



ОЦЕНКА НА МУТАНТНИ ФОРМИ ДОМАТИ И ТЕХНИТЕ ИЗХОДНИ ЛИНИИ ЧРЕЗ КЛАСТЕРЕН И ФАКТОРЕН АНАЛИЗ

И. ИВАНОВА, С. ГРОЗЕВА, В. РОДЕВА

ASSESSMENT OF TOMATO MUTANT FORMS AND THEIR INITIAL LINES BY CLUSTER AND FACTOR ANALYSIS

I. IVANOVA*, S. GROZEVA**, V. RODEVA**

* AGRICULTURAL UNIVERSITY – PLOVDIV

** INSTITUTE OF VEGETABLE CROPS „MARITZA” – PLOVDIV

E-mail: iren50@abv.bg

Abstract

There were estimated 7 newly obtained in vitro tomato lines and 4 initial parental forms by 10 determining morphological and biochemical indices aimed to ascertaining of genetic differences between the both groups of genotypes.

On the basis of performed cluster analysis they were classified in 3 clusters with different genetic distances.

In addition factor analysis was made to establish the indices with the highest influence on distribution of the genotypes in the received clusters.

Key words: Tomato mutant forms, In vitro, Cluster analysis, Factor analysis

УВОД

Известно е, че успехът на всяка селекционна програма зависи от включеното в работния процес изходно генетично разнообразие и неговото предварително и цялостно познаване. Проучването на новополучен растителен материал и различията му от родителските генотипове се основава на изследването на морфологични, качествени и генетически показатели с цел анализ и планираното му използване в отделните етапи от селекционния процес. Чрез прилагане на кластерния анализ е възможно определянето на различията между нееднократно включвани в селекцията родителски генотипове и получените от тях променени линии, притежаващи ценни характеристики, както и обособяването на кластери на базата на степента на генетическо сходство [1, 5]. С прилагането на факторен анализ се определят показателите с най-силно влияние върху разпределението на генотипите в получените кластери [3].

ЦЕЛ

С настоящето изследване си поставихме следните цели:

1) Да определим наличието на различия между получени *in vitro* мутантни линии домати и техните изходни родителски линии при групирането им по морфологични и биохимични показатели чрез кластерен анализ;

2) С прилагане на факторен анализ да се редуцира броят на изследваните показатели чрез обединението им и да се определят показателите с най-силно влияние върху разпределението на генотипите в получените кластери.

МАТЕРИАЛ И МЕТОДИ

В проучването участват седем новополучени мутантни генотипове (895, 896, 975, 985, 1527, 1624 и 1681) и четири изходни родителски линии (108, 120, XXIVA и 84602046), представители на *Solanum lycopersicum* L. Изследвани са следните морфологични и биохимични характеристики: индекс (1), маса (2), съдържание на сухо вещество (3), киселини (4), общи багрила (5), ликопин (6), бета-каротин (7) на плодовете, продуктивност (8) и височина на растението (9), ранозрялост (10). Данните са средни за периода 2007-2009г.

За определяне на еднородните групи по популационни средни е приложен кластерен анализ. Кластеризацията на обектите в групи е направена чрез йерархичен анализ и прилагане на метода на междугруповото свързване [2,6]. Като мярка за сходство е използвано Евклидовото разстояние. Данните са предварително стандартизирани, за да се избегне влиянието на различните дименсии. Групирането в кластери графично е представено с дендрограма, определяща последователността на обединяване на обектите и формираните кластери. Тъй като между много от изследваните характеристики съществува корелация, с цел да се намали размерността на пространството като подходящ подход е използван факторен анализ [3,4]. Той позволява да се определят обобщени фактори, които са по-малко на брой от изследваните характеристики, но всеки от тях съдържа в себе си едновременно свойствата на няколко от тях. Факторизирането е извършено чрез метода на анализ на главните компоненти. Подобен подход е използван и в [5]. При този метод факторите, които са линейни комбинации от корелиращи променливи, се получават с нулеви линейни корелации помежду си. Това дава възможност да ги интерпретираме като независими. Броят на главните компоненти се определя от броя на собствените значения на корелационната матрица, които са по-големи от 1 (критерий на Кайзер). Собствените значения показват приноса на съответния фактор при обясняване на общата дисперсия в наблюдаваните променливи. Факторният модел се определя най-вече от факторните тегла, които представляват корелационните коефициенти между съответните наблюдавани характеристики и факторите. Тъй като е трудно да бъдат интерпретирани в този си вид, за намиране на удобни за интерпретация фактори сме приложили допълнително ротирание на факторите чрез т.н. варимаксна трансформация. Така е възможна интерпретацията на всеки от факторите, както и определянето на

характеристиките, влияещи най-силно върху разпределението на генотиповете в кластери. Обработката е извършена чрез статистическата програма SPSS.

Резултати и обсъждане

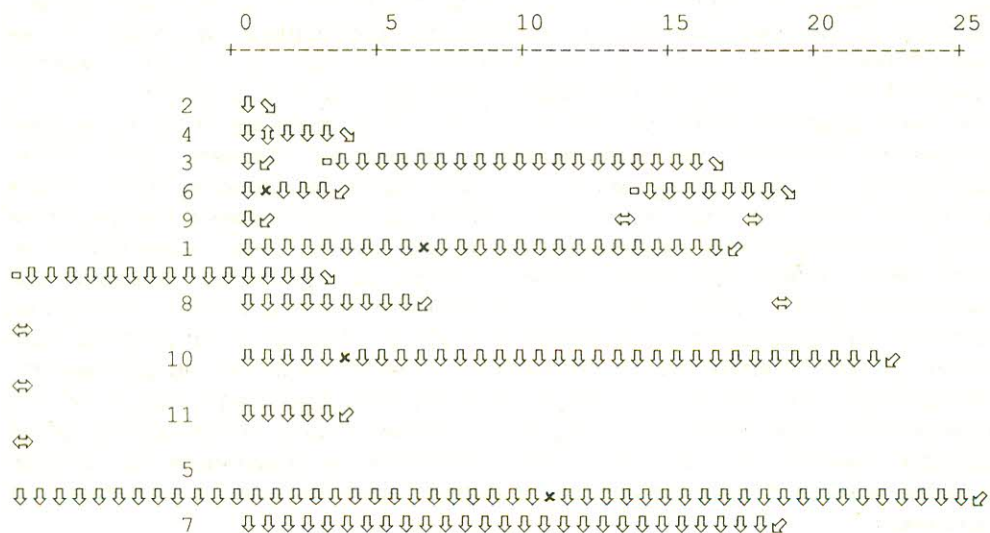
Групирането на проучваните 11 генотипа домати в отделни кластери е показано чрез дендрограмата на фиг.1, а междугруповите разстояния са посочени в табл 1. В резултат на проведения анализ са оформени 3 кластера. Първият е по-хомогенен и обединява линии 2 (985), 4 (1527), 3 (975), 6 (895) и 9 (1624), близки по комплекса от проучваните характеристики. Това е групата на мутантните линии, получените след *in vitro* облъчване на експланти от изходните линии 1 (120), 8 (108) и 5 (84602046). Резултатите потвърждават различията между изходните и получените мутантни генотипове. Вторият обединява изходните линии 1 (120) и 8 (108), а в третия попада както изходната линия 10 (XXIVA), така и получената от нея мутантна линия 11 (1681), което показва отсъствието на съществени различия в комплекса от проучваните характеристики. Линии 5 (84602046) и 7 (896) са с твърде голямо разстояние както помежду си, така и по отношение на другите линии, поради което не могат да се присъединят към оформените вече кластери.

Графично резултатите от проведения факторен анализ чрез метода на главните компоненти са представени на фиг.2. От участващите фактори само първите три са със собствени значения, по-големи от 1. Въз основа на този резултат броят на главните компоненти се определя на 3. Във факторната матрица е посочен процентът на общото разсейване, дължащ се на съответния фактор (табл. 2). Резултатите показват, че първият фактор изчерпва 38,7 % от общото разсейване, вторият – 32,9 %, а третият – 12,7 % или трите фактора обхващат 84,3% от общото разсейване и потвърждават графичния резултат, че първите три фактора са главните компоненти, достатъчни за факторния модел. В таблицата са дадени изчислените факторни тегла на характеристиките. Това са корелационни коефициенти между наблюдаваните характеристики и факторите. Тези характеристики, които имат големи факторни тегла ($> 0,55$), са най-силно свързани със съответния фактор. Допълнителното ротиране по метода Varimax позволява по-точна интерпретация на факторите (табл.3). Фактор 1 се свързва най-силно с продуктивността и обединява характеристиките продуктивност, ранозрялост, сухо вещество и маса, като връзката със сухото вещество е с отрицателен знак. Фактор 2 свързва с големи тегла показателите общи багрила, ликопин и киселини, като най-силно е влиянието на общите багрила. Фактор 3 включва каротин и индекс, като по-силно е влиянието на каротина.

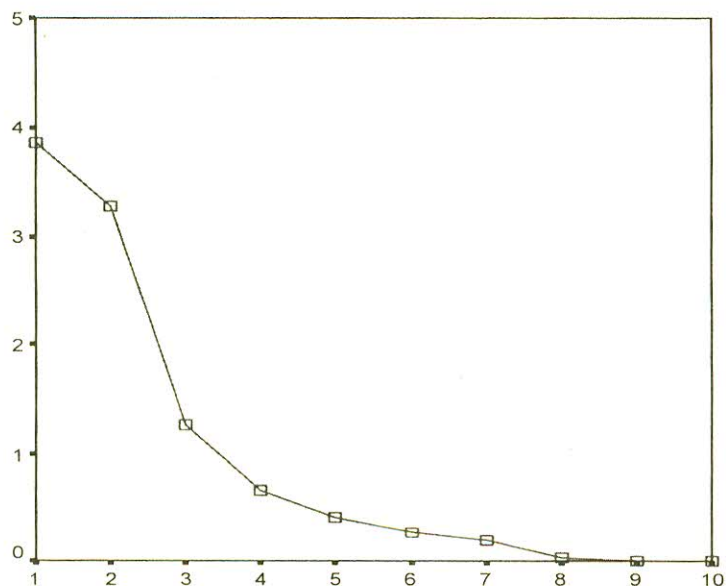
Таблица 1

Комбиниране на кластерите и междугрупови разстояния

Стъпки	Комбиниране на кластерите		Коефициенти
	Кластер 1	Кластер 2	
1	2	4	0,316
2	2	3	0,320
3	6	9	0,329
4	10	11	0,566
5	2	6	0,648
6	1	8	0,865
7	1	2	1,717
8	5	7	1,862
9	1	10	2,243
10	1	5	3,173



Фиг. 1. Йерархичен кластер анализ. Дендрограма на базата на средните междугрупови евклидови разстояния



Фиг.2. Факторен анализ чрез метода на главните компоненти. Стойности на собствените значения на корелационната матрица.

Таблица 2
Факторна матрица, получена чрез метода на анализ на главните компоненти

№	Показатели	Главни компоненти		
		1	2	3
1.	Индекс	-0,768	0,017	0,570
2.	Маса	0,624	0,621	-0,145
3.	Продуктивност	0,663	0,429	0,511
4.	Сухо в-во	-0,329	-0,781	-0,140
5.	Киселини	0,769	0,421	0,144
6.	Общи багрила	-0,333	0,877	-0,082
7.	Ликопин	-0,514	0,810	0,045
8.	Каротин	0,824	-0,136	0,714
9.	Ранозрялост	0,715	0,474	0,171
10.	Височина	0,425	-0,223	0,572
Процент от общото вариране (%)		38,7	32,9	12,7

Таблица 3

**Ротирана факторна матрица, получена чрез варимакс
трансформация на главните компоненти**

№	Показатели	Главни компоненти		
		1	2	3
1.	Индекс	-0,200	-0,121	-0,892
2.	Маса	0,694	-0,274	0,489
3.	Продуктивност	0,950	0,217	-0,005
4.	Сухо в-во	-0,758	0,398	-0,072
5.	Киселини	0,396	0,575	0,418
6.	Общи багрила	0,310	-0,864	-0,209
7.	Ликопин	0,222	-0,834	-0,421
8.	Каротин	0,157	0,235	0,939
9.	Ранозрялост	0,813	0,023	0,324
10.	Височина	0,147	0,390	-0,107

Изводи

1. Резултатите от кластерния анализ доказват възможността чрез *in vitro* методите да се получат мутантни линии, носители на нови селекционни признаци и значително различаващи се от изходните генотипове.

2. Чрез факторния анализ продуктивността, ранозреლობта и биохимичните признаци са определени като най-силно влияещи върху разпределението на генотиповете в получените кластери.

Литература

1. Ганева Д., И. Иванова, Г. Певичарова, 2006. Идентифициране на детерминантни F₁ хибриди чрез използване на кластер анализ. Сб. доклади „Екологични подходи при производството на безопасни храни“, Пловдив, 205-210.
2. Дюран, Б., Оделль П., 1977. Кластерный анализ, М., Изд. Статистика.
3. Gorsuch R., 1983. Factor Analysis, Second Edition, Lawrence Erlbaum Assoc., Inc., Publishers New Jersey
4. Kline P., 1994. An easy guide to factor analysis, Routledge, London.
5. Krasteva L., D. Dimova. 2004. Evaluation of a canning determinate tomato collection using cluster analysis and principal component analysis., Proc. III Balkan Symposium on Vegetables and potatoes, Acta horticulturae. 729. ISHS. 89-93.
6. Ward J. H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association. 58. 236-244.