

森林総合研究所
交付金プロジェクト研究 成果集 25

ポプラ等樹木の完全長cDNA塩基配列
情報の充実

独立行政法人 森林総合研究所
2010.1

序 文

ヒト、シロイヌナズナ、イネなど様々な生物のゲノム研究プロジェクトが活発に推進されている。これらゲノム情報がもたらす成果は、基礎生物学の分野にとどまらず、産業や地球環境問題といった応用分野においても革命的な影響を及ぼすと予想されている。「総合科学技術会議分野別推進戦略（ライフサイエンス分野）」や「バイオテクノロジー戦略大綱」では、ゲノム研究・ポストゲノム研究・バイオインフォマティクス（生物情報科学）研究の重要性を指摘している。林野庁は、「森林・林業分野における遺伝子組換え技術に関する研究開発の今後の展開方向について」（平成19年8月31日）の中で、地球温暖化軽減に貢献する組換え樹木の開発のためにはゲノム情報の基盤整備等が必要であるとしている。また、BT戦略推進官民会議は「ドリームBTジャパン：バイオテクノロジーによるイノベーション促進に向けた抜本的強化方策」（平成20年12月11日）を取りまとめ、バイオテクノロジー研究で得られた情報のデータベース化・生物遺伝資源の保存により、国民共有の財産としての研究や医療、農業等に活用していくための研究基盤整備の重要性を指摘している。

遺伝子の構造や機能の確定にはゲノムの塩基配列情報だけでは不完全であり、ゲノム上の遺伝子から発現するmRNA（cDNA）との構造比較が必要である。平成18年9月に、樹木では初めてポプラゲノムの塩基配列の概要解読結果が公開されたが、ゲノムの構造解析だけでは十分とは言えない。森林総合研究所は、交付金プロジェクト「ポプラ等樹木の完全長cDNA塩基配列情報の充実」（平成18～20年度）を推進し、ポプラ完全長cDNAを約20,000種収集した。この数値は、ポプラゲノムの概要解読から予測される発現遺伝子総数の約40%に相当し、樹木のポストゲノム研究の進展に十分に貢献できるバイオリソースと言える。樹木では初めてのポプラ完全長cDNAの塩基配列情報を理化学研究所植物科学研究センターと共同して公表するとともに、理化学研究所バイオリソースセンターへ寄託し、希望者へ完全長cDNAを配付している。国際ポプラゲノムコンソシアムからも依頼され、これらの塩基配列情報を提供している。また、スギ雄花の完全長cDNAを約10,000種収集することにも成功しており、これらの中には既知の花粉アレルゲンと類似する遺伝子、実験植物の雄ずいや花粉で特異的に発現する遺伝子、転写因子の遺伝子等重要な機能を持つ遺伝子が多数含まれていた。そこで、森林生物遺伝子データベース（ForestGEN）を構築し、外部から自由に利用できる形で公開した。さらに、ポプラDNAマイクロアレイによる網羅的発現解析を進め、環境ストレス処理に応答して発現が変動する遺伝子を同定した。こうした情報は、樹木の環境ストレス応答の解明にとっての足がかりとなるだけでなく、ストレス耐性を保持した遺伝子組換え樹木の開発に必要な導入遺伝子の探索に役立つはずである。

本研究プロジェクトの推進により、ポプラ完全長cDNA情報の充実やスギ雄花完全長cDNAの収集に成功したが、以上の研究成果がポストゲノム時代を迎えた樹木研究の進展の一助になることを期待している。

平成 22 年 1 月

独立行政法人 森林総合研究所
理事長 鈴木 和夫

研究課題：ポプラ等樹木の完全長cDNA塩基配列情報の充実

目 次

研究の要約	1
第1章 スギ完全長 cDNA 塩基配列情報の大規模収集	5
第2章 ポプラ等樹木の完全長 cDNA 塩基配列情報の充実	18
1 ポプラ完全長 cDNA の構造解析と機能分類	18
2 ポプラ環境ストレス応答性遺伝子の特定	22

研 究 の 要 約

I 研究の年次及び予算区分

平成18～20年度（3か年）
森林総合研究所交付金プロジェクト

II 主任研究者

主査：篠原健司（生物工学研究領域・領域長）
取りまとめ責任者：篠原健司（生物工学研究領域・領域長）

III 研究場所

森林総合研究所生物工学研究領域・森林遺伝研究領域

IV 研究目的

地球温暖化軽減に貢献する遺伝子組換え樹木の開発、スギ花粉症対策や遺伝子組換え樹木の商業化に必須の花成制御技術や不稔化技術等を開発するためには、樹木の遺伝子情報を大規模収集し、バイオリソースを整備することが急務である。一方、遺伝子の構造や機能を確定するためには、ゲノムの塩基配列情報を解読するだけでは不十分であり、ゲノム上の各遺伝子から発現するmRNA（cDNA）との構造比較が必要である。平成18年9月に、樹木では初めてポプラゲノムの塩基配列の概要解読結果が公開されたが、塩基配列の羅列にすぎないゲノムの構造解析だけでは十分とは言えない。完全長cDNAには、完全長のタンパク質を合成できる、遺伝子の発現を制御するプロモーター領域を正確に同定できる、ゲノム解析に正確な情報を提供できる、といった利点がある。本研究では、スギ完全長cDNAライブラリーの作製と多数のcDNAの構造解析を進めるとともに、ポプラ完全長cDNA情報の充実を図り、環境ストレス耐性等を付与したスーパー樹木の開発、花成制御技術や不稔化技術等の開発を加速化させることを目的とする。

V 研究方法

発達段階の異なるスギ雄花のmRNAを鋳型にして、ビオチン化キャップトラッパー法により完全長cDNAを選択し、均一化処理を行った後、完全長cDNAライブラリーを作製した。この完全長cDNAを約2万選抜し、両末端の塩基配列を解析した。ポプラやシロイヌナズナ等の塩基配列データベースやアレルゲンデータベース等の情報と比較し、完全長cDNAの機能を推定した。

ポプラ（セイヨウハコヤナギ、*Populus nigra* var. *italica*）の葉、茎、根や花の各器官、また乾燥、高塩濃度や低温などの環境ストレス処理をした葉から調製したmRNAを鋳型にして、ビオチン化キャップトラッパー法により完全長cDNAを選択し、均一化処理を行った後、完全長cDNAライブラリーを作製した。約4万個のポプラcDNAの両末端の塩基配列を解析し、完全長cDNAの機能を推定した。

無菌栽培したポプラの雌株の本葉に、乾燥、高塩濃度、低温及びアブシジン酸の各種ストレス

処理を行った。処理前後のサンプルから RNA を抽出し、DNA マイクロアレイを用いストレス処理により発現が有意に変動する遺伝子を解析した。

VI 研究結果

様々な発達段階のスギの雄花から調製したmRNAを鋳型にして、発現遺伝子の重複が極めて少ないスギ完全長cDNAライブラリーを作製し、最終的に10,463種類の完全長cDNAを収集した。これらの中には既知の花粉アレルゲンと類似している遺伝子、実験植物の雄ずいや花粉で特異的に発現する遺伝子、転写因子の遺伝子等重要な機能を持つ遺伝子が多数含まれていた。さらに、得られたcDNA情報を既知のスギEST情報と共にカタログ化して、森林生物遺伝子データベース（ForestGEN）から公開した。

ポプラから19,841種類の完全長cDNAを収集した。この数値は、ポプラゲノムの概要解読から予測される発現遺伝子総数の約40%に相当し、樹木のポストゲノム研究の進展に十分に貢献できるバイオリソースである。また、理化学研究所と共同してポプラ完全長cDNAクローンの配布と塩基配列情報の公開を開始した。さらに、ポプラDNAマイクロアレイによる網羅的発現解析を行った結果、ストレス処理に応答して発現が変動する2,214種類の遺伝子を同定した。各処理特異的に発現が上昇したものは、乾燥で56種、高塩濃度で329種、低温で49種、アブシジン酸で110種であった。各処理特異的に発現が下降したものは、乾燥で36種、高塩濃度で348種、低温で24種、アブシジン酸で163種であった。複数のストレス処理で共通して発現変動が確認された遺伝子は相当数あった。これらの知見は、樹木の環境ストレス応答の解明にとっての足がかりとなるだけでなく、環境ストレス耐性樹木の開発のための候補遺伝子探索に役立つ。

VII 成果の利活用

森林生物遺伝子データベース（ForestGEN）を構築し、外部から自由に利用できる形で公開した。本プロジェクトにより得られたスギEST情報と既知のEST情報を統合したクラスタ配列情報を提供している。相同性検索等の機能も充実しており、裸子植物の基盤的バイオリソースとして利用でき、将来の花粉症対策のための新規アレルゲンの探索、花成制御遺伝子や雄性不稔遺伝子の解析に役立つ。また、ポプラ完全長cDNAとそれらの塩基配列情報も、品質・規模ともに世界に誇れるものである。DNAマイクロアレイによる環境ストレス応答性遺伝子の探索は、高環境ストレス耐性組換え樹木の創出に基礎的知見を提供するものである。なお、ポプラ完全長cDNAの大規模収集は樹木で最初の報告であり、スギ完全長cDNAの大規模収集は針葉樹で初めてのものである。

VIII 今後の問題点

スギ完全長cDNA塩基配列情報の収集に関しては、林野庁事業「遺伝子組換えによる花粉発生制御技術等の開発事業」の中で継続し、さらに大規模な収集を進める。本研究では、DNAマイクロアレイによるポプラの環境ストレス応答性遺伝子を探索したが、個々の遺伝子の機能解析は十分とは言えない。今後は、環境ストレス応答性遺伝子の機能解析だけでなく、高環境ストレス耐性組換え樹木の開発に取り組む必要がある。

IX 研究発表

- 1) Futamura, N., Ujino-Ihara, T., Nishiguchi, M., Kanamori, H., Yoshimura, K., Sakaguchi, M., Shinohara, K. (2006) Characterization of expressed sequence tags from the pollen of *Cryptomeria japonica*. 8th International Congress of Plant Molecular Biology, Book of Abstracts 119.
- 2) Futamura, N., Ujino-Ihara, T., Nishiguchi, M., Kanamori, H., Yoshimura, K., Sakaguchi, M., Shinohara K. (2006) Analysis of expressed sequence tags from *Cryptomeria japonica* pollen reveals novel pollen-specific transcripts. Tree Physiology 26: 517-1528.
- 3) Futamura, N., Kusunoki, Y., Mukai, Y., Shinohara, K. (2007) Characterization of genes for a pollen allergen, Cry j 2, of *Cryptomeria japonica*. International Archives of Allergy and Immunology 143: 59-68.
- 4) Fujimura, T., Futamura, N., Midoro-Horiuti, T., Togawa, A., Yasueda, H., Saito, A., Shinohara, K., Masuda, K., Kurata, K., Sakaguchi, M. (2007) Isolation and characterization of native Cry j 3 from Japanese cedar (*Cryptomeria japonica*) pollen. Allergy 62: 547-553.
- 5) Nanjo, T., Sakurai, T., Totoki, Y., Toyoda, A., Nishiguchi, M., Futamura, N., Igasaki, T., Kado, T., Seki, M., Sakaki, Y., Shinozaki, K., Shinohara, K. (2007) Collection and characterization of full-length enriched expressed sequence tags in *Populus nigra*. IUFRO Tree Biotechnology 2007 Abstract SIII.15p.
- 6) Nanjo, T., Sakurai, T., Totoki, Y., Toyoda, A., Nishiguchi, M., Kado, T., Igasaki, T., Futamura, N., Seki, M., Sakaki, Y., Shinozaki, K., Shinohara, K. (2007) Functional annotation of 19,841 *Populus nigra* full-length enriched cDNA clones. BMC Genomics 8: 448.
- 7) Shiokawa, T., Yamada, S., Futamura, N., Osanai, K., Murasugi, D., Shinohara, K., Kawai, S., Morohoshi, N., Katayama, Y., Kajita, S. (2008) Isolation and functional analysis of the CjNdly gene a homolog in *Cryptomeria japonica* of *FLORICAULA/LEAFY* genes. Tree Physiology 28: 21-28.
- 8) Ogura, Y., Komatsu, A., Zikihara, K., Nanjo, T., Tokutomi, S., Wada, M., Kiyosue, T. (2008) Blue light diminishes interaction of Per-ARNT-Sim/light, oxygen, or voltage proteins, putative blue light receptors in *Arabidopsis thaliana*, with their interacting partners. J. Plant Research 121: 97-105
- 9) 篠原健司 (2008) 樹木のポストゲノム研究. 林木の育種 227: 1-6.
- 10) Futamura, N., Totoki, Y., Toyoda, A., Igasaki, T., Nanjo, T., Seki, M., Sakaki, Y., Mari, A., Shinozaki, K., Shinohara, K. (2008) Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. BMC Genomics 9: 383.
- 11) 二村典宏、楠城時彦、篠原健司 (2008) 樹木のゲノム解析の現状と展望－完全長cDNAを用いた遺伝子の機能解析－. 北海道の林木育種 51: 28-31.
- 12) Takata, N., Saito, S., Tanaka-Saito, C., Nanjo, T., Shinohara, K., Uemura, M. (2009) Molecular phylogeny and expression of poplar circadian clock genes, *LHY1* and *LHY2*. New Phytologist 181: 808-819.

X 研究担当者

1. スギ完全長cDNA塩基配列情報の大規模収集：
二村典宏（生物工学・樹木分子生物研）・津村義彦・伊原徳子（遺伝・森林遺伝研）・篠原健司（生物工学領域）
実施年度：H18－H19－H20
2. ポプラ等樹木の完全長cDNA塩基配列情報の充実：
楠城時彦（生物工学・ストレス応答研）・二村典宏（生物工学・樹木分子生物研）・

篠原健司（生物工学）
実施年度：H18－H19－H20

（篠原健司）

第1章 スギ完全長cDNA塩基配列情報の大規模収集

ア 研究目的

スギやマツ等の裸子植物は一般にゲノムサイズが大きく、塩基配列解読のコストが下がりつつある現在でも全ゲノムの解読は困難である。近年、ポプラや果樹でのゲノム概要が解読されたが、裸子植物のゲノム解読は進んでいない。海外では、効率よく遺伝子情報を収集するため、各遺伝子から発現する mRNA をもとに人工的に合成する cDNA を大規模収集するプロジェクトが進められてきた。cDNA から解読する断片的な塩基配列情報のことを EST (Expressed Sequence Tag) という。北米の代表的な経済林用樹種であるテーダマツ (*Pinus taeda*) やカナダトウヒ (*Picea glauca*) では、既に 30 万前後の EST 情報が登録されている。日本の主要な人工林用樹種であるスギにおいても EST 収集を進める必要がある。

裸子植物でこれまで収集された EST 情報は、大部分が遺伝子の全長をカバーしていない cDNA に由来している。我が国は、完全長の cDNA を合成する技術で世界をリードし、ヒト、マウス、シロイヌナズナなどで完全長 cDNA ライブラリーが作製され、解析が進められてきた。完全長 cDNA の配列を明らかにすることにより、遺伝子の正確な機能が予測できるだけでなく、遺伝子の転写を調節するプロモーター領域を推測することも可能になる。また、完全長 cDNA の塩基配列情報をもとにして完全なタンパク質を合成することもできる。(図 1-1)。

本研究では、スギ雄花を材料として、裸子植物で初となる完全長 cDNA ライブラリーを作製し、EST 情報を大規模に収集することを目的とする。スギ完全長 cDNA の情報は、有用遺伝子の探索や DNA マーカーの開発など様々な用途への活用が期待できる。さらに、裸子植物の基盤的バイオリソースとしての有用性を示し、樹木研究の進展に貢献するため、本研究により得られたスギ完全長 cDNA に由来する EST 情報を一般に公開する。

イ 研究方法

1) スギ完全長 cDNA ライブラリーの作製

発達段階の異なるスギ雄花から全 RNA を調製した。スギ全 RNA から mRNA を精製し、これを鋳型としてビオチン化キャップトラッパー法により完全長 cDNA を選択し、均一化処理を行った後、完全長 cDNA ライブラリーを作製した。

2) スギ完全長 cDNA の末端塩基配列の解析と機能分類

スギ雄花完全長 cDNA ライブラリーから約 2 万の cDNA クローンを選抜し、両末端の塩基配列 (EST) を解析した。得られた EST 情報をクラスタリングした後、ポプラやシロイヌナズナ等の塩基配列データベースやアレルゲンデータベース等の情報と比較し、完全長 cDNA の機能を推定した。

3) スギ完全長 cDNA に由来する EST 情報のデータベース化

スギ完全長 cDNA に由来する EST 情報を公的データベースへ登録すると共に、インターフェイスに優れたデータベースを構築して一般に公開した。

ウ 結果

1) スギ完全長 cDNA ライブラリーの作製と品質の検証

発達段階の異なる雄花から抽出した mRNA を用いて、発現遺伝子の重複の少ない均一化した完全長 cDNA ライブラリーを作製した。予備的に 91 個の cDNA についてサイズを検定した結果、平均鎖長は 1.6kbp であり、5kb 以上の長い鎖長の cDNA も含まれていた (図 1-2)。この結果は、完全長 cDNA ライブラリーが平均鎖長の比較的長い優れた品質であることを示すものである。

スギ雄花に由来する完全長 cDNA ライブラリーから、19,968 クローンを選抜し、両端の塩基配列を解析した。その結果、19,437 の cDNA クローンに由来する 36,011 の EST 情報を得た。EST 情報を整理した結果、これらの EST 情報は 10,463 種類の転写産物に相当することが明らかになった (表 1-1)。均一化した cDNA ライブラリーを材料とすることにより、重複を少なくして効率的に多種類の転写産物に由来する EST を取得できた。

完全長 cDNA の検証のため、既に遺伝子上流の DNA 配列が明らかになっているスギ花粉アレルゲン遺伝子 *Cryj 2* について、DNA 配列と完全長 cDNA の配列を比較した (図 1-3)。完全長 *Cryj 2* cDNA はタンパク質全長の情報を有し、通常の手法で合成された cDNA よりも 5' 非翻訳領域が長かった。DNA 配列との比較により、*Cryj 2* 遺伝子の転写開始点を明らかにできた。完全長 cDNA とゲノム DNA との比較解析は、転写開始点の特定やプロモーター領域の解析に有効である。

シロイヌナズナ遺伝子と高い相同性を示した cDNA について 5' 端の塩基配列情報を解析した結果からは、約 8 割がタンパク質全長の情報を有していると推測された。本研究で解析した cDNA クローンの大部分は完全長 cDNA であることが示された。

2) スギ完全長 cDNA の機能分類

10,463 種類の転写産物について、タンパク質データベース (Uniprot)、シロイヌナズナ及びイネのタンパク質データ、マツ (*Pinus*)・トウヒ (*Picea*)・ポプラ (*Populus*) の EST データとの比較を行った。スギ転写産物は、被子植物であるシロイヌナズナやイネ、ポプラよりも、針葉樹のマツやトウヒの遺伝子と高い相同性を示した (図 1-4)。スギ転写産物の約 2 割は既知の配列との相同性を示さなかった。各転写産物について機能付けした結果、7,369 (70.4%) の転写産物について分類・同定することができた (図 1-5)。

シロイヌナズナの雄ずいで特異的に発現する 1,145 遺伝子¹⁾及び花粉で特異的に発現する 1,274 遺伝子²⁾と比較解析した結果、雄ずい特異的遺伝子の約 66%、花粉特異的遺伝子の約 48%がスギ転写産物と高い相同性を示した (表 1-2)。180 のスギ転写産物は、既知の花粉アレルゲンと相同性を示した (表 1-3)。また、形態形成等に重要な働きをすることが知られている転写因子を探索した結果、207 転写産物が転写因子に特徴的な配列を有していた (表 1-4)。

3) スギ cDNA 情報の公開

スギ完全長 cDNA に由来する EST 情報は、公的データベースに登録した³⁾。さらに、本研究により得られた塩基配列情報を基盤的バイオリソースとして活用するため、インターフェイスに優れたデータベースを構築し、森林生物遺伝子データベース (ForestGEN) と名付けて森林総合研究所のウェブサイトより公開した (図 1-6)。

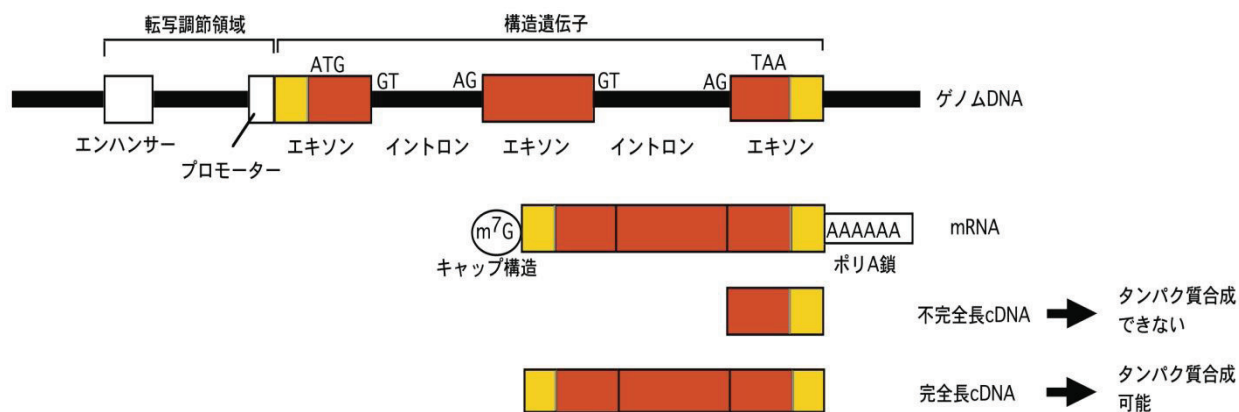


図 1 - 1 完全長 cDNA の模式図

キャップ構造を有する mRNA を選択して作られる完全長 cDNA からは、完全な機能を持つタンパク質の合成が可能となる。また、ゲノム配列情報と比較することにより、構造遺伝子を決定し、プロモーター領域を推定することができる。

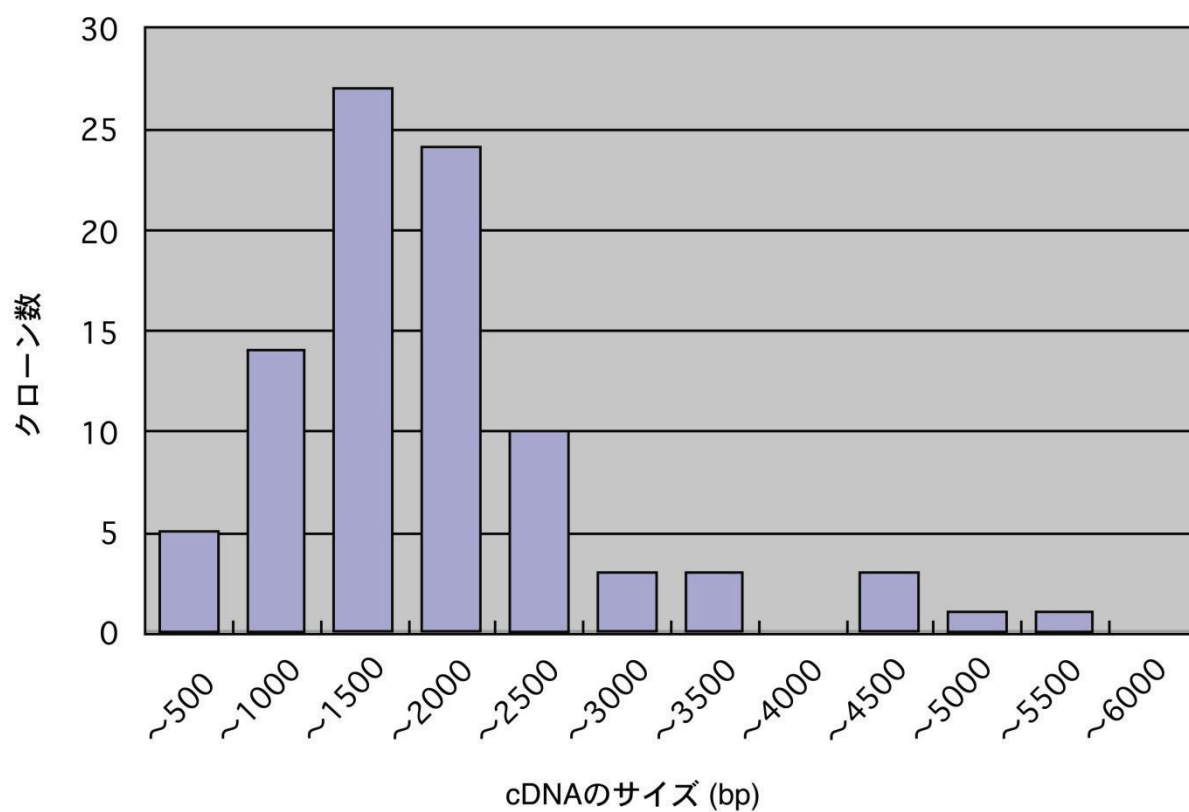


図 1-2 完全長 cDNA のサイズを検証した予備試験

無作為に抽出した 96 個の完全長 cDNA クローンについて、プラスミド DNA を精製し、制限酵素により cDNA インサートを切り出して、1%アガロースゲルにて電気泳動解析を行った。制限酵素による切断が不完全なサンプルや解像度の悪いサンプルを除いた結果、有効解析数は 91 であった。

表 1-1 スギ完全長 cDNA ライブラリーに由来する EST と転写産物の概要

総 EST 数	36,011
5'端配列数	18,843
3'端配列数	17,168
cDNA クローン数	19,437
コンティグ ^a 数	7,686
シングレット ^b 数	15,972
転写産物 ^c 数	10,463
単一の転写産物を構成する cDNA クローン数による分類	
1 クローンで構成される転写産物数	6,320
2 クローンで構成される転写産物数	2,078
3〜5 クローンで構成される転写産物数	1,686
6〜10 クローンで構成される転写産物数	353
11〜20 クローンで構成される転写産物数	21
21 クローン以上で構成される転写産物数	5

^a PHRAP プログラムによる EST 情報の整理により、2 つ以上の EST が重なり合っ
てできた配列。

^b PHRAP プログラムによる解析により、重なり合わないことが確認された EST。

^c コンティグとシングレットを BLASTN プログラムによりグループ化した後、
同一の cDNA クローンに由来する 5'端配列と 3'端配列をまとめた。

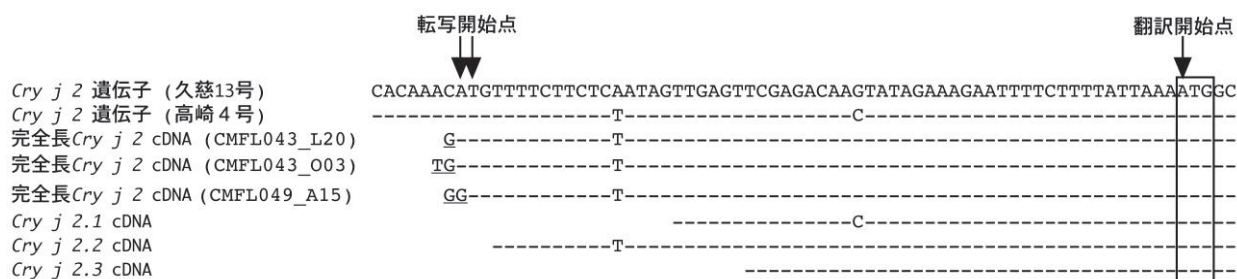


図 1-3 *Cry j 2* 遺伝子の転写開始点の特定

Cry j 2 遺伝子上流 DNA 配列 (久慈 13 号及び高崎 4 号)、完全長 *Cry j 2* cDNA (CMFL043_L20、CMFL043_O03、CMFL049_A15)、通常の方法で作製した *Cry j 2* cDNA (*Cry j 2.1*～*Cry j 2.3*) の塩基配列を比較した。バーは久慈 13 号の *Cry j 2* 遺伝子と一致した塩基配列。下線は完全長 cDNA で付加された塩基配列。ボックスは *Cry j 2* タンパク質 N 端アミノ酸に相当する塩基配列。完全長 cDNA を遺伝子上流の DNA 配列と比較することにより、通常の方法で作製した cDNA との比較では不明であった *Cry j 2* 遺伝子の転写開始点が明らかになった。

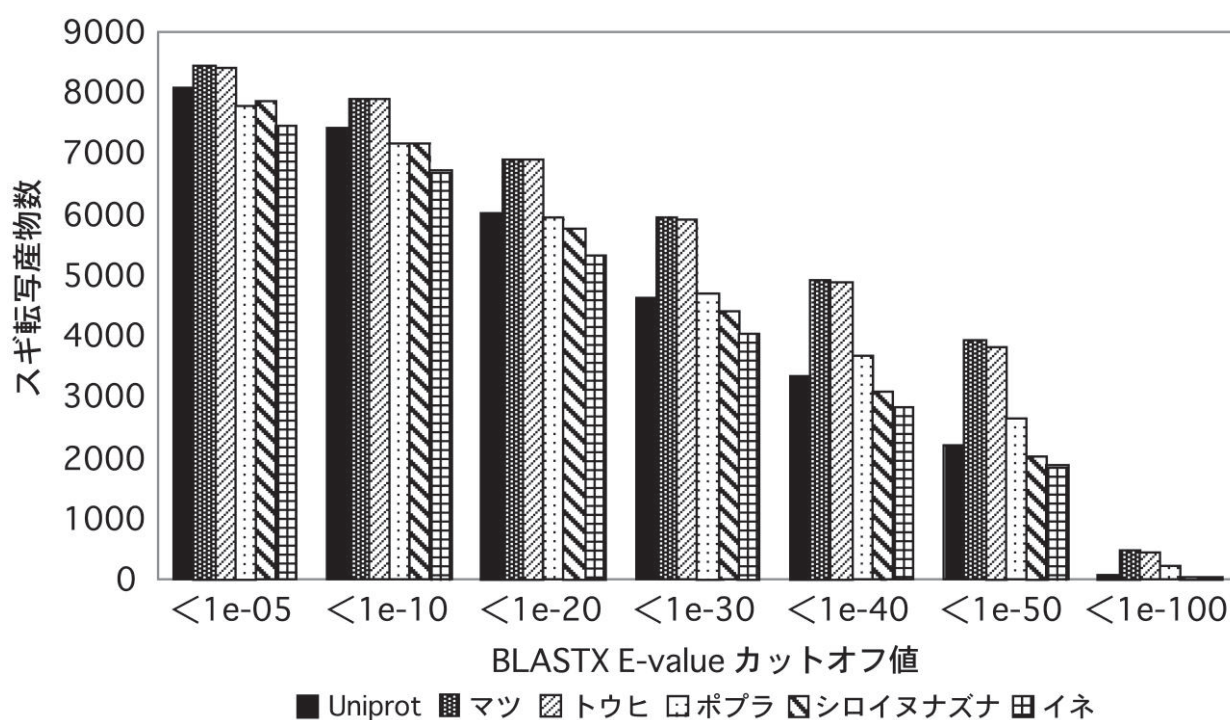


図 1-4 スギ転写産物の配列類似性

BLASTX プログラムを使用して、スギ転写産物から予想されるアミノ酸配列を Uniprot データベース、マツ・トウヒ・ポプラの EST 配列、シロイヌナズナ及びイネのタンパク質と比較した。E-value の値が低いほど相同性が高いことを示す。

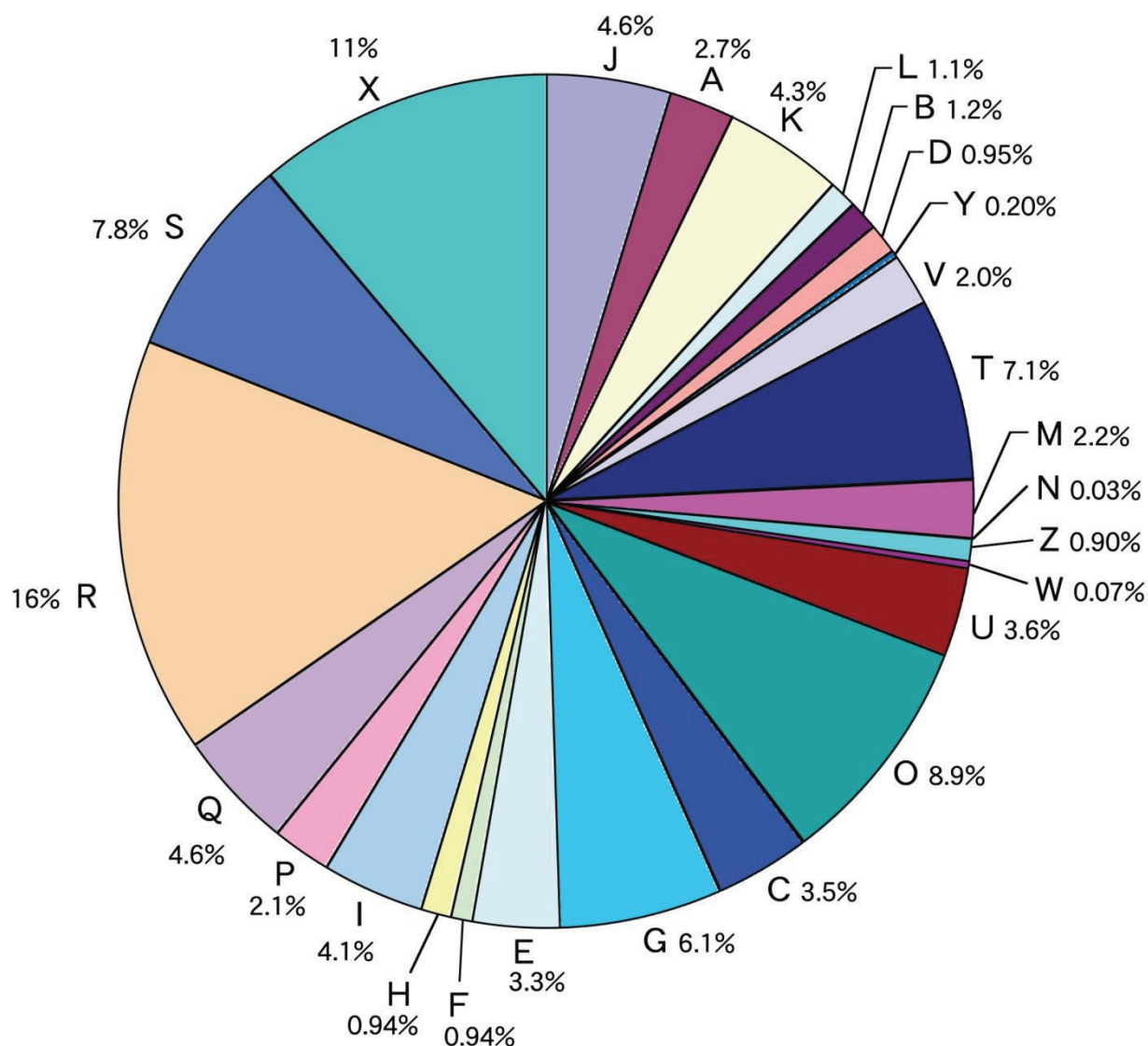


図 1-5 スギ転写産物の機能分類

スギ転写産物をそれぞれの塩基配列がコードするタンパク質の機能によって分類した。KOG、TWOG、LSE データベースとの相同性解析の結果、7,369 転写産物 (70.4%) について機能分類できた。J：タンパク質の翻訳、A：RNA のプロセッシングと修飾、K：転写、L：DNA の複製・組換え・修飾、B：クロマチンの構造変換、D：細胞周期・細胞分裂の制御、Y：核構造、V：防御機構、T：シグナル伝達、M：細胞壁・細胞膜の合成、N：細胞運動、Z：細胞骨格、W：細胞外構造、U：細胞内輸送・分泌、O：タンパク質の翻訳後修飾・代謝、C：エネルギー生産・変換、G：炭水化物の輸送・代謝、E：アミノ酸の輸送・代謝、F：核酸の輸送・代謝、H：補酵素の輸送・代謝、I：脂質の輸送・代謝、P：無機イオンの輸送・代謝、Q：二次代謝成分の合成・輸送、R・S・X：不明。

表 1-2 シロイヌナズナ雄ずい或いは雄性配偶体特異的遺伝子と相同性を示したスギ転写産物

タンパク質の機能 (Pfam ドメイン)	Pfam アク セッション 番号	スギ転 写産物 数	シロイヌ ナズナ雄 ずい特異 的遺伝子 ¹⁾ 数	シロイヌ ナズナ雄 性配偶体 特異的遺 伝子 ²⁾ 数
Protein kinase domain	PF00069	132	35	39
Plant invertase/pectin methylesterase inhibitor	PF04043	12	22	15
Protein tyrosine kinase	PF07714	19	16	12
GDSL-like lipase/acylhydrolase	PF00657	44	15	4
Pectinesterase	PF01095	14	15	11
ABC transporter	PF00005	20	13	5
Cytochrome P450	PF00067	120	13	3
Glycosyl hydrolase family 28	PF00295	10	12	11
Sodium/hydrogen exchanger family	PF00999	6	9	10
ABC-2 type transporter	PF01061	8	9	3
Oleosin	PF01277	4	9	4
No apical meristem (NAM) protein	PF02365	12	9	2
RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	PF00076	78	8	8
Sugar (and other) transporter	PF00083	33	8	3
Calcineurin-like phosphoesterase	PF00149	25	8	6
E1-E2 ATPase	PF00122	4	7	4
Multicopper oxidase	PF00394	4	7	4
Haloacid dehalogenase-like hydrolase	PF00702	19	7	3
Multicopper oxidase	PF07731	14	7	4
Multicopper oxidase	PF07732	17	7	4
Glycosyl hydrolase family 1	PF00232	20	6	4
Pectate lyase	PF00544	5	6	3
FAD-binding domain	PF01565	6	6	2
Galactose-binding lectin domain	PF02140	5	6	4
MtN3/saliva family	PF03083	14	6	1

¹⁾ Wellmer ら(2004)の報告をもとに抽出した 1,145 遺伝子

²⁾ Honys と Twell (2004)の報告をもとに抽出した 1,274 遺伝子

表 1-3 既知の花粉アレルゲンと相同性を示したスギ転写産物

アレルゲン	種名	タンパク質の機能	E-value ^a	転写産物数 ^b
Cry j 1	<i>Cryptomeria japonica</i> (スギ)	Pectate lyase	1E-127	6
Cry j 2	<i>Cryptomeria japonica</i> (スギ)	Polymethylgalacturonase	1E-119	10
Cry j 3.8	<i>Cryptomeria japonica</i> (スギ)	PR-5 protein	3E-98	16
CJP-4	<i>Cryptomeria japonica</i> (スギ)	Class IV chitinase	1E-112	4
CJP-6	<i>Cryptomeria japonica</i> (スギ)	Isoflavone reductase family	4E-87	7
Jun o 4	<i>Juniperus oxycedrus</i> L. (ビャクシン)	Calcium-binding protein	2E-67	16
Amb a 3	<i>Ambrosia artemisiifolia</i> L. (ブタクサ)	Plastocyanin-like protein	2E-8	2
Cat r 1	<i>Catharanthus roseus</i> (L.) G. Don (ニチニチソウ)	Cyclophilin	9E-79	10
Che a 1	<i>Chenopodium album</i> L. (シロザ)	Trypsin inhibitor	3E-28	1
Cor a 1.04	<i>Corylus avellana</i> L. (セイヨウハシバミ)	PR-10 protein	2E-14	1
Cor a 10	<i>Corylus avellana</i> L. (セイヨウハシバミ)	Luminal-binding protein	1E-107	10
Cro s 1	<i>Crocus sativus</i> L. (サフラン)	LAT52 protein	3E-13	3
Cyn d 22	<i>Cynodon dactylon</i> (L.) Pers. (ギョウギシバ)	Enolase	9E-25	1
Cyn d 24	<i>Cynodon dactylon</i> (L.) Pers. (ギョウギシバ)	PR-1 protein	2E-33	3
Hum j Profilin	<i>Humulus japonicus</i> Siebold & Zucc. (カナムグラ)	Profilin	6E-63	3
Hum j 1	<i>Humulus japonicus</i> Siebold & Zucc. (カナムグラ)	Uncharacterized protein	1E-10	3
Lol p 1	<i>Lolium perenne</i> L. (ホソムギ)	Expansin	6E-14	11
Ole e 5	<i>Olea europaea</i> L. (オリーブ)	Superoxide dimutase	7E-66	5
Ole e 9	<i>Olea europaea</i> L. (オリーブ)	β -1,3-glucanase	2E-46	10
Ole e 10	<i>Olea europaea</i> L. (オリーブ)	Glycosyl hydrolase	4E-23	3
Sal k 1.03	<i>Salsola kali</i> L. (ノハラヒジキ)	pectin esterase	2E-35	2
Sal k 2	<i>Salsola kali</i> L. (ノハラヒジキ)	Protein kinase	6E-47	53

^a 最も高い相同性を示したスギ EST の E-value^b 各アレルゲンと相同性を示したスギ転写産物数 (BLASTX スコア > 50 及び E-value < 10⁻⁷)

表 1-4 転写因子と推定されたスギ転写産物

タンパク質の機能 (Pfam ドメイン)	Pfam アクセション番号	スギ転写産物数
Zinc finger, C3HC4 type (RING finger)0	PF00097	48
Myb-like DNA-binding domain	PF00249	43
AP2 domain	PF00847	18
No apical meristem (NAM) protein; NAC domain	PF02365	12
Homeobox domain	PF00046	11
PHD-finger	PF00628	11
SRF-type transcription factor; MADS box	PF00319	10
Histone-like transcription factor (CBF/NF-Y) and archaeal histone	PF00808	7
Helix-loop-helix DNA-binding domain	PF00010	6
HSF-type DNA-binding	PF00447	6
B-box zinc finger	PF00643	6
Dof domain, zinc finger	PF02701	4
WRKY DNA -binding domain	PF03106	3
Response regulator receiver domain	PF00072	2
bZIP transcription factor	PF00170	2
GATA zinc finger	PF00320	2
B3 DNA binding domain; ABI3/VP1 transcription factor	PF02362	2
SBP domain	PF03110	2
GRAS family transcription factor	PF03514	2
ZF-HD protein dimerisation region	PF04770	2
CCT motif; CO-like protein	PF06203	2
Auxin response factor	PF06507	2
ARID/BRIGHT DNA binding domain	PF01388	1
CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B	PF02045	1
TCP family transcription factor	PF03634	1
CG-1 domain; CAMTA protein	PF03859	1
YABBY protein	PF04690	1
Plant protein of unknown function; BZR1/LAT61 family	PF05687	1
Whirly transcription factor	PF08536	1

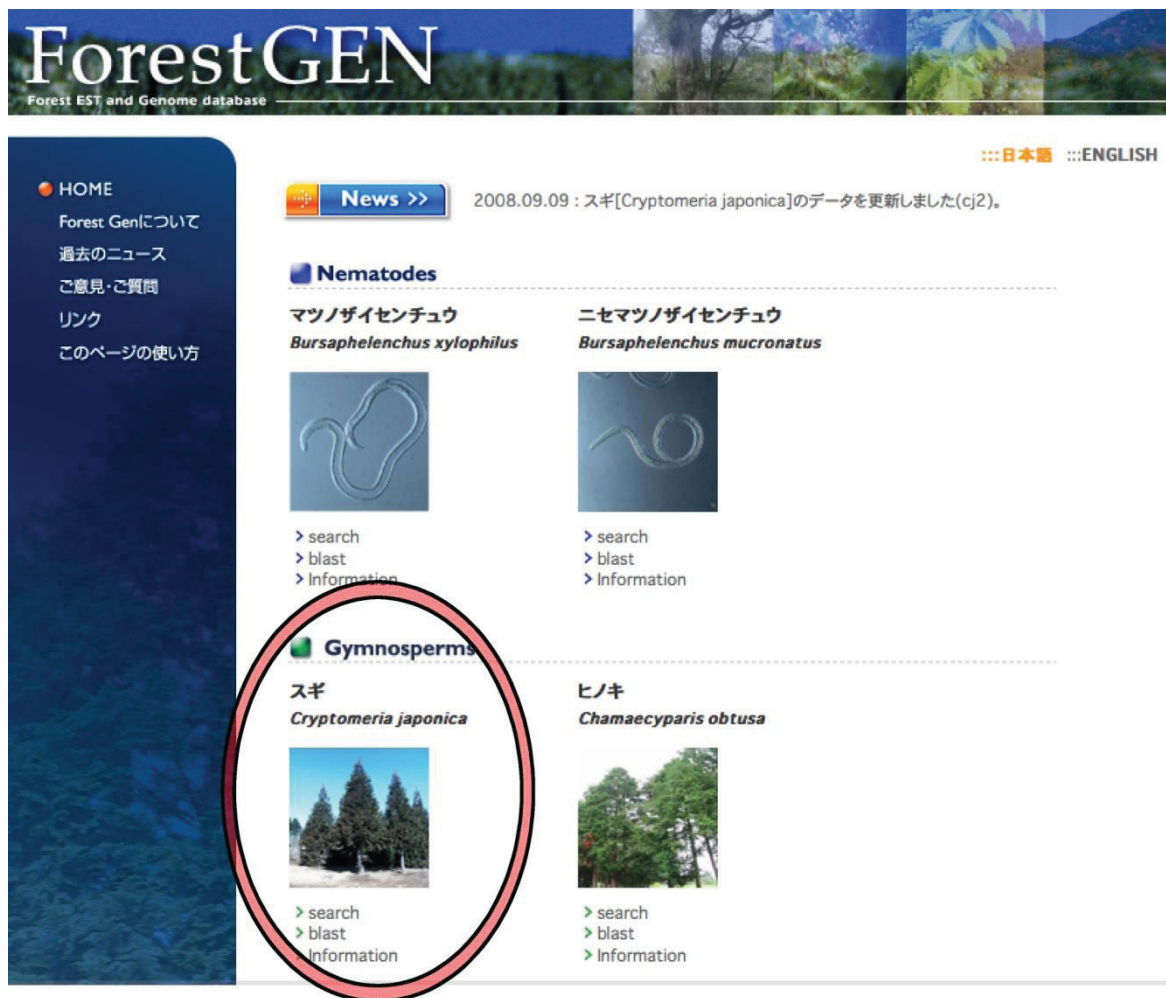


図 1-6 森林生物 EST データベース ForestGEN の画面

円内で示したところから、スギの EST に関する相同性検索やキーワード検索を行うことができる。

エ 考察

スギの有用遺伝子の探索や DNA マーカー開発のため、雄花に由来する RNA から完全長 cDNA ライブラリーを作製し、EST の収集を進めた。作製した完全長 cDNA ライブラリーは、完全長率が高く、重複が少ないという優れた特長を有していた。約 2 万クローンを解析し、最終的に 10,463 種類の転写産物に相当する EST を得た。これらの転写産物は様々な機能を有するタンパク質をコードしており、スギに特異的な遺伝子も含まれることを明らかにした。実験植物の雄ずいや花粉で特異的に発現する遺伝子やアレルゲン遺伝子、転写因子の遺伝子など重要な機能を持つ遺伝子も多数確認できた。さらに、森林生物遺伝子データベース (ForestGEN) を構築し、本プロジェクトにより得られた塩基配列情報を既知のスギ EST と共にカタログ化して、外部から自由に利用できる形で公開した。ForestGEN では、統合した EST 情報を示すことにより、重複なく整理された遺伝子配列情報を取得できる。相同性検索やキーワード検索等の機能も充実しており、裸子植物の基盤的バイオリソースとして活用できる。新規アレルゲンや花成制御遺伝子、雄性不稔遺伝子等の有用遺伝子の探索に役立つと期待される。なお、スギ完全長 cDNA の大規模収集は針葉樹で初めてのものである。

本研究の成果は学術誌や学会に発表された他、ForestGEN に関しては平成 19 年に、スギ完全長 cDNA に関しては平成 20 年にプレスリリースされた。ForestGEN には現在も月平均 1000 件程度のアクセスがあり、有効に活用されている。スギ完全長 cDNA に関する論文は、オープンアクセス雑誌に掲載され、発表後 5 ヶ月間で 1000 件を超えるアクセスがあった。

オ 今後の問題点

有用遺伝子の探索や DNA マーカーの開発のためには、さらに cDNA 情報の充実を図る必要がある。林野庁事業「遺伝子組換えによる花粉発生制御技術等の開発事業」の中で、針葉や雌花を対象に完全長 cDNA の大規模収集を進める予定である。今後、針葉樹のゲノム解析や有用遺伝子の活用を図っていくためには、更なる要員の確保等が必要である。

カ 要約

スギ雄花の完全長 cDNA を 10,464 種類収集し、機能付けを行った。実験植物の雄ずいや花粉で特異的に発現する遺伝子やアレルゲン遺伝子、転写因子の遺伝子など重要な機能を持つ遺伝子が多数含まれていることを明らかにした。収集した完全長 cDNA は、森林総合研究所のウェブサイトに構築した森林遺伝子データベース (ForestGEN) より一般に公開した。

キ 引用文献

- 1) Wellmer, F., Riechmann, J.L., Alves-Ferreira, M., Meyerowitz, E.M. (2004) Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. *Plant Cell* 16: 1314-1326.
- 2) Honys, D. and Twell, D. (2004) Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* 5: R85.
- 3) Futamura, N., Totoki, Y., Toyoda, A., Igasaki, T., Nanjo, T., Seki, M., Sakaki, Y., Mari, A., Shinozaki, K., Shinohara, K. (2008) Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. *BMC Genomics* 9: 383.

(二村典宏)

第2章 ポプラ完全長 cDNA 塩基配列情報の充実

1. ポプラ完全長 cDNA の構造解析と機能分類

ア 研究目的

哺乳類や高等植物を含む様々な生物種を対象にしたゲノムプロジェクトが進められる中、2006年に樹木で初めてポプラゲノムが解読された。しかし、遺伝子の構造や機能を理解するためには、ゲノムの塩基配列情報だけではもとより不十分であり、ゲノム上の塩基配列をもとにして発現する機能遺伝子 (mRNA/cDNA) の構造解析が不可欠である。さらに、細胞内で実際に機能するタンパク質をコードする各遺伝子の役割をより正確に知るためには、完全長の遺伝子を単離して塩基配列を解析する必要がある。生物工学研究領域は、2003年から「ポストゲノムとしてのポプラ完全長 cDNA ライブラリーコレクションの整備」を推進し、世界にさきがけ樹木で初めてポプラの完全長 cDNA を収集し遺伝子情報を公開した。しかし、さきがけ的に行った当プロジェクトで得られた完全長 cDNA の総数は、ポプラゲノムの構造から推定される発現遺伝子数の 10% 程度にすぎなかった。このため本研究課題では、樹木のポストゲノム研究の一環として、さらなるポプラ完全長 cDNA の集積と塩基配列情報の公開を目指した。

イ 研究方法

ポプラ (セイヨウハコヤナギ、*Populus nigra* var. *italica*) の葉、茎、根や花の各器官、また乾燥、高塩濃度や低温などの環境ストレス処理をした葉から調製した mRNA を鋳型にして、ビオチン化キャプトラッパー法により完全長 cDNA ライブラリーを作製した。さらに、発現遺伝子の重複を抑えるため、cDNA ライブラリー作製過程で均一化処理を行った。続いて、このライブラリーに含まれる約 4 万個のポプラ cDNA の 5' 末端および 3' 末端の塩基配列をサンガー法により決定した。これら cDNA の部分塩基配列は、EST (expressed sequence tags) と呼ばれる。本課題では、2005 年までに解析した約 3 万個の EST と合わせて 11 万個以上のポプラ EST を対象として重複を除く処理を行った。

ウ 結果

得られたポプラ EST から重複を除いて整理した結果、様々な機能を持つ 19,841 種類の遺伝子を同定した (図 2-1-1)。これは、ポプラのゲノムの概要解読から推定される遺伝子数の約 40% に相当する。ポプラ完全長 cDNA の中には多数の重要な遺伝子が含まれていた。また、収集した遺伝子は 19 本の染色体上にほぼ均等に分布することも明らかになった (図 2-1-2)。

一方、19,841 種類のポプラ完全長 cDNA を理化学研究所バイオリソースセンターに提供し、公的配布を委託した。また、ポプラ遺伝子群の塩基配列情報を理化学研究所植物科学研究センターに提供した。

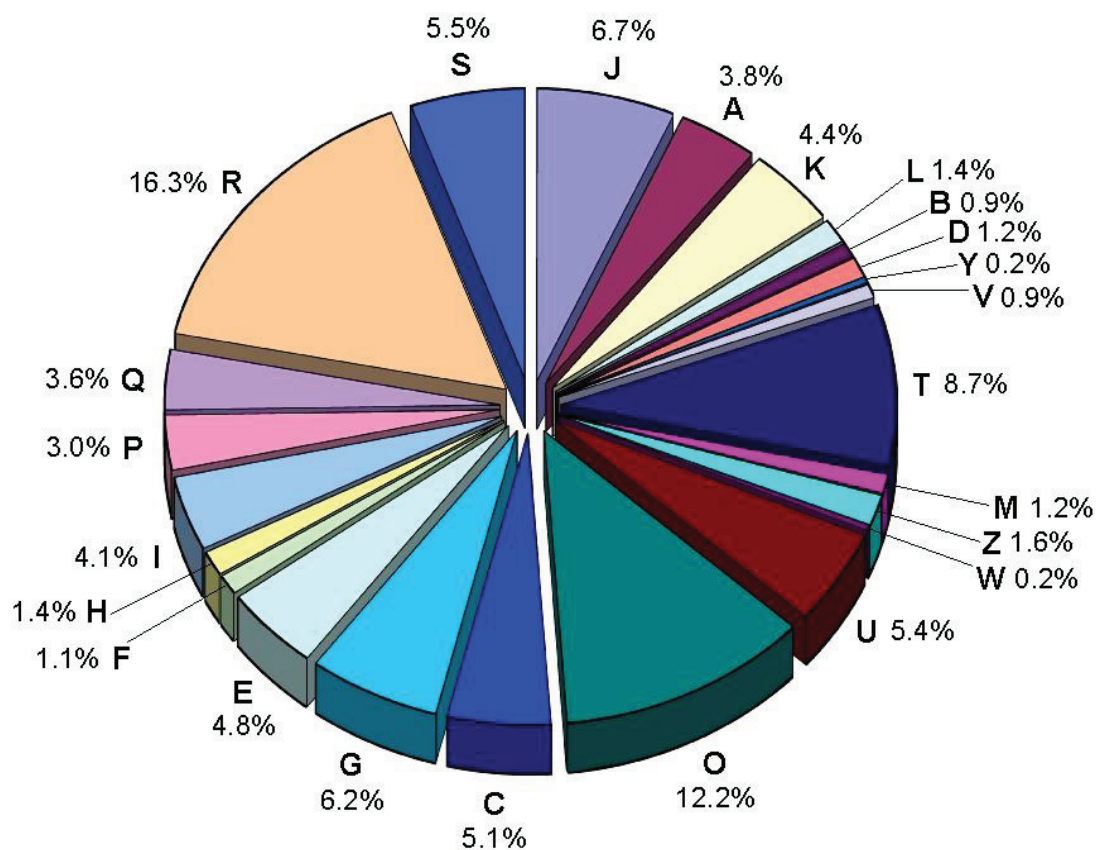


図 2-1-1 ポプラ完全長 cDNA の機能分類

約 2 万種類のポプラ完全長 cDNA をそれぞれの塩基配列がコードするタンパク質の機能によって分類した。J: タンパク質の翻訳, A: RNA のプロセッシングと修飾, K: 転写, L: DNA の複製・組換え・修復, B: クロマチンの構造変換, D: 細胞周期・細胞分裂の制御, Y: 核構造, V: 防御機構, T: シグナル伝達, M: 細胞壁・細胞膜の合成, Z: 細胞骨格, W: 細胞外構造, U: 細胞内輸送・分泌, O: タンパク質の翻訳後修飾・代謝, C: エネルギー生産・変換, G: 炭水化物の輸送・代謝, E: アミノ酸の輸送・代謝, F: 核酸の輸送・代謝, H: 補酵素の輸送・代謝, I: 脂質の輸送・代謝, P: 無機イオンの輸送・代謝, Q: 二次代謝産物の合成・輸送, R・S: 不明

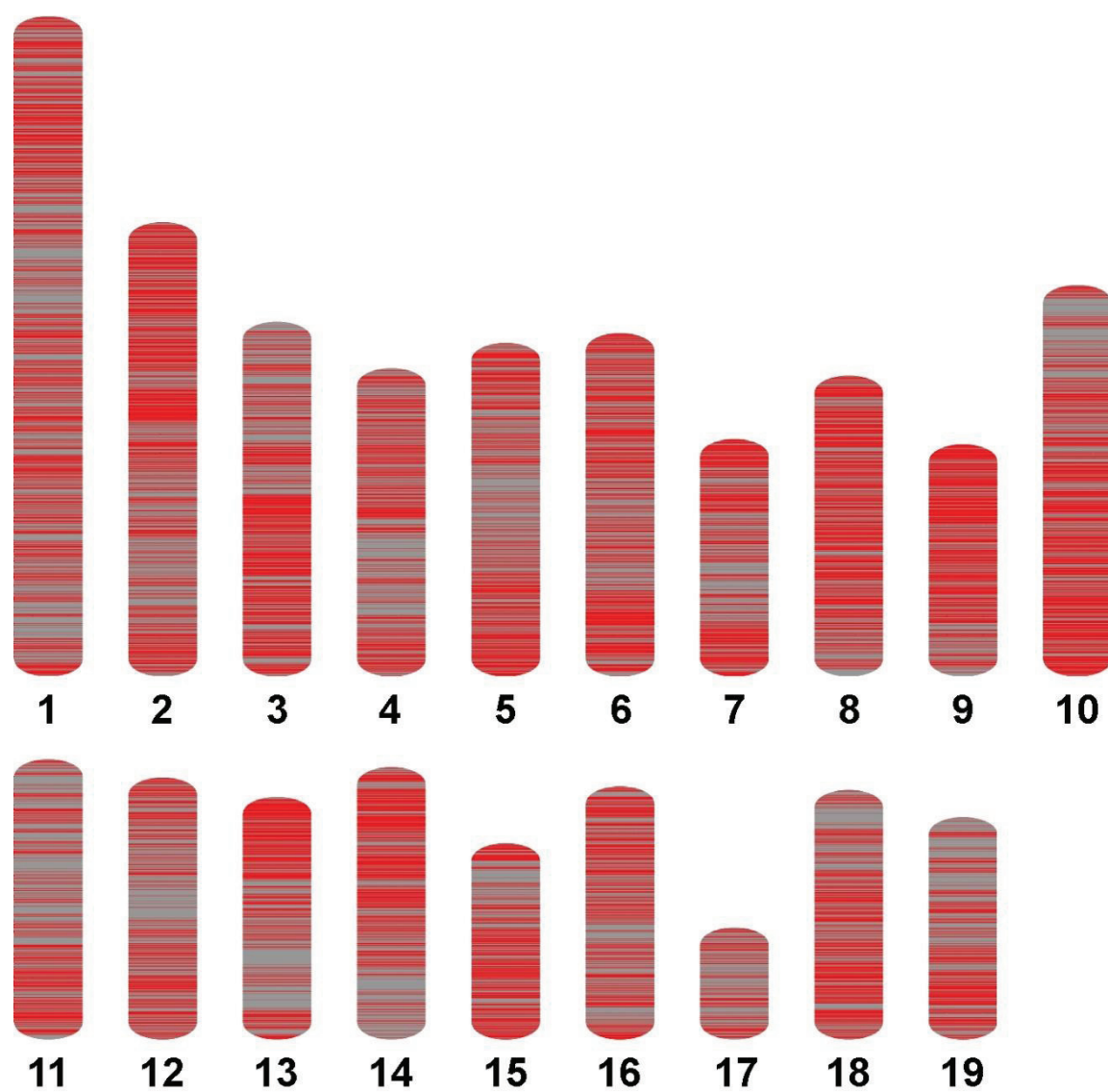


図 2-1-2 ポプラ完全長 cDNA の染色体上での位置

ポプラのゲノム情報をもとに、われわれが収集した約 2 万種類のポプラ完全長 cDNA を染色体上にマッピングした。棒状の図形は各染色体の模式図（染色体番号 1～19）であり、赤い部分がポプラ完全長 cDNA の位置を示している。

エ 考察

完全長 cDNA は、遺伝子の機能解析や組換え体の作出にとって有用なバイオリソース（生物資源）である。ポプラ完全長 cDNA は、樹木の構造ゲノム学的研究だけでなく、機能ゲノム学的研究の進展に大きく貢献すると考えられる。また、本課題で得られたポプラ遺伝子は、遺伝子組換えによる有用樹木の開発や、遺伝子マーカーとして用いることで環境耐性樹木の選抜などにも利用可能である。

オ 今後の問題点

ポプラ完全長 cDNA は、学術的にもまた応用研究にとっても非常に重要な公的バイオリソースである。現在ポプラのリソースについては、理研バイオリソースセンターに配布を寄託しているが、今後我々は、バイオリソースの恒久的な公開や提供を目指すための方策を再考しなければならない。

カ 要約

19,841 種類のポプラ (*Populus nigra*) 完全長 cDNA を収集した。これは、ポプラの全遺伝子数の約 40%に相当する。ポプラ cDNA を理化学研究所バイオリソースセンターに提供し、公的配布を寄託した。また、ポプラ遺伝子群の塩基配列情報を理化学研究所植物科学研究センターに提供し、データベースを通じて樹木の遺伝子情報発信に貢献した。ポプラ完全長 cDNA は、樹木の生理機能の解明や有用な組換え樹木の創出への応用が期待される。

キ 引用文献

- 1) 楠城時彦、篠崎一雄、篠原健司 (2004) ゲノム科学的手法を用いた樹木の環境ストレス応答機構の解明. 日本林学会誌 86: 69-73.
- 2) Nanjo, T., Futamura, N., Nishiguchi, M., Shinozaki, K., Shinohara, K. (2004) Characterization of full-length enriched expressed sequence tags of stress-treated poplar leaves. *Plant Cell Physiol.* 45: 1738-1748.
- 3) Nanjo, T., Sakurai, T., Totoki, Y., Toyoda, A., Nishiguchi, M., Kado, T., Igasaki, T., Futamura, N., Seki, M., Sakaki, Y., Shinozaki, K., Shinohara, K. (2007) Functional annotation of 19,841 *Populus nigra* full-length enriched cDNA clones. *BMC Genomics* 8: 448.

(楠城時彦)

2. ポプラ環境ストレス応答性遺伝子の特定

ア 研究目的

樹木の環境ストレス応答機構や耐性機構の解明は、地球温暖化の軽減等持続的な環境保全を考える上で大変重要な課題である。本研究では、乾燥、高塩濃度や低温等の環境ストレスに対する樹木の分子応答機構を解析するために、DNA マイクロアレイによるポプラ遺伝子の網羅的発現解析を行った。DNA マイクロアレイを使うと、数万種から数十万種におよぶ遺伝子の発現を一回の実験で解析することが可能である。本研究は、ポプラの DNA マイクロアレイを使って、環境ストレスを加えた時に発現が変動する遺伝子を同定することを目的とした。

イ 研究方法

無菌栽培したポプラ（セイヨウハコヤナギ、*Populus nigra* var. *italica*）雌株の本葉に、乾燥、高塩濃度、低温及びアブシジン酸（ABA）の各種環境ストレス処理を 2 時間もしくは 10 時間行った。処理前後のサンプルからトータル RNA を抽出し、アガロース電気泳動により品質を確認した。得られた RNA とアフィメトリックス社製ポプラジーンチップ（DNA マイクロアレイ）をハイブリダイズさせた。対照区（処理前サンプル）とストレス処理後の間で、発現が有意に変動する遺伝子を解析した（図 2-2-1）。

ウ 結果

DNA マイクロアレイ解析の結果、乾燥、高塩濃度、低温、ABA いずれかの環境ストレス処理により発現が変動する遺伝子は 2,214 種であった。これらの遺伝子群を発現パターンが似た遺伝子ごとに分類した（図 2-2-2）。これにより、ポプラの遺伝子の発現パターンが鳥瞰できるだけでなく、機能が明らかな遺伝子が属するクラスターの情報をもとに、そのクラスターに含まれる新規遺伝子の機能を推定することができる。

また、処理前後で発現変動が見られた 2,214 種の遺伝子をストレス処理種別に分類した（図 2-2-3）。ストレス処理により発現が上昇したものが 1,382 種、ストレス処理により発現が下降したものが 843 種であり、ストレス種により発現の上昇と下降の両方が確認されたものが 11 種であった（例えば乾燥処理では上昇するが低温処理では下降するもの）。

さらに、環境ストレス処理により発現が上昇したポプラ遺伝子から推定したタンパク質の機能を調べた。その結果、細胞内の浸透圧調節に関わるオリゴ糖合成酵素やアミノ酸（プロリン）合成酵素、ストレスシグナルの伝達に関わるプロテインキナーゼや転写因子（DNA 結合タンパク質）、植物ホルモンの代謝に関わる ABA 合成酵素やチトクローム P450、耐凍性に関わる脂肪酸不飽和化酵素、組織内の物質輸送に関わる水輸送タンパク質やイオン輸送タンパク質をコードする遺伝子の発現が顕著に上昇していた（表 2-2-1）。

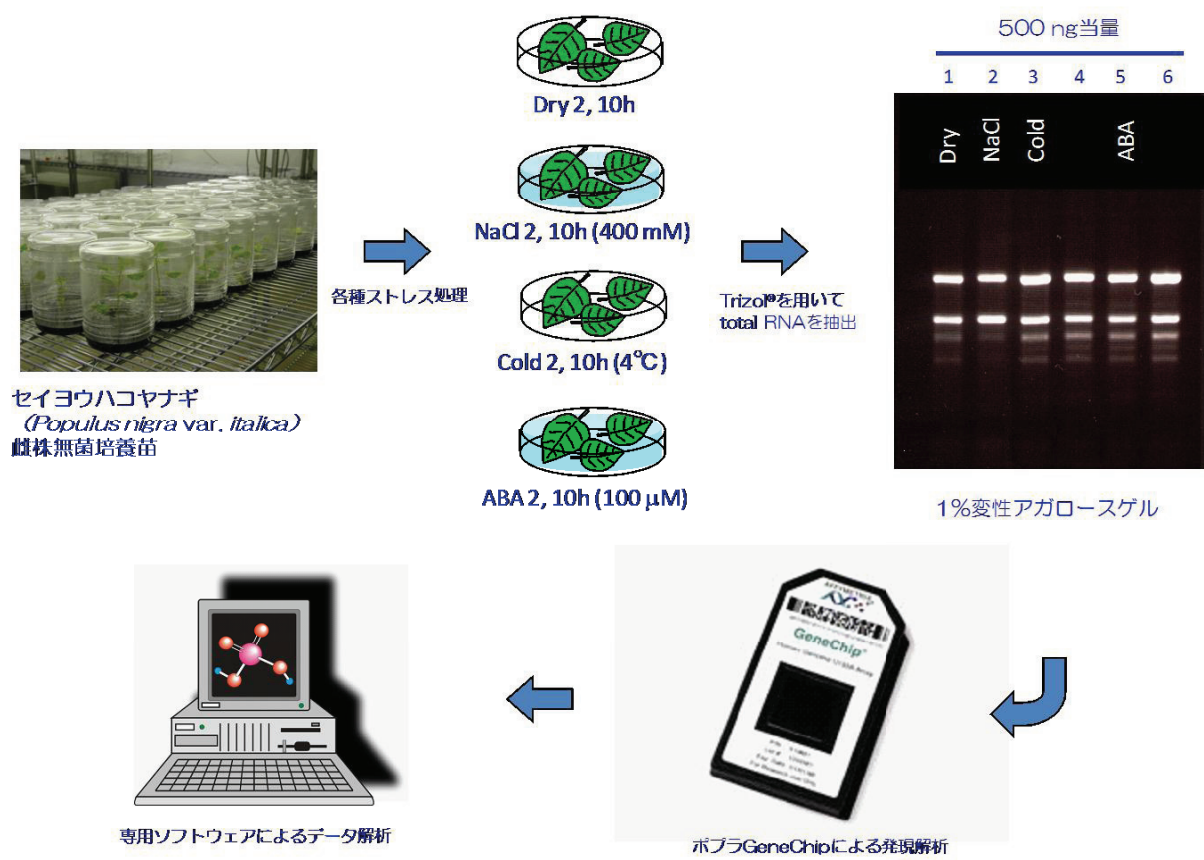


図 2-2-1 DNA マイクロアレイ解析の流れ

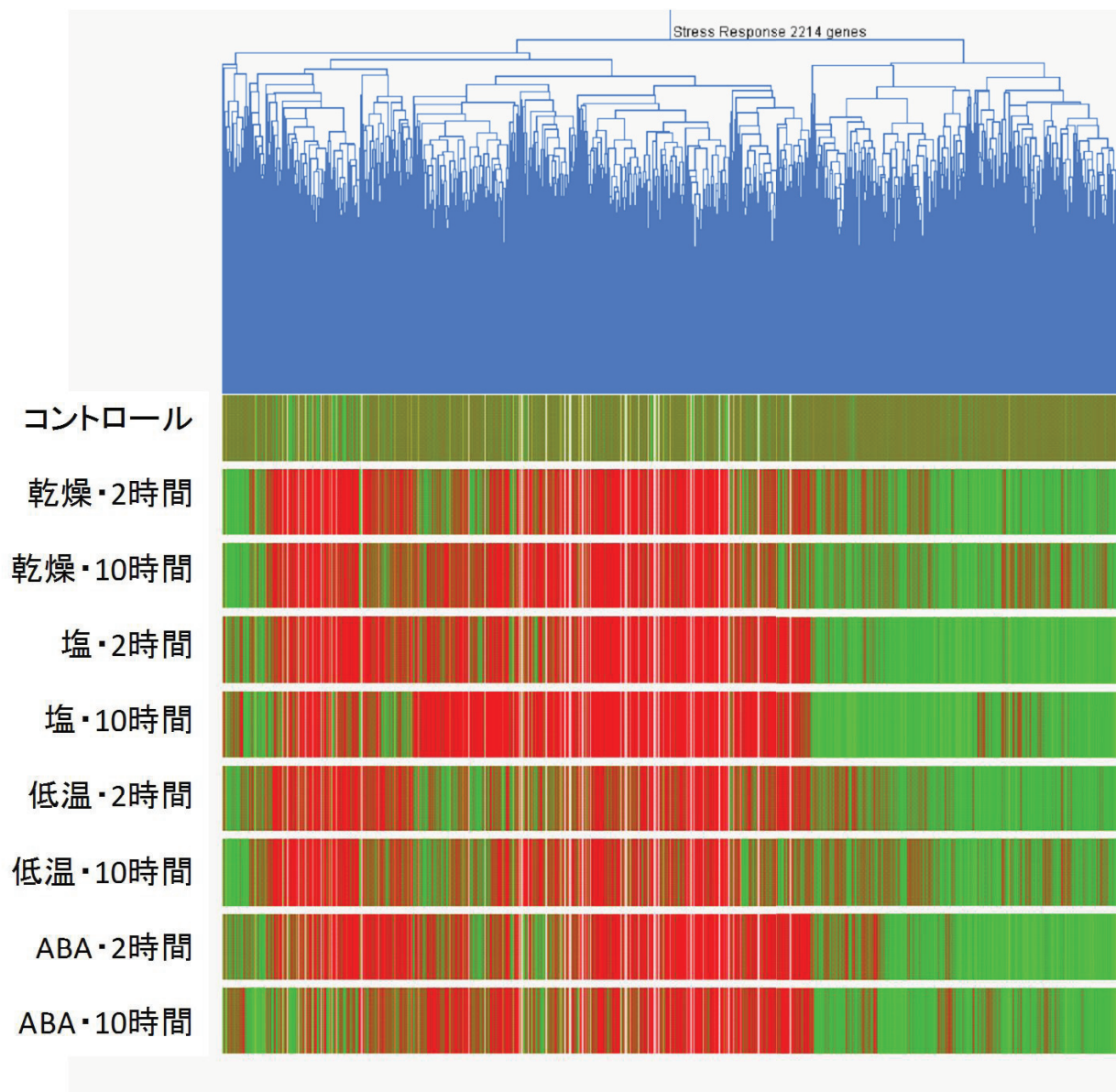


図 2-2-2 環境ストレス応答性遺伝子の階層的クラスタリング

赤色で示したものは処理により発現が上昇した遺伝子であり、緑色で示したものは発現が下降した遺伝子である。また、茶色で示したものは処理の前後で発現変動が見られなかったものである。

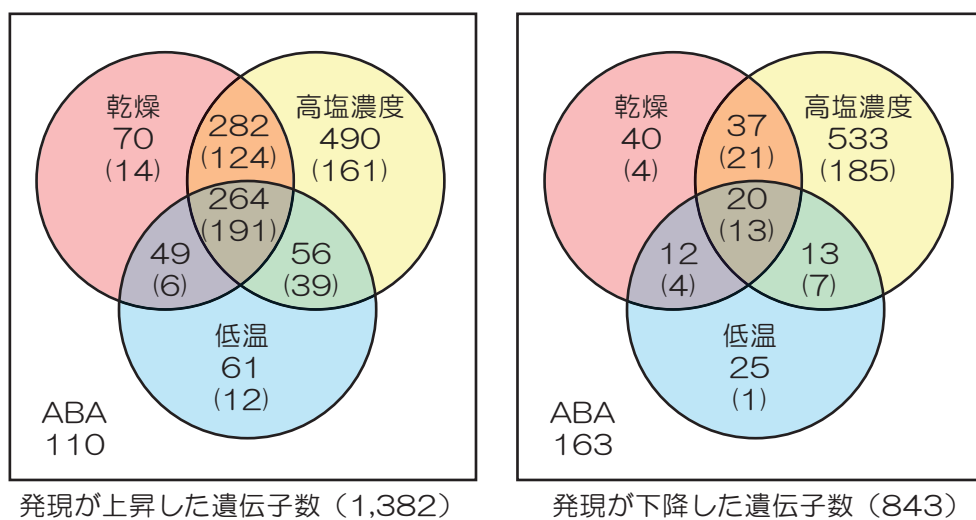


図 2-2-3 環境ストレス応答性遺伝子のグルーピング

乾燥、高塩濃度、低温、アブシジン酸（ABA）の各処理により発現が変動した遺伝子数。左がストレス処理により発現が上昇した遺伝子グループ、右がストレス処理により発現が下降した遺伝子グループ。括弧内は各処理区のうち ABA 処理による発現変動が見られたもの（内数）を示す。

表 2-2-1 ストレス処理により発現が上昇したポプラ遺伝子から推定したタンパク質の機能

タンパク質の機能	ストレス処理
オリゴ糖合成酵素	乾燥、高塩濃度、低温、アブシジン酸 (ABA)
チトクローム P450	乾燥、高塩濃度、低温、アブシジン酸 (ABA)
プロテインキナーゼ	乾燥、高塩濃度、低温、アブシジン酸 (ABA)
転写因子 (DNA 結合タンパク質)	乾燥、高塩濃度、低温、アブシジン酸 (ABA)
水輸送タンパク質	乾燥、高塩濃度
アブシジン酸合成酵素	乾燥
脂肪酸不飽和化酵素	低温
イオン輸送タンパク質	高塩濃度
アミノ酸 (プロリン) 合成酵素	高塩濃度

エ 考察

ストレス処理により発現レベルが変動した遺伝子は、いずれも樹木の環境ストレス応答に深く関わっている可能性がある。本研究で得られた成果は、樹木の環境ストレス応答機構解明の足がかりとなるだけでなく、樹木独自のストレス誘導性プロモーターの開発、環境ストレス耐性遺伝子組換え樹木の開発、さらに耐性品種の DNA 選抜マーカーの開発にも役立つ。

オ 今後の問題点

DNA マイクロアレイは、ある生理現象に関与する遺伝子群の発現解析を一挙に行える点で革新的な技術である。しかし、個々の遺伝子の機能を解明するためには、発現様式をさらに高精度の実験により詳細に解析する必要がある。今後は、マイクロアレイ解析により特定された有用候補遺伝子群の絞り込みおよび機能解明を通して、応用研究等に発展させていく必要があるだろう。

カ 要約

DNA マイクロアレイによる網羅的発現解析を進め、乾燥、高塩濃度、低温及びアブシジン酸の各処理によって発現が変動するポプラの環境ストレス応答性遺伝子を特定した。これらの遺伝子は、高環境ストレス耐性組換え樹木の開発に役立つ。

キ 引用文献

- 1) 楠城時彦、出村拓、大山幸子、篠原健司 (2009) ポプラ環境ストレス応答性遺伝子群の網羅的発現解析. 日本森林学会大会学術講演集 120: Pb3-31.

(楠城時彦)

「交付金プロジェクト」は、平成13年度に森林総合研究所が独立行政法人となるにあたり、これまで推進してきた農林水産技術会議によるプロジェクト研究（特別研究など）の一部、および森林総合研究所の経費による特別研究調査費（特定研究）を統合し、研究所の運営費交付金により運営する新たな行政ニーズへの対応、中期計画の推進、所の研究基盤高揚のためのプロジェクト研究として設立・運営するものである。

この冊子は、交付金プロジェクト研究の終了課題について、研究の成果を研究開発や、行政等の関係者に総合的且つ体系的に報告することにより、今後の研究と行政の連携協力に基づいた効率的施策推進等に資することを目的に、「森林総合研究所交付金プロジェクト研究成果集」として刊行するものである。

ISSN 1349-0605

森林総合研究所交付金プロジェクト研究 成果集 25
「ポプラ等樹木の完全長 cDNA 塩基配列情報の充実」

発 行 日 平成 22 年 1 月 14 日

編 集・発 行 独立行政法人 森林総合研究所

〒 305-8687 茨城県つくば市松の里 1 番地

電話. 029-873-3211 (代表)

印 刷 所 株式会社デジタル印刷

