New Phytol. (1984) **96**, 631–634

COMPUTER CONSTRUCTION AND TYPESETTING OF IDENTIFICATION KEYS

By R. W. PAYNE

Statistics Department, Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ UK

(*Accepted 9 November 1983*)

Summary

The ability to produce computer files containing keys, together with the commands necessary to typeset them directly from the program that constructs them, greatly simplifies the production of high quality identification keys.

Key words: Computer typesetting, diagnostic keys, identification keys.

Introduction

Computer programs for constructing identification keys have been available since the early 1970's (e.g. Pankhurst, 1970; Morse, 1971; Dallwitz, 1974; Payne, 1975, 1978). These programs greatly simplify the production of good keys; however they provide only part of the process of publishing a key. The computer output must still be converted into a publishable form.

The obvious, straightforward, way of publication would be to treat the computer line-printer output in exactly the same way as the other (typewritten) text of the paper. It would then be typeset (by hand) and proof read in the usual way. However long keys can be tedious and expensive to typeset and accurate proof-reading is crucial as any mistake in the index numbers (printed down the right-hand side) would cause the key to identify incorrectly. (The effect is exactly the same as the incorrect observation of a character as discussed by Payne & Preece, 1977).

Consequently many authors have decided to treat the line-printer output of the key as camera-ready copy. This is not altogether satisfactory; usually there is a limited choice of fonts (often only capital letters) and computer output may not photo-reproduce particularly well (cf. Payne, Walton & Barnett, 1974). Even where both lower and upper-case letters can be printed (e.g. Payne, Yarrow & Barnett, 1982), the line printer output will not look as attractive as typeset text. In the fonts used on line-printers, all characters occupy an equal width; these are not as easy to read as those used in typesetting, in which letters like 'i' occupy a smaller width than, for example, 'w'. Also, in many keys there is the need for italic and Greek fonts.

These disadvantages can be overcome by the use of computer typesetting techniques. In Australia, Dallwitz used the typesetting equipment of CSIRO Division of Computing Research to produce keys and descriptions for Australian grass genera (see Dallwitz, 1984). The equivalent facility for British academics is the Lasercomp Typesetter at Oxford University Computing Service. This was used by Barnett, Payne & Yarrow (1983) to produce keys, tables and descriptions for 473 species of yeast.

0028-646X/84/040631+04 $03.00/0

COMPUTER TYPESETTING

In computer typesetting, the text to be printed is entered into the computer, together with special instructions to specify typefaces, size of print, page widths, tabulation, etc. Keys, like those in Tables 1 and 2, tend to have a fairly regular format, so it is possible to modify programs that construct keys to insert the necessary typesetting instructions automatically into the text of the key.

Table 1. *A key to yeasts that utilize methanol*

		Negative	Positive
1	Nitrate growth	2	17
2	Erythritol growth	3	8
3	Salicin growth	4	6
4	Galactitol growth	5	*Hansenula ofunaensis*
5	D-Glucose fermentation	*Candida maris*	*Pichia pastoris*
6	L-Arabinitol growth	*Hansenula nonfermentans* *Pichia lindneri*	7
7	L-Rhamnose growth	*Candida sonorensis*	*Pichia lindneri*
8	Galactitol growth	9	14
9	Sucrose growth	10	13
10	Salicin growth	11	12
11	0.1% Cycloheximide growth	*Candida pinus*	*Pichia trehalophila*
12	0.1% Cycloheximide growth	*Pichia kodamae* *Pichia pini*	*Pichia methanolica* *Pichia pini*
13	L-Rhamnose growth	*Candida entomophila*	*Pichia naganishii*
14	D-Glucose fermentation	*Candida nemodendra*	15
15	D-Glucosamine growth	16	*Candida succiphila*
16	0.1% Cycloheximide growth	*Pichia kodamae*	*Pichia methanolica*
17	Erythritol growth	18	26
18	D-Glucosamine growth	19	24
19	L-Arabinitol growth	20	21
20	0.1% Cycloheximide growth	*Hansenula henricii*	*Hansenula nonfermentans*
21	L-Rhamnose growth	22	23
22	Salicin growth	*Candida pignaliae*	*Hansenula minuta*
23	0.1% Cycloheximide growth	*Hansenula henricii*	*Hansenula glucozyma*
24	Galactitol growth	25	*Hansenula ofunaensis*
25	Salicin growth	*Candida nitratophila*	*Candida methanosorbosa*
26	Sucrose growth	27	34
27	L-Rhamnose growth	28	30
28	D-Glucose fermentation	*Hansenula philodendra*	29
29	Salicin growth	*Candida boidinii*	*Hansenula capsulata*
30	D-Glucosamine growth	31	*Hansenula capsulata*
31	Salicin growth	32	33
32	D-Glucose fermentation	*Hansenula wickerhamii*	*Candida methylica*
33	0.1% Cycloheximide growth	*Hansenula henricii*	*Hansenula glucozyma*
34	D-Glucosamine growth	*Hansenula polymorpha*	*Candida cariosilignicola*

This has been done in the computer program Genkey (Payne, 1975, 1978) that constructed the yeast keys, mentioned above. One of these keys, that for the yeasts that can utilize methanol, is shown in Table 1. This is printed in one of the new representations of Payne *et al.* (1974). Genkey can also typeset keys in the conventional 'bracketed' and 'indented' forms. An example, part of a key for Australian species of the genus *Solanum*, is shown in Table 2. The four fonts used in these keys greatly enhance their appearance and the fact that much of the work has been done automatically makes them very much cheaper to produce.

Table 2. *Part of a key to Australian species of the genus Solanum*

1(0) Hairs stellate;	2
Hairs not stellate;	238
2(1) Hairs simple;	*Solanum hystrix*
Hairs not simple;	3
3(2) Hairs dense below;	4
Hairs sparse below;	215
4(3) Hairs glandular;	5
Hairs not glandular;	11
5(4) Prickles found on calyx; Berry not bony; Anther length 4-6 mm;	6
Prickles not found on calyx; Berry bony; Anther length 6-8 mm;	10
6(5) Hairs dense above; Form shrubby or small tree;	7
Hairs sparse above; Form short lived or herbaceous perennial;	*Solanum adenophorum*
7(6) Leaves entire;	8
Leaves lobed;	9
8(7) Leaf length 4-6 cm; Seed colour pale bone or browny buff;	*Solanum gabrielae*
Leaf length 8-10 cm; Seed colour black;	*Solanum campanulatum*
9(7) Berry 1/2 or more in calyx; Seed colour black; Seed length 1-2 mm;	*Solanum campanulatum*
Berry less than 1/2 in calyx; Seed colour pale bone or browny buff; Seed length 2-3 mm;	*Solanum cookii*
10(5) Leaf width 2-3 cm; Seed colour pale bone; Berry diameter 15-20 mm; Seed length 2-3 mm;	*Solanum lachnophyllum*
Leaf width >4 cm; Seed colour browny buff; Berry diameter 10-15 mm; Seed length 1-2 mm;	*Solanum ashbyae*
11(4) Leaves entire;	12
Leaves lobed;	125
12(11) Hairs dense above;	13
Hairs sparse above;	86
13(12) Prickles present;	14
Prickles absent;	57
14(13) Prickles found on calyx;	15
Prickles not found on calyx;	41
15(14) Prickles found on upper leaf;	16
Prickles not found on upper leaf;	32
16(15) Prickles scattered or sparse;	17
Prickles abundant;	24
17(16) Leaf width 1-2 cm;	18
Leaf width 2-3 cm;	20
Leaf width 3-4 cm;	23
18(17) Aspect rusty; Stigma bifed; Berry 1/2 or more in calyx;	*Solanum dioicum*
Aspect not rusty; Stigma not bifed; Berry less than 1/2 in calyx;	19
19(18) Flowers andromonoecious; Berry red, orange or yellow; Seed colour black; Anther length 6-8 mm;	*Solanum eburneum*
Flowers not andromonoecious; Berry green, green flushed purple or purple-black; Seed colour pale bone or browny buff; Anther length 4-6 mm;	*Solanum dianthophorum*
20(17) Petiole length 1-2 cm;	21
Petiole length 2-3 cm;	22
Petiole length 3-4 cm;	*Solanum ellipticum*
21(20) Flowers several; Flowers andromonoecious; Berry red, orange or yellow; Seed colour black; Anther length 6-8 mm;	*Solanum phlomoides*
Flowers solitary; Flowers not andromonoecious; Berry green, green flushed purple or purple-black; Seed colour pale bone or browny buff; Anther length 2-4 mm;	*Solanum horridum*
22(20) Anther length 2-4 mm;	*Solanum ellipticum* *Solanum horridum*
Anther length 4-6 mm;	*Solanum dianthophorum*

Much text for papers and books is now word-processed on microcomputers; such text can also be computer typeset – indeed the text for Barnett *et al.* (1983) was produced in just this way.

CONCLUSION

The use of computer typesetting techniques greatly simplifies the production of high quality identification keys. The typesetting facilities in Genkey are appropriate only for the Oxford University Lasercomp system; however similar concepts are used in other typesetters and conversion to their syntax should not be difficult. Details about other facilities and availability of Genkey can be obtained from the author and details about use of the Oxford University Lasercomp Typesetter can be obtained from: Lasercomp Service (External), Oxford University Computing Service, 13 Banbury Road, Oxford OX2 6NN.

REFERENCES

BARNETT, J. A., PAYNE, R. W. & YARROW, D. (1983). *Yeasts; Characteristics and Identification*. Cambridge University Press, Cambridge.

DALLWITZ, M. J. (1974). A flexible computer program for generating identification keys. *Systematic Zoology*, **23**, 50–57.

DALLWITZ, M. J. (1984). Automatic typesetting of computer-generated keys and descriptions. In: *Databases in Systematics, Systematics Association special vol. no. 26*, pp. 279–290. Academic Press, London.

MARRIOT, I. (1982). *LASERCHECK Users' Manual*. Oxford University Computing Service, Oxford.

MORSE, L. E. (1971). Specimen identification and key construction with time-sharing computers. *Taxon*, **20**, 269–282.

PANKHURST, R. J. (1970). A computer program for generating diagnostic keys. *Computer Journal*, **13**, 145–151.

PAYNE, R. W. (1975). Genkey; a program for constructing diagnostic keys. In: *Biological Identification with Computers, Systematics Association special vol. no. 7*, pp. 65–72. Academic Press, London.

PAYNE, R. W. (1978). *Genkey: a Program for Constructing and Printing Identification Keys and Diagnostic Tables*. Rothamsted Experimental Station, Harpenden.

PAYNE, R. W., WALTON, E. & BARNETT, J. A. (1974). A new way of representing diagnostic keys. *Journal of General Microbiology*, **98**, 595–597.

PAYNE, R. W. & PREECE, D. A. (1977). Incorporating checks against observer error into identification keys. *New Phytologist*, **79**, 203–209.

PAYNE, R. W., YARROW, D. & BARNETT, J. A. (1982). The construction by computer of a diagnostic key to the genera of yeasts and other such groups of taxa. *Journal of General Microbiology*, **128**, 1265–1277.