

SISTEMA DE PROGRAMAS PARA LA CARACTERIZACION BIOQUIMICA Y PROTEICA DEL BANCO DE DATOS DE GERMOPLASMAS

I. Vinardell, E. Alvarez, L. Sastre, A. Ruíz y R. Mitchel

Centro Nacional de Investigaciones Científicas y Facultad de Cibernética-Matemática, Universidad de La Habana, Cuba

ABSTRACT. This programm system permit the creation of the computer bank of the biochemical and proteical characterization of the sugar cane. It includes a data base and a pattern recognition system which permit to identify new varieties and varietal mixture, and thus produce an important economical saving. The germplasm pattern recognition is based in the qualitative analysis of the peroxidase, alpha-esterase and beta-esterase isoenzymes in the sugar cane, with the objective to obtain a characterization of the electrophoretic patterns of this isoenzymes in the varieties of the cuban sugar cane and with these patterns offer an easy, quick, and not expensive production mass method for the varietal identification. The program system was programmed in Turbo Pascal and it has been introduced in the practice.

RESUMEN. Este sistema de programas permite la creación de un banco computarizado de las caracterizaciones bioquímicas y proteicas de variedades de caña de azúcar. Incluye una base de datos y un sistema de reconocimiento de patrones que permite identificar nuevas variedades y mezcla de variedades, lo cual posibilita ahorros económicos importantes. El reconocimiento de patrones de germoplasmas se basa en el análisis cualitativo de la electroforesis de las isoenzimas peroxidazas, alfaesterazas y betaesterazas en la caña de azúcar. El paquete de programas permite lograr una caracterización de los patrones electroforéticos de estas isoenzimas en las variedades de la caña de azúcar cultivadas en el país y mediante estos patrones brindar un método en la producción que sea sencillo, rápido, masivo y de bajo costo para la identificación de las variedades. El sistema fue programado en Turbo Pascal y está en fase de introducción en la práctica.

INTRODUCCION

Uno de los problemas que confronta la agricultura cañera en la actualidad es la mezcla de variedades debido principalmente a las dificultades para su identificación.

Esta situación conduce a la siembra de variedades proscritas que se confunden con las comerciales. Las variedades proscritas son de baja productividad y sensibles a enfermedades, por lo cual pueden provocar bajos rendimientos y constituir focos de infección.

Hasta hace poco en Cuba la identificación de las variedades se hacía generalmente sobre la base de características fenológicas como el color del tallo, grosor del entrenudo, características de las yemas, hábito de crecimiento, etcétera.

Sin embargo, este método tiene serias limitaciones, como son:

- la apreciación es subjetiva,
- las características fenológicas pueden variar según el tipo de suelo, clima, etcétera,
- no es posible identificar las plantas juveniles,
- requiere de larga experiencia,
- no es posible hacer el análisis masivo de las plantas.

Es por ello que se ha comenzado a trabajar para lograr un método que permita la identificación de las variedades de caña¹ de forma confiable y rápida.

DESCRIPCION DEL SISTEMA

El sistema de programas que se describe en el presente trabajo se basa en el

reconocimiento de patrones del banco de datos de germoplasmas, utilizando el análisis cualitativo de la electroforesis de las enzimas peroxidazas, α -alfaesterazas y betaesterazas. Los objetivos que se persiguen son: caracterizar los patrones electroforéticos de las isoenzimas especificadas anteriormente de las variedades de la caña de azúcar cultivadas en Cuba y brindar un método a la producción sencillo, rápido, masivo y de bajo costo para la identificación de variedades de caña, mediante el empleo de los patrones isoenzimáticos.

El análisis utilizado en este sistema se basa en la aplicación de una fórmula de distancia que brinda, desde el punto de vista matemático y biológico, la relación entre las distintas variedades.

Cada una de las isoenzimas que intervienen en el análisis está compuesta por 19 bandas como máximo. A cada una de estas bandas se le hace corresponder un valor entero entre 0 y 3 en dependencia de su amplitud e intensidad.

Para determinar la relación entre variedades se utilizan los siguientes criterios: dos muestras son iguales si y sólo si coinciden en las tres isoenzimas; dos muestras son semejantes si y sólo si son semejantes en las tres isoenzimas; dos muestras son diferentes si y sólo si son diferentes al menos en una isoenzima.

La fórmula de distancia³ que se utiliza es:

$$S_k = \sum_{i=1}^{19} \frac{|X_{ik} - Y_{ik}|}{X_{ik} + Y_{ik} + 1} \quad k = 1, \dots, 3$$

$$0 \leq x, y \leq 3$$

Esta fórmula, desde el punto de vista matemático, mide la distancia entre dos puntos, cada uno de los cuales representa una muestra. Para determinar el grado de cercanía entre ellos se fija un α ($\alpha = 0,5$). Este α se compara con S_k para las tres isoenzimas.

Si $S_k = 0$ entonces la muestra es igual a una determinada variedad en la prueba k .

Si $0 < S_k \leq \alpha$ entonces la muestra es semejante a una variedad en la prueba k .

Si $S_k > \alpha$ entonces la muestra es diferente en la prueba k .

Desde el punto de vista biológico que S_k sea igual a cero significa que las muestras tienen características genéticas iguales, si $0 < S_k \leq \alpha$ significa que las muestras tienen características genéticas semejantes y si $S_k > \alpha$ significa que las muestras tienen características genéticas diferentes.

En la entrada de datos al sistema se realiza un chequeo de la información. El nombre de la variedad tiene que ser una cadena de caracteres. Los valores de las bandas, los cuales se introducen para cada prueba, deben ser números enteros entre 0 y 3. En caso de que haya algún error en la información de entrada el usuario tiene la posibilidad de rectificarlo.

Una vez suministrada la información sobre una muestra, se procede a determinar si es igual, semejante o diferente a una o varias muestras de la base de datos.

El sistema permite listar por la impresora o por el display los resultados del análisis realizado, obtener el listado de una familia de muestras, cambiar el nombre de cualquier muestra que esté en la base de datos verificando siempre que el nombre no se repite, incluir una muestra en la base.

El sistema ha sido programado en Turbo Pascal⁴ y se encuentra en fase de introducción en la práctica.

CONCLUSIONES

El procedimiento empleado en el estudio electroforético de las isoenzimas de la caña de azúcar por su bajo costo, es factible de ser extendido a la producción en todo el país. La extensión de este método puede contribuir sustancialmente a eliminar las mezclas varietales que conducen a grandes pérdidas en la economía y la industria azucarera. El análisis cuantitativo de estos marcadores moleculares brinda otras posibilidades de estudio genético de interés para la selección de nuevas variedades.

El sistema de programas que posibilita este estudio, se ha confeccionado de forma tal que le sea útil al especialista que lo va a emplear y sea de fácil manejo.

BIBLIOGRAFIA

1. Frias A.M., Antonic H. and Lozzin M.E. 79, 1975.
2. Waldron J.C. and Glasziou K.T. Proceeding of the 14th Congress I.S.S.C.T. Louisiana. *Breeding and Genetics*, 249, 1972.
3. Ralston A. Introducción al análisis numérico. Editorial Limusa, México, 1978.
4. Borland International Inc. Turbo Pascal V 3.0, Reference Manual, 1986.