Evidence of Positively Selected Sites in Mammalian α-Defensins

David J. Lynn,† Andrew T. Lloyd,*‡ Mario A. Fares,§ and Cliona O'Farrelly*‖*

*Education and Research Centre, St. Vincent's University Hospital, Dublin, Ireland; †Department of Medicine, University College Dublin, Dublin, Ireland; ‡Department of Genetics, Trinity College Dublin, Dublin, Ireland; §Biology Department, National University of Ireland, Maynooth, Ireland; and ‖Conway Institute, University College Dublin, Dublin, Ireland

α-Defensins are a family of mammalian antimicrobial peptides that exhibit variable activity against a panel of microbes, including bacteria, fungi, and enveloped viruses. We have employed a maximum-likelihood approach to detect evidence of positive selection (adaptive evolution) in the evolution of these important molecules of the innate immune response. We have identified 14 amino acid sites that are predicted to be subject to positive selection. Furthermore, we show that all these sites are located in the mature antimicrobial peptide and not in the prepropeptide region of the molecule, implying that they are of functional importance. These results suggest that mammalian α-defensins have been under selective pressure to evolve in response to potentially infectious challenges by fast-evolving microbes.

Introduction

Defensins are cationic antimicrobial peptides that show a broad spectrum of antimicrobial activity against gram-negative and gram-positive bacteria, fungi, and enveloped viruses. They are generally thought to act by disrupting the membrane integrity of microbes (Kagan et al. 1990; Satchell et al. 2003). In mammals, α-defensins and β-defensins are two structurally distinct groups, which differ in size and in their spacing of a six-cysteine motif. α-Defensins were first discovered as a family of peptides in rabbit macrophages (Selsted, Szklarek, and Lehrer 1984) and were subsequently identified in the neutrophils of humans, macaques, rats, rabbits, guinea pigs, and hamsters and in human and mouse small intestine Paneth cells (Raj and Dentino 2002).

In humans, six α-defensins have been identified and studied to date. Human α-defensins 1 to 4 are localized in azurophilic granules of neutrophils, which has led to them being referred to as human neutrophil peptides (HNP1 to HPN4). The HNPs are the most abundant protein in neutrophils and contribute to the oxygen-dependent killing of phagocytosed microorganisms (Ganz et al. 1985; Singh et al. 1988; Bateman et al. 1991). Recently, HNP1 to HPN3 have been shown to be the major components of a soluble factor secreted from CD8 T-lymphocytes that suppresses HIV-1 replication (Zhang et al. 2002). Human α-defensins (HAD) 5 and 6 are primarily located in the secretory granules of Paneth cells within the crypts of Lieberkühn in the small intestine (Jones and Bevins 1992, 1993; Porter et al. 1997; Salzman et al. 2003*b*) where they have been implicated in mucosal host defense. Further evidence for the important role these peptides play in mucosal host defense has been shown in a recent study in which transgenic mice expressing HAD5 were completely resistant to *Salmonella typhimurium* (Salzman et al. 2003*b*). HAD5 has also been detected in female reproductive tract (Svinarich et al. 1997; Quayle et al. 1998) and in bronchial and nasal epithelia (Frye et al. 2000). Peptides homologous to HAD5 and HAD6 have been found in mouse Paneth cells where they are termed cryptdins (Ouellette et al. 1992, 1994).

α-Defensins are encoded as propeptides that require proteolytic processing to become activated. It has been suggested that the proregion is cytoprotective because addition of it to in vitro assays inhibit activity (Ouellette et al. 1999; Wu et al. 2003). Paneth cell trypsin has been identified as the processing enzyme for HAD5, but the processing enzyme for HAD6 has yet to be identified (Ghosh et al. 2002). In mice, activation of cryptdins is mediated by matrix metalloproteinase-7 (MMP-7) (Wilson et al. 1999; Ayabe et al. 2002). The signal peptide and proregion in α-defensins exhibit more amino acid conservation than does the mature antimicrobial peptide region. Indeed only the six cysteine residues and a glycine residue at position 18 of the mature peptide are conserved between all α-defensins. These residues are important for the structure of the molecule (Hill et al. 1991), whereas the other residues are free to vary.

Most molecular variation either within species or between species is caused by the random fixation of mutations that are neutral. Mutations that are deleterious to an organism are removed by purifying selection. Occasionally, mutations confer selective advantage to the organisms having them, and, therefore, these mutations are fixed in the population by positive selection (adaptive evolution) at a higher rate than expected under neutral evolution. One of the most stringent methods of detecting adaptive evolution is to compare the rate of nonsynonymous substitutions (d_N) with the rate of synonymous substitution (d_S). The ratio between these rates ($\omega = d_N/d_S$) is then a reliable measure of the selective pressure acting in a protein-coding gene. If amino acid changes are neutral, they will be fixed at the same rate as synonymous mutations ($\omega = 1$). If mutations (d_N) are deleterious, $d_N/d_S = 0$, and the mutations will be removed by selection. If nonsynonymous mutations are slightly deleterious ω is less than 1, and the coefficient of selection acting against the mutation will depend on the population size. Finally, if the amino acid changes are selectively advantageous, they will be fixed at a higher rate ($\omega > 1$). A number of methods have been developed to calculate d_N and d_S among lineages (see for review, Yang [2002]). Maximum-likelihood (ML) methods are most statistically satisfactory because they employ an explicit model of evolution, taking into account the effects of unequal transition and transversion rates, unequal base

Key words: Adaptive evolution, antimicrobial peptide, α-defensin.

E-mail address: cliona.ofarrelly@ucd.ie.

Mol. Biol. Evol. 21(5):819–827. 2004
DOI:10.1093/molbev/msh084
Advance Access publication February 12, 2004

Table 1
Accession Numbers of the 34 α-Defensin Genes Used in the Analyses

Gene	Species	Accession Number
HNP1	*Homo sapiens*	NM_004084
HNP3	*Homo sapiens*	NM_005217
HNP4	*Homo sapiens*	NM_001925
HAD5	*Homo sapiens*	NM_021010
HAD6	*Homo sapiens*	NM_001926
MNP1	*Macaca mulatta*	AF184159
MNP1A	*Macaca mulatta*	AF184160
MNP2	*Macaca mulatta*	AF184161
Mdef1	*Macaca mulatta*	AF188268
Mdef3	*Macaca mulatta*	AF188269
Mdef4	*Macaca mulatta*	AF188271
Mdef6	*Macaca mulatta*	AF188272
Mdef8	*Macaca mulatta*	AF188270
CRP1	*Mus musculus*	NM_010031
CRP2	*Mus musculus*	U02996
CRP3	*Mus musculus*	NM_007850
CRP4	*Mus musculus*	NM_010039
CRP5	*Mus musculus*	NM_007851
CRP6	*Mus musculus*	NM_007852
NP1	*Rattus norvegicus*	U16686
NP2	*Rattus norvegicus*	NM_173329
NP3	*Rattus norvegicus*	U50353
NP3B	*Rattus norvegicus*	U50354
ED	*Rattus norvegicus*	AF115768
NP4	*Rattus norvegicus*	NM_173299
NP3A	*Oryctolagus cuniculus*	M64599
NP4	*Oryctolagus cuniculus*	M64601
NP5	*Oryctolagus cuniculus*	M64602
MCP1	*Oryctolagus cuniculus*	M28883
MCP2	*Oryctolagus cuniculus*	M28072
DEF1A	*Cavia porcellus*	D14119
DEF1B	*Cavia porcellus*	D14118
DEF2	*Cavia porcellus*	X63676
DEF1A	Cavia cutleri	X57705

NOTE.—Where RefSeq (http://www.ncbi.nlm.nih.gov/RefSeq/) accession numbers have been assigned, these have been used, otherwise, GenBank accession numbers are given.

and codon frequencies, and variable ω values among lineages in a phylogeny (Yang 1998; Yang and Nielsen 1998). As most amino acid sites are expected to be conserved to maintain key structural and functional characteristics of a protein, and only a small number of sites in a protein are likely to be affected by adaptive evolution, methods that average substitution rates over all sites in a sequence have little power in detecting cases of positive selection. Apart from models that detect positive selection among lineages, models that can detect amino acids sites under adaptive evolution have also been developed (Nielsen and Yang 1998; Yang et al. 2000).

Because α-defensins are a critical component of the innate immune response in an "arms race" against fast-evolving microbes, it is possible that these molecules are subject to adaptive evolution. It has been previously demonstrated that ω is greater than 1 in the antimicrobial peptide region of α-defensins, indicative of positive selection (Hughes and Yeager 1997). The Hughes and Yeager study was, however, unable to predict which particular amino acid sites were under positive selection, as appropriate models were undeveloped at that time. In this first comprehensive study of all available mammalian α-defensins, we implement ML models to test for positive selection among evolutionary lineages and among amino acid sites.

Materials and Methods

Coding sequences and corresponding protein sequences for 34 α-defensins from seven different species, *Homo sapiens* (human), *Macaca mulatta* (rhesus monkey), *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Oryctolagus cuniculus* (rabbit), *Cavia porcellus* (domestic guinea pig), and *Cavia cutleri* (wild guinea pig) were downloaded from GenBank (table 1). These sequences represent all the full-length α-defensins currently available in the database. Molecules annotated as α-defensin–related were not included in this study, as they contain a different cysteine motif in the mature peptide and do not align well in this region with other α-defensins. Inclusion of these molecules could bias the results by overestimating d_N. The protein sequences were aligned using the T-Coffee program (fig. 1 in Supplementary Material online) (Notredame, Higgins, and Heringa 2000). A Neighbor-Joining (NJ) tree was inferred from the protein alignment using MEGA version 2.1, with the gamma distribution model implemented to account for heterogeneity among sites (Kumar et al. 2001). The shape parameter of the gamma distribution (α) was estimated using the BASEML program (implementing the REV model) from the PAML package version 3.12 (Yang 1997). One-thousand bootstrap replicates were carried out to test the significance of each node in the tree. Branches leading to nodes with a bootstrap value of less than 500 were collapsed. The topology of the tree was used as the input tree for CODEML and CODEMLSITES programs, also from the PAML package version 3.12 (Yang 1997). To construct an alignment of the coding sequences, the protein alignment was used as a template, and a copygaps Perl script was used to align the DNA, maintaining the gaps that were present in the protein alignment. Any columns in the DNA alignment that had more than three gaps were removed.

Variable Selective Pressures Among Lineages

Models of variable ω ratios among lineages were fitted by ML to the alignment of the 34 α-defensin sequences. The one-ratio model assumes an equal ω ratio for all branches in the phylogeny. The free-ratios model assumes an independent ω ratio for each branch. The two models can be compared by a likelihood ratio test (LRT). Twice the log-likelihood difference between the two models is compared with a χ^2 distribution with N-1 degrees of freedom where N is the number of branches. Posterior Bayesian probabilities were estimated for codon substitutions in each branch of the phylogenetic tree.

Variable Selective Pressures Among Amino Acid Sites

Models of variable ω ratios among sites were used to test for the presence of sites under diversifying selection ($\omega > 1$) and to identify them. Five models for the ω distribution implemented in the CODEMLSITES program of the PAML package were tested. Model M1 (neutral) assumes two classes of sites in the protein: the conserved

sites (ω = 0) and the neutral sites (ω = 1). Model M2 (selection) adds a third class of site, with ω as a free parameter, allowing for sites with ω greater than 1. Under the discrete model M2, the proportion of sites under purifying selection (p_0) and proportion of sites under neutrality (p_1) are estimated from the data. Model M3 (discrete) uses a general discrete distribution with three classes of site, with the proportions (p_0, p_1, and p_2) and the ω ratios (ω_0, ω_1, and ω_2). Model M7 (beta) uses a beta distribution, which, depending on parameters p and q, can take different shapes in the interval (0, 1). Model M8 (beta and ω) adds an extra class of sites to the beta (M7) model, with the proportion and the ω ratio estimated from the data, thus allowing for sites with ω greater than 1. From these models, three LRTs compare M0 (one ratio) with M3 (discrete), M1 (neutral) with M2 (selection), and M7 (beta) with M8 (beta and ω), respectively. Models M2, M3, and M8 are tests of positive selection among sites. Posterior Bayesian probabilities of site classes were calculated for each amino acid site. If the ω ratios for some site classes are greater than 1, sites with high posterior probabilities for those classes are likely to be under positive selection.

Sliding-Window Analysis to Detect Selective Constraints by Maximum Parsimony

Several studies have shown that maximum-likelihood methods are sensitive to the violation of assumptions made in models to detect adaptive evolution and that false positive results could be obtained under certain conditions (Suzuki and Nei 2002). We, therefore, applied a maximum-parsimony method to test for adaptive evolution in our sequences. We applied the Kimura-based model of Li (1993) using a sliding-window procedure (Fares et al. 2002). Briefly, the method infers a statistically optimum codon-window size and slides it along the alignment. We then test in each sliding step the significance of the nonsynonymous nucleotide substitutions (d_S), synonymous substitutions (d_N), and the nonsynonymous-to-synonymous rate ratio (ω). The mathematical approach used is based on the maximum-parsimony method of Suzuki and Gojobori (1999). The main difference however resides in the fact that the window size is selected under a statistical procedure (Fares et al. 2002). Another advantage of using this method is that the number of synonymous substitutions is tested for significance, and, hence, saturated synonymous sites, if any, can be highlighted and removed from the analysis.

Results

An NJ phylogenetic tree was reconstructed from the amino acid alignment using MEGA version 2.1 (fig. 1). The gamma distribution model was implemented to account for heterogeneity among sites (Kumar et al. 2001). The gamma shape parameter was estimated using the BASEML program to be $\alpha = 1.39$. This tree topology was used in the subsequent analyses to detect adaptive evolution. To ensure that the topology of the tree was not dependent on the sites under positive selection or the invariant sites, the phylogenetic analysis was repeated without these sites. The topology of this tree was essentially the same as the tree in figure 1, with some of the resolution lost (see Supplementary Material online). The CODEMLSITES analysis was repeated using this topology as input and resulted in the detection of the same sites as being subject to positive selection (data not shown).

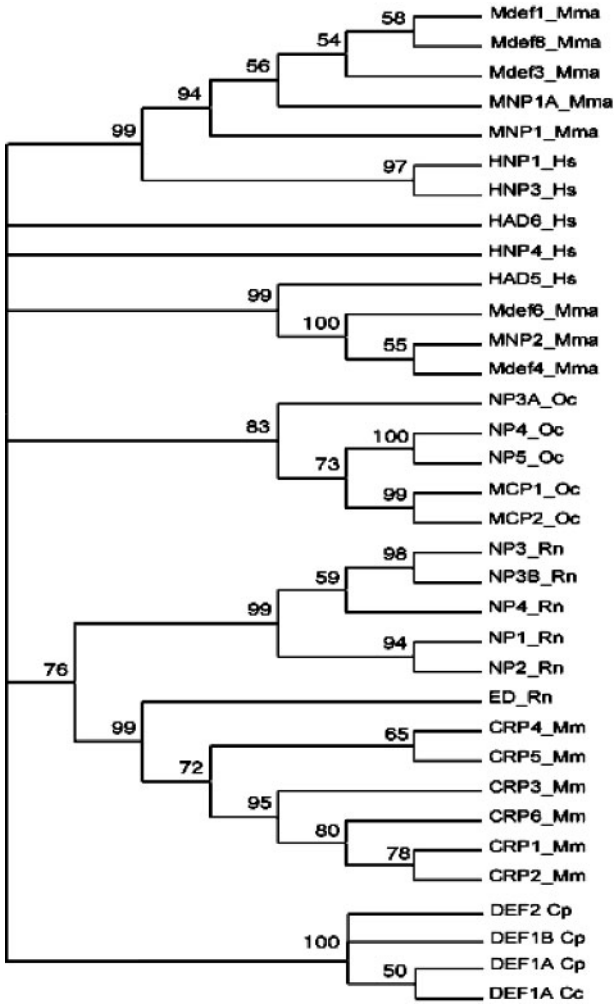

FIG. 1.—Neighbor-Joined tree of mammalian α-defensins. Constructed using MEGA version 2.1 (gamma distribution model, 1,000 bootstrap replicates). Branches with less than 50% bootstrap support have been collapsed. Mm = *Mus musculus* (mouse), Rn = *Rattus norvegicus* (rat), Hs = *Homo sapiens* (human), Mma = *Macaca mulatta* (rhesus monkey), Oc = *Oryctolagus cuniculus* (rabbit), Cp = *Cavia porcellus* (domestic guinea pig), and Cc = *Cavia cutleri* (guinea pig).

Multigene families whose members have the same function may evolve in a concerted fashion that homogenizes the sequences of the member genes by interlocus recombination or gene conversion, such that sequences within a species are more similar to each other than those between species (Liao 1999). The physical clustering typical of α-defensin families within species may imply concerted evolution. There are, however, two exceptions to phylogenetic clustering in our tree (fig. 1), which may mean that the birth-and-death model of evolution is more appropriate (Nei, Gu, and Sitnikova 1997). Neither the primates nor the rodents show species-specific clades of α-defensins. In fact, in the case of the mouse and rat, the

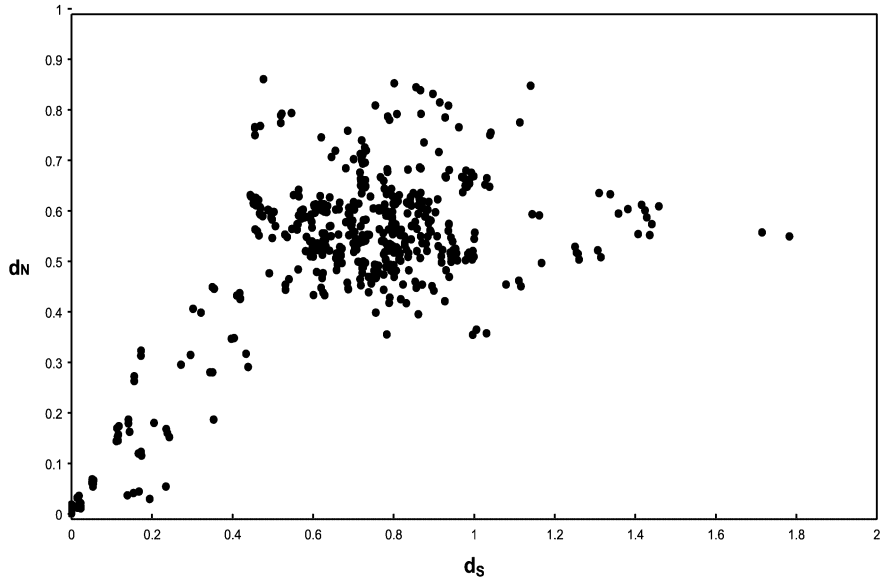

FIG. 2.—Correlation between d_N and d_S values as estimated by maximum likelihood.

phylogenetic clustering appears to be caused by a functional relationship between the different α-defensins. Mice do not have neutrophil defensins (Eisenhauer and Lehrer 1992) and most of the rat α-defensins are expressed in neutrophils; however, the one rat α-defensin expressed in the intestine (ED_Rn) in a similar manner to the mouse cryptdins, clusters with the mouse α-defensins. Furthermore, the level of sequence divergence, even at the protein level (fig. 1 in supplementary material available online), is inconsistent with the homogenization of sequences expected under the concerted evolution model. In the birth-and-death model of evolution, duplicate genes are produced in a gene family. Some of the genes functionally diverge, others may be lost from the genome, and still others become pseudogenes (Liao 1999). Consistent with this model, we have detected multiple α-defensin pseudogenes in the human and mouse genomes (unpublished data).

Examining Saturation of Synonymous Sites

Saturation of synonymous sites is an important issue in the detection of positive selection by the criterion that d_N/d_S is greater than 1, as saturation will lead to the underestimation of d_S and an inflation of the d_N/d_S ratio. To ensure that saturation was not an issue in our data set, pairwise d_N and d_S values were calculated using the ML method implemented by CODEML (Goldman and Yang 1994). A correlation analysis between d_N and d_S was performed using SPSS version 11.0 (fig. 2). If synonymous sites are saturated, we might expect a quadratic model with an increasing slope to be better fit to the data than a linear model. If synonymous sites are not saturated, a linear model should fit the data better than a quadratic model. A quadratic model with a decreasing slope implies that when average values of nonsynonymous nucleotide substitutions are examined, most of them show strong purifying selection and no saturation of synonymous sites exists in our data that could inflate d_N/d_S values.

Variable Selective Pressures Among Lineages

To test for variable ω ratios among lineages, the one-ratio model (Goldman and Yang 1994), which assumes the same ω ratio for all lineages, was compared with the free-ratio model (Yang 1998), which assumes an independent ω ratio for each branch, using the LRT. The log-likelihood value for the one-ratio model is $\ell_0 = -3888.76$, while the value for the free-ratio model is $\ell_1 = -3862.20$. Comparison of $2\Delta\ell = 2(\ell_1 - \ell_0) = 53.12$ (df = 60) reveals that the free-ratio model is not significantly better than the one-ratio model ($P > 0.5$). The free-ratio model does, however, predict variable ω values among lineages, some of which are greater than 1 (fig. 3). Because the LRT did not reveal a significant difference, we cannot conclude that there is evidence of positive selection among the α-defensin lineages. The relatively low ω values in some of the mouse lineages may be a result of the fact that all the mouse α-defensins are expressed only in the intestine and not in the neutrophils (Eisenhauer and Lehrer 1992), and as a group they may not be subject to the same selective pressures as the other species that have both neutrophil and intestine α-defensins.

Variable Selective Pressures Among Amino Acid Sites

To test for positive selection at individual amino acid sites, LRTs were carried out between model M0 (one ratio) and M3 (discrete), M1 (neutral) and M2 (selection), and M7 (beta) and M8 (beta and ω) (table 2). All 3 models (M2, M3 and M8) that allow for selection (table 3) are significantly favored over the other models ($P < 0.001$) in all cases (table 3). The sites predicted to be under positive selection with posterior probabilities greater than 0.95 are in agreement between models M2, M3 and M8, except for

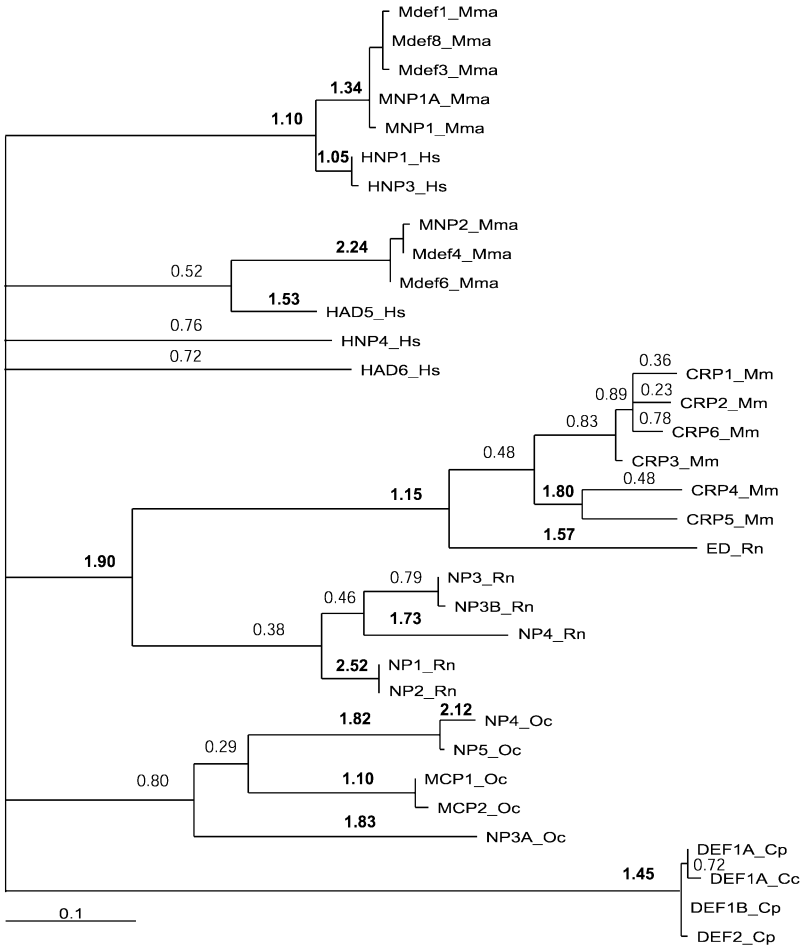

FIG. 3.—Phylogeny of mammalian α-defensins. The abbreviations used are the same as in figure 1. Branch lengths were estimated by maximum likelihood under the free-ratio model, which assumes an independent ω value for each branch. Branches with no ω values shown had values = ∞. ω values greater than 1 are shown in bold. The topology and number of branches is the same as in figure 1 (some branches are very short and are not visible on this tree).

sites 25 and 62. Neither of these sites is predicted by M8, and site 25 is only predicted by M3.

α-Defensins are encoded as prepropeptides that are proteolytically cleaved to release the C-terminal antimicrobial peptide. All of the sites predicted to be under positive selection under model M8, the most stringent model, are located in the mature antimicrobial peptide and not in the prepropeptide region (fig. 4*a*). Comparison of the distribution of positively selected sites to a Poisson distribution revealed that the clustering of these sites in the antimicrobial peptide was statistically significant ($P < 0.001$). Adaptive evolution of the mature peptide is likely to have driven the considerable amino acid variation in this region compared with the prepropeptide region.

The sites predicted to be subject to positive selection have been displayed on the three-dimensional structure of HNP3 (fig. 4*b*). Although sites under positive selection occur throughout the molecule, it is notable that almost all of the sites in the coil region (residues 7 to 14) are subject to adaptive evolution. The only site not predicted is Glu14, which is known to form a salt bridge with Arg6 and as such is relatively invariable (Hill et al. 1991). It is likely that this region has an important part to play in the function and activity of the antimicrobial peptide.

Many of the sites detected to be under positive selection using ML-based models were also detected using the sliding-window–based method. Some others, however (amino acid sites 62, 63, 68, 69, 78, and 82), did not show significant differences between d_S and d_N. Among the positive-selected amino acid sites, the average ω value was 2.972, being significantly higher than 1 and higher than the expectation under neutrality, even after correcting for multiple tests (multiple sliding-window tests) using Bonferroni correction ($Z = 11.47$; $P < 0.001$). We have not detected saturation of synonymous sites and hence ω values are not inflated by this bias. We have to stress, however, that maximum-parsimony methods are very conserved and may be subject to the problem of possible convergences. Despite this fact, our results using maximum parsimony are quite coincident with those using ML methods and do not affect the final conclusions of this study.

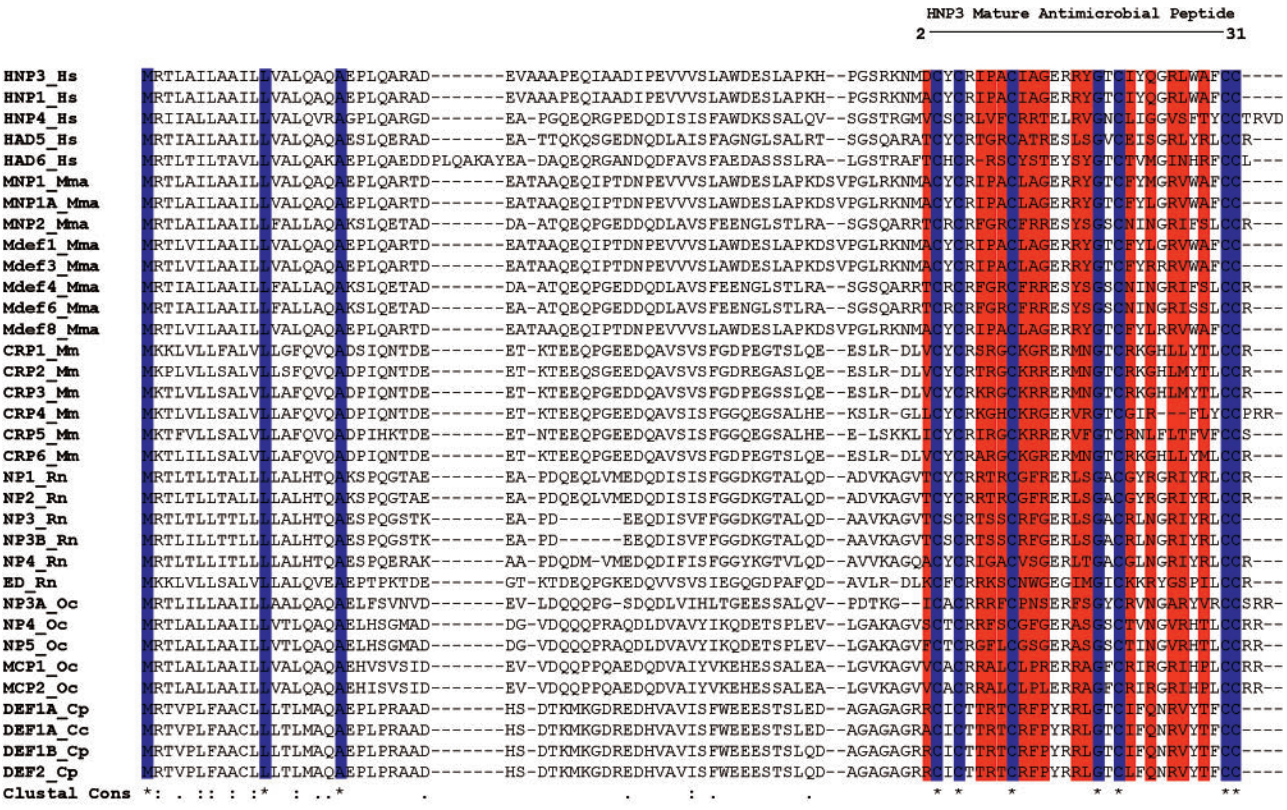

FIG. 4.—Sites predicted to be under positive selection in mammalian α-defensins. (*a*) Sites predicted to be under positive selection are highlighted in the multiple sequence alignment and (*b*) the dimer structure of the mature HNP3 peptide (PDB entry = 1DFN). Sites shown in red are those sites predicted to be under positive selection (model 8). Posterior probabilities for these sites are all greater than 0.95. Sites shown in blue are the sites that are 100% conserved across all OTUs. The mature antimicrobial peptide for HNP3 is highlighted. The mature peptides for other α-defensins may differ in size slightly at either terminus. The numbering of residues in the mature peptide of HNP3 in (*a*) corresponds to the numbering assigned to the HNP3 structure in the Protein Data Bank (PDB) database (http://www.rcsb.org/pdb/) (*b*). Note that residue 2 in (*a*) and (*b*) corresponds to residue 63 in table 3, and so on. The structure of HNP3 was displayed using RasMol version 2.7.2.1(http://www.openrasmol.org/software/rasmol/).

Table 2
Evidence of Adaptive Evolution Among Sites in Mammalian α-Defensins

Model	P	Parameters	ℓ	d_N/d_S	Positively Selected Sites
M0: one ratio	1	$\omega = 0.9342$	−3888.76	$= \omega$	
M1: neutral	1	$p_0 = 0.108$, $\omega_0 = 0$ $p_1 = 0.892$, $\omega_1 = 1$	−3837.45	0.8918	
M2: selection	3	$p_0 = 0.107$, $\omega_0 = 0$ $p_1 = 0.670$, $\omega_1 = 1$ $p_2 = 0.223$, $\omega_2 = 4.263$	−3751.88	1.6211	62,63,68,69,70,72,73,74,77,78,82,84,86,87,89
M3: discrete	5	$p_0 = 0.188$, $\omega_0 = 0.075$ $p_1 = 0.585$, $\omega_1 = 0.821$ $p_2 = 0.227$, $\omega_2 = 3.386$	−3740.19	1.264	25,62,63,68,69,70,72,73,74,77,78,82,84,86,87,89
M7: beta	2	$p = 0.479$, $q = 0.273$	−3797.38	0.6374	
M8: beta and ω	4	$p_0 = 0.791$, $p = 0.452$, $q = 0.230$ $p_1 = 0.209$, $\omega = 3.513$	–3738.80	1.258	63,68,69,70,72,73,74,77,78,82,84,86,87,89

Discussion

In this study, we have detected positive selection at several amino acid sites located in the active antimicrobial peptide region of α-defensins. It is likely that as mammals evolved to occupy new niches, they were faced with a new range of microbial pathogens. Evolution of antimicrobial peptides with new sensitivities capable of targeting novel infectious agents would confer a selective advantage. Our results are consistent with adaptive evolution of the α-defensins in response to such a challenge and the evolution of specificity for different microbes. There is experimental evidence that α-defensin peptides are diverse in their potency against different pathogens. The human neutrophil defensins HNP1, HNP2, and HNP3 are not active against the oral gram-negative bacterium *Actinobacillus actinomycetemcomitans* (Miyasaki et al. 1990*a*), whereas rabbit NP-1 shows strong activity (Miyasaki et al. 1990*b*). Trophozoites of *Giardia lamblia* are very sensitive to mouse cryptdins 2 and 3 but not to cryptdins 1 and 6 (Aley et al. 1994). Notably, the giardicidal activity has been attributed to the presence of a positively charged arginine residue at position 15 in the active peptide (Ouellette and Bevins 2001). In our study, this site is predicted to be under adaptive evolution. Rat and rabbit defensins have also been shown to have variable activities against a panel of gram-positive and gram-negative bacteria (Kohashi et al. 1992).

Small changes in the primary structure of these molecules can have enormous effect on their potency. HNP1 and HNP3 differ only by one residue at the N-terminus of the mature peptide, and, yet, HNP1 exhibits potent activity against *Candida albicans*, whereas HNP3 has little effect. (Lehrer et al. 1988; Raj, Antonyraj, and Karunakaran 2000). This site was also predicted to be subject to positive selection in our study. Similarly, the activity of mouse cryptdin 4, the most potent of the mouse α-defensins against a broad spectrum of microbes (Ouellette et al. 1994), is dependent on the presence of one or two residues at the N-terminus of the active peptide. Removal of these residues can totally eliminate the antimicrobial activity of cryptdin 4 (Ouellette et al. 2000).

It is probable that the evolution of different mechanisms of interaction with microbial membranes has been important in the evolution of specificity of particular α-defensins for particular microbes. Human HNP2 forms stable multimeric pores in model membranes (Wimley, Selsted, and White 1994), whereas rabbit NP-1 does not, but permeabilizes membranes by creating large, short-lived defects (Hristova, Selsted, and White 1996). These alternative mechanisms are likely to have varying effects, depending on the constitution of the membrane, which is variable among microbes.

Many microbes have evolved mechanisms that attempt to evade and subvert the actions of antimicrobial molecules (Ganz 2001). It is likely that this ongoing "arms race" with microbes has been a significant force driving the adaptive evolution of the α-defensins. For example, the human pathogens, *Pseudomonas aeruginosa, Enterococcus faecalis*, and *Streptococcus pyogenes* release dermatan sulphate, a compound that binds to and neutralizes HNP1 (Schmidtchen, Frick, and Bjorck 2001). Furthermore, the pathogen *Salmonella enterica* can elicit a decrease in the expression of mouse cryptdins (Salzman et al. 2003*a*).

In this study we have provided evidence that several amino acid sites in the active peptide of mammalian α-defensins are under positive Darwinian evolution. This work will assist in the design of in vitro analyses of functionally and structurally relevant sites. By synthetically changing the residues at sites predicted to be under positive selection, which are the sites most likely to be of functional importance, it is possible to alter the activity of these molecules against particular pathogens and gain insight into their mechanisms of activity.

Positive selection in the mature antimicrobial region of other antimicrobial peptide families has also been demonstrated. Evidence of positive selection has been detected in primate β-defensins (Boniotto et al. 2003; Semple, Rolfe, and Dorin 2003), murine β-defensins (Morrison et al. 2003), the *Drosophila* andropin antibac-

Table 3
Likelihood Ratio Test (LRT) to Detect Adaptive Evolution

Models	$2\Delta\ell$	χ^2 Value	df	*P*-value
M1 versus M2	2(–3837.45 −3751.88)	171.14	2	<0.001
M0 versus M3	2(–3888.76 −3740.19)	297.14	4	<0.001
M7 versus M8	2(–3797.38 −3738.80)	117.16	2	<0.001

terial peptide (Date-Ito et al. 2002), and in a number of amphibian antimicrobial peptides (Duda, Vanhoye, and Nicolas 2002). The detection of positive selection in these important modulators of the innate immune response provides evidence that despite the evolution of the adaptive immune response in jawed vertebrates more than 450 MYA, the innate immune response continues to function as a critical element in host defense.

Acknowledgments

We would like to thank two anonymous reviewers for helpful comments. This research was supported by the Food Institutional Research Measure (F.I.R.M.) Grant 01/R&D/D/135 from the Irish Department of Agriculture, Food and Rural Development.

Literature Cited

Aley, S. B., M. Zimmerman, M. Hetsko, M. E. Selsted, and F. D. Gillin. 1994. Killing of *Giardia lamblia* by cryptdins and cationic neutrophil peptides. Infect. Immun. **62**:5397–5403.

Ayabe, T., D. P. Satchell, P. Pesendorfer, H. Tanabe, C. L. Wilson, S. J. Hagen, and A. J. Ouellette. 2002. Activation of Paneth cell alpha-defensins in mouse small intestine. J. Biol. Chem. **277**:5219–5228.

Bateman, A., A. Singh, C. Shustik, W. M. Mars, and S. Solomon. 1991. The isolation and identification of multiple forms of the neutrophil granule peptides from human leukemic cells. J. Biol. Chem. **266**:7524–7530.

Boniotto, M., A. Tossi, M. DelPero, S. Sgubin, N. Antcheva, D. Santon, J. Masters, and S. Crovella. 2003. Evolution of the beta defensin 2 gene in primates. Genes Immun. **4**:251–257.

Date-Ito, A., K. Kasahara, H. Sawai, and S. I. Chigusa. 2002. Rapid evolution of the male-specific antibacterial protein andropin gene in *Drosophila*. J. Mol. Evol. **54**:665–670.

Duda, T. F. Jr., D. Vanhoye, and P. Nicolas. 2002. Roles of diversifying selection and coordinated evolution in the evolution of amphibian antimicrobial peptides. Mol. Biol. Evol. **19**:858–864.

Eisenhauer, P. B., and R. I. Lehrer. 1992. Mouse neutrophils lack defensins. Infect. Immun. **60**:3446–3447.

Fares, M. A., S. F. Elena, J. Ortiz, E. Barrio, and A. Moya. 2002. A sliding window-based method to detect selective constraints in protein-coding genes and its application to RNA viruses. J. Mol. Evol. **55**:509–521.

Frye, M., J. Bargon, N. Dauletbaev, A. Weber, T. O. Wagner, and R. Gropp. 2000. Expression of human alpha-defensin 5 (HD5) mRNA in nasal and bronchial epithelial cells. J. Clin. Pathol. **53**:770–773.

Ganz, T. 2001. Fatal attraction evaded: how pathogenic bacteria resist cationic polypeptides. J. Exp. Med. **193**:F31–F34.

Ganz, T., M. E. Selsted, D. Szklarek, S. S. Harwig, K. Daher, D. F. Bainton, and R. I. Lehrer. 1985. Defensins: natural peptide antibiotics of human neutrophils. J. Clin. Invest. **76**:1427–1435.

Ghosh, D., E. Porter, B. Shen, S. K. Lee, D. Wilk, J. Drazba, S. P. Yadav, J. W. Crabb, T. Ganz, and C. L. Bevins. 2002. Paneth cell trypsin is the processing enzyme for human defensin-5. Nat. Immunol. **3**:583–590.

Goldman, N., and Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. **11**:725–736.

Hill, C. P., J. Yee, M. E. Selsted, and D. Eisenberg. 1991. Crystal structure of defensin HNP-3, an amphiphilic dimer: mechanisms of membrane permeabilization. Science **251**:1481–1485.

Hristova, K., M. E. Selsted, and S. H. White. 1996. Interactions of monomeric rabbit neutrophil defensins with bilayers: comparison with dimeric human defensin HNP-2. Biochemistry **35**:11888–11894.

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

Jones, D. E., and C. L. Bevins. 1992. Paneth cells of the human small intestine express an antimicrobial peptide gene. J. Biol. Chem. **267**:23216–23225.

———. 1993. Defensin-6 mRNA in human Paneth cells: implications for antimicrobial peptides in host defense of the human bowel. FEBS Lett. **315**:187–192.

Kagan, B. L., M. E. Selsted, T. Ganz, and R. I. Lehrer. 1990. Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. Proc. Natl. Acad. Sci. USA **87**:210–214.

Kohashi, O., T. Ono, K. Ohki et al. (8 co-authors) 1992. Bactericidal activities of rat defensins and synthetic rabbit defensins on *Staphylococci, Klebsiella pneumoniae* (Chedid, 277, and 8N3), *Pseudomonas aeruginosa* (mucoid and nonmucoid strains), *Salmonella typhimurium* (Ra, Rc, Rd, and Re of LPS mutants) and *Escherichia coli*. Microbiol. Immunol. **36**:369–380.

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**:1244–1245.

Lehrer, R. I., T. Ganz, D. Szklarek, and M. E. Selsted. 1988. Modulation of the in vitro candidacidal activity of human neutrophil defensins by target cell metabolism and divalent cations. J. Clin. Invest. **81**:1829–1835.

Li, W. H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol **36**:96–99.

Liao, D. 1999. Concerted evolution: molecular mechanism and biological implications. Am. J. Hum. Genet. **64**:24–30.

Miyasaki, K. T., A. L. Bodeau, T. Ganz, M. E. Selsted, and R. I. Lehrer. 1990*a*. In vitro sensitivity of oral, gram-negative, facultative bacteria to the bactericidal activity of human neutrophil defensins. Infect. Immun. **58**:3934–3940.

Miyasaki, K. T., A. L. Bodeau, M. E. Selsted, T. Ganz, and R. I. Lehrer. 1990*b*. Killing of oral, gram-negative, facultative bacteria by the rabbit defensin, NP-1. Oral Microbiol. Immunol. **5**:315–319.

Morrison, G. M., C. A. Semple, F. M. Kilanowski, R. E. Hill, and J. R. Dorin. 2003. Signal sequence conservation and mature peptide divergence within subgroups of the murine beta-defensin gene family. Mol. Biol. Evol. **20**:460–470.

Nei, M., X. Gu, and T. Sitnikova. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc. Natl. Acad. Sci. USA **94**:7799–7806.

Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics **148**:929–936.

Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. **302**:205–217.

Ouellette, A. J., and C. L. Bevins. 2001. Paneth cell defensins and innate immunity of the small bowel. Inflamm. Bowel Dis. **7**:43–50.

Ouellette, A. J., D. Darmoul, D. Tran, K. M. Huttner, J. Yuan, and M. E. Selsted. 1999. Peptide localization and gene structure of cryptdin 4, a differentially expressed mouse Paneth cell alpha-defensin. Infect. Immun. **67**:6643–6651.

Ouellette, A. J., M. M. Hsieh, M. T. Nosek, D. F. Cano-Gauci, K. M. Huttner, R. N. Buick, and M. E. Selsted. 1994. Mouse

Paneth cell defensins: primary structures and antibacterial activities of numerous cryptdin isoforms. Infect. Immun. **62**:5040–5047.

Ouellette, A. J., S. I. Miller, A. H. Henschen, and M. E. Selsted. 1992. Purification and primary structure of murine cryptdin-1, a Paneth cell defensin. FEBS Lett. **304**:146–148.

Ouellette, A. J., D. P. Satchell, M. M. Hsieh, S. J. Hagen, and M. E. Selsted. 2000. Characterization of luminal Paneth cell alpha-defensins in mouse small intestine: attenuated antimicrobial activities of peptides with truncated amino termini. J. Biol. Chem. **275**:33969–33973.

Porter, E. M., L. Liu, A. Oren, P. A. Anton, and T. Ganz. 1997. Localization of human intestinal defensin 5 in Paneth cell granules. Infect. Immun. **65**:2389–2395.

Quayle, A. J., E. M. Porter, A. A. Nussbaum, Y. M. Wang, C. Brabec, K. P. Yip, and S. C. Mok. 1998. Gene expression, immunolocalization, and secretion of human defensin-5 in human female reproductive tract. Am. J. Pathol. **152**:1247–1258.

Raj, P. A., K. J. Antonyraj, and T. Karunakaran. 2000. Large-scale synthesis and functional elements for the antimicrobial activity of defensins. Biochem. J. **347**(pt3):633–641.

Raj, P. A., and A. R. Dentino. 2002. Current status of defensins and their role in innate and adaptive immunity. FEMS Microbiol. Lett. **206**:9–18.

Salzman, N. H., M. M. Chou, H. de Jong, L. Liu, E. M. Porter, and Y. Paterson. 2003*a*. Enteric salmonella infection inhibits Paneth cell antimicrobial peptide expression. Infect. Immun. **71**:1109–1115.

Salzman, N. H., D. Ghosh, K. M. Huttner, Y. Paterson, and C. L. Bevins. 2003*b*. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. Nature **422**:522–526.

Satchell, D. P., T. Sheynis, Y. Shirafuji, S. Kolusheva, A. J. Ouellette, and R. Jelinek. 2003. 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.

Schmidtchen, A., I. M. Frick, and L. Bjorck. 2001. Dermatan sulphate is released by proteinases of common pathogenic bacteria and inactivates antibacterial alpha-defensin. Mol. Microbiol. **39**:708–713.

Selsted, M. E., D. Szklarek, and R. I. Lehrer. 1984. Purification and antibacterial activity of antimicrobial peptides of rabbit granulocytes. Infect. Immun. **45**:150–154.

Semple, C. A., M. Rolfe, and J. R. Dorin. 2003. Duplication and selection in the evolution of primate beta-defensin genes. Genome Biol. **4**:R31.

Singh, A., A. Bateman, Q. Z. Zhu, S. Shimasaki, F. Esch, and S. Solomon. 1988. Structure of a novel human granulocyte peptide with anti-ACTH activity. Biochem. Biophys. Res. Commun. **155**:524–529.

Suzuki, Y., and T. Gojobori. 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. **16**:1315–1328.

Suzuki, Y., and M. Nei. 2002. Simulation study of the reliability and robustness of the statistical methods for detecting positive selection at single amino acid sites. Mol. Biol. Evol. **19**:1865–1869.

Svinarich, D. M., N. A. Wolf, R. Gomez, B. Gonik, and R. Romero. 1997. Detection of human defensin 5 in reproductive tissues. Am. J. Obstet. Gynecol. **176**:470–475.

Wilson, C. L., A. J. Ouellette, D. P. Satchell, T. Ayabe, Y. S. Lopez-Boado, J. L. Stratman, S. J. Hultgren, L. M. Matrisian, and W. C. Parks. 1999. Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. Science **286**:113–117.

Wimley, W. C., M. E. Selsted, and S. H. White. 1994. Interactions between human defensins and lipid bilayers: evidence for formation of multimeric pores. Protein Sci. **3**:1362–1373.

Wu, Z., A. Prahl, R. Powell, B. Ericksen, J. Lubkowski, and W. Lu. 2003. 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.

Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**:555–556.

———. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. **15**:568–573.

———. 2002. Inference of selection from multiple species alignments. Curr. Opin. Genet. Dev. **12**:688–694.

Yang, Z., and R. Nielsen. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J. Mol. Evol. **46**:409–418.

Yang, Z., R. Nielsen, N. Goldman, and A. M. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155**:431–449.

Zhang, L., W. Yu, T. He et al. (13 co-authors). 2002. Contribution of human alpha-defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor. Science **298**:995–1000.

Narvya Saitou, Associate Editor

Accepted December 10, 2003