Mechanisms for a Novel Immune Evasion Strategy in the Scabies Mite *Sarcoptes Scabiei*: A Multigene Family of Inactivated Serine Proteases

Deborah C. Holt, Katja Fischer, George E. Allen, Danny Wilson, Peter Wilson,* Robert Slade,* Bart J. Currie,† Shelley F. Walton,† and David J. Kemp

Queensland Institute of Medical Research and The Australian Center for International and Tropical Health and Nutrition, Brisbane, Australia; *Australian Genome Research Facility, University of Queensland, Brisbane, Australia; †Menzies School of Health Research, Darwin, Australia

Parasitic infestation of the skin by the mite *Sarcoptes scabiei* is a significant problem worldwide, particularly in socially disadvantaged communities. A multigene family of at least 24 homologs of a serine protease allergen have been identified in *S. scabiei*. Surprisingly, the products of all but one of these genes are predicted to be catalytically inactive, due to mutations at a critical triad of amino acids at the active site. We discuss the possibility that these genes for inactivated proteases have been conserved because they mediate a novel host defense evasion strategy that the mite has evolved as an adaptation to parasitism of the epidermis. The identification of this family, and elucidation of its value to the parasite, may present an unanticipated approach to protective vaccination. ***Key words: group 3 allergens/protease activated receptors. J Invest Dermatol 121:1419–1424, 2003***

Scabies is a parasitic infestation of skin by the mite *Sarcoptes scabiei* (Arlian, 1989; Green, 1989; Davis and Moon, 1990; Elgart, 1990). Despite the availability of pharmaceutical control measures approximately 300 million people worldwide are afflicted primarily in socially disadvantaged populations, such as indigenous communities (Currie *et al*, 1994), immunodeficient patients, and residents of nursing homes (Taplin *et al*, 1991). Classical scabies presents as a generalized pruritis, which is frequently more intense at night. An infestation occurs when the adult female mite, approximately 0.3 mm in length, forms a burrow in the skin often invisible to the naked eye. The severe itching and papular rash of the primary infestation take 4 to 6 wk to develop, during which time mite numbers can reach between 10 and 50. Thus transmission to others often occurs before treatment is sought. In contrast, reinfestation produces an immediate hypersensitivity reaction. Pruritic scabies lesions are often accompanied by streptococcal infections with significant sequelae (cellulitis, septicemia, and glomerulonephritis) (Currie and Carapetis, 2000). Scabies is also an important veterinary problem (Davis and Moon, 1990). The current wide prevalence and decreasing treatment efficacy for scabies indicate that development of novel control strategies, in the form of vaccination or immunotherapy, are warranted. A vaccine against the cattle tick *Boophilus microplus* has been developed, targeted against the gut wall antigen Bm86, which results in the dramatic lowering of tick numbers (Willadsen and Kemp, 1989). We have previously discussed the possibility of developing an analogous vaccine against scabies (Kemp *et al*, 2002). The feasibility of this goal is supported by studies where immunization with crude house dust mite extracts decreased the scabies mite burden in rabbits (Arlian *et al*, 1995). Furthermore, there is convincing epidemiologic evidence that clinical immunity to scabies develops in humans (Mellanby, 1944). Little is known, however, about the immunologic response to scabies infestation.

In contrast, allergens of the related astigmatid house dust mites have been studied extensively as they have been implicated among elicitors of the allergic response in asthma. Several of the allergens are hydrolytic enzymes secreted into the gut of the mites and are presumably involved in digestion (Chua *et al*, 1988; Stewart *et al*, 1992, 1994; Smith *et al*, 1994; Thomas and Smith, 1999). Anti-sera to allergens from the house dust mites *Dermatophagoides* cross-react with extracts of *S. scabiei* (Arlian *et al*, 1991). It is possible that immunization with a cocktail of the scabies mite homologs of house dust mite gut allergens would inhibit mite digestion and hence protect against a scabies mite infestation. In addition, scabies mites dying in burrows might release these molecules in a form accessible to the immune system and hence might boost the immune response. Certainly antibodies to scabies antigens are detectable in patients, although the identity of the antigens is not known (Falk and Bolle, 1980).

Group 3 house dust mite allergens are abundant serine proteases that are secreted into the gut of the mite. Group 3 allergens, typified by Der p 3 from *Dermatophagoides pteronyssinus* are serine proteases of the S1 family, related to chymotrypsin and trypsin (Rawlings and Barrett, 1994). They are excreted in fecal pellets, which can become aerosols and hence are an important cause of asthma. A large proportion of asthma patients have a strong IgE reaction to group 3 allergens (Ando *et al*, 1993).

A major limitation for molecular studies on scabies has been the difficulty of obtaining mites in sufficient numbers, because the parasite burden is generally very low and no *in vitro* culture system for propagating mites is available. cDNA libraries, however, have recently been prepared from *S. scabiei* var. *vulpes*

Manuscript received April 8, 2003; revised July 3, 2003; accepted for publication July 8, 2003

Address correspondence and reprint requests to: David J. Kemp, Queensland Institute of Medical Research, Post Office Royal Brisbane Hospital, Herston, Qld, 4029, Australia. Email: daveK@qimr.edu.au

Abbreviations: SMIPP, Scabies mite inactivated protease paralogues; PAR, protease activated receptor; SMART, simple modular architecture research tool; BLAST, basic local alignment search tool.

0022-202X/03/$15.00 ·

mRNA (Mattsson *et al*, 2001) and from var. *hominis* mRNA from mites in skin shed into the bedding of crusted scabies patients (Harumal *et al*, 2003). Homologues of the house dust mite allergens glutathione S-transferase, paramyosin, and M-177 were identified during initial characterization of clones for an expressed sequence tag (EST) library (Fischer *et al*, 2003a). We describe here the identification of multiple contigs encoding scabies homologs of group 3 allergen sequences in an expressed sequence tag library.

MATERIALS AND METHODS

S. scabiei var. *hominis* cDNA libraries Yv4, Yv5, and Yv6 in bacteriophage λ vector λZAP Express have been described (Fischer *et al*, 2003a). Libraries YvT, Yv7, Yv8, and Yv9 were derived from them by normalization utilizing reassociation kinetics (Fischer *et al*, 2003b). Excised phagemid clones were sequenced in 96 well plate format from the T3 primer using ABI PRISM BigDye terminator and an ABI PRISM 3700 capillary DNA analyzer (Applied Biosystems, Foster City, California) under contract by the Australian Genome Research Facility. The program *phred* (Ewing *et al*, 1998.) was used for automatic base calling and trimming of vector and low-quality sequences (*phred* score < 20). Contigs were assembled using the program *phrap* (Gordon *et al*, 1998) using the default values except for minscore (60) and minmatch (28) to increase the stringency of the contig assembly. Further sequencing of clones was undertaken as necessary from the T7 primer using ABI PRISM BigDye terminator and an ABI PRISM 377 DNA sequencer. These sequences were manually trimmed of vector and low-quality sequence and aligned to the assembled contigs.

Sequences were compared with the GenBank nonredundant databases using the basic local alignment search tool (BLAST) (Altschul *et al*, 1990). Signal sequences were predicted using the program SignalP-NN version 2.0.b2 (Nielsen *et al*, 1997a, b) (http://www.cbs.dtu.dk/services/SignalP). Sequences were analyzed for homology to known protein domain families using the simple modular architecture research tool (SMART) (Schultz *et al*, 1998, 2000), at the web site of the European Molecular Biology Laboratory (http://smart.embl-heidelberg.de).

The predicted mature protein sequences of the scabies mite contigs, the house dust mite group 3 allergens and the house dust mite group 6 allergen Der f 6 (accession number AAF28423) (Kawamoto *et al*, 1999) were aligned using ClustalW (Thompson *et al*, 1994). The alignment was bootstrapped using Seqboot (100 replicates), trees inferred using the parsimony algorithm and a consensus tree drawn using Consense (Felsenstein, 1989).

DNA from individual mites was extracted in 50 μL of PrepMan Ultra (Applied Biosystems). Polymerase chain reactions were conducted using 1 μL of the extracted DNA with 50pn (picomol) of each primer specific for contig Yv4005G12 (forward 5′-ATCTAAGCTTCATAATGACTAG-3′, reverse 5′-CTTTGTATCCACTCCCATAAAAG-3′) or Yv6023A04 (forward 5′-CGTTATCGTATTGTTTCGGTG-3′, reverse 5′-GTATCCATGAGACATAAGATC-3′). The reactions were initially heated to 94°C for 10 min followed by 35 cycles of 94°C for 40 s, 55°C for 40 s and 72°C for 2 min. Products were purified using a Qiaquick kit (QIAGEN, Clifton Hill, Victoria, Australia) and sequenced using the relevant forward and reverse specific primers, using ABI PRISM BigDye terminator and an ABI PRISM 377 DNA sequencer.

RESULTS

Sequences from 19,488 clones were examined for homology with sequences in the GenBank nonredundant database using BLAST. Fifty-four of the cDNA clones sequenced fell into 24 distinct contigs that each have house dust mite group 3 allergens as their top BLASTx match. In all cases, the next closest matches after house dust mite group 3 allergens were serine proteases of the S1 family. Clones from non-normalized libraries (9216 sequence runs) represented 0.34% of the total cDNA sequenced. Among the 24 contigs, sequences of the entire coding region of 15 contigs and of the mature protein of a further two contigs was completed. The sequences of these 17 contigs were submitted to GenBank and were assigned the accession numbers AY333071 to AY333087. The relationships of the 17 contigs and their house dust mite counterparts are shown in the ClustalW alignment of the predicted mature enzyme sequences in **Fig 1** and the consensus tree inferred from a bootstrapped alignment using the parsimony algorithm is shown in **Fig 2**.

The sequence exhibiting the closest match, Yv7016G03, appeared by all criteria available to be a homolog of the group 3 allergens. It was designated Sar s 3, while recognizing that it is not yet known if it is an allergen and hence appropriately described by this nomenclature. The house dust mite group 3 allergens are predicted to have a signal peptide that is cleaved after amino acid 18 in Der p 3 (accession number AAA19973) and Eur m 3 (accession number AAD10712) and at the homologous site (amino acid 16) in Der f 3 (accession number BAA09920), generating a proenzyme in each case. The proenzyme has been shown to be cleaved after amino acid 29 in Der p 3 (Stewart *et al*, 1992) to generate the mature enzyme. Examination of Sar s 3 using SignalP predicted that the first 21 amino acids of the protein function as a signal sequence. Analysis of the sequence using SMART revealed that the region after amino acid 29 was identified as being a member of the family of trypsin-like serine proteases (SMART family, SM0020; Pfam, PF00089). Amino acids 29 to 32 are identical to the first four amino acids of the mature enzyme Der p 3 so presumably this corresponds to the N-terminus of the mature Sar s 3 protease.

The predicted mature length of Sar s 3 is 231 amino acids, differing by only three insertions and four deletions from the house dust mite proteins (**Fig 1**). Der p 3, Der f 3, and Eur m 3 show 81% to 82% identity to each other, whereas the putative mature Sar s 3 protein shows 43% to 44% identity to these house dust mite proteins. The six cysteine residues involved in intramolecular disulfide bonding of serine proteases of family S1 (Rawlings and Barrett, 1994) are conserved in the mature house dust mite sequences and in the Sar s 3 protein. The histidine (position 46), aspartic acid (position 93), and serine (position 202) of the catalytic triad are all present in Sar s 3, and residue 196 is an aspartic acid, suggesting a trypsin-like specificity (**Fig 1**).

Examination of the other *S. scabiei* contigs using SignalP strongly predicted the presence of signal sequences on all 14 for which the full sequence was available. In all but one of these sequences, the predicted signal cleavage site is at the position corresponding to the site of cleavage of the proenzyme to form the mature protease in the house dust mite group 3 allergens (**Fig 1**). Hence it is likely that these putative proteins are secreted as the mature proteins. For the remaining sequence, Yv4005B08, the cleavage site predicted by SignalP is between residues 25 and 26. Analysis of Yv4005B08 protein sequence using SMART indicated that homology with the family of trypsin-like serine proteases begins at residue 24, which corresponds to the start of the mature house dust mite proteins. Thus this position was used as the putative start of the mature protein for this molecule (**Fig 1**). The putative N-terminus of the cleaved molecules is that which is most commonly found at the N-terminus of the chymotrypsin family, isoleucine, except in three of the proteins, Yv9017F05, Yv4005B08, and Yv5027C11, in which it is leucine (**Fig 1**).

These *S. scabiei* contigs are much more divergent than the house dust mite group 3 allergens (**Figs 1 and 2**). The predicted mature proteins range from 219 to 248 amino acid residues in length. The closest pair Yv7016C10 and YvT003F10 show 84% similarity to each other, whereas the most distant pair Yv4031D03 and Yv5001A04 show only 43% similarity. Apart from Sar s 3, Yv4005G12 shows the highest identity to Der p 3 at 36%.

In contrast to Sar s 3, all other scabies mite homologs examined lack at least two of the six conserved cysteine residues. In all of these putative proteins, the cysteine residues that flank the active serine in the house dust mite proteins and the putative Sar s 3, and that are generally disulfide bonded together in proteases of

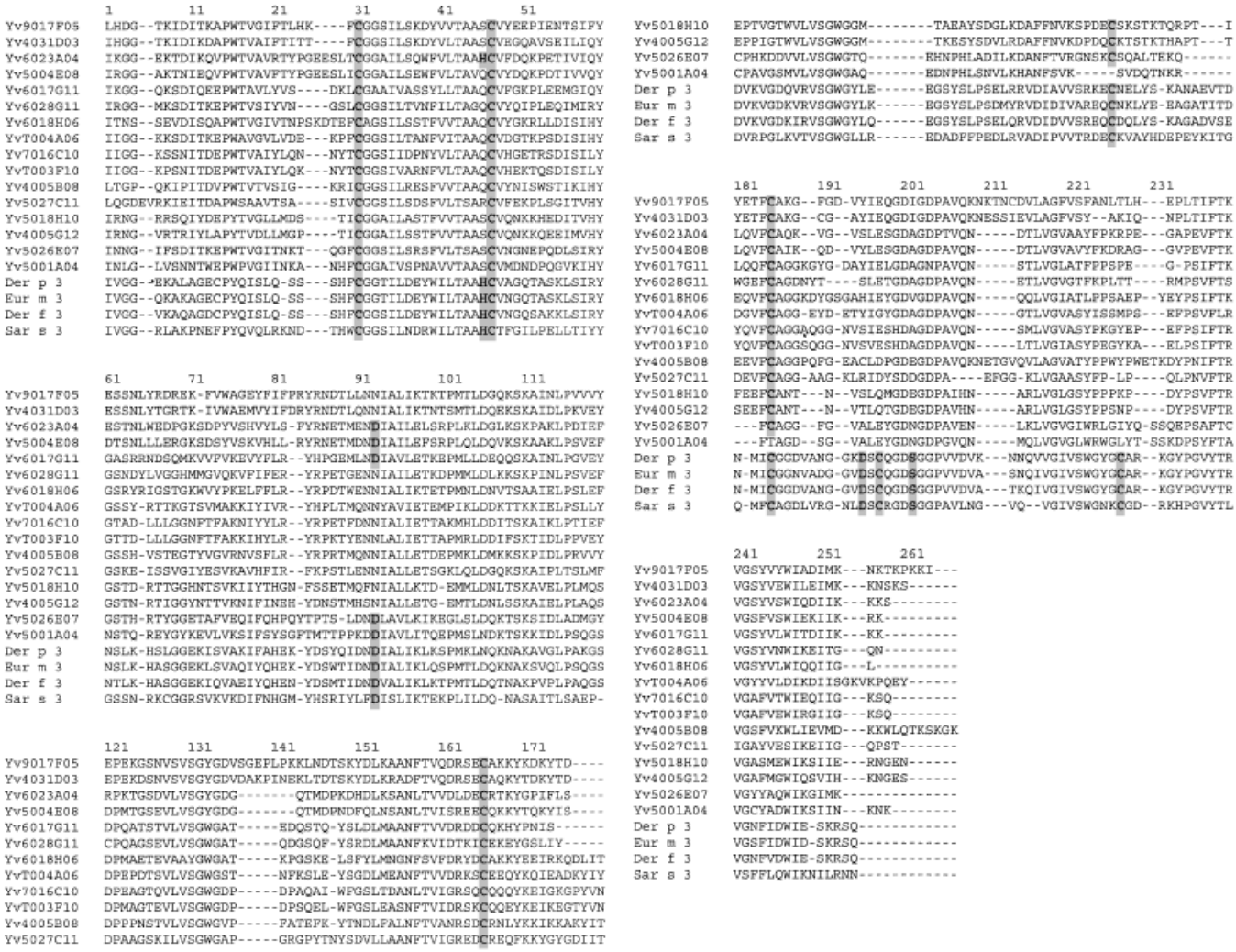

Figure 1. ClustalW alignment of the mature protein sequences of house dust mite group 3 allergens and their scabies mite homologs. The order of the sequences shown has been changed to correspond with that of the dendrogram in **Fig 2**. Conserved cysteine residues and the residues important for catalytic activity and specificity are shaded. The sequences described have been submitted to GenBank and have been assigned the accession numbers AY333071 to AY333087.

the S1 chymotrypsin family (Rawlings and Barrett, 1994) are absent. Yv5001A04 is more divergent as a further pair of cysteine residues are also absent (**Fig 1**). A remarkable feature of the contigs, however, is that with the exception of Sar s 3, not one of them encodes a putative protein with an intact catalytic triad (**Fig 2**). In all 16 other contigs the active serine is substituted, most commonly by alanine. Other substitutions for the active serine were isoleucine, glutamic acid, asparagine, aspartic acid, and valine. By itself this rules out the possibility that any of the members of the gene family with this substitution could act as proteases by a known mechanism. Equally remarkably, with the exception of Yv6023A04, which has only the active serine replaced, members of this gene family have two or, more commonly, all three amino acids of the catalytic triad mutated. Hence it is quite clear that they are no longer a family of active proteases.

Determination of the number of synonymous versus nonsynonymous mutations within the mature protein region revealed that in each possible comparison synonymous changes exceeded nonsynonymous changes by 1.8-fold. This is consistent with the proteins having functions that can be deleteriously affected by nonsynonymous mutations and hence are subject to "purifying selection" (Graur and Li, 2000). We conclude that they indeed have a function, presumably derived from their original functions as proteases.

Analysis of the phylogenetic relationships of these proteins suggests that inactivation of the capacity for proteolysis may have occurred in two different ancestral genes, which were both derived by successive duplications of the gene primordial to group 3 allergens. Phylogenetic trees generated by a variety of methods were similar and bootstrapping demonstrated high confidence levels for the major branches, reflecting the considerable degree of sequence diversity. The house dust mite proteins and Sar s 3 formed a clade (clade 4, **Fig 2**) considerably less divergent than the other sequences. In trees produced both by parsimony (**Fig 2**) and maximum likelihood methods (data not shown), Yv50018H10/Yv4005G12 and Yv5001A04/Yv5026E07 are related outliers (clades 2 and 3), with the remaining sequences forming a clade of more closely related sequences. As this gene family apparently contains at least 24 expressed genes, presumably expansion of the gene family has arisen by a succession of such recombination events. At least some of these genes appear to have undergone successive duplications and divergence after they ceased to encode functional proteases. We conclude that they must be under selection for some other function. We have termed them Scabies Mite Inactivated

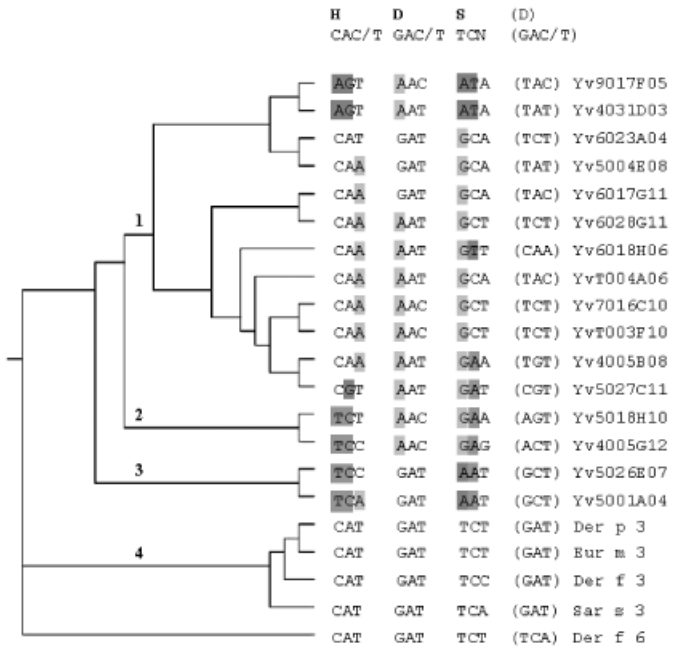

Figure 2. Phylogenetic relationships of the predicted mature protein sequences of house dust mite group 3 allergens and their scabies mite homologs. The consensus dendrogram of a bootstrapped alignment was generated by the parsimony method. The sequences separated into four distinct clades (1–4). The sequence of the group 6 allergen Der f 6 (Kawamoto *et al*, 1999) was included but not specified as an outgroup.

Protease Paralogue (*smipp*) genes. As these proteins are members of the serine protease family and other classes of proteases have been identified in *S. scabiei* (our unpublished observations), we will refer to the genes as *smipp-s* and their putative products as SMIPP-S.

Polymerase chain reaction products from chromosomal DNA of several individual mites were generated with primers specific for contigs Yv4005G12 (clade 2) and Yv6023A04 (clade 1). Alignment with the respective cDNA sequences revealed a 60 bp sequence in Yv4005G12 and a 54 bp sequence in Yv6023A04 with the features of an intron located at the same point in each sequence. This confirms that the cDNA are derived from expressed and spliced mRNA, and hence are not contaminating chromosomal fragments. The sequences of Yv4005G12 and Yv6023A04 determined from three individual mites, were all identical to the cDNA except that they contained the intron, which was identical in each case. Four additional mite genomic sequences of Yv6023A04 differed from the cDNA sequence by the same five single nucleotide polymorphisms, which included both synonymous and nonsynonymous changes; however, such differences meet the criteria for inclusion of a sequence within a contig. One of these generated a non-sense codon in the coding sequences at position 38, whereas none of the 54 cDNA clones contained non-sense codons. Hence it appears that if pseudogenes arising within the large family of *smipp-s* genes produce mRNA, it is rapidly degraded. As this pseudogene contained an intron it presumably arose by gene duplication rather than by retrotransposition (Mighell *et al*, 2000).

DISCUSSION

Identification of scabies homologs of house dust mite allergens as an initial approach to understanding the immunology of scabies at a molecular level, has recently been successful with the isolation of homologs of house dust mite allergen groups 8, 11, and 14 (Mattsson *et al*, 2001; Fischer *et al*, 2003a; Harumal *et al*, 2003). Clone Yv7016G03 described here appears to encode the scabies mite homolog of house dust mite group 3 allergens and hence it has been designated Sar s 3. Only one cDNA with this sequence was found among 19,488 sequences. Screening a library of house dust mite cDNA with a group 3 allergen DNA probe identified clones at a similar frequency (Smith and Thomas, 1996). The transcripts encoding other members of this gene family are collectively much more abundant. It has not yet been established whether Sar s 3 acts as an allergen and the large number of related sequences may complicate this question experimentally as they may cross-react serologically.

We have identified 54 scabies cDNA clones grouped into 24 contigs, all of which have house dust mite group 3 allergens as their closest known counterpart in the current nonredundant sequence database. The genome of *Drosophila melanogaster* encodes about 199 trypsin-like serine proteases and 178 chymotrypsin-like serine proteases (Rubin *et al*, 2000) and mite genomes would be expected to be similar in complexity. The *D. pteronyssinus* group 3 allergen Der p 3 has been considered to be the product of a single gene (Smith and Thomas, 1996). Of seven available Der p 3 sequences, four were identical in the coding region and two were closely similar (98.7% identity). The remaining sequence of 84.4% identity was clearly a separate gene. Further, only limited polymorphisms of Der p 1 (Chua *et al*, 1993) and Der p 2 (Chua *et al*, 1996) are found between house dust mites in the environment. Hence the size of the SMIPP-S gene family found here was not anticipated.

It is not yet established whether different *smipp-s* all represent distinct genes that are present in each mite or whether some represent sequence divergence within the mite population, as the library here was made from about 650 mites from a single patient (Fischer *et al*, 2003a). Because the group 3 allergens are gut molecules the Sar s 3 paralogues may also be secreted into the gut, although this remains to be established. Caution is necessary as there are numerous examples in which gene duplication and divergence has led to quite different functions of related gene products in different cellular locations (Graur and Li, 2000). *smipp-s* could be developmentally regulated and hence each may be expressed during only a limited part of the life cycle. These should be represented in the library as the mite population pooled for preparation of the mRNA contained large and small mites, including male and female mites of various instars, including gravid females.

Could these proteins still be active proteases? Whereas it could be hypothesized for any one of the *smipp-s* that it has mutated into a protease that functions by a previously unknown catalytic mechanism, the data argue strongly against this. We have identified a total of 10 different mutant forms of the catalytic triad. As most known proteases fall into a very limited number of classes of catalytic mechanism, and the cysteine proteases appear to have evolved separately at least twice, it is unrealistic to suggest that another 10 catalytic mechanisms have evolved in this family. Whether these different forms have relevance to the new functions is beyond speculation.

There was considerably more divergence between the *smipp-s* than between the house dust mite genes and Sar s 3, even though this is a comparison of intraspecies variation with interspecies variation. It could reflect less constraint on evolution of a family of dysfunctional proteases. The genes that now encode human chymotrypsins A and B, which show 79% amino acid similarity, are considered to have duplicated and commenced divergence about 270 million years ago (Graur and Li, 2000). Whereas we have no accurate measure of the rate of evolution of these genes in an organism with such a short life cycle, clearly separation of the most divergent members was an ancient event. If it was simply inactivity that promoted scabies survival, however, introduction of stop codons rather than specific mutations of the catalytic triad would be frequent. Whereas we have observed one example

of a stop codon in chromosomal sequences, clearly there are at least 24 genes without any. Further, if inactivity alone promoted survival, this would have selected for loss of expression rather than amplification of the *smipp-s* family up to at least 24 genes that are expressed.

Whereas being members of the S1 chymotrypsin protease family, as is trypsin, group 3 allergens have been considered to be trypsin-like enzymes. An important amino acid contributing to specificity for basic amino acids in trypsin is an aspartic acid at the position corresponding to 196 in **Fig 1**, whereas this is a serine in chymotrypsin. House dust mite group 3 allergens have an aspartic acid in this position (**Fig 1**) as does the Sar s 3 protein. This position shows considerable variation among the inactivated proteases; of the 21 other *S. scabiei* sequences available in this region, seven have a serine in this position, six tyrosine, two alanine, two cysteine, two arginine, one histidine, and one threonine. Notably, acidic amino acid residues are not represented among them.

Whereas Sar s 3 apparently has a proenzyme eight amino acids longer than the mature form, another unexpected observation was that the predicted signal cleavage point for each of the *smipp-s* for which the sequence is available, is at the site homologous to the predicted cleavage point of the proenzyme to form the mature house dust mite proteases. Genetic inactivation of the proteases would preclude self-cleavage, but should not affect cleavage by a different protease. Hence we presume that cleavage at this site by the signal peptidase is a secondary adaptation allowing production of molecules with structures equivalent to mature proteases in the absence of the ability to self-cleave. Whereas some critical cysteine residues mediating the disulfide bond structure of the chymotrypsin protease clan are wholly conserved, the disulfide bond from residues 198 to 227 across the active site has been lost. Taken together, all these features of *smipp-s* gene products must mean that they now function in another way. We hypothesize that this may reflect a critical step during adaptation to a parasitic lifestyle in the scabies lineage.

What could that function be? The most obvious feature of scabies infestation is that it elicits a delayed-type hypersensitivity reaction. As the gut proteases of house dust mites are potent allergens, it would be surprising if their homologs in scabies mites were not involved in the host immunologic response. Preliminary studies by immunoblotting have not revealed reactions of IgE from scabies patients with an expressed house dust mite group 3 paralogue (unpublished data) but this may well require molecules in native conformation, which will be difficult to ascertain in the absence of catalytic activity.

It is possible that in some way, inactivation of the catalytic triad is an adaptation enabling parasitic mites to survive against the onslaught of the host defense mechanisms. SMIPP-S may well retain the capacity to bind peptides, the initial step in catalysis, and this may mediate their present function. This hypothesis will be experimentally testable with recombinant SMIPP-S. If so, they may be antagonists of other proteases because they bind substrates; however, other properties may be more relevant. They may provide a web of immunologically cross-reactive sequences that interfere with affinity maturation of the host antibody response and thereby protect Sar s 3 against inactivation by antibodies imbibed into the gut of the mite; this may be critical for digestion. Alternatively, they may bind to serine protease inhibitors, thereby protecting Sar s 3 from inactivation.

Ixodid ticks, which imbibe blood, produce proteins that interfere with the blood clotting cascade, thus ensuring that the blood remains in a fluid state and hence is available for digestion (Lehane, 1994). Serine proteases of the mammalian blood coagulation cascade are under tight and co-ordinated control of serine protease inhibitors (serpins), including anti-thrombin, protein C inhibitor, heparin cofactor II, protease nexin 1, and plasminogen activator inhibitor-1. In the gut of the feeding scabies mite this complex proteolytic/serpin cascade presumably must be disrupted without proteolytic digestion disrupting. One possibility is that SMIPP-S may bind to mammalian serine protease inhibitors, thereby protecting Sar s 3 from inactivation. Whereas this has not been observed previously, the converse of parasite serpins interacting with mammalian proteases has (Zang *et al*, 1999). Absence of the active serine would prevent formation of the "suicide" covalent serpin protease complexes that are the end point of this interaction; however, it is notable that artificial substitution of an alanine for the active serine in trypsin has been shown to result in a very stable noncovalent Michaelis complex with bovine pancreatic trypsin inhibitor (Ye *et al*, 2001).

Another important function of proteases that may be relevant is interaction with protease-activated receptors (PAR), which modulate cell behavior (Macfarlane *et al*, 2001). Keratinocytes are known to possess PAR-2 receptors (Shpacovitch *et al*, 2002), which are trypsin rather than thrombin sensitive. Recently, it has been demonstrated that house dust mite proteases Der p 3 and Der p 9 (also a serine protease) cleave PAR-2 at the activation site and stimulate human pulmonary epithelial cells, inducing cytokine release (Sun *et al*, 2001). It was concluded that house dust mite proteases may induce a nonallergic inflammatory response through the release of pro-inflammatory cytokines mediated by PAR-2, in addition to their role as allergens (King *et al*, 1998; Sun *et al*, 2001). It has been postulated that such active interactions are the reason why house dust mite hydrolytic enzymes are such potent allergens. A counter argument is the recent discovery of an allergen from the German cockroach *Blattella germanica* with homology to aspartic proteases but with an inactive catalytic site (Pomes *et al*, 2002). The Sar s 3 protease of *S. scabiei* that we have identified could have such an effect in epidermis, which could be detrimental to mite survival. This may be the reason other members of the family have been genetically inactivated. It may be an important adaptation to parasitism, which is not relevant to house dust mites, which feed on dead skin fragments. We hypothesize therefore that the SMIPP-S bind to keratinocyte PAR-2 but cannot cleave it, and indeed may block it from activation by other proteases, including Sar s 3, therefore inhibiting the inflammatory response and protecting the mites. This should be testable using recombinant molecules for inhibition of proteolytic stimulation of keratinocytes and pulmonary epithelial cells via PAR-2 as described (King *et al*, 1998; Sun *et al*, 2001). It is possible that divergence of the SMIPP-S represents specialization for interaction with different PAR. These can all be activated by serine proteases but at differing concentrations of different individual proteases. We note that molecules that could inhibit PAR activation may have important implications for asthma therapeutics.

PAR-2 is expressed on keratinocytes but not melanocytes and has recently been ascribed importance in the regulation of pigmentation by phagocytosis of melanosomes (Seiberg *et al*, 2000; Paine *et al*, 2001). Topical application of the peptide SLIGRL (corresponding to the N-terminus of PAR-2 after cleavage), a peptide known to activate PAR-2 in the absence of cleavage, resulted in visible darkening of skin (Seiberg *et al*, 2000). Agents that inhibit serine proteases were shown to result in depigmentation of skin (Paine *et al*, 2001). Scabies can cause depigmentation, notably in severe crusted scabies. Hence there is circumstantial evidence that scabies produces molecules that result in inhibition of PAR-2.

Are these SMIPP-S vaccine candidates? If their function is indeed to mediate host defense evasion then interfering with that function may facilitate clearance of the parasite by normal host defense mechanisms. This may present an unanticipated avenue for protective vaccination.

We thank David Fairlie of the Institute for Molecular Biosciences, University of Queensland for suggesting that Protease Activated Receptors may be relevant. We would also like to thank John Davis, Vicky Hewitt and Matt Johnson from the Australian Genome Research Facility (AGRF) for their excellent technical assistance. The AGRF was established in 1997 by a Federal Government Major National

Research Facility grant. This work was supported by Australian NHMRC Medical Genomics grant 219175, by NHMRC project grant 137205 and by an NHMRC Senior Principal Research Fellowship to DJK. The sequencing program was carried out by contract at the AGRF.

REFERENCES

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

Ando T, Homma R, Ino Y, *et al*: Trypsin-like protease of mites: Purification and characterization of trypsin-like protease from mite faecal extract Dermatophagoides farinae. Relationship between trypsin-like protease and Der f III. *Clin Exp Allergy* 23:777–784, 1993

Arlian LG: Biology, host relations, and epidemiology of *Sarcoptes scabiei*. *Ann Rev Entomol* 34:139–161, 1989

Arlian LG, Vyszenski-Moher DL, Ahmed S, Estes SA: Cross-antigenicity between the scabies mite, *Sarcoptes scabiei*, and the house dust mite. *Dermatophagoides pteronyssinus* 96:349–354, 1991

Arlian LG, Rapp CM, Morgan MS: Resistance and immune response in scabies-infested hosts immunised with *Dermatophagoides* mites. *Am J Trop Med Hyg* 52:539–545, 1995

Chua KY, Stewart GA, Thomas WR, Simpson RJ, Dilworth RJ, Plozza TM, Turner KJ: Sequence analysis of cDNA coding for a major house dust mite allergen, *Der p I*. Homology with cysteine proteases. *J Exp Med* 167:175–182, 1988

Chua KY, Kehal PK, Thomas WR: Sequence polymorphisms of cDNA clones encoding the mite allergen *Der p I*. *Int Arch Allergy Immunol* 101:364–368, 1993

Chua KY, Huang CH, Shen HD, Thomas WR: Analysis of sequence polymorphism of a major mite allergen. *Der p 2*. *Clin Exp Allergy* 26:829–837, 1996

Currie B, Carapetis J: Skin infections and infestations in Aboriginal communities in northern Australia. *Australas J Dermatol* 41:139–143, 2000

Currie BJ, Connors CM, Krause VL: Scabies programs in Aboriginal communities. *Med J Aust* 161:636–637, 1994

Davis DP, Moon RD: Dynamics of swine mange: A critical review of the literature. *J Med Entomol* 27:727–737, 1990

Elgart ML: Scabies. *Dermatol Clin* 8:253–263, 1990

Ewing B, Hillier L, Wendl M, Green P: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–185, 1998

Falk ES, Bolle R: *In vitro* demonstration of specific immunological hypersensitivity to scabies mite. *Br J Dermatol* 103:367–373, 1980

Felsenstein J: PHYLIP—Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164–166, 1989

Fischer K, Holt D, Harumal P, Currie B, Walton S, Kemp D: Generation and characterisation of cDNA clones from *Sarcoptes scabiei* var. *hominis* for an expressed sequence tag library: Identification of homologues of house dust mite allergens. *Am J Trop Med Hyg* 68:61–64, 2003a

Fischer K, Holt DC, Wilson P, *et al*: Normalisation of a library of *Sarcoptes scabiei* cDNAs clones into λZAP by a long PCR and cDNA reassociation procedure. *Biotechniques* 34:250–254, 2003b

Gordon D, Abajian C, Green P: Consed: A graphical tool for sequence finishing. *Genome Res* 8:195–202, 1998

Graur D, Li W-H: *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates, Inc., 2000

Green MS: Epidemiology of scabies. *Epidemiol Rev* 11:126–150, 1989

Harumal P, Morgan M, Walton S, *et al*: Identification of a homologue of a house dust mite allergen in a cDNA library from *Sarcoptes scabiei* var. *hominis* and evaluation of its vaccine potential in a rabbit/*S. scabiei* var. *canis* model. *Am J Trop Med Hyg* 68:54–60, 2003

Kawamoto S, Mizuguchi Y, Morimoto K, *et al*: Cloning and expression of Der f 6, a serine protease allergen from the house dust mite, *Dermatophagoides farinae*. *Biochim Biophys Acta* 1454:201–207, 1999

Kemp D, Walton S, Harumal P, Currie B: The scourge of scabies. *Biologist* 49:19–24, 2002

King C, Brennan S, Thompson PJ, Stewart GA: Dust mite proteolytic allergens induce cytokine release from cultured airway epithelium. *J Immunol* 161: 3645–3651, 1998

Lehane M: Digestive enzymes, haemolysins and symbionts in the search for vaccines against blood-sucking insects. *Int J Parasitol* 24:27–32, 1994

Macfarlane SR, Seatter MJ, Kanke T, Hunter GD, Plevin R: Proteinase-activated receptors. *Pharmacol Rev* 53:245–282, 2001

Mattsson JG, Ljunggren EL, Bergstrom K: Paramyosin from the parasitic mite *Sarcoptes scabiei*. cDNA cloning and heterologous expression. *Parasitol* 122:555–562, 2001

Mellanby K: The development of symptoms, parasitic infection and immunity in human scabies. *Parasitology* 35:197–206, 1944

Mighell AJ, Smith NR, Robinson PA, Markham AF: Vertebrate pseudogenes. *FEBS Lett* 468:109–114, 2000

Nielsen H, Engelbrecht J, Brunak S, von Heijne G: Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10:1–6, 1997a

Nielsen H, Engelbrecht J, Brunak S, von Heijne G: A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* 8:581–599, 1997b

Paine C, Sharlow E, Liebel F, Eisinger M, Shapiro S, Seiberg M: An alternative approach to depigmentation by soybean extracts via inhibition of the PAR-2 pathway. *J Invest Dermatol* 116:587–595, 2001

Pomes A, Chapman MD, Vailes LD, Blundell TL, Dhanaraj V: Cockroach allergen Bla g 2: Structure, function, and implications for allergic sensitization. *Am J Respir Crit Care Med* 165:391–397, 2002

Rawlings ND, Barrett AJ: Families of serine peptidases. *Meth Enzymol* 244:19–61, 1994

Rubin GM, Yandell MD, Wortman JR, *et al*: Comparative genomics of the eukaryotes. *Science* 287:2204–2215, 2000

Schultz J, Milpetz F, Bork P, Ponting CP: SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci USA* 95:5857–5864, 1998

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P: SMART: A web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* 28: 231–234, 2000

Seiberg M, Paine C, Sharlow E, Andrade-Gordon P, Costanzo M, Eisinger M, Shapiro SS: The protease-activated receptor 2 regulates pigmentation via keratinocyte–melanocyte interactions. *Exp Cell Res* 254:25–32, 2000

Shpacovitch VM, Brzoska T, Buddenkotte J, *et al*: Agonists of proteinase-activated receptor 2 induce cytokine release and activation of nuclear transcription factor kappaB in human dermal microvascular endothelial cells. *J Invest Dermatol* 118:380–385, 2002

Smith WA, Thomas WR: Sequence polymorphisms of the Der p 3 house dust mite allergen. *Clin Exp Allergy* 26:571–579, 1996

Smith WA, Chua KY, Kuo MC, Rogers BL, Thomas WR: Cloning and sequencing of the *Dermatophagoides pteronyssinus* group III allergen, Der, p. III. *Clin Exp Allergy* 24:220–228, 1994

Stewart GA, Ward LD, Simpson RJ, Thompson PJ: The group III allergen from the house dust mite *Dermatophagoides pteronyssinus* is a trypsin-like enzyme. *Immunology* 75:29–35, 1992

Stewart GA, Kollinger MR, King CM, Thompson PJ: A comparative study of three serine proteases from *Dermatophagoides pteronyssinus* and *D. farinae*. *Allergy* 49:553–560, 1994

Sun G, Stacey M, Schmidt M, Mori L, Mattoli S: Interaction of mite allergens Der P3 and Der P9 with protease-activated receptor-2 expressed by lung epithelial cells. *J Immunol* 167:1014–1021, 2001

Taplin D, Porcelain SL, Meinking TL, Athey RL, Chen JA, Castillero PM, Sanchez R: Community control of scabies: A model based on use of permethrin cream. *Lancet* 337:1016–1018, 1991

Thomas WR, Smith W: Towards defining the full spectrum of important house dust mite allergens. *Clin Exp Allergy* 29:1583–1587, 1999

Thompson JD, Higgins DG, 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

Willadsen P, Kemp DH: Novel vaccination for control of the *Babesia* vector, *Boophilus microplus*. *Trans R Soc Trop Med Hyg* 83:107, 1989

Zang X, Yazdanbakhsh M, Jiang H, Kanost MR, Maizels RM: A novel serpin expressed by blood-borne microfilariae of the parasitic nematode *Brugia malayi* inhibits human neutrophil serine proteinases. *Blood* 94:1418–1428, 1999

Ye S, Cech AL, Belmares R, *et al*: The structure of a Michaelis serpin-protease complex. *Nat Struct Biol* 8:979–983, 2001