

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 18 日現在

機関番号：10101

研究種目：挑戦的萌芽研究

研究期間：2012～2013

課題番号：24658130

研究課題名(和文) 樹木の機能ゲノム学分野の創出に資するユニバーサルな網羅的遺伝子解析法の開発

研究課題名(英文) Development of universal transcriptome analysis method toward functional genomics in woody plants

研究代表者

齋藤 秀之 (SAITO, Hideyuki)

北海道大学・(連合)農学研究科(研究院)・講師

研究者番号：70312395

交付決定額(研究期間全体)：(直接経費) 3,100,000円、(間接経費) 930,000円

研究成果の概要(和文)：非モデル植物であるブナ目樹木の機能ゲノミクス研究の基盤整備のため、DNAマイクロアレイ法の汎用性の検討と、ブナ(*Fagus crenata*)のドラフトゲノム構築を行った。ブナのDNAマイクロアレイ法は、ブナの地理的な遺伝変異に関わらず汎用性を持ってゲノム網羅的発現遺伝子解析を行える手法であることが確認された。構築したブナのドラフトゲノムは442Mb(カバー率が約80%)で、推定された遺伝子の数は48,618個であった。このブナのドラフトゲノムはブナ目で最初のドラフトゲノム情報であり、今後のブナ目樹木の機能ゲノミクス研究に役立つ基盤情報であると考えられた。

研究成果の概要(英文)：To advance the functional genomics of non-model woody plant, *Fagus crenata* (Fagaceae), an ecologically important tree species, two studies was conducted as follows; (1) DNA microarray created from a *F. crenata* cDNA sequence data-set is examined if versatile tool to study transcriptome comparing among intraspecific variations or not, and (2) *F. crenata* draft genome was generated. The transcriptome analysis by using the *F. crenata* DNA microarray was available for three populations, Kuromatsunai in Hokkaido Island, Tanzawa and Fuji in Honshu Island, suggesting the method is versatile in *F. crenata*. A draft genome of *F. crenata* was generated with 442 Mbp (including gap), allowing 48,618 predicted genes. This genomic information of *F. crenata* would contribute to accelerate the functional genomics of woody plants in family Fagaceae.

研究分野：農学

科研費の分科・細目：森林学・森林科学

キーワード：機能ゲノム学 DNAマイクロアレイ トランスクリプトーム ドラフトゲノム ブナ

1. 研究開始当初の背景

樹木の成長は遺伝効果と環境効果で規定される。育苗、育種、育林など造林学の課題では、遺伝変異を含んだ樹木集団を扱うため、遺伝 vs. 環境効果を区別した統合的評価を必要とするが、実際には学術的・技術的に困難である。遺伝 vs. 環境効果の統合的評価という生物の根源的な理解へ迫る研究課題に対して、樹木でも取り組む必要がある。

近年、遺伝 vs. 環境効果の統合的理解を目指す研究分野として機能ゲノム学(ポストゲノム学とも呼ばれる)が生まれた。機能ゲノム学は、ゲノム(遺伝子群の総体)の構造と機能に基づき遺伝子の発現とシグナル伝達・ネットワーク解析から成長制御や生理機能の環境応答を解き明かす学問分野である。

機能ゲノム学にはゲノム網羅的な発現遺伝子解析が欠かせない。その解析手法の一つに DNA マイクロアレイ法がある。DNA マイクロアレイ法はゲノム全体を対象に遺伝子の数や構造変異、発現量を解析できる実験手法で、ゲノム網羅的に遺伝子解析ができる。しかし、その解析には対象樹種の遺伝子の塩基配列情報が必要なため、多くの樹木では DNA マイクロアレイ法を利用できない現状にある。

そこで申請者らはブナの発現遺伝子の網羅的な塩基配列解読を実施して、真正双子葉類バラ目群の中で塩基配列の変異量を調べたところ、約 9 割が共通配列であることを明らかにした。したがって、バラ目群に属する樹種間では直系遺伝子(オーソログ)に共通配列が比較的高頻度に存在し、様々な種内の地理的遺伝変異やさらに近縁種などに関わらず、共通の DNA マイクロアレイを利用して発現遺伝子の解析が可能であると予想するに至った。

さらに最近、ゲノム網羅的な発現遺伝子の定量法の新規技術として次世代シーケンサーを用いた RNA-seq 法が台頭してきた。RNA-seq の長所は、DNA マイクロアレイ法に比べて発現遺伝子の判定の解像度が高いことなどが挙げられる。一般には、RNA-seq 法の長所の一つにゲノム基盤情報を持たない生物種でもゲノム網羅的な発現遺伝子解析が可能である事が挙げられているが、しかしながら、厳密にはドラフトゲノム配列情報が必要である。

2. 研究の目的

研究目的は次の 3 点である。

(1) DNA マイクロアレイ法によるゲノム網羅的な発現遺伝子解析法が野生集団であるブナの地理的遺伝変異に対応して汎用性のある解析が可能であることを示すこと

(2) ブナのゲノム網羅的な発現解析手法である DNA マイクロアレイ法や RNA-seq 法の精緻化のため、ブナのドラフトゲノムを構築すること

(3) ブナのゲノム情報を活用するためのデー

タベースを構築すること

3. 研究の方法

(1) ブナ DNA マイクロアレイ法の地理的遺伝変異に対する汎用性

調査地は北海道黒松内、神奈川県丹沢山系、静岡県富士山の 3 地域のブナ天然林であった。黒松内は DNA マイクロアレイのプロープ設計に用いられた cDNA 配列解析個体の地域である。また、丹沢山系および富士山のブナ林は、黒松内ブナ林に対して異なる遺伝的系統に位置する集団である。供試木は、各ブナ林の林冠を構成するブナ成木であり、供試葉は日当たりのよい陽樹冠の葉を対象にした。全 RNA は cTAB 法により抽出した。DNA マイクロアレイはアジレント社製のシステムを用いた。DNA マイクロアレイのプロープはブナ専用開発されたものを用いた(43,803 遺伝子で機能推定済み遺伝子が 12,446 遺伝子)。ゲノムワイドな全遺伝子を対象に発現量の正規生を調べ、DNA マイクロアレイによる遺伝子発現量の定量性について検討した。さらにゲノム網羅的な発現パターンの傾向を捉えるために、高温・乾燥・酸ストレス応答性遺伝子 1118 個(高温:477 遺伝子、乾燥:195 遺伝子、酸性:446 遺伝子)を対象に主成分分析により全体の傾向を検討した。

(2) ブナのドラフトゲノムの構築

ゲノム DNA は、黒松内ブナ天然林で選抜されたブナのゲノム解析木から採取した葉から抽出した。ライブラリーにはメイトペアで異なるサイズのインサージョンのライブラリーを作成した。インサージョンのサイズは、200bp, 500 bp, 800 bp, 2 kbp, 5 kbp, 10 kbp であった。塩基配列のシーケンスは HiSeq2000 (イルミナ社製)を用いた。HiSeq2000 で得られた塩基配列のクリーン・データから、さらに Trimmomatic を用いて品質コントロールを行った。ゲノムアセンブルにはゲノムアセンブルの専用ソフトウェア Augustus を用いて行った。遺伝子予測に用いた遺伝子モデルはシロイヌナズナであった。

(3) ブナ・ドラフトゲノムの Web 公開用データベース構築

JBrowse (development version, The Evolutionary Software Foundation 2013) を用いてシーケンスデータ・アセンブルによるゲノム配列データ、予測遺伝子、遺伝子注釈結果、遺伝子発現量のデータ等をデータベース化して、Web による公開利用可能なツール作成を行った。

4. 研究成果

(1) ブナ DNA マイクロアレイ法の地理的遺伝変異に対する汎用性

3 集団の全ての個体におけるゲノム網羅的な遺伝子発現パターンは、ゲノムワイドに同様な正規分布を示し、ブナの地理的遺伝変異

に関わらず DNA マイクロアレイ法がゲノム網羅的発現遺伝子解析に対して有効に定量できたことを示した。

主成分分析では、各集団ごとに特徴的なストレス応答の傾向を示しながらも、ブナ全体としては統一的な分布パターンを示し、ブナ DNA マイクロアレイ法がブナの地理的遺伝変異に対して汎用性があることを示唆した。

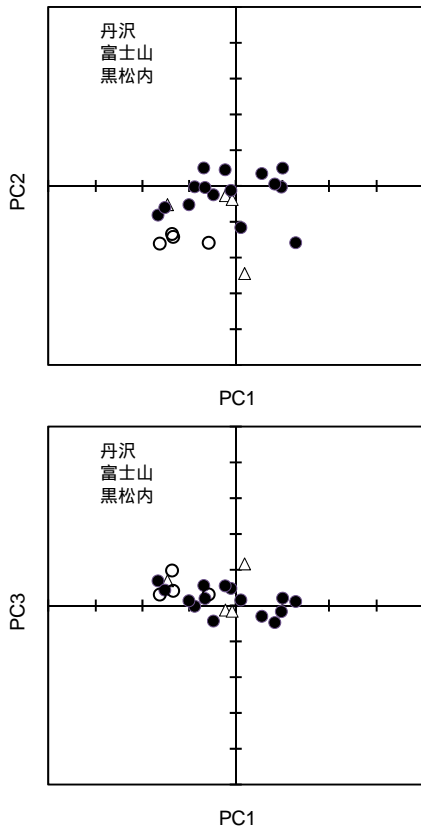


図 - 1 遺伝子発現パターンに基づく主成分分析

PC1, PC2, PC3 はそれぞれ高温・酸・土壌乾燥ストレスに対する評価軸を意味する。

(2) ブナのドラフトゲノムの構築

次世代シーケンスにより取得して品質コントロールされたクリーンデータは、200 bp insertion が 65896919 本、500 bp insertion が 50970997 本、800 bp insertion が 53507157 本、2 kbp insertion が 141451938 本、5 kbp insertion が 116922658 本、10 kbp insertion ライブラリーについては、ユニーク配列が 14.6% (クリーンデータ 56,513,144 本中 8,245,017 本) で大変少なかったためにゲノムアセンブルに反映させることが出来なかった。シーケンス・データをゲノム・アセンブルした結果、contig の全長が 366.78 Mb で、442 Mb (gap 込み) のドラフトゲノムが完成した (Scaffold の N50 値が 151 kb; Contig の N50 値が 7.4 kb)。GC 含有率は 35.7% であった。このブナのドラフトゲノムはブナの推定ゲノムサイズ (550 Mbp) の約 80% をカバーしていた。

推定された遺伝子の数は 48,618 個であった。この遺伝子数は、ゲノムモデルに用いたシロイヌナズナの 25,500 個よりも 2 倍近く多いが、ブナの属するブナ目と近縁のバラ目バラ科のチュウゴクナシの 42812 個やキントラノオ目ヤナギ科に属するブラックコットンウッド (ポプラ) の 45,555 個と近い数であった (表 - 1)。また、ブナの GC 含有率 (35.7%) は、シロイヌナズナの 34.7% (The Arabidopsis Genome Initiative, Nature 408,2000 から算出) と同程度であった。

表 - 1 全ゲノムの概要が解読された木本植物とその成果

和名	学名	分類	ゲノム研究の位置づけ
広葉樹 (被子植物)			
1 ブナ	<i>Fagus crenata</i>	ブナ目ブナ科	生態
2 ポプラ	<i>Populus trichocarpa</i>	キントラノオ目ヤナギ科	モデル樹木
3 セイヨウトネリコ	<i>Fraxinus excelsior</i>	シソ目モクセイ科トネリコ	生態
4 ブドウ	<i>Vitis vinifera</i>	ブドウ目ブドウ科	果樹
5			
6 リンゴ	<i>Malus domestica</i>	バラ目バラ科	果樹
7 モモ	<i>Prunus persica</i>	バラ目バラ科	果樹
8 ウメ	<i>Prunus mume</i>	バラ目バラ科	果樹
9 チュウゴクナシ	<i>Pyrus bretschneideri</i>	バラ目バラ科	果樹
10 エゾヘビイチゴ	<i>Fragaria vesca</i>	バラ目バラ科	果樹
11 オレンジ	<i>Citrus sinensis</i>	ムクロジ目ミカン科	果樹
12 パパイア	<i>Carica papaya</i>	アブラナ目パパイア科	果樹
13 カカオ	<i>Theobroma cacao</i>	アオイ目アオイ科	食料
14 ナツメヤシ	<i>Phoenix dactylifera</i>	ヤシ目ヤシ科	食料
15 ユーカリ	<i>Eucalyptus camaldulensis</i>	フトモモ目フトモモ科	環境・林業
16 ユーカリ	<i>Eucalyptus grandis</i>	フトモモ目フトモモ科	環境・林業
17 ジャコブ 苳り	<i>Jatropha curcas</i>	トウダイクサ目トウダイクサ科	バイオ燃料
18 トウゴマ	<i>Ricinus communis</i>	トウダイクサ目トウダイクサ科	油脂原料・薬
19 パラゴムノキ	<i>Hevea brasiliensis</i>	トウダイクサ目トウダイクサ科	天然ゴム・林業
針葉樹 (裸子植物)			
20 テーダマツ	<i>Pinus taeda</i>	マツ目マツ科	針葉樹・林業
21 グラウカトウヒ	<i>Picea glauca</i>	マツ目マツ科	針葉樹・林業
22 ドイツトウヒ	<i>Picea abies</i>	マツ目マツ科	針葉樹・林業
モデル植物 (被子植物)			
23 シロイヌナズナ	<i>Arabidopsis thaliana</i>	アブラナ目アブラナ科	モデル植物

推定ゲノムサイズ (Mb)	ゲノム解読サイズ (Mb)	カバー率 (%)	コーティング率 (%)	N50 遺伝子数	N50 Scaffold	N50 Contig
広葉樹 (被子植物)						
1	550	442	80.0	48,618	151	7.4
2	550	485	88.2	45,555		
3		875			98,766	5.516
4	475	487	100.0	30,434		
5	504.6	477.1	94.6	29,585		
6	742.3	603.9	81.4	57,386		16.2
7	265	224.6	84.8	27,852		
8	280	237	84.6	31,390	577.8	31.8
9	527	512	97.1	42,812	540.8	35.7
10	240	201.9	84.1	25,050	1360	
11	367	320.5	87.3	29,445	1690	49.89
12	372	271	73.0	28,629		
13	430	326.9	76.0	28,798	473.8	19.8
14	-658	381	57.9	25,059	30.5	6.4
15	650	654.9	96	118,501	18024	
16	-	691	-	36,376		
17	410 / 380	285.9	70 / 75	40,929		
18	-320	350	100	31,237		
19	-2,150	1119	-78	68,955		
針葉樹 (裸子植物)						
20	21600	-	7.5	-		
21	20000	20800	100	5,895		
22	19600	12000	61.2	28,354	7.7	
モデル植物 (被子植物)						
23	-	119,186	100	27,516		

表 - 1 に示した通り、他の木本植物のドラフトゲノム情報と本研究成果であるブナ・ドラフトゲノム情報を比較したところ、情報量において遜色はなく、有用なドラフトゲノムが完成したと考えられた。このブナのドラフトゲノムは、ブナ目で初めて構築された本格

的なドラフトゲノム情報である。

ゲノム全体の一塩基置換 (SNPs) の出現頻度は 12.34 SNPs/kbp であり、これまでドラフトゲノムとして研究報告されてきた果樹 (ブドウ、リンゴ、ミカン) と比べて高い SNPs 出現頻度を示し、ブナの野生集団としての特徴を示したものと考えられた。ブナのゲノム解析で得た SNPs は、ブナの遺伝マーカーの開発に貢献できるものと考えられた。

以上から、本ドラフトゲノムは、温帯林を優占するブナ科樹木のゲノミクス解析に有用な情報を提供すると考えられる。

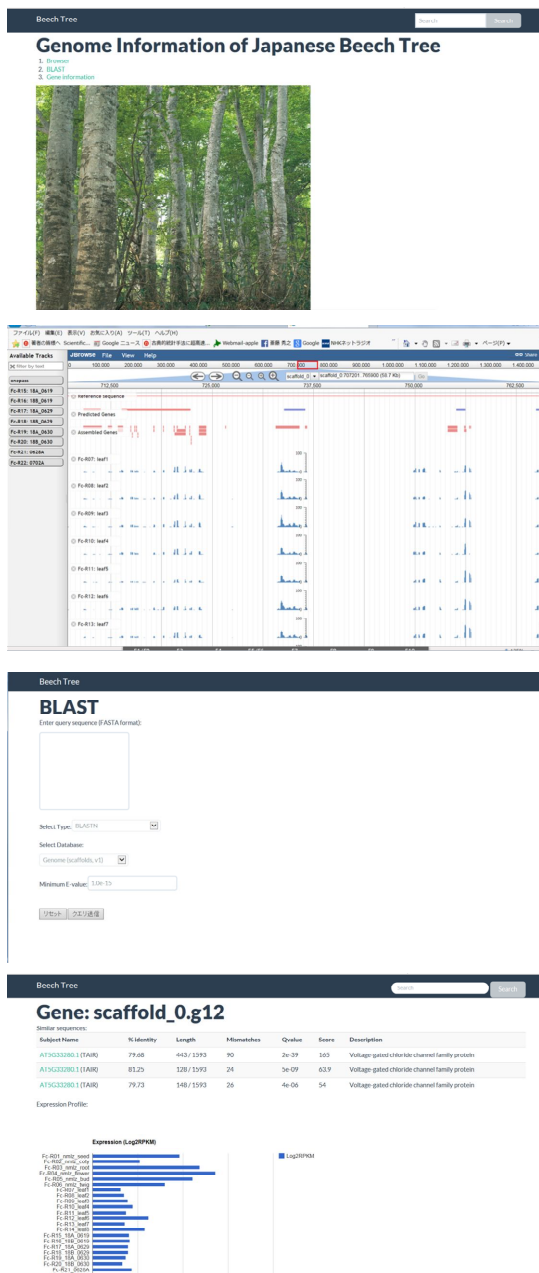


図 - 2 ブナのドラフトゲノム情報データベースの Web 画面

(パネル上からトップ画面, 素データ配列画面, Blast 画面, 予測遺伝子情報画面)

(3) ブナ・ドラフトゲノムの Web 公開用データベース構築

ブナゲノムのデータベースを構築して、Web によりアクセス・活用が可能なホームページ・ツールを作成した (Genome Information of Japanese Beech Tree, <http://fagus.seselab.org/>)。Web では、ゲノムアセンブルされた塩基配列の素データの閲覧、Blastn および Tblastn による Blast 検索の機能、推定された遺伝子の基本情報、シロイヌナズナに対する機能注釈情報およびブナでの発現量データの閲覧機能、キーワード検索、以上の主に 4 つの機能をもつ (図 - 2)。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

(雑誌論文)(計 4 件)

斎藤秀之 (2013) 木本植物のゲノム解読の研究動向. 北海道の林木育種, 56 (1) 18-24. (査読有)

小倉俊治・斎藤秀之 (2012) 針葉樹のフロリゲンをコードする遺伝子の探索に関する研究動向. 北海道の林木育種, 55 (2) 33-36. (査読有)

小倉俊治・斎藤秀之 (2012) 木本植物のフロリゲンをコードする遺伝子の発現調節に関する研究動向. 北海道の林木育種, 55 (1) 23-29. (査読有)

斎藤秀之 (2012) ブナ林のたどっている道 - 分布域北限の動向 -. 北方林業 64(2), 38-41. (査読有)

(学会発表)(計 7 件)

斎藤秀之 (2014) DNA マイクロアレイ vs. RNA-seq ~ 樹木の遺伝子発現解析の経験から ~ .次世代シーケンサー活用術~ 微生物から動物、樹木まで ~ 2014年3月31日, BGI セミナー, 北海道大学農学部, 札幌市
斎藤秀之・神村章子・瀬々潤・齋藤央嗣・谷脇徹・清水(稻継)理恵・清水健太郎 (2014)ゲノム網羅的な発現遺伝子によるブナ葉の環境ストレス評価~ 丹沢ブナ林の事例から ~ 第 125 回日本森林学会大会. 2014年3月27日~30日, 大宮ソニックシティ, 埼玉県 (ポスター発表)

斎藤秀之・瀬々潤・神村章子・山田宰靖・小倉淳・清水(稻継)理恵・清水健太郎 (2014) DNA マイクロアレイによるブナ林の環境影響評価法. 山地森林域の生物・環境モニタリング第 8 回ワークショップ, 2014年3月13日-14日, 広島県総合技術研究所保健環境センター, 広島市

小倉俊治・斎藤秀之 (2013) ブナの花の数と性比に関する FT 遺伝子の量的な発現調節, 第 62 回北方森林学会, 2013年11月13日, 札幌コンベンションセンター, 札幌市 (ポスター発表)

齋藤秀之・神村章子・瀬々 潤・清水(稲
継)理恵・清水健太郎(2013)ブナの葉の
高温ストレス前歴を指標する発現遺伝子
のゲノム網羅的解析.第124回日本森林学
会大会,2013年3月25日~28日,岩手大
学,盛岡市(ポスター発表)

齋藤秀之・小倉俊治・門松昌彦(2012)ブ
ナの花芽分化と花成遺伝子群の発現調節,
第61回北方森林学会,2012年11月12日,
札幌コンベンションセンター,札幌市(ポ
スター発表)

小倉俊治・齋藤秀之・鈴木慎吾・荒川圭
太・松浦英幸・門松昌彦(2012)ブナ葉の
糖濃度と花成遺伝子群の発現,第61回北
方森林学会,2012年11月13日,札幌コ
ンベンションセンター,札幌市(ポスター
発表)

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

Genome Information of Japanese Beech Tree
(<http://fagus.seselab.org/>)

6. 研究組織

(1)研究代表者

齋藤 秀之(SAITO Hideyuki)

北海道大学・大学院農学研究院・講師

研究者番号:70312395

(2)研究分担者

瀬々 潤(SESE Jun)

東京工業大学・情報理工学研究科・准教授

研究者番号:40361539

(3)連携研究者

なし