



Allergens as eukaryotic proteins lacking bacterial homologues

Cecilia Emanuelsson^{a,*}, Michael D. Spangfort^b

^a Department of Biochemistry, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, S-221 00 Lund, Sweden

^b ALK-Abelló, Bøge Allé 6-8, DK-2970 Hørsholm, Denmark

Received 5 December 2006; received in revised form 16 January 2007; accepted 17 January 2007

Abstract

Only a small number of protein homologues cause the majority of allergies. There is no consensus structure or other obvious common denominator discriminating the few proteins that are allergens from thousands of other, non-allergenic proteins. By database sequence homology searching, we here show that to date known allergen sequences have no or few bacterial homologues, in contrast to randomly selected control protein sequences. This finding suggests a novel common denominator for allergens of potential use for allergen prediction programs. A possible interpretation of this finding is that allergens are proteins which are exposed to the immune system and which lack bacterial homologues. This interpretation is discussed in relation to the many observations that allergies coincide with a delayed establishment of infant gut flora.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Allergen; Allergen prediction; Homology searching; Hygiene hypothesis

1. Introduction

A small number of protein homologues are responsible for the majority of the allergies (Aalberse et al., 2001; Breiteneder and Ebner, 2000; Mills et al., 2003; Vieths et al., 2002). With regard to non-occupational Type I allergy, most allergens are either inhaled (pollen proteins, dust mite proteins, or proteins in animal fur, dander or saliva), or ingested with food. Allergens derived from food are typically proteins that are either homologous to inhaled allergen proteins, or very stable proteins that resist complete degradation in the intestine due to the presence of abundant disulfide bonds. The Allergen nomenclature database maintained by the International Union of Immunological Societies (<http://www.allergen.org>) contains approximately 500 documented allergen proteins that are known to date. More allergens may remain to be discovered and characterized with increased incidence rates of allergy, yet the situation is such that among hundreds of thousands of proteins only a fraction is allergenic. The common denominator(s) among this very limited number of proteins which evoke allergies is still not fully understood (Bredehorst and David, 2001; Huby et al., 2000). There is no consensus allergen structure (Aalberse, 2006).

Here we describe a novel common denominator for allergens, based on homology searches in protein sequence databases. Whereas all randomly selected control proteins showed hundreds of bacterial homologues, the allergens showed no, or very few, bacterial homologues. To interpret this finding, we put forward the idea that allergens are eukaryotic proteins which lack bacterial homologues and which therefore may become allergens if presented to the human immune system.

2. Experimental

2.1. Allergen sequences

In the protein sequence database UniProt Knowledgebase (Release 8.4 25-Jul-2006/ Swiss-Prot Release 50.4: 230,133 entries, <http://www.expasy.org/sprot>), 391 protein sequence entries were annotated as “Allergen”.

2.2. Homology searching

Of these 391 allergen protein sequences, 30 were randomly selected (by taking every 12th sequence from the list) for homology searching, using the BLAST program (Altschul et al., 1990) at ExPASy/SIB to search the subsection Bacteria + Archae in the non-redundant UniProtKB/Swiss-Prot (Release 50.4: 230,133 entries) and the larger computer-annotated supplement

* Corresponding author. Tel.: +46 46 222 4872; fax: +46 46 222 4534.

E-mail address: Cecilia.Emanuelsson@biochemistry.lu.se
(C. Emanuelsson).

Table 1

Protein sequences annotated as allergens subjected to database sequence homology searching to evaluate number of possible bacterial homologues

Species	Allergen ^a			Control protein, randomly selected from the same species	
	Sp acc nr	Name	# Bact homol ^a	Sp acc nr + name	# Bact homol ^b
<i>Fagopyrum esculentum</i>	Q9XFM4	Fag e 1, seed storage prot 3	1	P00072 CYC_FAGES	81
<i>Equus caballus</i>	P81216	Dander allergen Equ c 2	0	P08907 AATM_HORSE	>100
<i>Triticum aestivum</i>	P81496	Allergen C-C wheat	0	Q43199 APT1_WHEAT	>100
<i>Bos taurus</i>	Q28133	Dander maj allerg Bos d 2	0	P34955 A1AT_BOVIN	>100
<i>Aspergillus fumigatus</i>	O60024	Allergen Asp f 4, fungi	0	P78746 CHSD_ASPFU	69
<i>Dermatophagoides farinae</i>	Q26456	Mite allergen Der f 7	0	(No control) ^c	– ^c
<i>Blatella germanica</i>	P54958	Bla g II, Asp protease	0	P54961 HMC1_BLAGE	>100
<i>Betula verrucosa</i>	P43180	Maj allergen Bet v 1-G	0	P51075 CHSY_BETVE	>100
<i>Ctenocephalides felis</i>	Q94424	Allergen Cte f 1 cat flea	0	P29872 COX2_CTEFE	>100
<i>Cladosporium herbarum</i>	P40108	Allergen Cla h 3, fungi	>100	(No control) ^c	– ^c
<i>Hevea brasiliensis</i>	Q7Y1X1	Latex allergen Hev b 13	7	P52704 HNL_HEVBR	50
<i>Felis silvestris</i>	P30439	Maj allergen Fel d 1-A, cat	0	P48892 COX3_FELCA	>100
<i>Apis mellifera</i>	Q08169	Api m 2, honey bee	0	P19039 EF1A_APIME	>100
<i>Ligustrum vulgare</i>	O82015	Major pollen allerg Lig v 1	0	Q05991 RBL_LIGVU	>100
<i>Malassezia furfur</i>	P56578	Allergen Mal f 3 yeast	>100	Q2VG90 Q2VG90_MALFU	51
<i>Phalaris aquatica</i>	P56165	Maj pollen allergen Pha a 5	0	(No control) ^c	– ^c
<i>Poa pratensis</i>	P22285	Pollen allergen Poa p 9	0	Q32880 NU5C_POAPR	>100
<i>Ambrosia artemisiifolia</i>	P00304	Pollen allergen Amb a 3	0	(No control) ^c	– ^c
<i>Lolium perenne</i>	P14946	Pollen allergen Lol p 1	0	O22380 CADH_LOLPR	>100
<i>Rattus norvegicus</i>	P02761	Rat n 1Alpha-2u-globulin	0	P06686 A1A2_RAT	>100
<i>Prunus armeniaca</i>	P81651	Major allergen Pru ar 3	0	O50003 RL12_PRUAR	59
<i>Polistes annularis</i>	Q9U6W0	Allergen Pol a 1, wasp	0	(No control) ^c	– ^c
<i>Phleum pratense</i>	Q8H6L7	Pollen allergen Phl p 11	0	(No control) ^c	– ^c
<i>Malus sylvestris</i>	Q9XF40	Pollen allerg Mal d 4 apple	0	Q9XHG2 FLS_MALDO	>100
<i>Betula verrucosa</i>	P25816	Pollen allerg Bet v 2 birch	0	P27783 NIA_BETVE	>100
<i>Oryza sativa</i>	Q01881	Seed allerg en RA5, rice	0	P27932 AM3A_ORYSA	>100
<i>Solanum tuberosum</i>	P30941	Allergen Sola t 4, potato	0	P14673 ADH1_SOLTU	>100
<i>Homarus americanus</i>	O44119	Tropomyosin alle Hom a 1	>100	P00357 G3P_HOMAM	>100
<i>Vespula flavopilosa</i>	P35783	Allergen Ves f 1, wasp	0	(No control) ^c	– ^c
<i>Arachis hypogaea</i>	P43237	Allergen Ara h 1	0	P20178 THS1_ARAHY	>100

^a In UniProt Knowledgebase Release 8.4 25-Jul-2006, 391 entries in Swiss-Prot were annotated as “Allergen”. Of these, 30 of the Swiss-Prot sequences were randomly selected for BLAST searching at ExPASy/SIB in the subsection of Swiss-Prot + TrEMBL called “Bacteria + Archae”, and the number of bacterial homologues (# bact homol) was determined as the number of matches with $E < 0.02$.

^b For each allergen sequence, another entry from the same species was randomly selected and subjected to a similar homology searching, and the number of bacterial homologues (# bact homol) was determined as the number of matches in Bacteria + Archae with $E < 0.02$.

^c (no control) means that the only entries in Swiss-Prot for this species were allergen protein sequences, hence no control protein sequence could be run as comparison.

UniProtKB/TrEMBL (Release 33.4: 3051,654 entries). Default settings were used (gapped alignment, low-complexity-region filter on, Autoselect matrix, i.e. Blosum62 if query length > 85 aa, Blosum80 if 50–5 aa), except that fragment sequences were excluded. The number of bacterial homologues was determined as the number of matches with $E < 0.02^1$, i.e. matches due to homology.

2.3. Non-allergen control sequences

For each allergen sequence, a non-allergen control sequence was randomly selected by taking the first non-allergen sequence on the species-specific entries sorted after name. For example, for the allergen sequence P43180, Bet v 1-G, the link *Betula verrucosa* shows a list with 24 entries, mostly isoforms of this allergen and other allergens identified in birch pollen (18

entries). Of the other six sequences, which were non-allergens, the uppermost (P51075, Chalcone synthase) was taken as control sequence, and subjected to homology searching accordingly. In some cases a particular species was represented in the database by allergen sequences only, as e.g. in the case of the dust mite *Dermatophagoides farinae*, hence no control protein sequence could be run as comparison.

3. Results

In the protein sequence database Swiss-Prot, close to 400 out of approximately 200,000 entries (0.2 %) are annotated as “Allergens”. Pairwise sequence alignment to other sequences showed that 25 of 30 randomly selected allergen sequences in the database matched no bacterial homologues (Table 1). Contrastingly, control sequences from non-allergen proteins matched a large number of bacterial homologues.

A few exceptions were observed: Five of the 30 allergen sequences did match bacterial homologues according to Table 1.

¹ The expected frequency value E indicates the confidence level and describes the number of matches expected to occur by chance when searching the database.

Table 2

Protein sequences of selected common major allergens from pollen, dust mite and cat subjected to database sequence homology searching to evaluate number of possible bacterial homologues

Source	Name	UniProtID	aa	No of bacterial homologues ($E < 0.02$)	No of human homologues ($E < 0.02$)
<i>Betula verrucosa</i>	Bet v 1	P15494	159	0	0
<i>Phleum pratense</i>	Mature Phl p 1	Q40967	240	1 ^a	0
<i>Phleum pratense</i>	Phl p 5	O65320	276	0	0
<i>Dermatophagoides pteronyssinus</i>	Mature Der p 1	P08176	222	42 ^b	0
<i>Dermatophagoides pteronyssinus</i>	Mature Der p 2	P49278	129	0	0
<i>Ambrosia artemisiifolia</i>	Mature Amb a 1	P27759	371	76 ^c	0
<i>Artemisia vulgaris</i>	Mature Art v 1	Q84ZX5	132	0	0
<i>Felis domesticus</i>	Fel d 1, mature chain 1	P30438	70	0	1 ^d
<i>Felis domesticus</i>	Fel d 1, mature chain 2	P30440	92	0	0

Nine of the most clinically relevant allergens were selected for BLAST searching as in Table 1 at ExPASy/SIB in the subsections of UniProt Knowledgebase called “Bacteria + Archae” and “Homo sapiens”.

^a Sequence identity 50/185 = 27%.

^b Sequence identities ranging from 81/250 = 32% to 50/221 = 22%.

^c Sequence identities ranging from 67/228 = 29% to 40/178 = 22%.

^d Sequence identity 20/70 = 28%.

Two (Q9XFM4 and Q7Y1X1) matched only a few bacterial sequences (1 and 7, respectively), and the remaining three (P40108, P56578 and O44119) matched a large number of bacterial homologues. The latter allergen, tropomyosin from lobster, is an autoallergen like albumin, profilin and MnSOD, which have >50% sequence identity to human homologues, their immune response evoked in a not fully understood manner by the parts that are not identical to the human homologue (Aalberse et al., 2001).

The randomly retrieved allergen sequences listed in Table 1 contains several grass allergens, 3 mammalian lipocalins and 2 profilins. Therefore one might object that these observations are not entirely independent. The two venom allergens (Pol a, Ves f) become injected into human body and may have different action mechanisms than the airborne allergens which are inhaled. Some well-known and common major airborne allergens were therefore selected for similar homology searching.

Pollen-, dust mite- and cat-allergy are frequent allergies, caused by a limited number of airborne, clinically important allergens. The sequences for these allergens were also subjected to homology searching after removal of their signal/pre/pro-peptides which are not encountered by the human immune system. Also these selected major allergens largely lack bacterial, as well as human, homologues (Table 2). There were some exceptions, since three of the common allergens did match bacterial homologues. One showed only 1 match. All three had a sequence identity with the matched homologues around 30% or less. The antibody cross-reactivity between these common allergens and the bacterial homologues is therefore expected to be low.

4. Discussion

We here show that allergen sequences lack bacterial homologues, whereas randomly selected control sequences have hundreds of bacterial homologues. A protein can become allergenic only provided that it does come into contact with

the human immune system. Even more appropriate control sequences would therefore be other proteins that – like the allergens – are available in a soluble form in the allergen source, and which do become presented to the immune system under natural conditions.

There is however no trivial way to retrieve such relevant control sequences because of lack of annotation in the sequence databases. In case of for example pollen allergens, not very much is known about other proteins (non-allergens) expressed in pollen, except for a few recent expression analysis studies on rice and *Arabidopsis thaliana*. For the latter it is estimated that 6000 out of 26 000 genes are expressed in pollen, and a minor fraction of these proteins have been identified by proteomics showing 135 proteins of all kinds of functional properties (Holmes-Davis et al., 2005). Out of the 25 most abundantly expressed of these pollen proteins, we found by homology searching that 80% do have bacterial homologues (data not shown).

That allergen sequences lack bacterial homologues is to our knowledge a novel finding, that should be further analyzed by and taken into consideration as a possible novel common denominator for allergen proteins in bioinformatics-based allergen prediction programs (Holt and Thomas, 2005). Web-tools such as Allermatch (Bairoch et al., 2005), WebAllergen (Riaz et al., 2005) and AlgPred (Saha and Raghava, 2006) rely on sequence similarity to known allergens combined with depletion of motifs and signatures present in both allergens and non-allergens (Bjorklund et al., 2005). Because the thresholds for false positives and true allergens are interrelated there is a risk of missing potentially allergenic proteins. Even small pieces of new information on features typical for allergens can potentially improve prediction programs.

That allergens lack bacterial homologues also has implications for the question of why such a small set of proteins are allergenic (Aalberse, 2006; Bredehorst and David, 2001; Huby et al., 2000), or rather, why most proteins are not allergenic. One interpretation of our finding is that only proteins that are inhaled and presented to the immune system and which

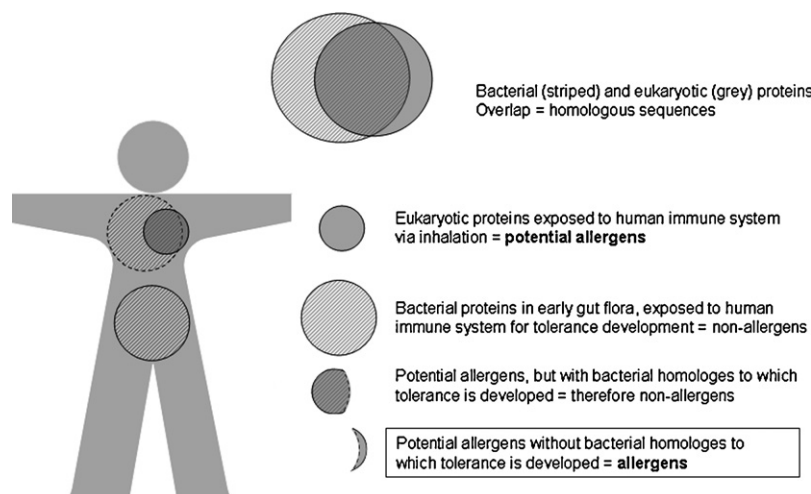


Fig. 1. Allergens are eukaryotic proteins which lack bacterial homologues – an interpretation in terms of tolerance development against commensal bacterial proteins. Shared homologues between proteomes are represented as the cross-section between eukaryotic (grey) and bacterial (striped) proteomes. The *potential allergens* are only the small part of all eukaryotic proteins which becomes presented to the human immune system, mainly by inhalation. Potential allergens which are homologous to the bacterial proteins to which tolerance is developed (cross-section, between eukaryotic (grey) and bacterial (striped) proteomes), therefore become *non-allergens*, assuming that tolerance is developed to commensal bacterial proteins in early gut flora. Only a fraction of potential allergens (grey, outside cross-section), that are not homologous to the bacterial proteins to which tolerance is developed, become *allergens*. The property of lacking bacterial homologues is in itself not enough for a protein to become an allergen. Probably most eukaryotic proteins which lack bacterial homologues are not allergens. Among the potential allergens, presumably the majority has bacterial homologues. All the randomly selected eukaryotic sequences in Table 1 had bacterial homologues.

lack bacterial homologues become allergens, as outlined in Fig. 1.

An early exposure to bacterial proteins could thus create tolerance against later inhaled homologous proteins. Exposure to bacterial proteins is known to be important in order not to develop allergy, and a delayed establishment of gut flora in newborn infants correlates with atopy and allergy (Bjorksten et al., 1999; Bjorksten et al., 2001; Holt et al., 1995, 1997; Kemp and Bjorksten, 2003). Possibly, a systemic tolerance could be developed against bacterial proteins in the non-pathogenic commensal bacterial gut flora (Hooper and Gordon, 2001; Tlaskalova-Hogenova et al., 2004; Xavier and Podolsky, 2000), such that B-cells or T-cells directed to the bacterial proteins become depleted from the repertoire if regarded as auto-reactive. Bacteria may also induce low-affinity polyreactive antibodies that cross-react with bacterial homologues (Notkins, 2004; Wang et al., 2001). Antigen presentation of bacterial proteins could also occur in the context of simultaneous stimulation of innate immunity mechanisms by bacterial agents with strong effects on regulatory T-cells (Elson et al., 2007).

That allergy correlates with delayed establishment of gut flora and a low pathogen burden, i.e. the hygiene hypothesis (Ring et al., 2004; Strachan, 1989; Wold, 1998; Yazdanbakhsh et al., 2002), has also been explained by induction of regulatory T-cells by bacteria (Ostman et al., 2006) or parasitic worms (Smits et al., 2005b), which in turn produce down-regulatory molecules such as IL-10 (Smits et al., 2005a).

The conclusion that allergens lack bacterial homologues based on homology searching is dependent on whether the sequences in the databases are representative or not. The homology searching with the allergen sequences retrieved from Swiss-Prot database was made in the 10 times larger TrEMBL database, a computer-annotated supplement which contains

also all the translations of EMBL nucleotide sequence entries not yet integrated to Swiss-Prot. Most sequences get into the sequence databases after full genome sequencing and more bacterial than eukaryotic species are sequenced, therefore more than 50% of the sequences are bacterial (121,552 in Bacteria plus Archae, out of in total 231,434, UniProtKB/Swiss-Prot Release 50.6 of 05-Sep-2006). Therefore it seems unlikely that the lack of bacterial homologues displayed by the allergen sequences in Tables 1 and 2 should be due to absence of bacterial sequences in the databases.

From an immunological perspective the $E < 0.02$ cut-off used in Tables 1 and 2 certainly includes bacterial homologues to the non-allergen sequences with no prospect of antibody cross-reactivity (usually requires >40% sequence identity). It remains to be determined how much and what parts of a eukaryotic protein sequence that needs to be aligned to a bacterial sequence for the eukaryotic protein not to become an allergen. Probably not only structural homology is of importance but also T-cell epitopes (Quilot et al., 2004; Reche et al., 2005). A closer investigation would be needed to evaluate the sequence identity between each of the non-allergens and their >100 detected bacterial homologues, and to address the question of how large the sequence identity must be for tolerance development.

There are many unresolved issues regarding the increased incidence of sensitization against allergens (Holt and Thomas, 2005). The finding, that allergens as opposed to non-allergens are eukaryotic proteins without bacterial homologues interpreted in terms of tolerance development against bacterial proteins, offers a possible additional explanation to why promoting a sound gut flora in newborn infants (Bjorksten, 2005; Matricardi et al., 2003) can be helpful to ameliorate and prevent atopic diseases and allergies.

Acknowledgements

We thank Kåre Meno at ALK-Albello and Patrick Holt and Wayne Thomas at Telethon Institute for Child Health Research in Perth for critical reading of the manuscript, and Nick Sirijovsky at Biochemistry department at Lund University for comments and linguistic advice. C.E. acknowledges the funding from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Dnr 225-2004-1790).

References

- Aalberse, R.C., 2006. Structural features of allergenic molecules. *Chem. Immunol. Allergy* 91, 134–146.
- Aalberse, R.C., Akkedaas, J., van Ree, R., 2001. Cross-reactivity of IgE antibodies to allergens. *Allergy* 56, 478–490.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., Yeh, L.-S.L., 2005. The Universal Protein Resource (UniProt) 10.1093/nar/gki070. *Nucl. Acids Res.* 33, D154–D159.
- Bjorklund, A.K., Soeria-Atmadja, D., Zorzet, A., Hammerling, U., Gustafsson, M.G., 2005. Supervised identification of allergen-representative peptides for in silico detection of potentially allergenic proteins. *Bioinformatics* 21, 39–50.
- Bjorksten, B., Naaber, P., Sepp, E., Mikelsaar, M., 1999. The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin. Exp. Allergy* 29, 342–346.
- Bjorksten, B., 2005. Evidence of probiotics in prevention of allergy and asthma. *Curr. Drug Targets Inflamm. Allergy* 4, 599–604.
- Bjorksten, B., Sepp, E., Julge, K., Voor, T., Mikelsaar, M., 2001. Allergy development and the intestinal microflora during the first year of life. *J. Allergy Clin. Immunol.* 108, 516–520.
- Breddehorst, R., David, K., 2001. What establishes a protein as an allergen? *J. Chromatogr. B Biomed. Sci. Appl.* 756, 33–40.
- Breiteneder, H., Ebner, C., 2000. Molecular and biochemical classification of plant-derived food allergens. *J. Allergy Clin. Immunol.* 106, 27–36.
- Elson, G., Dunn-Siegrist, I., Daubeuf, B., Pugin, J., 2007. Contribution of toll-like receptors to the innate immune response to Gram-negative and Gram-positive bacteria. *Blood* 109, 1574–1583.
- Holmes-Davis, R., Tanaka, C.K., Vensel, W.H., Hurkman, W.J., McCormick, S., 2005. Proteome mapping of mature pollen of *Arabidopsis thaliana*. *Proteomics* 5, 4864–4884.
- Holt, P.G., O'Keeffe, P., Holt, B.J., Upham, J.W., Baron-Hay, M.J., Suphioglu, C., Knox, B., Stewart, G.A., Thomas, W.R., Sly, P.D., 1995. T-cell “priming” against environmental allergens in human neonates: sequential deletion of food antigen reactivity during infancy with concomitant expansion of responses to ubiquitous inhalant allergens. *Pediatr. Allergy Immunol.* 6, 85–90.
- Holt, P.G., Sly, P.D., Bjorksten, B., 1997. Atopic versus infectious diseases in childhood: a question of balance? *Pediatr. Allergy Immunol.* 8, 53–58.
- Holt, P.G., Thomas, W.R., 2005. Sensitization to airborne environmental allergens: unresolved issues. *Nat. Immunol.* 6, 957–960.
- Hooper, L.V., Gordon, J.I., 2001. Commensal host-bacterial relationships in the gut. *Science* 292, 1115–1118.
- Huby, R.D., Dearman, R.J., Kimber, I., 2000. Why are some proteins allergens? *Toxicol. Sci.* 55, 235–246.
- Kemp, A., Bjorksten, B., 2003. Immune deviation and the hygiene hypothesis: a review of the epidemiological evidence. *Pediatr. Allergy Immunol.* 14, 74–80.
- Matricardi, P.M., Bjorksten, B., Bonini, S., Bousquet, J., Djukanovic, R., Dreborg, S., Gereda, J., Malling, H.J., Popov, T., Raz, E., Renz, H., Wold, A., 2003. Microbial products in allergy prevention and therapy. *Allergy* 58, 461–471.
- Mills, E., Madsen, C., Shewry, P., Wichers, H., 2003. Food allergens of plant origin - their molecular and evolutionary relationships. *Trends Food Sci. Technol.* 14, 145–156.
- Notkins, A.L., 2004. Polyreactivity of antibody molecules. *Trends Immunol.* 25, 174–179.
- Ostman, S., Rask, C., Wold, A.E., Hultkrantz, S., Teleme, E., 2006. Impaired regulatory T cell function in germ-free mice. *Eur. J. Immunol.* 36, 2336–2346.
- Quilot, B., Genard, M., Kervella, J., Lescourret, F., 2004. Analysis of genotypic variation in fruit flesh total sugar content via an ecophysiological model applied to peach. *Theor. Appl. Genet.* 109, 440–449.
- Reche, P.A., Zhang, H., Glutting, J.P., Reinherz, E.L., 2005. EPIMHC: a curated database of MHC-binding peptides for customized computational vaccinology. *Bioinformatics* 21, 2140–2141.
- Riaz, T., Hor, H.L., Krishnan, A., Tang, F., Li, K.B., 2005. WebAllergen: a web server for predicting allergenic proteins. *Bioinformatics* 21, 2570–2571.
- Ring, J., Kramer, U., Behrendt, H., 2004. A critical approach to the hygiene hypothesis. *Clin. Expe. Allergy Rev.* 4, 40–44.
- Saha, S., Raghava, G.P., 2006. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.* 34, W202–W209.
- Smits, H.H., Engering, A., van der Kleij, D., de Jong, E.C., Schipper, K., van Capel, T.M., Zaat, B.A., Yazdanbakhsh, M., Wierenga, E.A., van Kooyk, Y., Kapsenberg, M.L., 2005a. Selective probiotic bacteria induce IL-10-producing regulatory T cells in vitro by modulating dendritic cell function through dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin. *J. Allergy Clin. Immunol.* 115, 1260–1267.
- Smits, H.H., Hartgers, F.C., Yazdanbakhsh, M., 2005b. Helminth infections: protection from atopic disorders. *Curr. Allergy Asthma Rep.* 5, 42–50.
- Strachan, D., 1989. Hay fever, hygiene, and household size. *Br. Med. J.* 299, 1259–1260.
- Traskalova-Hogenova, H., Stepankova, R., Hudcovic, T., Tuckova, L., Cukrowska, B., Lodinova-Zadnikova, R., Kozakova, H., Rossmann, P., Bartova, J., Sokol, D., 2004. Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. *Immunol. Lett.* 93, 97–108.
- Vieths, S., Scheurer, S., Ballmer-Weber, B., 2002. Current understanding of cross-reactivity of food allergens and pollen. *Ann. N.Y. Acad. Sci.* 964, 47–68.
- Wang, Z., Chen, Z.J., Wheeler, J., Shen, S., Notkins, A.L., 2001. Characterization of murine polyreactive antigen-binding B cells: presentation of antigens to T cells. *Eur. J. Immunol.* 31, 1106–1114.
- Wold, A., 1998. The hygiene hypothesis revised: is the rising frequency of allergy due to changes in the intestinal flora? *Allergy* 53, 20–25.
- Xavier, R.J., Podolsky, D.K., 2000. Microbiology: how to get along-friendly microbes in a hostile world. *Science* 289, 1483–1484.
- Yazdanbakhsh, M., Kremsner, P.G., van Ree, R., 2002. Allergy, parasites, and the hygiene hypothesis. *Science* 296, 490–494.