GENOME-WIDE IDENTIFICATION OF MAMMALIAN

DEFENSINS AND FUNCTIONAL ANALYSIS OF

A NOVEL DEFENSIN-RELATED PEPTIDE

By

AMAR AJIT PATIL

Bachelor of Veterinary Science & Animal Husbandry
Dr. P.D.K.V. Agricultural University
Akola, Maharashtra, India
2000

Master of Veterinary Science in Animal Biotechnology
Indian Veterinary Research Institute
Izatnagar, Uttar Pradesh, India
2002

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2006

GENOME-WIDE IDENTIFICATION OF MAMMALIAN

DEFENSINS AND FUNCTIONAL ANALYSIS OF

A NOVEL DEFENSIN-RELATED PEPTIDE

Dissertation Approved:

Guolong Zhang
Dissertation Adviser

Clinton Krehbiel
Chair

Haobo Jiang

Lin Liu

A. Gordon Emslie
Dean of the Graduate College

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER V

CHAPTER I

INTRODUCTION

Antimicrobial peptides (AMPs) are found in virtually every living being as a part of the innate immune system. In vertebrates, there are two major families of AMPs, namely defensins and cathelicidins. They kill a broad range of pathogens ingested by phagocytes and also help fight off pathogens trying to enter through mucosal surfaces and skin. The antibacterial action of AMPs is mediated primarily through binding and disruption of microbial membranes.

In addition to the antimicrobial activity, they exert various modulatory effects on the innate and adaptive immune system. Many AMPs are directly chemotactic for neutrophils, monocytes, and lymphocytes. They are capable of inducing the production of an array of proinflamatory cytokines and chemokines from immune and epithelial cells. They also play a role in differentiation and maturation of dendritic cells, which are professional antigen-presenting cells of the immune system. Paradoxically, some AMPs also have anti-inflammatory effects by preventing the release of tumor necrosis factor-α from the immune cells induced by lipopolysaccharide (LPS), a major cell wall component of Gram-negative bacteria. Due to their pleotropic effects on the immune system, AMPs are also being called host defense peptides (HDPs).

1

Recently, there is a sharp rise in the incidence of antibiotic-resistant infections with some bacteria even showing resistance to multiple antibiotics. Due to their potent antimicrobial and immunomodulatory effects, AMPs are being actively explored as potential antibacterial drugs especially against antibiotic-resistant bacteria. It is extremely difficult for bacteria to gain resistance to these peptides as opposed to traditional antibiotics. Several of such peptides are in clinical trials to treat both topical and systemic infections.

In this report, we present the latest advances on defensins and cathelicidins pertaining to their pleotropic activities and potential as antibacterial drugs (Chapter II). We have discovered the complete repertoire of the α- and β-defensin gene families in various mammalian species including the human, rat, mouse, chimpanzee and dog through a comprehensive bioinformatics approach (Chapter III & IV). Furthermore, we characterized *in vitro* antibacterial properties of rattusin, a novel α-defensin-related peptide in rat (Chapter V). Overall findings are summarized and future prospect are also discussed (Chapter VI).

CHAPTER II

REVIEW OF LITERATURE

Discovery of penicillin and a host of other antibiotics in the middle of last century led to a decrease in infectious disease mortality by 20-fold from year 1900 to 1980 (170). Since then, widespread and sometimes misuse of antibiotics in clinical practice and in the food animals has caused bacteria to quickly acquire resistance to antibiotics (31, 41, 79, 167). These days, >70% hospital-acquired bacterial infections show resistance to at least one antibiotic (www.niaid.nih.gov/factsheets/antimicro.htm). For example, resistant strains have been found with *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Enterococci*, *Neisseria gonorrhoeae*, *Salmonella enterica* serovar Typhi (http://www.cdc.gov/drugresistance/diseases.htm), and certain bacteria have become resistant to multiple antibiotics (125).

Furthermore, antibiotic resistance increases the cost of treatment through longer duration of hospitalization and use of expensive drugs to treat infection. The estimated cost of treating antibiotic-resistant infections in the US is $30 billion (101). This cost keeps on rising as the number of bacterial strains resistant to antibiotics are increasing and the pharmaceutical industry is unable to bring novel antibiotics to treat resistant infections (174). Therefore, discovery of novel antimicrobial drugs with a very low risk of resistance development are urgently required. Cationic host defense peptides (HDPs)

constitute a potential novel class of antimicrobials by virtue of their bactericidal as well as immunomodulatory properties (32, 87, 88).

These HDPs are small (12 to 50 amino acids), cationic (+ 2 to + 10 charge due to the presence of multiple arginine and lysine residues), and amphipathic. They are capable of killing Gram-positive and Gram-negative bacteria and fungi and inhibiting the replication of enveloped viruses like influenza A virus, human papilloma virus, vesicular stomatitis virus and human immunodeficiency virus (58, 63). Antibacterial activity of HDPs is exerted by physical disruption of bacterial membranes mainly (15). Due to this physical mechanism of action, cationic HDPs kill a broad range of bacteria including antibiotic-resistant strains with a low risk of developing resistance. This property combined with their modulatory effects on innate and adaptive immune system, have made them very attractive candidates for antibacterial drug development, particularly against antibiotic-resistant strains (43, 44, 87, 197). This review will focus on the antibacterial and immunomodulatory properties of vertebrate HDPs.

CLASSIFICATION OF HOST DEFENSE PEPTIDES

Vertebrates mainly produce two major families of HDPs called defensins and cathelicidins (36, 72, 140, 162, 194, 195). Based on the structure difference, HDPs can be classified into distinct classes such as α-helical, β-sheet, cyclic and extended peptides (87). Defensins are amphipathic molecules characterized by the presence of six conserved cysteine residues, and a number of hydrophobic and cationic residues. Defensins are

4

further divided into three different subclasses, namely α-, β- and θ-defensins (36, 72, 140). This classification is based on the spacing of cysteine residues and the pattern of disulfide bonds. α-defensins consist of three disulfide bonds between Cys1-Cys6, Cys2-Cys4 and Cys3-Cys5, whereas in β-defensins, cysteines are connected between Cys1-Cys5, Cys2-Cys4 and Cys3-Cys6. α-defensins are mainly produced by leukocytes and Paneth cells of small intestine, whereas β-defensins are expressed mainly by mucosal epithelial cells and skin (36, 72, 140). In humans, there are four neutrophil α-defensins [human neutrophil peptide (HNP)-1 to 4] and two enteric α-defensins [human defensin (HD)-5 and 6] (113, 124, 127). In rhesus monkey, myeloid as well as intestinal α-defensins have been identified (154, 155). While multiple enteric α-defensins are present in mice, neutrophil α-defensins are absent (30, 114). Rat neutrophils express three α-defensins, and mRNA for six additional α-defensins have been identified in small intestine (120). Over 35 β-defensins have been identified in humans and rodents (121, 123).

Tertiary structures of several α- and β-defensins are available (4, 48, 51, 52, 59, 69, 86, 119, 135, 137, 145, 204). All defensins possess triple stranded antiparallel β-sheet with a hairpin conformation stabilized by three disulfide bonds. In the case of human β-defensin (HBD)-1 to 3, mouse β-defensin (MBD)-7, and 8 and penguin sphenecin, there is an additional N-terminal α-helical segment involving the first cysteine (4, 51, 52, 69). However, such a segment is absent in bovine neutrophil β-defensin (BNBD)-12 (204). α-defensins also lack the N-terminal helical segment. Cryptdin-4, which is a potent mouse Paneth cell α-defensin, has a different orientation of β-hairpin and lower β-sheet content

than other defensins. These unique structural features with its high cationicity have been suggested to be responsible for high antibacterial potency of cryptdin-4 (59). All α- and β-defensins possess a conserved GXC motif involving cysteine 4, which forms a bulge essential for correct folding of defensins (182, 191).

θ-defensins are cyclic peptides expressed by leukocytes of old world primates (156) and possess activity against a wide spectrum of bacteria, fungi, and viruses (156, 172, 190). They are produced by ligation of two different α-defensin-like peptides with premature termination of the open reading frame between the third and fourth conserved cysteine (156). Each peptide thus contributes three cysteines and cyclizes in a head-to-tail manner with formation of three disulfide bonds (156). NMR structure of rhesus θ-defensin (RTD)-1 shows the presence of two flexible β-strands connected by a hairpin at both ends (165). Unlike other classes of host defense peptides, RTD-1 is not amphipathic (165). In humans and chimpanzee, θ-defensin genes are pseudogenes with a stop codon in the signal sequence leading to early termination of translation (100, 121).

Cathelicidins constitute another major family of HDPs in vertebrates and are characterized by the presence of an N-terminal conserved pro-sequence called cathelin (162, 194, 196). This cathelin pro-sequence is connected to a 20-80-residue, diversified cationic segment at the C-terminus (162, 194, 196). Both cathelin and C-terminal mature peptide play a role in host defense (195). Cathelicidins are stored in the granules of neutrophils. In addition, they are also expressed by skin and various mucosal epithelial surfaces. To date, cathelicidins have been found in a number of vertebrate species

including primates, glire, canine, porcine, bovine, ovine, caprine, equine, chicken and fish (21, 22, 162, 180). In cattle, sheep, and pigs, multiple cathelicidins are present, whereas in primates and glire only one member is present in each species (162, 200).

Even though all cathelicidins are amphipathic and cationic, they adopt either α-helical, β-sheet, or extended structures. Examples of α-helical peptides include human LL-37, mouse CRAMP, sheep SMAP-29, and chicken fowlicidins (110, 149, 151, 180, 192). Guinea pig CAP-11 possesses a single cysteine towards the C-terminus and exists in the dimeric form due to the formation of an intermolecular disulfide bond (98). Some cathelicidins are rich in certain amino acids, which render them extended structures. Extended peptides include bovine indolicidin (Trp and Pro rich), porcine PR-39 (Pro and Arg rich), and porcine prophenin (Pro and Phe rich) (18, 45, 68). Porcine protegrins and bovine and ovine dodecapeptides are looped β-sheet peptides due to the presence of intramolecular disulfide bonds (66, 177). Protegrins consist of four cysteines forming two intramolecular disulfide bonds, whereas bovine and ovine dodecapeptides have two cysteines each forming one disulfide bond.

PLEOTROPIC EFFECTS OF HOST DEFENSE PEPTIDES

Defensins and cathelicidins are either produced constitutively or induced in response to infection or injury (194, 195). All these released HDPs exert a host defense role by directly killing a broad spectrum of bacteria. Recent evidence also shows a novel way of antibacterial action, in which macrophages acquire granules of apoptotic

neutrophils, thereby acquiring HDPs to kill intracellular *Mycobacterium tuberculososis* (152).

Studies involving knock-out and transgenic mice have provided convincing evidence showing an indispensable role of HDPs in protection against infections. Mice deficient in *matrix metalloproteinase 7* (*MMP-7)*, which is responsible for conversion of inactive precursors of enteric α-defensins into active forms, become more susceptible to a lethal *S. typhimurium* infection (175). Similarly, mice deficient in *β-defensin-1* (*Defb1*) show decreased clearance of *Staphylococci* in the bladder (91) and exhibit enhanced colonization of *H. influenzae* in the lung (92). Role of cathelicidin (mouse CRAMP) in protecting skin infection from Group A *Streptococci* is evident by the presence of large and persistent skin lesions in mice lacking *mCRAMP* (105). Conversely, transgenic mice possessing porcine cathelicidin *PR-39* display increased resistance to skin infections by Group A *Streptococci* (71). Transgenic mice expressing human α-defensin HD-5 are resistant to a lethal challenge of *S. typhimurium* infection (132).

Different models have been proposed to explain the mechanism of antibacterial action of HDPs (15). Initial interaction of cationic peptide occurs with negatively charged molecules on bacteria such as lipopolysacharide (LPS) of Gram-negative bacteria and lipoteichoic acid of Gram-positive bacteria, followed by formation of either transient or stable pores across membranes. Rabbit neutrophil defensins form short-lived pores, whereas human neutrophil defensins form stable channels (60, 77, 176). The formation of stable channels correlates with oligomerization of human neutrophil defensins in contrast

to the monomeric form adopted by rabbit neutrophil defensins. Mouse cryptdin-4 interacts with phospholipid head groups through charge-charge interactions between the cationic peptide and negatively charged membrane lipids and disrupts bilayer head groups (133, 134). Rhesus θ-defensin-1 (RTD-1), which lacks an amphipathic structure, has been shown to induce membrane thinning and formation of pores across membranes (17, 173). Additionally, defensins can also interfere with intracellular processes. Human neutrophil peptides 1-3 (HNP 1-3) completely inhibit the synthesis of DNA, RNA and protein at sub-lethal concentrations (73). Killing of bacteria by defensins can also occur through activation of cell-wall lytic enzymes (131).

In the case of cathelicidins, the mechanism of action on model membranes is studied in more detail than that for defensins. Human LL-37 disrupts model membranes through formation of transmembrane torroidal pores lined by polar lipid head groups and peptides (47). Porcine protegrin-1, a β-sheet cathelicidin, also permeabilizes model membranes through formation of torroidal pores (82, 83, 183). Ovispirin, a peptide derived from ovine cathelicidin SMAP-29, permeabilize bacterial membranes through the "carpet" mechanism by binding to lipid bilayer in parallel, leading to the formation of micelles when reaching the threshold concentrations (184). Cathelicidins also kill bacteria by interfering with intracellular processes. Porcine PR-39 kills bacteria by inhibition of bacterial DNA and protein synthesis (8). PR-26, an N-terminal version of PR-39, causes filamentation of *S. typhimurium*, suggesting the inhibition of cell division (142). Bovine indolicidin also inhibits DNA synthesis resulting in filamentation of

Escherichia coli (150). Indolicidin has been shown to directly bind HIV-1 DNA through covalent cross-linking, resulting in inhibition of integrase activity (84).

In addition to direct bacterial killing, HDPs exert important modulatory activities on the innate and adaptive immune system. Defensins and cathelicidins possess chemotactic activities on various immune cells. HNPs are chemotactic for monocytes, T cells, and immature dendritic cells (24, 158, 185). The identities of receptors used by HNPs are not clear. HBD-2 is chemotactic for immature dendritic cells, memory T cells, TNF-α-treated neutrophils, and mast cells through interactions with CCR-6 chemokine receptor (102, 186). HBD-3 and HBD-4 are chemotactic for monocytes (37, 38). Similarly, murine β-defensins are also capable of attracting immature dendritic cells and help induce antitumor immunity (6). LL-37 is chemotactic for neutrophils, monocytes, T cells and mast cells (29, 103). PR-39 is a chemoattractant for neutrophils (55). Bovine-bactenecin derived peptide, Bac2A, is a chemoattractant for monocytes (12). Mouse CRAMP is a chemoattractant for neutrophils and monocytes (67). Cathelicidins induce chemotaxis by binding to formyl peptide receptor. Recruitment of immune cells to the site of infection leads to elicitation of cell-mediated and humoral immune responses to fight off infection.

HDPs are also able to induce the release of cytokines and chemokines from immune cells and epithelial cells. Human neutrophil defensins up-regulate chemokine IL-8 production by human airway epithelial cells (168). Cryptdin-3, but not cryptdin-4, induces secretion of IL-8 from human intestinal cells (76). In peripheral blood

mononuclear cells (PBMCs), HBD-2 induces IL-8, IL-6, IL-10, MCP-1, MIP-1β, RANTES, IL-1β, ENA-78, GRO, IGFBP-3 and EGF, whereas HBD-1 induces IL-8, IL-10, IL-6 and MIP1β, and HBD-3 stimulates IL-8, MCP-1 and IGFBP-3 production (9). HBD-2, 3 and 4, but not HBD-1, induces expression of IL-6, IP-10, IL-10, RANTES, MCP-1, MIP-3α at both mRNA and protein levels in skin keratinocytes (104). LL-37 induces secretion of IL-8, IL-6, IL-1β, and TNF-α in human corneal endothelial cells (56) and IL-1α, IL-6, IL-8, TNF-α, and GM-CSF in keratinocytes (13). LL-37 also enhances IL-8 release from human airway smooth muscle cells (205). Thus, chemokines and proinflammatory cytokines released by different cell types under influence of HDPs might help recruit immune cells and modulate the microenvironment, ultimately leading to robust immune responses.

Given the immunostimulatory activities exhibited by HDPs, one would expect that they have adjuvant-like effects on the immune system. Indeed, both defensins and cathelicidins potentiate immune responses to particularly non-immunogenic antigen. HNPs administered together with an antigen induce antigen-specific IgG response in mice (75). HNPs help induce tumor-specific IgG and T cell responses, thereby increasing resistance of mice to a tumor challenge (157). HBD-1 also induces antigen-specific IgG response in mice (16). A DNA vaccine encoding a fusion protein of non-immunogenic lymphoma antigen with mouse β-defensin-2 and 3 generates tumor-specific protective humoral and cell-mediated immunity (6). Similarly, LL-37, when fused with a tumor antigen and administered as a DNA vaccine, generates humoral and cellular immune responses and increases survival against a tumor challenge (3). However, the mechanism

by which these HDPs exert adjuvant-like effects remains elusive (10). It is likely that defensin and cathelicidins act as a carrier and target the fusion partner to antigen-presenting cells by binding to chemokine receptors. It is also possible that these HDPs alter the micro-environment by inducing cytokines and chemokines, thereby influencing differentiation and maturation of dendritic cells indirectly. The third scenario is that HDPs directly interact with dendritic cells and modulate their differentiation. Nevertheless, further research is needed to study the immunomodulatory effect of HDPs to facilitate the development of HDPs as adjuvants for vaccine use (5).

Conversely, some HDPs are also able to suppress the production of proinflammatory cytokines such as TNF-α and IL-1β in response to LPS stimulation. Cathelicidins, including human LL-37, bovine indolicidin, rabbit CAP18, and chicken fowlicidins, are thought to neutralize the effects of LPS through direct binding to and competing with LPS for its receptor, CD14 (128, 180). However, recent evidence indicates that LL-37 and bovine BMAP-27 suppress LPS-induced responses by inhibiting the transcription of proinflammatory genes (89, 90). Both peptides reduce LPS-induced translocation of NF-κB subunits, p65 and p50, into nucleus. Thus, overall neutralization of LPS effects seems to occur through multiple mechanisms (89). Protection against LPS-induced sepsis by cathelicidins has been demonstrated in animal models (26, 33, 39, 40). In contrast, only one human β-defensin, DEFB-123, has been found to prevent LPS-induced TNF-α secretion *in vitro* and protect mice against LPS-induced septic shock (93).

In addition to endotoxins, several defensins were recently shown to be able to neutralize several types of exotoxins. HNPs neutralize anthrax lethal toxin and protect mice against lethal effects of the toxin (61). HNP-1 inhibits lethal toxin by inhibiting the cleavage of mitogen-activated protein kinase kinase by toxin. HNPs neutralize diphtheria toxin and *Pseudomonas* exotoxin A by competitively inhibiting toxin-mediated ADP-ribosylation of eukaryotic elongation factor 2 (62). θ-defensins also neutralize anthrax lethal toxin, in addition to killing bacilli and germinating spores of *Bacillus anthracis* (171).

HDPs also play a role in the resolution of inflammation and wound healing. HNPs stimulate proliferation of epithelial and fibroblast cells (1, 96) and enhance wound closure of the airway (2). HBD-2, -3 and -4 promote wound healing by promoting keratinoycte migration and proliferation through binding to epidermal growth factor receptor (EGFR) (104). Cathelicidin LL-37 promotes angiogenesis and neo-vacularization of endothelial cells (65) and is involved in re-epithelization of skin wounds (46). LL-37 induces keratinocyte migration through transactivation of EGFR (161). LL-37 also stimulates airway epithelial cell proliferation and wound closure through EGFR (141). Rat CRAMP promotes healing of gastric ulcers by inducing epithelial cell proliferation through TGF-α-dependant transactivation of EGFR (189). Porcine PR-39 facilitates the formation of new functional blood vessels in a mouse model through inhibition of proteasome-mediated degradation of hypoxia induced factor 1-α protein (74). PR-39 also helps in wound healing through induction of syndecan-1 and 4, which

are important in regulating cell proliferation and migration in response to heparin binding growth factors (35).

STRUCTURE-ACTIVITY RELATIONSHIPS

Structural and physicochemical properties such as disulfide bond, charge, length, helicity, hydrphobicity and amphipathicity play a critical role in antibacterial, immunomodulatory, and cytotoxic activities of HDPs. It is possible to design peptides with improved antibacterial and immunomodulatory activities and reduced cytotoxicity for therapeutic application by varying one or more of the above properties.

In the case of defensins, conserved disulfide bonds are dispensable for the antibacterial activity. Defensin variants lacking disulfide bonds or having a varying number of disulfide bonds display similar antibacterial activities to the parent peptide (53, 64, 80, 81, 178). Absence of disulfide bonds also does not affect the cytotoxicity of defensins (64). However, the disulfide motif is required for *in vivo* stability of defensins. Absence of disulfide bonds renders cryptdin-4 susceptible to MMP-7, a protease which converts cryptdin-4 precursors into active forms (80, 143). Another function of disulfide bonds in defensins is to maintain the chemotactic activity as observed for HBD-3 (178). Furthermore, even with intact disulfide bonds, mutations in N-terminal helix containing primarily hydrophobic residues of HBD-1 result in a significant reduction of its binding to the CCR-6 receptor, underscoring the importance of these residues in chemotaxis (122). A conserved salt bridge in HNP-2, which is dispensable for the antibacterial

activity, is necessary to maintain its stability against protease degradation (179). The same conserved salt bridge in cryptdin-4 is not required for the antibacterial activity or proteolytic stability (129).

Antibacterial activity of defensins is mainly dictated by net positive charges. In the case of cryptdin-4, charge reversal from cationic arginines to anionic aspartic acids abolishes the antibacterial activity, irrespective of the location of arginine (153). There is a direct positive correlation between the number of charged residues and antibacterial activity as seen with HBD-1 (122). A combination of strong cationicity and moderate hydrophobicity is necessary for the antibacterial activity of HBD-3, whereas high hydrophobicity results in an increase in cytotoxicity to mammalian cells (64). Arginine residues in defensins also have a role in modulating *in vivo* functions. ADP ribosylation of arginines of HNP-1 in the respiratory tract results in a decreased cytotoxicity and antibacterial activity without altering the chemotactic and IL-8-inducing ability (116).

In the case of α-helical peptides, an appropriate balance of the net positive charge, α-helical content and amphipathicity dictate the antibacterial activity (28, 163). Recently, this was demonstrated with a group of *de novo* synthesized peptides based on the template obtained by comparing naturally occurring α-helical peptides (198, 199, 201). These synthetic peptides show reduced potency when the net positive charge is decreased to +3 or below. Increasing the charge above +6 fails to enhance the antibacterial potency. Additionally, peptides that are both α-helical and amphipathic have broad-spectrum antibacterial activity. However, peptides with a rigid α-helix possess a strong hemolytic

activity and a poor antibacterial activity. An increase in the flexibility of α-helix by introducing a glycine or proline is generally associated with a decrease in cytotoxicity. Peptides also display a decreased cytotoxicity with moderate hydrophobicity, and a further decrease in hydrphobicity reduces the antibacterial activity.

Most of the above described principles are also true for naturally occurring α-helical cathelicidins. In the case of ovispirin, which is the N-terminal α-helical segment of SMAP-29 with strong hemolytic and cytotoxic activities (136), a replacement of either isoleucine 10 with glycine or isoleucine 7 with threonine enhances the flexibility and substantially reduces hemolytic and cytotoxic activities (136). In contrast, despite being a rigid helical peptide, CAP-18 shows a low hemolytic activity. Antibacterial activity of CAP-18-derived peptides with moderate to high helicity correlates well with the net positive charge than helicity (23, 166). CAP-18-derived peptides with very low helicity exhibit reduced antibacterial activity, despite introductions of additional positively charged residues (164). Some peptides contain a hinge around the center of α-helix consisting of PXXP or PXXXP motif. This hinge divides the peptide into the N- and C-terminal helices. When the N-terminal helix binds to membrane, the hinge bends and allows the insertion of the C-terminal helix into lipid bilayer (187). In the case of highly hydrophobic BMAP-27 and 28, maintenance of the N-terminal amphipathic α-helix while reducing hydrphobicity diminishes the cytotoxicity without changing the antibacterial activity (146). Removal of the four N-terminal hydrophobic residues of LL-37 gives rise to a less cytotoxic derivative without affecting membrane permeabilization or the antibacterial activity (110).

Successful attempts have been made to identify short fragments of naturally occurring α-helical peptides with improved antibacterial and immunomodulatory activities and less cytotoxicity. LL-37-derived peptides with N-terminal hydrophobic amino acids removed, display similar antibacterial, LPS-neutralizing, and chemotactic activities as LL-37, but with significantly reduced cytotoxicity (25, 99, 144). Recently, a large panel of LL-37-derived short fragments were compared for their antibacterial and immunostimulatory activities (13). Two LL-37 fragments, namely RK-31 and KS-30, possess a dramatically improved antibacterial activity with little capacity to stimulate the release of IL-8 from keratinocytes. These two fragments also display slightly reduced toxicity. Interestingly, both fragments have been reported to be present in human sweat, suggesting this gain of function demonstrated *in vitro* has *in vivo* implications (95).

For guinea pig CAP-11 peptide, the antibacterial activity of the N-terminal fragment of 18 residues can be further improved by increasing amphipathicity (109). Interestingly, the C-terminal α-helix of chicken fowlicidin-1 is the major region for antibacterial, cytotoxic and LPS-neutralizing activities with three hydrophobic residues just before N-terminal α-helix also sharing cytotoxic and LPS-binding activities. Removal of these three hydrophobic residues results in an almost 10-fold reduction in cytotoxicity and 60-fold reduction in hemolysis without significantly affecting LPS-neutralizing and antibacterial activities (181). Angiogenic and NF-κB-inhibiting activities of PR-39 have been attributed to the N-terminal 11 residues with three arginines being essential for both activities (34). Furthermore, the N-terminal arginines of PR-39 are critical for the antibacterial activity, induction of syndecans and interaction with the SH3

domain (20), whereas the C-terminal part of PR-39 is important only for antibacterial activity but not for interactions with the SH3 domain in mammalian cells (20).

STRATEGIES FOR PEPTIDE DESIGN AND SYNTHESIS

Although HDPs have broad-spectrum antibacterial activities, the issues related to antibacterial potency, stability, toxicity, cost and salt sensitivity need to be addressed in order to develop them into therapeutic drugs. A number of strategies for peptide design and synthesis are evolving. Besides structure-based rational design as depicted in the previous section, other strategies include amidation, acylation, cyclization, dimerization, D-enatiomerization, development of peptide mimetics, and high throughput synthesis and screening of synthetic peptides.

Various types of modifications of peptide termini, backbone or amino acid residues are being tried to improve the antibacterial activity and stability of peptides. Amidation of the C-terminus has been shown to increase the antibacterial activity of porcine cathelicidin, tritrpticin (138, 188). N-terminal acylation or C-terminal amidation also protects peptide from being cleaved by exopeptidases. Such *de novo* designed peptides with both termini modified retain the antibacterial activity (99). Peptide backbone can be modified by introducing carbamate bonds or reduced amide bonds (70, 107). Reduced amide bond increases the net positive charge and antibacterial activity. Introduction of carbamate bonds reduces helicity and the hemolytic activity. Another modification of backbone is the inclusion of peptoid residues, which are helix-disrupting

amino acids. Such modifications can decrease the cytotoxicity of HDPs with increased activity towards bacteria (147, 203). Backbone modification also increases half lives of peptides in serum as they become less susceptible to proteolysis (107). Incorporation of unnatural amino acids into synthetic antibacterial peptides has been tried successfully to increase the resistance of peptides to serum proteases without affecting the antibacterial activity (108).

Incorporation of D-amino acids into HDPs is another promising approach to achieve desirable biological properties. Diastereomeric synthetic peptides containing D-amino acids tend to be less hemolytic and cytotoxic while retaining the antibacterial activity (115, 117, 148, 201). Consistently, diastereomeric peptides are incapable of inserting into the lipid bilayer mimicking mammalian cell membranes, whereas peptides containing only L-amino acids do so readily (118). D-amino acids modulate peptide cytotoxic activity by changing helicity and hydrophobicity. Furthermore, diastereomers maintain the antibacterial activity in serum and are resistant to degradation by proteases (117). Such diastereomers, but not L-amino acid peptides, protect mice against antibiotic-sensitive as well as -resistant infections, suggesting improved stability of diastereomers *in vivo* (14).

Cyclization of certain peptides has been shown to increase the stability of peptides *in vivo* (42). Cyclic peptides with alternating D- and L-amino acids are resistant to proteases and highly stable when administered *in vivo* (27). Similarly, cyclo CP-11, a peptide derived from bovine indolicidin, has a prolonged half-life in the presence of

trypsin as compared to linear CP-11 (130). Cyclization also enhances the antibacterial activity and reduces the hemolytic activity in several cases. A cyclic peptide with D- and L-amino acids shows reduced hemolytic activity and broad spectrum antibacterial activity relative to the linear analog (111). It correlates with an increased ability of the cyclic peptide to interact with and permeabilize negatively charged membranes and reduced affinity for zwitterionic membranes. Cyclization also renders peptides resistant to salt. The most convincing examples of salt-insensitive cyclic peptides are θ-defensins. These peptides, as discussed in the previous section, are naturally occurring cyclic defensins of old world primates. Antibacterial activity of cyclic RTD-1 is insensitive to salt. However, such an activity is completely lost in the presence of physiological concentrations of salt when peptide is linearized (156). *De novo* designed cyclic peptides, resembling θ-defensins and porcine protegrins, have also been shown to act in a salt-insensitive manner (94). Even naturally occurring HDPs can be cyclized to acquire salt-insenstivie antibacterial activities. For example, end-to-end cyclized rabbit neutrophil peptide-1 (NP-1) is able to maintain its activity in the presence of salt, whereas wild-type NP-1 is inactivated (193).

Dimerization of peptides through covalent linkage by disulfide bonds may also increase antibacterial potency and render peptides salt-insensitive. Evidence for this argument comes from naturally occurring peptide dimers. Dimerization of cryptdin-related CRS peptides in mice increases antibacterial potency (54). Dimeric Defr1, a mouse β-defensin-related peptide, retains the antibacterial activity in the presence of physiological salt concentrations (19). Cathelicidins, PMAP-36 and CAP-11, also exist

in dimeric forms in neutrophils; however, dimerization does not increase antibacterial potency (109, 139), although it modulates the kinetics of killing and membrane permeabilization in the case of PMAP-36 (139). Increasing evidence suggests a positive association between dimerization and the protease stability of peptides (126). Cross linking of two tryptophans in indolicidin confers protease resistance without altering the antibacterial activity (112).

Synthetic combinatorial libraries (SCL) have also been employed to identify small peptides (hexapeptides) with the antibacterial activity (7). However, constructing such libraries is cost-prohibitive especially for longer peptides. Another approach to produce a large number of peptides is high throughput, *in situ* peptide synthesis on a cellulose support by pipetting robots (50). Peptides are subsequently cleaved from the support and the antibacterial activity is studied in a high throughput manner against *Peudomonas aeruginosa* expressing a luciferase reporter (50). It is estimated that 100,000 individual peptides can be produced by this approach using two robots and screened for antibacterial activities in a year. Using this approach, bovine bactenecin derivatives with increased antibacterial potency have been obtained (49). A recent report described the rational design of antimicrobial peptides based on a linguistic model using naturally occurring HDPs (78). This language is based on grammars that use common sequence patterns found in natural peptides. Using this strategy, a large number of *de novo* peptides can be synthesized and screened for the antibacterial activity.

Abiotic oligomeric and polymeric compounds are also being synthesized to mimic the structural properties of HDPs. Such compounds are resistant to proteases. For example, oligomeric compounds like acrylamide and oligourea and synthetic polymers like phenylene ethynylene are able to fold into amphipathic structures. These compounds possess antibacterial activities against a broad range of bacteria including antibiotic-resistant bacteria (106, 159, 160, 169). Theses abiotic compounds are selectively active against bacterial membranes. They insert themselves into the bacterial model membranes and are less selective towards mammalian model membranes (57). Additional investigations are warranted to study the *in vivo* toxicity and antibacterial efficacy of these compounds.

CURRENT THERAPEUTIC APPLICATION AND CLINICAL DEVELOPMENT

Due to a range of beneficial effects conferred by HDPs including broad-spectrum antibacterial activity and immunomodulatory effects, efforts are ongoing to develop them as novel antibiotics to treat infections. The companies involved in their commercial development are mostly small biotechnology companies. Their efforts have been met with mixed success. A number of peptides are being developed in various stages of clinical trials and are yet to be approved by FDA. The diseases against which peptides are developed include ventilator-associated pneumonia, in-dwelling catheter-associated infections, oral candidiasis, sepsis, cystic fibrosis, acne, and bacterial meningitis (85, 202).

The most successful peptide to date is bovine indolicidin-based omiganan pentahydrochloride 1% gel, also known as MX-226, which has been shown to reduce catheter-associated infections by 49% in a phase III clinical trial (58, 85). Another indolicidin-based peptide, MX-594AN, significantly reduced severity of acne lesions in a phase II trial (58, 85) (www.migenix.com). Pexiganan, a cationic peptide based on frog magainin, was the first peptide to enter clinical trials for treatment of diabetic foot ulcers. However, FDA rejected its approval despite its effectiveness, since it did not confer any additional benefits than treatments currently available to treat diabetic foot ulcers (85). Iseganan, a peptide based on porcine protegrin, did not provide obvious benefits in ventilator-associated pneumonia and was withdrawn from clinical trials (85).

Compared to other HDPs, not even a single defensin has reached clinical trials. The disadvantages associated with defensins are low potency, salt sensitivity, and significant costs in producing disulfide-containing peptides on a large scale. However, plectasin, a novel defensin from fungus *Pseudoplectania nigrella*, can be produced and purified on an industrial scale using a fungal expression system (97). Plectasin has a half life of 50 min in serum, presumably due to increased stability towards proteases provided by disulfide bonds. Furthermore, it protects mice from experimental peritonitis and pneumonia (97).

Mixed successes are not deterring companies from developing HDPs as antibacterial therapeutics. However, the trend is to develop smaller versions of peptides, which are cheaper to synthesize. For example, Helix BioMedix is developing short

hexapeptides to treat bacterial and fungal infections (202). Also, antibacterial activities of HDPs can be separated from immunomodulatory activities as discussed in the previous section. It will not be surprising if companies start utilizing immunomodulatory fragments, rather than antibacterial fragments for clinical development. Inimex pharmaceutical has identified a lead compound based on LL-37. This lead peptide lacks antibacterial activity, but is able to protect against infection in animal models due to its immunomodulatory activities (11). Helix BioMedix is developing HB107, a peptide based on cecropin, for wound healing rather than anti-infective purposes (202). Some companies are even developing peptide mimetics and oligomeric compounds to treat infections, since such drugs show more stability in serum than HDPs due to their resistance to proteases (202).

CONCLUSION AND FUTURE PROSPECT

HDPs possess antimicrobial as well as immunomodulatory activities. They have tremendous potential as novel anti-infective drugs. However, there is a need to address issues such as antibacterial potency, toxicity, stability, salt sensitivity, and cost of production. A variety of strategies are emerging to tackle these issues including structure-based rational design, amidation, cyclization, dimerization, amino acid modification, and peptide mimetics. Another important issue is the length of peptide as it is directly related to the production cost. Efforts should be directed to make HDPs as small as possible without compromising their biological activities. In addition to recombinant production, making flocks of transgenic animals producing a large amount of peptides is a practical

approach to produce HDPs. Taken together, identification of novel HDPs, study of their structure activity relationship, and optimal peptide design will help drive anti-infective peptides into clinics.

REFERENCES

1. **Aarbiou J, Ertmann M, van Wetering S, van Noort P, Rook D, Rabe KF, Litvinov SV, van Krieken JH, de Boer WI, and Hiemstra PS.** Human neutrophil defensins induce lung epithelial cell proliferation in vitro. *J Leukoc Biol* 72: 167-174, 2002.

2. **Aarbiou J, Verhoosel RM, Van Wetering S, De Boer WI, Van Krieken JH, Litvinov SV, Rabe KF, and Hiemstra PS.** Neutrophil defensins enhance lung epithelial wound closure and mucin gene expression in vitro. *Am J Respir Cell Mol Biol* 30: 193-201, 2004.

3. **An LL, Yang YH, Ma XT, Lin YM, Li G, Song YH, and Wu KF.** LL-37 enhances adaptive antitumor immune response in a murine model when genetically fused with M-CSFR (J6-1) DNA vaccine. *Leuk Res* 29: 535-543, 2005.

4. **Bauer F, Schweimer K, Kluver E, Conejo-Garcia JR, Forssmann WG, Rosch P, Adermann K, and Sticht H.** Structure determination of human and murine beta-defensins reveals structural conservation in the absence of significant sequence similarity. *Protein Sci* 10: 2470-2479, 2001.

5. **Biragyn A.** Defensins--non-antibiotic use for vaccine development. *Curr Protein Pept Sci* 6: 53-60, 2005.

6. **Biragyn A, Surenhu M, Yang D, Ruffini PA, Haines BA, Klyushnenkova E, Oppenheim JJ, and Kwak LW.** Mediators of innate immunity that target immature, but not mature, dendritic cells induce antitumor immunity when genetically fused with nonimmunogenic tumor antigens. *J Immunol* 167: 6644-6653, 2001.

7. **Blondelle SE and Houghten RA.** Novel antimicrobial compounds identified using synthetic combinatorial library technology. *Trends Biotechnol* 14: 60-65, 1996.

8. **Boman HG, Agerberth B, and Boman A.** Mechanisms of action on Escherichia coli of cecropin P1 and PR-39, two antibacterial peptides from pig intestine. *Infect Immun* 61: 2978-2984, 1993.

9. **Boniotto M, Jordan WJ, Eskdale J, Tossi A, Antcheva N, Crovella S, Connell ND, and Gallagher G.** Human beta-defensin 2 induces a vigorous cytokine response in peripheral blood mononuclear cells. *Antimicrob Agents Chemother* 50: 1433-1441, 2006.

10. **Bowdish DM, Davidson DJ, and Hancock RE.** Immunomodulatory properties of defensins and cathelicidins. *Curr Top Microbiol Immunol* 306: 27-66, 2006.

11. **Bowdish DM, Davidson DJ, Lau YE, Lee K, Scott MG, and Hancock RE.** Impact of LL-37 on anti-infective immunity. *J Leukoc Biol* 77: 451-459, 2005.

12. **Bowdish DM, Davidson DJ, Scott MG, and Hancock RE.** Immunomodulatory activities of small host defense peptides. *Antimicrob Agents Chemother* 49: 1727-1732, 2005.

13. **Braff MH, Hawkins MA, Di Nardo A, Lopez-Garcia B, Howell MD, Wong C, Lin K, Streib JE, Dorschner R, Leung DY, and Gallo RL.** Structure-function

relationships among human cathelicidin peptides: dissociation of antimicrobial properties from host immunostimulatory activities. *J Immunol* 174: 4271-4278, 2005.

14. **Braunstein A, Papo N, and Shai Y.** In vitro activity and potency of an intravenously injected antimicrobial peptide and its DL amino acid analog in mice infected with bacteria. *Antimicrob Agents Chemother* 48: 3127-3129, 2004.

15. **Brogden KA.** Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* 3: 238-250, 2005.

16. **Brogden KA, Heidari M, Sacco RE, Palmquist D, Guthmiller JM, Johnson GK, Jia HP, Tack BF, and McCray PB.** Defensin-induced adaptive immunity in mice and its potential in preventing periodontal disease. *Oral Microbiol Immunol* 18: 95-99, 2003.

17. **Buffy JJ, McCormick MJ, Wi S, Waring A, Lehrer RI, and Hong M.** Solid-state NMR investigation of the selective perturbation of lipid bilayers by the cyclic antimicrobial peptide RTD-1. *Biochemistry* 43: 9800-9812, 2004.

18. **Cabiaux V, Agerberth B, Johansson J, Homble F, Goormaghtigh E, and Ruysschaert JM.** Secondary structure and membrane interaction of PR-39, a Pro+Arg-rich antibacterial peptide. *Eur J Biochem* 224: 1019-1027, 1994.

19. **Campopiano DJ, Clarke DJ, Polfer NC, Barran PE, Langley RJ, Govan JR, Maxwell A, and Dorin JR.** Structure-activity relationships in defensin dimers: a novel link between beta-defensin tertiary structure and antimicrobial activity. *J Biol Chem* 279: 48671-48679, 2004.

20. **Chan YR, Zanetti M, Gennaro R, and Gallo RL.** Anti-microbial activity and cell binding are controlled by sequence determinants in the anti-microbial peptide PR-39. *J Invest Dermatol* 116: 230-235, 2001.

21. **Chang CI, Pleguezuelos O, Zhang YA, Zou J, and Secombes CJ.** Identification of a novel cathelicidin gene in the rainbow trout, Oncorhynchus mykiss. *Infect Immun* 73: 5053-5064, 2005.

22. **Chang CI, Zhang YA, Zou J, Nie P, and Secombes CJ.** Two cathelicidin genes are present in both rainbow trout (Oncorhynchus mykiss) and atlantic salmon (Salmo salar). *Antimicrob Agents Chemother* 50: 185-195, 2006.

23. **Chen C, Brock R, Luh F, Chou PJ, Larrick JW, Huang RF, and Huang TH.** The solution structure of the active domain of CAP18--a lipopolysaccharide binding protein from rabbit leukocytes. *FEBS Lett* 370: 46-52, 1995.

24. **Chertov O, Michiel DF, Xu L, Wang JM, Tani K, Murphy WJ, Longo DL, Taub DD, and Oppenheim JJ.** Identification of defensin-1, defensin-2, and CAP37/azurocidin as T- cell chemoattractant proteins released from interleukin-8-stimulated neutrophils. *J Biol Chem* 271: 2935-2940, 1996.

25. **Ciornei CD, Sigurdardottir T, Schmidtchen A, and Bodelsson M.** Antimicrobial and chemoattractant activity, lipopolysaccharide neutralization, cytotoxicity, and inhibition by serum of analogs of human cathelicidin LL-37. *Antimicrob Agents Chemother* 49: 2845-2850, 2005.

26. **Cirioni O, Giacometti A, Ghiselli R, Bergnach C, Orlando F, Silvestri C, Mocchegiani F, Licci A, Skerlavaj B, Rocchi M, Saba V, Zanetti M, and Scalise**

G. LL-37 protects rats against lethal sepsis caused by gram-negative bacteria. *Antimicrob Agents Chemother* 50: 1672-1679, 2006.

27. **Dartois V, Sanchez-Quesada J, Cabezas E, Chi E, Dubbelde C, Dunn C, Granja J, Gritzen C, Weinberger D, Ghadiri MR, and Parr TR, Jr.** Systemic antibacterial activity of novel synthetic cyclic peptides. *Antimicrob Agents Chemother* 49: 3302-3310, 2005.

28. **Dathe M and Wieprecht T.** Structural features of helical antimicrobial peptides: their potential to modulate activity on model membranes and biological cells. *Biochim Biophys Acta* 1462: 71-87, 1999.

29. **De Y, Chen Q, Schmidt AP, Anderson GM, Wang JM, Wooters J, Oppenheim JJ, and Chertov O.** LL-37, the neutrophil granule- and epithelial cell-derived cathelicidin, utilizes formyl peptide receptor-like 1 (FPRL1) as a receptor to chemoattract human peripheral blood neutrophils, monocytes, and T cells. *J Exp Med* 192: 1069-1074, 2000.

30. **Eisenhauer PB and Lehrer RI.** Mouse neutrophils lack defensins. *Infect Immun* 60: 3446-3447, 1992.

31. **Finch RG.** Antibiotic resistance: a view from the prescriber. *Nat Rev Microbiol* 2: 989-994, 2004.

32. **Finlay BB and Hancock RE.** Can innate immunity be enhanced to treat microbial infections? *Nat Rev Microbiol* 2: 497-504, 2004.

33. **Fukumoto K, Nagaoka I, Yamataka A, Kobayashi H, Yanai T, Kato Y, and Miyano T.** Effect of antibacterial cathelicidin peptide CAP18/LL-37 on sepsis in neonatal rats. *Pediatr Surg Int* 21: 20-24, 2005.

34. **Gaczynska M, Osmulski PA, Gao Y, Post MJ, and Simons M.** Proline- and arginine-rich peptides constitute a novel class of allosteric inhibitors of proteasome activity. *Biochemistry* 42: 8663-8670, 2003.

35. **Gallo RL, Ono M, Povsic T, Page C, Eriksson E, Klagsbrun M, and Bernfield M.** Syndecans, cell surface heparan sulfate proteoglycans, are induced by a proline-rich antimicrobial peptide from wounds. *Proc Natl Acad Sci USA* 91: 11035-11039, 1994.

36. **Ganz T.** Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 3: 710-720, 2003.

37. **Garcia JR, Jaumann F, Schulz S, Krause A, Rodriguez-Jimenez J, Forssmann U, Adermann K, Kluver E, Vogelmeier C, Becker D, Hedrich R, Forssmann WG, and Bals R.** Identification of a novel, multifunctional beta-defensin (human beta- defensin 3) with specific antimicrobial activity. Its interaction with plasma membranes of Xenopus oocytes and the induction of macrophage chemoattraction. *Cell Tissue Res* 306: 257-264, 2001.

38. **Garcia JR, Krause A, Schulz S, Rodriguez-Jimenez FJ, Kluver E, Adermann K, Forssmann U, Frimpong-Boateng A, Bals R, and Forssmann WG.** Human beta-defensin 4: a novel inducible peptide with a specific salt- sensitive spectrum of antimicrobial activity. *FASEB J* 15: 1819-1821, 2001.

39. **Ghiselli R, Cirioni O, Giacometti A, Mocchegiani F, Orlando F, Silvestri C, Licci A, Della Vittoria A, Scalise G, and Saba V.** The cathelicidin-derived tritrpticin enhances the efficacy of ertapenem in experimental rat models of septic shock. *Shock* 26: 195-200, 2006.

40. **Giacometti A, Cirioni O, Ghiselli R, Mocchegiani F, Del Prete MS, Viticchi C, Kamysz W, E LE, Saba V, and Scalise G.** Potential therapeutic role of cationic peptides in three experimental models of septic shock. *Antimicrob Agents Chemother* 46: 2132-2136, 2002.

41. **Gonzales R, Malone DC, Maselli JH, and Sande MA.** Excessive antibiotic use for acute respiratory infections in the United States. *Clin Infect Dis* 33: 757-762, 2001.

42. **Hancock RE.** Cationic peptides: effectors in innate immunity and novel antimicrobials. *Lancet Infect Dis* 1: 156-164, 2001.

43. **Hancock RE and Lehrer R.** Cationic peptides: a new source of antibiotics. *Trends Biotechnol* 16: 82-88, 1998.

44. **Hancock RE and Patrzykat A.** Clinical development of cationic antimicrobial peptides: from natural to novel antibiotics. *Curr Drug Targets Infect Disord* 2: 79-83, 2002.

45. **Harwig SS, Kokryakov VN, Swiderek KM, Aleshina GM, Zhao C, and Lehrer RI.** Prophenin-1, an exceptionally proline-rich antimicrobial peptide from porcine leukocytes. *FEBS Lett* 362: 65-69, 1995.

46. **Heilborn JD, Nilsson MF, Kratz G, Weber G, Sorensen O, Borregaard N, and Stahle-Backdahl M.** The cathelicidin anti-microbial peptide LL-37 is involved in re-epithelialization of human skin wounds and is lacking in chronic ulcer epithelium. *J Invest Dermatol* 120: 379-389, 2003.

47. **Henzler Wildman KA, Lee DK, and Ramamoorthy A.** Mechanism of lipid bilayer disruption by the human antimicrobial peptide, LL-37. *Biochemistry* 42: 6545-6558, 2003.

48. **Hill CP, Yee J, Selsted ME, and Eisenberg D.** Crystal structure of defensin HNP-3, an amphiphilic dimer: mechanisms of membrane permeabilization. *Science* 251: 1481-1485, 1991.

49. **Hilpert K, Elliott MR, Volkmer-Engert R, Henklein P, Donini O, Zhou Q, Winkler DF, and Hancock RE.** Sequence requirements and an optimization strategy for short antimicrobial peptides. *Chem Biol* 13: 1101-1107, 2006.

50. **Hilpert K, Volkmer-Engert R, Walter T, and Hancock RE.** High-throughput generation of small antibacterial peptides with improved activity. *Nat Biotechnol* 23: 1008-1012, 2005.

51. **Hoover DM, Chertov O, and Lubkowski J.** The structure of human beta-defensin-1: new insights into structural properties of beta-defensins. *J Biol Chem* 276: 39021-39026, 2001.

52. **Hoover DM, Rajashankar KR, Blumenthal R, Puri A, Oppenheim JJ, Chertov O, and Lubkowski J.** The structure of human beta-defensin-2 shows evidence of higher order oligomerization. *J Biol Chem* 275: 32911-32918, 2000.

53. **Hoover DM, Wu Z, Tucker K, Lu W, and Lubkowski J.** Antimicrobial characterization of human beta-defensin 3 derivatives. *Antimicrob Agents Chemother* 47: 2804-2809, 2003.

54. **Hornef MW, Putsep K, Karlsson J, Refai E, and Andersson M.** Increased diversity of intestinal antimicrobial peptides by covalent dimer formation. *Nat Immunol* 5: 836-843, 2004.

55. **Huang HJ, Ross CR, and Blecha F.** Chemoattractant properties of PR-39, a neutrophil antibacterial peptide. *JLeukocBiol* 61: 624-629, 1997.

56. **Huang LC, Petkova TD, Reins RY, Proske RJ, and McDermott AM.**
Multifunctional roles of human cathelicidin (LL-37) at the ocular surface. *Invest Ophthalmol Vis Sci* 47: 2369-2380, 2006.

57. **Ishitsuka Y, Arnt L, Majewski J, Frey S, Ratajczek M, Kjaer K, Tew GN, and Lee KY.** Amphiphilic poly(phenyleneethynylene)s can mimic antimicrobial peptide membrane disordering effect by membrane insertion. *J Am Chem Soc* 128: 13123-13129, 2006.

58. **Jenssen H, Hamill P, and Hancock RE.** Peptide antimicrobial agents. *Clin Microbiol Rev* 19: 491-511, 2006.

59. **Jing W, Hunter HN, Tanabe H, Ouellette AJ, and Vogel HJ.** Solution structure of cryptdin-4, a mouse paneth cell alpha-defensin. *Biochemistry* 43: 15759-15766, 2004.

60. **Kagan BL, Selsted ME, Ganz T, and Lehrer RI.** Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. *Proc Natl Acad Sci U S A* 87: 210-214, 1990.

61. **Kim C, Gajendran N, Mittrucker HW, Weiwad M, Song YH, Hurwitz R, Wilmanns M, Fischer G, and Kaufmann SH.** Human alpha-defensins neutralize anthrax lethal toxin and protect against its fatal consequences. *Proc Natl Acad Sci U S A* 102: 4830-4835, 2005.

62. **Kim C, Slavinskaya Z, Merrill AR, and Kaufmann SH.** Human alpha-defensins neutralize toxins of the mono-ADP-ribosyltransferase family. *Biochem J*, 2006.

63. **Klotman ME and Chang TL.** Defensins in innate antiviral immunity. *Nat Rev Immunol* 6: 447-456, 2006.

64. **Kluver E, Schulz-Maronde S, Scheid S, Meyer B, Forssmann WG, and Adermann K.** Structure-activity relation of human beta-defensin 3: influence of disulfide bonds and cysteine substitution on antimicrobial activity and cytotoxicity. *Biochemistry* 44: 9804-9816, 2005.

65. **Koczulla R, von Degenfeld G, Kupatt C, Krotz F, Zahler S, Gloe T, Issbrucker K, Unterberger P, Zaiou M, Lebherz C, Karl A, Raake P, Pfosser A, Boekstegers P, Welsch U, Hiemstra PS, Vogelmeier C, Gallo RL, Clauss M, and Bals R.** An angiogenic role for the human peptide antibiotic LL-37/hCAP-18. *J Clin Invest* 111: 1665-1672, 2003.

66. **Kokryakov VN, Harwig SS, Panyutich EA, Shevchenko AA, Aleshina GM, Shamova OV, Korneva HA, and Lehrer RI.** Protegrins: leukocyte antimicrobial peptides that combine features of corticostatic defensins and tachyplesins. *FEBS Lett* 327: 231-236, 1993.

67. **Kurosaka K, Chen Q, Yarovinsky F, Oppenheim JJ, and Yang D.** Mouse cathelin-related antimicrobial peptide chemoattracts leukocytes using formyl peptide receptor-like 1/mouse formyl peptide receptor-like 2 as the receptor and acts as an immune adjuvant. *J Immunol* 174: 6257-6265, 2005.

68. **Ladokhin AS, Selsted ME, and White SH.** Bilayer interactions of indolicidin, a small antimicrobial peptide rich in tryptophan, proline, and basic amino acids. *BiophysJ* 72: 794-805, 1997.

69. **Landon C, Thouzeau C, Labbe H, Bulet P, and Vovelle F.** Solution structure of spheniscin, a beta-defensin from the penguin stomach. *J Biol Chem* 279: 30433-30439, 2004.

70. **Lee KH and Oh JE.** Design and synthesis of novel antimicrobial pseudopeptides with selective membrane-perturbation activity. *Bioorg Med Chem* 8: 833-839, 2000.

71. **Lee PH, Ohtake T, Zaiou M, Murakami M, Rudisill JA, Lin KH, and Gallo RL.** Expression of an additional cathelicidin antimicrobial peptide protects against bacterial skin infection. *Proc Natl Acad Sci U S A* 102: 3750-3755, 2005.

72. **Lehrer RI.** Primate defensins. *Nat Rev Microbiol* 2: 727-738, 2004.

73. **Lehrer RI, Barton A, Daher KA, Harwig SS, Ganz T, and Selsted ME.** Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity. *J Clin Invest* 84: 553-561, 1989.

74. **Li J, Post M, Volk R, Gao Y, Li M, Metais C, Sato K, Tsai J, Aird W, Rosenberg RD, Hampton TG, Sellke F, Carmeliet P, and Simons M.** PR39, a peptide regulator of angiogenesis. *Nat Med* 6: 49-55, 2000.

75. **Lillard JW, Jr., Boyaka PN, Chertov O, Oppenheim JJ, and McGhee JR.** Mechanisms for induction of acquired host immunity by neutrophil peptide defensins. *Proc Natl Acad Sci USA* 96: 651-656, 1999.

76. **Lin PW, Simon PO, Jr., Gewirtz AT, Neish AS, Ouellette AJ, Madara JL, and Lencer WI.** Paneth cell cryptdins act in vitro as apical paracrine regulators of the innate inflammatory response. *J Biol Chem* 279: 19902-19907, 2004.

77. **Lohner K, Latal A, Lehrer RI, and Ganz T.** Differential scanning microcalorimetry indicates that human defensin, HNP-2, interacts specifically with biomembrane mimetic systems. *Biochemistry* 36: 1525-1531, 1997.

78. **Loose C, Jensen K, Rigoutsos I, and Stephanopoulos G.** A linguistic model for the rational design of antimicrobial peptides. *Nature* 443: 867-869, 2006.

79. **Lu K, Asano R, and Davies J.** Antimicrobial resistance gene delivery in animal feeds. *Emerg Infect Dis* 10: 679-683, 2004.

80. **Maemoto A, Qu X, Rosengren KJ, Tanabe H, Henschen-Edman A, Craik DJ, and Ouellette AJ.** Functional analysis of the alpha-defensin disulfide array in mouse cryptdin-4. *J Biol Chem* 279: 44188-44196, 2004.

81. **Mandal M and Nagaraj R.** Antibacterial activities and conformations of synthetic alpha-defensin HNP-1 and analogs with one, two and three disulfide bridges. *J Pept Res* 59: 95-104, 2002.

82. **Mani R, Cady SD, Tang M, Waring AJ, Lehrer RI, and Hong M.** Membrane-dependent oligomeric structure and pore formation of a {beta}-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR. *Proc Natl Acad Sci U S A*, 2006.

83. **Mani R, Waring AJ, Lehrer RI, and Hong M.** Membrane-disruptive abilities of beta-hairpin antimicrobial peptides correlate with conformation and activity: a 31P and 1H NMR study. *Biochim Biophys Acta* 1716: 11-18, 2005.

84. **Marchand C, Krajewski K, Lee HF, Antony S, Johnson AA, Amin R, Roller P, Kvaratskhelia M, and Pommier Y.** Covalent binding of the natural antimicrobial peptide indolicidin to DNA abasic sites. *Nucleic Acids Res*, 2006.

85. **Marr AK, Gooderham WJ, and Hancock RE.** Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Curr Opin Pharmacol* 6: 468-472, 2006.

86. **McManus AM, Dawson NF, Wade JD, Carrington LE, Winzor DJ, and Craik DJ.** Three-dimensional structure of RK-1: a novel alpha-defensin peptide. *Biochemistry* 39: 15757-15764, 2000.

87. **McPhee JB and Hancock RE.** Function and therapeutic potential of host defence peptides. *J Pept Sci* 11: 677-687, 2005.

88. **McPhee JB, Scott MG, and Hancock RE.** Design of host defence peptides for antimicrobial and immunity enhancing activities. *Comb Chem High Throughput Screen* 8: 257-272, 2005.

89. **Mookherjee N, Brown KL, Bowdish DM, Doria S, Falsafi R, Hokamp K, Roche FM, Mu R, Doho GH, Pistolic J, Powers JP, Bryan J, Brinkman FS, and Hancock RE.** Modulation of the TLR-mediated inflammatory response by the endogenous human host defense peptide LL-37. *J Immunol* 176: 2455-2464, 2006.

90. **Mookherjee N, Wilson HL, Doria S, Popowych Y, Falsafi R, Yu JJ, Li Y, Veatch S, Roche FM, Brown KL, Brinkman FS, Hokamp K, Potter A, Babiuk LA, Griebel PJ, and Hancock RE.** Bovine and human cathelicidin cationic host defense peptides similarly suppress transcriptional responses to bacterial lipopolysaccharide. *J Leukoc Biol*, 2006.

91. **Morrison G, Kilanowski F, Davidson D, and Dorin J.** Characterization of the mouse beta defensin 1, Defb1, mutant mouse model. *Infect Immun* 70: 3053-3060, 2002.

92. **Moser C, Weiner DJ, Lysenko E, Bals R, Weiser JN, and Wilson JM.** beta-Defensin 1 contributes to pulmonary innate immunity in mice. *Infect Immun* 70: 3068-3072, 2002.

93. **Motzkus D, Schulz-Maronde S, Heitland A, Schulz A, Forssmann WG, Jubner M, and Maronde E.** The novel beta-defensin DEFB123 prevents lipopolysaccharide-mediated effects in vitro and in vivo. *Faseb J* 20: 1701-1702, 2006.

94. **Muhle SA and Tam JP.** Design of Gram-negative selective antimicrobial peptides. *Biochemistry* 40: 5777-5785, 2001.

95. **Murakami M, Lopez-Garcia B, Braff M, Dorschner RA, and Gallo RL.** Postsecretory processing generates multiple cathelicidins for enhanced topical antimicrobial defense. *J Immunol* 172: 3070-3077, 2004.

96. **Murphy CJ, Foster BA, Mannis MJ, Selsted ME, and Reid TW.** Defensins are mitogenic for epithelial cells and fibroblasts. *J Cell Physiol* 155: 408-413, 1993.

97. **Mygind PH, Fischer RL, Schnorr KM, Hansen MT, Sonksen CP, Ludvigsen S, Raventos D, Buskov S, Christensen B, De Maria L, Taboureau O, Yaver D, Elvig-Jorgensen SG, Sorensen MV, Christensen BE, Kjaerulff S, Frimodt-Moller N, Lehrer RI, Zasloff M, and Kristensen HH.** Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. *Nature* 437: 975-980, 2005.

98. **Nagaoka I, Tsutsumi-Ishii Y, Yomogida S, and Yamashita T.** Isolation of cDNA encoding guinea pig neutrophil cationic antibacterial polypeptide of 11 kDa (CAP11) and evaluation of CAP11 mRNA expression during neutrophil maturation. *J Biol Chem* 272: 22742-22750, 1997.

99. **Nell MJ, Tjabringa GS, Wafelman AR, Verrijk R, Hiemstra PS, Drijfhout JW, and Grote JJ.** Development of novel LL-37 derived antimicrobial peptides with LPS and LTA neutralizing and antimicrobial activities for therapeutic application. *Peptides* 27: 649-660, 2006.

100. **Nguyen TX, Cole AM, and Lehrer RI.** Evolution of primate theta-defensins: a serpentine path to a sweet tooth. *Peptides* 24: 1647-1654, 2003.

101. **Niederman MS.** Impact of antibiotic resistance on clinical outcomes and the cost of care. *Crit Care Med* 29: N114-120, 2001.

102. **Niyonsaba F, Iwabuchi K, Matsuda H, Ogawa H, and Nagaoka I.** Epithelial cell-derived human beta-defensin-2 acts as a chemotaxin for mast cells through a pertussis toxin-sensitive and phospholipase C-dependent pathway. *Int Immunol* 14: 421-426, 2002.

103. **Niyonsaba F, Iwabuchi K, Someya A, Hirata M, Matsuda H, Ogawa H, and Nagaoka I.** A cathelicidin family of human antibacterial peptide LL-37 induces mast cell chemotaxis. *Immunology* 106: 20-26, 2002.

104. **Niyonsaba F, Ushio H, Nakano N, Ng W, Sayama K, Hashimoto K, Nagaoka I, Okumura K, and Ogawa H.** Antimicrobial Peptides Human beta-Defensins Stimulate Epidermal Keratinocyte Migration, Proliferation and Production of Proinflammatory Cytokines and Chemokines. *J Invest Dermatol*, 2006.

105. **Nizet V, Ohtake T, Lauth X, Trowbridge J, Rudisill J, Dorschner RA, Pestonjamasp V, Piraino J, Huttner K, and Gallo RL.** Innate antimicrobial peptide protects the skin from invasive bacterial infection. *Nature* 414: 454-457, 2001.

106. **Nusslein K, Arnt L, Rennie J, Owens C, and Tew GN.** Broad-spectrum antibacterial activity by a novel abiogenic peptide mimic. *Microbiology* 152: 1913-1918, 2006.

107. **Oh JE, Hong SY, and Lee KH.** Design, synthesis and characterization of antimicrobial pseudopeptides corresponding to membrane-active peptide. *J Pept Res* 54: 129-136, 1999.

108. **Oh JE and Lee KH.** Synthesis of novel unnatural amino acid as a building block and its incorporation into an antimicrobial peptide. *Bioorg Med Chem* 7: 2985-2990, 1999.

109. **Okuda D, Yomogida S, Tamura H, and Nagaoka I.** Determination of the antibacterial and lipopolysaccharide-neutralizing regions of guinea pig neutrophil cathelicidin peptide CAP11. *Antimicrob Agents Chemother* 50: 2602-2607, 2006.

110. **Oren Z, Lerman JC, Gudmundsson GH, Agerberth B, and Shai Y.** Structure and organization of the human antimicrobial peptide LL-37 in phospholipid membranes: relevance to the molecular basis for its non- cell-selective activity. *BiochemJ* 341 (Pt 3): 501-513, 1999.

111. **Oren Z and Shai Y.** Cyclization of a cytolytic amphipathic alpha-helical peptide and its diastereomer: effect on structure, interaction with model membranes, and biological function. *Biochemistry* 39: 6103-6114, 2000.

112. **Osapay K, Tran D, Ladokhin AS, White SH, Henschen AH, and Selsted ME.** Formation and characterization of a single Trp-Trp cross-link in indolicidin that confers protease stability without altering antimicrobial activity. *J Biol Chem* 275: 12017-12022, 2000.

113. **Ouellette AJ and Bevins CL.** Paneth cell defensins and innate immunity of the small bowel. *InflammBowelDis* 7: 43-50, 2001.

114. **Ouellette AJ, Hsieh MM, Nosek MT, Cano-Gauci DF, Huttner KM, Buick RN, and Selsted ME.** Mouse Paneth cell defensins: primary structures and antibacterial activities of numerous cryptdin isoforms. *Infect Immun* 62: 5040-5047, 1994.

115. **Pacor S, Giangaspero A, Bacac M, Sava G, and Tossi A.** Analysis of the cytotoxicity of synthetic antimicrobial peptides on mouse leucocytes: implications for systemic use. *J Antimicrob Chemother* 50: 339-348, 2002.

116. **Paone G, Wada A, Stevens LA, Matin A, Hirayama T, Levine RL, and Moss J.** ADP ribosylation of human neutrophil peptide-1 regulates its biological properties. *Proc Natl Acad Sci U S A* 99: 8231-8235, 2002.

117. **Papo N, Oren Z, Pag U, Sahl HG, and Shai Y.** The consequence of sequence alteration of an amphipathic alpha-helical antimicrobial peptide and its diastereomers. *J Biol Chem* 277: 33913-33921, 2002.

118. **Papo N and Shai Y.** Effect of drastic sequence alteration and D-amino acid incorporation on the membrane binding behavior of lytic peptides. *Biochemistry* 43: 6393-6403, 2004.

119. **Pardi A, Zhang XL, Selsted ME, Skalicky JJ, and Yip PF.** NMR studies of defensin antimicrobial peptides. 2. Three-dimensional structures of rabbit NP-2 and human HNP-1. *Biochemistry* 31: 11357-11364, 1992.

120. **Patil A, Hughes AL, and Zhang G.** Rapid evolution and diversification of mammalian alpha-defensins as revealed by comparative analysis of rodent and primate genes. *Physiol Genomics* 20: 1-11, 2004.

121. **Patil AA, Cai Y, Sang Y, Blecha F, and Zhang G.** Cross-species analysis of the mammalian beta-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract. *Physiol Genomics* 23: 5-17, 2005.

122. **Pazgier M, Prahl A, Hoover DM, and Lubkowski J.** Studies of the biological properties of human beta -defensin 1. *J Biol Chem*, 2006.

123. **Pazgiera M, Hoover DM, Yang D, Lu W, and Lubkowski J.** Human beta-defensins. *Cell Mol Life Sci* 63: 1294-1313, 2006.

124. **Porter EM, Liu L, Oren A, Anton PA, and Ganz T.** Localization of human intestinal defensin 5 in Paneth cell granules. *Infect Immun* 65: 2389-2395, 1997.

125. **Quinn T, O'Mahony R, Baird AW, Drudy D, Whyte P, and Fanning S.** Multi-drug resistance in Salmonella enterica: efflux mechanisms and their relationships with the development of chromosomal resistance gene clusters. *Curr Drug Targets* 7: 849-860, 2006.

126. **Raimondo D, Andreotti G, Saint N, Amodeo P, Renzone G, Sanseverino M, Zocchi I, Molle G, Motta A, and Scaloni A.** A folding-dependent mechanism of antimicrobial peptide resistance to degradation unveiled by solution structure of distinctin. *Proc Natl Acad Sci U S A* 102: 6309-6314, 2005.

127. **Rice WG, Ganz T, Kinkade JM, Jr., Selsted ME, Lehrer RI, and Parmley RT.** Defensin-rich dense granules of human neutrophils. *Blood* 70: 757-765, 1987.

128. **Rosenfeld Y and Shai Y.** Lipopolysaccharide (Endotoxin)-host defense antibacterial peptides interactions: Role in bacterial resistance and prevention of sepsis. *Biochim Biophys Acta* 1758: 1513-1522, 2006.

129. **Rosengren KJ, Daly NL, Fornander LM, Jonsson LM, Shirafuji Y, Qu X, Vogel HJ, Ouellette AJ, and Craik DJ.** Structural and functional characterization of the conserved salt bridge in mammalian paneth cell alpha-defensins: solution structures of mouse CRYPTDIN-4 and (E15D)-CRYPTDIN-4. *J Biol Chem* 281: 28068-28078, 2006.

130. **Rozek A, Powers JP, Friedrich CL, and Hancock RE.** Structure-based design of an indolicidin peptide analogue with increased protease stability. *Biochemistry* 42: 14130-14138, 2003.

131. **Sahl HG, Pag U, Bonness S, Wagner S, Antcheva N, and Tossi A.** Mammalian defensins: structures and mechanism of antibiotic activity. *J Leukoc Biol* 77: 466-475, 2005.

132. **Salzman NH, Ghosh D, Huttner KM, Paterson Y, and Bevins CL.** Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 422: 522-526, 2003.

133. **Satchell DP, Sheynis T, Kolusheva S, Cummings J, Vanderlick TK, Jelinek R, Selsted ME, and Ouellette AJ.** Quantitative interactions between cryptdin-4 amino terminal variants and membranes. *Peptides* 24: 1795-1805, 2003.

134. **Satchell DP, Sheynis T, Shirafuji Y, Kolusheva S, Ouellette AJ, and Jelinek R.** Interactions of mouse Paneth cell alpha-defensins and alpha-defensin precursors with membranes. Prosegment inhibition of peptide association with biomimetic membranes. *J Biol Chem* 278: 13838-13846, 2003.

135. **Sawai MV, Jia HP, Liu L, Aseyev V, Wiencek JM, McCray PB, Jr., Ganz T, Kearney WR, and Tack BF.** The NMR structure of human beta-defensin-2 reveals a novel alpha-helical segment. *Biochemistry* 40: 3810-3816, 2001.

136. **Sawai MV, Waring AJ, Kearney WR, McCray PB, Jr., Forsyth WR, Lehrer RI, and Tack BF.** Impact of single-residue mutations on the structure and function of ovispirin/novispirin antimicrobial peptides. *Protein Eng* 15: 225-232, 2002.

137. **Schibli DJ, Hunter HN, Aseyev V, Starner TD, Wiencek JM, McCray PB, Jr., Tack BF, and Vogel HJ.** The solution structures of the human beta-defensins lead to a better understanding of the potent bactericidal activity of HBD3 against Staphylococcus aureus. *J Biol Chem* 277: 8279-8289, 2002.

138. **Schibli DJ, Nguyen LT, Kernaghan SD, Rekdal O, and Vogel HJ.** Structure-function analysis of tritrpticin analogs: potential relationships between antimicrobial activities, model membrane interactions and their micelle-bound NMR structures. *Biophys J*, 2006.

139. **Scocchi M, Zelezetsky I, Benincasa M, Gennaro R, Mazzoli A, and Tossi A.** Structural aspects and biological properties of the cathelicidin PMAP-36. *Febs J* 272: 4398-4406, 2005.

140. **Selsted ME and Ouellette AJ.** Mammalian defensins in the antimicrobial immune response. *Nat Immunol* 6: 551-557, 2005.

141. **Shaykhiev R, Beisswenger C, Kandler K, Senske J, Puchner A, Damm T, Behr J, and Bals R.** Human endogenous antibiotic LL-37 stimulates airway epithelial cell proliferation and wound closure. *Am J Physiol Lung Cell Mol Physiol* 289: L842-848, 2005.

142. **Shi J, Ross CR, Chengappa MM, Sylte MJ, McVey DS, and Blecha F.** Antibacterial activity of a synthetic peptide (PR-26) derived from PR- 39, a proline-arginine-rich neutrophil antimicrobial peptide. *Antimicrob Agents Chemother* 40: 115-121, 1996.

143. **Shirafuji Y, Tanabe H, Satchell DP, Henschen-Edman A, Wilson CL, and Ouellette AJ.** Structural determinants of procryptdin recognition and cleavage by matrix metalloproteinase-7. *J Biol Chem* 278: 7910-7919, 2003.

144. **Sigurdardottir T, Andersson P, Davoudi M, Malmsten M, Schmidtchen A, and Bodelsson M.** In silico identification and biological evaluation of antimicrobial peptides based on human cathelicidin LL-37. *Antimicrob Agents Chemother* 50: 2983-2989, 2006.

145. **Skalicky JJ, Selsted ME, and Pardi A.** Structure and dynamics of the neutrophil defensins NP-2, NP-5, and HNP-1: NMR studies of amide hydrogen exchange kinetics. *Proteins* 20: 52-67, 1994.

146. **Skerlavaj B, Gennaro R, Bagella L, Merluzzi L, Risso A, and Zanetti M.** Biological characterization of two novel cathelicidin-derived peptides and identification of structural requirements for their antimicrobial and cell lytic activities. *J Biol Chem* 271: 28375-28381, 1996.

147. **Song YM, Park Y, Lim SS, Yang ST, Woo ER, Park IS, Lee JS, Kim JI, Hahm KS, Kim Y, and Shin SY.** Cell selectivity and mechanism of action of antimicrobial model peptides containing peptoid residues. *Biochemistry* 44: 12094-12106, 2005.

148. **Song YM, Yang ST, Lim SS, Kim Y, Hahm KS, Kim JI, and Shin SY.** Effects of L- or D-Pro incorporation into hydrophobic or hydrophilic helix face of amphipathic alpha-helical model peptide on structure and cell selectivity. *Biochem Biophys Res Commun* 314: 615-621, 2004.

149. **Storici P, Scocchi M, Tossi A, Gennaro R, and Zanetti M.** Chemical synthesis and biological activity of a novel antibacterial peptide deduced from a pig myeloid cDNA. *FEBS Lett* 337: 303-307, 1994.

150. **Subbalakshmi C and Sitaram N.** Mechanism of antimicrobial action of indolicidin. *FEMS MicrobiolLett* 160: 91-96, 1998.

151. **Tack BF, Sawai MV, Kearney WR, Robertson AD, Sherman MA, Wang W, Hong T, Boo LM, Wu H, Waring AJ, and Lehrer RI.** SMAP-29 has two LPS-binding sites and a central hinge. *Eur J Biochem* 269: 1181-1189, 2002.

152. **Tan BH, Meinken C, Bastian M, Bruns H, Legaspi A, Ochoa MT, Krutzik SR, Bloom BR, Ganz T, Modlin RL, and Stenger S.** Macrophages acquire neutrophil granules for antimicrobial activity against intracellular pathogens. *J Immunol* 177: 1864-1871, 2006.

153. **Tanabe H, Qu X, Weeks CS, Cummings JE, Kolusheva S, Walsh KB, Jelinek R, Vanderlick TK, Selsted ME, and Ouellette AJ.** Structure-activity determinants in paneth cell alpha-defensins: loss-of-function in mouse cryptdin-4 by charge-reversal at arginine residue positions. *J Biol Chem* 279: 11976-11983, 2004.

154. **Tanabe H, Yuan J, Zaragoza MM, Dandekar S, Henschen-Edman A, Selsted ME, and Ouellette AJ.** Paneth cell alpha-defensins from rhesus macaque small intestine. *Infect Immun* 72: 1470-1478, 2004.

155. **Tang YQ, Yuan J, Miller CJ, and Selsted ME.** Isolation, characterization, cDNA cloning, and antimicrobial properties of two distinct subfamilies of alpha-defensins from rhesus macaque leukocytes. *Infect Immun* 67: 6139-6144, 1999.

156. **Tang YQ, Yuan J, Osapay G, Osapay K, Tran D, Miller CJ, Ouellette AJ, and Selsted ME.** A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. *Science* 286: 498-502, 1999.

157. **Tani K, Murphy WJ, Chertov O, Salcedo R, Koh CY, Utsunomiya I, Funakoshi S, Asai O, Herrmann SH, Wang JM, Kwak LW, and Oppenheim JJ.** Defensins act as potent adjuvants that promote cellular and humoral immune responses in mice to a lymphoma idiotype and carrier antigens. *IntImmunol* 12: 691-700, 2000.

158. **Territo MC, Ganz T, Selsted ME, and Lehrer R.** Monocyte-chemotactic activity of defensins from human neutrophils. *JClinInvest* 84: 2017-2020, 1989.

159. **Tew GN, Clements D, Tang H, Arnt L, and Scott RW.** Antimicrobial activity of an abiotic host defense peptide mimic. *Biochim Biophys Acta* 1758: 1387-1392, 2006.

160. **Tew GN, Liu D, Chen B, Doerksen RJ, Kaplan J, Carroll PJ, Klein ML, and DeGrado WF.** De novo design of biomimetic antimicrobial polymers. *Proc Natl Acad Sci U S A* 99: 5110-5114, 2002.

161. **Tokumaru S, Sayama K, Shirakata Y, Komatsuzawa H, Ouhara K, Hanakawa Y, Yahata Y, Dai X, Tohyama M, Nagai H, Yang L, Higashiyama S, Yoshimura A, Sugai M, and Hashimoto K.** Induction of keratinocyte migration via transactivation of the epidermal growth factor receptor by the antimicrobial peptide LL-37. *J Immunol* 175: 4662-4668, 2005.

162. **Tomasinsig L and Zanetti M.** The cathelicidins--structure, function and evolution. *Curr Protein Pept Sci* 6: 23-34, 2005.

163. **Tossi A, Sandri L, and Giangaspero A.** Amphipathic, alpha-helical antimicrobial peptides. *Biopolymers* 55: 4-30, 2000.

164. **Tossi A, Scocchi M, Skerlavaj B, and Gennaro R.** Identification and characterization of a primary antibacterial domain in CAP18, a lipopolysaccharide binding protein from rabbit leukocytes. *FEBS Lett* 339: 108-112, 1994.

165. **Trabi M, Schirra HJ, and Craik DJ.** Three-dimensional structure of RTD-1, a cyclic antimicrobial defensin from Rhesus macaque leukocytes. *Biochemistry* 40: 4211-4221, 2001.

166. **Travis SM, Anderson NN, Forsyth WR, Espiritu C, Conway BD, Greenberg EP, McCray PB, Jr., Lehrer RI, Welsh MJ, and Tack BF.** Bactericidal activity of mammalian cathelicidin-derived peptides. *Infect Immun* 68: 2748-2755, 2000.

167. **van den Bogaard AE and Stobberingh EE.** Antibiotic usage in animals: impact on bacterial resistance and public health. *Drugs* 58: 589-607, 1999.

168. **van Wetering S, Mannesse-Lazeroms SP, van Sterkenburg MA, and Hiemstra PS.** Neutrophil defensins stimulate the release of cytokines by airway epithelial cells: modulation by dexamethasone. *Inflamm Res* 51: 8-15, 2002.

169. **Violette A, Fournel S, Lamour K, Chaloin O, Frisch B, Briand JP, Monteil H, and Guichard G.** Mimicking helical antibacterial peptides with nonpeptidic folding oligomers. *Chem Biol* 13: 531-538, 2006.

170. **Walsh C and Wright G.** Introduction: antibiotic resistance. *Chem Rev* 105: 391-394, 2005.

171. **Wang W, Mulakala C, Ward SC, Jung G, Luong H, Pham D, Waring AJ, Kaznessis Y, Lu W, Bradley KA, and Lehrer RI.** Retrocyclins Kill Bacilli and Germinating Spores of Bacillus anthracis and Inactivate Anthrax Lethal Toxin. *J Biol Chem* 281: 32755-32764, 2006.

172. **Wang W, Owen SM, Rudolph DL, Cole AM, Hong T, Waring AJ, Lal RB, and Lehrer RI.** Activity of alpha- and theta-defensins against primary isolates of HIV-1. *J Immunol* 173: 515-520, 2004.

173. **Weiss TM, Yang L, Ding L, Waring AJ, Lehrer RI, and Huang HW.** Two states of cyclic antimicrobial peptide RTD-1 in lipid bilayers. *Biochemistry* 41: 10070-10076, 2002.

174. **Wenzel RP.** The antibiotic pipeline--challenges, costs, and values. *N Engl J Med* 351: 523-526, 2004.

175. **Wilson CL, Ouellette AJ, Satchell DP, Ayabe T, Lopez-Boado YS, Stratman JL, Hultgren SJ, Matrisian LM, and Parks WC.** Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. *Science* 286: 113-117, 1999.

176. **Wimley WC, Selsted ME, and White SH.** Interactions between human defensins and lipid bilayers: evidence for formation of multimeric pores. *Protein Sci* 3: 1362-1373, 1994.

177. **Wu M and Hancock RE.** Improved derivatives of bactenecin, a cyclic dodecameric antimicrobial cationic peptide. *Antimicrob Agents Chemother* 43: 1274-1276, 1999.

178. **Wu Z, Hoover DM, Yang D, Boulegue C, Santamaria F, Oppenheim JJ, Lubkowski J, and Lu W.** Engineering disulfide bridges to dissect antimicrobial and chemotactic activities of human beta-defensin 3. *Proc Natl Acad Sci U S A* 100: 8880-8885, 2003.

179. **Wu Z, Li X, de Leeuw E, Ericksen B, and Lu W.** Why is the Arg5-Glu13 salt bridge conserved in mammalian alpha-defensins? *J Biol Chem* 280: 43039-43047, 2005.

180. **Xiao Y, Cai Y, Bommineni YR, Fernando SC, Prakash O, Gilliland SE, and Zhang G.** Identification and functional characterization of three chicken cathelicidins with potent antimicrobial activity. *J Biol Chem* 281: 2858-2867, 2006.

181. **Xiao Y, Dai H, Bommineni YR, Soulages JL, Gong Y-X, Prakash O, and Zhang G.** Structure-activity relationships of fowlicidin-1, a cathelicidin antimicrobial peptide in chicken. *FEBS J* 273: 2581-2593, 2006.

182. **Xie C, Prahl A, Ericksen B, Wu Z, Zeng P, Li X, Lu WY, Lubkowski J, and Lu W.** Reconstruction of the conserved beta-bulge in mammalian defensins using D-amino acids. *J Biol Chem* 280: 32921-32929, 2005.

183. **Yamaguchi S, Hong T, Waring A, Lehrer RI, and Hong M.** Solid-state NMR investigations of peptide-lipid interaction and orientation of a beta-sheet antimicrobial peptide, protegrin. *Biochemistry* 41: 9852-9862, 2002.

184. **Yamaguchi S, Huster D, Waring A, Lehrer RI, Kearney W, Tack BF, and Hong M.** Orientation and dynamics of an antimicrobial peptide in the lipid bilayer by solid-state NMR spectroscopy. *Biophys J* 81: 2203-2214, 2001.

185. **Yang D, Chen Q, Chertov O, and Oppenheim JJ.** Human neutrophil defensins selectively chemoattract naive T and immature dendritic cells. *J Leukoc Biol* 68: 9-14, 2000.

186. **Yang D, Chertov O, Bykovskaia SN, Chen Q, Buffo MJ, Shogan J, Anderson M, Schroder JM, Wang JM, Howard OM, and Oppenheim JJ.** Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science* 286: 525-528, 1999.

187. **Yang ST, Jeon JH, Kim Y, Shin SY, Hahm KS, and Kim JI.** Possible role of a PXXP central hinge in the antibacterial activity and membrane interaction of PMAP-23, a member of cathelicidin family. *Biochemistry* 45: 1775-1784, 2006.

188. **Yang ST, Shin SY, Hahm KS, and Kim JI.** Design of perfectly symmetric Trp-rich peptides with potent and broad-spectrum antimicrobial activities. *Int J Antimicrob Agents* 27: 325-330, 2006.

189. **Yang YH, Wu WK, Tai EK, Wong HP, Lam EK, So WH, Shin VY, and Cho CH.** The cationic host defense peptide rCRAMP promotes gastric ulcer healing in rats. *J Pharmacol Exp Ther* 318: 547-554, 2006.

190. **Yasin B, Wang W, Pang M, Cheshenko N, Hong T, Waring AJ, Herold BC, Wagar EA, and Lehrer RI.** Theta defensins protect cells from infection by herpes simplex virus by inhibiting viral adhesion and entry. *J Virol* 78: 5147-5156, 2004.

191. **Yount NY and Yeaman MR.** Multidimensional signatures in antimicrobial peptides. *Proc Natl Acad Sci U S A* 101: 7363-7368, 2004.

192. **Yu K, Park K, Kang SW, Shin SY, Hahm KS, and Kim Y.** Solution structure of a cathelicidin-derived antimicrobial peptide, CRAMP as determined by NMR spectroscopy. *J Pept Res* 60: 1-9, 2002.

193. **Yu Q, Lehrer RI, and Tam JP.** Engineered salt-insensitive alpha-defensins with end-to-end circularized structures. *J Biol Chem* 275: 3943-3949, 2000.

194. **Zanetti M.** Cathelicidins, multifunctional peptides of the innate immunity. *J Leukoc Biol* 75: 39-48, 2004.

195. **Zanetti M.** The role of cathelicidins in the innate host defenses of mammals. *Curr Issues Mol Biol* 7: 179-196, 2005.

196. **Zanetti M, Gennaro R, Skerlavaj B, Tomasinsig L, and Circo R.** Cathelicidin peptides as candidates for a novel class of antimicrobials. *Curr Pharm Des* 8: 779-793, 2002.

197. **Zasloff M.** Antimicrobial peptides of multicellular organisms. *Nature* 415: 389-395, 2002.

198. **Zelezetsky I, Pacor S, Pag U, Papo N, Shai Y, Sahl HG, and Tossi A.** Controlled alteration of the shape and conformational stability of alpha-helical cell-lytic peptides: effect on mode of action and cell specificity. *Biochem J* 390: 177-188, 2005.

199. **Zelezetsky I, Pag U, Sahl HG, and Tossi A.** Tuning the biological properties of amphipathic alpha-helical antimicrobial peptides: rational use of minimal amino acid substitutions. *Peptides* 26: 2368-2376, 2005.

200. **Zelezetsky I, Pontillo A, Puzzi L, Antcheva N, Segat L, Pacor S, Crovella S, and Tossi A.** Evolution of the Primate Cathelicidin. Correlation between structural variations and antimicrobial activity. *J Biol Chem* 281: 19861-19871, 2006.

201. **Zelezetsky I and Tossi A.** Alpha-helical antimicrobial peptides--using a sequence template to guide structure-activity relationship studies. *Biochim Biophys Acta* 1758: 1436-1449, 2006.

202. **Zhang L and Falla TJ.** Antimicrobial peptides: therapeutic potential. *Expert Opin Pharmacother* 7: 653-663, 2006.

203. **Zhu WL, Lan H, Park Y, Yang ST, Kim JI, Park IS, You HJ, Lee JS, Park YS, Kim Y, Hahm KS, and Shin SY.** Effects of Pro --> Peptoid Residue Substitution on Cell Selectivity and Mechanism of Antibacterial Action of Tritrpticin-Amide Antimicrobial Peptide. *Biochemistry* 45: 13007-13017, 2006.

204. **Zimmermann GR, Legault P, Selsted ME, and Pardi A.** Solution structure of bovine neutrophil beta-defensin-12: the peptide fold of the beta-defensins is identical to that of the classical defensins. *Biochemistry* 34: 13663-13671, 1995.

205. **Zuyderduyn S, Ninaber DK, Hiemstra PS, and Rabe KF.** The antimicrobial peptide LL-37 enhances IL-8 release by human airway smooth muscle cells. *J Allergy Clin Immunol* 117: 1328-1335, 2006.

CHAPTER III

RAPID EVOLUTION AND DIVERSIFICATION OF MAMMALIAN

α-DEFENSINS AS REVEALED BY COMPARATIVE ANALYSIS OF RODENT

AND PRIMATE GENES

Amar Patil[1], Austin L. Hughes[2], and Guolong Zhang[1]

[1]Department of Animal Science, Oklahoma State University, Stillwater, OK 74078;

[2]Department of Biological Sciences, University of South Carolina, Columbia, SC 29208.

ABSTRACT

Mammalian α-defensins constitute a family of cysteine-rich, cationic antimicrobial peptides produced by phagocytes and intestinal Paneth cells, playing an important role in innate host defense. Following comprehensive computational searches, here we report the discovery of complete repertoires of the α-defensin gene family in the human, chimpanzee, rat, and mouse with new genes identified in each species. The human genome was found to encode a cluster of 10 distinct α-defensin genes and pseudogenes expanding 132 kb continuously on chromosome 8p23. Such α-defensin loci are also conserved in the syntenic chromosomal regions of chimpanzee, rat, and mouse. Phylogenetic analyses showed formation of two distinct clusters with primate α-defensins forming one cluster and rodent enteric α-defensins forming the other cluster. Species-specific clustering of genes is evident in non-primate species, but not in the primates. Phylogenetically distinct subsets of α-defensins also exist in each species with most subsets containing multiple members. In addition, natural selection appears to have acted to diversify the functionally active mature defensin region but not signal or prosegment sequences. We concluded that mammalian α-defensin genes may have evolved from two separate ancestors originated from β-defensins. The current repertoire of the α-defensin gene family in each species are primarily a result of repeated gene duplication and positive diversifying selection after divergence of mammalian species from each other, except for the primate genes, which were evolved prior to the separation of the primate species. We argue that the presence of multiple, divergent subsets of α-defensins in each

54

species may help animals to better cope with different microbial challenges in the ecological niches which they inhabit.

INTRODUCTION

The defensins of vertebrates constitute a large family of cationic antimicrobial peptides that are characterized by the presence of six cysteine residues at defined positions. According to the spacing pattern of cysteine residues, these defensins are further classified into three subfamilies, namely α-, β- and θ-defensins with each consisting of a unique six-cysteine motif and a disulfide bonding pattern (6, 16, 21, 29). For example, the consensus α-defensin motif is $C-X_1-C-X_{3-4}-C-X_9-C-X_{6-10}-C-C$, where C1-C6, C2-C4 and C3-C5 form three intramolecular disulfide bridges. While β-defensins have been found in most vertebrate species with a much wider tissue expression pattern, α-defensins are specific to mammals and are mainly produced by leukocytes of myeloid origin and Paneth cells of small intestine (6, 16, 21, 29). On the other hand, θ-defensins are believed to have evolved from α-defensins and have only been discovered in leukocytes of primates (6, 16, 20).

Defensins possess potent antimicrobial activity against a broad range of bacteria, fungi, and enveloped viruses (6, 16, 21, 29). In addition, certain defensins are also capable of inducing maturation of dendritic cells and spermatocytes, inducing secretion of chloride and proinflammatory cytokines from epithelial cells, and chemoattracting immune and inflammatory cells (6, 16, 40). The mechanism by which defensins kill

55

microbes involves initial electrostatic interactions of cationic peptides with negatively charged phospholipids on the microbial membranes, followed by permeabilization of bacterial membrane and cell death (6). In contrast to most conventional antibiotics that kill microbes by certain biochemical mechanisms, such a physical mechanism of action confers on defensins a broad-spectrum antimicrobial activity with equal efficacy against antibiotic-resistant strains and little risk of developing resistance (8, 43). Consequently, defensins are being explored as a new class of antimicrobials mainly to control antibiotic-resistant microorganisms.

To date, a number of α-defensins have been discovered in human, rhesus macaque, rabbit, guinea pig, mouse, and rat (6, 16, 21). For example, three myeloid α-defensin genes (*DEFA1/2, DEFA3* and *DEFA4)* and two enteric α-defensin genes (*DEFA5/HD5* and *DEFA6/HD6*) have been found in humans, and a group of 19 highly homologous enteric α-defensins (cryptdins) have also been reported in mice. All α-defensins are synthesized from 80-105 amino acid precursors with each composed of a short N-terminal signal sequence (~20 amino acids), an anionic prosegment (40-50 amino acids), and a C-terminal cationic mature peptide (30-35 amino acids) (6, 16, 21). The α-defensins are stored as either inactive proforms or mature active forms in the cytoplasmic granules of phagocytes and Paneth cells and released (and processed if necessary) in response to microbial infection or cholinergic stimulation (6, 16, 21). The enzyme responsible for processing and activation of intestinal α-defensins is metalloproteinase matrilysin (MMP7) in mice (38) and trypsin in humans (7), respectively. The contribution of α-defensins to intestinal mucosal immunity has been best demonstrated by the

evidence that transgenic mice expressing human *DEFA5* or *HD5* became fully resistant to oral lethal infection with *Salmonella typhimurium* (27) and that *MMP7*-deficiency rendered mice more susceptible to *S. typhimurium* infection (38).

The genomes of several evolutionarily divergent vertebrate species have been sequenced, including zebrafish, Japanese pufferfish (*Fugu rubripes*), chicken, dog, rat, mouse, chimpanzee and human. Availability of such a large amount of sequence information provides a timely opportunity to search for the possible existence of α-defensins in these species. We hypothesized that identification of the entire repertoire of α-defensin genes across a range of animal species will reveal the origin and evolutionary relationships of this important gene family. This also serves as a first step to study the role of α-defensins in host immunity and explore their potential as novel antimicrobials. Following comprehensive genome-wide screening, here we report identification of the complete repertoires of α-defensin genes in the human, chimpanzee, mouse, and rat with novel sequences being discovered in each species. However, none of α-defensins have been found in any species other than glires (mouse, rat, guinea pig, and rabbit) and primates (human, chimpanzee, olive baboon, and rhesus macaque). In addition, we provide strong evidence showing that a rapid duplication and positive diversifying selection of α-defensin genes have occurred following divergence of mammalian species.

MATERIALS AND METHODS

Computational search for novel α-defensins

      To identify potential novel sequences in all vertebrate species with whole genome sequences available, all known α-defensin peptide sequences were individually queried against expressed sequence tags (EST), nonredundant sequences (NR), unfinished high throughput genomic sequences (HTGS), and whole genome shotgun sequences (WGS) in the GenBank by using the TBLASTN program (1) with the default settings on the NCBI web site (http://www.ncbi.nlm.nih.gov/BLAST). The vertebrate species examined included zebrafish, Japanese pufferfish, chicken, dog, rat, mouse, chimpanzee, and human. All potential hits were then examined for presence of the characteristic α-defensin motif or highly conserved signal/prosegment sequence. For every novel α-defensin sequence identified, additional iterative BLAST searches were performed as described above until no more novel sequences can be revealed. Because mammalian defensins tend to form clusters (17, 30, 31), all genomic sequences containing α-defensins were also retrieved from the GenBank in order to discover potential novel sequences with distant homology. The nucleotide sequences between two neighboring defensin genes were translated into six open reading frames and individually compared with the two defensin peptide sequences for the presence of α-defensin motif or signal/prosegment sequence by using the BLASTP (1) on the NCBI web site (http://www.ncbi.nlm.nih.gov/blast/bl2seq/) and/or ClustalW program (version 1.82) (35) on the European Bioinformatics Institute web site (http://www.ebi.ac.uk/clustalw).

Prediction of full-length coding sequences and genomic structures of α-defensins

All known α-defensin precursors are encoded in two separate exons separated by a short intron of less than 2 kb with one exon encoding signal/prosegment sequence and the other exon encoding the mature peptide containing the six-cysteine α-defensin motif (22). If either the signal/prosegment sequence or α-defensin motif of a novel gene was missing in a genomic sequence within a 2-kb distance, an approximately 5-kb sequence flanking the α-defensin motif or signal/prosegment sequence was retrieved to identify the full-length coding sequence and to derive the structural organization of that novel α-defensin gene by using a combination of GenomeScan (41), GENSCAN (4) and/or GeneWise (3). In the case of the rat, all computational predictions of coding sequences, except for the pseudogenes, have also been confirmed by cloning and sequencing of their respective RT-PCR products amplified from appropriate tissues.

Identification and characterization of α-defensin gene clusters

To determine the relative position and orientation of each defensin in the genome, individual defensins were searched against the assembled genomes of human (NCBI Build 35), chimpanzee (NCBI Build 1 Version 1), mouse (NCBI Build 33), and rat (BCM Version 3.1) released in May 2004, November 2003, May 2004, and June 2003, respectively. The BLAT program (14) was used for gene mapping through the UCSC Genome Browser (http://genome.ucsc.edu). The chromosomal locations of the α-defensin gene clusters of human, mouse, and rat were revealed by using the Map Viewer Program (http://www.ncbi.nlm.nih.gov/mapview).

Sequence alignment and molecular evolutionary analysis

Multiple sequence alignments were carried out by using the ClustalW program (version 1.82) (35) (http://www.ebi.ac.uk/clustalw). The neighbor-joining method (26) was used to construct the phylogenetic tree by calculating the proportion of nucleotide or amino acid differences (*p*-distance), and the reliability of each branch was tested by 1000 bootstrap replications. Pairwise comparisons of nucleotide sequences at the codon level was carried out by using the method of Nei and Gojobori (19) to estimate the number of nonsynonymous substitutions per nonsynonymous site (d_N) and the number of synonymous substitutions per synonymous site (d_S) with the Jukes-Cantor correction for multiple substitutions. Construction of the phylogenetic tree and pairwise comparison of nucleotide or amino acid sequences were carried out by using the MEGA software version 2.1 (15).

RT-PCR analysis of tissue expression patterns of α-defensins

A total of 26 different tissue samples were harvested from healthy, 2-month old Sprague-Dawley rats. Bone marrow progenitor cells were collected from femur, followed by hypotonic lysis of erythrocytes as described (44, 45). Only lineages of white blood cells were used for RNA isolation. Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA). A panel of 12 human tissue RNA, including bone marrow, small intestine, kidney, trachea, lung, liver, heart, spleen, skeletal muscle, testis, prostate, and uterus, were also purchased from BD Biosciences Clontech (Palo Alto, CA) and used for evaluating the expression pattern of *DEFA7*. For each human and rat RNA, 4 μg were reverse transcribed with random hexamers and Superscript II reverse transcriptase by

using a first-strand cDNA synthesis kit (Invitrogen) according to the instructions. The subsequent PCR was carried essentially as described (44, 45). Briefly, 1/40 of the first-strand cDNA from each tissue were used to amplify α-defensins and glyceraldehydes 3-phosphate dehydrogenase (GAPDH) with gene specific primers. Every pair of primers were designed from less conservative regions located on different exons to aid in specific amplification of target genes and in distinguishing PCR products amplified from cDNA versus genomic DNA (Table 1). The PCR program used was: $94^{o}C$ denaturation for 2 min, followed by different cycles of $94^{o}C$ denaturation for 20 sec, $55^{o}C$ annealing for 20 sec, and $72^{o}C$ extension for 40 sec, followed by a final extension at $72^{o}C$ for 5 min. The sequence of primers and number of PCR cycles used are shown in Table 1. The PCR products were analyzed by electrophoresis on 1.2% agarose gels containing 0.5 µg/ml ethidium bromide. The specificity of each PCR product was confirmed by cloning of the PCR product into T/A cloning vector, followed by sequencing of the recombinant plasmid.

GenBank accession numbers

All cDNA sequences of new α- and θ-defensins described in this manuscript have been submitted to the GenBank. Accession numbers are as follows: rat, *Defa6-Defa11*(AY623750-AY623755), *Defa12-ps-Defa14-ps* (AY746422-AY746424), *Defa-rs1*(AY623756); mouse, *Defcr20-Defcr25* (AY746425-AY746430), *Defcr26* (AY766470), *Defa-ps1* (AY746431), CRS1C-2 (AY761183), CRS1C-3 (AY761184), CRS4C-6 (AY761185); human, *DEFA7P-DEFA11P* (AY746432-AY746436);

chimpanzee, *PTAD1-PTAD11-ps* (AY746437-AY746445), *PTTD1-ps-PTTD2-ps* (AY746446-AY746447); and olive baboon, *PAAD1-PAAD3* (AY746448-AY746450).

RESULTS

Genome-wide identification of α-defensins in rodents and primates

To identify possible existence of novel α-defensin sequences in the vertebrate species whose genomes have been sequenced, all known α-defensin peptide sequences were individually queried against EST, NR, HTGS, and WGS sequences of zebrafish, Japanese pufferfish, chicken, dog, rat, mouse, chimpanzee, and human in the GenBank by using the TBLASTN program (1). For every novel α-defensin sequence identified, additional iterative BLAST searches were performed until no more novel sequences could be revealed. All genomic sequences containing α-defensins were also retrieved from the GenBank and translated into six open reading frames and curated for the presence of the α-defensin motif or signal/prosegment sequence to reveal potential sequences with distant homology. Whenever necessary, a combination of gene prediction programs, including GenomeScan (41), GENSCAN (4) and/or GeneWise (3), were used to predict full-length coding sequences of novel α-defensins from genomic sequences.

As a result, a number of sequences have been discovered in the human, chimpanzee, rat and mouse. In the case of human, five novel α-defensin genes were found in addition to five previously characterized ones, *DEFA1-6* (6, 21). Among new α-defensin genes identified is *DEFA7*, which consists of a canonical six-cysteine α-defensin

62

motif and an N-terminal signal/prosegment sequence highly homologous to known human α-defensins (Fig. 1A).

Because *DEFA7* lacks the start codon in the consensus position that initiates the translation of all other α-defensins, we sought to confirm if it is a functional gene. However, RT-PCR analysis revealed no expression of *DEFA7* in any of 12 human tissues examined, including small intestine, bone marrow, and kidney, three of which are the most dominant sites to produce α-defensins in other mammalian species. Although it is still possible to be transcribed in other places, *DEFA7* is most likely to be a pseudogene and therefore, is designated as *DEFA7P*. Such results are also consistent with the fact that there is no single EST sequence for *DEFA7P* in the GenBank, despite of nearly 5.8 million ESTs being deposited as of September 1, 2004.

In addition to *DEFA7P*, four α-defensin pseudogenes, namely *DEFA8P-11P*, have also been discovered in the human (Fig. 1A). Without a premature stop codon, each pseudogene would otherwise encode a peptide that is highly similar to a typical α-defensin (Fig. 1A). Identification of additional five α-defensin genes in the extensively studied human genome highlighted the power of our computational search strategy.

A total of 11 novel α-defensins, including four pseudogenes, have also been identified in the chimpanzee genome. These genes, which were termed *Pan troglodytes* α-defensins (*PTAD*) 1-11, are all human α-defensin orthologs (Fig. 1A), and therefore are expected to show a similar tissue expression pattern to human genes. *PTAD1* and *PTAD4*

are predicted to be neutrophil-specific, whereas *PTAD5* and *PTAD6* are expected to be expressed mainly in Paneth cells of small intestine. Interestingly, *PTAD7,* the chimpanzee ortholog of *DEFA7P*, appears to be a functional gene with the start codon at the consensus position (Fig. 1A), but its expression site(s) remain(s) to be studied. Four orthologs of human α-defensin pseudogenes are also present in the chimpanzee, namely *PTAD8-ps* to *PTAD11-ps*.

In addition to α-defensins, two pseudogenes for θ-defensin, *PTTD1-ps* and *PTTD2-ps*, have also been found in the chimpanzee (Fig. 1A). Both PTTD1-ps and PTTD2-ps are highly homologous to each other and to known human θ-defensin pseudogene, DEFT1P (6, 16) (Fig. 1A), but they differ with DEFT1P in the locations of premature stop codons. DEFT1P contains two stop codons at the amino acid positions 17 and 77; however, only the former stop codon is preserved in PTTD1-ps and PTTD2-ps.

Three new α-defensins, namely *PAAD1-3*, were also identified in two bacterial artificial chromosome (BAC) clones (accession no. AC116558 and AC116559) of olive baboon (*Papio anubis*) (Fig. 1A). *PAAD1-2* are orthologous to *DEFA5*, whereas *PAAD3* is a *DEF7P* homolog. However, unlike *DEF7P* that contains a single nucleotide mutation at the start codon, *PAAD3* is composed of a stretch of completely different nucleotides around the start codon (See supplementary Fig. 1), and therefore it is also predicted to be a pseudogene. This prediction is further supported by the fact that, despite of extensive experimental screening efforts (33, 34), no *PAAD3*-like sequences have been found to be produced in either jejunum or leukocytes of rhesus macaque, which is a member of

64

cercopithecids (Old World Monkeys) as olive baboon. Presumably, such *PAAD3*-like gene(s) are also inactive in rhesus macaque; however, it remains to be experimentally verified.

In the rat, a total of 14 α-defensin genes have been discovered, including four known ones (*RatNP1/2, RatNP3, RatNP4,* and *RD5*) (5, 42). Among ten novel rat α-defensins (*Defa*), six genes (*Defa6-11*) encode peptide sequences varying 87-104 amino acid residues in length with a canonical six-cysteine α-defensin motif (Fig. 1B). Also identified were three α-defensin pseudogenes, namely *Defa12-ps*, *Defa13-ps*, and *Defa14-ps*, with each having a premature stop codon at amino acid positions 34, 29, and 88, respectively (Fig. 1B). In addition, one α-defensin related sequence (*Defa-rs1*) was also found in the rat. Despite of the conservation in the N-terminal signal/prosegment sequence, Defa-rs1 has only five cysteine residues at the C-terminus with no characteristic α-defensin motif (Fig. 1B). Similarly, two groups of cryptdin-related sequences (CRS1C and CRS4C) with the signal/prosegment sequences highly homologous to α-defensins, but with a different number and spacing pattern of cysteines were reported earlier in mice (13). Such CRS peptides were shown recently to form covalent homo- or heterodimers and are capable of killing bacteria (9). Therefore, it is tempting to speculate that Defa-rs1, the only α-defensin -related sequence in the rat, associates with itself by forming homodimers and participates in innate defense in the gut as other enteric α-defensins.

Among 19 α-defensins (cryptdins) reported in mice, 17 belong to the highly homogeneous cryptdin 1 (Defcr1) subgroup with less than 5 amino acids in difference in the mature sequence and many are believed to be derived from allelic variants of the same gene (22). *Defcr4* and *Defcr5* represent distant cryptdin family members and their paralogous genes have not been reported in mice. Among eight novel α-defensins that we have identified in the mouse, five sequences (*Defcr23-Defcr26* and *Defcr-ps1*) can be grouped to the *Defcr1* subgroup, whereas the other three (*Defcr20-Defcr22*) are similar to *Defcr4* (Fig. 1C). In addition, three novel cryptdin-related sequences were also found in mice (Fig. 1D). Among them, CRS1C-2 and CRS1C-3 are highly homologous to CRS1C-1, which contains a total of 11 cysteines with eight forming C-X-Y triplet repeats. On the other hand, CRS4C-6 belongs to a distant member of the CRS4C subgroup. Instead of consisting seven C-X-Y repeats, CRS4C-6 has an extra cysteine located close to the N-terminal end of the mature sequence. Unlike other CRS4C peptides that form intermolecular covalent dimmers (9), CRS4C-6 is likely to function as a monomer or non-covalently associated dimmers/oligomers. Interestingly, none of α-defensin sequences have been found in the species other than glires and primates. This is perhaps not surprising, given the fact that no α-defensins have been reported in any other vertebrate species despite of the extensive searching efforts in the past (6, 21).

Tissue expression patterns of rat α-defensins

A semi-quantitative RT-PCR was used to analyze tissue expression patterns of rat α-defensins. As shown in Fig. 2, all rat α-defensins exhibited two distinct patterns of expression. One group, which include *RatNP1/2*, *RatNP3,* and *Defa12-ps,* are

66

preferentially expressed in bone marrow with or without expression in small intestine (Fig. 2). *Defa7*, *Defa10*, and *Defa11* also belong to this group with little but specific expression in bone marrow (data not shown). In contrast, the other group consisting of *RD5*, *Defa6*, *Defa8*, *Defa9*, and *Defa-rs1* are specifically expressed in small intestine (Fig. 2). These genes are highly transcribed in jejunum and ileum with a low level of expression in duodenum. Interestingly, such a grouping is also consistent with the degree of conservation in the signal peptide sequences, which are more homologous within each group (top panel vs. bottom panel, Fig. 1B). It is surprising to see the expression of *RatNP3* and *Defa12-ps* in both bone marrow and small intestine, given the fact that all known α-defensins are produced either by myeloid or Paneth cells, but not by both cell lineages (21). The cells that express rat α-defensins in the intestinal tract are presumably Paneth cells, which produce enteric α-defensins in humans and mice (21). It is worthy noting that *Defa-rs1* is expressed in the jejunum and ileum at the comparable levels to most other enteric defensins (Fig. 2), reinforcing the notion that it most likely will function *in vivo*.

Molecular evolutionary analyses of mammalian α-defensins

　　　　To analyze the phylogenetic relationships among all known α-defensins in glires and primates including the newly identified ones, the proportion of amino acid differences (*p*-distance) of full-length peptides was calculated by the neighbor-joining method (26) (See supplementary Fig. 2 for alignment of peptide sequences of 82 α-defensins used). As shown in Fig. 3A, two major clusters are evident with α-defensins in the primates with rabbit and guinea pig forming one cluster supported by a bootstrap

value of 65%, and mouse and rat enteric α-defensins forming the other cluster with a bootstrap value of 98%. Rat α-defensins appear to be unique in that the genes of myeloid origin formed a separate cluster from the genes of enteric origin. Rat myeloid α-defensins tend to group with primate, rabbit, and guinea pig α-defensins.

Perhaps the most striking feature of the phylogenetic tress is that α-defensins formed species-specific clusters in the non-primate species (rat, mouse, guinea pig, and rabbit), implying that these genes may have undergone repeated duplication after these species diverged from a common ancestor. In contrast, no species-specific clustering was demonstrated among four primate species examined (human, chimpanzee, rhesus macaque, and olive baboon) (Fig. 3A). Instead, several distinct subclusters of α-defensins exist across the primate species. For example, chimpanzee and human are in a complete analogy with α-defensins present in all subclusters. The α-defensins of Old World Monkeys are located in three subclusters with human and chimpanzee *DEFA1-3, DEFA5,* and *DEFA7P,* respectively. These results suggested that many of the primate genes were likely evolved before the divergence of these species. The primate-specific θ-defensins, which exist in both cercopithecids and hominids, are clearly clustered together with the *DEFA10P* gene lineage, reinforcing the notion that θ-defensins were evolved from α-defensins after divergence of the primates from other mammalian species (6, 16, 20, 21).

To minimize possible biases of selection pressure exerted on exon (and therefore peptide) sequences, the evolutionary relationships of the α-defensin genes were further evaluated by analysis of the intron sequences within the open reading frame. The *p-*

distance of the introns of 59 α-defensins with known gene sequences was calculated by the neighbor-joining method (the intron sequences used will be available upon request) (26). Consistent with the results obtained from the full-length peptide sequences, a similar two major clusters of genes were obtained, in which mouse and rat enteric α-defensins were separated from others, supported by a bootstrap value of 100%, and hence were most likely evolved from a different ancestor (Fig. 3B). On the other hand, rat myeloid gene introns are clearly clustered together with primate α-defensins and therefore were originated from the same ancestral gene. As for rat and mouse enteric genes, there are two distinct subclusters with each being supported by a bootstrap value of 100%, implying that mouse cryptdins (*Defcr*) share the same ancestor with rat *RD5*, *Def13-ps*, and *Defa14-ps*, whereas mouse cryptdin-related sequences (*CRS1C* and *CRS4C*) were derived from the same primordial gene as rat *Defa6*, *Defa8*, and *Defa9*, and *Defa-rs1* (Fig. 3B). It is obvious that each subcluster has undergone a repeated duplication and diversification to give rise to multiple members in each species after these mammalian species diverged from a common ancestor.

Despite of significant sequence conservation in the first exon (Fig. 1C and 1D) and intron (data not shown), substantial differences exist in the cysteine motif encoded by the second exon between canonical α-defensins and three subgroups of related sequences in rats and mice, namely Defa-rs1, CRS1C, and CRS4C (Fig. 1C and 1D). To further address the origin and evolution of the cysteine motif of α-defensin-related sequences, nucleotide sequences of the second exon were compared with rat enteric α-defensins (*Defa6, Defa8*, and *Defa9*) in the same subcluster (Fig. 3B). It is evident that *Defa-rs1*

shares a high homology with three rat α-defensins throughout the entire coding region in the second exon (Fig. 4), suggesting that, in addition to the first exon and intron, the second exon of *Defa-rs1* was also originated from the same ancestral gene as rat enteric α-defensins. In fact, the change in the number and pattern of cysteines in Defa-rs1 was mainly because of nucleotide deletions (Fig. 4).

As for mouse CRS1C and CRS4C, both subgroups are highly homologous in the second exon (Fig. 4), and therefore share the same ancestor. The difference in the cysteine motif of CRS1C and CRS4C peptides were primarily because of deletions and insertions rather than nonsynonymous substitutions (Fig. 4). Although CRS peptides share a less homology with rat enteric α-defensins and *Defa-rs1*, patches of identical nucleotides are evident among them throughout the entire second exon (Fig. 4), implying that, similar to Defa-rs1, the changes in the second exon of CRS genes were likely to be a result of sequence diversification, instead of exon shuffling, which commonly occurred during the evolution of many gene families (23). Coupled with the fact that both intron and the first exon are highly conserved among classic α-defensins and related sequences, the full-length CRS genes most likely were derived from the same ancestral genes for rat enteric α-defensins and *Defa-rs1*. Apparently, CRS genes and *Defa-rs1* were independently evolved from classic α-defensins in mice and rats, respectively, following divergence of these two rodent species from each other.

Collectively, our results suggested that all mammalian α-defensins may have evolved from two ancestral genes with each giving rise to one major cluster of daughter

genes. The absence of α-defensins in non-mammalian species clearly indicated that α-defensins have evolved after the divergence of mammals from other vertebrate species. Formation of species-specific clusters, particularly in non-primate species, strongly suggested that many α-defensins have evolved independently after these mammalian species diverged from a common ancestor. It is worthy noting that, because of the lack of complete repertoires of α-defensin genes in the rabbit and guinea pig, the origin of α-defensins in these two species remains inconclusive. Analysis of coding sequences placed them in the same cluster as primate genes (Fig. 3A). However, comparison of intron sequences indicated that guinea pig myeloid defensins (*GNCP1A, GNCB1B* and *GNCP2*) and rabbit kidney defensins (*RK1* and *RK2*) may share the same ancestor with rat and mouse enteric genes, while rabbit myeloid defensins (*NP1* and *NP2*) were likely evolved from the ancestral genes for primates (Fig. 3B).

To further understand the driving force for sequence divergence of α-defensins during evolution, we tested whether positive Darwinian selection has occurred by estimating the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N) substitution for different regions of mammalian α-defensins using the method of Nei and Gojobori (19). We computed these values for 11 pairs of closely related genes, which were phylogenetically and thus statistically independent of each other (24). Consistent with earlier results (10, 12), mean d_N was significantly greater than mean d_S in the mature peptide, but there was no significant difference between mean d_S and mean d_N in the signal peptide or prosegment (Table 2). These results support the hypothesis that natural

71

selection has acted to diversify the functionally active mature defensin region but not other portions of the molecules (10, 12).

Genomic organization of mammalian α-defensin gene clusters

Searching through HTGS and WGS sequences in the GenBank led to identification of a number of genomic sequences containing α-defensin genes in the human, chimpanzee, rat and mouse genome (data not shown). To determine the relative position and orientation of each defensin on the chromosome, individual defensins were searched against the assembled human, chimpanzee, rat and mouse genomes by using the BLAT program (14) through the UCSC Genome Browser. In the case of human, all 11 known and newly identified α- and θ-defensins genes were found to form a continuous cluster expanding 132 kb on the chromosome 8p23 (Fig. 5A). All human defensins are transcribed from the same direction from centromere to telomere with the θ-defensin pseudogene (*DEFT1P*) residing in the center. The chimpanzee genome encodes an α-defensin cluster of 10 distinct genes, which expanded 117 kb in the proximal region of chromosome 7 (Fig. 5A). The α-defensin clusters in the chimpanzee and human are nearly in perfect synteny. The only difference is the presence of two copies of *DEFA1* and *DEFT1P* in the human rather than a single copy in the chimpanzee. Such syntenic α-defensin loci are also present in rats and mice. In the case of rat, all 14 α-defensin genes are clustered within a 311-kb distance on the chromosome 16q12.4-q12.5 (Fig. 5B), whereas six mapped mouse α-defensin genes are located continuously on chromosome 8A1.3 (Fig. 5C).

It is noted that all α-defensin genes in both rodents and primates reside within a β-defensin cluster (data not shown) and are adjacent to β-defensin 1 (*Defb1*), which is evolutionarily conserved across the mammalian species (Fig. 5). Therefore, the syntenic location and physical proximity of α-defensins in these species provided additional evidence to support the conclusion that α-defensins have arisen from common ancestor(s) by gene duplication followed by diversification.

Comparison of the open reading frames with genomic sequences of all known and newly identified α-defensins (including α-defensin -related sequences) revealed a highly conserved gene structure, which is composed of two exons separated by a 500-700 bp intron, except for *DEFA5/PTAD5, DEFA6/PTAD6, RK1, RK2, PAAD1* and *PAAD2*, whose introns vary from 848-1646 bp (data available upon request). Consistent with earlier findings, signal sequence and most of the prosegment of α-defensins are encoded in the first exon, whereas the second exon primarily encodes the mature α-defensin sequences. However, based on α-defensin genes studied thus far, it appears that genes of myeloid origin consist of an additional exon encoding 5'-untranslated region, whereas enteric genes are composed of only two exons (21). The lack of full-length cDNA sequences for these newly identified novel sequences prevented us from predicting the genomic locations of 5'- and 3'-untranslated regions and thus possible existence of additional introns.

DISCUSSION

Sequence diversity of mammalian α-defensins

Following systemic, genome-wide screenings, the complete repertoires of the α-defensin gene family, including a number of novel genes, have been identified in the human, chimpanzee, rat, and mouse (Fig. 1). Although it is highly unlikely, we could not rule out the possibility that additional α-defensin genes with distant homology might be uncovered in these species by different computational methods such as the use of Hidden Markov models (18, 30, 31). Alignment of all known α-defensin peptide sequences revealed the conservation of signal and negatively charged prosegment sequences as well as the characteristic six-cysteine-containing α-defensin motif, with exception of pseudogenes and α-defensin-related sequences (Fig. 1). Given the significance of six cysteines in maintaining the spatial structure and biological functions, most pseudogenes have at least one mutation at the conserved cysteine position (Fig. 1), consistent with their dysfunctional nature.

Among all newly identified genes in the primates, chimpanzee *PTAD7* is unique in that it represents a distant, putative functional gene. Unlike a typical α-defensin that is positively charged, the putative mature sequence of PTAD7 contains two arginines (R) and two glutamic acids (E), and therefore is predicted not to have a net positive charge under physiological conditions. However, its human ortholog, *DEFA7P*, appears to be a pseudogene because of an absence of the start codon and hence lack of transcription in a wide range of tissues. Interestingly, *PAAD3*, a *DEFA7P* ortholog in olive baboon, also

appears to be a pseudogene with no start codon at the canonical position (see supplementary Fig. 1). The unique sequence and evolutionary dynamics of the *PTAD7* gene lineage in different primary species implied that PTAD7 might function differently in the chimpanzee from other α-defensins. Consistent with its novelty, PTAD7 is the only known putative functional α-defensin with a nonsynonymous mutation to serine at the fourth canonical cysteine position (Fig 1A).

As opposed to many β-defensins, most α-defensins contain no or only one amino acid following the last two cysteines (Fig. 1). However, Defa8, Defa9 and Defa11 in the rat share a unique feature of unusually long C-terminal tail (9-14 amino acids) rich in charged and polar uncharged amino acids. Such C-terminal tails, which are predicted to be exposed to the surface to influence the overall charge and amphipathicity, might confer a different antimicrobial spectrum and/or efficacy on those peptides, as the net charge and amphipathicity are strongly correlated with antimicrobial activity of defensins (25, 28, 32). Further studies are warranted to assess the importance of such long C-terminal tails in antimicrobial activity as well as in other biological functions.

Evolution of mammalian α-defensins

All α-defensins have been found only in certain mammals, but not in any other vertebrate species that we examined, suggesting that α-defensin genes have appeared following mammalian divergence. The presence of α-defensin loci in syntenic chromosomal regions of different mammalian species (Fig. 5) is indicative of a common ancestor. Phylogenetic analyses of both full-length peptide and intron sequences of α-

defensins revealed two distinct clustering (Fig. 3), implying that they may have independently evolved from two separate ancestral genes. One ancestral gene has undergone significant duplication and diversification giving rise to enteric-specific α-defensins in the rat and mouse, whereas the other ancestral gene has evolved to the genes in the primates as well as in rat myeloid cells.

Phylogenetic analysis of primate α-defensins revealed an interesting evolutionary pattern and pointed out possible presence of subgroups of α-defensins that are specific in the hominid lineage. Despite extensive searches for the repertoire of α-defensins in both leukocytes (34) and small intestines (33), neither *DEFA4- nor DEFA6*-like genes have been found in rhesus macaques (Fig. 1A and Fig. 3). Such genes are also absent in nearly 1000 olive baboon genomic sequences deposited in the GenBank as of September 1, 2004. Therefore, it is likely that *DEFA4/PTAD4* and *DEFA6/PTAD6* lineages are hominid-specific genes. On the other hand, albeit of a single copy of *DEFA5/PTAD5* gene encoded in human and chimpanzee genomes, multiple members of the *DEFA5*-like genes are present in both rhesus macaque (*RED1-6, RMAD4/5,* and *RMAD6/7*) and olive boboon (*PAAD1* and *PAAD2*) (Fig. 3A). Apparently, the ancestral gene for *DEFA5* has undergone a different evolutionary pattern: It duplicated and expanded in cercopithecids, but remained a single copy in hominids. In addition, the *DEFA5* ancestral gene has diverged in rhesus macaques to acquire a new ability for certain offspring genes to express in myeloid cells (e.g., *RMAD4/5,* and *RMAD6/7*), in addition to intestinal Paneth cells (e.g., *RED1-RED6*) as in humans.

With regard to the evolution of mammalian α-defensins, Bevins *et al.* proposed a model, which was based on dot matrix sequence comparisons of introns of a few mouse and human genes, that α-defensins were derived from two ancestral genes for human enteric *DEFA5* and *DEFA6*, and that a subsequent homologous unequal meiotic crossover of *DEFA5* and *DEFA6* generated a hybrid gene that was further evolved to present day myeloid α-defensins (2). Although such a model might be true in primate species (2), our results clearly argued against its generalization to non-mammalian species, because the mouse genome apparently encodes neither *DEFA5/DEFA6*-like genes nor myeloid α-defensins. Therefore, mouse α-defensin genes were apparently not derived from homologous crossover of the primordial genes for *DEFA5* or *DEFA6*. Furthermore, because of presence of multiple *DEFA5* genes expressed in both small intestine and leukocytes in cercopithecids, but an absence of *DEFA6* lineages, the validity of Bevins' evolutionary model in primate species remains to be examined with availability of additional primate genomes.

As for the origin of mammalian α-defensins, they were most likely evolved from β-defensins after mammals diverged from other vertebrates, primarily because of the physical proximity on the chromosome as well as a similarity in spatial structure and biological activities (6, 39). As suggested by Nguyen *et al.* (20), appearance and further diversification of α- and θ-defensins from already a large β-defensin gene family is probably because of the need for antiviral defense in certain mammalian and/or primate species, as these newly evolved α- and θ-defensins have acquired novel lectin-like activity and are capable of inhibiting the entry of viruses (such as HIV) to host cells (36,

37). It is possible that divergence of α-defensin-related sequences from canonical α-defensins in rodents could lead to some novel activities beyond antimicrobial. Further functional characterization of these α-defensins will shed light on the significance of their diversification during evolution and facilitate their development as a new class of antimicrobial agents.

Our earlier comparative analysis of β-defensins in the chicken, rodents, and human revealed that most gene lineages are conserved across mammalian species and thus were evolved before the divergence of mammals from each other (39). In contrast, α-defensins tend to form species-specific clusters particularly in non-primate species (Fig. 3). Distinct subsets of α-defensin genes also exist with most subsets containing multiple members in each species. It is even true with the rat and mouse (Fig. 3), which only diverged 17-25 million years ago. It is rather unique that most α-defensin genes duplicated very recently in the evolutionary time and have undergone significant, but independent expansion in different species. Clearly, the β-defensin gene family is evolutionarily older than α-defensins. As compared to the former, α-defensin genes duplicated and expanded much more rapidly. Calculation of the rates of synonymous vs. nonsynonymous substitutions of 11 pairs of representative α-defensins revealed that positive Darwinian selection appears to have acted to diversify α-defensins, particularly in the mature peptide region (Table 2).

Gene duplication followed by positive selection indeed has been observed in several gene families involved in immune responses (11). Divergence of these immune

genes often leads to either an additional layer of functional redundancy or acquisition of functional novelties, both of which conceivably help the hosts cope more effectively with a broad range of pathogens. In fact, α-defensins have been shown to exhibit selectivity against varied microorganisms, and a modest difference in the primary sequence could have a significant impact on the antibacterial spectrum and/or potency (22). Therefore, as different mammals live in quite different ecological niches, the production of species-specific α-defensins would presumably allow them to better respond to the specific microbial challenges that they face. The presence of α-defensin-related sequences in rodents and functional divergence of θ-defensins from α-defensins in certain primate species provide additional evidence supporting this notion.

ACKNOWLEDGEMENTS

GRANTS

REFERENCES

1. **Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ.** Basic local alignment search tool. *J Mol Biol* 215: 403-410, 1990.

2. **Bevins CL, Jones DE, Dutra A, Schaffzin J, and Muenke M.** Human enteric defensin genes: chromosomal map position and a model for possible evolutionary relationships. *Genomics* 31: 95-106, 1996.

3. **Birney E, Clamp M, and Durbin R.** GeneWise and Genomewise. *Genome Res* 14: 988-995, 2004.

4. **Burge C and Karlin S.** Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268: 78-94, 1997.

5. **Condon MR, Viera A, D'Alessio M, and Diamond G.** Induction of a rat enteric defensin gene by hemorrhagic shock. *Infect Immun* 67: 4787-4793, 1999.

6. **Ganz T.** Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 3: 710-720, 2003.

7. **Ghosh D, Porter E, Shen B, Lee SK, Wilk D, Drazba J, Yadav SP, Crabb JW, Ganz T, and Bevins CL.** Paneth cell trypsin is the processing enzyme for human defensin-5. *Nat Immunol* 3: 583-590, 2002.

8. **Hancock RE and Lehrer R.** Cationic peptides: a new source of antibiotics. *Trends Biotechnol* 16: 82-88, 1998.

9. **Hornef MW, Putsep K, Karlsson J, Refai E, and Andersson M.** Increased diversity of intestinal antimicrobial peptides by covalent dimer formation. *Nat Immunol* 5: 836-843, 2004.

10. **Hughes AL.** Evolutionary diversification of the mammalian defensins. *Cell Mol Life Sci* 56: 94-103, 1999.

11. **Hughes AL.** Natural selection and the diversification of vertebrate immune effectors. *Immunol Rev* 190: 161-168, 2002.

12. **Hughes AL and Yeager M.** Coordinated amino acid changes in the evolution of mammalian defensins. *J Mol Evol* 44: 675-682, 1997.

13. **Huttner KM and Ouellette AJ.** A family of defensin-like genes codes for diverse cysteine-rich peptides in mouse Paneth cells. *Genomics* 24: 99-109, 1994.

14. **Kent WJ.** BLAT--the BLAST-like alignment tool. *Genome Res* 12: 656-664, 2002.

15. **Kumar S, Tamura K, Jakobsen IB, and Nei M.** MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244-1245, 2001.

16. **Lehrer RI and Ganz T.** Defensins of vertebrate animals. *Curr Opin Immunol* 14: 96-102, 2002.

17. **Linzmeier R, Ho CH, Hoang BV, and Ganz T.** A 450-kb contig of defensin genes on human chromosome 8p23. *Gene* 233: 205-211, 1999.

18. **Lynn DJ, Higgs R, Gaines S, Tierney J, James T, Lloyd AT, Fares MA, Mulcahy G, and O'Farrelly C.** Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken. *Immunogenetics* 56: 170-177, 2004.

19. **Nei M and Gojobori T.** Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418-426, 1986.

20. **Nguyen TX, Cole AM, and Lehrer RI.** Evolution of primate theta-defensins: a serpentine path to a sweet tooth. *Peptides* 24: 1647-1654, 2003.

21. **Ouellette AJ.** Defensin-mediated innate immunity in the small intestine. *Best Pract Res Clin Gastroenterol* 18: 405-419, 2004.

22. **Ouellette AJ and Selsted ME.** Paneth cell defensins: endogenous peptide components of intestinal host defense. *FASEB J* 10: 1280-1289, 1996.

23. **Patthy L.** Genome evolution and the evolution of exon-shuffling--a review. *Gene* 238: 103-114, 1999.

24. **Piontkivska H and Hughes AL.** Between-host evolution of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1: an approach based on phylogenetically independent comparisons. *J Virol* 78: 11758-11765, 2004.

25. **Raj PA, Antonyraj KJ, and Karunakaran T.** Large-scale synthesis and functional elements for the antimicrobial activity of defensins. *Biochem J* 347 Pt 3: 633-641, 2000.

26. **Saitou N and Nei M.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425, 1987.

27. **Salzman NH, Ghosh D, Huttner KM, Paterson Y, and Bevins CL.** Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 422: 522-526, 2003.

28. **Schibli DJ, Hunter HN, Aseyev V, Starner TD, Wiencek JM, McCray PB, Jr., Tack BF, and Vogel HJ.** The solution structures of the human beta-defensins lead to a better understanding of the potent bactericidal activity of HBD3 against Staphylococcus aureus. *J Biol Chem* 277: 8279-8289, 2002.

29. **Schutte BC and McCray PB, Jr.** Beta-defensins in lung host defense. *Annu Rev Physiol* 64: 709-748, 2002.

30. **Schutte BC, Mitros JP, Bartlett JA, Walters JD, Jia HP, Welsh MJ, Casavant TL, and McCray PB, Jr.** Discovery of five conserved beta -defensin gene clusters using a computational search strategy. *Proc Natl Acad Sci USA* 99: 2129-2133, 2002.

31. **Semple CA, Rolfe M, and Dorin JR.** Duplication and selection in the evolution of primate beta-defensin genes. *Genome Biol* 4: R31, 2003.

32. **Tanabe H, Qu X, Weeks CS, Cummings JE, Kolusheva S, Walsh KB, Jelinek R, Vanderlick TK, Selsted ME, and Ouellette AJ.** Structure-activity determinants in paneth cell alpha-defensins: loss-of-function in mouse cryptdin-4 by charge-reversal at arginine residue positions. *J Biol Chem* 279: 11976-11983, 2004.

33. **Tanabe H, Yuan J, Zaragoza MM, Dandekar S, Henschen-Edman A, Selsted ME, and Ouellette AJ.** Paneth cell alpha-defensins from rhesus macaque small intestine. *Infect Immun* 72: 1470-1478, 2004.

34. **Tang YQ, Yuan J, Miller CJ, and Selsted ME.** Isolation, characterization, cDNA cloning, and antimicrobial properties of two distinct subfamilies of alpha-defensins from rhesus macaque leukocytes. *Infect Immun* 67: 6139-6144, 1999.

35. **Thompson JD, Higgins DG, and Gibson TJ.** 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: 4673-4680, 1994.

36. **Wang W, Cole AM, Hong T, Waring AJ, and Lehrer RI.** Retrocyclin, an antiretroviral theta-defensin, is a lectin. *J Immunol* 170: 4708-4716, 2003.

37. **Wang W, Owen SM, Rudolph DL, Cole AM, Hong T, Waring AJ, Lal RB, and Lehrer RI.** Activity of alpha- and theta-defensins against primary isolates of HIV-1. *J Immunol* 173: 515-520, 2004.

38. **Wilson CL, Ouellette AJ, Satchell DP, Ayabe T, Lopez-Boado YS, Stratman JL, Hultgren SJ, Matrisian LM, and Parks WC.** Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. *Science* 286: 113-117, 1999.

39. **Xiao Y, Hughes AL, Ando J, Matsuda Y, Cheng JF, Skinner-Noble D, and Zhang G.** A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. *BMC Genomics* 5: 56, 2004.

40. **Yang D, Biragyn A, Hoover DM, Lubkowski J, and Oppenheim JJ.** Multiple Roles of Antimicrobial Defensins, Cathelicidins, and Eosinophil-Derived Neurotoxin in Host Defense. *Annu Rev Immunol* 22: 181-215, 2004.

41. **Yeh RF, Lim LP, and Burge CB.** Computational inference of homologous gene structures in the human genome. *Genome Res* 11: 803-816, 2001.

42. **Yount NY, Wang MS, Yuan J, Banaiee N, Ouellette AJ, and Selsted ME.** Rat neutrophil defensins. Precursor structures and expression during neutrophilic myelopoiesis. *J Immunol* 155: 4476-4484, 1995.

43. **Zasloff M.** Antimicrobial peptides of multicellular organisms. *Nature* 415: 389-395, 2002.

44. **Zhang G, Hiraiwa H, Yasue H, Wu H, Ross CR, Troyer D, and Blecha F.** Cloning and characterization of the gene for a new epithelial beta-defensin. Genomic

structure, chromosomal localization, and evidence for its constitutive expression. *J Biol Chem* 274: 24031-24037, 1999.

45. **Zhang G, Wu H, Shi J, Ganz T, Ross CR, and Blecha F.** Molecular cloning and tissue expression of porcine beta-defensin-1. *FEBS Lett* 424: 37-40, 1998.

FIGURE LEGENDS

Fig. 1. Amino acid sequence alignment of novel α-defensins and α-defensin-related sequences in the primates (*A*), rat (*B*) and mouse (*C* and *D*). Panel *C* shows novel cryptdins identified in mice, whereas Panel *D* demonstrates new cryptdin-related sequences. The dashes are created to optimize the alignment, and canonical cysteines in the mature defensin regions are shaded. The letter 'X' in bold denotes the presence of premature stop codon in each pseudogene. Note that computational predictions of all novel rat α-defensin peptide sequences, except for the pseudogenes, have been confirmed by cloning and sequencing of their respective RT-PCR products amplified from appropriate tissues. Abbreviations: *DEFA/Defa*, α-defensin; *DEFT1P*, human θ-defensin 1 pseudogene; *PTAD*, chimpanzee (*Pan troglodytes*) α-defensin; *PTTD*, *Pan troglodytes* θ-defensin; *RMAD*, rhesus macaque myeloid α-defensin; *RED*, rhesus macaque enteric α-defensin; *RTD*, rhesus macaque θ-defensin; *PAAD*, *Papio anubis* α-defensin; *RatNP*, Rat neutrophil α-defensin; Defa-*rs1*, rat α-defensin-related sequence; *RD5*, rat α-defensin 5; *Defcr*, mouse α-defensin-related cryptdin; *CRS1C*, mouse cryptdin-related sequence 1C; *CRS4C*, mouse cryptdin-related sequence 4C; *P/ps*, pseudogene.

Fig. 2. Tissue expression patterns of novel rat β-defensins by RT-PCR. See Materials and Methods for details. The number of PCR cycles was optimized for each gene, and the specificity of each PCR product was confirmed by sequencing. The house-keeping gene, *GAPDH*, was used for normalization of the template input.

Fig. 3. Phylogenetic tree of mammalian α-defensins. The tree was constructed by calculating the proportion difference (*p*-distance) of aligned amino acid sites of full-length peptide sequences (*A*) or *p*-distance of aligned nucleotide sites of the entire intron sequences (*B*) according to the neighbor-joining method. The reliability of each branch was tested using 1000 bootstrap replications. Numbers on the branches indicate the percentage of 1000 bootstrap samples supporting the branch. Only branches supported by a bootstrap value of at least 40% are shown. Abbreviations: *NP/MCP*, rabbit neutrophil α-defensin/macrophage cationic peptide; *RK*, rabbit kidney α-defensin; *GNCP*, guinea pig neutrophil cationic peptide. See Fig. 1 legend for additional abbreviations.

Fig. 4. Alignment of the second exon sequences of rodent α-defensin-related sequences with three rat enteric α-defensins. For clarity, the last ~75 nucleotides of introns were also aligned and the exon/intron conjunction sites are indicated with a vertical line. Conserved regions are shaded and the stop codon of each gene boxed. See Fig. 1 legend for abbreviations.

Fig. 5. Genomic organizations of α-defensin clusters in the primates (*A*), rat (*B*), and mouse (*C*). The position of each gene is represented by a solid vertical bar and the width of each bar is proportional to the size of each gene. The direction of transcription is indicated by the triangle. The genes with solid triangles are transcribed in the direction opposite to the ones with open triangles. The gene clusters were derived from the assembled genomes of human (NCBI Build 35), chimpanzee (NCBI Build 1 Version 1), rat (BCM Version 3.1) and mouse (NCBI Build 33). Note that the α-defensin cluster is

located in a syntenic region adjacent to β-defensin 1 (*DEFB1*) in each species. In mice, a number of additional α-defensin genes are encoded in two other large but unmapped clusters, and are not indicated here for the sake of clarity. Abbreviations: *Tel*, telomere; *Cen*, centromere. See Fig. 1 legend for additional abbreviations.

Table 1. Primer sequences used for RT-PCR analysis of rat α-defensins

Gene	Primer Sequence		Product Size (bp)		No. of
	Forward	Reverse	cDNA	Genomic	Cycles
RatNP1/2	CCATGAGGACGCTCACTCTGCT	CAAATGGATTCTTCTTGGTCGG	316	833	35
RatNP3	CAAATGGATTCTTCTTGGTCGG	CTTCTTCTTCTTCTTTGTGCTCA	286	793	35
RD5	GAAGACACTTGTCCTCCTTTCTG	TGTTGCAGATCCCCATAATGCCT	245	861	35
Defa6	GAAGACACTTGTCCTCCTTTCTG	TATTCTTTAGGTCCGCAGTTTGC	228	815	35
Defa8	GAAGACACTTGTCCTCCTTTCTG	AATAGACCTGGACGACAGGACC	258	840	35
Defa9	GAAGACACTTGTCCTCCTTTCTG	TCATCCTTGGGGGCCATGTTTCT	305	890	35
Defa10	CAAATGGATTCTTCTTGGTCGG	ATCCGGCCATTGATTCGACACA	263	765	40
Defa11	CAAATGGATTCTTCTTGGTCGG	TTGTTCATAGTGTTCAGTGGTTG	293	810	40
Defa12-ps	CAAATGGATTCTTCTTGGTCGG	AGGGGGTAGGTGACGCCACCAG	269	769	40
Defa-rs1	GAAGACACTTGTCCTCCTTTCTG	ATGTGGACCTTGATAGCCGATG	261	834	35
GAPDH	GTGAAGGTCGGAGTGAACG	GAGATGATGACCCTCTTGGC	356	>390	25

Table 2. Demonstration of positive selection in 11 phylogenetically independent pairs of mammalian α-defensin genes [a].

Region	d_S	d_N
Signal peptide	0.031 ± 0.014	0.023 ± 0.008
Prosegment	0.066 ± 0.026	0.066 ± 0.022
Mature defensin	0.122 ± 0.041	0.182 ± 0.047[b]

[a] Mean (± S.E.) numbers of synonymous substitutions per synonymous site (d_S) and of nonsynonymous substitutions per nonsynonymous site (d_N) The following sister pairs of closely related genes were compared: from Primates, *DEFA5* vs. *RMAD4/5*, *DEFA1* vs. *DEFA3*, *RMAD1/2* vs. *RMAD8*; from rabbit *NP1/MCP1* vs. *NP2/MCP2*, *NP4* vs. *NP5*; from rodents, *GNCP1A* vs. *GNCP2*, *defcr5* vs. *defcr6*, *defcr20* vs. *defcr21*, *RatNP3* vs. *RatNP4*, rat *defa6* vs. rat *defa8*, rat *defa7* vs. rat *defa11*.

[b] Significant difference between d_S and d_N at P<0.01 level (paired t-test, two-tailed).

A

	Signal Peptide	Prosegment	Mature Defensin

```
                    10        20        30        40        50        60        70        80        90       100       110
DEFA1/2    MRTLAILAA-ILLVALQAQAEPLQARADE------VAAAPEQIAADIPEVVVSLAWDESSLAPKH---PGSR---KNMACYCRIPACIAGERRYGTCIYQG--RLWAFCC--------------
DEFA3      MRTLAILAA-ILLVALQAQAEPLQARADE------VAAAPEQIAADIPEVVVSLAWDESSLAPKH---PGSR---KNMDCYCRIPACIAGERRYGTCIYQG--RLWAFCC--------------
RMAD1/2    MRTLVILAA-ILLVALQAQAEPLQARTDE------ATAAQEQIPTDNPEVVVSLAWDESSLAPKDSV-PGLR---KNMACYCRIPACLAGERRYGTCFYLG--RVWAFCC--------------
RMAD3      MRTLVILAA-ILLVALQAQAEPLQARTDE------ATAAQEQIPTDNPEVVVSLAWDESSLAPKDSV-PGLR---KNMACYCRIPACLAGERRYGTCFYRR--RVWAFCC--------------
PTAD1      MRTLAILAA-ILLVALQAQAEPLQARADE------VAAAPEQIPADNPEVVVSLAWDESSLAPKH---PGSR---KNVACYCRIPACLAGERRYGTCIYQG--RLWAFCC--------------
DEFA8P     -RSLALLAA-ILLVALQARAEPLQAIADE-------ATAQEQPGADDQEVVDSFAWDERAPLQV---SGSR---RGLTCTCRXDSYSPGEHRRGTCSAPG--VRYPYCFS--------------
PTAD8-ps   -RSLALLAA-ILLVALQARAEPLQAMTDE-------ATAQEQPGADDQEVVDSFAWDERAALQV---SGSR---RGLTCTCRXDSCSPGEYRRGTCSAPG--VRYPYCFS--------------
DEFA10P    MRTLALLTATVLFVALQAQAEPLRARADE-------AAAQEXPGADDQEVSVSFTWDESAALPL--SGESG---RGLRCICRRGICHFFQRRFGSCALPG--RLYRICCC--------------
PTAD10-ps  MRTLVLLTATVLFVALQAQAEPLRARADE-------AAAEEQPGADDQEVSVSFTWDESAALPL---SGSG---RGLRCICRXGICHFFQCRFGSCALPG--RLYRICCC--------------
DEFT1P     MRTFALLTAMLLLVALXAQAEPLQARADE-------AAAQEQPGADDQEMAHAFTWHESAALPL---SDSA---RGLRCICGRGICRLLXRRFGSCAFRG--TLHRICCR--------------
PTTD1-ps   MRTFALLTAMLLLVALXAQAEPLQARADE-------AAAQEQPGADDQEMAHAFTWDESAALPL---SDSA---RGLRCIGGRGICGLLQRRFGSCAFRG--TLHRICCR--------------
PTTD2-ps   MTTFALLTAMLLLVALXAQAEPLQARADE-------AAAQEQPGADDQEMAHAFTWDESAALPL---SDSA---RGLRCIGGRGICGLLQRRFGSCAFRG--TLHRICCR--------------
RTD1a      MRTFALLTAMLLLVALHAQAEARQARADE-------AAAQQQPGTDDQGMAHSFTWPENAALPL---SESA---KGLRCICTRGFCRLLXCHLGSCAFRG--XLHRICCR--------------
RTD1b      MRTFALLTAMLLLVALHAQAEARQARADE-------AAAQQQPGADDQGMAHSFTRPENAALPL---SESA---RGLRCICRGVCQLLXRRLGSCAFRG--XLCRICCR--------------

DEFA4      MRIIAILAA-ILLVALQVRAGPLQARGDE-------APGQEQRGPEDQDISISFAWDKSSALQV---SGST---RGMVCSCRLVFCRRTELRVGNCLIGG--VSFTYCCTRVD----------
PTAD4      MRIIAILAA-ILLVALQVRAGPLQARGDE-------APGQEQRGPEDQDISISFAWDKSSALQV---SGST---RGMVCSCRLVFCRRTELRVGNCLIGG--VSFTYCCTRVD----------

DEFA5      MRTIAILAA-ILLVALQAQAESLQERADE-------ATTQKQSGEDNQDLAISFAGNGLSALRT---SGSQ---ARATCYCRTGRCATRESLSGVCEISG--RLYRLCCR-------------
PTAD5      MRTIAILAA-ILLVALQAQAESLQERADE-------ATTQKQSGEDNQDLAISFAGNGLSALRT---SGSQ---ARATCYCRIGHCTILESLSGVCEISG--RLYRLCCR-------------
PAAD1      MRTLAILAA-ILLVALQVQAESLQERADE-------ATTQEQPGEDDQDFAVSFEENGLSTLKA---SGSR---ARIICHCQIGRCSRPESYSRICILHG--QVARLCCRRAS----------
PAAD2      MRTLAILAA-ILLVALQAQAESLQERADE------VAT-QEQPGEDDQDLAVSFEENGLSTLRA---SGSQ---ARRTCYCRTGRCYTPEFHSGKCLFKG--LTYKLCCR-------------
RED1       MRTIAILAA-ILLFALLAQAKSLQETADE-------AATQEQPGEDDQDLAVSFEENGLSTLRD---SGSQ---ARRTCRCIRRCRGLESSFGNCILHG--QFAKLCCR--------------
RED3       MRTIAILAA-ILLFSLQAQAESLQERADE------VAT-QEQPGEDDQDFAVSFEENGLSTLRA---SGSQ---ARHTCYCRNKRCFTPEFHAGKCKVEG--RTYKLCCR-------------
RED5       MRTLTILAA-ILLVALQAQAESLQERADE-------AATQEQPAEDNQDFAVSFEENGLSTLKA---SGSQ---ARMICLCRIGRCSWREAHFGSCTKMG--QFAKICCRRASR---------
RMAD4/5    MRTIAILAA-ILLFALLAQAKSLQETADD-------AATQEQPGEDDQDLAVSFEENGLSTLRA---SGSQ---ARRTCRCRFGRCFRRESYSGSCNING--RIFSLCCR--------------

DEFA6      MRTLTILTA-VLLVALQAKAEPLQAEDDPLQAKAYEADAQEQRGANDQDFAVSFAEDASSSLRA---LGST---RAFTCHCRRS-CYSTEYSYGTCTVMG--INHRFCCL--------------
PTAD6      MRTLTILTA-VLLVALQAKAEPLQAEDEPLQAKAYEADAQEQRGANDQDFAVSFAEDASSSLRA---LGST---RAFTCHCRRS-CYSTEYSYGTCTVMG--INHRFCCL--------------
DEFA9P     MRTLTLLTA-ILLVPLQALSEPLQARADX-------AAAQEQLGADDQDFTISFTGDTSSGFRA---SGPR---TLLRCSCIRDYHLQERNYGTCTSGE--IHYKFCFL--------------
PTAD9-ps   MRTLTLLTA-ILLVPLQALSEPLQARADX-------AAAQEQLGADDQDFAISFTGDTSLGFRA---SGPR---TLLRCSCIRDYHLQERNYGTCTSGE--IHYKFCFL--------------
DEFA11P    MRTLTLLTA-ILLVPLQALSEPLQARADX-------AAAQEQLGADHQDFAISFTGDTSSGFRA---SGSA--TGLVCYCRISGCHFSKHHYGTCTNSGIMEHRGKCVHHLL----------
PTAD11-ps  MRTLSLLAA-ILLVPLQALSEPLQARADX-------AAAQEQLGADHQDFAISFTGHIRSGFRD---SGST--RGLVCYCRISGCHFSKHHYGTCTNSGIMEHRGKCVHHLL----------

DEFA7P     -RTLTLLSA-FLLVALQAWAEPLPARAHE-------MPAQKQPPADDQDVVLYFSGDDSCSLQV---PGST---KGLICHCRVLYCIFGEHLGGTCFILG--ERYPICCY-------------
PTAD7      MRTLTLLSA-FLLVALQAWAEPLQARADE-------MPAQKQPPADDQDVVIYFSGDDSSSLQV---PGST---KGLICHCRVLYCLFGEHLGGTSFIHG--ERYPICCY-------------
PAAD3      -RSLTLLSA-ILLGALQAWTEPLQARADE-------MPAQEQPPANDQDVVIYFSGDDSSSLQV---PGST---KGLSCHCRVLYCLFGEHLGGTSFIHG--ERSPICCY-------------
```
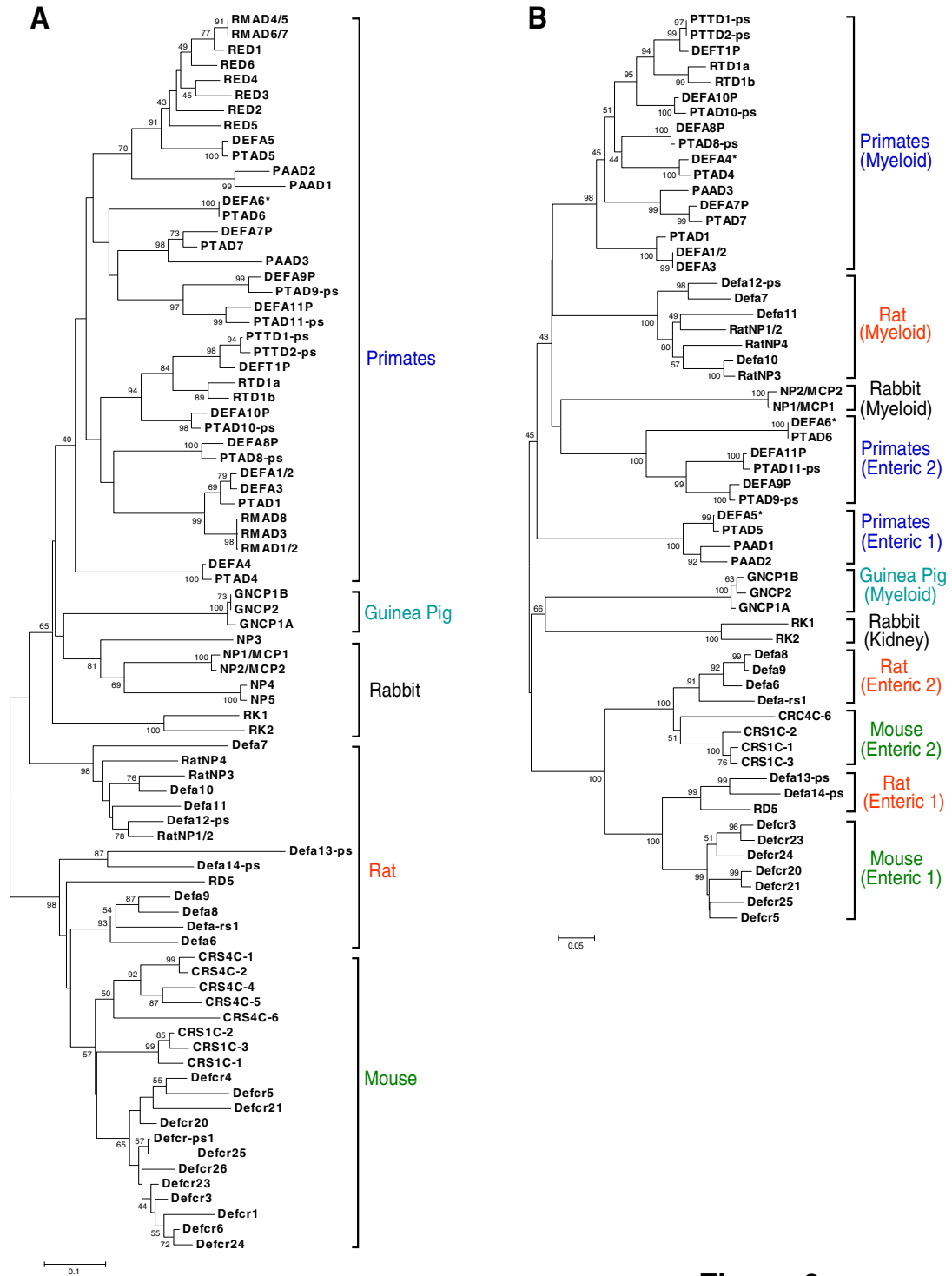
B

```
RatNP1/2   MRTLTLLTA-LLLLALHTQAKSPQGTAEE-------APDQEQLVMEDQDISISFGGDKGTALQD---ADVK---AGVTCYCRRTRCGFRERLSGACGYRG--RIYRLCCR-------------
RatNP3     MRTLTLLTT-LLLLALHTQAESPQGSTKE-------APD------EEQDISVFFGGDKGTALQD---AAVK---AGVTCSCRTSSCRFGERLSGACRLNG--RIYRLCC--------------
RatNP4     MRTLTLLIT-LLLLALHTQAESPQERAKA-------APDQD-MVMEDQDIFISFGGYKGTVLQD---AVVK---AGQACYCRIGACVSGERLTGACGLNG--RIYRLCCR-------------
Defa7      MRTLTLLTT-LLLLALHTQAESPQGSRSTEE-----ALDQEQLVMEDQDISISFGGDKGTALQD---ADVK---SGVTCYCRLLSCQFGERLAGSCRSGG--VTYPLCCH-------------
Defa10     MRTLTLLTA-LLLLALHTQAESPQGSPKE-------APDQEQLDMEDQDISVFFGGDKGTALQD---AA-----GSTGCRIGTCVSGEWLSWVCRING--RIYRLCCR--------------
Defa11     MGTLTLLTT-LLLLALHTQAESPQGCT-E---------DQKLLIMEDQDISISFTGDKGTALQD---ADVK---SGVIGNCPRGCSFKERVYGECGYNG--HILRLCCQPLNTMNKKSPSVI
Defa12-ps  MRTLTLLTA-LLLLALHTQAESPQGRAEE-------APDLEXLVMKDQDISISFGGDKGTALQN---ADVK---SGVTCYCRIS-CQFGERLSGSCRSGG--VTYPLCCR-------------

RD5        MKTLVLLSA-LVLLALQVQAEPTPKTDEG-------TKTDEQPGKEDQVVSVSIEGQGDPAFQD---AVLR----DLKQFCRAKSCNWGEGIMGICNKRY--GSLILCCR-------------
Defa6      MKTLVLLSA-LVLVAYQVQADPIQGAEEE-------TKTEEQPSDEDQDVSVSFEGPEASALQD---FEI---GRPVRRCRCRAN-CGPKEYATAFCAQGP-FKQFKFCCTA-----------
Defa8      MKTLVLLSA-LVLLAFQVQADHIQEAEEE-------TKTEEQPADEDQDMSVSFEGPEASALQN---LEI--GWPLKQCHCRKF-CRPYEKAEGSCRPGL-FIKRKICCIQQWTPGRT-----
Defa9      MKTLVLLSA-LVLLAFQIQADPIQEAEEE-------TKTEEQPADEDQDVSVSFEGPEPSALQN---LEI--RWPWKRCHCRSF-CRPYENATSFCAQGL-FKQHKFCCLETWPPRMK-----
Defa13-ps  MKTLLFLSA-LVLLAFQVQADPMQGADEE-------NGVAA---RERSQICLSPLEAQKALLFKMQILFVSSIARNQKCYYRARSCRNREGIIRTCNKRH--ESLILCCR-------------
Defa14-ps  MKTLLFLSA-LVLLAFQVQADPIQEADEE-------TKTEKKPGKEDPDVAVSFGGPQGSAFQDADNLFVSLVSRNQKCYYTARSYRNREGIMGTXTKRH--ESLILCCR-------------
Defa-rs1   MKTLILLSA-LVLLALQVQADPIQEAEEE-------TKTEEQPADEDQDVSVSFEGPEASAVQD---LRVR---RTLQCSCRR-VCR----NTCSCIRLS--RSTYAS---------------
```

C

```
Defcr1     MKKLVLLFA-LVLLGFQVQADSIQNTDEE-------TKTEEQPGEEDQAVSVSFGDPEGTSLQE---ESLR----DLVCYCRSRGCKGRERMNGTCRKGH--LLYTLCCR-------------
Defcr3     MKTLVLLSA-LVLLAFQVQADPIQNTDEE-------TKTEEQPGEDDQAVSVSFGDPEGSSLQE---ESLR----DLVCYCKRGCKRRERMNGTCRKGH--LMYTLCCR-------------
Defcr23    MKTLVLLSA-LILLAFQVQADPIQNTDEE-------TKTEEQPGEDDQAVSVSFGDPEGSSLQE---ESLR----DLVCYCRTRGCKRRERMNGTCRKGH--LIYTLCCR-------------
Defcr24    MKTLILLSA-LVVLAFQVQADPIQNTDEE-------TKTKEQPGEEDQAVSVSFGDPEGASLQE---ESLR----DLVCYCRARGCKGRERMNGTCSKGH--LLYMLCCR-------------
Defcr25    MKTLVLLSA-LALLAFQVQADPIQNRDEE-------SKIDEQPGKEDQAVSVSFGDPEGSSLQE-ECDALR----DLICYCRTRGCKRRERLNGTCRKGH-LMYMLWCC-------------
Defcr26    MKTLVLLSA-LFLLAFQVQADPIQNTDEE-------TNTEVQPQEDQAVSVSFGNPEGSSLQE---ESLR----DLGGCYCKRGCTRRERINGTCRKGH--LMYTLCCL-------------
Defcr-ps1  MKTLVLLSA-LALLAFQVQADPIQNTDEE-------TNTEEQPGKEDQAVSVSFGDPEGSSLQE---ESLR----DLIXYCRTRGCKRRERLNGTYRKGH--LMYMLCCC-------------

Defcr4     MKTLVLLSA-LVLLAFQVQADPIQNTDEE-------TKTEEQPGEEDQAVSISFGGQEGSALHE---KSLR---GLLCYCRKGHCKRGERVRGTCG----IRFLYCCPRR-----------
Defcr20    MKTLVLLSA-LVLLAFQVQADPIQNTDEE-------TNTEEQAGEEDQAVSVSFGGQEGSALHE---KSSR---DLICYCKGGCNRGEQVYGTCS-----GRLLFCCRRRHRH---------
Defcr21    MKTLVLLSA-LILLAYQVQTDPIQNTDEE-------TNTEEQPGEDDQAVSVSFGGQEGSALHE---KLSR---DLICLCRNRRCNRGELFYGTCA-----GPFLRCCRRRR----------
Defcr22    MKKLVLLSA-LVLLAYQVQTDPIQNTDEE-------TNTEEQPGEEDQAVSVSFGGQEGSALHE---KLSR---DLICRKRRCNRGELFYGTCA-----GPFLRCCRRRR----------
```

D

```
CRS1C-1    MKTLVLLSA-LALLALQVQADPIQNTDEE-------TKTQEQPGEEDQAVSVSFGGTEGSALQDV--AQRRFPWCRKCRVCQKCEVCQKCPVCPTCPQCPKQPLCKERQNKTAITTQAPNTHHKGC
CRS1C-2    MKTLVLLSA-LALLAFQVQADPIQNTDEE-------TKTEEQPEEEDQAVSVSFGGTEGSALQDV--AQRRFLWCRKCPVCQKCQVCQKCPVCPTCPQCPKQPLCEERQNKTAITTQAPNTQHKGC
CRS1C-3    MKTLVLLSA-LALLAFQVQADPIKNTDEE-------TKTGEQPGEEDQAVSVSFGGTEGSALQYV--AQRRFPWCRKCPVCQKCQVCQKCPVCPTCPQCPKLPLCKERQNKSAITTQAPNTQHKGC

CRS4C-1    MKKLVLLFA-LVLLAFQVQADSIQNTDEE-------TKTEEQPGEKDQAVSVSFGDPQGSALQD---AALG--WGRRCPQCPRCPSCPSCPRCPRCPRCK----CNPK-------------
CRS4C-2    MKKLVLLFA-LVLLAFQVQADSIQNTDEE-------TKTEEQQGEEDQAVSVSFGDPQGSGLQD---AALG--WGRRCPRCPPCRCSWCPRCPTCPRCN----CNPK-------------
CRS4C-4    MKKLVLLSA-FVLLAFQVQADSIQNTDEE-------TKTEEQPGEENQAMSVSFGDPEGSALQD---AAVG--MARPCPPCSCPSCPWCPMCPRCPSCK----CNPK-------------
CRS4C-5    MKKLVLLSA-FVLLAFQVQADSIQNTDEE-------IKTEEQPGEENQAVSISFGDPEGYALQD--AAAIR--RARRCPPCPSCLSCPWCPRCLRCPMCK----CNPK-------------
CRS4C-6    MKTLVLLSA-LVLLAFYVQADSTQETDEE-------TKTDDQPGEEDQGVSVSFEDPERYVLQV--SGLGKPPQCPKCPVCSKCPQCPQCPQCPGCPRCN----CMTK-------------
```

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Supplementary Fig. 1. Sequence alignment of the exon encoding the signal peptide and prosegments of three orthologous genes in the human (*DEFA7P*), chimpanzee (*PTAD7*), and olive baboon (*PAAD3*). The encoded peptide sequence for each exon is also indicated. Note that the start codon in *PTAD7* is mutated in both *DEFA7P* and *PAAD3* as highlighted in orange color, and that a string of sequences around the start codon (underlined) in *PAAD3* are completely different from those in *DEFA7P* and *PTAD7*. Abbreviations: *DEFA7P*, human α-defensin 7 pseudogene; *PTAD7*, *Pan troglodytes* α-defensin 7; *PAAD3*, *Papio anubis* α-defensin 3.

Supplementary Fig. 2. Amino acid sequence alignment of a total of 82 known and newly identified α-defensins and α-defensin-related sequences in the primates and glires. The dashes are created to optimize the alignment, and canonical cysteines in the mature defensin regions are shaded. The premature stop codon in each pseudogene is omitted. Note that computational predictions of all novel rat α-defensin peptide sequences, except for the pseudogenes, have been confirmed by cloning and sequencing of their respective RT-PCR products amplified from appropriate tissues. Abbreviations: *DEFA/Defa*, α-defensin; *DEFT1P*, human θ-defensin 1 pseudogene; *PTAD*, chimpanzee (*Pan troglodytes*) α-defensin; *PTTD*, *Pan troglodytes* θ-defensin; *RMAD,* rhesus macaque myeloid α-defensin; *RED*, rhesus macaque enteric α-defensin; *RTD*, rhesus macaque θ-defensin; *PAAD*, *Papio anubis* α-defensin; *RatNP*, Rat neutrophil α-defensin; Defa-*rs1*, rat α-defensin-related sequence; *RD5*, rat α-defensin 5; *Defcr*, mouse α-defensin-related

cryptdin; *CRS1C*, mouse cryptdin-related sequence 1C; *CRS4C*, mouse cryptdin-related

sequence 4C; *NP/MCP*, rabbit neutrophil α-defensin/macrophage cationic peptide; *RK*,

rabbit kidney α-defensin; *GNCP*, guinea pig neutrophil cationic peptide; *P/ps*,

pseudogene.

```
PTAD7    GTGACTACAGTTATGAGGACCCTCACCCTCCTCTCTGCCTTTCTCCTGGTGGCCCTTCAGGCCTGGGCAGAGCCGCTCCAGGCAAGAGCT    90
                      M   R   T   L   T   L   L   S   A   F   L   L   V   A   L   Q   A   W   A   E   P   L   Q   A   R   A

DEFA7P   GTGACTACAGTTAGGAGGACCCTCACCCTCCTCTCTGCCTTTCTCCTGGTGGCCCTTCAGGCCTGGGCAGAGCCGCTCCCGGCAAGAGCT    90
                      R   R   T   L   T   L   L   S   A   F   L   L   V   A   L   Q   A   W   A   E   P   L   P   A   R   A

PAAD3    GAGGCTCCTCTCTGCCGTTCCCTCACCCTCCTCTCTGCCATTCTCCTGGGGGCCCTCCAGGCCTGGACAGAGCCGCTCCAGGCAAGAGCT    90
                      C   R   S   L   T   L   L   S   A   I   L   L   G   A   L   Q   A   W   T   E   P   L   Q   A   R   A


PTAD7    GATGAGATGCCAGCCCAGAAGCAGCCTCCAGCAGACGACCAGGATGTGGTCATTTACTTTTCAGGAGATGACAGCTCCTCTCTTCAGGTTCAG   183
          D   E   M   P   A   Q   K   Q   P   P   A   D   D   Q   D   V   V   I   Y   F   S   G   D   D   S   S   S   L   Q   V   P

DEFA7P   CATGAGATGCCAGCCCAGAAGCAGCCTCCAGCAGATGACCAGGATGTGGTCCTTTACTTTTCAGGAGATGACAGCTGCTCTCTTCAGGTTCAG   183
          H   E   M   P   A   Q   K   Q   P   P   A   D   D   Q   D   V   V   L   Y   F   S   G   D   D   S   C   S   L   Q   V   P

PAAD3    GATGAGATGCCAGCCCAGGAGCAGCCTCCAGCAAACGACCAAGATGTGGTCATTTACTTTTCAGGAGATGACAGCTCCTCTCTTCAG---       177
          D   E   M   P   A   Q   E   Q   P   P   A   N   D   Q   D   V   V   I   Y   F   S   G   D   D   S   S   S   L   Q
```

Supplementary Fig. 1

```
                 1        15            30             45              60            75          90        105
Defa7       MRTLTLLTT-LLLLALHTQAESPQGSRSTEE-----ALDQEQLVMEDQDISISFGGDKGTALQD---ADVK-----SGVTCYCRLLS-CQFGERLAGSCRSGG---VTYI
RatNP1/2    MRTLTLLTA-LLLLALHTQAKSPQGTAEE-------APDQEQLVMEDQDISISFGGDKGTALQD---ADVK-----AGVTCYCRRTR-CGFRERLSGACGYRG---RIYI
Defa12-ps   MRTLTLLTA-LLLLALHTQAESPQGRAEE-------APDLE-LVMKDQDISISFGGDKGTALQN---ADVK-----SGVTCYCRIS--CQFGERLSGSCRSGG---VTYI
Defa11      MGTLTLLTT-LLLLALHTQAESPQGCT-E---------DQKLLIMEDQDISISFTGDKGTALQD---ADVK-----SGVICNCRPRG-CSFKERVYGECGYNG---HILI
RatNP3      MRTLTLLTT-LLLLALHTQAESPQGSTKE-------APD------EEQDISVFFGGDKGTALQD---AAVK-----AGVTCSCRTSS-CRFGERLSGACRLNG---RIYI
Defa10      MRTLTLLTA-LLLLALHTQAESPQGSPKE-------APDQEQLDMEDQDISVFFGGDKGTALQD---AGER-----PGSTCSCRIGT-CVSGEWLSWVCRING---RIYI
RatNP4      MRTLTLLIT-LLLLALHTQAESPQERAKA-------APDQD-MVMEDQDIFISFGGYKGTVLQD---AVVK-----AGQACYCRIGA-CVSGERLTGACGLNG---RIYI
DEFA4       MRIIALLAA-ILLLVALQVRAGPLQARGDE------APGQEQRGPEDQDISISFAWDKSSALQV---SGST-----RGMVCSCRLVF-CRRTELRVGNCLIGG--VSFI
PTAD4       MRIIAILAA-ILLLVALQVRAGPLQARGDE------APGQEQRGPEDQDISISFAWDKSSALQV---SGST-----RGMVCSCRLVF-CRRTELRVGNCLIGG--VSFI
DEFA8P      -RSLALLAA-ILLLVALQARAEPLQAIADE------ATAQEQPGADDQEVVDSFAWDERAPLQV---SGSR-----RGLTCTCRD-S-YSPGEHRRGTCSAPG---VRYI
PTAD8-ps    -RSLALLAA-ILLLVALQARAEPLQAMTDE------ATAQEQPGADDQEVVDSFAWDERAALQV---SGSR-----RGLTCTCRD-S-CSPGEYRRGTCSAPG---VRYI
DEFA7P      -RTLTLLSA-FLLLVALQAWAEPLPARAHE------MPAQKQPPADDQDVVLYFSGDDSCSLQV---PGST-----KGLICHCRVLY-CIFGEHLGGTCFILG---ERYI
PTAD7       MRTLTLLSA-FLLLVALQAWAEPLQARADE------MPAQKQPPADDQDVVIYFSGDDSSLQV---PGST-----KGLICHCRVLY-CLFGEHLGGTCFIHG---ERYI
PAAD3       -RFATLLAHA-LLSMLQAWTEPLQARADE-------MPAQEQPPANDQDVVIYFSGDDSSLQV---PGST-----KGLSCHCRVLY-CLFGEHLGGTCFIHG---ERSI
DEFA6*      MRTLTILTA-VLLLVALQAKAEPLQAEDDPLQAKAYEADAQEQRGANDQDFAVSFAEDASSSLRA---LGST-----RAFTCHCRRS--CYSTEYSYGTCTVMG---INHI
PTAD6       MRTLTILTA-VLLLVALQAKAEPLQAEDEPLQAKAYEADAQEQRGANDQDFAVSFAEDASSSLRA---LGST-----RAFTCHCRRS--CYSTEYSYGTCTVMG---INHI
DEFA9P      MRTLTLLTA-ILLVPLQALSEPLQARAD--------AAAQEQLGADDQDFTISFTGDTSSGFRA---SGPR-----TLLRCSCRIRD-YHLQERNYGTCT-SG--EIHYI
PTAD9-ps    MRTLTLLTA-ILLVPLQALSEPLQARAD--------AAAQEQLGADDQDFAISFTGDTSLGFRA---SGPR-----TLLRCSCKIRD-YHLQERNYGTCT-SG--EIHYI
DEFA11P     MRTLTLLAA-ILLVPLQALSEPLQARAD--------AAAQEQLGADHQDFAISFTGDISSGFRD---SGSA-----TGLVCYGRISG-CHFSKHRYGTCTNSGIMEHRGI
PTAD11-ps   MRTLSLLAA-ILLVPLQALSEPLQARAD--------AAAQEQLGADHQDFAISFTGHIRSGFRD---SGST-----RGLVCYCRISG-CHFSKHHYGTCTNSGIMEHRGI
DEFA1/2     MRTLAILAA-ILLLVALQAQAEPLQARADE-----VAAAPEQIAADIPEVVVSLAWDESLAPKH---PGSR-----KNMACYCRIPA-CIAGERRYGTCIYQG---RLWA
DEFA3       MRTLAILAA-ILLLVALQAQAEPLQARADE-----VAAAPEQIAADIPEVVVSLAWDESLAPKH---PGSR-----KNMDCYCRIPA-CIAGERRYGTCIYQG---RLWA
PTAD1       MRTLAILAA-ILLLVALQAQAEPLQARADE-----VAAAPEQIPADNPEVVVSLAWDESLAPKH---PGSR-----KNVACYCRIPA-CLAGERRYGTCIYQG---RLWA
RMAD3       MRTLVILAA-ILLLVALQAQAEPLQARTDE------ATAAQEQIPTDNPEVVVSLAWDESLAPKDS-VPGLR-----KNMACYCRIPA-CLAGERRYGTCFYRR---RVWA
RMAD8       MRTLVILAA-ILLLVALQAQAEPLQARTDE------ATAAQEQIPTDNPEVVVSLAWDESLAPKDS-VPGLR-----KNMACYCRIPA-CLAGERRYGTCFYLR---RVWA
RMAD1/2     MRTLVILAA-ILLLVALQAQAEPLQARTDE------ATAAQEQIPTDNPEVVVSLAWDESLAPKDS-VPGLR-----KNMACYCRIPA-CLAGERRYGTCFYLG---RVWA
DEFA10P     MRTLALLTATVLFVALQAQAEPLRARADE-------AAAQE-PGADDQEVSVSFTWDESAALPL---SGESG---RGLRCICRRGI-CHFFQRRFGSCALPG---RLYI
PTAD10-ps   MRTLVLLTATVLFVALQAQAEPLRARADE-------AAAEEQPGADDQEVSVSFTWDESAALPL---SGSG-----RGLRCICR-GI-CHFFQCRFGSCALPG---RLYI
PTTD1-ps    MRTFALLTAMLLLVAL-AQAEPLQARADE-------AAAQEQPGADDQEMAHAFTWDESAALPL---SDSA-----RGLRCIGGRGI-CGLLQRRFGSCAFRG---TLHI
PTTD2-ps    MTTTFALLTAMLLLVAL-AQAEPLQARADE------AAAQEQPGADDQEMAHAFTWDESAALPL---SDSA-----RGLRCIGGRGI-CGLLQRRFGSCAFRG---TLHI
DEFT1P      MRTFALLTAMLLLLVAL-AQAEPLQARADE------AAAQEQPGADDQEMAHAFTWHESAALPL---SDSA-----RGLRCICRGI-CRLLRR-FGSCAFRG---TLHI
RTD1a       MRTFALLTAMLLLVALHAQAEARQARADE-------AAAQQQPGTDDQGMAHSFTWPENAALPL---SESA-----KGLRCICTRGF-CRLL------|----------
RTD1b       MRTFALLTAMLLLVALHAQAEARQARADE-------AAAQQQPGADDQGMAHSFTRPENAALPL---SESA-----RGLRCLCRRGV-CQLL------|----------
PAAD2       MRTLAILAA-ILLVALQAQAESLQERADE-------VATQEQPGEDDQDLAVSFEENGLST-LRA--SGSQ-----ARRTCYCRTGR-CYTPEFHSGKCLFKG---LTYI
RED4        MRTLTILAA-ILLFTLQAQAESLQERADE-------VATQEQPGKDDQDLAVSFEENGLSTLRA---SGSQ-----ARRTCYCRTGR-CYTPEFHSGKCVFNG---RTYI
RED3        MRTIAILAA-ILLFSLQAQAESLQERADE-------VVDQEQPGEDDQDFAVSFEENGLSTLRA---SGSQ-----ARHTCYCRNKR-CFTPEFHAGKCKVEG---RTYI
RMAD4/5     MRTIAILAA-ILLFALLAQAKSLQETADD-------AATQEQPGEDDQDLAVSFEENGLSTLRA---SGSQ-----ARRTCRCRFGR-CFRRESYSGSCNING---RIFS
RMAD6/7     MRTIAILAA-ILLFALLAQAKSLQETADE-------AATQEQPGEDDQDLAVSFEENGLSTLRA---SGSQ-----ARRTCRCRFGR-CFRRESYSGSCNING---RISS
RED1        MRTIAILAA-ILLFALLAQAKSLQETADE-------AATQEQPGEDDQDLAVSFEENGLSTLRD---SGSQ-----ARRTCRCRIRR-CRGLESSFGNCILHG---QFAI
RED6        MRTLTILAA-ILLFALLAQAKSLQETADD-------AATQEQPGEDDQDLAVSFEENGLSTLRA---SGSQ-----ARRNCHCRIGH-CRRPAAPMGVCIIHG---QFGI
PAAD1       MRTLAILAA-ILLVALQVAESLQERADE-------ATTQEQPGEDDQDFAVSFEENGLST-LKA--SGSR-----ARIICHCQIGR-CSRPESYSRICILHG---QVAI
RED5        MRTLTILAA-ILLVALQAQAESLQERADE-------AATQEQPAEDNQDFAVSFEENGLSTLKA---SGSQ-----ARMICLCRIGR-CSWREAHFGSCTKMG---QFAI
RED2        MRTIAILAA-ILLVALQAQAESLQERADE-------ATTQEQPGEDDQDLAVSFEENGLSTPRA---SGPQ-----ARFTCHCRIGR-CSWFETRFRSCTLLG---LAAI
DEFA5       MRTIAILAA-ILLVALQAQAESLQERADE-------ATTQKQSGEDNQDLAISFAGNGLSALRT---SGSQ-----ARATCYCRTGR-CATRESLSGVCEISG---RLYI
PTAD5       MRTIAILAA-ILLVALQAQAESLQERADE-------ATTQKQSGEDNQDLAISFAGNGLSALRT---SGSQ-----ARATCYCRIGH-CTILESLSGVCEISG---RLYI
NP1/MCP1    MRTLALLAA-ILLVALQAQAEHVSVSIDE-------VVDQQPPQAEDQDVAIYVKEHESSALEA---LGVK-----AGVVCACRRAL-CLPRERRAGFCRIRG---RIHI
NP2/MCP2    MRTLALLAA-ILLVALQAQAEHISVSIDE-------VVDQQPPQAEDQDVAIYVKEHESSALEA---LGVK-----AGVVCACRRAL-CLPLERRAGFCRIRG---RIHI
NP4         MRTLALLAA-ILLVTLQAQAELHSGMADD-------GVDQQQPRAQDLDVAVYIKQDETSPLEV---LGAK-----AGVSCTCRRFS-CGFGERASGSCTVNG---VRHI
NP5         MRTLALLAA-ILLVTLQAQAELHSGMADD-------GVDQQQPRAQDLDVAVYIKQDETSPLEV---LGAK-----AGVFCTCRGFL-CGSGERASGSCTING---VRHI
GNCP1A      MRTVPLFAA-CLLLLTLMAQAEPLPRAADH-------SDTKMKGDREDHVAVISFWEEESTSLED---AGAG-----AGRRCICTTRT-CRFPYRRLGTCIFQN---RVYI
GNCP1B      MRTVPLFAA-CLLLLTLMAQAEPLPRAADH-------SDTKMKGDREDHVAVISFWEEESTSLQD---AGAG----AGRRCICTTRT-CRFPYRRLGTCIFQN---RVYI
GNCP2       MRTVPLFAA-CLLLLTLMAQAEPLPRAADH-------SDTKMKGDREDHVAVISFWEEESTSLQD---AGAG-----AGRRCICTTRT-CRFPYRRLGTCLFQN---RVYI
RK1         MRTFALLTA-IVLVALQAKVEPVPPSDEE-------IPDQEGTDVGGQVMAISFAEDETSALQD---PVSP--LRSQKMPCSCKKY--CDPWEVIDGSCGLFN----SKY
RK2         MRTFTLLTA-IVLVALQAKPLSPSDEE-------IPDQEQTDVGEQVMANSCAEDETSALQN---RGCP---QNLKPYCSCKWR--CGIGEEEKGICHKFP--IVTY
Defa9       MKTLVLLSA-LVLLAFQIQADPIQEAEEE-------TKTEEQPADEDQDVSVSFEGPEPSALQN---LEIR-----WPWKRCHCRS-F-CRPYENATSFCAQGL--FKQHI
Defa8       MKTLVLLSA-LVLLAFQVQADHIQEAEEE-------TKTEEQPGADDQDMSVSFEGPEASALQN---LEIG----WPLKQCHCRK-F-CRPYEKAEGSCRPGL--FIKRI
Defa6       MKTLVLLSA-LVLVAYQVQADPIQEAEEE-------TKTEEQPSDEDQDVSVSFEGPEASALQD---FEIGR----PVRRCRCRAN--CGPKEYATAFCAQGP--FKQFI
Defa-rs1    MKTLILLSA-LVLLALQVQADPIQEAEEE-------TKTEEQPADEDQDVSVSFEGPEASAVQD---LRVR----RTLQCSCRR-V-CR---NTCSCIR-LSRSTYAS
CRS1C-2     MKTLVLLSA-LALLAFQVQADPIQNTDEE-------TKTEEQPEEEDQAVSVSFGGTEGSALQDV--AQRRFLWCRKCPVCQ-K-CQVCQKCPV--CPTCPQCPKQP--
CRS1C-3     MKTLVLLSA-LALLAFQVQADPIKNTDEE-------TKTGEQPEEEDQAVSVSFGGTEGSALQYV--AQRRFPWCRKCPVCQ-K-CQVCQKCPV--CPTCPQCPKLP--
CRS1C-1     MKTLVLLSA-LALLALQVQADPIQNTDEE-------TKTQEQPGEEDQAVSVSFGGTEGSALQDV--AQRRFPWCRKCRVCQ-K-CEVCQKCPV--CPTCPQCPKQP--
CRS4C-1     MKKLVLLFA-LVLLAFQVQADSIQNTDEE-------TKTEEQQGEKDQAVSVSFGDQGSALQD---AALG--WGRRCPQCP-R-CPSCPSCPR-CPRCPRCK------
CRS4C-2     MKTLVLLSA-LVLLAFQVQADSIQNTDEE-------TKTEEQQGEENQAVSVSFGDPQGSGLQDA--AALG--WGRRCPRCP-P-CPRCSWCPR-CPTCPRCN------
CRS4C-4     MKKLVLLFA-FVLLAFQVQADSIQNTDEE-------TKTEEQPGEENQAMSVSFGDPEGSALQD---AAVG--MARPCPPCP-S-CPSCPWCPM-CPRCPSCK------
CRS4C-5     MKKLVLLSA-FVLLAFQVQADSIQNTDEE-------IKTEEQPGEENQAVSISFGDPEGYALQDA--AAIR--RARRCPPCP-S-CLSCPWCPR-CLRCPMCK------
CRS4C-6     MKTLVLLSA-LVLLAFYVQADSTQETDEE-------TKTDDQPGEDQGVSVSFEDPERYVLQVS---GLG--KPPQCPKCP-V-CSKCPQCPQ-CPQCPGCPR------
Defcr6      MKTLILLSA-LVLLAFQVQADPIQNTDEE-------TKTEEQPGEEDQAVSVSFGDPEGTSLQE---ESLR------DLVCYCRARG-CKGRERMNGTCRKGH---LLYI
Defcr24     MKTLILLSA-LVVLAFQVQADPIQNTDEE-------TKTKEQPGEEDQAVSVSFGDPEGASLQE---ESLR------DLVCYCRARG-CKGRERMNGTCSKGH---LLYI
Defcr1      MKKLVLLFA-LVLLGFQVQADSIQNTDEE-------TKTEEQPGEEDQAVSVSFGDPEGTSLQE---ESLR------DLVCYCRSRG-CKGRERMNGTCRKGH---LLYI
Defcr3      MKTLVLLSA-LVLLAFQVQADPIQNTDEE-------TKTEEQPGEEDQAVSVSFGDPEGSSLQE---ESLR------DLVCYCRKRG-CKRRERMNGTCRKGH---LMYI
Defcr23     MKTLVLLSA-LILLAFQVQADPIQNTDEE-------TKTEEQPGKEDQAVSVSFGDPEGSSLQE---ESLR------DLVCYCRTRG-CKRRERMNGTCRKGH---LIYI
Defcr26     MKTLVLLSA-LFLLAFQVQADPIQNTDEE-------TNTEVQPQEEDQAVSVSFGNPEGSDLQE---ESLR------DLGCYCRKRG-CTRRERINGTCRKGH---LMYI
Defcr-ps1   MKTLVLLSA-LALLAFQVQADPIQNTDEE-------TNTEQPGKEDQAVSVSFGDPEGSSLQE---ESLI-YCTRG-CKRRERLNGTYRKGH---LMYI
Defcr25     MKTLVLLSA-LALLAFQVQADPIQNRDEE-------SKIDEQPGKEDQAVSVSFGDPEGSSLQE-ECDALR------DLICYCRTRG-CKRRERLNGTCRKGH---LMYI
Defcr20     MKTLVLLSA-LVLLAFQVQADPIQNTDEE-------TNTEEQAGEEDQAVSVSFGDPEGSALHE---KSSR------DLICYCRKGG-CNRGEQVYGTCS------GRLI
Defcr21     MKTLVLLSA-LILLAYQVQTDPIQNTDEE-------TNTEEQPGEDQAVSVSFGGQEGSALHE---KLSR------DLICLCRNRR-CNRGELFYGTCA------GPFI
Defcr4      MKTLVLLSA-LVLLAFQVQADSIQNTDEE-------TKTEEQPGKEDQAVSISFGGQEGSALHE---KSLR------GLLCYCRKGH-CKRGERVRGTCG------IRFI
Defcr5      MKTFVLLSA-LVLLAFQVQADPIHKTDEE-------TNTEEQPGEEDQAVSISFGGQEGSALHE---ELSK------KLICYCRIRG-CKRRERVFGTCRNLF---LTFV
Defa13-ps   MKTLLFLSA-LVLLAFQIQADPMQGADEE-------NGVAA---RERSQICLSPLEAQKALLFKMQILFVSSIA--RNQKCYYRARS-CNREGIIRTCNKRH---ESLI
Defa14-ps   MKTLLFLSA-LVLLAFQVQADPTQEADEE-------TKTEKKPGKEDPDVAVSFGGPQGSAFQDADNLFVSLVS--RNQKCYYTARS-YRNREGIMGT-TKRH---ESLI
NP3         MRTLILLAA-ILLAALQAQAELFSVNVDE-------VLDQQQP-GSDQDLVIHLTGEESSALQV---PDTK------G-ICACRRRF-CPNSERFSGYCRVNG---ARYV
RD5         MKTLVLLSA-LVLLALQVQAEPTPKTDEG-------TKTDEQPGKEDQVVSVSIEGQGDPAFQDA---VLR------DLKCFCRAKS-CNWGEGIMGICNKRY---GSLI
```

CHAPTER IV

CROSS-SPECIES ANALYSIS OF THE MAMMALIAN B-DEFENSIN GENE FAMILY: PRESENCE OF SYNTENIC GENE CLUSTERS AND PREFERENTIAL EXPRESSION IN THE MALE REPRODUCTIVE TRACT

Amar A. Patil[1], Yibin Cai[1], Yongming Sang[2], Frank Blecha[2] and Guolong Zhang[1]

[1]Department of Animal Science, Oklahoma State University, Stillwater, OK 74078;

[2]Department of Anatomy and Physiology, Kansas State University, Manhattan, KS 66506.

ABSTRACT

Mammalian β-defensins are an important family of innate host defense peptides with pleiotropic activities. As a first step to study the evolutionary relationship and biological role of the β-defensin family, we identified their complete repertoires in the human, chimpanzee, mouse, rat, and dog following systemic, genome-wide computational searches. Although most β-defensins genes are composed of two exons separated by an intron of variable length, some contain additional one or two exons encoding an internal pro-sequence, a segment of carboxyl terminal mature sequences or untranslated regions. Alternatively spliced isoforms have also been found with several β-defensins. Furthermore, all β-defensin genes are densely clustered in 4-5 syntenic chromosomal regions with each cluster spanning <1.2 Mb across the five species. Phylogenetic analysis indicated that, although the majority of β-defensins are evolutionarily conserved across species, subgroups of gene lineages exist that are specific in certain species, implying that some β-defensins originated after divergence of these mammals from each other, while most others arose before the last common ancestor of mammals. Surprisingly, reverse transcriptase-PCR revealed that all but one rat β-defensin transcripts are preferentially expressed in the male reproductive tract particularly in epididymis and testis, except that *Defb4,* a human β-defensin-2 ortholog, is more restricted to the respiratory and upper gastrointestinal tracts. Moreover, most β-defensins expressed in the reproductive tract are developmentally regulated with enhanced expression during sexual maturation. Existence of such a vast array of β-defensins in the male reproductive tract suggests that these genes may play a dual role in both fertility and host defense.

INTRODUCTION

Defensins and defensin-like molecules comprise a diverse group of cationic antimicrobial peptides characterized by the presence of multiple cysteine residues and a highly similar tertiary structure known as the defensin motif (14, 28, 35, 46, 60). Defensin-like genes have been discovered in many species including plants, fungi, arthropods, mollusks, reptiles, birds, and mammals (13, 14, 28, 36, 47, 52, 55, 57). Defensins in vertebrate species are further classified into three families, namely α-, β-, and θ-defensins, based on the spacing pattern of six cysteine residues (10, 14, 28). Although the evolutionary relationship between vertebrate defensins and defensins in other species remains unclear, phylogenetic analysis revealed that a primordial β-defensin gene is the common ancestor for all vertebrate defensins (57). β-defensins are believed to evolve before the divergence of mammals from birds (57) and gave rise to α-defensins in glires and primates after they diverged from other mammalian species (36). θ-defensins further originated from α-defensins following separation of primates from other mammals (33, 36).

In contrast with α- and θ-defensins that are produced primarily in the granules of either leukocytes or intestinal Paneth cells, β-defensins are unique in that they are primarily produced by nongranular mucosal epithelial cells lining the respiratory, gastrointestinal, and genitourinary tracts (14, 28, 35, 46, 60). In addition to their broad-spectrum antimicrobial activity against bacteria, fungi, and certain enveloped viruses, β-defensins have been recognized recently to resemble chemokines both structurally and functionally (14, 28, 46, 60). Recent studies also revealed that certain β-defensins are actively involved in sperm maturation and capacitation (41, 65, 68, 70), physiological properties that are not related to host defense.

All β-defensin genes encode a precursor peptide that consists of a hydrophobic, leucine-rich signal sequence, a pro-sequence, and a mature six-cysteine defensin motif at the carboxyl terminus (14, 28). In most cases, β-defensin precursors are encoded in two separate exons separated by an intron of variable length with one exon encoding the signal and pro-sequence and the other encoding primarily the mature peptide (14, 28). It is generally believed that mature β-defensins need to be cleaved from the pro-sequences by proteolytic enzymes to become biologically active, although such proteases are yet to be identified.

A number of putative β-defensins were reported recently in humans and mice (47); however, later identification of a novel human sequence implied possible existence of additional β-defensins in those two species (42). In addition, the evolutionary relationships among β-defensin genes across mammalian species are poorly understood, although limited data are available on β-defensin evolution within species or among the primates (9, 22, 32, 48).

Recent availability of complete genome sequences of a number of phylogenetically distinct mammalian species provides an excellent opportunity to identify the entire β-defensin gene family members in these species. Here, we report the discovery of complete repertoires of β-defensins in the human, chimpanzee, mice, rat, and dog by using a systemic computational search strategy that we developed recently (36, 57). We showed that β-defensin genes form 4-5 syntenic clusters in these mammals. Although the majority of β-defensin genes are evolutionarily conserved across rodents, canines and primates, a few gene lineages exist only in certain species. We also examined the tissue expression patterns of the entire β-defensin gene family in the rat by reverse transcriptase-PCR. To our surprise, all but one of the β-defensins were found to be

preferentially expressed in the male reproductive system, particularly in testis and different segments of epididymis. The only exception, *Defb4*, is more restricted to the respiratory and upper gastrointestinal tracts with virtually no expression in the male reproductive system. Furthermore, we provide evidence that many β-defensins are developmentally regulated and the expression levels are elevated with age. These findings argue that the major biological functions of β-defensins may be related to reproduction and fertility, in addition to innate host defense.

MATERIALS AND METHODS

Computational search for novel β-defensins

To identify potential novel β-defensins in the human, mouse, rat, dog, and chimpanzee, systemic computational searches were performed essentially as we described (36, 57). In brief, all known β-defensin peptide sequences were individually queried against expressed sequence tags (EST), nonredundant sequences (NR), unfinished high throughput genomic sequences (HTGS), and whole genome shotgun sequences (WGS) in GenBank by using the TBLASTN program (1) with the default settings on the NCBI web site **(http://www.ncbi.nlm.nih.gov/BLAST)**. We then examined all potential hits for the presence of the characteristic β-defensin motif or conserved signal peptide and prosequence. Additional iterative BLAST searches were performed for every novel β-defensin sequence identified as described above until no more novel sequences were revealed. Because mammalian defensins tend to form clusters (47, 48), all genomic sequences containing β-defensins were also retrieved from GenBank to discover potential novel sequences with distant homology. The nucleotide sequences between two neighboring defensin genes were translated into six open reading frames

103

and individually compared with the two defensin peptide sequences for the presence of β-defensin motif or signal peptide/pro-sequence by using the BLASTP program (1) on the NCBI web site **(http://www.ncbi.nlm.nih.gov/blast/bl2seq/)** and/or ClustalW program (version 1.82) (53) on the European Bioinformatics Institute web site **(http://www.ebi.ac.uk/clustalw)**.

Prediction of full-length coding sequences and genomic structures of β-defensins

In most cases, signal and pro-sequence of β-defensins are encoded in a separated exon from the mature peptide-encoding exon. If either the signal/pro-sequence or β-defensin motif of a novel gene was missing in a genomic sequence, a 5-15 kb sequence flanking sequence was retrieved to identify the full-length coding sequence and to derive the structural organization of that novel β-defensin gene by using a combination of GenomeScan (61), GENSCAN (5) and/or GeneWise2 (4). All cDNA sequences of novel β-defensins described in this study have been submitted to GenBank and the accession numbers are listed in Supplementary Table 1.

Chromosomal mapping of β-defensin gene clusters

The BLAT program (25) was used to determine the relative position and orientation of each defensin in the genome through the UCSC Genome Browser **(http://genome.ucsc.edu)**. Individual defensins were searched against latest versions of the assembled genomes of human (NCBI Build 35 Version 1), chimpanzee (NCBI Build 1 Version 1), mouse (NCBI Build 32), rat (BCM Version 3.1), and dog (NCBI Build 1 Version 1) released in June 2004, November 2003, September 2004, June 2003, and August 2004, respectively.

Sequence alignment and phylogenetic analysis of β-defensins

Multiple sequence alignments were carried out by using the ClustalW program (version 1.82) (53). The neighbor-joining method (43) was used to construct the phylogenetic tree by calculating the proportion of amino acid differences (*p*-distance), and the reliability of each branch was tested by 1000 bootstrap replications.

Semi-quantitative RT-PCR analysis of tissue and developmental expression patterns of β-defensins

A total of 29 different tissue samples (see Fig. 6B) were harvested from healthy, 2-month old Sprague-Dawley rats. Total RNA was extracted using Trizol (Invitrogen). For each RNA sample, 4 μg were reverse transcribed with random hexamers and Superscript II reverse transcriptase by using a first-strand cDNA synthesis kit (Invitrogen) according to the instructions. Subsequent PCR was performed with DNA Engine (Model PTC-200, MJ Research) essentially as described (36, 57, 69). Briefly, 1/40 of the first-strand cDNA from each tissue were used to amplify β-defensins and glyceraldehydes 3-phosphate dehydrogenase (GAPDH) with gene-specific, exon-spanning primers (Supplementary Table 2). The PCR program used was: 94°C denaturation for 2 min, followed by different cycles of 94°C denaturation for 20 sec, 55°C annealing for 20 sec, and 72°C extension for 40 sec, followed by a final extension at 72°C for 5 min. The number of PCR cycles was optimized for each gene to ensure linear amplification (Supplementary Table 2). Specificity of each PCR product was confirmed by cloning the PCR product into the pGEM-T Easy vector (Promega), followed by sequencing of the recombinant plasmid.

To examine developmental expression of the reproductive tract-specific β-defensins, testes were harvested from Sprague-Dawley rats on 1, 15, 30, 60 and 90 days after birth; and epididymis were collected on 7, 15, 30, 60 and 90 days after birth with 2-4 animals per time point. RNA isolation and semi-quantitative RT-PCR were performed as described above. Signal intensity of each gene-specific PCR product was quantified by using Scion Imaging software (**http://www.scioncorp.com**), followed by normalization against the GAPDH signal amplified from the same cDNA sample, and the results were expressed as relative level of expression.

RESULTS

Discovery of mammalian β-defensin gene repertoires

Based on the conservation of the characteristic six-cysteine-containing β-defensin motif, an iterative, genome-wide homology search strategy was developed to identify complete repertoires of the β-defensin gene family in the human, chimpanzee, mouse, rat, and dog as we previously described (36, 57). Because such computational searches are biased toward the defensin motif, the nucleotide sequences identified were primarily composed of the exons encoding the cysteine-spanning region. To identify the first exon sequences, 5-15 kb of finished or unfinished genomic sequence containing each putative defensin gene were retrieved from GenBank and used for computational prediction by using GenomeScan (61), GENSCAN (5), and/or GeneWise2 (4). As a result, a number of novel, full-length defensin precursor sequences with conserved six cysteines and signal peptides have been identified in each species and submitted to the GenBank. All new genes were named by following the recommendations of the

HUGO and Mouse Gene Nomenclature Committees and kept in consistence with the published sequences.

In the case of rat, a total of 42 β-defensin (*Defb*) genes and a pseudogene (*Defb16-ps*) were discovered, including four known ones (24, 29, 41, 68) (Fig. 1). Among all rat genes, 12 β-defensins, namely *Defb1, 4, 14, 22, 24, 27, 29, 33, 36, 39, 51,* and *52,* were found in the EST database (data not shown), from which full-length peptide precursor sequences can be deduced (Fig. 1). All the remaining genes were identified from WGS and HTGS seuqences with first exons being predicted, except for *Defb19*. Consistent with earlier observations, the predicted signal sequences are highly homologous with known β-defensin genes in other mammalian species, rich in leucine and hydrophobic residues (Fig. 1). The accuracy of the exon-intron boundary predicted for each gene has also been confirmed by direct sequencing of RT-PCR product with exon-spanning primers (Supplementary Table 2). Inability to identify the signal sequence for *Def19* may be due to mis-annotation of genome sequences or presence of premature stop codons in its first exon.

It is noted that all rat β-defensin genes were named according to their mouse orthologs as previously described and approved by the Mouse Nomenclature Committee (47). Consequently, rat β-defensin-1 (24), β-defensin-2 (24), E-3/2D6 (41, 68), and *Bin-1b/EP2e* (29) have been renamed as *Defb1, Defb4, Defb22,* and *Spag11e*, respectively. An extra internal exon encoding an additional segment of pro-sequence has been predicted for *Defb12, Defb52,* and *Spag11c*, which have also been experimentally confirmed by sequencing of RT-PCR products in both

cases. Presence of an extra exon in *Spag11c* has also been shown previously in its orthologs (*HE2c/EP2c/SPAG11c*) in the human and rhesus monkey (11, 12).

A total of 43 novel β-defensins genes and pseudogenes were discovered in the canine genome (Supplementary Fig. 1). All sequences were named so as to maintain consistency with the nomenclature of human β-defensins. The canine β-defensin-1 (*CBD1*) gene that we identified earlier was re-named *CBD122* (44). Five other *CBD* genes with no human orthologs were named *CBD138-CBD142*. Among all canine sequences identified, eight (*CBD1, CBD102, CBD103, CBD108, CBD122*, *CBD124, CBD139,* and *CBD40,*) were also supported by at least one EST sequence (data not shown). First exons of all but two canine β-defensins were also predicted. While most *CBD* genes consist of two exons separated by an intron, *CBD105, CBD118, CBD122,* and *CSPAG11c* (canine ortholog of human *SPAG11c/HE2c*) were predicted to have an extra exon (data not shown).

Although a number of β-defensins have been described in humans and mice with the use of hidden Markov models (HMM) (47), later discovery of another novel sequence in the human (42) prompted us to re-screen the human and mouse genome sequences for possible presence of additional new sequences by using the strategy that we described above. Albeit of low throughput, our method yielded a total of 39 human and 52 mouse β-defensin genes and pseudogenes, including four novel human (*DEFB133, 134, 135,* and *136*) and six new mouse β-defensins (*Defb49, 50, 51, 52, 53*, and *54-ps*). All new β-defensin sequences consist of the characteristic signal sequence and a conserved defensin motif (Fig. 2A). Moreover, most of their orthologs are present in dogs and/or rodents except for mouse *Defb53 and Defb54-ps*, which

appear to be mouse-specific (Fig. 2B). To further support their authenticity, all new genes were found to locate inside the same genomic clusters as others (Fig. 3).

We also predicted first exons for most human (Supplementary Fig. 2) and mouse genes (Supplementary Fig. 3), which were lacking in the earlier publication (47). The predicted sequences and their intron-exon boundaries are consistent with the rat orthologous genes, which have been confirmed by direct sequencing. Such sequences are also in agreement with a few full-length human and mouse β-defensins as described in the follow-up studies (3, 32, 39, 42, 48, 58, 66).

We also searched the chimpanzee genome and found a repertoire of 37 β-defensin genes that share 91-100% identity to their human orthologs in each case (data not shown), except that *DEFB110* and *DEFB128* orthologs were not found the chimpanzee genomic sequences currently available in GenBank. Because of their conservation in dogs and rodents, there is little reason to suspect that *DEFB110* and *DEFB128* are missing in the chimpanzee, but instead it is most likely due to low coverage (4-fold) of the current genome sequences.

Alignment of all β-defensins revealed a high degree of diversity in the primary amino acid sequence in the mature region despite of the conservation of six cysteine residues, indicative of their possible nonredundant biological functions. Noticeably, the DEFB126/CBD126/Defb22 lineage (Fig. 1 and Supplementary Figs. 1, 2, and 3) and dog-specific CBD141 (Supplementary Fig. 1) contain additional four amino acids between the third and fourth cysteine as compared to most other β-defensins. Moreover, three canine- and rodent-specific β-defensin lineages, namely

CBD138/Defb33, CBD139/Defb52, and CBD140/Defb51, consist of 3-4 extra amino acids between the second and third cysteine but 2-3 fewer amino acids between the third and fourth cysteine (Fig. 1 and Supplementary Figs. 1 and 3). Conceivably, such variations are expected to alter their overall tertiary structures and possibly functional properties from conventional β-defensins.

Strikingly, in contrast to most characterized β-defensins that contain less than 5 amino acids following the last cysteine, the majority of the family members in fact have much longer and mostly linear C-terminal tails. The role of such long tails remains largely a mystery. However, several β-defensins were recently found to be heavily glycosylated in vivo in the tail region (30, 41, 65, 68). Furthermore, rat Defb22/E-3/2D6 and rhesus DEFB118/ESC42 contain lectin and trefoil-like motifs, respectively, which potentially facilitate oligomerization and interactions of proteins with target cell membrane (30, 41).

Chromosomal clustering of β-defensins

To identify genomic organization of β-defensins, individual β-defensin genes were searched against the most current versions of the assembled human, mouse, rat, chimpanzee, and dog genomes by using the BLAT program (25) through the UCSC Genome Browser. It is apparent that β-defensin genes in each species are clustered densely in 4-5 different chromosomal regions with each cluster spanning between 55-1118 kb (Fig. 3). All clusters are syntenic across species because orthologous genes are extensively conserved with the same order and orientation for most genes in each cluster (Fig. 3).

In the rat genome (BCM Version 3.1), all 43 β-defensin genes have been mapped and form four distinct clusters expanding 812 kb, 78 kb, 55 kb, and 193 kb continuously on chromosomes 16q12.5, 15p12, 9q12-q13, and 3q41, respectively (Fig. 3). Such syntenic clusters are also conserved in mice (NCBI Build 32) with four clusters spanning 1118 kb, 71 kb, 56 kb, and 164 kb on chromosomes 8qA1.3-A2, 14qC3, 1qA3, and 2qH1, respectively (Fig. 3). However, the genomic contig containing mouse *Defb33, Defb51,* and *Defb52* have not been assigned to a specific region on mouse chromosome 8. The location of mouse *Defb48-ps* currently remains unmapped. Mouse *Defb32* reported earlier (47) was found on chromosome 6. However, we suspect that *Defb32* is unlikely to be an authentic β-defensin gene, because of the facts that 1) only five instead of six cysteines are present in the putative mature region, 2) such a cysteine-containing region is negatively instead of positively charged, 3) no characteristic first exon sequence could be identified within a 15-kb upstream sequence, and 4) no orthologs can be found in other mammals. It is also noteworthy that we could not find mouse *Defb31* and *Defb47-ps* (47) in either EST, WGS, HTGS, or assembled genome sequences that are publicly available in GenBank.

The canine genome (NCBI Build 1 Version 1) also encodes four β-defensin clusters expanding 356 kb, 76 kb, 99 kb, and 373 kb on the syntenic regions on chromosomes 16, 25, 12, and 24, respectively (Fig. 3). However, human β-defensin genes in the most current genome assembly (NCBI Build 35 Version 1) were annotated in five separate clusters on three different chromosomes with two on chromosome 8p23.1 (Fig. 3A and 3B), one on 6p12.3 (Fig. 3C), and two on 20p13 and 20q11.21 (Fig. 3D). Apparently, two human β-defensin clusters on the *p* and *q* arms of chromosome 20 share a common ancestor with a continuous synthetic cluster in

rodents and dogs (Fig. 3D). In contrast, another two human gene clusters on 8p23.1 are only separated by less than 4 Mb, but their counterparts in rodents and dogs are located on two different chromosomes (Fig. 3A and 3B). The current chimpanzee genome (NCBI Build 1 Version 1) also encodes five major β-defensin clusters in almost perfect synteny with human genes (Supplementary Fig. 4). It is noteworthy that the α-defensin cluster is located within the largest β-defensin cluster in rodents and humans, but is apparently missing in the canine genome (Fig. 3A), consistent with the fact that α-defensins have only been discovered in glires and primates (36).

Earlier studies indicated that a varied copy number of α- and β-defensin genes are present on chromosome 8p23.1 of human individuals (19, 51). Because of such complexity, current annotation of several human β-defensins in this region, such as *DEFB107, 108, 109, 130*, and *131,* shows an obvious discrepancy in the location and order of their orthologous genes with other species (Fig. 3A and 3B). Thus, the accuracy of the current human genome assembly in this region remains to be refined and corrected in the future (51).

Nevertheless, our current version of human and mouse β-defensin genomic clusters manifests a significant improvement over an earlier attempt (47). For example, several genes in human 8p23.1 and their orthologous loci in mice have been re-ordered, which agrees with their syntenies in other species (Fig. 3A) and are also consistent with more recent studies (48, 51). Moreover, an ambiguous cluster containing *DEFB130* and *DEFB131* (47) has been mapped to human chromosome 8p23.1 with three additional novel genes (*DEFB134, 135,* and *136*) being

added (Fig. 3B). Another novel human β-defensin gene (*DEFB133*) has also been mapped to 6p12.3 and its orthologs have been found and localized in rodent and canine genomes (Fig. 3C).

Structural organizations and transcriptional flexibility of β-defensin genes

A comparison of all available EST with genomic sequences of β-defensins across mammalian species revealed that, similar to enteric α-defensins (35), most β-defensin genes are composed of two separate exons with the first exon encoding the signal and pro-sequence and the second exon encoding primarily the mature sequence containing the defensin motif. The intron sizes of β-defensins vary greatly, spanning 1-10 kb in most cases, in contrast with α- and θ-defensins consisting of introns of usually <1 kb (36). However, we also found additional five distinct patterns of the β-defensin gene structure (Fig. 4). For example, mouse *Defb21* (EST accession number AK076875), mouse *Defb44* (AK079042), rat *Defb29* (CK839317), rat *Defb51* (BM390654)*,* and canine *CBD102* (DN346173)*, CBD139* (CO687407)*,* and *CBD140* (CO692331) have EST sequences showing the presence of additional one or two exons encoding the 5'-untranslated region (UTR), reminiscent of myeloid α-defensin genes (35). On the other hand, mouse *Defb30* (AK078987) contains an additional exon encoding the 3'-UTR. Mouse *Defb28* (AV044615) and canine *CBD122* (BM537999) are unique in that they are composed of three exons with the last exon encoding a segment of the mature sequence downstream of the six cysteines, while the first two exons encoding the signal, pro- and, majority of mature sequences. A few genes, such as *SPAG11c /CSPAG11c/Spag11c* (11, 12) and *DEFB105/CBD105/Defb12* (data not shown)*,* exhibit a three-exon structure, containing an extra intron presumably encoding an additional segment of pro-sequence. Interestingly, rat *Defb52* (CR466423) is composed of four exons with three exons encoding the entire open reading frame and the fourth exon

encoding a part of 5'-UTR. Additional gene structures are expected to be revealed with availability of more mRNA and EST sequences for β-defensins.

It is well known that the *SPAG11* gene is capable of differentially utilizing different exons to produce multiple isoforms of β-defensin-like sequences (11, 12). Such alternative splicing now does not appear to be unique to the *SPAG11* gene. We have found two EST sequences (AK020304 and BB787829) that consist of the first exon of mouse *Defb17* and the second exon of *Defb16*, suggesting that these two genes are actually alternatively spliced isoforms. Such splicing is most likely evolutionarily conserved, because of the absence of signal sequence for its orthologous gene, *DEFB110/CBD110/Defb16* in the human, chimpanzee, dog, and rat. Therefore, primate *DEFB110* and *DEFB111*, canine *CBD110* and CBD*111*, and rat *Defb16-ps and Defb17* are likely to be two isoforms of the same gene. Similarly, human *DEB119* and *DEFB120* have been shown recently to share the same first exon encoding the signal sequence (39). *CBD122*, on the other hand, has three isoforms sharing identical signal, pro- and the majority of mature sequences with the only difference in the carboxyl terminal tail after the last cysteine (44). Collectively, acquirement of additional exons and frequent alternative splicing add additional layers of complexity in β-defensin gene transcription and regulation, which presumably allow the host to better cope with invading infections.

Phylogenetic analysis of β-defensins

To reveal the evolutionary relationships of all known β-defensins including newly identified ones in rodents, dogs, and primates, we sought to construct a phylogenetic tree by using the neighbor-joining method (43). Interestingly, genes within a genomic cluster tend to

form a separate clade from the genes in other clusters (data not shown), implying that many of them most likely evolved by gene duplication events. Indeed, a number of highly similar paralogous genes are present in each species and physically located adjacent to each other. Consistent with the genetic mapping data (Fig. 3), many β-defensins have obvious orthologs with minimum sequence variation across four mammalian species and are clustered together (Fig. 5), meaning that these lineages arose before the last common ancestor of rodents, canines, and primates. Some even form well-supported clusters with different chicken β-defensins, suggesting their existence prior to the divergence of mammals from birds (57). The conservation of these genes during evolution may be indicative of their functional significance.

Conversely, subgroups of β-defensins also exist that are specific in mice, rats, and dogs, implying that these genes originated after these species separated and have undergone different evolutionary patterns. For example, two well-supported clades of rodent-specific genes (Clade I: *Defb2, 9, 10,* and *11*; and Clade II: *Defb37, 38, 39, 40,* and 50) exist (Fig. 5). High sequence similarity within each clade supports the notion that these rodent-specific β-defensin genes arose from gene duplication and positive diversifying selection as previously described (32, 48). In contrast to the Clade II genes with the orthologs present in both rodent species, many mouse β-defensins in Clade I have no obvious orthologs in rats, implying that many Clade I genes were likely to have emerged only after the mouse-rat split about 41 million years ago (27), whereas all Clade II genes apparently appeared prior to the split. Another large cluster of rodent genes, namely *Defb3, 4, 5, 6, 7,* and *8,* form a common clade with human *DEFB4/hBD-2* gene (Fig. 5). Therefore, it is likely that such a gene lineage duplicated and expanded significantly in rodents, but remained unchanged in primates and was lost in the canine lineage. Alternatively, multiple

such ancestral gene lineages were lost in primates and canines, but retained in rodents. However, significant homology of murine genes in this cluster as well as their physical vicinity (Fig. 3A) supports the former conclusion.

Multiple gene lineages specific to dogs and primates are also present, including *DEFB104/CBD104,* *DEFB108/CBD108,* *DEFB114/CBD114,* *DEFB120/CBD120,* *DEFB121/CBD121, DEFB133/CBD133,* and *DEFB134/CBD134* (Figs. 3 and 5). Conversely, *CBD138/Defb33*, *CBD139/Defb52*, and *CBD140/Defb51* (Fig. 3A) only exist in rodents and dogs, but are absent in primates (Fig. 5). Dog-specific genes also exist, such as *CBD141* and *CBD142* (Fig. 3D). However, no primate-specific β-defensins were found, suggesting that gene duplication did not occur in the primates after their divergence. It is noteworthy that *CBD102* is not the canine ortholog of human *DEFB4,* murine *Defb2* or *Defb4*, but instead it is specific to the dog, paralogous to *CBD103* (Fig. 5 and Supplementary Fig. 1)*.* Therefore, the *CBD102* gene is most likely to evolve from duplication of the *CBD103* gene after separation of canines from other mammals.

Tissue expression pattern of rat β-defensins

We next examined the tissue expression patterns of the entire β-defensin gene family in the rat as a first step toward understanding their *in vivo* biological functions. Semi-quantitative RT-PCR was performed with all identified rat β-defensin genes by using a panel of 29 different tissues from healthy, 2 month-old Sprague-Dawley rats. Primers were designed from two different exons to differentiate between the amplicons from cDNA and those from genomic DNA and verify predicted exon-intron junctions (Supplementary Table 2). Different numbers of

PCR cycles were performed for each gene to ensure the linearity of amplification. Specificity of each amplicon was confirmed by direct sequencing of the RT-PCR product or recombinant plasmid containing the RT-PCR product.

To our surprise, nearly all β-defensins were found to be expressed preferentially in the male reproductive system particularly in epididymis and testis (Fig. 6A) with no or minimum expression in most other tissues (data not shown). More strikingly, the reproductive tract-specific genes are expressed differentially in three different regions of epididymis, namely caput (head), corpus (body), and cauda (tail) (Fig. 6A). For example, *Defb12/35, 15/34, 17, 18, 21, 25, 41, Spag11c,* and *Spag11e* are more abundantly expressed in the caput of epididymis, whereas *Defb9, 10, 11, 13,* and *40* are restricted to the cauda (Fig. 6A). On the other hand, *Defb24* and *Defb33* are highly specific to testis (Fig. 6A). Our results are consistent with earlier studies on the expressions of rat *Defb1/RBD-1* (8, 24) and *Spag11e/Bin-1b* (29) and also largely agree with the expressions of their human and mouse orthologs (39, 42, 58, 59, 66).

In addition to a canonical *Spag11c/EP2c* transcript (3, 12, 56), an alternatively spliced longer transcript was detected in the rat (Fig. 6A) and has been deposited in GenBank under accession number DQ012093. Such a transcript is unique in that it contains an extra 60 bp at the 3'-end of the first exon, resulting in a longer pro-sequence, which is not found in either human or rhesus monkey (11, 12). To further illustrate the difference in the *SPAG11/HE2/EP2* locus between primates and rodents, the *d* isoform, which is composed of the first two exons of the *c* form and the last exon of the *e* form, was not detected in the rat (data not shown), but is readily detectable in the primates (11, 12).

Although the majority of genes showed the most abundant expression in testis or epididymis of healthy adult rats, *Defb1* and *Defb42* are also expressed at considerable levels in kidney (Fig. 6A) and *Defb24* in spleen and ovary (data not shown). *Defb36* was found at a lesser extent on mucosal surfaces lining the respiratory, gastrointestinal, and reproductive tracts as well as on skin (Fig. 6B). The only obvious exception is *Defb4/RBD-2,* which shows virtually no expression in either testis or epididymis, but instead is more restricted to the respiratory tract, particularly in lung (Fig. 6B). *Defb4* transcript was also found at low levels in the upper gastrointestinal tract as well as in vagina (Fig. 6B), which is consistent with the previously published report (24). The biological significance of such concentrated, but apparently differential production of so many β-defensins in testis and different regions of epididymis remains largely unknown.

It is noted that we failed to detect the expression of *Defb3, Defb5,* and *Defb16-ps* following multiple attempts with different combinations of primer pairs. Although we cannot rule out possibility of wrong computational prediction of first exons, it is most likely that *Defb3* and *Defb5* are expressed at extremely low levels in healthy adult rats, but may be inducible in response to infection and inflammation. *Defb16-ps* could be a non-expressing pseudogene. Alternatively, these three genes could be expressed in the tissues other than the ones that we examined. Because of the inability to identify the first exon of *Defb19*, we could not detect its gene expression by RT-PCR either.

Developmental regulation of rat β-defensins

Semi-quantitative RT-PCR was performed to study developmental expression patterns of a few selected rat genes that are specific to epididymis and/or testis. Rat testes were collected on 1, 15, 30, 60, and 90 days after birth; and epididymis were collected on 7, 15, 30, 60, and 90 days after birth with 2-4 animals used per time point. As shown in Fig. 7A, expressions of all selected β-defensins (*Defb1, 15, 29, 30, 42*, and *49*) in epididymis are significantly elevated during development with virtually no expression on day 7 but peaked quickly at one month. *Defb42* is slightly different in that it shows little expression up to one month, but the expression level is dramatically enhanced at two months of age and increased up to three months after birth. Such a pattern is reminiscent of rat *Spag11e/Bin-1b*, whose transcript does not peak in epididymis until four months and remains elevated in rats of 2 years of age (29).

However, β-defensin genes in testis showed a different pattern. Although most epididymal β-defensin genes were not expressed on day 7, many genes in testis display considerable expression soon after birth, with some (*Defb24* and *Defb29*) expressed constitutively in rats of all ages studied and some (*Defb27 and Defb36*) were upregulated and peaked at one month of age (Fig. 7B). *Defb33*, on the other hand, behaves similar to epididymal *Defb42* (Fig. 7A) in that it is not detected in testis until one month, but continues to increase for at least three moths after birth (Fig. 7B). It is interesting to note that *Defb29* is developmentally regulated in the epididymis (Fig. 7A), but exhibits a constitutive pattern in the testis (Fig. 7B), implying the presence of certain epididymis-specific enhancer(s) or testis-specific inhibitory factor(s).

The constitutive expression of several β-defensins in testis, but not epididymis, of newborn animals is presumably required to protect testis and particularly germ cells from invading infections before the time of sexual maturation. Simultaneous production of such a large array of defensins in various segments of the male reproductive tract with different developmental expression patterns implies that these molecules may play a nonredundant role in the reproductive process.

DISCUSSION

Sequence diversity of mammalian β-defensins

Through systemic genome-wide screening, we have discovered a total of 39, 37, 43, 52, and 43 β-defensin genes and pseudogenes in the human, chimpanzee, dog, mouse, and rat, respectively. Alignment of all β-defensins revealed a high degree of conservation in the spacing pattern of six cysteines with a consensus pattern of $C-X_{5-8}-C-X_{3-7}-C-X_{5-13}-C-X_{4-7}-CC$. The invariantly spaced cysteines form a rigid, triple antiparallel β-sheet structure and are believed to be critically important in maintaining the functional activities (14, 28). However, we observed the absence of a canonical cysteine or presence of an extra cysteine in a few putative functional defensins. For example, human DEFB107 and mouse Defb8/Defr1 have a nonconservative mutation in the first cysteine, whereas mouse and rat Defb50 lack the second conserved cysteine (Supplementary Figs. 2 and 3). DEFB133/Defb49 lineage, on the other hand, contains a mutation in either the fifth or sixth cysteine in the mouse and human (Supplementary Figs. 2 and 3), but is fully conserved with six canonical cysteines in the rat (Fig. 1).

Extra cysteines have also been found within the six-cysteine motif of three peptides, namely human DEFB119 and DEFB132 and mouse Defb5 (Supplementary Figs. 2 and 3). In contrast, the instance of cysteines occurring outside the defensin motif appears to be much higher with multiple members being found in each species. Among 39 human β-defensins, seven such peptides are found (DEFB105, 106, 112, 117, 126, 127, and 133) (Supplementary Fig. 2). Interestingly, five are conserved across species, including DEFB105/CBD105/Defb12/Defb35, DEFB106/CBD106/Defb15/Defb34, DEFB112/CBD112, DEFB117/CBD117/Defb19, and DEFB126/CBD126/Defb22.

It is largely unknown how a missing or an extra cysteine would affect the tertiary structures and functional properties of β-defensins. Obviously, the existence of an odd number of cysteines will potentially oligomerize through intermolecular disulfide bridging. Indeed, such β-defensins as Defb8/Defr1 (6) and DEFB126/Defb22/ESP13.2/E-3/2D6 (41, 65, 68) have been shown to form covalently bounded homodimers in vitro and/or in vivo. Such dimerization in fact leads to enhanced antibacterial activities in the case of Defb8/Defr1 (6). Not surprisingly, a few β-defensins with six cysteines also have a tendency to form dimers or oligomers, which are believed to facilitate membrane interaction and permeabilization (20, 45).

It is noteworthy that our strategy yielded a total of 10 additional novel β-defensins in humans and mice as compared with the strategy based on the Hidden Markov Model (HMM) (47). Although we believe that the current list most likely represents the complete repertoires of β-defensins in these species, we could not rule out the possibility that additional β-defensin-related genes with remote similarity might be uncovered in these species. In fact, a group of

human SPAG11/HE2/EP2 isoforms were recently found to share identical signal and pro-sequences with similar antimicrobial properties with classic β-defensins, but differ significantly in the number and cysteine spacing pattern in the mature sequence (3, 56, 62, 63). Such α-defensin-related sequences have also been discovered in mice (21, 23) and rats (36). Nevertheless, the current study represents the most comprehensive attempt to identify mammalian canonical β-defensin gene family members with the characteristic six-cysteine motif.

Evolutionary relationships of mammalian β-defensins

β-defensins are the most ancient family of vertebrate defensins as compared to α- and θ-defensins (57). β-defensin-like sequences have been found in the rattlesnake (34), orange-spotted grouper (GenBank accession number AY129305), and zebrafish (Zhang, G., unpublished results), suggesting that the primordial gene for β-defensins appeared prior to the fish-primate split about 450 million years ago (27). In contrast, α-defensins were derived from β-defensins after divergence of glires and rodents from other mammals about 91 million years ago, whereas θ-defensins did not appear until primates separated from other mammals around 23 million years ago (27, 33, 36).

Our earlier study discovered a single cluster consisting of 13 β-defensin genes encoded in the entire chicken genome (57). Comparative analysis of the chicken and mammalian β-defensin gene clusters revealed that two clusters on human 8p23.1 and their orthologous loci in other mammalian species (Fig. 3A and 3B) are syntenic with the chicken gene cluster, implying that the ancestral genes in these two mammalian clusters evolved prior to the bird-mammal

separation (57). Many genes in these two ancient clusters were demonstrated to have undergone significant repeated duplication and positive diversification in humans and mice after their divergence (32, 48), which further gave rise to other β-defensin clusters in mammals during evolution, presumably as a result of chromosomal translocation and expansion of certain gene lineages.

Consistent with the evolutionarily active nature of the two ancient clusters, subgroups of genes exist in these two regions that are specific in rodents, dogs and/or primates (Fig. 5). This notion is further reinforced by the presence of multiple highly homologous α- and β-defensins and emergence of α- and θ-defensins in rodents and primates in these regions (Fig. 3A). In contrast, the remaining genomic clusters are rather static throughout evolution with the presence of orthologs across mammalian species, except that *CBD141* and *CBD142* on chromosome 24 are specific to dogs and was likely duplicated and diverged from adjacent genes (Fig. 3D) after the dog diverged from other mammals.

Collectively, these results suggest that the evolution of mammalian β-defensins is an extremely dynamic and active process. Individual gene lineages were derived at different evolutionary times. Although some β-defensins evolved prior to divergence of mammals from fish, many appear to be mammal-specific with a few emerging as recently as the rat-mouse split, which occurred only 41 million years ago (27). However, no gene duplication occurred prior to the separation of chimpanzee from humans at about 5.5 million years ago (27), because of the complete orthology of β-defensin genes in these two species (Supplementary Fig. 4).

What is the driving force for the presence of such a large number of divergent β-defensins in mammals? One plausible explanation is that sequence diversification may confer functional novelties on different β-defensins for each species to better cope with a variety of microbial threats from the environment. Consistent with this hypothesis, different β-defensins (15, 16, 18) and even orthologous genes in different primates (2) have been found to differ quite dramatically in their antimicrobial potency and spectrum. Furthermore, expression of a large array of β-defensins in the male reproductive tract is presumably to safeguard the reproductive process, enhance fertility, and sustain species survival, as detailed in the next section.

Dual role of region-specific expression of β-defensins in the male reproductive tract

Obviously, optimal reproductive function ensures survival of the species, which needs full protection of germ cell precursors in testis, sperm in different stages of maturation in epididymis and after deposition in the female reproductive tract. Infection of testis and epididymis may result in temporary or permanent loss of fertility by disrupting specialized environment of these organs conducive for sperm storage and maturation (37). As adaptive immunity is largely absent in the male reproductive system, the host must evolve alternative and effective mechanisms of protection of sperm from infectious agents.

Region-specific expression of all but one rat β-defensins in the male reproductive system (Fig. 6) suggests that these molecules may constitute an essential component in maintaining the normal reproductive process. In line with this, many testis- and epididymis-specific β-defensins were shown to be antimicrobial and capable of protecting sperm from infections (3, 8, 29, 56, 62, 64). In addition to being microbicides, β-defensins expressed in different regions of epididymis

124

appear to be actively involved in the reproduction. These defensins are mainly produced by epithelial cells lining the epididymal duct, regulated by androgens, secreted in luminal plasma, and bind preferentially to the surface of maturing but not immature sperm (29, 30, 41, 65, 68, 70). Rat Spag11e/Bin-1b has been shown to enhance sperm maturation by inducing Ca^{2+} uptake and subsequent motility and progressive movement of immature spermatozoa (70). Moreover, another epididymis-specific β-defensin, macaque DEFB126/ESP13.2, coats the entire ejaculated sperm and masks zona pellucida ligands on the sperm surface, but becomes dissociated when sperm are fully capacitated, suggesting that DEFB126 may be an important decapacitation factor on the sperm surface that needs to be removed prior to sperm-zona interaction and fertilization (54, 65).

Besides β-defensins, α-defensins have also been reported in the male reproductive tract (8, 17). Human cathelicidin LL-37/hCAP-18, which belongs to another important family of vertebrate antimicrobial peptides (40, 67), is expressed by epididymal epithelia and is abundantly present in the seminal fluid with ability to bind to sperm (31). Furthermore, the female reproductive tract also produces α- and β-defensins albeit at low levels (26, 38, 50). Coupled with the fact that many β-defensins in epididymis and testis are developmentally regulated with elevated expression at the time of puberty and sexual maturation (Fig. 7), these results strongly favor the argument that β-defensins, perhaps together with many other antimicrobial peptides, have a dual function in both innate host defense and fertility.

Defb4/RBD-2 represents the only β-defensin that is expressed most abundantly in lung with virtually no expression in the reproductive tract (Fig. 6B), and thus it may play an important

role in airway defense. Its human ortholog, DEFB4/hBD-2 was shown to be the predominant β-defensin in human neonatal lung and developmentally regulated (49) and its concentration is inversely correlated with the severity of lung disease in cystic fibrosis patients (7). Further studies are needed to identify the regulatory mechanisms that specify the expression of defensins in the male reproductive tract vs. respiratory tract. Functional divergence of β-defensins produced in the same and different regions of the reproductive tract also remains to be investigated and these molecules may have potential as fertility and contraceptive drugs.

ACKNOWLEDGEMENTS

GRANTS

REFERENCES

1. **Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ**. Basic local alignment search tool. *J Mol Biol* 215: 403-410, 1990.

2. **Antcheva N, Boniotto M, Zelezetsky I, Pacor S, Falzacappa MV, Crovella S, and Tossi A**. Effects of positively selected sequence variations in human and Macaca fascicularis beta-defensins 2 on antimicrobial activity. *Antimicrob Agents Chemother* 48: 685-688, 2004.

3. **Avellar MC, Honda L, Hamil KG, Yenugu S, Grossman G, Petrusz P, French FS, and Hall SH**. Differential expression and antibacterial activity of epididymis protein 2 isoforms in the male reproductive tract of human and rhesus monkey (Macaca mulatta). *Biol Reprod* 71: 1453-1460, 2004.

4. **Birney E, Clamp M, and Durbin R**. GeneWise and Genomewise. *Genome Res* 14: 988-995, 2004.

5. **Burge C, and Karlin S**. Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268: 78-94, 1997.

6. **Campopiano DJ, Clarke DJ, Polfer NC, Barran PE, Langley RJ, Govan JR, Maxwell A, and Dorin JR**. Structure-activity relationships in defensin dimers: a novel link between beta-defensin tertiary structure and antimicrobial activity. *J Biol Chem* 279: 48671-48679, 2004.

7. **Chen CI, Schaller-Bals S, Paul KP, Wahn U, and Bals R**. Beta-defensins and LL-37 in bronchoalveolar lavage fluid of patients with cystic fibrosis. *J Cyst Fibros* 3: 45-50, 2004.

8. **Com E, Bourgeon F, Evrard B, Ganz T, Colleu D, Jegou B, and Pineau C**. Expression of antimicrobial defensins in the male reproductive tract of rats, mice, and humans. *Biol Reprod* 68: 95-104, 2003.

9. **Crovella S, Antcheva N, Zelezetsky I, Boniotto M, Pacor S, Falzacappa MV, and Tossi A**. Primate beta-defensins - structure, function and evolution. *Curr Protein Pept Sci* 6: 7-21, 2005.

10. **Diamond G, and Bevins CL**. beta-Defensins: endogenous antibiotics of the innate host defense response. *Clin Immunol Immunopathol* 88: 221-225, 1998.

11. **Frohlich O, Ibrahim NM, and Young LG**. EP2 splicing variants in rhesus monkey (Macaca mulatta) epididymis. *Biol Reprod* 69: 294-300, 2003.

12. **Frohlich O, Po C, and Young LG**. Organization of the human gene encoding the epididymis-specific EP2 protein variants and its relationship to defensin genes. *Biol Reprod* 64: 1072-1079, 2001.

13. **Froy O, and Gurevitz M**. Arthropod and mollusk defensins--evolution by exon-shuffling. *Trends Genet* 19: 684-687, 2003.

14. **Ganz T**. Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 3: 710-720, 2003.

15. **Garcia JR, Jaumann F, Schulz S, Krause A, Rodriguez-Jimenez J, Forssmann U, Adermann K, Kluver E, Vogelmeier C, Becker D, Hedrich R, Forssmann WG, and Bals R**. Identification of a novel, multifunctional beta-defensin (human beta- defensin 3) with specific antimicrobial activity. Its interaction with plasma membranes of Xenopus oocytes and the induction of macrophage chemoattraction. *Cell Tissue Res* 306: 257-264, 2001.

16. **Garcia JR, Krause A, Schulz S, Rodriguez-Jimenez FJ, Kluver E, Adermann K, Forssmann U, Frimpong-Boateng A, Bals R, and Forssmann WG**. Human beta-defensin 4: a novel inducible peptide with a specific salt- sensitive spectrum of antimicrobial activity. *FASEB J* 15: 1819-1821, 2001.

17. **Grandjean V, Vincent S, Martin L, Rassoulzadegan M, and Cuzin F**. Antimicrobial protection of the mouse testis: synthesis of defensins of the cryptdin family. *BiolReprod* 57: 1115-1122, 1997.

18. **Harder J, Bartels J, Christophers E, and Schroder JM**. Isolation and characterization of human beta -defensin-3, a novel human inducible peptide antibiotic. *J Biol Chem* 276: 5707-5713, 2001.

19. **Hollox EJ, Armour JA, and Barber JC**. Extensive normal copy number variation of a beta-defensin antimicrobial gene cluster. *Am J Hum Genet* 73: 591-600, 2003.

20. **Hoover DM, Rajashankar KR, Blumenthal R, Puri A, Oppenheim JJ, Chertov O, and Lubkowski J**. The structure of human beta-defensin-2 shows evidence of higher order oligomerization. *J Biol Chem* 275: 32911-32918, 2000.

21. **Hornef MW, Putsep K, Karlsson J, Refai E, and Andersson M**. Increased diversity of intestinal antimicrobial peptides by covalent dimer formation. *Nat Immunol* 5: 836-843, 2004.

22. **Hughes AL**. Evolutionary diversification of the mammalian defensins. *Cell Mol Life Sci* 56: 94-103, 1999.

23. **Huttner KM, and Ouellette AJ**. A family of defensin-like genes codes for diverse cysteine-rich peptides in mouse Paneth cells. *Genomics* 24: 99-109, 1994.

24. **Jia HP, Mills JN, Barahmand-Pour F, Nishimura D, Mallampali RK, Wang G, Wiles K, Tack BF, Bevins CL, and McCray PB, Jr.** Molecular cloning and characterization of rat genes encoding homologues of human beta-defensins. *Infect Immun* 67: 4827-4833, 1999.

25. **Kent WJ**. BLAT--the BLAST-like alignment tool. *Genome Res* 12: 656-664, 2002.

26. **King AE, Critchley HO, and Kelly RW**. Innate immune defences in the human endometrium. *Reprod Biol Endocrinol* 1: 116, 2003.

27. **Kumar S, and Hedges SB**. A molecular timescale for vertebrate evolution. *Nature* 392: 917-920, 1998.

28. **Lehrer RI**. Primate defensins. *Nat Rev Microbiol* 2: 727-738, 2004.

29. **Li P, Chan HC, He B, So SC, Chung YW, Shang Q, Zhang YD, and Zhang YL**. An antimicrobial peptide gene found in the male reproductive system of rats. *Science* 291: 1783-1785, 2001.

30. **Liu Q, Hamil KG, Sivashanmugam P, Grossman G, Soundararajan R, Rao AJ, Richardson RT, Zhang YL, O'Rand MG, Petrusz P, French FS, and Hall SH**. Primate epididymis-specific proteins: characterization of ESC42, a novel protein containing a trefoil-like motif in monkey and human. *Endocrinology* 142: 4529-4539, 2001.

31. **Malm J, Sorensen O, Persson T, Frohm-Nilsson M, Johansson B, Bjartell A, Lilja H, Stahle-Backdahl M, Borregaard N, and Egesten A**. The human cationic antimicrobial protein (hCAP-18) is expressed in the epithelium of human epididymis, is present in seminal plasma at high concentrations, and is attached to spermatozoa. *Infect Immun* 68: 4297-4302, 2000.

32. **Morrison GM, Semple CA, Kilanowski FM, Hill RE, and Dorin JR**. Signal sequence conservation and mature peptide divergence within subgroups of the murine beta-defensin gene family. *Mol Biol Evol* 20: 460-470, 2003.

33. **Nguyen TX, Cole AM, and Lehrer RI**. Evolution of primate theta-defensins: a serpentine path to a sweet tooth. *Peptides* 24: 1647-1654, 2003.

34. **Nicastro G, Franzoni L, de Chiara C, Mancin AC, Giglio JR, and Spisni A**. Solution structure of crotamine, a Na+ channel affecting toxin from Crotalus durissus terrificus venom. *Eur J Biochem* 270: 1969-1979, 2003.

35. **Ouellette AJ**. Defensin-mediated innate immunity in the small intestine. *Best Pract Res Clin Gastroenterol* 18: 405-419, 2004.

36. **Patil A, Hughes AL, and Zhang G**. Rapid evolution and diversification of mammalian alpha-defensins as revealed by comparative analysis of rodent and primate genes. *Physiol Genomics* 20: 1-11, 2004.

37. **Purvis K, and Christiansen E**. Infection in the male reproductive tract. Impact, diagnosis and treatment in relation to male infertility. *Int J Androl* 16: 1-13, 1993.

38. **Quayle AJ, Porter EM, Nussbaum AA, Wang YM, Brabec C, Yip KP, and Mok SC**. Gene expression, immunolocalization, and secretion of human defensin-5 in human female reproductive tract. *Am J Pathol* 152: 1247-1258, 1998.

39. **Radhakrishnan Y, Hamil KG, Yenugu S, Young SL, French FS, and Hall SH**. Identification, characterization, and evolution of a primate beta-defensin gene cluster. *Genes Immun* 6: 203-210, 2005.

40. **Ramanathan B, Davis EG, Ross CR, and Blecha F**. Cathelicidins: microbicidal activity, mechanisms of action, and roles in innate immunity. *Microbes Infect* 4: 361-372, 2002.

41. **Rao J, Herr JC, Reddi PP, Wolkowicz MJ, Bush LA, Sherman NE, Black M, and Flickinger CJ**. Cloning and characterization of a novel sperm-associated isoantigen (E-3) with defensin- and lectin-like motifs expressed in rat epididymis. *Biol Reprod* 68: 290-301, 2003.

42. **Rodriguez-Jimenez FJ, Krause A, Schulz S, Forssmann WG, Conejo-Garcia JR, Schreeb R, and Motzkus D**. Distribution of new human beta-defensin genes clustered on chromosome 20 in functionally different segments of epididymis. *Genomics* 81: 175-183, 2003.

43. **Saitou N, and Nei M**. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425, 1987.

44. **Sang Y, Ortega MT, Blecha F, Prakash O, and Melgarejo T**. Molecular cloning and characterization of three beta-defensins from canine testes. *Infect Immun* 73: 2611-2620, 2005.

45. **Schibli DJ, Hunter HN, Aseyev V, Starner TD, Wiencek JM, McCray PB, Jr., Tack BF, and Vogel HJ**. The solution structures of the human beta-defensins lead to a better understanding of the potent bactericidal activity of HBD3 against Staphylococcus aureus. *J Biol Chem* 277: 8279-8289, 2002.

46. **Schutte BC, and McCray PB, Jr.** Beta-defensins in lung host defense. *Annu Rev Physiol* 64: 709-748, 2002.

47. **Schutte BC, Mitros JP, Bartlett JA, Walters JD, Jia HP, Welsh MJ, Casavant TL, and McCray PB, Jr.** Discovery of five conserved beta -defensin gene clusters using a computational search strategy. *Proc Natl Acad Sci USA* 99: 2129-2133, 2002.

48. **Semple CA, Rolfe M, and Dorin JR**. Duplication and selection in the evolution of primate beta-defensin genes. *Genome Biol* 4: R31, 2003.

49. **Starner TD, Agerberth B, Gudmundsson GH, and McCray PB, Jr.** Expression and activity of beta-defensins and LL-37 in the developing human lung. *J Immunol* 174: 1608-1615, 2005.

50. **Svinarich DM, Wolf NA, Gomez R, Gonik B, and Romero R**. Detection of human defensin 5 in reproductive tissues. *Am J Obstet Gynecol* 176: 470-475, 1997.

51. **Taudien S, Galgoczy P, Huse K, Reichwald K, Schilhabel M, Szafranski K, Shimizu A, Asakawa S, Frankish A, Loncarevic IF, Shimizu N, Siddiqui R, and Platzer M**. Polymorphic segmental duplications at 8p23.1 challenge the determination of individual

defensin gene repertoires and the assembly of a contiguous human reference sequence. *BMC Genomics* 5: 92, 2004.

52. **Thomma BP, Cammue BP, and Thevissen K**. Plant defensins. *Planta* 216: 193-202, 2002.

53. **Thompson JD, Higgins DG, and Gibson TJ**. 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: 4673-4680, 1994.

54. **Tollner TL, Yudin AI, Treece CA, Overstreet JW, and Cherr GN**. Macaque sperm release ESP13.2 and PSP94 during capacitation: the absence of ESP13.2 is linked to sperm-zona recognition and binding. *Mol Reprod Dev* 69: 325-337, 2004.

55. **Torres AM, and Kuchel PW**. The beta-defensin-fold family of polypeptides. *Toxicon* 44: 581-588, 2004.

56. **von Horsten HH, Derr P, and Kirchhoff C**. Novel antimicrobial peptide of human epididymal duct origin. *Biol Reprod* 67: 804-813, 2002.

57. **Xiao Y, Hughes AL, Ando J, Matsuda Y, Cheng JF, Skinner-Noble D, and Zhang G**. A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. *BMC Genomics* 5: 56, 2004.

58. **Yamaguchi Y, Nagase T, Makita R, Fukuhara S, Tomita T, Tominaga T, Kurihara H, and Ouchi Y**. Identification of multiple novel epididymis-specific beta-defensin isoforms in humans and mice. *J Immunol* 169: 2516-2523, 2002.

59. **Yamamoto M, and Matsui Y**. Testis-specific expression of a novel mouse defensin-like gene, Tdl. *Mech Dev* 116: 217-221, 2002.

60. **Yang D, Biragyn A, Hoover DM, Lubkowski J, and Oppenheim JJ**. Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense. *Annu Rev Immunol* 22: 181-215, 2004.

61. **Yeh RF, Lim LP, and Burge CB**. Computational inference of homologous gene structures in the human genome. *Genome Res* 11: 803-816, 2001.

62. **Yenugu S, Hamil KG, Birse CE, Ruben SM, French FS, and Hall SH**. Antibacterial properties of the sperm-binding proteins and peptides of human epididymis 2 (HE2) family; salt sensitivity, structural dependence and their interaction with outer and cytoplasmic membranes of Escherichia coli. *Biochem J* 372: 473-483, 2003.

63. **Yenugu S, Hamil KG, French FS, and Hall SH**. Antimicrobial actions of the human epididymis 2 (HE2) protein isoforms, HE2alpha, HE2beta1 and HE2beta2. *Reprod Biol Endocrinol* 2: 61, 2004.

64. **Yenugu S, Hamil KG, Radhakrishnan Y, French FS, and Hall SH**. The androgen-regulated epididymal sperm-binding protein, human beta-defensin 118 (DEFB118) (formerly ESC42), is an antimicrobial beta-defensin. *Endocrinology* 145: 3165-3173, 2004.

65. **Yudin AI, Tollner TL, Li MW, Treece CA, Overstreet JW, and Cherr GN**. ESP13.2, a member of the beta-defensin family, is a macaque sperm surface-coating protein involved in the capacitation process. *Biol Reprod* 69: 1118-1128, 2003.

66. **Zaballos A, Villares R, Albar JP, Martinez AC, and Marquez G**. Identification on mouse chromosome 8 of new beta-defensin genes with regionally specific expression in the male reproductive organ. *J Biol Chem* 279: 12421-12426, 2004.

67. **Zanetti M**. Cathelicidins, multifunctional peptides of the innate immunity. *J Leukoc Biol* 75: 39-48, 2004.

68. **Zanich A, Pascall JC, and Jones R**. Secreted epididymal glycoprotein 2D6 that binds to the sperm's plasma membrane is a member of the beta-defensin superfamily of pore-forming glycopeptides. *Biol Reprod* 69: 1831-1842, 2003.

69. **Zhang G, Hiraiwa H, Yasue H, Wu H, Ross CR, Troyer D, and Blecha F**. Cloning and characterization of the gene for a new epithelial beta-defensin. Genomic structure, chromosomal localization, and evidence for its constitutive expression. *J Biol Chem* 274: 24031-24037, 1999.

70. **Zhou CX, Zhang YL, Xiao L, Zheng M, Leung KM, Chan MY, Lo PS, Tsang LL, Wong HY, Ho LS, Chung YW, and Chan HC**. An epididymis-specific beta-defensin is important for the initiation of sperm maturation. *Nat Cell Biol* 6: 458-464, 2004.

FIGURE LEGENDS

Fig. 1. Amino acid sequence alignment of rat β-defensins. Dashes are inserted to optimize the alignment and conserved residues are shaded. Internal exon sequences of *Defb12, Defb52*, and *Spag11c* have been deleted to maximize the alignment. Long carboxyl terminal tails of several β-defensins as well as an additional segment of amino terminal sequence preceding the canonical start codon of *Defb17* are not shown for the sake of clarity. The letter 'X' denotes the presence of premature stop codon in the pseudogene. All computationally predicted novel rat β-defensin sequences except for *Defb3, Defb5*, and *Defb16p* have been confirmed by sequencing of their respective RT-PCR products amplified from appropriate tissues. The failure to amplify *Defb3, Defb5,* and *Defb16p* is likely due to no or extremely low expression in healthy animals. Abbreviations: *Defb*, β-defensin; *SPAG11,* sperm-associated antigen 11; *p*, pseudogene.

Fig. 2. Identification of novel β-defensins in humans and mice. *A,* Amino acid sequence alignment of novel sequences with DEFB1/Defb1. Dashes are inserted to optimize the alignment and conserved residues are shaded. The letter 'X' denotes the presence of premature stop codon in the pseudogene. *B,* Phylogenetic tree of novel β-defensins in the human and mice demonstrating their conservation in other mammalian species, except for mouse Defb53 and Defb54p. The tree was constructed by the neighbor-joining method based on the proportion difference (*p*-distance) of aligned amino acid sites of full-length peptide sequences. 1000 bootstrap replicates were used to test the reliability of each branch. Numbers on the branches indicate the percentage of 1000 bootstrap samples supporting the branch. Only branches supported by a bootstrap value of at least 50% are shown. Abbreviations: *DEFB/Defb*, β-defensin; *CBD*, canine β-defensin; *mDefb*, mouse β-defensin; r*Defb*, rat β-defensin; *P/p*, pseudogene.

Fig. 3. Genomic organization of syntenic mammalian β-defensin gene clusters. Four different β-defensin clusters (*A*, *B*, *C*, and *D*) are present in the rat, mouse, and dog, whereas five clusters exist in the human and chimpanzee due to the split of the Cluster *D* in the rodents and dog into two separate loci in the primates. The position of each gene is represented by a vertical bar. The direction of transcription is indicated by the triangle. The orthologous genes are linked across species and paralogous genes are also connected with dashed lines. However, in panel *A*, murine *Defb13, Defb15,* and *Defb35* that are orthologous to *DEFB107/CBD107, DEFB106/CBD106/Defb34,* and *DEFB105/CBD105/Defb12* gene lineages, respectively, are not linked for simplicity. Mouse *Defb2, 9, 10*, and *11* are not orthologs of the rat genes with corresponding numbers, and therefore they are not linked. The gene mapping was derived from

the assembled genomes of human (NCBI Build 35 version1), rat (BCM version 3.1), mouse (NCBI Build 32), and dog (NCBI Build 1 Version 1), respectively. The genomic cluster containing mouse *Defb33, 51,* and *52* is not mapped on chromosome 8, and therefore, is placed within slanted lines based on the locations of their orthologs on rats and dogs. Note that an α-defensin gene cluster is located within a β-defensin locus in the human, mouse, and rat, but is absent in the dog genome (*A*). Abbreviations: *Tel*, telomere; *Cen*, centromere; *P/p*, pseudogene; *11e/11c*, SPAG11e/SPAG11c.

Fig. 4. Structural organizations of mammalian β-defensin genes. Although most β-defensin genes are composed of two exons (boxes) separated by an intron (lines) of variable length, certain genes in mice (*Defb21* and *Defb44*), rats (*Defb29* and *Defb51*), and dogs (*CBD102, CBD139, and CBD140*) contain additional one or two exons encoding 5'-untranslated regions. On the other hand, mouse *Defb30* consists of an extra 3'-untranslated region. An extra exon encoding the carboxyl terminal mature sequence exists in mouse *Defb28* and canine *CBD122*, whereas an internal extra pro-sequence is present in *SPAG11c /CSPAG11c/Spag11c* and *DEFB105/CBD105/Defb12* gene lineages. In addition to an extra internal exon, rat *Defb52* comprises a fourth exon encoding the 5'-untranslated region. Solid boxes refer to the coding exons, whereas the open boxes refer to the noncoding exons.

Fig. 5. Phylogenetic tree of mammalian β-defensins. Only β-defensins in the rat chromosome 16q12.5 and its syntetic loci in the human, dog, and mouse are shown. The tree was constructed by the neighbor-joining method based on the proportion difference (*p*-distance) of aligned amino acid sites of full-length peptide sequences. 1000 bootstrap replicates were used to test the

reliability of each branch. Numbers on the branches indicate the percentage of 1000 bootstrap samples supporting the branch. Only branches supported by a bootstrap value of at least 30% are shown. Abbreviations: *DEFB,* human β-defensin; *CBD*, canine β-defensin; *mDefb*, mouse β-defensin; r*Defb*, rat β-defensin; *SPAG11/spag11*, sperm-associated antigen 11; *CSPAG11*, canine sperm-associated antigen 11; *P/p*, pseudogene;

Fig. 6. Tissue expression patterns of rat β-defensins by RT-PCR. Lanes: 1, kidney; 2, testis; 3, caput of epididymis; 4, corpus of epididymis; 5, cauda of epididymis; and 6, vas deferens. *Defb12/35* and *Defb15/34* have identical expression patterns and are therefore combined in a single panel. Notice that most β-defensins are expressed preferentially in the male reproductive system with no or minimum expression in other tissues (*A*), except that *Defb4* and *Defb36*, which have more universal distribution patterns (*B*). See Materials and Methods for details. The number of PCR cycles was optimized for each gene, and the specificity of each PCR product was confirmed by sequencing. The house-keeping gene, *GAPDH*, was used for normalization of the template input.

Fig. 7. Developmental regulation pattern of rat β-defensins in epididymis (*A*) and testis (*B*) by RT-PCR. The signal intensity of each gene-specific PCR product was digitized and quantitated, followed by normalization against the *GAPDH* signal amplified from the same cDNA sample. The results are expressed as means ± SEM of 2-4 animals per data point.

```
Defb23    ------------MKLLFPLFASLVLQYQVES----EFMVV---------KKCLMGF-GKCKDS---CLPEETQ----VQNCKSK-K--CCMGPKVTELIKSYLRHEIPH
Spag11c   ------------MKRRLLPSFVSLLLAALLLPEPGPNYQI------VNCKKNE-GFCQKY---CNFMETQ----VGYCSKK-KEACCLHPF---------------
Defb41    ------------MRSHSFLSALFLLVMMMPR---GKAGVVP------GEKQCILLK-GVCKDVS--CTSTDDT----IGVCND-EKKCCRRWWVFDPYPTPVPKG--K
Defb42    ------------MRLYLLLSTLLFLLGLLPR----VRSGLGA------AETHCVNLQ-GICRRDI--CKLIEDE----IGACRR-RWKCCRLWWVLLPIPTPVIFSDYQ
Defb9     ------------MRTLCSLLLICCLLFSYDT----PVVGELKHLGLKTEFEQCQRIR-GYCLNTY--CRFPTSH----VGSCYPE-KRYCCKNIR---------------
Defb10    ------------MKTLCSLLLIGCLLFSYDT----PVVGELKHLGMTAETEWCRLFE-GFCHDKN--CPPPTSH----VGSCHPE-KRSCCKDRR---------------
Defb11    ------------MRTLCSLLLIGCLLFSYDT----PVVGFLRRS--VSGFQECHSKG-GYCYRYY--CPRPHRR----LGSCYPY-AANCCRRRR---------------
Defb2     ------------MRTLCSLLLIGCLLFSYAT----PVAGILGPLRIQTDYHRCLREK-GFCLNAV--CPRSTLF----VGTCFPY-KFYCCKFKR-------
Defb1     ------------MKTHYFLLVMLFFLFSQME----LGAGILTSLGRRTDQYRCLQNG-GFCLRSS--CPSHTKL----QGTCKPD-KPNCCRS---
Defb37    ------------MKISCFLLLVLSLSCFQIN----PFAVL-------DTRVCIEKR-NTCHILQ--CPLFRDV----VGTCFEG-IGKCCHKYF-----
Defb39    ------------MKISCFLLLVLSLSCFQIN----SVSGI--------DSVKCFQKN-NTCHTIR--CPYFQDE----VGTCYEG-RGKCCQKRLLSIRVPKKKV-----
Defb40    ------------MKISCFLMLVLFLSCFQMN----SGAGL-------DTMKCVRGK-NNCHMHR--CPWFFVL----ISTCYSG-KGSCCQKRRWFTRSHVNNV-----
Defb38    ------------MKISCFLLLVLSLYLFQVN----QATDQ-------DTAKCVQKK-NVCYYFE--CPWLSIS----VSTCYKG-KAKCCQKRY-----
Defb4     ------------MRIHYLLFSFLLVLLSPLS----AFTQSIN---NPITCLTKG-GVCWGP--CTGGFRQ----IGTCGLP-RVRCCKKK---------
Defb5     ------------MRIHYLLFSFLLVLLSPLS----VFTQSIN---NPVSCVTHG-GICWGR--CPGSFRQ----IGTCGLG-KVRCCKKK---------
Defb3     ------------MRIHYLLFSFLLVLLSPLS-----AFSEKVY------NAVSCMTNG-GICWLK---CSGTFRE----IGSCGTR-QLKCCKKK-----
Defb14    ------------MRIHYLLFVFLILFLVPAPG--DTFIPKSL------RRFFCRVRG-GRCAILN--CLGKEEQ----IGRCSNR-GQKCCRKKK-------
Defb27    ------------MKTAVLTMVLLLLSQVIPG--------SPEKCWKSF-GICREE---CLRKEKF----YIFCWD--GSLCCVKPKNVPQWSQSSE-----
Defb36    ------------MKLLLLTLAALLLLSQLTPG----------DAQKCWNLH-GKCRHR---CSRKESV----YVYCTN--GKMCCVKPKYQPKPKPWMF-----
Defb30    ------------MGSLQLILVLFVLLSDVPPVRSGVNMYIRQ----IYDTCWKLK-GHCRNV---CGKKEIF----HIFCGT-QFLCCIERKEMPVLFVK-----
Defb43    ------------MRLLSILGVLTLLS-ILP----LARSFL------ANQECFSEY-RHCRMK---CKANEYA----IRYCAD--WTICCRVKKREAKKKIMW-----
Defb21    ------------MRLLLMALPLLALLPQVIP--------DYS------AEKRCLNRL-GHCKRK---CKAGEMV----METCKY--FQVCCVLDDNDYKQKASITRTMEK
Defb13    ------------MRIFSLIVAGLVLLIQLHP-----AKGTLY------RRFLCKRMK-GRCETA---CLSFEKK----IGTCRADLTPLCCKEKKKH-------
Defb33    ------------MRLLFLLFLLLVCLAQKTS----GRKRNT------KFRQCEKMG-GICKYQKTHGCSIL----PAECKSR-YKHCCRL---
Defb16p   ------------MKFHLLFFILLFSITILTV---RSLIDIKY------RIERCAKMX-GMCKAF---CDVDEYA----FGYCIKW-RNQCCI---
Defb17    ------------MKFHLLFFILLFSITILTGKRSYPEYGSLD------LRKECRMSK-GHCKLQ---CSENEIR----IAFCIRP-GTHCCI---
Defb49    ------------MKLPVLFLLFCFLDLKAEMK--------DTLFCFLKK-GKCRHV---CMNVEKR----VGPCTKL-NANCCIFVRDMRAIIPEDQRTVSI
Defb25    ------------MAKWILLIVALLVLGHVPS-------GST------EPFKRCWNGQ-GACRTY---CTRQEKF----IHLCPD-ASLCCLSYSLKASPHSRAGGV---
Defb12    --------MALSRGTFYFGLALFFIVVELPS--------GEFA------VCETCRLGR-GKCRRT---CLDSEKI----AGKCKLN-FFCCRERI---
Defb28    MLLTRSSTLSGHIRLWFLTLAVLVVLAQTSP-----EGWF------RTCFYGM-GKCRHV---CRANEKK----KERCGEN-SFCCLGETKSKLSNIPTNKGRKK
Defb24    ------------MKLVLLLLAIFVITELVMS-------GKNP------TLQCMGNR-GFCRPS---CKKGEQA----YFYCRTY-QICCLQSHVRISLTGVEDNTNWS
Spag11e   ------------MKVLLLFAVFFCLVQRNSGD--IPPGIR------NTVCFMQR-GHCRLFM--CRSGERK----GDICSDP-WNRCCVSSSIKNR-----------
Defb20    ------------MKLPQLLLILFVVLADSVQ----------PKRCFSNVAGYCRKR----CRLVEIS----EMGCLHG-KYCCVNELENKRHKKDTVVEQPM
Defb22    ------------MKSLLSALMIIMFLAHLVT-------GNWY------VRRCANKL-GTCRKT---CRKGEYQTDPATGKCSIG-KLCCILDLKLAGQCGGADGNQAA
Defb52    ------------MKLLYLLISVVLLISQVMA--------DYG------HHYCDSPT-SVCLRRRIN-CTRMP-------GLCPGR-SFCCVRT---
Defb18    ---MQSAMRLFFIFLIFVFSVSCGPSAPQMKTRDVLER------THKCFLVG-GECKSE---CSSWEYE----YVFCYTG--PCCVMREYKRVEKFSNTPKYTT
Defb44    ----------MDLHLLCLLLFLVTSLPEG---YCVIGNSGV------SFKPCTSEG-GYCFFG---CKLGWIW----ITYCNN-IMSCCKKDTKHSLPQTKGV-----
Defb29    --------MPVTKPYFVTVAVLLILVDKTTGGLFGLRSGKRRE-----PWVSCELYQ-SGCRNA----CQKYEIQ----YLTCPKK-KKCCLKFPMKITRV-------
Defb15    MDTVLTLTALQVMKTFLFLFAVFFFLDPAKN-------AFF--------DERCSRIN-GRCTES----CLKNEEL----IALCQKN-LKCCVTVQPCGRDKGDELDEDSG
Defb19    ---------------------------------------YRC-------RKSCWVIR-GHCRKH---CRSGERV----KRPCSNG-DYCCIAKKINTVPQAPKNAFSRN
Defb26    -----------MLFPTLAFIICGLLTQVTK----AAWRR-----MPPKCWKDNLGRCRVR----CRGDERY----IYLCRNK-ANCCILLTLAEDQSAHPKPTSVH
Defb50    ------------MKTLHLLLLISGLLSVFVKGVG-SHPGTVH------VRFKCIPKIAAVFGDN----CPFYGNV----DGLCNDK-KSVCCMVPVRLDNI----------
Defb51    ------------MRIHAFLALLAIFQVLHAITS-----ALN------FQRPCYLRG-GICLKQGTPNCEPF------RGPCRAF--TVCCKIRS----------
```

Fig. 1

A

```
DEFB1     - - M R T S Y L L L F T L C L L L S E M A S G G N F L T G L G H R S D H Y - N C V S S - - - - G
DEFB133   - - M K I H I F L F V L F F F L V P I A T - R - - - - G V K C A V K D T Y - S C F I V - - - - R
DEFB134   - - M K P L L V V F V F L F L W D P V L A - - - - - - G I N S L S S E M H K K C Y - K - - - - N
DEFB135   M A T R S V L L A L V V L N L L F Y V P P G R S - - - G P N V Y I Q K I F A S C W R L - - - - Q
DEFB136   - - M N L C L S A L L F F L V I L L P S G - K - - - - G M F G N D G V K V R T C T S Q - - - - K
Defb1     - - M K T H Y F L L V M I C F L F S Q M E P G V G I L T S L G R R T D Q Y - K C L Q H - - - - G
Defb49    - - M K L P V - L F L L F C F L D L - - - - - - - - - - K A N M K D T Y - F C F I K - - - - K
Defb50    - - M K T L C F L L L T S G L L Y L M V K G - - - - V G S H P G T F H V R I K C M P K - - - - M
Defb51    - - M R L H G F F L A L F A V F Q V L H A - - - - - - - - I S S A L N F E R P C Y L R - - - - G
Defb52    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Y E H H Y - - C D S P - - - - T
Defb53    - - M R I H Y L L F A F L L V L L S P C A - - - - - - - - - - - V F I Y I P A C I Q I I S S I R
Defb54p   - - M R I H Y L L F A F L L V L L C P L A - - - - - - A F S X K N Q Q S S - K F L M R - - - - G

DEFB1     G Q C L Y S - - - A C P I F T K I Q G T C Y R G K A K C C K A - - - - - - - - - - - - - - - - -
DEFB133   G K C R H - - - - E C H D F E K P I G F C T K L N A N C Y M - - - - - - - - - - - - - - - - - -
DEFB134   G I C R L - - - - E C Y E S E M L V A Y C - M F Q L E C C V K G N P A P - - - - - - - - - - - -
DEFB135   G T C R P - - - - K C L K N E Q Y R I L C D T I H L - C C V N P K Y L P I L T G K - - - - - - -
DEFB136   A V C F F - - - - G C P P G Y R W I A F C H N I L S - C C K N M T R F Q P P Q A K D P W V H - -
Defb1     G F C L R S - - - S C P S N T K L Q G T C K P D K P N C C K S - - - - - - - - - - - - - - - - -
Defb49    G K C R H - - - - V C S N V E M R V V A C T K L S A N F C I I M P E V Q S M V S I K I S N K D N
Defb50    T A V F G D - - - N C S F Y S S M G D L C N N T K S V C C M V P V R M D N I - - - - - - - - - - -
Defb51    G I C L K Q G T P N C - - - E P F R G P C R A F T V - C C K I R S - - - - - - - - - - - - - - -
Defb52    T V C L R - R K T K C - - - I R M P G F C P G R S F - C C M R T M T - - - - - - - - - - - - - -
Defb53    G V C W G - - - - Q C V S G T R E I G M S G L R S M K C C I R K - - - - - - - - - - - - - - - -
Defb54p   G K C Q Y - - - - K C I G N I R Q I G I C P V S L L K C X K K N T K I P - - - - - - - - - - - -
```

B

Fig. 2

Fig. 3

Fig. 4

Fig. 5

143

Fig. 6

144
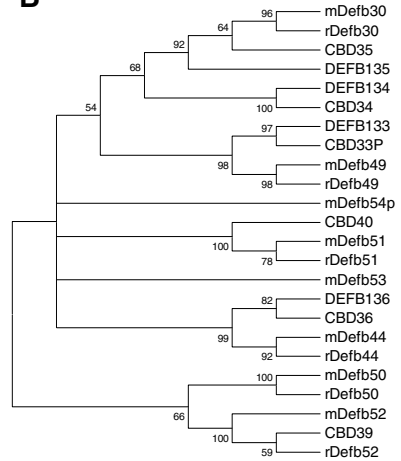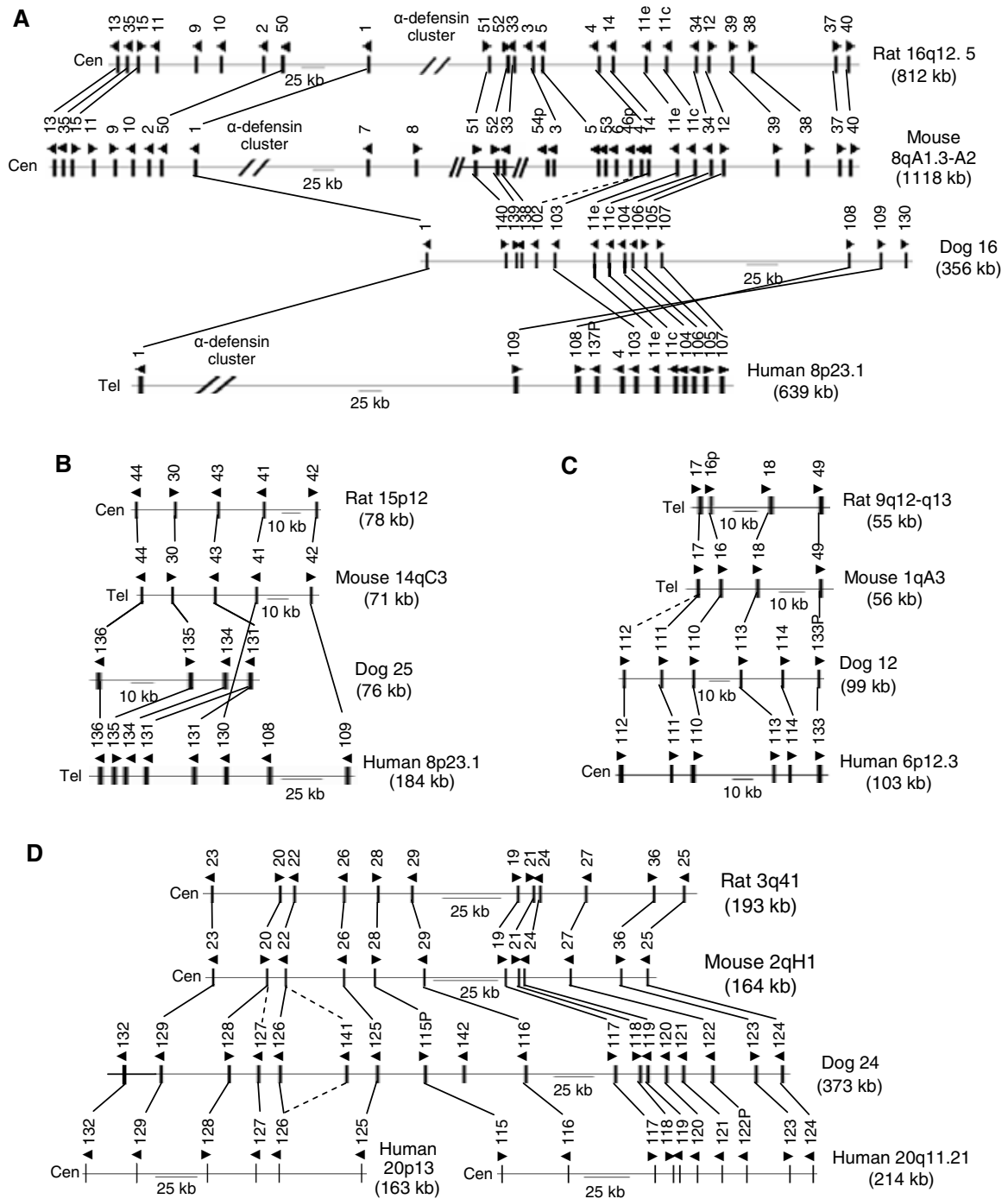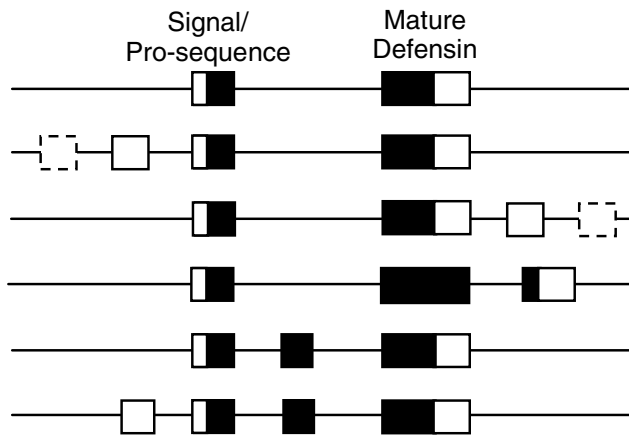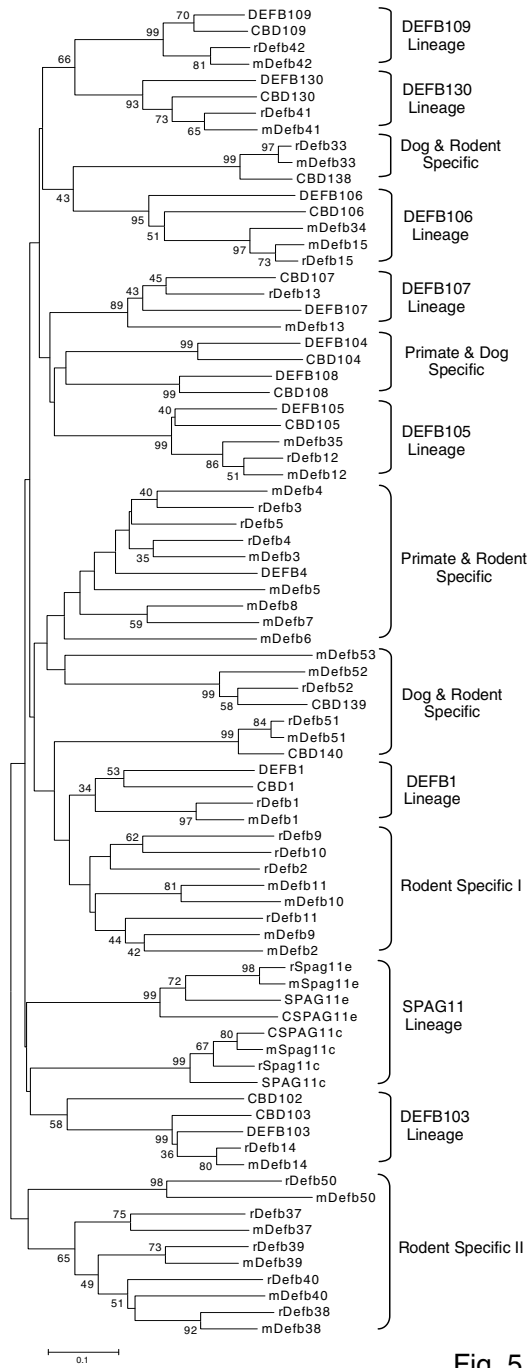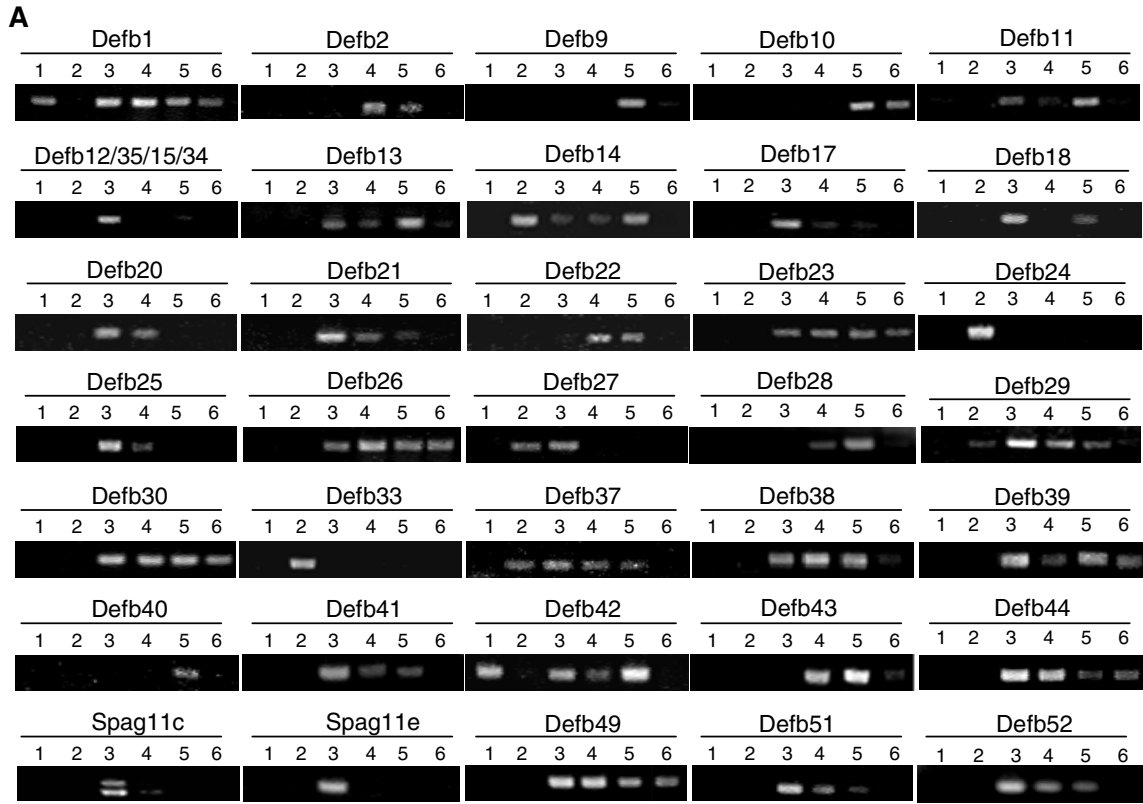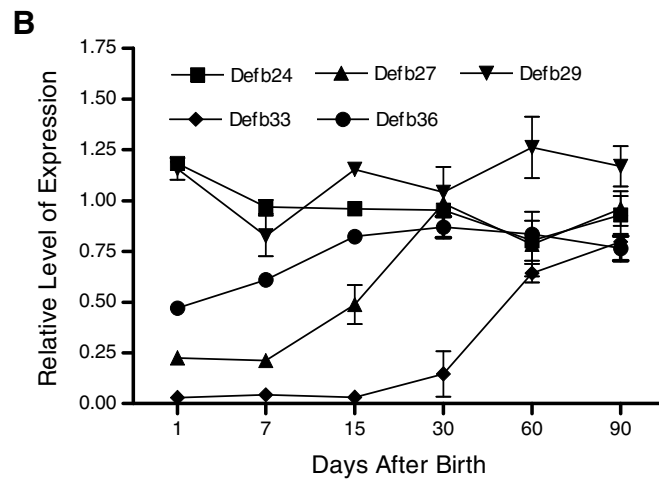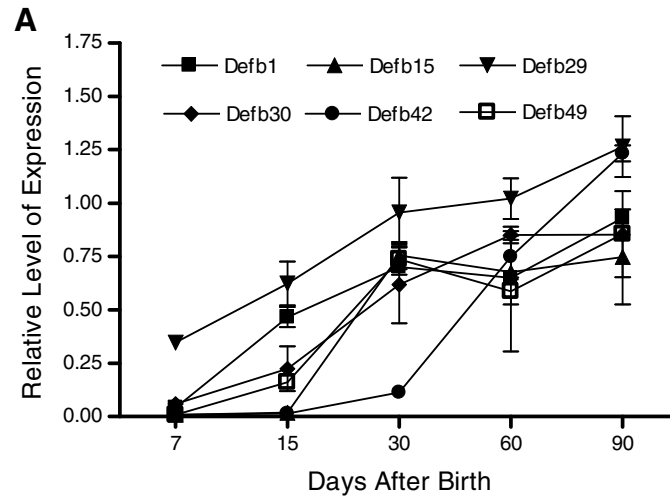
Fig. 7

Supplementary Fig. 1. Amino acid sequence alignment of the canine β-defensin gene family. Dashes are inserted to optimize the alignment and conserved residues are shaded. Internal exon sequences of *CBD105* and *CSPAG11c* as well as the last several amino acids of the first exon sequence of *CBD112* have been deleted to maximize the alignment. Long carboxyl terminal tails of several β-defensins are not shown for the sake of clarity. The letter 'X' denotes the presence of premature stop codon in the pseudogene. First exons for some β-defensins were not identified. Abbreviations: *CBD*, canine β-defensin; *CSPAG11*, canine sperm-associated antigen 11; *P*, pseudogene.

Supplementary Fig. 2. Amino acid sequence alignment of the human β-defensin gene family. Dashes are inserted to optimize the alignment and conserved residues are shaded. Internal exon sequences of *DEFB105* and *SPAG11c* as well as the last several amino acids of the first exon sequence of *DEFB112* have been deleted to maximize the alignment. Long carboxyl terminal tails of several β-defensins are not shown for the sake of clarity. The letter 'X' denotes the presence of premature stop codon in the pseudogene. First exons for some β-defensins were not identified. Abbreviations: *DEFB*, β-defensin; *SPAG11*, sperm-associated antigen 11; *P*, pseudogene.

Supplementary Fig. 3. Amino acid sequence alignment of the mouse β-defensin gene family. Dashes are inserted to optimize the alignment and conserved residues are shaded. Internal exon sequences of *Defb12, Defb52*, and *Spag11c* have been deleted to maximize the alignment. Long

146

carboxyl terminal tails of several β-defensins are not shown for the sake of clarity. The letter 'X' in bold denotes the presence of premature stop codon in the pseudogene. First exons for some β-defensins were not identified. Abbreviations: *Defb*, β-defensin; *p*, pseudogene.

Supplementary Fig. 4. Genomic organization of five syntenic β-defensin gene clusters in the human and chimpanzee. The position of each gene is represented by a vertical bar. The direction of transcription is indicated by the triangle. The orthologous genes are linked across species. The gene mapping was derived from the assembled genomes of human (NCBI Build 35 Version 1) and chimpanzee (NCBI Build 1 Version 1), respectively. It is noted that chimpanzee *DEFB130* and *DEFB131* are not conclusively mapped, whereas *DEFB110* and *DEFB128* are missing in the current genome assembly. Abbreviations: *Tel*, telomere; *Cen*, centromere; *P*, pseudogene; *11e/11c*, SPAG11e/SPAG11c.

```
CBD104    ----MR-TFLFLFAVFFL---LAPVTS----------EFETDKICGY--GTARCRRS----CKTEESSIG----ICPN-VY-ACCLKKMNTTH-------------------
CBD106    ----MR-TFLFLFAVFFL---LAPARN----------AFF-DEKCFK--LNGKCVHS----CQKNEELVA----LCQN-FL-KCCLTLQPCWMSKDDKSEDSGFLEGYK----
CBD105    MTLSRK-MSCFFFAFFFI---LAQFLSG---------EFAVCEPCWL--GRGKCRRI----CTEDEKVVG----NCKM-NF-FCCRRRIR----------------------
CBD102    ----MK-IHCLFFILPVL---FSSPTPGLSGR------VLFPLSCIG--SSGFCFPFR----CPHNREEIG----RCFFPIQ-KCCRRQK--------------------
CBD103    ----MR-IYYLLLLLPLL---FLMPVPGNGGI----INTLQRYYCRI--RSGRCALLS---CLPKEEQIG----RCSSTGR-KCCRRKK--------------------
CBD1      ----MRPLYLLLLLCLL---FSYLPPGAGFLTGI-GQRSDQYICAR--KGGTCNFSP---CPLFTRIDG----TCYRGKA-KCCMP-----------------------
CBD108    ----MR-PAGLLLALLFF---LSRVLAARGR-----F----KEICAH--PNGSCQEF----CLETEIQAG----RCLN-GH-ACCRPMVFESIIEPTTPKE-----------
CBD118    ----MR-LLFLTFAILVF---LPQVIPAYG---------GRKCWH--KSGYCRKT----CKADEVIKT----ACQH-YQ-VCCVPAQKFPITSPTHIYNFLFDITDTMVTS
CBD109    ----MR-LHVLF-FMLLF---LTLLSPVRSGL----AS--GENHCLN--LSGVCRRDI---CKMTEDQIG----ACKR-RW-VCCRAWWILIPIPTPLIYSDYQEPLKPKLK-
CBD130    ----MR-LHSLISVLLLF---VTIIPKARTGV----IP--GQKQCVL--LKGVCKDGG---CTTTDDTIG----ECND-EK-KCCRKWWVLFPYATPAPKAKSP---------
CBD138    ----MR-LLFLLFLLLVC---LVQMTSG-------REKRRKSLECER--MGGVCKHQKTHGCSILPA-------ECKSRNK-HCCRV-----------------------
CBD132    ----MR-LLLLVFASFGF---LVTPASG------------GGTSCGRK-YPGRCKLH----CHSQERSIF----ICDRYR--LCCMKNPIEGITASPVHRAKHPCRKTTESQN
CBD110    ----MK-IHLFFFILLFW---VTILPARSNFDPKY-----RFERCAK--VKGICKTF----CDDDEYDYG----YCIKWRN-QCCI-----------------------
CBD111    ----MK-IHLFFFILLFW---VTILPATKKY-PQY-GSLDLRRQCRK--GNGRCRLE----CHESEIRIA----FCMRPAT-HCCLQK---------------------
CBD112    ----MR-ILLSFIILSFG---VFIPPARSKE--HY-TVFNWWDACIK--LGSQCKNK----CGEKEFKMA----YCGRPTT-LCCMKECDHVELLKRSIHKKRTNIN------
CBD119    ----MR-FFLFFVILLAM---EPVISGKHH-----------ILRCMG--NTGICRPS----CRKTEQPYL----YCLNYQ--SCCLQS--YMRIS--ISGREEKDDWSQQNRW
CBD120    ----MR-FFLFFVILLAM---EPVIS--------------EECWM--N-GKCRLV----CKSDEDSVV----RCENRK--RCCVPSR-YLTVQPMTVERIEPETVPHTPKP
CBD128    ----MK-LFLVLIILQF----EVPTDTAR-----------PRKCFSN-ITGYCRKK----CKMGEIYEM----ACLNGK--LCCVNEDKNKKHQKAPESPLPLPQPKGKIDY
CBD134    ----MR-PLLVVLILLFFWDPVLAGLNPLSSEI--------HRKCYG---NG1CRLE----CFTSBMLVD----YCLFQL--ECCIKGNPEP--------------------
CBD122    ----MK-AFLLTLAALVL---LSQVTSGS-----------AEKCWN--LRGSCREK----CIKNEKLYI----FCTSGK--LCCLKPKFQPNMLQRNRKDNPKICLEPQKIL
CBD123    ----MK-LLWLTVAALLL---LTQLTPGG-----------TQRCWN--LHGKCRQK----CSRRERTYV----YCTNNK--LCCVKPKFQPRENLWPF-------------
CBD121    ----MK-LILLMLAVALL---LVQVTQ-------------AMSCWG--KLGRCRAT----CEKNEVFHI----LCTNEA--KCCVHPKHVPIGAGSSSAPESLG--------
CBD135    ----MR-SLLLVLVVLVL---LSCVPPVRSG-PNA-YIRRVFSTCWR--TKGVCKKS----CGKSEEYHI----LCDITR--LCCINKKYLPVQIGK--------------
CBD127    ----MK-LILIIAILLFQ---KSKVTEQ-----------LKRCWGEYIRGYCRKI----CRISEIREV----LCENGR--YCCLNIVELEARRKITKPPPPEPRTYAMTFP
CBD107    MSGAMR-IFFLVSAALIL---LAHIFSAHG-------AIHRRMQCQK--MDGRCEVE----CLSFEDKIG----GCRAELTPLCCKKRKNN--------------------
CBD113    ----MK-ILCIFLIFVLT---VSCGPSVSQKKTK--EDAGRKRECYL--VRGACKSS----CNSWEYIYN----YCSTEP---CCVVREYQKPVSKSINFTGDIL-------
CBD116    -MSVMK-PYLMTISFLLI---LVCKTQGDLFRSQYGKSQEPWIPCQL--DHGMCRNT----CRKNEIQYL----TCPNNQ--KCCLKLSVKIARSNNMKDYDPNSNLSVTNTS
CBD124    ----MTQLLLLIVALLA---LGHMQPGR---------SEFKRCWK--GQGACRTY----CTRQETYLH----MCPDAS--LCCLPYGSRPLSSKPEYV-----------
CSPAG11e  ----MKALFLFAVLFCLVQGNSGDMPPGI----------RDVICLT--QHGTCRLFF---CHFGERKAE----ICSDPWN-RCCLPSTEAEIKNKPETNDRNNKART-----
CBD114    ----MK-IFYYLLHFLCY---VTFILPAT-------CTLIDPDQCSN--KFGHCRRR----CFRGEKRID----ICFLPNK-ICCIERLYDED----------------
CBD131    ----MR-VLYSILGALAL---LSVVPQAR-------TFFLKDTCSL--EYLNCRMK----CNLDEHAIK----YCADWT--ICCKAKNTKFKRKKKW-------------
CBD133P   -------------------PPLTVKCAM----------KDTCSCFV--NRCKCRHK----CPVFEKTLG----FXAKLNA-KCCMFFVK--------------------
CBD141    ----MK-SLLFTFAIFFL---LIHLVSGH---------LFLRRCEN--KMGYCKKK----CISGELQKDYLKWRCDRHK--VCCVLATDNKIKPLSSSNKIFMPTVKNQQLS
CBD126    ----MK-SLLFTLSSFML---LTQLVSGN---------WYVRKCAN--KTGNCRSV----CRPGEMATVPPTGMCPKEK--MCCILSACGETTKTTNSAQASESSTTTPGPR
CSPAG11c  ---MKSPLSPLVSFLLVA--LMFPEPEPD---------FKIVNCKK--SEGYCQEY----CNYMETQVG----YCLKKKY-ACCLHQNVLV-----------------
CBD142    -----------------------RKTLKRFK--------TEELTKCYQ-GYGVCRER----CRKDETEVH----FCGNNR-RCCINSDSTILTRVVVKDIVWSKVITTSMPF
CBD140    ----MR-LLHVVLTAAFL---LSQVLPDSTG------AFNFERPCYL--RGGNCLKRGTPNC---EPFRG----PCRAFT--VCCKAKGSGRKRSSQGTLAPRAKARLPLSK-
CBD139    ----MK-LFSLLIFVVLF---ISQVMPGYR-------HGEVYHFCDS--QTSVCLRRKK-NC--MTRMPG----MCPGRS--FCCVRMK------------------------
CBD136    ----MR--LCISGLLFLL---LISLPSGNGLFGND---GVELRTCTA--LKGRCFFG----CRVGWTWVA----FCHNIL--SCCTKMKKNLPPQVKEP-------------
CBD129    ----MK-LLFTIFASLML---QYQVNTEYFG---------LRRCLM--GFGRCRDH----CAMAEKEIQ----KCKKKK---CCIGPKVVQMIKNYMQNEMSHTLEGSQEQL
CBD115P   -------------------------DGWVKRCSY--RIGRCRQS----RKKXEKKKE----KCGAKE--FCCIPDVKYIPSNPVEKQQTICEPKARPAKR
CBD125    ----MN-LMLTFIICGS---LNLVTKAG----------WETERCWKD-NIGHCRKR----CFHVERYKL----LCMNKL--ICCIPIKDDHGYTKMPPRLEPIKEEITTDVS
CBD117    -------------------------CRGQKSCWI--IKGHCRKN----CKSGEQVKK----PCKNGD--YCCILSKTDSQPHRLTQAPTGKSQILSTDLE
```

Supplementary Fig. 1

```
DEFB124   -----------MTQLLL-FLVALLVLGHVPSGRSE----------FKRCWKGQ---GACQTY-CTRQETYM----HLCPDASL--CCLSYALKPPPVPKHEYE------------------------
DEFB135   ----------MATRSVLL-ALVVLNLLFYVPPGRSGPNVYIQK--IFASCWRLQ---GTCRPK-CLKNEQYR----ILCDTIHL--CCVNPKYLPILTGK------------------------------
DEFB117   -----------------------------CR--SQKSCWIIK---GHCRKN-CKPGEQVK----KPCKNGDY--CCIPSNTDS-----------------------------------------------
DEFB118   -----------MKLLLL-ALPRLVLLP-QVIPA-------YS--GEKKCWNRS---GHCRKQ-CKDGEAVK----DTCKNLRA--CCIPSNEDHRRVPATSPTPLSDSTPGIIDDILTVRFTTDYFEVS
DEFB119   -----------MKLLYL-FLAILLAIEEPVISG-------KR--HILRCMGNS---GICRAS-CKKNEQFY----LYCRNCQS--CCLQSYMRISISGKEENTDWSYEKQWPRLP--------------
DEFB120   -----------MKLLYL-FLAILLAIEEPVIS--------VECWMD----GHCRLL-CKDGEDSI----IRCRNRKR--CCVPSRYLTIQPVTIHGILGWTTPQMSTTAPKMKTNITNR----
DEFB121   -----------MKLLLL-LLTVTLLLAQVTP------------VMKCWGKS---GRCRTT-CKESEVYY----ILCKTEAK--CCVDPKYVPVKPKLTDTNTSLESTSAV---------------
DEFB123   -----------MKLLLL-TLTVLLLLSQLTPG-----------GTQRCWNLY---GKCRYR-CSKKERVY----VYCINNKM--CCVKPKYQPKERWWPF----------------------------
DEFB116   --------MSVMKPCLM-TIAILMILAQKTPGGLFRSHNGKSREPWNPCELYQ---GMCRNA-CREYEIQY----LTCPNDQK--CCLKLSVKITSSKNVKEDYDSNSNLSVTNSSSYSHI--------
DEFB127   -----------MGLF---MIIAILLFQKPTVT-----------EQLKKCWNNYVQ-GHCRKI-CRVNEVPE----ALCENGRY--CCLNIKELEACKKITKPPRPKPATLALTLQDYVTIIENFPSLKT
DEFB128   -----------MKLF---LVLILLLFEVLTDG-----------ARLKKCFN-KVT-GYCRKK-CKVGERYE----IGCLSGKL--CCANDEEEKKHVSFKKPHQHSGEKLSVLQDYIILPTITIFTV--
DEFB122P  -----------------------------RIETCWNFR---GSCRDE-CLKNERVY----VFCVSGKL--CCLKPKDQPHLPQHIKN------------------------------
DEFB104   -----------MQRLVL-LLAVSLLLYQDLPVR-----SEFEL--DRICGYGT---ARCRKK-CRSQEYRI----GRCPNTYA--CCLRKWDESLLNRTKP-----------------------------
DEFB108   --MKPNFLLSSAMRIAVL-LFAIFFFMSQVLPAR----GKFK----EICERPN---GSCRDF-CLETEIHV----GRCLNSQP--CCLPLGHQPRIESTTPKKD---------------------------
DEFB105   -------MALIRKTFYF-LFAMFFIL-VQLPSG--EFAVC------ESCKLGR---GKCRKE-CLENEKPD----GNC-RLNFL-CCRQRI-----------------------------------------
DEFB133   -----------MKIHIF-LFVLFFFL-VPIATRGVKCAVKDT----YSCFIVR---GKCRHE-CHDFEKPI----GFCTKLNAN-CYM--------------------------------------------
DEFB110   -----------MKIQLF-FFILHFWV-TILPARSNFEPKYRF----ERCEKVR---GICKTF-CDDVEYDY----GYCIKWRSQ-CCV--------------------------------------------
DEFB111   -----------MKIQLF-FFILHFWV-TILPAKKKY-PEYGSLDLRRECRIGN---GQCKNQ-CHENEIRI----AYCIRPGTH-CCLQQ------------------------------------------
DEFB4     -----------MRVLYL-LFSFLFIFLMPLPG--VFGGIG--DPVTCLKSG---AICHPVFCPRRYKQI----GTCGLPGTK-CCKKP------------------------------------------
DEFB137P  -----------------------------XR----CVCVLNVCSTSLKQI----GTYGHDRIK-CCKK------------------------------------------------
DEFB1     -----------MRTSYL-LLFTLCLLLSEMASGGNFLTGLGHRSDHYNCVSSG---GOCLYSACPIFTKIQ----GTCYRGKAK-CCK-------------------------------------------
DEFB103   -----------MRIHYL-LFALLFLFLVPVPG----HGGIINTLQKYYCRVRG---GRCAVLSCLPKEEQI----GKCSTRGRK-CCRRKK-----------------------------------------
DEFB109   -----------MRLHLL-LLILLLFSILLSPV----RGGLGP--AEGHCLNLF---GVCRTDVCNIVEDQI----GACRR-RMK-CCRAWWILMPIPTPLIMSDYQEPLKPNLK---------------
DEFB130   -----------MKLHSL-ISVLLLFVTLIPKG----KTGVIP--GQKQCIALK---GVCRDKLCSTLDDTI----GICNE-GKK-CCRWWILEPYPTPVPKGKSP------------------------
SPAG11E   -----------MRVFFL-FAVLFCLVQTNSGD----VPPG---IRNTICRMQQ---GICRLFFCHSGEKKR----DICSDPWNR-CCVSNTDEEGKEKPEMDGRSGI-----------------------
DEFB131   -----------MRVLFF-VFGVLSLMFT------VPPGRSF-ISNDECPSEY---YHCRLK-CNADEHAI----RYCAD-FSI-CCKLKIIEIDGQKKW---------------------------
DEFB113   -----------MKILCI-FLTFVFTVSCGPSVPQKKTREVAE--RKRECQLVR---GACKPE-CNSWEYVY----YYC-NVNP--CCAVWEYQKPIINKITSKLHQK---------------------
DEFB115   MLPDHFSPLSGDIKLSVL-ALVVLVVLAQTAPDG----------WIRRCYYGT---GRCRKS-CKEIERKK----EKCGEKHI--CCVPKEKDKLSHIHDQKETSELYI--------------------
DEFB106   -----------MRTFLF-LFAVLFFLTPAKNAF-----------FDEKCNKLK---GTCKNN-CGKNEELI----ALCQKSLK--CCRTIQPCGSIID------------------------------
DEFB134   -----------MKPLLV-VFVFLFLWDPVLAGI-----NSLSSEMHKKCYKN----GICRLE-CYESEMLV----AYCMFQLE--CCVKGNPAP----------------------------------
DEFB114   -----------MRIFYY-LHFLCYVTFILPATC--------TLVNADRCTKRY---GRCKRD-CLESEKQI----DICSLPRKI-CCTEKLYEEDDMF-----------------------------
DEFB125   -----------MNILML-TFIICGLLTRVTKGS----------FEPQKCWKNNV--GHCRRR-CLDTERYI----LLCRNKLS--CCISIISHEYTRRPAFPVIHLEDITLDYSDVDSFTGSPVSMLND
DEFB132   -----------MKFLLL-VLAALGFLTQVIPAS-----------AGGSKCVSNTP-GYCRTC-CHWGETAL----FMCNASRK--CCISYSFLPKPDLPQLIGNHWQSRRRNTQRKDKKQQTTVTS---
DEFB126   -----------MKSLLF-TLAVFMLLAQLVSGN----------WYVKKCLNDV--GICRKK-CKPEEMHVKNGWAMCGKQRD--CCVPADRRANYPVFCVQTKTTRISTVTATTATTTLMMTTASMSS
DEFB129   -----------MKLLFP-IFASLMLQYQVNTEF----------IGLRRCLMGL---GRCRDH-CNVDEKEI----QKCKMKK---CCVGPKVVKLIKNYLQYGTPNVLNEDVQEMLKPAKNSSAVIQRK
DEFB107   -------MPGAMKIFVF-ILAALILLAQIFQAR----------TAIHRALISKRME--GHCEAE-CLTFEVKI----GGCRAELAPFCCKNRKKH---------------------------------
SPAG11C   -----------MRQRLLPSVTSLLLVALLFPEP--------ASDLKVVDCRRSE---GFCQEY-CNYMETQV----GYCSKKKDA-CCLH-----------------------------------------
DEFB112   -----------MKLLTT-ICRLKLEKMYSKTNT-RSEGHHITFSRWKSCTAIG---GRCKNQ-CDDSEFRI----SYCARPTTH-CCVTECDPTDPNNWIPKDSVGTQEWYPKDSRH-------------
DEFB136   -----------MNLCLSALLFFLVILLPSGKGM--FG--NDGVKVRTCTSQK---AVCFFG-CPPGYRWI----AFCHNILS--CCKNMTRFQPPQAKDPWVH-------------------------
```

Supplementary Fig. 2

Defb12   - - - - -MKNLPSNMALSREVFYFGFALFFIVVELP-SGEIA- - - - - - -VCETCRLG- - - -RGKCRR- - - -TCIESEK- - - -IAGWCKLNF-FCCRERI- -
Defb35   - - - - - - - - - - -MPQT- -FFVFCFLFFVFLQLFPGTGEIA- - - - - - -VCETCRLG- - - -RGKCRR- - - -ACIESEK- - - -IVGWCKLNF-FCCRERI- -
Defb27   - - - - - - - - - - -MK- - - -TAVLTMA-LLLLSQVI-PGTP- - - - - - - - - -ERCWKS- - - -FGVCRE- - - -ECAKKES- - - -FYIFCWNGK-LCCVKPKNVPLWSQNLD- - - - - - - - - - - - - - - - - - - -
Defb36   - - - - - - - - - - -MK- - - -LLLLTLAALLLVSQLT-PGDA- - - - - - - - - -QKCWNL- - -HGKCRH- - - -RCSRKES- - - -VYVYCTNGK-MCCVKPKYQPKPKPWMF- - - - - - - - - - - - - - - - - - - -
Defb30   - - - - - - - - - - -MG- - - -SLQLTLVLFVLLSYVP-PVRSGVNMYIKRIYDTCWKL- - - -KGICRN- - - -TCQKEEI- - - -YHIFCGIQS-LCCLEKKEMPVLFVK- - - - - - - - - - - - - - - - - -
Defb25   - - - - - - - - - - -MA- - - -KWILLIVALLVLSHVP-SGST- - - - - - - -EFKRCWNG- - - -QGACRT- - - -FCTRQET- - - -FMHLCPDAS-LCCLSYSFKPSRPSRVGDV- - - - - - - - - - - - - - - - - -
Defb23   - - - - - - - - - - -MK- - - -LLFPLFASLMLQSQVKSEFAT- - - - - - - - -VKRCLVG- - - -FGKCKD- - - -SCLADET- - - -QMQHCKAKK- -CCIGPFKVSELIKSYLRQEIPHIPDEDIVEMMRTKKNST
Spag11c - - - - - - - - - - - -MIP- - -RLLPFFASLLFAALLFPEPEPNY- - - - - -QIVNCKKS- - - -EGQCQE- - - -YCNFMET- - - -QVGYCSKKKEPCCLHPF- -
Defb21   - -MFYLPRAMR- - - - -FLLLVLPVLALLPQVTPDYGA- - - - - - - -KKRCLKI- - - -LGHCRR- - - -HCKDGEM- - - -DHGSCKYYR-VCCVPDLNSYNQNYAITWTTEETSTSEYDLSSDFIGVLF
Defb13   - - - - - - - - - - -MR- - - -IFSLIVAGLVLLIQLYPAWGTLY- - - - - -RRFLCKKM- - - -NGQCEA- - - -ECFTFEQ- - - -KIGTCQANF-LCCRKRKEH- -
Defb42   - - - - - - - - - - -MRLY-LLLSTLFFLLDLLPQVRSGMGAS- - - - - - -ETLCVSL- - - -QGTCRRD- - - -ICKLIED- - - -EIGACRRRW-KCCRLWWILIPIPTPVIFSDYQEPLQSKVK- - - - - - - - -
Defb41   - - - - - - - - - - -MRSR-SFLPVLFLFVMMMLRGKAGLVPG- - - - - - - - -EKQCISL- - - -KGICKDL- - - -ACTSSDD- - - -TIGVCNDVK-KCCRKWWVFEPYPTPVPKG- -KSP- - - - - - - - -
Defb17   - - - - - - - - - - -MK- - - -FHLFFFILLFGATILTAKKSYPEYGSLD-LRKECKMR- - - -RGHCKL- - - -QCSEKEL- - - -RISFCIRPGTHCCM- -
Defb16   - - - - - - - - - - -MK- - - -FHLFFFILLFGATILTARSHIDIKNGI- - - -ERCEKV- - - -RGMCKT- - - -VCDIDEY- - - -DYGYCIRWRNQCCI- -
Defb48p - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLSPLF-HLLAKKIN-SQLSCWMK- - - -XGICQY- - - -RCFGNTH- - - -KIDSCGAPFLKCCKR- -
Defb3   - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLSPP- -AAFSKKIN-NPVSCLRK- - - -GGRCWN- - - -RCIGNTR- - - -QIGSCGVPFLKCCKRK- -
Defb5   - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLSPLA-GVFSKTIN-NPVSCCMI- - - -GGICRY- - - -LCKGNIL- - - -QNGSCGVTSLNCCKRK- -
Defb54p - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLCPL- -AAFSXKNQ-QSSKFLMR- - - -GGKCQY- - - -KCIGNIR- - - -QIGICPVSLLKCXKKNTKIP- - - - - - - - - - - - - - - - - - - -
Defb8   - - - - - - - - - - -MR- - - -IHYLLFTFLL- -VLLSPL- -AAFSQKIN-DPVTYIRN- - - -GGICQY- - - -RCIGLRH- - - -KIGTCGSP-FKCCK- -
Defb7   - - - - - - - - - - -MR- - - -IHYVLFAFLL- -VLLSPF- -AAFSQDIN-SKRACYRE- - - -GGELQ- - - -RCIGLFH- - - -KIGTCNFR-FKCCKFQIPEKKTKIL- - - - - - - - - - - - - - - - - - -
Defb53   - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLSPC- -AVF-IYI- -P-ACIQISSIRGVCWG- - - -QCVSGTR- - - -EIGMSGLRSMKCCIRK- -
Defb6   - - - - - - - - - - -MK- - - -IHYLLFAFIL- -VMLSPL- -AAFSQLIN-SPVTCMSY- - - -GGSCQR- - - -SCNGGFR- - - -LGGHCGHPKIRCCRRK- -
Defb4   - - - - - - - - - - -MR- - - -IHYLLFTFLL- -VLLSPL- -AAFTQIIN-NPITCMTN- - - -GAICWG- - - -PCPTAFR- - - -QIGNCGHFKVRCCKIR- -
Defb46p - - - - - - - - - - -MR- - - -IHYLLFAFLL- -VLLSPY- -AG- - - - - - -ESLTCIAN- - - -RGFCWH- - - -SCIQGFQ- - - -LAGHCGHPKVRLLH- -
Defb14   - - - - - - - - - - -MK- - - -LHYLLFVFLI- -LFLVPAPGDAFLPKTL-RKFFCRIR- - - -GGRCAVL- - -NCLGKEE- - - -QIGRCSNSGRKCCRKKK- -
Defb34   - - - - - - - - - - -MK- - - -TFLFLFAVLF- -FWSQPRMHFFFF- - - - - -DEKCSRI- - - -NGRCTA- - - -SCLKNEE- - - -LVALCWKN-LKCCVTVQSCGRSKGNQSDEGSGHMGTRG- - - - - - - -
Defb15 MDTVLTLAALQVMK- - - -TFLFLFAVLF- -FLDPAKN- -AFF- - - - -DEKCSRV- - - -NGRCTA- - - -SCLKNEE- - - -LVALCQKN-LKCCVTVQPCGKSKSNQSDEGSGHMGTWG- - - - - - - -
Defb11   - - - - - - - - - - -MR- - -TLCSLLLICCLLFSYTTPAVGDLKHLILKAQLARCYKF- - - -GGFCYNS- - -MCPPHTK- - - -FIGNCHPDHLHCCINMKELEGST- - - - - - - - - - - - - - - - - - -
Defb10   - - - - - - - - - - -MR- - -TLCSLLLICCLLFSYTTPAVGDLKHLILKAQLTRCYKF- - - -GGFCHYN- - -ICPGNSR- - - -FMSNCHPENLRCCKNIKQF- - - - - - - - - - - - - - - - - - - -
Defb9   - - - - - - - - - - -MR- - -TLCSLLLICCLLFSYTTPAANSIIGV- - -SEMERCHKK- - - -GGYCY-F- - - -YCFSSHK- - - -KIGSCFPEWPRCCKNIK- - - - - - - - - - - - - - - - - - - -
Defb2   - - - - - - - - - - -MR- - -TLCSLLLICCLLFSYTTPAVGSLKSIGYEAELDHCHTN- - - -GGYCVRA- - -ICPPSAR- - - -RPGSCFPEKNPCCKYMK- - - - - - - - - - - - - - - - - - - -
Defb1   - - - - - - - - - - -MK- - -THYFLVMICFLFSQMEPGVGILTSLGRRTDQYRCLQH- - - -GGFCLRS- - - -SCPSNTK- - -LQGTCKPDKPNCCKS- -
Defb40   - - - - - - - - - - -MK- - - -ISCFLLMIFFLSCFQINPVAV- - - - - - -LDTIKCLQG- - - -NNNCHIQ- -KCPWFLL- - -QVSTCYKGKGRCCQRRWFARNHVYHV- - - - - - - - - - - - - - - -
Defb38   - - - - - - - - - - -MK- - - -ISCFLLLILSLYFFQINQAIG- - - - - - - -PDTKKCVQR- - - -KNACHYF- - - -ECPWLYY- - - -SVGTCYKGKGCCCQKR- - - - - - -Y- - - - - - - - - - - - - -
Defb39   - - - - - - - - - - -MK- - - -ISYFLLLILSLGSSQINPVSG- - - - - - - -DDSIQCFQK- - - -NNTCHTN- - -QCPYFQD- - -EIGTCYDRRGKCCQKRLLHIRVPRKKKV- - - - - - - - - - - - - - -
Defb37   - - - - - - - - - - -MK- - - -FSYFLLLLLSLSNFQNNPVAM- - - - - - -LDTIACIEN- - - -KDTCRLK- - -NCPRLHN- - -VVGTCYEGKGKCCHKN- - - - - - - - - - - - - - - - - - - -
Defb28 MLQTRSSTLSGHNK- - - -LWFLTLAVLVVLAQTSPEG- - - - - - - - - -WFRTCFYG- - -LGKCRR- - - -ICRANEK- - - -KKERCGERT-FCCLRETNSRLSHIPVTKDREKDESC- - - - - - - - - -
Defb22   - - - - - - - - - - -MK- - - -SLLSTLVIIMFLAHLVTGG- - - - - - - -WYVKKCANT- - - -LGNCRK- - -MCRDGEKQTEPATSKCPIGK-LCCVLDFKISGHCGGGGQNSDNLVTAGGDEGSSAKASTA
Defb33   - - - - - - - - - - -MR- - - -LLFLLFILLVCLAQTTSGRKRNS- - - - - -KFRPCEKM- - - -GGICKSQKTHGCSI- - - - - - -LPAECKSRYKHCCRL- - - - - - - - - - - - - - - - - - - -
Defb26   - - - - - - - - - - -MI- - - - - - -FLTLAFIVHGLLTQGNKAAWRR- - - - - -PIPKCWKNS- - -LGYCRV- - - -RCQEEER- - - -YIYLCKNKV-SCCIHRTLSEDKLPSGPPIHTVTRVIPSFNIPVNSSATF
Defb19   -RCSKACWVL- - - -RGHCRK- - - -HCRSGER- - - -VRKPCSNGD-YCCTASKMNHLPQAPKNALSRTGSVPVAVLKSDNK- - - -
Defb24   - - - - - - - - - - -MR- - - -LALLLLAILVATELVVSGKNP- - - - - - -ILQCMGN- - -RGFCRS- - - -SCKKSEQ- - - -AYFYCRTFQ-MCCLQSYVRISLTGVDDNTNWSYEKHWPRIP- - - - - - - -
Defb51   - - - - - - - - - - -MR- - - -LHGFFLALFAVFQVLHAISSALN- - - - - -FERPCYLR- - - -GGICLKQGTPNCEP- - - - - -FRGPCRAFT-VCCKIRS- -
Defb29   - - - -MPVTKSYFMT- - - - - - - -VVVVLILVDETTGGLFGFRSSKRQEPWIACELY- - -QGLCRN- - - -LCKYEI- - - -QYLSCPKTR-KCCLKYPRKITSF- - - - - - - - - - - - - - - - - - -
Defb20   - - - - - - - - - - -MK- - - - - -LLQVLLVLLFVALADGAQP- - - - - - - -KRCFSNV- - -EGYCRK- - - -KCRLVEIS- - -EMG-CLHGK-YCCVNELENKHKHKHSVVEETVKLQDKSKVQDYMILPTV
Defb43   - - - - - - - - - - -MR- - - -VLFSILGVLTLLSIVPLARSFLEN- - - - - -QDCSKH- - - -RHCRM- - - -KCKANEY- - - -AVRYCEDWT-ICCRVKKKESKKKKMW- - - - - - - - - - - - - - - - -
Spag11e - - - - - - - - - - -MK- - - - - -VLLLFAVFFCFVQGNSGDIPPG- - - - - -IRNTVCLMQ- - - -QGHCRLF- - - -MCRSGER- - - -KGDICSDPWNRCCVPYSVKDR- - - - - - - - - - - - - - - - -
Defb44   - - - - - - - - - - -MN- - -LRLSCLLLFILVTSLPAGRCSIGNKG- - -ISFETCTAI- - - -GLCFF- - - -GCKLGWV- - - -WIAYCNNIM-SCCRKDTDFVLPQTKGI- - - - - - - - - - - - - - - - -
Defb50   - - - - - - - - - - -MK- - -TLCFLLLTSGLLYLMVKGVGSHPGTFH- - -VRIKCMPK- - - -MTAVFGD- - -NCSFYSS- - -MGDLCNNTKSVCCMVPVRMDNI- - - - - - - - - - - - - - - - - - -
Defb49   - - -MKL- -PVLFLLFCFLDLKANMK- - - - - - - -DTYFCFIK- - - -KGKCRH- - - -VCSNVEM- - -RVVACTKLSANFCIIMPEVQSMVSIKISNKDN- - - - - - - - - - - - - - -
Defb18   - - - - - - -MQSTMK- - - - - -MFGIIILMVIFSVSCGPSAPQMKTREVAERTHKCSLV- - - -RGTCKS- - - -ECNSWEY- - -KYNYCHTEP- -CCVVREYKRMEKLLSTPKYTT- - - - - - - - -
Defb52   -YEHHYCDSP- - - -TTVCLRRKT-KCIR- - - - - - -MPGFCPGRS-FCCMRTMT- - - - - - - - - - - - - - - -

Supplementary Fig. 3
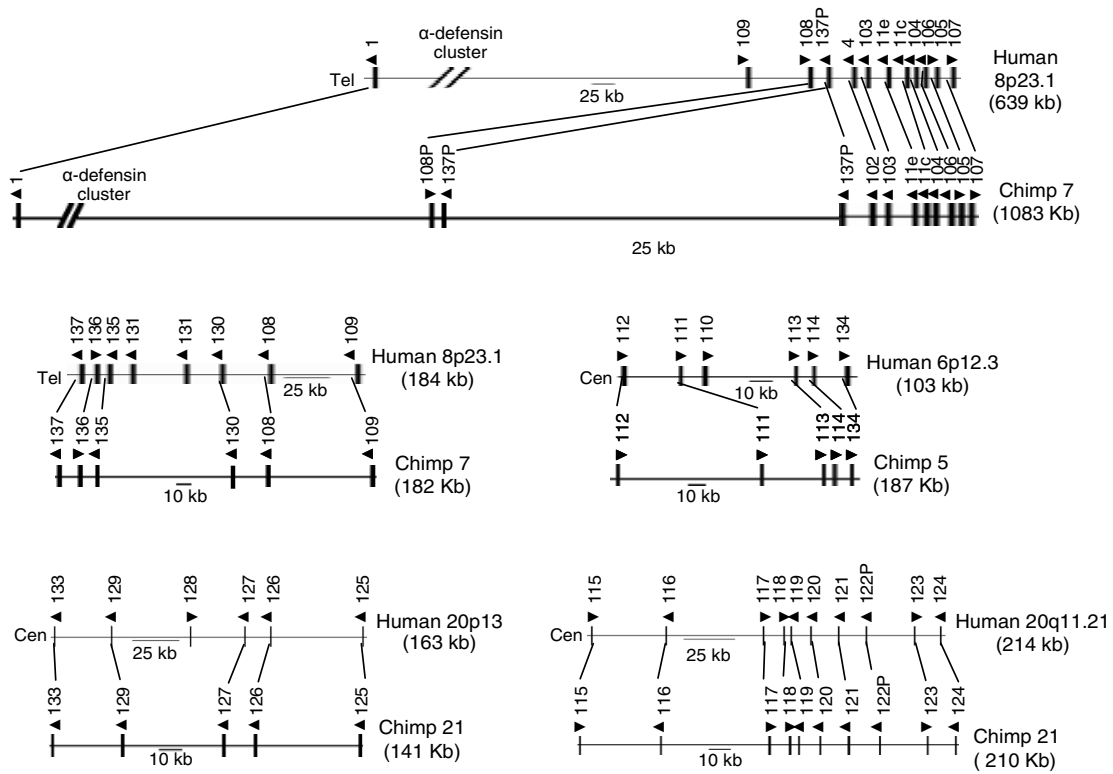
Supplementary Fig. 4

Supplementary Table 1. GenBank accession numbers of mammalian β-defensin cDNA sequences[a]

Human		Chimpanzee		Dog		Mouse		Rat	
β-defensin	Accession No.	β-defensin	Accession No.	β-defensin	Accession No.	β-defensin	Accession No.	β-defensin	Accession No.
DEFB1	*NM_005218*	*DEFB1*	*NM_001009151*	CBD1	DQ011970	*Defb1*	*NM_007843*	*Defb1*	*NM_031810*
DEFB4	*NM_004942*	*DEFB102*	*NM_001009076*	CBD102	DQ011971	*Defb2*	*NM_010030*	Defb2	AY621338
DEFB103	*NM_018661*	DEFB103	DQ012056	CBD103	DQ011972	*Defb3*	*NM_013756*	Defb3	AY621339
DEFB104	*NM_080389*	DEFB104	DQ012057	CBD104	DQ011973	*Defb4*	*NM_019728*	Defb4	AY621374
DEFB105	*NM_152250*	DEFB105	DQ012058	CBD105	DQ011974	*Defb5*	*NM_030734*	Defb5	AY621340
DEFB106	*NM_152251*	DEFB106	DQ012059	CBD106	DQ011975	*Defb6*	*NM_054074*	Defb9	AY621341
DEFB107	DQ119826	DEFB107	DQ012060	CBD107	DQ011976	*Defb7*	*NM_139220*	Defb10	AY621342
DEFB108	*NM_001002035*	DEFB108P	DQ012061	CBD108	DQ011977	*Defb8*	*NM_153108*	Defb11	AY621343
DEFB109	DQ012013	DEFB109	DQ012062	CBD109	DQ011978	*Defb9*	*NM_139219*	Defb12	AY621344
DEFB110	DQ012014	DEFB111	DQ012063	CBD110	DQ011979	*Defb10*	*NM_139225*	Defb13	AY621345
DEFB111	DQ012015	DEFB112	DQ012064	CBD111	DQ011980	*Defb11*	*NM_139221*	Defb14	AY621346
DEFB112	DQ012016	DEFB113	DQ012065	CBD112	DQ011981	*Defb12*	*NM_152802*	Defb15	AY621347
DEFB113	DQ012017	DEFB114	DQ012066	CBD113	DQ011982	*Defb13*	*NM_139223*	Defb16-ps	DQ012089
DEFB114	DQ012018	DEFB115	DQ012067	CBD114	DQ011983	*Defb14*	*NM_183026*	Defb17	AY621348
DEFB115	DQ012019	DEFB116	DQ012068	CBD115P	DQ011984	*Defb15*	*NM_139222*	Defb18	AY621349
DEFB116	DQ012020	DEFB117	DQ012069	CBD116	DQ011985	Defb16	DQ012027	Defb19	AY621350

DEFB117	DQ012021	DEFB118	DQ012070	CBD117	DQ011986	Defb17	DQ012028	Defb20	AY621351
DEFB118	*NM_054112*	DEFB119	DQ012071	CBD118	DQ011987	Defb18	DQ012029	Defb21	AY621352
DEFB119	*NM_153289*	DEFB120	DQ012072	CBD119	DQ011988	Defb19	DQ012030	Defb22	AY621353
DEFB120	*NM_153323*	DEFB121	DQ012073	CBD120	DQ011989	Defb20	DQ012031	Defb23	AY621354
DEFB121	*NM_001011878*	DEFB122P	DQ012074	CBD121	DQ011990	Defb21	DQ012032	Defb24	AY621355
DEFB122P	*AY122474*	DEFB123	DQ012075	CBD122	DQ011991	Defb22	DQ012033	Defb25	AY621356
DEFB123	*NM_153324*	DEFB124	DQ012076	CBD123	DQ011992	Defb23	DQ012034	Defb26	AY621357
DEFB124	DQ119827	DEFB125	DQ012077	CBD124	DQ011993	Defb24	DQ012035	Defb27	AY621358
DEFB125	*NM_153325*	DEFB126	DQ012078	CBD125	DQ011994	Defb25	DQ012036	Defb28	AY621359
DEFB126	*NM_030931*	DEFB127	DQ012079	CBD126	DQ011995	Defb26	DQ012037	Defb29	AY621360
DEFB127	*NM_139074*	DEFB129	DQ012080	CBD127	DQ011996	Defb27	DQ012038	Defb30	AY621361
DEFB128	*AF525930*	DEFB130	DQ012081	CBD128	DQ011997	Defb28	DQ012039	Defb33	AY621362
DEFB129	*NM_080831*	DEFB131	DQ012082	CBD129	DQ011998	*Defb29*	*AF515626*	Defb36	AY621363
DEFB130	DQ012022	DEFB132	DQ012083	CBD130	DQ011999	Defb30	DQ012040	Defb37	AY621364
DEFB131	*AF529414*	DEFB133	DQ012084	CBD131	DQ012000	Defb33	DQ012041	Defb38	AY621365
DEFB132	*NM_207469*	DEFB134	DQ012085	CBD132	DQ012001	*Defb34*	*NM_183035*	Defb9	AY621366
DEFB133	DQ012023	DEFB135	DQ012086	CBD133P	DQ012002	*Defb35*	*NM_139224*	Defb40	AY621367
DEFB134	DQ012024	DEFB136	DQ012087	CBD134	DQ012003	Defb36	DQ012042	Defb41	AY621368
DEFB135	DQ012025	DEFB137P	DQ012088	CBD135	DQ012004	*Defb37*	*NM_181683*	Defb42	AY621369
DEFB136	DQ012026	*SPAG11c*	*AF263553*	CBD136	DQ012005	*Defb38*	*NM_183036*	Defb43	AY621370

DEFB137P	DQ119828	*SPAG11e*	*AF263555*	CBD138	DQ012006	*Defb39*	*NM_183038*	Defb44	AY621371
SPAG11c	*NM_058203*			CBD139	DQ012007	*Defb40*	*NM_183039*	Defb49	AY621372
SPAG11e	*NM_058207*			CBD140	DQ012008	Defb41	DQ012043	Defb50	DQ012090
				CBD141	DQ012009	Defb42	DQ012044	Defb51	AY621373
				CBD142	DQ012010	Defb43	DQ012045	Defb52	DQ012091
				SPAG11c	DQ012011	Defb44	DQ012046	Spag11c	DQ012092
				SPAG11e	DQ012012	Defb46-ps	DQ012047	Spag11c-1	DQ012093
						Defb48-ps	DQ012048	*Spag11e*	*NM_145087*
						Defb49	DQ012049		
						Defb50	DQ012050		
						Defb51	DQ012051		
						Defb52	DQ012052		
						Defb53	DQ012053		
						Defb54-ps	DQ012054		
						Spag11c	DQ012055		
						Spag11e	*NM_153115*		

[a] The complete repertoires of β-defensins identified in the human, chimpanzee, dog, mouse, and rat are shown together with GenBank accession numbers for their mRNA sequences. Previously identified genes are italicized and listed with a reference accession number (*NM*) if available. The remaining genes were discovered in this study. *Spag11c-1* is an

alternatively spliced variant of *Spag11c* in the rat. Abbreviations: *DEFB/Defb*, β-defensin; *SPAG11*, sperm-associated

antigen 11; *P/ps*, pseudogene.

Supplementary Table 2. Primer sequences used for RT-PCR analysis of rat β-defensins[a]

Gene	Primer Sequence		Product Size (bp)		No. of
	Forward	Reverse	cDNA	Genomic	Cycles
Defb1	TGGACCCTGACTTCACCGACAT	CTTCAAACCACTGTCAACTCCTG	244	15238	32
Defb2	GACTCTCTGCTCTCTGCTGCT	CCCACAAACAAGGTAGAACGA	172	13785	40
Defb3	CTCTTCTCATTTCTCCTGGTGCT	GTGCCACAACTGCCAATCTCT	150	9905	40
Defb4	CGCATCCATTACCTTCTCTTCTC	ATCTCACTCTAGGCAGTCCACAA	173	2943	35
Defb5	CTCTTCTCATTTCTCCTGGTGCT	TCTCACTTTACCCAGTCCACAA	164	2125	40
Defb9	GACTCTCTGCTCTCTGCTGCTG	CACATGGGAGGTAGGAAAACGAC	163	2731	35
Defb10	GACTCTCTGCTCTCTGCTGCTG	GGTGACAGGAACCCACATGCGA	177	10013	35
Defb11	GACTCTCTGCTCTCTGCTGCTG	GTAACAGCTACCCAGTCTTCTG	169	6869	35
Defb12/35	AGCAGAGGGACGTTCTATTTTGG	CTCCCGACAGCAGAAGAAGTTC	219	2398/2413[b]	35
Defb13	CGGGCTTGTTCTTCTCATTCAAC	GTGTTAAGTCAGCTCGGCAGGT	143	1789	35
Defb14	TCATCTTGTTCTTGGTGCCTG	CTGCAGCATTTTTGACCTCTGT	160	814	35
Defb15/34	GACACCGTGCTGACCCTCACA	CATCCCCGAGTCCTGTTGTAG	272	2638/2631[b]	35
Defb16p	TGGTCATTGTTTCCCGACTAC	GCATACTCATCCACATCACAGAA	216	8567	40
Defb17	TGGTCATTGTTTCCCGACTAC	GCATTCTTTCCTCAAGTCCAAG	186	3422	35
Defb18	TTCTCTGTCCTGTGGTCCTTC	TATTCTCGCATCACACAGCAAG	173	819	35

Defb20	TCCTTATTTTGCTGTTTGTGGT	TTTCTTGTGCCTTTTGTTCTCC	181	2506	40
Defb21	GCTCTTCCTCTGCTTGCTCTT	GTTTGTAGTCGTTGTCGTCCAG	180	1235	40
Defb22	ATGAAGTCCCTATTGTCTGCTC	TTTCCCCCTGTTCCCTGTGTT	303	3922	40
Defb23	GCTTCAGTATCAGGTGGAATCA	GGTATGTGGGGTATTTCGTGTC	197	3915	35
Defb24	CCTGCTTCTTGCCATATTTGTG	TCTGGAAGTCTGTGCATGGCGT	268	6016	32
Defb25	GGCAAAGTGGATTCTGCTCATT	ATAGGACAGGCAGCACAGAGA	180	797	40
Defb26	CCTTCATTATCTGTGGGTTGCT	CCGAGGTAGGTTTTGGGTGTG	214	3224	35
Defb27	AACACCCTGGATTTCCCGAAGC	CTTCGTTTCTGTGGCCTTGGGA	258	4873	40
Defb28	ACATTAAGCTCTGGTTCCTGAC	ACTGTATGTGTCCCCGACTTG	285	1890	35
Defb29	TCTGCTGCTGTGATTTCTATCC	CTGGCGACATTCAAACCTGAG	359	1336	35/40/30
Defb30	GTTGATCCTTGTGCTCTTTGTC	ACGGGCATTTCCTTTCTTTC	221	1764	35
Defb33	GGCTGCTATTTCTTCTCTTTCTCC	GGCAGTGTTTGTATCGACTTTTGC	172	2473	35
Defb36	AGGCCCATCATGAAGCTCCTGC	TTCCGGCTTCCGGGCAGTTAG	227	13428	35/32
Defb37	TGCTTTCTGCTGCTGGTCCTCT	CAGGTGCCGACTACATCACGA	140	20543	40
Defb38	TGCTTTCTGCTGCTGGTCCTCT	CGCAGGTGCTAACAGAAATAGAC	145	3181	35
Defb39	TGCTTTCTGCTGCTGGTCCTCT	GGCTTTGGAACTCGGATAGACA	205	3712	35
Defb40	TGCTTTCTGCTGCTGGTCCTCT	GACTCCACGTAAACCAACGACG	201	2900	35
Defb41	GATCCCACTCCTTCCTGTCTG	CCCACCATCTCCTACAACATTTC	197	2471	35

Defb42	GTATCTGCTCCTCTCTACGCTAC	GCAGCATTTCCATCTTCTCC	180	6726	35
Defb43	GAGGCTGCTGCTTTCAATCC	CTTGGCTTCTCTCTTCTTCACC	201	5959	35
Defb44	ATGTTTACTGCTCTTCCTGGTG	ACACCTTTGGTTTGAGGAAGTG	219	1046	35
Spag11c	GACTCCTTCCTTCCTTTGTCAG	ATCCCTCATTTTTCTTGCAGTT	246/306	1852	35
Spag11e	CATCAGTCACATCTGCTTTCCTG	TTCCAGGTCTCTAGCTTCCAAG	315	1675	35
Defb49	CGAAGCTGCCTGTTCTTTTTC	ATGGCTCTCATGTCACGCACA	209	4608	35/30
Defb51	ATGAGAATCCACGCTTTCCTG	TACAGCAAACAGTGAAGGCTC	177	747	40
Defb52	CCTTCTCATCTCTGTTGTTCTCTT	TAATGGTGCCCGTAATCGTTG	147	1829	40
GAPDH	GTGAAGGTCGGAGTGAACG	GAGATGATGACCCTCTTGGC	356	>390	25

[a] RT-PCR reactions were performed with DNA Engine (Model PTC-200, MJ Research). Number of PCR cycles used for evaluating tissue and developmental expression patterns was the same for all relevant genes, except for *Defb29, Defb36*, and *Defb49*. For *Defb29* amplification, 35, 40, and 30 cycles were used for gene expression in the tissue panel, epididymis, and testis, respectively. For *Defb36* amplification, 35 cycles were used for the tissue panel with 32 cycles used for its developmental expression in testis. A total of 35 cycles were performed for tissue expression of *Defb49*, whereas 30 cycles were used for developmental expression in epididymis.

[b] Because of 100% identity in the coding sequences between *Defb12* and *Defb35* and between *Defb15* and *Defb34* in the rat, the same primers were used for each pair.

CHAPTER V

RATTUSIN, A NOVEL α-DEFENSIN-RELATED PEPTIDE WITH SALT-INSENSITIVE ANTIBACTERIAL ACTIVITY

ABSTRACT

Cationic antimicrobial peptides (CAMPs) are an essential component of the innate immune system. As a major family of CAMPs, defensins are expressed mainly by mucosal epithelial cells and neutrophils, which constantly come into direct contact with microbes. Defensins are capable of killing a broad spectrum of bacteria through physical disruption of membranes. However, all known defensins show a significantly reduced antibacterial activity in the presence of salt and/or divalent cations, thereby limiting their therapeutic potential for the treatment of systemic infections. A genome-wide computational screen of the entire rat genome led to the identification of a novel α-defensin-related gene, designated rattusin. Rattusin shares a highly conserved pro-sequence with mammalian α-defensins. However, unlike defensins with a canonical six-cysteine motif, rattusin consists of only five cysteines with a different spacing pattern. Furthermore, rattusin is preferentially expressed in distal small intestine with potent activities against a broad range of Gram-negative and Gram-positive bacteria, including antibiotic-resistant strains. In contrast to classical α-defensins, rattusin retains its activity in the presence of physiological concentrations of salt and divalent cations. To our knowledge, rattusin is the only defensin-related peptide with salt-insensitive antibacterial activities, making it an attractive therapeutic candidate for diseases like cystic fibrosis and Crohn's disease.

INTRODUCTION

The emergence of antibiotic-resistant pathogens has become a major health crisis worldwide (10, 15). Novel antimicrobial drugs against resistant microbes are urgently needed. Defensins and cathelicidins, an essential component of the vertebrate innate immune system, are capable of killing a broad spectrum of bacteria with an equal efficiency against both antibiotic-resistant and -susceptible strains and little chance of developing resistance (27). Therefore, these CAMPs are being actively explored as antibacterial drugs. However, the potential of defensins as therapeutics is dampened by a loss of activity in the presence of physiological concentrations of salt (25, 43). As a consequence, none of the defensin-based therapeutics except for a fungal defensin, plectasin, are under clinical development (57).

In vertebrates, there are two major families of antimicrobial peptides, namely defensins and cathelicidins (12, 25, 55, 56). Unlike cathelicidins that are mostly free of cysteines, defensins consist of six conserved cysteine residues in the mature region. Depending upon the disulfide bond connectivity, defensins are further divided into three classes known as α-, β-, and θ-defensins. α-defensins are mainly produced by intestinal Paneth cells and leukocytes, whereas β-defensins are mainly produced by skin keratinocytes as well as mucosal epithelial cells of the gastrointestinal, respiratory and urogenital tracts (4, 12, 25, 32, 34, 42, 43).

In humans, there are only two Paneth cell α-defensins, namely human α-defensin (HD)-5 and 6, whereas multiple Paneth cell α-defensins (known as cryptdins in mice) have been found in mice and rhesus monkey (4, 32-34). Matrilysin or matrix metalloproteinase-7 (MMP-7) is responsible for post-translational processing of inactive cryptdin precursors into antibacterially active peptide (48). Cryptdins are stored as the mature active form in Paneth cell granules and are released in response to bacterial products like lipopolysachharide (LPS), lipoteichoic acid, lipid A and muramyl dipeptide (3). In contrast, HD-5 is stored in Paneth cell granules as the precursor form and is processed by trypsin during or after secretion into intestinal lumen (13). Recently, a family of cysteine-rich cryptdin-related sequences (CRS) have been found in mouse intestine (19). Evidence for the essential role of α-defensins in protecting enteric infections has been demonstrated convincingly. Transgenic mice expressing *HD-5* were resistant to *Salmonella typhimurium* infection (39). Conversely, mice lacking the *Matrilysin* gene were unable to produce mature Paneth cell cryptdins and became highly susceptible to *S. typhimurium* infections (48).

Defensins possess pleotropic functions. In addition to broad spectrum antibacterial, antiviral and antifungal activities, certain defensins are chemotactic for dendritic cells, mast cells, monocytes, and T cells, and some can also induce maturation of dendritic cells and sperms. Defensins help in wound healing by inducing vascularization, promoting proliferation of epithelial and fibroblast cells, and augmenting wound closure (17, 25, 31, 43). In addition, several human neutrophil α-defensins were shown recently to be capable of neutralizing bacterial toxins (20, 21, 30).

We identified a α-defensin-related gene (*Defa-rs1*) recently through a computational screen of the rat genome (36), which we re-named rattusin in this report. Unlike classical mammlian defensins with a canonical six-cysteine motif, rattusin consists of five cysteines with a unique cysteine spacing pattern. To test whether rattusin is a *bona fide* antimicrobial peptide, we synthesized putatively mature rattusin based on its C-terminal sequence and compared its activity with mouse cryptdin-4 and human HD-5, which are the two most potent Paneth cell α-defensins that have been reported.

MATERIALS AND METHODS

RT-PCR analysis of gene expression pattern of rattusin

Different segments of the gastrointestinal tract were collected from 2-month-old Sprague-Dawley rats. Total RNA was extracted using TRIzol (Invitrogen, Carlsbad, CA). For each RNA sample, 4 μg were reverse transcribed using random hexamers and SuperScript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions. The subsequent PCR was carried out as described (36). Briefly, 1/40 of the first-strand cDNA was used to amplify rattusin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) with gene-specific primers (36). The PCR product was analyzed by electrophoresis on 1.2% agarose gel containing 0.5 μg/ml ethidium bromide. The identity of the rattusin PCR product was confirmed by cloning into T/A cloning vector and direct DNA sequencing.

Peptide synthesis and refolding

Since MMP-7 preferentially cleaves the sequence prior to leucine (2, 19, 44, 45), we rationalized that the mature rattusin is likely to start from the first leucine after pro-sequence. Therefore, putatively mature rattusin of 31 amino acids (LRVRRTLQCSCRR VCRNTCSCIRLSRSTYAS) was chemically synthesized in the reduced form using the standard solid-phase synthesis, purified by reverse phase-high pressure liquid chromatography (RP-HPLC) to >95% purity (Bio-synthesis, Lewisvile, TX). The mass of peptide was confirmed by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (MS) using Voyager DE-PRO (Applied Biosystems, Foster City, CA) housed in the Recombinant DNA/Protein Core Facility at Oklahoma State University.

To oxidize the reduced rattusin, 0.1 mg/ml of the peptide was exposed to O_2 gas for 5 min in 50 mM Tris, pH 8.0, and stirred gently for 48 h at room temperature with the cap open as described (19). Following air oxidation, refolded rattusin was purified by RP-HPLC on a 4.6 × 250 mm Vydac C18 column (Grace Vydac, Hesperia, CA) and BioLogic DuoFlow Liquid Chromatography System (Bio-Rad, Hercules, CA). Buffer A consisted of 5% acetonitrile and 0.18% trifluoroacetic acid (TFA). Buffer B consisted of 90% acetonitrile & 0.15% TFA. The gradient used was 0 to 60% buffer B over 90 minutes at a flow rate of 1 ml/min. Eluted peptide was lyophilized and stored at -80°C until use. The peptide was reconstituted in 0.01% acetic acid and quantified by measuring UV absorbance at 280 nm based on the extinction coefficients of tyrosine and cysteines present in rattusin (41). Recombinant cryptdin-4 was produced in bacteria as described

(41, 44). Plasmid encoding cryptdin-4 was provided by Dr. Andre Ouellette at University of California, Irvine. Synthetic HD-5 was kindly provided by Dr. Wuyuan Lu at University of Maryland, Baltimore (49). Cryptdin-4 and HD-5 were used as the reference peptides.

Recombinant expression and purification of pro-rattusin

Recombinant pro-rattusin was produced as a N-terminal 6X histidine-tagged fusion protein using the strategy described previously for cryptdin-4 (41, 44). Restriction sites for *Hind*III and *Bam*HI were included in forward and reverse primers, respectively, as underlined in primer sequences. Additional methionine-encoding *ATG* codon was introduced in the forward primer prior to the start of pro-rattusin nucleotide sequence to enable cyanogens bromide cleavage in order to remove tag sequence. Forward primer was 5'-CGC<u>GGATCC</u>*ATG*GATCCCATTCAAGAGGCAGAGG-3', and the reverse primer sequence was 5'-GCCC<u>AAGCTT</u>ATCTCATCCTTGGGGTCAT-3'. Pro-rattusin cDNA was amplified from rat ileum RNA by using random hexamers, SuperScript II reverse transcriptase, and high-fidelity Platinum *Pfx* DNA polymerase (Invitrogen). The resulting PCR product was gel purified, digested with *Hind*III and *Bam*HI, ligated into pET28a vector (Novagen, Madison, WI), and transformed into *E. coli* DH-5α. The presence of the correct insert was confirmed by DNA sequencing.

Recombinant plasmid was then re-transformed into *E. coli* BL21 (DE3)-Codon Plus-RIL cells (Stratagene, La Jolla, CA). For large-scale production of recombinant pro-rattusin, overnight bacteria were subcultured 1:100 in 2 L of LB broth till OD_{600} reached

0.4 to 0.6. Recombinant fusion protein was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM for 6 h. Bacteria were pelleted, resuspended in the buffer containing 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole and 6 M guanidine chloride, pH 7.4, and sonicated on ice. The lysate was further stirred for 1 h at room temperature to ensure complete solubilization of inclusion bodies.

Follwing centrifugation of the lysate at $30,000 \times g$ for 30 min at 4^oC, recombinant His-tagged pro-rattusin was purified from the supernatant using Ni Sepharose 6 fast flow resins (GE Healthcare Biosciences, Piscataway, NJ) according to the manufacturer's instructions. Fusion protein was then eluted using a buffer contained 20 mM sodium phosphate, 0.5 M NaCl, 0.5 M imidazole and 6 M guanidine chloride, pH 5.0. Following dialysis in 0.01% acetic acid using a Slide-A-Lyzer 3.5 MWCO dialysis cassette (Pierce Biotechnology, Rockford, IL), fusion protein was lyophilized and dissolved in 2.5% TFA at 2.5 mg/ml and incubated with 25 mg/ml CNBr in dark for 24 h to remove the fusion tag (50). The cleaved product was lyophilized, dissolved in Buffer A and subjected to RP-HPLC. Pro-rattusin fraction was then lyophilized, and its molecular weight was confirmed by MALDI-TOF mass spectrometry. The peptide concentration was quantified by measuring UV absorbance at 280 nm.

Bacterial culture and antibacterial assays

All bacterial strains including *Staphylococcus aureus* ATCC 25923, *Listeria monocytogenes* ATCC 19115, *S. aureus* ATCC BAA-39, *S. aureus* ATCC 43300,

Escherichia coli O157:H7 ATCC 700728, *Salmonella enterica* serovar Typhimurium ATCC 14028, *S. enetrica* serovar Typhimurium DT104 ATCC 700408, and *Klebsiella pnumoniae* ATCC 13883 were purchased from either ATCC (Manassas, VA) or MicroBiologics (St. Cloud, MN). Bacteria were grown in trypticase soy broth (TSB) overnight and subcultured for 3-4 h at 37°C in a shaking incubator to the mid-log phase. To study the antibacterial spectrum, a broth microdilution assay was used as described (53). Briefly, mid-log phase bacteria were washed with 25 mM sodium phosphate buffer, pH 7.4, and suspended to 5×10^5 CFU/ml in 25 mM sodium phosphate, pH 7.4, containing 5% TSB with and without 100 mM NaCl. Bacteria (90 μl) were then dispensed in a 96-well tissue culture plate, followed by addition in duplicate of serially diluted rattusin or cryptdin-4 (10 μl). After overnight incubation at 37°C, the minimum inhibitory concentration (MIC) of each peptide was determined as the lowest concentration that gave no visible bacterial growth.

To study the kinetics of bacterial killing, a standard colony counting assay was used as described (53). Rattusin, cryptin-4 and HD-5 were incubated with 90 μl of 5×10^5 CFU/ml *Staphylococcus aureus* or *E. coli* O157:H7 in 25 mM sodium phosphate buffer, pH 7.4, containing 1% TSB with and without 100 mM NaCl. Following incubation at 37°C for 10, 30, 60, 120 and 240 min, bacteria were diluted rapidly with ice-cold PBS and serially plated onto trypticase soy agar (TSA) plates. Viable bacteria were counted after overnight incubation at 37°C. The effect of Mg^{2+} on the antibacterial activity was studied by incubating rattusin and cryptdin-4 for 4 h with 90 μl of *S. aureus* (5×10^5 CFU/ml) at the concentration of 2 μM each or with *E. coli* O157:H7 at the

concentration of 4 μM each in 1% TSB, 25 mM sodium phosphate buffer, pH 7.4, containing 0, 1, 2 and 5 mM of $MgCl_2$.

Cytotoxicity Assay

The toxicity of rattusin to human Caucasian colon adenocarcinoma (Caco-2) cells (ATCC, Manassas, VA) was measured using alamarBlue dye (Biosource, Camarillo, CA) as described previously (53, 54). Briefly, Caco-2 cells were seeded into a 96 well plate at 5×10^4 cels/ml in Dulbecco's modified Eagle's medium (DMEM) with 10% FBS and grown overnight in a humidified 5% CO_2 incubator. Cells were washed once with DMEM, followed by addition of fresh DMEM containing 100 μM of rattusin, cryptdin-4 or chicken fowlicidin-1and in the presence and absence of 10% FBS. After 18 h incubation, 10 μl of alamarBlue dye was added and cells were further incubated for another 6 h at 37^oC. The plate was read with excitation at 545 nm and emission at 590 nm. The percentage of cell death was calculated as $[1-(F_{peptide}-F_{background})/ (F_{acetic\ acid}-F_{background})] \times 100$, where $F_{peptide}$ is the fluorescence of cells exposed to 100 μM peptide, $F_{acetic\ acid}$ is the fluorescence of cells exposed to 0.01% acetic acid only, and $F_{background}$ is the background fluorescence of 10% alamarBlue dye in cell culture medium without cells.

RESULTS

Rattusin, an α-defensin-related peptide with preferential expression in the distal small intestine

Following a comprehensive screening of the entire rat genome, we identified an α-defensin-related gene (36). Similar to α-defensins, this gene is composed of two exons with the first exon encoding the 5' untranslated region (UTR) and preprosegment and the second encoding a putatively mature peptide and 3' UTR (Fig. 1). The encoded peptide, named rattusin, shares a high degree of similarity in the signal and pro-sequences with α-defensins (Fig. 2). However, rattusin consists of five cysteines in the C-terminal mature region, instead of canonical six cysteines in α-defensins (36). Futhermore, the cysteine spacing patterns are quite different between rattusin and classical α-defensins (Fig. 2).

We revealed earlier that the rattusin mRNA was transcribed abundantly in the small intestine of rat (36). To further analyze the detailed expression pattern of rattusin, RT-PCR was performed using the RNA isolated from different segments of the gastrointestinal tract of 2-month-old healthy rats. As shown in Fig. 3, rattusin was highly expressed in distal jejunum and entire ileum, but not in other parts of the gastrointestinal tract. Cryptdin-4 and HD-5, Paneth cell-specific α-defensins, also show a similar expression pattern in small intestine (9, 35, 47). This suggested that, similar to enteric α-defensins, rattusin is likely to be produced by Paneth cells.

170

Antibacterial properties of rattusin

To test whether putatively mature rattusin is functional or not, we synthesized putatively mature rattusin in the reduced form and further oxidized it. Successful refolding was confirmed by RP-HPLC, showing a decrease in the retention time due to the formation of disulfide bonds and a change in the conformation (Fig. 4). Refolded rattusin was used to study antibacterial activity against several representative Gram-negative and Gram-positive bacteria by a modified broth microdilution assay with and without 100 mM NaCl (53). As shown in Table 1, rattusin exhibited antibacterial activity with the MIC value mostly in the range of 2-4 μM. In most cases, rattusin was as potent as cryptdin-4, the most efficacious mouse Paneth cell α-defensins that has been reported (32). Importantly, rattusin also displayed a similar antibacterial efficiency against two strains of methicillin-resistant *Staphylococcus aureus* (MRSA) and multi-drug resistant *Salmonella typhimurium* DT 104 (Table 1). Importantly, the activity of rattusin remained largely unchanged in the presence of NaCl, in sharp contrast to cryptdin-4, whose activity was abolished by salt.

To study the kinetics of bacterial killing, each representative strain of Gram-negative and Gram-positive bacteria was used in a standard colony counting assay. Rattusin, cryptdin-4 and HD-5 at 4 μM killed *E. coli* by three logs within 2 h (Fig. 5A). In the presence of 100 mM of NaCl, killing of bacteria by rattusin was slightly delayed with a complete killing occurring at 4 h. In contrast, Cryptdin-4 and HD-5 were completely inactivated in 100 mM NaCl (Fig. 5B). A similar trend also occurred with *S. aureus*. All three peptides at 2 μM killed *S. aureus* by two logs in 4 h (Fig. 6A). Killing

171

by rattusin was largely unaffected by 100 mM NaCl. However, cryptdin-4 and HD-5 were significantly inhibited in the presence of salt (Fig. 6B). Collectively, these results clearly suggested that rattusin with a broad-spectrum, salt-insensitive bactericidal activity is among the most potent defensins. To further study the effect of salinity, the antibacterial activity of rattusin was examined in the presence of increasing concentrations of NaCl. As shown in Fig. 7, rattusin maintained antibacterial activity against both *E. coli* and *S. aureus* in up to 200 mM NaCl.

Divalent cations such as Mg^{++} are usually present at 1 to 2 mM in most biological fluids (6, 7). At such low concentrations, divalent cations are known to inhibit antibacterial activities of cationic peptides (8, 11). To study the effect of Mg^{++} on the antibacterial activity, rattusin and cryptdin-4 were incubated with bacteria in the presence of increasing concentrations of $MgCl_2$. There was a dose-dependent loss of the activity of cryptdin-4 against *E. coli* with complete inactivation occurring at 2-5 mM of $MgCl_2$ (Fig. 8A). In contrast, the activity of rattusin was largely unaffected in the presence of $MgCl_2$ (Fig. 8A). Mg^{++} did not inhibit the activity of either rattusin or cryptdin-4 against *S. aureus* (Fig 8B). In the case of cryptdin-4, increased killing observed at 1 mM was not observed at higher concentrations of $MgCl_2$ (Fig 8B).

Low cytotoxicity of rattusin

The cytotoxic effect of rattusin to intestinal epithelial cells was studied using human Caco-2 cells. Fowlicidin-1 (100 μM) was used as a positive reference since it showed significant cytotoxicity towards mammalian cells (53). Following treatment of

cells for 24 h, rattusin exhibited no cytotoxicity even at 100 μM, similar to cryptdin-4 (Fig. 9).

Inhibition of the antibacterial activity of rattusin by anionic pro-segment

Recombinant pro-rattusin was produced as an N-terminal 6× histidine-tagged fusion protein followed by removal of the fusion tag by CNBr cleavage. The antibacterial activity of pro-rattusin was studied by a modified broth microdilution assay against two representative Gram-positive and Gram-negative bacteria. Pro-rattusin showed a MIC value of 16 μM against *S. aureus*, but no inhibition of growth was observed for *L. monocytogenes*, *E. coli* and *S. typhimurium* even at 32 μM, the highest concentration used (Table 2). This is reminiscient of cryptdin-4, whose anionic pro-segment also inhibited the antibacterial activity of C-terminal cationic peptide (44).

DISCUSSION

Rattusin shares a highly conserved signal and pro-sequence with classical mammalian α-defensins, but with a unique C-terminal region (Fig. 2). Our results demonstrate that this C-terminal peptide possesses potent, broad-spectrum antibacterial activities. Recently, a number of cysteine-rich cryptdin-related sequences (CRS) in mice were described, which contain a highly similar prepro-region to mice cryptdins (19). However, the cysteine-spacing patterns of these CRS peptides are different from rattusin.

Similar to cryptdin-4 and HD-5 (9, 35), rattusin is preferentially expressed in distal small intestine. It is tempting to speculate that preferential expression of certain

enteric defensins in the distal small intestine might prevent colonization of microflora from adjacent colon. Consistently, the number of microbes in small intestine is 10^4-10^6 times less than the number of microbes in colon (5). Thus, the number and species of microorganisms in small intestine might be dictated by Paneth cell defensins.

Desirably, rattusin maintains its antibacterial activity in the presence of physiological concentrations of NaCl, which is in sharp contrast to cryptdin-4, HD-5, and other defensins, which show reduced antibacterial activity in the presence of NaCl (12, 25, 38, 43). Maintenance of the antibacterial activity of rattusin in the presence of salt might be due to its distinct spacial structure conferred by the unique cysteine spacing pattern. Cryptdin-4 disrupts bilayer head groups through interfacial binding onto lipid bilayer rather than inserting into the hydrophobic core (40, 41). It interacts favorably with bacterial phospholipid vesicles containing negatively charged phospholipids than neutral phospholipids indicating ionic interaction between positive charged peptide and negative charged lipids (40, 41). Such interaction is more susceptible to inhibition by physiological concentrations of NaCl, which in rodent ileum is maintained at around 162 mM (29). Since rattusin is relatively resistant to NaCl, its mechanism of action is likely to be different from cryptdin-4.

Divalent cations such as Mg^{++} stabilize LPS of Gram-negative bacteria by binding to negatively charged phosphate and pyrophosphate groups of LPS, thereby preventing charge-charge repulsion (16, 28). Antimicrobial peptides must displace these divalent cations in order to interact with the outer membrane (16, 28). A number of defensins and

cathelicidins are inhibited at physiological concentrations of divalent cations (1, 8, 11, 24, 26). Rattusin kills bacteria in the presence of 2-5 mM Mg^{++}, whereas cryptdin-4 is inactivated. This suggested that rattusin is able to displace divalent cations.

Divalent cations are not inhibitors of the antibacterial activity of CAMPs against Gram-positive bacteria since they lack LPS. On the contrary, Mg^{++} potentiated the activity of rattusin against *S. aureus*. However, this finding is not surprising since Mg^{++} causes transcriptional suppression of *dlt* operon in *S. aureus* (22). Genes of *dlt* operon are involved in incorporation of D-alanine into lipoteichoic acid and wall teichoic acid, which introduces positively charged amino group and partially neutralizes negative charges on teichoic acid, thereby providing resistance to CAMPs (23). Therefore, suppression of *dlt* operon by Mg^{++} increases susceptibility of *S. aureus* to rattusin and cryptdin-4.

Similar to pro-cryptdin-4, pro-rattusin generally lacks antibacterial activity (Table 2). Pro-cryptdin-4 does not interact with phospholipid vesicles and lacks a typical β-sheet structure (41). It is likely that such inhibition involves neutralization of positive charges on the C-terminal peptide by the negative charge on the N-terminal pro-segment. This mechanism is further supported by evidence that a replacement of aspartate and glutamate by glycine confers bactericidal and membrane-interacting ability of CAMPs (46).

In conclusion, rattusin is unique α-defensin-related peptide, which is preferentially expressed in distal small intestine. It possesses broad-spectrum antibacterial activity insensitive to salt and divalent cations. The salt-insensitive activity of rattusin may be further exploited for the treatment of cystic fibrosis and Crohn's disease. Increased salt concentrations are believed to be responsible for inactivation of defensins in the airway of the cystic fibrosis patients (14). In Crohn's disease, a deficiency of intestinal Paneth cell defensins is associated with an increase in susceptibility to bacterial infections (37). Therfore, exogenous delivery of potent, salt-insensitive rattusin represents a promising therapeutic strategy to treat both topical and systemic antibiotic-resistant infections.

REFERENCES

1. **Anderson RC and Yu PL.** Factors affecting the antimicrobial activity of ovine-derived cathelicidins against E. coli 0157:H7. *Int J Antimicrob Agents* 25: 205-210, 2005.

2. **Ayabe T, Satchell DP, Pesendorfer P, Tanabe H, Wilson CL, Hagen SJ, and Ouellette AJ.** Activation of Paneth cell alpha-defensins in mouse small intestine. *J Biol Chem* 277: 5219-5228, 2002.

3. **Ayabe T, Satchell DP, Wilson CL, Parks WC, Selsted ME, and Ouellette AJ.** Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria. *Nat Immunol* 1: 113-118, 2000.

4. **Bevins CL.** Events at the host-microbial interface of the gastrointestinal tract. V. Paneth cell alpha-defensins in intestinal host defense. *Am J Physiol Gastrointest Liver Physiol* 289: G173-176, 2005.

5. **Bevins CL.** Paneth cell defensins: key effector molecules of innate immunity. *Biochem Soc Trans* 34: 263-266, 2006.

6. **Bowdish DM, Davidson DJ, and Hancock RE.** Immunomodulatory properties of defensins and cathelicidins. *Curr Top Microbiol Immunol* 306: 27-66, 2006.

7. **Bowdish DM, Davidson DJ, and Hancock RE.** A re-evaluation of the role of host defence peptides in mammalian immunity. *Curr Protein Pept Sci* 6: 35-51, 2005.

8. **Bowdish DM, Davidson DJ, Lau YE, Lee K, Scott MG, and Hancock RE.** Impact of LL-37 on anti-infective immunity. *J Leukoc Biol* 77: 451-459, 2005.

9. **Darmoul D and Ouellette AJ.** Positional specificity of defensin gene expression reveals Paneth cell heterogeneity in mouse small intestine. *AmJPhysiol* 271: G68-G74, 1996.

10. **Finch RG.** Antibiotic resistance: a view from the prescriber. *Nat Rev Microbiol* 2: 989-994, 2004.

11. **Friedrich C, Scott MG, Karunaratne N, Yan H, and Hancock RE.** Salt-resistant alpha-helical cationic antimicrobial peptides. *Antimicrob Agents Chemother* 43: 1542-1548, 1999.

12. **Ganz T.** Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 3: 710-720, 2003.

13. **Ghosh D, Porter E, Shen B, Lee SK, Wilk D, Drazba J, Yadav SP, Crabb JW, Ganz T, and Bevins CL.** Paneth cell trypsin is the processing enzyme for human defensin-5. *Nat Immunol* 3: 583-590, 2002.

14. **Goldman MJ, Anderson GM, Stolzenberg ED, Kari UP, Zasloff M, and Wilson JM.** Human beta-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis. *Cell* 88: 553-560, 1997.

15. **Gonzales R, Malone DC, Maselli JH, and Sande MA.** Excessive antibiotic use for acute respiratory infections in the United States. *Clin Infect Dis* 33: 757-762, 2001.

16. **Hancock RE and Scott MG.** The role of antimicrobial peptides in animal defenses. *Proc Natl Acad Sci USA* 97: 8856-8861, 2000.

17. **Hiemstra PS.** Defensins and cathelicidins in inflammatory lung disease: beyond antimicrobial activity. *Biochem Soc Trans* 34: 276-278, 2006.

18. **Hoenderop JG and Bindels RJ.** Epithelial Ca2+ and Mg2+ channels in health and disease. *J Am Soc Nephrol* 16: 15-26, 2005.

19. **Hornef MW, Putsep K, Karlsson J, Refai E, and Andersson M.** Increased diversity of intestinal antimicrobial peptides by covalent dimer formation. *Nat Immunol* 5: 836-843, 2004.

20. **Kim C, Gajendran N, Mittrucker HW, Weiwad M, Song YH, Hurwitz R, Wilmanns M, Fischer G, and Kaufmann SH.** Human alpha-defensins neutralize anthrax lethal toxin and protect against its fatal consequences. *Proc Natl Acad Sci U S A* 102: 4830-4835, 2005.

21. **Kim C, Slavinskaya Z, Merrill AR, and Kaufmann SH.** Human alpha-defensins neutralize toxins of the mono-ADP-ribosyltransferase family. *Biochem J*, 2006.

22. **Koprivnjak T, Mlakar V, Swanson L, Fournier B, Peschel A, and Weiss JP.** Cation-induced transcriptional regulation of the dlt operon of Staphylococcus aureus. *J Bacteriol* 188: 3622-3630, 2006.

23. **Kraus D and Peschel A.** Molecular mechanisms of bacterial resistance to antimicrobial peptides. *Curr Top Microbiol Immunol* 306: 231-250, 2006.

24. **Krishnakumari V, Singh S, and Nagaraj R.** Antibacterial activities of synthetic peptides corresponding to the carboxy-terminal region of human beta-defensins 1-3. *Peptides* 27: 2607-2613, 2006.

25. **Lehrer RI.** Primate defensins. *Nat Rev Microbiol* 2: 727-738, 2004.

26. **Lehrer RI, Ganz T, Szklarek D, and Selsted ME.** Modulation of the in vitro candidacidal activity of human neutrophil defensins by target cell metabolism and divalent cations. *J Clin Invest* 81: 1829-1835, 1988.

27. **Marr AK, Gooderham WJ, and Hancock RE.** Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Curr Opin Pharmacol* 6: 468-472, 2006.

28. **McPhee JB and Hancock RE.** Function and therapeutic potential of host defence peptides. *J Pept Sci* 11: 677-687, 2005.

29. **Miller DL, Hamburger SA, and Schedl HP.** Effects of osmotic gradients on water and solute transport: in vivo studies in rat duodenum and ileum. *Am J Physiol* 237: E389-396, 1979.

30. **Motzkus D, Schulz-Maronde S, Heitland A, Schulz A, Forssmann WG, Jubner M, and Maronde E.** The novel beta-defensin DEFB123 prevents lipopolysaccharide-mediated effects in vitro and in vivo. *Faseb J* 20: 1701-1702, 2006.

31. **Niyonsaba F and Ogawa H.** Protective roles of the skin against infection: implication of naturally occurring human antimicrobial agents beta-defensins, cathelicidin LL-37 and lysozyme. *J Dermatol Sci* 40: 157-168, 2005.

32. **Ouellette AJ.** Defensin-mediated innate immunity in the small intestine. *Best Pract Res Clin Gastroenterol* 18: 405-419, 2004.

33. **Ouellette AJ.** Paneth cell alpha-defensin synthesis and function. *Curr Top Microbiol Immunol* 306: 1-25, 2006.

34. **Ouellette AJ.** Paneth cell alpha-defensins: peptide mediators of innate immunity in the small intestine. *Springer Semin Immunopathol* 27: 133-146, 2005.

35. **Ouellette AJ, Darmoul D, Tran D, Huttner KM, Yuan J, and Selsted ME.** Peptide localization and gene structure of cryptdin 4, a differentially expressed mouse paneth cell alpha-defensin. *Infect Immun* 67: 6643-6651, 1999.

36. **Patil A, Hughes AL, and Zhang G.** Rapid evolution and diversification of mammalian alpha-defensins as revealed by comparative analysis of rodent and primate genes. *Physiol Genomics* 20: 1-11, 2004.

37. **Peyrin-Biroulet L, Vignal C, Dessein R, Simonet M, Desreumaux P, and Chamaillard M.** NODs in defence: from vulnerable antimicrobial peptides to chronic inflammation. *Trends Microbiol*, 2006.

38. **Porter EM, van Dam E, Valore EV, and Ganz T.** Broad-spectrum antimicrobial activity of human intestinal defensin 5. *Infect Immun* 65: 2396-2401, 1997.

39. **Salzman NH, Ghosh D, Huttner KM, Paterson Y, and Bevins CL.** Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 422: 522-526, 2003.

40. **Satchell DP, Sheynis T, Kolusheva S, Cummings J, Vanderlick TK, Jelinek R, Selsted ME, and Ouellette AJ.** Quantitative interactions between cryptdin-4 amino terminal variants and membranes. *Peptides* 24: 1795-1805, 2003.

41. **Satchell DP, Sheynis T, Shirafuji Y, Kolusheva S, Ouellette AJ, and Jelinek R.** Interactions of mouse Paneth cell alpha-defensins and alpha-defensin precursors with membranes. Prosegment inhibition of peptide association with biomimetic membranes. *J Biol Chem* 278: 13838-13846, 2003.

42. **Selsted ME.** Theta-defensins: cyclic antimicrobial peptides produced by binary ligation of truncated alpha-defensins. *Curr Protein Pept Sci* 5: 365-371, 2004.

43. **Selsted ME and Ouellette AJ.** Mammalian defensins in the antimicrobial immune response. *Nat Immunol* 6: 551-557, 2005.

44. **Shirafuji Y, Tanabe H, Satchell DP, Henschen-Edman A, Wilson CL, and Ouellette AJ.** Structural determinants of procryptdin recognition and cleavage by matrix metalloproteinase-7. *J Biol Chem* 278: 7910-7919, 2003.

45. **Turk BE, Huang LL, Piro ET, and Cantley LC.** Determination of protease cleavage site motifs using mixture-based oriented peptide libraries. *Nat Biotechnol* 19: 661-667, 2001.

46. **Weeks CS, Tanabe H, Cummings JE, Crampton SP, Sheynis T, Jelinek R, Vanderlick TK, Cocco MJ, and Ouellette AJ.** Matrix Metalloproteinase-7 Activation of Mouse Paneth Cell Pro-{alpha}-defensins: SER43{downarrow}ILE44 PROTEOLYSIS ENABLES MEMBRANE-DISRUPTIVE ACTIVITY. *J Biol Chem* 281: 28932-28942, 2006.

47. **Wehkamp J, Chu H, Shen B, Feathers RW, Kays RJ, Lee SK, and Bevins CL.** Paneth cell antimicrobial peptides: topographical distribution and quantification in human gastrointestinal tissues. *FEBS Lett* 580: 5344-5350, 2006.

48. **Wilson CL, Ouellette AJ, Satchell DP, Ayabe T, Lopez-Boado YS, Stratman JL, Hultgren SJ, Matrisian LM, and Parks WC.** Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. *Science* 286: 113-117, 1999.

49. **Wu Z, Ericksen B, Tucker K, Lubkowski J, and Lu W.** Synthesis and characterization of human alpha-defensins 4-6. *J Pept Res* 64: 118-125, 2004.

50. **Wu Z, Li X, de Leeuw E, Ericksen B, and Lu W.** Why is the Arg5-Glu13 salt bridge conserved in mammalian alpha-defensins? *J Biol Chem* 280: 43039-43047, 2005.

51. **Wu Z, Powell R, and Lu W.** Productive folding of human neutrophil alpha-defensins in vitro without the pro-peptide. *J Am Chem Soc* 125: 2402-2403, 2003.

52. **Wu Z, Prahl A, Powell R, Ericksen B, Lubkowski J, and Lu W.** From pro defensins to defensins: synthesis and characterization of human neutrophil pro alpha-defensin-1 and its mature domain. *J Pept Res* 62: 53-62, 2003.

53. **Xiao Y, Cai Y, Bommineni YR, Fernando SC, Prakash O, Gilliland SE, and Zhang G.** Identification and functional characterization of three chicken cathelicidins with potent antimicrobial activity. *J Biol Chem* 281: 2858-2867, 2006.

54. **Xiao Y, Dai H, Bommineni YR, Soulages JL, Gong Y-X, Prakash O, and Zhang G.** Structure-activity relationships of fowlicidin-1, a cathelicidin antimicrobial peptide in chicken. *FEBS J* 273: 2581-2593, 2006.

55. **Zanetti M.** Cathelicidins, multifunctional peptides of the innate immunity. *J Leukoc Biol* 75: 39-48, 2004.

56. **Zasloff M.** Antimicrobial peptides of multicellular organisms. *Nature* 415: 389-395, 2002.

57. **Zhang L and Falla TJ.** Antimicrobial peptides: therapeutic potential. *Expert Opin Pharmacother* 7: 653-663, 2006.

FIGURE LEGENDS

Fig.1. Nucleotide sequence of the *rattusin* gene. Numbering is based upon the transcription start site, which is numbered +1 with arrow showing the direction of transcription. Transcribed mRNA sequence is shown in uppercase letters. The open reading frame coding for the rattusin precursor is indicated below each triplet codon. TATA and CAAT boxes in the promoter region are underlined. Stop codon is denoted by an asterisk. The polyadenylation signal is underlined and italicized. The gene structure is

derived from rat EST sequences available in GenBank (Accession Nos. AI639089.1 and CR476138.1) and our previous analysis of *rattusin* gene (36)

Fig. 2. Alignment of rattusin with representative Paneth cell α-defensins, namely mouse cryptdin-4 and human defensin-5 (HD-5). Dashes are created for optimal alignment. Conserved amino acid residues are shaded. The net positive charge on each mature peptide, as underlined, is indicated in parenthesis and length of mature peptide is also indicated. Note that the signal peptide and pro-sequence region of all three peptides are highly homologous, whereas C-terminal mature peptide region are diversified. A conserved six-cysteine motif of cryptdin-4 and HD-5 is absent in rattusin.

Fig. 3. Expression pattern of the rattusin gene across gastrointestinal axis. RT-PCR was used to evaluate the expression of the rattusin transcript. The housekeeping gene, GAPDH, was used to normalize template input.

Fig. 4. RP-HPLC profile of reduced and oxidized rattusin. Reduced synthetic peptide was refolded by air oxidation in 50 mM Tris buffer, pH 8.0, for 48 h. Oxidized rattusin was purified to homogeneity by RP-HPLC. Note that there is a decrease in the retention time of oxidized rattusin due to refolding.

Fig. 5. Kinetics of killing of *E.* coli O157:H7 in the absence (A) and presence (B) of 100 mM NaCl by rattusin, cryptdin-4 and HD-5. *E. coli* O157:H7 ATCC 700728 was incubated with 4 μM rattusin, cryptdin-4, HD-5, or an equal volume of 0.01% acetic acid (no peptide) in duplicate in 25 mM sodium phosphate buffer, pH 7.4, 1% TSB with and without 100 mM NaCl for 10, 30, 60, 120 and 240 min. Surviving bacteria were plated and counted. Data shown are means ± SEM of two independent experiments.

Fig. 6. Kinetics of killing of *S. aureus* in the absence (A) and presence (B) of 100 mM NaCl by rattusin, cryptdin-4 and HD-5 . *S. aureus* ATCC 25923 was incubated with 2 μM rattusin, cryptdin-4, HD-5, or an equal volume of 0.01% acetic acid (no peptide) in duplicate in 25 mM sodium phosphate buffer, pH 7.4, 1% TSB with and without 100 mM

NaCl for 10, 30, 60, 120 and 240 min. Surviving bacteria were plated and counted. Data shown are means ± SEM of two independent experiments.

Fig. 7. Effect of salinity on the antibacterial activity of rattusin against *S. aureus* (A) and *E. coli* O157:H7 (B). *S. aureus* and *E. coli* were incubated with 2 and 4 μM of rattusin, or an equal volume of 0.01% acetic acid (no peptide) in 25 mM sodium phosphate buffer, pH 7.4, 1% TSB with increasing concentrations of NaCl, for 4 h. Surviving bacteria were plated and counted. Data shown are means ± SEM of two independent experiments.

Fig. 8. Effect of Mg^{++} on the antibacterial activity of rattusin and cryptdin-4 against *S. aureus* (A) and *E. coli* O157:H7 (B). *S. aureus* or *E. coli* were incubated with rattusin, cryptdin-4, or an equal volume of 0.01% acetic acid (no peptide) in 25 mM sodium phosphate buffer, containing 1% TSB with 0, 1, 2 and 5 mM $MgCl_2$, pH 7.4, for 4 h. Surviving bacteria were plated and counted. Both peptides were used at identical concentrations for each bacterial strain (2 and 4 μM against *S. aureus* and *E. coli*, respectively). Data shown are means ± SEM of 2-3 independent experiments.

Fig. 9. Absence of cytotoxicity of rattusin and cryptdin-4 to Caco-2 cells. Cells were incubated with 100 μM of rattusin, cryptdin-4 or fowlicidin-1 for 24 h in DMEM with and without 10 % fetal bovine serum for 24 hours. Cell viability was measured with an alamarBlue dye-based method. Data are representative of two independent experiments done in duplicate.

Table 1. Antibacterial spectrum of rattusin and cryptdin 4

Bacteria	ATCC Number	Rattusin 0 mM NaCl	Rattusin 100 mM NaCl	Cryptdin-4 0 mM NaCl	Cryptdin-4 100 mM NaCl
Gram-negative					
E. coli O157:H7	700728	2-4	2-4	2	>16
S. typhimurium	14028	4	8	8	>16
S. typhimurium DT104	700408	4	4-8	4-8	>16
K. pneumoniae	13883	2-4	4-8	8	>16
Gram-positive					
S. aureus	25923	4	4	1-2	≥16
L. monocytogenes	19115	4	4	8	>16
S. aureus (MRSA)	43300	4	4	2	8
S. aureus (MRSA)	BAA-39	4-8	4-8	1-2	>16

MIC (μM)

Table 2. Inhibition of the antibacterial activity of rattusin by the pro-segment

Bacteria	ATCC Number	MIC (μM)	
		Pro-rattusin	Rattusin
Gram-negative			
E. coli O157:H7	700728	>32	2-4
S. typhimurium	14028	>32	4
Gram-positive			
S. aureus	25923	16	4
L. monocytogenes	19115	>32	4

```
-154  ggaagccttagggacagaggaacatcaaagggatcctgagggcaacagctgggagcagtt

 -94  gccatcaatgagtgccttctctaagtgtggggcatgttctttgccctataaatgcaggct
                                         +1

 -34  gcctcctctctctccacacactggtctccagctcACCAATCCTCCAGGTGACTTCCAGCC

  27  ATGAAGACTCTTATCCTCCTCTCTGCTCTTGTCCTGCTGGCCTTGCAGGTCCAGGCTGAT
      Met Lys Thr  Leu Ile Leu  Leu Ser Ala  Leu Val Leu Leu  Ala Leu Gln  Val Gln Ala Asp

  87  CCCATTCAAGAGGCAGAGGAAGAGACTAAAACTGAGGAGCAGCCAGCAGATGAGGACCAG
       Pro Ile Gln  Glu Ala Glu  Glu Glu Thr  Lys Thr Glu Glu  Gln Pro Ala  Asp Glu Asp Gln

 147  GATGTGTCTGTCTCCTTTGAAGGCCCAGAAGCCTCTGCTGTTCAAGATTTACgtgagttc
      Asp Val Ser  Val Ser Phe  Glu Gly  Pro Glu Ala Ser Ala  Val Gln Asp  Leu

 207  tggtacacagaatgtttgaggcttggagtctcctgaggcgggaagggttgtagatgagcc

 267  ctgtagtcctgtcaggtacagtcttgagagccatcggatcttccctgtctcattcataca

 327  agtaacagtgagagtcaagccactgggcttgattttcttccttttaaggttttttgatcca

 387  acaatttatctatgaaaaggctttaaaatatagaacatattgattagacattttagatcc

 447  gattgataattttgttataggaagagatattcgttctgctttgaaaattagatttgggac

 507  ccaaatgccagtgtatgaaggtgttgggtcaggaaaacacaaaaatgctgatagaatgtg

 567  gtgtagatgtcaactcctatatgtgattaacacttgttaaacatcttatctccatgtgtt

 627  tggggttgaccctggtgctggccatgatgtcacccacacagcaaacctactctctaccat

 687  gcacaggatatcttcatagggttgttcactaccacacactgctggcctccttacttcatg

 747  cctgatgctttcttctttcctcagGCGTAAGAAGGACCTTGCAGTGCAGTTGCAGAAGAG
                              Arg Val Arg  Arg Thr Leu Gln  Cys Ser Cys Arg Arg  V

 807  TCTGCAGAAATACGTGTAGCTGCATTCGGCTATCAAGGTCCACATATGCATCATAAATTC
      al Cys Arg Asn Thr  Cys Ser  Cys  Ile Arg Leu Ser  Arg Ser Thr Tyr  Ala Ser *

 867  TGCTGCAGCTGAACATGGACCCCAAGGATGAGATAACCACATGCCCTGGGGCCTCACAGT

 927  CGGTCATTACCCTTGTGCTTGGCCTCGACTGTTTTTCCTTCTCCAATAAACTCCTTGCAG

 987  ATAAAGaatctgtgtttttgtttctttgatgctttttctatgtggcttagacgggactct
```

Fig. 1

188

```
                  Signal Peptide                    Prosequence
Rattusin    MKTLILLSALVLLALQVQADPIQEAEEETKTEEQPADEDQDVSVSFEGPEASA
Cryptdin-4  MKTLVLLSALVLLAFQVQADPIQNTDEETKTEEQPGEEDQAVSISFGGQEGSA
HD-5        MRTIAILAAILLVALQAQAESLQERADEATTQKQSGEDNQDLAISFAGNGLSA

Rattusin    VQDLRVRRTLQCSCRR-------VCRNTCSC-IRLSRSTYAS-   (+ 8)   31
Cryptdin-4  LHEKSLRG-LLCYCRKGHCKRGERVRGTCG--IR-FLYCCPRR   (+ 8)   32
HD-5        LRTSGSQARATCYCRTGRCATRESLSGVCEISGRLYRLCCR--   (+ 4)   32
```
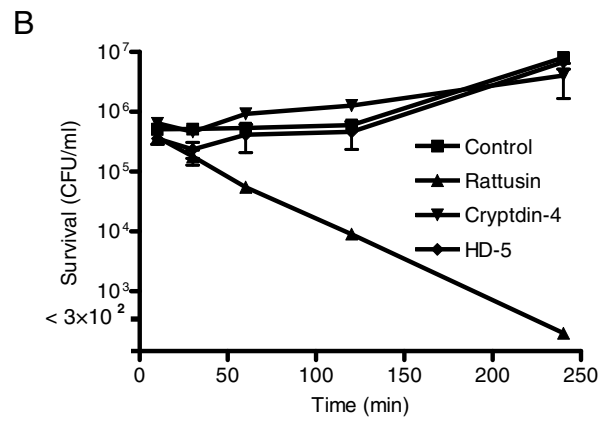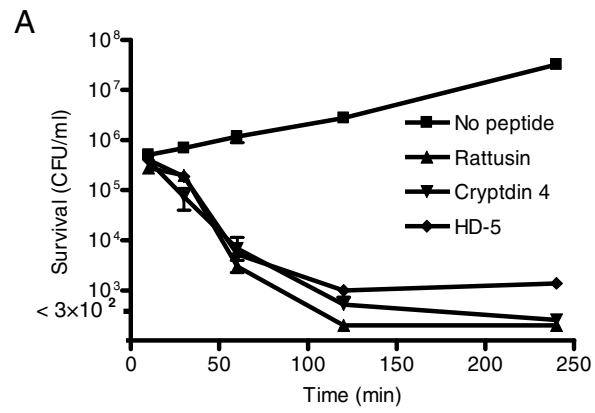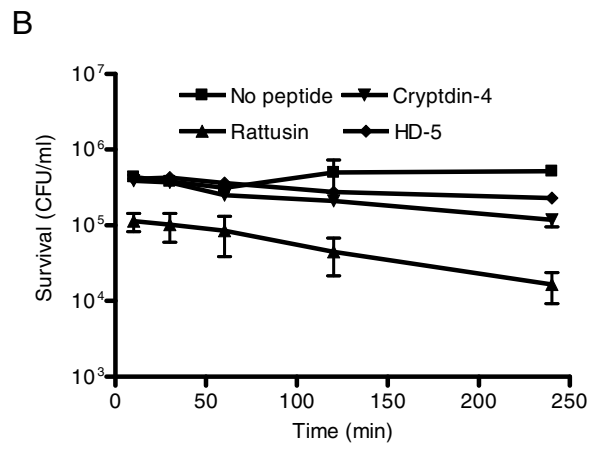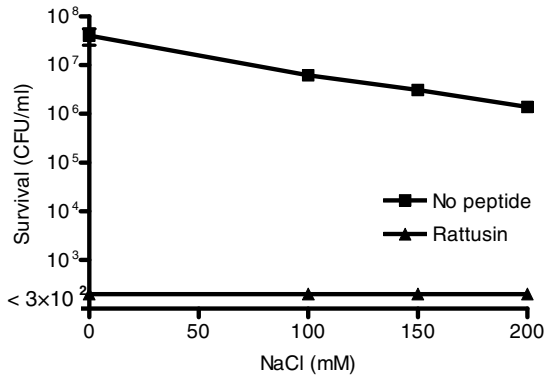
Fig. 2
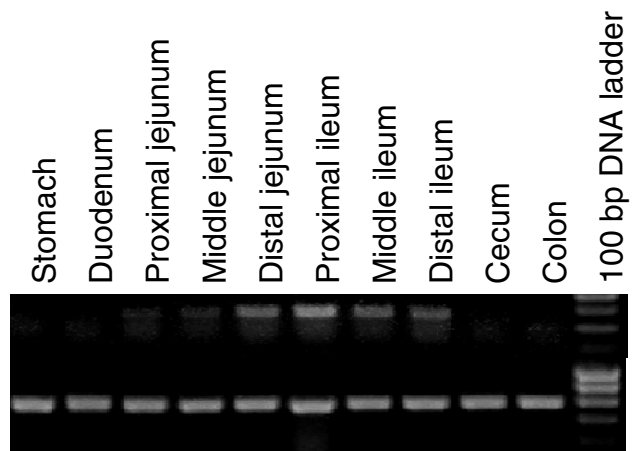
Fig. 3
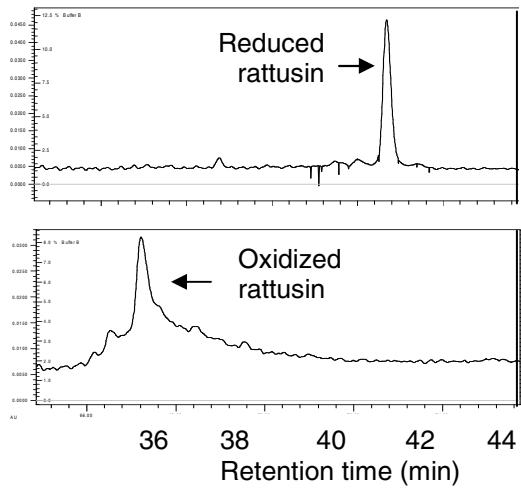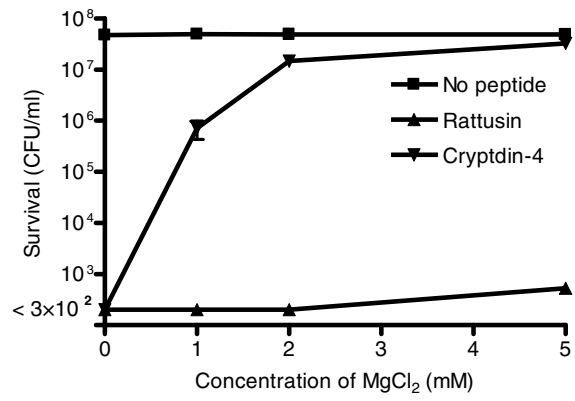
Fig. 4
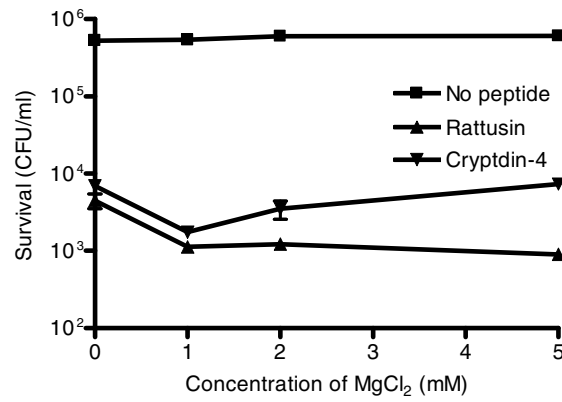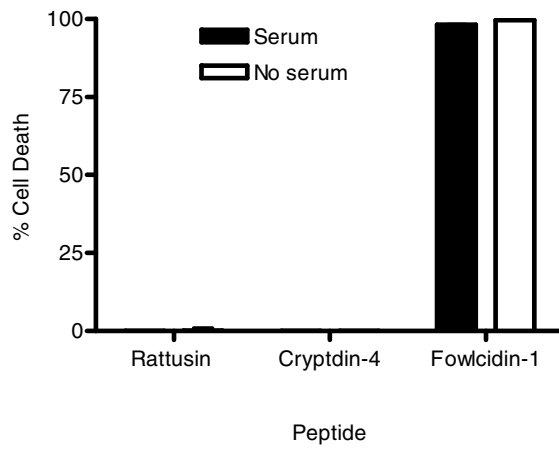
Fig. 5

Fig. 6

A

B

Fig. 7

194
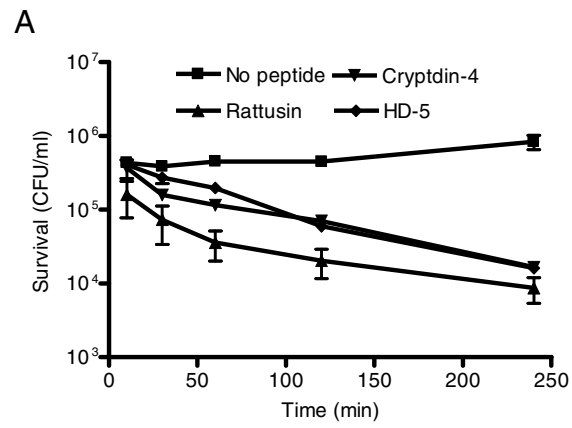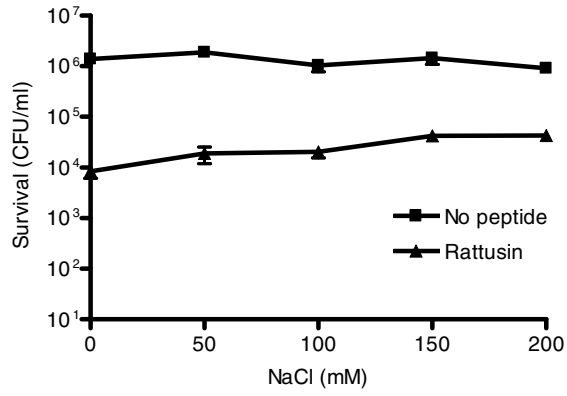
Fig. 8

Fig. 9

CHAPTER VI

SUMMARY AND FUTURE PROSPECT

Using a comprehensive computational search strategy, we identified the complete repertoire of the α-defensin gene family in the human, chimpanzee, rat, and mouse, including a number of novel genes. α-defensin genes were located on syntenic chromosomes in each species. Phylogenetic analysis revealed that α-defensin genes in rodents have undergone repeated duplication which gave rise to the current repertoire of genes. We also found that positive selection is acting on the α-defensin mature peptide coding region. Further, gene expression studies with rat α-defensin genes revealed that they are either expressed in bone marrow or small intestine.

We also identified the entire family members of β-defensins including a large number of novel genes in the human, chimpanzee, rat, mouse and dog. Each species possessed >35 β-defensin genes encoded in their genome. Such genes were located on 4-5 syntenic chromosomes across species. In contrast to most known genes with two-exon gene structure, several novel β-defensin genes had three exons. Phylogenetic analysis showed that most of the genes are conserved in all species. However, certain gene lineages were only present in certain species, indicating that these genes appeared by gene duplication after species separation. Gene expression analysis of rat β-defensins

revealed that all but one are expressed in the male reproductive tract especially in testis and epididymis. Further, this expression was developmentally regulated with the highest level occurring at sexual maturity. This indicates that defensins may have a role in both reproduction and antimicrobial protection.

A novel α-defensin-related gene, which we named rattusin, was identified in the rat and preferentially expressed in small intestine. The encoded peptide displayed an entirely different cysteine-spacing pattern, suggesting that it might have distinct antibacterial properties. Putatively mature rattusin possessed a broad-spectrum antibacterial activity with similar efficiency against both antibiotic-sensitive as well as - resistant bacteria. Further, this activity of rattusin was salt-insensitive although salt slightly delayed kinetics of bacterial killing, in contrast with human defensin-5 and mouse cryptdin-4, both of which are among the most potent defensins identified. Additionally, rattusin was active in the presence of 2-5 mM of Mg^{++}, whereas cryptdin-4 lost its activity completely. Rattusin also displayed little toxicity towards human colon epithelial cells.

To our knowledge, rattusin is the only defensin-related peptide with salt-insensitive antibacterial activities, and therefore, has great potential to become a novel anti-infective drug. Elucidation of its tertiary structure will help explain the structural-basis for such an antibacterial property. In addition, this study has also identified a large number of novel α- and β-defensins. Their specific functions in reproductive health and

fertility largely remain unknown. Further studies are needed to explain the biological significance for such a large number of defensin genes.

VITA

Amar Ajit Patil

Candidate for the Degree of

Doctor of Philosophy

Thesis: GENOME-WIDE IDENTIFICATION OF MAMMALIAN DEFENSINS AND FUNCTIONAL ANALYSIS OF A NOVEL DEFENSIN-RELATED PEPTIDE

Major Field: Animal Science (Molecular Biology)

Biographical:

    Personal Data: Born in Belgaum district of India, Son of Mr. Ajit and
                 Mrs. Surekha Patil

    Education

      2000-2002: Master of Veterinary Science (Animal Biotechnology)
             Indian Veterinary Research Institute, Izatnagar, India
      1995-2000: Bachelor of Veterinary Science & Animal Husbandry
             Nagpur Veterinary College, Nagpur
             Dr. P.D.K.V. Agricultural University, Akola, India
    Experience

      2003-2006: Graduate Research Assistant, Molecular biology lab,
             Dept. Animal Science, Oklahoma State University,
             Stillwater, Oklahoma
      2000-2002: Junior Research Fellow, National Biotech. Center
             Indian Veterinary Research Institute, Izatnagar, India
         2000: Veterinary Intern
             Various State Government veterinary hospitals, India

    Professional Memberships

    American Association for Advancement of Science (AAAS)

Name: Amar Ajit Patil   Date of Degree: December, 2006

Institution: Oklahoma State University   Location: Stillwater, Oklahoma

Title of Study: GENOME-WIDE IDENTIFICATION OF MAMMALIAN DEFENSINS
               AND FUNCTIONAL ANALYSIS OF A NOVEL DEFENSIN-RELATED
               PEPTIDE

Pages in Study: 199   Candidate for the Degree of Doctor of Philosophy

Major Field: Animal Science

Scope and Method of Study: The purpose of this study was to identify novel antimicrobial defensins encoded in the genomes of evolutionarily diverse species such as dog, rat, mouse, human and chimpanzee. Our hypothesis was that identification of novel genes will help us to understand the evolution of this important gene family and provide novel antimicrobial candidates. Genome-wide computational searches of various mammals were conducted to identify the defensin genes using known defensins as query. Further, a novel defensin-related sequence named rattusin with a unique cysteine-spacing pattern was synthesized and evaluated for its antibacterial activity to test the hypothesis that it possesses a unique antibacterial activity. Its precursor peptide was also produced recombinantly and tested against pathogenic bacteria. In addition, the gene expression of rattusin was studied across the entire gastrointestinal tract.

Findings and Conclusions: A large number of defensins were found. α-defensin genes were present only in primates and glires, whereas β-defensins genes were identified in all species studied. Computational mapping revealed that all α- and β-defensin genes were clustered on syntenic chromosomes across all species with 4 clusters in each species, except that humans have 5 clusters. Phylogenetic and evolutionary analysis revealed that α-defensin were subjected to rapid evolution and diversification. α-defensin genes were primarily expressed in bone marrow and small intestine, whereas β-defensin genes were mainly expressed in the male reproductive tract. A novel α-defensin-related gene *rattusin* was identified in rat. This gene displayed a preferential expression in the distal small intestine. Precursor of rattusin did not exhibit significant antibacterial activity, but the C-terminal putatively mature antibacterial peptide showed potent, broad spectrum, salt insensitive antibacterial activity in contrast to cryptdin-4 and human defensin-5. Rattusin also maintained antibacterial activity in the presence of divalent cation, Mg^{++}. Rattusin showed no cytotoxicity to intestinal epithelial cells. Thus, rattusin represents a potential antimicrobial drug to treat infections.

ADVISER'S APPROVAL: _____Guolong Zhang_____