

Charting Natural Product Chemical Space: Cheminformatics and Artificial Intelligence for Drug Discovery

José L. Medina-Franco

DIFACQUIM Research Group, Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico

medinajl@unam.mx

Abstract

Natural products (NPs) have long served as a rich source of bioactive molecules and privileged scaffolds for drug discovery, yet their systematic exploration remains challenging due to their structural complexity and diversity. In this plenary lecture, I will discuss how modern cheminformatics and artificial intelligence (AI) are transforming the exploration of NP chemical space at unprecedented scale. I will highlight recent efforts to curate, standardize, and analyze large NP collections, with particular emphasis on the Latin American Natural Product Database (LANaPDB), an open-access resource that integrates nearly 15,000 nonredundant compounds from multiple countries across the region. Through comprehensive chemoinformatic characterization—encompassing structural diversity, molecular complexity, synthetic feasibility, biological annotations, and ADMET-related properties—I will illustrate how NP libraries compare with approved drugs and global NP repositories. I will further discuss how machine learning and pattern-recognition approaches enable the identification of bioactive motifs, guide NP-inspired design, and support target selectivity. Overall, this lecture will showcase how scalable AI-driven workflows, combined with curated chemical knowledge, are redefining the role of natural products in contemporary computer-aided drug discovery.