同時発表：農政クラブ、農林記者会、科学記者会、

公益財団法人
かずさDNA研究所

ゲノム配列から同定した遺伝子の機能予測を最速で行うプログラムを開発しました

～Hayai-Annotation Plants～

５月１５日に Bioinformatics 誌にてオンライン公開

令和元年５月２７日

公益財団法人　かずさＤＮＡ研究所

- ✧ かずさＤＮＡ研究所は、植物のゲノム解析において重要な、遺伝子配列から遺伝子の働きを推定するプログラム「Hayai-Annotation Plants」を開発しました。
- ✧ ゲノムを解析し他の植物種と比較するには、配列データを解析してどこにどのような遺伝子があるかを注釈づける（アノテーション）作業が必要です。
- ✧ これまでアノテーション作業には大型の計算機が必要でしたが、開発したプログラムは通常のパソコンレベルのスペックで解析できます（現在は Mac 版 OS にのみ対応）
- ✧ 計算は従来法に比べて非常に高速で、さらに精度の高い結果を得ることができます。
- ✧ 研究成果は、国際科学雑誌 Bioinformatics に５月１５日付でオンライン公開されました。

＜報道に関すること＞
公益財団法人かずさＤＮＡ研究所　広報・研究推進グループ　TEL：0438-52-3930

＜研究に関すること＞
公益財団法人かずさＤＮＡ研究所　植物ゲノム・遺伝学研究室
室長　磯部　祥子（いそべ　さちこ）　TEL：0438-52-3935

1. 背景

次世代シークエンサーなどの DNA 配列解析技術の向上により、多様な植物種のゲノム配列が解読されています。かずさＤＮＡ研究所でも近年、トマトやイチゴなど多くの実用作物のゲノムの解析を行っています。実用植物の種・品種間のゲノム配列の違いによって多くの形質の差異が生じることから、それらのデータを比較して育種に利用するには、配列データを解析してどこにどのような遺伝子があるかを注釈づける(アノテーション)作業が必要です。

アノテーションには、構造アノテーションと機能アノテーションがあります。構造アノテーションとはゲノム配列から遺伝子として機能する配列の部分を見つけ出すことで、機能アノテーションとは見つけ出された遺伝子配列がどのような機能をもっているのかを推定することです。機能アノテーションでは、これまでに世界中で解析された様々な生物種の遺伝子のもつ情報が格納されているデータベースを用いて、配列の類似性などから調べたい遺伝子がどのような機能を持っているかを推定します。検索の対象となる配列数が膨大であることから、これまでは大型計算機サーバーを用いて計算を行う必要がありました。

今回開発した「Hayai-Annotation Plants」は、機能アノテーションをつけるためのプログラムで、植物を対象としています。各遺伝子に対して、1) タンパク質の名前、2) 遺伝子オントロジー（GO；遺伝子の機能や細胞内局在を示すタグのようなもの)：3 つの GO カテゴリーから推定されたもの、3) EC 番号（酵素の機能を分類するためのもの)、4) 機能推定の精度レベル 、5) 推定の根拠となる情報の由来、を速く、かつ高精度で付与することができます。これらのデータは主にゲノム解読の際に得られた遺伝子配列の機能を推定する際に利用されます。

かずさＤＮＡ研究所では、これらの情報も含めたさまざまな植物のゲノム情報を集めたポータルサイト、Plant GARDEN（ Genome And Resource Database Entry: https://plantgarden.jp/app/pc_01.html）を運営し、統合データベース PGDB（Plant Genome Database Japan: http://pgdbj.jp/index.html ）において検索に有効なリソースの整備を行っています。

2. 研究成果の概要と意義

① 高速な計算を実現するために、植物を対象としたアノテーションに特化させました。
② データベースの配列セットを分割し、検索を並列処理することによってラップトップ型パソコンでの処理を可能にしました。
③ 機能アノテーションは次の４段階で実施し、推定された機能の精度を評価することができます。1) 対象とする種でアミノ酸配列が存在している、2) 転写産物の配列が存在している、3) モデル植物で類似の遺伝子配列が存在している、4) アミノ酸配列の推定のみ、の順で評価し、これにより精度の高い結果を得ることができました。
④ シロイヌナズナの遺伝子配列データセット（約 3 万遺伝子）を用いたテストでは、およそ 5 分で解析が終了しました。

3. 将来の波及効果

① 大型計算機サーバーや特別な専門性がなくても遺伝子アノテーションを実施できることから、大学などのアカデミアのみならず、企業の研究開発部署においても多様な植物種の遺伝子機能推定が進むことが期待されます。

② 植物のゲノム解析が加速され、多様な植物種でゲノム情報をもとにした育種が可能になります。

この研究は、かずさ DNA 研究所、および、JST ライフサイエンスデータベース統合推進事業の助成によって行われました。

論文タイトル： Hayai-Annotation Plants: an ultra-fast and comprehensive functional gene annotation system in plants.

著者：Andrea Ghelfi, Kenta Shirasawa, Hideki Hirakawa and Sachiko Isobe.

掲載誌：*Bioinformatics*
DOI：10.1093/bioinformatics/btz380

参考資料