

ALIAS - A COMPUTER SYSTEM FOR
TRACKING BREEDING MATERIALS WITH SEVERAL NAMES

Don E. Riemenschneider 1/

Abstract.--Maintaining accurate records of plant materials that have several names (aliases) is critical in a cooperative breeding environment. This paper describes a computer system designed specifically for maintaining searching such records.

Inter-agency cooperation in tree improvement in the Lake States has been commonplace for at least the last three decades. This cooperation has, for the most part, been informal with little emphasis on administrative structures. Under this loose arrangement, little attention has been given to joint record keeping.

As our Region moves into a new era of cooperative breeding with the organization of LASTICC, the Lake States Tree Improvement Coordinating Council, there is a need for increased emphasis on accurately maintaining the identity of breeding materials. One major problem is that many plant materials have several names. It is possible that an organization could receive a seedlot from another agency, assign its own identification number, and later in response to a seed request return the same seedlot to the originating agency. The originating agency then possesses two identical seedlots, but due to poor record keeping they are thought to be different. Such errors in material identification can lead to genetic costs when new breeding populations are created, selections are made, and seed production plantations are established.

One of the first goals of the LASTICC Records Committee is to develop a system for tracking multiple names of breeding materials (aliases). The system will include measures to ensure accurate reporting of new names as material is transferred among organizations and will also include a method for updating, editing, and searching large lists of such data. This paper describes one component of the system -- a computerized data management system which can maintain and search such a data set.

In the following description the term "accession number" refers to an identification code given to a family, clone, or other set of plant material. If the same plant material has more than one accession number, the numbers are referred to as "aliases". The term "originator" refers to the agency that collected the seed originally.

DESCRIPTION OF THE SYSTEM

Data File Structure

All alias data are contained in disk files. Separate files may be maintained for each species or breeding program. One file record is allocated for each

1/Research Plant Geneticist, North Central Forest Experiment Station, Forestry Sciences Laboratory, Box 898, Rhinelander, WI.

unique entry (i.e., family, clone, seed source, etc.) with a maximum of six aliases per entry. Each alias field in the record contains the following information:

<u>Data Field</u>	<u>Data Type</u>	<u>Length (characters)</u>
1. Accession	Alpha-Numeric	8
2. Agency Code	Numeric	3
3. Originator	Alpha	1

The accession numbers (aliases) may be any combination of alpha or numeric characters such as "8002" or "MN105". The agency code is a three digit integer. LASTICC recommends that all cooperators use the agency codes currently recognized by the USDA Forest Service, Region Nine tree improvement program. The originator field is a simple indicator of which agency in the alias list made the original collection and in the data file is represented by a "*" character. The total length of each record is 72 bytes, for a storage requirement of 72 K bytes per 1000 records. This amount of storage is well within the capacity of the floppy disk drives on most microcomputers. The record length is fixed, which allows direct access to the data file for editing operations.

Software

All programming is FORTRAN-IV-PLUS (Digital Equipment Corporation. ^{2/}). The language is based on American National Standard (ANSI) FORTRAN X3.9-1966 but includes several enhancements. In this application the most important of these is direct access input/output. The system runs on a Digital Equipment Corporation PDP-11/34 minicomputer under the RSX-11M version 3.2 operating system.

The program is interactive and menu-driven. The user is asked if any of the following operations are desired: 1) create an alias file, 2) print a file, 3) add a record to a file, 4) edit a file, 5) search a file, or 6) terminate a session. Any function except print and search requires a password so the database cannot be corrupted by unauthorized users.

The search function has several options that will allow flexibility in extracting information. A user may obtain all records that contain a specific agency code. This search mode may be further controlled by specifying that only those records for which the requested agency is the originator be reported. Alternatively, the user may search for any alias number-agency code combination. This search mode will report only 1 record, if such an entry is present, with all alias information.

A sample search session is presented in Appendix 1. Copies of the FORTRAN code and supporting information are available from the author.

^{2/}Mention of trade names does not constitute an endorsement by the U.S. Department of Agriculture.

PERFORMANCE

Performance trials on search speed were run for files of several lengths (table 1). These search times represent the total time required for a sequential search of each file size. The times will be accurate only for the PDP-11/34 under conditions of no user competition or for machines of similar processor and disk input/output speeds.

Table 1.--Search times for files of various sizes.

File Size (Number of 72 byte records)	Search Time (seconds)
200	3.8
600	11.1
1200	22.0
2000	36.0

The data indicate a search time of .018 seconds per record or 18 seconds per 1000 records. These speeds are for "no hit" (no match found) searches only. The time for an actual search would be increased by an amount depending on the number of successful "hits" and the speed of the user's printer.

ADDITIONAL COMMENTS

For the computerized alias identification system to be truly effective, additional manual measures will need to be implemented. It will be the responsibility of all participating agencies to ensure that a proper reporting procedure is maintained. I suggest that each species working group in LASTICC select one member to be responsible for maintaining the alias records. The system described here is simple and should be compatible with a large range of small computer systems. Alternatively, one central working group could manage the records for all tree improvement programs in our region. The exact implementation of the system has yet to be determined.

The program presented in this paper is offered only as a partial solution to a larger impending problem; that of the need for a complete data management system for cooperative breeding programs. The most desirable solution would be to immediately develop a comprehensive data management system that would fill all the needs of a cooperative breeding program -- not just one aspect of it such as searching for aliases. The required software may be available and may require only modest adjustment to serve our needs. If planning for such a system is not begun soon, it is inevitable that additional single function systems will be developed as each need arises. This will only lead to duplication of effort and increased cost.

APPENDIX A

SAMPLE ALIAS SYSTEM SEARCH

)RUN SYSTEM

Welcome to ALIAS-Alias Identification System

What is your full name please? DON RIEMENSCHNEIDER

Function Menu:

1=Create Alias File
2=Print File
3=Add Record
4=Edit Record
5=Search File
6=Terminate Session

Enter function number: 5

Search file.

Enter file name: JTEST.AID

Search by agency only (.1) or by agency and accession number combination (.2) ? 2

Enter agency code: 960

Enter accession number: 795

APPENDIX A (Cont'd.)

SAMPLE ALIAS SYSTEM SEARCH

ALIAS REPORT

File record number 53.

<u>ALIAS</u>	<u>ACCESSION</u>	<u>AGENCY</u>	<u>ORIGINATOR</u>
1	597	37	
2	686	43	
3	795	960	*
4	650	900	

End of search. 1 record found.

Another search ? N

Function Menu:

```
1=Create Alias File
2=Print File
3=Add Record
4=Edit Record
5=Search File
6=Terminate Session
```

***** Session terminated. *****

Leaving ALIAS at 13:34:25 on 18-JUL-83.

TT10 -- STOP