

3 ゲノム予測技術を用いたサガンスギの無花粉化に関する研究

将来を見据えた新たな試験研究推進事業（開発事業）

（県単：R5（2023）～R7（2025））

江島淳 柴田晴矢

1 目的

近年、無花粉遺伝子を有するスギ品種が極まれに存在することが分かり、それらの無花粉遺伝子を利用することで、無花粉化できる。無花粉遺伝子はDNA上の1箇所の変異で決まる潜性（劣勢）遺伝子で、人工交配によりホモ型にすれば無花粉化でき、DNA分析により容易に同定が可能な質的な形質である。

一方、サガンスギなどが有する優れた成長や材質などの形質は、DNA上の多数の変異の影響により発現される量的な形質と考えられており、DNA情報からの予測が難しいとされてきた。しかし、2010年代に登場した次世代シーケンサーと呼ばれるDNAの高密度な判読技術により、約10,000箇所のDNA上の変異（SNP）を検出し、それらの多数の箇所のDNA上の変異と材質などの形質情報を関連付け、DNA情報だけから形質を予測するモデル（ゲノミック予測：通称GP：Genomic prediction）が構築可能となりつつある。しかしGPに必要な大量の一塩基多型（以下SNP：Single Nucleotide Polymorphism）の情報を得るRAD-seq解析を外注する場合、高額なコストがGPの導入を難しくしている。

本研究課題では、RAD-seq解析費用のうち、高額な費用がかかるDNAライブラリー調整技術を習得し、SNPデータを取得し、GPを試行する。1年目（2023年度）は、GPの理論的背景を整理し（令和5年度（2023）業務報告書を参照）、DNAライブラリー調整技術を研修により習得する。2年目（2024年度）は、DNAライブラリー調整に必要な試薬を調合し、佐賀県林業試験場の実験室でDNAライブラリー調整を行う。また、次世代シーケンス解析後、得られたSNPデータをもとに系統解析を行う。3年目（2025年度）は、DNAライブラリー調整技術を確実なものとし、シーケンス解析後、得られたSNPデータをもとにGPを試行し、GPを実用化するための基盤を整える。

2-1 実験内容

DNAライブラリー調整に必要な、以下の4つのカスタムプライマーを入手後、アダプターをアニーリングして、試薬を調整した。また、DNAの精製に使用するAMPureを作成し、実験に必要な試薬を全て整えた。佐賀県林業試験場の実験室において192個体のDNAを対象に、ライブラリー調整作業を実施後に、次世代シーケンサーによる解析を外注した。

CS1_PstI & CS1_3' overhang_common

CS2_SphI & CS2_3' overhang_common

2-2 SNP データ解析内容

前年の2023年度に次世代シーケンサーを用いて取得したリード（DNAの配列データ）を、2023年に発表されたスギの全ゲノム参照配列

(https://forestgen.ffpri.go.jp/jp/info_sugi1.html) に対してマッピングした。マッピングされたリード情報をもとに SNP を抽出した。その後、解析に利用できる品質の SNP だけにするため、フィルタリングという作業を実行した。これらの解析は、森林総合研究所樹木分子遺伝研究領域の協力のもと実行した。

2-3 系統解析の方法

得られた高品質の SNP から各個体の遺伝距離を算出し、個体間の遺伝距離（ユークリッド距離）をウォルド法により系統樹を作成した。その際、解析結果が正しいか確認するため、同一のサンプルに対して当試験場で品種の同定作業などに使用している4座を用いた SSR（Simple Sequence Repeat）解析による親子解析を実施し、SNP データをもとに作成した系統樹の各個体の配置と照合した。

3 解析結果

192 個体の SNP データを抽出した結果、175 個体において精度の高い SNP データが得られた。解析に用いることができる比較的品质の高い SNP 数は、10,389 座あり、スギ全ゲノム参照配列（遺伝地図）の11個の染色体上の特定の位置にマッピングされた（表-1）。この10,389座の SNP のうち、175 個体の中で1座でも欠測値を含むものを除外し、最終的に5,510座の SNP データを用いて系統樹を作成した。加えて、SSR 解析による親子関係の結果と、SNP データに基づいて得られた系統樹における個体の配置を照合した（図-1）。

図-1に示すように、交配の組み合わせごとに、個体は系統樹上の各ブロックに配置された。また、SSR 解析から外部花粉が混入したと判定された複数個体についても別ブロックに配置された。今回、B-74 と B-16 の雌雄逆の交配個体も解析対象としたが、雌雄逆でおこなった場合、多くの個体が別ブロックに配置された。この点については、今後、系統樹の偏りの原因となった可能性のある SNP 座について調査することで、原因を明らかにしていく予定である。

表－1 染色体別のSNP数

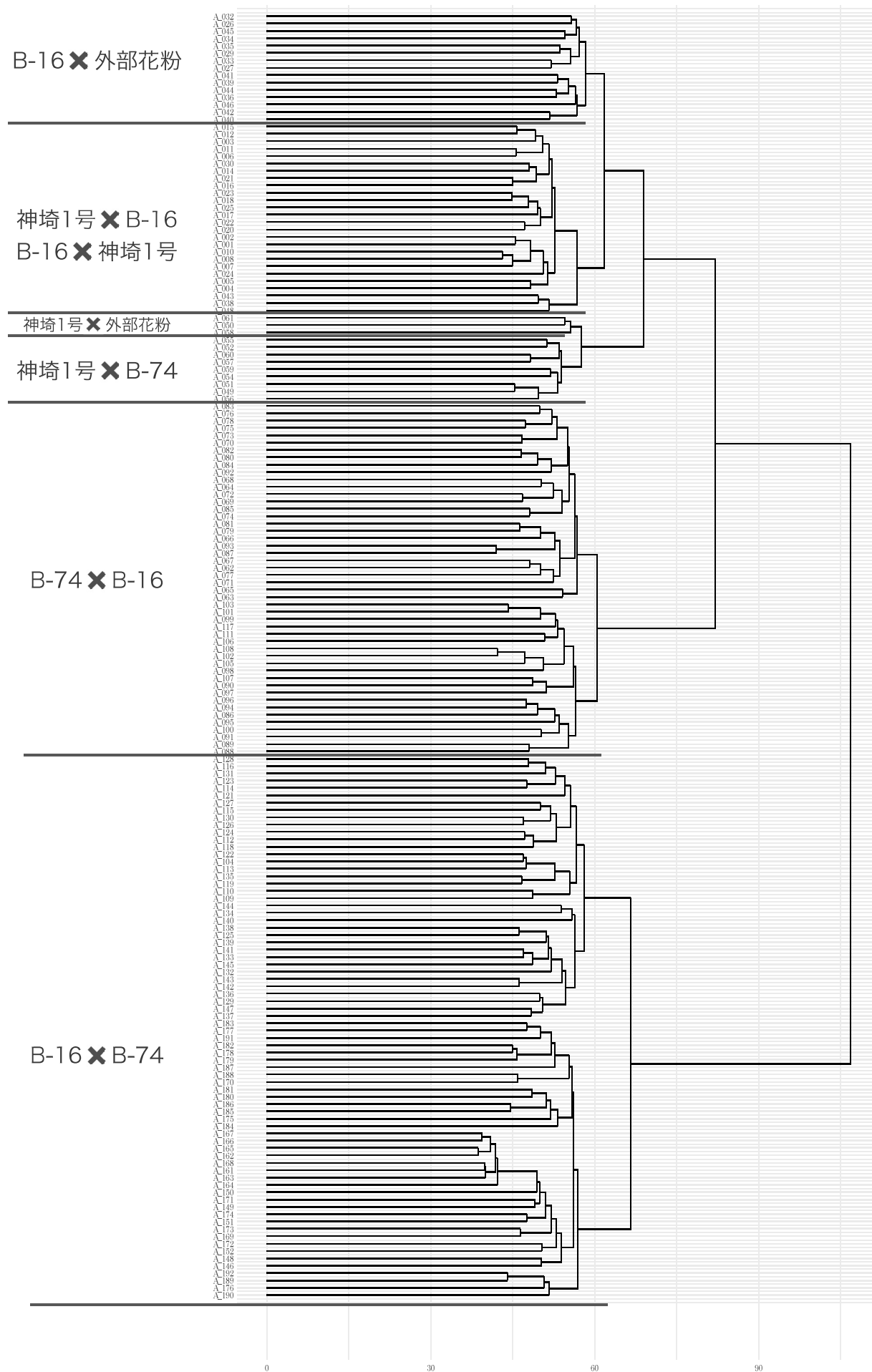
染色体番号	SNP
1	701
2	842
3	1,412
4	891
5	1,015
6	750
7	1,052
8	749
9	901
10	1,129
11	947
合計	10,389

4 考察

今年度は、前年の2023年度に森林総合研究所の実験室で作成したDNAライブラリをもとに、次世代シーケンサーにより得られた大量のリード（断片化されたDNA配列データ）を対象にSNP検出のための解析を行った。解析の結果、10,389座のSNPが同定され、そのうち欠測値が一つでもある座を除く5,510座のSNPデータを用いて、各個体間の遺伝距離を算出し、系統樹を作成した。その際、使用実績の多いSSRの親子解析結果と比較することで、ゲノミック予測に使用するRad-seq法によって得られたSNPデータの信頼性が確認された。これらの解析は、2023年に発表されたスギの全ゲノム参照配列を活用しており、その結果、得られた各SNP座は、遺伝地図上にマッピングすることが可能となった。

また、Rad-seq法のDNAライブラリ作成に必要な試薬と資材をそろえたことで、当試験場内でDNAの調整作業が行える基盤を整えることができた。

最終年度である2025年度は、DNAライブラリー調整技術を確実なものとし、得られたSNPデータをもとにGPを試行する計画である。



図ー1 5,510 SNP座の情報から作成したサンプル間の遺伝距離
各ブロックに示した交配組み合わせは、事前にSSR解析により確認した家系情報を表記している。