ID: 794

A Study on Chromatographic Analysis and Machine Learning in Strawberry Breeding

Maryam Akbari[1] İlbilge Oğuz[2]*, Halil İbrahim Oğuz[3], Nesibe Ebru Kafkas[2]

[1] *University of Tehran, Department of Agronomy and Plant Breeding Science, College of Aburaihan, Tehran, Iran*
[2]*Çukurova University, Faculty of Agriculture, Department of Horticulture, TR-01330, Adana, Türkiye.*
[3]*Adıyaman University, Faculty of Agriculture, Department of Horticulture, TR-02040, Adıyaman, Türkiye.*

Abstract

In recent years, there have been remarkable advances in determining the biochemical properties of berry fruits. In fruit breeding, GWAS and QTL mapping studies are fundamental statistical methods used to identify candidate genes that influence fruit quality parameters, particularly flavour and nutritional quality. As is well known, the cultivated strawberry (*Fragaria* × *ananassa*) is a plant with a highly complex genetic structure due to its octoploid constitution (2n=8x=56). For this reason, machine learning algorithms—particularly artificial neural networks—offer significant advantages in strawberry breeding research for modelling the complex, non-linear relationships between genotype and phenotype. Furthermore, artificial intelligence techniques offer significant benefits in interpreting model predictions, integrating genomic, transcriptomic, proteomic and metabolomic data, and more accurately identifying flavour mechanisms, thereby saving time and ensuring the right targets are met. For example, the rapid and accurate identification of the biochemical and genetic characteristics of strawberry fruits is of vital importance for breeding purposes. Also, the metabolic profiles—such as sugars (glucose, fructose, sucrose), organic acids (citric, malic, ascorbic) and volatile aroma compounds—which constitute each fruit's unique flavour profile, are particularly the most distinctive fruit quality parameters in strawberries. Additionally, the use of artificial neural networks (ANNs) to model the complex, non-linear relationships between genotype and phenotype in strawberry breeding has gained significance in breeding research in recent years. Moreover, convolutional neural networks (CNNs) offer revolutionary advantages in processing and analysing visual data, thereby facilitating the achievement of accurate and successful results in strawberry breeding.

The aim of this study is to explain how breeding efforts can be accelerated by using machine learning algorithms to digitise the phenotyping stages—which are time-consuming, costly and prone to human error in traditional breeding processes—using machine learning algorithms, thereby explaining how breeding efforts can be accelerated. This study therefore provides practical guidance on how to select high-quality, superior strawberry lines and parental combinations at an early stage with a high degree of accuracy, and on how this can improve the success rate of the breeding programme.

Keywords: *strawberry, fruit quality, strawberry breeding, machine learning*