T-cell epitope conservation across allergen species is a major determinant of immunogenicity

Luise Westernberg, PhD,[a] Véronique Schulten, PhD,[a] Jason A. Greenbaum, PhD,[a] Sara Natali,[c] Victoria Tripple, BS,[a] Denise M. McKinney,[a] April Frazier, PhD,[a] Heidi Hofer, PhD,[b] Michael Wallner, PhD,[b] Federica Sallusto, PhD,[c,d] Alessandro Sette, PhD,[a] and Bjoern Peters, PhD[a] *La Jolla, Calif, Salzburg, Austria, and Bellinzona, Switzerland*

Background: Patients with pollen allergies are frequently polysensitized. Pollens contain epitopes that are conserved across multiple species.
Objective: We sought to demonstrate that cross-reactive T cells that recognize conserved epitopes show higher levels of expansion than T cells recognizing monospecific epitopes because of more frequent stimulation.
Method: RNA was sequenced from 9 pollens, and the reads were assembled *de novo* into more than 50,000 transcripts. T-cell epitopes from timothy grass *(Phleum pratense)* were examined for conservation in these transcripts, and this was correlated to their ability to induce T-cell responses. T cells were expanded *in vitro* with *P pratense*–derived peptides and tested for cross-reactivity to pollen extracts in ELISpot assays.
Results: We found that antigenic proteins are more conserved than nonimmunogenic proteins in *P pratense* pollen. Additionally, *P pratense* epitopes that were highly conserved across pollens elicited more T-cell responses in donors with grass allergy than less conserved epitopes. Moreover, conservation of a *P pratense* peptide at the transcriptomic level correlated with the ability of that peptide to trigger T cells that were cross-reactive with other non–*P pratense* pollen extracts.
Conclusion: We found a correlation between conservation of peptides in plant pollens and their T-cell immunogenicity within *P pratense*, as well as their ability to induce cross-reactive T-cell responses. T cells recognizing conserved epitopes might be more prominent because they can be stimulated by a broader range of pollens and thereby drive polysensitization in allergic donors. We propose that conserved peptides could potentially be used in diagnostic or immunomodulatory approaches that address the issue of polysensitization and target multiple pollen allergies. (J Allergy Clin Immunol 2016;138:571-8.)

Key words: *T cell, epitope, timothy grass allergy, pollen allergy, cross-reactivity, RNA sequencing, sequence conservation, transcriptome*

Patients with pollen allergies are often polysensitized, as evidenced by positive RAST and/or skin prick test results to multiple pollen allergens. The relatively low frequency of monosensitizations can be explained by the presence of cross-reactive IgE epitopes conserved across multiple pollens, which result in immune reactivity to homologous regions in allergens to which the patient was not originally sensitized.[1] In the context of immunotherapy, the high degree of polysensitization in subjects suggests that a single allergen administered in therapeutic mode could be sufficient to induce tolerance. In fact, several investigators have suggested that immunotherapy with a single grass species, such as *Phleum pratense*, is sufficient to also treat allergies to other grass pollens caused by observed cross-reactivity at the IgE level.[2,3] On the other hand, it is firmly established that allergen-specific T cells play an important role in allergic inflammation[4] and that induction of antigen-specific regulatory T cells[5] or elimination of allergen-specific T_H2 cells might be a prerequisite for the induction of specific tolerance.[6] Yet evaluation of cross-reactivity at the T-cell level has been less documented.

A recent study using tetramer costaining of 6 different MHC-epitope complexes found limited cross-reactivity of these epitopes with homologs in other Pooidea grasses and concluded that multiple grass pollen species immunotherapy is likely to be more beneficial than single-species immunotherapy.[7] Although that study was limited to epitopes from 2 major allergens (Phl p 1 and 5), we have recently shown that a large fraction of *P pratense*–specific T cells target epitopes contained in timothy grass T-cell antigens (TGTAs) unrelated to the known *P pratense* allergens that are also major targets of T-cell responses but were initially identified based on their high IgE reactivity.[8] On the basis of these data, we reasoned that a broader evaluation of T-cell cross-reactivity including more epitopes and also those from TGTAs would be of interest, particularly because clinical studies have shown a good degree of success for single-species immunotherapy[9,10] contrary to what might be expected based on the data presented in the tetramer costaining study for a selected set of epitopes. To gain a comprehensive picture of conservation between different grass, weed, and tree pollens, we sequenced the

From [a]the La Jolla Institute for Allergy and Immunology; [b]the Department of Molecular Biology, University of Salzburg; and [c]the Laboratory of Cellular Immunology and [d]Center of Medical Immunology, Institute for Research in Biomedicine, University of Italian Switzerland, Bellinzona.

Supported in part by federal funds from the National Institutes of Allergy and Infectious Diseases (grant no. U19AI100275) and the European Research Council (grant no. 323183 PREDICT) in addition to funds from ALK-Abelló A/S (Hørsholm, Denmark). The Institute for Research in Biomedicine and the Center of Medical Immunology are supported by the Helmut Horten Foundation.

Disclosure of potential conflict of interest: L. Westernberg has received travel support from ALK-Abelló. A. Sette and B. Peters have received grants from the NIH and ALK-Abelló, have consultant arrangements with ALK-Abelló, and have a patent in collaboration with ALK-Abelló. S. Natali and F. Sallusto have received grants from the European Research Council. H. Hofer is an employee of the University of Salzburg. M. Wallner is employed by the University of Salzburg and has received grants from the University of Salzburg, the Austrian Science Fund (FWF), and the Austrian Federal Ministry of Science, Research, and Economy (BMWFW). The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication February 14, 2015; revised November 2, 2015; accepted for publication November 18, 2015.

Available online February 13, 2016.

Corresponding author: Bjoern Peters, PhD, Division of Vaccine Discovery, La Jolla Institute for Allergy & Immunology, La Jolla, CA 92037. E-mail: bpeters@liai.org.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749

© 2016 The Authors. Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.1016/j.jaci.2015.11.034

Abbreviations used
IUIS: International Union of Immunological Societies
SIT: Specific immunotherapy
TGTA: Timothy grass T-cell antigen

transcriptome of 9 allergenic pollen species and determined how the conservation between *P pratense* T-cell epitopes and other pollen transcriptomes related to their immunogenicity and their ability to elicit cross-reactive T-cell immune responses.

METHODS

Patient population

Study participants were recruited, as previously described.[11] We drew from a cohort of 55 patients between the ages of 19 and 62 years, and 28 of our patients were female. We only included participants with timothy grass allergy who had skin reactions with wheals of 5 mm or greater in diameter after skin prick testing or *P pratense*–specific IgE levels of 0.35 kU/L or greater, as determined by RAST, and had a history of seasonal grass pollen allergy symptoms. We did not control for other allergies, and many patients were polysensitized. We included patient samples collected both in and out of the season. Further details about the age, sex, and allergy status of the patients used in our studies can be found in Table E1 in this article's Online Repository at www.jacionline.org. Patients' PBMCs used in T cell cross-reactivity studies were further prescreened based on their reactivity to *P pratense* peptides after timothy grass culture.

RNA sequencing and *de novo* transcriptome assembly

Total RNA of the pollen extract was isolated, as previously described.[12] RNA was sequenced on a HiSeq 2500 Sequencer (Illumina, Sn Diego, Calif). Replicate samples were run across all lanes of the sequencer to generate paired reads of 100 bp in each direction. Before assembly, several preprocessing steps were performed: (1) reads not passing Illumina filters were discarded, (2) portions of reads matching adapter/primer sequences were trimmed, (3) 3′ regions of reads following a low-quality score ($Q < 20$) were discarded, and (4) remaining reads of less than 30 bp in length were discarded. These preprocessing steps were performed with a combination of FASTX-toolkit (0.0.13.2)[13] and cutadapt (1.3).[14] High-quality reads were assembled into transcripts by using Trinity (r2012-10-05),[15] specifying "–min_kmer_cov 2" to ensure each sequence was observed at least twice.

In vitro expansion of allergen-specific T cells from PBMCs

PBMCs were isolated from whole blood by means of density gradient centrifugation and cryopreserved, as previously described.[8] For *in vitro* expansion, cells were thawed and cultured with RPMI 1640 (Omega Scientific, Tarzana, Calif) supplemented with 5% human AB serum (Cellgro, Manassas, Va) at a density of 2×10^6 cells/mL in 24-well plates (BD Biosciences, San Jose, Calif) and stimulated with peptide (0.5 μg/mL). Cells were kept at 37°C in a 5% CO_2 atmosphere, and additional IL-2 (10 U/mL; eBioscience, San Diego, Calif) was added every 3 days after initial antigenic stimulation. On day 14, cells were harvested and screened for cytokine production by means of ELISpot after restimulation with peptides or pollen extracts.

ELISpot assays

Production of IL-5 from *in vitro*–expanded PBMCs in response to peptide pool or extract stimulation was measured with an ELISpot, as previously described.[16] Briefly, 1×10^5 cells per well were incubated with a peptide pool (5 μg/mL) or extract (50 μg/mL, except for oak, which was tested at 25 μg/mL). After 22 hours, cells were removed, and plates were washed and incubated with 2 μg/mL biotinylated anti-human IL-5 antibody (Mabtech, Cincinnati, Ohio) at 37°C. After 2 hours, plates were washed, and avidin-peroxidase complex was added (Vector Laboratories, Burlingame, Calif) for 1 hour at room temperature. Peroxidase-conjugated spots were developed with 3-amino-9-ethylcarvazole solution (Sigma-Aldrich, St Louis, Mo).

T-cell clones

PBMCs were labeled with carboxyfluorescein succinimidyl ester and cultured in complete RPMI 1640 supplemented with 2 mmol/L glutamine, 1% (vol/vol) nonessential amino acids, 1% (vol/vol) sodium pyruvate, penicillin (50 U/mL), streptomycin (50 μg/mL), and 5% human serum (Swiss Red Cross, Bern, Switzerland) at a density of 2×10^6 cells/mL in 24-well plates. Cells were stimulated with pools of peptides (0.5 μg/mL per peptide), and rIL-2 (10 U/mL) was added on day 3 after initial antigenic stimulation. At day 11, cells were stained with Pacific blue–labeled anti–inducible costimulator mAb (C398.4A; BioLegend, San Diego, Calif) and BV785-labeled anti-CD25 mAb (BC96; BioLegend). carboxyfluorescein succinimidyl ester–low, inducible costimulator–positive, $CD25^+$ cells were sorted and cloned by means of limiting dilution, as previously described.[17] T-cell clones were screened at day 20 after initial stimulation by culturing 3×10^4 T cells/well with autologous irradiated (95 Gy) EBV-transformed B cells (2×10^4) in the absence or presence of allergen extracts (50 μg/mL) or peptides (0.5 μg/mL). *Mycobacterium tubercolosis* lysate (5 μg/mL) was used as a negative control. Proliferation was measured on day 3 after a 16-hour pulse with tritiated thymidine (GE Healthcare, Fairfield, Conn).

RESULTS

Determining sequence conservation among a diverse selection of pollen species

To address the potential effect of sequence conservation and T-cell cross-reactivity on allergic responses, we selected 9 pollen species that represent the 3 major groups of pollen allergens (grasses, weeds, and trees). Specifically, we included 4 grasses (sweet vernal *[Anthoxanthum odoratum]*, rye *[Lolium perenne]*, Kentucky blue *[Poa pratensis]*, and Bermuda *[Cynodon dactylon]*), 3 trees (ash *[Fraxinus excelsior]*, olive *[Olea europaea]*, and oak *[Quercus alba]*), and 2 weeds (western ragweed *[Ambrosia psilostachya]* and English plantain *[Plantago lanceolata]*). Because no pollen transcriptomic data were available for any of these, we isolated and sequenced pollen RNA and assembled the reads *de novo* into transcripts, resulting in more than 50,000 transcripts for each pollen. Table E2 in this article's Online Repository at www.jacionline.org provides an overview of the read counts and assembly statistics.

As a quality control, for each pollen transcriptome, we examined whether the known IgE-reactive allergens listed by the International Union of Immunological Societies (IUIS) were reidentified in our analysis. A total of 26 allergens are listed by IUIS (minimum length, 50 residues), covering all pollen species we sequenced (except for sweet vernal grass). For 23 of these allergens, we identified transcripts that had more than 90% sequence identity for more than 50% of the length of previously described allergens (Table I). For 2 of the 3 remaining allergens (Poa p 5 and Que a 1), isoforms of the IUIS allergens that met these criteria were listed in Allergome.[18] Thus our transcriptomic analysis reidentified isoforms of all but 1 (Amb p 5) of the known allergens from the pollen species we sequenced.

Next, we wanted to determine the degree of conservation of the 10 IgE-reactive *P pratense* allergens listed by the IUIS. We

TABLE I. Recovery of known allergen sequences in transcriptomic analysis

Allergen	Match in transcriptome	Allergen	Match in transcriptome
Bermuda grass		Oak	
Cyn d 1	Yes	Que a 1	No
Cyn d 7	Yes	Ash	
Cyn d 12	Yes	Fra e 1	Yes
Cyn d 15	Yes	Olive	
Cyn d 23	Yes	Ole e 1	Yes
Cyn d 24	Yes	Ole e 2	Yes
Rye grass		Ole e 3	Yes
Lol p 1	Yes	Ole e 6	Yes
Lol p 2	Yes	Ole e 8	Yes
Lol p 3	Yes	Ole e 9	Yes
Lol p 4	Yes	Ole e 10	Yes
Lol p 5	Yes	Ole e 11	Yes
Lol p 11	Yes	English plantain	
Kentucky blue grass		Pla l 1	Yes
Poa p 1	Yes	Western ragweed	
Poa p 5	No	Amb p 5	No

All allergens from the 9 pollen species we sequenced for which there was a greater than 50-amino-acid residue protein sequence available from IUIS are listed. The "Match in transcriptome" column is set to yes if there was a match with greater than 90% sequence identity over more than 50% of the protein sequence. No sequences of sufficient length were available in IUIS for sweet vernal grass *(Anthoxanthum odoratum)*.

assessed the conservation of 15-mer peptides within each protein because this constitutes the typical length of an MHC class II–restricted T-cell epitope. Peptides were considered conserved if a transcript encoding a homolog with 2 or fewer residue mismatches was found. Fig 1, *A*, plots the percentage of peptides conserved in all the grasses we sequenced (pangrass) or even in all pollens (panpollen) for each of the IUIS allergens. A broad spectrum of conservation was detected. Less than 20% of the Phl p 2–, 4–, 5–, and 6–derived peptides were conserved across grasses. By contrast, Phl p 7, 11, and 12 peptides were highly homologous, with more than 60% of the peptides conserved across all grasses. Phl p 1, 3, and 13 exhibited intermediate conservation, with 30%, 34%, and 58% of their peptides conserved in pangrasses, respectively. A similarly broad range of conservation was observed for panpollens, with Phl p 11 and 12 showing greater than 60% conservation.

Sequence conservation contributes to immunogenicity

To test the hypothesis that the level of conservation of a protein across pollens can contribute to its immunogenicity, we examined proteins distinct from IUIS allergens that were recently identified in a transcriptomic and proteomic characterization of *P pratense* pollen. We assembled 2 sets of proteins. The first set contained 13 TGTAs (ie, all *P pratense* proteins distinct from the IUIS allergens for which we have detected T-cell responses in ≥20% of allergic patients in our original screen).[8] The second set of proteins consisted of 14 unreactive proteins identified in the same transcriptomic and proteomic analysis as the TGTAs but for which no T-cell or antibody responses were detected in any allergic donor. Fig E1 in this article's Online Repository at www.jacionline.org shows a breakdown of conservation for each of the proteins in the unreactive protein and TGTA protein sets analogous to Fig 1, *A*. These data were further condensed in Fig 1, *B* and *C*, which compare the pangrass and panpollen conservation of these protein sets to the IUIS IgE allergens. In both the pangrass and panpollen comparisons, the unreactive proteins showed the lowest degree of conservation. The median pangrass conservation of peptides from unreactive proteins was 11% compared with 32% for the IUIS allergens and 65% for the TGTAs. For the panpollen conservation, the medians were 0%, 23%, and 29% for unreactive proteins, IUIS allergens, and TGTAs, respectively. These differences were significant ($P < .05$, 1-tailed Mann-Whitney test) for both comparisons of unreactive proteins versus TGTAs and for the panpollen conservation comparison of unreactive proteins versus IUIS allergens. The pangrass conservation comparison of unreactive proteins versus IUIS allergens showed the same trend but was less than the significance threshold ($P = .087$).

In conclusion, the data presented in Fig 1 demonstrate that pollen proteins recognized by immune cells of allergic donors, either at the T-cell or IgE level, are more conserved compared with other pollen proteins and that this trend is more pronounced for T-cell allergens. This suggests that sequence conservation might be of particular relevance in the context of T-cell immunogenicity.

To determine whether conservation of the individual peptides correlates with the immunogenicity of a peptide, we studied previously generated IL-5 responses in donors with *P pratense* allergy. Originally, we tested 648 overlapping peptides from IUIS *P pratense* allergens[16] in 25 allergic donors. For 85 donor-peptide pairs (patients used are designated IUIS in Table E1), a significant IL-5 response was detected. Thus the likelihood of an individual peptide to be immunogenic in an individual donor was as follows:

$$85/(25 * 648) = 0.5\%.$$

In our follow-up study testing additional *P pratense* proteins based on transcriptomic and proteomic analysis, we screened 822 peptides for immune recognition in 20 donors (referred to as TGTA in Table E1) and identified 375 donor-peptide hits,[8] corresponding to a 2.3% likelihood for an individual peptide to be immunogenic in an individual donor. Given that the peptides in the second study were preselected based on predicted HLA-binding affinities, a higher hit rate was expected.

Next, we separated peptides into different sets based on the number of pollens in which they were conserved. For each set, we calculated the average frequency with which an individual peptide was immunogenic in an individual donor and normalized these values to an average of 1.0 within each study to make them comparable (Fig 2). Indeed, peptides from both the IUIS allergens and the TGTAs were more frequently immunogenic if they were more conserved. Peptides that were found to be conserved in only 1, 2, or 3 pollens were 49% less likely to be immunogenic than average, whereas peptides conserved in 8, 9, or 10 pollens were 47% more likely to be immunogenic than average. This correlation is highly significant ($r^2 = 0.796$, $P = .00056$). Thus conservation of a *P pratense* peptide across multiple pollens is correlated with an increased likelihood of the peptide being recognized by T cells in subjects with *P pratense* allergy.

Sequence conservation predicts allergic T-cell responses

We hypothesized that the more frequent recognition of conserved peptides might be a result of selective expansion of cross-reactive T cells by repeated stimulation with various allergen sources. Cross-reactive T cells that recognize epitopes

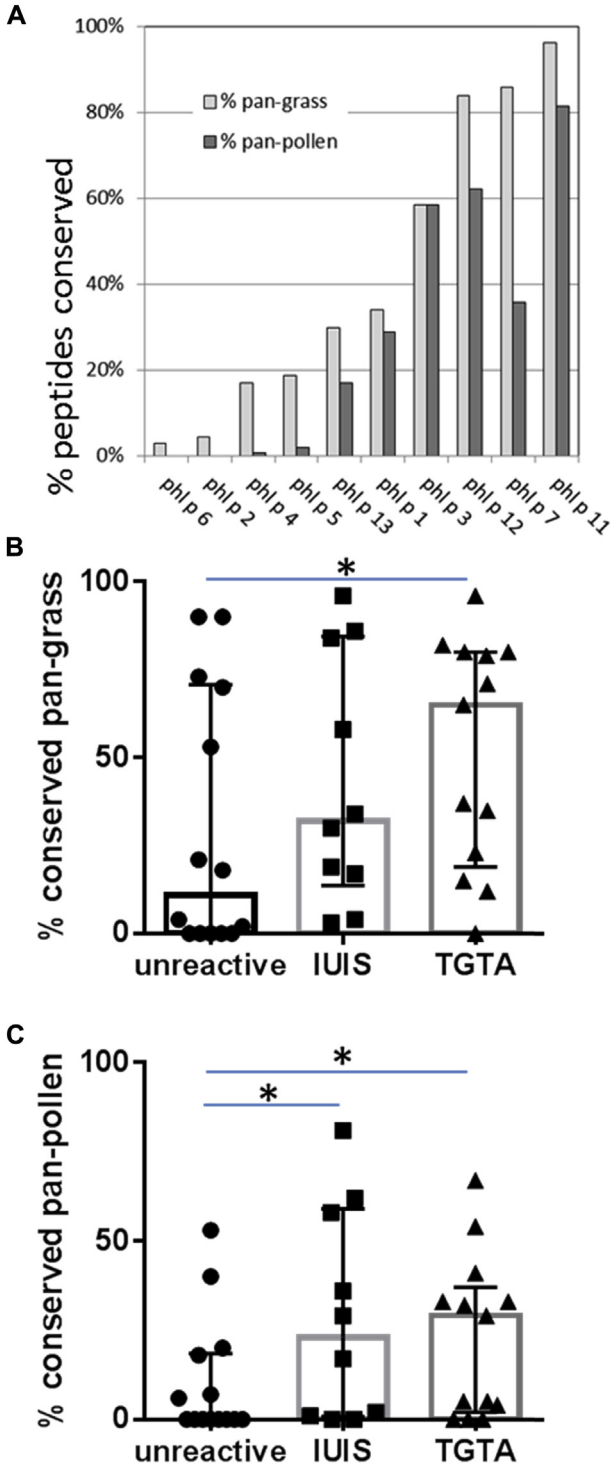

FIG 1. Conservation of timothy grass proteins across other pollens. **A,** Six hundred forty-eight peptides from IgE-reactive *P pratense* allergens listed by IUIS were examined for their conservation in other pollen species. A peptides was considered conserved in another pollen if its transcriptome encoded the peptide or a variant with up to 2 amino acid substitutions. The percentage of peptides conserved in all 4 additional grass transcriptomes (pangrass) is indicated by *light gray columns*, and the percentage of peptides conserved in all 10 pollens (panpollen) is indicated by *dark gray columns*. Antigens are sorted based on pangrass conservation from low to high. **B** and **C,** The pangrass (Fig 1, *B*) and panpollen (Fig 1, *C*) conservation of IUIS allergens was then compared with that of proteins identified in *P pratense* pollen based on a transcriptomic and proteomic analysis. Medians and quantile ranges are indicated by *boxes* and *error bars*. *P pratense* pollen proteins that were not recognized by either T-cell or B-cell responses were less conserved than both IUIS allergens and TGTAs. *Asterisks* indicate statistically significant differences: $P < .05$, 1-tailed Mann-Whitney test.

contained in multiple pollens might be stimulated more frequently than those recognizing epitopes exclusively found in a single source, which could result in dominance of the allergic response to conserved epitopes. To test this hypothesis, we expanded PBMCs from 19 *P pratense*–reactive donors with individual epitopes derived from TGTA or IUIS allergens (these patients are designated as CR in Table E1, and the sequence of the peptides used is described in Table E3 in this article's Online Repository at www.jacionline.org). After 14 days, IL-5 and IFN-γ release was measured to the epitope itself, *P pratense* extract, and extracts from other pollen. T cells from subjects with *P pratense* allergy released much more IL-5 than IFN-γ (see Fig E2 in this article's Online Repository at www.jacionline.org). Responses to extracts and peptide pools were expressed as the relative fraction epitope response to account for variation between patients (Fig 3, *A*). When we correlated cross-reactive responses to different extracts with transcriptomic conservation of the epitope in those extracts, a clear hierarchy was observed. Non–*P pratense* extracts in which the *P pratense* epitope is completely conserved (0 mismatches) showed the highest response, followed by non–*P pratense* extracts with 1 to 2 mismatches, and the lowest responses were observed for non–*P pratense* extracts with 3 or more mismatches. The exact same hierarchy was observed when analyzing peptides from IUIS allergens and TGTA peptides separately (see Fig E3 in this article's Online Repository at www.jacionline.org). Thus *P pratense* epitopes found to be conserved in other pollen transcriptomes were indeed more likely to induce T-cell responses that could cross-recognize different pollen species.

To examine how our results with polyclonal T-cell cultures were reflected on a single-cell level, we generated T-cell clones from PBMCs of patients with *P pratense* allergy (see the Methods section). Clones were *in vitro* expanded with the respective peptide used for T-cell clone generation (Fig 3) and then restimulated with the relevant peptide, an irrelevant control peptide, and extracts, as previously described.[17] In Fig 4, *A*, we show the mean spot-forming cells/million for the polyclonal T-cell line generated by culturing PBMCs with peptide P5. In Fig 4, *B-E*, we show clones generated by using the same peptide with the same patients' PBMCs. Additional clones with another peptide (P7) are shown in Fig E4 in this article's Online Repository at www.jacionline.org. Several of the T-cell clones generated were cross-reactive (Fig 4, *B* and *C*) but with different cross-reactivity patterns. We also found some monospecific T-cell clones that reacted only with *P pratense* or only with birch pollen (Fig 4, *D*). The overall cross-reactivity pattern of clones for peptides P5 and P7 are summarized in Table II and compared with T-cell culture cross-reactivity. We find that the diverse specificity of the clones, when averaged, starts to mimic the reactivity pattern of the T-cell culture, with all strongly cross-reactive extracts (>30% in the T-cell culture) being represented with cross-reactivity also at the clonal level. These data suggest that the cross-reactivity pattern observed at the polyclonal level

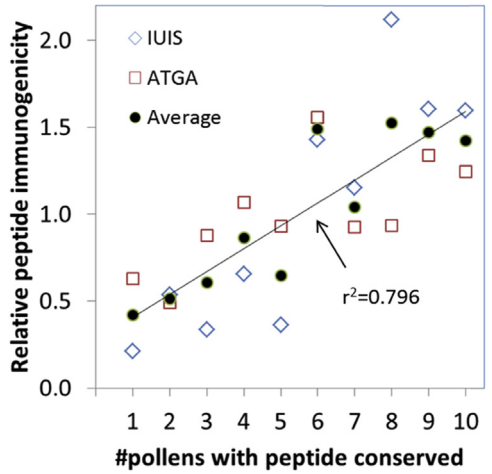

FIG 2. Peptide conservation correlates with immunogenicity. Panels of peptides from timothy grass proteins were previously tested for the ability to induce IL-5 responses in PBMCs from allergic patients after *in vitro* culture with timothy grass extract in 2 separate cohorts. Peptides from each study were separated into sets based on the number of pollen species in which they were conserved *(x-axis)*. For each set, the average frequency of T-cell responses was calculated and normalized to 1.0 for each study *(y-axis)*. Data for peptides derived from IUIS allergens are shown as *blue diamonds*, whereas peptides derived from TGTA antigens are shown as *red boxes*, and averages of the 2 are shown as *black circles*. The line depicts a linear correlation for the averaged data, which is highly significant with an r^2 value of 0.796.

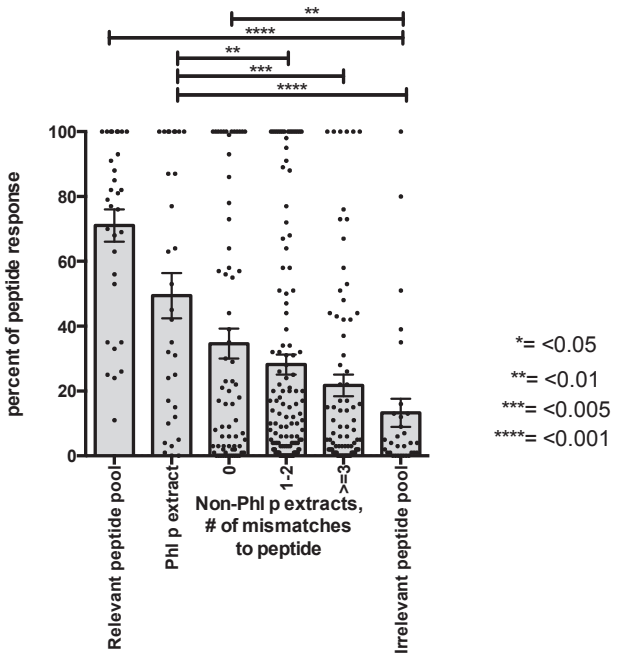

FIG 3. Conservation in the transcriptome predicts peptide cross-reactivity. For each peptide, donors with *P pratense* allergy were selected who reacted to the peptide after expanding PBMCs *in vitro* with *P pratense* extract. PBMCs were stimulated with individual peptides for 14 days, and IL-5 responses were measured by using an ELISpot to (1) the peptide itself, (2) *P pratense* extract, (3) the 9 other extracts for which the transciptomes were sequenced, and (4) peptide pools that did or did not contain the peptide as relevant and irrelevant controls. T-cell cultures that did not induce a robust response of greater than 200 spot-forming cells/million to the peptide itself were excluded (30% of cultures). Reponses to extracts and peptide pools are expressed as the relative fraction of the response to the peptide itself and capped at 100%. Each *gray bar* represents the average response ± SEM. *Asterisks* indicate *P* values of statistical significance, as indicated on the right according to 1-tailed Mann-Whitney tests.

is the result of a heterogeneous population of clones with varying cross-reactivity patterns.

DISCUSSION

The present study has generated the largest panel of pollen transcriptomes to date to perform an unbiased analysis of the effect of sequence conservation on shaping allergen-specific T-cell responses. We show that within and beyond the dominant IgE allergens, there is a substantial fraction of peptides and antigens that are highly conserved across pollens and that this conservation is positively correlated with their likelihood to elicit an immune response. Notably, we could predict the likelihood of *P pratense* peptides to induce a cross-reactive T-cell immune response to other pollens based on their degree of sequence conservation in other pollens.

On the basis of these observations, we hypothesize that cross-reactions at the level of T-cell responses might play a key role in polysensitization to different allergens. Specifically, as outlined in the model presented in Fig 5, we hypothesize that cross-reactive T cells elicited by allergen exposure will (1) be boosted and selectively expanded by exposure to additional allergens containing the conserved epitope and (2) generate help for any B cell specific for an allergen cross-reactive at the T-cell level through a classic antigen bridge–linked T-cell/B-cell help mechanism, regardless of whether the IgE response is cross-reactive.

The specific pattern of the cross-reactive responses and our data in T-cell clones suggest that this is an MHC/T-cell–dependent response and not a bystander effect. Benjaponpitak et al[19] observed a general increase in IL-4 cytokine production in response to a control allergen and tetanus toxoid during the escalation phase of an immunotherapy regimen. They describe this increase as evidence for a T_H2 bystander effect. If the effects in our culture were similarly nonspecific, we would have observed an increased cytokine release to all extracts tested in our culture and tested to the negative control pool. Instead, we observed a pattern that shows a preferential increase in response to extracts that contain conserved sequences. Additionally, our T-cell clone data show that cells derived from one clone can produce cytokines when stimulated with several extracts pointing more toward true cross-reactivity rather than a bystander effect. However, we observed some low-level cross-reactivity against *C dactylon* and *F excelsior* in the T-cell cultures with peptide 7 that were not recapitulated on the clonal level. Therefore we cannot completely exclude that some of the reactivity we observed in our culture system was due to bystander activation.

Allergic cross-reactive responses are relevant in a variety of clinical settings extending beyond pollen and grass allergies. Patients sensitized to shrimp, for example, are very likely to also have positive skin prick test responses to other crustaceans and even other bivalves.[20] Patients with oral allergy syndrome become sensitized to pollen proteins through inhalation and then experience an IgE-mediated allergic reaction against food antigens that are similar in structure.[21] Cross-reactivity of IgE antibodies is determined by means of structural homology of the epitope,[22] and the presence of cross-reactive IgE antibodies is typically reflective of clinical cross-sensitization.[23] The results

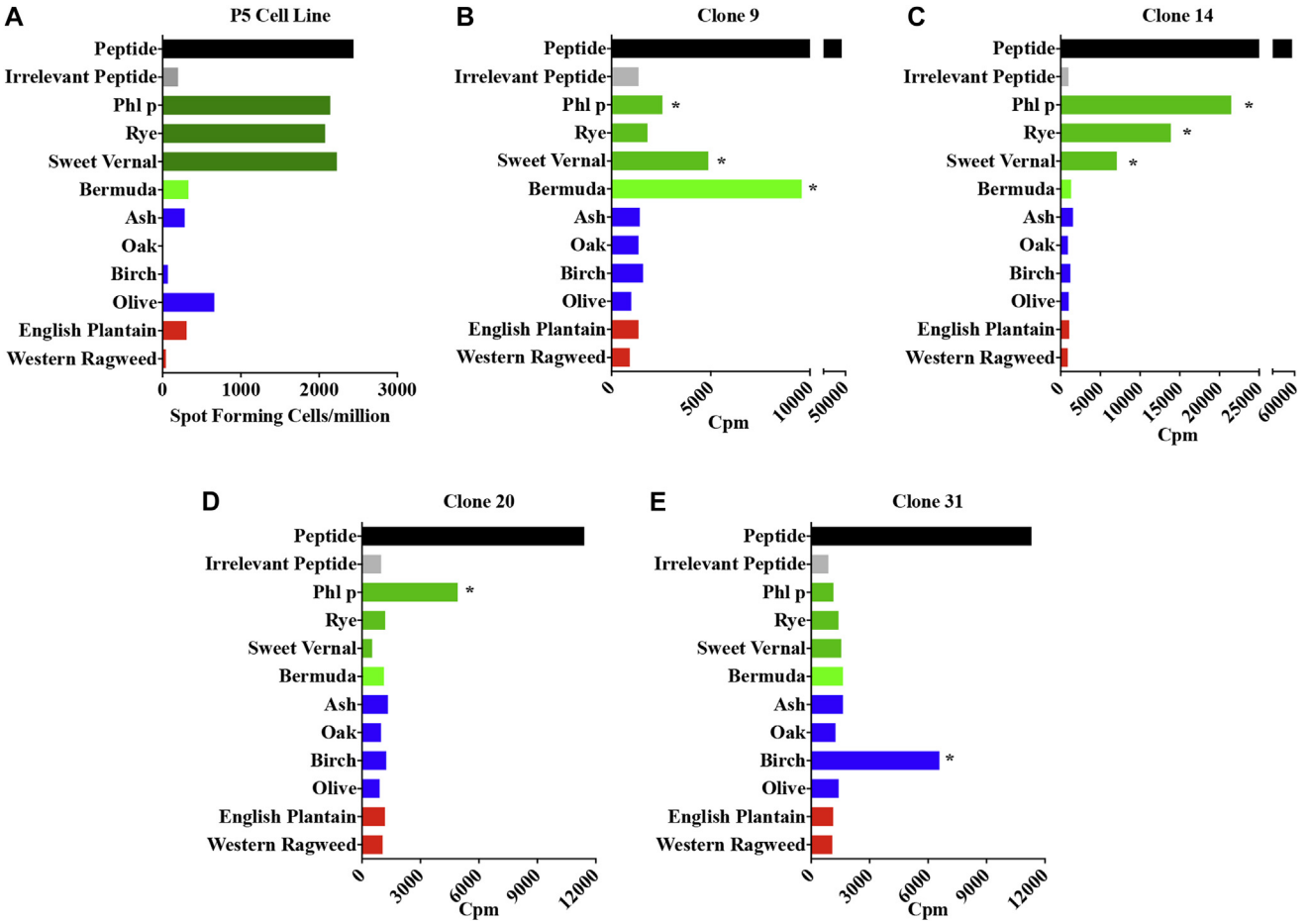

FIG 4. T-cell clone cross-reactivity. Donor 1583 was cultured with peptide P5 and then restimulated with that same (relevant) peptide, an irrelevant peptide control, and peptides from the 4 grasses *(green)*, 4 trees *(blue)*, and 2 weeds *(red)*. **A,** ELISpot data from the P5 cultured cell line. **B-E,** Proliferation of 4 representative T-cell clones derived from the same patient. SI was calculated by dividing cpm of antigen stimulated wells with cpm of unstimulated wells. Conditions with SI>2 were considered positive (*).

TABLE II. Comparison of T-cell clonal data with data from T-cell cultures

	Peptide 5		Peptide 7	
	Clone	T-cell culture	Clone	T-cell culture
Peptide	100%	100%	100%	100%
Irr	0%	8%	0%	22%
Phleum pratense	64%	88%	33%	30%
Lolium perenne	18%	85%	100%	62%
Poa pratensis	55%	91%	100%	45%
Cynodon dactylon	45%	14%	0%	16%
Fraxinus excelsior	0%	12%	0%	27%
Quercus alba	0%	0%	0%	0%
Betula verrucosa	9%	3%	0%	3%
Olea europaea	0%	27%	0%	3%
Ambrosia psilostachya	0%	2%	0%	5%
Plantago lanceolata	0%	13%	0%	13%

Reactivity from 11 clones for peptide 5 and 3 clones from peptide 7 was averaged and compared with the percentage of cross-reactivity generated by using T-cell culture methods.
Irr, Irrelevant control peptide.

of our present study suggest that allergen cross-reactivity at the T-cell level should also be explored in these settings.

In a more general context, the data we present here highlight how the interactions between the immune system and its environment are highly complex and do not fit a simple paradigm in which exposure to a single species of allergen elicits a species-specific response. Rather, our data suggest that T-cell responses are shaped by repeated exposure to related species that carry conserved and cross-reactive linear sequences. This phenomenon is not unique to allergies, and in this respect it is noteworthy that a recent study that examined the level of sequence conservation associated with T-cell cross-reactivity in different Dengue viral strains also derived a maximum of 2 substitutions as a threshold beyond which little or no cross-reactivity was observed.[24] Furthermore, we have recently shown that interspecies conservation plays a key role in shaping the repertoire of human T-cell responses to epitopes conserved in different herpes viruses, such as CMV and EBV,[25] and in different species of the *Mycobacterium* genus.[26]

It must be stressed that the observed correlation between the transcriptomic conservation of epitopes and the observed cross-reactivity is a statistical phenomenon not deterministically predictive for an isolated sequence. T-cell lines expanded with individual *P pratense* peptides that are completely conserved in the transcriptomes of other pollens nevertheless sometimes

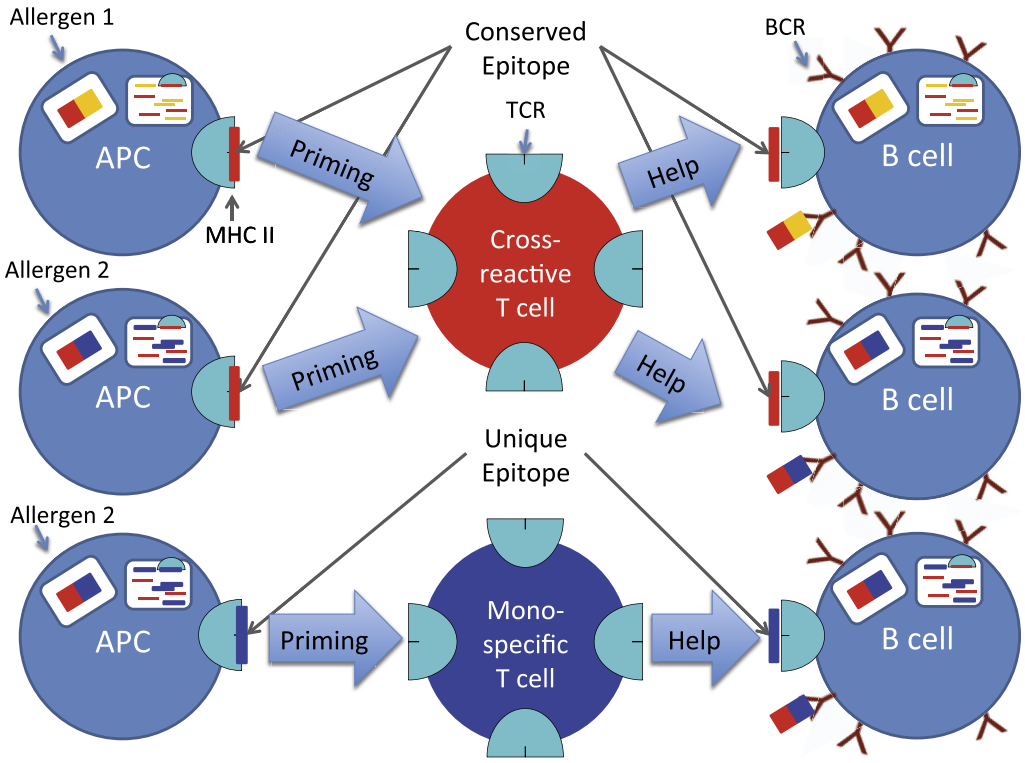

FIG 5. Schematic representation of allergen cross-reactive versus monospecific T-cell epitope recognition. Allergens 1 and 2 contain a conserved T-cell epitope. Therefore cross-reactive T cells that respond to this epitope will be primed more frequently (cross-reactive T-cell in *red, top*) than T cells that are specific for an epitope unique to a single allergen (allergen monospecific T cell in *blue, bottom*). Consequently, cross-reactive T cells can provide T-cell help to B cells that are specific to different allergens and promote antibody production/isotype switching to IgE. Importantly, both B cells that produce cross-reactive IgE and B cells that produce monospecific IgE can receive help from cross-reactive T cells as long as they process and present a conserved T-cell epitope. *APC,* Antigen-presenting cells.

show no reactivity to those pollen extracts, whereas other peptides with many substitutions elicit high responses. This is expected because factors other than the peptide sequence itself, notably abundance of the antigen in which it is embedded, can influence reactivity levels. Also, some substitutions are more likely to disrupt cross-reactivity than others, in particular nonconservative substitutions and those that target anchor residues for MHC binding and key residues for T-cell recognition. Although we can show a high polarization toward IL-5–secreting T cells in allergic subjects, we have not shown that cross-reactivity can be observed with other T_H2 cytokines. However, we chose IL-5 because Schulten et al[8] previously showed that it is representative of multiple T_H2 cytokines after extract and peptide pool stimulation.

Additional limitations to our study must be pointed out. First, donor samples were collected both in and out of the allergy season, and the study was not designed or powered to distinguish whether the exposure pattern at the time of sample collection could further affect the reactivity pattern. Future studies will have to test this potential effect of seasonality.

Second, it is possible that some of the cross-reactivity observed is due to bystander activation. Given that the cross-reactivity pattern we observed at the clonal level mimicked that of the polyclonal T-cell cultures, we do not think that bystander activation is a major effect in our system, but its contribution cannot be ruled out.

The fact that there are large numbers of peptides in *P pratense* that are conserved across pollens and are capable of inducing cross-reactive T-cell responses suggests that there is potential to design cocktails of peptides and antigens that could serve as diagnostic or immunotherapeutic reagents to simultaneously target multiple pollen allergens. It has previously been shown that immunotherapy with *P pratense* extract can induce cross-reactive responses to pollen extracts from other grass species.[2,3] Our work shows that different T-cell epitopes have different potential for eliciting cross-reactive responses and that this potential is predictable based on sequence conservation of the peptides between different pollens. Thus it should be possible to develop cocktails of peptides that specifically elicit cross-reactive immune responses, which could be used for pangrass or even panpollen immunotherapy. Importantly, it is not necessary that these are the same peptides that elicit T_H2 responses in allergic patients, but it is also conceivable to develop specific immunotherapy (SIT) approaches using conserved peptides present across pollens that elicit bystander suppression of the allergic responses.

In terms of diagnostics, the use of extracts or recombinant allergens that are at least partially conserved across different pollen species means that patients for whom there is a single clinically relevant allergen might still be classified as sensitized to multiple pollens because of cross-reactivity in skin tests or RASTs.[27,28]

It will be worth examining whether the use of T-cell epitopes uniquely conserved in different pollen species could more clearly delineate patient sensitization patterns. This could be useful both to modify and improve the SIT preparations[29] or to instruct patients on what allergens to avoid.

In conclusion, we have assembled a large panel of pollen transcriptomes and used these data to examine the interplay of sequence conservation and T-cell cross-reactivity. Thereby we established thresholds of sequence conservation for T-cell epitopes that have potential immunologic relevance and found that many epitopes exist that are highly conserved across grass, tree, and weed pollens. Our findings have potential relevance for the design of next-generation SIT treatments and development of pollen species–specific allergy diagnostics.

Clinical implications: Our data suggest that conserved epitopes that elicit highly immunogenic T cells could delineate patient sensitization patterns, might inform the design of therapeutics for polysensitized allergic subjects, and refine the understanding of monospecific versus cross-reactive allergic responses.

REFERENCES

1. Weber RW. Cross-reactivity of pollen allergens. Curr Allergy Asthma Rep 2004;4:401-8.
2. Martinez-Cocera C, Sastre J, Cimarra M, Quirce S, Fernández-Rivas M, Enríquez-Matas A, et al. Immunotherapy with a *Phleum pratense* allergen extract induces an immune response to a grass-mix allergen extract. J Invest Allergol Clin Immunol 2010;20:13-9.
3. Hejl C, Wurtzen PA, Kleine-Tebbe J, Johansen N, Broge L, Ipsen H. *Phleum pratense* alone is sufficient for allergen-specific immunotherapy against allergy to Pooideae grass pollens. Clin Exp Allergy 2009;39:752-9.
4. Kwok WW, Roti M, Delong JH, Tan V, Wambre E, James EA, et al. Direct ex vivo analysis of allergen-specific CD4+ T cells. J Allergy Clin Immunol 2010;125:1407-9.e1.
5. Palomares O, Yaman G, Azkur AK, Akkoc T, Akdis M, Akdis CA. Role of Treg in immune regulation of allergic diseases. Eur J Immunol 2010;40:1232-40.
6. Wambre E, DeLong JH, James EA, LaFond RE, Robinson D, Kwok WW. Differentiation stage determines pathologic and protective allergen-specific CD4+ T-cell outcomes during specific immunotherapy. J Allergy Clin Immunol 2012;129:544-51, e1-7.
7. Archila LD, DeLong JH, Wambre E, James EA, Robinson DM, Kwok WW. Grass-specific CD4(+) T-cells exhibit varying degrees of cross-reactivity, implications for allergen-specific immunotherapy. Clin Exp Allergy 2014;44:986-98.
8. Schulten V, Greenbaum JA, Hauser M, McKinney DM, Sidney J, Kolla R. Previously undescribed grass pollen antigens are the major inducers of T helper 2 cytokine-producing T cells in allergic individuals. Proc Natl Acad Sci U S A 2013;110:3459-64.
9. Dranitsaris G, Ellis AK. Sublingual or subcutaneous immunotherapy for seasonal allergic rhinitis: an indirect analysis of efficacy, safety and cost. J Eval Clin Pract 2014;20:225-38.
10. Roberts G, Hurley C, Turcanu V, Lack G. Grass pollen immunotherapy as an effective therapy for childhood seasonal allergic asthma. J Allergy Clin Immunol 2006;117:263-8.
11. Schulten V, Tripple V, Sidney J, Greenbaum J, Frazier A, Alam R, et al. Association between specific timothy grass antigens and changes in TH1- and TH2-cell responses following specific immunotherapy. J Allergy Clin Immunol 2014;134:1076-83.
12. Wallner M, Erler A, Hauser M, Klinglmayr E, Gadermaier G, Vogel L, et al. Immunologic characterization of isoforms of Car b 1 and Que a 1, the major hornbeam and oak pollen allergens. Allergy 2009;64:452-60.
13. Hannon G. FASTX-Toolkit. Cold Spring Harbor, NY: 2010.
14. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 2011;17:10-2.
15. 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;29:644-52.
16. Oseroff C, Sidney J, Kotturi MF, Kolla R, Alam R, Broide DH, et al. Molecular determinants of T cell epitope recognition to the common Timothy grass allergen. J Immunol 2010;185:943-55.
17. Messi M, Giacchetto I, Nagata K, Lanzavecchia A, Natoli G, Sallusto F. Memory and flexibility of cytokine gene expression as separable properties of human T(H)1 and T(H)2 lymphocytes. Nat Immunol 2003;4:78-86.
18. Allergome. Available at: www.allergome.org. Accessed February 1, 2016.
19. Benjaponpitak S, Oro A, Maguire P, Marinkovich V, DeKruyff RH, Umetsu DT. The kinetics of change in cytokine production by CD4 T cells during conventional allergen immunotherapy. J Allergy Clin Immunol 1999;103:468-75.
20. Pedrosa M, Boyano-Martinez T, Garcia-Ara C, Quirce S. Shellfish allergy: a comprehensive review. Clin Rev Allergy Immunol 2015;49:203-16.
21. Kondo Y, Urisu A. Oral allergy syndrome. Allergol Int 2009;58:485-91.
22. Jimenez-Lopez JC, Rodriguez-Garcia MI, Alche JD. Analysis of the effects of polymorphism on pollen profilin structural functionality and the generation of conformational, T- and B-cell epitopes. PLoS One 2013;8:e76066.
23. Aalberse RC, Akkerdaas J, van Ree R. Cross-reactivity of IgE antibodies to allergens. Allergy 2001;56:478-90.
24. Weiskopf D, Angelo MA, Sidney J, Peters B, Shresta S, Sette A. Immunodominance changes as a function of the infecting dengue virus serotype and primary versus secondary infection. J Virol 2014;88:11383-94.
25. Chiu C, McCausland M, Sidney J, Duh FM, Rouphael N, Mehta A, et al. Broadly reactive human CD8 T cells that recognize an epitope conserved between VZV, HSV and EBV. PLoS Pathog 2014;10:e1004008.
26. Lindestam Arlehamn CS, Paul S, Mele F, Huang C, Greenbaum JA, Vita R, et al. Immunological consequences of intragenus conservation of *Mycobacterium tuberculosis* T-cell epitopes. Proc Natl Acad Sci U S A 2015;112:E147-55.
27. Weber RW. Patterns of pollen cross-allergenicity. J Allergy Clin Immunol 2003;112:229-40.
28. Marcucci F, Sensi L, Di Cara G, Incorvaia C, Puccinelli P, Scurati S, et al. Which allergen extract for grass pollen immunotherapy? An in vitro study. Immunol Invest 2010;39:635-44.
29. Stringari G, Tripodi S, Caffarelli C, Dondi A, Asero R, Di Rienzo Businco A, et al. The effect of component-resolved diagnosis on specific immunotherapy prescription in children with hay fever. J Allergy Clin Immunol 2014;134:75-81.

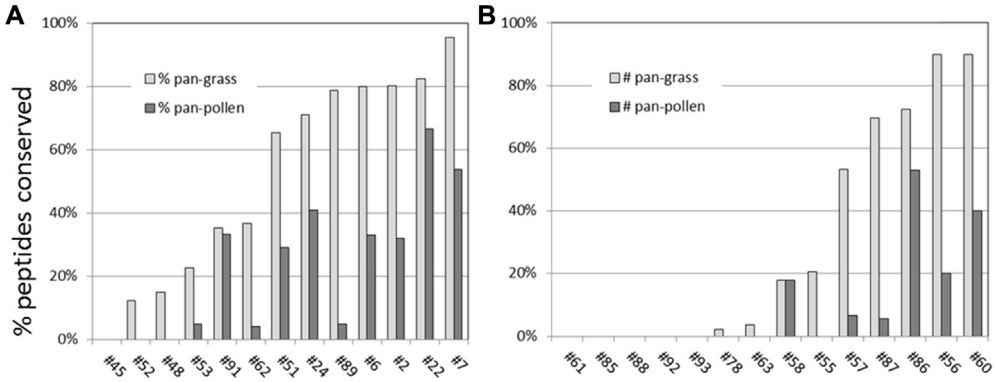

FIG E1. Conservation of peptides in different timothy grass pollen proteins. **A,** Conservation of the 13 TGTAs. **B,** Conservation of timothy grass pollen proteins that were unreactive based on both T-cell and B-cell responses.

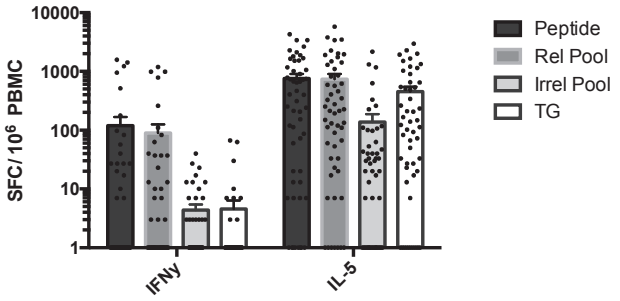

FIG E2. Cytokine release by T cells after 2 weeks of peptide culture. Shown are IFN-γ and IL-5 release by T cells restimulated either with the peptide with which they were originally cultured or restimulated with a relevant peptide pool or an irrelevant peptide pool or timothy grass extract. Each *bar* represents the average spot-forming cells/million *(SFC)* ± SEM.

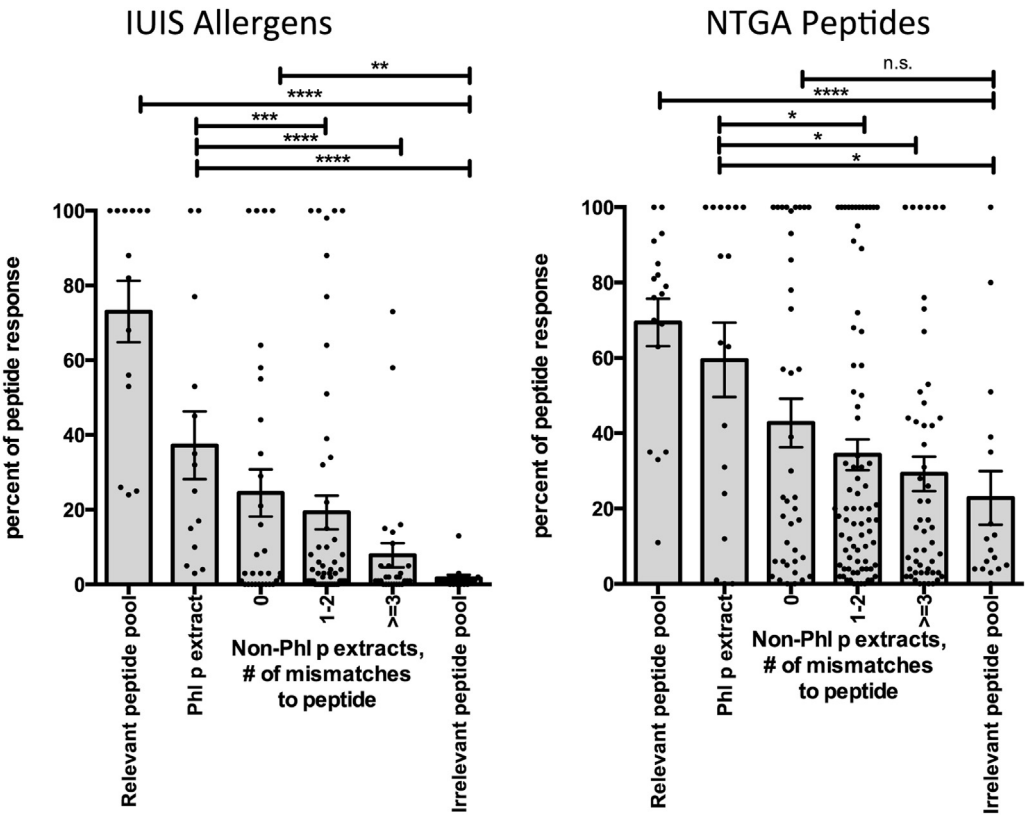

FIG E3. Separate analysis of correlation between cross-reactivity and conservation for IUIS allergen peptides and TGTA peptides. The analysis performed is the same as in Fig 3 but splits out peptides from known IUIS allergens versus TGTA proteins. **P* < .05, ***P* < .01, ****P* < .005, and *****P* < .001. *n.s.*, Not significant.

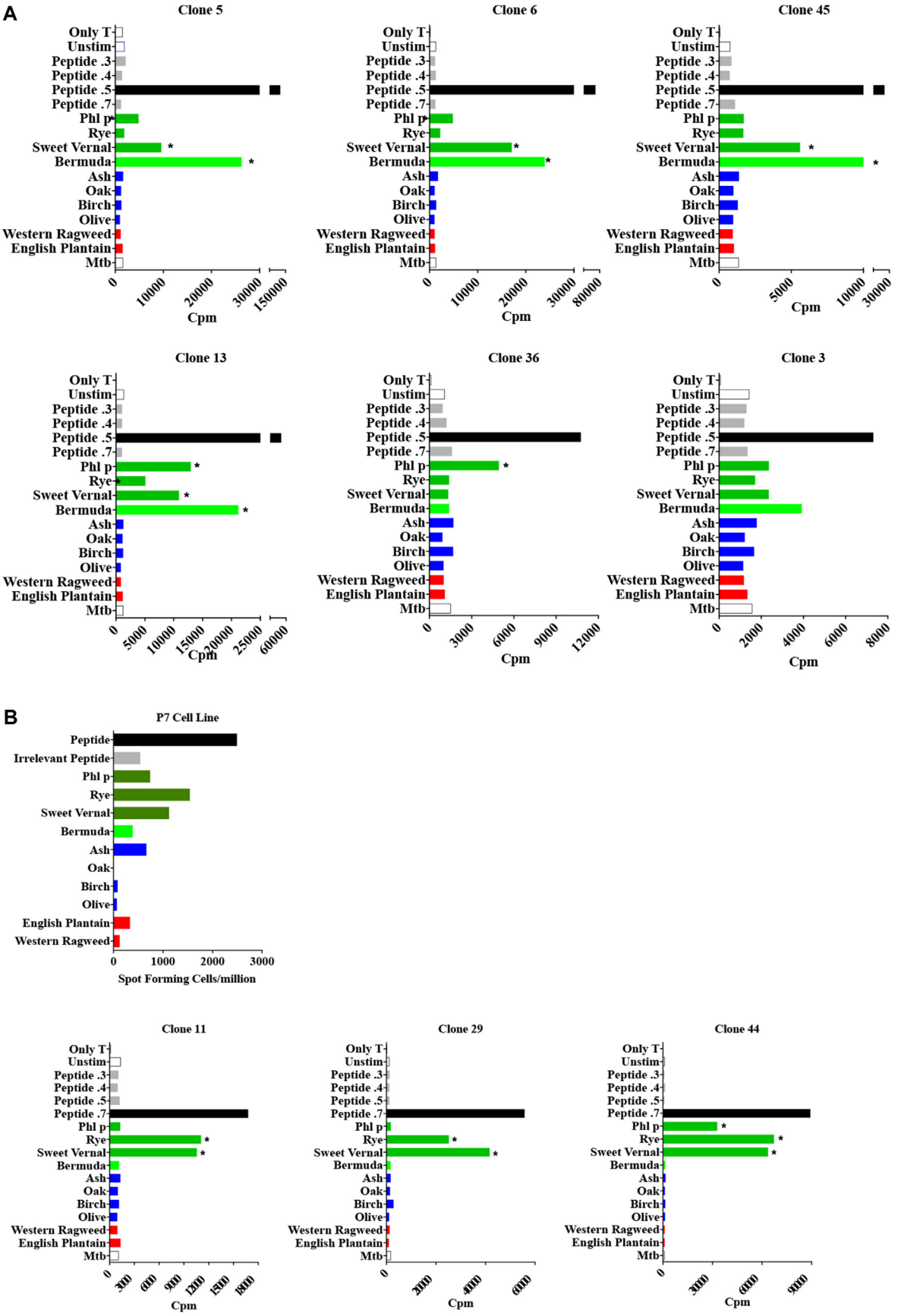

FIG E4. **A,** Additional T-cell clones generated from the same culture (with *P pratense* peptide P5) as shown in Fig 4. **B,** ELISpot data from the cell line *(top)* and proliferation from individual T-cell clones *(bottom)* derived from another culture with *P pratense* peptide P7. SI was calculated by dividing cpm of antigen stimulated wells with cpm of unstimulated wells. Conditions with SI>2 were considered positive (*).

TABLE E1. Donor information

				Grasses				Trees				Weeds			
Patient ID	**Sex**	**Age (y)**	***Phleum pratense***	***Cynodon dactylon***	***Poa pratensis***	***Lolium perenne***	***Anthoxanthum odoratum***	***Fraxinus excelsior***	***Betula verrucosa***	***Quercus alba***	***Olea europaea***	***Plantago lanceolata***	***Ambrosia psilostachya***	**Used in study**	**Season**
D00004	Female	25	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	IUIS	In
D00008	Male	28	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	IUIS	In
D00010	Male	50	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	IUIS	In
D00012	Female	45	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	IUIS	In
D00015	Female	40	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	IUIS	In
D00016	Male	54	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	TGTA	In
D00017	Female	62	Yes	No	No	Yes	No	Yes	No	No	No	No	No	TGTA	In
D00020	Female	40	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	IUIS	In
D00039	Male	31	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	IUIS	Out
D00041	Female	26	Yes	No	Yes	Yes	Yes	No	No	No	No	No	No	IUIS	Out
D00042	Male	42	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	IUIS	Out
D00045	Female	21	Yes	No	No	Yes	Yes	No	No	No	No	No	Yes	IUIS	Out
D00053	Female	50	Yes	No	Yes	Yes	Yes	No	No	No	No	No	No	IUIS	Out
D00056	Male	42	Yes	No	Yes	No	Yes	No	No	No	No	No	No	IUIS	Out
D00061	Female	52	Yes	No	Yes	Yes	Yes	Yes	No	No	No	No	No	IUIS	Out
D00062	Male	27	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	IUIS	Out
D00073	Female	44	Yes	No	Yes	No	Yes	No	Yes	No	No	Yes	No	TGTA	Out
D00078	Male	54	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	CR	Out
D00084	Male	30	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	TGTA, CR	Out
D00089	Female	62	Yes	No	No	Yes	No	Yes	No	No	No	No	No	CR	Out
D00090	Male	42	Yes	Yes	No	No	No	No	No	No	No	No	No	TGTA	Out
D00092	Female	39	Yes	Yes	Yes	No	No	No	No	No	No	No	No	TGTA	Out
D00102	Male	28	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	TGTA	In
D00104	Male	61	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	TGTA, CR	In
D00117	Female	53	Yes	No	Yes	Yes	No	No	No	No	No	No	No	TGTA	Out
U00001	Female	20	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes	IUIS	Out
U00013	Female	22	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	IUIS	Out
U00016	Female	34	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	IUIS	Out
U00022	Female	29	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	IUIS	Out
U00029	Female	22	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	IUIS	Out
U00032	Female	21	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	IUIS	Out
U00039	Male	29	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	IUIS	Out
U00043	Male	33	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	IUIS	Out
U00057	Male	56	Yes	No	Yes	Yes	Yes	No	No	No	No	No	No	IUIS	Out
U00058	Female	22	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	IUIS, TGTA, CR	Out
U00062	Female	22	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	IUIS	Out
U00095	Male	52	Yes	No	Yes	Yes	Yes	No	No	No	No	No	Yes	TGTA	Out
U00098	Male	22	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	TGTA, CR	Out
U00106	Male	19	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TGTA, CR	Out
U00125	Male	23	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	TGTA	Out
U00129	Male	22	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	TGTA	In
U00140	Male	29	Yes	No	Yes	Yes	Yes	No	No	No	No	Yes	No	TGTA	Out
U00147	Male	23	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TGTA, CR	Out
U00150	Female	19	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TGTA, CR	Out
U00151	Male	23	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TGTA, CR	Out
U00153	Female	21	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TGTA	In
1140	Male	21	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1249	Male	45	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1250	Female	47	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1370	Male	35	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1372	Female	43	Yes	NA	NA	NA	NA	NA	No	NA	NA	NA	NA	CR	Out
1373	Female	37	Yes	NA	NA	NA	NA	NA	No	NA	NA	NA	NA	CR	Out
1376	Female	37	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1378	Male	51	Yes	NA	NA	NA	NA	NA	Yes	NA	NA	NA	NA	CR	Out
1379	Female	51	Yes	NA	NA	NA	NA	NA	No	NA	NA	NA	NA	CR	Out
Summary	28 Female/ 27 male	19-62	55/55	33/46	42/46	42/46	38/46	21/46	24/55	20/46	19/46	21/46	27/46		16 In/ 39 out

CR, PBMCs from 19 *P pratense*–reactive donors with individual epitopes derived from TGTA or IUIS allergens; *NA*, skin prick test or RAST data are not available for this allergen in this patient; *No*, patient with a negative skin prick test response or RAST result with a specific IgE level of less than 0.35 kU/L in response to the allergen tested; *Yes*, patient with a positive skin prick test response or RAST result with a specific IgE level of 0.35 kU/L or greater in response to the allergen tested.

TABLE E2. Sequencing and assembly statistics

Common name	Sweet vernal grass	Western ragweed	Bermuda grass	Ash	Rye grass	Olive	English plantain	Kentucky blue grass	Oak
Abbreviation	AO	AP	CD	FE	LP	OE	PL	PP	QA
Reads (millions)	75.4	62.7	66.3	75.6	65.1	73.5	59.0	67.2	63.5
Bases (MB)	7,535	6,271	6,625	7,559	6,511	7,353	5,898	6,717	6,354
Assembled scaffolds	317,874	121,659	112,527	81,401	122,266	74,333	57,102	128,174	54,280
Median scaffold length	544	390	842	722	631	710	696	635	634
Maximum scaffold length	11,515	8,325	14,364	9,838	9,631	8,133	8,090	10,100	14,807

TABLE E3. Sequences for peptides used in cross-reactivity studies

IUIS	TGTA
P1 = EEWEPLTKKGNVWEV	P9 = ELRKTYNLLDAVSRH
P2 = NVWEVKSSKPLVGPF	P10 = AVMLTFDNAGMWNVR
P3 = KPPFSGMTGCGNTPI	P11 = IGSFFYFPSIGMQRT
P4 = STWYGKPTGAGPKDN	P12 = QVYPRSWSAVMLTFD
P5 = GELELQFRRVKCKYP	P13 = AAYLATRGLDVVDAV
P6 = SGIAFGSMAKKGDEQ	P14 = NFTVGRIIELFTAKG
P7 = AFKVAATAANAAPAN	P15 = APSGRIVMELYADVV
P8 = LAKYKANWIEIMRIK	P16 = HYKGSSFHRVIPGFM
	P17 = IIELFTAKGFTVQEM
	P18 = GEVLNALAYDVPIPG
	P19 = NGSQFFLCTAKTAWL
	P20 = VKLRRSSAAQVDGFY
	P21 = VVSRLLIPVPFDPPA
	P22 = GDLYIFESRAICKYA
	P23 = NPMTVFWSKMAQSMT
	P24 = CDASILIDPLSNQSA
	P25 = PRRWLRFCNPELSEI
	P26 = QYAKEIWGITANPVP
	P27 = LVSKLYEVVPGILTE