

# Genome Box を用いた 高速シーケンスデータのマッピングおよび 大腸菌株間のゲノム比較

Genome Box では、高速シーケンサーから出力された大量配列データを簡便にマッピングし、ゲノムブラウザ上にマップ結果を表示することが可能です。本アプリケーションノートでは、高速シーケンサーによる Whole Genome Sequencing により得られた、複数の大腸菌株のゲノムシーケンスデータを用いて、菌株間のゲノム比較を実施した例をご紹介します。

## 序論

高速シーケンサーの登場により、ラボレベルでも容易かつ比較的安価にゲノムシーケンスが実施できる環境が整ってきました。

リファレンスとなるゲノムが既に解読されている生物種、ないし近縁種・株においては、高速シーケンサーで Whole Genome Sequencing によって得られたゲノム配列を、比較対象となるリファレンスゲノム配列に対してマッピングすることにより、簡易的なゲノム比較解析を実施することが可能です。

本アプリケーションノートでは、米国 National Center for Biotechnology Information (NCBI) の The Sequence Read Archive (SRA) に登録されている複数の *Escherichia coli* 株の高速シーケンサーによる Whole Genome Sequencing データを、腸管出血性大腸菌 O-157 株をリファレンスゲノムとしてマッピングした結果を、Genome Box によって可視化することで、直感的に病原性遺伝子の有無と、当該大腸菌の病原性有無とが合致することが確認できました。

## 方法

SRA より、腸管出血性大腸菌株 O-157 株 (ERR018562)、O-104 株 (SRR292678) および非腸管出血性大腸菌 K-12 株 (SRR001666) の Whole Genome Sequencing データを取得し、*E. coli* O-157 株のゲノムおよびプラスミド配列 (BA000007、NC\_002127、NC\_002128) に対して、Genome Box 上で Burrows-Wheeler Aligner (BWA) によるマッピングを実施、結果をブラウザに表示しました。

腸管出血性大腸菌が産生する志賀毒素は大別して2種類 (Stx1、Stx2) 存在します。毒素遺伝子 *stx1* (ECs2973、ECs2974) および *stx2* (ECs1205、ECs1206) 領域を Genome Box 上で検索し、該当領域を目視により確認しました。

## 結果

### *stx1* 遺伝子領域 (図1)

*stx1* 領域 (赤枠領域) の O-157 株、O-104 株、K-12 株のリードが該当ゲノム領域において、マッピングされた depth を

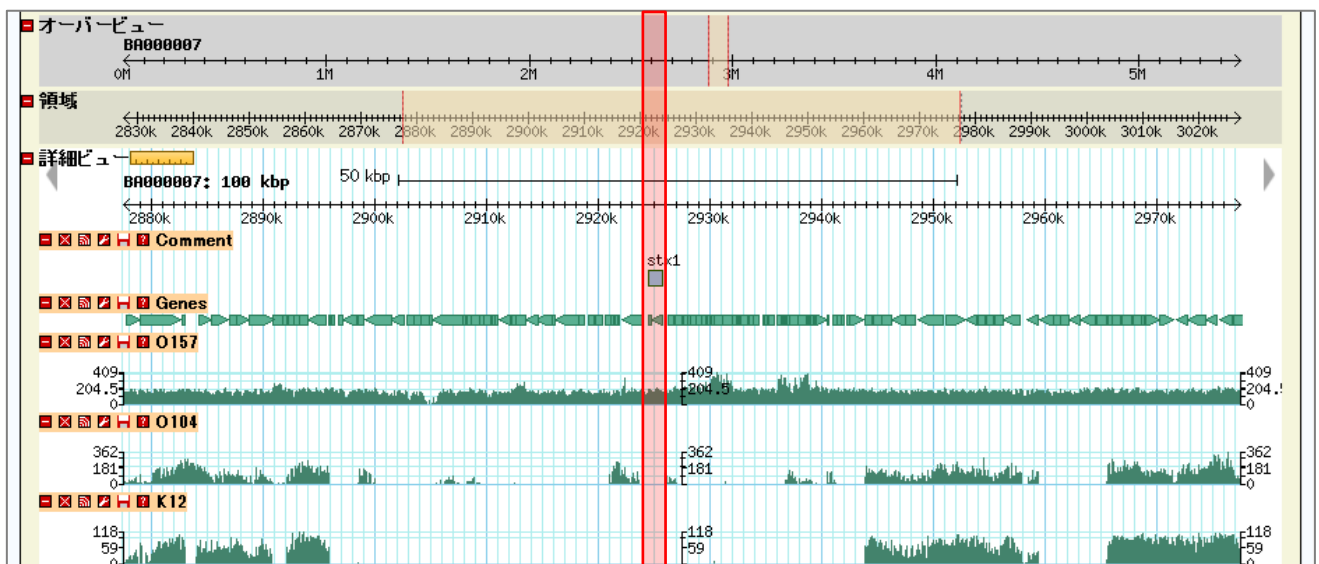


図1. *stx1* 遺伝子領域近傍のマッピング結果

# Application Note

グラフの高さで表わした結果です。O-157 株は *stx1* 領域を含め全体にわたってマッピングされていることが分かります。これに対して O-104 株および、K-12 株に関しては、*stx1* 領域にリードがマッピングされませんでした。

*stx1* 遺伝子は病原性株 O-157 には存在し、本解析に供した病原性株 O-104、非病原性株 K-12 には存在しないことが知られており、Genome Box 上で観察されるマッピング結果と合致することが見て取れます。(※*stx1* を持つ O-104 株の報告もあります)

## *stx2* 遺伝子領域 (図2)

*stx2* 領域 (赤枠領域) の O-157 株、O-104 株、K-12 株のリードが当該領域において、マッピングされた depth をグラフの高さであらわした結果です。O-157 株および O-104 株は *stx2* 領域にリードがマッピングされていることが分かります。これに対し K-12 株に関しては、*stx2* 領域にリードがマッピングされませんでした。

本解析に供した O-104 株の毒素は *stx2* によるものであることが知られており、Genome Box 上で観察されるマッピング結果と合致することが見て取れます。

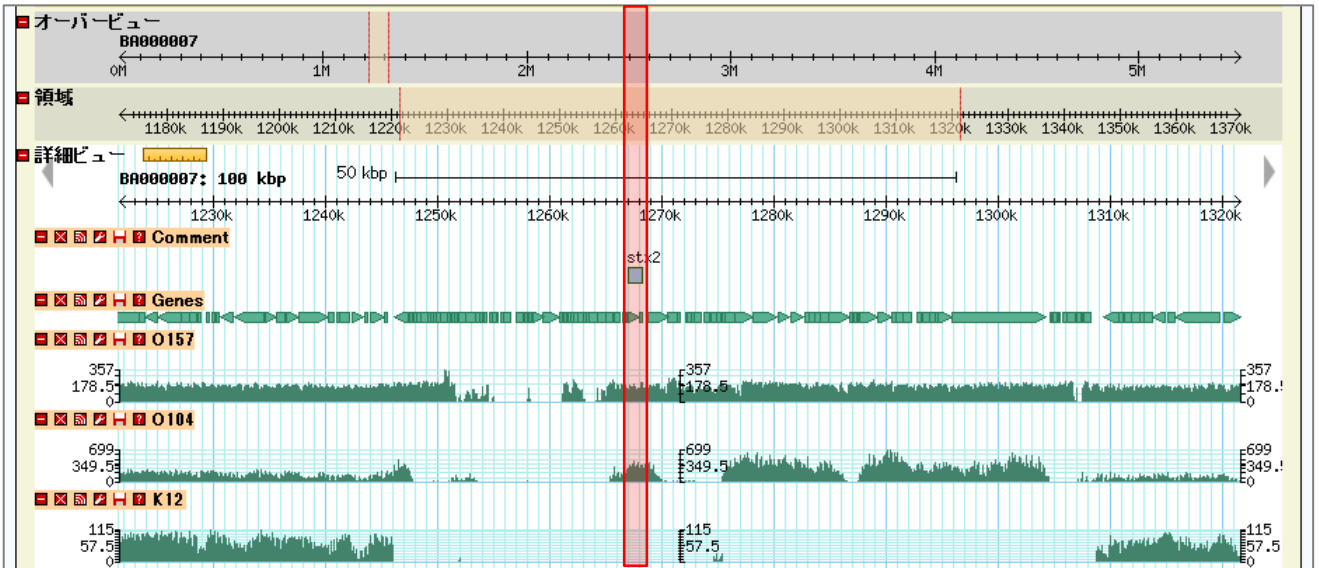


図2. *stx2* 遺伝子領域近傍のマッピング結果

## まとめ

ゲノムブラウザと簡易解析機能が統合されている Genome Box は、高速シーケンサーからの大量データであっても十分に適用が可能です。簡便なゲノム解析のために、大規模サーバのようなインフラへの投資をする必要が無く、比較的小規模のプロジェクトや、高速シーケンシングの委託による少数の結果を解析、閲覧するためには最適なプラットフォームとなっています。

## 参考文献

Hayashi T, et al. (2001). *DNA Res.*, 8, 11-22.  
Mellmann A, et al. (2011). *Plos One.*, 6, e22751.

## Genome Box シリーズについて

Genome Box シリーズに関する詳細は  
<http://www.genome-box.com/>  
をご参照下さい。

## ビット株式会社

Web : <http://www.bits.cc/> Mail : [hts@bits.cc](mailto:hts@bits.cc)

東京本社  
〒101-0023 東京都千代田区神田松永町5 第二砂川ビル201  
Tel/Fax : 03-3255-1715

三島研究所  
〒411-0943 静岡県駿東郡長泉町下土狩20-3 山光ビルA402  
Tel/Fax : 055-989-5560

代理店