

## Bivartect : リード直接比較による高精度・省メモリ使用量な変異検出法

事前のマッピングを必要とせず、リード直接比較によって遺伝子変異を検出する、簡便で正確な算出方法である”Bivartect”(bit-based variant detection)を開発した。メモリ使用量を削減して高速化を図るため、Bivartect では全ての文字列をビット列に変換し、ブレイクポイント検出の際にはリード接尾辞のごく一部だけをメモリ空間に保持している。さらに、共通の接頭辞を持つ、正常リードと変異リードの接尾辞の一部をソートすることで、ブレイクポイント候補を検出する方法を採用した。Bivartect は単一ヌクレオチドの変異(SNVs)だけでなく、挿入(Insertions)、欠損(Deletions)、反転(Inversions)やこれらの複合型も検出することができる。小規模のオミックスデータをシングルノードのコンピュータで解析できるように、高度なメモリ節約の仕組みのもとで高い予測精度を実現した。シミュレーションで生成したベンチマーク用のデータと実際のゲノム編集のデータを用いた実験では、Bivartect が出力した変異予測数はかなり少量であったにもかかわらず、SNV 検出の陽性的中率において最先端のバリエーションコーラーと同等であると示された。これらの結果は、リファレンスフリーの方法を用いた Bivartect が、生殖細胞の変異やゲノム編集にもたらされたオフターゲット部位を高精度で特定する際に有用であることを示している。

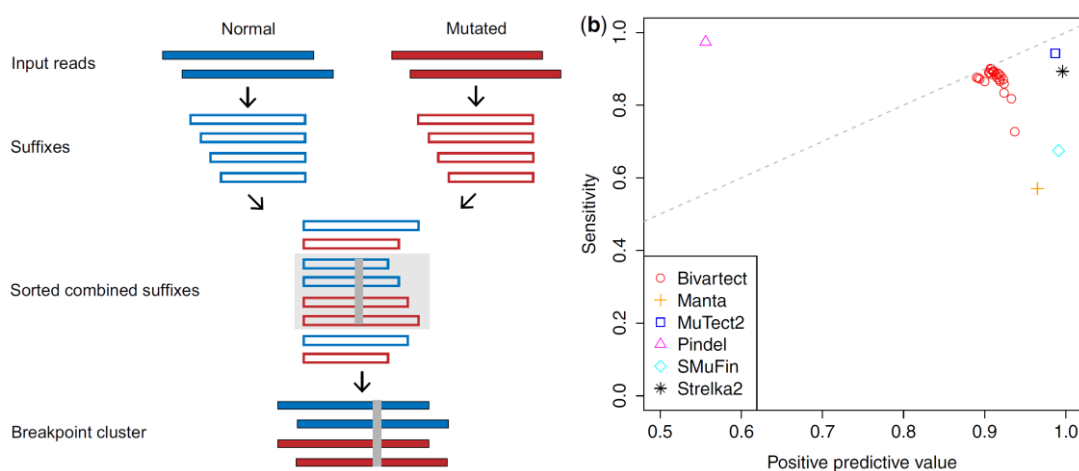


Fig. K.Shimmura et al, Bioinformatics, 2020