

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 21 日現在

機関番号：82603

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22510223

研究課題名（和文） ゲノムスキニングによる生物活性物質探索系の確立

研究課題名（英文） Development of a system to find bioactive molecules by genome scanning

研究代表者 石川 淳（ISHIKAWA JUN）

国立感染症研究所・生物活性物質部・室長

研究者番号：40202957

研究成果の概要（和文）：放線菌ゲノムに含まれる未知の生物活性物質生合成遺伝子クラスターの探索系を開発するために、*Nocardia brasiliensis* IFM 0406 株において次世代シーケンサーを用いたゲノムスキニングを行い、多くの生物活性物質生合成遺伝子クラスター候補を見出し、そのうちのひとつについては、調節遺伝子の強制発現によって代謝産物を検出することができた。この手法は新規な生物活性物質の探索に有効であると考えられる。

研究成果の概要（英文）：To develop a system to find bioactive molecules by genome scanning, *Nocardia brasiliensis* IFM 0406 was sequenced by a next-generation sequencer. Many candidates for biosynthetic gene clusters of bioactive molecules were found, and the metabolite of one of the clusters was detected by forced expression of positive regulatory genes. These results suggest that genome scanning facilitates the finding of novel bioactive molecules.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010 年度	1,200,000	0	1,200,000
2011 年度	1,100,000	0	1,100,000
1012 年度	1,100,000	0	1,100,000
総計	3,400,000	0	3,400,000

研究分野：微生物ゲノム科学

科研費の分科・細目：ゲノム科学・応用ゲノム科学

キーワード：ゲノム、次世代シーケンサー、抗生物質、応用微生物、創薬、生物活性物質

1. 研究開始当初の背景

現在我々が認知している生物活性物質は、放線菌の持つ能力の 1 割ほどでしかないことが放線菌ゲノム解析により明らかになった。残りの 9 割には新規な生物活性物質が多数含まれていると期待されているが、それらの解析は、現時点でゲノム解析が終了している少数の菌種に限られている。しかしながら、今後次世代シーケンサーによるゲノム解析が頻繁に行われるようになると予想されるため、次世代シーケンサーのデータを効率よく用いて、多くの未知ゲノムにアプローチす

るための手法の確立が望まれる。

2. 研究の目的

現在用いられている抗生物質や抗がん剤などの生物活性物質の大半は、放線菌と呼ばれる菌群により生産された化合物をもとにして作られている。近年、放線菌においても数種のゲノム配列が明らかにされ、多様な生物活性物質を生産する背景がゲノムレベルで解明されつつあり、その創薬への応用に期待が持たれている。放線菌は多くの場合、複数の化学構造の異なる生物活性物質を生産

する能力を持ち、生産物の数だけ生合成遺伝子セット（クラスター）を持っている。生合成遺伝子クラスターは数十キロベースから百キロベースを超えることもあり、そのような大きな遺伝子セットをゲノム中に複数保持していることだけでも興味深い。それら既知の生合成遺伝子クラスターは、放線菌ゲノム解析によって予測された全ての生合成遺伝子クラスター候補のわずか1割ほどではないことが明らかになった。残りの9割の生合成遺伝子クラスターは、その産物の化学構造は未知であるものの、ポリケチド合成酵素（PKS）やペプチド合成酵素（NRPS）などの既知の生合成酵素との配列相同性を用いて予測されているため、既知物質の類縁体であると想像される。しかしながら、生合成遺伝子あるいは酵素が明らかになっていない物質や、既知物質とは全く構造の異なる化合物の場合は、既知物質の生合成酵素とは相同性がないと考えられるため、そのようなものを含めれば、さらに多くの未知（新規）物質の生合成遺伝子が放線菌ゲノムの中に眠っていると考えられる。

このような状況の中、残り9割の未知生物活性物質を同定するための努力が国内外で行われているが、その対象は現時点でゲノム配列が明らかになっている小数の菌種に限られている。その理由のひとつは、ゲノムシーケンシングが高コストであり、しかもコストに応じた結果が得られるかどうかかわからないことである。しかしながら、いわゆる次世代シーケンサーの登場により、ゲノムシーケンシングのコストは従来の十分の一以下に低下してきているため、これを有効に利用することを考えた。次世代シーケンサーによって得られる配列は、従来のサンガー法によって得られる配列に比べ非常に短く、それをアセンブルして得られるコンティグ配列も相対的に短い。したがって、次世代シーケンサーを用いた場合には、生合成遺伝子クラスター全体が一度に明らかになるとは考えにくく、クラスターの一部であると予測される断片を探索することになる。

Nocardia brasiliensis IFM 0406 株は本邦の臨床分離株であり、これまでにブラジリカルジン（免疫抑制作用、Komaki et al., J. Antibiot. 52:13-19, 1999）、ブラジリノライド（免疫抑制作用および抗真菌作用、Tanaka et al., J. Antibiot. 50:1036-1041, 1997）、ノカルディオラクトン（抗菌作用、Mikami et al., Nat. Prod. Lett. 13:277-284, 1999）などの多様な生物活性物質を生産することが明らかにされてきた病原性放線菌の一種である。ブラジリカルジンについては先ごろ生合成遺伝子がクローニングされたが（Hayashi, Y., et al. J. Antibiot. 61:164-174, 2008）、他については試みられておらず、その生合成過程は不明である。そこ

で本研究では、次世代シーケンサーを用いて *N. brasiliensis* IFM 0406 株のゲノムシーケンシングを行い、種々の二次代謝産物の生合成遺伝子を探索し、コスミドライブラリーから発見した遺伝子を含むクローンを単離し、申請者らが開発した *Nocardia* 宿主・ベクター系（Chiba, K., et al. Jpn. J. Infect. Dis. 60:45-47, 2007）を用いて遺伝子を発現させ、生産物を特定することにより、生物活性物質の生合成過程や調節機構の解明につなげることを目的とした。

IFM 0406 株のゲノム配列ははまだ解析されていないことから、発見された生合成遺伝子クラスターが生産物が特定されれば、ゲノム配列未知の生物活性物質生産菌のゲノムシーケンシングによって生合成遺伝子を同定した初めての例（モデル）となり、この手法をより有用な物質の生産菌に適用することで、創薬研究に大きな進展をもたらすと期待された。

3. 研究の方法

N. brasiliensis IFM 0406 株は千葉大学真菌医学研究センターより分与された。Red-ET法に用いたプラスミドはエール大学 Coli Genetic Stock Center から入手した。ゲノムシーケンシングは Nextera (Epicentre) で作製したライブラリーをイルミナ社 GAIIX を用いて行った。遺伝子予測には MetaGeneAnnotator (Noguchi, H. et al. DNA Res. 15:387-396, 2008.) を用いた。タンパク質のドメイン検索には Pfam 26.0 (Punta, M., et al., Nucl. Acids Res. 40:D290-D301, 2012) を用いた。その他にも多数の in-house Perl スクリプトを開発して配列解析に用いた。

4. 研究成果

ゲノムアセンブリーの最適化：次世代シーケンサーデータのアセンブリーは発展途上であるため、様々なアセンブリーで、様々な条件でアセンブルを行い、好適と考えられるアセンブリーを得る必要がある。そこで、アセンブリーとして ABySS (Jared, T. et al. Genome Res. 19:1117-1123, 2009) および Velvet (Zerbino, D. R. & Birney, E., Genome Res.

表 1. 最適なアセンブリー

コンティグ数	27,255
≥1 kb	1,725
≥5 kb	206
≥10 kb	64
≥50 kb	1
最長	51.7 kb

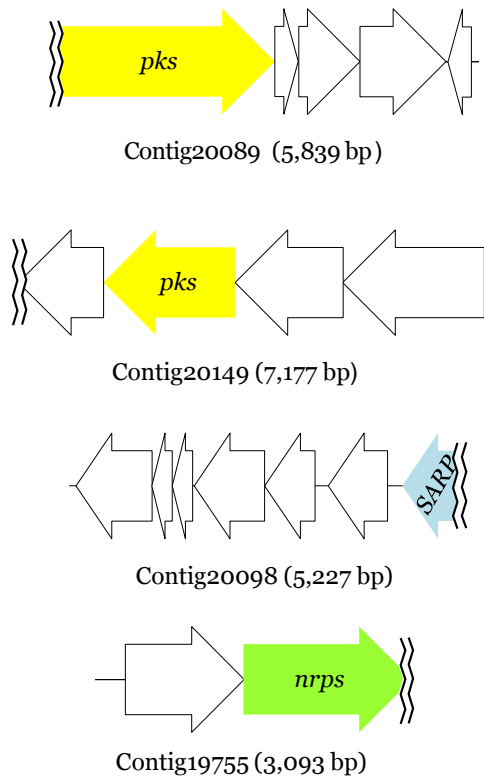


図 1. 生物活性物質生成遺伝子を含むコンティグの例。PKS (黄色)、NRPS (緑色)、SARP (水色)。

18:821-829, 2008) を用い、各アセンブラーで数 100 種類の条件でアセンブリーを作成した。アセンブリーの良し悪しは、コンティグ数、最長コンティグ、コンティグ総延長、さらにブラシリカルジン生成遺伝子クラスターを網羅できているかで判断した。その結果、アセンブリーの良し悪しはリード長が長ければ良いわけでも、多くのリードを用いれば良いわけでもなく、数 100 通りのアセンブリーを作成することで初めて明らかとなった。

最適なアセンブリー (表 1) は 1,725 本 (≧ 1kb) のコンティグからなり、最長コンティグは 51.7kb であった。(筆者注: 現在は次世代シーケンサーの性能がより向上しているため、コンティグ数は 500 前後、最長コンティグは 400kb 以上となっている。)

生成遺伝子候補の探索: 上述の最適なアセンブリーを用いて、生物活性物質の生成に密接に関わっているポリケチド合成酵素 (PKS)、ペプチド合成酵素 (NRPS) および糖転移酵素、さらに経路特異的調節因子 (SARP) などを BLAST、Pfam などを用いて探索した。その結果、PKS、NRPS あるいは SARP を含むコンティグをそれぞれ 20、32 あるいは 5 個見出した。

コンティグの連結: 生成遺伝子クラスターは、数 10 キロベース (kb) から 100kb を

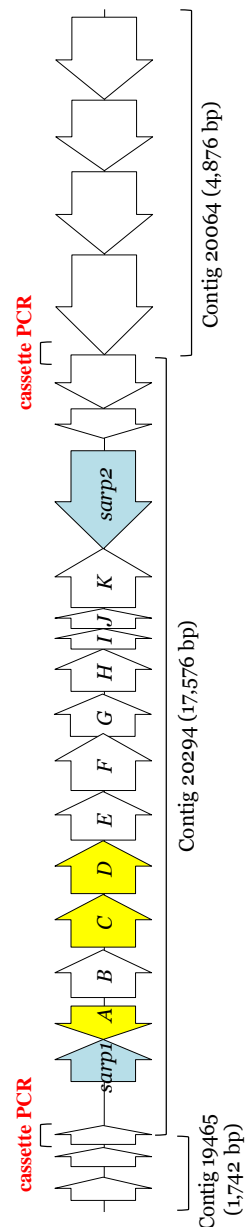


図 2. コンティグウォーキングにより明らかとなった *hsn* クラスターの構造。II 型 PKS (黄色)、SARP (水色)。

超えることも多いため、上で見出したコンティグのいくつかについて、その周辺配列をカセット PCR (Isegawa, Y. et al. Mol. Cell. Probes 6:467-475, 1992.、タカラバイオのキットを使用) を用いて解析し、その配列を含むコンティグをアセンブリー中に探索することによって、隣接するコンティグを明らかにした (コンティグウォーキング)。

生成遺伝子クラスター領域のクローニング: コンティグウォーキングによって明らかになった二次代謝産物生成遺伝子クラスター候補のうち、II 型 PKS および SARP を含むクラスター (*hsn* クラスターと命名) (図

2) をクローニングするために、コスミドライブラリーから Red-ET 法を用いてターゲットを含むコスミドをタグ付によって単離した。

SARP の強制発現による *hsn* クラスターの活性化と代謝産物の検出: *hsn* クラスターは二つの SARP 遺伝子 (*sarp1* および *sarp2*) を含むことから、*hsn* 遺伝子群は SARP 遺伝子の発現があって初めて発現すると考えられた。そこで、*sarp1* および *sarp2* を構成的に発現するプラスミドを作製し、IFM 0406 株に導入した (図 3)。その結果、*sarp1* および *sarp2* を強制的に発現させた株でのみ *hsnF* 遺伝子の発現が観察された。

続いて、HPLC を用いて代謝産物の解析を行った (図 4)。その結果、*sarp1* と *sarp2* を同時に発現させた時に新たなピークが検出された。これらのピークの UV スペクトルは互いに類似しており (データ示さず)、複数の代謝産物が生産されたわけではなく、ひとつの代謝産物とその中間体が検出されていると考えられた。

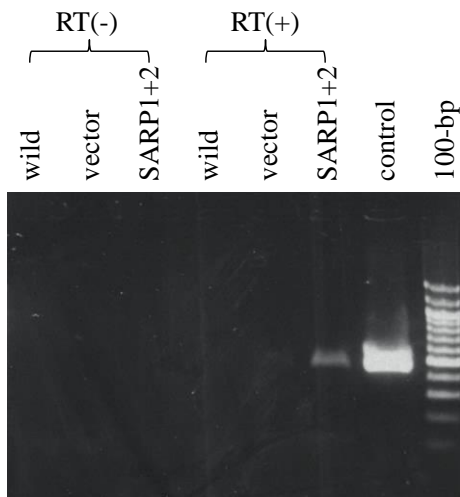


図 3. SARP 遺伝子の発現の有無による *hsnF* 遺伝子の発現変化。各株における *hsnF* の発現を RT-PCR 法により解析した。RT, reverse transcription

結論: 次世代シーケンサーを用いてシーケンシングを行い、最適なアセンブリーの作製、コンティグウォーキングなどを行うことにより、未知のゲノムをスキヤニングすることで、二次代謝産物の生合成遺伝子を効率よく見出すことができ、その遺伝子を含むコスミドクローンを Red-ET 法を用いて迅速に単離できた。さらに、調節遺伝子 (SARP ファミリー) を強制発現させることにより、通常の培養条件下では生産されていない代謝産物を生産させることができた。以上の戦略は、放線菌ゲノムの中に眠る生物活性物質生合成遺伝子を迅速に見出し、その代謝産物を

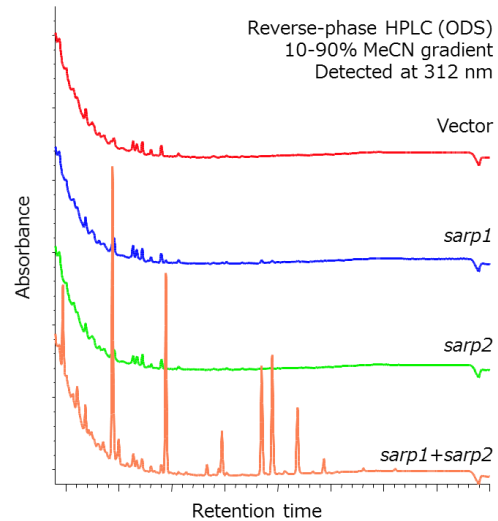


図 4. SARP 遺伝子の強制発現による代謝産物の変化。

同定するために有効であると考えられる。

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 2 件)

Hoshino, Y., Chiba, K., Ishino, K., Fukai, T., Igarashi, Y., Yazawa, K., Mikami, Y., and Ishikawa, J.: Identification of nocobactin NA biosynthetic gene clusters in *Nocardia farcinica*. J. Bacteriol. **193**:441-448 (2011) 査読有

Yamamura, H., Ohnishi, Y., Ishikawa, J., Ichikawa, N., Ikeda, H., Sekine, M., Harada, T., Horinouchi, S., Otoguro, M., Tamura, T., Suzuki, K., Hoshino, Y., Arisawa, A., Nakagawa, Y., Fujita, N., and Hayakawa, M.: Complete genome sequence of the motile actinomycete *Actinoplanes missouriensis* 431^T (= NBRC 102363^T). Stand. Genomic Sci. **7**:294-303 (2012) 査読有

[学会発表] (計 5 件)

石川 淳: 次世代シーケンサーによる *Micromonospora griseorubida* のゲノム解析、第 25 回日本放線菌学会大会、2010 年 9 月 2 日、東京

Hoshino, Y., Chiba, K., Ishino, K. and Ishikawa, J.: Nocobactin NA biosynthesis gene clusters in *Nocardia farcinica*. 16th Intl. Symp. on the Biology of Actinomycetes, Dec. 11-15, 2011, Puerto Vallarta, Mexico.

石川 淳：二次代謝産物研究における次世代シーケンサーの利用、日本農芸化学会東北支部シンポジウム（招待講演）、2012年6月30日、秋田

石川 淳：次世代シーケンサーによる「とりあえずゲノム」解析のすすめ、第56回日本医真菌学会総会（招待講演）、2012年11月10-11日、東京

石川 淳：*De novo* アセンブリーの最適化と二次代謝生合成遺伝子発見ツールの開発（シンポジスト）、第86回日本細菌学会総会、2013年3月18-20日、千葉

6. 研究組織

(1) 研究代表者

石川 淳 (ISHIKAWA JUN)

国立感染症研究所・生物活性物質部・室長
研究者番号：40202957