VERSITA

Biologia **64**/4: 819—825, 2009
Section Cellular and Molecular Biology
DOI: 10.2478/s11756-009-0117-4

Positive selection of three chitinase genes of the family 18 of glycoside hydrolases in mammals

Beibei He[1], Li Wang[2], Jinhong Wang[1], Gang Li[1]* & Shuyi Zhang[1]

[1]*School of Life Science, East China Normal University, Shanghai 200062, People's Republic of China; e-mail: li_gang1978@yahoo.com.cn*
[2]*College of Animal Science, Yunnan Agricultural University, Kunming 650201, People's Republic of China*

Abstract: The digestive enzyme chitinase degrades chitin, and is found in a wide range of organisms, from prokaryotes to eukaryotes. Although mammals cannot synthesize or assimilate chitin, several proteins of the glycoside hydrolase (GH) chitinase family GH18, including some with enzymatic activity, have recently been identified from mammalian genomes. Consequently, there is growing interest in molecular evolution of this family of proteins. Here we report on the use of maximum likelihood methods to test for evidence of positive selection in three genes of the chitinase family GH18, all of which are found in mammals. These focal genes are CHIA, CHIT1 and CHI3L1, which encode the chitinase proteins acidic mammalian chitinase, chitotriosidase and cartilage protein 39, respectively. The results of our analyses indicate that each of these genes has undergone independent selective pressure in their evolution. Additionally, we have found evidence of a signature of positive natural selection, with most sites identified as being subject to adaptive evolution located in the catalytic domain. Our results suggest that positive selection on these genes stems from their function in digestion and/or immunity.

Key words: protein evolution; glycoside hydrolase family GH18; chitinase; positive selection.

Abbreviations: CHIA, chitinase A; CHIT1, chitinase 1; CHI3L1, chitinase 3-like 1; AMCase, acidic mammalian chitinase; HC-gp39, human cartilage glycoprotein 39; OVGP1, oviduct glycoprotein 1.

Introduction

The polysaccharide chitin is widespread in nature, and is the major constituent of the exoskeletons of arthropods. Digestive enzymes that degrade chitin (chitinases) also occur widely, and have been found in taxa as diverse as bacteria, fungi, plants, fish and mammals (Flach et al. 1992; Tharanathan & Kittur 2003). Based on their sequence similarity, chitinases have been classified as members of two families of glycoside hydrolases (GHs; http://www.cazy.org/), GH18 and GH19, which have different structures and catalytic mechanisms (Henrissat 1991; Fukamizo 2000). For example, the GH18 family deals with chitins using a substrate, while GH19 family can degrade chitins directly, without the substrate. The human genome contains eight genes that encode proteins belonging to the family GH18: CHIA, CHIT1, CHI3L1, CHI3L2, LOC149620, OVGP1, CTBS and CHID1. With the exception of CHID1, which is found on chromosome 11p15.5, the others are all located on chromosome 1 (Henrissat 1999; Funkhouser & Aronson 2007). The chitotriosidase and acidic mammalian chitinase (AMCase), which are encoded by CHIA and CHIT1, respectively, are considered to be true chitinases, whereas several other highly homologous mammalian proteins, termed chi-lectins (Renkema et al. 1998; Houston et al. 2003; Bussink et al. 2006), lack chitinase activity due to substitutions in their key catalytic residues.

All these proteins, belonging to GH18 family, are characterised by a signal peptide, a glycohydrolase domain and a chitin-binding domain (Bleau et al. 1999). Additionally, members of the GH18 family usually contain a strongly conserved $(\beta/\alpha)_8$-barrel, that forms part of the glycohydrolase domain. The chitin-binding domain, classified as the carbohydrate-binding module family CBM14, is located in the carboxyl terminal region of the protein (Stam et al. 2005). Until now, the biological function of most human GH18 proteins has remained unclear, though multiple activities have been identified, including nutrient utilization and growth-related turnover of chitinous structures. This diversification of functionality might result from past gene duplication events, because the present study shows that GH18 family evolved by decline and expansion which is associated with the selective force in speciation. However, the GH18 family is known to have a long evolutionary history (Funkhouser & Aronson 2007). Indeed,

* Corresponding author

Fujimoto et al. (2002) found two active chitinases (AMCase and chitotriosidase) in the amphibian *Xenopus*, and, therefore, suggested that gene duplication must have occurred very early in the evolution of these proteins. By comparison, mutations that led to the evolution of chi-lectins are more likely to have occurred much more recently. Recent results indicate that mammalian chitinases play an important role in innate immunity, such as host defense (Kasprzewska 2003). Nei et al. (1997) suggested that members of the family GH18, like other genes implicated in the vertebrate immune system, have experienced an unstable evolutionary history of frequent duplications and losses, termed 'birth-death evolution'.

Although Swanson et al. (2001) reported positive natural selection in the chitolectin protein oviduct glycoprotein 1 (OVGP1) in mammals, this was linked to this protein's role in reproduction. In this study, we test for changes in selection pressure in the evolutionary history of three GH18 family genes that encode chitinase proteins in mammals and try to test the relationship between the natural selection and functional divergence, such as the adaption to the different immunizing antigen.

Material and methods

We obtained sequence data of the chitinase genes CHIA, CHIT1 and CHI3L1 from two databases. First, we used the tBlastn tool (Altschul et al. 1997) to search GenBank (Benson et al. 2008) for orthologous gene sequences of mammals, using the amino acid sequences of human as the input data. Secondly, we downloaded sequences from Ensembl (Flicek et al. 2008). Accession numbers are listed in Table 1.

Amino acid sequence alignments were undertaken with the software BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html), and reverse translated to give a nucleotide alignment. Phylogenetic tree reconstruction was undertaken using the software MEGA3.1 (Kumar et al. 2004) with the amino acid substitution model and Poisson correction. The robustness of the tree was tested by bootstrapping, and nodes with bootstrap values of less than 50% were treated as paraphyletic groups.

The three-dimensional structures of studied GH18 proteins were retrieved from the Protein Data Bank (Berman et al. 2000).

Comparisons of the number of synonymous substitutions per synonymous site (dS) versus the number non-synonymous substitutions per non-synonymous site (dN) are informative in inferring the processes driving molecular evolution. A dN/dS (ω) ratio of >1 indicates positive Darwin selection, whereas a ω ratio of <1 is typically considered evidence of purifying selection, and a ratio of around one denotes a neutral evolution. We estimated rates of substitution using the software CODEML in the package PAML 3.15 (Yang 1997). We tested both branch and site models. Firstly, we tested for evidence of variation in the rate of substitution among different branches by comparing the likelihood of a free ratio model versus a one ratio model. The former assumes an independent value of ω on each branch, while this is fixed in a one ratio model. These two models will present the different maximum likelihood values. We used the likelihood ratio test to compare these two models. The one ratio model should be null hypothesis, and free ratio was assumed as the alternative model. If the results of likelihood ratio test showed that the free ratio model is better than the one ratio model significantly, it can be concluded that the assumption of different branches in the phylogenetic tree owning the independent evolutional ratio fits the real evolution case better than the assumption of identical evolutional ratio for all taxon. Moreover, the comparison of free ratio and one ration can give the indication for the special accelerated evolution or purified selection for some interested lineages at the same time.

For the site models, we undertook two comparisons: model M0 (one ratio) versus model M3 (discrete), and model M7 (beta) versus model M8 (beta&ω). Model M0 assumes an equal ω ratio for all codons in a gene, while model M3 (with four parameters) captures heterogeneous ratios among sites, and can also detect positive selection if the ω ratio is >1. To compare model M0 and model M3, we used a log-likelihood test with four degrees of freedom. Model M7 and model M8 were compared to identify the codons under positive selection, with two degrees of freedom. These two comparisons can be used for testing whether the different amino acid sits under the different natural selection pressure and identify the locations of amino acid sites which were selected positively.

Table 1. The information of the sequences retrieved from GenBank and Ensembl.

Gene	Species	Accession numbers
CHI3L1	*Erinaceus europaeus*	ENSEEUT00000011096
	Myotis lucifugus	ENSMLUT00000010322
	Homo sapiens	ENST00000255409
	Macaca mulatta	XM_001103739
	Canis familiaris	ENSCAFT00000037736
	Bos taurus	ENSBTAT00000024253
	Mus musculus	ENSMUST00000082060
	Pan troglodytes	ENSPTRT00000003404
	Pongo pygmaeus	CR858592
	Capra hircus	AY081150
	Sus scrofa	NM_001080219
	Ovis aries	AY392761
	Rattus norvegicus	ENSRNOT00000004444
CHIT1	*Homo sapiens*	ENST00000255427
	Mus musculus	ENSMUST00000112256
	Macaca mulatta	ENSMMUT00000012171
	Rattus norvegicus	ENSRNOT00000035658
	Monodelphis domestica	ENSMODT00000001312
	Microcebus murinus	ENSMICT00000002309
	Oryctolagus cuniculus	ENSOCUT00000011025
	Ochotona princeps	ENSOPRT00000014806
	Pan troglodytes	ENSPTRT00000047221
CHIA	*Homo sapiens*	ENST00000343320
	Microcebus murinus	ENSMICT00000011956
	Tupaia belangeri	ENSTBET00000006966
	Erinaceus europaeus	ENSEEUT00000005347
	Rattus norvegicus	ENSRNOT00000039871
	Mus musculus	ENSMUST00000079132
	Bos taurus	BC102931
	Sus scrofa	AK238270
	Monodelphis domestica	ENSMODT00000001818
	Canis familiaris	XM_537030

Results

We examined the molecular evolution in mammals of three genes (CHIA, CHI3L1 and CHIT1) from the fam-

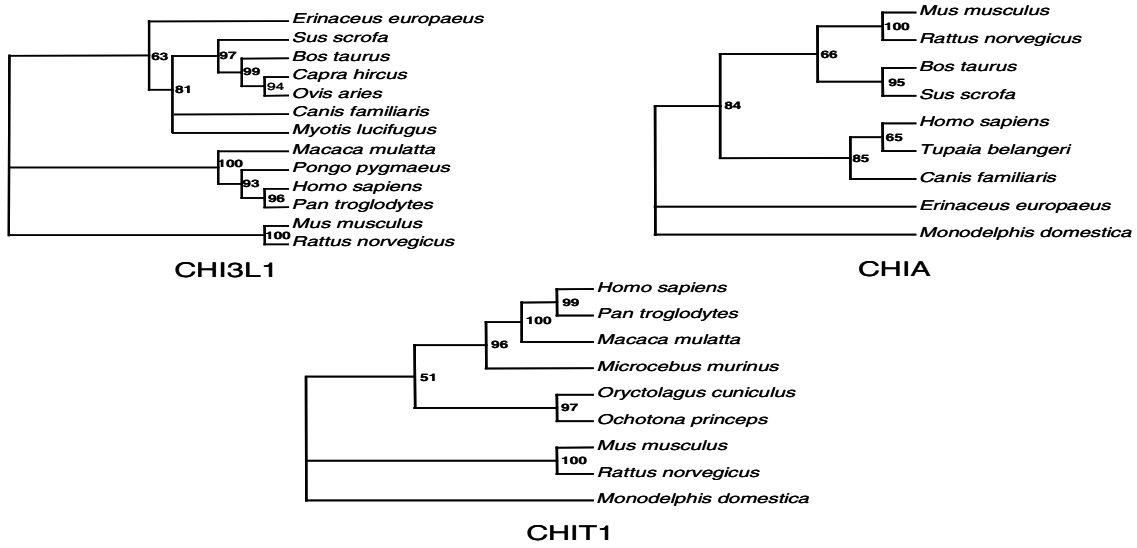

Fig. 1. The gene trees for three genes belonging to the GH18 family. The topologies of the trees were obtained using the software MEGA3.1 under amino acid substitution model Poisson correction. The numbers near each node are the values of bootstrap test. These trees were used as the initial tree for site model test.

Table 2. Estimated Parameters of the site-specific models.[a]

Gene	Model	ℓ	Parameters	Positively selected sites
CH3L1	one-ratio (M0)	−5460.27	ω= 0.191	
	M3: discrete (k=3)	−5369.00	P_0=0.608, P_1=0.338, (P_2=0.054) ω_0=0.049, ω_1=0.393, ω_2=1.407	**7Q**, 108K, *144R*, 162I, **167P**, **169K**, 171Q, **224R**, 242G, *291A*, **304R**, **312L**, **339Q**, 366Q, **369R**
	M7: beta	−5372.60	p=0.405, q=1.386	
	M8: beta&ω	−5368.81	P_0=0.955, (P_1=0.045) p=0.543, q=2.400, ω=1.467	7Q, **144R**, *162I*, 167P, **169K**, 224R, **291A**, 304R, 312L, 339Q, 369R
CHIT1	one-ratio (M0)	−6667.60	ω=0.235	
	M3: discrete (k=3)	−6519.42	P_0=0.540, P_1=0.374, (P_2=0.085) ω_0=0.054, ω_1=0.397, ω_2=1.698	36Q, 38E, 69T, 71W, *75T*, 85K, 125R, 150V, 176E, **186A**, 190Y, *241L*, 249Q, **286T**, **298G**, 300M, 311G, 314K, 366A, **389Y**, *390L*, 392S, 394T, **398E**, **401K**, **402P**, *407E*, **416Q**, 432R, 441A, 452T, 458N, **466N**
	M7: beta	−6531.30	p=0.413, q=1.072	
	M8: beta&ω	−6519.42	P_0=0.925, (P_1=0.075), p=0.639, q=2.427, ω=1.764	**69T**, **71W**, **75T**, **125R**, 186A, 190Y, 241L, 286T, **314K**, 389Y, 390L, 407E, 416Q, **432R**, **452T**, **458N**
CHIA	one-ratio (M0)	−6138.04	ω=0.154	
	M3: discrete (k=3)	−6056.31	P_0=0.519, P_1=0.442, (P_2=0.039) ω_0=0.015, ω_1=0.270, ω_2=1.533	16L, 25T, 62Q, *78Q*, **317G**, 343D, *381S*, 385K, 390Q, 406A, 407A, 418S, **420S**, *451V*
	M7: beta	−6062.73	p=0.347, q=1.595	
	M8: beta&ω	−6057.12	P_0=0.972, (P_1=0.028) p=0.499, q=2.906, ω=1.720	25T, *62Q*, **78Q**, 317G, *343D*, **381S**, **390Q**, **406A**, 420S, **451V**

[a] Substitutions are based on the human sequence as a reference. Sites are formatted based on their Bayesian posterior probabilities as follows: 0.5–0.7, unformatted; 0.7–0.9, in bold; 0.9–0.95, in italics; and >0.95, underlined.

ily GH18 (Henrissat 1991). For each gene, a phylogenetic tree based on amino acid sequences was used as the input in the analysis of molecular evolution (Fig. 1). Maximum-likelihood estimates of ω, based on a single ratio across all branches, were 0.19, 0.28 and 0.17 for CHIA, CHI3L1 and CHIT1, respectively. This indicates that all of these genes are subject to purifying selection. However, free-ratio models of evolution were found to be significantly more likely for each gene, suggesting that selective pressures vary across branches in the mammalian tree (data not shown).

In the discrete site model (M3), three classes of sites were classified, for which the ω ratio was estimated. This was used to test for codons under positive selection ($dN/dS > 1$), and was compared to model M0, which served as a null model (see Table 2 for results). Significant model improvement supports positive selection, with 5% of codons of the CHI3L1 gene assigned an estimated ω ratio of 1.4. Similarly, model M3 tests of the CHIT1 and CHIA genes also indicated a signature of positive selection on codon sites (Table 3). Around 8.5% and 4.4% of sites were identified as being under

Table 3. The Likelihood Ratio test results of different models comparison.

Gene	Models compared	$2\Delta\ell$	df	P value
CHI3L1	M3 (k=3) vs. M0	199.23	4	$P<0.001$
	M7 vs. M8	8.25	2	$0.01<P<0.05$
CHIT1	M3 (k=3) vs. M0	300.87	4	$P<0.001$
	M7 vs. M8	22.20	2	$P<0.001$
CHIA	M3 (k=3) vs. M0	163.4630	4	$P<0.001$
	M7 vs. M8	11.2236	2	$P<0.01$

positive selection for the CHIT1 ($\omega = 1.467$) and CHIA gene ($\omega = 1.377$), respectively.

A comparison of model M7 and model M8 can be used to identify amino acid sites under positive selection using the improved Bayes empirical Bayes method of estimating posterior probabilities (Yang et al. 2005). This was undertaken for all three genes, and all suggested that model M8 was a significantly better fit to the data than model M7 (proportions of codon sites of each gene identified as being under positive selection are given in Tables 2 and 3). Eleven amino acid residues sites of the human cartilage glycoprotein 39 (HC-gp39) were identified as positively selected. Glutamine on the 7^{th} site belongs to the signal peptide of the protein (Hakala et al. 1993; Fusetti et al. 2003; Houston et al. 2003), and other sites locate on the catalytic domain, including ten amino residues (144R, 162I, 167P, 169K, 224R, 291A, 304R, 312L, 339Q, 369R; the HC-gp39 sequence was used as the reference). Three out of ten (144R, 162I, 339Q) sites are from the α-helix and the other three sites (167P, 304R, 312L) are involved in the framework of the β-turn (Boot et al. 1995; Fusetti et al. 2002). Moreover, 16 amino acid sites with the positive selection signature of chitotriosidase included three amino acid residues (432R, 452T, 458N) located on the chitin-binding domain, as well as others (69T, 71W, 75T, 125R, 186A, 190Y, 241L, 286T, 314K, 389Y, 390L, 407E, 416Q) belonging to the catalytic domain. Similarly, the model M8 identified 10 amino acid sites in AMCase, 9 amino residues on catalytic domain (25T, 62Q, 78Q, 317G, 343D, 381S, 390Q, 406A, 420S) and only one (451V) on the chitin-binding domain (Saito et al. 1999; Boot et al. 2001).

Discussion

An increasing body of work suggests that proteins belonging to the GH18 family, especially those with enzymatic activity, have a range of important functions in mammals, from food digestion and innate immunity to the maintenance of natural physiological function (Elias et al. 2005; Suzuki et al. 2002). The results of phylogenetic analyses based on the nematodes, fruit flies and mammals have divided the members of this family into three groups: the chitobiases, chitinases/chitolectins, and stabilin-1 interacting chitolectins (Funkhouser & Aronson 2007). Moreover, mammalian-specific orthologous GH18 genes have been further classified into four clades: the acidic chitinases with related chitolectins (including CHIA, CHI3L4 and CHI3L3), chitotriosidase (CHIT1), the non-active chitinase-like protein CHI3L1, and the oviduct glycoprotein OVGP1 (Funkhouser & Aronson 2007).

To characterize the nature of sequence evolution in three representatives of the GH18 gene family in mammals (CHIA, CHIT1, CHI3L1), we applied a suite of phylogenetic and molecular evolution analyses. For each gene, the results of our maximum likelihood methods suggest that rates of substitution differ both among branches across the tree, and also among sites within the gene. Several sites were identified as potentially being under positive selection.

Members of GH18 family possess multidomain structure with a strongly conserved $(\beta/\alpha)_8$-barrel catalytic, a signal peptide and a chitin-binding domain. Usually, chitolectins lack the chitin-binding domain (Henrissat 1999; Stam et al. 2005). All the mammalian chitinases have a conserved sequence motif (DXXDXDXE) that spans the amino acid residues from 133 to 140, with 140E identified as the key residue (van Aalten et al. 2001) that, when substituted, leads to a loss of catalytic activity (Kzhyshkowska et al. 2006). The replacement of 140E can make the functional change because the chitinases protein sequences alignment indicated that this site remains in all the chitinase catalytic proteins, such as AMCase and chitotriosidase of human (data not shown). Such a loss of chitinase activity has previously been shown in CHI3L1, resulting from the replacement of glutamic acid by leucine (Hakala et al. 1993).

AMCase and chitotriosidase are considered to be true chitinase due to their chitinase activity (Kzhyshkowska et al. 2006). They have a relatively complete structure including a signal peptide of about 22 amino residues, a catalytic $(\beta/\alpha)_8$-barrel domain of about 400 amino residues and a chitin-binding domain of at least 49 amino residues (Tjoelker et al. 2000), with the key glutamic acid residue at the position 140 (van Aalten et al. 2001). Our results suggested that a total of 10 and 16 amino acid sites were under positive selection in the genes CHIA and CHIT1, respectively. In CHIA, only one of these sites was located in the chitin-binding domain, while the other 9 belonged to the catalytic domain. Similarly, in CHIT1, 3 positively selected sites were located in chitin-binding domain and the remaining 13 occurred in the catalytic domain (see Figure 2).

Although the functions of mammalian chitinases are not fully resolved, these two proteins have been proposed to play a role in degrading chitin-containing pathogens. AMCase activity has also been linked to the pathophysiology of asthma (Zhu et al. 2004), while chitotriosidases have been associated with several diseases (Zheng et al. 2005). There is some evidence that the existence of AMCase may compensate for the lack of a functional chitotriosidase (Boot et al. 1998, 2001).

Fusetti et al. (2002) found that mammalian chi-

Fig. 2. The simple structure of AMCase, chitotriosidase and HC-gp39, human sequence as a reference. The signal peptide (about 21 amino acids) is red, the catalytic domain is no color and the chitin-binding (about 50 amino acids) is yellow. The site we identified is simply marked by the short line.

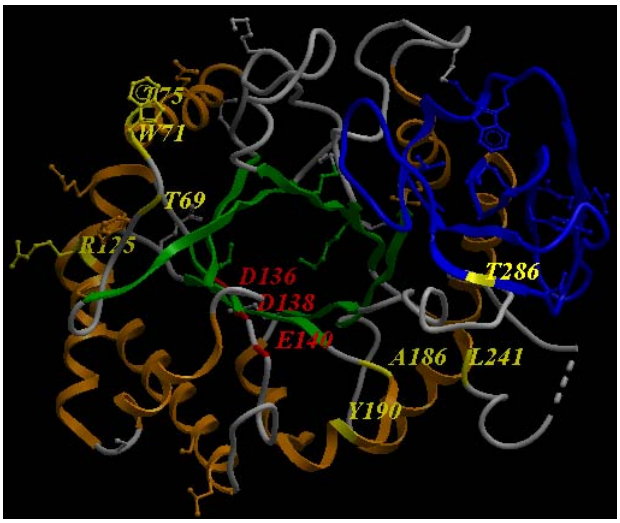

Fig. 3. Overview of human chitotriosidase structure (Fusetti et al. 2002) and the active site. The backbone is gray. The α-helices are shown in orange and the β-strands are shown in green. The α/β domain is in blue. D136, D138 and E140 are shown in red. The positive sites we identified are in yellow.

lectin proteins such as YM1 can only interact with monomeric carbohydrates, i.e. they lost the affinity to the polymeric carbohydrate. Apart from the mutation in the active site, the non-conservative substitutions (Fig. 3), such as Y24D, W72E, Y190V, W218K located in the chitin-binding cleft, are also an important reason. Through the results of the protein sequence alignment by Fusetti et al. (2002), it is possible that, like YM1, AMCase also have the substitutions in these sites (chitotriosidase as a reference). So we believe that these substitutions possibly caused the function changed, and we suggest that the sites identified in this study as being under possible positive selection are important in discerning and binding the different substrate in different loci to retain the chitinase activity. Due to the revealed multi-function of these two true chitinases, we conclude that the sites we identified in our work have a relationship with the proteins' functional diversity. It is likely that at least some of these substitutions will influence the spatial structure and function of the protein. Bussink et al. (2008) found that contrary to the chitotriosidase, most substituted

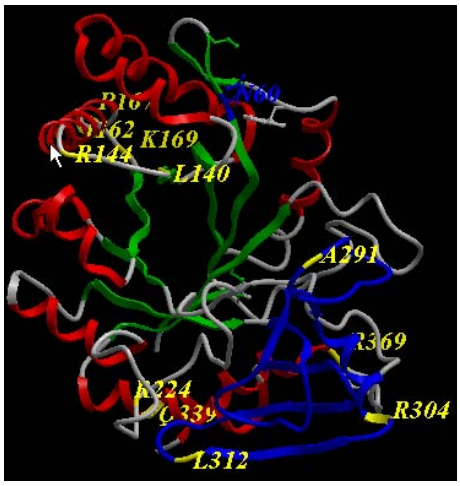

Fig. 4. Overview of HC-gp39 structure (Fusetti et al. 2003) and the active site. The backbone is gray. The α-helices are shown in orange and the β-strands are shown in green. The N-glycosylation at N60 is shown in blue and the α/β domain is also shown in blue. L140 and other positive sites we identified are shown in yellow.

residues in AMCase occur on the protein surface, leading to the suggestion that these are more likely related to adaptive stability rather than the reaction mechanism. Both interior and surface substitutions might also be important in modulating functions in immunity, particularly in the protection from fungal and other pathogens.

The CHI3L1 gene, located in the chromosome domain 1q32.1 (Rehli et al. 1997, 2003) of the human is expressed in several kinds of cells including arthritic chondrocytes (Hakala et al. 1993), activated neutrophils (Volck et al. 1998), macrophages during later stages of differentiation (Rehli et al. 1997, 2003) and the processes of inflammation, fibrosis and degeneration (Kirkpatrick et al. 1997; Boot et al. 1999; Johansen et al. 1999, 2000; Baeten et al. 2000; Volck et al. 2001; Junker et al. 2005). CHI3L1 has been shown to be linked to metabolism, tissue remodeling or fibrogenesis and schizophrenia (Johansen et al. 2000, 2001; Kawasaki et al. 2001; Volck et al. 1999, 2001; Yang et al. 2008).

The three-dimensional structure of HC-gp39 (Fig. 4) contains the eight-stranded $(\beta/\alpha)_8$-barrel catalytic domain and has the ability to bind the insoluble chitin with the aromatic residues lining in the carbohydrate-binding groove. Fusetti et al. (2003) found that among all these aromatic residues, 70E, 97G, 99W, 100N, 179S, 184T, 263R, 290E, 352W (Fig. 5) contribute to the carbohydrate binding with different loci of the N-acetylglucosamine.

We identified 11 amino acid sites (about 4.5% of all the residues) as likely to be under positive selection, and most were located in the carbohydrate-binding groove of the catalytic domain, with one more located in the signal peptide. The N-terminal of the signal peptide usually contains basic and hydrophobic amino acids, and the substation at the 7^{th} site occurs between these amino acid residues, including proline, leucine, glycine and histidine, suggesting a possible role in the signal

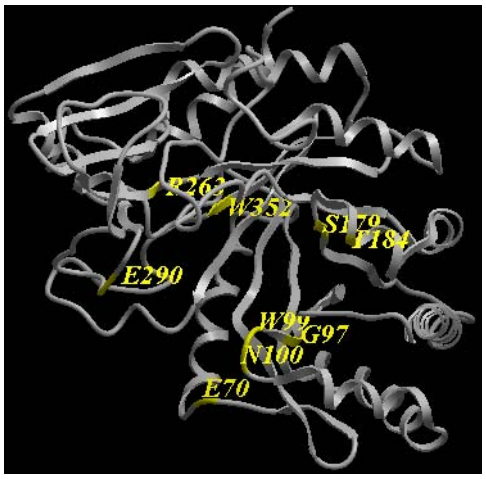

Fig. 5. Structure of binding site of HC-gp39 (Fusetti et al. 2003). The backbone is gray. The binding sites which contribute to the carbohydrate binding with different loci of the N-acetylglucosamine are shown in yellow.

peptide. Unlike AMCase and chitotriosidase, HC-gp39 lost its chitin-binding domain and distributed the binding sites in the carbohydrate-binding groove, so we suggested that the other 10 sites we identified are related to the function of binding carbohydrate. Recent research has also implicated the CHI3L1 gene in susceptibility to schizophrenia (Zhao et al. 2007).

The mammalian chitinase plays a role in immunity, several lines of evidence have indicated that genes involved in immunity and reproduction possess higher rates of evolution than other genes (Jansa et al. 2003). For example, mammalian oviduct glycoproteins, also of the GH18 family, appear to have undergone positive selection during the divergence of mammals, probably due to sperm competition and sexual selection (Swanson et al. 2001). Since the signal peptides and chitin-binding domain-coding sequences were omitted and lesser sequences were used in the protein sequences alignment, other results on the evolution of GH18 family genes in mammals have pointed to strong purifying selection (Bussink et al. 2007). Although a comparison of the catalytic domain across the whole family suggested the relative strong functional constraint during the evolution of mammals, it is still worthy to test for changes in selection pressure during the divergence of mammals.

More taxa and longer sequences were involved in our analyses. The more detailed data can provide the more precise estimates concerning the ancestral sequences of each node in the phylogenetic tree using maximum likelihood method. Applying the more data and the maximum likelihood method which can detect the evolution ratio for independent amino acid site, our results indicate the specific adaptation in each of the three genes studied, probably resulting from positive natural selection. In particular, we suggest that the catalytic domain is a hotspot for rapid change due to its vital function in a range of processes.

Acknowledgements

We thank Stephen J. Rossiter for his helpful comments on the manuscript. This work was funded by a grant under the Key Construction Program of the National "985" Project and Shanghai Priority Academic Discipline to S. Zhang.

References

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. & Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

Baeten D., Boots A.M., Steenbakkers P.G., Elewaut D., Bos E., Verheijden G.F., Berheijden G., Miltenburg A.M., Rijnders A.W., Veys E.M. & De Keyser F. 2000. Human cartilage gp-39+,CD16+ monocytes in peripheral blood and synovium: correlation with joint destruction in rheumatoid arthritis. Arthritis Rheum. **43:** 1233–1243.

Bleau G., Massicotte F., Merlen Y. & Boisvert C. 1999. Mammalian chitinase-like proteins. EXS **87:** 211–221.

Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J. & Wheeler D.L. 2008. GenBank. Nucleic Acids Res. **36 (Database issue):** D25–D30.

Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N. & Bourne P.E. 2000. The Protein Data Bank. Nucleic Acids Res. **28:** 235–242.

Boot R.G., Blommaart E.F., Swart E., Ghauharali-van der Vlugt K., Bijl N., Moe C., Place A. & Aerts J.M. 2001. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. J. Biol. Chem. **276:** 6770–6778.

Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J. & Aerts J.M. 1995. Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. J. Biol. Chem. **270:** 26252–26256.

Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J., de Meulemeester T.M., Mannens M.M. & Aerts J.M. 1998. The human chitotriosidase gene. Nature of inherited enzyme deficiency. J. Biol. Chem. **273:** 25680–25685.

Boot R.G., van Achterberg T.A., van Aken B.E., Renkema G.H., Jacobs M.J., Aerts J.M. & de Vries C.J. 1999. Strong induction of members of the chitinase family of proteins in atherosclerosis: chitotriosidase and human cartilage gp-39 expressed in lesion macrophages. Arterioscler. Thromb. Vasc. Biol. **19:** 687–694.

Bussink A.P., Speijer D., Aerts J.M. & Boot R.G. 2007. Evolution of mammalian chitinase(-like) members of family 18 glycosyl hydrolases. Genetics **177:** 959–970.

Bussink A.P., van Eijk M., Renkema G.H., Aerts J.M. & Boot R.G. 2006. The biology of the Gaucher cell: the cradle of human chitinases. Int. Rev. Cytol. **252:** 71–128.

Bussink A.P., Vreede J., Aerts J.M. & Boot R.G. 2008. A single histidine residue modulates enzymatic activity in acidic mammalian chitinase. FEBS Lett. **582:** 931–935.

Elias J.A., Homer R.J., Hamid Q. & Lee C.G. 2005. Chitinases and chitinase-like proteins in T(H)2 inflammation and asthma. J. Allergy Clin. Immunol. **116:** 497–500.

Flach J., Pilet P.E. & Jolles P. 1992. What's new in chitinase research? Experientia **48:** 701–716.

Flicek P., Aken B.L., Beal K., Ballester B., Caccamo M., Chen Y., Clarke L., Coates G., Cunningham F., Cutts T., Down T., Dyer S.C., Eyre T., Fitzgerald S., Fernandez-Banet J., Gräf S., Haider S., Hammond M., Holland R., Howe K.L., Howe K., Johnson N., Jenkinson A., Kähäri A., Keefe D., Kokocinski F., Kulesha E., Lawson D., Longden I., Megy K., Meidl P., Overduin B., Parker A., Pritchard B., Prlic A., Rice S., Rios D., Schuster M., Sealy I., Slater G., Smedley D., Spudich G., Trevanion S., Vilella A.J., Vogel J., White S., Wood M, Birney E., Cox T., Curwen V., Durbin R., Fernandez-Suarez X.M., Herrero J., Hubbard T.J., Kasprzyk A., Proctor G., Smith J., Ureta-Vidal A. & Searle S. 2008. Ensembl 2008. Nucleic Acids Res. **36 (Database issue):** D707–D714.

Fujimoto W., Kimura K. & Iwanaga T. 2002. Cellular expression of the gut chitinase in the stomach of frogs *Xenopus laevis* and *Rana catesbeiana*. Biomed. Res. **23:** 91–99.

Fukamizo T. 2000. Chitinolytic enzymes: catalysis, substrate binding, and their application. Curr. Protein Pept. Sci. **1:** 105–124.

Funkhouser J.D. & Aronson N.N., Jr. 2007. Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. BMC Evol. Biol. **7:** 96.

Fusetti F., Pijning T., Kalk K.H., Bos E. & Dijkstra B.W. 2003. Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. J. Biol. Chem. **278:** 37753–37760.

Fusetti F., von Moeller H., Houston D., Rozeboom H.J., Dijkstra B.W., Boot R.G., Aerts J.M. & van Aalten D.M. 2002. Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. J. Biol. Chem. **277:** 25537–25544.

Hakala B.E., White C. & Recklies A.D. 1993. Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. J. Biol. Chem. **268:** 25803–25810.

Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280:** 309–316.

Henrissat B. 1999. Classification of chitinases modules. EXS **87:** 137–156.

Houston D.R., Recklies A.D., Krupa J.C. & van Aalten D.M. 2003. Structure and ligand-induced conformational change of the 39–kDa glycoprotein from human articular chondrocytes. J. Biol. Chem. **278:** 30206–30212.

Jansa S.A., Lundrigan B.L. & Tucker P.K. 2003. Tests for positive selection on immune and reproductive genes in closely related species of the murine genus mus. J. Mol. Evol. **56:** 294–307.

Johansen J.S., Baslund B., Garbarsch C., Hansen M., Stoltenberg M., Lorenzen I. & Price P.A. 1999. YKL-40 in giant cells and macrophages from patients with giant cell arteritis. Arthritis Rheum. **42:** 2624–2630.

Johansen J.S., Christoffersen P., Moller S., Price P.A., Henriksen J.H., Garbarsch C. & Bendtsen F. 2000. Serum YKL-40 is increased in patients with hepatic fibrosis. J. Hepatol. **32:** 911–920.

Johansen J.S., Olee T., Price P.A., Hashimoto S., Ochs R.L. & Lotz M. 2001. Regulation of YKL-40 production by human articular chondrocytes. Arthritis Rheum. **44:** 826–837.

Junker N., Johansen J.S., Andersen C.B. & Kristjansen P.E. 2005. Expression of YKL-40 by peritumoral macrophages in human small cell lung cancer. Lung Cancer **48:** 223–231.

Kasprzewska A. 2003. Plant chitinases – regulation and function. Cell. Mol. Biol. Lett. **8:** 809–824.

Kawasaki M., Hasegawa Y., Kondo S. & Iwata H. 2001. Concentration and localization of YKL-40 in hip joint diseases. J. Rheumatol. **28:** 341–345.

Kirkpatrick R.B., Emery J.G., Connor J.R., Dodds R., Lysko P.G. & Rosenberg M. 1997. Induction and expression of human cartilage glycoprotein 39 in rheumatoid inflammatory and peripheral blood monocyte-derived macrophages. Exp. Cell. Res. **237:** 46–54.

Kumar S., Tamura K. & Nei M. 2004. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. Bioinform. **5:** 150–163.

Kzhyshkowska J., Mamidi S., Gratchev A., Kremmer E., Schmuttermaier C., Krusell L., Haus G., Utikal J., Schledzewski K., Scholtze J. & Goerdt S. 2006. Novel stabilin-1 interacting chitinase-like protein (SI-CLP) is up-regulated in alternatively activated macrophages and secreted via lysosomal pathway. Blood **107:** 3221–3228.

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

Rehli M., Krause S.W. & Andreesen R. 1997. Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation. Genomics **43:** 221–225.

Rehli M., Niller H.H., Ammon C., Langmann S., Schwarzfischer L., Andreesen R. & Krause S.W. 2003. Transcriptional regulation of CHI3L1, a marker gene for late stages of macrophage differentiation. J. Biol. Chem. **278:** 44058–44067.

Renkema G.H., Boot R.G., Au F.L., Donker-Koopman W.E., Strijland A., Muijsers A.O., Hrebicek M. & Aerts J.M. 1998. Chitotriosidase, a chitinase, and the 39–kDa human cartilage glycoprotein, a chitin-binding lectin, are homologues of family 18 glycosyl hydrolases secreted by human macrophages. Eur. J. Biochem. **251:** 504–509.

Saito A., Ozaki K., Fujiwara T., Nakamura Y. & Tanigami A. 1999. Isolation and mapping of a human lung-specific gene, TSA1902, encoding a novel chitinase family member. Gene **239:** 325–331.

Stam M.R., Blanc E., Coutinho P.M. & Henrissat B. 2005. Evolutionary and mechanistic relationships between glycosidases acting on α- and β-bonds. Carbohydr. Res. **340:** 2728–2734.

Suzuki M., Fujimoto W., Goto M., Morimatsu M., Syuto B. & Iwanaga T. 2002. Cellular expression of gut chitinase mRNA in the gastrointestinal tract of mice and chickens. J. Histochem. Cytochem. **50:** 1081–1089.

Swanson W.J., Yang Z., Wolfner M.F. & Aquadro C.F. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc. Natl. Acad. Sci. USA **98:** 2509–2514.

Tharanathan R.N. & Kittur F.S. 2003. Chitin – the undisputed biomolecule of great potential. Crit. Rev. Food Sci. Nutr. **43:** 61–87.

Tjoelker L.W., Gosting L., Frey S., Hunter C.L., Trong H.L., Steiner B., Brammer H. & Gray P.W. 2000. Structural and functional definition of the human chitinase chitin-binding domain. J. Biol. Chem. **275:** 514–520.

van Aalten D.M., Komander D., Synstad B., Gaseidnes S., Peter M.G. & Eijsink V.G. 2001. Structural insights into the catalytic mechanism of a family 18 exo-chitinase. Proc. Natl. Acad. Sci. USA **98:** 8979–8984.

Volck B., Johansen J.S., Stoltenberg M., Garbarsch C., Price P.A., Ostergaard M., Ostergaard K., Lovgreen-Nielsen P., Sonne-Holm S. & Lorenzen I. 2001. Studies on YKL-40 in knee joints of patients with rheumatoid arthritis and osteoarthritis. Involvement of YKL-40 in the joint pathology. Osteoarthritis Cartilage **9:** 203–214.

Volck B., Ostergaard K., Johansen J.S., Garbarsch C. & Price P.A. 1999. The distribution of YKL-40 in osteoarthritic and normal human articular cartilage. Scand. J. Rheumatol. **28:** 171–179.

Volck B., Price P.A., Johansen J.S., Sorensen O., Benfield T.L., Nielsen H.J., Calafat J. & Borregaard N. 1998. YKL-40, a mammalian member of the chitinase family, is a matrix protein of specific granules in human neutrophils. Proc. Assoc. Am. Physicians **110:** 351–360.

Yang M.S., Morris D.W., Donohoe G., Kenny E., O'Dushalaine C.T., Schwaiger S., Nangle J.M., Clarke S., Scully P., Quinn J., Meagher D., Baldwin P., Crumlish N., O'Callaghan E., Waddington J.L., Gill M. & Corvin A. 2008. Chitinase-3–Like 1 (CHI3L1) gene and schizophrenia: genetic association and a potential functional mechanism. Biol. Psychiatry **64:** 98–103.

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

Yang Z., Wong W.S. & Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. **22:** 1107–1118.

Zhao X., Tang R., Gao B., Shi Y., Zhou J., Guo S., Zhang J., Wang Y., Tang W., Meng J., Li S., Wang H., Ma G., Lin C., Xiao Y., Feng G., Lin Z., Zhu S., Xing Y., Sang H., St Clair D. & He L. 2007. Functional variants in the promoter region of chitinase 3–like 1 (CHI3L1) and susceptibility to schizophrenia. Am. J. Hum. Genet. **80:** 12–18.

Zheng T., Rabach M., Chen N.Y., Rabach L., Hu X., Elias J.A. & Zhu Z. 2005. Molecular cloning and functional characterization of mouse chitotriosidase. Gene **357:** 37–46.

Zhu Z., Zheng T., Homer R.J., Kim Y.K., Chen N.Y., Cohn L., Hamid Q. & Elias J.A. 2004. Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. Science **304:** 1678–1682.

Received July 29, 2008
Accepted December 31, 2008