

Clustering of allergenic pollen on the basis of skin responses of atopic patients by matrix analysis

Kosman E, Eshel A, Keynan N, Waisel Y. Clustering of allergenic pollen on the basis of skin responses of atopic patients by matrix analysis. *Allergy* 1994; 49: 502–507. © Munksgaard 1994.

The responses of 148 atopic patients to some 43 different extracts of allergenic pollen were tested by prick tests. The measure of dissimilarity was introduced and calculated for all pairs of allergens. The investigated allergens were clustered into groups, according to their unbiased greatest similarity, by a matrix-structuring method. Results indicate that subgroups of allergens can be distinguished even within groups of closely related pollen allergens that were believed to be fully cross-reactive. A few cases are demonstrated for various varieties of olives, pecans, date palms, and turf grasses and for some wild chenopods and amaranths. The usefulness of the suggested solution for allergy research and for clinical practice is discussed.

E. Kosman¹, A. Eshel², N. Keynan², and Y. Waisel²

¹ School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, and ² George S. Wise Faculty of Life Sciences, Department of Botany, Tel Aviv University, 69978 Tel Aviv, Israel

Key words: allergens; clustering; cross-reactivity; pollen.

Prof. Yoav Waisel
Department of Botany
Tel Aviv University
Tel Aviv 69978
Israel

Accepted for publication 18 June 1993

Anemophilous plants release large quantities of pollen into the air. Such pollen reaches the epithelium of the human nose, eyes, and respiratory system, and evokes typical allergic responses in atopic patients (12, 13). The same symptoms can be caused by spores of fungi and ferns. The effects of the pollen and of the spores are direct and specific. However, because pollen and spores are carried by winds over long distances and because they appear in a large assortment, it is not easy to identify the specific plant species that causes the allergic response in a patient. This is further complicated by the phenomenon of cross-allergenicity, in which closely related plant species cause similar immunologic responses.

Cross-allergenicity was investigated in detail only for a small number of taxa; e.g., members of the *Amaranthaceae*, *Chenopodiaceae*, *Cupressaceae*, *Oleaceae*, *Poaceae*, and *Urticaceae* families. It can be identified and confirmed by RAST inhibition or by other related methods, but because of the high costs, this can be done for only a small number of responders. In most cases, cross-reactivity can be demonstrated by positive skin responses to extracts of the involved pollen. Nonetheless, the precise causes of cross-allergenicity are only partly understood (1–3, 5, 7, 13, 15). One of the reasons for that

is the difficulty in analyzing the immunoresponses of a large number of patients and discerning the precise spectrum of pollen characteristics that are responsible for cross-allergenicity. This is particularly complicated when large populations of patients are tested for responses to a large number of allergens (14), and in the cases where RAST analysis is either too expensive or not available.

The purpose of the following study was to apply a mathematic tool to the analysis of the results of large sets of data of skin tests, in order to determine the similarity and dissimilarity of the examined allergens as revealed by the responses of the patients to them. The obtained information should improve our understanding of the proximity among various plant allergens and confirm their status regarding cross-reactivity. Such information is of value for a better identification of allergens and for improvement of immunotherapy.

Material and methods

Population responses to various pollen allergens were surveyed in eight rural settlements in Israel. Skin prick tests were performed with commercially available pollen extracts (produced by Dome/

Hollister-Stier and by Center Laboratories) and with locally made pollen extracts. The extracts were prepared from pollen that was collected either from individual trees (date palms) or from cloned plants of well-defined cultivars (pecans, olives, turf grasses, etc.). The pollen grains were sonicated in Coca's solution (4, 8). The extracts were filtered (0.45- μ m mesh), and their protein content was determined according to Lowry et al. (11). The protein content of these extracts was adjusted to the same level as that of the corresponding commercial extracts: 2.0 mg/ml for date palms, 2.0 mg/ml for pecans, 5.0 mg/ml for olives, and 3.0 mg/ml for grasses. The extract was then mixed with glycerol (1:1). All the used extracts were prepared from the same pollen batch. Altogether, 43 allergens were tested with most of the patients. These allergens are common in the outdoor environment of the investigated sites. The list of all the tested allergens is presented in Table 1.

The skin prick tests were made on both arms of the patients, and the developing wheals were classified into five grades: 0 – wheal diameter smaller than 2 mm; 1 – wheal diameter 2–4 mm; 2 – wheal diameter 4–10 mm; 3 – wheal diameter 10–15 mm; and 4 – wheals larger than 15 mm, mostly with developed pseudopodia. Positive and negative controls of saline glycerol and histamine were used.

A total of 148 atopic patients, suffering from allergic rhinitis, allergic asthma, conjunctivitis, or allergic dermatitis, were skin tested with the pollen extracts. Most of the patients had lived in their present locations for several years. Results were

based on the frequency of responses among the patients and on the intensity of the responses (size of the wheal).

The data consisted of patient responses R_i to $A_1 \dots A_n$ allergens, as determined by the skin tests and rated in increasing order: $R_i = i$ for $i = 0 \dots 4$, as described above. The responses of a patient to two allergens were used to define the measure of dissimilarity between those allergens. The arbitrary function ρ of difference between responses, as determined by our data, had to meet the following conditions:

- 1) $\rho(R_i, R_j) = \rho(R_j, R_i) > 0$ when $i \neq j$
- 2) $\rho(R_i, R_i) = 0$
- 3) $\rho(R_i, R_j) < \rho(R_i, R_k)$ for any triple i, j, k under the conditions $i < j < k$.

In this study such function ρ was determined as follows:

- 1) $\rho(R_i, R_j) = |j - i|$ for $i, j = 1, 2, 3, 4$
- 2) $\rho(R_0, R_j) = j + \delta$ for $j > 0$, where $\delta \geq 0$.

Values of $\delta = 0, 1, 2, 3$ were used, but their effect on the results was very small; therefore, a value of $\delta = 0$ was adopted.

The data were arranged in a two-way table, including the responses of every patient to each allergen, in which columns denote the allergens and rows denote the patients. In the present study, not all patients were tested for all the allergens; therefore, there were missing data. Moreover, the data were very heterogeneous in terms of the number of measurements taken for every allergen.

The approach used for clustering allergens according to the patient's responses was based on the proven equivalence of the matrix-structuring method and the well-known "traveling salesman problem" (TSP), which is a method of minimizing accumulative distances (9). This approach enabled the clustering of allergens by solving TSP for the corresponding matrix of dissimilarity among them.

Let K_{rs} be the number of patients who were tested for their reaction to allergens A_r and A_s , and M_{rs} of those who responded positively to one or to both allergens. Two measures of dissimilarity between this pair of allergens can be applied:

$$d_1(A_r, A_s) = \frac{N}{K_{rs}} \sum_{t=1}^{K_{rs}} \rho(R_{rt}, R_{st})$$

and

$$d_2(A_r, A_s) = \frac{N}{M_{rs}} \sum_{t=1}^{M_{rs}} \rho(R_{rt}, R_{st})$$

where R_{rt} and R_{st} are the responses of patient t to allergens A_r and A_s , respectively, and N is the total number of patients tested. The choice between the two measures of difference – d_1 or d_2 – depends on the significance of the fact that a certain patient did

Table 1. List of allergens used in this investigation

1. <i>Phoenix dactylifera</i> Mehola # 1	25. <i>Amaranthus</i> sp.*
2. <i>P. dactylifera</i> Mehola # 2	26. <i>Salsola kali</i> *
3. <i>P. dactylifera</i> Degania # 1	27. <i>Chenopodium album</i> *
4. <i>P. dactylifera</i> Degania # 2	28. <i>Atriplex wrightii</i> *
5. <i>P. dactylifera</i> Degania # 3	29. <i>Cynodon dactylon</i> *
6. <i>P. dactylifera</i> Yotveta # 1	30. <i>Zoysia</i> sp. cv. Emerald
7. <i>P. dactylifera</i> Yotveta # 2	31. <i>Paspalum vaginatum</i>
8. <i>P. dactylifera</i> Yotveta # 3	32. <i>Pennisetum caldestinum</i>
9. <i>P. sylvestris</i> Yotveta # 1	33. <i>Stenotaphrum secundatum</i>
10. <i>P. sylvestris</i> Yotveta # 2	34. <i>Olea europaea</i> cv. Barnea
11. <i>P. sylvestris</i> Yotveta # 3	35. <i>O. europaea</i> cv. Muhassan
12. <i>P. sylvestris</i> Yotveta # 4	36. <i>O. europaea</i> cv. Manzanillo
13. <i>Chamaerops humilis</i>	37. <i>O. europaea</i> cv. Uovo di Piccione
14. <i>P. dactylifera</i> *	38. <i>O. europaea</i> cv. Souri
15. <i>Carya pecan</i> cv. Western Shley	39. <i>O. europaea</i> cv. Kadesh
16. <i>Carya pecan</i> cv. Mohawk	40. <i>O. europaea</i> cv. Koronaiki
17. <i>Carya pecan</i> cv. Choctaw	41. <i>O. europaea</i> cv. Kalamata
18. <i>Carya pecan</i> cv. Wichita	42. <i>O. europaea</i> cv. Merhavia
19. <i>Carya pecan</i> cv. Desirable	43. <i>O. europaea</i> *
20. <i>Carya pecan</i> cv. Apache	
21. <i>Carya pecan</i> cv. Delmas	
22. <i>Carya pecan</i> cv. Burkett	
23. <i>Carya pecan</i> cv. Mahan	
24. <i>Carya pecan</i> *	

* Commercial extract.

not respond to both allergens. With respect to dissimilarity among allergens, this fact is not relevant. In this study, which included allergens of different origins, d_2 was the appropriate distance. Therefore, only cases in which the patient responded to at least one allergen were included in the computation of the measure of dissimilarity between the members of this pair. The rounded values of d_2 were used. The approximation of the TSP solution for the matrix of distance $D_2 = ||d_2(A_r, A_s)||_{r,s=1}^n$ was found. Clusters were generated in accordance with obtained solution.

This method was designed to reveal the presence of clusters of allergens, if such clusters really did exist. The smaller the differences between the allergens, the more uniform the cluster. No a priori assumptions or conditions were imposed on the solution.

The investigation was approved by the institutional ethics review board.

Results

The results illustrate the clustering of the responses of 148 patients to 43 allergens, according to the calculated differences among them (Fig. 1). Degrees of difference among the allergens are presented by the arranged symmetric matrix. In this figure, the actual values of the matrix were grouped and presented by symbols rather than by the actual numbers. The order of columns and lines of the matrix was determined by the TSP solution. The circular order of the data, which was obtained by the TSP solution, was transformed into the linear one, by “cutting” between allergens with greatest value of dissimilarity. The points along the diagonal represent a “zero difference”, i.e., the zero dissimilarity of each allergen to itself. This forms the axis of symmetry of the matrix. The structure of the matrix for any given group of allergens depends upon the degree of dissimilarity between each pair in the group. Therefore, there might be some difference between the analysis which is made for the full set of data and one in which only a specific subgroup is included. This explains why the order of the allergens within specific subgroups (Figs. 2 and 3) may appear to be somewhat different from that of the full set of data (Fig. 1).

The matrix of all the allergens tested, arranged according to the responses of the patients, is presented in Fig. 1. The discrete groups of allergens shown in this figure resulted from an unbiased matrix analysis. It is evident that allergens 15–24, i.e., the pollen extracts of the various cultivars of pecan trees, form a distinct cluster which is clearly different from all the other allergens. In this clustering, the commercial extract of pecan pollen (allergen no. 24) located at the center of the group shows a similar

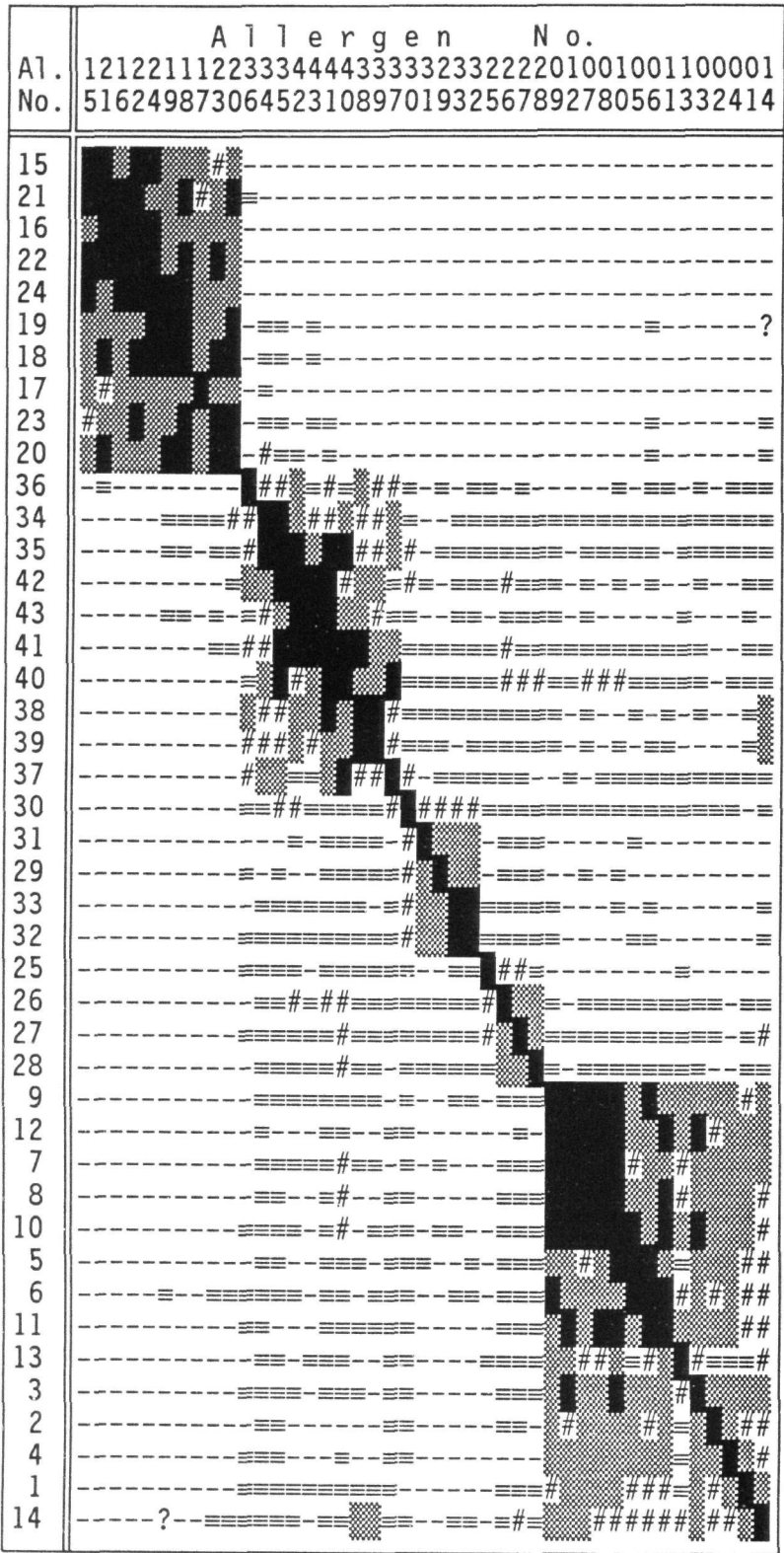


Fig. 1. Matrix of dissimilarities among all allergens investigated. Distances were grouped and replaced by symbols as follows: 0–100 [solid black], 100–150 [cross-hatched], 150–200 [diagonal lines], 200–300 [horizontal lines], > 300 [empty]. See Table 1 for identification of plant sources.

difference from all other pecan extracts, and thus seems to be a good representative of the various cultivars. This is of interest although not critical for pecan allergens, because the responses of the patients to pecan extracts were always very strong and most patients had reacted to the pollen of all the tested cultivars.

The remaining 33 allergens seem to be detached from the pecan. Nonetheless, they form four apparent and distinct clusters (Fig. 1). These include one cluster of the various cultivars of olive (nos. 34–43), one cluster of the grasses (nos. 29–33), one cluster

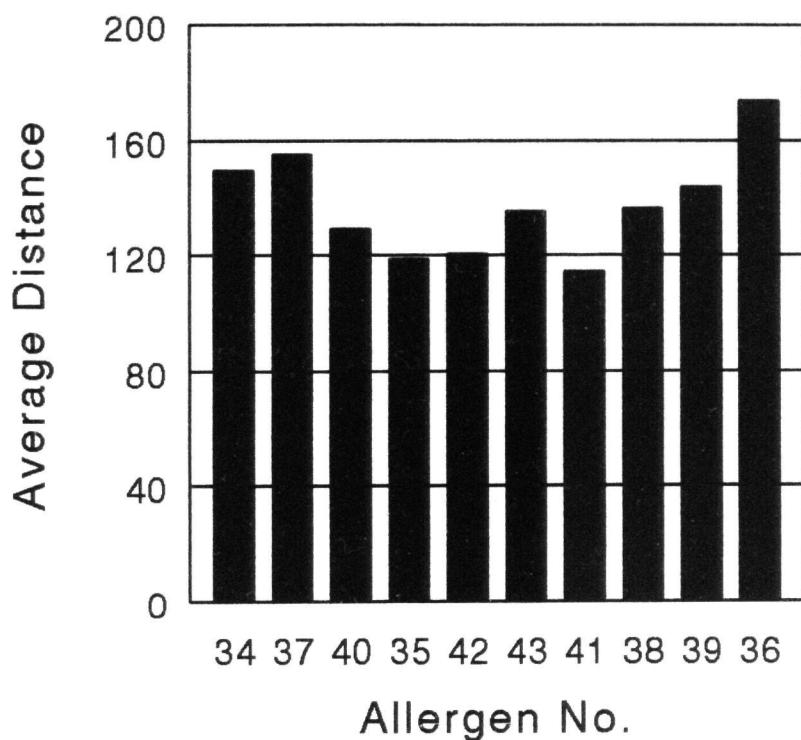


Fig. 2. Average dissimilarities, within group of olive allergens, ordinated by matrix-structuring method. See Table 1 for identification of plant sources.

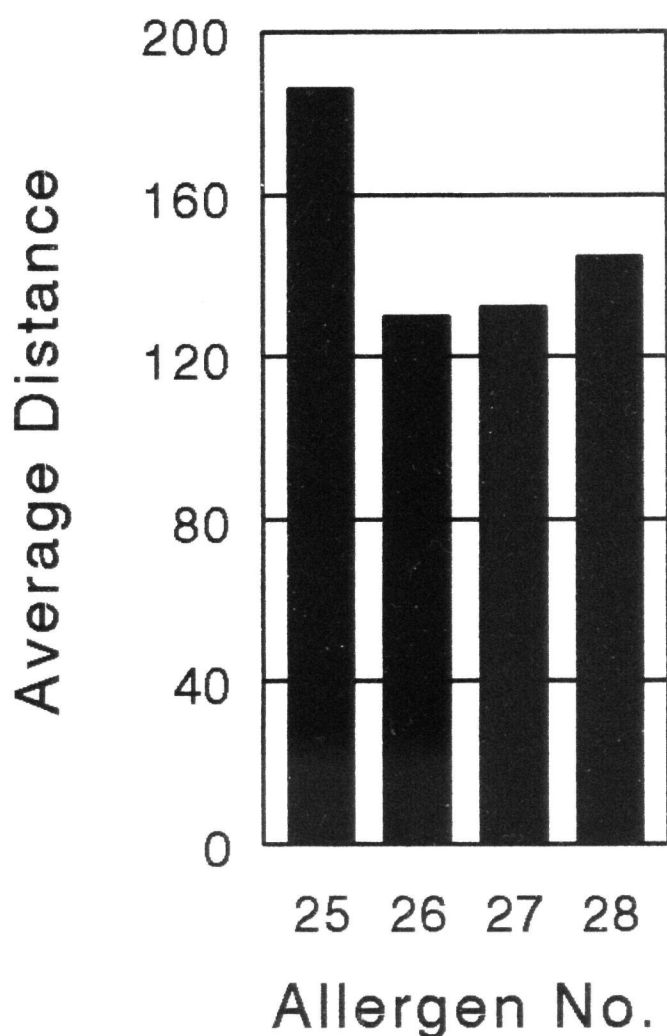


Fig. 3. Average dissimilarities within group of allergens comprised of chenopods and one amaranth, ordinated by matrix-structuring method. See Table 1 for identification of plant sources.

comprising three Chenopodiaceae species and one amaranth species (no. 25), and one cluster of the palms (nos. 1–14).

The allergens of the various cultivars of olive (nos. 34–43) comprise a distinct group. However, the variability within this cluster is rather high. A consequent analysis of the variability within this group (Fig. 2) confirms that the average dissimilarities within the cluster range between 114 units for the Kalamata (no. 41) and 173 for the Manzanillo olive cultivar (no. 36). Apparently, the latter is odd among the tested olives, having the largest average dissimilarity. The commercial extract of olive (no. 43) is located in the center of the group (Figs. 1 and 2), with an average dissimilarity of 135 units. It thus presents an acceptable representative of at least eight of the tested varieties. It is interesting to note that a certain degree of similarity was found between one olive cultivar (cv. Kalamata (no. 41)), the three chenopods, and three of the palms from the Yotveta collection (*Phoenix dactylifera* (nos. 7 and 8) and *Phoenix sylvestris* (no. 10)) (Fig. 1).

Each of the two small clusters displays a clear dissimilarity of one of the constituents. Four of the five investigated grasses (nos. 29, 31–33) are closely related and form a conspicuous subgroup (Table 2). The average dissimilarity among those four ranges between 138 and 156. The fifth grass, *Zoysia* cv. Emerald (no. 30), is distinctly dissimilar, its average difference from the other four allergens being 190 units. Thus, dissimilar species can be found even among the highly cross-reactive grasses.

Another cluster is formed by the responses of the patients to pollen extracts of the three members of the Chenopodiaceae (*Salsola kali* (no. 26), *Chenopodium album* (no. 27), and *Atriplex wrightii* (no. 28)) and to that of *Amaranthus* (no. 25) (Fig. 1). Within this cluster, a subgroup of the three chenopods can be ascertained, with lower average dissimilarity than that of *Amaranthus* (Fig. 3). Thus, in spite of the commonly accepted cross-reactivity between the pollen allergens of those two families (10), a significant dissimilarity between them is revealed by the presented solution. The difference between the two plant families is further corroborated by the similarity of the responses of the patients to one of the olive al-

Table 2. Matrix of dissimilarities among the five grasses. See Table 1 for allergen specifications. Subgroup of the four similar allergens is separated by dashed line

Allergen no.	Allergen no.				
	32	33	29	31	30
32	0	95	135	142	181
33	95	0	133	142	195
29	135	133	0	148	191
31	142	142	148	0	194
30	181	195	191	194	0

lergens (no. 40) and to those of the three chenopods, but not to the allergens of *Amaranthus* (Fig. 1).

Variability was relatively small among the various pollen sources of the palms, except for those of *Chamaerops humilis* (no. 13) (Fig. 1). Yet, variability within the date palms (*P. dactylifera* (nos. 1–8)) was still higher than that of the forest palm (*P. sylvestris* (nos. 9–12)). The responses of the patients to the allergens of *Chamaerops humilis* greatly deviated from the response to all other palm extracts.

The commercial extract of date palm pollen (no. 14) leans toward one end of the entire group. Apparently, the source of this allergen is closer genetically to some of the investigated varieties of *P. dactylifera*. However, it seems to be dissimilar to other examined sources of *P. dactylifera* as well as to the allergens of *P. sylvestris*.

Discussion

The presented data clearly demonstrate that the clustering by means of the TSP solution provides an effective tool for the analysis of: 1) the information obtained from skin tests of large populations of patients and 2) the ranking of dissimilarity among closely related allergens. The TSP method does not replace the precise analysis of antigen similarity, based on RAST. Nevertheless, it is a convenient method for fast analysis of information regarding similarity or dissimilarity in responses of a large population of patients to a large number of allergens.

The results show that analysis of skin test data, which are relatively easy to obtain for a large number of cases, can reveal the existence of distinct groups of allergens within which a certain degree of cross-reactivity may be expected.

One of the questions often asked by allergologists is, what is the relevance of commercial pollen extracts for identification of the causes of allergy in different countries, i.e., for testing patients that are exposed to different plant sources? Is it proper to use pollen extracts, of unknown botanic origin, for skin tests and for immunotherapy? Do the commercially available pollen extracts, whose sources are probably far apart and diverse, give a fair representation of the considered local allergens? It should be stressed again that we are dealing with plant species that are very heterogeneous. This fact is reflected not only by the diversity of the plants' characteristics but also by their different allergic effects. How can one distinguish between relevant and irrelevant test materials? The matrix presented herein yields a sensible, cheap, and convenient solution to such problems. When the commercial extract is located within a relatively homogeneous group, and when its average measure of dissimilarity to the other cross-reactive pollen allergens is relatively small, it should be considered to be

representative. However, in cases where such an extract deviates from the group of local allergens, one should use a more specific extract made from locally collected pollen. The decision of which of the extracts of olives, pecans, date palms, etc. represent the allergenicity of the local varieties in the best way should be left to the allergologists of each country.

A second question that was tackled in the present paper is the question of cross-reactivity. "Full cross-reactivity" among the allergens of plants of the Chenopodiaceae and the Amaranthaceae is frequently reported in the literature (6, 10). Are the pollens of both sources really equally allergenic? The analysis suggested herein has depicted the minor nuances in allergenicity of the tested amaranth and the chenopods. The same is true for the minor differences in allergenicity that exist among various grasses. Turf grasses constitute a major source of allergenic pollen all over the world. Hence, the capability to detect the lower allergenicity of some turf-grass species, e.g., of *Zoysia* cv. Emerald, serves as an excellent example of the practical advantages that are embodied by the suggested solution.

Acknowledgments

The support of the Ministry of Immigrant Absorption, Israel, and of the Colton Foundation to E. Kosman is gratefully acknowledged. We are much indebted to Drs Z. Dolev, M. Finelt, A. Goldberg, S. Kivity, I. Panasoff, D. Nusem, and E. Shahar for participating in this investigation and for performing the skin tests. Sincere thanks are due to Mr Z. Bernstein of Kinneret and to Mr G. Ziv of Yotveta for the supply of date pollen. This investigation was supported by the Committee for Research in Occupational Safety and Health, Ministry of Labour and Social Affairs, Israel.

References

1. BERZOFKY JA, SCHECHTER AN. The concept of cross-reactivity and specificity in immunology. *Mol Immunol* 1981; 18: 751–63.
2. BOUSQUET J, GUERIN B, HEWITT B, LIM S, MICHEL FB. Allergy in the Mediterranean area. III. Cross reactivity among Oleaceae pollens. *Clin Allergy* 1985; 15: 439–48.
3. BOUSQUET J, HEWITT B, GUERIN B, DHIVERT H, MICHEL FB. Allergy in the Mediterranean area. II. Cross reactivity among Urticaceae pollen (*Parietaria* and *Urtica*). *Clin Allergy* 1986; 16: 57–64.
4. COCA H. Studies in hypersensitiveness. V. The preparation of fluids, extracts and solutions for use of the diagnosis and treatment of the allergies with notes on the collection of pollen. *J Immunol* 1922; 7: 163–73.
5. CORBI AL, PELAEZ A, ERRIGO E, CARREIRA J. Cross reactivity between *Parietaria judaica* and *Parietaria officinalis*. *Ann Allergy* 1985; 54: 142–7.
6. D'AMATO G, MULLINS J, NOLARD N, SPIEKSMATHM, WACHTER R. City spore concentrations in the European Economic Community (EEC). VII. Oleaceae (*Fraxinus*, *Ligustrum*, *Olea*). *Clin Allergy* 1988; 18: 541–7.
7. ESCH RE, KLAPPER DG. Isolation and characterization of a major cross-reactive grass group. I. Allergenic determinant. *Mol Immunol* 1989; 26: 557–61.

8. KEYNAN N, GELLER-BERNSTEIN C, WAISEL Y, BEJER-ANO A, SHOMER-ILAN A, TAMIR R. Positive skin tests to pollen extracts of four species of *Pistacia* in Israel. Clin Allergy 1987; 17: 243-9.
9. KOSMAN EG. Matrix structuring. Sov Math Dokl 1990; 41: 152-5.
10. LOMBARDEO M, DUFFORT O, SELLES JG, HERNANDEZ J, CARREIRA J. Cross-reactivity among Chenopodiaceae and Amaranthaceae. Ann Allergy 1985; 54: 430-6.
11. LOWRY OH, ROSEBROUGH NJ, FARR AL, RANDALL RJ. Protein measurement with the Folin phenol reagent. J Biol Chem 1951; 193: 265-75.
12. MIDDLETON E, REED CE, ELLIS EF, eds. Allergy: principles and practice. Toronto: CV Mosby, 1983: 43-74, 271-94.
13. ROITT I, BROSTOFF J, MALE D, eds. Immunology. 2nd ed. London: Gower Medical, 1989.
14. WAISEL Y, MEKORI J. Environmental pollution with pollen and its effect on the frequency of allergic rhinitis and asthma among agricultural workers. First Annual Report. Committee for Research and Prevention in Occupational Safety and Health, Ministry of Labour and Social Affairs, Israel: 1992.
15. WEBER R. Cross reactivity among pollens. Ann Allergy 1981; 46: 208-15.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.