Use of a Spreadsheet Template and Microcomputer to Calculate Gene Frequencies of Multi-allele Systems

M.A. Helgeson, L. Wiessner and H.F. Polesky

Memorial Blood Center of Minneapolis, 2304 Park Ave. So., Minneapolis, MN 55404, USA

INTRODUCTION

Using Lotus 123^{TM}, we have developed a macro driven template with simplified data entry and menu driven functions to analyze population data for multi-allelic systems. Four systems, with 4, 7, 26 and 66 alleles respectively were evaluated. In designing the template, various approaches of analyzing data were evaluated to determine the best method for giving expedient results and using the least amount of memory.

METHODS

Timed studies were run on sample data using two versions of Lotus - with and without speed-up, and the Quattro^{TM} program. The built-in Data Distribution function was used to count allele numbers and gene frequency was calculated using cell formulas. Two methods of determining the frequency of phenotypes for the 4 DNA systems were evaluated. Each system was set-up in a two-way table format, then calculated, using the program's built-in Data Table function and with a programmed macro using the spreadsheet's command language. The same formula, @DCOUNT(DATABASE,OFFSET,CRITERIA), was used for both methods. Both spreadsheet programs and methods were initially tested on two microcomputers, each having a different central processing unit (CPU) and speed.

RESULTS

Results of the timed studies on the above variables are shown in Table 1. When using Lotus 123^{TM} the programmed macro was considerably faster than the program's internal Data Table function on both models of computers tested. Quattro^{TM} produced results more rapidly with the internal Data Table function versus the programmed macro, but was still slower than the faster method using Lotus.

Table 2 shows a comparison of time tests using the final template design and Lotus 123^{TM} with speed-up, to process a 500 and 1000 record database on three computer models. As would be expected, the more powerful computer is able to complete the task in a more expedient manner.

SUMMARY

The program chosen to process the final template was Lotus 123^{TM} with speed-up. The Data Distribution function and cell formulas were used to determine gene frequencies. Macro driven commands were used to determine phenotype frequencies using a two-way table format. The template incorporates custom menus enabling persons with minimal computer experience to enter data, perform calculations and generate printed reports. The template will accommodate up to 2000 records using conventional memory (640K). Additional records can be added by using expanded memory. Although the template can be used on a PC (8088,4.77MHz) computer, a more powerful model processes records in a much faster time. The template will also run with-out

Advances in Forensic Haemogenetics 3
Edited by H. F. Polesky and W. R. Mayr

modification with the Quattro spreadsheet program. A template such as this could also be used in conjunction with templates designed to perform statistical analysis.

While dedicated software could probably perform analysis of multi-allelic systems more expediently than this template, these experiments demonstrate that analysis of small amounts of data can be accomplished using a well designed spreadsheet template and "state of the art" microcomputer system.

Table 1. Time studies of Macro vs Data Table function for determining phenotypes.^{a}

		Lotus 123$^{\text{TM}}$				Quattro$^{\text{TM}}$	
		-Speed-up		+Speed-up			
Computer Model	Table^b Size	Macro	Data Table	Macro	Data Table	Macro	Data Table
80286,10	4x4	:04	:47	:04	:08	:10	:07
	7x7	:10	2:17	:10	:19	:28	:15
	26x26	2:22	30:39	2:22	5:12	5:53	4:21
	66x66	14:58	114:05	14:59	51:21	52:20	39:31
80386,20	4x4	:01	:10	:01	:02	:04	:04
	7x7	:03	:29	:04	:06	:10	:07
	26x26	:49	6:47	:49	1:35	2:20	2:02
	66x66	5:13	46:46	5:12	21:25	14:14	8:26

[a] Times are shown in Min:Sec.

[b] Table sizes represent a 4, 7, 26 and 66 allele system respectively.

Table 2. Time for calculating frequencies of 500 and 1000 records using 3 computers.^{a}

	Computer Models					
	8088-4.77MHz		80286-10MHz		80386-20MHz	
	500	1000	500	1000	500	1000
Gene frequencies^b	:17	:47	:10	:12	:04	:06
Phenotypes:						
4 alleles	:13	:23	:04	:06	:01	:02
7 alleles	:37	1:01	:10	:16	:04	:08
26 alleles	4:40	16:00	2:22	3:30	:49	1:22
66 alleles	26:57	95:31	14:59	20:35	5:12	7:16

[a] Times are shown in Min:Sec.

[b] Time given is the total time to calculate 4 systems with 4, 7, 26 and 66 alleles.

REFERENCES

Gasteiger D (1987) Fast two-way tables. Lotus 3(5):29-32.