

## **Diversidade genética de *Capsicum* spp. por meio de aprendizado de máquina**

**Ana Izabella Freire<sup>1</sup>; Anabela L Mazivile<sup>2\*</sup>; Anderson P de Paiva<sup>1</sup>; Filipe B M de Souza<sup>2</sup>; Tiago M de Azevedo<sup>1</sup>; Valter C de Andrade Júnior<sup>2</sup>**

<sup>1</sup>UNIFEI – Câmpus de Itajubá, CEP: 37500-903, Itajubá – MG, Brasil; anaizabellinha2014@gmail.com; andersonppaiva@unifei.edu.br; tiago.deazevedo@yahoo.com.br; <sup>2</sup>UFLA – Câmpus de Lavras, CEP: 37200-000, Itajubá – MG, Brasil; anabela.mazivile@estudante.ufla.br; fbmsouza@yahoo.com.br; valter.andrade@ufla.br

**\* Apresentador do trabalho no 57º CBO**

### **RESUMO**

O estudo de diversidade genética de *Capsicum* spp. é fundamental para atender às demandas do mercado e o uso de descritores morfoagronômicos é uma das formas utilizadas. Técnicas de aprendizado de máquina podem auxiliar na seleção de caracteres que permitem a escolha de genótipos divergentes. O objetivo deste trabalho foi desenvolver um algoritmo de classificação de descritores de *Capsicum frutescens* e *C. chinense* por meio de técnicas de aprendizado de máquina. A metodologia de Data Mining é o principal suporte para a elaboração do modelo de previsão de diferenciação dos caracteres de *Capsicum* spp. Avaliaram-se 39 genótipos, através de 13 descritores morfológicos multicategóricos. A análise incluiu a estimativa da distância de Mahalanobis, da Distância Média Euclidiana, da similaridade e dissimilaridade, bem como o agrupamento de Tocher e a rede de correções. No caso das variedades de *C. frutescens*, os descritores peso de fruto (PF), massa fresca (MF), comprimento do fruto (CF) e espessura de pericarpo (EF) contribuíram com 89,95% da divergência genética, enquanto para *C. chinense*, essa contribuição foi de 68,72%. As correlações mais relevantes para os caracteres de plantas foram altura de planta e longitude de caule. Para os caracteres de fruto, destacaram-se PF, MF, massa seca, diâmetro do fruto e CF. Por meio do método de otimização de Tocher, os genótipos foram subdivididos em 10 grupos. Os caracteres selecionados contribuem para a discriminação da diversidade genética em *Capsicum* spp. O uso de algoritmos de classificação é eficiente na avaliação de caracteres morfoagronômicos e na seleção de genótipos divergentes em programas de melhoramento.

**PALAVRAS-CHAVE:** *Capsicum chinense*, *Capsicum frutescens*, descritores, algoritmo, diversidade.

### **AGRADECIMENTOS**

Os autores agradecem à FAPEMIG, CNPq e Capes.