# Inconclusive Evidence for or against Positive Antigen Selection in the Shaping of Human Immunoglobulin E Repertoires: A Call for New Approaches.

Levin, Mattias; Ohlin, Mats

*Published in:*
International Archives of Allergy and Immunology

*DOI:*
10.1159/000345421

2013

SHORT COMMUNICATION

# Inconclusive evidence for or against positive antigen selection in the shaping of human IgE repertoires – a call for new approaches

Mattias Levin and Mats Ohlin *

Dept. of Immunotechnology, Lund University, Lund, Sweden

**Short title:** Antigen selection shaping IgE

**Keywords:** allergen; IgE; mutation; repertoire; selection

* **Corresponding author:** Dept. of Immunotechnology, Lund University, BMC D13, S-22184 Lund, Sweden; telephone: +46-46-2224322; telefax: +46-46-2224200; e-mail: mats.ohlin@immun.lht.se



Published as: Int Arch Allergy Immunol 2013;161:122–126

(http://dx.doi.org/10.1159/000345421)

Supplementary information is available at Publisher's web site.

## Abstract

**Background:** The mechanisms driving the development of IgE antibody repertoires are a matter of debate. Alternatives to the classical view on antibody development, involving somatic mutation and antigen driven selection of high-affinity variants in germinal centers, have been proposed.

**Methods:** We have re-analyzed the pattern of mutations in previously isolated and characterized human clonally unrelated IgE-encoding transcripts using the validated focused binomial methodology to find evidence in such genes of antigen-specific selection.

**Results:** As expected there is a selection against replacement mutations in IgE framework regions. In contrast, in all examined cases but one (assessing IgE repertoires of parasite-infected individuals) there was no evidence in favor of either positive or negative selection in complementarity determining regions. Importantly, however, the validated method also failed to detect selection for replacement mutations in two, non-IgE, hypermutated antibody population targeting tetanus toxoid and vaccinia virus, respectively.

**Conclusions:** Current methodology is unable to define with certainty, using commonly assessed IgE repertoire sizes, if antigen selection is or is not a major driving force in the establishment of human IgE. New approaches are needed to address this matter.

## Introduction

Antibodies of the IgE isotype are well-described, key components of the defense against parasites but also of type 1-hypersensitivity reactions. The mechanisms behind the development of diversity in IgE antibody repertoires, however, are a matter of substantial debate. The classical view of development of hypermutated antibodies is that selection of B cells producing such antibodies occurs in germinal centers. To ensure development towards a high affinity antibody this process is believed to favor replacement mutations targeting the complementarity determining regions (CDR) rather than silent mutations. In contrast replacement mutations are counter-selected in the framework regions (FR), as they are likely to be accompanied by unacceptable structural effects on the folded protein. The end product of this antigen selection process should be a pool of B cells carrying genes encoding antigen-specific, high-affinity antibodies. In the case of IgE-repertoire development alternative mechanisms, not necessarily involving selection by antigen [1], have however been proposed. For instance, extensive polyclonal expansion of B-cells producing IgE has been suggested to be a major contributor to the establishment of such antibodies [1-2]. Microbial toxins might act as superantigens, exhibiting excessive polyclonal mitogenic activity on B-cell populations. Such repertoires may be dominated by sequences that carry a lower than expected number of substitutions in the CDR as there will be no selective advantage in creating an optimal binding site towards a conventional antigen. Indeed some [2-5] but not all [6] studies have indicated evidence of mutations in IgE not compatible with selection by antigens.

What is then the basis for the differences between different studies of the involvement of selection in the shaping of IgE repertoires? Firstly the studies have assessed IgE in different clinical conditions and it is not unreasonable to expect that IgE develop in different ways in different diseases. Secondly, the analytical approach is a source of error. Identification of antigen-directed selection in antibody evolution is problematic for a number of reasons including but not limited to high background of neutral mutations, different codon usage in different parts of antibody sequences and differences in the mutagenability of different codons [7-11]. Furthermore, unequivocal assignment of germline gene origin is difficult as there is generally no knowledge of the germline gene repertoire that generated the IgE repertoires. Consequently miss-assignment is likely occurring in several cases adding irrelevant noise to the data set.

To minimize the influence of methodological differences we have now re-assessed the pattern of mutations in several IgE-encoding gene populations using a single, extensively

validated methodology [10-12] to define whether or not it is possible to pin-point antigen selection as a major driving force shaping the development of IgE repertoires as they are described in today's literature.

## Materials and Methods

### Study material

Antibody heavy chain variable domain-encoding sequences (Table 1, Supplementary Table 1) were derived from IgE-producing cells found in peripheral blood of patients diagnosed with allergic rhinitis [13-14] and in sinus mucosa of patients diagnosed with chronic rhinosinusitis or non-allergic fungal eosinophilic sinusitis [15], the latter being a condition devoid of systemic allergy but with signs of local reactions similar to allergic inflammation. We also included sequences encoding IgE derived from children with allergic asthma [6] and parasitized subjects [5] in this study. In addition, two repertoires of antibodies, likely mostly IgG, specific for tetanus toxoid and vaccinia virus of recently boosted individuals were included for comparison [16-17].

### Analysis

To avoid repeated analysis of frequently occurring rearrangements, only one randomly picked sequence of the different clonotypes (a clonotype is defined as described by the authors or as a set of clones with an origin in a given rearrangement (with an origin in a common IGHV gene and a nucleotide identity within the heavy chain CDR3 of >80%) that has occurred in a single individual) was used in the study. To be included, a sequence from each clonotype must also be accessible from GenBank. The sequence must furthermore be complete across the investigated sequence, must not show evidence (as assessed by the V-QUEST algorithm [18]) of mutations involving insertions or deletions, and must not contain unspecified bases. The IGHV gene/allele that had been used in the process that generated a rearranged heavy-chain encoding sequence was defined using the IMGT/V-QUEST online tool [18]. As PCR had been used to amplify sequences bases comprising the first codons were not included in the analysis in order to avoid an influence of primer design on the assignment of germline gene origin and mutation. Consequently, codons 1 to 8 were always excluded and in the case of sequences described by Kerzel *et al.* [6] codons 1-26 were excluded. All these gene populations represent diverse sets of sequences in terms of V gene subgroup and gene

usage (Table 1). All sets also carry high frequencies of mutations at the nucleotide level (Table 1).

The analysis of selection was performed using the extensively validated focused binomial test (http://clip.med.yale.edu/sel/) as described by Hershberg *et al.* [11] and as modified by Uduman *et al.* [12]. We used the default setting of the online tool, i.e. CDR were defined as codons 27-38 and 56-65 in accordance with the unique IMGT CDR definition rules and the IMGT numbering system [19]. These residues create a surface directed towards an antigen sitting in a binding site and avoids analysis of residues in the C'' β-strand of the folded heavy chain variable domain as included in the Kabat definition of the second heavy chain CDR. Codons beyond residue 104 were not included in the analysis to eliminate confounding effects caused by sequence alterations associated with VDJ rearrangement process, i.e. modifications that are unrelated to the somatic hypermutation and the selection process. The Benjamini–Hochberg false discover method was employed, using an on-line tool (http://sdmproject.com/utilities/?show=FDR), to correct for the fact that multiple statistical tests were performed.

## Results and Discussion

In agreement with prior studies of mutated immunoglobulin sequences [5,7], the vast majority of sequences did not show statistical significant evidence of positive or negative selection on their own (data not shown). All populations showed strong negative selection against mutations in immunoglobulin FR (Table 1), in agreement with expectations (assuming that many mutations in FR will detrimentally influence the stability of the V domains) and past experience [10]. Furthermore there was little evidence of positive selection in the CDR of IgE-encoding sequence populations. Only IgE derived from parasitized individuals [5] showed statistically significant selection in CDR, in this case negative selection. This is in agreement with past studies of this population of sequences that showed less evidence of antigen selection as compared to IgG [5]. The study by Kerzel et al. [6], a study that involved sequencing of more than 1000 IgE-encoding genes from 13 lymphocyte donors, has, however, previously reported evidence of antigen selection in IgE. That repertoire, however, did not show evidence of positive selection as assessed using the focused binomial test. Importantly, though, the tetanus toxoid-specific and the vaccinia virus-specific antibody population also did not show evidence of positive selection in CDR using the focused binomial test system.

This finding suggests that an extensively validated analytical approach like the one employed in this study cannot with ease detect positive selection in human antibody populations highly suspected to have undergone positive selection. This points to the limitation of the approach itself. It has previously even been suggested that none of the frequently used tests, approaches like the binomial test [8], the multinomial test [20] or the focused binomial test [11], are able to detect selection by antigen in the CDR of immunoglobulin-encoding transcripts [10]. This highlights the complexity of designing such algorithms with high enough specificity without reducing sensitivity to levels were selection is very difficult to detect. The lack of selection was not accompanied by a low level of mutation in the sequences used for analysis as the frequency of mutation was 5.6-8.4 % (Table 1). Thus, sample sizes like those used in this study, a size common to most studies of IgE-encoding gene sequences, are likely insufficient to establish evidence of positive selection in CDR using those tests available today.

The failure to detect evidence for positive selection in sets of antibody-encoding genes of sizes common to many studies of IgE repertoires suggest that new approaches need to be taken to address this issue. Firstly one has to realize that it by no means is certain that antigen selection translates into a mathematically detectable enhancement of substitutions in CDR. Indeed many affinity-enhancing mutations do not reside in CDR [21], a fact that confounds this analytical approach. It has also been argued that false-positive results are frequent in particular when not applying an analytical strategy based on a focused binomial methodology and, even more damaging, that one should not even expect to find mathematically detectable signs of selection in the CDR of antigen-binding site of antibodies [10].

A problem if we ever are going to be able to detect subtle increases in substitution in CDR as a consequence of antigen selection is that even rather extensive conventional sequencing efforts (exemplified by Kerzel *et al.* [6]) generates a relatively small number of independent clones due to the oligoclonal nature of IgE repertoires [13]. It is expected that larger, carefully controlled studies to clarify the existence, or not, of positive and negative selective pressure during the development of the different human IgE repertoires will be required to resolve this matter. It is anticipated that high throughput sequencing of genes derived from large numbers of donors may aid in this endeavor. Firstly such methodology provides larger data sets that may be required to resolve this matter with statistical significance. Secondly the methodology can be used to deduce the germline repertoire of each individual [22] thereby facilitating correct germline gene assignments eliminating background noise in the data set. It is anticipated that such efforts may eventually permit us to address this matter.

Ultimately the resolution of the matter of whether or not selection for higher affinity occurs in human IgE may have to await a very laborious analysis of mutated, allergen-specific IgE clones in comparison to their unmutated germline counterparts. Current technology has, as far as we are aware, not been able to identify gene sequences encoding native combinations of heavy and light chain variable domains of IgE with known allergen specificity, for instance through sorting of cells of the B cell lineage by flow cytometry in combination with cloning of the corresponding variable domain-encoding genes. Such technology in combination with determination of the affinity for the immunizing agent has in the past been used to decipher the extent of affinity maturation in the much more commonly occurring cells of the B cell lineage that encode isotypes other than IgE [16]. However, even when access to such human IgE-producing clones can be ensured in the future, many IgE responses pose a very specific problem in the context of analysis of affinity maturation. Many allergens have cross-reacting counterparts in other species or are even represented by a multitude of more or less similar isoallergens and isoforms within a given species [23]. In most cases it will likely be impossible to know the allergen form(s) that originally induced the response and that was/were the driver in the affinity maturation process that eventually resulted in the population of IgE-producing B cells observed in a given allergic individual. Consequently it will likely be very difficult to define with certainty the extent of affinity maturation that has occurred in vivo unless one is able to assess a response involving few cross-reacting allergens/allergen isoforms or a response focused on a conserved epitope. Until these conceptual matters have been resolved and given the inability of even a highly validated statistical methodology to detect positive selection in IgG-encoding antibody populations, we suggest that claims of the presence or lack of selection in IgE responses, and its relation to the mechanism of IgE-mediated disease, have to be treated with caution.

## Acknowledgements

This study was supported by grants from the Swedish Research Council (grant numbers: 521-2008-3614 and 521-2011-3282) and Alfred Österlunds stiftelse.

## References


1. Gould HJ, Takhar P, Harries HE, Chevretton E, Sutton BJ The allergic march from Staphylococcus aureus superantigens to immunoglobulin E. Chem Immunol Allergy 2007;93:106-136.
2. Coker HA, Harries HE, Banfield GK, Carr VA, Durham SR, Chevretton E, Hobby P, Sutton BJ, Gould HJ: 2005. Biased use of VH5 IgE-positive B cells in the nasal mucosa in allergic rhinitis. J Allergy Clin Immunol 2005;116:445-452.
3. Dahlke I, Nott DJ, Ruhno J, Sewell WA, Collins AM: Antigen selection in the IgE response of allergic and nonallergic individuals. J. Allergy Clin. Immunol. 2005;117:1477-1483.
4. Pratt E, Collins AM, Sewell WA, Harvey RJ: Antigen selection in IgE antibodies from individuals with chronic rhinosinusitis with nasal polyps. Am J Rhinol Allergy 2010;24,:416-421.
5. Wang Y, Jackson KJL, Chen Z, Gaeta BA, Siba PM, Pomat W, Walpole E, Rimmer J, Sewell WA, Collins AM: IgE sequences in individuals living in an area of endemic parasitism show little mutational evidence of antigen selection. Scand J Immunol 2011;73:496-504.
6. Kerzel S, Rogosch T, Struecker B, Maier RF, Zemlin M: 2010. 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;185:2253-2260.
7. Bose B, Sinha S: Problems in using statistical analysis of replacement and silent mutations in antibody genes for determining antigen-driven affinity selection. Immunology 2005;116:172-183.
8. 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 Today1994; 15:367-373.
9. Dunn-Walters DK, Spencer J: Strong intrinsic biases towards mutation and conservation of bases in human IgVH genes during somatic hypermutation prevent statistical analysis of antigen selection. Immunology 1998;95:339–345.
10. MacDonald CM, Boursier L, D'Cruz DP, Dunn-Walters DK, Spencer J: Mathematical analysis of antigen selection in somatically mutated immunoglobulin genes associated with autoimmunity. Lupus 2010;19:1161-1170.

11. Hershberg U, Uduman M, Shlomchok MJ, Kleinstein SH: Improved methods for detecting selection by muatation analysis of Ig V region sequences. Int. Immunol. 2008;20:683-694.
12. Uduman M, Yaari G, Hershberg U, Stern JA, Shlomchik MJ, Kleinstein SH: Detecting selection in immunoglobulin sequences. Nucleic Acids Res 2011;39:W499-W504.
13. Andréasson U, Flicker S, Lindstedt M, Valenta R, Greiff L, Korsgren M, Borrebaeck CAK, Ohlin M: The human IgE-encoding transcriptome to assess antibody repertoires and repertoire evolution. J Mol Biol 2006;362:212-227.
14. Persson H, Karbalaei Sadegh M, Greiff L, Ohlin M: Delineating the specificity of an IgE-encoding transcriptome. J Allergy Clin Immunol 2007;120:1186-1192.
15. Levin M, Tan LW, Baker L, Wormald PJ, Greiff L, Ohlin M: Diversity of IgE-encoding transcripts in sinus mucosa of subjects diagnosed with non-allergic fungal eosinophilic sinusitis. Clin. Exp. Allergy 2011;41:811-820.
16. Poulsen TR, Meijer PJ, Jensen A, Nielsen LS, Andersen PS: Kinetic, affinity and diversity limits of human polyclonal antibody responses against tetanus toxoid. J Immunol 2007;179:3841-3850.
17. Lantto J, Haahr Hansen M, Rasmussen SK, Steinaa L, Poulsen TR, Duggan J, Dennis M, Naylor I, Easterbrook L, Bregenholt S, Haurum J, Jensen A: Capturing the natural diversity of the human antibody response against vaccinia virus. J Virol 2011;85:1820-1833.
18. Brochet X, LeFranc MP, Guidicelli V: IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucl Acids Res 2008;36:W503-W508.
19. Lefranc MP, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, Lefranc G: IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev Comp Immunol 2003;27:55-77.
20. Lossos IS, Tibshirani R, Narasimhan N, Levy R: The interferance of antigen selection on Ig genes. J Immunol 2000;165:5122-5126.
21. Daugherty PS, Chen G, Iverson BL, Georgiou G: Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. Proc Natl Acad Sci USA 2000;97:2029-2034.
22. Boyd SD, Gaëta BA, Jackson KJ, Fire AZ, Marshall EL, Merker JD, Maniar JM, Zhang LN, Sahaf B, Jones CD, Simen BB, Hanczaruk B, Nguyen KD, Nadeau KC, Egholm M, Miklos DB, Zehnder JL, Collins AM: Individual variation in the germline Ig gene

repertoire inferred from variable region gene rearrangements. J Immunol 2010;184:6986-6992.

23. Chapman MD: Allergen nomenclature. Clin Allergy Immunol 2008;21:47-58.

**Table 1. Characteristics of investigated antibody repertoires.**

| Donor condition | Lymphocyte source | Isotype | Number of donors | Number of analyzed IGHV sequences | Number of IGHV germline[a] | | Frequency of mutations in IGHV gene (%) | Focused binomial test P-value[b] | | Corrected focused binomial test P-value[b,c] | | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | gene subgroups | genes | | CDR | FR | CDR | FR | |
| Chronic rhinosinusitis | Sinus mucosa | IgE | 4 | 24 | 4 | 14 | 6.1 | 0.16 | $-2.4 \times 10^{-7}$ | 0.19 | $-4.8 \times 10^{-7}$ | 15 |
| Non-allergic fungal eosonophilic sinusitis | Sinus mucosa | IgE | 4 | 34 | 4 | 19 | 6.6 | 0.37 | $-1.2 \times 10^{-14}$ | 0.40 | $-2.8 \times 10^{-14}$ | |
| Seasonal allergic rhinitis | Peripheral blood | IgE | 2 | 70 | 5 | 20 | 5.6 | 0.056 | $-6.4 \times 10^{-20}$ | 0.087 | $-1.8 \times 10^{-19}$ | 13 |
| Allergic asthma (children) | Peripheral blood | IgE | 13 | 102 | 5 | 28 | 6.8 | -0.11 | $-7.4 \times 10^{-25}$ | -0.14 | $-5.2 \times 10^{-24}$ | 6 |
| Parasitized | Peripheral blood | IgE | 14 | 53 | 6 | 28 | 8.4 | -0.015 | $-8.3 \times 10^{-29}$ | -0.026 | $-1.2 \times 10^{-27}$ | 5 |
| Immunized with tetanus toxoid | Peripheral blood | IgG[d] | 2 | 57 | 5 | 23 | 8.2 | -0.42 | $-4.4 \times 10^{-20}$ | -0.42 | $-1.5 \times 10^{-19}$ | 16 |
| Immunized with vaccinia virus | Peripheral blood | IgG[d] | 5 | 85 | 7 | 27 | 7.6 | 0.081 | $-1.1 \times 10^{-20}$ | 0.11 | $-5.1 \times 10^{-20}$ | 17 |

[a] In addition, several different allelic variants of these germline genes were used in some repertoires

[b] Test was performed as described by Uduman et al [12]. Positive and negative P-values suggest positive and negative selection, respectively. Repertoires showing significant evidence for selection are shaded in grey.

[c] Corrected for the influence of multiple comparisons using the Benjamini–Hochberg false discover method.

[d] The isotype of this repertoire was not described in detail. However, the 3'-primer used for amplification of the heavy chain variable domain-encoding genes shows perfect match with sequences encoding IgG but 11, 8, 7 and 3 mismatches with sequences encoding IgM, IgA, IgD and IgE, respectively. Furthermore, transcripts encoding IgE are likely to be very rare in this material. Altogether this suggests that the majority of isolated genes encoded IgG.