

平成22年 6月21日現在

研究種目：若手研究（スタートアップ）
 研究期間：2008～2009
 課題番号：20880041
 研究課題名（和文） イネにおいて遺伝子発現制御に関わるシス配列情報整備のための基盤研究
 研究課題名（英文） Study for developing information of cis-elements related to regulation of gene expression in rice
 研究代表者 佐藤 豊（SATO YUTAKA）
 独立行政法人 農業生物資源研究所・ゲノムリソースセンター・任期付研究員

研究者番号：905106944

研究成果の概要（和文）：

遺伝子発現制御に関わるシス配列探索ツール作成を目的とした基盤情報を整備するため、イネで予測された遺伝子の転写開始点より上流 1kb と 2kb の配列情報を抽出し、さらに同じ発現パターンを示す遺伝子群の情報を大規模な共発現解析によって得た。また、実験によって転写因子の結合配列を効率良く特定できる方法を検討するため Protein Binding Microarray 用のプローブ配列を設計し、カスタムアレイを作成した。

研究成果の概要（英文）：

In order to construct an analysis tool for identifying cis-elements related to regulation of gene expression, we extracted the sequences of putative promoter regions of all annotated genes in rice, which correspond to the 1 kb and 2 kb from the start point of transcription for each gene, and performed a wide range of co-expression analyses. Furthermore, we designed probes for protein binding microarray in order to extract the sequence, with which a transcription factor directly interact.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,340,000	402,000	1,742,000
2009年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	2,540,000	762,000	3,302,000

研究分野：農学

科研費の分科・細目：育種学

キーワード：遺伝子発現、シス配列

1. 研究開始当初の背景

イネの全塩基配列が 2004 年に決定され、

イネゲノム上には約 32,000 の遺伝子情報が Rice Annotation Database (RAP-DB) によって公開されていたが、マッピングされた遺伝

子の中には機能未知なものが多数存在し、それらの機能解析を進めることは、植物の持つ潜在能力をさらに引き出すために重要な課題であった。また、遺伝子機能解明のために様々な順遺伝学的、逆遺伝学的アプローチが確立されていたが、それを強力にサポートする網羅的遺伝子発現情報の整備がシロイヌナズナと比較してイネでは遅れていた。そのため、研究開始当初大規模な遺伝子発現情報の整備を目的とし、その情報収集が開始された。これらの情報は単に遺伝子がいつ、どこで発現するかということを知るだけでなく、共発現解析等のデータマイニングによって、遺伝子の発現制御に関わるシス配列探索に極めて有用な情報であると考えられた。

2. 研究の目的

インフォマティクスに精通していない実験研究者にとって大量のデータ解析を行い、目的とする必要な情報を得ることは非常に難しい。そのため、イネの大規模マイクロアレイ解析によって蓄積する遺伝子発現情報を用いて、シス配列探索が行えるような汎用性の高い基盤整備を行う必要がある。また、インフォマティクスによって得られた情報は極めて有用であると考えられる一方で、間違った情報も多く含まれると予想される。そこで、本研究ではバイオインフォマティクスの手法によるシス配列探索ツール作成のための基盤研究を行うとともに、実験によって遺伝子発現制御配列を特定できるような方法も探索する。

3. 研究の方法

マイクロアレイ解析等の網羅的遺伝子発現解析によって得られた遺伝子発現情報から、同じ発現パターンを示す遺伝子群を選抜しそれらのプロモーター領域の配列比較から保存された配列（推定シス因子）を見つけることが可能である。そこで、イネで生産された大量のマイクロアレイデータを用いて、様々な実験条件に依存した、あるいは依存しない共発現解析を実施した。また転写因子が結合する配列を **high-throughput** にスクリーニングできるようなシステムを検討した。

4. 研究成果

(1) イネ品種「日本晴」で予測された遺伝子のうち転写開始点より上流 1kb と 2kb の配列を抽出可能なもの(29,389 遺伝子)に関して

は、その情報を抽出し目的の配列を選択して利用することができるようにした。

(2) (1)により得られた 29,389 遺伝子の上流 1kb 配列すべてについて、PLACE というデータベースで公開されている 469 件の既知のシス配列の存在頻度をカウントした。その結果、解析したすべての遺伝子でモチーフが見られないものが 86 件あり、それ以外では 1~209288 まで幅広く観察された(図 1)。これらの情報は、共発現遺伝子群で有意に蓄積されている既知シス配列情報を得るのに有用である。

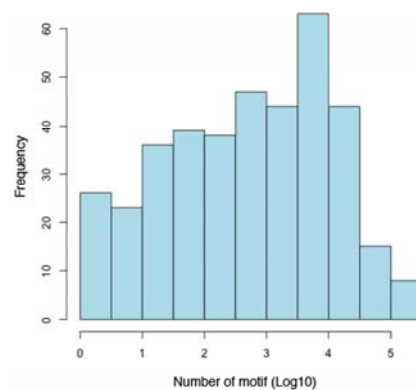


図1. イネの29,389遺伝子の上流1kbに存在する既知のシス配列の頻度分布

(3) イネにおける大規模遺伝子発現プロファイルから栄養生長期から生殖生長期への移行前に特異的に発現変動する遺伝子群を共発現解析(階層型クラスター解析)によって抽出した(実験条件に依存した共発現解析)。(2)で得られた情報を用いて抽出された遺伝子群の上流 1kb 領域に有意に蓄積している既知のシス配列を探索した。その結果、この遺伝子群には **GNATATNC** というモチーフが有意に蓄積していることが明らかになった(χ^2 test, $p=0.0003$)。これは **PHR1** という転写因子が結合する配列として単離されたもので、リン代謝制御に重要であるシス配列である。つまり、抽出した遺伝子群の中にはリン代謝に関わる遺伝子が多く含まれていると考えられた。さらに、(2)のような基盤情報整備はインフォマティクスにあまり精通していない実験研究者にとって有用なツールであると考えられた。

(4) イネにおける大規模遺伝子発現プロファイリングによって得られたマイクロアレイデータを使って実験条件に依存しない共発現解析を実施した。具体的な方法は、マイクロアレイデータ毎に中央値と MAD を求め、中央値が 0、MAD が 1 になるように正規化し、その後すべての遺伝子間でピアソンの積率相関係数を算出した。さらに、相関係数の順位をもとに Mutual Rank (MR) を計算し、相関係数とともに共発現関係の強さを示す指標として用いた。

① 得られた共発現データの信頼性を確認するために、既知のジテルペン型ファイトアレキシンであるモミラクトンやファイトカサン合成に関与しおり、第 2 染色体と第 4 染色体で gene cluster を形成していることが知られている *OsKSL4*, *OsKSL5*, *OsKSL6*, *OsKSL7*, *OsCPS2*, *OsCPS4* と共発現する遺伝子群をそれぞれ抽出した (MR < 300)。さらに視覚的にその情報を整理するためにネットワーク図を作成した (図 2)。その結果、ネットワーク図では大きく二つのクラスターが確認でき、一つは *OsKSL5*, *OsKSL6* を含むクラスター、もう一つが *OsKSL4*, *OsKSL7*, *OsCPS2*, *OsCPS4* を含むクラスターである。興味深いことに、前者はキチンエリシターで誘導されないもの、後者はキチンエリシターで誘導されるという報告がなされているものである。このことは、キチンエリシター処理条件によるマイクロアレイデータを含んでいないデータセットによる共発現解析でもその発現制御系の違いを識別できる可能性があることが示唆される。さらに、*OsCPS4* や *OsKSL4* の発現を制御することが報告されている転写因子 *OsTGAP1* が同じクラスターに含まれており、またこの転写因子以外にも、bHLH, NAC, Myb, WRKY などの転写因子の存在が確認された。これらの結果を考慮すると、シス配列や転写因子などのトランス因子を特定するために共発現情報は非常に有用性が高いことが示唆される。またここに示したジテルペン型ファイトアレキシン合成系はシロイヌナズナには存在しないため、先行しているシロイヌナズナの共発現解析結果は利用することができない。またイネは農業上重要な作物であることから、イネによる情報基盤の整備は重要であると考えられ、今後その需要は高まると予想される。

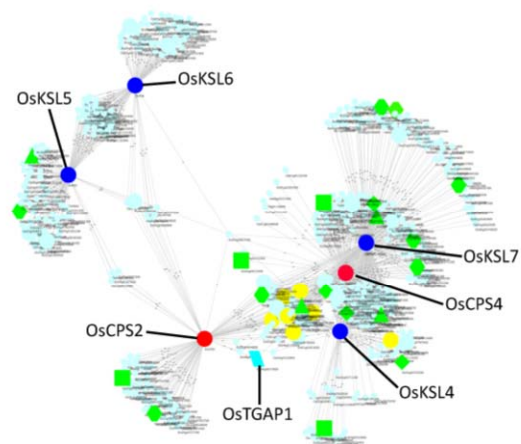


図2. ジテルペン型ファイトアレキシン合成系の遺伝子間ネットワーク

■ bHLH ▲ NAC ● Myb ◆ WRKY

② 視覚的に表現型を観察できる形質で、かつ関係する遺伝子が報告されているクロロフィル代謝系に着目して in silico の解析を実施した。まず、クロロフィル分解のポジティブレギュレーターであると考えられている *SGR* と共発現する TOP300 の遺伝子情報を抽出した。しかし、その中には直接クロロフィル代謝に関係のないものも多く含まれており、ノイズを除く必要があると考えられた。クロロフィル代謝系は植物の持つ基本的な代謝系であることからその制御機構は高度に保存されている可能性があることが予想された。そこで、シロイヌナズナの共発現データベースである ATTED-II から同様に *AtSGR* と共発現する TOP300 の遺伝子の情報、さらにそれらの遺伝子のイネにおけるホモログ情報を得て、イネの共発現解析結果と積集合の関係にある 26 の遺伝子を抽出してきた。その中には、*SGR* と同じくクロロフィルの代謝に関係する酵素をコードする *NYC3*, *OsPAO* が含まれていたことから、この情報の信頼性は高いと考えられた。26 の遺伝子の中に二つの 2 つの転写因子 (NAC, NF-YA) が含まれており、そのうち、NAC 転写因子の中には老化を制御していると報告のあるものが存在するため、特定された NAC 転写因子がクロロフィル代謝系の一部を制御している可能性が考えられた。この転写因子については今後過剰発現体の表現型を調べる予定である。

(5) 転写因子と結合する配列を high-throughput に解析できる Protein binding microarray (PBM) の検討を行った。PBM はマイクロアレイスライド上の 1 本鎖

DNA を 2 本鎖にするためのプライマー結合配列(不変領域: 26 mer)、転写因子結合配列を特定するための可変領域 (34 mer)の二つの領域が必要である。前段階として 44,000 のプローブにおいて可変領域の配列を決定した。決定した配列間で完全一致がないことを、構築した配列に対して BLASTN 検索を行って確認し、可変領域と不変領域を組み合わせてプローブとしカスタムアレイを作成した。現在実験系の確立はまだ十分ではないが、今後実験条件の検討を行いPBMのシステムを構築することで、上述のバイオインフォマティクス的手法と組み合わせて、より信頼性の高いシス配列探索が行えるようになると考えている。

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 0 件)

[学会発表] (計 0 件)

6. 研究組織

(1) 研究代表者

佐藤 豊 (SATO YUTAKA)

(独)農業生物資源研究所・ゲノムリソースセンター・任期付研究員

研究者番号: 905106944

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: