

## 論文 (Original article)

# スギの遺伝的地域性識別のための SNP パネルの開発と利用

内山 憲太郎<sup>1)\*</sup>、松本 麻子<sup>1)</sup>

### 要旨

スギは日本の最も重要な林業樹種であり、その植栽面積は日本の人工林の44%に及ぶ。これまでの遺伝解析から、スギには4つの系統地理学的なグループが存在することが明らかになっている。一般的に、造林時にはそれぞれの地域環境に適応している地域性種苗の利用が遺伝子汚染の観点からも推奨される。しかしながら、スギは集団間の遺伝的分化度が極めて低く、地域の遺伝的グループを明らかにするために、ゲノム内からランダムに選んだ場合には数千のDNAマーカーが必要であった。そこで本研究では、このスギの4つの遺伝的グループを識別可能な最小セットの一塩基多型(SNP)パネルを作成することを目的とする。スギのゲノム上での位置が明らかな2500座のSNPのうち、ゲノム全体から均等に288SNPを選抜し、SNPパネルを構築した。作成したパネルを用いて天然林234個体18集団を解析した結果、そのうちの257座において遺伝子型が決定できた。また、スギの4つの遺伝的グループを明瞭に識別でき、識別能が数千のSNPマーカーを用いた場合と遜色ないことが示された。開発したSNPパネルにより屋久島に植栽された由来不明の人工林のサンプルを解析した結果、屋久島と同じ遺伝的グループに属することを明らかにできた。これらの結果から、今回開発したSNPパネルは、スギの地理的変異の識別や、由来不明な個体の地理的由来を確認することに有用であると考えられた。

キーワード：ゲノムワイドSNPパネル、遺伝子汚染、遺伝構造、遺伝的保全

### 1. はじめに

固着性である植物は、長い年月を通して、それぞれの生育環境へと遺伝的に適応していることが多い(Hufford and Mazer 2003, Leimu and Fischer 2008)。樹木においても、相互移植試験を通して、生育環境への適応が生じている例が多数報告されている(Alberto et al. 2013)。スギでは全国規模での産地試験の例はないが、古くから地域変異の存在が指摘されてきた。これまでに、針葉の形態(Murai 1947)、ジテルペン成分(Yasue et al. 1987)、DNA変異(Takahashi et al. 2005, Tsumura et al. 2007, 2012, 2014, Kimura et al. 2014)などについての研究が行われ、スギは日本海側と太平洋側で形態的にも遺伝的にも分化していることが報告されている。また、最近のスギ天然林の詳細なDNA解析では、太平洋側、日本海側の他に、南北の端の集団も異なる遺伝的クラスターに分類されることが示され、スギには4つの遺伝的なグループがあることが明らかとなっている(Kimura et al. 2014, Tsumura et al. 2014)。このような遺伝的グループは、過去の分布変遷や隔離後の遺伝的浮動によって生じると考えられるが、地域環境からの自然選択の影響も受けている。特にスギで検出された4つの遺伝的グループ間では環境が大きく異なっており、スギの遺伝的分化に地域環境への適応が寄与している可能性が指摘でき

る。スギでは林業種苗法により種苗配布区が定められており、植生や気候などの環境条件に応じて、種苗の移動が制限されている。これは、環境の大きく異なった地域に種苗が移動されることによる成長不良や不成績造林地の成立を防ぐのに役立ってきたと考えられる。しかしながら、種苗の移動による遺伝子汚染を防ぎ、地域に適した種苗を利用するためには、地域に応じた遺伝的変異や由来不明な植栽木の地域性を評価する必要がある。しかしながら、他の針葉樹と同様に、風媒で風散布種子をつけるスギは、集団間の遺伝子流動が活発であり、集団間の遺伝的分化の程度が低いため、その評価は難しい。過去の研究例では、スギの4つの遺伝的なグループを検出するために、千を超える多数のDNAマーカーを解析する必要があり、多大な時間とコストが必要である。このため、短期間かつ安価に遺伝的地域性を評価解析可能なプラットフォームの作成が必要である。

近年のシーケンス技術の向上と、データの蓄積により、1塩基多型(Single nucleotide polymorphism: SNP)の情報が多数利用できるようになった。SNPは、1座あたりの情報量は少ないが、多数の座を一度に解析可能な様々なプラットフォームが利用可能である(Chen and Sullivan 2003)。例えば、蛍光色素で標識されたSNP特異的なプローブとPCRを組み合わせてSNPを蛍光検出

原稿受付：平成29年8月14日 原稿受理：平成30年3月6日

1) 森林総合研究所 樹木分子遺伝研究領域

\* 森林総合研究所 樹木分子遺伝研究領域 〒305-8687 茨城県つくば市松の里1

する方法 (McGuigan and Ralston 2002) や、ハイブリダイゼーションを利用する方法 (Akhunov et al. 2009)、塩基間の質量の違いを検出する方法 (Gabriel et al. 2009) などが代表的なものである。

スギは日本の樹木の中で最もゲノム情報が充実している樹種であり、これまでに数多くの DNA マーカーが開発されている (Iwata et al. 2001, Moriguchi et al. 2003, Tani et al. 2004, Uchiyama et al. 2012, Ueno et al. 2012)。また、発現遺伝子配列情報と連鎖解析による高密度の基盤連鎖地図も構築されており、現在までに 2,500 を超える遺伝子について、ゲノム上での位置が明らかになっている (Moriguchi et al. 2012, 2016)。本研究ではこれらの情報を活用し、スギの遺伝的な地域性を簡便に明らかにできる SNP パネルの開発を目的とした。また、開発した SNP パネルを用いて、屋久島の由来不明な人工林の遺伝的な地域性を明らかにし、その有用性を評価した。スギの分布の南端である屋久島はその生態系の特異性から 1993 年には世界自然遺産に登録されており、遺伝的な変異の保全が重要な課題のひとつである。屋久島のスギは、過去の研究例から他の地域と遺伝的に異なることが明らかになっており (Takahashi et al. 2005, Kimura et al. 2014)、他地域からの種苗の持ち込みが起こった場合、遺伝的攪乱をもたらす可能性が高いと考えられる。

## 2. 材料と方法

### 2.1 SNP パネルの開発

スギの基盤連鎖地図に座乗した SNP マーカーの情報 (Moriguchi et al. 2012, 2016) を利用し、スギのゲノム全体の挙動を把握するために、11 本の染色体全体から 300 座の SNP マーカーを選抜した。具体的には、基盤連鎖地図の各連鎖群 (染色体) の地図距離に応じて、300 座のマーカーを均等に割り振り、連鎖群内でのマーカー間の距離を最低 0.5cM 以上離すようにした。また、中立な遺伝マーカーを用いた遺伝構造の推定時には自然選択の影響を受けている遺伝子座が問題となるため、過去のスギ天然林の解析 (Tsumura et al. 2012, 2014) において自然選択の痕跡が認められた遺伝子座は除外した。その結果、各連鎖群内でのマーカー間の平均地図距離は 5.48cM となった。スギのゲノムサイズ 10.8Gb (Hizume et al. 2001) と、基盤連鎖地図の総地図距離 1266.2cM (Moriguchi et al. 2016) から計算すると、最低地図距離として設定した 0.5cM は約 4.3Mbp (430 万塩基対) に相当する。スギの連鎖不平衡は遺伝子間領域では 100Kbp を超える例が報告されているが (Moritsuka et al. 2012)、今回のマーカー間距離は、少なくともそれより 40 倍以上長く、マーカー間の挙動はほぼ独立していると考えられる。SNP のタイピングには Fluidigm 社の EP1 genotyping system を用いた。これはマイクロ流路を搭載したチップを用いた SNP タイピング手法であり、

ごく微量の DNA および試薬により、一度に 47 サンプルについて 48SNP の遺伝子型を決定することが可能である。上述の 300SNP の候補から、48SNP、6 セット分の 288SNP に関してプライマーを設計し、解析に使用した。

### 2.2 SNP パネルの識別能の確認

森林総合研究所で保有している全国のスギ天然林 18 集団、計 234 個体の DNA について、作成した SNP パネルにより遺伝構造を評価し、遺伝的な地域性の識別能を確認した。これらの集団はスギの天然分布の北端となる青森県鱒ヶ沢から南端にあたる鹿児島県屋久島集団までを含み、日本におけるスギ天然林の遺伝的な特徴を評価するための代表的な集団のセットとみなすことができる (Fig. 1)。まず、使用した SNP の多型頻度を確認するため、天然林の遺伝子型から各 SNP のマイナーアレル頻度 (minor allele frequency: MAF) を算出した。また、集団間の遺伝的分化の程度について、 $F_{ST}$  (Weir and Cockerham 1984) を FSTAT v2.9.4 (Goudet 2005) を用いて算出した。続いて天然集団の遺伝的な地域性を評価するために、個体単位での遺伝的クラスタの推定手法である STRUCTURE 解析 (Pritchard et al. 2000) をソフトウェア STRUCTURE v2.3.4 を用いて行った。STRUCTURE 解析は集団の遺伝構造の評価、個体の集



Fig. 1. 解析に用いたスギ天然林 18 集団の位置図

団への割り振り、雑種形成地域の検出、移住の識別、個体内での複数集団の混合の検出などに応用可能とされる (Pritchard et al. 2000)。STRUCTURE 解析では共通祖先を持つ遺伝的クラスターを仮定して、各クラスター内でのハーディーワインベルグ不平衡ならびに連鎖不平衡が最小になるようにクラスタリングを行い、同時に各個体がそれぞれのクラスターに割り振られる確率が計算される。クラスター数 10 までの独立試行を各 10 回行い、1 試行あたり 5 万回を burn-in period (初期値に依存する期間) として除き、10 万回のマルコフ連鎖モンテカルロ (MCMC) を行った。解析条件は、集団の採取地情報を含めた Hubisz et al. (2009) に従った。得られた結果から  $\Delta K$  (Evanno et al. 2005) を算出し、各試行の事後確率 (LnP (D)) とともに最適クラスター数を評価した。

得られた天然林集団の遺伝構造について、解析遺伝子座数の少ないものから 148 座の CAPS マーカー (Tsumura et al. 2007) で解析された例、さらに、1026 座 (Tsumura et al. 2012) および、3930 座の SNP マーカー (Tsumura et al. 2014) で解析されたと比較した。

### 2.3 由来不明のスギ個体の地域性の確認

今回開発した SNP パネルを用い、由来不明のスギのサンプルに対して遺伝構造の解析を行い、天然林との比較からスギの 4 つの遺伝的グループのどこに属するかの確認を試みた。材料には、鹿児島県熊毛郡屋久島町の小杉谷に 1980 年代に植栽された由来不明のスギ人工林のサンプルを用いた。調査地は安房川の上流、安房森林軌道沿いにあり、代表的な屋久杉として有名な縄文杉まで 1.5km の距離にあたる。Fig. 2 に調査地位置図を示す。人工林 10 箇所からそれぞれ 5-14 個体、合計 94 個体を採取し解析対象とした。

採取したスギの本葉は冷蔵輸送し、森林総合研究所にて DNA の抽出を行なった。改変 CTAB 法を用いて

DNA 抽出を行い (Tsumura et al. 1995)、本研究で開発した SNP パネルを用いて遺伝子型データを決定した。地域性確認のため、全国のスギ天然林のデータと合わせて STRUCTURE 解析を行った。解析条件は天然林と同様とした。また、集団間の遺伝的な関係を明らかにするために、集団間の総当たりでの遺伝距離 (Nei's standard genetic distance:  $D_s$  (Nei 1972)) を算出し、プログラム Splitstree (Huson and Bryant 2006) を用いてネットワーク図を作成した。

## 3. 結果および考察

### 3.1 SNP パネルの開発

連鎖地図に基づいて選んだ 300 座の SNP のうち、290 座において Fluidigm 社の SNP タイピングシステム用のプライマーを設計することができた。残りの 10 座では、対象とする SNP の近傍に別な SNP が存在している座および SNP 近傍の塩基の GC 含量が高いためにプライマーの設計が困難であった。さらに 48 座を単位とする Fluidigm 社のプラットフォームに合わせるため、290 座からマーカー間の距離なるべく均等になるように 288 座を選び、SNP タイピングを行った。288 座のマーカーは全て DDBJ (アクセッション No. AB874696-AB889383) および TreeGenes database (<http://dendrome.ucdavis.edu/cmap/>, accession: 1438-1448) に登録されている。天然林 234 個体を用いた解析では、288 座中 257 座において遺伝子型を決定できた。連鎖群毎の SNP 座数を Table 1 に示す。タイピングができなかった 31 座では、ホモ接合体とヘテロ接合体の判別が不明瞭であったり、少数の個体の特定のアレルで蛍光の検出ができなかった座などが存在した。これらの理由として、プライマー部位に変異を持った個体が存在していた可能性、また想定したアレル以外の第 3 のアレルを持った個体が存在した可能性が考えられた。

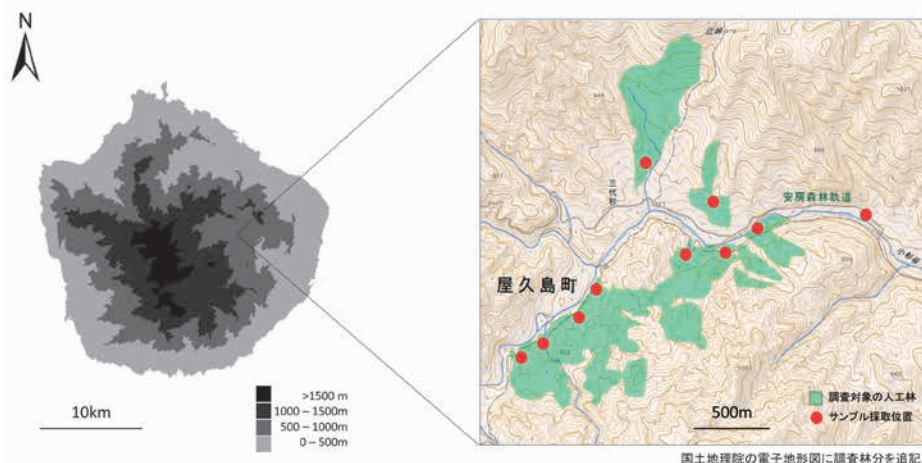


Fig. 2. 屋久島小杉谷の調査対象の人工林およびサンプル採取位置



**Table 1.** 各連鎖群に座乗している SNP マーカーの数

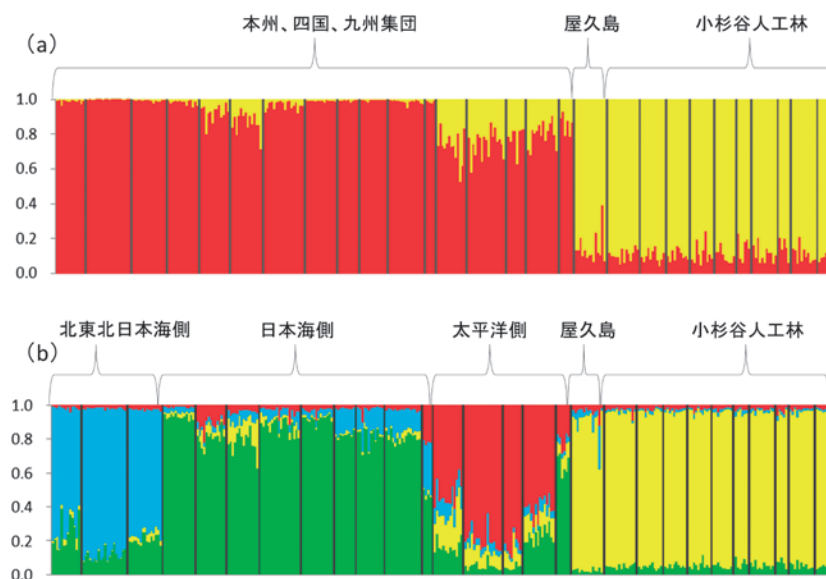
連鎖群	座数
LG01	23
LG02	20
LG03	35
LG04	25
LG05	27
LG06	16
LG07	23
LG08	22
LG09	15
LG10	27
LG11	24

### 3.2 SNP パネルの識別能の確認

明瞭なタイピングが行えた 257 座は全ての座で多型が検出され、MAF の平均は 0.28 であった。一般の遺伝解析で有用とされるコモンアレルは MAF が 0.05 もしくは 0.10 以上とされるが (Altshuler et al. 2005)、今回作成した 257 座のうち 231 座 (90%) で  $MAF \geq 0.1$  となり、多型性が高く情報量の多い SNP パネルが作成できたと考えられた。集団全体での  $F_{ST}$  は 0.027 となり、他の SNP での研究例における 0.039 (Tsumura et al. 2012)、0.050 (Tsumura et al. 2014)、0.062 (Tsumura et al. 2007) よりも低い値となった。他の解析においては自然選択を受けたと予想される高い  $F_{ST}$  (>0.2) を示すような遺伝子座も一定数含まれているため、全体の  $F_{ST}$  が高くなっ

ていると考えられる。事実、中立なマイクロサテライトマーカー 11 座での解析では本結果と同程度に低い  $F_{ST}$  が報告されている ( $F_{ST}=0.028$ : Takahashi et al. 2005)。

257 座を用いた天然林の Structure 解析の結果、最適な遺伝的クラスター数は  $\Delta K$  の基準では 2、 $\text{LnP}(D)$  では 4 となった。各クラスター数での結果を Fig. 3 に示す。なお、後述の小杉谷のサンプルを加えた場合でも同様の結果であったため、小杉谷サンプルも含んだ結果を示す。クラスター数を 2 とした場合、屋久島とその他の遺伝的クラスターに分かれた。一方、クラスター数を 4 とした場合は、それぞれ北東北日本海側の 3 集団、日本海側の 9 集団、太平洋側の 5 集団、屋久島の 1 集団からなる 4 つの遺伝的クラスターに分かれた。この結果は 3930SNP のゲノムワイド SNP での結果 (Tsumura et al. 2014) とほぼ一致する。異なった点は、クラスター数 2 の場合に、今回の SNP パネルでは屋久島とその他に分かれたのに対し、3930SNP の例では太平洋側、日本海側がまず分かれたことである。3930SNP の例では、環境の大きく異なる太平洋側と日本海側で高い分化度を示す SNP が多く含まれていたことも影響していると考えられる。一方、クラスター数 3 および 4 の結果は完全に一致した。スギが持つ 4 つの遺伝的クラスターを評価できることが、地域性の評価では特に重要である。1026 座、3096 座の SNP マーカーでの集団構造解析と同様に本研究でも同様にこれら 4 つのクラスターを識別することができたことから、今回開発した SNP パネルの識別能は 3930SNP を用いた結果と遜色ないものと考えられた。140 座の解析では識別が不明瞭であったのに対し、今回 257SNP マーカーでスギの遺伝的クラスターを評価できた



**Fig. 3.** ベイジアンクラスタリング解析により検出された遺伝的クラスター  
a: クラスター数が 2 の場合、b: クラスター数が 4 の場合

理由として、ゲノム全体に均等にマーカーを配置したことで、ゲノムの様々な領域に痕跡として残っている集団の履歴を、より高い精度で検出できたためと考えられた。

### 3.3 由来不明のスギ個体の地域性の確認

開発したパネルを用いて小杉谷の94個体を解析したところ、天然林と同様に、257遺伝子座の遺伝子型を明瞭に決定できた。天然林データと合わせてSTRUCTURE解析を行った結果、最適クラスター数は天然林のみでの解析と同様に2と4となった。いずれのクラスター数の場合でも、小杉谷の94個体は平均クラスター頻度0.9以上で全て屋久島クラスターに分類された (Fig. 3)。このことより、今回の小杉谷の人工林は屋久島内の個体から採取された種子もしくは挿し穂に由来するものと考えられた。九州では挿し木林業が盛んに行われているが、今回の小杉谷サンプルに関してはクローン個体が検出されなかったことから、種子由来の苗の可能性が高いと考えられた。

集団間の遺伝距離からネットワーク図を作成した結果を Fig. 4 に示す。STRUCTURE 解析と同様に小杉谷集団は明瞭に屋久島集団と同じグループに分類され、今回の小杉谷由来不明なサンプルは屋久島島内に由来すると考えられた。

屋久島のスギが他の地域のスギと遺伝的に離れていることに関しては、最終氷期時の逃避地であった可能性や、海により隔離されて年月を経た結果、遺伝的浮動により変異が蓄積した可能性などが指摘されている (Takahashi et al. 2005)。一方で、分布域の南端にあたる屋久島の多雨で湿潤な環境による自然選択で特徴的な遺

伝的変異が蓄積した可能性も考えられ (Tsumura 2011, Kimura et al. 2014)、屋久島のスギの遺伝的変異は、スギの重要な遺伝資源と考えられる。今回の調査林分は屋久島のスギの標高分布の中心付近である標高 800 ~ 900m 前後にあり、国の特別天然記念物に指定されている屋久島スギ原始林にも隣接している。周囲の天然林の個体とも活発に遺伝子流動が起こっていると予想され、林分の大きさと個体数の多さから、もし屋久島以外の地域のスギが持ち込まれていた場合には、遺伝子汚染の大規模な汚染源となっていたと考えられる。今回の解析において、調査対象とした小杉谷の人工林のサンプルは全て屋久島の天然林のサンプルと変わらない遺伝的組成を示したことから、屋久島内で採取された種子によって造林された林分であると確認され、本林分による遺伝子汚染の可能性は否定できると考えられた。

## 4. まとめ

今回開発した SNP パネルは、スギの自然集団が持つ4つの遺伝的グループを明瞭に識別でき、このパネルを用いることで個体の地理的な由来を短期間に安価に明らかにできるようになった。由来不明であった屋久島の小杉谷の人工林の解析の結果、全ての個体が屋久島に特異的な遺伝的クラスターに属し、当時の造林が屋久島由来の種苗によって行われていたことが示されたと考えられる。今回開発した SNP パネルは、ゲノム全体をカバーしていることから、スギの地理的な変異の解析のみならず QTL 解析や形質予測などの様々な遺伝解析への応用も期待できると考えられる。

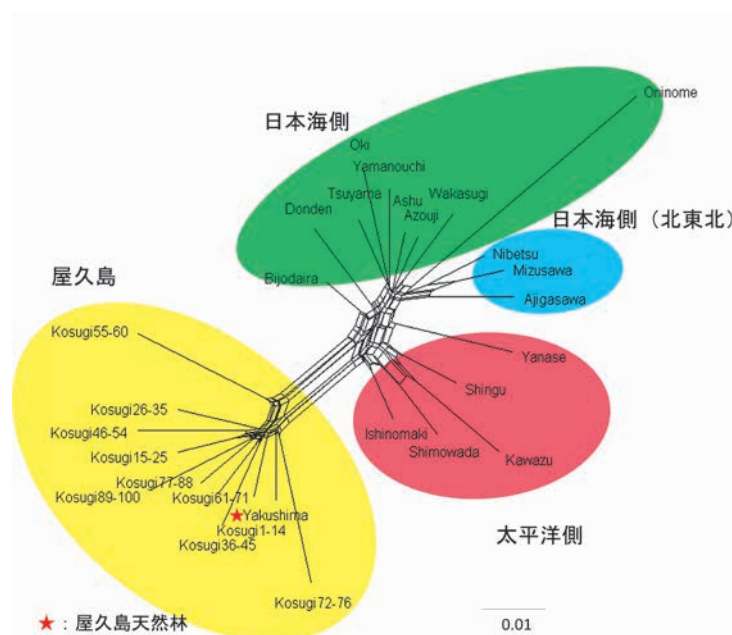


Fig. 4. 集団間の遺伝距離 (Nei's  $D_s$  (Nei 1972)) を用いたネットワーク図色のついた円は STRUCTURE 解析の4つのクラスターに対応

## 謝 辞

本研究の遂行にあたり、屋久島の小杉谷周辺の情報の収集には森林総合研究所九州支所の金谷整一氏にご尽力いただいた。新潟大学の森口喜成准教授にはスギの連鎖地図情報を提供いただいた。また、スギ天然林集団の長年にわたる収集および管理には、筑波大学の津村義彦教授、森林総合研究所実験林の皆様、同樹木遺伝研究室の皆様にご尽力いただいた。ここに厚くお礼申し上げる。本課題は林野庁九州森林管理局「平成27年度屋久島世界遺産地域等における森林生態系に関するモニタリング調査等に関わる業務」の一環で行われた。

## 引用文献

- Akhunov, E., Nicolet, C. and Dvorak, J. (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor. Appl. Genet.*, 119, 507-517.
- Alberto, F. J., Aitken, S. N., Alía, R., González - Martínez, S. C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S. and Whetten, R. (2013) Potential for evolutionary responses to climate change—evidence from tree populations. *Global Change Biol.*, 19, 1645-1661.
- Altshuler, D., Brooks, L., Chakravarti, A., Collins, F., Daly, M. and Donnelly, P. (2005) A haplotype map of the human genome. *Nature*, 437, 1299-1320.
- Chen, X. and Sullivan, P. F. (2003) Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *Pharmacogenomics J.*, 3, 77-96.
- Evanno, G., Regnaut, S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.*, 14, 2611-2620.
- Gabriel, S., Ziaugra, L. and Tabbaa, D. (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr. Protoc. Hum. Genet.*, 60, 2.12.1-2.12.18.
- Goudet, J. (2005) FSTAT (version 2.9. 4): a program to estimate and test population genetics parameters, Updated from Goudet (1995).
- Hizume, M., Kondo, T., Shibata, F. and Ishizuka, R. (2001) Flow cytometric determination of genome size in die Taxodiaceae, Cupressaceae *sensu stricto* and Sciadopityaceae. *Cytologia*, 66, 307-312.
- Hubisz, M. J., Falush, D., Stephens, M. and Pritchard, J. K. (2009) Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.*, 9, 1322-1332.
- Hufford, K. M. and Mazer, S. J. (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends Ecol. Evol.*, 18, 147-155.
- Huson, D. H. and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.*, 23, 254-267.
- Iwata, H., Ujino-Ihara, T., Yoshimura, K., Nagasaka, K., Mukai, Y. and Tsumura, Y. (2001) Cleaved amplified polymorphic sequence markers in sugi, *Cryptomeria japonica* D. Don, and their locations on a linkage map. *Theor. Appl. Genet.*, 103, 881-895.
- Kimura, M. K., Uchiyama, K., Nakao, K., Moriguchi, Y., San Jose-Maldia, L. and Tsumura, Y. (2014) Evidence for cryptic northern refugia in the last glacial period in *Cryptomeria japonica*. *Ann. Bot.*, 114, 1687-1700.
- Leimu, R. and Fischer, M. (2008) A meta-analysis of local adaptation in plants. *PloS one*, 3, e4010.
- McGuigan, F. E. and Ralston, S. H. (2002) Single nucleotide polymorphism detection: allelic discrimination using TaqMan. *Psychiatr. Genet.*, 12, 133-136.
- Moriguchi, Y., Iwata, H., Ujino-Ihara, T., Yoshimura, K., Taira, H. and Tsumura, Y. (2003) Development and characterization of microsatellite markers for *Cryptomeria japonica* D. Don. *Theor. Appl. Genet.*, 106, 751-758.
- Moriguchi, Y., Uchiyama, K., Ueno, S., Ujino-Ihara, T., Matsumoto, A., Iwai, J., Miyajima, D., Saito, M., Sato, M. and Tsumura, Y. (2016) A high-density linkage map with 2560 markers and its application for the localization of the male-sterile genes *ms3* and *ms4* in *Cryptomeria japonica* D. Don. *Tree Genet. Genom.*, 12, 1-12.
- Moriguchi, Y., Ujino-Ihara, T., Uchiyama, K., Futamura, N., Saito, M., Ueno, S., Matsumoto, A., Tani, N., Taira, H., Shinohara, K. and Tsumura, Y. (2012) The construction of a high-density linkage map for identifying SNP markers that are tightly linked to a nuclear-recessive major gene for male sterility in *Cryptomeria japonica* D. Don. *BMC Genomics*, 13, 95.
- Moritsuka, E., Hisataka, Y., Tamura, M., Uchiyama, K., Watanabe, A., Tsumura, Y. and Tachida, H. (2012) Extended linkage disequilibrium in noncoding regions in a conifer, *Cryptomeria japonica*. *Genetics*, 190, 1145-1148.
- Murai, S. (1947) Major forestry tree species in the Tohoku region and their varietal problems, In Aomori-rinyukai (eds.) "*Kokudo Saiken Zourin Gijutsu Kouenshu*", Aomori, Japan. 131-151.
- Nei, M. (1972) Genetic distance between populations. *The American Naturalist*, 106, 283-292.
- Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

- Takahashi, T., Tani, N., Taira, H. and Tsumura, Y. (2005) Microsatellite markers reveal high allelic variation in natural populations of *Cryptomeria japonica* near refugial areas of the last glacial period. *J. Plant Res.*, 118, 83-90.
- Tani, N., Takahashi, T., Ujino-Ihara, T., Iwata, H., Yoshimura, K. and Tsumura, Y. (2004) Development and characteristics of microsatellite markers for sugi (*Cryptomeria japonica* D. Don) derived from microsatellite-enriched libraries. *Ann. For. Sci.*, 61, 569-575.
- Tsumura, Y. (2011) *Cryptomeria*, In Kole C. (eds.) "Wild Crop Relatives: Genomic and Breeding Resources." Springer, Berlin. 49-63.
- Tsumura, Y., Kado, T., Takahashi, T., Tani, N., Ujino-Ihara, T. and Iwata, H. (2007) Genome scan to detect genetic structure and adaptive genes of natural populations of *Cryptomeria japonica*. *Genetics*, 176, 2393-2403.
- Tsumura, Y., Uchiyama, K., Moriguchi, Y., Kimura, M. K., Ueno, S. and Ujino-Ihara, T. (2014) Genetic differentiation and evolutionary adaptation in *Cryptomeria japonica*. *G3: Genes| Genomes| Genetics*, g3. 114.013896.
- Tsumura, Y., Uchiyama, K., Moriguchi, Y., Ueno, S. and Ujino-Ihara, T. (2012) Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity*, 109, 349-360.
- Tsumura, Y., Yoshimura, K., Tomaru, N. and Ohba, K., (1995) Molecular phylogeny of conifers using RFLP analysis of PCR-amplified specific chloroplast genes. *Theor. Appl. Genet.*, 91, 1222-1236.
- Uchiyama, K., Ujino-Ihara, T., Ueno, S., Taguchi, Y., Futamura, N., Shinohara, K. and Tsumura, Y. (2012) Single nucleotide polymorphisms in *Cryptomeria japonica*: their discovery and validation for genome mapping and diversity studies. *Tree Genet. Genom.*, 8, 1213-1222.
- Ueno, S., Moriguchi, Y., Uchiyama, K., Ujino-Ihara, T., Futamura, N., Sakurai, T., Shinohara, K. and Tsumura, Y. (2012) A second generation framework for the analysis of microsatellites in expressed sequence tags and the development of EST-SSR markers for a conifer, *Cryptomeria japonica*. *BMC Genomics*, 13, 136.
- Weir, B. S. and Cockerham, C. C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358-1370.
- Yasue, M., Ogiyama, K., Suto, S., Tsukahara, H., Miyahara, F. and Ohba, K. (1987) Geographical differentiation of natural *Cryptomeria* stands analyzed by diterpene hydrocarbon constituents of individual trees. *J. Jpn. For. Soc*, 69, 152-156.

## Development and utilization of a SNP panel for identification of the geographic origin of *Cryptomeria japonica*

Kentaro UCHIYAMA<sup>1)\*</sup> and Asako MATSUMOTO<sup>1)</sup>

### Abstract

Sugi, *C. japonica* D. Don, is the most important commercial forestry species in Japan; it has been widely planted across the whole of the country and constitutes 44% of all plantation forests. Previous large-scale genome wide analysis have detected four phylogeographic groups in this species. These genetic groups have considerable relevance to *C. japonica* silviculture, for example, in ensuring the use of local seed sources, that are considered to be best adapted to local environments, in afforestation and to reduce the risk of any deleterious gene flow from planted forests into remnant natural stands. In *C. japonica*, however, thousands of DNA markers are required to reliably detect these phylogeographic genetic groups because of its low range-wide genetic differentiation. The aim of this study was to develop a small set of SNPs, more evenly spread across the genome, which could more efficiently distinguish between the four phylogeographic groups of *C. japonica*. From the 2500 mapped markers of *C. japonica*, we selected 288SNP, which were roughly evenly spaced across the *C. japonica* genetic map. Conversion of these SNPs to a cost-effective and flexible throughput SNP typing system generated successful assays for 257SNPs. In order to test the utility of the resulting SNP set, genotyping was performed to determine levels of polymorphism within 234 trees representing 18 natural populations of *C. japonica*. The results indicate samples were clustered into the four genetic groups consistent with their geographic origin, and that SNP set developed in this study can successfully identify population structure. Furthermore, we successfully identified the origin of samples of unknown geographic origin from artificial forests in Yakushima islands. We anticipate that the genome-wide SNP set reported here will be useful for evaluating the species range-wide genetic structure and identification of the geographic origin of unknown samples.

**Key words :** genome-wide SNP panel, genetic structure, genetic pollution, genetic conservation

---

Received 14 August 2017, Accepted 6 March 2018

1) Department of Forest Molecular Genetics and Biotechnology, Forestry and Forest Products Research Institute (FFPRI)

\* Department of Forest Molecular Genetics and Biotechnology, FFPRI, 1 Matsunosato, Tsukuba, Ibaraki, 305-8687 JAPAN; e-mail: kruchiyama@affrc.go.jp