

教師なし機械学習型 AI によるゲノムビッグデータからの想定外の発見: 「Mb-level CpG and TFBS islands」 とそれらの核内配置における役割

池村 淑道

長浜バイオ大学 客員教授・名誉教授

国立遺伝学研究所・総合研究大学院大学 名誉教授

生命科学の多様なデータがビッグデータ化している。その解析には、モデルや予備知識なしに様々な情報を抽出できる「教師なし機械学習」が威力を発揮する。また、機械学習により、社会的に重要な課題（例えば強毒性ウイルスに対する対策）への応用も可能になってきた。池村先生は、ゲノム配列上の連続塩基（例えば4連や5連）の出現頻度に着目した機械学習を行うことで、これまでに様々な発見をされ、約30報の論文として発表してこられた。今回は、ヒトゲノムに見出した「Mb-level CpG and TFBS islands」の特徴、並びにそれらの核内配置やクロマチンにおける機能について紹介して頂く。