

# Systemic study of the evolution of flowers

Beatriz Moura Kfoury de Castro<sup>1</sup>, Tetsu Sakamoto<sup>2</sup>, Carlos Alberto Xavier Gonçalves<sup>1</sup>, José Miguel Ortega<sup>2</sup>,

*1 UFMG*

*2 Universidade Federal de Minas Gerais, Laboratório de Biodados*

## Abstract

Flowers are recent innovations in the evolutionary history of plants on the geological timescale of plant diversification. They are the reproductive structures of angiosperms (flowering plants), which constitute the most diverse and cosmopolite group of plants. In order for flower to be formed, a complex gene regulatory networks control the floral development, involving several genes. The main purpose of this work was to collect the current knowledge about the molecular basis of floral development and assemble it in the format of a metabolic pathway. Since flowers are recent innovation among angiosperms, we also investigated the evolutionary origin of flowering genes, verifying if they have originated recently along with the appearance of flowering plants. To retrieve genes associated to flowering process, we collected 1000 scientific articles depicting the molecular biology of flowering. The list with articles of PubMed IDs (PMIDs) resulting from search were submitted to text-mining tools, to assist us on determining genes and biointeractions described on them. After this selection, construction of the metabolic pathway was done with manual curation. To analyze the evolutionary origin of those genes retrieved from the text-mining, their sequences were retrieved from Uniprot database, their orthologs from other species were retrieved using SeedServer and the Lowest Common Ancestor (LCA) was determined and used to infer the gene origin. Phylogenetic inference was also applied to analyze in more depth the evolution of some genes in the pathway. As result, we found 80 genes to be linked to flowering process. Among them, 20 belong to MADS-box family, demonstrating its relevance on shaping the flower. Gene origin inference indicated that genes comprising the floral development pathway appeared on different clades during the evolutionary history of the plants. However, some proteins of MADS-box family seem to have originated during the conquest of the terrestrial environment by the plants, well before the appearance of the flower structure. However, some genes were found exclusively on the clade of the angiosperms (Magnoliophyta) and non-basal angiosperms (Mesangiospermae). A more in-depth investigation of the evolution of MADS-box genes participating in flower development showed that recent and multiple duplications have taken place on this family among the flowering plants. Since the period of the expansion of this family coincides with the origin of angiosperm, it suggests that these events have contributed greatly to the appearance of the floral structure.

Funding: UFMG