

2002年3月22日  
独立行政法人 理化学研究所

## シロイヌナズナ完全長 cDNA 約 14,600 種を同定

### - 高等植物において世界で初めて完全長 cDNA 情報を大量公開 -

理化学研究所（小林俊一理事長）は、機能アノテーション（機能注釈）<sup>※1</sup> および制御領域（プロモーター）情報を付与したシロイヌナズナ<sup>※2</sup> の完全長 cDNA <sup>※3</sup> 情報を公開します。高等植物において完全長 cDNA の情報が大量に公開されるのは、今回が世界で初めてです。理研横浜研究所ゲノム科学総合研究センター（和田昭允センター所長）植物ゲノム機能情報研究グループの篠崎一雄プロジェクトディレクター、関原明上級研究員、および遺伝子構造・機能研究グループの林崎良英プロジェクトディレクターらの研究グループによる成果です。

研究グループでは、世界に先駆けて「シロイヌナズナ遺伝子エンサイクロペディア（百科事典）」としてシロイヌナズナ完全長 cDNA（約 14,600 クローン）の解析を行い、機能注釈を付ける作業を進めてきました。本研究の大きな成果として下記の 3 点があげられます。

- 1) タンパク質の機能解析や遺伝子組み換え作物の作成に利用可能な完全長 cDNA を 14,600 種（全遺伝子の約 60%）を同定し、機能注釈をつけました。
- 2) この約 14,600 種の遺伝子のうち、約 2,400 種はこれまでに明らかになっていなかった新しい遺伝子であることがわかりました。
- 3) 遺伝子の発現を制御するために重要な、14,000 種におよぶ大量のシロイヌナズナ遺伝子の制御領域（プロモーター）情報を明らかにしました。

同定された遺伝子、および遺伝子（またはプロモーター）の配列情報は、将来の有用作物の生産性向上、乾燥、低温、塩、強光などの環境ストレスに耐性をもつ有用作物の作成、耐病性、耐害虫性有用作物の作成、食品工学への応用、新薬の創製、有用物質の生産、さらには環境保全に役立つ植物の育成などへの応用が期待されます。

植物ゲノム機能情報研究グループでは、本研究で得られたシロイヌナズナ完全長 cDNA を用いて国内外の研究者との植物遺伝子の機能解析のための共同研究を進めています。本研究成果は、米国の科学雑誌『Science』のインターネット版（3月22日号）で公開されるとともに4月5日号に掲載されます。また、cDNA情報は、同研究グループのホームページ（<http://www.gsc.riken.go.jp/Plant/index.html>）でも公開されます。

### 1. 背景

植物は、食料、環境、産業など私たちが生活を営むにはなくてはならない存在です。植物のさまざまな生理機能のことを詳しく知ることは、より優れた特徴を持った有用作物を作るうえで極めて重要です。2000年12月には日本の研究機関を含む国際研究チームによって、高等モデル植物の一つ「シロイヌナズナ」の全ゲノムが解析され、植物の機能解析が飛躍的に進歩しました。シロイヌナズナは、“イネ、

コムギ、トウモロコシ、大豆など主要穀物の遺伝子と基本的に類似している”、“植物の基本的ライフサイクルのすべてのステージ、または種々の環境条件で発現している遺伝子の機能が解析できる”といった特徴を持ち、さらに“実験室で扱いやすいこと”から研究材料として多くの研究者に利用されています。

理研ゲノム科学総合研究センターの植物ゲノム機能情報研究グループを中心とした研究グループでは、シロイヌナズナ遺伝子（完全長 cDNA）のすべてを取り出し、その塩基配列情報などを体系的に整理した辞書となる「シロイヌナズナ遺伝子エンサイクロペディア(百科事典)」の作成を目指して、機能情報のアノテーション（注釈付け）を行っています。さらに、シロイヌナズナの種々の変異体を多数作成し、個体レベルでの遺伝子の機能解析を進めています。シロイヌナズナのゲノム上にある情報を網羅的に解析を行い、およそ 26,000 ともいわれる遺伝子の機能を明らかにし、植物が持っているさまざまな能力を理解することで、生物学、農学といったさまざまな分野での貢献が期待されます。

## 2. 研究成果

研究グループでは、種々の発生段階の植物体、種々の植物組織、乾燥、低温などの種々の環境ストレス、植物ホルモンおよび光処理したシロイヌナズナ植物体を出発材料として 19 個の完全長 cDNA ライブラリーを作成しました。この完全長 cDNA ライブラリーより、これまでに約 14,600 種のシロイヌナズナ完全長 cDNA を単離しました。この約 14,600 種の完全長 cDNA について 5'末端および 3'末端の端読み塩基配列情報の解読に加え、“それらが既知の遺伝子なのか、新規の遺伝子なのか”、“既知の遺伝子とはどのような類似性を有するのか”、“どのような性質の遺伝子なのか”などの機能情報を注釈付け（機能アノテーション）を行いました。その結果、約 14,600 種の遺伝子のうち、約 2,400 種の遺伝子は新規の遺伝子でした。これらの中には環境ストレス耐性、生産性などに関連する有用な新規遺伝子が存在すると思われれます。

また、約 14,000 種のシロイヌナズナ遺伝子について 5'末端の端読み配列データをゲノム上にマッピングさせることにより、遺伝子の発現制御に働くプロモーター配列情報を獲得し、プロモーター・データベース (DB) を作成しました。作製したプロモーター・DB は、マイクロレイ解析結果と組み合わせることにより、転写因子が結合する新規なシス制御配列の同定や、既知の制御遺伝子の下流で機能する機能遺伝子群の同定などへの利用が期待されます。今回、単離したシロイヌナズナ完全長 cDNA の 5'末端および 3'末端の端読み塩基配列情報およびプロモーター情報などの公開により、すでに公開されているシロイヌナズナゲノム配列情報と組み合わせることで、高等植物遺伝子の機能解明などの研究が飛躍的に促進することが期待されます。

なお、単離した約 14,600 種のシロイヌナズナ完全長 cDNA の 5'末端および 3'末端の端読み塩基配列データは、日本 DNA データバンク (DDBJ) や GenBank などのデータベースに登録し 3 月 22 日にはアクセスが可能になる予定です。

## 3. 研究成果の意義

植物科学においてもシロイヌナズナ、イネなど膨大なゲノム情報（DNA 塩基配列）が急激に解読されつつあるなか、それらの情報をもとに、革新的な有用作物、

食品、医薬品、農薬などの開発につなげるための、ゲノム機能研究の推進が重要となっています。近年、有用作物、食品、医薬品開発に直結するゲノム機能研究の主要な課題として、遺伝子から合成されるタンパク質（生体反応を担う主役）に関する機能・構造研究が重要視されており、激しい国際競争が展開されつつあります。本研究の対象である完全長 cDNA（遺伝子の本体）は、このタンパク質を合成するための設計情報をすべて有し、タンパク質そのものを合成することができることから、ゲノム機能研究における非常に重要な「基盤ツール」となります。

本研究で約 14,600 種という多量のシロイヌナズナ完全長 cDNA について、5'末端および 3'末端の端読み配列情報、機能情報だけでなく、プロモーター情報をも付加することは、「基盤ツール」としての価値を飛躍的に高めるとともに、今後のゲノム機能研究の効率的推進に大きく貢献するものと期待されています。

また、完全長 cDNA のほかのメリットとして、遺伝子組換え作物の作成(特に有用遺伝子の過剰発現植物を作成する場合)に用いる導入遺伝子としても完全長 cDNA であれば、開始コドンから終止コドンまでのタンパク質をコードする領域をカバーしているため、その遺伝子の機能を正確に個体で発現することができます。これに対しゲノム配列から予想される遺伝子は、あくまでも予想にしかすぎず、完全長 cDNA から決められた遺伝子と異なる場合がしばしば見られます。

#### 4. 今後の展開

研究グループでは、これまでに単離した約 14,600 種のシロイヌナズナ完全長 cDNA を含む cDNA マイクロアレイの作製を進めています。それらマイクロアレイを作成後、国内外の植物研究者との共同研究によりシロイヌナズナの発現遺伝子の網羅的な発現情報の基本データベースを作成し、新規なシス制御配列を同定したり、既知の制御遺伝子の下流で機能する機能遺伝子を同定するなどして、高等植物における遺伝子発現制御ネットワークを解明していく予定です。また、単離されたシロイヌナズナ完全長 cDNA を過剰発現させたり、遺伝子発現を抑制させたトランスジェニック植物を作成するなどして遺伝子の機能解析を進めて行く予定です。

植物ゲノム機能情報研究グループでは、単離したシロイヌナズナ完全長 cDNA を活用した遺伝子機能解析等に関する共同研究に関して国内外の研究機関と積極的に推進していきます。さらに応用面でも、植物の乾燥や低温など環境ストレス耐性に関わる遺伝子、植物の光合成能の向上や生産性の向上に関わる遺伝子、脂質や糖質など種々の物質生産に関係する遺伝子、植物の生長制御に関係する遺伝子など有用な遺伝子の探索に役立てていきたいと考えています。

(問い合わせ先)

独立行政法人理化学研究所 横浜研究所 ゲノム科学総合研究センター  
植物ゲノム機能情報研究グループ 植物変異開発研究チーム

プロジェクトディレクター 篠崎 一雄

Tel : 0298-36-4359 / Fax : 0298-36-9060

上級研究員 関 原明

Tel : 0298-36-4359 / Fax : 0298-36-9060

横浜研究所 研究推進部

反町 耕記

Tel : 045-503-9121 / Fax : 045-503-9113

(報道担当)

独立行政法人理化学研究所 広報室

嶋田 庸嗣

Tel : 048-467-9272 / Fax : 048-462-4715

## <補足説明>

### ※1 機能アノテーション

機能アノテーションとは、解析対象である遺伝子について機能の注釈付けを行うこと。例えば、既知遺伝子塩基配列情報とのホモロジー（類似性）解析、遺伝子産物であるタンパク質のモチーフ（2次構造あるいはそれらの組み合わせの構造）解析、遺伝子の属性分類等の方法が用いられる。

### ※2 ロイヌナズナ

学名は *Arabidopsis thaliana* L. Heynh. でアラビドプシスともいう。全長約 30～40 cm のアブラナ科 1 年生草本植物。食用でも観賞用でもない、いわゆる雑草である。北半球のほぼ全域の冷温帯にかけて広く分布している。自家和合性を持ち、基本的に自家受粉で次世代の種子をつくるが、人工交配による他家受粉も可能である。通常、秋に発芽して冬を越した後、春から秋にかけて日が長くなると花が咲く長日植物である。実験室内では約 2 か月で次世代の種子をつけることが可能である。二倍体で、5 対（合計 10 本）の染色体をもつ。日米欧の 6 グループ（日本はかずさ DNA 研究所の田畑哲之のグループ）が進めたゲノムプロジェクトにより、2000 年 12 月に高い精度で全ゲノム配列が決定され、ゲノムサイズは約 1 億 2,500 万 DNA 塩基対、全遺伝子数は約 26,000 個存在することが予想された。

シロイヌナズナはモデル高等植物として植物の遺伝子の機能解析に幅広く用いられている。それは、以下に示すように研究材料として非常に優れているという特徴だけでなく、シロイヌナズナを用いた研究環境が国際的に非常にオープンで種々の研究材料や情報が利用できる点にあると考えられる。シロイヌナズナはモデル高等植物としての特色として、1)、植物体が小さく、生育が容易で、生活環が 2 か月と短いなど遺伝学的研究に利用しやすい、2)、ゲノムサイズが高等植物で最小である、3)、国際的共同研究体制が確立しており、ゲノム、遺伝子、変異体材料などのリソースセンターが存在する一などがあげられる。

### ※3 完全長 cDNA

cDNA とは、ゲノム DNA の中から不要な配列を除き、タンパク質をコードする配列のみに整理された遺伝情報物質である mRNA（メッセンジャー RNA）を鋳型にして作られた DNA のこと。完全長 cDNA は、断片 cDNA と異なり、タンパク質を合成するための設計情報をすべて有しているため、タンパク質を合成することができる。この完全長 cDNA を効率的に合成するためには、非常に高い技術を必要とし、わが国が世界に先んじている。

#### ※4 プロモーター

遺伝子の上流に存在する DNA 領域で、遺伝子がいつ、どのような状況下で、どの程度の量を発現させるのかが決定される。

#### ※5 シス制御配列

プロモーター領域には、多数のシス制御配列が存在しており、それを認識する多様な転写因子の作用によって遺伝子の転写活性が制御されている。シス制御配列の中には、細胞内外のシグナルに応答した転写調節に関与するものなどがある。

## モデル植物としてのシロイヌナズナの特徴



世代時間が短い (6-8 週間)

植物体が小さく、実験室で多くの植物を生育させられる



ゲノムサイズが植物の中で最小

全塩基配列が明らかとなっており、全 26,000 個の遺伝子情報を得ることができる



モデル植物として解析が進んでおり植物の中で最も解析が進んでいる

様々な植物種が持つ遺伝子をほとんど持ち合わせているため様々な作物や樹木に応用することが出来る

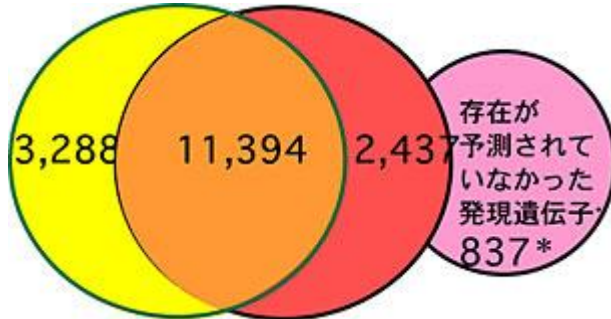


## シロイヌナズナ研究で報告された遺伝子

単離された発現遺伝子: 17,956 個

全ゲノム塩基配列決定後

予測された遺伝子: 26,285 個



既に報告済みの遺伝子: 14,682 個

研究グループが単離した RAFL cDNA 遺伝子:

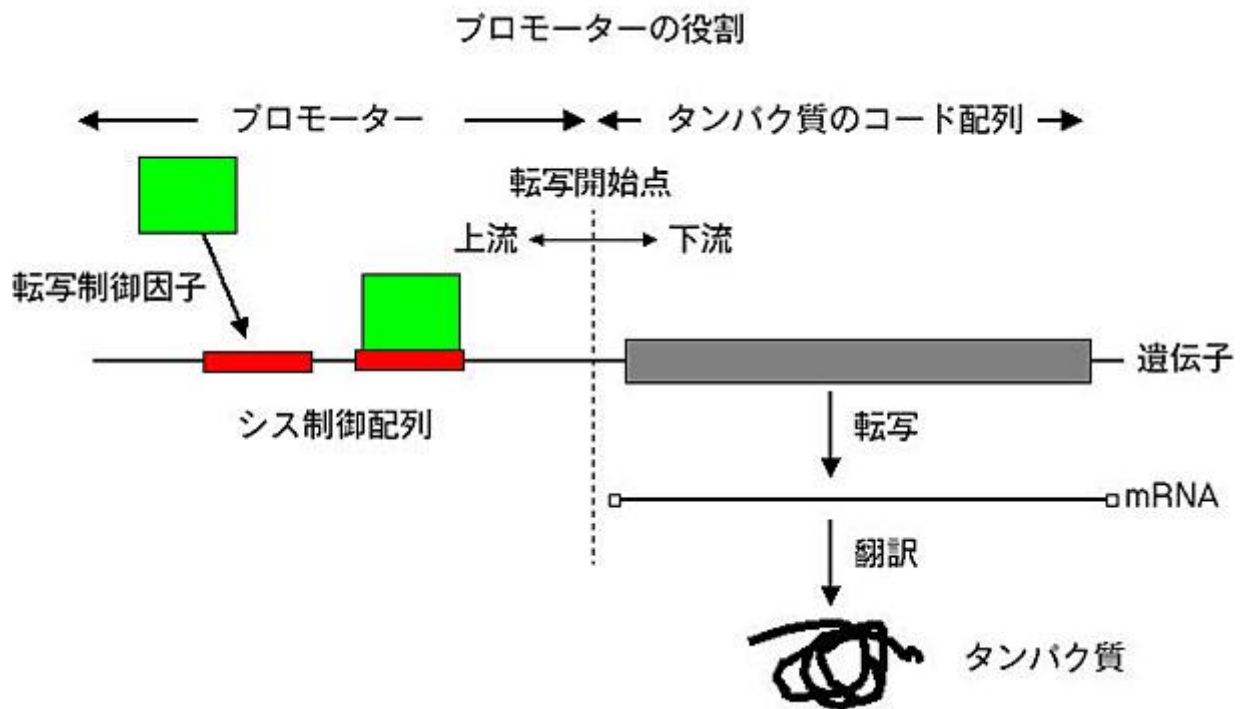
14,668 個 (全発現遺伝子の 80% をカバー、新規な 2437 個を含む)

※RAFL: RIKEN Arabidopsis Full Length

\* 今までのゲノム配列をもとにした遺伝子解析プログラムでは予測されていなかった発現遺伝子

## 研究グループが単離した cDNA 遺伝子の機能分類

機能分類	RAFL cDNA遺伝子の数
代謝	521
エネルギー生産	98
細胞成長、細胞分裂、DNA合成	54
転写	331
タンパク質合成	145
タンパク質輸送、選別	320
膜輸送	89
細胞構築	107
シグナル応答	262
傷害応答	172
イオンホメオスタシス	4
細胞骨格	255
発生、分化	40
機能未知	12270
プロテインキナーゼ	506
転写因子	1087



遺伝子のタンパク質をコードする配列の上流に存在する DNA 領域で、遺伝子がいつ、どのような状況下で、どの程度の量を発現させるのか、の決定に関与する領域をプロ



モーターという。一般に、プロモーター領域にあるシス制御配列（シス因子）と呼ばれる発現調節に関わる部位には、転写制御因子タンパク質（トランス因子）が結合して遺伝子の転写発現の調節を行う。

