is differentially regulated along the proximal–distal axis of the intestinal tract. These results suggest that these genes may play nonoverlapping roles in the development of the intestinal tract.

Gsdm family genes are expressed in an epithelial cell type-specific manner

We showed above that *Gsdmd* and *Gsdmc* cluster genes are expressed in a cell-type-specific manner, by Northern blot, and we wanted to characterize further the expression of all *Gsdm* family genes by in situ hybridization using cryosections of mouse GI tract. In the cardiac region of the stomach, *Gsdma1* was expressed in the transitional cell layer as well as in the esophageal epithelium (Fig. 6A and data not shown), but *Gsdmc*-expressing cells contact the epithelial basal layer (Fig. 6B). The pattern of *Gsdma1* expression suggests that it is not expressed in the stem cells of the squamous cell layer. The differentiation of the gastric glandular epithelium is bidirectional [19] in both humans and mice, and stem cells are present in the middle of the gastric mucosa. In the mouse glandular stomach, *Gsdma2*-expressing cells were present from the middle to the upper region of the gastric mucosa (Fig. 6C), and expression of *Gsdmc* cluster genes also occurred in the middle of the gastric mucosa (Fig. 6D). The cells expressing *Gsdma2* seem destined to become surface mucosal cells, while those expressing *Gsdmc* cluster genes are mucus neck cells. In addition, elevated levels of *Gsdma2* were observed from the fundus to the pylorus in the glandular stomach, while expression of *Gsdmc* cluster genes declined in a proximal to distal fashion. High expression of *Gsdmc* cluster genes was seen in the fundus linked to the esophageal–cardiac (EC)

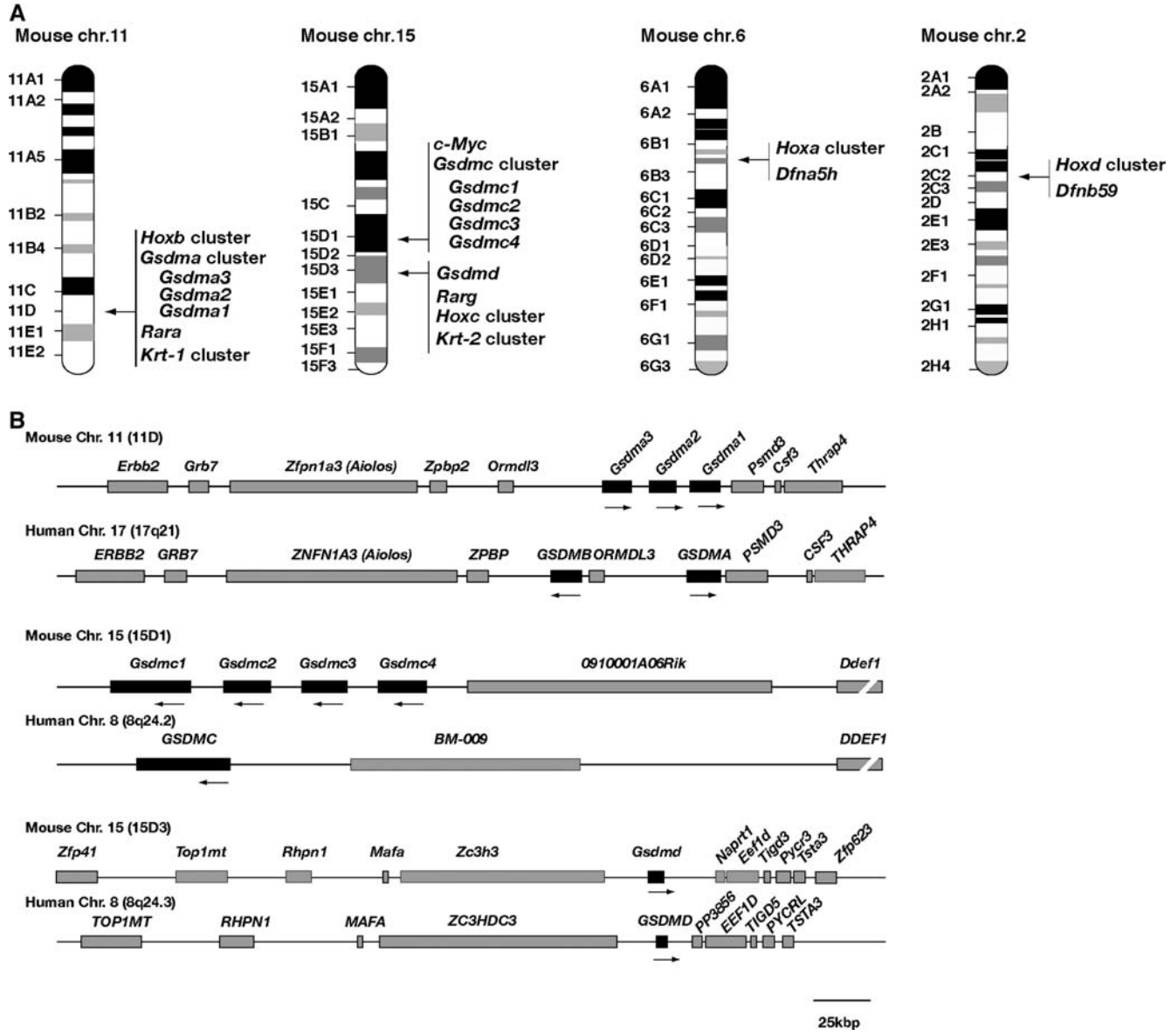


Fig. 3. Chromosomal locations and genome structures of the *Gsdm* family genes. (A) Chromosomal locations of the mouse *Gsdm* family genes, *Dfna5h*, and *Dfnb59* are shown together with the *Hox* clusters and surrounding genes. (B) Genome structures around the mouse *Gsdm* family genes and the human syntenic regions were constructed based on the NCBI genome database (Mouse build 35 and Human build 35). Transcription orientation of each gene is represented by an arrow.

junction (Fig. 6D). However, we detected low levels of *Gsdmc* cluster gene expression in the fundus and pylorus excluding the EC junction in the glandular stomach (data not shown). In the small intestine, *Gsdmd* and *Gsdmc* cluster genes were expressed in the villi and crypts, respectively. Cells expressing the *Gsdmc* cluster genes occurred from the middle to the upper crypt (Figs. 6E and 6F), and these expression patterns were maintained from the small intestine to the caecum and colon (Figs. 6G and 6H).

To examine whether the expression of *Gsdmc* cluster genes overlaps with that of the *Gsdma* cluster and *Gsdmd,* we carried out double-labeled in situ hybridization using digoxigenin- and fluorescein-labeled cRNA probes. The majority of *Gsdma* cluster-expressing cells did not also express *Gsdmc* cluster genes, but some cells in the fore and glandular stomach expressed members of both *Gsdm* clusters (Figs. 6I and 6J). Similarly, few cells coexpressing *Gsdmd* and *Gsdmc* cluster genes were seen in the small intestine and colon (Figs. 6K and 6L).

In situ hybridization analysis indicated that *Gsdma1* and *Gsdmd* are not expressed in the proliferating and stem cells in squamous epithelium and the gastrointestinal tract. However, the localization of *Gsdmc* cluster gene expression was consistent with the distribution of stem cells, and we wished to examine this possibility further by costaining cells of the stomach and lower intestinal tract with a stem cell marker, proliferating cell nuclear antigen (PCNA). However, *Gsdmc* cluster genes were expressed in the differentiating epithelium cell layer near, but not overlapping, the PCNA-expressing cells

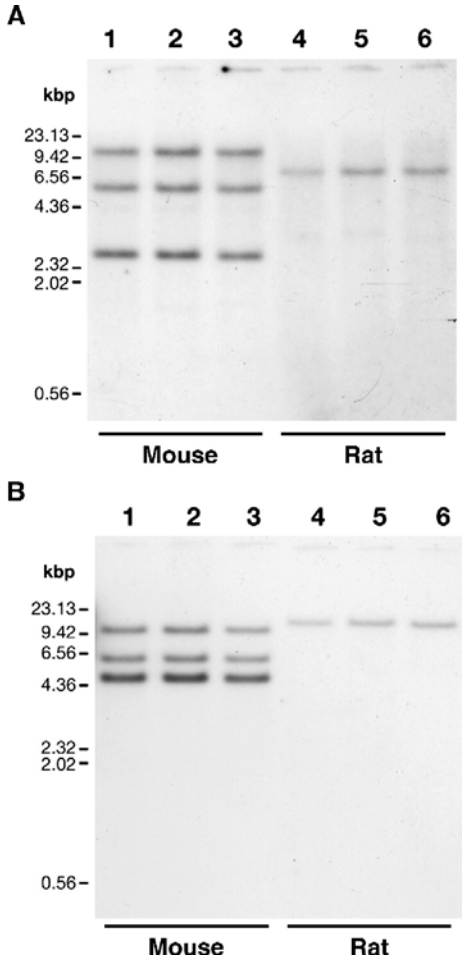


Fig. 4. Southern blot analysis of the *Gsdma* and *Gsdmc* genes of the mouse and rat genomes. (A) Southern blot analysis of *Gsdma.* Genomic DNA samples extracted from three mouse strains (C57BL/6J, 129/SvJ, and BALB/cAnN), two strains (WKAN/Msfw and Sprague–Dawley) of *R. norvegicus,* and one strain of *R. rattus* (*R. rattus kandianus*) were blotted onto nylon membranes and hybridized with labeled short genome fragment of mouse *Gsdma1.* Three bands were detected in the mouse genome, but only one band in the rat genome. (B) Southern blot analysis of *Gsdmc.* DNAs were extracted as in (A) and hybridized with a DNA fragment including a single exon (exon 12) of *Gsdmc4.* Note a single copy of *Gsdmc* in the rat genome, while the mouse genome contains four *Gsdmc* genes. Because of the high degree of similarity, the signals for *Gsdmc2* and *Gsdmc3* (approximately 4.5 kb) overlap. All of the genome samples were digested with the restriction enzyme *Xba*I. DNA size is indicated on left. Lanes 1–6 (in A and B), C57BL/6J, 129/SvJ, BALB/cAnN, WKAN/Msfw, Sprague–Dawley, and *R. rattus kandianus*.

(Figs. 6M–6O). These data indicate that all *Gsdm* family genes are expressed in differentiated cells, not proliferating and stem cells.

Discussion

Naming of the Gsdm family

Gasdermin was originally named based on its expression profile; it is expressed in the upper gastrointestinal tract and epidermis [8]. Our study has confirmed the appropriateness of this name as all identified members of the *Gsdm* family are expressed in a tissue-restricted manner in epithelial tissues throughout the gastrointestinal tract and epidermis. The *Gsdm* family now consists of eight mouse and four human genes, and a comparative analysis of the genome structure and amino acid sequences of the *Gsdm* family genes revealed several novel motifs. Thus, when considered together, the results of this paper indicate that the *Gsdm* family represents a novel set of genes with common structural and functional features.

Expression of the Gsdm family genes

The most striking feature of the *Gsdm* family is the unique expression pattern of the different family members. They are exclusively expressed in differentiated cells, and the expression of each member is highly tissue specific. In the gastrointestinal tract, their expression is strictly regulated along the proximal–distal and basal–apical axes. In mammals, genes of no other family are expressed in the gastrointestinal tract of adult animals in such a restricted fashion. *Hox* family genes are expressed in the gastrointestinal tract of developing embryos [20,21]. For example, *Hoxa-9, Hoxd-9,* and *Hoxc-9* expression is restricted to the caecum and colon, whereas other *Hox* family genes are diffusely expressed throughout the gastrointestinal tract. However, the expression of *Hox* family genes is not similarly restricted in the gastrointestinal tract of adult mice.

Gsdm1, Gsdm2, or *Gsdm3* expression was previously seen at the bottom of the gastric mucosa, probably in chief cells, using a polyclonal antibody [12]. In contrast, we clearly detected *Gsdma2* expression in the glandular stomach specifically and *Gsdma2*-expressing cells in the neck and pit region of the gastric mucosa (faveolar differentiation), but not in the glandular differentiation cell lineage (Fig. 6C). Because stem cells in the gastric mucosa differentiate bidirectionally, the luminal cells (surface mucosa and mucosa neck cells) and the basal cells (parietal and chief cells) arise from different cell lineages. This apparent discrepancy could be explained by the specificity of the antiserum used by Lunny et al. This antiserum reacts with Gsdma3 in addition to Gsdma1, but it is not known if it also reacts with Gsdma2. Alternatively, if the novel motifs have transmembrane domain activity and the *Gsdm* genes encode secretory proteins, mRNA expression and protein localization could easily differ on a cellular and anatomic level.

Although our understanding of the expression profiles of the *Gsdm* family members in mouse embryos and the human gastrointestinal tract remains limited, the promoters controlling the tightly restricted *Gsdm* expression in the gastrointestinal tract are good candidates for incorporation into gene therapy strategies targeting diseases of the digestive tract.

DFNA5 and DFNB59

Mutations in *DFNA5* were found in one Chinese and two Dutch families with sensorineural hearing impairment that led to the skipping of exon 8 with a resulting frameshift and protein truncation [13,22,23]. Additionally, point mutations (T54I, R183W) in *DFNB59* were identified in two Iranian families

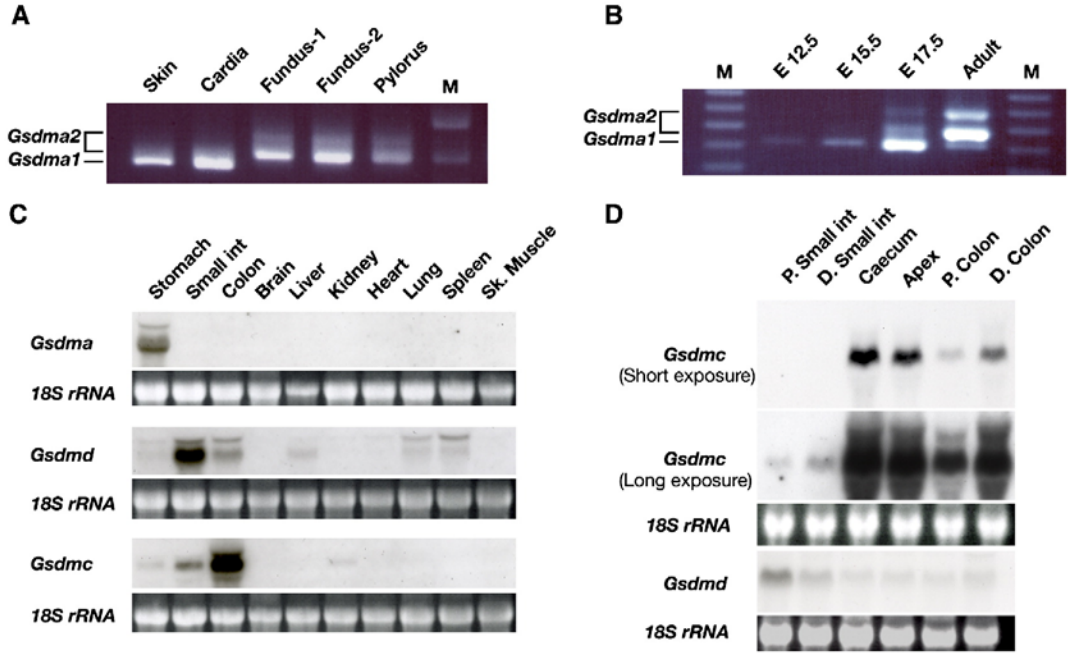


Fig. 5. Expression patterns of *Gsdm* family genes. (A and B) Expression of *Gsdma1* and *Gsdma2.* RT-PCR was used to detect the expression of these genes in the skin and stomach of adult mice (A) and in the mouse embryonic stomach (B). Fundus-1 and -2 indicate boundary regions between the fundus and the cardia and between the fundus and the pylorus, respectively. The expression of *Gsdma1* and *Gsdma2* was detected with the same primer set, but they can be distinguished by their sizes. (C and D) Northern blot analysis of the *Gsdm* family genes. The expression of the *Gsdm* family genes in various tissues of adult mice was examined. Total RNA (15 μg) prepared from several tissues of adult mice was analyzed. Blots were hybridized with cDNA probes for *Gsdma1, Gsdmc4,* and *Gsdmd. Note: Gsdma1* and *Gsdmc4* probes detect mRNAs of three *Gsdma* and four *Gsdmc* cluster genes, respectively. Before transfer, gels were stained with ethidium bromide to determine the quantity and quality of each RNA preparation (bottom in C and D). *Gsdma, Gsdmc,* and *Gsdmd* mRNAs were detected preferentially in the gastrointestinal tract (C). Proximal–distal axis-dependent expression of *Gsdmc* cluster genes and *Gsdmd* was detected in the intestinal tract (D). Exposure times are indicated in parentheses. M, DNA marker; Sk. muscle, skeletal muscle; P. Small int, proximal small intestine; D. Small int, distal small intestine; P. Colon, proximal colon; D. Colon, distal colon.

causing auditory disturbance (autosomal recessive, nonsyndromic bilateral, prelingual sensorineural hearing impairment) [17]. Therefore, both *DFNA5* and *DFNB59* play roles in sensorineural function, and mutations in these genes can lead to genetic forms of deafness in humans.

The genomic organization of *DFNA5* and *DFNB59* shares several similarities with that of the *Gsdm* family genes. The exon organization is well conserved in both groups (data not shown). Furthermore, both groups are located at chromosomal regions tightly linked to the *Hox* clusters. Thus, *DFNA5* and *DFNB59* and the *Gsdm* family genes most likely evolved from the same ancestor and diverged into the present forms through two rounds of duplications of the vertebrate genome. However, both *DFNA5* and *DFNB59* are found in diverse species from teleost fishes to human, but the *Gsdm* family genes are found only in birds to mammals. It is conceivable that the ancestral *Gsdm* genes that existed before teleost and mammal divergence were lost in the teleosts, amphibian, and reptile lineages.

The expression patterns of *DFNA5* and *DFNB59* differ from those of the *Gsdm* family genes. *DFNA5* is expressed predominantly in the placenta and weakly in the brain, heart, and kidney [13]. The mouse *DFNA5* homologue *Dfna5h* is expressed in the cochleae [13]. *Dfnb59* transcripts are detected in brain, eye, inner ear, heart, lung, kidney, liver, intestine, testis, and skeletal muscle of adult mice [18]. Thus, despite relatively similar sequences and genomic arrangements, the divergent expression and phenotypes associated with different mutants suggest that *DFNA5, DFNB59,* and *Gsdm* family genes have evolved distinct molecular functions.

Diversification of the Gsdm family genes

The mouse genome lacks a *GSDMB* homologue. This gene is found in primates and ruminants, but not rodents or carnivores, suggesting that a *GSDMB* ancestral gene appeared in the ruminants, but was deleted during the evolution of the carnivores and rodent lineage. A recently generated mammalian phylogenetic analysis placed rodents and primates in the same clade, exclusive of other mammals [24]. Indeed, based on the extensive published molecular biological data, rodents and primates are more closely related to each other than either group is to carnivores [24–28]. This implies that the deletion of *GSDMB* in the carnivores and rodents occurred independently.

Among the species examined, only the mouse genome contains *Gsdma* and *Gsdmc* gene clusters. In humans, *GSDMA* and *GSDMC* are solitary genes (Fig. 3B). Even *R. norvegicus,* which also belongs to the subfamily *Murinae,* has single *Gsdma, Gsdmc,* and *Gsdmd* genes. The ancestral gene of the mouse *Gsdmc* cluster is likely *Gsdmc1,* because all mammals have *Gsdmc1.* There is 97% amino acid similarity between

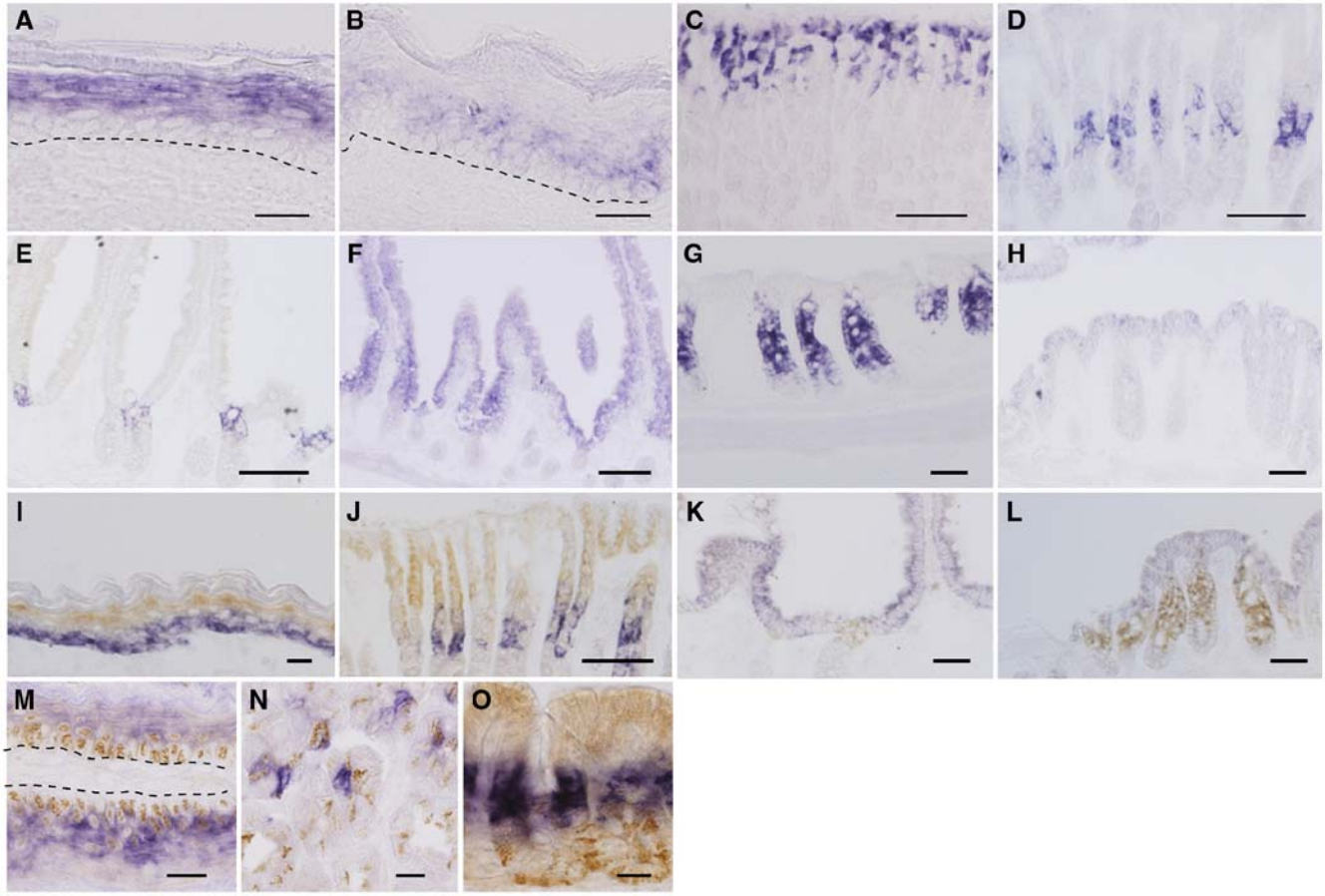


Fig. 6. The expression of *Gsdma, Gsdmc,* and *Gsdmd* genes in the gastrointestinal tract of adult mice. Expression of (A) *Gsdma1,* (D) *Gsdma2,* (B, G, J) *Gsdmc* cluster genes, and (H, K) *Gsdmd* was analyzed by section in situ hybridization. Sections of fore stomach (A, B), glandular stomach (D, E), small intestine (G, H), and colon (J, K) were analyzed. *Gsdma1* and *Gsdmc4* antisense cRNA probes cross-react with *Gsdma* and *Gsdmc* cluster transcripts, respectively. Control hybridization using sense-strand probes for *Gsdma1, Gsdmc4,* and *Gsdmd* gave no significant signal (data not shown). Broken lines in A and B indicate the basement membrane. Expression of two different *Gsdm* family genes was simultaneously analyzed in fore stomach (C; *Gsdma1,* orange staining; *Gsdmc* cluster genes, purple staining), glandular stomach (F; *Gsdma2,* orange staining; *Gsdmc* cluster genes, purple staining), small intestine (I; *Gsdmc* cluster genes, orange staining; *Gsdmd,* purple staining), and colon (L; *Gsdmc* cluster genes, orange staining; *Gsdmd,* purple staining) by double-labeled in situ hybridization. Sections of (M) cardia, (N) fundus, and (O) caecum of adult mouse were double labeled with *Gsdmc4* antisense cRNA probe (purple staining) and PCNA antibody (orange nuclear staining). Scale bars: A, B, I, M, N, O, 20 μm; C, D, E, F, J, 100 μm; G, H, K, L, 50 μm.

Gsdmc2 and *Gsdmc3,* suggesting that the duplication of these two genes occurred very recently.

Although *DFNA5* and *DFNB59* genes were mapped to different chromosomes in the human genome, the human *GSDMB* gene was located in reverse orientation and on the near side of the *GSDMA* gene. As in the case of *GSDMA* and *GSDMB,* human *GSDMC* and mouse *Gsdmc* cluster genes are located in reverse orientation and on the same chromosome as human *GSDMD* and mouse *Gsdmd,* respectively. These facts suggested that *Gsdm* family genes evolved from a single gene, which is the same ancestor of *DFNA5* and *DFNB59,* and were developed by chromosomal duplication and local segmental rearrangements. A local rearrangement has given rise to *GSDMB* and *GSDMC* from *GSDMA* and *GSDMD,* respectively, or *GSDMA* and *GSDMD* from *GSDMB* and *GSDMC.*

The *Gsdma* and *Gsdmc* loci appear to be the genomic regions frequently rearranged during mammalian evolution. Notably, the chromosomal regions syntenic to the mouse *Gsdma* and *Gsdmc* clusters overlap amplicons that are frequently amplified in human cancer cells. *GSDMA* overlaps the 17q21 amplicon, which is amplified in carcinomas of the skin, breast, stomach, ovary, bladder, and uterine cervix [29–32], and *GSDMC* overlaps the 8q24 amplicon found in acute myeloid leukemia [33]. Therefore, these genomic regions might be highly unstable and contain fragile sites.

Prospective function of the Gsdm family genes

Our ability to understand the function and mechanism of *Gsdm* family members is hampered by their unique primary sequence; the characteristics of the motifs and domains common to these proteins are unknown. *Gsdma3* is the causative gene for several mouse skin mutations, which exhibit hyperproliferation and altered differentiation of the epidermis [11,12]. These observations suggest that *Gsdm* family members, including *Gsdma3,* are involved in the regulation of cell growth and differentiation in the skin and/or hair follicle epithelium, as well as the epithelium of the gastrointestinal tract. A

combination of *Gsdm* family genes may specify the fine segmentation, achieving and/or maintaining the final differentiation state of epithelial cells.

The conserved sequences we identified in this study might be functionally important. *Re*^{den} has a mutation in the conserved sequence $LX_3LLX_5EX_{6-7}LX_4GL$ (VIII in Fig. 2A). In this mutation, duplication of the two amino acids 411Glu–412Ala (EA) immediately preceding the highly conserved 413Leu changed the distance between leucine-399 and leucine-413 in *Gsdma3,* a distance that is highly conserved in this sequence. In addition, five skin mutations, *Bsk, Re*^{den}*, Rco2, Dfl,* and *Fgn,* are all caused by mutations in the C-terminus in *Gsdma3* [11,12]. When considered together, these data suggest that the conserved C-terminal domain plays an essential role in the function of the *Gsdm* family genes, possibly in the regulation of development and/or homeostasis of epithelial cells.

Although further study is needed to understand fully the functions of *Gsdm* family genes, the study of *Gsdm* could reveal an important missing link in the molecular mechanisms that specify and maintain epithelial cell growth and fate in the skin and gastrointestinal tract.

Materials and methods

Mice

C57BL/6J strain mice were purchased from The Jackson Laboratory (Bar Harbor, ME, USA) and maintained at Genetic Strains Research Center, National Institute of Genetics (NIG). The animal experiments were approved by the Animal Care and Use Committee of NIG.

RNA isolation and RT-PCR

Total RNA was extracted from adult tissues of 3-month-old C57BL/6J males at several different stages with Isogen (Nippon Gene) and from C57BL/6J embryos with the RNeasy minikit according to the manufacturer's protocol. The extracted RNA was treated with RQ1 DNase (Promega). Two micrograms of total RNA was reverse-transcribed using Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's protocol.

PCRs were performed as follows: initial 1-min denaturing step at 94 °C followed by 25-30 cycles of 94 °C for 20 s, 56-58 °C for 30 s, 72 °C for 30 s. All reactions were performed in 50-μl volume. Primer sequences for detection of *Gsdma1* and *Gsdma2* were as follows: A1A2F1, 5′-ACAGAACGAGCTCTGGTTCC-3′; A1A2R1, 5′-TTGTGCAAGTTCTCCAGAGC-3′.

Bioinformatics

Homology searches of the GenBank (NCBI; http://www.ncbi.nlm.nih.gov) EST database and mouse and human genomic databases were conducted using BLAST. Evidence Viewer and Map Viewer tools were used for establishing the genomic structure of *Gsdm* family genes and for identifying mouse and human orthologues. The genome database versions we used were as follows: mouse, *Mus musculus* build 35.1; human, *Homo sapiens* build 35.1; rat, *Rattus norvegicus* build 3.1; cow, *Bos taurus* build 2.1; dog, *Canis familiaris* build 2.1; chicken, *Gallus gallus* build 1.1; chimpanzee, *Pan troglodytes* build 1.1. Alignments were performed using the program ClustalW with default values [34]. The sequence alignment shown in Fig. 2B was manually edited for display, but not for phylogenetic analysis or amino acid identity calculations. The phylogenetic tree was constructed using the neighbor-joining method [35] of the ClustalW multiple alignment package. Reliability of the tree topology was evaluated using bootstrap analysis [36] with 1000 iterations to provide confidence levels.

Northern blot analysis

Total RNA was extracted from adult mouse tissues with Isogen (Nippon Gene) or RNeasy minikit (Qiagen) according to the manufacturer's specifications. A sample of 15 μg of total RNA was separated in 1% agarose gel in the presence of 6.3% formaldehyde and Mops buffer (20 mM morpholinepropane sulfonic acid, 50 mM sodium acetate, 10 mM EDTA, pH 7.0). RNA was blotted on Hybond-N+ nylon membranes (Amersham Biosciences) using the upward capillary transfer method [37]. Digoxigenin (DIG)-labeled probes were synthesized by PCR according to the manufacturer's protocol (Roche). The open reading frame (ORF) sequences of *Gsdma1, Gsdmc4,* and *Gsdmd* were used as templates to prepare the probes. Due to high similarity among the clustering genes, the *Gsdma1* and *Gsdmc4* probes cross-reacted with the transcripts of the entire *Gsdma* and *Gsdmc* clusters, respectively. After Northern hybridization, signals were detected using alkaline phosphatase (AP)-coupled sheep anti-DIG antiserum, CDP-Star (Roche) reagent, and X-ray film (Amersham Biosciences).

Southern blot analysis

Genomic DNA from three mouse strains (C57BL/6J, 129/SvJ, and BALB/cAnN), two *R. norvegicus* strains (WKAN/Msfw and Sprague–Dawley), and one *R. rattus* strain (*R. rattus kandianus*) were obtained from the genome DNA collection at the Genetic Strains Research Center of NIG. After digestion with *Xba*I, 3 μg of each genomic DNA sample was transferred to Hybond-N+ nylon membrane after electrophoresis and hybridized with mouse *Gsdma1* or *Gsdmc4* genome fragment as a probe. Digoxigenin-labeled probes were synthesized by PCR according to the manufacturer's protocol (Roche). Primer sequences were as follows: Gsdma genome F1, 5′-GAATTGTGGTGGAACGGTAG-3′; Gsdma genome R1, 5′-CAGGACATCTTTGGGGAGTGC-3′; Gsdmc genome F1, 5′-GGTAAAGAGCATCTTACAGC-3′; Gsdmc genome R1, 5′-TATGCTATGGTGCAATATCT-3′. Signals were detected using AP-coupled anti-DIG antiserum and CDP-Star (Roche) reagent as described above.

In situ hybridization

Adult mouse GI tract samples fixed in 4% paraformaldehyde–phosphate-buffered saline (PFA/PBS) for 2–4 h were washed with 30% sucrose, embedded in OCT compound (Sakura Finetek), and frozen at −80 °C until sectioning. DIG- and fluorescein (Flu)-labeled cRNA probes were transcribed in vitro with ORFs of *Gsdma1, Gsdmc4,* and *Gsdmd.* Ten-micrometer sections were placed on MAS-coated glass slides (Matsunami) and dried immediately. Samples were refixed in 4% PFA/PBS. After treatment with proteinase K (Merck; 50 μg/ml in 50 mM Tris–HCl, pH 7.5, 5 mM EDTA) for 6 min at room temperature, the samples were acetylated in triethanolamine–HCl to reduce nonspecific hybridization of the probe. Prehybridization was performed in hybridization buffer (50% formamide, 5× SSC, 1 mg/ml yeast tRNA, 100 μg/ml heparin, 1× Denhardt's solution, 0.1% Tween 20, 0.1% Chaps, 5 mM EDTA) for 3 h at 60 °C, followed by hybridization overnight at 60 °C using DIG-labeled and/or Flu-labeled riboprobes. After hybridization, the slides were washed once in 1× SSC/0.3% Chaps at 60 °C for 10 min, 1.5× SSC/0.3% Chaps at 60 °C for 10 min, and twice in 2× SSC/0.3% Chaps at 37 °C for 20 min. To reduce background due to nonhybridized probes, the sections were treated with RNase (20 μg/ml RNase A in 2× SSC/0.3% Chaps at 37 °C for 30 min), followed by washing twice in 0.2 × SSC/0.3% Chaps for 30 min each at 60 °C and twice in PBS/0.1% Tween 20/0.3% Chaps for 10 min at 60 °C. Finally, the sections were washed in PBT (PBS, 0.1% Triton X-100, 2 mg/ml BSA) for 15 min at room temperature. For detection of DIG-labeled probes, the samples were incubated with alkaline phosphatase-conjugated anti-DIG Fab fragments (diluted 1:2000; Roche) in blocking solution. Bound antibodies were washed with PBT buffer three times for 15 min at room temperature. Alkaline phosphatase was activated in NTMT buffer (0.1 M Tris–HCl, pH 9.5, 0.1 M NaCl, 50 mM $MgCl_2$, 0.1% Tween 20) for 10 min at room temperature. Color development was accomplished with 5-bromo-4-chloro-3-indolyl phosphate (BCIP)/nitroblue tetrazolium (Roche) in NTMT buffer according to the manufacturer's instructions. Before being mounted in glycerol, the sections were washed in PBS and fixed in 4% PFA for 10 min. To visualize mRNA expression of two distinctive genes, the sections were hybridized

simultaneously with a Flu-labeled cRNA and a different DIG–labeled cRNA. First detection of DIG-labeled probe was performed as described above, and the sections were then washed three times in PBS for 5 min each. To block the AP activity of anti-DIG-AP, the sections were incubated twice in TE buffer (100 mM Tris–HCl, pH 7.5, 50 mM EDTA) at 85 °C for 10 min each. Visualization of the second Flu-labeled probe was performed by an overnight incubation with AP-coupled anti-Flu antiserum diluted to 1:1000 in 20% sheep serum/PBT. After being washed four times in PBT for 15 min each and three times in AP buffer for 5 min each, the hybridized antisense cRNA was detected by iodonitro tetrazolium/BCIP solution (Roche). In some experiments, Flu-labeled riboprobes were detected prior to visualization of DIG-labeled probes.

Expression of *Gsdmc* cluster genes and proliferating cells was determined by a combination of in situ hybridization and immunohistochemistry. After in situ hybridization, sections were blocked with 1% skim milk in PBS followed by incubation with PCNA rabbit polyclonal antiserum (Santa Cruz Biotechnology) diluted 1:50 in PBS. Slides were then reacted with biotin-labeled anti-rabbit IgG and incubated with preformed avidin–biotin peroxidase complex. The antigen site was visualized by incubating the sections with the substrate 3,3′-diaminobenzidine (DAB). Negative control sections showed no nuclear staining with DAB.

Acknowledgments

We thank K. Hiratsuka, K. Aida, F. Murofushi, Y. Yamazaki, and A. Sugiyama for maintaining the mouse strains; Y. Kato, J. Iwata, Y. Mizushina, H. Kobayashi, N. Sakurai-Yamatani, H. Yamamoto, and A. Okagaki for excellent technical assistance and valuable advice. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. This paper is contribution No. 2505 from the National Institute of Genetics, Japan.

References

[1] S. Ohno, Evolution by Gene Duplication, Allen & Unwin, London, 1970.
[2] A. Wagner, The fate of duplicated genes: loss or new function? BioEssays 20 (1998) 785–788.
[3] M. Lynch, J.S. Conery, The evolutionary fate and consequences of duplicate genes, Science 290 (2000) 1151–1155.
[4] A. Force, et al., Preservation of duplicate genes by complementary, degenerative mutations, Genetics 151 (1999) 1531–1545.
[5] A. Wagner, Birth and death of duplicated genes in completely sequenced eukaryotes, Trends Genet. 17 (2001) 237–239.
[6] S.L. Pollard, P.W. Holland, Evidence for 14 homeobox gene clusters in human genome ancestry, Curr. Biol. 10 (2000) 1059–1062.
[7] V.E. Prince, The hox paradox: more complex(es) than imagined, Dev. Biol. 249 (2002) 1–15.
[8] N. Saeki, Y. Kuwahara, H. Sasaki, H. Sato, T. Shiroishi, Gasdermin (Gsdm) localizing to mouse chromosome 11 is predominantly expressed in upper gastrointestinal tract but significantly suppressed in human gastric cancer cells, Mamm. Genome 11 (2000) 718–724.
[9] H. Sato, et al., A new mutation *Rim3* resembling *Re*^{den} is mapped close to retinoic acid receptor alpha (*Rara*) gene on mouse chromosome 11, Mamm. Genome 9 (1998) 20–25.
[10] M. Katoh, M. Katoh, Evolutionary recombination hotspot around GSDML–GSDM locus is closely linked to the oncogenomic recombination hotspot around the PPP1R1B–ERBB2–GRB7 amplicon, Int. J. Oncol. 24 (2004) 757–763.
[11] F. Runkel, et al., The dominant alopecia phenotypes Bareskin, Rex-denuded, and Reduced Coat 2 are caused by mutations in gasdermin 3, Genomics 84 (2004) 824–835.
[12] D.P. Lunny, et al., Mutations in gasdermin 3 cause aberrant differentiation of the hair follicle and sebaceous gland, J. Invest. Dermatol. 124 (2005) 615–621.
[13] L. Van Laer, et al., Nonsyndromic hearing impairment is associated with a mutation in DFNA5, Nat. Genet. 20 (1998) 194–197.
[14] K. Watabe, et al., Structure, expression and chromosome mapping of *MLZE,* a novel gene which is preferentially expressed in metastatic melanoma cells, Jpn. J. Cancer Res. 92 (2001) 140–151.
[15] Y. Yu, et al., Gene expression profiling in human fetal liver and identification of tissue- and developmental-stage-specific genes through compiled expression profiles and efficient cloning of full-length cDNAs, Genome Res. 11 (2001) 1392–1403.
[16] M. Katoh, M. Katoh, Identification and characterization of human DFNA5L, mouse Dfna5l, and rat Dfna5l genes in silico, Int. J. Oncol. 25 (2004) 765–770.
[17] S. Delmaghani, et al., Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy, Nat. Genet. 38 (2006) 770–778.
[18] M.H. Kaufman, The Atlas of Mouse Development, Academic Press, London, 1992.
[19] K. Theiler, The House Mouse. Atlas of Embryonic Development, Springer-Verlag, 1989.
[20] T. Sekimoto, et al., Region-specific expression of murine Hox genes implies the Hox code-mediated patterning of the digestive tract, Genes Cells 3 (1998) 51–64.
[21] Y. Kawazoe, et al., Region-specific gastrointestinal Hox code during murine embryonal gut development, Dev. Growth Differ. 44 (2002) 77–84.
[22] C. Yu, et al., A 3-nucleotide deletion in the polypyrimidine tract of intron 7 of the DFNA5 gene causes nonsyndromic hearing impairment in a Chinese family, Genomics 82 (2003) 575–579.
[23] A.M. Bischoff, et al., A novel mutation identified in the DFNA5 gene in a Dutch family: a clinical and genetic evaluation, Audiol. Neurootol. 9 (2004) 34–46.
[24] W.J. Murphy, et al., Resolution of the early placental mammal radiation using Bayesian phylogenetics, Science 294 (2001) 2348–2351.
[25] O. Madsen, et al., Parallel adaptive radiations in two major clades of placental mammals, Nature 4096 (2001) 610–614.
[26] W.J. Murphy, et al., Molecular phylogenetics and the origins of placental mammals, Nature 4096 (2001) 614–618.
[27] U. Arnason, et al., Mammalian mitogenomic relationships and the root of the eutherian tree, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 8151–8156.
[28] A. Reyes, et al., Congruent mammalian trees from mitochondria and nuclear genes using Bayesian methods, Mol. Biol. Evol. 21 (2004) 397–403.
[29] H. Sasano, et al., Protooncogene amplification and tumor ploidy in human ovarian neoplasms, Hum. Pathol. 21 (1990) 382–391.
[30] L.M. Coombs, et al., Amplification and over-expression of c-erbB-2 in transitional cell carcinoma of the urinary bladder, Br. J. Cancer 63 (1991) 601–608.
[31] T. Kihana, et al., Prognostic significance of the overexpression of c-erbB-2 protein in adenocarcinoma of the uterine cervix, Cancer 73 (1994) 148–153.
[32] R. Wolber, B. Dupuis, M.R. Wick, Expression of c-erbB-2 oncoprotein in mammary and extramammary Paget's disease, Am. J. Clin. Pathol. 96 (1991) 237–243.
[33] C.T. Storlazzi, et al., Identification of a commonly amplified 4.3 Mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies, Hum. Mol. Genet. 13 (2004) 1479–1485.
[34] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[35] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425.
[36] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.
[37] J. Sambrook, D.W. Russell, Molecular Cloning: a Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, New York, 2001.