doi: 10.1111/j.1365-3083.2011.02525.x

IgE Sequences in Individuals Living in an Area of Endemic Parasitism Show Little Mutational Evidence of Antigen Selection

Y. Wang*, K. J. L. Jackson*†, Z. Chen†, B. A. Gaëta†, P. M. Siba‡, W. Pomat‡, E. Walpole*, J. Rimmer§, W. A. Sewell¶** & A. M. Collins*

*School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia; †School of Computer Science and Engineering, University of New South Wales, Sydney, NSW, Australia; ‡Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea; §St. Vincent's Clinic, Darlinghurst, NSW, Australia; ¶Institute of Laboratory Medicine, St Vincent's Hospital, Darlinghurst, NSW, Australia; and **St Vincent's Clinical School, University of New South Wales, Darlinghurst, NSW, Australia

Received 20 October 2010; Accepted in revised form 25 January 2011

Correspondence to: Dr A. M. Collins, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia. E-mail: a.collins@unsw.edu.au

Abstract

Patterns of somatic mutation in IgE genes from allergic individuals have been a focus of study for many years, but IgE sequences have never been reported from parasitized individuals. To study the role of antigen selection in the evolution of the anti-parasite response, we therefore generated 118 IgE sequences from donors living in Papua New Guinea (PNG), an area of endemic parasitism. For comparison, we also generated IgG1, IgG2, IgG3 and IgG4 sequences from these donors, as well as IgG1 sequences from Australian donors. IgE sequences had, on average, 23.0 mutations. PNG IgG sequences had average mutation levels that varied from 17.7 (IgG3) to 27.1 (IgG4). Mean mutation levels correlated significantly with the position of their genes in the constant region gene locus (IgG3 < IgG1 < IgG2 < IgG4). Interestingly, given the heavy, life-long antigen burden experienced by PNG villagers, average mutation levels in IgG sequences were little different to that seen in Australian IgG1 sequences (19.2). Patterns of mutation provide clear evidence of antigen selection in many IgG sequences. The percentage of IgG sequences that showed significant accumulations of replacement mutations in the complementarity determining regions ranged from 22% of IgG3 sequences to 39% of IgG2 sequences. By contrast, only 12% of IgE sequences had such evidence of antigen selection, and this was significantly less than in PNG IgG1, IgG2 and IgG4 subclass sequences ($P < 0.01$). The anti-parasite IgE response therefore has the reduced evidence of antigen selection that has previously been reported in studies of IgE sequences from allergic individuals.

The IgE response is often considered to be fundamentally deleterious, because IgE-mediated allergic disease is a significant burden on the community, particularly in developed countries [1]. IgE antibodies may, however, offer some protection against parasitic infections, and such antibodies remain a conspicuous part of the humoral response of most individuals in rural parts of the developing world [2, 3]. Studies of parasite infections of humans and animals have therefore informed our understanding of IgE antibodies in allergic disease, and together these conditions have provided insights into the biology of IgE more generally [4].

While many aspects of IgE-mediated effector function have now been well characterized [4, 5], the rarity of IgE-committed cells has made it difficult to elucidate the developmental pathway of IgE-producing B cells [6]. Details of this pathway have emerged, however, by the application of molecular techniques to the study of IgE antibody gene sequences [7].

Antibody gene sequences provide glimpses of B cell clonal history, through the somatic point mutations that they may carry. These mutations are believed to accumulate through the germinal centre reaction [8]. Within the germinal centres, rapidly expanding clones of antigen-specific B cells accumulate mutations within their rearranged antibody genes at the rate of about one mutation per cell division. Interactions between these cells and antigen on the surface of follicular dendritic cells leads to the selection of cells that have accumulated beneficial mutations, and to the death of cells that have accumulated deleterious mutations [9]. This process of antigen selection should be reflected in a tendency for

SCANDINAVIAN JOURNAL OF Immunology AN INTERNATIONAL MONTHLY JOURNAL

replacement mutations to accumulate in the complementarity determining regions (CDRs) of the immunoglobulin variable region genes [10]. By analysing the distribution of mutations within sequences, a number of early studies highlighted an apparent lack of antigen selection in the evolution of allergic IgE gene sequences [11, 12]. More recently, we have also reported that IgE sequences from an individual suffering from atopic dermatitis lacked evidence that mutations accumulated under the pressure of antigen selection [13]. Kerzel *et al.* [14] saw similar patterns in four patients with atopic dermatitis, but they reported evidence of antigen selection in IgE sequences from 13 asthmatic patients.

IgE antibodies are conspicuous in the response to helminths [15] and other parasites [16, 17], but we are unaware of any study that has examined IgE cDNA transcripts from parasitized individuals. In the light of the accumulating evidence of the unique mutational features of the IgE response in some conditions, we undertook immunogenetic studies of IgE antibodies in a community from the highlands of Papua New Guinea (PNG), where helminth infections are endemic, malaria is increasingly common, but allergic disease is almost unknown [18]. Here, we describe an analysis of sequences derived from 14 rural PNG villagers. To provide suitable data sets for comparison, we also amplified IgG sequences from both PNG and Australian individuals, and because of the possibility that different IgG subclasses could display varying patterns of mutation, we generated IgG sequences using subclass-specific PCR primers.

The average number of mutations in the IgE sequences of Papua New Guineans was very high and was broadly similar to the number of mutations seen in IgG sequences from the same individuals. Although the extent of IgE mutations was significantly higher than has been reported from studies of allergic individuals, the mean level of IgG mutations reported here is little different to previous reports of IgG sequences from the developed world. The distribution of replacement and silent mutations between framework regions (FRs) and CDRs suggest the involvement of antigen selection in the development of responses of each IgG subclass, but there was little evidence of selection in the IgE response.

Materials and methods

Sample processing. After informed consent, and with the approval of both the UNSW Human Research Ethics Committee and the Papua New Guinea Medical Advisory Council, peripheral blood was collected from 14 life-long residents of Masilakaiufa village, Eastern Highlands Province. The donors had no clinical symptoms or history of allergic disease and were aged between 22 and 53 years. Peripheral blood was also collected from 14 residents of Sydney, Australia. Serum was prepared from peripheral blood samples and frozen at −70 °C for later testing. Mononuclear cells were also prepared by density gradient centrifugation and frozen at −70 °C.

Serum Ig determination. Total IgE concentrations were determined for each PNG serum sample by enzyme immunoassay on a UniCAP® 100 system (Pharmacia, Uppsala, Sweden). On initial testing, one IgE sample was out of range (>5000 kU/l). It was re-measured after 1:5 dilution in an Australian serum sample known to have very low IgE (<5 kU/l). Total IgG and IgG subclasses were determined using a BN ProSpec® (Dade Behring, Sydney, Australia) nephelometer. Reference ranges were based on data from healthy adult residents of Sydney, Australia.

Sequence amplification. Mononuclear cells were prepared from peripheral blood samples by density gradient centrifugation, and total cellular RNA was extracted using TRIzol® Reagent (Invitrogen, Carlsbad, CA, USA). Synthesis of cDNA was performed using Superscript® III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. IgE and IgG heavy chain gene rearrangements were then amplified using an isotype-specific PCR. PCR amplification was performed with 100–200 ng cDNA or aliquots of the PCR1 product as templates, 0.2 μM of each primer, 200 μM of each dNTP, 1.25 units PFU polymerase (Promega, Madison, WI, USA) and a buffer supplied by the manufacturer. Details of the primers used are shown in Table 1. Specific primers for the three large IGHV gene families (VH1F, VH3F and VH4F) were

Table 1 Primers used for PCR amplification of immunoglobulin gene sequences.

Primer name	Sequences (5′–3′)	Description
VH1F	CACTCCCAGGTBCAGCTGGTGCAGTCTGG	IGHV1 family forward primer
VH3F	GTCCAGTGTSAGGTGCAGCTGGTGGAGTCTGG	IGHV3 family forward primer
VH4F	GTCCTGTCCCAGSTGCAGCTRCAGSAGTSGGG	IGHV4 family forward primer
G1	CCCAGAGGTGCTCTTGGAGGA	Reverse primer for IgG1
G2/G4IN	GGCTGTGCTCTCGGAGGTGCT	Reverse primer for IgG2 and IgG4 PCR2
G3OUT	TGTGTCACCAAGTGGGGTTTTGAGC	Reverse primer for IgG3 PCR1
G3IN	CAGAGGTGCTCCTGGAGCA	Reverse primer for IgG3 PCR2
G4OUT	CAYGGGGGACCATATTTGGACT	Reverse primer for IgG4 PCR1
IGEOUT	GCTGAAGGTTTTGTTGTCGACCCAGTC	Reverse primer for IgE PCR1
IGEIN	GGGGAAGACGGATGGGCTCTGTGTGG	Reverse primer for IgE PCR2

used as forward primers in separate reactions. IgG1 and IgG2 were amplified by standard PCR using appropriate isotype specific primers (G1 and G2/G4IN) as reverse primers. Reactions times for this PCR were 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 61 °C for 30 s, 72 °C for 4 min and then a final extension at 72 °C for 5 min. Semi-nested PCR were used for IgG3 (reverse primers: G3OUT and G3IN), IgG4 (G4OUT and G2/G4IN) and IgE (IGEOUT and IGEIN) sequence amplifications. PCR1 conditions used were initial denaturation at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s, 72 °C for 4 min and a final extension at 72 °C for 5 min. For PCR2, the only changed condition from those of PCR1 was the annealing temperature, which was 65 °C for IgE and IgG4, and 61.7 °C for IgG3. PCR2 was run for 25 cycles. All PCR were run on a Tpersonal 48 cycler (Biometra, Gottingen, Germany). PCR products were then cloned and sequenced at the Ramaciotti Centre for Gene Function Analysis, University of New South Wales, as previously described [13].

Bioinformatic analysis. Rearranged VDJ sequences were aligned against the germline repertoire using the iHMMune-align program [19] and the UNSWIg repertoire of germline genes [20] (http://www.ihmmune.unsw.edu.au/unswig.html). This repertoire was updated with a number of IGHV polymorphisms that we have identified in the PNG population and have submitted to GenBank (accession numbers HM855272–HM855948), as well as putative polymorphisms that have been identified in previous studies [20, 21]. Evidence in support of the existence of these putative polymorphisms within rearranged VDJ genes can be found at http://cgi.cse.unsw.edu.au/~ihmmune/IgPdb/. The number of mismatches between the germline IGHV genes and each rearranged sequence was noted. Sequences with more than 45 mismatches were removed from the data set because of the likelihood they included sequencing errors. Clonally related sequences were identified on the basis of shared IGHV, IGHD and IGHJ genes, as well as shared N regions and shared point mutations. Where clonally related sequences were identified, they were removed from the data set after a representative sequence was generated by averaging the mutation numbers seen in the different FR and CDR regions of each sequence in the set. Where there were sequences associated with two or more isotypes in a set, averages sequences were generated for each isotype.

To investigate the role of antigen selection in the evolution of patterns of mutation within the IgE sequences, the proportion of replacement mutations within the CDR1 and CDR2 of each sequence was calculated. Broad definitions of CDR1 and CDR2 were used, incorporating the CDR regions of both Kabat [22] and IMGT [23], and analysis was made with reference to a random model of mutations as previously described [13]. In this model, the probability that a random mutation would introduce a replacement mutation in the CDR was estimated to be 0.26, based upon patterns of mutation and hotspots in a data set of non-productive sequences [13]. Analysis showed that this estimate was appropriate for all IGHV sequences, for there is little variation in the mutability of different IGHV genes (data not shown). Using the binomial distribution, the estimate was then used to establish 95% confidence limits for the proportion of the total mutations that would be replacement mutations in the CDR (R_{CDR}), if the mutation process targeted hotspots, but if these mutations were not subject to antigen selection pressure. Proportions were calculated for varying numbers of total IGHV mutations (Mv). The upper limit (97.5%) was used to distinguish sequences that showed evidence of antigen selection from sequences that lacked such evidence.

Results

Serum antibody concentrations

Total serum immunoglobulin concentrations were determined for all PNG samples, and the results are summarized in Table 2. Concentrations of serum IgE antibodies were all above the laboratory reference range for healthy Sydney adults, and the mean IgE concentration of the serum samples was 2465 kU/l. IgG subclass concentrations are also shown in Table 2. IgG1 and IgG4 concentrations were particularly high. Nine of the 14 PNG individuals had IgG1 concentrations above the laboratory reference range for healthy Sydney adults, while all but one of the individuals studied had serum IgG4 concentrations that were above the Laboratory Reference Range. In Western populations, IgG4 is typically the least abundant IgG subclass, but IgG4 in these PNG samples was seen at substantially higher concentrations than IgG3.

Sequence analysis

Sequences were aligned against the germline IGHV, IGHD and IGHJ gene repertoires using the iHMMune-

Table 2 Serum immunoglobulin concentrations of 14 Papua New Guineans. Mean values and the range of values are indicated, along with laboratory reference ranges for healthy individuals, based on data from adult residents of Sydney, Australia.

	Mean	Range	Reference range
IgE (kU/l)	2465	343–12,485	0–180
IgG1 (g/l)	12.2	9.9–14.8	2.4–10.8 (male) 3.4–11.2 (female)
IgG2 (g/l)	3.7	1.6–5.4	1.2–5.5 (male) 1.5–5.2 (female)
IgG3 (g/l)	0.7	0.15–1.7	0.3–1.3 (male) 0.2–1.1 (female)
IgG4 (g/l)	2.4	0.8–5.7	0.08–0.9 (male) 0.07–0.9 (female)

align program, while IGHG gene identity was confirmed by BLAST. PCR error rates were determined by analysis of errors within the IGHG constant region genes and were shown to vary from 0.9‰ (IgG2) to 1.2‰ (IgG4). The amplified constant region of the IgE sequences was too short for such a calculation.

Alignments against the germline IGHV genes were examined, and after the removal of sequences with more than 45 mismatches to the germline, there were 1108 unique PNG sequences, including 125 IgE sequences, 482 IgG1 sequences, 288 IgG2 sequences, 59 IgG3 sequences and 154 IgG4 sequences. These sequences were submitted to GenBank and were assigned the accession numbers HM773966–HM775073. One hundred and sixty-two IgG1 sequences were also amplified from Australian samples.

A number of VDJ sequences were found that aligned to a recently identified germline IGHV3 gene (HM855939). The IGHV3-NL1*01 gene was seen in seven VDJ rearrangements (accession numbers HM773984, HM774108, HM774124, HM774201, HM774302, HM774729, and HM774738). One of these is an IgG3 sequence (HM774124) that contains no somatic point mutations.

Alignments were also seen to 12 other recently identified IGHV allelic variants, including IGHV1-8*02 (HM855457), IGHV1-18*03 (HM855463), IGHV3-7*03 (HM855666), IGHV3-9*02 (HM855577), IGHV3-11*06 (HM855329), IGHV3-21*03 (HM855323), IGHV3-21*04 (HM855688), IGHV3-33*06 (HM855436), IGHV3-48*04 (HM855336), IGHV3-53*04 (HM855453), IGHV4-59*11 (HM855471) and IGHV7-4-1*04 (HM 855485). In total, alignments were seen to 91 different IGHV genes and allelic variants. Despite the use of primers specific for the VH1, VH3 and VH4 gene families, many sequences were also amplified that utilized the IGHV5 family genes. In fact, the IGHV5 family genes as well as IGHV1-69 alleles were over-represented in all data sets, when compared with previously reported rearrangement frequencies [21].

Analysis of the VDJ junctions showed the mean CDR3 lengths of PNG IgG sequences to vary between 14.9 (IgG2) and 16.6 amino acids (IgG3), while the IgE sequences had a mean length of 15.4. These differences were not statistically significant. Within the junctions, all previously reported functional IGHD genes were observed. Alignments were also seen to one or more allele of each IGHJ gene, including both IGHJ3*01 and IGHJ3*02. IGHJ3*01 was originally reported as part of a haplotype that includes IGHJ4*01 and IGHJ5*01. In an earlier bioinformatic study of VDJ rearrangements, we failed to find convincing evidence for the existence of these three alleles [24]. The alignments seen in this study confirm the existence of IGHJ3*01, although no convincing alignments were observed to IGHJ4*01 or IGHJ5*01.

In the PNG data sets, 64 sets of clonally related sequences were seen, involving a total of 175 sequences. Forty-four sets contained two sequences, 12 sets contained 3 sequences, 3 sets contained 4 sequences, 2 sets contained 5 sequences and 3 additional sets contained 6, 7 and 16 sequences, respectively. Seven sets contained clonally related sequences from different isotypes, including three sets of mixed IgG1/IgG2 sequences, three set of IgG1/IgG4 sequences and one set of IgG1/IgE sequences.

Clonally related sequences were particularly common amongst the IgG4 sequences. Of the 154 IgG4 sequences, 55 (35.7%) sequences were related to other IgG4 or IgG1 sequences. In contrast, only 69 of the 482 IgG1 sequences, 23 of the 288 IgG2, 16 of the 59 IgG3 and 12 of the 125 IgE sequences were members of clonally related sets. The number of clonally related sequences was significant higher amongst the IgG4 sequences (χ^2 test: $P < 0.001$). No clonally related sequences were identified in the Australian samples.

For subsequent mutation analyses, clonally related sequences were removed from the data sets. After their removal, 1004 unique PNG sequences remained, including 118 IgE sequences, 445 IgG1 sequences, 276 IgG2 sequences, 49 IgG3 sequences and 116 IgG4 sequences. The average mutation count for the IgE-associated IGHV genes was 23.0. The average number of mutations seen in PNG sequences associated with the different IgG subclasses correlated with the position of the various constant region gamma genes in the constant region locus. IgG3, which is encoded by the most 5′ IGHG gene, had the lowest number of mutations (mean: 17.7). The IGHG1 gene is located downstream of the IGHG3 gene, and IgG1 sequences had an average 21.0 mutations. The IGHG2 gene is found downstream of IGHG1, and IgG2 sequences had an average 22.0 mutations. IgG4, which is encoded by the most 3′ IGHG gene, had the highest number of mutations (mean: 27.1). Differences between PNG isotypes were significant (one-way ANOVA: $P < 0.001$) with IgG4 being significantly higher than all other isotypes including IgE (Dunn multiple comparison: $P < 0.05$). Perhaps surprisingly, there was no significant difference seen between the level of mutations in the Australian IgG1 sequences (mean: 19.2) and in the PNG IgG1 sequences. Mean numbers of mutations for PNG IgG subclasses and IgE are shown as Fig. 1, and the frequency distributions of IGHV mutation numbers are shown as Fig. 2A–F. Chi-squared analysis of the frequency distribution of IGHV mutations showed a significant difference between isotypes ($P < 0.01$). Striking differences were seen in the proportion of sequences that were relatively unmutated (<10 mutations). Eight per cent of IgE sequences had fewer than 10 mutations, but very few IgG4 sequences were relatively unmutated, with only two of 116 IgG4 sequences having fewer than 10 mutations. In contrast, 31% of IgG3 sequences carried fewer than 10 mutations, with two sequences having no

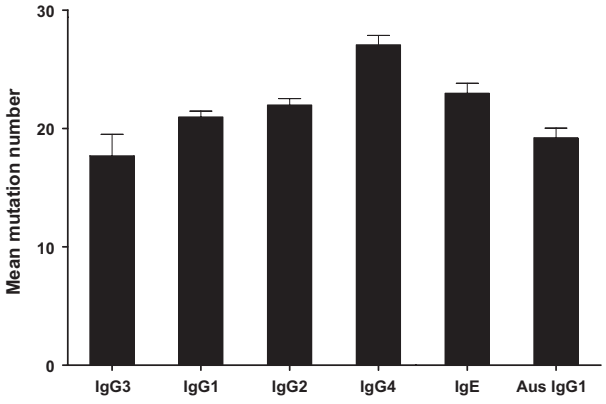

Figure 1 Mutations in IGHV genes of different isotypes of Papua New Guinea immunoglobulin sequences and of Australian IgG1 sequences. Data are presented as mean ± SEM.

mutations at all. These differences between IgG4 and the other isotypes, including differences between IgG4 and IgE, were all significant (χ^2 tests; in all cases $P < 0.05$).

Inference of antigen selection

The percentages of PNG sequences in each sequence data set that showed evidence for selection are shown in Fig. 3, and plots of replacement mutations in the CDR (R_{CDR}) against total IGHV mutations (Mv) are shown for IgE and the IgG subclasses as Fig. 4. The IgE sequences showed evidence of antigen selection in only 12% of sequences, which was significantly less than in the IgG sequences (χ^2 test: $P < 0.001$). Amongst the IgG sequences, the percentage of sequences showing evidence of antigen selection were 28% (IgG1), 39% (IgG2), 22% (IgG3) and 27% (IgG4). All subclasses showed signifi-

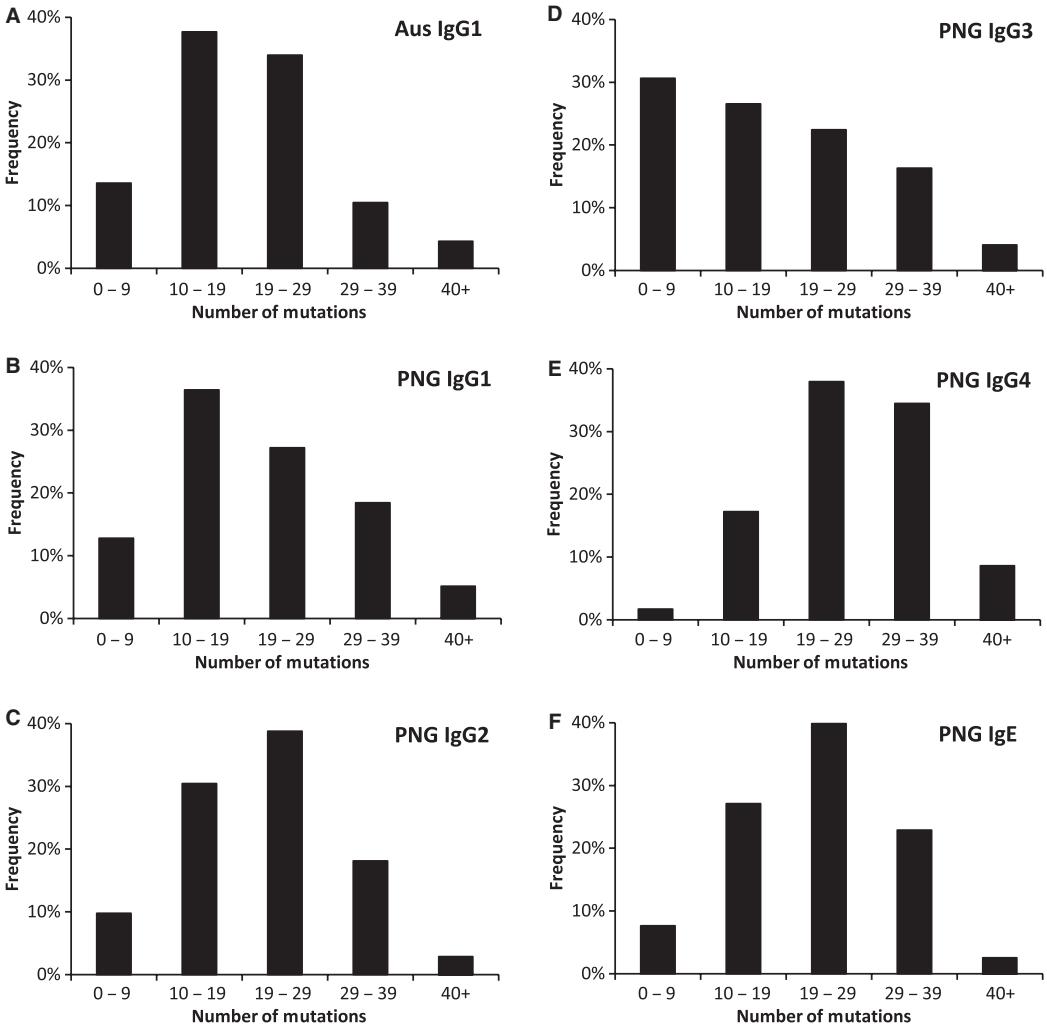

Figure 2 Frequency distribution of point mutations in IGHV genes. Data are presented for (A) Australian IgG1 sequences and for Papua New Guinean (B) IgG1, (C) IgG2, (D) IgG3, (E) IgG4 and (F) IgE.

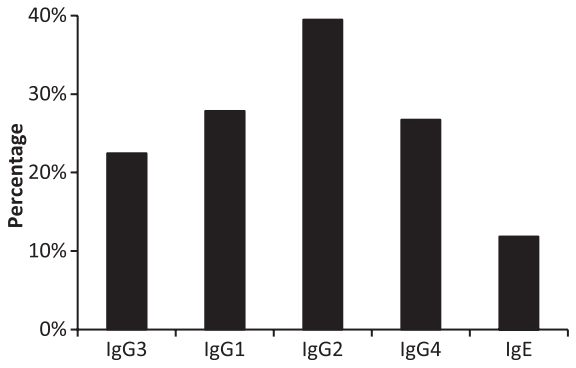

Figure 3 Percentage of unique immunoglobulin IGHV gene sequences, sourced from Papua New Guinean individuals, which showed evidence of antigen selection. Sequences that included more replacement mutations within the complementarity determining regions (CDR1 and CDR2) than the 97.5% confidence limit of a random model of mutation were deemed to show evidence of antigen selection.

cantly elevated levels of selection in comparison with IgE ($P < 0.01$ in each case), except for the small set of IgG3 sequences.

Discussion

The antigen-induced clustering of cell surface IgE is a key activation pathway for mast cells, basophils and eosinophils, and these cells are all conspicuous players in response to parasite infections. A detailed understanding of the fine specificity of IgE antibodies is therefore essential if we are to properly understand the biology of these critical effector cells. Much of our understanding of IgE antibodies is drawn from more general studies of humoral immunity, for it has been widely assumed that the IgE response develops in parallel with the IgG response. That is, it has been thought that the IgE response develops within germinal centres where, guided by antigen selection, and in the presence of T follicular helper cells, clonal proliferation and mutation lead to the emergence of high-affinity antibodies and the development of both plasma cells and memory cells. Recent work has challenged this view. It has been proposed, for example, that IgE-switched cells may be early emigrants from the germinal centre reaction [6]. It has also been proposed that the IgE response could be driven by superantigen-like stimulation [14]. Indirect evidence that may help us clarify these fundamental aspects of the biology of IgE comes from studies of IgE sequences and the point mutations that accumulate in these genes.

To investigate the IgE response in circumstances other than allergic disease, we conducted the present study of individuals from a community in which parasite infections are endemic [25]. The prevalence of allergic disease was investigated in this population in the 1980s, and it was shown to be almost entirely absent [18]. Although epidemiological studies have not recently been conducted in the area, none of the subjects in this study reported any symptoms indicative of allergic disease. All the individuals, however, had very high serum IgE concentrations. Although the specificities of the IgE antibodies remain unknown, it is reasonable to suppose that most of the IgE was generated as a consequence of parasite infection. The very high serum IgG4 concentrations seen are also typical of the response to persistent parasite infections [26].

Patterns of gene usage have been a focus of many studies of IgE sequences. An over-representation of genes of the IGHV5 family in IgE VDJ rearrangements has been reported by some [11, 12] but not all studies of IgE sequences [13, 14], and this has been taken as evidence of superantigen-driven responses [14]. In this study, biased usage of IGHV1-69 genes and genes of the IGHV5 family were seen in sequence sets of all isotypes and in both Australian and PNG IgG sequences. This suggests that the bias seen is likely to be a consequence of the variable efficiency of the amplification of different IGHV genes by the family-specific degenerate PCR primers used in this study. Previously reported biases could also be artefactual.

We have previously identified large numbers of clonally related IgE sequences that utilized the IGHV6-1 gene, in a study of an individual with atopic dermatitis [27]. This is likely to be the result of the thorough sampling of a highly restricted portion of sequence space that is revealed by PCR amplification using a forward primer specific for IGHV6-1. In contrast, the relatively low proportion of clonally related sequences seen in this study suggests that the IgE response, in parasitized individuals, may be highly diverse. The varying proportions of clonally related sequences seen in association with different IgG subclasses may also point to varying levels of diversity in these responses, although analysis is confounded by the unequal numbers of different IgG subclass transcripts obtained from different individuals. Certainly, the high proportion of clonally related IgG4 sequences suggests a lack of diversity which might be expected if this subclass response was restricted to the minor set of the most persistent antigens.

Further insights into the IgE anti-parasite response come from analysis of somatic point mutations, and to better interpret our observations of mutations in IgE sequences, we amplified IgG-associated VDJ gene sequences, using IgG subclass-specific reverse PCR primers. The mean mutation levels seen in these 886 unique IgG sequences varied substantially between subclasses and correlated with the position of the constant region genes within the constant region gene locus (IgG3 < IgG1 < IgG2 < IgG4). Although unexpected, this is in accord with the reports of low-affinity IgG3 being seen early in a response [25], and high-affinity IgG4 emerging after long periods of persistent

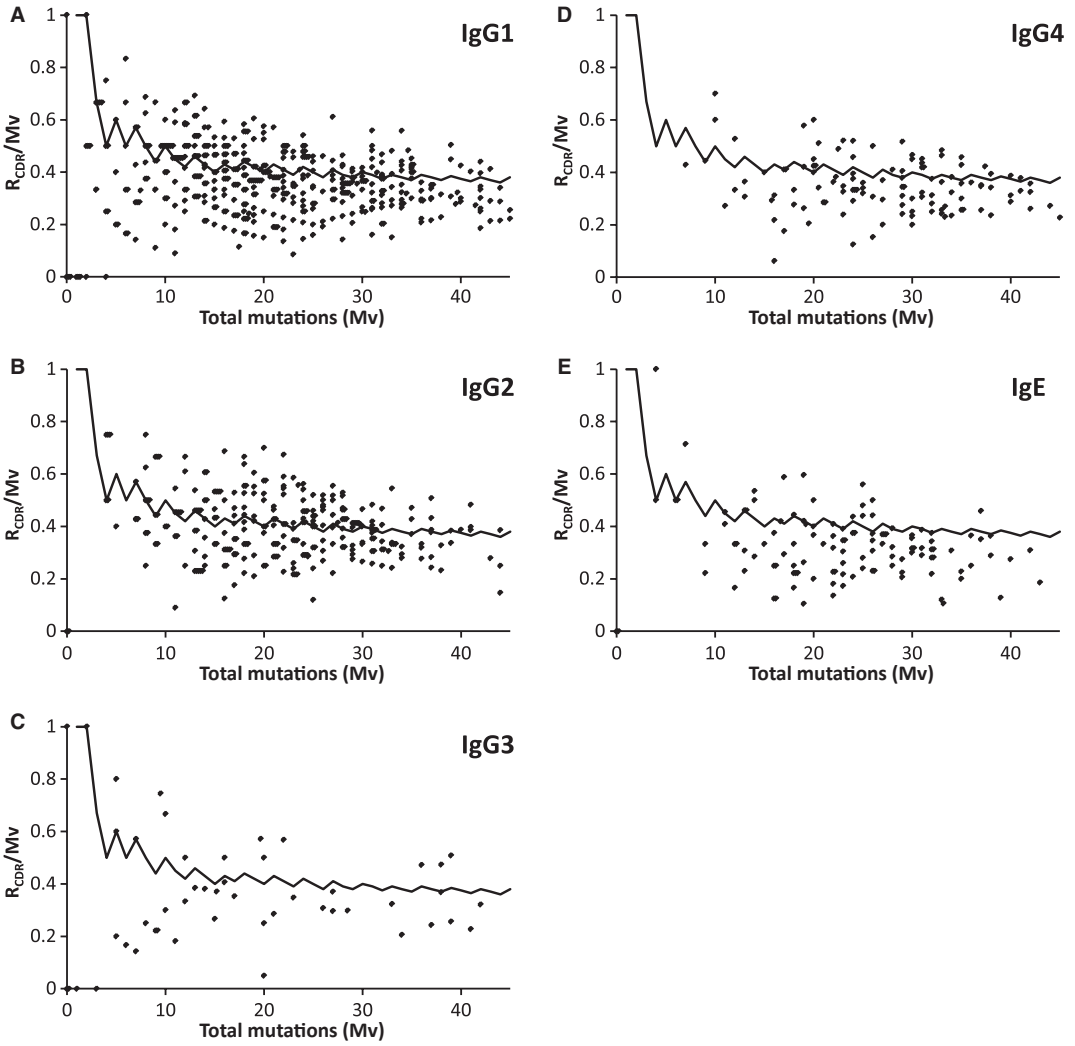

Figure 4 The proportion of total IGHV mutations (Mv) that are replacement mutations within the CDR1 and CDR2 regions (R_{CDR}). The proportions of Mv that are R_{CDR} mutations are plotted against the total numbers of mutations in (A) 445 IgG1 sequences, (B) 276 IgG2 sequences, (C) 49 IgG3 sequences, (D) 116 IgG4 sequences and (E) 118 IgE sequences. The solid lines show the 97.5% confidence limits for the R_{CDR}/MV ratio in a model of random mutation, where $p(R_{CDR})$ equals 0.26. Data points have been adjusted to highlight clusters of overlaid values. CDR, complementarity determining region.

antigen stimulation [28]. These studies are consistent with the concept that B cells only switch to IgG4 after multiple rounds of cell division, during which the VDJ sequences accumulate high numbers of mutations [29]. On the other hand, it does not imply that IgG class-switching progresses inevitably by a series of sequential downstream steps, for cells may switch to IgG4 both directly from IgM and indirectly via other constant region genes, as is also known to occur in the IgE response [30, 31].

Interestingly, despite IgE also being associated with persistent stimulation, and the IGHE gene being downstream of the IGHG genes, the level of mutation in IgE sequences was similar to that of IgG1 and IgG2 sequences and was significantly less than that of IgG4 sequences. The average number of mutations seen in the IgE-associated VDJ gene sequences was 23.0, which is substantially higher than we previously reported for IgE sequences from individuals with atopic dermatitis, whose mean mutation counts were 14.7 and 15.7 [13, 27]. Higher counts have been seen in individuals with seasonal rhinitis and allergy to grass pollen [32], with a reported median count of 21.

While an average of 19.2 mutations were seen in IgG1 sequences derived from PBMC of Australian individuals, the PNG IgG1 sequences had an average of 21.0 mutations. This broad similarity in the extent of mutations between IgG sequences from PNG villagers and sequences from urban residents of developed nations is surprising. It might be expected that mutation numbers would reflect an individual's history of antigen exposure.

Individuals from developed nations could therefore be expected to have substantially fewer mutations than individuals who have lived in the less hygienic circumstances of the developing world. Our observation may be explained by recent studies of the memory response. It has been shown that in a recall response, IgG^+ memory cells rapidly give rise to plasma cells, but IgM^+ memory cells re-enter the germinal centre reaction [33]. As $CD27^+$ IgM^+ memory cells carry few mutations in their immunoglobulin genes [34, 35], the extent of mutations generated in the germinal centre reaction of a recall response is likely to be little different from that seen as a result of the earlier exposure to the antigen. Repeated exposure to common microbial antigens, which is a likely feature of village life in developing countries, would therefore be likely to lead to a relatively slow rise in mean mutation levels with age.

As expected, many IgG sequences displayed a significantly higher proportion of replacement mutations within the CDRs than is seen in a model of random mutation. This can be taken as evidence that antigen selection guided the evolution of these sequences. The percentage of such sequences ranged between 22% (IgG3) and 39% (IgG2). The majority of sequences do not show evidence of antigen selection. This is not because most IgG sequences develop in the absence of antigen selection, but rather it likely reflects the underlying random nature of the mutational process, which makes it impossible to see clear evidence of antigen selection in more than a fraction of all selected sequences. In contrast to the IgG sequences, only 12% of IgE sequences showed evidence of antigen selection. This is in line with previous observations of allergic IgE sequences. We and others have reported an absence of antigen selection, and therefore presumably the absence of affinity maturation in allergic IgE sequences [13, 36, 37].

Kerzel *et al.* [14] recently used the same kind of comparison with a random model of mutation in a study of antigen selection and mutations in allergic IgE sequences. In their study, a different probability of mutation was used, as different definitions of the CDRs were also used. The use of these different definitions and probabilities do not alter the conclusions of the present analysis.

The relative lack of antigen selection in the evolution of IgE sequences in parasitized individuals could be the result of early departure of IgE-committed cells from the germinal centre reaction and the continuing accumulating of mutations at other sites where follicular dendritic cells and follicular helper T cells, that are essential to the antigen selection process, are absent [6, 38]. Although the IgE response could continue to develop at extra-germinal centre sites within the lymph nodes, it is also possible that clonal expansion and mutation continue outside the lymph nodes. The mutation process is initiated by the enzyme activation-induced cytidine deaminase (AID) [39] and there is, for example, evidence of AID expression by B cells at mucosal surfaces [40]. Intestinal helminths could perhaps drive proliferation and mutation of IgE-committed B cells within these tissues, giving rise to antibodies of low affinity or even of uncertain specificity.

A number of studies have reported that the specific IgE response is accompanied by an approximately 10-fold higher production of non-specific or 'bystander' IgE [41, 42]. The source and molecular features of this non-specific response have not been reported, but it has been suggested that 'bystander' IgE arises through non-cognate interactions between activated T cells and antigen non-specific B cells [41]. This hypothesis is not consistent with the patterns of mutation seen in the IgE sequences in this study. On the other hand, an essentially random mutation process and the accumulation of IgE-switched B cells of lower affinity or even of altered specificity offers an alternative explanation for the phenomenon of 'bystander' IgE that is consistent with the mutational evidence. If we are to understand the efficacy of these antibodies, and if we are to harness the IgE response, through vaccination, against those parasite infections that remain a burden on the health of much of the world's people, the biological processes that lead to these patterns of mutation must be explored.

Acknowledgment

This work was supported by a grant from the National Health and Medical Research Council. We thank Kate McGill of the Immunopathology Department, St Vincent's Pathology, St Vincent's Hospital Sydney, for performing the serology tests. We thank Sauli Bebes and the people of Masilakaiufa village for their assistance and participation in this study.

References

1 Gupta R, Sheikh A, Strachan DP, Anderson HR. Time trends in allergic disorders in the UK. *Thorax* 2007;62:91–6.
2 Erb KJ. Helminths, allergic disorders and IgE-mediated immune responses: where do we stand? *Eur J Immunol* 2007;37:1170–3.
3 Cooper PJ, Ayre G, Martin C, Rizzo JA, Ponte EV, Cruz AA. Geohelminth infections: a review of the role of IgE and assessment of potential risks of anti-IgE treatment. *Allergy* 2008;63:409–17.
4 Gould HJ, Sutton BJ, Beavil AJ *et al.* The biology of IgE and the basis of allergic disease. *Annu Rev Immunol* 2003;21:579–628.
5 Prussin C, Metcalfe DD. IgE, mast cells, basophils, and eosinophils. *J Allergy Clin Immunol* 2003;111:S486–94.
6 Erazo A, Kutchukhidze N, Leung M *et al.* Unique maturation program of the IgE response *in vivo*. *Immunity* 2007;26:191–203.
7 Davies JM, O'Hehir RE. Immunogenetic characteristics of immunoglobulin E in allergic disease. *Clin Exp Allergy* 2008;38:566–78.
8 Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intraclonal generation of antibody mutants in germinal centres. *Nature* 1991;354:389–92.
9 Kimoto H, Nagaoka H, Adachi Y *et al.* Accumulation of somatic hypermutation and antigen-driven selection in rapidly cycling

surface Ig+ germinal center (GC) B cells which occupy GC at a high frequency during the primary anti-hapten response in mice. *Eur J Immunol* 1997;27:268–79.
10 Chang B, Casali P. The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. *Immunol Today* 1994;15:367–73.
11 Snow RE, Chapman CJ, Frew AJ, Holgate ST, Stevenson FK. Pattern of usage and somatic hypermutation in the V(H)5 gene segments of a patient with asthma: implications for IgE. *Eur J Immunol* 1997;27:162–70.
12 van der Stoep N, van der Linden J, Logtenberg T. Molecular evolution of the human immunoglobulin E response: high incidence of shared mutations and clonal relatedness among epsilon VH5 transcripts from three unrelated patients with atopic dermatitis. *J Exp Med* 1993;177:99–107.
13 Dahlke I, Nott DJ, Ruhno J, Sewell WA, Collins AM. Antigen selection in the IgE response of allergic and non-allergic individuals. *J Allergy Clin Immunol* 2006;117:1477–83.
14 Kerzel S, Rogosch T, Struecker B, Maier RF, Zemlin M. IgE transcripts in the circulation of allergic children reflect a classical antigen-driven B cell response and not a superantigen-like activation. *J Immunol* 2010;185:2253–60.
15 Flohr C, Quinnell RJ, Britton J. Do helminth parasites protect against atopy and allergic disease? *Clin Exp Allergy* 2009;39:20–32.
16 Duarte J, Deshpande P, Guiyedi V *et al.* Total and functional parasite specific IgE responses in *Plasmodium falciparum*-infected patients exhibiting different clinical status. *Malar J* 2007;6:1.
17 Furuta T, Imajo-Ohmi S, Fukuda H, Kano S, Miyake K, Watanabe N. Mast cell-mediated immune responses through IgE antibody and Toll-like receptor 4 by malarial peroxiredoxin. *Eur J Immunol* 2008;38:1341–50.
18 Turner KJ, Dowse GK, Stewart GA, Alpers MP. Studies on bronchial hyperreactivity, allergic responsiveness, and asthma in rural and urban children of the highlands of Papua New Guinea. *J Allergy Clin Immunol* 1986;77:558–66.
19 Gaëta B, Malming HR, Jackson KJL, Bain ME, Wilson P, Collins AM. iHMMune-align: hidden Markov model-based alignment and identification of germline segments in immunoglobulin gene sequences. *Bioinformatics* 2007;23:1580–7.
20 Boyd SD, Gaëta BA, Jackson KJ *et al.* Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. *J Immunol* 2010;184:6986–92.
21 Wang Y, Jackson KJL, Sewell WA, Collins AM. Many human immunoglobulin heavy-chain IGHV gene polymorphisms have been reported in error. *Immunol Cell Biol* 2008;86:111–5.
22 Kabat EA, Wu TT, Perry HM, Gottesman SK, Foeller C. *Sequences of Proteins of Immunological Interest*. Bethesda, MD: U.S. Dept of Health and Human Services, 1991, ISBN-10: 094137565X.
23 Lefranc MP, Pommie C, Ruiz M *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* 2003;27:55–77.
24 Lee CEH, Jackson KJL, Sewell WA, Collins AM. Use of IGHJ and IGHD gene mutations in analysis of immunoglobulin sequences for the prognosis of chronic lymphocytic leukemia. *Leuk Res* 2007;31:1247–52.
25 Wilson KM, Di Camillo C, Doughty L, Dax EM. Humoral immune response to primary rubella virus infection. *Clin Vaccine Immunol* 2006;13:380–6.
26 Hoerauf A, Satoguina J, Saeftel M, Specht S. Immunomodulation by filarial nematodes. *Parasite Immunol* 2005;27:417–29.
27 Edwards MR, Brouwer W, Choi CHY, Ruhno J, Ward RL, Collins AM. Analysis of IgE antibodies from a patient with atopic dermatitis: biased V gene usage and evidence for polyreactive IgE heavy chain complementarity-determining region 3. *J Immunol* 2002;168:6305–13.
28 Devey ME, Bleasdale-Barr KM, Bird P, Amlot PL. Antibodies of different human IgG subclasses show distinct patterns of affinity maturation after immunization with keyhole limpet haemocyanin. *Immunology* 1990;70:168–74.
29 Hasbold J, Lyons AB, Kehry MR, Hodgkin PD. Cell division number regulates IgG1 and IgE switching of B cells following stimulation by CD40 ligand and IL-4. *Eur J Immunol* 1998;28:1040–51.
30 Jung S, Siebenkotten G, Radbruch A. Frequency of immunoglobulin E class switching is autonomously determined and independent of prior switching to other classes. *J Exp Med* 1994;179:2023–6.
31 Niederberger V, Niggemann B, Kraft D, Spitzauer S, Valenta R. Evolution of IgM, IgE and IgG(1–4) antibody responses in early childhood monitored with recombinant allergen components: implications for class switch mechanisms. *Eur J Immunol* 2002;32:576–84.
32 Davies JM, O'Hehir RE. VH gene usage in immunoglobulin E responses of seasonal rhinitis patients allergic to grass pollen is oligoclonal and antigen driven. *Clin Exp Allergy* 2004;34:429–36.
33 Dogan I, Bertocci B, Vilmont V *et al.* Multiple layers of B cell memory with different effector functions. *Nat Immunol* 2009;10:1292–9 Epub 25 October 2009.
34 Klein U, Rajewsky K, Kuppers R. Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J Exp Med* 1998;188:1679–89.
35 Ma CS, Pittaluga S, Avery DT *et al.* Selective generation of functional somatically mutated IgM+CD27+, but not Ig isotype-switched, memory B cells in X-linked lymphoproliferative disease. *J Clin Invest* 2006;116:322–33.
36 Snow RE, Chapman CJ, Frew AJ, Holgate ST, Stevenson FK. Analysis of Ig VH region genes encoding IgE antibodies in splenic B lymphocytes of a patient with asthma. *J Immunol* 1995;154:5576–81.
37 Janezic A, Chapman CJ, Snow RE, Hourihane JO, Warner JO, Stevenson FK. Immunogenetic analysis of the heavy chain variable regions of IgE from patients allergic to peanuts. *J Allergy Clin Immunol* 1998;101:391–6.
38 King C, Tangye SG, Mackay CR. T follicular helper (TFH) cells in normal and dysregulated immune responses. *Annu Rev Immunol* 2008;26:741–66.
39 Martin A, Scharff MD. AID and mismatch repair in antibody diversification. *Nat Rev Immunol* 2002;2:605–14.
40 Coker HA, Durham SR, Gould HJ. Local somatic hypermutation and class switch recombination in the nasal mucosa of allergic rhinitis patients. *J Immunol* 2003;171:5602–10.
41 McMenamin C, Girn B, Holt PG. The distribution of IgE plasma cells in lymphoid and non-lymphoid tissues of high-IgE responder rats: differential localization of antigen-specific and 'bystander' components of the IgE response to inhaled antigen. *Immunology* 1992;77:592–6.
42 Mitchell AJ, Moss ND, Collins AM. The biological activity of serum IgE changes over the course of a primary response. *Scand J Immunol* 2002;55:33–43.