

Cross-reactions vs co-sensitization evaluated by *in silico* motifs and *in vitro* IgE microarray testing

P. Pfiffner¹, B. M. Stadler¹, C. Rasi², E. Scala² & A. Mari^{2,3}

¹University Institute of Immunology, University of Bern, Bern, Switzerland; ²Center for Molecular Allergology, IDI-IRCCS, Rome; ³Allergy Data Laboratories s.c., Latina, Italy

To cite this article: Pfiffner P, Stadler BM, Rasi C, Scala E, Mari A. Cross-reactions vs co-sensitization evaluated by *in silico* motifs and *in vitro* IgE microarray testing. *Allergy* 2011; DOI: 10.1111/j.1398-9995.2011.02743.x.

Keywords

allergens; bioinformatics; cross-reaction; IgE; sequence motifs.

Correspondence

Beda M. Stadler, University Institute of Immunology, Sahli Haus 2, Inselspital, 3010 Bern, Switzerland.
Tel.: +41 (0)31 632 35 21
Fax: +41 (0)31 381 57 35
E-mail: beda.stadler@iib.unibe.ch

Accepted for publication 2 October 2011

Edited by: Reto Cramer

DOI:10.1111/j.1398-9995.2011.02743.x

Abstract

Background and objective: Using an *in silico* allergen clustering method, we have recently shown that allergen extracts are highly cross-reactive. Here we used serological data from a multi-array IgE test based on recombinant or highly purified natural allergens to evaluate whether co-reactions are true cross-reactions or co-sensitizations by allergens with the same motifs.

Methods: The serum database consisted of 3142 samples, each tested against 103 highly purified natural or recombinant allergens. Cross-reactivity was predicted by an iterative motif-finding algorithm through sequence motifs identified in 2708 known allergens.

Results: Allergen proteins containing the same motifs cross-reacted as predicted. However, proteins with identical motifs revealed a hierarchy in the degree of cross-reaction: The more frequent an allergen was positive in the allergic population, the less frequently it was cross-reacting and vice versa. Co-sensitization was analyzed by splitting the dataset into patient groups that were most likely sensitized through geographical occurrence of allergens. Interestingly, most co-reactions are cross-reactions but not co-sensitizations.

Conclusions: The observed hierarchy of cross-reactivity may play an important role for the future management of allergic diseases.

We have proposed a method to predict cross-reactivity from protein sequence (1). The method uses sequence motifs to identify evolutionary conserved structural domains, which possibly present highly similar surface patches. Hence if two proteins contain the same motif, they are predicted to cross-react. We previously applied this method to assess specific IgE determinations based on allergen extracts in view of cross-reactivity. Interestingly, we found a linear correlation between the number of positive extracts and the number of recognized motifs in a 3 : 1 ratio (2). These findings suggested that cross-reactivity may pose a greater problem than previously assumed. However, we were unable to conclude that these cross-reactions can reliably be predicted through allergen motifs because our study was based on crude allergen extracts, containing many proteins and often several motifs.

With the advent of allergen microarray systems using highly purified natural or recombinant proteins, it became possible to determine specific IgE in a serum directed against many different allergens (3, 4). These data allowed for the

first time a direct comparison between observed co-reactivity at the level of a single allergen protein. Our results show that the predicted cross-reactivity coincides with wet laboratory data and is not because of co-sensitization with other allergens.

Materials and methods

In our study we used specific IgE data of 3142 serum samples analyzed with the Immuno Solid-phase Allergen Chip (ISAC®; Phadia Multiplexing Diagnostics GmbH, PMD, Vienna, Austria). The samples were collected at two sites. Two thousand and five hundred samples were received from patients predominantly living in Italy and 642 samples from patients predominantly living in Austria. No selection criteria other than having at least one positive reaction on the allergen chip were applied and no clinical or demographic information on the donors was included. The two centers used different ISAC chip generations differing in two proteins

only. Each chip contained 103 different, highly purified natural or recombinant allergen proteins immobilized in triplicate onto glass slides. These systems quantitate specific IgE content in ISAC Standardized Units (ISU), a semi-quantitative unit normalized against calibration serum. We collected 311 058 specific IgE values of 105 different allergen proteins. The cut-off value to test positive for a protein was set to 0.1 ISU in order to filter eventual noise. Normalization was achieved by scaling the IgE values between zero and 100. The average of the five highest values per protein and collection site was assumed to be the highest value and set to 100.

The terms *co-reaction*, *cross-reaction* and *co-sensitization* were used as follows: if we observed two proteins positive in the same serum, the term *co-reaction* was used. This expresses a neutral standpoint without implying a mechanism. If co-reacting proteins contained the same motif, that is, our approach predicted these proteins to cross-react, the term *cross-reaction* was used. Co-sensitization was assumed for co-reacting proteins that occur in the same organism.

To connect our three data types – ISAC data, allergen sequences and allergen motifs – a MySQL database was created (MySQL 5.0, obtained from <http://www.mysql.com/>). The allergen sequences to be used in the motif-discovery process were downloaded from Allergome (<http://www.allergome.org> as of December 1, 2010) and imported into the database. MEME 4.4.0 (5), cd-hit 3.1.1 (6) and pftools 2.3.4 (7) were used for the iterative allergen-motif discovery. Data were handled using scripts written in Perl 5.10.0 (<http://www.perl.org/>) and PHP 5.3+ (<http://www.php.net/>) languages. Statistical evaluations were performed with R 2.11.0 (8).

We used our published iterative motif discovery approach to identify allergen motifs (1). Three modifications were applied to our original approach: first, the 2708 protein sequences were preclustered by cd-hit into 1218 clusters with at least 90% sequence identity. Only the longest ('representative') sequence per cluster was subsequently submitted to the motif discovery process; second, MEME was allowed to select protein motifs with a variable length of 35–70 amino acid residues – in our earlier studies we used a fixed length of 50 amino acids; and third, each protein was allowed to contain more than one motif. This was achieved by scanning all proteins against all found motifs instead of assigning only the motif identified during the discovery process.

Results

We had previously demonstrated that the majority of the available allergen sequences can be grouped according to allergen motifs (1). Since then, the number of allergen sequences has more than tripled. Here we have used the same iterative approach to determine the number of motifs. Fig. 1 shows that the number of motifs continued to increase over time, but the motif number seemed to plateau earlier despite the continuous increase of allergen sequences in the Allergome database. After July 2010, we modified motif detection parameters because motif quality and number were occasionally variable as described later. Our modification explains the increase in the number of four to five motifs (data not

shown). Additionally, new motifs were found, which results in an increase of 13 motifs compared to the previous calculation. In this study, motifs were identified in the sequences available from Allergome as of December 2010. Within the 2708 allergen protein sequences, we identified 115 allergen motifs; 311 sequences did not contain a motif. Identified motifs are also displayed in each allergen monograph on Allergome.

Initially, we chose an arbitrary stretch of 50 amino acids for reasons discussed earlier (1). Here we used as a measure of motif quality the *P*-value describing the difference between protein pairs carrying the same motif and the pairs not carrying the same motif in regard to their co-reactivity as determined by the commercial array system. The *P*-value data are shown in Fig. 4. The reason why we modified our previous standard motif length of 50 amino acids will be depicted in Fig. 3. Fig. 2 shows the results using different length approaches. An amino acid residue stretch over 30–35 resulted in a low probability of co-reaction as did stretches above 100 amino acids. Thus we confirmed the previously defined motif length, but 40 or 60 amino acids produced motifs of slightly higher quality. Motif lengths above 70 amino acids resulted in lower motif coverage; thus fewer allergens would match a motif, and the risk of generating multi-domain motifs increased. Therefore we modified motif-detection parameters and used a variable length between 35 and 70 amino acids, as represented by a horizontal line in the figure. The median of the variable length motifs used throughout our studies was 68.5 residues.

As shown in Fig. 3A, some motifs identified in earlier datasets showed a variable correlation coefficient with a group of correlating (0.73–0.96) and a second group of noncorrelating (–0.04–0.14) protein pairs. Here, the correlation coefficient between two proteins was used to describe the correlation between the two proteins' ISU values, that is, if one protein

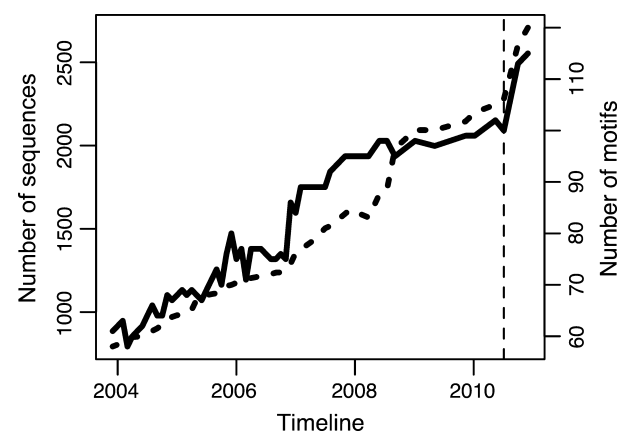


Figure 1 Motif timeline. Number of motifs (continuous line) compared to the number of proteins (dashed line), as gathered from the Allergome database, over time. Motifs have been calculated using a fixed width of 50 amino acid residues. After July 2010, indicated by the dashed vertical line, we changed algorithm parameters to allow any motif length between 35 and 70 amino acids.

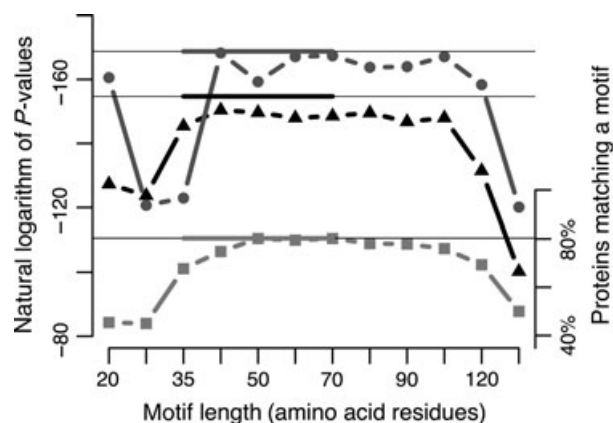


Figure 2 Identification of motifs using different motif lengths. *P*-values are used as explained and depicted in Fig. 4 to assess motif performance. Round (ISAC Standardized Units correlation) and triangular (co-reactivity) symbols apply to the Y-axis on the left, rectangular to the right. Horizontal lines represent results for motifs with a variable motif length of 35–70 amino acids.

had a high ISU level in any given serum, the other also had a high ISU level, and vice versa. Each circle represents one protein pair, and all possible pairs are shown. Using newer datasets, the proteins originally matching motif 2, as shown in Fig. 3A, were separated into two motifs, now containing only highly correlating pairs (0.89–0.96), shown in Fig. 3B. This illustrates that more protein sequences lead not necessarily to more motifs, but motifs may become of higher quality.

Next we assessed the frequency of co-reaction between proteins defined by the same motif. The protein pairs were considered for evaluation only if both proteins were positive in at least 10 different patient sera, which was the case for

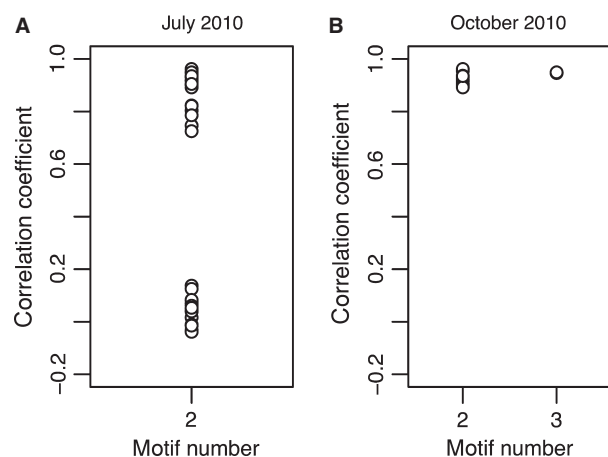


Figure 3 Improvement in motif quality. (A) shows the ISAC Standardized Units (ISU) correlation coefficient for the six proteins matching motif 2 determined in July 2010. Using the dataset of a later time point (October 2010, B), the same proteins are distributed over the previous motif 2 (four proteins) and a new motif 3 (two proteins), both with high ISU correlation coefficients.

6272 pairs. Fig. 4A shows that among the proteins on the chip, 218 pairs each contained the same motif. Protein pairs with a common motif tested significantly more often positive in the same serum (median: 63.6%) than proteins without a common motif (median: 15.7%, $P < 4.9153 \times 10^{-70}$, Mann–Whitney *U*-test). Data distribution was visualized using a box-plot: a gray box was drawn for results between the first and third quantile (0.25–0.75), thus representing 50% of all values, and an individual circle was drawn for all outliers (> 1.5 interquartile ranges distance from the median).

We further analyzed whether specific IgE levels of two proteins also correlated if the proteins shared a motif. Fig. 4B shows that indeed the correlation coefficients of proteins with

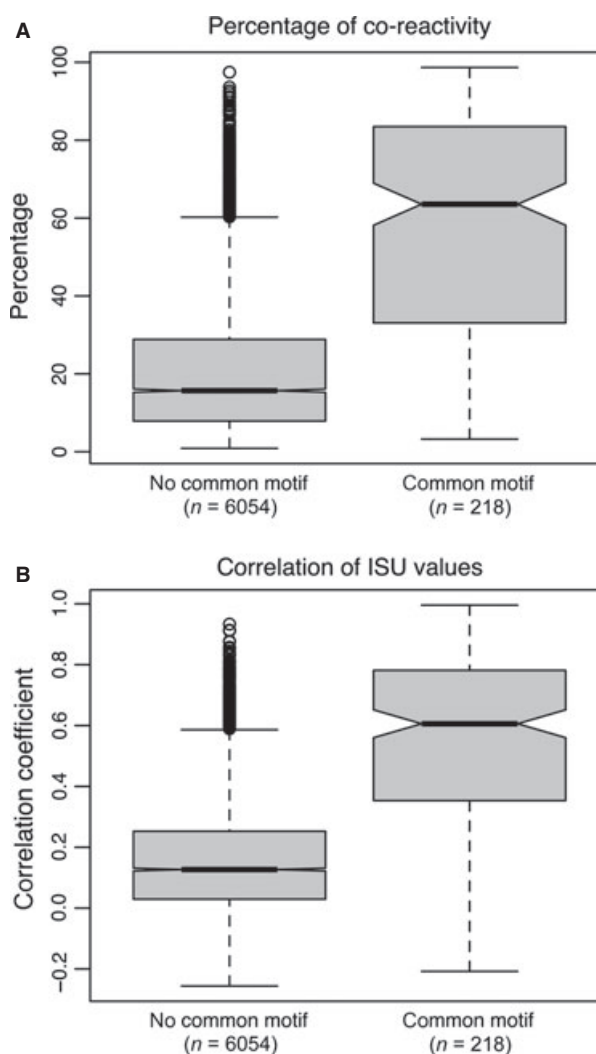


Figure 4 Co-reaction of 105 recombinant allergens on a microarray system (ISAC®; PMD, Austria) as Box-Whisker-plot. Pairs of allergens were grouped according to the presence of a common motif. Motifs were present in 86 of the allergens on the chip. Panel (A) compares the frequency of co-reactivity between the groups while panel (B) compares the correlation coefficient of signal intensity (level of specific IgE).

a common motif were significantly higher (median: 0.61) than the correlation of ISU levels for proteins without a common motif (median: 0.13, $P < 5.0666e-74$, Mann–Whitney U -test). Thus higher specific IgE levels produced also a stronger cross-reacting signal.

The analysis so far has shown that proteins characterized by the same motif co-react at a statistically highly significant level. The crucial question, however, is how often such co-reactions, presumably cross-reactions, would occur within a given motif group. Fig. 5 addresses this question and depicts the results for a selection of motifs where we found three or more allergen proteins sharing the motif. The selection is further based on a minimum of 100 positive sera for at least one of the allergens reacting

on the chip. Interestingly, we observed a hierarchy among the proteins within one motif group, namely each motif group has one representative that produces the maximum of positive reactions and cross-reacting proteins with decreasing frequencies of positivity. The gray shade shows the same value as depicted percent-wise, meaning the least frequent cross-reacting allergen within a motif group showed the most stringent cross-reaction toward allergens in the same group. Figures S1 and S2 provide a breakdown of Fig. 5 by serum origin.

To address the question whether a certain degree of co-reaction might be due to co-sensitization with allergenic organisms containing several motifs, we split the dataset into groups according to the origin of the patient sera (Table 1).

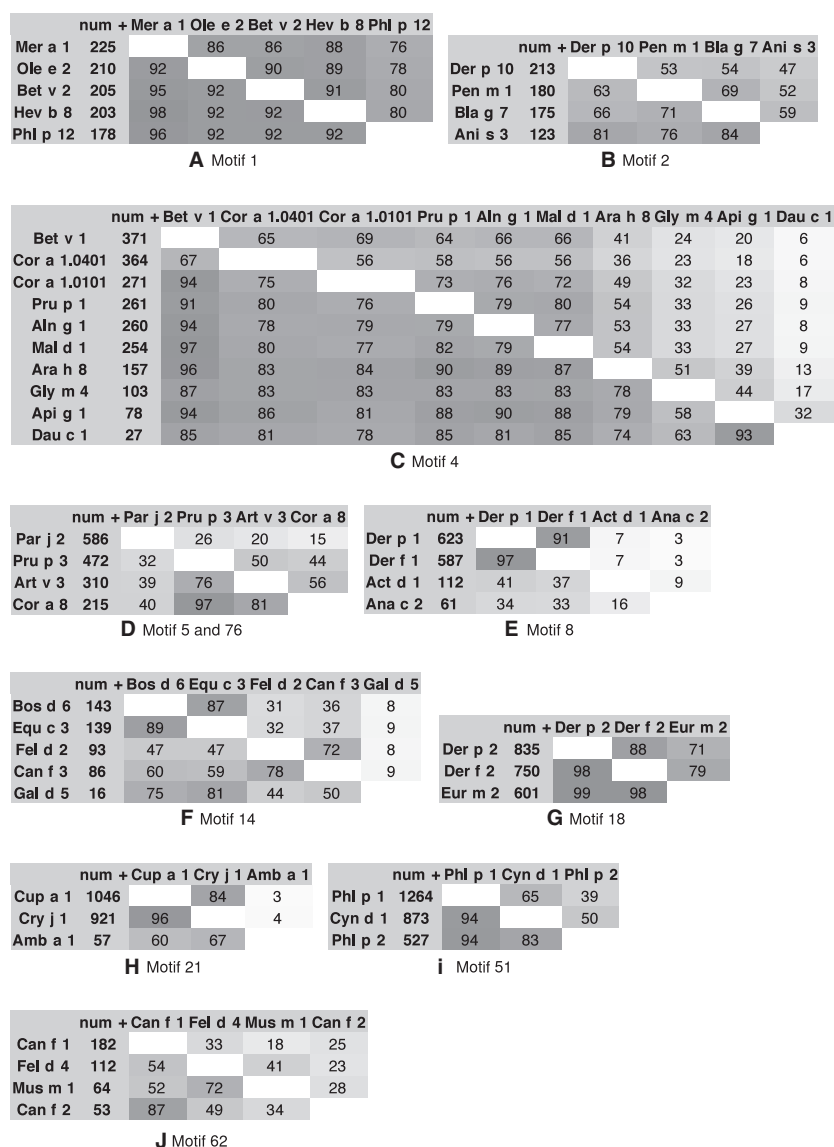


Figure 5 Matrices for all motifs contained in three or more proteins and having at least 100 positive sera. The numbers depict

which percentage of sera positive against the Y-axis protein is also positive against the X-axis protein.

Patient sera from the dataset 'Austria,' based on the serum collection in the more northern hemisphere, were presumably from patients who had encountered birch pollen. These sera recognized Bet v 1 more frequently (22.3%) than sera contained in the dataset 'Italy' (9.1%). The surroundings of Rome can be considered as mostly free from birch tree and birch pollen (9, 10). Unexpectedly, patients from Italy seemed to have higher anti-Bet v 1 IgE levels than patients in the dataset 'Austria', despite the assumption that they were less frequently exposed to birch tree pollen. As there may be differences between laboratories performing the assay or differences in batches delivered to different centers of the microarray chips, we normalized the IgE values to exclude a laboratory bias. Still, patient sera from Italy showed higher normalized specific IgE values than those from Austria. Thus, cross-reactions may deliver a higher diagnostic IgE level than the putative sensitizing allergen.

Complementary to these findings, patients from Austria were more often (49.8%) positive against the German cockroach allergen Bla g 4 than patients from the Italian dataset (3.2%). Bla g 4 is the only protein matched by motif 35.

Table 1 depicts another argument that co-sensitization is less important for *in vitro* diagnostics than cross-reaction. Namely, the two datasets show that allergens such as Bet v 2 (motif 1) produce cross-reactions and almost no co-sensitizations, as predicted.

Discussion

There exist different approaches to determine allergenicity using bioinformatics (11–19). Our initial algorithm was based on the assumption that similar sequences will lead to similar surface structures that may be recognized by a given antibody (1). Since then, we and others have been capable of showing that such an approach is a valid tool for allergenicity prediction (20, 21). Nevertheless the validation of the computational approach with wet laboratory data were ham-

pered for a long time as most specific IgE determinations were based on allergen extracts (2). When comparing allergenicity prediction with wet laboratory data, obtained by diagnostic procedures using allergen extracts, we could confirm the predicted motifs but were left with the question whether some of the observed cross-reactions were true immunological cross-reactions at the protein level or only reactions occurring with nonrelated proteins derived from the same organism within the allergen extract (2).

The situation changed with the availability of new *in vitro* diagnostic procedures such as allergen microchips (22). Now it became possible to extensively compare cross-reactions at the allergen protein level. Before doing so, we re-evaluated our algorithm for defining motifs on a large dataset obtained by such new microchips. As we show here, the number of motifs continuously increased in parallel with the number of new allergen sequences that were added to the Allergome database during the first years. Recently, a plateau of allergen motifs seemed to be reached. Our initial motif length of 50 amino acids was computationally verified; however, we now had diagnostic data at hand. Therefore we re-evaluated our initial approach and we could demonstrate that by using a variable length, a higher motif quality can be achieved, as shown in Fig. 2. As a consequence, the number of motifs rose to a higher level. Future determinations will show whether this slightly higher number of motifs actually represents a plateau again, meaning that most of the potential allergens are currently known because no further common structures in the form of motifs can be detected.

Using the dataset based on the allergen-chip, we show here that cross-reactions were as frequent as previously reported when cross-reactions were determined by allergen extracts. However, there are still some cross-reactions within the group of allergens that have no common motif. On the one hand, this may be because we do not yet have sufficient allergen sequences to construct a motif for every known allergen. On the other hand, observed co-reactions are not true cross-reac-

Table 1 Breakdown of sera positive against Bet v 1

		IgE level				
		Number of positive sera (%)	Mean ISU	Normalized	Median ISU	Normalized
Dataset 'Italy' (total 2500 sera)						
A	Positive against Bet v 1 (motif 4)	228 (9.1)	11.45	18.01	6.05	9.52
B	From A positive against other proteins of motif 4	215 (94.3)	12.12	19.06	6.78	10.66
C	From A positive against Bet v 2 (motif 1)	50 (21.9)	15.8	24.85	9.74	15.32
D	From A positive against Bet v 4 (motif 15)	14 (6.1)	8.32	13.09	1.83	2.88
E	From A not in B, C or D	11 (4.8)	0.31	0.49	0.22	0.35
Dataset 'Austria' (total 642 sera)						
A	Positive against Bet v 1 (motif 4)	143 (22.3)	6.66	14.62	2.26	4.96
B	From A positive against other proteins of motif 4	117 (81.8)	7.97	17.50	3.26	7.16
C	From A positive against Bet v 2 (motif 1)	16 (11.2)	18.55	40.73	10.56	23.18
D	From A positive against Bet v 4 (motif 15)	9 (6.3)	20.58	45.18	17.01	37.35
E	From A not in B, C or D	25 (17.5)	0.76	1.67	0.49	1.08

Breakdown of sera positive against Bet v 1 by dataset origin. Normalized ISAC Standardized Units (ISU) values represent ISU values scaled from 0 to 100 where 100 is the average of the highest 5 values for the respective protein in the respective set.

tions but increased IgE-levels because of the fact that a patient was sensitized with an organism containing several allergens. In other words, our new dataset allowed addressing the question of co-sensitization.

Fortunately, the dataset could be divided into two groups of sera depending on the laboratories that performed the microchip test. One set had been collected by the last author from the greater region of Rome and was termed 'Italy'. The other set was provided by the microchip producer and contains mostly sera from within Austria. As expected, the number of positive sera against Bet v 1 (22.3%) is higher in the dataset 'Austria' than in the dataset 'Italy' (9.1%). This remains true if both datasets were normalized to avoid manufacturer or operator biases. Interestingly, the great number of patients in the vicinity of Rome positive for Bet v 1 cannot be attributed to co-sensitization as birch pollen are not commonly found in this region. To further test the question of co-sensitization, we compared the reaction of patients positive against Bet v 1 to other allergens normally present in birch pollen, namely Bet v 2 and Bet v 4. Indeed, Bet v 2 (motif 1) and Bet v 4 (motif 15) were far less often positive in these patients than other motif-4-containing allergens, again excluding co-sensitization.

Most interestingly, our motif approach in conjunction with wet laboratory evaluation produced a hierarchy of allergens within motif groups. The greater the number of allergens within one group the clearer a hierarchy was apparent. At present it can only be speculated what such a hierarchy might mean for future specific immunotherapy. Many of the 2500 patients from the Italian dataset recognized Bet v 1 probably without being sensitized by this allergen. On the other hand, the allergen that was least frequently positive within a motif group seemed to be most stringently cross-reacting with other allergens higher up in the hierarchy. For example, in the motif 4 group we found 371 sera positive against Bet v 1, but only 27 against Dau c 1 (carrot). Thus testing positive against Dau c 1 happens infrequently; however, if you test positive, your chances of reacting against each other allergen in the motif group is higher than 63%. This opens the question whether the most frequently cross-reacting allergen or an allergen that most stringently cross-reacts should be used for specific immunotherapy.

An interesting observation is our finding that often, the least frequently cross-reacting allergen within one motif group is a food allergen. This favors speculations that oral tolerance might play a role in the development of cross-reactions.

In summary, our highly significant correlations of cross-reactions are a strong argument that the common protein structure of an allergen motif rather than the actual allergen should be considered for future IgE diagnostics or therapy.

Acknowledgments

We are grateful to VBC Genomics, Austria, for their dataset. We thank Michael Stadler for discussion and review of this manuscript. This work was supported by grant no 8803.1 from the Commission of Technology and Innovation, Switzerland, and by grants from the Italian Ministry of Health, Current research funding 2008–2009.

Contributions

Pascal Pffner has performed all experiments except IgE data gathering, created the figures and tables and has written the manuscript. Beda Stadler was involved in the study design, experiment setup and result discussions. He has corrected and edited the manuscript. Chiara Rasi was responsible for the allergen sequence data and the integration of chip data with Allergome sequence data. Enrico Scala collected most patient serum data of the Italian dataset. Adriano Mari was involved in the study design and result discussions. He was responsible for the Allergome data integration and has written parts of the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article found at: <http://www.wileyonlinelibrary.com>

Figure S1. Fraction matrices, dataset 'Italy'. Matrices for the motif groups from Fig. 5 with data only from the dataset 'Italy'.

Figure S2. Fraction matrices, dataset 'Austria'. Matrices for the motif groups from Fig. 5 with data only from the dataset 'Austria'.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1. Stadler MB, Stadler BM. Allergenicity prediction by protein sequence. *FASEB J* 2003;17:1141–1143.
2. Pffner P, Truffer R, Matsson P, Rasi C, Mari A, Stadler BM. Allergen cross-reactions: a problem greater than ever thought? *Allergy* 2010;65:1536–1544.
3. Scala E, Alessandri C, Bernardi ML, Ferrara R, Palazzo P, Pomponi D et al. Cross-sectional survey on immunoglobulin E reactivity in 23,077 subjects using an allergenic molecule-based microarray detection system. *Clin Exp Immunol* 2010;40:911–921.
4. Mari A, Alessandri C, Bernardi ML, Ferrara R, Scala E, Zennaro D. Microarrayed allergen molecules for the diagnosis of allergic diseases. *Curr Allergy Asthma Rep* 2010;10:357–364.
5. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 1994;2:28–36.
6. Li W, Jaroszewski L, Godzik A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* 2001;17:282–283.

7. Bucher P, Karplus K, Moeri N, Hofmann K. A flexible motif search technique based on generalized profiles. *Comput Chem* 1996;**20**:3–23.
8. R Core Development Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2008. ISBN 3-900051-07-0, <http://www.R-project.org>.
9. Caiola MG, Mazzitelli A, Capucci E, Travaglini A. Monitoring pollinosis and airborne pollen in a Rome university. *Aerobiologia* 2002;**18**:267–275.
10. Mari A, Wallner M, Ferreira F. Fagales pollen sensitization in a birch-free area: a respiratory cohort survey using Fagales pollen extracts and birch recombinant allergens (rBet v 1, rBet v 2, rBet v 4). *Clin Exp Allergy* 2003;**33**:1419–1428.
11. Li KB, Issac P, Krishnan A. Predicting allergenic proteins using wavelet transform. *Bioinformatics* 2004;**20**:2572–2578.
12. Saha S, Raghava GPS. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 2006;**34**:W202–W209.
13. Zhang ZH, Koh JL, Zhang GL, Choo KH, Tammi MT, Tong JC. AllerTool: a web server for predicting allergenicity and allergic cross-reactivity in proteins. *Bioinformatics* 2006;**23**:504–506.
14. Cui J, Han LY, Li H, Ung CY, Tang ZQ, Zheng CJ et al. Computer prediction of allergen proteins from sequence-derived protein structural and physicochemical properties. *Mol Immunol* 2007;**44**:514–520.
15. Martinez Barrio A, Soeria-Atmadja D, Nistér A, Gustafsson MG, Hammerling U, Bongcam-Rudloff E. EVALLER: a web server for in silico assessment of potential protein allergenicity. *Nucleic Acids Res* 2007;**35**:W694–W700.
16. Tong JC, Tammi MT. Prediction of protein allergenicity using local description of amino acid sequence. *Front Biosci* 2008;**13**:6072–6078.
17. Lim SJ, Tong JC, Chew FT, Tammi MT. The value of position-specific scoring matrices for assessment of protein allergenicity. *BMC Bioinformatics* 2008;**9**(Suppl 12):S21.
18. Muh HC, Tong JC, Tammi MT. AllerHunter: a SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. *PLoS ONE* 2009;**4**:e5861.
19. Ivanciuc O, Schein CH, Garcia T, Oezguen N, Negi SS, Braun W. Structural analysis of linear and conformational epitopes of allergens. *Regul Toxicol Pharmacol* 2009;**54**(3 Suppl):S11–S19.
20. Thomas K, Bannon G, Hefle S, Herouet C, Holsapple M, Ladics G et al. In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report, 23–24 February 2005. *Toxicol Sci* 2005;**88**:307–310.
21. Schein CH, Ivanciuc O, Braun W. Bioinformatics approaches to classifying allergens and predicting cross-reactivity. *Immunol Allergy Clin North Am* 2007;**27**:1.
22. Ferrer M, Sanz ML, Sastre J, Bartra J, del Cuvillo A, Montoro J et al. Molecular diagnosis in allergology: application of the microarray technique. *J Invest Allergol Clin Immunol* 2009;**19**(Suppl 1):19–24.

num + Mer a 1 Ole e 2 Bet v 2 Hev b 8 Phl p 12						
Mer a 1	198		87	87	88	78
Ole e 2	188	91		89	88	78
Bet v 2	183	94	92		91	80
Hev b 8	179	97	92	93		82
Phl p 12	161	96	91	91	91	

(a) Motif 1

num + Der p 10 Pen m 1 Bla g 7 Ani s 3					
Der p 10	196		51	53	44
Pen m 1	165	61		69	49
Bla g 7	164	63	70		56
Ani s 3	108	80	75	85	

(b) Motif 2

num + Bet v 1 Cor a 1.0401 Cor a 1.0101 Aln g 1 Pru p 1 Mal d 1 Ara h 8 Gly m 4 Api g 1 Dau c 1											
Bet v 1	228		75	86	70	68	66	40	28	19	07
Cor a 1.0401	222	77		71	64	66	64	39	28	19	07
Cor a 1.0101	210	93	75		71	70	67	42	30	19	07
Aln g 1	174	91	82	86		76	72	50	34	24	09
Pru p 1	174	89	84	84	76		77	51	34	24	09
Mal d 1	159	95	89	88	79	84		52	36	26	10
Ara h 8	97	94	90	92	90	92	86		57	36	13
Gly m 4	76	83	83	83	79	79	76	72		37	14
Api g 1	48	90	88	81	88	88	85	73	58		35
Dau c 1	19	79	79	79	79	84	84	68	58	89	

(c) Motif 4

num + Par j 2 Pru p 3 Art v 3 Cor a 8					
Par j 2	581		26	20	14
Pru p 3	455	33		49	44
Art v 3	278	41	80		60
Cor a 8	209	40	97	80	

(d) Motif 5 and 76

num + Der p 1 Der f 1 Act d 1 Ana c 2					
Der p 1	602		91	07	02
Der f 1	560	98		07	02
Act d 1	108	41	36		07
Ana c 2	46	30	28	17	

(e) Motif 8

num + Bos d 6 Equ c 3 Can f 3 Fel d 2 Gal d 5						
Bos d 6	131		89	34	27	08
Equ c 3	130	89		33	28	08
Can f 3	66	67	65		71	09
Fel d 2	59	61	61	80		08
Gal d 5	14	71	79	43	36	

(f) Motif 14

	num +	Der p 2	Der f 2	Eur m 2
Der p 2	791		89	74
Der f 2	713	99		82
Eur m 2	594	99	98	

(g) Motif 18

	num	+ Cup a 1	Cry j 1	Amb a 1
Cup a 1	1027		84	03
Cry j 1	895	97		03
Amb a 1	47	60	66	

(h) Motif 21

num + Phl p 1 Cyn d 1 Phl p 2				
Phl p 1	1130		66	39
Cyn d 1	798	94		50
Phl p 2	467	94	85	

(i) Motif 51

num + Can f 1 Fel d 4 Mus m 1 Can f 2					
Can f 1	131		21	11	21
Fel d 4	58	48		34	21
Mus m 1	36	42	56		19
Can f 2	31	90	39	23	

(j) Motif 62

num + Mer a 1 Hev b 8 Ole e 2 Bet v 2 Phl p 12										
Mer a 1	27		89	81	81	63				
Hev b 8	24	100		88	83	67				
Ole e 2	22	100	95		91	73				
Bet v 2	22	100	91	91		77				
Phl p 12	17	100	94	94	100					

num + Der p 10 Ani s 3 Pen m 1 Bla g 7									
Der p 10	17		82	76	65				
Ani s 3	15	93		80	73				
Pen m 1	15	87	80		73				
Bla g 7	11	100	100	100					

(k) Motif 1

(l) Motif 2

num + Bet v 1 Cor a 1.0401 Mal d 1 Pru p 1 Aln g 1 Cor a 1.0101 Ara h 8 Api g 1 Gly m 4 Dau c 1											
Bet v 1	143		50	66	58	60	43	42	21	19	06
Cor a 1.0401	142	51		44	44	42	32	30	18	15	05
Mal d 1	95	100	65		79	79	59	57	29	28	07
Pru p 1	87	95	72	86		83	60	61	31	30	08
Aln g 1	86	100	70	87	84		65	60	33	30	08
Cor a 1.0101	61	100	74	92	85	92		70	39	38	10
Ara h 8	60	100	72	90	88	87	72		45	42	12
Api g 1	30	100	83	93	90	93	80	90		57	27
Gly m 4	27	100	81	100	96	96	85	93	63		22
Dau c 1	8	100	88	88	88	88	75	88	100	75	

(m) Motif 4

num + Art v 3 Pru p 3 Cor a 8 Par j 2						num + Der f 1 Der p 1 Ana c 2 Act d 1					
Art v 3	32		38	19	16	Der f 1	27		78	26	07
Pru p 3	17	71		35	24	Der p 1	21	100		33	10
Cor a 8	6	100	100		33	Ana c 2	15	47	47		13
Par j 2	5	100	80	40		Act d 1	4	50	50	50	

(n) Motif 5 and 76

(o) Motif 8

num + Fel d 2 Can f 3 Bos d 6 Equ c 3 Gal d 5						
Fel d 2	34		59	24	24	06
Can f 3	20	100		40	40	10
Bos d 6	12	67	67		67	17
Equ c 3	9	89	89	89		22
Gal d 5	2	100	100	100	100	

num + Der p 2 Der f 2 Eur m 2				
Der p 2	44		75	16
Der f 2	37	89		19
Eur m 2	7	100	100	

(p) Motif 14

(q) Motif 18

num + Cry j 1 Cup a 1 Amb a 1							
Cry j 1	26		62	27			
Cup a 1	19	84		32			
Amb a 1	10	70	60				

num + Phl p 1 Cyn d 1 Phl p 2							
Phl p 1	134		51	43			
Cyn d 1	75	91		55			
Phl p 2	60	97	68				

(r) Motif 21

(s) Motif 51

num + Fel d 4 Can f 1 Mus m 1 Can f 2									
Fel d 4	54		59	48	26				
Can f 1	51	63		35	35				
Mus m 1	28	93	64		39				
Can f 2	22	64	82	50					

(t) Motif 62