

博士論文番号：

大腸菌全遺伝子の発現クローンセット作製の研究

北川正成

平成18年7月11日提出

目次

第1章	緒言	3
第2章	材料と方法.....	9
	1. 使用菌株と培地	
	2. ベクターDNA の調製	
	3. PCR primers	
	4. PCR 増幅反応	
	5. PCR 増幅断片の精製	
	6. pCA24N ベクターへの各 ORF のクローン化	
	7. 制限酵素 <i>Bgl</i> Ⅰ, <i>Bam</i> HI, <i>Sfi</i> Ⅰ による確認	
	8. GFP 発現コロニーの観察	
	9. クローンのシーケンス解析	
	10. 融合タンパク質発現の確認と精製	
	11. GFP 局在の観察	
第3章	結果と考察.....	20
	1. 大腸菌全遺伝子のクローン化と発現用のプラスミド pCA24N の構築	
	2. ORF の PCR 増幅	
	3. pCA24N へのクローニング	
	4. 産生された GFP 融合タンパク質の細胞毒性	
	5. さまざまなリソースとしての利用	
	1) DNA マイクロアレイの鋳型への利用	
	2) <i>Sfi</i> Ⅰ によるサブクローン化とさまざまな受け皿ベクターの利用	
	3) <i>Not</i> Ⅰ による GFP 領域除去	
	4) タグつきタンパク質のアフィニティ精製	
	5) プルダウンアッセイによるタンパク質の相互作用解析	
	6) タンパク質の構造解析への利用	
	7) GFP 融合タンパク質の局在性観察	
	6. システムバイオロジーのための遺伝子機能解析ツールとしての今後	
	謝辞.....	56
	参考文献.....	57

第1章 緒言

過去半世紀以上にわたって、大腸菌は遺伝学、生化学、分子生物学的な研究の好材料として研究され、膨大な意義深いデータの蓄積をわれわれに与えてくれた¹。

この大腸菌の代表的な実験株であるK12株のゲノムには、その全塩基配列から、約 4,300 個のタンパク質をコードする遺伝子が存在していると推定されている^{2,3}。

大腸菌はこれらの遺伝子群により単細胞原核生物としてこの地球上に生き続けているが、「どのようにこれらの遺伝子を使って生き抜いているのか?」、「各遺伝子の役割は何か?」という疑問に対して、その答にもっとも近いモデル生物の一種と考えられる。ゲノム研究がポストゲノム解析の時代へと進んだ現在、大腸菌は、システムバイオロジーの分野において、最適な材料を提供してくれる価値ある生物であり、ひとつの生命体としての総合理解にもっとも近い生物の候補であることは間違いない⁴。

大腸菌 K12 株には、約 4,300 個の遺伝子が存在し、時には周りの環境変化への適応などに迫られながら、それらすべての遺伝子が協調しながら複雑な相互作用を展開して働くと考えられる。一方、これらの遺伝子群の中には、構造（配列）的には互いによく似ているものもあり、微妙な役割分担を行っている場合もある。これらの協調および相互作用や役割分担は、近年、「ネットワーク」という言葉で言い表され、その研究は、ゲノム解析が容易でなかった当時から、分子遺伝学的な解析や生化学的な解析技法を駆使して進められてきたものである。

一例として、筆者自身が研究に取り組んでいた熱ショック応答に関して紹介

する。大腸菌を 30℃以下の温度環境から 42℃程度の高温ストレスにさらすと、熱ショックタンパク質と呼ばれる一群のタンパク質の発現が一過的に誘導される⁵。この誘導は σ^{32} と呼ばれるRNA polymeraseの σ サブユニットの1種による制御を受けていることが知られている。この熱ショックタンパク質のなかには、分子シャペロンやプロテアーゼといった、タンパク質のフォールディングや分解などの品質管理に関与するタンパク質が多く知られており、それぞれさまざまな役割分担や共同作用を行いながら、細胞の恒常性の維持やストレスで損傷した機能の回復などに貢献していると考えられる。この熱ショック応答の全体像のような複雑な機構を理解するには、そこに登場する役者たちを網羅的に解析する必要があった。

筆者は、そのような観点にたって、ストレスに応答する大腸菌遺伝子を網羅的に見つけるために、熱ショック応答性の転写プロモーターのスクリーニングを行い、その中から、*clpB*遺伝子の熱ショック誘導性のプロモーターを発見した⁶。大腸菌にはそのパラログ遺伝子として*clpA*という相同な遺伝子も存在しており、その産物ClpAタンパク質は、ATP依存性プロテアーゼのATPaseサブユニットであり、プロテアーゼ活性の制御因子として機能することがわかっていたが^{7,8}、一方のClpBタンパク質は分子シャペロンとしての機能を持つことがその後明らかにされ、このClpファミリータンパク質は別名HSP100ファミリータンパク質として広く生物に保存されていることがわかっている⁹。

また、大腸菌をはじめとする原核生物にはポリシストロニックな転写単位が存在し、多くの場合、そのような同一の転写単位上の遺伝子群は、共同してひとつの代謝機能や生理機能を担っていることが知られている¹⁰。しかし、このようなオペロンを形成している遺伝子のオペロンである必要性に関しては、実際には意味不明なケースもよく見られ、なぜこのような転写単位を持っているの

か、想像が難しいケースも少なくない。たとえば、大腸菌にはリポプロテインという膜タンパク質がいくつかあり、その分泌の過程でシグナルペプチドを切断する役割を持つペプチダーゼとして、リポプロテインシグナルペプチダーゼ（遺伝子 *lspA*）が知られている。その *lspA* 遺伝子とイソロイシン tRNA 合成酵素の遺伝子 *ileS* がオペロンを形成していることを筆者は明らかにしたが¹¹、そのオペロン構造形成の意味は未だに不明である。

このような遺伝子の多様な協調や役割分担、あるいはその遺伝子構造の意味を理解するためには、そのシステムの中の構成遺伝子それぞれを独立に用意しながら、それらの相互作用関係（ネットワーク）について、遺伝子発現量の変化やタンパク質機能発揮の状況など、各段階の解析を網羅的に進める必要があると考えられる。

大腸菌では、古くから網羅的なリソースの作製とその利用が行われてきた歴史がある。たとえば、クラーク・カーボンのコレクションと呼ばれる可動性のプラスミドベクターにゲノムの断片をランダムにクローン化したライブラリは有名なものである¹²。これは、大腸菌の変異株の表現型を相補するような遺伝子スクリーニングするために古くから重宝されてきた。

また、小原らが作製した、ラムダファージをベクターとするゲノムライブラリは、大腸菌ゲノムの物理地図の作成に利用され、約 3000 個のファージクローンからゲノムの精密な制限酵素地図が作られた¹³。大腸菌ゲノム全体を最少数でカバーするように、そこから選ばれた 476 個のファージは小原クロンのミニマムセットとして、大腸菌 K12 株の派生株 W3110 の正確なゲノムシーケンス解析をはじめ、大腸菌ゲノム研究のリソースとして活用されてきた。

その他、ファージやコスミドをベクターにしたライブラリが、ゲノム配列決定以前から研究のリソースとして作製された歴史があり^{14,15,16}、大腸菌を使用し

での網羅的な研究がさまざまな角度からアプローチされてきたが、ゲノム配列全体が決定されたことにより、よりシステムティックな解析が可能となった。

大腸菌のゲノム配列決定は、日本とアメリカのグループによって、それぞれ独自に進められた。

日本の大腸菌ゲノム配列決定プロジェクトの開始は、小原クロンの発表から間もなくの 1989 年に始まった。当時はまだDNAシーケンサーも開発途上にあり、ほとんど実用的なレベルになかったため、結局ラジオアイソトープを利用した従来技術が主として用いられ、大規模配列解析と言える状況ではまっくくなかった。また、機能不明な領域の塩基配列決定の意味自体についての議論も不十分だったため、組織的に配列決定を推進する体制の構築も難しく、企業に委託するという形で配列決定が進められた（筆者はその業務を委託され、担当した一人であった）。そのような状況でも、当時としては相当に長い、111.4kb の連続した領域（染色体地図上 0-2.4 分）の配列がDNAデータベースに登録された¹⁷。その後さらに解析は進められ、染色体地図上 4.1 分までこの全長は延ばされた¹⁸。この時のプロジェクトの成果は単に長さにあるのではなく、意味のわからない領域のDNA配列の決定から、間違いなく価値のある情報を発掘できることを実証したことであった。たとえば、*fix*遺伝子群という、*Azorhizobium caulinodans*が持つ窒素固定に機能する遺伝子に高い相同性を持つ遺伝子のオペロンなど、新たな興味深い遺伝子が見つかったことも一例であろう。その他、外膜構築に関与する遺伝子群の大きなオペロン構造が種を越えてよく保存されて存在することや、BIMEと呼ばれる反復配列の分布状態の特徴（タンパク質のコード領域には存在しない）などが示唆された。一方、それと同時に、それまでにGenbankなどの公共データベースに登録されていた大腸菌の配列データに

は、正確性という点で多くの問題（登録者の人為的ミスによるベクター配列等の混入など）が潜んでいたことも明らかとなった。

この成果を礎に、大腸菌のプロジェクトは第 2 期と呼ばれる時期に入り、1992 年からの 2 年間で 400kb に及ぶ配列が「分担作業」によって決定された。ちょうどそのころ、DNA シーケンサーの技術革新が大きく発展し、それまではとても不可能に思われていた、大腸菌のゲノム全体を配列決定するという作業も、「分業体制」という徹底的な効率的体制を研究者組織で構成し、TIGR のインフルエンザ菌解析¹⁹の例などを見習って配列決定の最新技術とシステムも取り入れながら進められ、大腸菌のプロジェクトは第 3 期と呼ばれる時期を迎えた。そこでは、小原クローンから大腸菌のゲノム断片の部分だけを long-PCR によって増幅し、それをショットガン法による配列決定に供していく作業を、日本各地の大腸菌研究者の研究室で分業して進められた（筆者自身もこのプロジェクトに参加した）。解析は冗長度 8 以上を目標にし、小原クローンひとつあたり 250 解析のシーケンスの結果とギャップとして残った部分の解析が行われ、染色体地図上 70 分までの領域が配列決定された^{20,21,22,23}。その配列に、すでにアメリカのグループが決定していた 70 分以降のデータを取り込む形で、大腸菌ゲノムの配列が一本の環状染色体として「ひとつの輪」につながったのは 1997 年の初めのことであった。

本論文では、このように日本のグループがゲノム配列解析を行った大腸菌 K12 株の派生株 W3110 のゲノム配列をもとに、そこから見出された Open reading frame (ORF : Genbank 形式のファイルで用いられる Coding sequence を意味する CDS という呼び方もあるが、本論文ではあえて ORF と呼ぶ) を対象として、推定遺伝子も含めたすべてのタンパク質コード領域 4,276 個の ORF のほぼすべてを、独自に工夫を施した発現ベクター pCA24N プラスミドを用いて、

Hisタグと蛍光タンパク質Green fluorescent protein(GFP)の両者に方向性と読み枠を考慮した形でクローン化することに成功した過程を示す。

また、これを材料にして、

- ・ DNA マイクロアレイの実用化に際し、鋳型調製の材料として活用されていること。
- ・ 蛍光タンパク質 GFP との融合タンパク質の観察から、各タンパク質の細胞内での局在性解析の材料として実用可能であること。
- ・ His タグを利用して、網羅的にタンパク質の精製が可能であり、それを利用してタンパク質間相互作用を網羅的に解析できること。
- ・ レアな認識配列酵素である制限酵素 *SfiI* によって方向性と読み枠を保持したままに容易にサブクローン化して、さまざまな目的の機能解析に供することができること。
- ・ ORF 産物の誘導発現による宿主大腸菌の生育阻害の評価。

等を論ずる。

このクローンセット”ASKA library (A Complete Set of *E. coli* K-12 ORF Archives)” は、遺伝子機能の網羅的解析にとって格好の素材として、すでに多数の多様な研究に活用されているが、今後もシステムバイオロジーにおけるきわめて有用な実験材料として、さらなる活用が期待される。

第2章 材料と方法

1. 使用菌株と培地

大腸菌 K-12 株の派生株 AG1 [*recA1 endA1 gyrA96 thi-1 hsdR17(rK-mK+) supE44 relA1*](Stratagene, Inc.) をクローン作製の宿主として使用した。

大腸菌の培養は L-broth medium (1% Bacto Tryptone (BD Diagnostic Systems), 0.5% yeast extract (BD Diagnostic Systems), 0.5% NaCl) に必要に応じて、50 μ g/ml の ampicillin(Amp) (Meiji Seika Kaisha, Ltd.), 30 μ g/ml の kanamycin(Km) (Wako Pure Chemical Industries, Ltd.) あるいは chloramphenicol(Cm) (Nacalai Tesque, Inc.) をそれぞれ添加した。クローン化した遺伝子の誘導発現には Isopropyl-beta-D-thiogalactoside (IPTG) を 0.1mM の濃度で添加した。アーカイブプラスミドの調製用の培養には、30 μ g/ml の Cm を添加した 2xYT 培地(1.6% Bacto Tryptone (BD Diagnostic Systems), 1% yeast extract (BD Diagnostic Systems), 0.5% NaCl)を使用した。

2. ベクターDNAの調製

ORF断片のクローン化には、新規構築のpCA24Nを使用した(**Fig.1**)。このプラスミド上の *StuI* 部位にはさむ形でKm耐性遺伝子のカセット配列を持たせた pCA24N-Km から、制限酵素 *StuI* (TAKARA BIO INC.) 切断によってKm耐性遺伝子領域を切り出したうえで直鎖化し、アルカリフォスファターゼ (TAKARA BIO INC.) による処理を 50°C、2 時間行って末端を脱リン酸化後、Km カセット配列断片と *StuI* による切れ残り断片を除くために、アガロース電気泳動により、直鎖化したベクターDNAを分離精製した。

一方、Km 耐性カセット配列を別途 PCR 増幅 (プライマーの 5'末端は、後

述の ORF の増幅時と同様に、あらかじめリン酸化済み)したインサート断片(コントロール実験用インサート断片)を用意し、上記で調製したベクター断片とそのインサート断片の **ligation** 反応を行って、その反応物による大腸菌への形質転換によって得られる **Cm** 耐性(ベクターのマーカー)獲得の形質転換効率を調べた。インサート断片なしの場合は、相対的に 1,000 分の 1 以下のコロニーしか出てこないことと、得られた形質転換コロニーがすべて **Km** 耐性も同時に獲得していることを確かめることによって、きわめてバックグラウンドの低い直鎖化ベクターDNAであることを確認した上で、それを用いて、すべてのクローニング操作を行った。

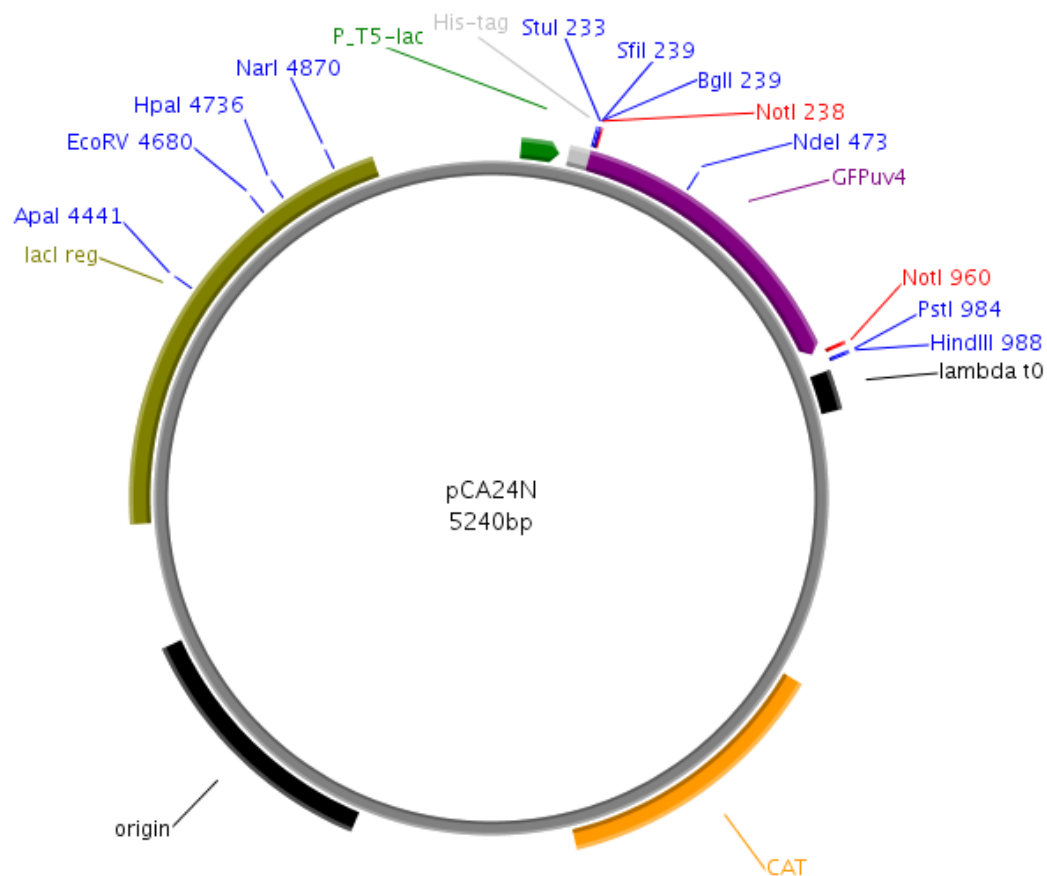


Fig.1 クローニングベクタープラスミド pCA24N の構造

ORFのクローニングに使用したプラスミドpCA24Nの構造を示したプラスミド機能地図を示す(PlasMapper Ver.2.0 (<http://wishart.biology.ualberta.ca/PlasMapper/>) により描画)。“P_{T5-lac}”はIPTG誘導性の転写プロモーターである。灰色の“His tag”は、Histidineを6個連なったタグ配列領域である。紫の線で示した“GFPuv4”は、今回採用したGFP遺伝子をコードしている領域を示している。“CAT”はchloramphenicol耐性遺伝子の産物 chloramphenicol acetyltransferaseの遺伝子を示している。“lambda t0”は転写終結シグナルの位置を示す。“origin”はプラスミドの複製起点の領域である。“lacI reg”が示している領域は*lacI*の遺伝子とそのコード領域を示している。このプラスミドにとって重要な制限酵素認識部位である*StuI*と*NotI*の位置のほか、1ヶ所切断の制限酵素サイトのいくつかを表示した。このプラスミドの配列情報は、GenBankで公開されており、Acc.No.はAB052891である。

3. PCR primers

PCRによって大腸菌のORFをコードするDNA断片を増幅するとき、ORFのN末側とC末側の2個のPCR primerの5'末端にGCCあるいはCCの余分な配列を、それぞれ付加した。

例えば、

5' -ATG CGC ATA... GTA CGT TAA-3'
3' -TAC GCG TAA... CAT GCA ATT-5'

という配列のORFがあったとき、開始コドンATGと終止コドンTAAを除き、

N末側のプライマーは、GCCCGCATA...

C末側のプライマーは、CCACGTAC...

のような配列になるようにした。長さは23塩基が基本であるが、それでPCRによる増幅が芳しくなかったものについては30から35塩基に長さを伸ばしたものを用いた。それらのプライマーの配列はインターネットを通じて公開している (http://ecoli.naist.jp/gb6/Resource_download.html)。

これらのプライマーペアでのPCR増幅によって増幅されるDNA断片を二本鎖で表すと、

5' -GCC CGC ATA...GTA CGT GG-3'
3' -CGG GCG TAA...CAT GCA CC-5'

という配列のDNA断片が増幅されることになる。

ベクターのクローニング部位配列には、

5' -GGCCCTGAGGCCTATACGGCC-3'

のように制限酵素 *StuI* (**AGGCCT**) の切断配列の前後に、それと重ねて制限酵素 *SfiI* 認識配列 **GGCCNNNNNGGCC** で、中心3個の塩基の 3' -突出で切断) とその変形の配列 (中の **N** の数が4塩基) が存在する。この *StuI* の切断位置に平滑末端結合によって前記の PCR 産物をクローニングすると、目的とする方向 (転写) 翻訳方向が ORF のコード方向と一致する向き) でクローニングされた場合、その配列は、

5' -GGCC CTGAG GGCC-CGC ATA...GTA CGT GGCC TATGC GGCC-3'

となり、ORF を含む断片が制限酵素 *SfiI* で切り出される (制限酵素 *BglII* でも切り出される) ような配列構造を持つようになる。

一方、逆向きにクローニングされた場合は、

5' -GGCC CTGA GGCC-acg tac...tat gcg-GGCC CTATGC GGCC-3'

のようになり (ORF 部分の相補配列を小文字で表示)、*SfiI* の配列はできないので切れない (*BglII* でも切れない) ことになる。このようにして、*SfiI* (および *BglII*) で切れるか切れないかを指標に、クローンの向きの正しさを検証し、

また切り出される断片の大きさによりクローンの区別をつけることができるように工夫を施した。

プライマー配列のmodifyについて

以下のようなケースの場合、対象 ORF の配列によって、プライマー配列がゲノム配列とは違う塩基に置換されている。

ケース 1. アーカイブベクターにクローン化後、人工的に *Sfi*I 配列ができてしまっただ都合が悪い場合、アミノ酸を変えないように一部置換した。

具体的には、ORF の配列が以下のような時、

(N-terminal) ATG NNN NNG GCC N...

または

....N NNN GGC CNN NNN TAA (C-terminal)

そこからできる PCR 断片の配列はそれぞれ、

(N-terminal) GCC NNN NNG GCC N...

または

....N NNN GGC CNN NNN GG (C-terminal)

のようになり、ベクターにクローン化後の配列はそれぞれ、

..AGGGCC NNN NNG GCC N...N NNN GGC CNN NNN GGCCT....

となる。つまり、

(N-terminal) ..AGGGCC NNN NNG GCC N...

または

....N NNN GGC CNN NNN GGCCT.... (C-terminal)

となるため、余計な *SfiI* 部位が ORF 内部にできてしまう。

そこで、それを避けるため、

N 末側が該当する場合は、ORF の 3rd コドン NNG をアミノ酸の変化しないコドンに変えることにした。なお、3rd アミノ酸が Met あるいは Trp の場合は、代替できるコドンがなく変えられないので、その場合は 4th コドン GCC を GCG に変える必要があったが、実際にはそのようなケースはひとつもなかった。

C 末側が該当した場合は、ORF の後ろから数えて 3 番目のコドン GGC を GGT に変えることによって、余計な *SfiI* 部位の生成を避けた。

ケース 2. ORF の最後のコドンの 3rd letter が G の場合。

ORF の最後のコドンの 3rd letter が G のとき、プライマー配列の最初が CCC... となるので、それを避けつつ、しかしアミノ酸は変わらないように、プライマーの 3 番目（つまり ORF の終止コドン直前で、できあがる *SfiI* 配列の直前）の塩基を置換した。

具体的には、ORF の配列が以下のような時である。

ATG NN...NN NNG TAA

この置換の理由は、PCR 産物のクローン化では、増幅断片の端の塩基が脱落することが起こる場合があるからである。通常の ORF の配列なら、*SfiI* で切れなくなるのでそういうものを区別できるが、例のような場合に限って、ひとつの塩基の欠失が起きた場合でも *SfiI* 配列が生じてしまうため、制限酵素によるチェックで気づかない可能性があり、それを避けるために、アミノ酸は変えないような C 以外の塩基に置換した。

ただし、該当するアミノ酸が Met(ATG)か Trp(TGG)の場合、変える余地がなかったなので、そのままにしておき、実際の配列確認によって問題ないことを別途確かめた。

この置換は、クローン化作業の後半の ORF にのみ行なっていて、先に作業した 2,000 個の ORF については置換されておらず、該当 ORF について問題ないことは別途配列確認することで確かめた。

4. PCR 増幅反応

全てのPCR primerは、T4 kinase (TAKARA BIO INC.) により 5'末端のリン酸化反応に供した後に、その反応液をそのままPCR反応液に投入した。その場合、リン酸化反応液から持ち込まれるMgCl₂を考慮して、次のPCR反応液に添加するMgCl₂を、各PCR酵素の推奨濃度になるように調整した。

PCR 増幅は 96 ウェルフォーマットのプレートチューブで行った。増幅の酵素としては、正確性に優れている KOD DNA polymerase (TOYOBO CO., LTD.) を採用し、以下の反応系（総容量 50μl）で増幅した。

1unit of KOD DNA polymerase

20-30 ng of Kohara phage clone DNA or *E. coli* W3100 genomic DNA

1.0 uM of each primer

200 uM dNTPs

PCR プログラム : 25 cycles of at 95°C for 15 s, 64°C for 15 s, 72°C for 4 min followed by a final incubation at 72°C for 5 min.

全体の約90%の遺伝子（4217個）については、KOD DNA polymeraseによって

増幅に成功したが、増幅できなかったものの多くは、長さが2kbを超えるものであった。それらの中の90%程度は、KOD DNA polymeraseに近い正確性を持つと言われるDynaZyme (Finzymes Oy)を用いることによって増幅できた。残り（全体のおよそ1%弱）については LA-Taq DNA polymerase (TAKARA BIO, INC.)によってのみ増幅が可能であった。

Dynazyme および LA-Taq DNA polymerase による増幅の場合も同様の反応を行ったが、これらの酵素で PCR を行った場合は、増幅断片の末端を平滑化するために、反応液を MagExtractor kit(TOYOBO CO., LTD.) で精製後、KOD DNA polymerase によって末端平滑化反応を行った。

すべての反応は全自動のシーケンス反応ロボットである PRIZM877 (Applied BioSystems Inc.製) を用いて行い、操作過程の記録を取り、人的な操作ミスによる検体の取り違いなどが起きないように注意した。

5. PCR 増幅断片の精製

すべての PCR 増幅反応物は、その断片長に応じて、0.7 から 1.5% のゲル濃度のアガロースゲル電気泳動によって増幅を確認しながら、目的の大きさの DNA 断片を切り出し、MagExtractor kit によって精製した。

多数の DNA 断片切り出し作業において、作業中の UV 光による DNA 断片への損傷を防ぐために、ゲルからの DNA の切り出しには、一般的な EtBr 染色と UV トランスイルミネーターではなく、SyberGOLD (Molecular Probes, Inc.) 染色と、480 nm の青色ライトボックスを用いて検出し、オレンジのフィルターを通して観察した（なお、健康のためにも、大量の DNA 切り出し作業の作業中に UV 光に曝されることはよくないという判断もあった）。

6. pCA24N ベクターへの各 ORF のクローン化

ligation 反応と 大腸菌への形質転換も、96 ウェルフォーマットのプレートチューブで、以下の反応系で行った。

1μl of DNA fragments

1μl(0.1μg) of vector pCA24N DNA

1μl of A solution of ligation kit (TAKARA BIO, INC.).

1 6℃で 1 時間の保温後、氷中に移し、30μl のコンピテントセル²⁴（大腸菌AG1 株）を添加した。

上記の混合液を氷上で 30 分、42℃で 45 秒おいた後、全量を 0.27ml のSOC 培地（2% Bacto Tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose）をあらかじめ分注しておいた 1.5ml 試験管に移し、1 時間の回復培養を 30℃で行い、全量をCm添加済みLB プレートに、直径 2mmの滅菌済みガラスビーズとともにひろげ、プレートにスプレッドした（多数の形質転換作業を効率よくこなすため）。

30℃で一晩保温したプレートから得られた形質転換コロニーから、各 8 個ずつコロニーを選び、1.2ml の LB 培地の入った 96 ウェルフォーマットの DeepWell 型プレートに接種し 37℃で一晩培養した。プラスミド DNA の調製はマルチスクリーン（Millipore Co. Ltd.）を用いて行った。

7. 制限酵素 *Bgl*Ⅰ, *Bam*HI, *Sfi*Ⅰによる確認

プラスミド DNA 溶液 2μl を *Bgl*Ⅰ, *Bam*HI, *Sfi*Ⅰ（TAKARA BIO INC.）の 3 種の制限酵素でそれぞれ消化することによって、その切断パターンから、インサート長が目的断片のクローンと一致するか、その向きが転写方向と同じか逆

か、挿入断片が 1 個のみか（並列に複数が入入されてないか）を検証した。

8. GFP 発現コロニーの観察

GFP を発現している大腸菌コロニーの観察には、エムエス機器製の青色ライトボックスによる 480nm の光の照射とオレンジフィルターを使用した。

9. クロンのシーケンス解析

制限酵素切断によって確認されたクロンの各インサートの配列を確認するために、以下の配列のプライマーペア、

pCA-F primer : 5' - GGCGTATCACGAGGCCCTTTTCGTCTTCACC - 3'

pCA-R3 primer : 5' - TTGCATCACCTTCACCCTCTCCACTGACAG - 3'

で PCR 増幅されたインサート領域の DNA を、ダイレクトにシーケンス読取を行った。読み取りには、以下のプライマーを使用した。

F-CA primer : 5' - CATTAAAGAGGAGAAATTAACATGAGAGG - 3'

R-CA primer : 5' - CATCTAATTCAACAAGAATTGGGACAACATC - 3'

シーケンス解析は Big-dye terminator kit version 1 (Applied BioSystems Inc.)で反応し、3700 シーケンサー(Applied BioSystems Inc.製)を使用した。

10. 融合タンパク質発現の確認と精製

各クロン大腸菌を、Cm 入り LB 培地で、30℃で培養し、濁度(OD600)が 0.3 に達したところで、終濃度 0.1mM となるように IPTG を添加し、さらに 2 時間の培養を続けた。

培養液から遠心分離(10000 rpm x 3 min at 4°C)により菌体を集め、400μl の冷やした buffer-I(50 mM Sodium phosphate [pH7.0], 200 mM NaCl, Proteinase inhibitor [Hoffmann-La Roche Ltd])に懸濁した。超音波処理 (5 x 5 秒, level 3, Astrason ultrasonic processor) により細胞を破壊し、遠心上清 (16000 rpm x 15 min)を粗抽出液とした。この 30 μlを Nickel (Ni²⁺)-column (QIAGEN, Inc) またはTALONレジン (Clontech Inc.) にかけて、製品添付のマニュアルにしたがって精製した。検体を通したカラムは、1 ml のbuffer-II (Buffer-I contains 20mM imidazole, 0.05% n-octyl-glucoside [Pierce Biotechnology, Inc]) で3回洗浄し、buffer-III (50mM Tris-HCL [pH 6.8], 2% SDS, 0.1% bromophenol blue, 10% glycerol, 100mM DTT, 6M Urea, 250mM imidazole)で抽出した。抽出タンパク質検体は、7.5%から 15 % SDS-PAGEの濃度勾配SDS-PAGEゲル(BIOCRAFT Co.,Ltd.)により分離し、Coomasie Brilliant Blue(Wako Pure Chemical Industries, Ltd.)により染色観察した。

11. GFP 局在の観察

細胞内で発現している GFP 融合タンパク質の観察には、位相差蛍光顕微鏡 Axioscope (Carl Zeiss 製) を用いた。

12. IPTG 誘導による生育阻害効果の評価

NotI 断片を除く方法で GFP 部分を取り除いたプラスミドを持つ各大腸菌を、LB 培地を用いて 37°Cで培養し、O.D.(600nm)の値が約 0.5 程度に達した時点で IPTG を終濃度 0.1mM 添加し、その 2 時間後の O.D.値を測定した。

第3章 結果と考察

1. 大腸菌全遺伝子のクローニングと発現用のためのプラスミド pCA24N の構築

ORF領域のクローニングのために、有用な特徴をいくつか持つプラスミドベクターを設計構築した。**Fig.2**に、このプラスミドpCA24Nの構築の履歴を示した。

プラスミド pCA24N の特徴は、以下の通りである。

- 1) Cm 耐性遺伝子と高コピー複製起点 (pUC 系) を持つ。
- 2) IPTGによる発現誘導が可能なプロモーターP_{T5}-*lac*を持つ。
- 3) *lacI^q*遺伝子 (*lac*レプレッサー過剰発現変異遺伝子) をシスにコードしている。
- 4) N 末側に His タグを持つ融合タンパク質を合成する。
- 5) C 末側に GFP が付いた融合タンパク質を合成する。
- 6) 正しい方向にクローニングされたときのみ、ORF 領域の両側に *Sfi*I 切断配列ができる。
- 7) *Not*I で切断することにより GFP 領域を取り除くことができる。

GFPとの融合タンパク質は、細胞内の局在性の解析のためであり、またHisタグは融合タンパク質の精製を簡便に行い、生化学的解析をハイスループットに進めるためである。そのため、このGFPにはできるだけ明るく光り、かつ大腸菌の一般的培養温度である 37°Cにおいても良好に光るものとして、伊藤らが改良したGFPuv4 を選んだ²⁵。

さらに、このプラスミド構造のもうひとつの工夫点は、精製などのためのタグタンパク質に方向性と読み枠を考慮して ORF をクローニングするとともに、

一旦クローニングされた **ORF** 断片を一種類の制限酵素で切り出して、他の目的（新しい機能未知遺伝子の解析にあたって、その機能解析にはそれぞれの目的に応じたベクターへクローニングを行うことが必要である）のベクターにも、方向性と読み枠を考慮して容易にサブクローニングできるようにすることであり、それには、制限酵素 *Sfi*I の特徴を利用した（後述）。

また、薬剤耐性マーカーには、一般的な発現ベクターや機能解析ベクターが **Amp** 耐性であるものが多いことなど、サブクローニングの利便性を考慮し、**Amp** 耐性ではなく **Cm** 耐性を採用した。

クローン化遺伝子の転写プロモーターについては、発現制御が厳しくコントロールできるようにするため、**T5** ファージのプロモーター配列と *lac* 遺伝子のオペレーター配列を持ち (**P_{T5}-lac**)、同じプラスミド上に *lacI* 遺伝子をコードさせ、発現を **IPTG** で誘導可能にした。また下流には *lambda t0* の転写ターミネーター配列を持たせた。

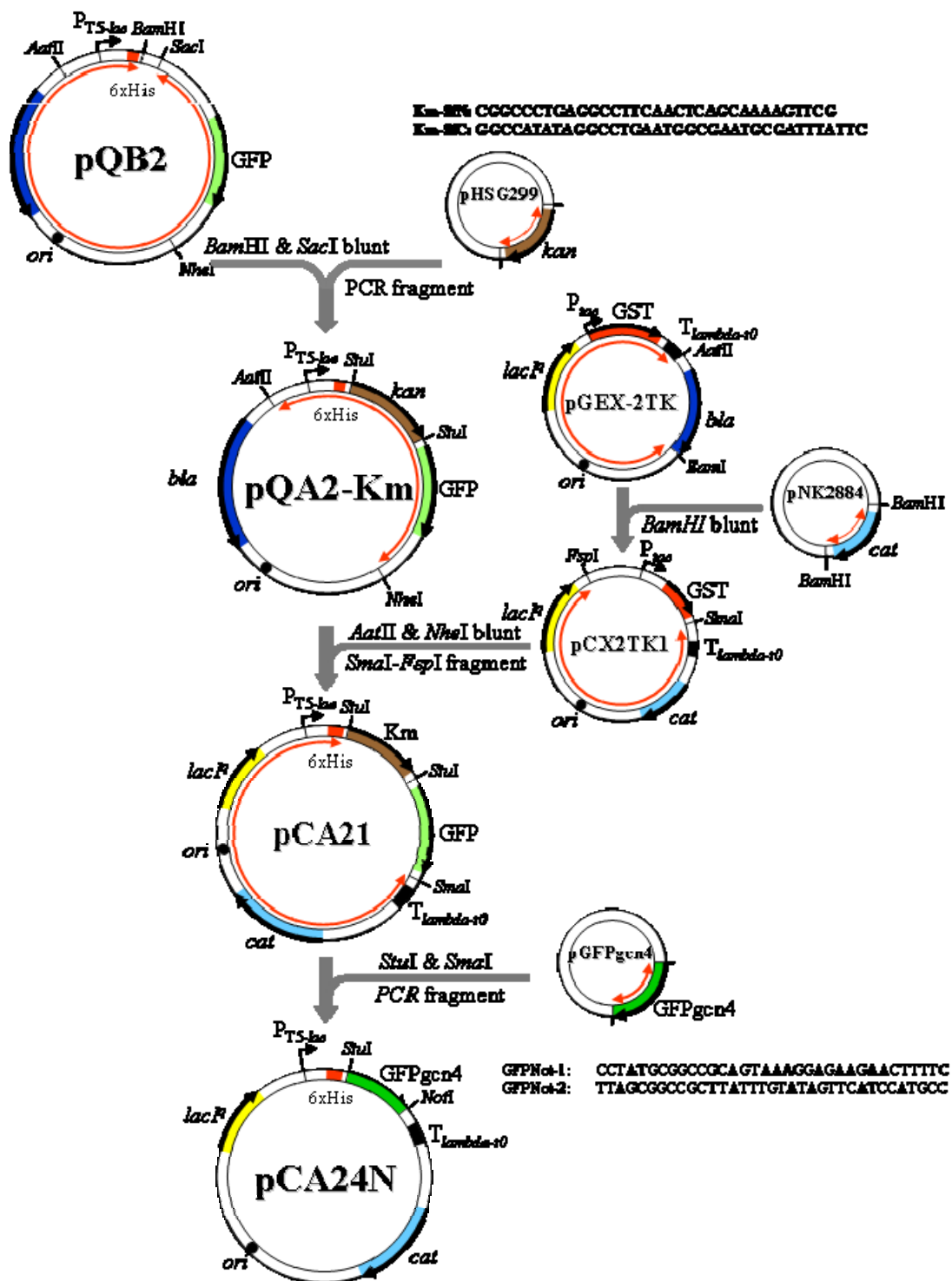


Fig.2 クローニングベクターの構築

プラスミドpQB2のBamHI-SacI断片4.1kbの切断末端を平滑化し、プラスミドpHSG299からPCR増幅したKanamycin耐性遺伝子の断片と連結したプラスミドpQA2-Kmを作製した。PCRに用いたプライマー配列は、

5' -CGGCCCTGAGGCCTTCAACTCAGCAAAAGTTTCG-3' と

5' -GGCCATATAGGCCTGAATGGCGAATGCGATTTATTC-3' である。

一方、pGEX-2TKの*Aat*II-*Eam*I断片の末端を平滑化したものから*lac*I遺伝子と転写ターミネーターのDNAを用意し、プラスミドpNK2884の*Bam*HI消化断片の平滑末端化したDNA上のChloramphenicol耐性遺伝子のDNAと連結して作製したpCX2TK1を作製し、これから調製した*Sma*I-*Fsp*I断片と、プラスミドpQA2-Kmから調製した*Aat*IIと*Nhe*I消化後の末端平滑化断片を連結してプラスミドpCA21を構築した。pCA21上のGFPを、より蛍光の明るい変異体GFPであるGFPuv4の遺伝子に置き換えるため、pCA21の*Stu*I-*Sma*I断片と、pGFPgcn4からPCRを用いて増幅したGFPコード断片を連結して、pCA24Nを構築した。GFPuv4の増幅に使用したオリゴDNAの配列は、

5' -CCTATGCGGCCGCGCAGTAAAGGAGAAGAACTTTTC-3' および、

5' -TTAGCGGCCGCTTATTTGTATAGTTCATCCATGCC-3' で、末端に*Not*I認識配列を持たせておき、ORFクローン化後にGFP領域を取り除くことが容易になるようにしておいた。このように構築されたpCA24Nの複製起点は、pGEX-2TK由来であり、Hisタグの部分はpQE31(pQB2の元になったプラスミド) 由来である。

このプラスミドpCA24Nの制限酵素*Stu*Iの切断部位(AGGCCTの真ん中で平滑末端を生成)に、目的とする方向にORF断片が挿入されたとき、IPTG誘導性プロモーターであるP_{T5}-*lac*のプロモーターからの転写の下、Hisタグ配列と翻訳フレームを合わせてORF配列とGFPの融合タンパク質が発現できる構造とした。

Fig.3は、転写方向と一致する方向にクローン化した後のORF両側の翻訳配列を示したものである。このように、HisタグからORFの間には7残基のアミノ酸がスペーサーとなり、ORFからGFPの間には5残基のアミノ酸がスペーサーとなっている構造をとっている。

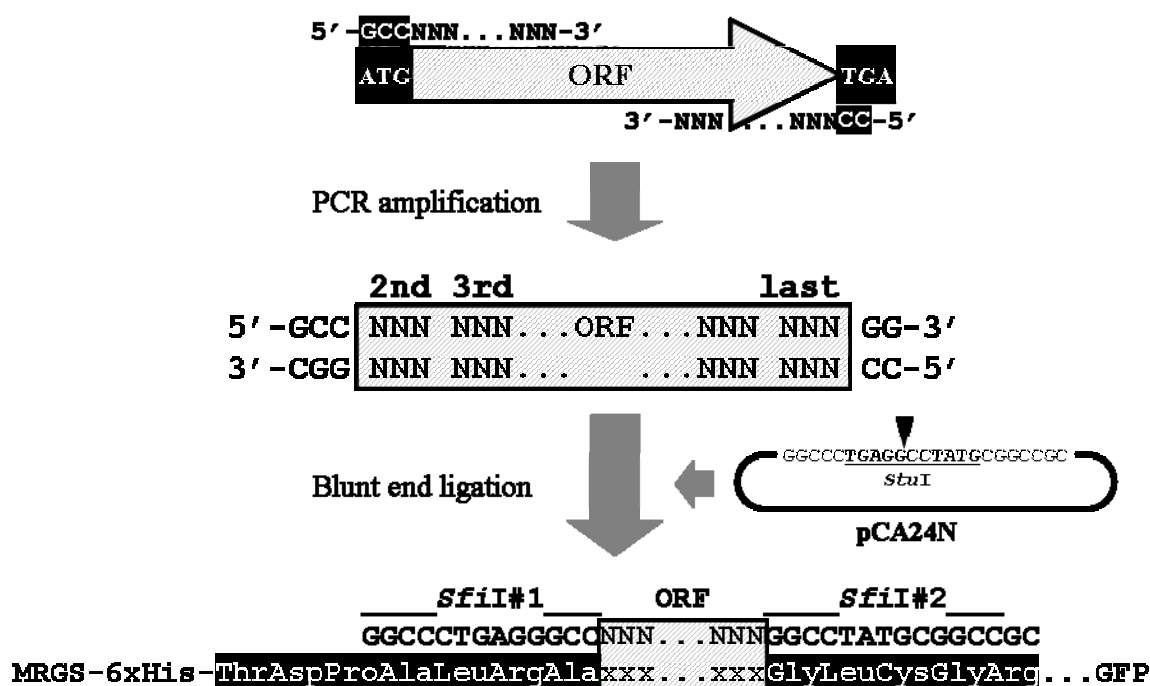


Fig.3 クローン化した ORF の前後に構築されるアミノ酸配列

矢印は、クローン化対象のORFと増幅に用いたPCRプライマーを示している。下に示した塩基配列とアミノ酸配列は、クローン化後のORF前後の配列である。その産物であるタンパク質のN末側には6個のHistidineと7個のアミノ酸、C末側には5個のアミノ酸がそれぞれ付加され、その後にGFPが続く形になる。すなはち、

MetArgGlySer-6xHis-ThrAspProAlaLeuArgAlaXXX...XXXGlyLeuCysGlyArg...GFP

という構造になる（XXX...XXXの部分はORFの2番目のアミノ酸から最後のアミノ酸までを表す）。

Fig.4は、この *StuI* 部位へPCR増幅断片がクローン化された際の、正逆両方向のケースでの境界部分の配列を示したものであるが、転写翻訳方向と一致するような目的の方向に断片が挿入された場合のみ *SfiI* 切断可能な配列が完成し、逆方向での挿入の場合では、*SfiI* 配列にはならないことを示している。

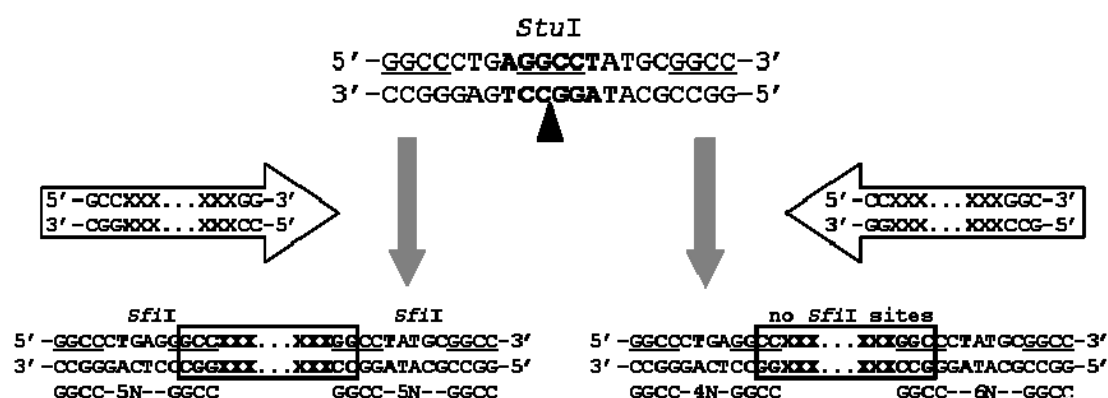


Fig.4 目的方向へのクローン化によってのみ実現される *SfiI* 認識配列の構築

N側に3塩基(GCC)、C側に2塩基(GG)を付加したプライマーでORFを増幅し、pCA24Nの *StuI* サイトに対して、そのORF断片が転写翻訳方向と一致する向きに挿入された場合のみ、ORFの前後に *SfiI* 認識配列が構築される（左のケース）。逆向きに挿入が起きた場合（右のケース）は、*SfiI* 配列にはならない。

一方、このプラスミド pCA24N 自身は、このままでは His タグと GFP の翻訳フレームが合っていないため、GFP を発現できない。そこで、*StuI* 部位に 8mer の *Bam*HI リンカーオリゴ（配列"CGGATCCG"）を挿入し、ORF 部分なしで「His タグつき GFP」のみが発現するプラスミド pCA24N+B を対照実験のプラスミドとして作製した。この場合には、"GCC-(ORF)-GG"となる PCR 断片の配列が"CGGATCCG"となって、His タグと GFP の間に 13 個のアミノ酸配列を持つペプチドが発現されるようになっている。

実際にこのプラスミドを持たせた大腸菌で IPTG 誘導をかけた場合に His タグつき GFP が発現されていることを、プレート培養によるコロニーの色の観察、液体培養による菌体の蛍光色の観察、His タグ抗体および GFP 抗体によるウェスタンブロッティング、GFP 蛍光の測定のほか、後に示すように、アフィニティレジンによる His タグつき GFP の精製、蛍光顕微鏡による GFP 蛍光の観察で確かめ、このベクターにおける、IPTG 誘導性のタンパク質発現（と非誘導時

の発現の程度)、および His タグと GFP の機能性を確認した。

2. ORF の PCR 増幅

クローン化対象としたのは、IS 因子を除く、大腸菌 K12 株の全 ORF4276 個であったが、実際に目論見どおりのクローン化に成功したのは、4,267 個であった。4270 個については PCR による増幅まで成功したが、6 個 (*eaeH* (7,104bp), *yheB* (2,694bp), *yhiH* (2,685bp), *yagF* (1,968bp), *ydbD* (2,313bp) , *btuB* (1845bp)) については目的の ORF 断片を調製できなかった。その場合は、ORF 内部の部分的な配列のうち、N 末側のコード領域約 330 アミノ酸の部分のみを増幅するように C 末側のプライマーを再設計し、再度増幅させて、3 個については部分断片を取得し、同様にクローン化した。これらの部分的な ORF を持つクローンはタンパク質として機能解析には不適だが、DNA マイクロアレイ用の鋳型としては十分に利用可能である。

3. pCA24N へのクローニング

クローン化は、このようにして PCR 増幅された ORF 断片すべてをアガロースゲル電気泳動により分離精製後、末端を脱リン酸化した pCA24N の *StuI* 切断 DNA とライゲーションし、大腸菌 AG1 を形質転換して行った。

30℃で一晩保温したプレートから得られた形質転換コロニーは、平均的に約 500 個であったが、多いものでは 1,000 個以上、少ないものは数個と、かなりのばらつきが見られたものの、一回に得られたコロニーの数が 10 個以下だったものはわずかで、全体の 1%以下であった。このように形質転換コロニーが少ないものの中にはあり、これらは何らかのクローン化バイアスがかかっている可能性もないとは言えない。しかし、今回のクローン作製においては、平滑末端ど

うしの連結によるため、逆向きに挿入されてくる断片をあえて排除しないようにした。それは、クローン化遺伝子の性質がクローン化バイアスとして表現されるケースでも、逆向きの挿入であれば、そのようなバイアスもかかりにくく、クローン化されてくる可能性が高いと考えたからである。したがって、たとえ作業の上で形質転換の効率にばらつきが見られても、逆向きのクローンが取得される数との相対比較（逆向きクローンが 1 種のコントロール実験となる）によって、得られたクローンに対するバイアスの程度を推察することができると考えた。実際、ほとんどの遺伝子について逆向きのクローンが得られており、その時の正方向と逆方向のクローンが取得されてくる頻度に大きな偏りがあるケースは稀であった。また、コロニーの形態で見る限り、IPTG による誘導をかけなければ、そのクローン化方向の区別は全くつかない状態であった。以上のようなことから、得られたクローンに変異などが入っているといったクローン化バイアスがかかっている可能性はきわめて低いと考えられ、それは同時に、機能を保持しているタンパク質の遺伝子としてクローン化されている可能性が高いことを示唆するものである。このことは His タグと GFP との融合による機能喪失が起きていることを否定できるものではないので、クローンの利用に際してはその点を留意しておく必要はあるものの、クローン化バイアスによる変異等の懸念は非常に小さいと考えてよい。

クローンの確認作業としては、各形質転換クローンから 8 個ずつを選び、プラスミド調製用の培養と同時に、IPTG を含む LB プレートに植菌して、各形質転換コロニーの IPTG 感受性、およびコロニーの色を観察にすることによる GFP の発現を調べた。すると、8 個の中のいくつかには生育できないクローンが見られたケースがあるとともに、GFP が持つ蛍光色を示すコロニーが見られた。Fig.5 に示したのはその一例であるが、実際、IPTG 添加 LB プレートでのコロニーに、

480nmの波長の青い光を当てて、はっきりと緑色の蛍光シグナルが見られたクローンは 1,935 個であった。残りのORFクローンでは、蛍光が見られなかったが、そのうちの 1,003 個は膜タンパク質と推定され、975 個は生育自体が見られなかったものである。

多くの場合、このように生育が IPTG 感受性を示すか、または GFP の蛍光色コロニーが見られたものは、目的の断片が目的の方向にクローン化されたものであった。GFP の蛍光の明るさは、各 ORF によって差が見られ、また、IPTG 感受性が強く出た場合には、コロニーの形成そのものがほとんど見られなかったため、GFP の蛍光を観察することができない場合も多かった。しかし、この GFP の蛍光は ORF が期待されるフレームで読まれていることを示唆してくれるものであり、これを指標にすることにより、多くのケースで正しいクローンの選択に役に立った。GFP 融合タンパク質として発現させる最大の目的は、各タンパク質の細胞内での局在部位を調べ上げることであったが、それだけでなく、このように GFP が発現されていることを指標に、クローン化した ORF 領域の翻訳枠が期待とずれていないかの判断も可能であった。

制限酵素切断によるプラスミドの構造確認のほかに、以上のような「コロニー間の表現型の同一性」をもとに、各 ORF につき 4 つ以上の独立したクローンを選択し、また、PCR エラーによる変異が入ったクローンのみの選択保管を避けるため、それらの DNA をプールしてひとつの一次アーカイブクローンとした。

PCRで目論見どおりの長さの断片が増幅できた中でも、クローン化には成功しなかったものが 3 個 (*yhcQ(aaeA)*, *yibP(envC)*, *aidB*) あった。これらは 16 個以上の独立なコロニーを調べたが、すべてが P_{T5-lac} プロモーターの転写方向とは反対向きの挿入クローンであった。

結果的に、4,276 個の ORF を対象として、PCR で全長が増幅できた 4270

個（ORF断片が2個以上並列にクローン化されたもの2個（*pheM*, *thrL*）、前述の反対向きしかクローン化できなかったもの3個を含む）について全長をクローン化し、全長の増幅がかなわなかった6個については、そのうちの3個（*eaeH* homolog, *yheB*, *yhiH*）については、N末側の330アミノ酸分を部分的にクローン化し、これら全体を「ORF クローンのアーカイブス」として、“ASKA library (A Complete Set of *E. coli* K-12 ORF Archives)” と命名した。結局、全くクローン化断片を得られなかったのは、3個（*yagF*, *btuB*, *ydbD*）であった。

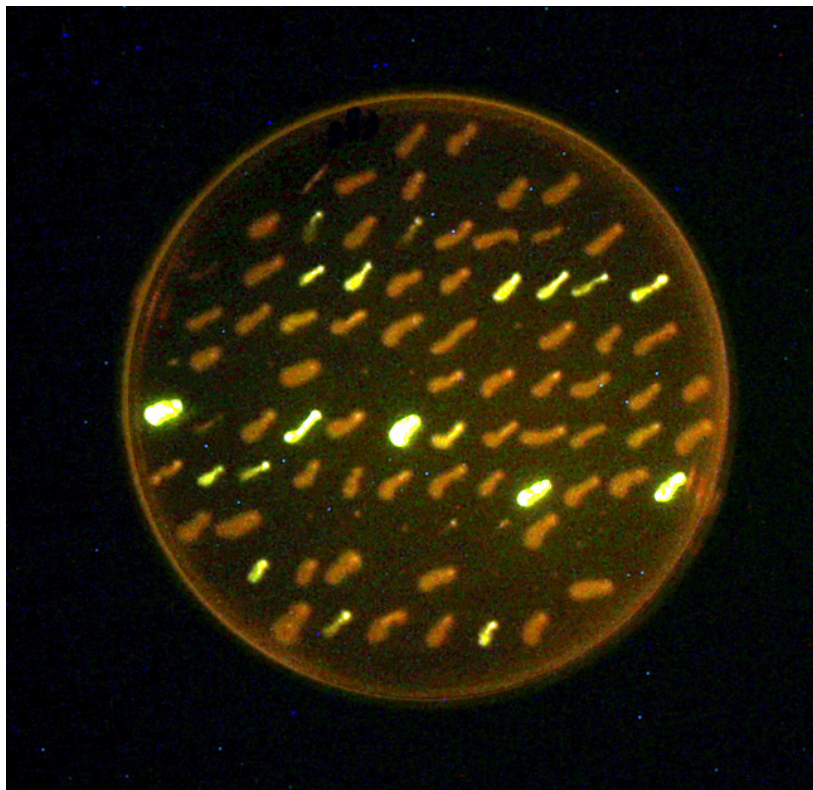


Fig.5 コロニーによるGFP発現と生育（IPTG 感受性）の確認例

形質転換で得られた候補クローンを、各8個ずつ拾い、0.1mM の IPTG を含む LB 培地のプレートに植菌して、37℃で培養し、480nm の青色ライトボックスで光をあて、オレンジ色のフィルターを通して見られる蛍光を観察した。緑色の蛍光色に光るコロニーや生育の見られない（IPTG 感受性の）コロニーが見えている。

4. 産生された GFP 融合タンパク質の細胞毒性。

大腸菌では、マルチコピープラスミドへのクローン化が困難な遺伝子や DNA 断片も知られているが、“ASKA library”でクローン化された ORF4,267 個には、クローン化が極めて困難とされてきた遺伝子も含まれている。

そのようなクローンについては、本来の機能を喪失させるような変異が入って、細胞への毒性を失ったためにクローン化できた可能性も否定できない。しかし、前述のとおり、逆向きクローンが取得できていて、その頻度において、正方向クローンの取得頻度との差があまり認められないケースがほとんどであったということは、発現の抑制が非常に厳密にできている系であるためとも考えられる。中には、正方向のクローンが得られず、逆向きのクローンしか得ることができなかった遺伝子もあり、IPTG 無添加の状態でもわずかに融合タンパク質が漏れて発現し、それが細胞に毒性を与えたと思われるケースがあるが、そのようなケースは、前述のとおり、わずか 3 例であった。

一方、先に述べたように、クローンを選択する過程での IPTG 存在下での増殖チェックによって、各クローンプラスミドを持つ菌の、全体の 51% は GFP 融合タンパク質の発現によりコロニー形成の著しい阻害がみられ、28% の中にも若干の生育阻害が観察されるものがあり、GFP 融合タンパク質の過剰発現が細胞に毒性を与えることが非常に多いことがわかった (Fig.6)。このように IPTG による誘導発現が細胞に致死的な影響を与える一方で、誘導がなければ生育にはほとんど影響を与えないという点は、このクローンセットの重要な特徴であり、逆に、毒性の強いタンパク質でも、発現抑制が効いているためにクローン化に成功していることを示唆するものである。

なお、クローン化の過程で、ORF断片が目論見の発現方向とは逆向きにクロ

ーン化されたものが得られてきたが、それらの中にも、IPTG含有培地で増殖阻害の起きるものがいくつかあった。その理由としては、これらが増殖に必須の遺伝子であり、そのアンチセンスRNAの転写によってゲノム上に元から存在する遺伝子の発現が抑制された可能性なども考えられるが、そのような遺伝子に対応するKeioコレクション（大腸菌K12株の全遺伝子の欠失株のコレクション）の欠失株は得られているので²⁶、何らかのペプチドが翻訳発現されて、その毒性が出ているのかもしれない、さらに詳細な解析が必要である。

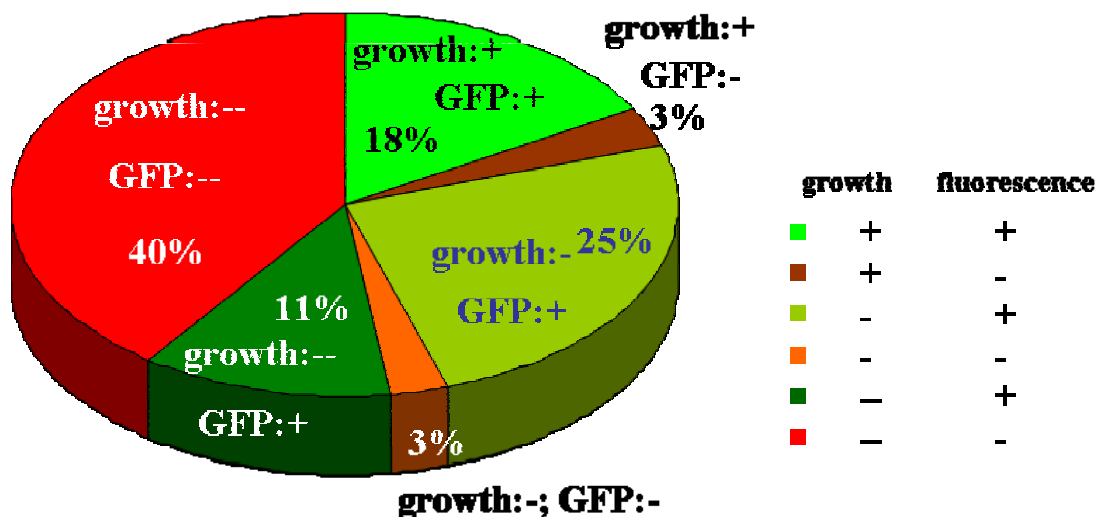


Fig.6 GFP融合タンパク質の誘導発現による生育への影響

IPTG 含有プレートで観察された、コロニー形成能とそのコロニーの蛍光色の程度の観察結果による、各表現型を示すクローンの割合をまとめたパイグラフを示す。コロニー形成については、ほとんど生育できなかったものは”--”、生育の阻害が見られているが、コロニーの形成があったものは”-”、それ以外のほぼ正常に生育が見られたものは、”+”で表現した。一方、GFP 発現による蛍光色は、生育がほとんど見られなくても、光の見えたものは”+”で、まったく光の見えなかったものは”-”で表した。

5.さまざまなリソースとしての利用

このような多目的な発現プラスミドの設計においては、以下の網羅的および個別機能解析実験への適用を目的とした。

- ・クローン化断片を利用した DNA マイクロアレイ作製のための鋳型調製
- ・多様な解析に供するための容易なサブクローン化
- ・変異株に対する相補試験による変異遺伝子の同定
- ・タンパク質間相互作用の網羅的な検出による相互作用地図の作成
- ・誘導発現タンパク質の精製による生化学的解析と高次構造決定
- ・タンパク質の細胞内局在性解析
- ・大腸菌アルカリフォスファターゼ **PhoA** 融合タンパク質を利用した膜タンパク質のトポロジー解析

以上の計画を基に、多様な機能解析ツールとしてクローンセットを設計し、提供した。以下に、実際に利用された例などをあげて、今後のさらなる活用の可能性について考察する。

1) DNA マイクロアレイの鋳型への利用

大腸菌 K12 株の遺伝子を網羅するマイクロアレイの作製に、このクローンセットが実際に利用されて、すでに多くの成果が発表されている^{27,28,29,30,31}。

現在タカラバイオから市販されている大腸菌の DNA マイクロアレイは、このクローンセットが鋳型になっている。

2) *SfiI* によるサブクローン化とさまざまな受け皿ベクターの利用

制限酵素 *SfiI* で ORF 配列断片が切り出されたときの末端の塩基配列の形状は各切断配列特異的な 3 個の塩基の 3'-突出末端（N 末端側配列：TCA、C 末端側配列：ATG）であるため、その特異的な「のりしろ配列」を活用して、様々な目的に応じたベクターに特異的な *SfiI* 認識配列を持たせ、すべての ORF 断片を方向性と読み枠を合わせてサブクローニングすることが可能である。

ただし、ORF 内部に *SfiI* 切断部位を持つようなものは使えない場合もある

ので、必要に応じて個別に対応しなければならない場合もあるが、その可能性を持つ ORF は、大腸菌において 28 個しか存在しない（しかも、そのうちの 4 個は *dcm* メチラーゼによるメチル化の影響を受けるので切断されない）。また、もし内部に *SfiI* 部位があっても、それぞれ特異的な「のりしろ」の配列を持つので、サブクローン化に際しての懸念はないと考えられる。

この特性を利用して、酵母Two-hybrid systemなどへの容易な移行や、他のタグの位置をN末側や、C末側につけ換えること、あるいは、GFPやHisタグを持たない、Nativeな遺伝子産物に限りなく近い産物の生産も可能である。そのようなサブクローン化用のベクターのクローニングサイトに適応する配列例を Fig.7 に示す。

Archives (pCA24N)

```
GGCCCTGAGGGCCNNN...NNNGGCCTATGCGGCCGC
6xHis...LeuArgAlaXXX...XXXGlyLeuCysGlyArg...GFP
```

Non-fusion (Ex. pNTR30-SD)

```
AGGAGGTGGCCATGAGGGCCNNN...NNNGGCCTATGAGGCC
SD          MetArgAlaXXX...XXXGlyLeuTer
```

N-terminal fusion (Ex. Two-Hybrid vectors)

```
GGCCCTGAGGGCCNNN...NNNGGCCTATGAGGCC
Gal4DB...LeuArgAlaXXX...XXXGlyLeuTer
Gal4AD
```

Fig.7 多様な目的に供するためのサブクローニング用ベクターのクローニングサイト配列

アーカイブクローンから *SfiI* によって切り出された断片をサブクローン化する際に、その受け手側のベクターに必要なクローニングサイトの配列とサブクローン化された場合の付加アミノ酸配列を示す。“Non-fusion”の場合は、開始コドンを除けば、前後に 2 個ずつのアミノ酸の付加しかないものができる。

たとえば、Two-hybrid 用のベクターなど目的に応じたベクターに、このような特異的な *Sfi*I 部位を持たせておけば、方向性と読み枠を合わせてサブクローニングすることができる。

つまり、

N末端側 : 5' -GGCCATGAGGGCC...

C末端側 : ...GGCCTATGAGGCC-3'

という配列を持たせれば、開始コドン ATG と終止コドン TGA が付加される。

一方、

N末端側 : 5' -GGCCCTGAGGGCC...

なら、開始コドンは復活せず、

C末端側 : ...GGCCTATGIGGCC-3'

なら、終止コドンは復活しない。

これらを使えば、タグタンパク質との融合の位置を、N 末端側でも C 末端側でも自由に選ぶことができる。また、このような *Sfi*I 部位のカセット DNA 配列を準備しておけば、様々なベクターに容易にこれらの配列が導入できる。たとえば、融合タンパク質としない場合は、余分なアミノ酸が、前後それぞれにわずか 2 個 (N 末端側 : Arg-Ala, C 末端側 : Gly-Leu) しか付加されないようにできる。

西村らとの共同研究では、実際にこの系を活用して、接合性プラスミドにすべての ORF のクローン化を行い、数多くの温度感受性細胞分裂関連変異遺伝子の同定が行われた³²。この結果は、"ASKA library" の ORF が機能を持っていることを示すものである。

このような変異株の相補遺伝子探索のソースとして、あるいはある特定の酵

素活性を持つタンパク質のスクリーニングなどの解析にも、このクローンセットは利用可能である。Yakuninらのグループは、このクローンを利用して、Hypotheticalなタンパク質であったYfbRが、*p*-nitrophenyl phosphateを基質とするフォスファターゼ活性を有する酵素であることを見つけ、そこからさらに詳細な解析を進め、Co²⁺依存的な 5'-deoxyribonucleotidase活性を持つことを示し、それが細胞内のdNTP蓄積量の調節に関与していることを示唆する結果を得ている³³。このタンパク質は、真正細菌だけでなく、始原菌からヒトを含む真核生物にまで相同遺伝子が見出されるもので、HD ドメインファミリーと呼ばれるスーパーファミリータンパク質の一種である。そこから、金属イオン依存的なphosphohydrolase活性を持つと推定されていたのだが、このようにゲノム情報から推定された機能を検証する素材として、ASKA libraryが便利なツールであることが認められたよい例である。

そのほか、サブクローン化の受け手側には、いろいろな機能解析用のベクターを想定でき、たとえば酵母Two-hybrid系によるタンパク質間相互作用の網羅的解析や³⁴、アルカリフォスファターゼ(PhoA)との融合タンパク質作出の系による膜タンパク質の配向性を調べる系³⁵、あるいはいろいろなタグ配列との融合タンパク質作出で精製を容易にする系など、様々な応用の可能性が考えられる。

3) *NotI*によるGFP領域除去

プラスミド pCA24N の GFP の領域の前後には *NotI* 切断配列があり、プラスミドを *NotI* で切断後、自己環状化させると、GFP 領域の除去が可能である。

ただし、この場合、C-末側の塩基配列は、

NNN GGC CTA TGC GGC CGC TAA
Gly Leu Cys Gly Arg

となるので、**Fig.8** に示すように、GFP 領域を取り除いたあとにできるタンパ

ク質の C 末端側の翻訳配列は、ORF の後ろに、余分なアミノ酸が 5 個 (GlyLeuCysGlyArg) 付加されることになる。なお、すでにこのような GFP を取り除いた全クローンのセットを作製済みである。

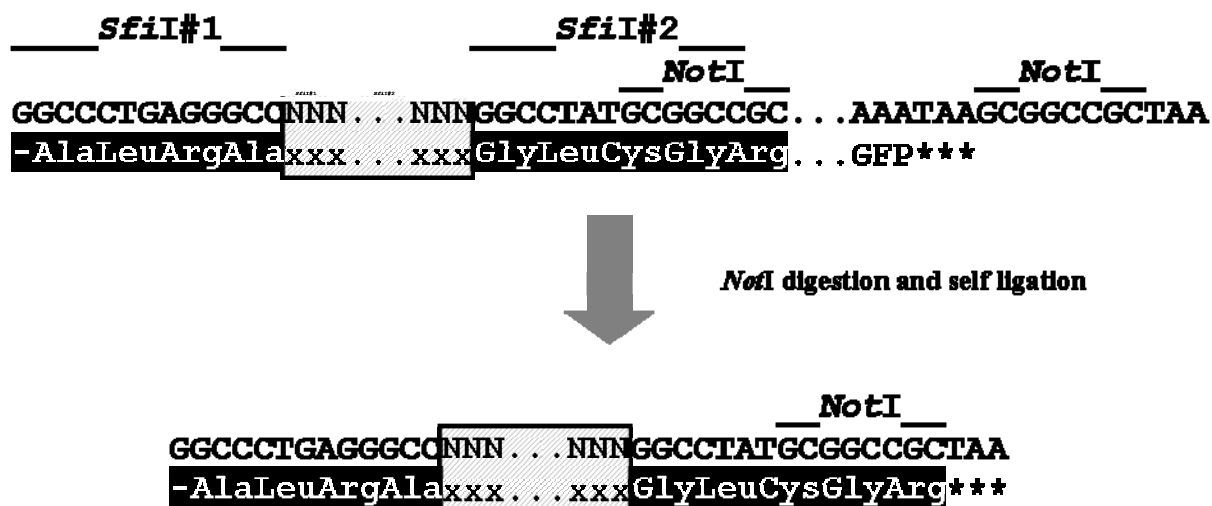


Fig.8 *NotI*切断によるGFP領域の簡便な除去方法

GFP領域の前後に*NotI*認識配列が配置されているので、各クローンプラスミドを*NotI*で切断後、ライゲーションにより自己環状化してできるプラスミドからは、GFPの除かれたタンパク質が発現される。

4) タグつきタンパク質のアフィニティ精製

ASKA library の各クローンからは、His タグ付きのタンパク質を IPTG によって発現誘導できる。各クローンから誘導発現されるタンパク質は N 末端側に His-tag を持っている構造なので、読み枠が合って全長のタンパク質ができていれば、GFP 融合タンパク質は NTA 樹脂あるいは TALON 樹脂でトラップされるはずである。逆に、読み枠が合っていない場合や途中から読まれている場合、あるいはプロセッシングを受ける場合などは、GFP タンパク質は、たとえ発現していても NTA 樹脂などでトラップされない。GFP をマーカーにこれを調べ

ば、プロセッシングを受ける分泌タンパク質の場合などの例外を除いて、予想されたORF全長を含む融合タンパク質ができているかどうかを調べることも可能かもしれない。

実際に、発現されたタンパク質を、そのタグを頼りに精製できるかどうか検証を行った事例を、**Fig.9**に示す。これはクローンプラスミドを*NotI*消化によってGFP部分を取り除いた方のクローンによる発現および精製の例であるが、Ni-(or Co-)NTA樹脂カラムを用いて、容易に精製タンパク質を調製することができている。このシステムでは、残念ながら、膜タンパク質の多くや、いくつかの細胞質タンパク質について、過剰発現時の生育の問題および可溶化の問題、カラムへの吸着の問題などで精製できないケースがあった。そのような場合については、タグの位置をC末側に変えたり、タグの種類を変えるなどが必要かと思われるが、前述のように、ベクター等の変更が容易にできるのがこの系のひとつの特徴である。

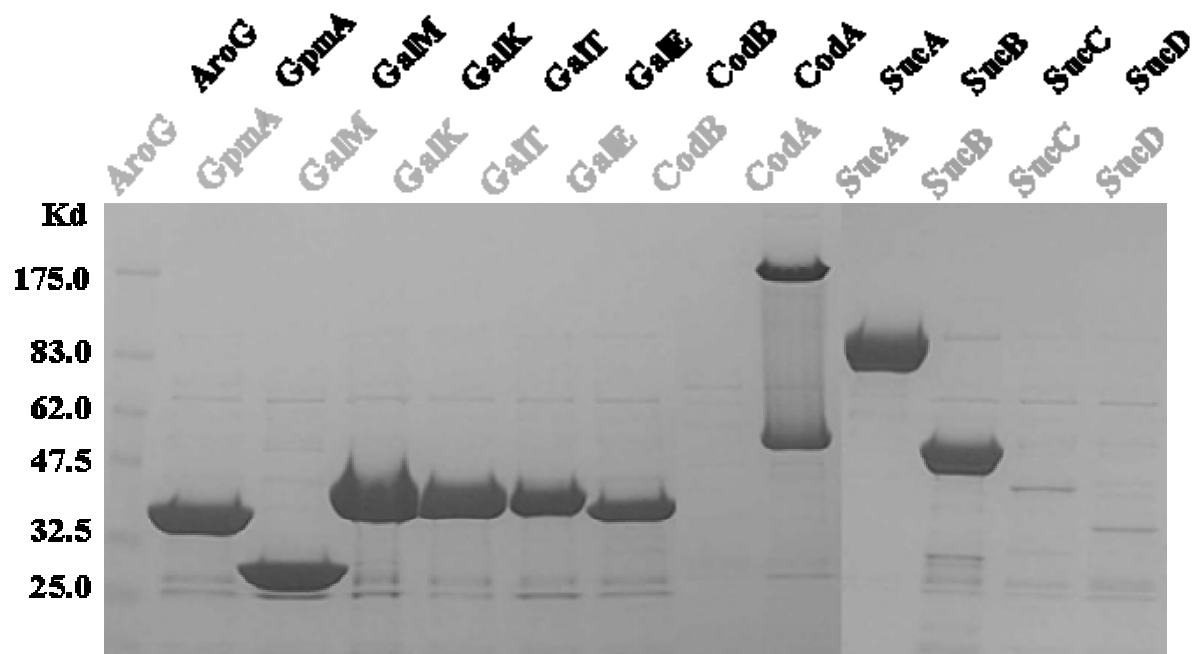


Fig.9 Ni-NTA樹脂カラムによるIPTG誘導タンパク質の精製

GFPを取り除いた形態のASKAライブラリの各クローンから過剰発現されたHistidineタ

グ付きタンパク質をNi-NTA樹脂カラムにより精製した。培養は3mlのスケールで、発現の誘導は0.1mM IPTG添加後2時間起こった。全溶出タンパク質を、SDS-PAGE後、Coomassie Brilliant Blueによって染色して分析した。いくつか収量が非常に低いタンパク質(CodB、SucC、およびSucD)もあった。

5) プルダウンアッセイによるタンパク質の相互作用解析

この精製手法の応用として、Ariffzzamanらは、野生型大腸菌W3110 にクローンプラスミド導入した株で、プルダウンアッセイによるタンパク質の相互作用を網羅的に調べ上げた³⁶。そこでは、Hisタグのついたタンパク質を誘導発現させた株の菌体ライゼートを、Ni-NTA樹脂カラムにアプライし、共精製されてくる精製物を質量分析 (Matrix-Assisted Laser Desorption Ionization Time Of Flight Mass Spectrometry : MALDI-TOF MS) にかけることにより、相互作用因子候補の網羅的同定を行った。大腸菌W3110 のゲノム配列更新によって追加更新されたORFクローンも含む 4,339 個をベイトタンパク質の発現クローンとしてこの解析を行い、精製に成功した 2,667 個のタンパク質のうちの 2,337 個に関して、共精製される相手が見出だされている。残念ながら、1,568 個のタンパク質に関しては、過剰発現あるいはアフィニティ精製ができなかったが、これらのうちの 928 個は膜タンパク質と推定されるもので、残りの細胞質タンパク質と考えられるものも、発現誘導時に細胞毒性が発揮されて菌体量の回収が足りなかったり、発現されたタンパク質が凝集してしまって非変性状態での解析が困難という状況であった(**Fig.10**)。

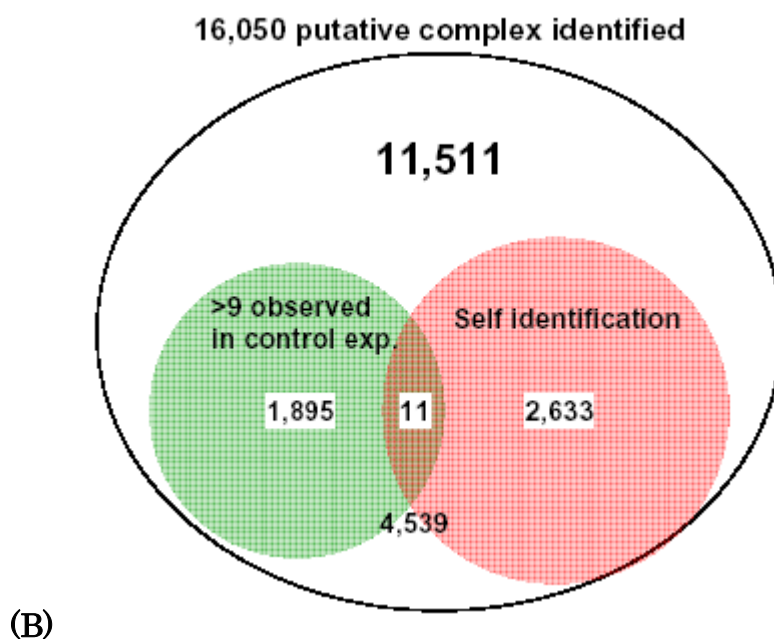
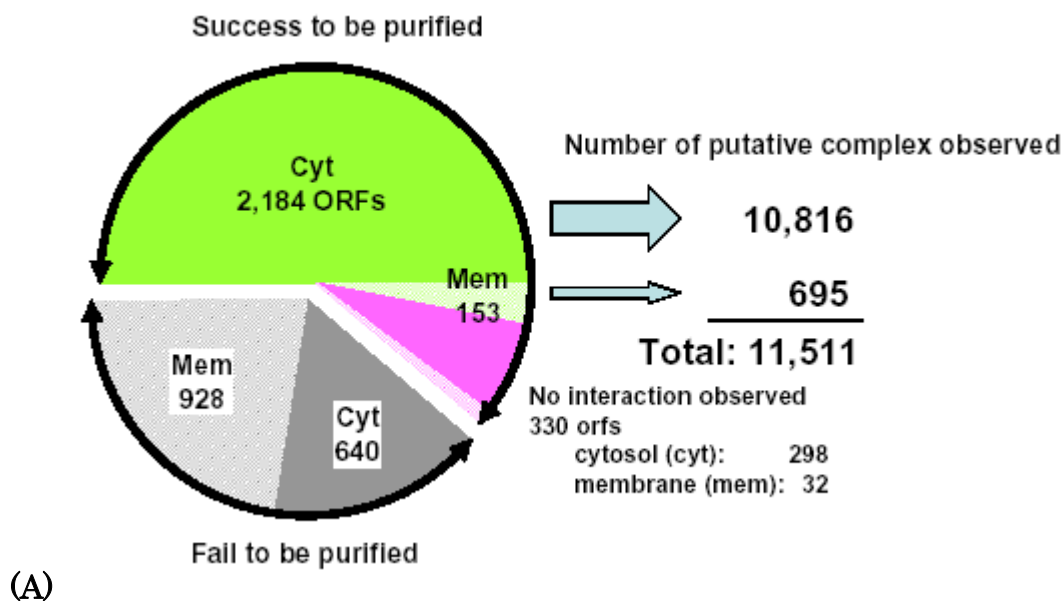


Fig.10 推定されるタンパク質タンパク質相互作用の統計値（文献42、Ariffzzaman et al. より引用）

(A) 左側の円の緑色の領域は、うまく精製された細胞質性のbaitタンパク質の数を表す。ピンクの領域はタンパク質間相互作用のパートナーを全く持っていなかったbaitタンパク質の数を表す。灰色の領域は、精製できなかったbaitタンパク質を示す。扉で示した領域(mem)は既知膜タンパク質もしくは膜タンパク質と推定されるbaitの数を示している。右側の数字は今回観測されたタンパク質間相互作用(PPIs)数を表す。

(B) 今回特定されたタンパク質間相互作用の総数は16,050個であった。すべての相互作用

の中の、左の小円(薄緑)は、対照実験に基づいてNi²⁺-NTAにアフィニティを持っているかもしれないと判断されるタンパク質相互作用の数を示している。右の円(薄赤)は、タンパク質が、baitとしたHisタグ付きタンパク質と同じものがpreyとして同定された数で、これらを合わせた4,539個の相互作用については、不確かなものとして除き、残りの11,511個のタンパク質相互作用をさらなる分析の必要なものとした。

このような網羅的なタンパク質間相互作用の解析では、擬陽性の相互作用との区別が問題となるが、コントロールプラスミドのみの場合に得られてくるものや、過剰発現したタンパク質自身の分解物などによるノイズを慎重に除いて、11,511 個という膨大な数の相互作用を見出だすことができた。このような成果は、従来の手法では容易には得ることができないものであり、遺伝子を「しらみつぶし」するような材料を用意したからこそ可能となったものである。

この結果の評価として、実験的に検証された相互作用のカatalogをまとめているデータベース、DIP database³⁷と比較してみると、その中で公開されている 447 個の相互作用 (Protein-Protein Interactions :PPIs) のうちの 372 個のタンパク質ペアの一方もしくは両方で、結果の一致が見られており、確かに信頼性の高いデータが取得できていると思われる。

大腸菌において、このような網羅的なタンパク質間相互作用の解析が行われた例としては、TAP-あるいはSPA-のタグを用いた手法による研究がある³⁸。彼らの系はASKA libraryの系とはかなり異なり、大腸菌の染色体から、もともとの発現プロモーターによる転写環境での解析を行っているので、より自然な状態に近い相互作用を見ているという主張は正しいかもしれないが、実際には、検出感度の点で限界があると考えられる。結果を比較してみると、彼らと共通で調べられたベイトタンパク質は 521 個あって、彼らの方法では、それに対する 5,030 個の相互作用が検出され、ASKA libraryの系では 3,088 個の相互作用が検出されているが、そこで共通していたのは 218 個 (それぞれの 4%と 7%)

しかなかった。しかし、このような差は、網羅的な相互作用解析の系では、しばしば見られており、たとえば酵母でもいくつかの系でタンパク質間相互作用の網羅的な解析がなされているが、それぞれで共通して検出されている相互作用の数は、この大腸菌での場合と同程度である³⁹。

また、このような網羅的な相互作用解析の実験では、そこで検出された相互作用の候補に対して、過去に蓄積されたさまざまな実験データも取り入れて検証しながら、得られた結果を評価していく必要があるが、重要な点は、生物の全体像を理解するのに、統一された条件の下、ある一定の実験手法で、均一な品質レベルで比較可能なデータを取得することであり、それこそが一種のシステムバイオロジーという新しい研究手法である。そのような観点で、この ASKA library という統一された材料を利用した実験は非常に有意義であると言えるだろう。

今回の相互作用候補を、KEGG パスウェイデータベース⁴⁰でまとめられている知見等を基にして機能分類を行い、意味のある相互作用の抽出を進め、「相互作用地図」を作成した。一例として、RNA polymerase を中心にした相互作用を見てみると (Fig.11)、コアエンザイムのサブユニットである α 、 β 、 β' (それぞれ RpoA, RpoB, RpoC) の相互作用を中心にして、それらと多種類ある σ 因子のほとんど (σ^s を除く) との相互作用を検出できており、その他の重要な転写制御因子、たとえば NusA, NusG, Rho といった転写終結に関与する因子^{41,42}が検出されている。加えて、RhlE (エネルギー依存的 RNA helicase)、VacB (3'-5' exonuclease) といった、いくつかの新規な相互作用因子も検出された。これらは、RNA の成熟に関与していることが示唆されている^{43,44}。また、大腸菌の Cold-shock protein である CspA⁴⁵ と RpoB および RpoC との相互作用が見えており、これまでの知見と合わせて考えれば、これは CspA が転写制御因子として作用することを示唆するものである。

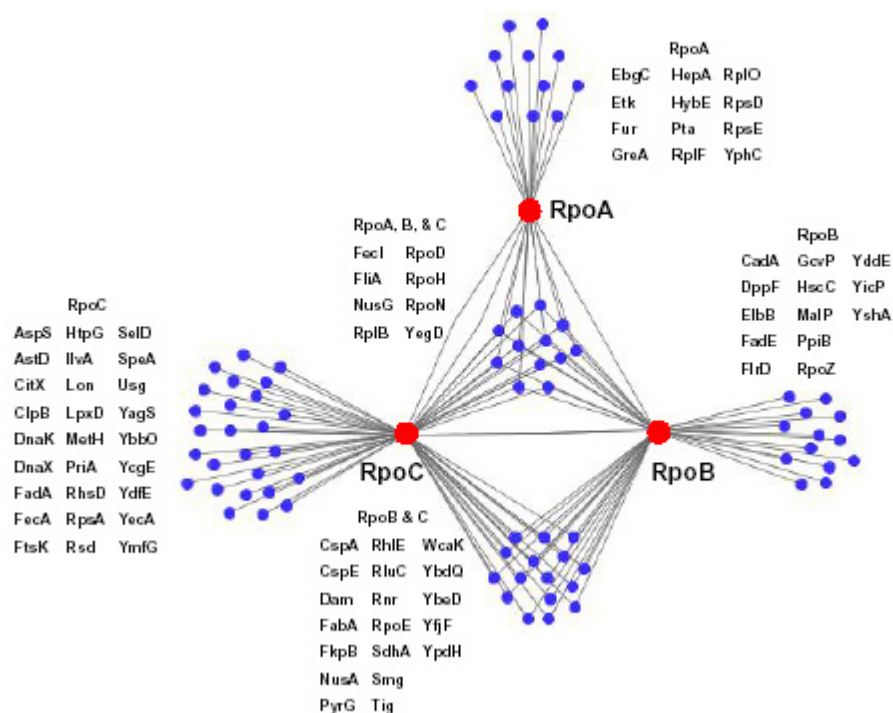


Fig.11 RNAポリメラーゼを中心とするタンパク質間相互作用（文献42、Ariffzzaman et al. より引用）

（青いノード）直接的にそれぞれの RNA polymerase subunits と相互作用するタンパク質のネットワーク。（赤いノード）RpoA、RpoB、および RpoC を通して相互作用するタンパク質のネットワーク。

なお、多くの転写因子と RNA polymerase との相互作用が検出できていないが、おそらくは、それらの因子の相互作用が一時的な作用であるために共精製の濃度が低くなってしまうのか、あるいはそれらの因子が結合するターゲットの DNA エlement との相互作用が強いため共精製ができていないのかもしれない。

一方、RNA polymerase と DnaK, ClpB, Tig (Trigger factor) といった分子シャペロンとの相互作用が検出されている点は、非常に興味深いことである。RNA polymerase のような高次な構造を持つ分子のアセンブリーに、それぞれどのような分子シャペロンが関わるのか (分子シャペロンの特異性と役割分担) といったことにも、このような網羅的な解析手法が有効であることを示唆している。実際、これらの解析結果

から、**Fig.12** のように、分子シャペロンの総合的な相互作用地図を描いてみると、DnaK, GroL(GroEL), Tigの間での役割分担を示唆するような結果を読み取ることができる。また、第1章で触れたClpBタンパク質の相互作用に注目してみると、相互作用因子として 43 個の候補が得られており、そのうちの 11 個はDnaKと、13 個はGroLとも相互作用していることが見えている。つまりClpBが分子シャペロンとしてDnaKおよびGroLと共同作用していることを示唆するとも考えられる結果である。さらに、YbdQという機能未知のタンパク質との相互作用している候補は 83 個も得られているが、これらの多くは、DnaK, DnaJ, GroLとの相互作用分子と重なっており、このYbdQが、Universal stress proteinと呼ばれる UspA⁴⁶と相同なタンパク質であることなどと合わせて考えると、やはり分子シャペロンとして機能するタンパク質であることを示唆していると思われる。

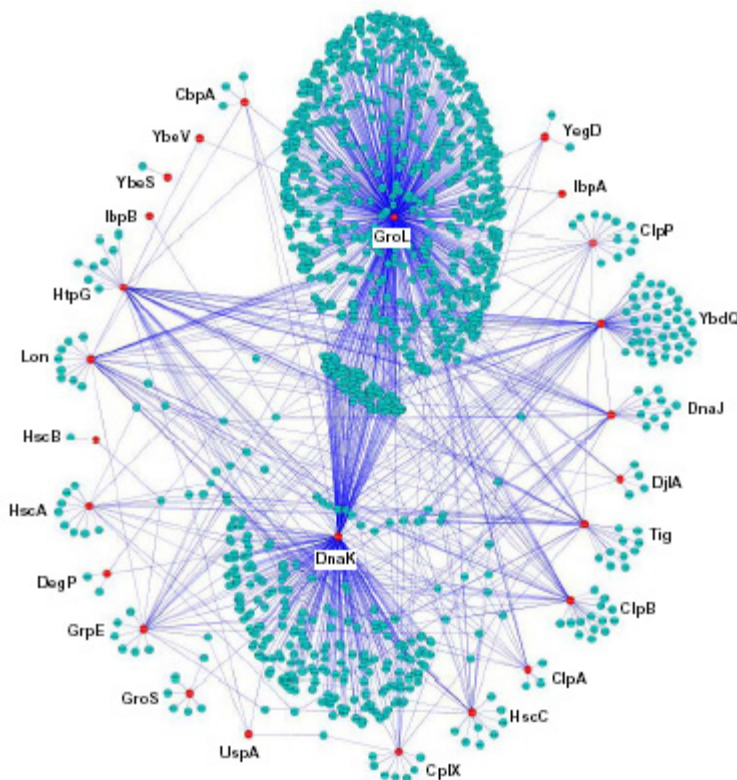


Fig.12 分子シャペロンとそのターゲットタンパク質間の相互作用（文献42、Ariffzzaman et al. より引用）

分子シャペロンとそれらのターゲットの相互作用関係をネットワーク図として表現し

た。赤いノードは主要な付分子シャペロンを表し、青いノードはそれらのターゲット分子を表している。最も大きいクラスタはGroL相互作用ネットワークとDnaKの相互作用ネットワークを中心に示している。

また、これまで予測 ORF として機能未知の遺伝子のうちから 798 個について、この解析で相互作用因子として見つかったものがあつたが、つまりこれらは、大腸菌で実際に発現されていることが証明されたことを意味し、そのような点においても、この解析手法の有用性があると考えられた。

6) タンパク質の構造解析への利用

この系で精製されたタンパク質が大きな意義を発揮できる分野のひとつに、タンパク質の構造解析研究があげられる。この系を使えば、タグを用いた精製の他にも、タグを除かれた形で、ほとんど余分なアミノ酸配列の付加のない形に発現系を改変することも容易である。最近では、タンパク質の構造解析リソースとして、このような網羅的な発現ライブラリのセットが作られるケースが増加しているが、ASKA library はそれらに先駆けて作製され実用に供されてきた。大腸菌のように、個々の機能解析が深くかつ詳細に進んでいる生物由来のタンパク質の構造が解明されることは、構造と機能の相関関係を知る上で、非常に有意義なものである。

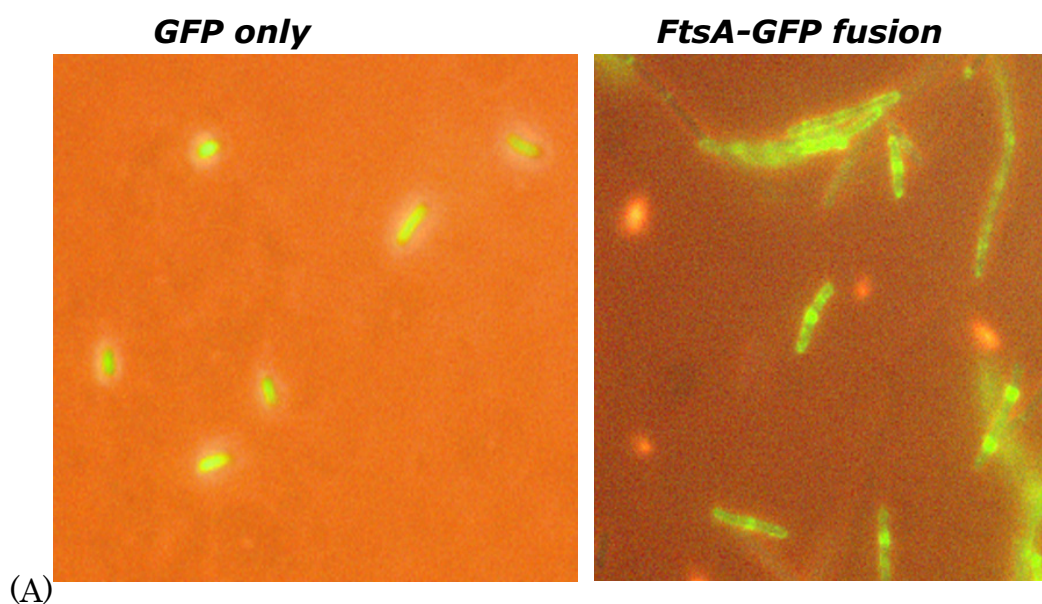
7) GFP 融合タンパク質の局在性観察

タンパク質が細胞内のどこで機能を発揮しているか、つまりその局在場所の同定は、その機能の理解にとって大変重要なことは言うまでもない。

GFP などの蛍光タンパク質をタグにして、各種の生物でタンパク質の機能する組織や細胞内の位置を同定することは、広く行われている研究手法である。

そこで、このクローンセットをそのような目的に利用可能かどうか、菌体を蛍光顕微鏡観察し、**GFP** の細胞内局在性の検証を行った。

先に述べたように、クローン化された**ORF**の発現自体が、かなりの多くの場合、大腸菌自体の生育を阻害する。また、誘導発現された融合タンパク質が封入体と呼ばれる粒子構造を形成してしまうものも多いので、それらの影響をできるだけ少なくするためには、**1nM**程度のきわめて低濃度の**IPTG**で誘導直後か、もしくは、**IPTG**添加なしの状態でもれて発現する**GFP**融合タンパク質の観察が結果的には有効であった。**Fig.13 (A)** は、**FtsA**タンパク質の局在性を検証したものであり、細胞分裂のセプタム形成の位置に局在していることが確認できている⁴⁷。



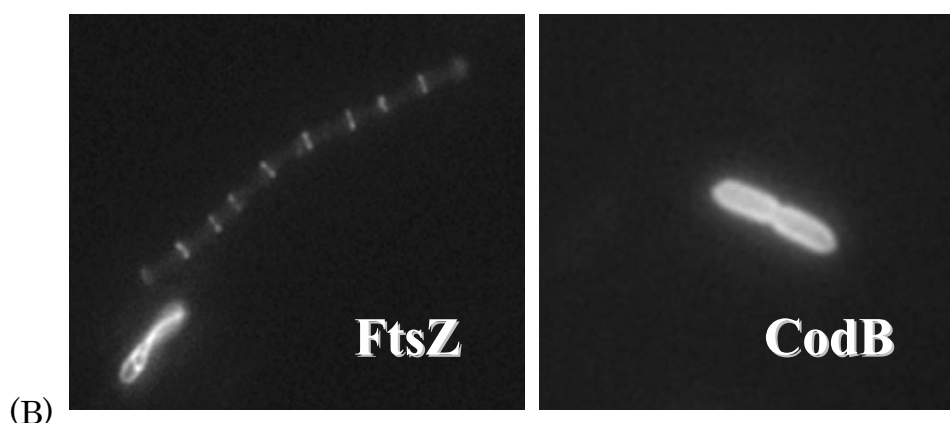


Fig.13 GFPとGFP融合タンパク質の大腸菌での局在性検証例

- (A) ORF の入っていないクローンおよび FtsA-GFP 融合タンパク質の発現クローンについて、1nM の低濃度 IPTG によって 10 分間の発現誘導後に、スライドガラス上にアガロース抱埋して、100 倍対物レンズによって位相差観察を行い、そこに GFP 検出系の光源とフィルターのセットで観察された像を冷却 CCD カメラで撮影した（擬似カラー）。
- (B) 同様に、FtsZ-GFP および CodB-GFP 融合タンパク質の発現クローンについて、IPTG による誘導なしで観察した写真（仁木宏典博士より提供）。

仁木らは、実際にこのクローンセットを用いて、網羅的な局在位置解析を進めており、すでに多くのクローンについての観察結果が **GenoBase** (http://ecoli.naist.jp/GFP/gfp_top.jsp) 上で公開されている。その一例を **Fig.13 (B)**に示す。

その結果では、各顕微鏡観察像をいくつかの特徴的な局在パターンが観察されており、細胞全体が蛍光を示す場合(whole cell)、細胞分裂部位に局在するように見え、細胞の中央部で細胞表面に「リング」状の形を示すもの (ring)、局所的に集積して分布しているように見えるもの (focus)、膜など細胞表層局在するように見え、細胞周辺部に蛍光が観察されるもの (membrane)、核様体の部分に局在しているように細胞の中心部に蛍光が見られ、なもの (nucleoid)、およびそれぞれの複合的 (membrane_focus,nucleoid_focus) なもの、といった具

合に、局在パターン別に分類が行われている。

タンパク質の局在に関しては、SOSUI などのコンピュータソフトウェアを使えば、膜タンパク質かそうでないかなどはある程度推定可能である。その推定の結果と、今回の観察結果を比べてみると、“membrane”に分類されたものの中の 83.5%が Integral membrane protein と推定されるものであった。また、“whole cell”と分類されたものの中の 85%は細胞質またはペリプラズムのタンパク質と推定されていた。一方、転写因子のファミリータンパク質と同定または推定されているものの中で、このような蛍光観察が可能であったもののうち、47.5%は、“focus”および“nucleoid”または“nucleoid_focus”に分類されていた。

このことから、この蛍光観察による手法で見られる局在性が、実際の局在性を反映している確度は相当に高い。したがって、この手法によって、機能未知のタンパク質の機能する「場」をつかんでおくことは、非常に有意義であると考えられる。

6. システムバイオロジーのための遺伝子機能解析ツールとしての今後

遺伝子の生理学的機能の同定など、目的遺伝子の破壊株や欠失株を作製し、その表現型の解析がよく行われるが、それと同様に、遺伝子産物の過剰発現による表現型変化も、多くの機能的示唆を与えてくれるのは間違いない。

先に記したように、GFP 融合タンパク質の過剰発現は、生育阻害をもたらしたものもある。ただしこれは、GFP が C 末に融合していることに主原因がある可能性があり、遺伝子の機能との関連を考えるには、GFP 融合の影響を除いた系で検証する必要がある。そこで、GFP を除いたクローンにおいてタンパク質を IPTG 誘導により過剰生産させ、それぞれの大腸菌の増殖について、濁度を指標に調べてみた。Fig.14 は、*NotI* 断片を除く方法で GFP 部分を取り除いた

プラスミドを持つ大腸菌を、LB 培地で培養し、O.D.(600nm)の値が約 0.5 程度に達した時点で IPTG を終濃度 0.1mM 添加し、その 2 時間後までの O.D.値の対数を取り、IPTG 添加時の O.D.値の対数との差を求め、その値（対数 O.D.値差）の小さいものから順にプロットしたグラフで、対数 O.D.値差の分布状況をみたものである。このグラフを見ると、対数 O.D.値差の値が 0.3 を越えるか超えないかという点付近に、分布の変化点を見ることができる。また、ORF をクローン化していないクローンの値は 0.3 から 0.4 に分布した。そこでこれを生育への影響が見られた状態変化のひとつの指標として、対数 O.D.値差が 0.3 以上あった場合を"Non"（ほとんど影響なし）、0 以上-0.3 未満の場合を"Moderate"（抑制あり）、0 未満の場合を"Severe"（厳しく抑制）と分類を試みた。

その閾値で見えてみると、GFP 融合時ではコロニー形成能が著しく阻害されていたクローンでも、そのおよそ半数については、GFP を除いた場合には十分に生育の伸びが観察された部類に入っていた (**Fig.15**)。それらのタンパク質については、GFP との融合が生育阻害の主原因だった可能性が高い。

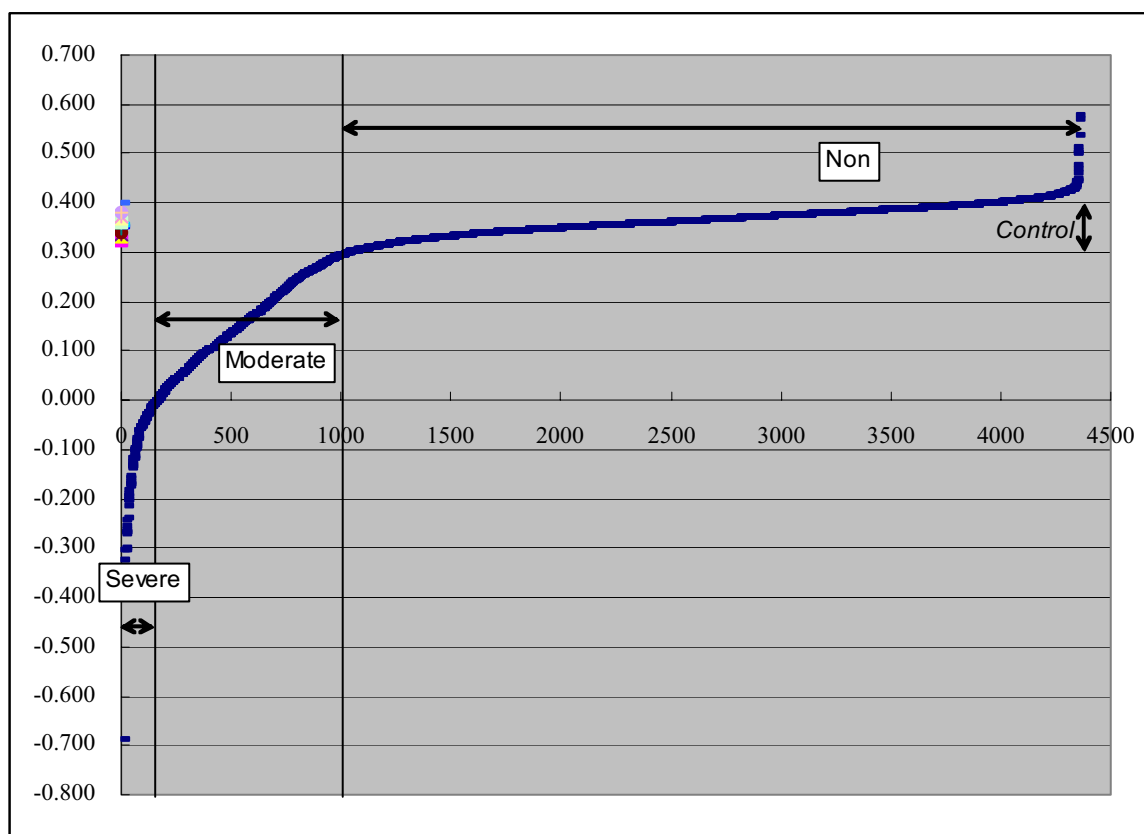


Fig.14 GFP 部分を取り除いたクローンでの IPTG 誘導発現時の生育阻害程度の分布

GFP部分を取り除いたクローンの生育阻害程度の分布

NotI 断片を除く方法で GFP 部分を取り除いたクローンを、LB 培地で培養し、O.D.(600nm)の濁度測定で生育を測った。O.D.値が約 0.5 に達した時点で IPTG を添加し、その 2 時間後までの O.D.値の対数を取り、対数 O.D.値差が小さいものから順にプロットした。対照として GFP をコードしていないクローンでの値をグラフ左端にプロットした。

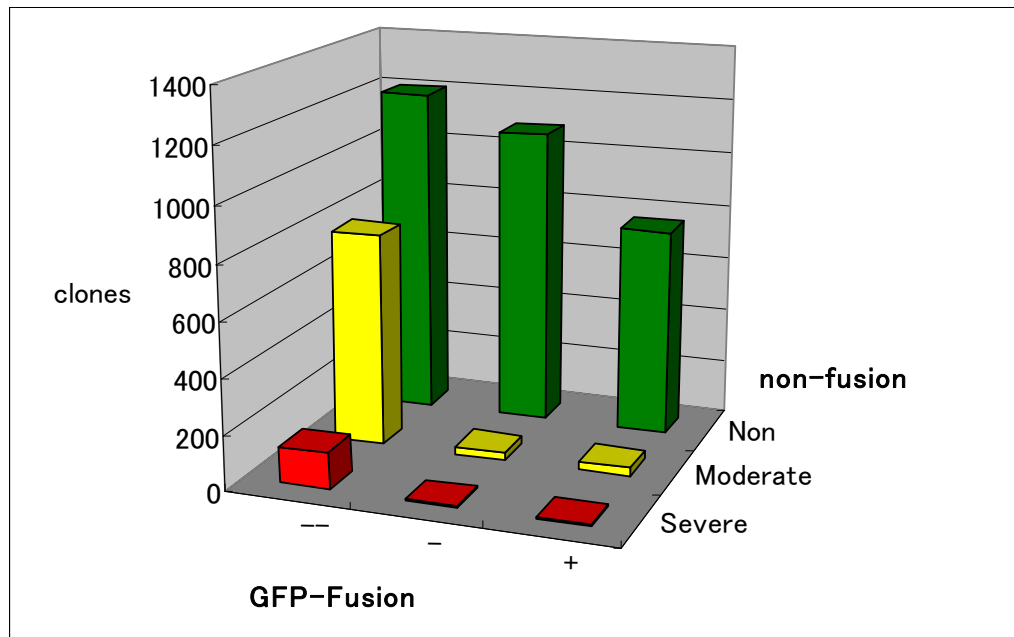


Fig.15 GFP 融合タンパク質と GFP 部分を持たないタンパク質の誘導発現時の生育阻害比較

対数O.D.値差から、増殖への影響の大きさを見積もり、0.3 以上あった場合を"Non"（ほとんど影響なし）、0.03 以上-0.3 未満の場合を"Moderate"（抑制あり）、0.03 未満の場合を"Severe"（厳しく抑制）と分類した。その分類と、GFP融合タンパク質の誘導発現時に見られたコロニー形成を指標とした生育への影響（Fig.6で示したGrowthへの影響の分類）について、それぞれに分類されるクローンの数を比較した。

一方、GFP を取り除いてもなお、その過剰な誘導発現が生育に著しく阻害的に働くものの内訳を、細胞内の局在性の予測結果(ソフトウェア SOSUI 使用)と照合してみると、多くは膜タンパク質あるいは分泌性のタンパク質などの細胞表層のタンパク質であると推定された (Fig.16)。つまり、膜タンパク質や分泌タンパク質の過剰発現は、大腸菌細胞にとって、一般的に毒性が強いと考えられる。また、GFP を除けば、過剰発現しても生育に著しい阻害をもたらさないように見える遺伝子は全体の 75%ほどもあり、その 80%以上は細胞質に存在する可溶性のタンパク質であると推定された。

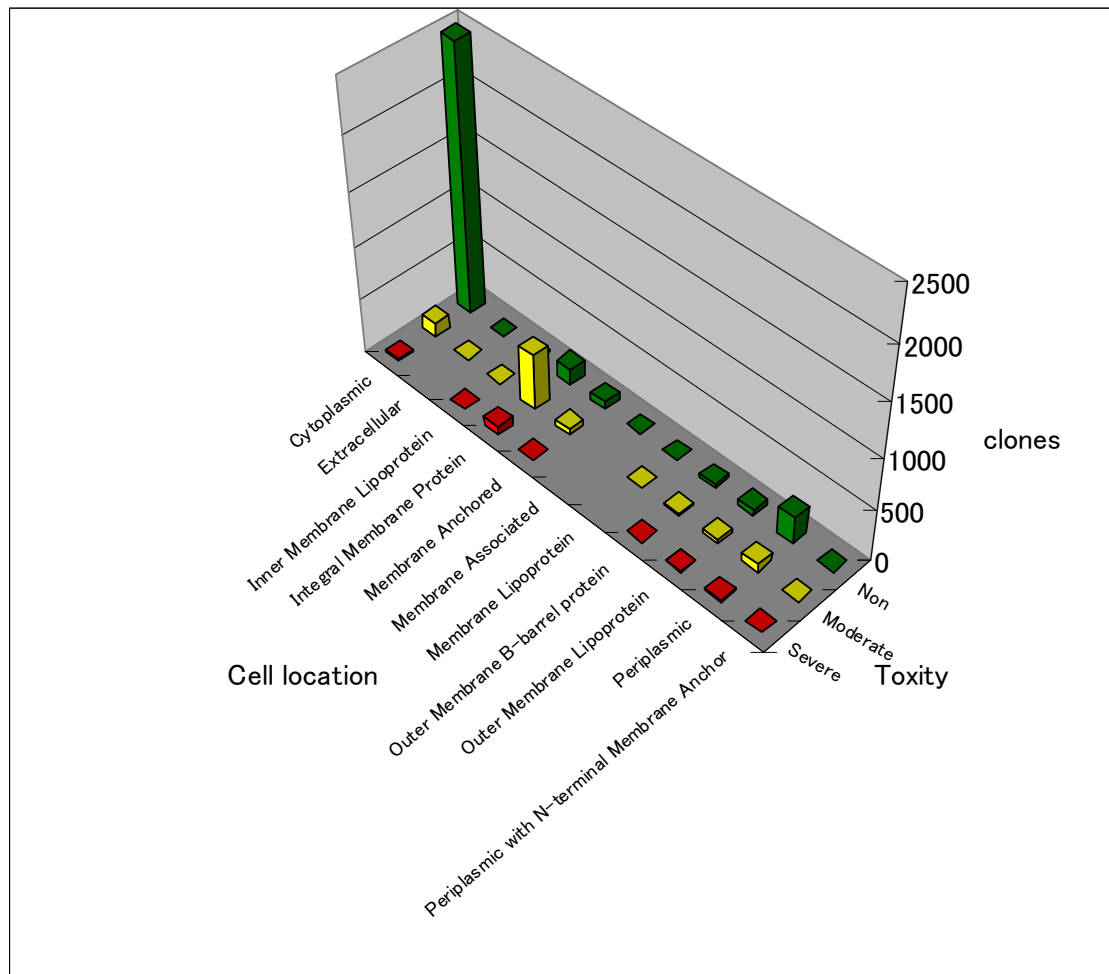


Fig.16 生育阻害効果とタンパク質の局在性の相関

GFP 部分を取り除いたクローンにおいて、IPTG による誘導発現が生育に与える影響度合いを、細胞内の局在性の予測結果(ソフトウェア SOSUI 使用)ごとに集計した。

生育の阻害とは、つまり細胞システムに異常をきたした結果と考えれば、もともと大腸菌が持っていた遺伝子の産物が過剰発現され、それによってシステム異常が起きたということであり、そういう部分は、システムとして脆弱性(Flagility)が高いと言えるかもしれない。逆に、過剰発現によっても生育に支障をきたさない部分は、多少のバランスのずれにも対応できるという意味では、システムとして頑健性(Robustness)が高いと考えられないだろうか。

このような観点からこれらの結果を見ることによって、たとえば代謝系の中

の脆弱性（頑健性）の推定について考えてみた。それに際し、膜タンパク質や分泌性のタンパク質の生育阻害効果は著しいが、それは部分的なシステム異常というより膜構造全体などへの毒性の影響が強いと考えられるので、細胞質タンパク質と推定されるものに限って、生育効果の著しい遺伝子について注目してみると、以下の 8 つの遺伝子について、特に顕著な生育の阻害が見られた。

relE : Qin prophage; toxin of the RelE-RelB toxin-antitoxin system

dicB : Qin prophage; cell division inhibition protein

ydfD : Qin prophage; predicted protein

yjfP : CP4-57 prophage; predicted GTP-binding protein

yicF : Definition putative DNA ligase

fabA : beta-hydroxydecanoyl thioester dehydrase

argB : acetylglutamate kinase

nadB : L-aspartate oxidase

最初の 4 つは、プロウイルスの遺伝子と思われるもので、特に *relE* 遺伝子の産物は、細菌に対する毒性を持つタンパク質であることが知られており、そのアンチトキシン遺伝子である *relB* 遺伝子の欠失株は Keio コレクションにおいて「致死」の表現型を示すことが報告されている。したがって、この生育阻害は、RelE タンパク質の過剰発現によりトキシン・アンチトキシンのバランスが崩れて、細胞毒性が発揮された結果であると推察される。後半の 4 つの遺伝子産物の過剰発現が、生育を強く阻害した原因は不明であるが、そのうちの 3 つ (*fabA*, *argB*, *nadB*) については、KEGG データベース上に対応するパスウェイを抽出することができた。

Table1 は、その対応パスウェイについて、全く影響の見られなかった遺伝子のパスウェイとともに示したものである。このような解析により、頑健性の高

い代謝系の候補と脆弱な代謝系の候補を抽出できれば、そのような候補について、実験によってシステムとしての頑健性や脆弱性を調べていくことは、そこをターゲットするような「病因」やそれに対する「創薬」の候補探索につながる可能性がある。

High Robustness	Pathway	
<i>citG</i> , <i>ybdT</i> ; 2'-(5"-triphosphoribosyl)-3-dephosphor-CoA synthase [EC:2.7.8.25] [SP:CITG_ECOLI]	eco02020	Two-component system - General
<i>dppF</i> , <i>dppE</i> ; dipeptide transport protein (ABC superfamily, atp_bind) [SP:DPPF_ECOLI]	eco02010	ABC transporters - General
	eco02011	ABC transporters - Organism-specific
<i>glgA</i> ; glycogen synthase [EC:2.4.1.21] [SP:GLGA_ECOLI]	eco00500	Starch and sucrose metabolism
<i>ilvN</i> ; acetolactate synthase I, small subunit [EC:2.2.1.6] [SP:ILVN_ECOLI]	eco00290	Valine, leucine and isoleucine biosynthesis
	eco00650	Butanoate metabolism
	eco00660	C5-Branched dibasic acid metabolism
	eco00770	Pantothenate and CoA biosynthesis
<i>xdhB</i> , <i>ygeT</i> ; putative xanthine dehydrogenase subunit, FAD-binding domain [EC:1.17.1.4] [SP:XDHB_ECOLI]	eco00230	Purine metabolism
<i>ygfH</i> ; propionyl-CoA:succinate-CoA transferase [EC:2.8.3.-] [SP:YGFH_ECOLI]	eco00624	1- and 2-Methylnaphthalene degradation
High Fragility	Pathway	
<i>fabA</i> ; beta-hydroxydecanoyl thioester dehydrase (trans-2-decenoyl-ACP isomerase) [EC:4.2.1.60] [SP:FABA_ECOLI]	eco00061	Fatty acid biosynthesis
<i>argB</i> , <i>argA</i> ; acetylglutamate kinase [EC:2.7.2.8] [SP:ARGB_ECOLI]	eco00220	Urea cycle and metabolism of amino groups
<i>nadB</i> , <i>nic</i> , <i>nicB</i> ; L-aspartate oxidase [EC:1.4.3.16] [SP:NADB_ECOLI]	eco00252	Alanine and aspartate metabolism
	eco00760	Nicotinate and nicotinamide metabolism

Table.1

生育への影響の度合い特に強いもの（対数 OD 値差が 0.43 以上あるもの：High Robustness）と弱いもの（対数 OD 値差が負となるもの：High Fragility）のうち、その局在性が SOSUI の予測で"Cytoplasmic"とされるものに関して、それが含まれる KEGG の大腸菌 Pathway map を抽出した。

ポストゲノムシーケンス後の機能未知遺伝子の機能解明は、ジグソーパズルを順々に完成させていく過程にしばしば例えられる。配列決定による全遺伝子構造の解読は、全ピースが裏向きで出揃った状態を意味し、個々の遺伝子の機能解析は各ピースを表に向けることであり、表にできればピース同士を組み合わせる（遺伝子間のネットワークを結ぶ）ことも易しくなる。その中で、大腸菌はこれまでの研究の歴史が長く、知識の蓄積は質・量ともに極めて多い生物種である。一時は研究し尽くされたという感もあって、「重箱の隅をつつく」研究と揶揄された時期もあったかもしれないが、ゲノム研究の進展によって、ひとつの生命体の全体像を理解する材料として、非常に重要な素材となった。大腸菌は、多くのピースがすでに表を向いており、そのピースが組み合わせられた多数の島ができている状態と言えなくもない。今後の網羅的な機能解析によって、さらに島と島を結んでいく作業が進めば、真の全体像の理解につながるものと考えられる。実際、この遺伝子セットとともに、全遺伝子の破壊株セットを利用したシステムバイオロジーの計画も提案され⁴⁸、推進されている。また、**capillary electrophoresis electrospray ionization mass spectrometry (CE-MS)**などを活用したメタボロームの手法への応用も期待される^{49,50}。そのような研究によって、生きるために必要な遺伝子セットとその機能ネットワークの統合体としての解明を進め、究極的には細胞内の生命活動のシミュレーションによる再現も目前にあるかもしれない。解析リソースの充実は、その具体的な道筋を示してくれるものであろう。

すでに数々の利用に供されているこのクローンセット（ASKA library）であるが、その利用先は大腸菌研究者にとどまらず、他の微生物やヒトも含めた高等生物の研究にも活用されている。

リソースとしての活用には、その情報の整備と更新が必要である。それにつ

いては奈良先端大のWEB上で公開されているGenoBase

(<http://ecoli.naist.jp/GB6/search.jsp>) によくまとめられており、ゲノム配列の更新やアノテーション情報の更新にも対応していくなどの体制整備

(http://ecoli.naist.jp/Resource/reclone_list.html) も進められ、配布体制に関しては、奈良先端大だけでなく、国立遺伝学研究所においても整備されている。

今後も新たな研究のリソースとして広く供されることを期待するものである。

謝辞

本研究を進めるにあたり、終始ご支援ご指導をいただきました、奈良先端科学技術大学院大学の森浩禎教授と森研究室の皆さんに心より感謝いたします。

クローンセットの作製については、森研究室の中道朋子さん、稲本英次さん、豊永宏美さんに多大なご助力をいただきました。この論文は、私が奈良先端科学技術大学院大学を退職後も、森研究室の方々が 5 年以上にわたって継続してくださった研究成果の集大成の意味も大きいものであると感じています。

クローンの設計には、荒武氏に助けていただきました。

Md. Arifuzzaman 氏には、多くのデータの取得で助けていただきました。

国立遺伝学研究所の仁木宏典先生には、GFP の局在性解析データの取得でご支援いただきました。

GFPuv4 の遺伝子は、埼玉大学伏見研究室の伊藤洋一博士より供与いただきました。

私にゲノム研究とストレス研究の機会を与えてくださった由良隆先生と和田千恵子先生、および由良研究室同窓の皆様と、HSP 研究所時代の同僚の皆様に深く感謝いたします。

大腸菌ゲノムプロジェクトで叱咤激励いただいた堀内嵩先生、三木健良先生、磯野克己先生、平賀壮太先生、井口八郎先生、山本義弘先生、西村昭子先生、小椋光先生、松田秀雄先生、加藤潤一先生らの諸先生方に深く感謝いたします。

私を大腸菌の研究に導いてくださった、故水島昭二先生と名古屋大学農学部時代の水島研究室の山田寿美先生ら皆様に深く感謝いたします。

私のことを頼りにしてくれ、いつも迷惑をかけているタカラバイオ・ドラゴンジェノミクスセンターの皆様にも深く感謝します。

最後に、いつも応援してくれた家族に深く感謝いたします。

参考文献

- ¹ Neidhardt, F. C., ed., *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Am. Soc. Microbiol., Washington, DC) (1996).
- ² Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. The Complete Genome Sequence of *Escherichia coli* K-12., (1997) *Science*, **277**, 1453-1462.
- ³ Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B., Mori, H., and Horiuchi, T. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110, (2006) *Mol Syst Biol*, **2**, doi:10.1038/JA - Mol Syst Biol
- ⁴ Mori, H., Isono, K., Horiuchi, T., and Miki, T. Functional genomics of *Escherichia coli* in Japan., (2000) *Res Microbiol*, **151**, 121-128.
- ⁵ Yura, T., Nagai, H., and Mori, H. Regulation of the heat-shock response in bacteria., (1993) *Annu Rev Microbiol*, **47**, 321-350.
- ⁶ Kitagawa, M., Wada, C., Yoshioka, S., and Yura, T. Expression of ClpB, an analog of the ATP-dependent protease regulatory subunit in *Escherichia coli*, is controlled by a heat shock sigma factor (sigma 32)., (1991) *J Bacteriol*, **173**, 4247-4253.
- ⁷ Wickner, S., Gottesman, S., Skowrya, D., Hoskins, J., McKenney, K., and Maurizi, M. R. A molecular chaperone, ClpA, functions like DnaK and DnaJ., (1994) *Proc Natl Acad Sci U S A*, **91**, 12218-12222.
- ⁸ Thompson, M. W., and Maurizi, M. R. Activity and specificity of *Escherichia coli* ClpAP protease in cleaving model peptide substrates., (1994) *J Biol Chem*, **269**, 18201-18208.
- ⁹ Squires, C., and Squires, C. L. The Clp proteins: proteolysis regulators or molecular chaperones?, (1992) *J Bacteriol*, **174**, 1081-1085.
- ¹⁰ JACOB, F., PERRIN, D., SANCHEZ, C., and MONOD, J. Operon: a group of genes with the expression coordinated by an operator., (1960) *C R Hebd Seances Acad Sci*, **250**, 1727-1729.
- ¹¹ Yamada, H., Kitagawa, M., Kawakami, M., Mizushima, S. The gene coding for lipoprotein signal peptidase (*lspA*) and that for isoleucyl-tRNA

-
- synthetase (*ileS*) constitute a cotranscriptional unit in *Escherichia coli*, (1984) *FEBS Lett*, **171**, 245-248.
- ¹² **Clarke, L., and Carbon, J.** A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome., (1976) *Cell*, **9**, 91-99.
- ¹³ **Kohara, Y., Akiyama, K., and Isono, K.** The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library., (1987) *Cell*, **50**, 495-508.
- ¹⁴ **Tabata, S., Higashitani, A., Takanami, M., Akiyama, K., Kohara, Y., Nishimura, Y., Nishimura, A., Yasuda, S., and Hirota, Y.** Construction of an ordered cosmid collection of the *Escherichia coli* K-12 W3110 chromosome., (1989) *J Bacteriol*, **171**, 1214-1218.
- ¹⁵ **Knott, V., Blake, D. J., and Brownlee, G. G.** Completion of the detailed restriction map of the *E. coli* genome by the isolation of overlapping cosmid clones., (1989) *Nucleic Acids Res*, **17**, 5901-5912.
- ¹⁶ **Birkenbihl, R. P. and Vielmetter, W.** Cosmid-derived map of *E. coli* strain BHB2600 in comparison to the map of strain W3110., (1989) *Nucleic Acids Res.*, **17**, 5057-5069.
- ¹⁷ **Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K., Mizobuchi, K., and Nakata, A.** Systematic sequencing of the *Escherichia coli* genome: analysis of the 0-2.4 min region., (1992) *Nucleic Acids Res*, **20**, 3305-3308.
- ¹⁸ **Fujita, N., Mori, H., Yura, T., and Ishihama, A.** Systematic sequencing of the *Escherichia coli* genome: analysis of the 2.4-4.1 min (110,917-193,643 bp) region., (1994) *Nucleic Acids Res*, **22**, 1637-1639.
- ¹⁹ **Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., Venter, C.,** Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd., (1995) *Science*, **269**, 496-512.
- ²⁰ **Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K., Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nashimoto, H., Nishio, Y., Saito, N., Horiuchi, T.** A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7-28.0 min region on the linkage map., (1996) *DNA Res*, **3**, 137-155.

-
- ²¹ Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kasai, H., Kashimoto, K., Kimura, S., Kitakawa, M., Kitagawa, M., Makino, K., Miki, T., Mizobuchi, K., Mori, H., Mori, T., Motomura, K., Nakade, S., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Horiuchi, T. A 570-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 28.0-40.1 min region on the linkage map., (1996) *DNA Res*, **3**, 363-377.
- ²² Itoh, T., Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Kasai, H., Kimura, S., Kitakawa, M., Kitagawa, M., Makino, K., Miki, T., Mizobuchi, K., Mori, H., Mori, T., Motomura, K., Nakade, S., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Horiuchi, T. A 460-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 40.1-50.0 min region on the linkage map., (1996) *DNA Res*, **3**, 379-392.
- ²³ Yamamoto, Y., Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kimura, S., Kitagawa, M., Makino, K., Miki, T., Mitsuhashi, N., Mizobuchi, K., Mori, H., Nakade, S., Nakamura, Y., Nashimoto, H., Oshima, T., Oyama, S., Saito, N., Sampei, G., Satoh, Y., Sivasundaram, S., Tagami, H., Horiuchi, T. Construction of a contiguous 874-kb sequence of the *Escherichia coli* K-12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features., (1997) *DNA Res*, **4**, 91-113.
- ²⁴ Sambrook, J., Fritsch, E. F., Maniatis, T. (1998) *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor, New York Cold Spring Harbor Laboratory
- ²⁵ Ito, Y., Suzuki, M., and Husimi, Y. A novel mutant of green fluorescent protein with enhanced sensitivity for microanalysis at 488 nm excitation., (1999) *Biochem Biophys Res Commun*, **264**, 556-560.
- ²⁶ Baba, T., Ara, T., Okumura, Y., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H. Construction of *Escherichia coli* K-12 in-frame, single-gene knock-out mutants—the Keio collection., (2006) *Mol Syst Biol* doi:10.1038/msb4100050
- ²⁷ Oshima, T., Aiba, H., Masuda, Y., Kanaya, S., Sugiura, M., Wanner, B. L., Mori, H., and Mizuno, T. Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12., (2002) *Mol Microbiol*, **46**, 281-291.
- ²⁸ Oshima, T., Wada, C., Kawagoe, Y., Ara, T., Maeda, M., Masuda, Y., Hiraga, S., and Mori, H. Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*,

(2002) *Mol Microbiol*, **45**, 673-695.

²⁹ Nakahigashi, K., Kubo, N., Narita, S., Shimaoka, T., Goto, S., Oshima, T., Mori, H., Maeda, M., Wada, C., and Inokuchi, H. From the Cover: HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and *hemK* knockout induces defects in translational termination., (2002) *Proc Natl Acad Sci U S A*. **99**, 1473-1478.

³⁰ Kodama, K., Kobayashi, T., Niki, H., Hiraga, S., Oshima, T., Mori, H., and Horiuchi, T. Amplification of Hot DNA segments in *Escherichia coli*., (2002) *Mol Microbiol*, **45**, 1575-1588.

³¹ Minagawa, S., Ogasawara, H., Kato, A., Yamamoto, K., Eguchi, Y., Oshima, T., Mori, H., Ishihama, A., and Utsumi, R. Identification and Molecular Characterization of the Mg²⁺ Stimulon of *Escherichia coli*., (2003) *J. Bacteriol.*, **185**, 3696-3702.

³² Saka, K., Tadenuma, M., Nakade, S., Tanaka, N., Sugawara, H., Nishikawa, K., Ichiyoshi, N., Kitagawa, M., Mori, H., Ogasawara, N., Nishimura, A. A complete set of *Escherichia coli* open reading frames in mobile plasmids facilitating genetic studies., (2005) *DNA Res*, **12**, 63-68.

³³ Proudfoot, M., Kuznetsova, E., Brown, G., Rao, N. N., Kitagawa, M., Mori, H., Savchenko, A., Yakunin, A. F. General enzymatic screens identify three new nucleotidases in *Escherichia coli*. Biochemical characterization of SurE, YfbR, and YjgG., (2004) *J Biol Chem*, **279**, 54687-54694.

³⁴ Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P., and Vidal, M. Towards a proteome-scale map of the human protein-protein interaction network., (2005) *Nature*, **437**, 1173-1178.

³⁵ Ehrmann M, Boyd D, Beckwith J. (1990) Genetic analysis of membrane protein topology by a sandwich gene fusion approach., (1990) *Proc Natl Acad Sci U S A*. **87**(19):7574-7578.

³⁶ Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R.,

-
- Ara, T., Nakahigashi, K., Huang, H. C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., Mori, H. Large-scale identification of protein-protein interaction of *Escherichia coli* K-12., (2006) *Genome Res*, **5**, 686-691.
- ³⁷ Xenarios, I., Rice, D. W., Salwinski, L., Baron, M. K., Marcotte, E. M., and Eisenberg, D. DIP: the database of interacting proteins., (2000) *Nucleic Acids Res*, **28**, 289-291.
- ³⁸ Butland, G., Peregrin-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. Interaction network containing conserved and essential protein complexes in *Escherichia coli*., (2005) *Nature*, **433**, 531-537.
- ³⁹ Ito, T., Ota, K., Kubota, H., Yamaguchi, Y., Chiba, T., Sakuraba, K., and Yoshida, M. Roles for the two-hybrid system in exploration of the yeast protein interactome., (2002) *Mol Cell Proteomics*, **1**, 561-566.
- ⁴⁰ Kanehisa, M., Goto, S., Kawashima, S., and Nakaya, A. The KEGG databases at GenomeNet., (2002) *Nucleic Acids Res*, **30**, 42-46.
- ⁴¹ Nehrke, K. W., Zalatan, F., and Platt, T. NusG alters rho-dependent termination of transcription in vitro independent of kinetic coupling., (1993) *Gene Expr*, **3**, 119-133.
- ⁴² Kainz, M., and Gourse, R. L. The C-terminal domain of the alpha subunit of *Escherichia coli* RNA polymerase is required for efficient rho-dependent transcription termination., (1998) *J Mol Biol*, **284**, 1379-1390.
- ⁴³ Li, Z., Pandit, S., and Deutscher, M. P. Polyadenylation of stable RNA precursors *in vivo*., (1998) *Proc Natl Acad Sci U S A*, **95**, 12158-12162.
- ⁴⁴ Cheng, Z. F., Zuo, Y., Li, Z., Rudd, K. E., and Deutscher, M. P. The *vacB* gene required for virulence in *Shigella flexneri* and *Escherichia coli* encodes the exoribonuclease RNase R., (1998) *J Biol Chem*, **273**, 14077-14080.
- ⁴⁵ Yamanaka, K., Fang, L., and Inouye, M. The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation., (1998) *Mol Microbiol*, **27**, 247-255.
- ⁴⁶ Nystrom T., Neidhardt F.C. Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest., (1994) *Mol Microbiol*. **11**(3):537-544.
- ⁴⁷ Ma, X., Ehrhardt, D. W., and Margolin, W. Colocalization of cell division

proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein, (1996) *PNAS*, **93**, 12998-13003.

⁴⁸ **Arita, M., Robert, M., and Tomita, M.** All systems go: launching cell simulation fueled by integrated experimental biology data., (2005) *Curr Opin Biotechnol*, **16**, 344-349.

⁴⁹ **Itoh, A., Ohashi, Y., Soga, T., Mori, H., Nishioka, T., and Tomita, M.** Application of capillary electrophoresis-mass spectrometry to synthetic in vitro glycolysis studies., (2004) *Electrophoresis*, **25**, 1996-2002.

⁵⁰ **Saito, N., Robert, M., Kitamura, S., Baran, R., Soga, T., Mori, H., Nishioka, T., and Tomita, M.** Metabolomics approach for enzyme discovery., (2006) *J Proteome Res*, **5**, 1979-1987

推薦教員	森浩禎 教授		
氏名	北川正成	提出	平成18年 7月11日
題目	大腸菌全遺伝子の発現クローンセット作製の研究		
要旨 過去半世紀以上にわたって、大腸菌は遺伝学、生化学、分子生物学的な研究の好材料として研究され、膨大な意義深いデータの蓄積をわれわれに与えてくれた。 この大腸菌の代表的な実験株である K12 株のゲノムには、その全塩基配列から、約 4,300 個のタンパク質をコードする遺伝子が存在していると推定されている。 大腸菌はこれらの遺伝子群により単細胞原核生物としてこの地球上に生き続けているわけだが、「どのようにこれらの遺伝子を使って生き抜いているのか?」、「各遺伝子の役割は何か?」という疑問に対して、答にもっとも近いモデル生物の一種が大腸菌であり、ゲノム研究がポストゲノム解析の時代へと進んだ今、システムバイオロジーの分野においても、最適な材料を提供してくれる価値のある生物であり、ひとつの生命体としての総合理解にもっとも近い生物であることは間違いない。 本研究では、日本のグループがゲノム配列解析を行った大腸菌K12 株の派生株W3110 のゲノム配列をもとに予測された遺伝子のタンパク質コード領域 4276 個のOpen reading frame(ORF)をクローン作製の対象として、独自に工夫を施したマルチパーパスな発現ベクターpCA24Nプラスミドを用いて、Hisタグと蛍光タンパク質Green fluorescent protein(GFP)の両者に方向性と読み枠を考慮して各ORFをクローニングしたセット”ASKA library (A Complete Set of <i>E. coli</i> K-12 ORF Archives)”を作製した。 このようなマルチパーパスな発現プラスミドの設計において、 ・クローン化断片を利用した DNA マイクロアレイ作製のための鋳型調製 ・多様な解析に供するための容易なサブクローン化 ・変異株に対する相補試験による変異遺伝子の同定 ・タンパク質間相互作用の網羅的な検出による相互作用地図の作成 ・誘導発現タンパク質の精製による生化学的解析と高次構造決定 ・タンパク質の細胞内局在性の解析 ・膜タンパク質のトポロジー解析 などの網羅的および個別機能解析実験への適用を目的とし、実際に様々な応用が可能であることを示した。 一方、作製過程においては、クローン化バイアスに対する懸念をできるだけ除くため、			

あえて平滑末端でのクローン化を行い、**ORF**断片が逆向きに挿入されてくることを排除せず、双方向にクローン化されることをひとつの指標とすることにした。その結果、この系において、正方向と逆方向のクローンが取得されてくる頻度に大きな偏りがあるケースは稀であり、また、コロニーの形態で見る限り、**IPTG**による誘導をかけなければ、そのクローン化方向の区別は全くつかない状態であることを確認できた。これは、得られたクローンに変異などが入っているといったクローン化バイアスの低さを示唆するものである。

また、各クローンプラスミドから **GFP** 部分を容易に除ける系を用いて、**GFP** を持たない形のタンパク質を、それらのクローンから誘導発現させた時の、宿主大腸菌の生育へ与える影響を検討し、生育の阻害を細胞システムに異常をきたした結果と捉え、システムとしての脆弱性(**Flagility**)と頑健性(**Robustness**)についての可能性を示した。

このクローンセットは、大腸菌の遺伝子機能の解析にとって格好の材料であり、すでに上記のような様々な網羅的な解析実験に供与され、数々の実績を上げている。今後もさらなる活用が期待され、とくにシステムバイオロジーにおけるきわめて有用な実験材料としてきわめて有用である。

論文目録

推薦教員	森浩禎教授		
氏 名	北川正成	提 出	平成18年 7 月11 日
学位論文の主たる部分を公表した論文 (題名、全著者名、公表時期、雑誌名、巻、ページ) (1) Complete set of ORF clones of <i>Escherichia coli</i> ASKA library (A Complete Set of <i>E. coli</i> K-12 ORF Archive): Unique Resources for Biological Research. <u>Kitagawa, M.</u>, Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., Mori, H., (2005) <i>DNA Res</i>, 12, 291-299. 参考論文 (題名、全著者名、公表時期、雑誌名、巻、ページ) (1) Large-scale identification of protein-protein interaction of <i>Escherichia coli</i> K-12. Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H. C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., <u>Kitagawa, M.</u>, Tomita, M., Kanaya, S., Wada, C., Mori, H., (2006) <i>Genome Res</i>, 5, 686-691. (2) A complete set of <i>Escherichia coli</i> open reading frames in mobile plasmids facilitating genetic studies. Saka, K., Tadenuma, M., Nakade, S., Tanaka, N., Sugawara, H., Nishikawa, K., Ichiyoshi, N., <u>Kitagawa, M.</u>, Mori, H., Ogasawara, N., Nishimura, A., (2005) <i>DNA Res</i>, 12, 63-68. (3) General enzymatic screens identify three new nucleotidases in <i>Escherichia coli</i>. Biochemical characterization of SurE, YfbR, and YjjG.. Proudfoot, M., Kuznetsova, E., Brown, G., Rao, N. N., <u>Kitagawa, M.</u>, Mori, H., Savchenko, A., Yakunin, A. F., (2004) <i>J Biol Chem</i>, 279, 54687-54694. (4) The genome sequence of silkworm, <i>Bombyx mori</i>. Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Namiki, N., <u>Kitagawa, M.</u>, Yamashita, H., Yasukochi, Y., Kadono-Okuda, K., Yamamoto, K., Ajimura, M., Ravikumar, G., Shimomura, M., Nagamura, Y., Shin-I, T., Abe, H., Shimada, T., Morishita, S., Sasaki, T., (2004) <i>DNA Res</i>, 11, 27-35.			

- (5) Second transmembrane segment of FtsH plays a role in its proteolytic activity and homo-oligomerization. **Makino, S., Makino, T., Abe, K., Hashimoto, J., Tatsuta, T., Kitagawa, M., Mori, H., Ogura, T., Fujii, T., Fushinobu, S., Wakagi, T., Matsuzawa, H.,** (1999) *FEBS Lett*, **460**, 554-548.
- (6) Heat shock regulation in the ftsH null mutant of *Escherichia coli*: dissection of stability and activity control mechanisms of sigma32 *in vivo*. **Tatsuta, T., Tomoyasu, T., Bukau, B., Kitagawa, M., Mori, H., Karata, K., Ogura, T.,** (1998) *Mol Microbiol*, **30**, 583-593.
- (7) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. **Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura, T.,** (1998) *Appl Environ Microbiol*, **64**, 1694-1699.
- (8) Construction of a contiguous 874-kb sequence of the *Escherichia coli* -K12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features. **Yamamoto, Y., Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kimura, S., Kitagawa, M., Makino, K., Miki, T., Mitsuhashi, N., Mizobuchi, K., Mori, H., Nakade, S., Nakamura, Y., Nashimoto, H., Oshima, T., Oyama, S., Saito, N., Sampei, G., Satoh, Y., Sivasundaram, S., Tagami, H., Horiuchi, T.,** (1997) *DNA Res*, **4**, 91-113.
- (9) A 570-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 28.0-40.1 min region on the linkage map. **Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kasai, H., Kashimoto, K., Kimura, S., Kitakawa, M., Kitagawa, M., Makino, K., Miki, T., Mizobuchi, K., Mori, H., Mori, T., Motomura, K., Nakade, S., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Horiuchi, T.,** (1996) *DNA Res*, **3**, 363-377.
- (10) A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7-28.0 min region on the linkage map. **Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K., Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nashimoto, H., Nishio, Y., Saito, N., Horiuchi, T.,** (1996) *DNA Res*, **3**, 137-155.
- (11) A 460-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 40.1-50.0 min region on the linkage map. **Itoh, T., Aiba, H., Baba, T., Hayashi, K., Inada, T., Isono, K., Kasai, H.,**

- Kimura, S., Kitakawa, M., Kitagawa, M., Makino, K., Miki, T., Mizobuchi, K., Mori, H., Mori, T., Motomura, K., Nakade, S., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Horiuchi, T., (1996) *DNA Res*, **3**, 379-392.**
- (12) Nucleotide sequence of the arginine deiminase gene of *Mycoplasma hominis*. **Harasawa, R., Koshimizu, K., Kitagawa, M., Asada, K., Kato, I., (1992) *Microbiol Immunol*, **36**, 661-5.**
- (13) Expression of ClpB, an analog of the ATP-dependent protease regulatory subunit in *Escherichia coli*, is controlled by a heat shock sigma factor (sigma 32). **Kitagawa, M., Wada, C., Yoshioka, S., Yura, T., (1991) *J Bacteriol*, **173**, 4247-4253.**
- (14) The gene coding for lipoprotein signal peptidase (*lspA*) and that for isoleucyl-tRNA synthetase (*ileS*) constitute a cotranscriptional unit in *Escherichia coli*. **Yamada, H., Kitagawa, M., Kawakami, M., Mizushima, S., (1984) *FEBS Lett*, **171**, 245-248.**