

## 【特集】ゲノム情報に基づく林木育種の可能性 —スギを対象として—

## 発現遺伝子解析による材関連遺伝子マーカー開発の可能性

三嶋 賢太郎<sup>\*1</sup>・藤原 健<sup>2</sup>・井城 泰一<sup>1</sup>・黒田 克史<sup>2</sup>・山下 香奈<sup>2</sup>・田村 美帆<sup>3</sup>・藤澤 義武<sup>1</sup>・渡辺 敦史<sup>3</sup>

## はじめに

材質形質は針葉樹における最も重要な育種目標の一つであり、森林総合研究所林木育種センターではスギ精英樹の材質に関わる諸形質を順次評価してきた。これに加え、ゲノム情報を利用した早期選抜手法についても模索している。しかし、スギの材形成関連遺伝子の集積は十分ではないことから、材形質の早期選抜及び材質育種を行うためにはゲノム情報の基盤整備を行う必要がある。そこで、材形質の評価済みの精英樹から数クロンを選び、材形成を担う形成層帯から時系列に沿ってサンプリングを行い、形成層帯部位特異的 cDNA ライブラリー (cDNA ライブラリー: メッセンジャーRNA から逆転写反応によって合成された DNA をプラスミド等に組み込んだもの) の作成及びマイクロアレイ解析 (マイクロアレイ解析: DNA 断片をガラス板状に高密度に多数配置しておいたものに対象のサンプルをハイブリダイズすることによって一度に多数の遺伝子の発現情報を得る解析手法) を行ってきた。本稿では、これらの過程で得られた知見を既に報告した Mishima et al. (2014) から一部を引用する形で報告したい。

## 形成層帯からの網羅的遺伝子収集

現在までに材形成に関わる発現遺伝子 (EST: Expressed Sequence Tags の略) の配列情報の集積とその発現パターンに関する研究は、*Pinus*、*Populus*、*Picea* 等の主要な育種対象樹種で先行して行われてきている (Allona et al

1998; Sterky et al 1998; Pavy et al 2008)。スギにおける材形成関連の EST は、ForestGen (Forest EST and Genome database) に内樹皮や辺材、Yoshida et al. (2012) による白線帯 (移行帯) の EST が登録されているにすぎず、欧米での取り組みとは大きな開きがあった。これは、日本においては「スギ花粉症」が注目されてきたために雄花の EST の蓄積が優先されてきたことも一因と考えられる。我々は、日上市にある林木育種センター場内に植栽されている数クロンを対象として、3月9日、4月9日、5月10日、6月1日、6月24日、7月16日、8月16日、9月19日、9月29日、10月29日の合計10回、定時に形成層を含む短冊状の木片を採取毎に同一クロンの別個体から採取した。尚、これらと同時期に組織観察用のサンプル片も採取した。採取した木片を形成層帯付近で割り、割った内側の両面を削ぎ落とした細胞/組織を RNA 抽出に用いた。また、RNA 抽出後のサンプル片を顕微鏡観察し、用いた組織が師部と形成中木部 (拡大帯) を含む形成層帯であることも確認した。抽出した10の RNA サンプルを混合した後、標準化することなく、次世代シーケンサー (Roche454) を利用して発現遺伝子の塩基配列情報、約30万リードを取得した。これらのリードをアセンブル後、平均長1,069bpの14,616 isotigs (isotig: リードのアセンブルの結果得られた contig をさらに連結して得られた配列) と40,435 singlets (singlet: オーバーラップ配列がなく、アセンブルされなかったリード) の合計55,051の配列が得られた (表-1)。

取得した55,051配列は、NCBI (National Center for Biotechnology Information) non-redundant protein database、TAIR (The Arabidopsis Information Resource)、ForestGen 及

\*E-mail: mishimak@affrc.go.jp

<sup>1</sup> みしま けんたろう、いき たいいち、ふじさわ よしたけ 森林総合研究所林木育種センター<sup>2</sup> ふじわら たけし、くろだ かつし、やました かな 森林総合研究所<sup>3</sup> たむら みほ、わたなべ あつし 九州大学大学院農学研究院

表-1 得られた配列とアセンブルの概要

	配列数	塩基数 (Mbp)
配列		
総リード数	308,542	125.1
平均リード長	405.29 bp	
アセンブル		
使用不能なリード数	9764	
アセンブルに使用したリード数	298,778	121.2
平均リード数	405.69 bp	
Contig		
総 Contigs		
contig としてアセンブルされたリード数	241,696	98.6
contig 数 (100bp 以上)	15,521	12.7
Contigs (500bp 以上)		
contig 数	11,022	11.2
平均 contig 数	1,014	
最長 contig 長	9,656	
50%以上の contig 長	1,102	
Isotigs		
isotig 数	14,616	15.6
平均 isotig 長	1,069	
最長 isotig 長	9,656	
50%以上の isotig 長	1,261	
Isotig を形成する contig の平均数	1.7	
Singletons	40,435	
総配列数	55,051	

び ForestGen の材関連データベース (inner bark and sapwood ライブラリー)、The Gene Index の *Pinus*、*Populus* や *Picea* における木部 (Xylem) 及び形成層 (Cambium) に関する既知のライブラリーとの塩基配列の類似性を BLAST 検索により解析した。当然ではあるが、ForestGen の登録配列と最も類似性が高いことに加え、広葉樹である *Populus* よりも針葉樹との類似性の方が高いという結果が得られた。また、木部及び形成層に関する既知のライブラリーとの比較では、今回取得した配列は ForestGen の材関連データと最も類似性が低く、むしろ既に良く知られている *Pinus*、*Populus* 及び *Picea* との類似性の方が高かった。これらの結果は、今回取得した配列が、ForestGen における材関連の EST データの不足を補うとともに、他器官で取られた EST との共通性を示唆するものと考えられる (図-1A、B)。

### 形成層帯における発現遺伝子情報の収集

上記によって得られた 55,051 配列のうち、高発現していると考えられる 14,616 isotigs に加え、ForestGen の既知の材関連データベースの配列を加えて、18,082 プローブを 3-4 反復した 72K マイクロアレイスライドガラスを設計した。設計したマイクロアレイスライドガラスに形成層帯の異なる日付の cDNA サンプルをハイブリダイズすれば、複数時期の 18,082 遺伝子の発現挙動を一度に捉えることができる。

供試した時期については成層帯の活動期における細胞の変化を一通り捉えられるように 3月24日、4月27日、6月22日、8月27日、10月7日の5時期を選び形成層帯の組織を観察した (図-2A、B)。半径方向の形成層細胞は3月24日に採取したサンプルでは細胞分裂が見られず、組織観察上ではまだ休眠状態であると判断できた。

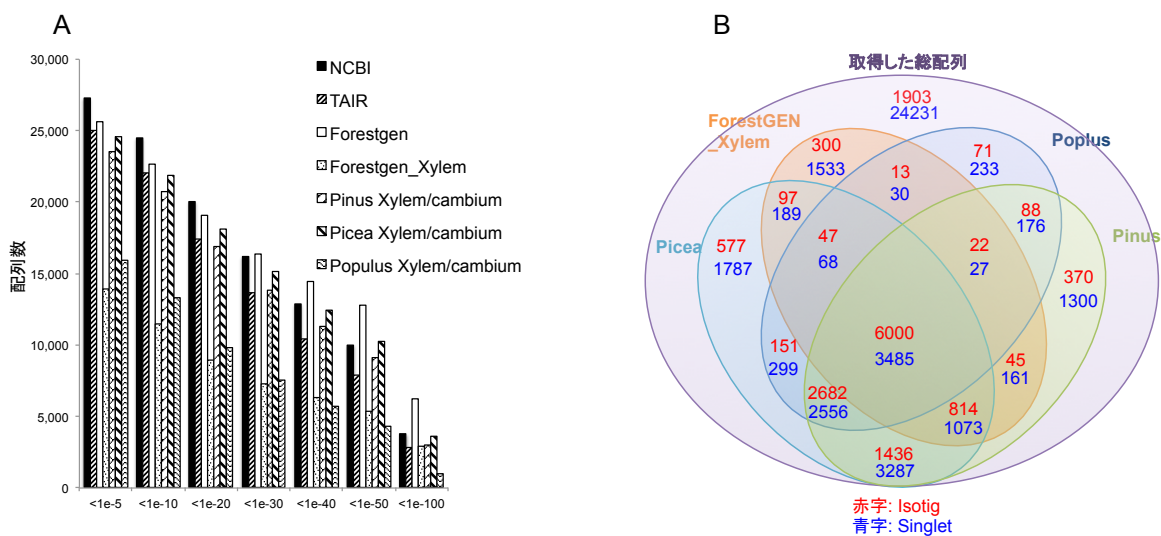


図-1 配列の類似性。A) BLASTx/tBLASTx *E*-value cutoff 値による NCBI、TAIR、ForestGen、ForestGen\_Xylem、pine、spruce and poplar データベースと取得した配列と類似する配列数。B) tBLASTx 検索 (*E*-values < 1E-5) による ForestGen\_Xylem、pine、spruce and poplar データベースと取得した配列間の類似する配列数間のベン図。

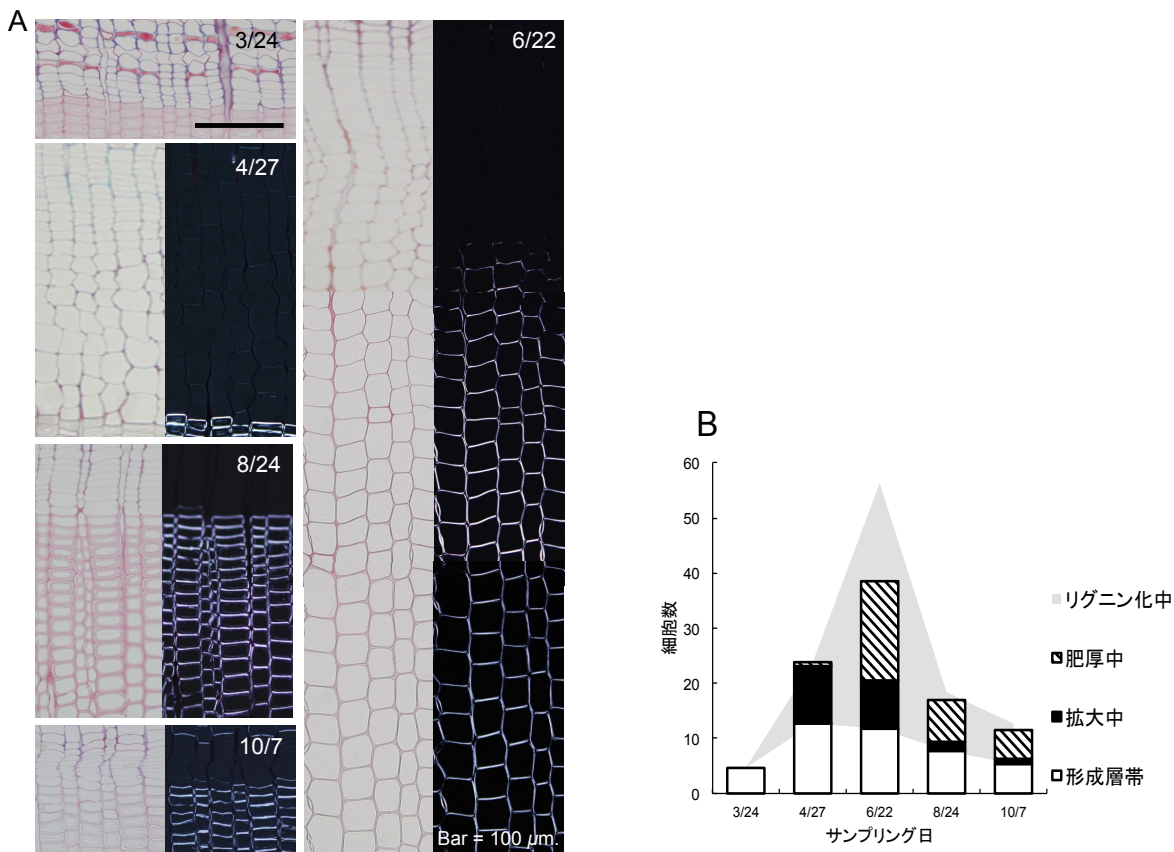


図-2 クロスセクションにおける形成層および分化中木部の組織。A) 光学顕微鏡下で観察されたクロスセクションにおける通常 (左) と、偏光 (右) 画像 (3/24、4/27、6/22、8/24、10/7)。B) 形成層帯と分化中木部の細胞数。分化中木部の細胞における拡大中、肥厚中、リグニン化中の細胞は光学顕微鏡下での通常/偏光観察によって判断した。

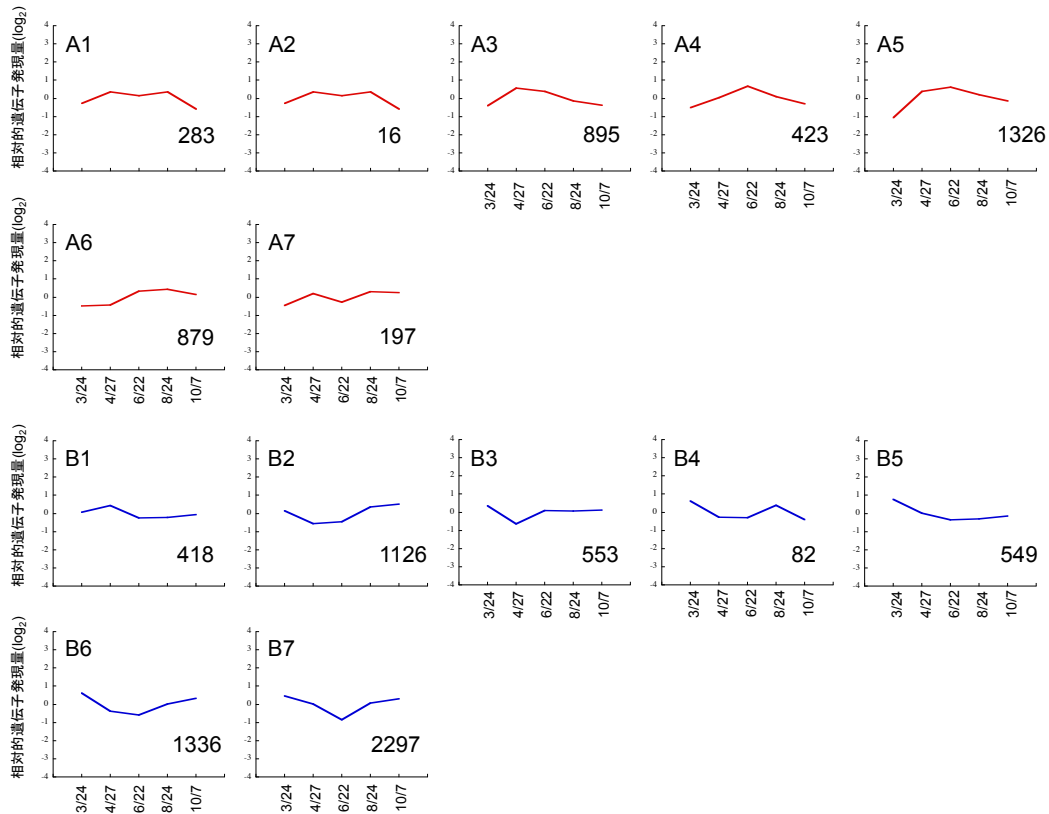


図-3 材形成期間における遺伝子の類似発現パターン。10380 遺伝子の発現パターンは、ピアソン相関係数により 14 のグループに分けられた。グラフは、各クラスターの遺伝子の発現量の平均を  $\log_2$  変換して示す。遺伝子発現パターンは、材形成期間中に A : 高発現、B : 低発現として示した。

3月24日の平均形成層細胞数(4.8個)は、4月( $p < 0.01$ )、6月( $p < 0.01$ )、8月( $p < 0.05$ )のサンプルで見られるよりも有意に少なかった。拡大中仮道管については4月27日のサンプルにおいて最も広く、仮道管数も最大となり、形成層から派生した細胞の仮道管形成が最も盛んであると考えられた。また、4月27日のサンプルにおいて一部の仮導管には木化が見られ、二次壁形成はこの頃から始まると考えられた。6月22日のサンプルでは、二次壁形成とリグニン化の段階にある仮導管数が最大となり、他の時期より有意に多くなるため( $p < 0.01$ )、6月22日のサンプルは、細胞分化及び二次壁形成を含む木部形成の活動のピークであると思われた。6月22日に拡大中の仮道管数の平均数が8.8細胞であったものが、8月24日には1.7細胞大幅に減少しており( $p < 0.01$ )、形成層活動が初期と比べて減退していることを示していると判断できた。従って、8月24日及び10月7日には、細胞の分裂及び拡大というよりも二次壁形成とリグニン化が主に行われていると考えられる(図-2A、B)。

上記のように、組織観察によって年間の形成層活動を定義づけた5時期のサンプルを用いてマイクロアレイにより発現遺伝子の経時的発現パターンを調べた。その結果、18,082 遺伝子中 10,380 遺伝子が材形成期間に有為異なる挙動を見せた( $p < 0.05, q < 0.2$ )。発現パターンは、形成層活動開始時期及び材形成のピーク時に発現が上昇する 4,019 遺伝子とそれとは反対に分裂の減少期及び成長停止期に発現が上昇する 6,361 遺伝子の大きく2つに分かれた。さらにそれぞれのパターンは、7ずつに細かく分かれ、合計 14 のパターン(クラスター)に分割できた(図-3)。このうち、クラスターA5は、形成層活動開始時期及び材形成のピーク時に最も多くの遺伝子が高発現しているクラスターとなった。同様に、分裂の減少期及び成長停止期には、クラスターB2、B6、B7といったクラスターに属する遺伝子が高発現しているという結果となった。

各クラスターにおける全ての遺伝子の配列情報は、COGs (clusters of orthologous groups) database に対して、BLSTX 検索により類似性を解析し、遺伝子の機能を分類

表-2 各クラスターとCOGデータベースの分類に基づく発現遺伝子数

	A1	A2	A3	A4	A5	A6	A7	B1	B2	B3	B4	B5	B6	B7
[A] RNA processing and modification	3	0	6	3	11	21	0	3	23	9	4	15	60	123
[B] Chromatin structure and dynamics	1	0	7	0	13	2	4	8	10	1	1	5	18	78
[C] Energy production and conversion	20	0	25	26	37	29	0	8	31	18	0	6	32	10
[D] Cell cycle control, cell division, chromosome partitioning	1	0	8	4	23	5	0	3	12	3	0	2	20	21
[E] Amino acid transport and metabolism	7	2	32	11	39	28	4	10	23	16	1	7	8	53
[F] Nucleotide transport and metabolism	0	0	2	0	12	11	0	1	8	2	2	4	6	15
[G] Carbohydrate transport and metabolism	9	0	65	26	113	41	9	29	32	23	2	19	69	47
[H] Coenzyme transport and metabolism	11	0	2	7	19	7	1	1	8	2	0	1	3	5
[I] Lipid transport and metabolism	6	0	34	10	53	33	9	13	37	18	1	18	17	40
[J] Translation, ribosomal structure and biogenesis	7	0	28	1	11	8	0	35	9	5	0	4	13	42
[K] Transcription	15	2	19	7	30	23	3	12	59	24	2	20	91	98
[L] Replication, recombination and repair	0	0	8	5	8	4	2	7	16	4	1	5	18	35
[M] Cell wall/membrane/envelope biogenesis	2	1	29	9	65	17	5	11	19	3	0	3	14	38
[N] Cell motility	0	0	0	2	1	0	0	0	1	0	0	0	1	2
[O] Posttranslational modification, protein turnover, chaperones	18	0	30	37	51	84	29	21	44	33	10	20	62	110
[P] Inorganic ion transport and metabolism	9	0	24	8	11	15	5	7	8	8	1	9	12	48
[Q] Secondary metabolites biosynthesis, transport and catabolism	8	2	17	11	50	31	9	18	68	36	1	7	36	48
[R] General function prediction only	16	0	104	52	133	121	26	45	159	60	7	50	132	263
[S] Function unknown	18	0	61	23	87	39	5	22	37	50	4	33	68	69
[T] Signal transduction mechanisms	23	4	97	22	126	40	32	51	77	40	10	88	144	435
[U] Intracellular trafficking, secretion, and vesicular transport	5	0	48	26	45	17	3	20	12	6	0	8	25	28
[V] Defense mechanisms	2	0	8	8	35	19	7	19	20	15	3	55	64	227
[W] Extracellular structures	1	0	2	1	22	2	1	0	0	0	0	0	1	3
[X] Unassigned	21	2	70	25	123	67	24	19	87	36	5	43	98	146
[Y] Nuclearstructure	1	0	3	0	1	0	1	5	1	0	0	3	0	9
[Z] Cytoskeleton	4	0	32	19	41	9	0	14	5	1	0	9	15	8

化した。その結果、形成層活動開始時期及び材形成のピーク時に高発現すると考えられ、材形成遺伝子としてよく知られている機能グループである「炭水化物の輸送と代謝 (Carbohydrate transport and metabolism)」、「細胞壁・膜組織等の生合成 (Cell wall/membrane/envelope biogenesis) や「細胞骨格 (Cytoskeleton)」がクラスターA5 の遺伝子群において顕著に見られた。一方、分裂の減少期及び成長停止期では様々な環境状態に対する耐性や細胞プロセスに関連すると考えられる「転写後の加工過程と修飾 (RNA processing and modification)」、「シグナル伝達機構 (Signal transduction mechanisms)」や「防御機構 (Defense mechanisms)」等の機能グループが顕著に見られた (表-2)

### 終わりに

本稿では、形成層活動期の形成層帯からの EST 集積とそれに基づくマクロアレイ解析による発現遺伝子解析の概要を述べた。詳細なマイクロアレイ解析から個々の遺伝子の挙動については、原著論文 Mishima et al. (2014) を参照されたい。

現在、本稿で書いた形成層活動期に引き続き、形成層活動休眠期及び形成層帯以外の木部 (心材含む) についても EST 集積及び全てのライブラリーを網羅したマイクロアレイの設計も終わっており、さらに詳細にスギ材形成に関わる遺伝子の季節及び器官における挙動等を明らかにしていく予定である。

一方で、この EST ライブラリーをリファレンスとしてリシーケンスを行うことによって、約 30 万以上の SNP (一塩基多型) 情報も取得しており、マッピング集団を用いた QTL 解析及び精英樹を用いたアソシエーション解析にも取り組んでいる。今後は、SNP 情報を用いた順遺伝学的方法 (Forward genetics : 各個体のゲノム情報や個体間の血縁関係の情報等) を利用し、それらの個体から得られる後代の遺伝的パフォーマンスを予測的に推定する遺伝学的手法) にマイクロアレイ解析で明らかになる発現

情報を加味することによって、可能な限り効率的かつ効果的に材質育種のための選抜マーカーを得たいと考えている。

### 引用文献

- Allona I, Quinn M, Shoop E, Swope K, St Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 95: 9693–9698
- Mishima K, Fujiwara T, Iki T, Kuroda K, Yamashita K, Tamura M, Fujisawa Y, Watanabe A (2014) Transcriptome sequencing and profiling of expressed genes in cambial zone and differentiating xylem of Japanese cedar (*Cryptomeria japonica*). *BMC genomics* 15: 219
- Pavy N, Boyle B, Nelson C, Paule C, Giguère I, Caron S, Parsons LS, Dallaire N, Bedon F, Bérubé H, Cooke J, Mackay J (2008) Identification of conserved core xylem gene sets: conifer cDNA microarray development, transcript profiling and computational analyses. *New Phytologist* 180: 766–786
- Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarreal R, Van Montagu M, Sandberg G, Olsson O, Teeri TT, Boerjan W, Gustafsson P, Uhlén M, Sundberg B, Lundeberg J (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags. *Proceedings of the National Academy of Sciences of the United States of America* 95: 13330–13335
- Yoshida K, Futamura N, Nishiguchi M (2012) Collection of expressed genes from the transition zone of *Cryptomeria japonica* in the dormant season. *Journal of Wood Science* 58: 89–103