

ゲノムネットワークプロジェクトの  
課題の報告

プログラム名		(1) <b>ゲノム機能情報の解析</b> (2) ヒトゲノムネットワークプラットフォームの構築 (3) 次世代ゲノム解析技術の開発 (4) 個別生命機能の解析 (5) 動的ネットワーク解析技術開発					
課題の分類		(1) 中核機関 (2) <b>指定課題</b> (3) 公募課題					
課題名 (和英併記)		ヒト全遺伝子レトロウイルス型siRNAライブラリの構築 (Development of Retrovirus-Type siRNA Library for Human Genome)					
代表研究者名 (所属機関・職名)		秋山 徹 (国立大学法人東京大学・教授)					
年度別研究費 (千円)		16年度	17年度	18年度	19年度	20年度	合計
		170,000	177,500	69,690	68,000	60,000	545,190
		研究者名	所属機関・部局・職			専門分野	
研究 組 織 図	代表機関	☆ 秋山 徹	東京大学・分子細胞生物学研究所・教授			分子情報学	
		程 久美子	東京大学・大学院理学系研究科・准教授			ゲノム情報生物学	
	分担機関	☆ (代表研究者)					
		(分担研究者)					
<p><b>【研究目的】</b></p> <p>ヒト全遺伝子に相当する約2万種類の遺伝子に対するsiRNAをコードするプラスミド(あるいは産生大腸菌)を完成させ、ゲノムネットワークコンソーシアム・メンバーに配布し、ゲノムネットワークプロジェクトによる遺伝子発現ネットワークの解明を目標とする研究に資する事を目的とする。</p>							
<p><b>【課題の概要】</b></p> <p>(1) ヒト遺伝子レトロウイルス型siRNA発現ライブラリの構築 ヒト遺伝子配列依存的に遺伝子機能破壊を惹起するsiRNAをコードし、かつレトロウイルス粒子に取り込み可能なRNAを発現するDNAプラスミドを、約20,000種のヒト遺伝子に対して構築し、遺伝子のシステムティックな解析に資する。</p> <p>(2) siRNAを生産するレトロウイルスの調製とライブラリの評価 上記(1)で作製したライブラリの検定、評価を行う。</p>							

## 【平成20年8月末までの研究成果】

### (1) ヒト遺伝子レトロウイルス型及びレンチウイルス型 siRNA 発現ライブラリの構築 (1-1) レトロウイルス型ライブラリ (pSUPERretro)

siRNA 発現コンストラクト構築に際しては、研究開始当初に有用性が明らかで、最も広く使用されていた shRNA をコードする方法 (shRNA 法) を採用し、pSUPERretro ベクターを用いた。ヒト遺伝子のうち、明確に配列が特定できていた、およそ 20,000 のヒト個別遺伝子のうち siRNA 発現クローンを容易に作製できると考えられたものは、検討の結果、約 14,000 遺伝子であった。そこでまず約 14,000 遺伝子を Gene Ontology に従って分類し、転写制御因子を中心に、優先度の高いものから順次 siRNA 発現クローンを作製していった。その結果、平成 17 年度末までには、予定を上回る約 15,000 遺伝子に対するレトロウイルス型 (pSUPERretro) のコンストラクションを完成することができた。平成 17 年度までに構築したライブラリには、配列設計の時点で、転写 variants が複数あり、それらすべてを同時につぶすことが困難である遺伝子に対する siRNA 発現コンストラクト、及び使用していた発現ベクターの性質から、有効と考えられる siRNA を発現できるコンストラクトを作製することが不可能であったものは含まれていなかった。しかし、RNAi 法はヒト遺伝子機能の解析を飛躍的に容易にするので、完成度の高いライブラリの構築は、コンソーシアムメンバーの研究に大きく貢献すると考えられたため、残りのヒト約 5,000 遺伝子に対するコンストラクションの追加作製が、平成 18 年度途中より開始された。Variants が複数あるためにコンストラクトを作製することができなかった遺伝子群については、すべての variants はカバーできなくても、できるだけ多くの variants をつぶすことが可能な siRNA 発現コンストラクトを作製し、平成 19 年度末までには、ヒトのほぼ全遺伝子をカバーする約 19,000 遺伝子に対する siRNA ライブラリが完成した。これらの全てのコンストラクトは、シーケンズ解析を行い、間違いがないことを確認した。作製したコンストラクトは、そのままプラスミドとしてトランスフェクションして使用することもできるし、レトロウイルスとして使用することもできる。レトロウイルスとして使用する場合には、作製コンストラクトをパッケージング細胞 (レトロウイルス産生細胞) にトランスフェクションして培養上清に産生されるレトロウイルスを目的の細胞に感染させればよい。この場合、目的の遺伝子の発現を持続的に抑制することができる。

### (1-2) レンチウイルス型ライブラリ (pENTR. H1)

プロジェクト推進中に、ES 細胞や神経細胞にも適用可能なレンチウイルス型ベクターの改良が進み、良好な RNAi 効果が得られることが明らかになってきた。そこで当初の予定にはなかったが pSUPERretro に加えてレンチウイルス型コンストラクト (pENTR. H1) の作製を行い、平成 18 年度までに、ヒト約 2,100 遺伝子 (主に転写関連遺伝子) に対する shRNA 発現コンストラクトを完成した。平成 20 年度には、主要な遺伝子を中心として、ヒト約 3,000 遺伝子に対するレンチウイルス型ライブラリの追加構築を行っている。また、RNAi 法では、標的とする遺伝子以外の遺伝子も抑制してしまう “off-target 効果” と呼ばれる副作用的効果が起こり、特定の遺伝子機能解析の妨げとなる場合がある。これまではコンピュータシオナルな解析により、できる限り標的以外の遺伝子とは相補性が少なく、ミスマッチの多い siRNA 配列を用いることによって off-target 効果を最少とする方法を用いてきた。しかし、この方法では、ミスマッチの場所によっては必ずしも off-target 効果を完全に避けることはできない。我々は off-target 効果が起こる機構を解析し、off-target 効果がない、あるいは減弱できる siRNA 配列を選択できる手法を開発した。この方法を用いたレンチウイルス型コンストラクトも追加構築中のコンストラクトに含まれている。

### (1-3) polIII promoter 型ライブラリ (pENTR. CMV)

前述のレトロウイルス型及びレンチウイルス型 siRNA 発現コンストラクトは、RNA polymerase III プロモータによって shRNA が転写される。この場合、プロモータの性質から、転写できる shRNA には配列上の制限があつて、発現コンストラクトを作製することができないものがあつた。そのため、プロジェクト進行中にどのような配列でも使用できる RNA polymerase II promoter によってドライブされる shRNA 発現ベクターを開発した。さらに、効く siRNA の配列規則性を検討したことによって、利用できる siRNA の配列の制限が大幅に緩和され、99.5%のヒト遺伝子に対して、有効と予想される siRNA の配列設計が可能となった。平成 18 年度までに約 2,200 遺伝子（主に転写関連遺伝子、ミトコンドリア関連遺伝子）に対するコンストラクトを完成した。約 1,100 遺伝子についてはヒト/マウス共通配列を用いた。さらに、平成 19 年度までに、転写調節因子について約 900 の追加構築合成を完了した。

### (2) siRNA を生産するレトロウイルスの調製とライブラリの評価

完成した 約 19,000 遺伝子に対するレトロウイルス型コンストラクトのうちから転写関連遺伝子約 462 を選択し、プラスミドとして HeLa 細胞及び RKO 細胞にトランスフェクションして実際に RNAi 効果を示すかどうかを検討した。その結果、約 75%の遺伝子について RNAi 効果を確認することができた。したがって、作製されたレトロウイルス型コンストラクトは、実際の研究に使用可能な、レベルの高い品質を保持していると考えられた。また、レトロウイルスとして使用した場合にも、同様に標的遺伝子の発現を抑制できることが明らかになった。さらに、レンチウイルス型コンストラクトによりレンチウイルスを産生して RNAi 効果を見た場合にも、同様に標的遺伝子の発現を抑制できることが明らかになった。pol II promoter 型についても、同様に遺伝子発現抑制効果があることが確認できた。さらに、癌化などの情報伝達経路における解析のツールとしての有効性も検討している。

## 【GNPとの関連・GNPへの貢献】

### a) 転写・制御の解明に関連する研究成果

ヒト全遺伝子レトロウイルス型 siRNA ライブラリを構築して、共有リソースとしてプロジェクトの参加者に配布し、転写・制御を中心とする遺伝子機能のシステムティックな解析に資することを目的として研究を進めてきた。主要な転写因子については、1 遺伝子に対して 2 種類のレトロウイルス型 shRNA 発現コンストラクトを作製した。さらに、pol II promoter 型も含めると、3 種類の shRNA 発現コンストラクトを利用できる転写因子群もある。

RNAi 法は、ヒト遺伝子の機能解析を飛躍的に容易にするので、ヒト遺伝子解析を行うコンソーシアムメンバーの研究にも大きな貢献をするものと推測される。特に、転写制御因子のような遺伝子カスケードに関わる因子の解析では、“二重変異”、“三重変異”の解析が必須であり、このような研究には欠かせない道具と考えられるため、研究全体が大きく前進すると期待される。転写調節因子を中心として、生命活動を成立させているネットワークを明らかにし、疾患の発症機構の解明や新しい治療法の開発につながる成果を挙げることを目的とするゲノムネットワークプロジェクトの遂行に、大きく貢献できる重要プロジェクトであると言える。

## b) 縦軸研究（または横軸研究）への貢献に繋がる研究成果

### ・ shRNAライブラリソースの知的財産権に関する整理及び他機関への配布

本プロジェクトでは、いくつかの他のグループの知的財産に関わる成果が利用されているので、その取り扱いには配慮が必要と考えられる。少なくとも、使用している2つの siRNA 発現ベクターについては、知的財産に関する問題があり、現行では以下のように整理されている。

(i) pSUPER. retro. puro : CRT (Cancer Research Technology) 社がライセンスを保有しているレトロウイルス型ベクターであるが、東京大学との契約により、GNP の同意書・依頼書に従って、プロジェクト内での配布が行われている。

(ii) pENTR. U6 : Invitrogen 社のレンチウイルス型ベクターで、ゲートウェイシステムが利用できる。我々は、このベクターの siRNA 発現領域を、H1 プロモータ、及び CMV プロモータ/ポリ A などに改変している。Invitrogen 社と東京大学との契約により、GNP の同意書・依頼書に従って、プロジェクト内での配布が行われている。

プロジェクト終了後は、コンソーシアム内の各研究機関の研究者と東京大学との MTA の定めにより、すでに配布した siRNA ライブラリは処分（または権利者から直接買取）する必要があり、このまま継続して使用することはできない。プロジェクト終了後も siRNA ライブラリを維持し、広く国内の研究者も利用可能とするためには、権利者と東京大学（あるいは企業を含めた委託機関）が新たな契約を締結し、配布体制を維持する必要がある。そのためには、別途経費を確保する必要があるが、現在のところ、目処は立っていない状況である。

平成 17 年度までに構築した、約 15,000 遺伝子に対する siRNA ライブラリについては、平成 18 年度からコンソーシアムメンバーへの配布を開始した。残りの約 5,000 遺伝子については平成 19 年度から配布を開始した。これまでに、レトロウイルス型ライブラリは、延べ 16 研究機関に約 101,000 クローンの配布を行った。また、レンチウイルス型及び我々が開発した RNA polymerase II promoter を使用したレンチウイルス型ベクターを用いたライブラリの配布も行っている。

### ・ プラットフォームへのデータの開示状況

情報公開は、現在はゲノムネットワークプラットフォームから東京大学へリンクする形で行っている。構築した shRNA クローンの標的遺伝子情報及び siRNA 配列が全て掲載されており、結果のフィードバックも可能となっている。また、ライブラリ配布のリクエストも、コンソーシアムメンバーが各自の ID とパスワードを利用して、当サイトにアクセスすることによって行っている。

【当該研究に関して活用した研究データ・リソース】

該当なし。

【主な研究論文とその概要】

1. Naito Y, Saigo K, Ui-Tei K. (2007) Evaluation of published rational siRNA design algorithms using firefly luciferase gene as a reporter. RNA interference research progress.
2. Ui-Tei K., Naito Y, Saigo K. (2007) Guidelines for the selection of effective short-interfering RNA sequences for functional genomics. Methods Mol. Biol. 361, 201-216.
3. Ui-Tei K., Naito Y, Saigo K. (2006) Essential notes regarding the design of functional siRNAs for efficient mammalian RNAi. J. Biomed. Biotechnol. 2006, 65052.

上記の論文は、有効なsiRNA配列選択の規則性及び他のsiRNA配列設計アルゴリズムとの比較、さらにはRNAi効果の評価法についても記載している。Naito et al. (2008)の論文では、現在、主として使用されている3つのアルゴリズムによって設計したsiRNAによるRNAi効果の比較を行っており、本手法が最も効率が良かったという結果を示している。

【特許出願数の実績】

該当なし。

ゲノムネットワークプロジェクトの  
課題の報告

プログラム名	(1) <u>ゲノム機能情報の解析</u> (2) ヒトゲノムネットワークプラットフォームの構築 (3) 次世代ゲノム解析技術の開発 (4) 個別生命機能の解析 (5) 動的ネットワーク解析技術開発					
課題の分類	(1) 中核機関 (2) <u>指定課題</u> (3) 公募課題					
課題名 (和英併記)	酵母ツーハイブリッド法による転写因子間の相互作用の解明と補助因子の探索・同定 (Large-Scale Analysis of Human Transcription Factor Networks using High-quality Yeast-Two-Hybrid System)					
代表研究者名 (所属機関・職名)	大友 純 (株式会社日立製作所・主任研究員)					
年度別研究費 (千円)	16年度	17年度	18年度	19年度	20年度	合計
	200,000	200,000	190,000	150,000	80,000	820,000
	研究者名	所属機関・部局・職			専門分野	
研究組織 図	代表機関	☆ 大友 純	株式会社日立製作所 中央研究所・主任研究員			分子生物学
	分担機関	☆ 久原 哲	九州大学 大学院農学研究院・教授			バイオインフォマティクス
		☆ 森田 良治	理化学研究所・オミックス基盤研究領域・ユニットリーダー			分子生物学、配列解析
<b>【研究目的】</b>						
<p>本研究は、日立の高精度ハイスループット酵母ツーハイブリッド(Y2H)系を用いて、ヒト転写因子と相互作用する種々タンパク質を大規模に同定することにより、転写調節におけるタンパク質間ネットワーク全体像の解明とともに、新規調節機構の発見・解析、さらに創薬基盤研究への応用を目指すものである。</p> <p>得られた Y2H データに関しては、実験による検証とバイオインフォマティクスによる検証の両面よりバリデーションを実施する。また、日立 Y2H 系の特長であるドメイン間相互作用データの詳細な解析に取り組む。さらに、得られている大規模データから疾患関連ネットワークを抽出し、創薬・診断分野への産業応用の基盤となる解析を目指す。</p> <p>当該研究のため、Y2H 系並びに実験によるバリデーションにおける塩基配列解析を独立行政法人理化学研究所が担当し、相互作用ドメインの詳細な解析を国立大学法人九州大学が担当し、日立が全体を取り纏めてこれらの 2 機関と共同で実施する。</p>						

### 【課題の概要】

本研究においては、ヒト転写調節因子を解析対象タンパク質(ベイト)とし、ヒト各種臓器・細胞由来ライブラリーをスクリーニングすることにより、相互作用するタンパク質(プレイ)を同定する。転写調節因子群は、縦軸研究機関からのリクエストにより決定し、スクリーニングするヒト各種臓器・細胞由来ライブラリーも、各個別テーマに合うものを選定して、縦軸機関の研究に貢献する。また、中核機関である理化学研究所との連携によっても、解析する転写調節因子を決定し、結果について検討する。さらに、ゲノムネットワークプラットフォームへ順次データ提供を進め、国立遺伝学研究所によるプラットフォーム構築と運営にも貢献するとともに、コンソーシアム内ならびに一般に対してのデータ公開を積極的に行う。

産生したタンパク質間相互作用データに関しては、実験とバイオインフォマティクスによるバリデーションを実施し、精度の評価を行う。また、得られたネットワークデータの中から、ヒト各種疾患に関わるパスウェイを抽出し、創薬基盤研究への応用を目指す。

### 【平成20年8月末までの研究成果】

#### 1. プラットフォーム公開状況(日立)

ゲノムネットワークプラットフォームを通じて、これまでに10回のコンソーシアム内データ公開と、7回の一般公開を実施した。コンソーシアム公開データの内訳は以下の表1の通り。また、全相互作用データのネットワーク図を図1に示した。

表1 日立Y2Hデータのプラットフォーム登録実績

データ種類	データ数
相互作用データ(主に断片対断片)[タンパク質対タンパク質]	2,042 [1,414]
$\beta$ -galアッセイによる相互作用評価	1,489
ベイトの塩基・アミノ酸配列(主に断片)	261
プレイの塩基・アミノ酸配列(主に断片)	1,874
スクリーニングしたライブラリーの種類(臓器・細胞)	14

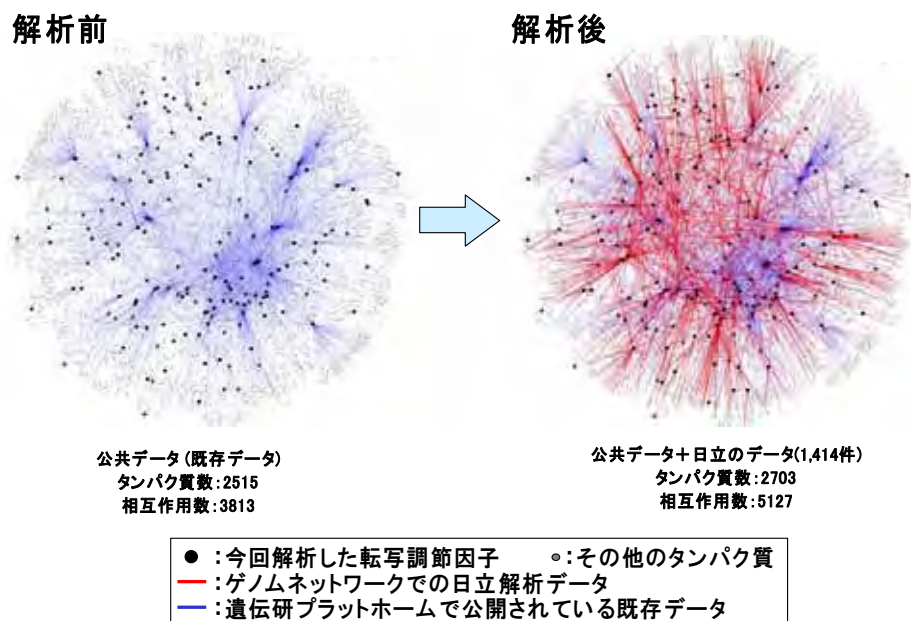


図1 日立Y2H解析により得られた新規転写調節因子ネットワーク