

MCBJ 法を用いた DNA シーケンシングによる連続塩基数識別法の開発

Discrimination of homo polymer base length with DNA sequencing by using of MCBJ method

阪大産研¹ ○小本 祐貴¹, 大城 敬人¹, 谷口 正輝¹

Osaka Univ., ISIR¹, °Yuki Komoto¹, Takahito Ohshiro¹, Masateru Taniguchi¹

E-mail: komoto@sanken.osaka-u.ac.jp

【緒言】

Mechanically Controllable Break Junction (MCBJ)法は、金属細線を描画した弾性基板を押し曲げ、金属細線を破断させることにより生じたナノギャップを流れる電流を計測し、ナノギャップ中に存在する分子の伝導度を計測する手法である。MCBJ 法は、単一分子を介した電流を測定できるために、微小分子の検出・識別技術として注目を集めている。MCBJ 法を用いた単分子計測の最も注目を集めている実用例は、DNA シーケンサーとしての利用である。MCBJ 法を用いた単分子 DNA シーケンサーは、直接 1 塩基を検出できるため、高速かつ、従来の手法では検出できないメチル化などのエピゲノムのシーケンスの可能性が期待されている。これまでに、MCBJ 法を用いた DNA、RNA シーケンスの結果が報告されてきた。従来の解析手法では、計測した電流値から対応する塩基を読み取る。この手法では連続した同一塩基の塩基数を決定することは不可能で、連続塩基を含む DNA, RNA の配列を識別、決定することができない。生物の多くは、mRNA にポリアデニン鎖と呼ばれる、連続したアデニン塩基の配列を有し、ポリアデニン鎖の長さが RNA の翻訳に関与している。したがって、従来の解析手法では判別の難しかった連続塩基の塩基数を判別することを可能にする必要がある。そこで本研究では、連続した異なる長さの dAMP 鎖の識別を目的に実験・解析を行った。

【実験・解析手法】

識別対象は dAMP を連続して 6 塩基、及び 7 塩基有する配列の DNA、A6, A7 とした。A6, A7 の水溶液の MCBJ 法による単分子計測を行い、それぞれのシグナルを得た。これらのシグナルの一部を学習用データ、一部をテスト用データとして機械学習を用いて分類した。

【結果・考察】

得られたシグナルを 10 分割し、それぞれの平均電流値を特徴量として識別を行った結果を Fig.1 に示す。機械学習を用いた識別の結果、精度 0.60 での識別に成功した。機械学習により、長さの差異による挙動の差異を学習し、識別することができたと考えられる。本研究では、長さの異なる連続塩基数の検出法の開発に成功した。

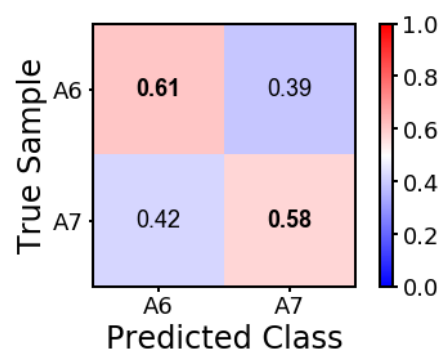


Fig.1 Confusion matrix of classifying between A6 and A7 DNA homo-polymers