

ORIGINAL ARTICLE

Basic and Translational Allergy Immunology

Diverse and highly cross-reactive T-cell responses in ragweed allergic patients independent of geographical region

Peter A. Würtzen¹  | Ilka Hoof¹ | Lars H. Christensen¹ | Zuzana Váczy² |
Helene Henmar¹ | Guillermo Salamanca¹ | Claus Lundegaard¹ | Lise Lund¹ |
Tatiana Kráľova³ | Edwards G. Brooks⁴ | Peter S. Andersen¹

¹Global Research, ALK, Hørsholm, Denmark

²Louis Pasteur University Hospital, Košice, Slovak Republic

³Immunology and Allergology, Košice, Slovak Republic

⁴UT Health San Antonio, San Antonio, TX, USA

Correspondence

Peter A. Würtzen, Global Research, ALK, Boege Alle 6-8, 2970, Hørsholm, Denmark.
Email: pawdk@alk.net

Funding information

ALK

Abstract

Background: Ragweed frequently causes seasonal allergies in North America and Europe. In the United States, several related ragweed species with diverse geographical distribution cause allergic symptoms. Cross-reactivity towards related ragweed species of IgE and treatment-induced IgG₄ has been demonstrated previously. However, less is known about the underlying T-cell cross-reactivity.

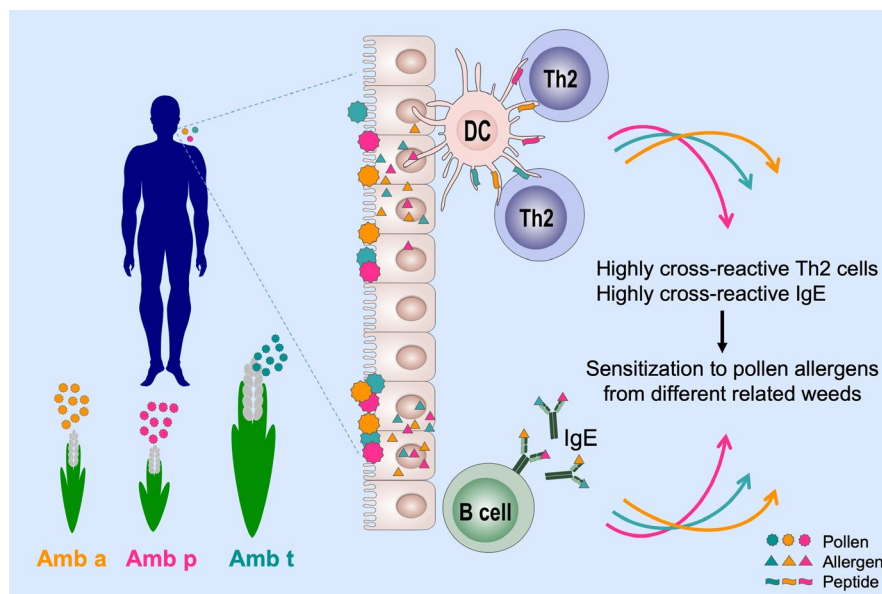
Methods: The allergen content of ragweed extracts was determined by mass spectrometry and related to T-cell epitopes of Amb a allergens (group 1, 3, 4, 5, 8 and 11) in 20 American ragweed allergic patients determined by FluoroSpot and proliferation assays. T-cell responses to 50 frequently recognized Amb a-derived T-cell epitopes and homologous peptides from western ragweed (Amb p), giant ragweed (Amb t) and mugwort (Art v) were investigated in an additional 11 American and 14 Slovakian ragweed allergic donors.

Results: Ragweed extracts contained all known allergens and isoallergens thereof. Donor T-cell responses were diverse and directed against all Amb a 1 isoallergens and to most minor allergens investigated. Similar response patterns were seen in American and Slovakia donors. Several epitopes were cross-reactive between isoallergens and ragweed species, some even including mugwort. T-cell cross-reactivity generally correlated with allergen sequence homology.

Conclusion: T-cell epitopes of multiple allergens/isoallergens are involved in the diverse T-cell responses in ragweed allergic individuals. T-cell lines were highly cross-reactive to epitopes of related ragweed species without any apparent geographical response bias. These data support that different ragweed species can be considered an allergen homology group with Amb a as the representative species regarding diagnosis as well as allergy immunotherapy.

KEYWORDS

cross-reactivity, homologous groups, immunotherapy, ragweed allergy, T-cell epitopes



GRAPHICAL ABSTRACT

T-cell epitopes of multiple major and minor ragweed allergens/isoallergens are involved in the diverse T-cell responses in ragweed allergic individuals. The T-cells cross-react to epitopes of related ragweed species and the same epitopes are important for American and European individuals. Extensive T-cell and IgE cross-reactivity support that related ragweed species can be considered an Amb a allergen homology group regarding diagnosis and immunotherapy treatment.

1 | INTRODUCTION

Ragweed allergy represents a major health problem in North America as well as in central Europe, and the prevalence is increasing worldwide.¹ The most widely distributed ragweed species in North America are short ragweed (*Ambrosia artemisiifolia*; Amb a), giant ragweed (*Ambrosia trifida*; Amb t) and western ragweed (*Ambrosia psilostachya*; Amb p), whereas Amb a dominates in Europe.² Knowledge of the allergens involved and patterns of immunological cross-reactivity towards pollen of different weed species is crucial for diagnostics and for selection and formulation of allergy immunotherapy (AIT) vaccines.³ Amb a 1 is the key major allergen in ragweed with IgE prevalence up to 95% reported in some studies. Five isoallergens have been described for Amb a 1 with sequence identity from 62% to 87%.⁴ Nine other Amb a allergens have been identified so far with IgE prevalence in the 10%-30% range for Amb a group 5, 6, 7, 9 and 10 allergens and around 40%-60% for group 3, 4, 8⁵ and 11.⁶ The presence of homologous allergens in different ragweed species leads to extensive IgE cross-reactivity between Amb a, p and t pollen extracts. This was demonstrated in US ragweed allergic patients participating in a trial addressing the clinical effect of ragweed SLIT-tablet AIT,⁷ and the IgE cross-reactivity was even extended to seven less widespread ragweed species. Further, sublingual AIT with a single ragweed species (Amb a) induced IgG₄ antibodies fully cross-reactive with Amb p and Amb t extracts.⁷ Regarding T cells, more than 20 Amb a 1 T-cell epitopes including 3 "dominant" epitope areas has been reported^{8,9} including examples of T-cell cross-reactivity between Amb a 1 isoforms.¹⁰ However, little is known about T-cell cross-reactivity between different species of ragweed.

It appears reasonable to assume that recognition of specific T- and B-cell epitopes plays a key role in mediating clinical effect of AIT. It is generally accepted that AIT performed with allergen extracts causes changes to the allergen-specific immune response involving induction of blocking non-IgE antibodies as well as shifting the balance between Th1/Th2 and regulatory B and T cells.¹¹ Regarding clinical cross-reactivity or cross-protection, allergen extract from a single grass species (*Phleum pratense*),¹²⁻¹⁴ a single ragweed species (*Ambrosia artemisiifolia*)¹⁵⁻¹⁹ or a single tree species (*Betula verrucosa*)²⁰ administered sublingually is clinically effective in areas where patients are exposed to multiple grass species, where patients are exposed to several related ragweed species or when patients are exposed to pollen of either birch or white oak, respectively. In contrast, cross-protection is not observed between unrelated allergen sources such as ragweed/grass²¹ or grass/birch.²² Thus, within closely related species, known as homologous allergen groups,²³ immunological cross-reactivity is an important factor in causing not only allergic symptoms but also in securing clinical benefit of AIT.

The aim of the current study was to investigate to what extent T cells in individual ragweed allergic patients from Texas, USA (US), and Slovakia (SK) are cross-reactive towards homologous allergens from related weed species and to correlate T-cell responses to IgE sensitization and allergen sequence homology.

2 | METHODS

Essential information is included in this section, and further details can be found in the online repository Appendix S1.

TABLE 1 Patient gender, age, symptoms and sensitizations. Serum IgE was analysed by ImmunoCAP, and sensitizations to mite (Der p), cat (Fel d), grass (Phl p/Cyn d), birch (Bet v) and mountain cedar (Jun s) are indicated in percentages of the individual populations. Polysensitized indicates percentage of patients with more than one sensitization

	Avg age	Gender	Asthma (+/-)	Amb a IgE (kU/L)	Amb a 1 IgE (kU/L)	Der p	Feld	Phl p/Cyn d	Bet v	Jun s	Poly sensitized
US (n = 31)	38.5	14f/17m	9+/22-	9.32	9.52	58%	39%	77%	58%	84%	100%
SK (n = 14)	37.9	7f/7m	6+/8-	12.40	7.75	29%	36%	79%	71%	36%	86%
All (n = 45)	38.3	21f/24m	15+/30-	10.27	8.97	49%	38%	78%	62%	67%	96%

2.1 | Donors

20 US donors were used for the initial epitope mapping, and 11 US and 14 SK donors were used for investigation of responses to selected T-cell epitopes and homologous peptides. Conclusive T-cell responses were obtained for 31 US and 14 SK donors (Amb a IgE > 0.6 kU/L, ImmunoCAP), and the data were used to investigate linkage between the different immune parameters investigated. See Table 1 for additional patient characteristics.

2.2 | Allergen extracts, allergens and peptides

SDS-PAGE gels of allergen extracts, and recombinant and purified allergens are seen in Figure S1, online.

Amb a extract (ALK) was gel-filtered by size-exclusion chromatography before use for establishment of T-cell lines, and additional experimental extracts of Amb p, Amb t and Artemisia vulgaris (Art v) were used for T-cell activation assays.

Endotoxins were removed from all recombinant proteins and measured to <0.01 EU/μg. Endotoxin content of all purified allergens was below 20 EU/mg except nAmb a8, which was 53 EU/mg.

Overlapping 20 mer peptides, selected peptides 20-23 mers, homologous 20-23 mers and full-length Amb a5 were custom-made by Genscript (purity ≥ 95%).

2.3 | Allergome profiling of Amb a extract

Amb a extract was denatured, reduced, alkylated and trypsin-digested before analysis by LC-MS/MS on a ESI-QTOF mass spectrometer.

2.4 | Data analysis

MS/MS spectra were searched (via MASCOT search engine, Matrix Science) against all available protein sequences of IUIS Amb a allergens and isoforms thereof. The sum of the relative and semi-quantitative (emPAI) scores²⁴ of all Amb a allergens was set to 100%, and the relative percentage (molar %) of each allergen was calculated.

2.5 | Generation of T-cell lines, FluoroSpot, proliferation and cytokine production

Amb a-specific T cells from cryopreserved PBMC were expanded in vitro, and specific responses were measured by FluoroSPOT²⁵ at day 14 and by thymidine incorporation²⁵ and cytokine production²⁵ at day 24 or later. In these assays, the T-cell lines were stimulated with the different peptides (2 μg/mL), full-length allergens (2 μg/mL), allergen extracts (5-10 μg/mL) and controls (PHA or medium alone).

2.6 | Quantitative assessment of T-cell epitope cross-reactivity

The level of cross-reactivity was evaluated on the basis of T-cell proliferation data, and all T-cell lines with a response to the original Amb

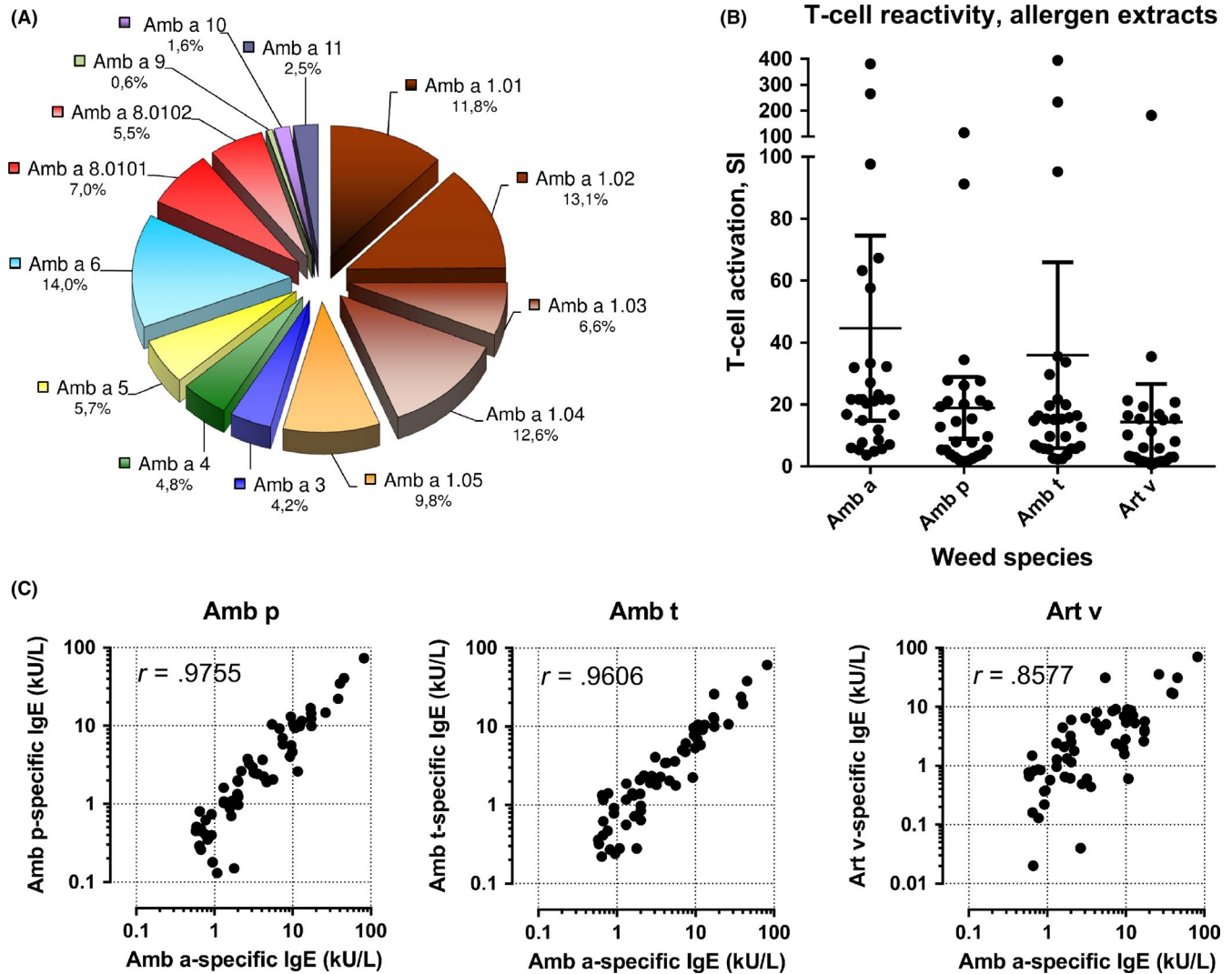


FIGURE 1 T-cell responses and serum IgE concentrations to weed extracts. Relative content (molar %) of individual Amb a allergens including Amb a 1 and Amb a 8 isoforms in Amb a extract as determined by mass spectrometry (A), T-cell proliferation (SI = cpm stim/cpm medium control, mean $\pm$ 95% conf.) to weed pollen allergen extracts detected by thymidine incorporation (B), and correlation of serum IgE concentrations towards weed pollen allergen extracts (C)

a peptide were included in the analysis. T-cell proliferation of similar strength (SI) should be obtained for a number of T-cell lines, and Pearson's correlation coefficient was used to quantify cross-reactivity. T-cell responses against a given homologous peptide pair were defined as cross-reactive if the correlation was significant ($P < .05$) and the correlation coefficient was above .5.

2.7 | HLA typing and predicted population coverage

Genotyping of the HLA class II locus to 4-digit resolution was performed by ProImmune Ltd using a PCR-based assay. HLA-peptide binding predictions were performed using NetMHCIIpan 3.0.^{26,27}

2.8 | Transcriptomes and homology

RNA was isolated from Amb p and Amb t pollen (ALK). Library preparation and RNA-seq (Illumina HiSeq 2500) were performed

by Eurofins Genomics. After sequences trimming identification of homologous sequences, the resulting transcriptomes were used to identify Amb a homologue allergens in Amb p and Amb t extracts through LC-MS/MS analyses.

All allergen sequences used for selection of homologous peptides are listed in Figure S2, online. Table S1 (online) includes the level of identity for the ragweed group 1 allergens and Art v 6.²⁸ Art v homologues for Amb a 3, 4 and 8 available from previous publications²⁸ were also used for generation of homologous peptides (Figure S2, online).

3 | RESULTS

3.1 | T-cell response and IgE reactivity to ragweed allergen extracts and Amb a 1 isoallergens

The relative allergen composition of the Amb a extract determined using mass spectrometry (Figure 1A) showed that the five Amb a

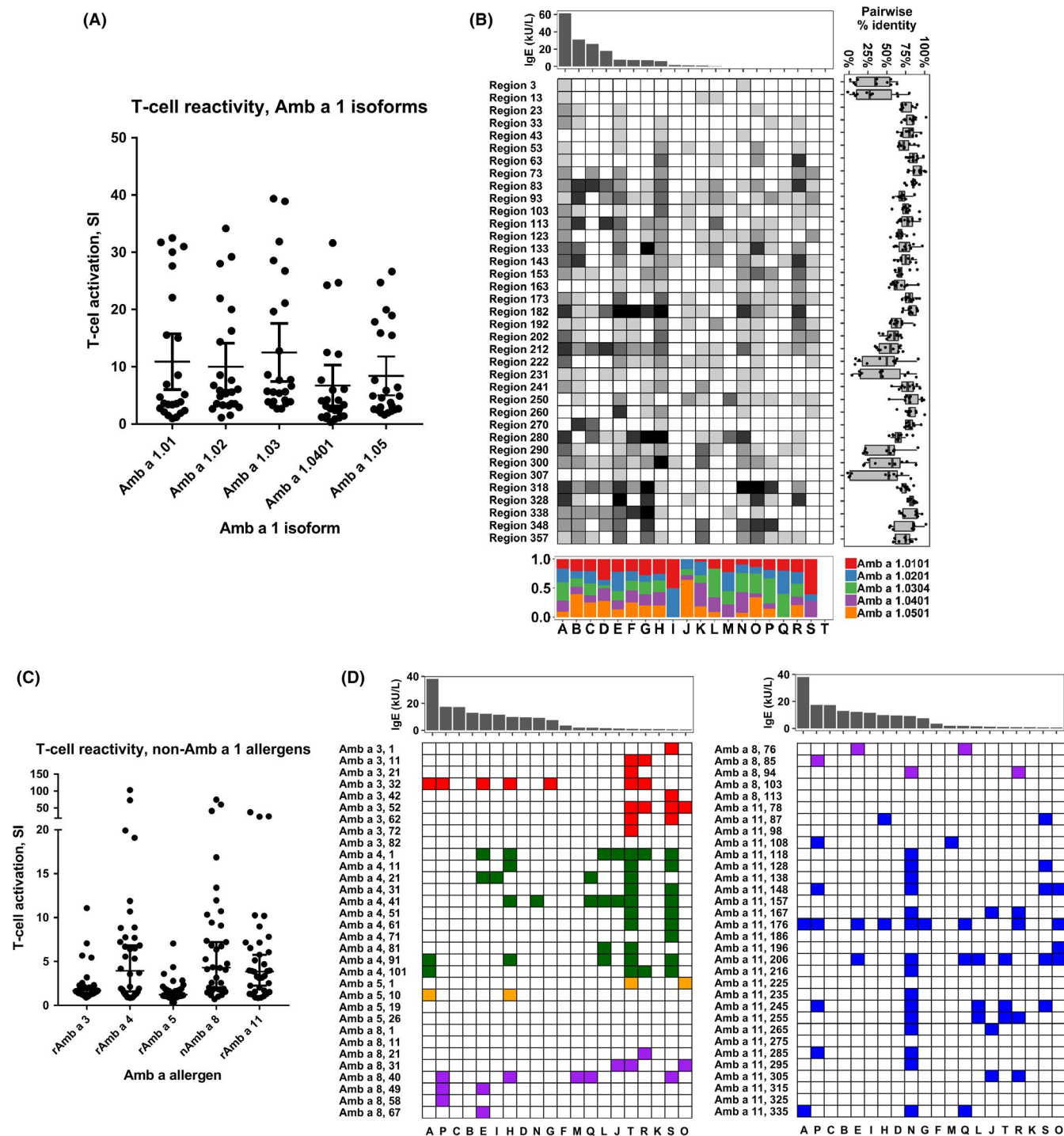


FIGURE 2 (A,B) T-cell proliferation (SI = cpm stim/cpm medium control, median $\pm$ 95% conf.) induced by purified or recombinant Amb a 1 isoallergens (A). T-cell responses (SI > 3) to individual peptides from the five Amb a 1 isoallergens (B, central panel) and relative responses to each isoallergen based on the number of peptides recognized (B, bottom panel). Serum IgE ImmunoCAP values (B, top panel) and sequence identity between isoallergens for the epitope regions (B, right panel). The colour intensity (B, central panel) indicates the number of homologous peptides recognized (from 1 to 5). Patients are ordered on the basis of serum IgE concentration towards Amb a 1 (B). (C,D) T-cell proliferation (SI = cpm stim/cpm medium control, mean $\pm$ 95% conf.) induced by purified or recombinant Amb a allergens (C). T-cell responses (SI > 3) to individual peptides from the Amb a 3, Amb a 4, Amb a 5, Amb a 8 and Amb a 11 isoallergens and isoforms (D, central panel). Patients are ordered on the basis of serum IgE ImmunoCAP values (D, top panels). Patients are ordered on the basis of serum IgE concentration towards Amb a extract (D)

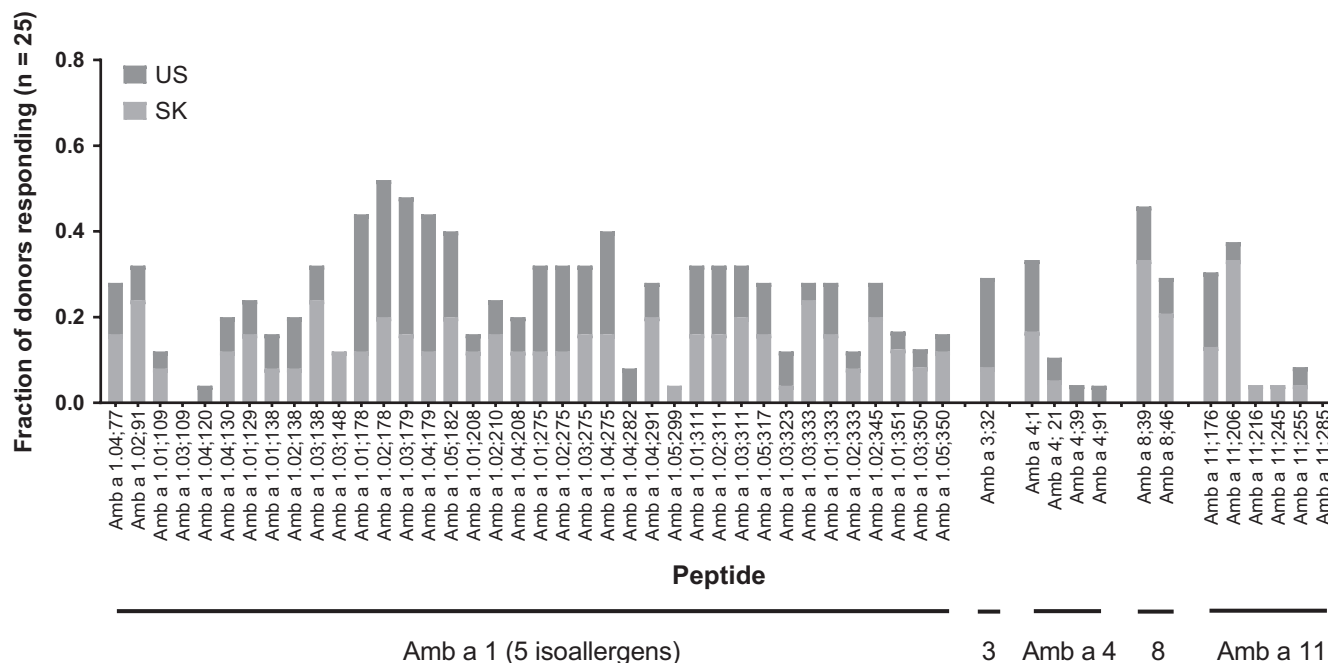


FIGURE 3 Frequency of T-cell responses (SI > 3) towards selected peptides from Amb a allergens in new donor (n = 25). Columns are split into responses from US (n = 11, dark grey) and SK (n = 14, light grey) donors for each peptide. Dashed lines indicate homologous peptides from different Amb a 1 isoforms

1 isoallergens were all present in the extract adding up to around 50% of the total allergen content. Minor Amb a allergens of known sequence were also detected in relative quantities ranging from 0.6% (Amb a 9) to 14% (Amb a 6) (Figure 1A). T-cell lines generated using Amb a extract responded equally to extracts from two other ragweed species Amb t and Amb p and to a lesser extent to extract from the more distantly related species Art v (Figure 1B). Similarly, most T-cell lines responded strongly to all five Amb a 1 isoallergens (Figure 2A). T-cell responses were dominated by IL-5 production, whereas IFN- γ production and IL-10 production were seen only in a few cell lines indicative of a strong Th2 polarization (data not shown). The IgE reactivity towards the individual weed pollen allergen extracts was highly correlated (Figure 1C) especially for Amb a/p and Amb a/t, whereas a more scattered pattern was seen for Amb a/Art v.

3.2 | T-cell epitope mapping and reactivity towards Amb a 1 isoallergens

T-cell lines were stimulated with overlapping 20-mer peptides covering the entire amino acid sequence of all the five known isoallergens of Amb a 1 (Figure 2B, central panel) to further understand the reactivity towards Amb a 1. T-cell epitopes were found throughout the individual isoallergen sequences with the exception of the N-terminal region (approx. aa 1-70) and a central part (approx. aa 220-270), which mainly contained few discrete epitopes. A part of these regions were also found to have less sequence homology between the isoallergens (Figure 2B, right panel), but enzymatic digestion in the endosomes and HLA binding affinity may also affect the availability of these epitope regions. Some homologous

peptides were recognized frequently in all isoallergens, some in a few isoallergens, and some seemed to be isoallergen-specific. A group of donors recognized multiple epitopes throughout the Amb a 1 sequences, and other responded to relatively few or even no (donor T) Amb a 1 T-cell epitopes. T-cell reactivity to multiple peptides seemed to be more likely in patients with high Amb a 1 IgE titres (Figure 2B, upper panel), even though there were exceptions (eg, donor N). The majority of donors responded to most Amb a 1 isoallergens (Figure 2B, lower panel) indicating no obvious immunodominance of one particular Amb a 1 isoallergen.

3.3 | T-cell epitope mapping of Amb a minor allergens

The role of the Amb a minor allergens in mediating allergic T-cell responses was also investigated using full-length proteins (Figure 2C) and overlapping peptides covering the sequence of Amb a 3, 4, 5, 8 and 11 (Figure 2D). The responses to minor Amb a allergens varied and were found throughout their aa sequences but with lower frequency than Amb a 1 epitopes (Figure 2). Four (Amb a 3), 10 (Amb a 4), 2 (Amb a 5), 5 (Amb a 8) and 13 (Amb a 11) peptides were recognized by more than 2 donors (10%), and between zero (Amb a 5) and three (Amb a 4 and 11) of the peptides contained epitopes recognized by 25% or more of these US donors (Figure 2D). Patients with low Amb a IgE titres (Figure 2B, upper panel) were able to respond to a variety of Amb a 3, 4, 8 and 11 T-cell epitopes indicating no association between Amb a-specific IgE levels and the number of minor allergen T-cell epitopes. Further, donor T, without T-cell reactivity to Amb a 1 peptides, responded to multiple minor allergen epitopes.

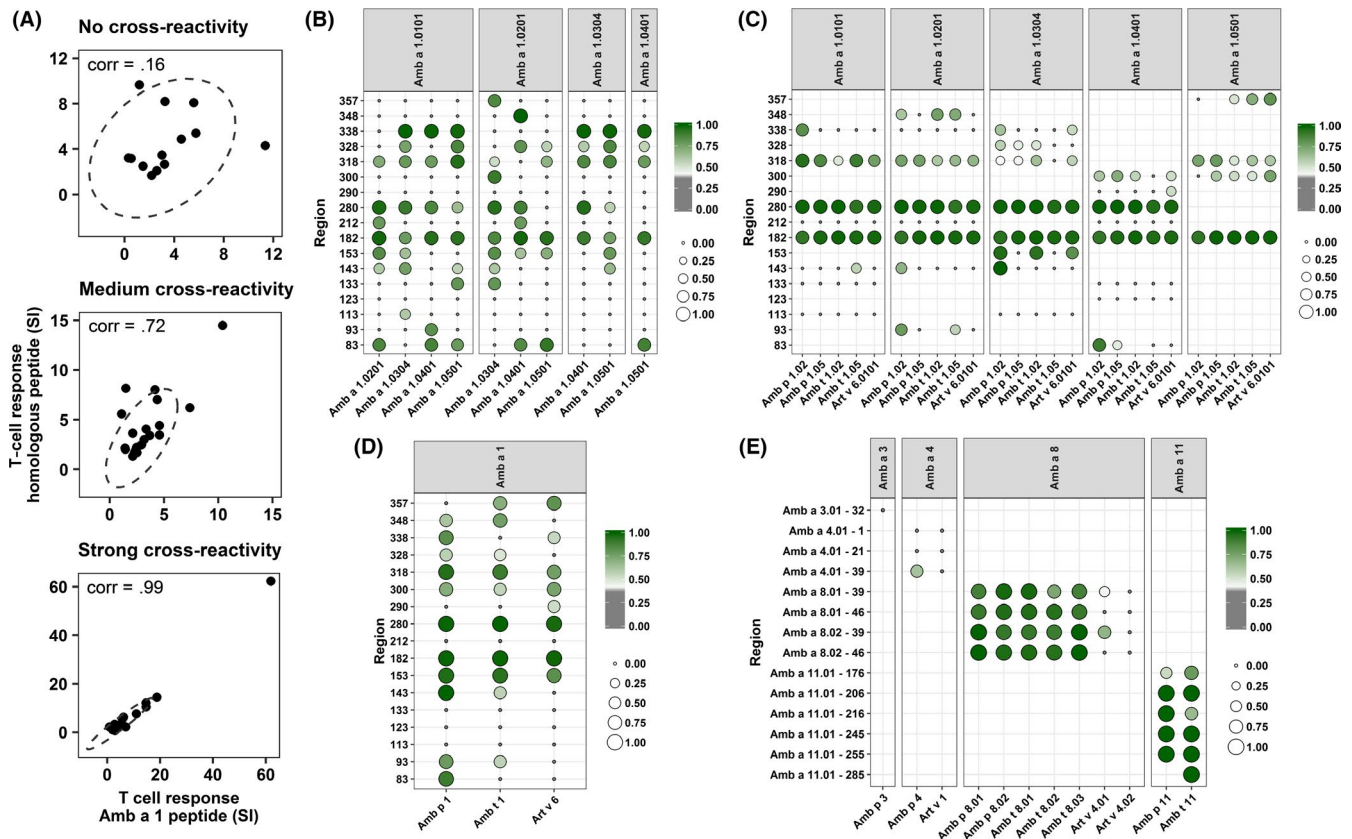


FIGURE 4 Correlations were investigated for homologous sequence regions by plotting each peptide pair, and the Pearson correlation coefficients and significance levels were determined as described in Methods sections (examples in A). Significant correlations are depicted in B–E by increasing size and darkness of green with increasing correlation coefficient. Correlations within Amb a 1 isoforms (B), correlations between group 1/6 allergens from different weed species for each Amb a 1 isoallergens (C), correlations between group 1/6 allergens irrespective of isoallergen (D) and correlations between non-Amb a 1 allergens (and isoallergens/isoforms) between weeds (E)

3.4 | Geographical variation in ragweed T-cell epitopes

To replicate the findings from the first US donor cohort, 50 immunodominant peptides (Figure 2) were selected from Amb a 1, 3, 4, 8 and 11 and investigated in a second cohort consisting of 11 novel US donors and 14 SK donors. T-cell activation was observed towards all but two peptides (Figure 3), 25 peptides were recognized in more than 25% of the donors, and peptides from all five allergens were represented among these. In addition, both US and SK donors contributed to the response frequencies for most peptides.

3.5 | Quantitative comparison of intra-species and inter-species T-cell epitope cross-reactivity

The cross-reactivity pattern of Amb a 1 epitopes relative to the homologous epitopes of the other isoallergens of Amb a 1 was depicted (Figure 4B) using single numerical values as a measure of cross-reactivity (Figure 4A). This revealed that some epitope regions (182, 280 and 318) were recognized throughout the five isoforms. For one epitope region (328), T-cell cross-reactivity was more scattered among individual isoallergens, suggesting that particular

amino acid substitutions affected T-cell activation negatively. Other epitope regions showed T-cell cross-reactivity restricted to a few isoallergens or individual isoallergen pairs and only two of the investigated epitope regions (123, 290) were not involved in T-cell cross-reactivity.

Extending this analysis into homologous regions of allergens from Amb t, Amb p and Art v (Figure 4D) revealed that in some cases, the observed cross-reactivity towards Amb a 1 isoallergens (intra-species) extended further into related weed species (inter-species). Three epitope regions with consistent intra-species T-cell cross-reactivity (region 182, 280 and 318) were also recognized throughout all 4 weed species with significant correlation between the T-cell responses to almost all investigated peptide pairs. Interestingly, several regions with restricted or no cross-reactivity among Amb a 1 isoallergens showed cross-reactivity to one (83, 290), two (93, 143, 348 and 357) or all weed species investigated (153, 300). When differences between the isoallergens of Amb a 1/p 1/t 1/Art v 6 were disregarded to evaluate the overall T-cell cross-reactivity (Figure 4D), the response to Amb p 1 showed the highest cross-reactivity with 5 strong and 6 medium/low cross-reactive peptides. The T-cell cross-reactivity to Amp t 1 was slightly less pronounced with 3 highly cross-reactive and seven medium/low epitopes, and for Art v 6, strong cross-reactivity was

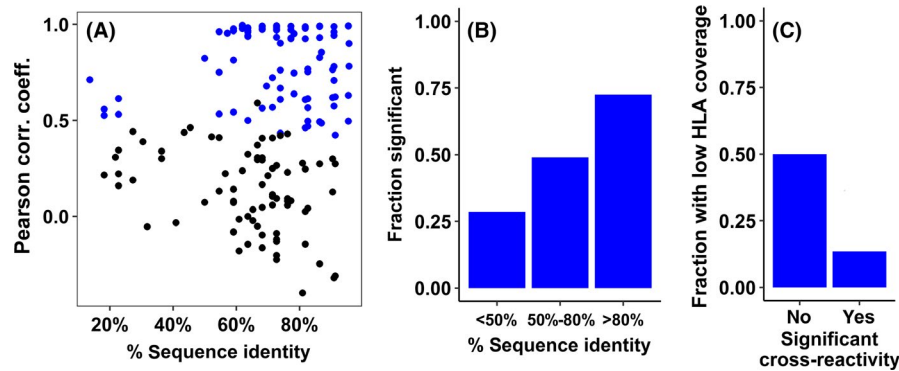


FIGURE 5 Correlation coefficients for each peptide pair are plotted in A according to the sequence identity between the peptide pair (significant correlations in blue). The same data are presented in B with sequence identities grouped <50%, 50%-80% and >80% to illustrate the increase in the fraction of significant correlations found for peptide pairs within the increasing sequence identity cut-offs. In C, the fraction of peptide pairs with low HLA coverage (low promiscuity) is depicted for the >80% sequence identity subset of peptide pairs

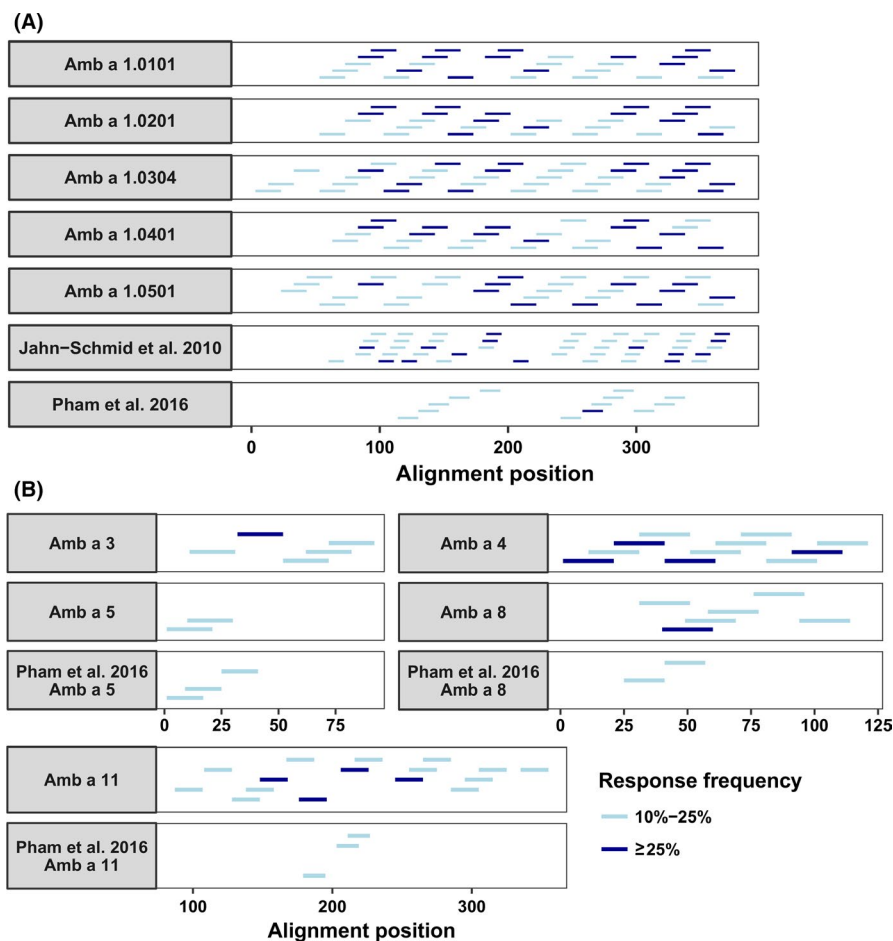


FIGURE 6 Overview of T-cell epitopes in Amb a 1 isoallergens 0.01-0.05 with previously published data on T-cell epitopes in Amb a 1.0304 (Jahn-Schmid et al) and Amb a 1.0101 (Pham et al) included in the lower panels (A). A similar overview of T-cell epitopes in Amb a 3, Amb a 4, Amb a 5, Amb a 8 and Amb a 11 with previously published data on T-cell epitopes in Amb a 5, Amb a 8 and Amb a 11 (Pham et al) included in the lower panels is seen in B. All data are depicted with T-cell epitopes recognized by 25% or more of the donors investigated indicated in dark blue and by 10%-25% indicated in light blue

observed for two peptides and medium/low cross-reactivity for seven peptides.

A similar analysis of the epitope regions of the minor allergens (Figure 4E) demonstrated that most Amb a 8 and Amb a 11 peptides contributed to the T-cell cross-reactivity towards Amb p and Amb t. Cross-reactivity towards Art v 4 was also observed for Amb a 8 epitope region 39, whereas cross-reactivity for Amb a 3 and 4 peptides was limited to one Amb a 4 epitope region. Cross-reactivity was most pronounced for group 11 ragweed allergens followed by group 8/Art

v 4 and group 4/Art v 1, which likely reflects the decreasing level of sequence conservation within these groups in combination with the lower number of donors responding.

3.6 | Correlation of epitope sequence homology and T-cell cross-reactivity

To investigate whether there was a link between the observed cross-reactivity and sequence identity between pairs of homologous

peptides, Spearman's correlation coefficients were plotted as a function of percentage sequence identity (Figure 5A). Significant correlations are clustered in the upper right corner with high correlation coefficients $> .5$ (blue dots) and high sequence identity ($>50\%$). The association between the two parameters is highlighted in Figure 5B where peptide pairs with $<50\%$, $50\%-80\%$ and $>80\%$ sequence identity show increasing cross-reactivity with increasing sequence identity. Sequence identity above 80% does not necessarily lead to cross-reactivity, and this may be related to the HLA binding capacity of the homologous peptides (Figure 5C), whereas virtually all peptides with significant T-cell cross-reactivity do have a broad predicted HLA coverage.

4 | DISCUSSION

Immunological cross-reactivity is an important factor in the diagnosis and treatment of allergic disease. Here, we present an in-depth analysis of the cross-reactivity of T-cell responses in ragweed allergic donors and associate these responses to IgE titres, allergen homology between weed species and geographical origin of the donors. The MS data presented here (Figure 1A) demonstrate that all Amb a allergens of known sequence are present in the aqueous pollen extracts in relative amounts ranging from 0.6% to 14% of total allergen content and the sum of Amb a 1 isoallergens constituted around 50% .

T-cell responses appeared to mirror the composition in the extract because all the investigated allergens (Figure S1) contained T-cell epitopes (Figure 2). Early investigations of allergen-specific T-cell reactivity included HLA²⁹ and V β ³⁰ restriction of responses to Amb a 5 (antigen Ra5), and during the 1990s, different isoallergens of Amb a 1 (antigen E) were shown to induce comparable T-cell responses in ragweed allergic patients.³¹ Recent reports include more detailed analysis of additional allergens and Amb a 1 isoallergens. For Amb a 1 (Amb a 1.0304), more than 20 T-cell epitopes including 3 "dominant" epitopes have been described⁸ (Figure 6A). The dominant epitopes were confirmed in a later study⁹ (Figure 6A), and both studies report responses dominated by Th2 cytokines in line with the current investigation. Most of the previously identified major and minor allergen epitope regions were replicated in the current study, which is seen in Figure 6 depicting the peptide recognition frequencies for all 31 US and 14 SK donors analysed. In addition, we identified several novel T-cell epitopes mainly in Amb a 3, 4, 8 and 11. T-cell reactivity to multiple Amb a 1 epitopes seemed to be more likely in patients with high Amb a 1 IgE (>5 kU/L) in line with the cut-off for strong T-cell responses towards mite allergen epitopes (>10 kU/L) suggested by Hinz et al³². The micro-heterogeneity of the T-cell responses in individual donors may explain why various attempts to develop next-generation AIT vaccines based on a single recombinant Amb a 1 isoform³³ or short peptides derived from such a molecule^{34,35} have failed to progress beyond phase II clinical trials. It may also predict similar challenges for the suggested hybrid molecule combining parts of Amb a 1 and Art v 6.³⁶ Moreover, one

donor responded only to minor allergen epitopes and other donors responded to multiple epitopes in the minor allergens, indicating that T-cell responses to minor allergens can be important factors in allergic disease. This is in line with the observation that the minor grass allergen, profilin (pHl p 12), can lead to T-cell responses comparable to major grass allergens in a subset of profilin IgE-sensitized grass allergic patients.²⁵ Regarding the major allergen, Wolf et al¹⁰ suggested that Amb a 1.01 and 1.03 are immunodominant isoforms based on data from 5 Amb a 1-specific T-cell clones and PBMC from seven ragweed allergic donors. Our results based on T-cell lines from 45 ragweed allergic individuals do not support this hypothesis because we observed diverse T-cell responses to all Amb a 1 isoforms; many homologous peptides from different isoforms induced similar T-cell responses in a substantial number of patients (Figure 2; lower panel).

The extensive epitope mapping of Amb a allergens made it possible to select 50 frequently recognized peptides from Amb a 1, 3, 4, 8 and 11 for further investigation. When responses of a second patient cohort of 11 US and 14 Slovakian donors were analysed, we observed T-cell responses towards all selected peptides except two. Twenty-five of the 50 selected peptides were recognized by $>25\%$ of the donors confirming the immunological relevance of the selected peptides (Figure 3). The highly diverse T-cell response towards various Amb a allergens in both US and SK cohorts expressing various HLA-DRB1 alleles (Figure S3) indicates that the geographical differences in exposure to pollens from individual ragweed species do not introduce any obvious bias in the T-cell response.

To address T-cell cross-reactivity to related ragweed species, Amb p and Amb t transcriptomes were searched to identify sequences homologous to Amb a group 1 3, 4, 8 and 11 allergens. The sequences of these homologues allergens from Amb p, Amb t and Art v (published sequences) were used to identify peptide homologues for all the 50 selected Amb a T-cell epitope-containing peptides, adding up to almost 200 peptides with up to six homologues for each Amb a peptide.

To allow for a combined analysis of such huge amounts of cross-reactivity data, we developed a novel analytical method for quantitative assessment of epitope cross-reactivity. When T-cell lines are stimulated in separate experiments with homologous pairs of peptides, they will respond with similar signal strength if there is full cross-reactivity. For each homologous peptide pair, this allows for a linear fit of the response data (see Figure 4A) and the Pearson correlation coefficient will indicate how similar the responses are (Figure 4B-E). From these plots, it was evident that several epitopes mediated T-cell cross-reactivity to a broad range of Amb a 1 isoallergens and weed species, whereas a more scattered cross-reactivity pattern was observed for other epitopes. Interestingly, epitopes inducing T-cell cross-reactivity between weed species were found in all investigated Amb a 1 isoallergens and minor allergens, indicating that the clinically observed cross-reactivity or cross-protection^{7,15} is, in part, mediated by both major and minor allergens and isoallergens thereof.

Having a single quantitative measure of T-cell cross-reactivity allowed for evaluation of the general correlation between epitope sequence homology and T-cell cross-reactivity (Figure 5). The correlation observed supports that allergen sequence homology between weed species is a key factor. However, epitopes that were predicted to bind relative few HLA alleles (low population coverage) were less cross-reactive even though they were more than 80% homologous between species. This supports the "multiple hit" hypothesis formulated by Sette and co-workers, who proposed that epitope conservation in combination with HLA binding is the underlying mechanisms selecting the cross-reactive T-cell repertoire in allergic patients.⁹ Interestingly, the combined analysis of cross-reactivity also indicated that there is approximately a 50% chance of cross-reactivity for epitopes having <50% sequence identity. This allows for T-cell receptor recognition to play a key role, even among relatively distantly related species known to cause clinically relevant allergic reactions. T-cell cross-reactivity between Amb a and the more distantly related Art v was investigated in detail by Jahn-Schmid et al³⁷ in Austrian ragweed allergic patients where Amb a 1 was concluded to contain a broader spectrum of T-cell epitopes than its mugwort homologue Art v 6 in line with the current data (Figure 4D). In addition, very limited T-cell cross-reactivity was observed for Amb a 4/Art v 1 and Amb a 8/Art v 4 (Figure 4E), suggesting that only a subset of Art v-specific T cells will be affected by challenge or treatment with Amb a allergen extract.

We believe the present study has clinical implications both for diagnosis and for allergen immunotherapy of ragweed allergic patients. The data indicate that in ragweed allergic patients, the T-cell-mediated late-phase response upon allergen challenge cannot discriminate between different ragweed species. With respect to AIT, the IgE cross-reactivity demonstrated previously and the T-cell cross-reactivity observed in the current study fully support that both antibody and T-cell responses to Amb a, p and t will be modified when ragweed allergic patients are treated with Amb a extract AIT as observed in the recent clinical trials with ragweed SLIT tablets.^{15,17-19,38} Whether this is also the case for more distantly related weed species such as mugwort has to be addressed in dedicated clinical trials because the IgE^{37,39} cross-reactivity and T-cell (Figure 6)³⁷ cross-reactivity demonstrate substantial variation among individual ragweed- and/or mugwort-sensitized patients.

In summary, we here present evidence that T-cell responses in ragweed allergic patients are heterogeneous with no apparent influence of geographical location. Both major and minor allergens appeared to be important factors in overall immunogenicity and in mediating cross-reactivity to other ragweed species. These data support that different ragweed species can be considered an allergen homology group with Amb a as the representative species regarding diagnosis as well as AIT.²³

ACKNOWLEDGMENTS

The experiments were funded by ALK. In this context, the authors would like to thank the staff of translational Research, Biochemistry

and BioAnalytics for their excellent technical assistance. In addition, the authors would like to thank Patty Pompa-Aguilar and Joseph Diaz, MD, of Allergy and Asthma Research Center, PA, San Antonio, Texas, for recruitment of US ragweed allergic donors.

CONFLICT OF INTEREST

IH, LHC, HH, GS, CL, LL, PSA and PAW were all employed by ALK (a Pharma company producing allergy vaccines) at the time of this investigation. ZV, TK and EGB received research grants from ALK related to this investigation.

ORCID

Peter A. Würtzen  <https://orcid.org/0000-0003-1551-2415>

REFERENCES

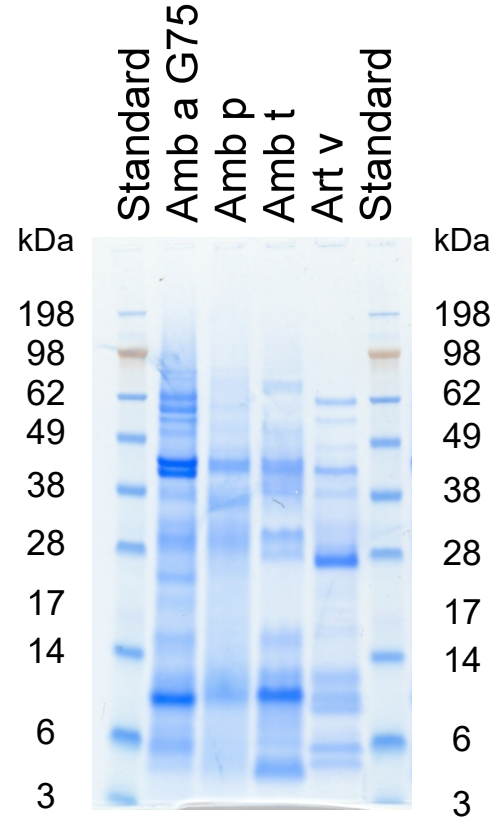
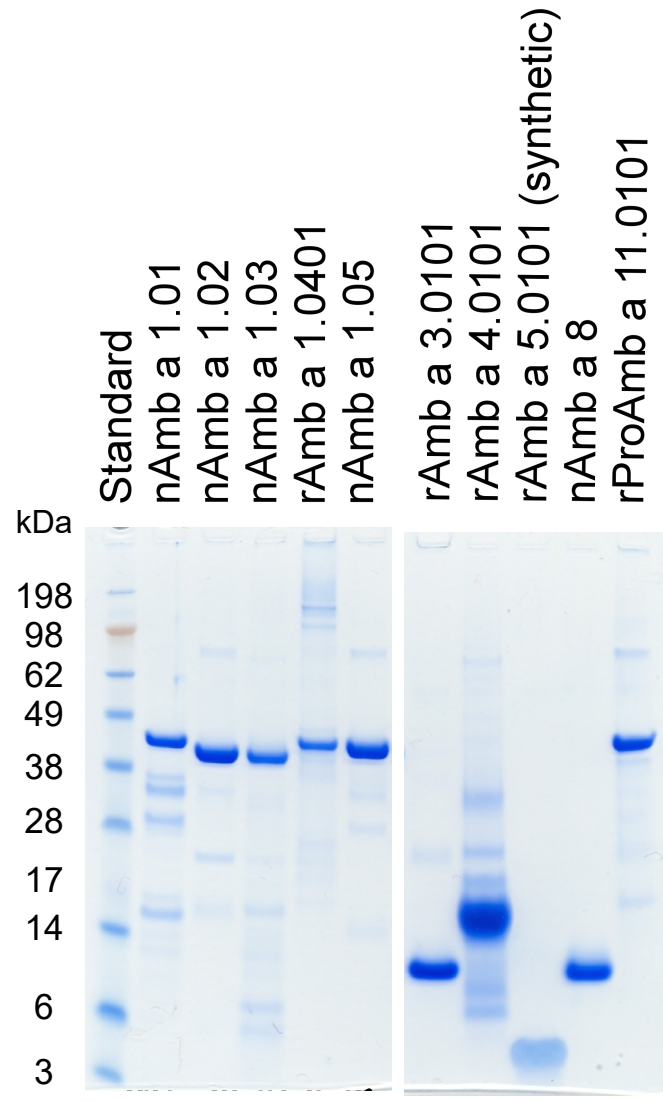
- Chen KW, Marusciac L, Tamas PT, Valenta R, Panaitescu C. Ragweed pollen allergy: burden, characteristics, and management of an imported allergen source in Europe. *Int Arch Allergy Immunol*. 2018;176(3-4):163-180.
- White JF, Bernstein DI. Key pollen allergens in North America. *Ann Allergy Asthma Immunol*. 2003;91(5):425-435.
- Weber RW. Patterns of pollen cross-allergenicity. *J Allergy Clin Immunol*. 2003;112(2):229-239.
- Radauer C, Nandy A, Ferreira F, et al. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. *Allergy* 2014;69(4):413-419.
- Gadermaier G, Hauser M, Ferreira F. Allergens of weed pollen: an overview on recombinant and natural molecules. *Methods*. 2014;66(1):55-66.
- Bouley J, Groeme R, Le Mignon M, et al. Identification of the cysteine protease Amb a 11 as a novel major allergen from short ragweed. *J Allergy Clin Immunol*. 2015;136(4):1055-1064.
- Christensen LH, Ipsen H, Nolte H, et al. Short ragweeds is highly cross-reactive with other ragweeds. *Ann Allergy Asthma Immunol*. 2015;115(6):490-495.
- Jahn-Schmid B, Wopfner N, Hubinger G, et al. The T-cell response to Amb a 1 is characterized by 3 dominant epitopes and multiple MHC restriction elements. *J Allergy Clin Immunol*. 2010;126(5):1068-1071.
- Pham J, Oseroff C, Hinz D, et al. Sequence conservation predicts T cell reactivity against ragweed allergens. *Clin Exp Allergy*. 2016;46(9):1194-1205.
- Wolf M, Twaroch TE, Huber S, et al. Amb a 1 isoforms: unequal siblings with distinct immunological features. *Allergy*. 2017;72(12):1874-1882.
- Shamji MH, Durham SR. Mechanisms of allergen immunotherapy for inhaled allergens and predictive biomarkers. *J Allergy Clin Immunol*. 2017;140(6):1485-1498.
- Dahl R, Kapp A, Colombo G, et al. Sublingual grass allergen tablet immunotherapy provides sustained clinical benefit with progressive immunologic changes over 2 years. *J Allergy Clin Immunol*. 2008;121(2):512-518.
- Durham SR, Emminger W, Kapp A, et al. SQ-standardized sublingual grass immunotherapy: confirmation of disease modification 2 years after 3 years of treatment in a randomized trial. *J Allergy Clin Immunol*. 2012;129(3):717-725.
- Durham SR, Yang WH, Pedersen MR, Johansen N, Rak S. Sublingual immunotherapy with once-daily grass allergen tablets: a randomized

- controlled trial in seasonal allergic rhinoconjunctivitis. *J Allergy Clin Immunol*. 2006;117(4):802-809.
15. Creticos PS, Maloney J, Bernstein DI, et al. Randomized controlled trial of a ragweed allergy immunotherapy tablet in North American and European adults. *J Allergy Clin Immunol*. 2013;131(5):1342-1349.
 16. Nolte H, Amar N, Bernstein DI, et al. Safety and tolerability of a short ragweed sublingual immunotherapy tablet. *Ann Allergy Asthma Immunol*. 2014;113(1):93-100.
 17. Creticos PS, Pfaar O. Ragweed sublingual tablet immunotherapy: part I - evidence-based clinical efficacy and safety. *Immunotherapy*. 2018;10(7):605-616.
 18. Pfaar O, Creticos PS. Ragweed sublingual tablet immunotherapy: part II - practical considerations and pertinent issues. *Immunotherapy*. 2018;10(7):617-626.
 19. Nelson HS. Ragweed allergy immunotherapy tablet MK-3641 (Ragwitek(R)) for the treatment of allergic rhinitis. *Expert Rev Clin Immunol*. 2018;14(12):1003-1011.
 20. Couroux P, Ipsen H, Stage BS, et al. A birch sublingual allergy immunotherapy (SLIT) tablet educates rhinoconjunctivitis symptoms when exposed to birch and oak and induces IgG4 to allergens from all trees in the birch homologous group. *Allergy*. 2018;74(2):361-369.
 21. Norman P, Lichtenstein L, Tignall J. The clinical and immunologic specificity of immunotherapy. *J Allergy Clin Immunol*. 1978;61(6):370-377.
 22. Ellis AK, Tenn MW, Steacy LM, et al. Lack of effect of Timothy grass pollen sublingual immunotherapy tablet on birch pollen-induced allergic rhinoconjunctivitis in an environmental exposure unit. *Ann Allergy Asthma Immunol*. 2018;120(5):495-503.
 23. Lorenz AR, Luttkopf D, May S, Scheurer S, Vieths S. The principle of homologous groups in regulatory affairs of allergen products—a proposal. *Int Arch Allergy Immunol*. 2009;148(1):1-17.
 24. Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*. 2005;4(9):1265-1272.
 25. Lund G, Brand S, Ramos T, et al. Strong and frequent T-cell responses to the minor allergen Phl p 12 in Spanish patients IgE-sensitized to Profilins. *Allergy*. 2018;73(5):1013-1021.
 26. González-Galarza FF, Takeshita L, Santos E, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res*. 2015;43(D1):D784-D788.
 27. Karosiene E, Rasmussen M, Blicher T, Lund O, Buus S, Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics*. 2013;65(10):711-724.
 28. Wopfner N, Gadermaier G, Egger M, et al. The spectrum of allergens in ragweed and mugwort pollen. *Int Arch Allergy Immunol*. 2005;138(4):337-346.
 29. Huang SK, Zwollo P, Marsh DG. Class II major histocompatibility complex restriction of human T cell responses to short ragweed allergen, Amb a V. *Eur J Immunol*. 1991;21(6):1469-1473.
 30. Burton DA, Shepherd GM, Siskind GW, Posnett DN. T cell receptor variable gene expression: analysis in ragweed-sensitive patients during allergen exposure. *Int Arch Allergy Appl Immunol*. 1990;91(3):306-310.
 31. Bond JF, Garman RD, Keating KM, et al. Multiple Amb a I allergens demonstrate specific reactivity with IgE and T cells from ragweed-allergic patients. *J Immunol*. 1991;146(10):3380-3385.
 32. Hinz D, Oseroff C, Pham J, Sidney J, Peters B, Sette A. Definition of a pool of epitopes that recapitulates the T cell reactivity against major house dust mite allergens. *Clin Exp Allergy*. 2015;45(10):1601-1612.
 33. Creticos PS, Schroeder JT, Hamilton RG, et al. Immunotherapy with a ragweed-toll-like receptor 9 agonist vaccine for allergic rhinitis. *N Engl J Med*. 2006;355(14):1445-1455.
 34. Norman PS. Therapeutic potential of peptides in allergic disease. *Ann Allergy*. 1993;71(3):330-333.
 35. Klimek L, Pfaar O, Worm M. New opportunities for allergen immunotherapy using synthetic peptide immuno-regulatory epitopes (SPIREs). *Expert Rev Clin Immunol*. 2016;12(10):1123-1135.
 36. Sancho AI, Wallner M, Hauser M, et al. T cell epitope-containing domains of Ragweed Amb a 1 and Mugwort Art v 6 modulate immunologic responses in humans and mice. *PLoS ONE*. 2017;12(1):e0169784.
 37. Jahn-Schmid B, Hauser M, Wopfner N, et al. Humoral and cellular cross-reactivity between Amb a 1, the major ragweed pollen allergen, and its mugwort homolog Art v 6. *J Immunol*. 2012;188(3):1559-1567.
 38. Nolte H, Hébert J, Berman G, et al. Randomized controlled trial of ragweed allergy immunotherapy tablet efficacy and safety in North American adults. *Ann Allergy Asthma Immunol*. 2013;110(6):450-456.
 39. Asero R, Wopfner N, Gruber P, Gadermaier G, Ferreira F. Artemisia and Ambrosia hypersensitivity: co-sensitization or co-recognition? *Clin Exp Allergy*. 2006;36(5):658-665.

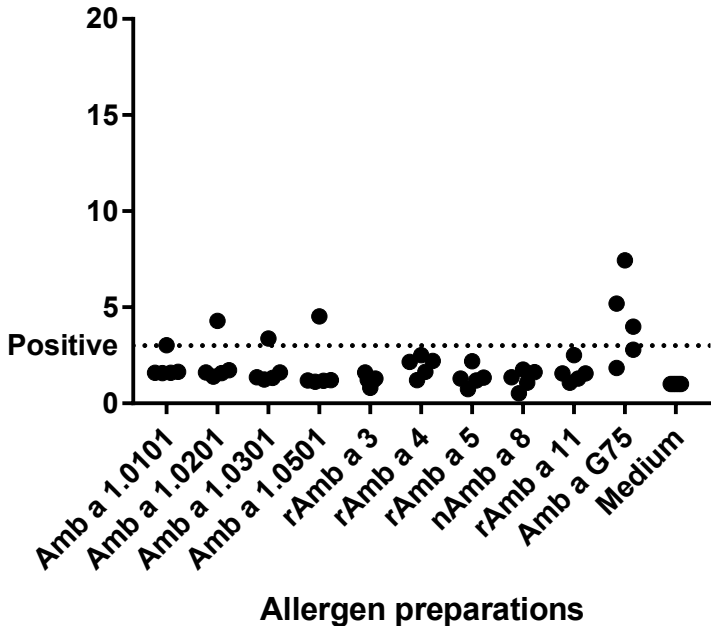
SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Würtzen PA, Hoof I, Christensen LH, et al. Diverse and highly cross-reactive T-cell responses in ragweed allergic patients independent of geographical region. *Allergy*. 2020;75:137-147. <https://doi.org/10.1111/all.13992>

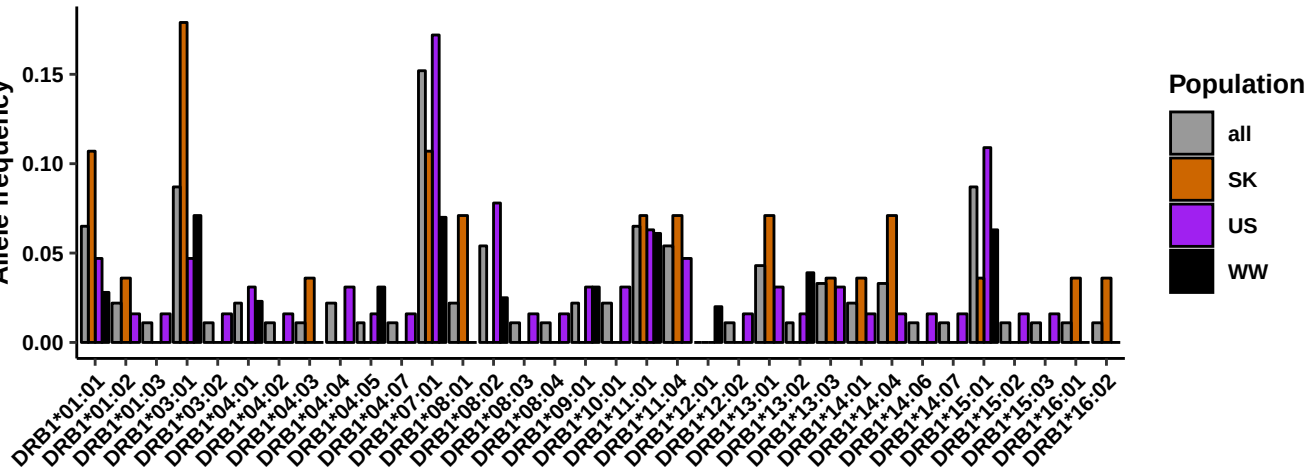


SI



Section 1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Amb a 1.0101 (P27759)	(1)	1	10	20	30	40	50	60	70	80	90	100	110	126																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
Amb a 1.0201 (P27760)	(1)	--AED	LE	ILF	--VNET	RL	ITTS	GA	YNI	ID	GC	WRG	KAD	WAEN	NK	ALAD	CAQ	GG	F	KGT	Y	GGK	D	GDI	Y	T	V	T	S	L	DD	D	V	A	N	P	K	E	G	T	L	R	F	G	A	A	Q	N	R	P	L	W	I	I	F	E	R	D	M	V	I	R	L	D	K	E	M	V	N	S	D	K	T	I	D	G	R	G	A	K	V	E																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Amb a 1.0304 (E1XUL4)	(1)	--AEG	VE	ILF	S	VNET	--SL	QACE	AL	NI	ID	GC	WRG	KAD	WAEN	NK	ALAD	CAQ	GG	F	KGT	Y	GGK	W	GDI	Y	T	V	T	S	L	DD	D	V	A	N	P	K	E	G	T	L	R	F	G	A	A	Q	N	R	P	L	W	I	I	F	K	N	D	M	V	I	N	L	N	Q	E	L	V	N	S	D	K	T	I	D	G	R	G	V	K	V	N																																																																																																																																																																																																																																																																																																																																																																																																																																																					
Amb a 1.0401 (P28744)	(1)	--AED	L	Q	ILF	S	--ANET	RS	IT	TC	T	CG	YNI	ID	GC	WRG	KAD	WAEN	NK	ALAD	CAQ	GG	F	KGT	Y	GGK	D	GDI	Y	T	V	T	S	L	DD	D	V	A	N	P	K	E	G	T	L	R	F	G	A	A	Q	N	R	P	L	W	I	I	F	E	R	D	M	V	I	R	L	D	R	E	L	A	N	K	D	K	T	I	D	G	R	G	A	K	V	E																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Amb a 1.0501 (P27762)	(1)	--GEE	V	ILF	S	P	ND	TRRS	L	Q	CE	AHN	I	D	K	C	WR	C	P	D	WAEN	R	QAL	GN	CAQ	GG	F	K	G	A	T	H	GG	K	W	G	D	I	Y	M	V	T	S	D	Q	DD	V	N	P	K	E	G	T	L	R	F	G	A	T	Q	D	R	P	L	W	I	I	F	Q	R	D	M	I	Y	L	Q	E	M	V	T	S	D	K	T	I	D	G	R	G	A	K	V	E																																																																																																																																																																																																																																																																																																																																																																																																																																											
Amb p-homolog of Amb a 1.02	(1)	--AEE	V	ILF	S	ANET	RRSL	K	CE	AHN	I	D	K	C	WR	C	KAD	WEN	N	RQAL	AD	CAQ	GG	F	KGT	H	GG	K	W	G	D	I	Y	M	V	T	S	D	K	DD	D	V	A	N	P	K	E	G	T	L	R	F	G	A	A	Q	N	R	P	L	W	I	I	F	K	R	D	M	V	I	N	L	N	Q	E	L	V	N	S	D	K	T	I	D	G	R	G	V	K	V	N																																																																																																																																																																																																																																																																																																																																																																																																																																													
Amb t-homolog of Amb a 1.05	(1)	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Allele frequency



	Amb a 1.0101	Amb a 1.0201	Amb a 1.0304	Amb a 1.0401	Amb a 1.0501	Amb p-homolog of Amb a 1.02	Amb p-homolog of Amb a 1.05	Amb t-homolog of Amb a 1.02	Amb t-homolog of Amb a 1.05	Art v 6.0101
Amb a 1.0101	100	75	76	82	66	75	67*	77*	64*	61
Amb a 1.0201	75	100	87	73	69	88	71*	82*	68*	65
Amb a 1.0304	76	87	100	72	70	88	71*	78*	69*	65
Amb a 1.0401	82	73	72	100	62	70	61*	72*	58*	58
Amb a 1.0501	66	69	70	62	100	72	90*	67*	90*	61
Amb p-homolog of Amb a 1.02	75	88	88	70	72	100	72*	79*	69*	67
Amb p-homolog of Amb a 1.05	67*	71*	71*	61*	90*	72*	100	66*	87*	73*
Amb t-homolog of Amb a 1.02	77*	82*	78*	72*	67*	79*	66*	100	63*	67*
Amb t-homolog of Amb a 1.05	64*	68*	69*	58*	90*	69*	87*	63*	100	61*
Art v 6.0101	61	65	65	58	61	67	73*	67*	61*	100

Diverse and highly cross-reactive T cell responses in ragweed allergic patients independent of geographical region

Würtzen, P.A.¹, Hoof, I.¹, Christensen, L.H.¹, Váczy, Z.², Henmar, H.¹, Salamanca, G.¹, Lundegaard, C.¹, Lund, L.¹, Kráľova, T.³, Brooks, E.G.⁴, and Andersen, P.S.¹.

¹Global Research, ALK, Hørsholm, Denmark; ²Louis Pasteur University Hospital, Košice, Slovak Republic;

³Immunology and Allergology, Košice, Slovak Republic; ⁴UT Health San Antonio, Texas, USA

Online Repository materials

Methods in detail

- Study populations

Amb a allergic individuals (71 US and 31 Slovakian (SK) donors) were recruited for the study on basis of previously performed SPT and a clinical history of ragweed allergy. After subsequent screening for IgE responses to Amb a, 38 US and 17 SK donors were included for T-cell experiments based on an IgE response of 0.6 kU/L or above if sufficient PBMC were available. Conclusive T-cell responses were obtained for 31 US and 14 SK donors. 20 US donors were used for the initial epitope mapping while 11 US and 14 SK donors were used for investigation of responses to selected T-cell epitopes and homologous peptides. All donors with conclusive T-cell data were used to investigate linkage between the different immune parameters investigated. The study was approved by the UT Health San Antonio (US) institutional review board (HSC20150773H) as well as by the local ethical committee of Kosice, Slovakia (4451/2016/OSVaZ-20811). The ragweed allergic patient population analyzed for T-cell reactivity were 47% females, and the average age was 38.3 years. All patients suffered from rhinitis and 33.3% had physician diagnosed asthma. Most patients (96%) were sensitized to other allergens than ragweed. None of the enrolled patients had received AIT prior to blood collection. Full patient characteristics can be found in table 1.

- IgE measurements

IgE responses to Amb a, Amb a 1, Amb p, Amb t, Art v, Art v 1, Phl p, Cyn d, Der f, Der p, Fel d, Bet v, Que a, Cry j, and Jun s were measured by ImmunoCAP (ThermoFisher Scientific, Uppsala, Sweden)

- Allergens and peptides

Amb a extract (in-house ALK production scale extract) was gel-filtered by size exclusion chromatography using a HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences cat no.:17-1068-01) coupled to an ÄKTA explorer FPLC system (GE Healthcare Life Sciences), to remove small molecules that could potentially trigger non-specific T-cell activation during establishment of T-cell lines. Mass spectrometry analysis was applied to assess the allergome profile before and after gel-filtration as described further below.

The resulting extract preparation contained all Amb a allergens (with known sequences) recognized by IUIS including the 5 isoform groups of Amb a 1.

Experimental extracts of Amb p, Amb t and Art v were made as follows: Pollen from each species (30 g) was extracted in 300 mL of NH₄HCO₃ (0.125mol/L, pH 8.3) at 5 °C for 2 hours followed by dialysis (MW cut-off 3.5K), filtration (0.2 mm), and freeze-drying.

Recombinant allergens: rAmb a 1.0401, rAmb a 3.0101, rAmb a 4.0101 and rproAmb a 11.0101 were expressed in *E. coli* and purified via a N-terminal His-tag (custom-made by Genscript NJ, USA). Endotoxins were removed and measured to <0.01 EU/ug.

The gel-filtrated Amb a extract (see above) was used for purification of Amb a 1 isoforms (Amb a 1.01, Amb a 1.02, Amb a 1.03 and Amb a 1.05). The first separation step was a Cu-chelate column (GE Healthcare Life Sciences HiTrap Chelating HP 5 ml cat: 17-0409-03), followed by a combination of ion exchange (GE Healthcare Life Sciences HiTrap SP HP 5 ml cat: 17-1152-01 (cation)) and hydrophobicity (GE Healthcare Life Sciences HiTrap Phenyl HP 5 ml cat: 17-5195-01). The last step was a size exclusion chromatography using a HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences cat no.:17-1068-01). nAmb a 8 was purified from the gel-filtrated Amb a extract following a CNBr activatedTM 4B sepharose (GE Healthcare Bio Sciences LB cat:17-0430-02) coupled with Poly-L-Proline (Sigma-Aldrich Product no. P2129) column. The final step was size exclusion chromatography using a HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences cat no.:17-1068-01).

All steps to obtain the purified natural allergens were performed on an ÄKTA explorer FPLC system (GE Healthcare Life Sciences). The purity of the recombinant allergen preparations were above 90% apart from Amb a 4, which was 85% according to the certificate of Analyses included with the delivery. Natural purified Amb a 1 isoallergens and Amb a 8 were estimated to be at least 95% pure based on SDS-PAGE and mass spec analysis.

Endotoxin content of all purified allergens were below 20 EU/mg except nAmb a 8 which was 53 EU/mg. In our hands endotoxin levels above 20 EU/mg will cause unspecific T-cell proliferation at an allergen concentration of 10 ug/ml, so we use 0.2 EU/ml as the limit of endotoxin content for all allergen preparations used for stimulation. Amb a 8 was used in 2 ug/ml leading to an endotoxin level of 0.11 EU/ml.

Peptides: Overlapping 20mers, selected peptides 20-23mers, homologous 20-23mers and Amb a 11 derived 20mers were custom-made by Genscript NJ, USA with a purity of ≥ 95 %.

- LC-MS/MS allergome profiling of Amb a extract used for T-cell stimulation

5 µl Amb a extracts (before and after gel-filtration) were denatured (6 M urea, 0.3 M NH_4HCO_3), reduced (9 mM DTT, 56 °C for 15 min), alkylated (17 mM Iodoacetamide) and trypsin digested (37 °C, overnight). Subsequently, tryptic peptides were separated by reverse phase liquid chromatography (Ultimate 3000 RSLC nano, Thermo Fischer Scientific) via a C18 analytical column (flow rate 300 nl/min, 4 hour gradient from 4-55 % acetonitrile). Eluting peptides were sprayed into an ESI-QTOF mass spectrometer (Maxis, Bruker), spectra acquired in the mass range of 50-2200 m/z at 2 Hz, and MS/MS sequencing at a spectral rate of 4-16 Hz. Data analysis: MS/MS spectra were searched (via MASCOT search engine, Matrix Science) against an in-house allergen database that included the protein sequences of all IUIS Amb a allergens and isoforms thereof, except for Amb a 7 that has not yet been cloned/sequenced. The sum of the relative and semi-quantitative Exponentially Modified Protein Abundance Index (emPAI) scores⁽¹⁾ of all Amb a allergens were set to 100%, and the relative percentage (molar %) of each allergen was calculated.

- T-cell activation⁽²⁾

Peripheral blood mononuclear cells (PBMC) were isolated from heparinized blood according to standard procedure and stored in liquid nitrogen. Amb a-specific T-cells were expanded *in vitro*, and specific responses were measured by Fluorescence based ImmunoSpot (FluoroSPOT) at day 14 and by thymidine incorporation at day 24 or after additional culture cycles with non-specific stimulation. Briefly, frozen PBMCs were thawed, CD8 removed using magnetic CD8 Dynabeads (cat 111.47 D, Dynal), stimulated with gel-filtered Amb a extract (0.5 ug/ml) and expanded for 14 days with addition of IL-2 on days 5, 6, and 7 (Proleukin, Chiron). On day 14, the T-cells were isolated and re-stimulated (2×10^5 /ml) in 24 well plates (1 ml/well) with gel-filtered Amb a extract (0.5 ug/ml) and PHA (0.5 μ g/ml, supplier) in the presence of autologous PBMC (10^6 /ml, irradiated 3000 Rad), and IL-2 was added on days 3-5 of this 10 day culture period. Excess T-cells were frozen and stored in liquid nitrogen. If needed, further culture of T-cells at day 24 or T-cells thawed from storage was performed in 24 well plates ($2-4 \times 10^4$ T-cells/ml, 1 ml/well) by un-specific stimulation with PHA (0.5 μ g/ml) using a mixture of JY EBV B-cells (1×10^5 /ml, irradiated 5000 Rad), and allogeneic PBMC (1×10^6 /ml, irradiated 3000 Rad) as feeder cells. IL-2 was added at day 3-5 of these 10 day culture periods and all data were obtained within a 6-7 week culture period for the individual T-cell lines.

Using this protocol T-cell lines with reproducible responses to the Amb a allergen preparations were almost exclusively observed in T-cell cultures from Amb a SPT and IgE positive donors. The responses of 5 allergic Amb a IgE negative donors to amb extract and allergens resulted in marginal responses for a few allergens for one donor and towards the Amb a G75 extract preparation in 3 out of 5 IgE negative donors (figure E1B). all responses were at a very low level compared to the Amb a responses from ragweed allergic donors shown in figure 1B. Based on these observations, we are convinced that the T-cell responses are specific and related to the ragweed allergic status of the donors investigated in the current manuscript.

- FluoroSPOT⁽²⁾

On day 14, the T-cell lines were tested for cytokine production by stimulation with 2 μ g/ml of the different peptides, full-length allergens (2 μ g/ml), allergen extracts (5-10 μ g/ml) and controls (PHA or medium alone), at 1×10^5 cells/well (100 ul/well, triplicates) in FluoroSpot plates: IL-5/IFN γ FluoroSPOT (Mabtech FS-0108-10), for 16-20 h at 37°C. Fluorospot was developed and measured according to the manufacturer's protocol.

Criteria for positive responses were determined relative to cells without stimuli (medium): Data with significant positive responses (Student's t-test, $p < 0.05$), with at least 20 spot forming cells (SFC) per 10^6 PBMCs above baseline, and with a stimulation index (SI) > 2 were regarded as positive and the actual SFC used for data analysis. All other measurements were set to zero. Donor coverage was calculated from the number of positive responses relative to the number of valid (Amb a positive) T-cell lines.

- Proliferation⁽²⁾

At day 24 or later, T-cell activation by peptides or allergens was measured by incorporation of tritiated thymidine, and cell supernatants were collected for cytokine measurements (below) as described previously (ref). Briefly, T-cells were tested for specific activation by stimulation with the different peptides (2 μ g/ml), full-length allergens (2 μ g/ml), allergen extracts (5-10 μ g/ml) and controls (PHA or medium alone), at 3×10^4 cells/well (100 ul/well, triplicates) in 96 well U-bottom plates (Nunc, Roskilde, DK), for 72 h at 37°C. Autologous PBMC (2×10^4 /ml, irradiated 3000 Rad) were added as APC. Cell supernatants were

harvested after 48 hours, and ^3H -thymidin was added for the last 16-20 hours of the culture period before determination of thymidine incorporation by scintillation counting.

Criteria for positive responses were determined relative to cells without stimuli (medium): Data with significant positive responses (Student's t-test, $p < 0.05$) and with a stimulation index > 3 were regarded as positive and the actual CPM values used for further data analysis. All other measurements were set to zero. Donor response frequencies were calculated from the number of positive responses relative to the number of valid (Amb a positive) T-cell lines.

- Measurements of cytokine secretion⁽²⁾

Cytokine content in cell supernatants were assessed by multiplex electrochemiluminescent immunoassays (MesoScale Discovery, Gaithersburg, MD, USA). T-cell derived cytokines were measured using a kit for measurement of IFN- γ , IL-10, IL-12p70, IL-13, IL-2, IL-4 and IL-5. All reagents were provided with the kit and analyses followed the manufacturer's guidelines. The electrochemical-luminescence of the secondary ruthenium-conjugated antibody was read in a MesoScale Discovery Sector (MSD) Imager plate reader and analyzed using the associated workbench software. Cytokine data were used to further substantiate the T-cell activation determined in the thymidine incorporation assay, and only cytokine data from stimulations with positive T-cell proliferation were used for analysis.

- Quantitative assessment of T cell epitope cross-reactivity

The level of cross-reactivity was evaluated on the basis of T-cell proliferation data, and all T-cell lines with a response to the original Amb a 20-mer peptide were included in the analysis. The criterion for cross-reactivity was that T-cell proliferation of similar strength (SI) should be obtained for a number of T-cell lines. To assess this, we used the Pearson correlation coefficient to quantify cross-reactivity, defining T cell responses against a given homologous peptide pair as cross-reactive if the correlation coefficient was significant ($p < 0.05$) and the correlation coefficient was above 0.5.

- HLA typing and predicted population coverage

Genotyping of the HLA class II locus to 4-digit resolution was performed by ProImmune Ltd using a PCR-based assay. We calculated the predicted population coverage of the tested peptides in order to get an estimate of what fraction of a population would be expected to carry one or more HLA alleles that are predicted to bind and present a given peptide. The predicted population coverage C_p of peptide p was defined as $C = 2 * Sf - (Sf * Sf)$ where Sf is the sum of the frequencies of the HLA-DRB1 alleles that are predicted to bind peptide p .

Allele frequencies were derived from <http://allelefrequenciest.net/>⁽³⁾ using the populations "USA Colorado Univ Cord Blood Bank Hispanic", "USA Colorado Univ Cord Blood Bank Caucasian", "USA African American pop 4", and "USA Asian pop 2" in order to cover the four major ethnicities in the US. HLA-DRB1 alleles with an observed frequency > 0 in any of the four ethnic groups were considered relevant. HLA-peptide binding predictions were performed using NetMHCIIpan-3.0⁽⁴⁾. A given HLA-DRB1 allele was defined as binding a given peptide if the predicted affinity was < 300 nM and the rank score $< 30\%$.

- Transcriptomes and homology:

RNA was isolated from Amb p and Amb t pollen (ALK). Library preparation and RNA-seq (Illumina HiSeq 2500) was performed by Eurofins Genomics. All sequences were trimmed with respect to sequencing quality using fastq_quality_trimmer in the FASTX Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/). Bases with a quality of less than 20 were trimmed from both ends, and reads were discarded if they were shorter than 50 bp after trimming. For de novo assembly of the transcriptomes, the software Trinity v2.1.1⁽⁵⁾ was used. Homologous sequences of Amb a allergens in the transcriptomes of Amb p and Amb t were identified with blastp.

Identification of Amb a homolog allergens derived from Amb p and Amb t by LC-MS/MS

LC-MS/MS analysis of Amb p and Amb t was made similar to Amb a as described above.

MS/MS spectra were searched (via MASCOT search engine, Matrix Science) against an in-house database that included Amb p and Amb t homolog sequences of known Amb a allergens, derived from the Amb p and Amb t transcriptomes described above.

All allergen sequences used for selection of homologous peptides are listed in figure E2. The level of identity for the ragweed group 1 allergens was generally high similar to the identity between Amb a 1 Isoforms whereas the identity to Art v 6⁽⁶⁾ was somewhat lower (Table E1). One or more Art v homologues for Amb a 1, 4, and 8 were available from previous publications⁽⁶⁾ and these were also used for generation of homologous peptides (figure E2).

- Data processing and visualization

Data processing and visualization was performed using R 3.5.0⁽⁷⁾. For the creation of most figures, the R package ggplot2 (version 2.2.1)⁽⁸⁾ was used.

Results:

All full length allergen and extract preparations used in the study are included in the SDS-PAGE gel seen in figure E1. Clear dominant bands are seen for all purified, recombinant, or synthetic proteins and multiple bands in the range from 3 to 96 kD was observed for the allergen extracts.

Table E1 shows the amino acid sequence identity of the Amb a, p, t 1 and Art v 6 sequences used in the study. Sequence identities between 62% and 87% were seen for Amb a 1 isoallergen groups and identities between 58% and 90% were seen when the Amb p and t 1 sequences identified in the transcriptomes and the known Art v 6 sequence were included in the comparison. Amb p and t 1 names were composed on basis of the Amb a 1 sequences with the maximum sequence homology. The exact sequences are listed in figure E2, where the sequences for Amb p and t homologous to the minor allergens investigated are included and the names are composed similar to the strategy for the group 1 allergens.

The DRB1 allele frequencies of the 31 Us and 14 Slovakian donors are seen in figure E3 where the world wide allele frequencies of DRB1 alleles with available frequency data are included for comparison. Overall, allele frequencies of the total donor population is similar to the world wide allele frequencies and no major allele frequency bias is seen between US and Slovakian donors even though some discrepancies are seen in the frequency for individual alleles.

Figure legends:

Table E1: Sequence identity between all Amb a 1, Amb p 1, Amb t 1, and Art v 6 molecules investigated. Amb p 1 and Amb t 1 sequences identified in the Amb p and Amb t transcriptome/allergome analyses were named on basis of the maximal sequence identity to known Amb a 1 isoform groups (blue circles). * indicate truncated sequences (% identity within coding region).

Figure E1: SDS-PAGE analysis of all full length protein and pollen allergen extracts used for T-cell activation (A). Amb a G75 was used to initiate all T-cell cultures and the rest of the preparations were used for FluoroSPOT and proliferation assays as described in the methods section. T-cell responses of 5 Amb a IgE negative donors (B). T-cell proliferation (SI=cpm stim/cpm medium control, mean +/- 95% conf.) to ragweed allergens, selected group 1 isoallergens and Amb a allergen extract detected by thymidine incorporation.

Figure E2A: Sequence alignment between all Amb a 1, Amb p 1, Amb t 1, and Art v 6 molecules investigated. Amb p 1 and Amb t 1 sequences identified in the Amb p and Amb t transcriptome/allergome analyses were named on basis of the maximal sequence identity to known Amb a 1 isoform groups (Table E1, blue circles).

Figure E2B: Sequence alignment between all Amb a/p 3, Amb a 4/Art v 1, Amb a/p/t 8/Art v 4, and Amb a/p/t 11 molecules investigated. Amb p and Amb t allergen sequences identified in the Amb p and Amb t transcriptome/allergome analyses were named on basis of the maximal sequence identity to known Amb a allergens.

Figure E3: HLA-DRB1 allele frequencies in subgroups. DRB1 allele for all 45 donors (grey), 31 US (purple) and 14 SK (orange) were determined by genotyping of the HLA class II locus by PCR. World wide frequencies (black) are obtained from <http://allelefrequencies.net/> and included for comparison.

References

- 1 Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005; 4(9):1265-72.
- 2 Lund G, Brand S, Ramos T, Jimeno L, Boissy P, Vega F et al. Strong and frequent T-cell responses to the minor allergen Phl p 12 in Spanish patients IgE-sensitized to Profilins. *Allergy* 2018; 73(5):1013-21.
- 3 Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res* 2015; 43(Database issue):D784-D788.

- 4 Karosiene E, Rasmussen M, Blicher T, Lund O, Buus S, Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics* 2013; 65(10):711-24.
- 5 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011; 29(7):644-52.
- 6 Wopfner N, Gadermaier G, Egger M, Asero R, Ebner C, Jahn-Schmid B et al. The spectrum of allergens in ragweed and mugwort pollen. *Int Arch Allergy Immunol* 2005; 138(4):337-46.
- 7 R Core Team. R: A language and environment for statistical computing. 2018.
Ref Type: Online Source
- 8 Wickham H. ggplot2: Elegant graphics for data analysis. 2009.
Ref Type: Online Source