A DATA STORAGE AND RETRIEVAL SYSTEM FOR FORMOSAN POLLEN GRAINS[1,2]

CHANG-FU HSIEH[3] and TSENG-CHIENG HUANG[4]

(Received for publication October 20, 1982 and in revised form February 19, 1983)

Abstract: A palynological data storage and retrieval system is briefly described. It is currently being applied to 900 records of Formosan pollen grains. First, procedures which were used for encoding the morphological features of a pollen grain are outlined. Next, a storage and retrieval system is discussed, which organizes and updates reference pollen records, catalogs them, and enables retrieval of those stored records which share the same features with an unknown specimen. The whole system has been proved useful and effective. The ways in which the variable characters are treated for the sake of obviating encoding errors are highlighted.

INTRODUCTION

The pollen grains, with their diversified and complex characters, are found to possess the potential usefulness in such research fields as aeroallergens, palaeoecology and stratigraphy. They are also of high value in taxonomic and phylogenetic studies. As a result students working in these fields are faced inevitablely with the necessity of identifying pollen grains by comparison with a very large variety of reference preparations, published illustrations, and descriptions. Because of the stenopalynous and eurypalynous natures of the pollen grains, it is usually very difficult to answer the question to which of a number of taxa an unknown pollen should be allotted. In order to serve the needs of researches and to perform a traditional task more efficiently, a palynological information retrieval system has been developed.

There is clearly abundant scope for the profitable application of computers to help with the identification of plant specimens, and several approaches have been explored. These cover construction of diagnostic keys (Bower and Barnett 1971, Hall 1970, Morse 1971, Pankhurst 1971, Watson and Milne 1972, Dallwitz 1974, Češka and Trumpour 1979) and multi-entry keys (Boughey *et al.* 1968, Goodall 1968, Morse 1974, Pankhurst and Aitchison 1975, Johnston 1980), matching or comparison methods (Walker *et al.* 1968), probabilistic methods related to Bayers' theorem (Lapage *et al.* 1971, Baum and Lefkovitch 1972), and taxonomic information systems (Krauss 1973), Despite the central role that studies of identification by computer play in taxonomy and microbiology, relatively few palynologically useful data systems have been realized on other than a pilot or demonstration basis (Walker *et al.* 1968).

Traditionally, identifying an unrecognized pollen grain encountered in honey, atmosphere,

(1) This article is a condensed and improved version of the *Character coding and identification systems for Formosan pollen grains* (Hsieh 1981), a doctoral dissertation under the guidance of the second author.

(2) Portions of this research have been supported by National Science Council (R.O.C.) grants to the second author (NSC-67B-0201-02(1)) and the first author (NSC-68B-0201-02(10)).

(3) 謝長富, Associate Professor, Department of Botany, National Taiwan University.

(4) 黃增泉, Professor, Department of Botany, National Taiwan University.

or fossil-bearing deposits, has largly been achieved with the aid of a key. A key in its usual form, certainly, has some disadvantages. It's requirement is that the characters should be considered in the order specified by the writer of the key. This usually make the user encounter an insuperable difficulty where important distinctions are based upon characters only fragmentally observable. Comparison, either with named specimens or illustrations, is another common method in palynology, even though it suffers from some faults. There is usually a great deal of work involved in checking procedure. This is even more the case as the number of reference pollen species increases to the order of thouands. Moreover, complete and accurate results may not be expected when more than one species possess the same characters or pollen type. These considerations have led some palynologists to recognize the potential advantage of multiple-entry key (Faegri *et al.* 1964:199), in which the choice of a subset of characters and the sequence in which they are used may be at the discretion of the user, in the light of the material available. This kind of key usually exists in the form of edge-punched cards. Each card represents a taxon, and each hole positions a certain charater. In this article we report on the development of a palynological data storage and retrieval system in which the advantages of the multiple-entry key can be combined with the speed and virtually unlimited capacity of computer. The system attempts to store and update a pollen record, catalog it, and enable on-line retrieval of the stored data by any combination of characters. A partially inverted file organization has accordingly been designed specifically for these purposes. The whole system comprises encoding procedure, file construction (including key directory, inverted list, and data files), file update and maintenance, and identifying procedure.

SOURCE MATERIALS

The original reference pollen preparations, upon which was based the ***Pollen Flora of Taiwan*** (Huang 1972), were used for microscopic examinations. The pollen slides were made either by acetolysis method (Erdtman 1952) or by Ikuse (1956) method. Apart from this, relevant information has been gathered from studies by Wodehouse (1935), Erdtman (1952, 1966, 1969), Thomson and Pflug (1953), Kuyl *et al.* (1955), Ikuse (1956), Faegri *et al.* (1964), Kremp (1965), Germeraad and Muller (1970), Huang (1972), and Nilsson *et al.* (1977). The pollen slides and other data are deposited at the Palynology Laboratory of the National Taiwan University. All the poorly representative or imperfectly preserved slides were omitted, and in consequence the input data embodied a total of 900 pollen records, covering 419 genera and 136 families of dicotyledons, 60 genera and 22 families of monocotyledons, and 16 genera and 9 families of gymnosperms. Airborne or widely distributed pollen taxa were included as much as possible.

KINDS OF CHARACTERS

For the purpose of expressing the shapes and basic dimension measurements, three kinds of axes are suggested to the main types of the pollen grains (Fig. 1). In the case of saccate grains, the measurements of the entire grain, sacci, and corpus are given in the order of A-, B-, and C-axes. In grains without apertures or without showing a specific polarity, the A-axis is taken to coincide with the longest axis. In monocolpate or monoporate grains, the equatorial diameter (A-axis) is followed by two polar measurements (B- and C-axes). For grains with two equal apertures at opposite ends, the A-axis lies in a plane perpendicular to the B-axis connecting the centres of both apertures. In radio-symmetric grains the polar diameter (A-axis) is followed by the equatorial diameter (B-axis). In the case of tetrads or polyads, like nonaperturate grains, the A-axis is parallel to the longest axis.

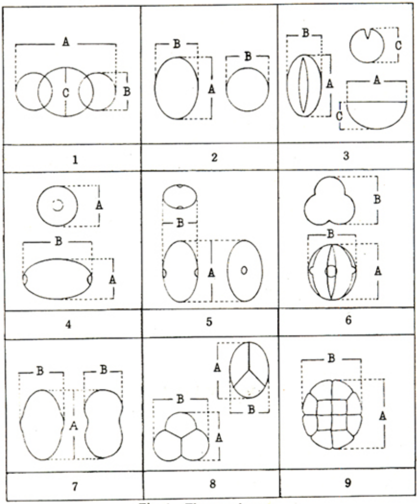

Fig. 1. Three main axes.

The characters recorded for each pollen taxon may be so diversified that they can be distinguished into a number of states or alternatives, therefore, are of great dignostic value to make the process of identification easier. On the other hand, they may be of constant nature and are seldom of use for recognition purposes. Some characters show consistency, while others are subject to wide variation within a taxon, or even a specimen, and consequently of little value as primary diagnosis characters. After examining most of the pollen slides and considering the requirements of the data system, it was decided that the only feasible approach was to group all the characters into two main classes, the key and non-key characters.

1. Key characters

The primary requirement of the key characters is that they should include a number of alternatives, and provide in most cases an excellent means of recognition. They are also necessary for file maintenance and query or retrieval operations. Two kinds of key characters can be delimited, namely discrete key characters and variable key characters.

(1) Discrete key characters

These characters are selected because they are fairly consistent and lacking in plasticity in most of the pollen species. Pollen types (Fig. 2), ornamentation (Fig. 3), sculptures in side view (Fig. 4) are examples for them.

(2) Variable key characters

Sometimes a character is subject to wider variation in a certain pollen species, that means it covers two or more character states and is, consequently, not the most useful for identification. On the other hand, the same character may be quite consistent and merely falls into a specific state in other species, and therefore affords a reliable key for separation. Such kind of character includes shapes in equatorial view (Fig. 5), shapes in polar view (Fig. 6), types of ora (Fig. 7), and size classes.

2. Non-key chayacters

In most situations such characters are quantitative and can be assessed directly by length, number, and size. Some of them may be of qualitative nature, but with few character states. It follows that they are of little value as primary identifying characters, although in some cases they may be of equal or even greater importance in the identification of pollen grains. Based on their properties, the following categories shall be made.

(1) Quantitative non-key characters

Thirteen categories involve measurements and are, therefore, with a conventional indication of range and extremes. Most often their variations are so wide that more than ten individual grains are selected in order to allow an adequate typification of a reference pollen preparation. These characters are enumerated as: A-axis length, B-axis length, C-axis length, exine thickness, excrescence heigth, element width, distance between elements, L-axis length of a pore, W-axis length of a pore, furrow length, furrow width, margo (or annulus) width, and aperture number in one hemisphere.

(2) Qualitative non-key characters

These characters are either of few states or of restricted occurrence, and as a result they can serve to separate species in one group, but are of no value for such a purpose in other groups. There are four of such kinds, namely aperture types, aperture types in side view, reticulum structure, and evenness of exine thickness.

(3) Subsidiary non-key characters

These characters comprise flowering period, altitude, and distribution area. It was considered that they might be as valid for diagnostic purpose as any others, especially when surveys were conducted in such fields as aeropalynology and melitopalynology, and therefore should not be underrated.

CODING SYSTEM

As noted above, the characters used in this system are built by a number of states. The term state has been employed in two distinct senses. It may be phenetically discrete and can be easily recognizable. On the other hand, it may comprise more than one state which is found continuously varying within or between species, and therefore is not easy in arriving at a discrimination. In many cases, it is not uncommon for a species to possess a character with its states varying along two or even more lines. For example, the varying lines of the shape states 24–25–26–27 and 26–36 in Fig. 5 may occur concurrently. This raises some problems when encoding pollen taxa. In dissected situation, it becomes necessary to represent the taxon by five separate pollen recods, each with an independent state of that character. For the sake of simplifying and facilitating, efforts should be made to find the most frequent varying line to which a series of continuous code numbers is allocated,

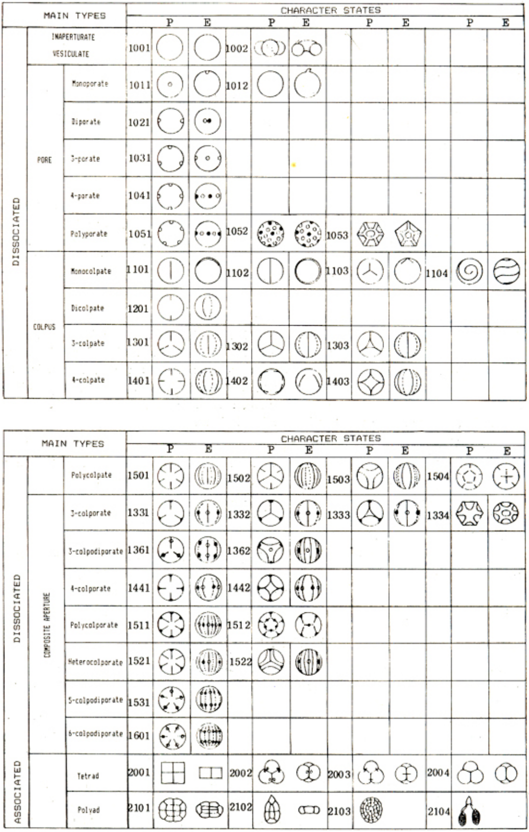

Fig. 2. Pollen types (P pollar view; E equatorial view).

and then leave the other lines in completely discontinuous coding codition. As a result two records are enough to exhibit the whole varying range of the character just exemplified, with one indicated as 24–27 and the other as 36. The coding system explained below is primarily designed for data obtained by optical microscopy. Each state of the key character is exemplified by a taxon in which the typical pattern is shown.

1. Families and species

To facilitate the handling of scientific names, numbers have been assigned to all families, species, and infraspecific taxa. Each family was given a number with three digits. Species and infraspecific taxa were designated numbers of five and one digits respectively. They were all arranged and numbered in accordance with the sequence provided by the *Flora of Taiwan* (Vol. VI, 1979).

2. Discrete key characters

(1) Pollen types (Fig. 2)

There are 20 types comprising 47 states. Each state was given a number with four digits. The first digit denotes dissociation or association state. The second and third digits mostly relate to the numbers of furrows and pores respectively. The last one is merely serial number.

⟨1001⟩ Inaperturate (*Aristolochia*)
⟨1002⟩ Vesiculate (*Pinus*)
⟨1011⟩ Monoporate (*Cynodon*)
⟨1012⟩ Monoporate with papillate projection (*Cryptomeris*)
⟨1021⟩ 2-porate (*Itea*)
⟨1031⟩ 3-porate (*Myrica*)
⟨1041⟩ 4-porate (*Adenophora*)
⟨1051⟩ Stephanoporate (*Alnus*)
⟨1052⟩ Periporate (*Amaranthus*)
⟨1053⟩ 12-porate and dodecahedral (*Telanthera*)
⟨1101⟩ Monocolpate (*Crinum*)
⟨1102⟩ Zonisulculate (*Nymphaea tetragona*)
⟨1103⟩ Trichotomosulcate (*Dianella*)
⟨1104⟩ Spiraperturate (*Berberis*)
⟨1201⟩ 2-colpate (*Dioscorea*)
⟨1301⟩ 3-colpate (*Arabis*)
⟨1302⟩ 3-syncolopate (*Trapa*)
⟨1303⟩ 3-parasyncolpate (*Hyphear*)
⟨1401⟩ 4-colpate (*Impatiens*)
⟨1402⟩ 4-pericolpate
⟨1403⟩ 4-parasyncolpate (*Hyphear*)
⟨1501⟩ Polycolpate (*Mesona*)
⟨1502⟩ Polycolpate (*Schizandra*)
⟨1503⟩ Polycolpate (*Passiflora edulis*)
⟨1504⟩ Polypericolpate (*Mollugo*)
⟨1331⟩ Tricolporate (*Acer*)
⟨1332⟩ Syncolporate (*Eucalyptus robusta*)
⟨1333⟩ Parasyncolporate (*Syzygium formosanum*)
⟨1334⟩ Fenestrate (*Lactuca*)
⟨1361⟩ 3-colpodiporate

⟨1362⟩ 3-syncolporate with two additional pores on poles (*Caesalpinia crista*)
⟨1441⟩ 4-colporate (*Mussaendra*)
⟨1442⟩ 4-parasyncolporate
⟨1511⟩ Polycolporate (*Polygala*)
⟨1512⟩ Pericolporate (*Phyllanthus*)
⟨1521⟩ Heterocolpate (*Melastoma*)
⟨1522⟩ Heterocolpate and parasyncolpate (*Dicliptera*)
⟨1531⟩ 5-colpodiporate (*Breynia*)
⟨1601⟩ 6-colpodiporate (*Chloranthus*)
⟨2001⟩ Tetragonal tetrad (*Philydrum*)
⟨2002⟩ Tetrahedral tetrad formed by 3-porate grains (*Ludwigia*)
⟨2003⟩ Tetrahedralal tetrad formed by 3-colp(or)ate grains (*Rhododendron*)
⟨2004⟩ Tetrahedral tetrad (*Drosera*)
⟨2101⟩ Polyad (*Acacia*)
⟨2102⟩ Polyad (*Calliandra*)
⟨2103⟩ Polyad (*Spiranthes*)
⟨2104⟩ Polyad (*Asclepiadaceae*)

(2) Ornamentation (Fig. 3)

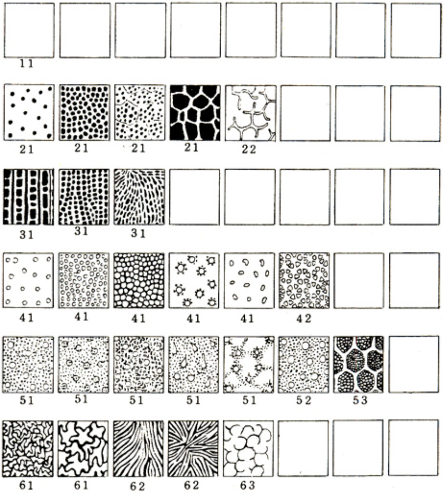

Fig. 3. Ornamentation.

This deals with the sculpture of the exine. Twelve states were selected, each of which was allocated a number of two digits.

〈11〉 Smooth or obscure (*Begonia*)
〈21〉 Foveolate to reticulate (*Aralia, Aristolochia, Phaseolus*)
〈22〉 Retiulate, with interrupted muri (*Boerlagiodendron*)
〈31〉 Striato-reticulate (*Goldfussia, Rhus, Bridelia*)
〈41〉 Granulate (*Clerodendron cyrtophyllum, Ilex, Dicliptera, Parachampionella rankanensis, Deeringia*)
〈42〉 Croton pattern (*Croton*)
〈51〉 Two kinds of excrescences different in size (*Fatsia, Melothria, Asrum, Ipomoea gracilis, I. acuminata*)
〈52〉 Grana and depressions appearing together (*Annona montana*)
〈53〉 Grana appearing within lumina (*Polygonum*)
〈61〉 Rugulate (*Actinidia, Corydalis*)
〈62〉 Striate (*Fragaria, Sedum*)
〈63〉 Narrow ditches connected to form a interrupted net-like arrangement (*Ludwigia*)

(3) Sculptures in side view (Fig. 4)

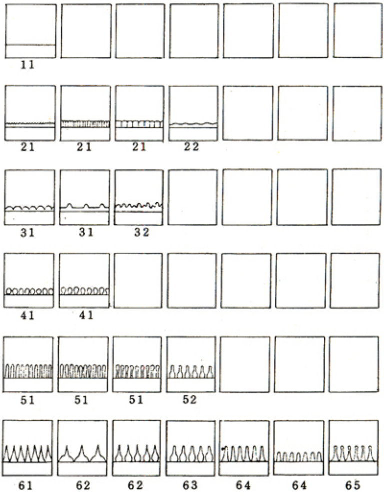

Fig. 4. Sculptures in side view.

The sculptures used here merely comprise the elements which contribute directly the external geometrical features without reference to their internal construction. Two digits were assigned to each of the thirteen states.

⟨11⟩ Psilate (*Zea*)
⟨21⟩ Scabrate or acinate (*Hemigraphis*, *Euphorbia*, *Annona*)
⟨22⟩ Undulate (*Pogostemon*)
⟨31⟩ Verrucate or tuberose (*Gentiana scabrida*, *Goldfussia*)
⟨32⟩ Irregularly verrucate (*Blechnum*)
⟨41⟩ Gemmate (*Stephania*, *Ilex*)
⟨51⟩ Columnate or clavate (*Cardamine*, *Ilex goshiensis*, *Euonymus echinatus*)
⟨52⟩ Columnate, with broad base (*Stictocardia*)
⟨61⟩ Echinate (*Blumea*)
⟨62⟩ Echinate, with broad base and sharp tip (*Abutilon*, *Malvastrum*)
⟨63⟩ Echinate, with broad base and blunt tip (*Ipomoea acuminata*)
⟨64⟩ Echinate, with blunt tip (*Hibiscus taiwanensis*)
⟨65⟩ Echinate, with enlarged tip (*Hibiscus tiliaceus*)

3. Variable key characters

(1) Shapes in equatorial view (Fig. 5)

The shapes of the grains have been instituted on the basis of two criteria, each of which was represented by one digit. The first digit denotes the edge conditions: ⟨2⟩ convex; ⟨3⟩ straight; ⟨4⟩ concave in equatorial region; ⟨5⟩ concave in polar region. The second digit refers to the class of the A-axis/B-axis length ratio (R): ⟨1⟩ $R<0.50$ (peroblate); ⟨2⟩ $0.50<R<0.75$ (oblate); ⟨3⟩ $0.75<R<0.90$ (suboblate); ⟨4⟩ $0.90<R<1.10$ (spheroidal and subspheroidal); ⟨5⟩ $1.10<R<1.33$ (subprolate); ⟨6⟩ $1.33<R<2.0$ (prolate);

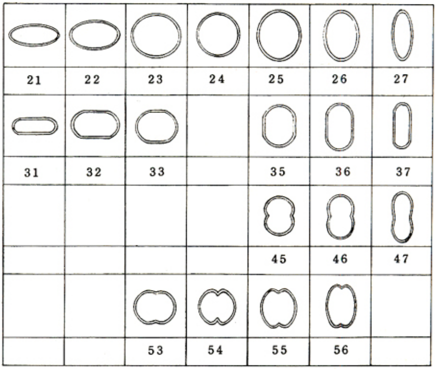

Fig. 5. Shapes in equatorial view.

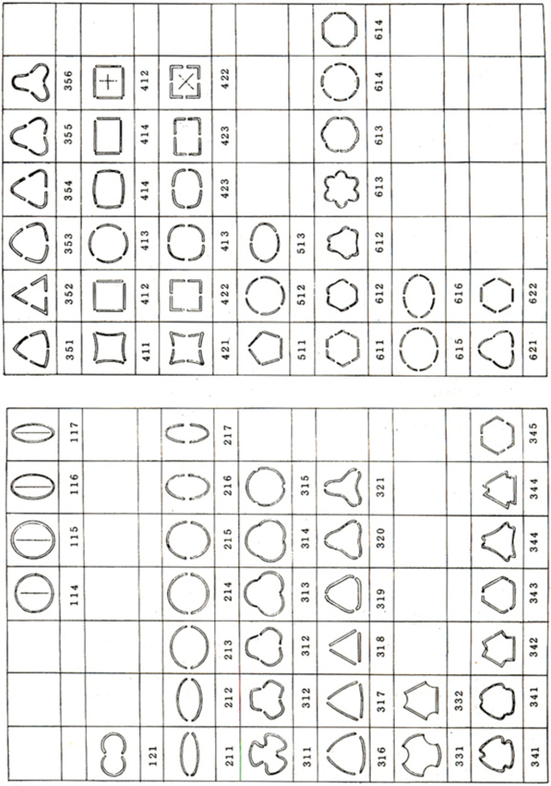

Fig. 6. Shapes in polar view.

⟨7⟩ $2.0<R$ (perprolate). The representative examples are as follows:

⟨21⟩ *Bifaria*
⟨22⟩ *Eucalyptus*
⟨23⟩ *Alnus*
⟨24⟩ *Cordia*
⟨25⟩ *Tetrapanax*
⟨26⟩ *Grewia*
⟨27⟩ *Sanicula*
⟨32⟩ *Impatiens*
⟨33⟩ *Basella*
⟨35⟩ *Justicia quadrifaria*
⟨36⟩ *Castanopsis hystrix*
⟨37⟩ *Foeniculum*
⟨46⟩ *Cynoglossum*
⟨47⟩ *Foeniculum*
⟨53⟩ *Aeschynomene*
⟨54⟩ *Aeschynomene*
⟨56⟩ *Arabis*

(2) Shapes in polar view (Fig. 6)

It is always the case that identifying an unknown pollen grain should be carried out merely based on those limited characters observed in polar view. Therefore an elaborated classifying system should be suggested in order to reinforce their diagnostic value. In all, 55 states were selected, each of them was coded by a number of three digits. The same criterion (the second digit) used for the shape classes described above was applied here to describe the shapes of the grains with one furrow or two apertures. In most cases, the first digit indicates the numbers of the pores, furrows, or composite apertures, although few exceptions are made to some classes. Since it is quite difficult to find an appropriate term to match each of the states, only representative taxa, if present, are enumerated here.

⟨114⟩ *Cynodon*
⟨115⟩ *Crinum*
⟨116⟩ *Aneilema*
⟨121⟩ *Stauntonia*
⟨211⟩ *Itea*
⟨213⟩ *Alyxia*
⟨214⟩ *Broussonetia*
⟨215⟩ *Justicia*
⟨311⟩ *Solanum nigrum*
⟨312⟩ *Bryophyllum*
⟨313⟩ *Meliosma*
⟨314⟩ *Castanopsis*
⟨315⟩ *Fatsia*
⟨316⟩ *Fragaria*
⟨317⟩ *Potentilla*
⟨318⟩ *Trapa*
⟨319⟩ *Euphoria*
⟨320⟩ *Eucalyptus robusta*
⟨321⟩ *Bifaria*
⟨331⟩ *Aeschynomene*
⟨341⟩ *Sloanea, Parnassia*
⟨342⟩ *Vitis*
⟨343⟩ *Aralia*
⟨344⟩ *Koelreuteria, Leea*
⟨345⟩ *Ixeris*
⟨353⟩ *Apium*
⟨354⟩ *Angelica*
⟨355⟩ *Bombax*
⟨411⟩ *Symplocos paniculata*
⟨412⟩ *Claoxylon, Hiptage*
⟨413⟩ *Carpinus, Adenophora*
⟨414⟩ *Impatiens*
⟨422⟩ *Tabemaemontona, Basella*
⟨423⟩ *Nerium*
⟨511⟩ *Alnus*
⟨512⟩ *Mucuna*
⟨611⟩ *Barthea*
⟨612⟩ *Ehretia, Terminalia*
⟨613⟩ *Messerschmidia, Breynia*
⟨614⟩ *Achyranthes, Alternanthera*
⟨615⟩ *Mosla*
⟨616⟩ *Hyptis*
⟨621⟩ *Melastoma*
⟨622⟩ *Prunella*

(3) Types of ora (Fig. 7)

Each state was given a number with two digits. The first one indicates the shapes of the ora: ⟨1⟩ ora circumscribed by the colpi; ⟨2⟩ ora transversely extended (lalongate); ⟨3⟩ lalongate ora with both ends pointed. The representative example of each state is as follows:

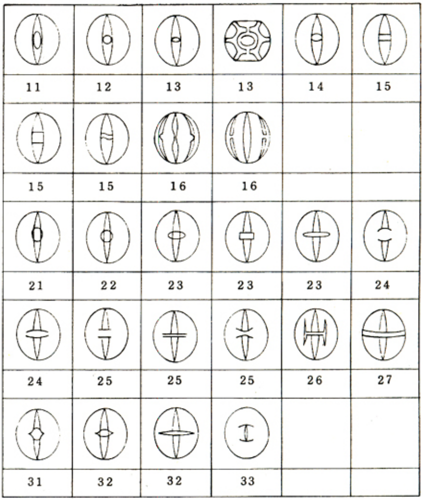

Fig. 7. Types of ora.

⟨11⟩ *Bauhinia*
⟨12⟩ *Leontopodium*
⟨13⟩ *Viburnum, Lactuca*
⟨14⟩ *Rubus shinkoensis*
⟨15⟩ *Mazus, Ilex*
⟨16⟩ *Actinidia, Hemiboea*
⟨21⟩ *Melia*
⟨22⟩ *Fagus*
⟨23⟩ *Ehretia, Rhaphiolepis*
⟨24⟩ *Cordia, Diospyros*
⟨25⟩ *Hedera, Tetrapanax, Lumnitzera*
⟨26⟩ *Schefflera*
⟨27⟩ *Polygala*
⟨31⟩ *Lagenophora*
⟨32⟩ *Gynura, Ixora*
⟨33⟩ *Lonicera*

(4) Size classes

The following size classes, based on the length of the B-axis, have been suggested: ⟨1⟩ <10; ⟨2⟩ 11-20; ⟨3⟩ 21-30; ⟨4⟩ 31-40; ⟨5⟩ 41-50; ⟨6⟩ 51-60; ⟨7⟩ 61-70; ⟨8⟩ 71-80; ⟨9⟩ >81.

4. Quantitative non-key characters

These are concerned with measurements and are, therefore, with an indication of their minimum and maximum range. The unit of the measurement is micron. Sometimes it should be multiplied by 10, thus eliminating any decimal point. A number of two or three digits was assigned to each of the characters.

(1) A-axis length

Of three digits.

(2) B-axis length

Of three digits.

(3) C-axis length

Of two digits.

(4) Exine thickness

Including sexine and nexine; multiplied by 10 to give a number up to three digits. If it is unevenly thickened, then measurement should be taken in the equatorial region.

(5) Excrescence height

Only concerned with the height of the outmost layer of the sculpturing elements, when double or multiple layers are present. The height should be multiplied by 10. The number is of three digits the most. Generally we measure those of the most prevailing ones, while exclude the extremes.

(6) Element width

Of two digits at most, after being multiplied by 10. When two kinds of the elements are mixed up, we measure the larger one. For those elements of elongate shape, we measure their greatest width.

(7) Distance between elements

Multiplied by 10, the largest number is of two digits. The distance is recorded as the measurement of the most prevailing ones. When two kinds of the elements are present, we base the measurement on the larger one.

(8) L-axis length of a pore

Should be multiplied by 10, the largest number is of three digits. The L-axis of a pore is its shortest diameter. For ora the L-axes are parallel to the longitudinal axes of the forrows.

(9) W-axis length of a pore

The same as L-axis, except that the measurement should be taken from the longest diameter of a pore or the axes perpendiculer to the L-axis of the ora.

(10) Furrow length

Comprising furrows of colpate, colporate, hetero- and colpodiporate grains. The largest number is of two digits.

(11) Furrow width

Multiplied by 10, giving a number of three digits the most.

(12) Margo (or annulus) width

When a margo and annulus appear together, the width relates to margo only, Should be multiplied by 10, the largest number is of two digits.

(13) Aperture number in one hemisphere

Of two digits, dealing with the pore (furrow) numbers of polyporate, polycolpate, or polycolporate grains.

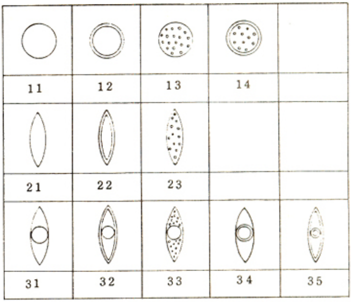

Fig. 8. Aperture types.

5. Qualitative non-key characters

(1) Aperture types (Fig. 8)

This character describes the manner of the aperture being a pore, furrow, or a composite one; the edges of apertures, whether they are thickened or not; and the appearance or lack of the sculpturing elements distributed over the surface of the aperture membrane. Twelve states were chosen, each of which was coded by a number of two digits. The typical examples are as follows:

〈11〉 *Buxus* 〈12〉 *Carpinus*
〈13〉 *Ipomoea* 〈14〉 *Stellaria*
〈21〉 *Cuscuta* 〈22〉 *Dysosma*
〈23〉 *Corydalis* 〈31〉 *Acer*
〈32〉 *Mangifera* 〈33〉 *Cayratia*
〈34〉 *Cephalanthes* 〈35〉 *Elaegnus*

(2) Aperture types in side view (Fig. 9)

The types chosen were from Thomson and Pflug (1953, pp. 34–35), which will not be

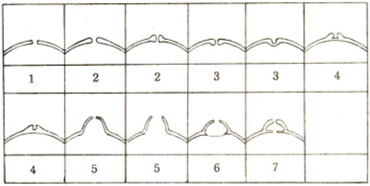

Fig. 9. Aperture types in side view.

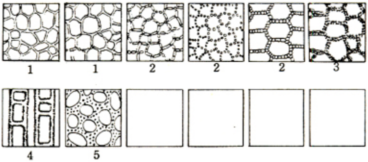

Fig. 10. Reticulum structure.

further discussed here. Only examples are given as follows:

〈1〉 *Boehmeria* 〈2〉 *Nerium*
〈3〉 *Adenophora* 〈4〉 *Pyrola*
〈5〉 *Ludwigia* 〈6〉 *Myrica*
〈7〉 *Alnus*

(3) Reticulum structure (Fig. 10)

This is a character concerned with the patterns of the ridges. Five states were definec

〈1〉 Reticula with continuous ridges (*Phaseolus*, *Viburnum*)
〈2〉 Reticula with segmented or rod-like ridges (*Limonium*, *Gentiana*, *Philoxerus*)
〈3〉 Reticula with multiple rod-like ridges (*Polygonum*)
〈4〉 With some depressions on the margins of the ridges (*Goldfussia*)
〈5〉 Reticula with spinose or granular ridges (*Alternanthera*)

(4) Evenness of exine thickness

〈1〉 Evenly thickened (*Quercus*)
〈2〉 Thickened in equatorial region (*Arachi*
〈3〉 Thickened in polar region (*Aconitum*)

6. Subsidiary non-key characters

(1) Flowering period

This indicates the month period.

(2) Altitude

Using meter as an unit. Four digits are needed.

(3) Distribution area

This character only refers to those topographically restricted or highly localized species. For practical convience, we subdivided Taiwan Region into ten areas based upon its topographical features, and each of them was coded by a number of one or two digits.

〈1〉 Northern area 〈2〉 North-western area
〈3〉 North-eastern area 〈4〉 Western area
〈5〉 Central area 〈6〉 Eastern area
〈7〉 South-western area 〈8〉 Southern area
〈9〉 Lanyu and Lutao Islands 〈10〉 Penghu Island

Any species whose distribution areas fall into one or two of the above designed categories should be assigned with those area codes.

POLLEN IDENTIFICATION SYSTEM

1. File structure (Fig. 11)

The principal aims of this system are to identify unrecognized pollen grains and operate update transaction within a very short time. To meet these requirements, partially inverted file system (Lefkovitz 1972), common in computerized information system, is accordingly adopted, in which all pollen records containing a given character state will have their addresses (sequence numbers when records are stored in the disk) listed in a monotonic sequence within a variable length record, the address of which is an elemnt of the key directory for that state. That is three linked files, namely pollen data file, list file, and key directory file, are contained in this identification system.

(1) Pollen data file

Pollen data file is composed of reference pollen records, each of which contains identity numbers of family, species and infraspecific taxon; coded key and non-key characters; scientific name, major localities, ecological and other miscellaneous information in alphanumeric form. As all records are written onto disk, sequential address is assigned to each of them.

(2) List file

As mentioned above, a list is a set of addresses linked to certain corresponding records in the pollen data file, which possess the same character state. Therefore, each state of the key characters may produce a list. The length of each list is variable, and cannot be stored as a single record. As a result, one or more blocks of a specified length are assigned in the disk for each list. In other words, each list may occupy one or several blocks. The same as the pollen data file, addresses are assigned in sequence to all of the blocks. At the end of each block, a reserve space should be left so that a link address indicating the next linked block, if more than one are needed, can be stored.

(3) Key directory file

Key directory is a triplet containing character state, head address of list, and list length. By means of the key directory, we can retrieve all addresses of pollen records which contain a specific character state.

2. File construction

(1) Input form

The input of the reference pollen records can be carried out either by punching on standard 80-column cards or by seting up on a display console, with a single digit, letter or other symbol occupying each column or field. The items involved in each input are listed in Table 1. Since size classes can be generated automatically from B-axis length, it is not necessary to include them here. More frequent are instances when certain characters or states are lacking or not detectable, with the result that comparable columns should be remained as blanks. For a character with lower and upper limits, it is always the case that they should be of the same value if no variation occurs. Should a taxon be variable for a discrete key character or show two trends of variation for a variable key character, then it would become necessary to represent the taxon by two or more records.

(2) Construction procedure

The procedure is broken into seven steps:

〈1〉 Storing all the input pollen records onto a disk area, each of them possesses an address. This forms the pollen data file.

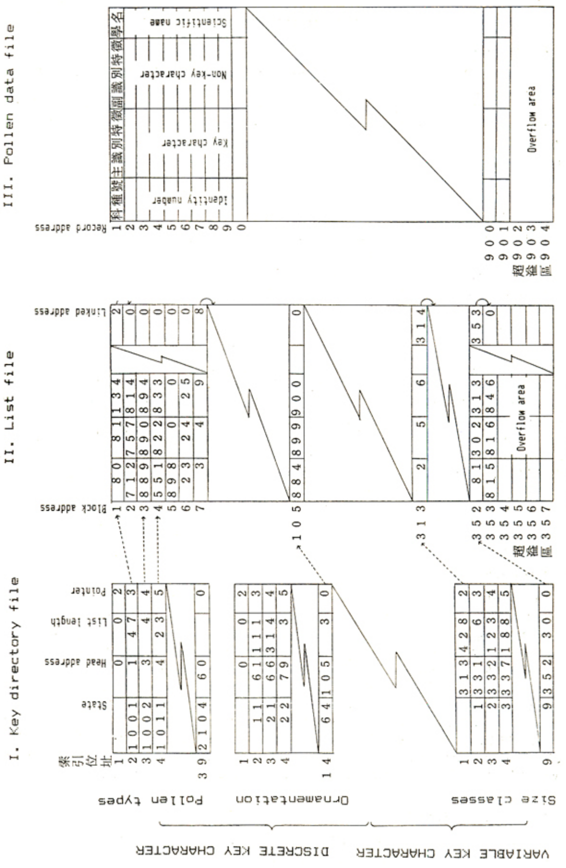

Fig. 11. File structure.

Table 1. Input form

Columns	Data
1	For deletion use
2— 4	Family identity number
5— 9	Species identity number
10	Identity number for infraspecific taxon
11— 14	Pollen types
15— 16	Ornamentation
17— 18	Sculptures in side view
19— 20	Shapes in equatorial view: lower limit
21— 22	Shapes in equatorial view: upper llmit
23— 25	Shapes in polar view: lower limit
26— 28	Shapes in polar view: upper limit
29— 30	Types of ora: lower limit
31— 32	Types of ora: upper limit
33— 35	A-axis length: lower limit
36— 38	A-axis length: upper limit
39— 41	B-axis length: lower limit
42— 44	B-axis length: upper limit
45— 46	C-axis length: lower limit
47— 48	C-axis length: upper limit
49— 51	Exine thickness: lower limit
52— 54	Exine thickness: upper limit
55— 57	Excrescence height: lower limit
58— 60	Excrescence height: upper limit
61— 62	Element width: lower limit
63— 64	Element width: upper limit
65— 66	Distance between elements: lower limit
67— 68	Distance between elements: upper limit
69— 71	L-axis length of a pore: lower limit
72— 74	L-axis length of a pore: upper limit
75— 77	W-axis length of a pore: lower limit
78— 80	W-axis length of a pore: upper limit
81— 82	Furrow length: lower limit
83— 84	Furrow length: upper limit
85— 87	Furrow width: lower limit
88— 90	Furrow width: upper limit
91— 92	Margo (or annulus) width: lower limit
93— 94	Margo (or annulus) width: upper limit
95— 96	Aperture number in one hemisphere: lower limit
97— 98	Aperture number in one hemisphere: upper limit

Table 1. Input form (Continued)

Columns	Data
99—100	Aperture types
101	Aperture types in side view
102	Reticulum structure
103	Evenness of exine thickness
104—105	Beginning of flowering period
106—107	End of flowering period
108—111	Altitude: lower limit
112—115	Altitude: upper limit
116—117	First distribution area
118—119	Second distribution area
120—159	Scientific name in text form
160—193	Major localities and ecological data

〈2〉 Generating character state/address pairs for all pollen records, begining with the first key character.

〈3〉 The pairs are sorted in address within state sequence, and the result is that all addresses appearing in a given state subsequence are arranged in a monotonic sequence. This forms a list for each state of a key character.

〈4〉 Storing a list in one or more blocks of a definite length, based on the list length.

〈5〉 Repeating step 〈4〉 for all states of a key character.

〈6〉 Repeating step 〈2〉-〈5〉 for all key characters, and the result is the forming of the list file.

〈7〉 Finally generating character state/head address of list/list length triplet in state sequence for all key characters, and this forms the key directory file.

As mentioned above, any variable key character can be variable or constant. We merely selected those records with constant character states for making list and key directory files and put the others in a blank list. This means that each variable key character possesses a blank list made up by addresses of those records that contain two or more states, in addition to those lists of non-blank states. Under such condition we shall not miss any record address included in the blank lists when identfication is executed. For UNIVAC 1100 computer, the whole construction procedure takes about 6.5 minutes.

3. File maintenance

File maintenance can be put into three categories: record addition, record modification, and record deletion. All of these can be executed in on-line mode.

(1) Record addition

It is always the case that new pollen records must be added to the system already erected. Coding process needs to be carried out in exactly the same way as described above. The records added can be one or more. Right after input, each record is assigned an address following the last record address of the pollen data file. Then the new addresses must be inserted in sequence into the corresponding lists. If the insertion of the addresses causes a list to overflow the allocated block in disk, another block (always following the last block of the list file) is attached and the link address inserted into the last reserve space of the previous block. The key directories are then updated by incrementing the list length, or sometimes by insertion of new character states.

(2) Record modification

This is conerned with the updates of the key characters, non-key characters, and other alphanumeric data. The modification of a key character means to delete a wrong character state and then add a desired one, and consequently all files involved in this system much be updated. If the data to be updated are not key characters, then only the pollen data file is involved.

The input form is quite simple, just filling out the desired data into the corresponding columns, while leaving the other columns as blanks, and then followed by an address number of the record. The data to be updated can be one digit, one letter, one or more coded character states, or even the whole record.

(3) Record deletion

The deletion of the whole record is not often encountered. There are two alternatives to attain this purpose. Method one is identical to a whole record modification as noted above. Another approach is to set a record deletion column at the beginning of the pollen record to 1 without having to affect list and key directory files. This means that the pollen record retrieved after identification must be skipped over, if it contains 1 in its first column.

4. Identification of unrecognized pollen grains

(1) Procedure

〈1〉 An unrecognized pollen is coded in almost a similar manner to that described above for reference pollen taxa except that columns 2–10 are used for filling identity code of that grain, that is of any alphanumeric combination, and column 1 may be scored to indicate the printing out of those unmatched character numbers if necessary. Columns 116–117 are used for storing area code from where the specimen is collected. If it happens to be near the boundary of two areas, then both columns 116–117 and 118–119 may be used.

〈2〉 By means of consulting key directory file, we can retrieve all lists from list file, which contain the sequential addresses of all pollen records that contain the same character states found in unknown pollen grain. Of course blank list of each variable key character, if present, must be included.

〈3〉 Merging of the blank list into non-blank list for each variable key character.

〈4〉 Finding intersection addresses from all lists.

〈5〉 Based upon the addresses intersected, it is possible to retrieve all records from the pollen data file, which should therefore satisfy a logical combination of those key characters owned by the unknown specimen.

〈6〉 Comparison of non-key characters is made between each pollen record retrieved and unknown specimen. Any record whose qualitative non-key characters do not match altogether, or range of any other non-key character does not include any of the alternatives or covers the range specified for the unknown specimen is rejected. Consequently, the records retained being one or several are results of the whole identification procedure.

(2) Output forms

The output resulting from an identification may fall into one of the following forms, each of which carries an identity code of the unknown grain.

〈1〉 When an unknown grain contains new character states which are not found in the key directory file, the result of the output is shown as follows:

TA001 (***) SPECIES WITH NEW KEY CHARACTER STATES

〈2〉 When an unknown grain contains new combination of character states, then the result of the output is:

TA002 (***) NEW SPECIES

〈3〉 When all characters of the unknown grain are fitted to those of the retrieved reference records, the result is as follows:

TA003 1. (464) TRAPA NATANS VAR. JAPONICA
AQUATIC
TA003 2. (496) BARRINGTONIA RACEMOSA
N. & S. TAIWAN, COASTAL
TA003 3. (505) CASSIA FISTULA
CULTIVATED

Each identity code is followed by a serial number, address of reference record, scientific name, and other information.

〈4〉 Sometimes it may be desirable to printout those reference records with well fitted key characters but not all of the non-key characters, then a series of numbers indicating the unmatched non-key characters will be appended to them.

TA004 1. (1) BLECHUM PYRAMIDATUM
UNMATCHED NON-KEY CHARACTERS: 4. 5. 9.

(3) Access timing

The time to process an identification is a function of the list length, list number and the number of the pollen records accessed. This means that the more the key characters are involved or lists intersected, the more is the time consumed. Consequently it takes 6–8 seconds for an unknown grain of tricolporate type, but only 2–3.5 seconds for other types, if more than ten unknown are inputed at the same time. The average access time is about 4.5 seconds.

DISCUSSION

1. The characters

It has been pointed out (Nair 1965) that in establishing the phylogenetic dicta for pollen morphology, apertures are considered as primary (most conservative), exine ornamentation as secondary and other characters (e. g. size, shape) as tertiary, according to their degree of importance. The arrangements of the characters in the present system, the discrete key characters, variable key characters, and non-key characters, are roughly fitted into these dicta. Besides, in consideration of practical reason, the key characters especially the pollen types which are mainly concerned with the apertures showing a great diagnostic value based upon their diversified and fairly constant manners are, therefore, very important as compared with other characters.

Owing to the successive occurrence of parallel and convergent evolution in the long geological period (Kuprianova 1969), the distribution of the states is not even within each character. For instance, approximately 40% of the pollen records are allocated to tricolporate type, and 30% to reticulate sculpturing according to the key directory file. These highly uneven frequencies of the character states greatly lessen their diagnostic value and consequently make it very difficult to identify an unknown from some groups of the grains. On the other hand, variations of pollen size and shape within the same species are very common. For example, 51% and 43% of the pollen records contain two or up to four states of the shape in equatorial and polar views respectively. These further aggravate the difficulties in identification. To overcome these, several approaches might be helpful. In the first place, a refinement of the sculpturing pattern should be made through the incorporation of the features observable by scanning electron microscopy. Secondly, a further evaluation of some non-key characters should be made, for example the furrow length and width, in order to realize their real diagnostic value. If it is necessary, other kinds of characters or expression manners may be adopted to reinforce the value of the non-key characters. Statistical treatments of quantitative characters such as multidimension methods may prove rewarding in groups with much overlap, nevertheless laboriousness must be involved and consequently they are able to handle only a certain limited number of the taxa if on-line query is necesary.

The erection of the subsidiary non-key characters is another approach already made. Of course, this should be based upon a fully understanding of the local flora. The real advantages of these characters in identification is very apparent. For example, if the altitude columns are filled with 0–300 (meters), then all those species of median and high mountain origin will be taken away. Nevertheless, altitude cannot be used without regard to the topographic features of Taiwan. When an air-borne pollen survey is conducted in Hwalien of the eastern coast area, seven miles aside from where is the Eastern Mountain Range attaining an average height of 2500 meters, then the altitude columns of the unrecognized grains should be left as blanks otherwise (e. g. 0–50 meters) it may eliminate those pollen taxa dispersed from high mountains. Another character deserving further discussion is the distribution area which may be used in two different senses, namely narrowly restricted area for reference pollen records and collecting area for unknown specimens. The result is that when we identify an unknown specimen from the northern area of Taiwan, strictly localized or topographically isolated species such as those found in southern area, Penghu Island, or other areas will be eliminated. This may raise a problem concerning long distance transport of the pollen grains, which has been occasionally mentioned (Ritchie and Lichti-Federovick 1967). Nevertheless, it is self-evident that in most cases the greatest quantities are within several kilometers (Colwell 1951, Wright 1953, Potter and Rowley 1960, Allessio and Rowley 1966, Huang and Chung 1973, Chen and Huang 1980, Tsou and Huang 1982). This is especially true in regard to bee-carried pollen grains which are conveyed to a distance largely within two kilometers (Crane 1976).

2. The system

The principle behind the matching method (e. g. Walker *et al.* 1968) is simply to calculate a measure of similarity between an unknown and each of the members of the reference pollen records, and to pick out those records which score the highest. This method has the advantage that it is not wrecked by a few mistakes in observation of the characters. However, it becomes inefficient when a lot of pollen records are accumulated. Multiple entry key is another method of identification, that allow one to select the characters for use in identifying each unknown specimen, taking his choices from some character set and repeating an elimination process until a tentative identification is made. The present system takes full advantage of the merits suggested by both methods and combines with the high speed of computer operation. We can identify an unknown pollen grain primarily by inputing any key character or subset of key characters, in the light of the material available. Then comparisons are made between non-key characters of each record retrieved and those of the unknown in order to test their agreement. Usually only the completely fitted records are printed. We have already seen that non-key characters are in most cases variable and sometimes the data prepared may not be representative of them. For this reason the system also offer an option for the user to retrieve those records partially matched in certain non-key characters and printout the unmatched characters, the palynologists may check these and decide which of the records should be skipped over. Although the present system cannot be regarded as a perfect one, it has already proved quite effective and further refinement may be necessary as its use continues.

3. Sources of errors

There are two main sources of errors that affect the correctness of identication: poor sampling of the pollen grains and error in reccording of character states. It is clear that poor sampling is easier to tackle simply by an amassment of a great number of relevant data and sufficient inclusion of intrataxon variation. As described above, it is acceptable in the present system to code a character with blanks if it is missed, and the result is that no comparison is made if either side of a compared character pair is coded as blanks. Therefore, too many characters missed in reference pollen records will bring about a number of records after identification, among which may contain some irrelevant ones. Consequently, the number of actually existing characters coded as blanks should be kept at a minimum. Another source of error that should be given more attention is the observational error. Owing to the difficulty of making accurate examination under optical microscopy, discrepancies may be introduced between the user and system designer in their choice of the character states. This is especially true if ornamentation of fine structure is involved. One solution to this problem is to find the variation trends for a variable key character. Thus if the user cannot make a clear discrimination, he will do well to use several states to represent that character (e. g. 313–314 in Fig. 6). Another solution is to use the same code number for those states whose delimitation is not clear-cut when preparing the coding system. This is frequently so in coding the discrete key characters. As a last resort, the user may use alternative states for input and then check the different results by direct comparising with the reference materials.

4. The use of a personal computer

The present storage and retrieval system was initially written in FORTRAN to run on UNIVAC 1100 computer. Since the acceptance of the personal, affordable microcomputer has rapidly spread throughout the world, a new version of the system was written in BASIC for Apple II computer. The computer system includes the 48 K microcomputer itself, a terminal which combines a keyboard and display, three 5¼″ floppy diskette drives, and a printer (not necessary).

After running the first program of file construction, it was found that several hours should be needed to complete the whole procedure. A lot of time was spent in sorting character states and address. The resulting pollen data file occupied 690 sectors (each sector stores 256 ASCII characters), list file 197 sectors, and key directory file 9 sectors. Since each diskette contains 496 sectors, two drives are the least requirements. If 8″ or 5¼″ of double sided diskette drive is adopted, one is far enogh to store all the files. As it spent a lot of time transferring information between the computer and the diskette drives, one of the variable key characters, the size classes, was discarded in order to improve the speed of identification. Consequently, 0.5–4 minutes were needed to identify an unknown specimen. It seems rather inefficient as compared with a large or median computer, but with the appearance of the high speed microprocessor and diskette drive, its potentiality cannot be underrated.

5. Progress and prospects

It is apparent that the identification procedure described above should involve a great deal of efforts, especially in examining and coding the reference pollen records. We may therefore anticipate a real automated method that can be introduced for extracting information directly from pollen grains and converting it into character states. This may be done by instruments such as scanning electron microscopes coupled to computers. Only through this way can we obtain images that show detailed patterns of the pollen surfaces. Then the processes of pattern recognition should be carried out and the resulting information, incorporated with scientific name and other data, be stored in the disk. The following data storage and retrieval system would be in the similar way as detailed above. There is no doubt that the foundamental problem underlying such kind of automatic identification is its expense. Moreover, the results of pattern recognition may not be as precise as what derived from human judgements. In these respects, semi-automatic identification with refined and man judged character coding system will concern us most in the near future, even though pattern recognition methods have been employed effectively for identification in some other fields.

REFERENCES

ALLESSIO, M. L. and J. R. ROWLEY, 1966. Atmospheric pollen in Amherst, Massachusetts. Bot. Gaz. **127**: 35-40.

BAUM, B. R. and L. P. LEFKOVITCH, 1972. A model for cultivar calssification and identification with reference to oats (avena). II. Canadian Journ. Bot. **50**: 131-138.

BELL, C. R., 1959. Mineral nutrition and flower to flower pollen size variation. Amer. Journ. Bot. **46**: 621-624.

BOUGHEY, A. S., K. W. BRIDGES and A. G. IKEDA, 1968. An automated biological identification key. Museum of Systematic Biology, Univ. of Cal., Irvine, research series No. 2 June.

ČEŠKA, A. and A. D. TRUMPOUR, 1979. Computer editing of serial and indented identification keys. Taxon **28**: 329-335.

CHENG, S. H. and T. C. HUANG, 1980. Aeropalynological study of Taipei Basin, Taiwan. Grana **19**: 147-155.

COLWELL, R. N., 1951. The use of radioactive isotopes in determing spore distribution patterns. Amer. Journ. Bot. **38**: 511-523.

CRANE, E., 1976. Honey, a comprehensive survey. Morrison & Gibb Ltd., London and Edingburgh, 608 pp.

DALLWITZ, M. J., 1980. A flexible program for generating identification keys. Systemic Zoology **23**: 50-59.

ERDTMAN, G., 1952. Pollen morphology and plant taxonomy, Angiosperms. Almquist & Wiksells, 539 pp.

________, 1966. Pollen morphology and taxonomy. Angiosperm. Hafner Publishing Co., New York and London, 533 pp.

________, 1969. Handbook of palynology. Hafner Publishing Co., New York, 486 pp.

FAEGRI, K., J. IVERSEN and H. T. WATERBOLK, 1964. The textbook of pollen analysis. Hafner Publishing Co., New York, 237 pp.

GERMERAAD, J. H. and J. MULLER, 1970. A computer-based numerical coding system for the description of pollen grains and spores. Rev. Palaeobot. Palynol. **10**: 175-202.

GOODALL, D. W., 1968. Identification by computer. Bioscience **18**: 485-488.

GOWER, J. C. and J. A. BARNETT, 1971. Selecting tests in diagnostic keys with unknown responses. Nature **232**: 491-493.

GREGORY, P. H., 1961. The microbiology of the atmosphere. Leonard Hill, London, 251 pp.

HALL, A. V., 1970. A computer—based system for forming identification keys. Taxon **19**: 12-18.

HSIEH, C. F., 1981. Character coding and identification systems for Formosan pollen grains. Dissertation, National Taiwan Univ. 169 pp. (in Chinese).

HEBDA, R. J. and N. A. LOTT, 1973. Effects of different temperatures and humidities during growth on pollen morphology; an SEM study. Pollen et Spores 15: 39-571.

HUANG, T. C., 1972. Pollen flora of Taiwan. National Taiwan University, Taipei, 297 pp. 10 text-figures, 177 plates.

________, 1981. Spore flora of Taiwan. Private publication, Taipei, 114 pp. 12 text-figures, 120 plates.

________ and T. F. CHUNG, 1973. Aeropalynological study of Taiwan (1)—Chu-shan station. Taiwania 18: 160-168.

IKUSE, M., 1956. Pollen grains of Japan. Hirokawa Publishing Co., Tokyo, 303 pp.

JOHNSTON, B. C., 1980. Computer programs for constructing polyclave keys from data matrices. Taxon 29(1): 47-51.

KRAUSS, H. M., 1973. The use of generalized information processing systems in the biological sciences. Taxon 22(1): 3-18.

KREMP, G. O. W., 1965. Morphologic encyclopedia of palynology. The University of Arizona Press, Tucson, 263 pp.

KUPRIANOVA, L. A., 1969. On the evolutionary levels in the morphology of pollen grains and spores. Pollen et Spores 11: 333-351.

KUYL, O. S., J. MULLER, and H. T. WATERBOLK, 1955. The application of palynology to oil geology with reference to Western Venezuela. Geol. Mijnbouw 17(3): 49-76.

LAPAGE, S. P., S. BASCOMB, W. R. WILLCOX and M. A. CURTIS, 1971. Computer identification of bacteria in automation, mechanization and data-handling in microbiology, Academic Press.

LEFKOVITZ, D., 1972. File structures for on-line systems. Computer Command and Control Company, Philadelphia, 215 pp.

MIKKELSEN, W. M., 1949. Has temperature any influence on pollen size? Physiologia Plantarum. 2: 323-324.

MORSE, L. E., 1968. Construction of identification keys by computer. Amer. Jour. Bot. 55: 737.

________, 1969. Time-sharing computers as aids to identification of plant specimens. Demonstration at XI Int. Bot. Cong., Seattle.

________, 1971. Specimen identification and key construction with time—sharing computers. Taxon 20: 269-282.

________, 1974. Computer programs for specimen identification, key construction and description printing using taxonomic data matrices. Michigan State Univ. Mus. Biol. Ser. 5: 1-128.

NAIR, P. K. K., 1965. Pollen grains of Western Himalayan plants. Asia Publishing House, New York. 102 pp.

________, 1967. Pollen morphology with reference to the taxonomy and phylogeny of the Monochlamydeae. Rev. Palaeobot. Palynol. 3: 81-91.

NILSSON, S., 1967. Notes on pollen morphological variation in Gentianaceae-Gentianinae. Pollen et Spores 9: 4-58.

________, J. PRAGLOWSKI and L. NILSSON, 1977. Atlas of airborne pollen grains and spores in Northern Europe. 1st ed. Natur och Kultur, Stockholm, 159 pp.

PANKHURST, R. J., 1970. A computer program for generating diagnostic keys. Comp. Jour. 13: 145-151.

________, 1971. Botanical keys generated by computer. Watsonia 8: 357-368.

________, and R. R. AITCHISON, 1975. A computer program to construct polyclaves. *In*: R. J. Pankhurst (ed.), *Biological Identification with Computers.* Systematics Association Special Volume 7: 73-78.

POTTER, L. D. and J. ROWLEY, 1960. Pollen rain and vegetation, San Augustin plains, New Mexico. Bot. Gaz. 122: 1-25.

REITSMA, TJ., 1970. Suggestions towards unification of descriptive terminology of angisperm pollen grains. Rev. Palaeobot. Palynol. 10: 39-60.

RITCHIE, J. C. and S. LICHTI-FEDEROVICK, 1967. Pollen dispersal phenomena in arctic—subarctic Canada. Rev. Palaeobotan. Palynol. 3: 255-266.

SCHOCH-BODMER, H., 1940. The influence of nutrition upon pollen grain size in *Lythrum salicaria.* Journ. Genetics 40: 393-402.

THOMSON, P. W. and H. PFLUG, 1953. Pollen und Sporen des mitteleuropäischen Tertiärs. Palaeontographica, Abt. B, v. 94, Liefg. 1-4, p. 1-138.

TSOU, C. H. and T. C. HUANG, 1982. Aeropalynological study of Taipei Suburban (New Flower Garden City). Taiwania **27**: 9-29.

WALKER, D., P. MILNE, J. GUPPY and J. WILLIANS, 1968. Computer assisted storage and retrieval of pollen morphological data. Pollen et Spores **10**: 251-262.

WATSON, L. and P. MILNE, 1972. A flexible system for automatic generation of special-purpose dichotomous keys, and its application to Australian grass genera. Australian Journ. Bot. **20**: 331-352.

WEBER, W. A. and P. NELSON, 1973. Random acess key to the genera of Colorado mosses. Bryologist **76**(1): 222-223.

WOODEHOUSE, R. P., 1935. Pollen Grains. Hafner Publishing Co., New York, 574 pp.

WRIGHT, J. W., 1953. Pollen-dispersion studies: some practical applications. Jour. Forestry **51**: 114.

The editorial committee, 1975-1978. Flora of Taiwan, Vol. 1-6. Epoch Publishing Co., Ltd., Taipei, Taiwan.

臺灣植物花粉的資料儲存及識別系統

謝 長 富　　黃 增 泉

摘　　要

花粉具有萌芽口、雕紋、形狀等特徵以及不易損毀之外壁，致使其成爲研究植物分類學、古植物地理學、空中孢粉學、考古學與地層學所不可缺之一環。由於花粉特徵之變異性大，同時許多不同種類能產生類似之花粉型，加上經常受到觀察角度之限制，使得花粉之鑑別困難重重。因此本論文之目的卽在尋求一迅速有效之方法，以辨識花粉之種類。全文分爲兩部分。第一部分將易於分辨、且具有識別價值之特徵及特徵類別儘量予以蒐集。然後有系統地加以歸類編碼。歸類後之花粉特徵可歸納爲主、副兩識別特徵。前者依其變異性之大小復分爲固定、變異兩類；後者則依其性質劃分爲甲、乙、丙、丁四羣。所有特徵計 31 項，而特徵類別則超過 168 種 。 此種歸類方式，不僅適應花粉之特性，同時亦符合查詢系統設立之需求。第二部分在設立一套查詢識別之系統。整個系統由五項目所組成：⑴建立適合快逐查詢之花粉檔案。此檔案包括資料、索引及串列三部門；⑵新花粉記錄之加塡；⑶錯誤記錄之修正；⑷錯誤記錄之删除；⑸未知花粉之查詢。