

細菌コロニーの呈色反応を定量するシステムの開発および新規バイオフィルム合成関連遺伝子探索への応用に関する研究

Development of monitoring and quantification method of bacterial colony color and its application on the biofilm formation of *Escherichia coli*.

眞喜志 一

奈良先端科学技術大学院大学

バイオサイエンス研究科 システム微生物学講座

(森 浩禎 教授)

平成30年 2月16日提出

要旨	4
第一章 バクテリアのコロニーの呈色度を測定する手法の開発	6
I. 序論	6
バクテリアを用いた呈色実験の歴史	6
バイオフィルムの概要	9
バイオフィルム合成ネットワークを調べる意義	14
本研究の目的	14
II. 材料と方法	15
菌株	15
培地	16
培養条件	16
SDS-PAGEとウェスタンブロット用の各種試薬	17
CsgAタンパク質の量を測定するウェスタンブロットまでの流れ	20
呈色したコロニーの撮影と平均のRGB値の算出	20
RGB色空間からHSV色空間への変換	21
III. 結果	22
RGB色空間における色の表現方法	22
実験の流れ	24
コロニーの実際のRGB値の検出	25
RGB値からHSV値への変換	27
他の色の検出と呈色値の算出	30
IV. 考察	33
第二章 ハイスループット化への検討とデータの信頼性	35
I. 序論	35
ハイスループットスクリーニング (High-throughput screening) 技術とは	35
大規模表現型解析を見据えたゲノムワイドなクローンセットの作製と表現型解析のため の手法開発	36
Colony-liveの開発	38
Colony-liveの課題	40
バイオフィルム合成ネットワークを網羅的に観測する意義	40
本研究の目的	40
II. 材料と方法	41
菌株	41
培地	41
培養条件	42
経時的なスキャニング	42
画像の明るさ調整	43
コロニーのエッジ検出と呈色値の算出	43
寒天培地上におけるコロニー面積のブレと補正	44
増殖解析	45
III. 結果	46
各コロニーの増殖曲線のモニタリング	46
データの再現性	47
各コロニーの生育ステージを考慮した呈色の比較	50
呈色スコアの算出	51
結果のvalidation	55
IV. 考察	56

第三章 呈色が著しく変化する遺伝子欠失株に関する情報の整理および新規 バイオフィーム形成関連遺伝子探索に向けた検討	59
I. 序論	59
低温ストレス、stationary phase突入によるバイオフィームの形成誘導	59
浸透圧によるバイオフィームの形成誘導	60
カタボライトリプレッションによるバイオフィームの形成阻害	60
Cyclic di-GMPによるバイオフィームの形成誘導	61
その他のバイオフィーム形成する要因	61
DAVIDによる候補株のアノテーション	64
本研究の目的	64
II. 材料と方法	65
菌株	65
培地	65
クリスタルバイオレット法によるバイオフィーム量の測定	65
呈色が顕著に変化した株のリスト	66
DAVID使用したtermとcategoryの紹介	66
DAVIDのクラスタリングの方法 (kappa係数に基づく類似度計算)	67
III. 結果	69
IV. 考察	87
謝辞	92
参考文献	93

要旨

細菌コロニーの呈色の違いは、種の同定や、細胞の生理状態の指標として有用な情報である。寒天培地に特定の基質や色素を加える方法は、簡便で感度も良く広く活用されてきた。しかし、ほとんどの場合、色の有る無しによる細胞の選択に利用されるなど、呈色の評価は目視のみで行われており、定量性を欠いた評価法に留まっていた。

染色剤の一種であるコンゴレッドは、バイオフィーム形成時特異的に合成される特殊な線毛 (Curli線毛) を赤く染める特性を持つ。コンゴレッド染色法はこの特性を利用し、主に細菌のバイオフィーム形成能を評価するために利用される手法である (Hammar, M., et.al, 1996)。バイオフィームは、自然環境下での細菌細胞の生存戦略の一つであり、微生物汚染や慢性的な感染症の原因のひとつであることから、医療分野で重要な研究対象である (Niba, E. T., et.al, 2007)。また、バイオフィームを用いて有用微生物を固体表面に固定化する技術の開発など、生物工学的な応用を目指した研究も進められており、その形成メカニズム解明の意義は大きい。Curli線毛の合成経路には多数の遺伝子が関与していると考えられており、バイオフィーム形成に向かう細胞内機能ネットワーク解明の意義は大きい。

本研究は、コロニーの画像情報から色を検出し、その呈色を定量する手法の確立を行い、その方法を用いてバイオフィーム合成経路の制御機能ネットワーク解明に取り組んだ。

最初に、コロニー呈色度の定量化手法を開発した。コンゴレッドを含んだ寒天培地 (Congo red agar : CRA) 上のコロニーが赤いほどバイオフィーム量が多い。「赤さ」の定量表現手法の検討を重ねた結果、以下の3つの手法を確立した。(1)市販のスキヤナでコロニーを反射光撮影し、各コロニーの重心座標を中心とした含む5ピクセル四方の領域 (25ピクセル) 内の平均RGB値を計算した。(2) 算出した平均RGB値を3つの色チャンネルに分解し、色相 (Hue)、彩度 (Saturation)、明度 (Value) からなるHSV色空間に変換した。(3) 彩度Sと明度Vの値を用いて、CRA上のコロニーの呈色度を数値化した。

計算値と、肉眼による呈色評価との相関を確認するため、既知の Curli線毛形成に関与する遺伝子 (csgA、csgD、cpxR)を用いた検証を行なった。CRAにおいて Δ csgA及び Δ csgDはバイオフィームの合成量が低下し (Blanco LP, et.al, 2012)、 Δ cpxR (Dudin O, et.al, 2014)では上昇することが知られているが、本手法による定量値は良い相関を示した。

明度Vは、色の強さだけでなく、その色の明るさ、暗さを示すことから、明度の違いによる遺伝子の分類が可能となった。

コンゴレッドはアミロイド状のCsgAタンパクに結合する事で呈色することから、得られた呈色定量値とCsgAタンパク質量と相関をウェスタンブロットにより解析したところ、非常に強い相関を示した。

本手法をゲノムワイドな呈色度定量解析手法への拡張を行った。大腸菌は、欠失された遺伝子により、その生育に大きなばらつきがあることが分かっている。コロニーの呈色は、細胞の生育に依存する為、呈色を細胞の生育による標準化手法の開発が必要である。CRAに高密度で大腸菌一遺伝子欠失株ライブラリー (Keio collection) を植菌 (1536株/plate) し、1時間ごとにコロニーを反射光撮影することでそれぞれの株の経時的な呈色度変化を記録し、以下の3点を実現する手法の確立を行った。(1)各コロニーの面積を計算し、ゴンペルツ曲線に回帰することで生育曲線を得る、(2)反射光によるコロニー表面の反射部分をノイズとして除去する方法の開発、(3)プレート上に

おけるコロニーの位置情報を考慮した測定データの標準化手法、の各方法を確立した。この手法により、生育不良な一部の株を除く3902株（全株の94.3%）の高密度に殖菌されたコロニースキャン画像からの生育曲線取得と呈色度曲線から、コロニーの増殖ステージを考慮した呈色定量手法を確立した。定常状態到達細胞量を100とした場合の90%到達時(t90%)において、各コロニーの呈色は飽和していたことから、t90%における呈色値を用い標準解析手法を確立した。本手法をNibaらにより報告されているCRAでの呈色結果との比較を行った結果、92%の株（13株中12株）において良い相関が得られた(Niba, E. T., et.al, 2007)。

本研究により、バイオフィルムの構成成分の一つであるCurli線毛の量を、コロニーの呈色定量値から高効率に取得可能にした。本研究で得られた新規バイオフィルム合成関連遺伝子の候補株の機能解析により、バイオフィルム合成機構の更なる解明に繋がるものである。

第一章 バクテリアのコロニーの呈色度を測定する手法の開発

I. 序論

バクテリアを用いた呈色実験の歴史

種の同定や、細胞の生理状態を観察するために、細胞の染色は広く用いられる手法である。古くは、1884年にデンマークの細菌学者であったGramが細菌を色素によって染め分けるグラム染色法を考案している(図1 A)。細菌がもつ細胞壁の構造の違いにより染色性が異なり、これを利用して細菌をグラム陰性菌とグラム陽性菌の2種に大別することができる。これまでに複数の変法が考案されてはいるが、現在においても細菌を分類する上で基礎となる技術である。また、寒天培地に特定の基質や色素を加えてコロニーの呈色の有無を評価する実験系は、手順が簡便であることから多くの場面で採用されている。DAPI(4',6-diamidino-2-phenylindole)は細胞内のDNAを染色に用いられる蛍光色素であり、膜透過性があるため生細胞にも使用できる(図1 B)。米国のbiolog社が発売したPhenotype Microarray (PM) は酸化還元反応試薬として tetrazolium と、様々な炭素源があらかじめ充填された96穴のマイクロタイタープレートを提供している。プレートのウェルに菌を植菌し培養すると、菌の増殖を tetrazolium の発色によってモニターすることが出来る(図1 C)。

1905年にイギリスの細菌学者MacConkeyが発表したマッコンキー培地は腸内細菌を分離するために開発された(Macconkey, 1905)。同培地は胆汁酸およびクリスタルバイオレットによる菌生育の抑制作用を利用しており、大腸菌群以外は生育できないように設計されている。さらに培地組成中にはラクトースとpH指示薬としてニュートラルレッドが含まれている。ラクトースを発酵分解できる大腸菌群由来のコロニーは、pH低下によるニュートラルレッドの反応で赤く呈色するが、サルモネラ属はラクトースを分解できないためコロニーは呈色しない。この仕組みを利用して、検体内のグラム陰性菌を呈色度合いによって分類することが可能である(図1 D)。このようにコロニー呈色反応の評価系は初期の細菌分類学や臨床の現場で主に利用されてきたが、分子遺伝学の発展に伴い遺伝子工学の場面でも用いられてくるようになってくる。1985年にMessingらによって構築されたpUC19(Yanisch-Perron, Vieira, & Messing, 1985)は*lacZ*遺伝子の配列中にMulti Cloning Site (MCS) をもつクローニングベクターである。MCS中への目的遺伝子導入が成功していればX-galおよびIPTGを含む培地で白色のコロニー、導入失敗の場合は青色のコロニーとなる。この手法はBlue-White selectionと呼ばれており、遺伝子工学には欠かせない。*lacZ*遺伝子はタンパク質相互作用を調べる手法であるtwo-hybrid systemにも使われており、X-gal・IPTG存在下のコロニーの呈色の有無によって目的タンパク質とのタンパク質相互作用やその強度を目視で確認できる(図1 E)。なお、X-galとよく似た構造のいわゆるX-gal variantsは

これまでに多数開発されており、目的にあわせて検出する色を変えたり、目的遺伝子の発現量が少ない場合はX-galよりもsensitivityが高いvariantを選択することもできる(Heuermann & Cosgrove, 2001)。近年ではlycopene合成遺伝子である*crtEBI*をレポーター遺伝子として用いた細菌の代謝ネットワーク改変技術が報告されており(Alper, Miyaoku, & Stephanopoulos, 2005)、呈色反応を指標にした実験系の開発は盛んである。コンゴレッド染色法もこうしたコロニーの呈色反応を利用した実験系の一つであり、細菌学においては主にバイオフィーム形成能を判定するために用いられる。この手法は、染色剤の一つであるコンゴレッドがバイオフィーム形成時特異的に合成される特殊な線毛(Curli線毛)を赤く染色する特性に基づいている。このため、細菌をコンゴレッド培地(Congo red agar : CRA)上で生育させると、Curli線毛の合成量に応じてコロニーが赤く呈色する。本手法はバイオフィーム形成能を簡便に評価できることから、臨床の現場を中心にごく一般に用いられている。

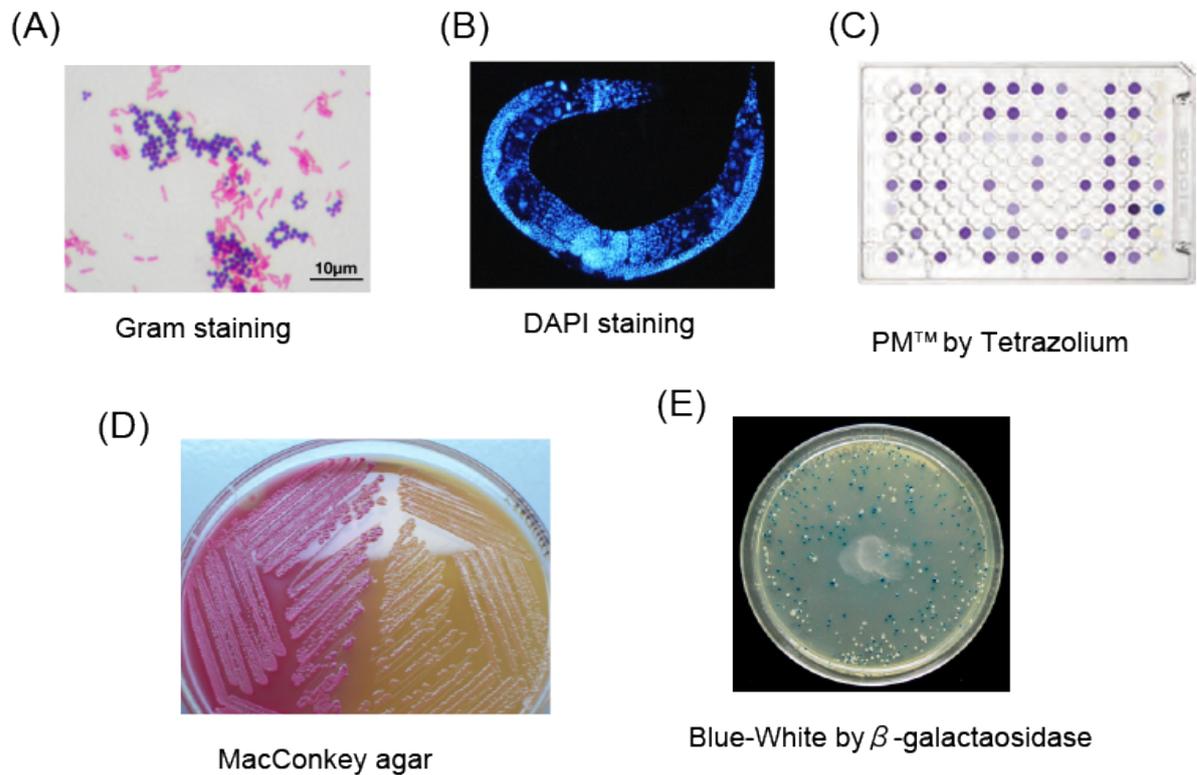


図1 生物学における様々な染色の例。A: グラム染色法によるグラム陽性菌とグラム陰性の鑑別。紫色に染色されているものがグラム陽性菌（黄色ブドウ球菌）、赤色に染色されているものがグラム陰性菌（大腸菌）。B: DAPI(4',6-diamidino-2-phenylindole)による*C. Elegans*のDNA染色。膜透過性があるため、生細胞でも染色することができる。C: biolog社が開発したPhenotype Microarray。Tetrazoliumは細胞が呼吸時する際の電子伝達系から作用を受けることで還元され紫に呈色する。D: MacConkey agarによる大腸菌群の鑑別。培地中にはラクトースとpH指示薬のニュートラルレッドが含まれている。大腸菌群がラクトースを分解し周囲のpHが低下すると赤く呈色する。E: Blue-White selectionによる遺伝子導入の成否判定。*lacZ*遺伝子の配列中にMulti Cloning Site (MCS) をもつpUC19ベクターが広く用いられている。*lacZ*遺伝子は β -galactosidaseをコードしており、寒天培地中のX-galを加水分解する。加水分解されたX-galは青色を呈する。目的の遺伝子がMCSに挿入されていれば*lacZ*遺伝子配列が途中で分断され、寒天培地中のX-galを代謝できずコロニーは白色となる。

バイオフィルムの概要

バイオフィルムとは、細菌が環境中の固体表面へ付着する際に形成する細胞外の高次構造体のことを指す。バイオフィルムは細菌にとってある種のシェルターのような役割を果たしており、以前から抗生物質や抗菌剤に対する耐性に寄与している (Costerton, Stewart, & Greenberg, 1999)。また、バイオフィルム内では多数の細菌種が入り交じって活動しており、これら細菌間での遺伝子の水平伝搬が指摘されている (Madsen, Burmølle, Hansen, & Sørensen, 2012)。より離れた生物間とのやり取りとしては、*Vibrio fischeri*とイカ的一种である*Euprymna scolopes*間のバイオフィルムを介した共生関係が研究されている (Shibata, Yip, Quirke, Ondrey, & Visick, 2012)。種によって様々だが、バイオフィルムは主に次の4つからなる構造物であると理解されている (図2)

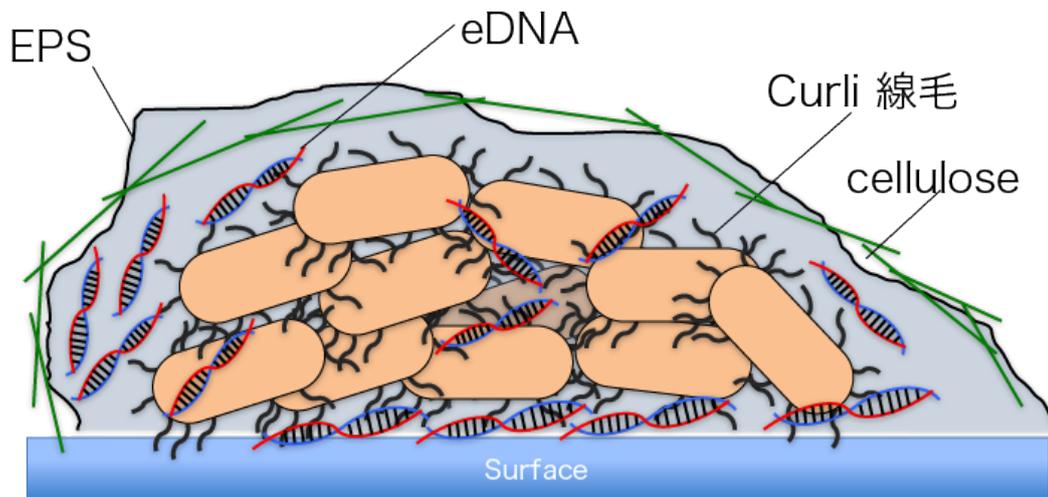
1) exopolysaccharides (EPS) は菌体外に分泌されるゲル状の多糖類の総称で、いくつかの種類が確認されている。アルギン酸は緑膿菌 (*Pseudomonas aeruginosa*) でよく知られるEPSの一つである。遺伝性疾患である嚢胞性繊維症 (Cystic Fibrosis: CF) 患者はこの菌が原因で細菌感染を引き起こすとされており、臨床の観点から研究が進められている (Wozniak et al., 2003)。大腸菌 (*Escherichia coli*) においてはcolanic acid、Poly-N-acetyl glucosamine (PGA) がそれぞれ知られている。colanic acidはgalactose、fructose、glucose、glucuronic acidからなる糖鎖で、グラム陰性菌細胞壁外膜構成成分であるLipopolysaccharide (LPS) の一部を担う。歴史的にはmucoidな表現型を示す株から発見されたことからM-antigenとも呼ばれ (Grant, Sutherland, & Wilkinson, 1969)、19遺伝子から構成される*wca(cps)* clusterによって合成される (Stevenson, Andrianopoulos, Hobbs, & Reeves, 1996)。またこれら遺伝子群は*rcs* (*regulator of capsule synthesis*) 遺伝子群に発現が調節されていること、さらにLonおよびClpYQプロテアーゼによる速やかな分解によって、細胞内のRcsAタンパク質の量が低く制御されることで、colanic acidの過度な生合成を抑制していることが分かっている。colanic acidはbiofilmの構成成分であるが、細胞の表面への付着、あるいは細胞同士の付着には寄与していない。PGAはN-acetylglucosamineが β -1,6結合で重合化したもので、もともとは*Staphylococcus epidermidis*や*S. Aureus*において細胞間同士の接着物質 (PIA: polysaccharide intercellular adhesin) として研究が進んでいた。後に大腸菌においても細胞間接着および表面への付着に直接寄与している物質であることが明らかになっている。PGAの生合成には*pgaABCD* オペロンが関与しており、その発現はCsr (carbon storage regulatory) によって負に制御されている。特にCsrAは*pgaABCD* のmRNA配列上の6カ所に直接結合することで転写を阻害しPGA生合成プロセスを抑制する。一方でCsrはglycolysis、motility、acetate metabolism、バイオフィルムの分解を正に制御し

ていることから、Csrは細胞がplanktonicな状態からバイオフィーム形成へシフトする際の重要な因子であると言えよう。

2) extracellular DNA(eDNA)は細胞外にあるDNAで、バイオフィーム中で絡まるようにして存在している。種によって様々だが、eDNAはバイオフィームを構成する他の物質と相互作用しており、EPSやDNA結合タンパク質を介した細胞表面、Type IV piliとの結合が確認されている。eDNAは様々な菌由来のバイオフィームに広く存在していることから、eDNAをターゲットとしたバイオフィーム形成阻害を目指した試みが進められており(Okshevsky, Regina, & Meyer, 2015)少なくとも大腸菌では、DNaseを利用したバイオフィーム形成阻害の実験が2例報告されている(Nijland, Hall, & Burgess, 2010; Tetz & Tetz, 2010)。eDNAはバイオフィーム内の溶菌細胞に由来するものや、種によっては細胞外に積極的にeDNAを分泌するものも知られている(Sanchez-Torres, Maeda, & Wood, 2010)。こういった細胞外核酸がバイオフィーム内で水平伝搬の一翼を担っている可能性は十分にあるかと思われる。

3) celluloseはglucoseが β -1,4-結合と呼ばれる結合様式で連結して出来ており、その生合成には*bcs* (*bacterial cellulose synthesis*) 遺伝子群が関与している。*bcs*遺伝子群は*yhjR-bcsQABZC*と*bcsEFG*の2つのオペロンに分類される。特に、*bcsAB*はinner membrane上で複合体を形成しウリジン二リン酸グルコース (UDP-glucose) からcelluloseを合成する機能を果たす。

4)細胞外タンパク質として主要な位置を占めているのがCurli線毛と呼ばれる特殊な線毛で、培養温度が30℃以下においてバイオフィーム形成時特異的に生合成されることが知られている。線毛を構成しているタンパク質の遺伝子群はオペロンを形成していて、*csgBAC*と*csgDEFG*の2つに分かれている(Hammar, Arnqvist, Bian, Olsén, & Normark, 1995) (図3,4A)。細胞外でCsgBを核にCsgAが自己重合化することで繊維状になる。これらタンパク質の細胞外への輸送はCsgGが担っている。CsgDは転写因子で、*csg*オペロンの他バイオフィームに関与する遺伝子群の発現調節に広く関わるタンパク質である。CsgDは菌が定常期に突入した際に発現が上昇するシグマ因子のRpoSの制御下にある。また、two component systemの一つであるCpxRによって通常負の制御がされている。Curli線毛の構造解析 β -sheet richなアミロイド様繊維の特徴をもつ(Wang, Smith, Jones, & Chapman, 2007; Wang, Zhou, Ren, Hammer, & Chapman, 2010)。染色剤の一つであるコンゴレッドはアミロイド繊維の β -sheet構造に優先的に結合し、Curli線毛を赤く染める特徴を有している。このことを利用して、Curli線毛の生合成の有無をコンゴレッドによる呈色の程度から言及する研究例が存在する(Eisert, Felau, & Brown, 2006; Khurana, Uversky, Nielsen, & Fink, 2001) (図3、図4 BC、図5)。



EPS: Extracellular PolySaccharides
 eDNA: Extracellular DNA

図2 バイオフィルムを構成する成分の例。細菌は環境中の表面に付着した後、自らが凝集するために、Curli線毛とよばれる短い線毛を合成する。また、Curli線毛の間を埋めるようにEPS (Extracellular PolySaccharides) とよばれるゲル状の物質を分泌する。これはバイオフィルム全体を乾燥から防ぐ役目も果たす。eDNA(Extracellular DNA)はバイオフィルム内で足場の役目を果たす。celluloseはバイオフィルム外部に展開し、機械的強度をもたらす。

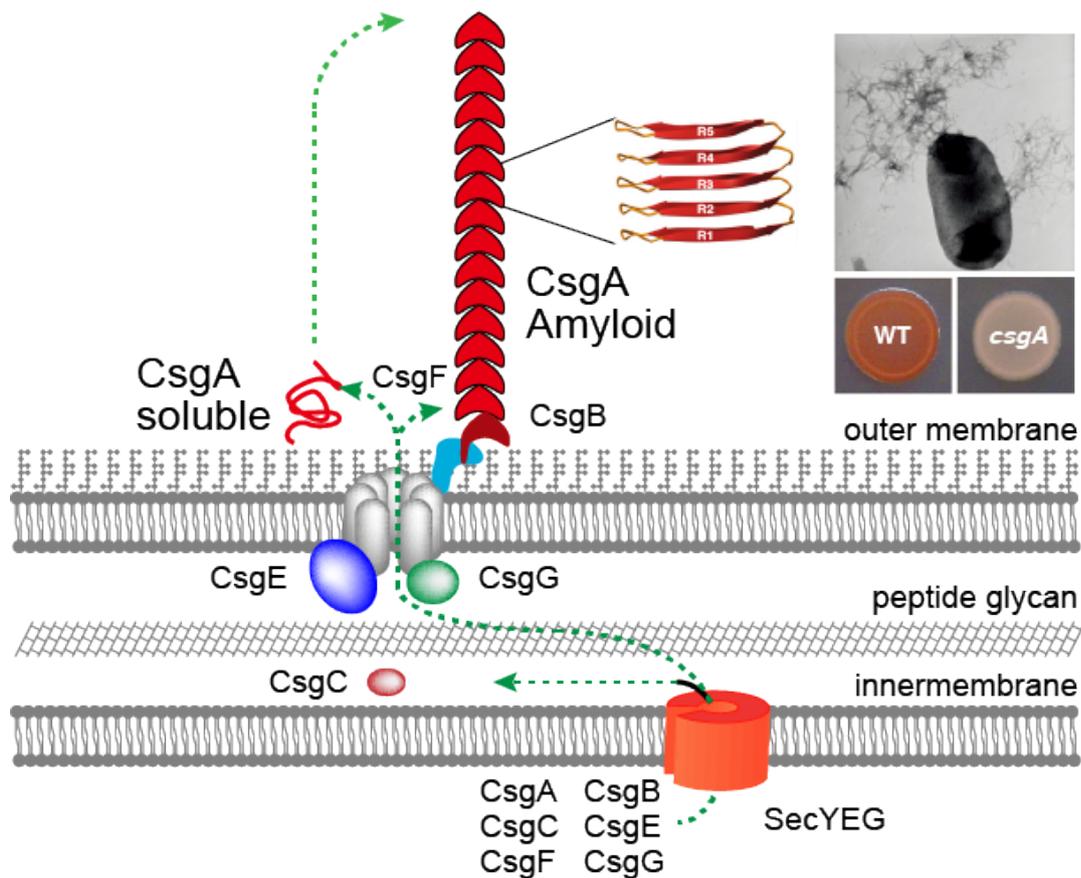


図3 Curli線毛の形成のモデル図。Curli線毛の主成分はCsgAタンパク質であり、細胞外で自己重合化することでアミロイド様の繊維を形成する。また、この重合開始にはCsgBタンパク質が必要で、CsgAポリマーの核となっている。

Curli線毛のマスターレギュレーターであるCsgDを除いて、全てのCsgタンパク質はSec依存性の移行シグナルを持つ。Curli線毛の構成成分の細胞外分泌は、独自のouter membrane secretion systemで達成されている。CsgGはporeをouter membraneに形成し、CsgEおよびCsgFと直接結合した状態で存在する。CsgEはペリプラズム側に局在し、可溶性CsgAタンパク質のCsgGを介した分泌に寄与している。一方、CsgFは細胞表面に露出しており、CsgBのouter membraneへのアンカリングに寄与している。

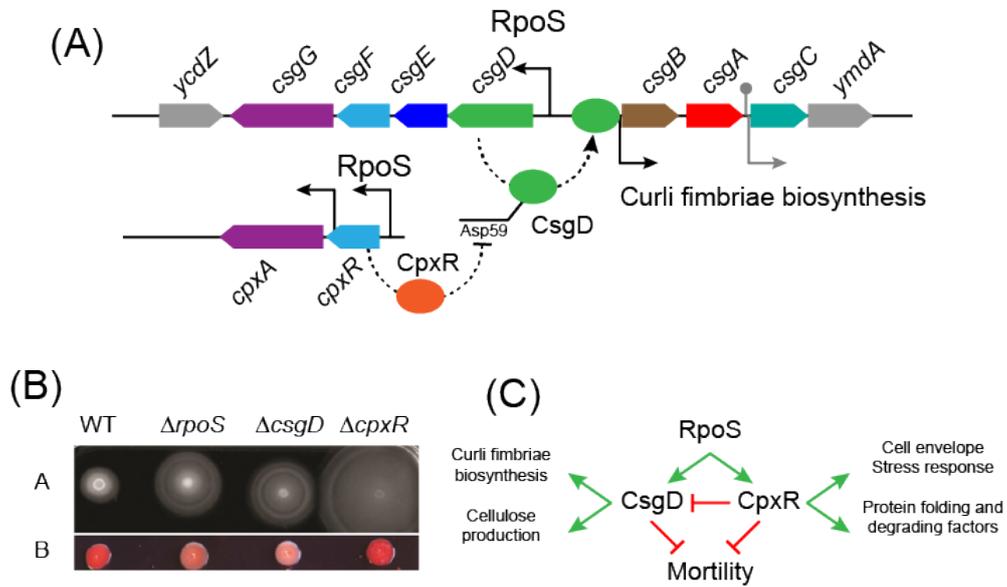


図4 *csg*遺伝子群とその制御関係の図。A: *csg*遺伝子群が構成するオペロン。*csg*遺伝子群は7つの遺伝子からなり、*csgDEFG*および*csgBAC*の2つのオペロンを構成する。*CsgD*はCurli線毛合成のマスターレギュレーターであり*csgBAC*オペロンの転写に必要である。*CsgD*は菌が定常期に突入した際に発現が上昇するシグマ因子のRpoSの制御下にある。また、two component systemの一つであるCpxRによって通常負の制御がされている。B、C:Curli線毛合成に関与する遺伝子破壊株のmotilityとコンゴレッド存在下における呈色の違い。

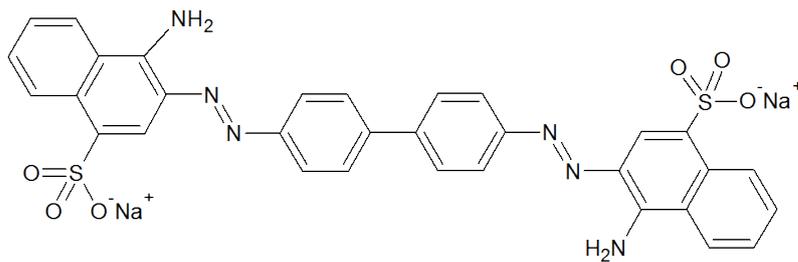


図5 コンゴレッドの化学式。Curli線毛の主成分であるCsgAのβシート構造に吸着し赤く染める特徴をもつ。臨床の現場においては内臓器官や脳に沈着したアミロイド繊維を染色する病理診断等に用いられる。

バイオフィーム合成ネットワークを調べる意義

バイオフィームを形成することで、細菌は薬剤や宿主免疫応答に対して抵抗性を示す。この特性上、バイオフィームは医療器具の微生物汚染や慢性的な感染症の一因と認識されているほか、水処理施設ではろ過装置の膜に発生するバイオフィームの除去が課題になっている。一方、バイオフィームを利用し有用微生物を固体表面に固定化させ、産業に活用しようとする動きがある。近年では、汚染された環境の修復技術である、バイオレメディエーションへの積極的な応用展開が議論される(ref)など、バイオフィームは形成の防止・促進の両視点から注目を集めている。

本研究の目的

コロニーの呈色状態の観察は研究者の肉眼に頼ったものであり、客観性を欠く。また、それが理由となって説得力に欠ける実験系に留まっていた。これはひとえに、呈色度を数値化するための有効なユーザーインターフェースの開発が遅れていたことに尽きる。一方で、遺伝子発現や細胞の生理状態のモニターにコンゴレッド法をはじめとした、コロニーの呈色や発光が利用されてきた歴史は長い。したがって、呈色実験の結果を肉眼ではなく計算機を用いて解析する環境を開発する研究はこれまで行われてきた実験系の価値を高めるという意味でインパクトがあると考えられる。こうした点を背景とし画像解析技術に基づくコロニーの呈色度の数値化に関する研究を着想した。その具体例として臨床的・工業的に研究の価値が高いと思われるバイオフィーム形成の有無を呈色度から明らかにすることを目的として、まずはその基礎的な系の構築を目指した。

II. 材料と方法

菌株

遺伝子欠失株は全てKeio collectionを用いた。また、野生株 (WT) としてKeio collection(Baba et al., 2006)の親株であるBW25113株を使用した (表1)。

Strains	Description	Pigmentation	Source
BW25113	Keio Collection WT(<i>rrnB3</i> Δ <i>lacZ4787</i> <i>hsdR514</i> Δ <i>araBAD567</i> Δ <i>rhaBAD568</i> <i>rph-1</i>)	+	Baba et al., 2006)
<i>csgA</i>	BW25113 Δ <i>csgA</i>	-	Baba et al., 2006)
<i>csgD</i>	BW25113 Δ <i>csgD</i>	-	Baba et al., 2006)
<i>cpxR</i>	BW25113 Δ <i>cpxR</i>	-	Baba et al., 2006)
<i>rpoS</i>	BW25113 Δ <i>rpoS</i>	++	Baba et al., 2006)

表1 本研究で使用した株と呈色の有無。これら遺伝子はバイオフィーム形成に関与していることが既に報告されている。CRA上での呈色の有無はWTを基準に+と-で表現されている。*cpxR*はWTよりも呈色がより亢進することが分かっているため、++とした。

青色色素タンパク質の発現実験にはpACTRham-aeBlue(株式会社堀場製作所 中山秀喜 博士より譲渡)を保持するBW25113株を用いた。pACTRham-aeBlueはイソギンチャク的一种*Actinia equina*由来の青色色素タンパク質をコードしたaeBlueがクローニングされており、同遺伝子はラムノースプロモーターによって発現制御されている (図6)。

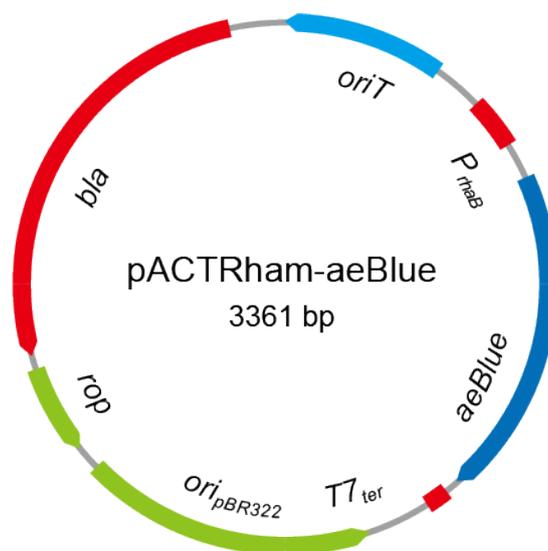


図6 pACTRham-aeBlueの構造。

培地

本章において研究に使用した培地組成を以下に示す。全ての培地は121°C、20分の高圧滅菌処理後、終濃度30 ug/L カナマイシン、または終濃度50 ug/Lアンピシリンを添加した。LB寒天培地の場合は7.5g/Lの寒天を混ぜて使用した。

• LB培地

Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	10 g/L

• YESCA寒天培地(Hammar, Bian, & Normark, 1996)

Cazamino acid	10 g/L
Yeast Extract	1 g/L
Agar	20 g/L

• CRA(Congo red agar) plate

滅菌後のYESCA培地に終濃度0.004 % コンゴレッドを添加	
HEPES pH7.4	25 mM

培養条件

スポットティング実験に用いる株の前培養は全て5 ml LB培地を用い、30°Cで好氣的に16 - 18時間行った。培養後、培養液を96-well plateに100 µlずつ分注し、RoToR (Singer Instrument Co.Ltd.、英国) を用いてCRA plateに培養液をスポットし、28°Cで好氣的に32時間培養した。

青色色素タンパク質の発現実験に用いた株は37°Cで好氣的に16-18時間前培養し、培養後の菌液をラムノースの濃度勾配をつけたLB培地に2 uLずつスポットした。スポット後、37°Cで好氣的に24時間培養した。

SDS-PAGEとウェスタンブロット用の各種試薬

・ SDS-PAGE用バッファー

1M Tris-HCl(pH6.8、 pH8.8)

Tris	121.2 g/L
HCl	適宜 (pH調整用)
H ₂ O	1Lにfill up

・ 10% SDS溶液(500 ml)

SDS	50g (終濃度 10% w/v)
滅菌水	500ml にfill up

・ 30% Acrylamide ・ 0.8%Bis混合液 (300 ml)

Acrylamide	90g(終濃度 30% w/v)
Bis (N,N'-Methylenebisacrylamide)	2.4g(終濃度 0.8%)
滅菌水	300 mlにfill up

・ 10%APS溶液 (10 ml)

APS (Ammonium persulfate)	1g(終濃度 10% w/v)
滅菌水	10ml にfill up

- 2xSDS Sample buffer(50 ml)

0.25M Tris-HCl(pH6.8)	25ml
SDS	2g (終濃度 4% w/v)
Sucrose	5g(終濃度 10%w/v)
1%BPB	125 μ l (適宜変更可)
超純水	45 mlにfill up

- 1%BPB溶液 (10 ml)

BPB (bromo phenol blue)	0.1g (終濃度 1% w/v)
100% EtOH	10mlにfill up

- 10x Running buffer

Tris	30g
Glycine	144g
10%SDS	100ml
超純水	1Lにfill up

- 15% Separating Gel(30 ml、large gel 2枚用)

1M Tris-HCl(pH8.8)	12 ml
60%Glycerol	3ml
超純水	0.2ml
10%SDS	0.3ml
30%AA/Bis	15 ml
10% APS	0.225 ml
TEMED	0.035 ml

- Stacking Gel(16.3 ml)

1M Tris-HCl(pH6.8)	1ml
超純水	12ml
10%SDS	0.16ml
30%AA/Bis	0.12ml
TEMED	0.02ml

<ul style="list-style-type: none"> ・ウェスタンブロット用buffer 	
10x Transfer buffer(pH8.6)	
Glycine	144g(終濃度1920mM)
Tris	30g(終濃度250mM)
超純水	900 mlにfill up
<ul style="list-style-type: none"> ・ 1x Transfer buffer 	
10x Transfer buffer	100ml
MeOH(またはEtOH)	100ml(終濃度 10%v/v)
超純水	800ml
<ul style="list-style-type: none"> ・ 10xTBS(pH7.4) 	
NaCl	80g(終濃度 1.38M)
KCl	2g (終濃度 0.027M)
Tris	30g(終濃度 0.25M)
超純水	1Lにfill up
HCl	適宜 (pH調整用)
<ul style="list-style-type: none"> ・ 1xTBS・T 	
10xTBS	100ml
Tween20	1ml
超純水	1Lにfill up
<ul style="list-style-type: none"> ・ Kpi buffer(pH7.2、 500 ml) 	
KH ₂ PO ₄	1.97g(終濃度28.9mM)
K ₂ HPO ₄	1.84g (終濃度21.1mM)
超純水	500mlにfill up
<ul style="list-style-type: none"> ・ 0.25% CBB溶液 (300 ml) 	
CBB (Coomassie Brilliant Blue R-250)	750mg(終濃度0.25% w/v)
MeOH	90ml (終濃度30% v/v)
Acetic Acid	30ml(終濃度10% v/v)
超純水	300mlにfill up

・脱色液 (500ml)

MeOH	150ml(終濃度30% v/v)
Acetic Acid	50ml(終濃度10%v/v)
超純水	500mlにfill up

CsgAタンパク質の量を測定するウェスタンブロットまでの流れ

あらかじめCRA上で32時間培養した菌のコロニーを回収し、1ml Kpi bufferに懸濁し、OD600が1.0に揃えた。その後、遠心 (15,000 rpm、3分) し、上清をすて、70 μ lの90%ギ酸を加えた。再懸濁後10分間室温で静置した。その後SpeedVacで遠心濃縮した (45°C、30分)。濃縮後、150 μ lの2xSDS Sample buffer に再懸濁し、98°C、20分でインキュベーションした。その後、用意していた15%Acrylamide Gel(large gel)し、サンプルをアプライしてSDS電気泳動を行った (100mA低電流、110分)。泳動後、セミプロッターを用いてPVDF膜に転写した (30分、24V低電圧)。
バンド検出にはCan Get Signal(東洋紡)とBCIP-NBT溶液キット (ナカライテスク) のマニュアルに従った。1次抗体にはrabbit anti-CsgA polyclonal antibody、およびrabbit anti-ftsZ polyclonal antibody (慈恵医大・医学部・細菌学講座 杉本真也講師より譲渡)、2次抗体にgoat anti-rabbit IgG-AP (Santa Cruz Biotechnology社、米国) をそれぞれ5000倍希釈で用いた。

呈色したコロニーの撮影と平均のRGB値の算出

培養後のCRA plateを市販のスキャナとスキャナ専用ソフト (スキャナの型番 : GT-970、専用ソフト : EPSON scan、共にセイコーエプソン株式会社) で反射光撮影した。スキャナで取得した画像から各コロニーの重心座標を含む $5 \times 5 = 25$ pixels領域内の平均RGB値を算出した。

RGB色空間からHSV色空間への変換

算出した平均のRGB値を下に、HSV値（色相、彩度、明度）に変換した(図7 A)。変換には以下の式に従った(図7 B)。Hは色相環において赤を0度としたときの角度で表現される。Hは360段階、SとVは256段階で表す。色相環の中心を彩度ゼロの点として半径256の円盤を考え、それが明度の段階ごとに256枚並んだ円柱を考えるとわかりやすい。MaxとMinをそれぞれRGBの3つの最大値と最小値とすると、VはMax - Minと定義される。Vが大きいほど明るいことを意味する。また、最大値と最小値の差を最大値で割った値がSになる。Sが大きいほど色が鮮やかになる。Hは「何色っぽい」のかを360段階で表す。計算方法はVに入るRGBの値によって3通りある。

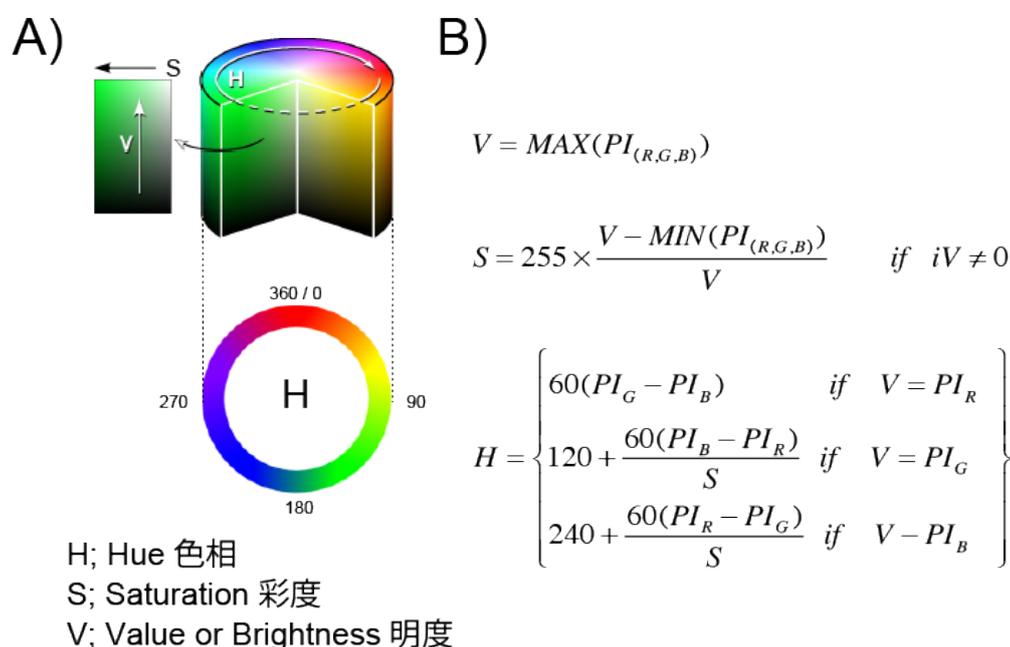


図7 HSV色空間の概念図とRGB色空間からHSV色空間への変換式。

III. 結果

RGB色空間における色の表現方法

一般にコンピュータ上では色を表現する方法として24ビットカラーと呼ばれる方式が採用されている。この方式では、1ピクセルに24ビットの色情報を持たせており最大で16,777,216色を表現することができる。また、1ピクセルは赤・緑・青の3色チャンネルが8ビットずつ割り振っており、チャンネルごとに0から255までの256階調で三原色を表現している（図8 A）。この数値のことをピクセル輝度（Pixel Intensity: PI）と呼ぶ。ある物体が赤いということは、赤の波長が反射され、緑・青の波長は吸収されていることを意味する。これをコンピュータ上の24ビットカラー方式で表現すると赤チャンネルのPIが255、他のチャンネルが0、すなわちRGB : 255,0,0となる。あるいは赤色に近い色は赤チャンネルに対して他の2つの色チャンネルのPIが低いことになる。全てのチャンネルのPIが255だった場合、白（RGB : 255,255,255）となる。色を重ねると光の量が増えて明るさが増すので、このような色の混ぜ方を加法混色という（図8 B）。

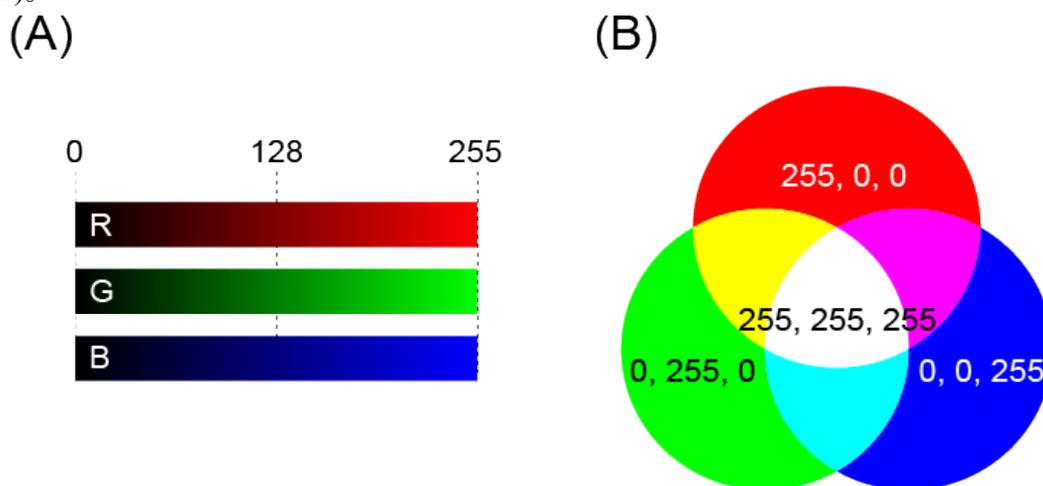


図8 RGB色空間に基づいたコンピュータにおける色情報の扱い。A:コンピュータ上での色情報の扱い方。24ビットカラー方式は3つの色チャンネル(R、G、B)が各0-255までの256段階の階調をもつ。B:RGBの発色方式。光の3原色である赤 (Red)、緑 (Green)、青 (Blue) によって色を表現する。赤のPixel Intensity(PI)は[255,0,0]、緑は[0,255,0]、青が[0,0,255]で表現される。すべての色が混ざった色が白となりPIは[255,255,255]となる。このような色の混ぜ方を加法混色といい、色が混ざるほどに明るさが増し、白に近づく。

色検出におけるRGB色空間の問題点とHSV色空間へ変換する利点

RGB色空間による色の表現は直感的に理解しやすいが、色検出を行おうとする際は煩雑であるため、その問題点について検討した。RGB色空間では色を重ねると光の量が増え明るくなる加法混色に従っている。つまり、同じ色であっても明るさや鮮やかさが異なる色を表現するためにはRGB全ての光の強さを細かく調整しなければならない(図9)。一方で、色検出するために都合の良い色空間としてHSV(色相、彩度、明度)色空間と呼ばれるものがある(図6)。HSV色空間では同じ色(色相H)における色の变化を数字で捉えることが簡便になる。したがって、呈色の強度が比べやすいと考えられた。そこで、呈色したコロニーが生じている寒天培地をスキャナで取り込み、得られたRGB値をHSVに変換して呈色の強度を調べることにした。

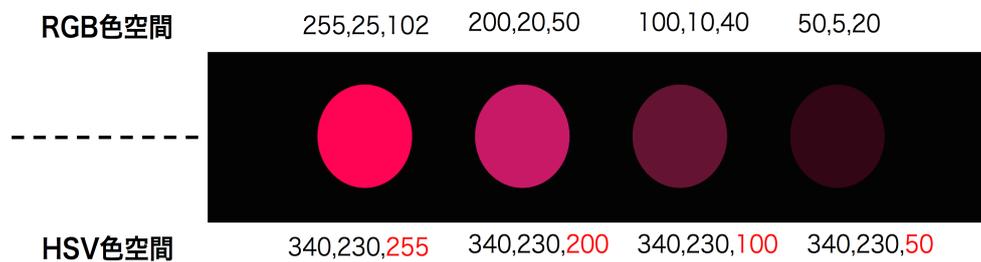


図9 色検出におけるRGB色空間とHSV色空間の違い。左から右にかけて明るさのみが異なる赤色のプロット。上がRGB色空間におけるR,G,Bそれぞれのピクセル輝度。下がHSV色空間に変換した際のH,S,Vの値。RGB色空間では3つのピクセル輝度が変動しているのに対し、HSV色空間では明度Vのみが変動している。

実験の流れ

コロニーの呈色を数値化するために、次の流れで実験を行った。図10に大まかな流れを示す。コンゴレッド寒天培地(CRA: CongoRedAgar)上に生じたコロニーは市販のスキヤナによって反射光撮影した。スキヤナによって検出される色のシグナルはRGB値である。すなわち、光源からの光を反射したコロニーの色はスキヤナに内蔵されたセンサーによってRGBの色チャンネルごとにそれぞれ検出される。検出されたRGB値の合成色が実際のコロニーの色としてコンピュータ上で再現される。HSV値でコロニーの呈色を評価するにあたり、まずスキヤナで検出されたRGB値を解析した。

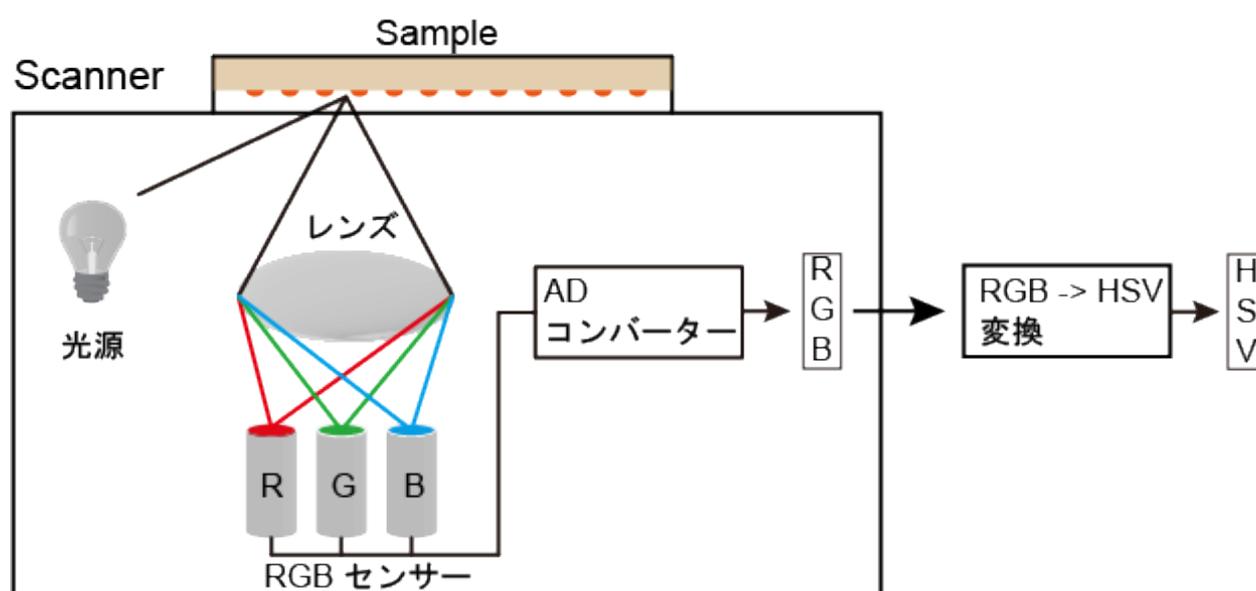


図10 スキヤナを用いたコロニー呈色の撮影とHSV値取得までの流れ。

コロニーの実際のRGB値の検出

24ビットカラーで撮影されたスキャナ画像は色分解することで、それぞれの色チャンネルのPIを調べることができる。色分解後の3つの色チャンネルにおけるPIのパターンを確認することで、実際のコロニーの写真との関係について調べた。実験には既知のバイオフィルム合成に関与している遺伝子の破壊株を4株選定した(表1)。WTとともにこれら遺伝子欠失株を培養液をCRAに5スポットずつ行い、32時間培養した後にスキャナを用いて培地上に生じたコロニーを撮影した(図11 A-a)。各コロニーの重心座標を中心に $5 \times 5 = 25$ ピクセルずつを切り取り、色チャンネルごとのPIを計算することでコロニーごとの色情報を計算した。計算された色情報は遺伝子欠失株ごとにまとめ、欠失株の平均の色を再構成した。再構成した色は元画像と一致しており、抽出したPIに問題がないことを確認できた(図11 A-b)。

各欠失株の色チャンネルの平均PIをチャンネル別にプロットした(図11 B、表2)。赤チャンネルのPIは全てのコロニーにおいてそれほど変化は見られなかった。一方、緑・青チャンネルのPIは同じような傾向を示していた。5つの株で最も緑・青チャンネルのPIが強かったのが*csuD*であった。赤チャンネルと緑・青チャンネルのPIの差が小さいということは、コロニーが白に近い色を呈していることを示す。元画像でもコロニーは白いことが確認できる。反対に、最も緑・青チャンネルのPIが低かったのは*cpxR*だった。赤チャンネルに対して緑・青チャンネルのPIが低いということは、観察している画像領域が赤色であることを示している。実際、*cpxR*のコロニーは5つの株のなかで最も赤色の呈色が亢進していた。

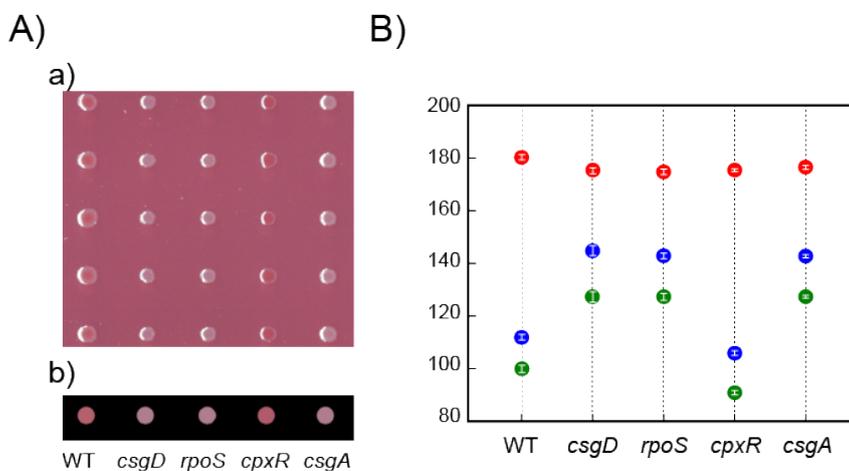


図11 培地上に生じたコロニーから検出されたRGB値。(A-a) Curli線毛合成関連遺伝子の破壊株のコロニー。各株（左からWT : BW25113、*csgD*、*rpoS*、*cpxR*、*csgA*）の培養液2 μ lを寒天培地上に5つスポットし、32時間28°Cで培養した。培養後スキャナで反射光撮影した。(A-b) 各株の平均のRGB値で再構成されたコロニーのイラスト。(B) 各株の平均の色チャンネル別のピクセル輝度 (Pixel Intensity: PI)。赤 : RチャンネルにおけるPI、緑 : GチャンネルにおけるPI、青 : BチャンネルにおけるPI。エラーバーは各チャンネルにおけるPIの標準偏差を示す。

Table RGB values of control strains

gene	Mean			SD		
	R	G	B	R	G	B
WT	179.238	98.497	111.125	0.989	1.395	1.086
<i>csgD</i>	174.122	125.588	143.575	1.098	2.003	1.875
<i>rpoS</i>	173.708	125.578	141.615	1.094	1.273	1.043
<i>cpxR</i>	174.337	89.457	104.523	0.562	0.809	0.864
<i>csgA</i>	175.112	126.062	141.523	0.733	0.503	0.598

SD; standard deviation

表2 測定された各遺伝子欠失株の平均RGB値と標準偏差。各遺伝子欠失株ごとに5つのコロニーの平均RGB値を算出している。

RGB値からHSV値への変換

得られた欠失株のRGB値をHSV値（色相、彩度、明度）に変換し、コロニーの呈色を解析した。まず、コロニーの色相Hについて確認した。色相は色の種類を説明するパラメータである。HSV色空間において、色相Hは0から360°の円状に展開される。観察した全てのコロニーの色相Hは約30°のあたりに集中していた(図12 A)。赤色の色相は300~60°の範囲内とされているので、観察した全てのコロニーの色は赤色の成分をもっており、呈色の違いは彩度Sと明度Vの違いによるものだということがわかった。

次に、彩度Sと明度Vの関係について調べた。彩度Sは色の鮮やかさを説明するパラメータ、明度Vは色の明るさを説明するパラメータである。WTと比して、*csgA*、*csgD*、*rpoS*ではいずれも低い彩度Sを認めた（*csgA* : 0.622、*csgD* : 0.621、*rpoS* : 0.616、いずれもWTを1.0としたとき）。一方で、*cpxR*では彩度Sの値は1.079と、WTよりもわずかに高い値を示した（図12 B、表3）。これらの値はいずれもコロニーの写真から確認できる赤色の強弱関係を上手く表していた。また、明度Vについて見てみると、*cpxR*は0.973と、WTに対してわずかであるが低い値であった。これはWTと比べ、より暗い赤色であることを示している。過去の文献で、CRA上におけるWTと*cpxR*の呈色の強さはそれぞれred、およびdark redといった表現で報告されており、今回の結果と概ね一致していた。

呈色の強さとCurli線毛の主成分であるCsgAタンパク質の量の間を調べた。呈色しているコロニーをKpi bufferに懸濁し、OD600 = 1.0になるように調整したのち、遠心を行って上清を取り除きペレットのみを回収した。CsgAタンパク質を単量体にするため、ギ酸で処理したのちSDS-PAGEで電気泳動を行いウェスタンブロットウェスタンブロット法によるCsgAタンパク質の検出を行った。インターナルコントロールとして、FtsZを検出している。彩度Sおよび明度Vの値と、CsgAタンパク質の量を比較したところ、特に彩度Sと実際のCsgAタンパク質の量が一致していることを確認した。これら結果から、検出したコロニーの呈色の値は本実験でターゲットとしているCsgAタンパク質の発現量の変化を捉えていると考えられた。

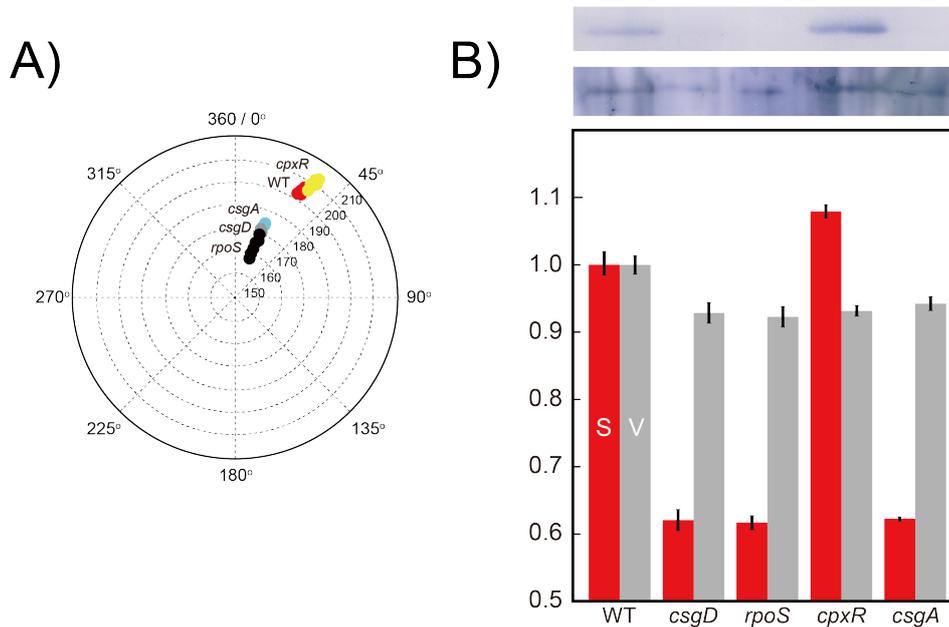


図12 RGB値をHSV値に変換した後の値。A: 各遺伝子欠失株の色相Hのと彩度Sの分布。300~60°の範囲内であればそのコロニーの色相が「赤」であることを意味している。彩度Sは中心からの距離で表現されている。B: 各遺伝子欠失株の彩度Sおよび明度Vおよびコロニーから検出されたCsgAタンパク質とFtsZタンパク質。いずれの値もWTを1.0としたときの値に変換されている。ウェスタンブロットによるCsgAの検出のため、各欠失株のコロニーをKpi bufferに懸濁後、懸濁液OD600が1.0にそろえるように調整した。CsgAは細胞外で重合化することでCurli線毛を形成していることから、そのまま泳動すると正しい分子量(単量体の場合、15.049 kDa)が分からないため、ギ酸によって単量体にした。エラーバーは各欠失株における彩度Sおよび明度Vの標準偏差を示す。

Table S and V values of control strains

gene	nS	nV	stdS	stdV
WT	1.000	1.000	0.018	0.006
<i>csgD</i>	0.621	0.971	0.016	0.006
<i>rpoS</i>	0.616	0.969	0.010	0.006
<i>cpxR</i>	1.079	0.973	0.010	0.003
<i>csgA</i>	0.622	0.977	0.003	0.004

nS; normalized S value

nV; normalized V value

stdS; standard deviation of normalized S

stdV; standard deviation of normalized V

表3 測定された各遺伝子欠失株の平均の彩度S、明度Vおよびそれぞれの標準偏差。各遺伝子欠失株ごとに5つのコロニーの平均のS、V値を算出している。

他の色の検出と呈色値の算出

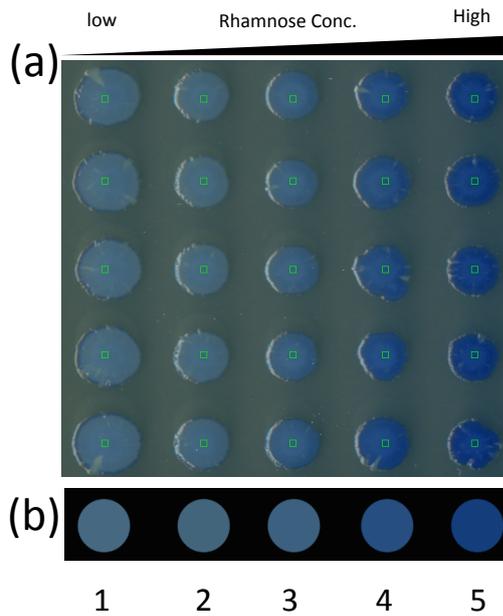
本手法が他の色の検出にも応用可能か検証するため、イソギンチャク (*Actinia equina*) 由来の青色色素タンパク質の発現量をラムノースで制御できるようにした大腸菌株 (BW25113株) を作製し、青色の強度をRGB色空間およびHSV色空間に基づいて評価した。

各コロニーの重心座標を中心に $5 \times 5 = 25$ ピクセルを切り出し、RGB色空間における赤・青・緑のPIを得た(図13 A-a,b)。色チャンネルごとにPIをプロットしたところ、全てのコロニーで青チャンネルのPIの変動はみられなかった。一方、赤チャンネルと緑チャンネルのPIは同じ傾向を示した(図13 B)。青チャンネルに対して、赤・緑チャンネルのPIが低いということは、観察している画像領域が青色であることを示していると考えられた。

次に、各色チャンネルのPIをHSV色空間に変換してコロニーの色の評価を行った。青色色素タンパク質を発現した全てのコロニーの色相H値は約 200° のあたりに集中していた(図14 A)。「青色」を示す色相は $180 \sim 300^\circ$ の範囲内とされており、観察したコロニーの色が青色の成分を持っていることが分かった。

彩度Sと明度Vの関係を調べたところ、特に彩度Sがラムノースの濃度依存的に変動しており、写真上のコロニーの色の変化をよく表現していた(図14 B)。このことから、本手法が赤色以外の別の色を検出できることが確認できた。

(A)



(B)

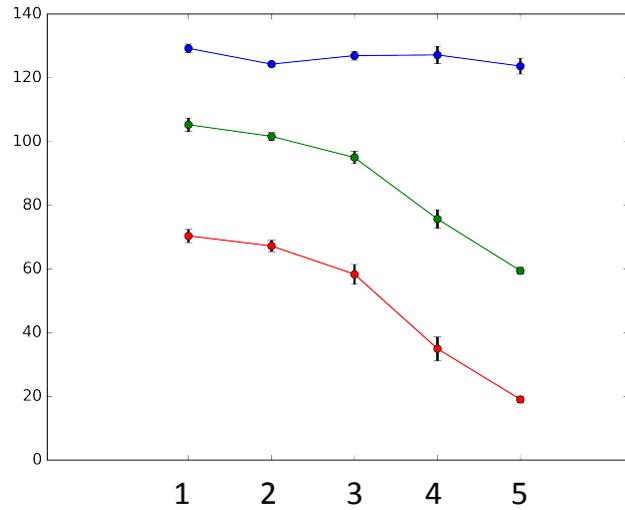


図13 培地上に生じたコロニーから検出されたRGB値。(A-a)青色色素タンパク質を発現した株のコロニー。同一の寒天培地で右に行くほどラムノースの濃度が濃くなるようにした。培養液2 μ lを寒天培地上に5つスポットし、24時間37°Cで培養した。培養後スキャナで反射光撮影した。(A-b) 縦5つのコロニーを同一のラムノース濃度で呈色したコロニーとみなし、平均のRGB値を計算した。得られた平均RGB値で再構成されたコロニーのイラスト。(B) 各株の平均の色チャンネル別のピクセル輝度 (Pixel Intensity: PI)。赤: RチャンネルにおけるPI、緑: GチャンネルにおけるPI、青: BチャンネルにおけるPI。エラーバーは各チャンネルにおけるPIの標準偏差を示す。

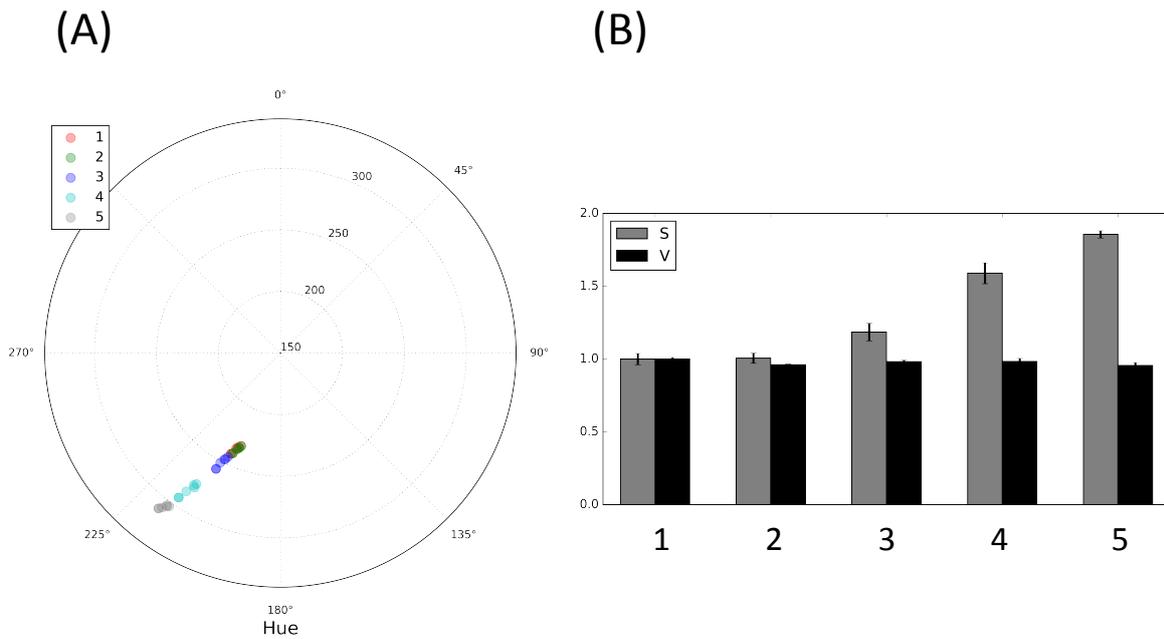


図14 RGB値をHSV値に変換した後の値。A: 各遺伝子欠失株の色相Hのと彩度Sの分布。180~300°の範囲内であればそのコロニーの色相が「青」であることを意味している。彩度Sは中心からの距離で表現されている。B: コロニーの色情報から計算された彩度Sおよび明度V。いずれの値もWTを1.0としたときの値に変換されている。

IV. 考察

Curli線毛の主成分はCsgAタンパク質であり、*csgA*遺伝子領域は*csgB*とともに*csgBA*オペロンを形成している。また、*csgD*はこの*csgBA*オペロンの転写因子として、当該オペロンの発現調整を司る。RpoSはシグマ因子 σ sとも呼ばれ、栄養飢餓やストレス、または細胞がstationary phaseに突入すると細胞内に蓄積することが知られている。また、RpoSと強調して働くsmall proteinとしてCrlがある。Curliの名前に由来するこのタンパク質は30°以下の培養条件で菌がstationary phaseに突入した際、細胞内に蓄積するが37°Cではそれが見られない(Bougourd, Lelong, & Geiselman, 2004)。蓄積したCrlはRpoSに結合し、その活性を亢進させる。Crl-RpoSによって発現が制御されている遺伝子は少なくとも40 遺伝子が確認されており(Pratt & Silhavy, 1998)、CsgDも直接RpoSの正の制御を受けている。

今回の実験で得られた5つの株の彩度S、明度Vの値を用いて散布図を描くことで各株の呈色の違いを可視化した(図15 A)。*csgA*、*csgD*、*rpoS*株はいずれもCRAにおいてWTと比べて著しく低い呈色を示すことは既に報告されていたことであったが、今回の測定系で改めて確認することができた(*csgA*:0.622、*csgD*:0.621、*rpoS*:0.616、WTの彩度Sを1.0としたときの各遺伝子欠失株の彩度S)。

*cpxAR*は膜ストレスに応答するtwo component signal transduction pathwayの一つである。CpxAは、外膜上のOmpAからシグナルを受け取り、細胞内のCpxRをリン酸化(CpxR-P)する。CpxR-Pは通常CsgDの発現を負に制御している(Ma & Wood, 2009)。したがって、*cpxR*ではCpxR-PによるCsgDの発現抑制がきかないためWTよりも呈色が亢進したと考えられる。実際に、今回の測定系でも*cpxR*の彩度Sは1.079と、WTと比べ亢進していたことを確認できた。また、今回はRGB値をHSV値に変換したことで、色の明るさ(明度V)をパラメータとして得ることが出来た。特に、*cpxR*の明度VはWTと比べ0.973とわずかに低い値を示した。過去の文献でCRA上におけるWTと*cpxR*の呈色の強さはそれぞれred、およびdark redといった表現で報告されている(図15 B)ことから、WTと*cpxR*の彩度Sと明度Vの値は観察者の色彩感覚を上手く説明している値であると言えるだろう。ウェスタンブロットによるCsgAタンパク質の量と彩度Sの値は一致しており、呈色の強さが目的のタンパク質の量を表す数値であることを確認できた。

HSV色空間に変換することで様々なコロニーの呈色に対応可能かを示すために、青色色素タンパク質を発現する株を用いて呈色の強度を測定した(図13、14)。色素タンパク質を用いた実験においても、本測定法によって色の種類と色の強度を数字にすることができた。HSV色空間という同一の指標で様々なコロニーの呈色を評価することは、例えば土壌や海水由来の雑多な環境からスクリーニングで呈色の異なる複数の菌種をオートマチックに検出するような用途への応用が期待できる。

これらの結果から、スキャナの写真からコロニーの呈色の強さを定量化することは十分可能であると判断できた。今回の実験ではコロニー表面で生じる光の反射の処理が課題として残った。光の反射はコロニーの呈色を正しく判定する上で障害となりうる。ゲノムワイドな呈色データの取得を目指すためには本測定系をハイスループット化することが必須条件だが、そのためにはコロニー表面の反射領域を適切に処理する必要があると思われた。そこで、次章ではこの点を念頭においた上で本測定系に基づいたゲノムワイド呈色データの取得を試みた。

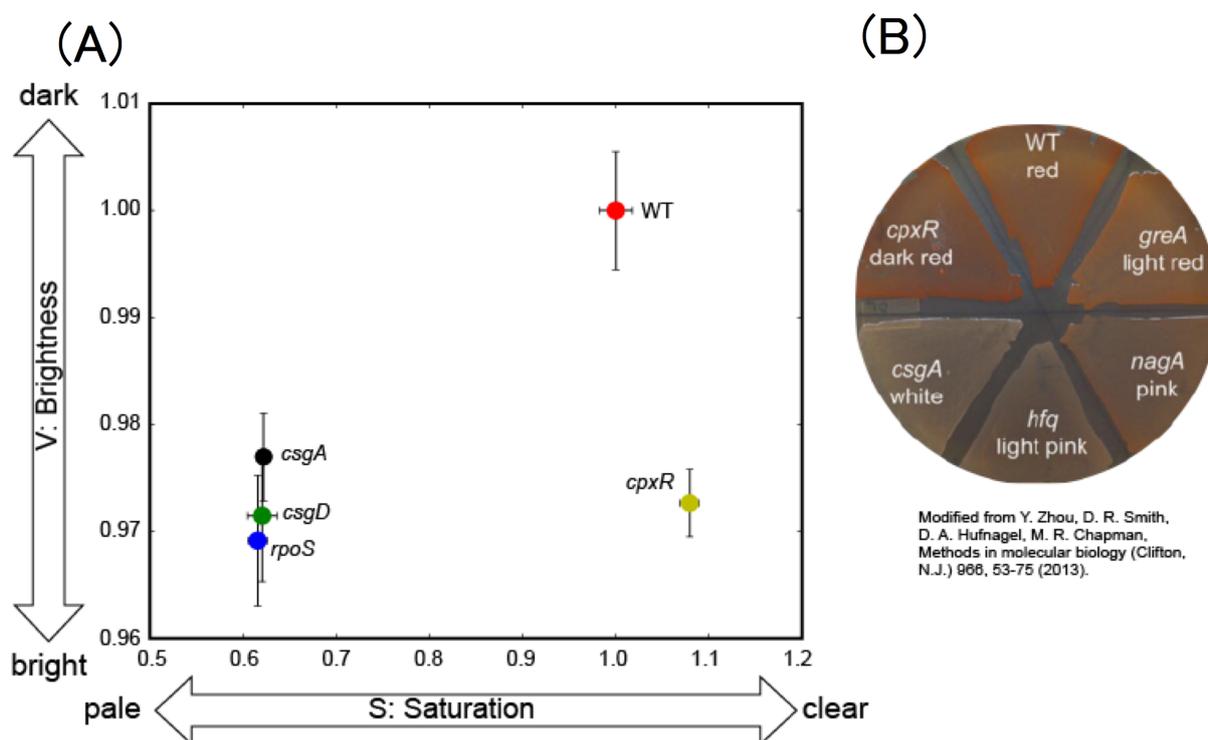


図15 コロニーの彩度Sおよび明度Vの分布。A: スポット実験に用いた株のコロニーの平均の彩度Sおよび明度Vの2次元プロット。x軸が彩度S、y軸が明度Vの分布を表す。WTと*cpxR*はそれぞれ赤く呈色したが (WTの彩度S値を1.0としたとき、*cpxR*が1.079)、両株の暗さは異なった (WTの明度V値を1.0としたとき、*cpxR*は0.973)。B: 過去の文献における呈色の違いを言及した例。WTはred、*cpxR*はdark redとそれぞれ表現されているが、今回の実験結果でもその表現型の違いを彩度Sと明度Vの数値として表現することができた。

第二章 ハイスループット化への検討とデータの信頼性

I. 序論

ハイスループットスクリーニング (High-throughput screening) 技術とは

ハイスループットスクリーニング技術とは、文字通り一度の実験で数万から数十万の検体の動態を観察することで目的の機能をもった検体を選択する手法であり、元々は製薬企業が創薬を目的として積極的に採用してきた技術である。創薬の現場では、生理活性が認められる既知の化合物を出発として、理論的にその構造の一部を改変していくことで、化合物の薬理作用の改良を目指した研究が日々行われている。こうした技術をドラッグデザインといい、ドラッグデザインの出発となる化合物はリード化合物とよばれる。例えば、1928年にイギリスの細菌学者であるAlexander Flemingによって、世界初の抗生物質であるペニシリンが発見され、以降ペニシリンをリード化合物として様々な派生化合物が化学合成されている。こうした派生化合物は一般に誘導体とよばれ、ペニシリン誘導体またはペニシリン系抗生物質は半合成および全合成されたものも含めて数多く存在する。ペニシリン同様、1943年Selman Abraham WaksmanとAlbert Schatzがストレプトマイシン、1945年Benjamin minge duggarがクロルテトラサイクリンを天然由来の抗生物質として発見された。これら抗生物質はそれ自身がリード化合物となって、その後の様々な誘導体の開発につながった。新規に創薬するにあたり、こういったリード化合物の発見は大変重要で各製薬企業が行っているハイスループットスクリーニングではしばしばヒット化合物として見いだされるが、その工程のほとんどが多数のサンプルからわずかな数のヒット化合物、あるいはそれになりうる候補を選択する作業に他ならない。つまり、スクリーニングの成功はかなりの割合で偶発的なものであり、運に支配されているとっていい。しがたって、スクリーニングの成功確率を向上させるためにはスクリーニング自体の工程を少しでも体系立てて、合理的に作業をすすめることが求められる。先に述べた抗生物質の発見事例も含め、スクリーニングの成功例の影には研究者がこれまで蓄積していた知識に基づく、ある種の勘による部分が多い。こうした勘に頼ることなく、目的に適合したヒット化合物をより高い確率で発見するためには、スクリーニングに供するサンプルの数を膨大にし、かたよりなく網羅性を担保する必要がある。ハイスループットスクリーニング技術とは、こうした点を背景として生まれ、「スクリーニング」というプロセスの合理化を主眼にして発展してきた技術と言える。これを達成するためにはスクリーニング専用開発された機械の使用、およびスクリーニングの結果を計算機で解析することが必須である。言い換えれば、これまで手作業で行ってきた実験あるいは作業を可能な限りバッチ処理できるように改変する必要がある。生物学においては、ヒトをはじめとした様々な生物種のゲノムが明らかになってきたこともあり、ゲノムワイドな規模の実験をデザインする基調が広がりを見せている。これに伴い、

ハイスループットスクリーニングに必要な機械やソフトウェア、それを使いこなすノウハウは製薬会社に留まらず、生物学の基礎研究の分野においても必要なものとして認識されるようになってきた。

大規模表現型解析を見据えたゲノムワイドなクローンセットの作製と表現型解析のための手法開発

大腸菌に先駆け、酵母 (*Saccharomyces cerevisiae*) の全ゲノム情報は1996年に決定された(Goffeau et al., 1996)。大腸菌のゲノムは1997年に決定されている(Winzler et al., 1999)。大腸菌と酵母のゲノム情報が決定されて以降、微生物を対象にした研究現場では、2000年代の初頭にかけて様々なモデル微生物におけるゲノムワイドなクローンセットの構築が精力的に行われた。構築されたセットを表4にまとめる ((Barker, Farha, & Brown, 2010)の表より改変)。大腸菌では2005年に全予測遺伝子のクローニングセットであるcloning library (Kitagawa et al., 2005)、次いで2006年に全予測遺伝子の一遺伝子欠失ライブラリーであるKeio collection(Baba et al., 2006) (図16)、gfp-promoter fusionライブラリーがそれぞれ構築されている。ゲノムワイドなクローンセットの完成は、それまで各研究者が個別で作製していた株に基づかない、菌株が統一された状態での議論を提供した。また、複数の菌種でシステムチックな遺伝子欠失株が作製された結果、生命維持のために必要最低限な遺伝子セット、いわゆる必須遺伝子の概念が産まれた。こうした点は、個別の遺伝子を対象にして進められてきた従来のアプローチでは得ることのできなかったことである。薬剤のターゲット遺伝子の同定や作用機序の解明を進めるために行われるchemical genomics研究においてもゲノムワイドなクローンセットが強力なツールとなった。薬剤が特定の物質の働きを阻害することと、遺伝子破壊による摂動は互いに相補的な関係であることを利用し、クローンセットに対して作用機序の不明な化合物を用いたゲノムワイドなchemical genomicsが報告されている。膨大な数にのぼる表現型の解析をシステムチックに進める手法の開発も併せて進められてきた。米国のBIOLOG社が発売したPhenotype Microarray (PM) は微生物を扱う分野における網羅的表現型解析のきっかけとなった。PMはあらかじめ95種類の炭素源と酸化還元反応試薬としてtetrazoliumが充填された96穴のマイクロタイタープレートを提供している。プレートのウェルに菌を植菌し、もしその菌がウェル内の炭素源を代謝できることが出来れば、増殖に伴いtetrazoliumが発色し紫色を呈する(図1C)。こうして、注目している微生物の資化性を様々な炭素源に対して検証することを実現している。呈色の様子は専用のプレートリーダーがモニターしており、実験者は各ウェルの増殖曲線を数値情報として取得することが出来るため、得られた数値データを基に計算機で更なる解析が可能となっている。一方でマイクロタイタープレートではゲノムワイドな解析に展開するにはスループットが低いこともあり、固体培地に生じたコロニーの大きさを指標にした表現型解析の手法も報告されている(Collins, Schuldiner, Krogan, & Weissman, 2006; A. Liu et al., 2010; Tamae et al., 2008; Tong et al.,

2001)。しかしながら、これら手法はいずれも培養開始から一定時間後のコロニーの大きさを一度だけ計測するのみで、増殖曲線の情報が欠落した手法と成っている。我々研究室ではこの点に着目し、当研究室は多数の大腸菌の変異株を1枚の寒天培地に植菌し、経時的なコロニーの成長変化を自動でモニタリングするシステムを構築した(Takeuchi et al., 2014)。

Organism	Reference
Mutant libraries	
<i>Bacillus subtilis</i>	(Kobayashi et al., 2003)
<i>Candida albicans</i>	(Xu et al., 2007)
<i>Escherichia coli</i>	(Baba et al., 2006)
<i>Helicobacter pylori</i>	(Salama, Shepherd, & Falkow, 2004)
<i>Pseudomonas aeruginosa PA01</i>	(Jacobs et al., 2003)
<i>Pseudomonas aeruginosa PA14</i>	PMID:16477005
<i>Saccharomyces cerevisiae</i>	(Giaever et al., 2002; Winzeler et al., 1999)
<i>Staphylococcus aureus</i>	(Donald et al., 2009)
Plasmid close libraries	
<i>Escherichia coli</i>	(Kitagawa et al., 2005)
<i>Neisseria gonorrhoeae</i>	(Brettin et al., 2005)
<i>Pseudomonas aeruginosa</i>	(Labaer et al., 2004)
<i>Saccharomyces cerevisiae</i>	(Gelperin et al., 2005; Hudson et al., 1997)
<i>Staphylococcus aureus</i>	(Brandner et al., 2008)

表4 これまでに構築されたゲノムワイドなクローンセット。

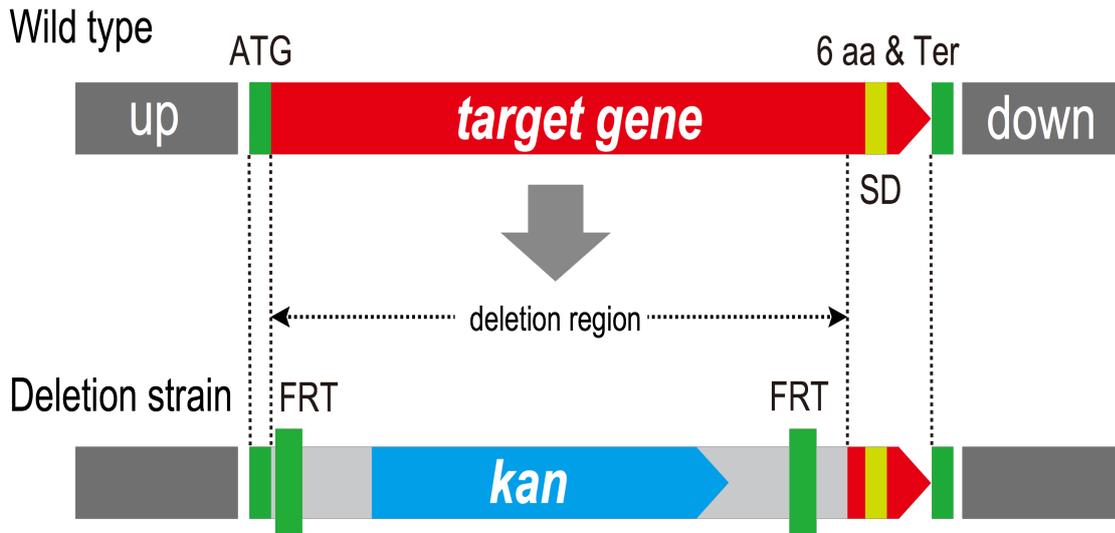


図16 Keio Collectionの基本構造。 *E.coli* K12 BW25113株を親株としてオープンリーディングフレーム領域は両端にFLP認識部位（FRT）をもつカナマイシン耐性カセットで置換されている。

Colony-liveの開発

Colony-liveと呼ばれている本システムは寒天培地1枚あたり1536株の密度でKeio collectionの株を植菌し、これら遺伝子欠失株のコロニーを最大48の環境条件で同時観測する(図17)。観測を行うために市販のスキャナ12台がインキュベータ内に設置されており、プログラムでコントロールされたスキャナが30分置きに寒天培地を撮影する。画像解析処理によって、撮影された画像から各コロニーの増殖曲線が算出される。この増殖曲線から3つのパラメーター（lag phase、 log phase、 stationary phase）が算出される。これらパラメーターを菌の増殖特性として扱うことでシステムティックな表現型解析を行うことが出来る。1536株を寒天培地に精密に植菌するために、植菌専用ロボット（シンガー社）を用いている。一回の実験で得られるデータ数は1536種類×48条件におよび、約21万のデータを一度に観測することが可能である。また、変異株の成長過程を数値化することもできる。我々研究室は、現在このシステムを運用することで大腸菌を用いた全遺伝子間の遺伝的相互作用（Genetic Interaction: GI）解析を進めている段階である。

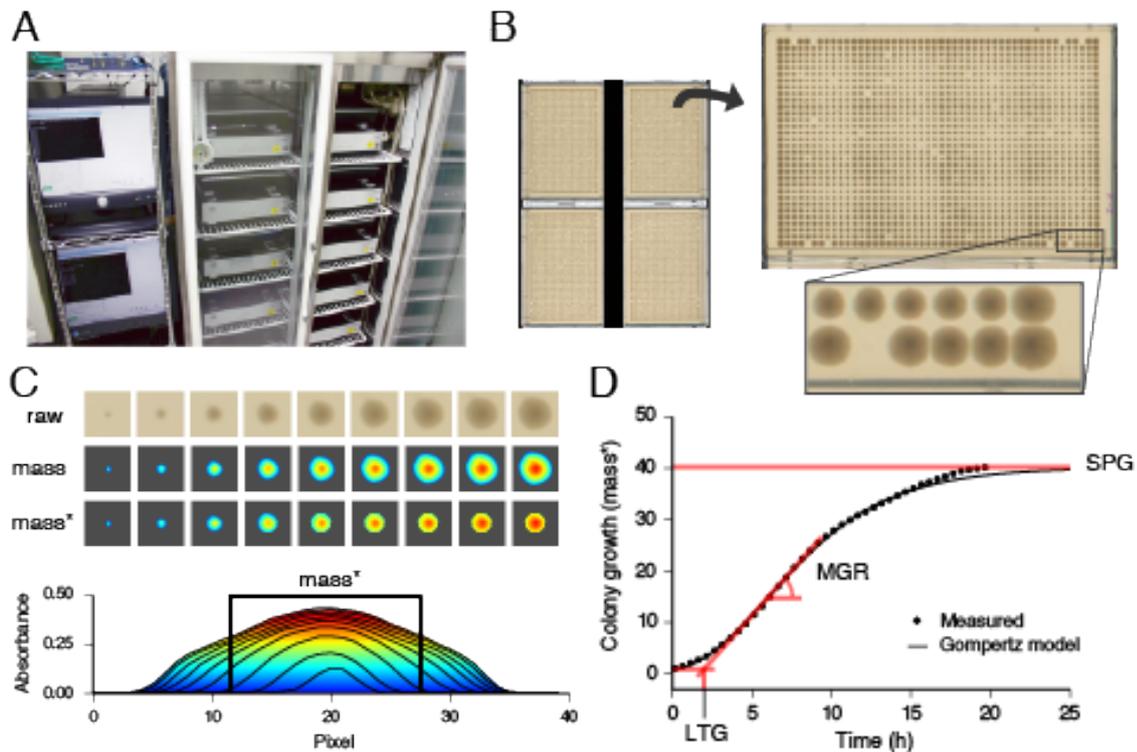


図17 Colony-liveの概要。A: Colony-liveを構成する機材。Colony-liveでは市販のスキヤナ12台が大型インキュベータの中にセットされており、これらスキヤナは30分おきに透過光撮影するように制御されている。この制御は2台のLinuxマシンとPythonで実装された制御プログラムによって行われている。B: Colony-liveによって撮影された寒天培地。Colony-liveを構成する各スキヤナには高密度（1536株/plate）に植菌された寒天培地4枚がセットされる。寒天培地はスキヤナによって30分おきに透過光撮影され、自動的に4分割される。分割された画像中の各コロニーを自動認識しコロニーの体積を算出する。C: Colony liveを用いたコロニーの細胞量の測定。透過光スキヤナを用いて、各コロニーの細胞量を数値化する。撮影時にコロニーに光を当て、どの程度コロニーを光が透過したか（absorbance）を調べることで細胞量を決定する。D:経時的なコロニーの増殖曲線。得られた増殖曲線をゴンペルツ曲線にfittingし、3つの増殖特性（Lag Time of Growth: LTG、Maximum Growth Rate: MGR、Satulation Point of Growth: SPG）を得ることができる。

Colony-liveの課題

Colony-liveに使用されているスキャナはRGB値の情報を持ったカラー画像で撮影する。しかしながら、現状のシステムではコロニーの生育度を算出する際に画像を二値化（白黒）処理を施している。したがって、各コロニーの色情報は解析対象とされていない。もし、この色情報を抽出し数値化することが出来れば、これまで3つに留まっていたコロニーの増殖特性値を拡張することが可能である。データの次元が増えることで解析性の向上が期待できる。

バイオフィルム合成ネットワークを網羅的に観測する意義

*csgA*遺伝子の発現は多数の遺伝子群の支配下にあることが知られており、これらの遺伝子が複雑に関与し合った結果、バイオフィルム形成に至る。したがって、バイオフィルム形成をより深く理解するためには、*csgA*遺伝子を含む一群の遺伝子の挙動を網羅的に捉えることが重要であり、それを念頭においた実験および技術開発が望ましい。こうした背景から、現状のColony-liveの設備を拡張する形で、ハイスループットにコロニーの呈色度を測定する系の確立はバイオフィルム合成ネットワークをターゲットとした研究において意義があるものと思われる。

本研究の目的

ハイスループットにコロニーの呈色度を測定する系を開発した上で、**Keio collection**を本研究に活用し、網羅性を担保しつつ大腸菌のバイオフィルム形成に関与する遺伝子破壊株のプロファイルを取得する。併せて、取得した呈色値データが統計的および生物学的に信頼性のあるものかを検証する。

II. 材料と方法

菌株

遺伝子欠失株は全てKeio collectionを用いた。また、野生株 (WT) としてKeio collection(Baba et al., 2006)の親株であるBW25113株を使用した。全てのKeio collectionの株は96ウェルディーププレートを用いてLB培地で一晚培養した後、グリセロールストック化し384 ウェルプレートの形で-80°Cで保管した。Keio collectionの全株はおよそ4000株あるので、384 ウェルプレートの形で作製したグリセロールストックを12枚用いて全ての株を網羅した。

培地

本章において研究に使用した培地組成を以下に示す。全ての培地は121°C、20分の高圧滅菌処理後、終濃度30 µg/L カナマイシンを添加した。LB寒天培地の場合は7.5g/Lの寒天を混ぜて使用した。

• LB培地

Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	10 g/L

• YESCA寒天培地(Hammar et al., 1996)

Cazamino acid	10 g/L
Yeast Extract	1 g/L
Agar	20 g/L

• CRA(Congo red agar) plate

滅菌後のYESCA培地に終濃度0.004 % コンゴレッドを添加	
HEPES pH7.4	25 mM

培養条件

1 プレートにつき1536 コロニーを集積させるため3日に分けて培養した。

1) 1日目：-80℃、384 ウェルプレートの形で保存していたKeio collectionの菌株のグリセロールストックを室温で溶かした後、RoToR (Singer Instrument Co.Ltd.、英国)を用いてLB寒天培地に384個/plateの密度でコロニーを植菌した。植菌後37℃にて16-18時間培養した。

2) 2日目：RoToRを用いて生えてきたコロニーをピックアップし、LB寒天培地にコロニーを1536個/プレートの密度で集積化した。植菌後、30℃で12時間培養した。

3) 3日目：RoToRを用いて生えてきたコロニーをピックアップし、YESCA寒天培地にコロニーを1536個/プレートの密度で継代培養した。植菌後、28℃条件下のスキヤナ上 (スキヤナの型番：GT-970、セイコーエプソン株式会社) に4枚のYESCA寒天培地を置き、48時間培養した。培養中、1時間ごとにスキヤンをすることでコロニーの生育をモニターした。

経時的なスキヤニング

経時的にスキヤニングするために自動でスキヤンをする環境を開発した。全ての開発環境はLinux Debian 7.8上で行った。Linuxに接続されたスキヤナを自作のプログラム上からコントロールするためにEPSONが提供しているLinux用のスキヤンソフト (iscan) を利用した。作製したプログラムは全てPython 2.7.11を用いて行った。作製したプログラムと簡単な説明を以下に示す。

- device.py LinuxマシンにUSB接続されたスキヤナを認識するプログラム。scanning.pyの内部で呼び出される。
- scanning.py 経時的にスキヤナを動作させるためのプログラム。内部でdevice.pyが認識したスキヤナを呼び出し、1時間ごとに寒天培地を撮影する。
- cropping.py 撮影されたスキヤナ画像から寒天培地が移っている領域のみを切り取り、実験のIDごとに保存するプログラム。
- pict2plate.py スキヤナ画像の加工に必要な処理をまとめたプログラム。cropping.pyの内部で呼び出されるようになっている。
- plate2colony.py 切り取られた寒天培地の画像から、1536個のコロニーの領域を認識して、呈色値を計算するための処理をまとめたプログラム。
- main.py 寒天培地上のコロニーの呈色値を計算し、結果をcsvファイルとして出力するプログラム。内部でplate2colony.pyを呼び出す。
- config.ini 切り出す寒天培地の座標情報やコロニーの座標を記した設定ファイル。

画像の明るさ調整

スキャンされた画像の明るさを調節するために、ターミナル上でImagemagick 6.8.9-7を用いた。行った命令を以下に示す。

```
$ convert -contrast-stretch 1.8% -rotate +90 input_img output_img | echo input_img
```

コロニーのエッジ検出と呈色値の算出

スキャナによって反射光撮影された画像からコロニーのエッジを検出するため、次の処理を行った(図18)。まず、測定開始(0h)から48hまでのコロニーが生じた部分を切り取った。次に、切り取った画像をRGBチャンネルに分解し、0hの時点での画像をバックグラウンドとして、ピクセルごとの輝度の絶対値を計算することで、背景差分を行った。0hの画像と比べ、よりピクセル輝度が増加している部分、つまりコロニーが生じている部分ほど赤く見え、培地の部分は青く見える。次に三つの差分画像を重ね合わせて合成画像を得た。得られた合成画像には、ごま塩ノイズが認められたので、画像平滑化処理(中央値フィルタリング)を行ってノイズを除去した。得られた画像を二値化処理し、白いピクセルの総数をコロニーの面積とした。また、コロニー面積には白とびの部分が存在したので、適切な閾値の設定によって白とびの領域を検出した。白とびの部分から4ピクセル左にずらした $5 \times 5 = 25$ ピクセル内の平均のRGB値を算出しHSV変換した。得られたHSV値を各測定時におけるコロニーの呈色データとした。

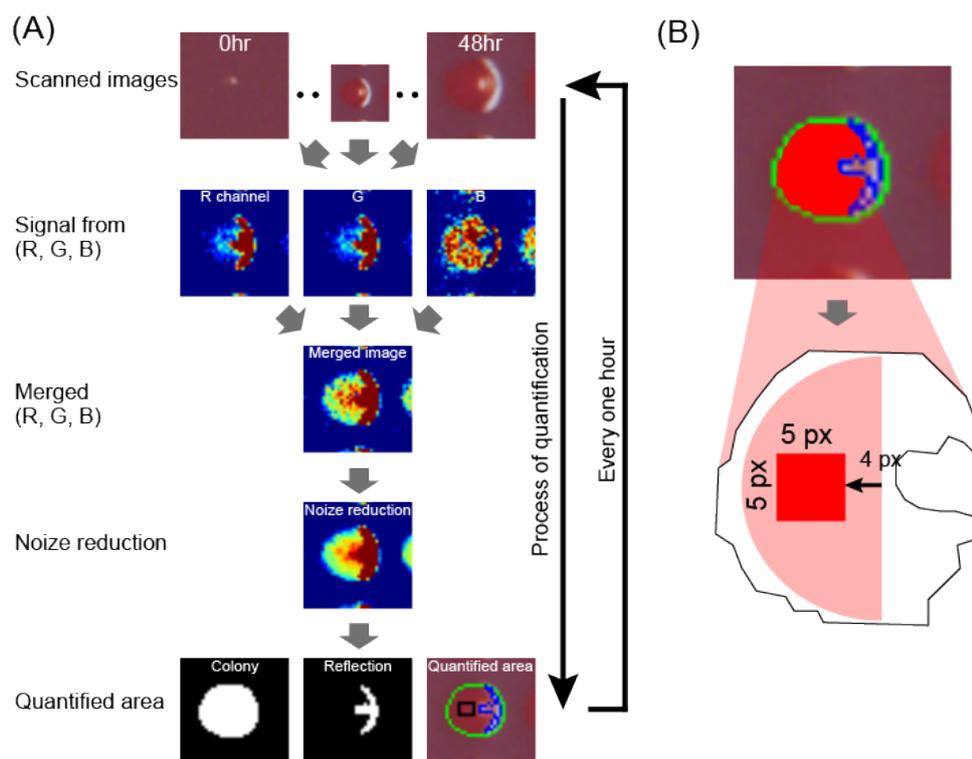


図18 自動でコロニーのエッジを検出するアルゴリズム。各測定時刻におけるコロニーの写真と植菌直後(0h)の差分を取ることでバックグラウンドを除去した。この処理を各色チャンネルごとで行い、得られた画像を合成した。合成画像にノイズが含まれていたため適切なフィルタで除去し、二値化処理した。得られた二値化画像の白いピクセルの部分のコロニーの面積とした。また、反射している部分は別に検出した。

寒天培地上におけるコロニー面積のブレと補正

寒天培地の端に生じたコロニーは内側のコロニーと比べ栄養の競合が少なく、生育が良くなる。これはedge effectと呼ばれる現象で、測定値のブレになってしまう。そこで、測定されたコロニーの面積をプレートの場合順に並べ、測定値のブレを補正した。

増殖解析

一時間おきに観測されたコロニー面積の経時データをゴンペルツ成長モデルに非線形回帰し、3つのパラメータ (k, b, c) にまとめた。これらパラメータを用いてゴンペルツ生育モデルは次の式(式1)で表される。

ここで、 k は最大到達点、 b は $x=0$ のときの y 、 c は傾きのパラメータである。

このとき、 $y=\alpha k$ ($0 < \alpha \leq 1$) となる x は式2となる。ゴンペルツ生育モデルに回帰させて得られた k をコロニーの増殖曲線の推定される飽和点 (100% growth) とした際の、90 % growthは式3で求めた。

$$y = kb e^{-cx} \quad \text{式1}$$

$$x = \frac{1}{c} \times \ln \left(\frac{\ln(b)}{\ln(\alpha)} \right) \quad \text{式2}$$

$$x = \frac{1}{c} \times \ln \left(\frac{\ln(b)}{\ln(0.9)} \right) \quad \text{式3}$$

III. 結果

各コロニーの増殖曲線のモニタリング

寒天培地上にKeio collectionの株を高密度（1536株/plate）に植菌後、48時間後まで1時間ごとにスキャナを用いて寒天培地を透過光撮影した。その後、画像解析によってコロニーのエッジを自動検出し、コロニー面積の経時データを得た（図19 A）。得られたコロニー面積の経時データにはノイズや、欠損値が含まれていた。得られた経時データをゴンペルツ生育モデル非線形回帰させた曲線をコロニーの生育曲線とした（図19 B）。非線形回帰させることで、植菌直後にコロニーが小さすぎて観測が出来なかった部分の推定や、欠損値の推定を行うことができた。また、得られた生育曲線のほとんどは飽和点まで達していなかったが、回帰によって飽和点（100% growth）を推定することができた。

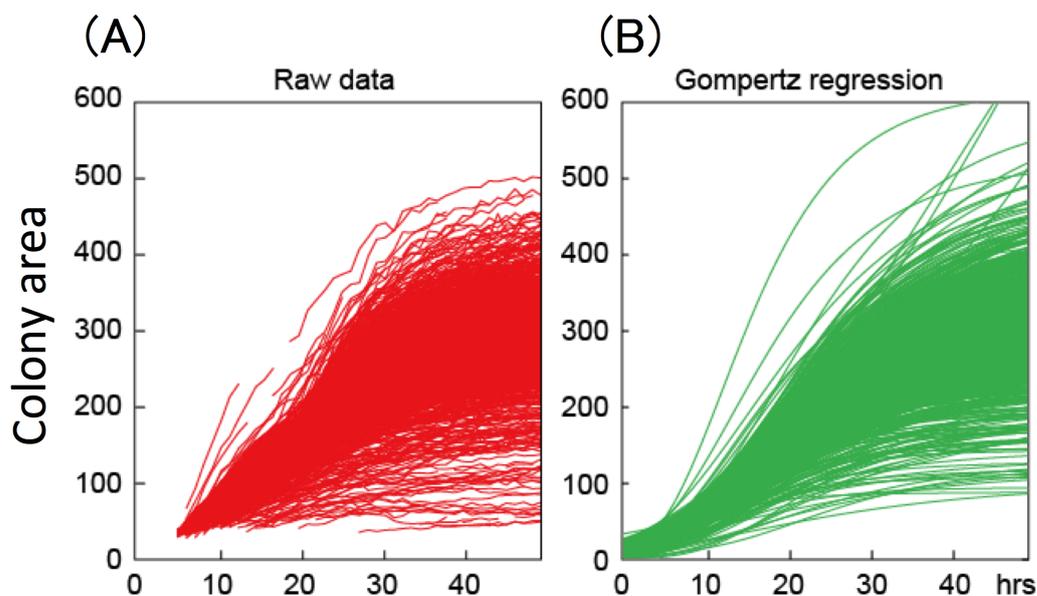
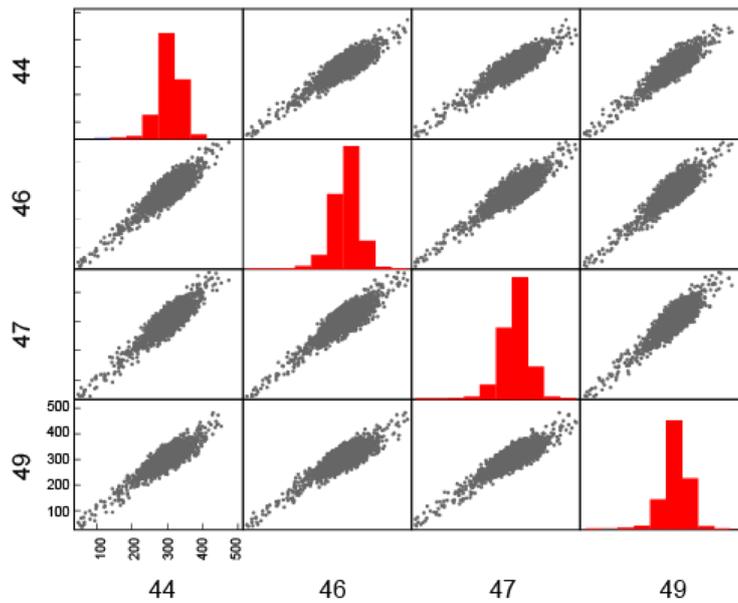


図19 各コロニー面積の経時データ。A: 1時間ごとに測定されたコロニー面積の経時データ。植菌直後や一部データに欠損値が含まれていた。B: ゴンペルツ生育モデルに非線形回帰させて得られた各コロニーの生育曲線。欠損値だった部分の補完や、測定値のノイズを除去することができている。

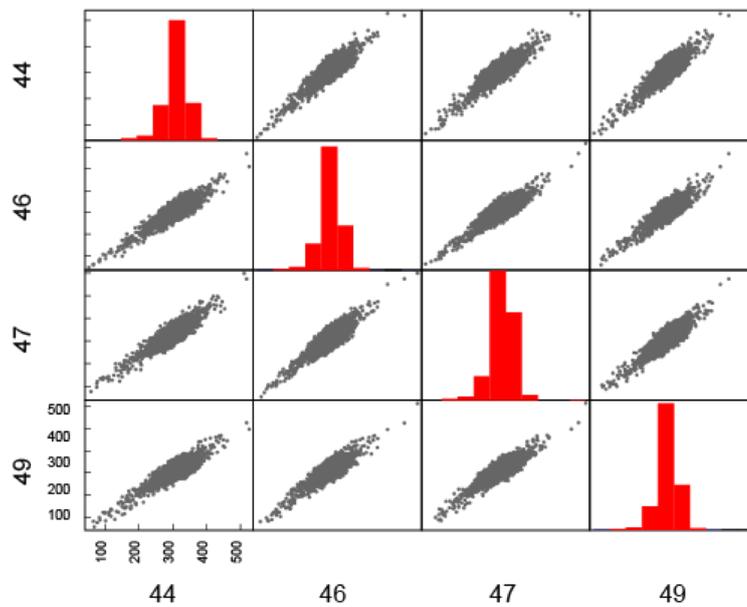
データの再現性

実験の精度を調べるため4の繰り返し実験における各欠失株の各測定値の再現性を散布図行列で確認した。実験間において測定値の傾向は一致していることが確認できた (図20)。

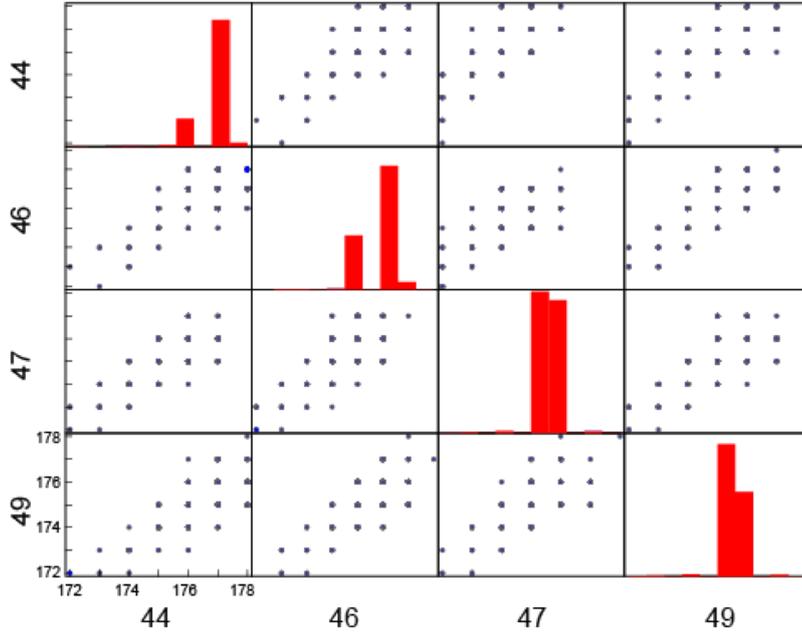
Area



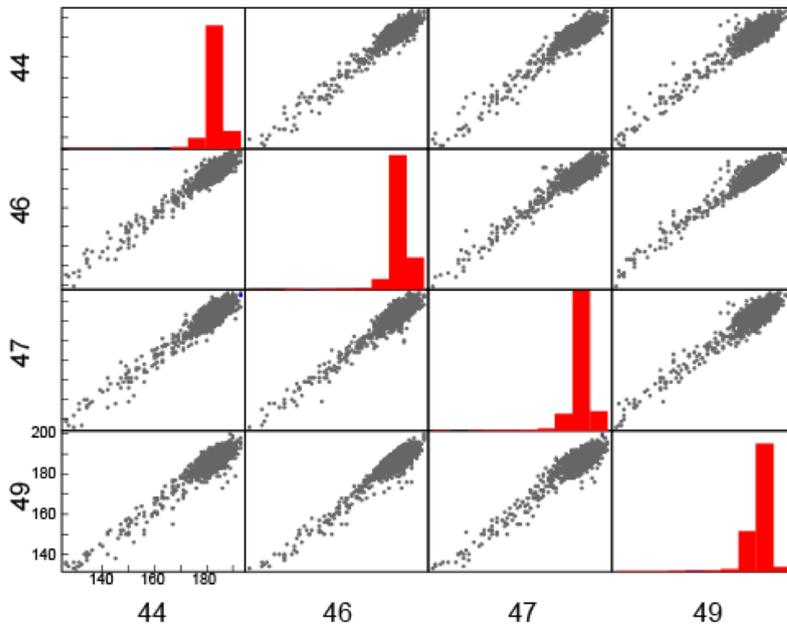
Gompertz



H



S



V

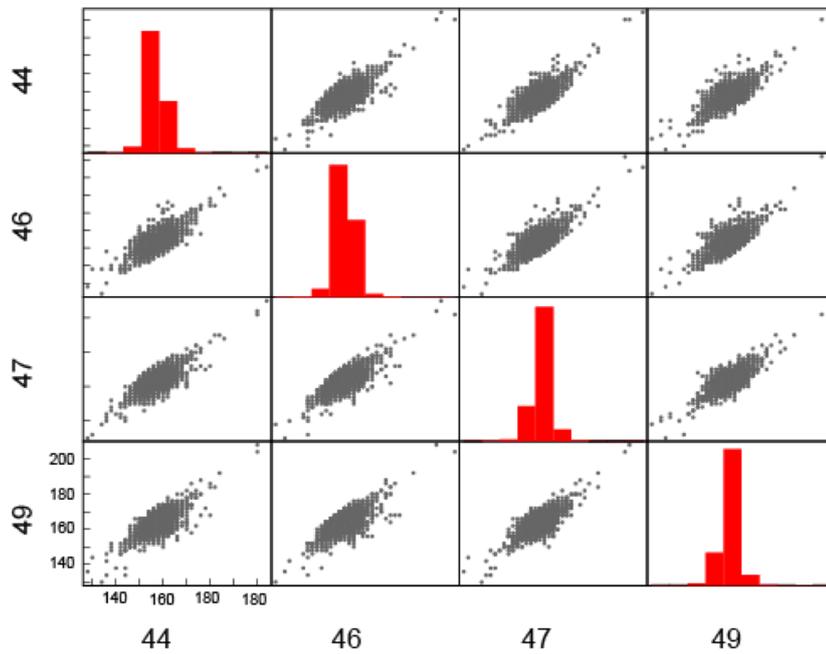


図20 繰り返し実験の再現性を表す散布図行列。培養開始から48時間後における各測定値を実験間で比較した (Area: 非線形回帰前のコロニー面積、Gompertz: 非線形回帰後のコロニー面積、H:色相H、S: 彩度S、V: 明度V)。44 - 49 は行った実験の通し番号。右下がり対角線状に並ぶのは得られた測定値の密度を表す。

各コロニーの生育ステージを考慮した呈色の比較

得られた増殖曲線のデータには、増殖が早いものや遅いものが含まれていた。バイオフィームは細菌が増殖ステージの後半（early stationary phase）にかけて形成されることが指摘されていたため、呈色の違いを比べるためには、コロニーの増殖ステージをそろえることが望ましいと思われた。そこで、非線形回帰によって得られた推定される増殖曲線の飽和点（100% growth）を基準に、各生育曲線の90% growthを算出し、その時点でのコロニーの呈色を測定することで、全てのコロニーの生育ステージを揃えた呈色の比較を行った。一部生育不良なコロニー（185株）を除き、3953株の生育曲線を得ることが出来た。この内、90% growth点が48時間以内だったのは3902株だった。最終的に、全体の93.4%(4140株中3902株)を呈色を測定するための解析対象株とした（図21 A）。解析対象となったほとんどの株の色相Hは348~354°の範囲内に収まっていたことから、測定された色が赤色であることを確認した（図21 B）。彩度Sと明度Vの値をプロットすると、ほぼ全ての株が平均0付近に集中しており、対象株の呈色が同じであることが分かった（図21 C）。一方で、一部株に関して彩度Sや明度Vの値が著しく変化した株も含まれており、遺伝子破壊によって著しくCurli線毛の合成量が増加した株の集団だと考えられた。

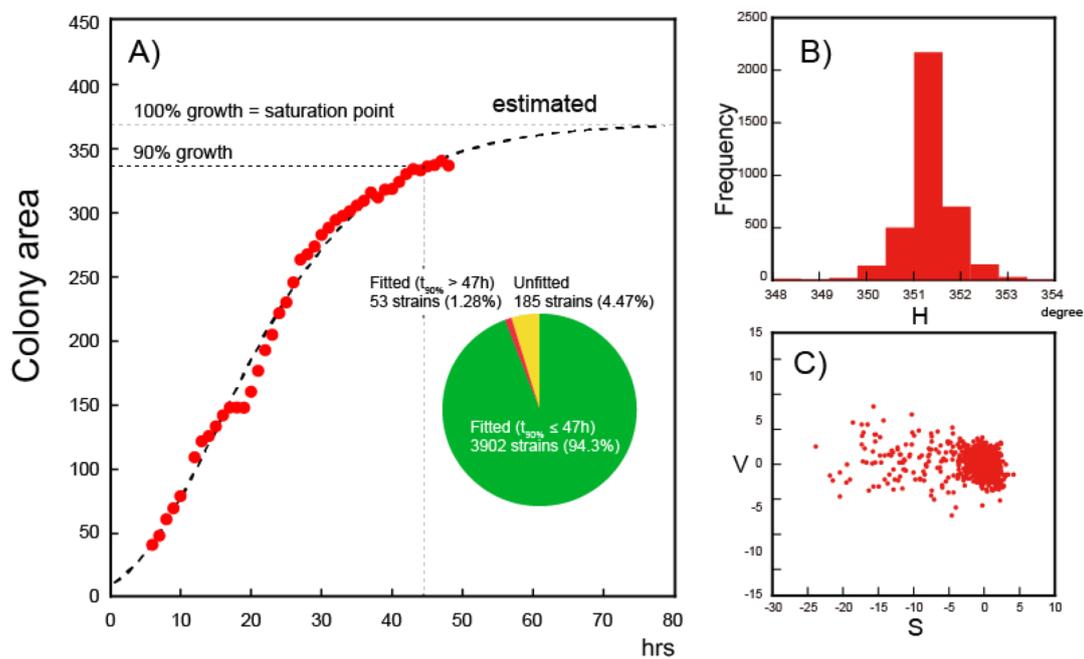


図21 コロニー面積から得られた生育曲線の詳細。A: 生育曲線から推定される飽和点（100% growth）と90% growth点、および90%growthに達した際の時間 $t_{90\%}$ 。

呈色スコアの算出

彩度Sと明度Vの関係について詳しく見てみると両値には緩やかな逆相関が見られ、プロットの密度分布からデータにゆがみがあるように見てとれた。(図22 A)。そこで、データを最もよく説明できる軸の再設定を行うため、主成分分析(PCA: Principal Component Analysis)を行った。主成分分析後、新たに設定した軸に射影した。また、今後の解析を簡便化するため、第1主成分に射影した値を呈色スコアとして、呈色値を1次元化した(図22 B)。得られた呈色スコアのランキングを各株の呈色の強さとした。

PCA analysis

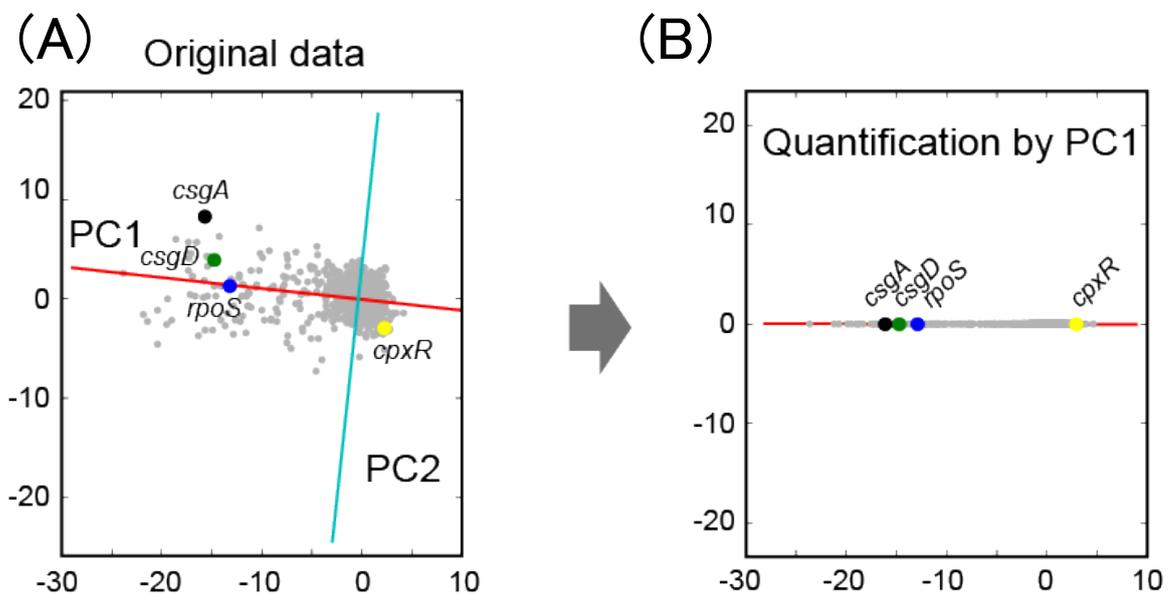


図22 主成分分析による呈色スコアの算出。A: 彩度Sおよび明度Vの散布図に主成分分析で計算された第1主成分と第2主成分を重ねた図。赤が第1主成分(PC1)。水色が第2主成分(PC2)。B: 呈色スコアの算出。全てのデータを第1主成分に射影することで、一次元のスコアを作成し、これを呈色スコアとし、スコアが高いほど呈色が亢進した株とした。

候補株の選定 ($-2\sigma \sim 2\sigma$ の範囲内)

呈色スコアとコロニー面積の関係を調べるため、まずコロニー面積の分布について調べた。コロニーの面積をZ変換によって標準化した上で、どのコロニーの大きさの範囲内を解析対象にするかを検討した。正常と見なすかを正規Q-Q (quantile-quantile)plotを用いて検討した (図23)。これはデータが正規分布に従っているかを簡易的に調べることが出来る手法で、理論的な正規分布と比べたいデータを比較する。今回は標準化したコロニーの面積が -2σ から 2σ の範囲であれば正規分布に従っているとして、コロニー面積の閾値とした (図24 A)。

標準化したコロニー面積と呈色スコアの間を見ると、面積が極端に小さいものは呈色スコアも下がる傾向にあった (図24 B)。今回は $-2\sigma \sim 2\sigma$ の範囲内のコロニー面積の株 (生育曲線が得られた3902株中の3605株) のみの呈色スコアを解析することにし、グループ2とした。また、面積が小さい (-2σ 以下) 株をグループ1、面積が大きい (2σ 以上) 株をグループ3とした。グループ2にカテゴリズされた3605株の呈色スコアの上位5%と下位5%を計算するとそれぞれ1.659および-1.273であった。そこで、呈色スコアが1.659以上だった170株と、-1.273以下の167株をそれぞれ呈色が著しく亢進した株、呈色が著しく下がった株としてリストアップした (図24 C)。

Probability Plot

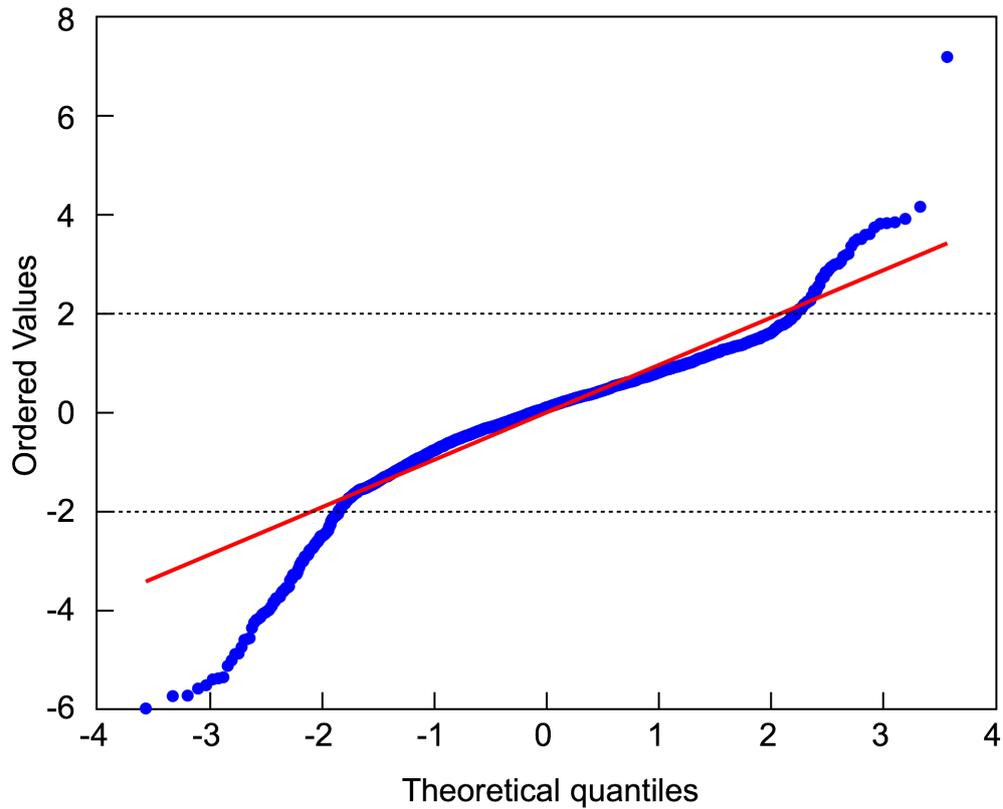


図23 正規Q-Q plotを用いた、コロニー面積の閾値の検討。X軸に正規分布のパーセンタイル、Y軸に標準化したコロニーの面積をプロットした。もしコロニーの面積が完全に正規分布に従った分布をとっていれば直線となる。今回の場合はコロニー面積が -2σ ~ 2σ の範囲内であれば正常とみなした。赤線は25パーセンタイルと75パーセンタイルを直線で結んでいる。

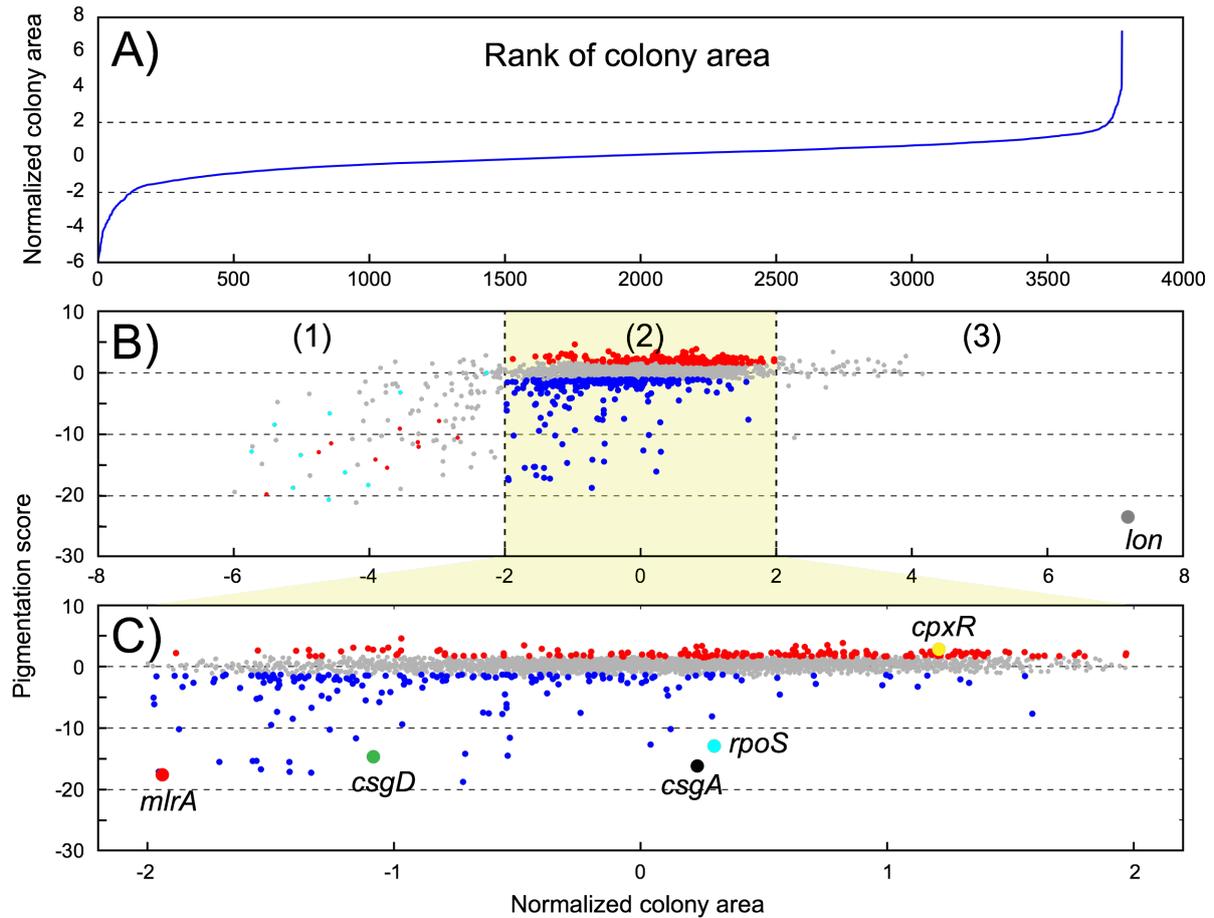


図24 コロニー面積と呈色スコアの関係性。A: 正規化後のコロニー面積のランキング。破線は -2σ および 2σ の部分を示す。B: 正規化後のコロニー面積と呈色スコアのプロット。正規化後のコロニー面積が -2σ 以下をグループ1(1)、 -2σ から 2σ の範囲内をグループ2(2)、 2σ 以上をグループ3(3)とした。呈色スコアは上位5% (赤) と下位5% (青) でそれぞれ閾値を設定した (上位5%: 1.659 、下位5%: -1.273)。C: コロニー面積が $-2\sigma \sim 2\sigma$ の範囲内に収まった株のみの呈色スコア。バイオフィーム形成に関与していると判明しているいくつかの株を選定してプロットしている。

結果のvalidation

まず、1章のスポット実験用いている呈色が既知の株（表1、図11）の呈色スコアについて確認した。その結果、予想通り *cpxR* は呈色が亢進した候補株の群、*csgA*、*csgD*、*rpoS* は呈色しない候補株の群にそれぞれリストされた（図24 C）。過去に報告された (Niba, Naka, Nagase, Mori, & Kitakawa, 2007) 呈色しない22株のデータと本実験の結果を比較した。報告されていた22株についてコロニー面積を確認すると、6株が -2σ 以下（グループ1）、13株が解析対象内（グループ2）、2株が 2σ 以上（グループ3）、1株が検出不可であった(表5)。解析対象内の13株について呈色スコアを確認したところ、12株が呈色しない候補株群にリストされた。

Gene	Judge	Group
<i>csgA</i>	Y	2
<i>csgB</i>	Y	2
<i>csgD</i>	Y	2
<i>csgE</i>	Y	2
<i>csgG</i>	Y	2
<i>cyaA</i>	-	-
<i>hsrA</i>	N	2
<i>ihfB</i>	Y	2
<i>lon</i>	-	3
<i>mlrA</i>	Y	2
<i>nlpD</i>	Y	2
<i>ompR</i>	-	1
<i>rpoS</i>	Y	2
<i>rfaG</i>	Y	2
<i>crp</i>	-	3
<i>gcvA</i>	Y	2
<i>hfq</i>	Y	2
<i>nifU</i>	-	1
<i>tolB</i>	-	1
<i>tolR</i>	-	1
<i>yciM</i>	-	1
<i>ydaM</i>	-	1

表5: 過去の報告された呈色が極端に低くなる株の一覧と今回の結果。(Niba et al., 2007)を参考にを参考に作成した。赤文字が今回呈色スコアの解析対象になった13株（グループ2の株）。グループ1（6株）およびグループ3（2株）は黒文字で示す。*cyaA*に関しては本実験では生育曲線が得られなかった。本実験で呈色が低い（呈色スコア下位5%以下と判定）とされた株をY、そうでないものをN、解析対象外を-で表す。解析対象内の13株中、12株が本実験でも呈色が下がった株として判定されている。

IV. 考察

本章では第1章で検証したコロニーの呈色度を測定する手法をハイスループットな系に展開することで、バイオフィーム形成に関与する遺伝子破壊株の網羅的な呈色データを得ることを目的とした。その過程で、寒天培地上に生じた大量のコロニーを自動で検出することで、コロニーの面積や表面の反射領域を避け、適切な呈色値を算出できる測定系を新たに開発した。経時的にコロニーの面積や呈色の亢進をモニターすることが出来る点は本測定系の特徴の一つである。経時データであれば欠損値を含んでいても比較的容易に推定することが可能である。また、適切な数理モデルを採用することで、多少ノイズな観測データであってもノイズを取り除くことができるほか、観測区間外であっても推定値として算出することが可能である。実際、1時間ごとにコロニー面積の経時データをゴンペルツ生育モデルに非線形回帰させることで、植菌直後の小さなコロニーの面積や、増殖曲線の理論上の飽和点を推定することができた（図19、21）。得られた増殖曲線のデータには、増殖が早いものや遅いものが含まれていた。バイオフィームは細菌が増殖ステージの後半（early stationary phase）にかけて形成されることが知られている。したがって、呈色の違いを比べるためには、増殖ステージがそろっていることが望ましいと思われた。これまでは培養開始から一定時間経過した後の呈色の度合いしか評価されてこなかったが、これは菌の増殖ステージを厳密には考慮しているとは言えなかった。本章ではこの点に着目し、観測結果から推定される増殖曲線を基に、菌の増殖を揃えた上での呈色の比較を行った。植菌した株の内94.3% (4140株中3902株)の生育曲線と呈色のデータを得ることに成功し、今回は90% growthの時点での呈色を比べることにした。解析対象となったコロニーの色相Hの分布は赤色の範囲内に収まっており、彩度Sと明度Vの間には緩やかな逆相関が認められた。

呈色度が2次元（彩度Sと明度V）に分かれたままでは全株の呈色度を比較する上で煩雑であったため、一次元のスコアにしてランキングすることにした。PCA解析の結果から第1主成分に射影した値を呈色度のスコアとして用いた。得られた呈色スコアとコロニー面積の関係を見ると、小さいコロニーほど呈色スコアが低かった（図24A、B）。これまで、CRAを用いたバイオフィーム形成能のスクリーニングでは呈色の有無のみを示すことがほとんどだったが、呈色の強度に菌の増殖が関係していることを本実験で確認することができた。本実験ではZ変換後、 $-2\sigma \sim 2\sigma$ の範囲内のコロニー面積の株（生育曲線が得られた3902株中の3605株）のみの呈色スコアを解析することにした（図24 B）。3602株の呈色スコアの上位5%と下位5%の値はそれぞれ1.659および-1.273であった。この値を閾値として、呈色が亢進した170株と呈色しない167株の計337株を著しく呈色が変化した候補株とした。得られた呈色スコアのデータが正しいかを確かめるため、まず1章でコントロールとして用いていた呈色が亢進する株

(*cpxR*) と呈色しない株 (*csgA*, *csgD*, *rpoS*) の値を確認したところ、予想通り *cpxR* は呈色が亢進した候補株の群、*csgA*, *csgD*, *rpoS* は呈色しない候補株の群にそれぞれリストされた (図24 C)。次に、過去に報告された呈色しない22株のデータと本実験の結果を比較した。報告されていた22株についてコロニー面積を確認すると、6株が -2σ 以下 (グループ1)、13株が解析対象内 (グループ2)、2株が 2σ 以上 (グループ3)、1株が検出不可であった (表 5)。解析対象内の13株について呈色スコアを確認したところ、12株が呈色しない候補株群にリストされた。CRAスクリーニングの場合、本来は株が赤くなるような実験系であるため、呈色しなくなる株の情報が目立ち、赤くなる株の情報は *cpxR* を除き手に入らなかった。しかしながら、呈色しない株の情報が過去の報告と概ね一致していることから、呈色が亢進する候補株の群も信頼性があるデータであると考えられる。これまでのCRAスクリーニングにおいては、観察が容易であるという点から、呈色しない株のみがスクリーニングしてきた傾向がある。本章ではこれまでの報告と一貫性のある呈色しない株の情報に加え、通常よりも呈色が亢進する株の情報に関する数値データとして得ることができた。これら呈色が亢進する株の情報は本研究における特色の一つであり、何らかの生物学的な共通性を含んだデータであると思われる。

本章ではコロニー面積が極端に変化した遺伝子欠失株については解析対象外とした。その内の一つ、コロニー面積が小さい欠失株群 (グループ1) の遺伝子機能をみたところ、エネルギーや核酸合成に関与する遺伝子破壊株が目立った。特に、呼吸鎖に関係する *nuo* (NADH:ubiquinone oxidoreductase) 遺伝子の欠失株はコロニー面積が極端に小さかった。 *nuo* 遺伝子群は13の遺伝子からなるオペロンを形成しているが、このうち10遺伝子がグループ1に含まれていた。また、プリン合成経路に関与する *pur* 遺伝子群の欠損株も増殖が著しく悪く、呈色スコアも低かった (図25 A)。こういった生育が極端に悪くなる遺伝子欠失株の呈色スコアは今回測定することが出来なかった。この結果は、菌の生育状況が呈色に影響していることを意味している。これまでのCRAスクリーニングでは呈色の強さのみが報告されていた。しかしながら今回の結果は、従来のスクリーニングの結果の一部には生育不良によって呈色なしと判定された株も含まれていることを示唆している。

一方、通常よりコロニー面積が大きい株の一部には *lon* などムコイドになる株がいた (図25 B)。Lon protease は通常、EPS (ExtraPolySaccaride) の一つである colanic acid の生合成に関与している RcsA を細胞内で速やかに分解している。lon の欠失によって細胞内の RcsA が過剰になり、過度な colanic acid の生合成につながったため、ムコイドの表現型になったと思われる。 *rfaD* は LPS (lipopolysaccharide) 生合成に関与しており、欠失株がムコイドになることが知られている (Karow, Raina, Georgopoulos, & Fayet, 1991)。したがって、コロニー面積が大きくなっていく株の一部はムコイド状の表現型によって、見かけ上面積が大きくなって測定されてしまったと考えられる。EPS がコロニー表面に過剰に分泌してしまうと、本来の呈色スコアを測定することは出来な

いと考えられたため、したがって今回の解析からは除外した。他方、ムコイド状を示したコロニーの表面は通常よりも反射領域が大きかった。本測定系では反射領域も検出していることから、今後この反射領域のデータに注目してムコイドになるコロニーを検出できる可能性を見いだすことができた。

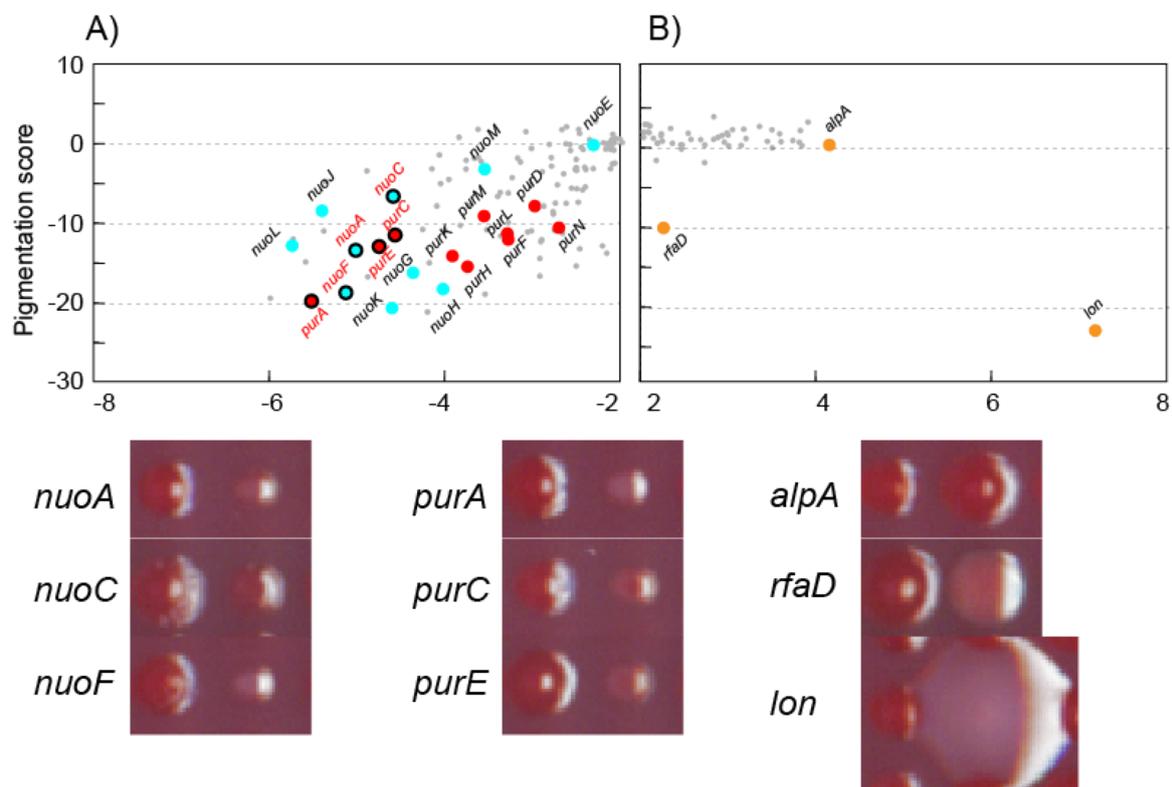


図25 グループ1およびグループ3におけるコロニー面積が極端に異なる株。左が通常のコロニーの大きさと右が各遺伝子欠失株。

第三章 呈色が著しく変化する遺伝子欠失株に関する情報の整理および新規バイオフィーム形成関連遺伝子探索に向けた検討

I. 序論

バイオフィーム合成に関連する遺伝子としては、Curli線毛の構成タンパク質をコードする*csg*遺伝子群の他に、様々な調節遺伝子が明らかになっている。その中心となっているのがCsgDである。CsgDはCurli線毛生合成のマスターレギュレーターであり*csgBAC*オペロンの転写に必要である。これまで多くの研究によってCsgDを中心としたCurli線毛合成の調節メカニズムが明らかにされてきた。また、Curli線毛の合成には様々な環境ストレスによってコントロールされている。二章で網羅的検出により得られたバイオフィーム合成関連遺伝子群には、*csg*遺伝子群及びそれらを直接的または間接的に調節する遺伝子群が含まれているものと期待される。これまでに明らかになっているバイオフィーム合成関連機構について記す。また、*csg*遺伝子を構成する2つのオペロンに対して制御関係が明らかになっている因子に関しては表6にまとめた。

低温ストレス、stationary phase突入によるバイオフィームの形成誘導

大腸菌やサルモネラでは低温ストレス時（25℃） σ 因子であるRpoSとsmall RNAであるDsrAによる遺伝子の発現調節が存在することが知られている。低温時のlog phaseにおいて、RpoSの細胞内濃度はDsrAを介して上昇することがLB培地と最小培地(M63)の両培地で確認されている(Sledjeski, Gupta, & Gottesman, 1996)。*dsrA*の転写は低温によって上昇し、DsrAは*rpoS*のmRNAに結合することでその二次構造を変える。これにより*rpoS* mRNAの翻訳効率が上昇し、log phaseにおける細胞内のRpoSの濃度が上昇する(Lease & Belfort, 2000; Olsén, Arnqvist, Hammar, Sukupolvi, & Normark, 1993)。これに伴い、*csgB*、*csgA*、*csgD*、および*dsrB*の発現が上昇することが確認されている(Sledjeski et al., 1996)。また、低温時に発現されるCrlは30℃以下の培養条件で菌がstationary phaseに突入した際に蓄積し、RpoSに結合することでRpoSの活性を向上させる(Bougdour et al., 2004)。同様に、stationary phaseに突入における細胞内のRpoS濃度は*csgD*の発現を促す。

浸透圧によるバイオフィルムの形成誘導

浸透圧は様々な微生物のバイオフィルム形成に関与している外的要因である。多くの場合、高浸透圧下ではバイオフィルムの形成が抑制されるが、その程度は種によってばらつく。例えば*P. fluorescens*は培地にNaCl またはsucroseを添加することでバイオフィルムの形成が阻害される。*S. typhimurium*の場合、培地に高濃度のNaClを添加すると*csgD*の転写が抑制される(U Römling, Sierralta, Eriksson, & Normark, 1998)。大腸菌でも同じような現象が観察されており、100 mM NaClによってCurli関連遺伝子の転写が抑制される。この制御にはOmpR、CpxRが関与していること、また高濃度のsucroseではCpxRによるCurli関連遺伝子の転写抑制は見られず、代わりにH-NSが働いていたことが報告されている(Jubelin et al., 2005)。その一方で、200 mM NaClはPoly-N-acetyl glucosamine(PGA)の合成に関与している*pga*オペロンの転写を促進する(Goller, Wang, Itoh, & Romeo, 2006)。

カタボライトリプレッションによるバイオフィルムの形成阻害

環境中のglucoseがない場合、細胞内にセカンドメッセンジャーであるcyclic adenosine 3',5'-monophosphate (cAMP)の濃度が上昇する。また、これに伴いcAMPはcAMP受容タンパク質 (CRP) と結合し複合体を形成することで、グルコース以外の糖を代謝する。この抑制をカタボライトリプレッションといい、バイオフィルム形成においてもこの抑制が関与している。cAMP-CRPはsmall RNAの*mcaS*の発現を促す。McaS (Multi-cellular adhesive sRNA) は*csgD*のmRNAの5'-UTRに直接結合することで*csgD* mRNAが分解されるのを促進する(Thomason, Fontaine, De Lay, & Storz, 2012)。これら遺伝子はそれぞれPGA合成遺伝子および細胞のmotilityに関与している遺伝子である。McaSはその他多くのsmall RNAと同様にRNAシャペロンであるHfqやCsrAを介してmRNAと結合することで翻訳抑制を行う。Hfqが存在しない場合、細胞内におけるMcaSの半減期は著しく下がるが、CsgAはMcaSの安定性に寄与していない(Jørgensen, Thomason, Havelund, Valentin-Hansen, & Storz, 2013)。

Cyclic di-GMPによるバイオフィルムの形成誘導

CsgDによるバイオフィルム関連遺伝子の発現上昇の一部はbis-(3',5')-cyclic diguanylic acid (cyclic di-GMP) を介して達成されている。cyclic di-GMPはcyclic di-GMP合成酵素 (diguanylate cyclase) によって作られており細胞内のセカンドメッセンジャーとして、バイオフィルム形成を促進する。cyclic di-GMP合成酵素の一つである*adrA*がCsgDによって正に発現調節されているのは、その例の一つである。逆に、cyclic di-GMPがCsgD自身の発現も促進している。*rpoS*の活性化によるYdaMの発現は、細胞内のcyclic di-GMP濃度を上昇させ、*csgD*の発現が活性化される。cyclic di-GMP合成酵素は活性部位のGGDEFという特徴的なドメインをもつ。反対にcyclic di-GMP分解酵素 (cyclic di-GMP phosphodiesterase) は活性部位としてEALというドメインを共通にもつ。ドメインベースの配列解析から、大腸菌ではGGDEFドメインを持つタンパク質が12、EALドメインを持つタンパク質が10、そして興味深いことに両ドメインをいずれももつタンパク質が7つあることが確認されている(Hengge et al., 2015)。このようにバイオフィルムの形成には複数の遺伝子や外部からの刺激が関わっている。細胞内のcyclic di-GMPレベルの上昇は*csgD*の発現上昇と共に*ycgR*の発現を促す。YcgRは細胞が動き回るために必要なflagellarを構成するFliGおよびFliMに結合することで、その働きを阻害する(Fang & Gomelsky, 2010; Paul, Nieto, Carlquist, Blair, & Harshey, 2010)。これにより細胞がplanktonicな状態になるのを抑制する。バクテリアの約85%はcyclic di-GMPを持っていることから、バイオフィルム内では種を超えたコミュニティを形成してmotilityとバイオフィルム形成の切り替えを行っていると予想される(Galperin, 2004; Hengge, 2009; Ute Römling & Amikam, 2006)。

その他のバイオフィルム形成する要因

バイオフィルム形成のきっかけとなっているその他の要因として栄養飢餓、酸素分圧、pH、Integration Host Factor(IHF)、H-NS、などが報告されている。これらの刺激のほとんどは転写因子である*csgD*の発現上昇につながる。CsgDはCurli線毛やcelluloseといったバイオフィルムの構成成分に関与する遺伝子の発現調節に関わっており、バイオフィルム形成ネットワークのハブと捉えることができる。

Regulator	P _{csgBAC}	P _{csgDEFG}	Evidence	Reference(s)
<i>Transcriptional regulators</i>				
<i>RpoS</i>	+	+	Expression analysis, consensus binding site	(Arnqvist, Olsén, & Normark, 1994; Bougdour et al., 2004; Olsén et al., 1993; Pratt & Silhavy, 1998)
<i>Crl</i>	+	+	Interacts with RpoS and stabilizes binding to <i>csg</i> promoters	(Arnqvist, Olsén, Pfeifer, Russell, & Normark, 1992; Bougdour et al., 2004)
<i>CsgD</i>	+	+	Expression analysis, consensus binding site	(Hammar et al., 1995; Ogasawara, Yamamoto, & Ishihama, 2011)
<i>MlrA</i>		+	Expression analysis, consensus binding site	(Brown et al., 2001; Ogasawara, Yamamoto, & Ishihama, 2010)
<i>Cra (FruR)</i>		+	Expression analysis, consensus binding site	(Reshamwala & Noronha, 2011)
<i>Crp</i>		+	Expression analysis	(Zheng, Constantinidou, Hobman, & Minchin, 2004)
<i>RcdA</i>		+	Expression analysis, consensus binding site	(Shimada et al., 2012)
<i>DNA modifying enzymes</i>				
IHF		+	Expression analysis, consensus binding site	(Gerstel, Park, & Römling, 2003; Ogasawara, Yamada, Kori, Yamamoto, & Ishihama, 2010)
H-NS		-/+ (<i>E. coli/S. typhimurium</i>)	Expression analysis, consensus binding site	(Arnqvist et al., 1994; Gerstel et al., 2003; Ogasawara, Yamada, et al., 2010; Olsén et al., 1993)
<i>Two-component systems</i>				

<i>CpxA/R</i>	–	–	Expression analysis, consensus binding site	(Dorel, Vidal, Prigent-Combaret, Vallet, & Lejeune, 1999; Jubelin et al., 2005; Ogasawara, Yamada, et al., 2010)
<i>EnvZ/OmpR</i>		+	Expression analysis, consensus binding site	(Gerstel, Kolb, & Römling, 2006; Jubelin et al., 2005; Ogasawara, Yamada, et al., 2010; Ute Römling, Bian, Hammar, Sierralta, & Normark, 1998; Vidal et al., 1998)
<i>RcsA/B</i>	–	–	Expression analysis	(Ferrieres & Clarke, 2003; Jubelin et al., 2005; Vianney et al., 2005)
<i>RstB/A</i>		–/+ (basic/acidic conditions)	Expression analysis, consensus binding site	(Ogasawara, Yamada, et al., 2010; Ogasawara et al., 2007)
<i>ArcA/B</i>		+	Expression analysis, consensus binding site	(X. Liu & De Wulf, 2004)
<i>BasS/R</i>		+	Expression analysis, consensus binding site	(Ogasawara, Shinohara, Yamamoto, & Ishihama, 2012)
<i>Small regulatory RNAs</i>				
<i>OmrA/B</i>		–	Antisense binding to 5'UTR	(Holmqvist et al., 2010)
<i>McaS</i>		–	Antisense binding to 5'UTR	(Jørgensen et al., 2013; Thomason et al., 2012)
<i>GcvB</i>		–	Antisense binding to 5'UTR	(Jørgensen et al., 2013)
<i>RprA</i>		–	Antisense binding to 5'UTR	(Jørgensen et al., 2013; Mika et al., 2012)
<i>ArcZ</i>		+	Expression analysis post-transcriptional regulation	(Monteiro et al., 2012)
<i>SdsR</i>		+	Expression analysis, transcriptional regulation	(Monteiro et al., 2012)

表6 *PcsgBAC*および*PcsgDEFG*に対する制御が明らかになっている因子。各プロモーターに対して正の制御であれば+、負の制御であれば-とした ((Evans & Chapman, 2014)より一部改変)。

DAVIDによる候補株のアノテーション

機能未知の遺伝子配列やタンパク質に生物学的概念を持つ用語（term）を関連づけることで、具体性を付加する作業のことをアノテーションという。アノテーション作は遺伝子機能がまとめられたデータベースを利用して進めていくのが一般的である。今回はその中でもDAVID(The Database for Annotation, Visualization and Integrated Discovery <https://david.ncifcrf.gov/>)に基づくアノテーションを行った。DAVIDとはDAVIDはNational Institute of Allergy and Infectious Diseaseが提供するアノテーション用のデータベースで、2017年7月現在でversion 6.8としている。マイクロアレイやゲノムワイドなスクリーニング実験から得られた候補遺伝子に対し、複数の異なるデータベースによって管理されているtermを引用して関連づけることで具体性のあるアノテーションを行うことが出来る点が特徴的なwebベースのツールである。また、アノテーション結果を簡便に可視化できる機能も提供している。

本研究の目的

上述の通り、バイオフィーム形成に関わる制御ネットワークは複数の遺伝子と環境要因が関与しており複雑である。このうち、どの遺伝子が今回開発した測定系で検出できているのかを明らかにすることで、データ全体の生物学的な妥当性を検討することができると考えられる。また、ゲノムワイドなデータにはしばしば既知情報にはない機能未知の遺伝子が含まれる。こうした遺伝子を既知情報と照らし合わせながら整理することは機能推定を行う最初のきっかけにつながると思われる。以上を背景とし本研究では、前章までに得られた呈色データをクラスタリングツールを用いて分類し、整理することにした。併せて、今後個別の研究につながりうる新規バイオフィーム形成遺伝子の探索を行った。

また、2章で得られた候補株のうち、著しく呈色が変化した一部の株に関しては実際にクリスタルバイオレット法に基づくバイオフィーム量の測定を行い、コンゴレッド法のスクリーニングと比較した。

II. 材料と方法

菌株

バイオフィルム量の測定に用いた遺伝子欠失株は全てKeio collectionを用いた。また、野生株 (WT) としてKeio collection(Baba et al., 2006)の親株であるBW25113株を使用した。2章で算出した呈色スコアが高い順に2以上の87株、呈色スコアが低い順に87株をそれぞれ選定し、96ウェルディーププレートを用いてLB培地で一晚培養した後、グリセロールストック化した。その後-80°Cディープフリーザーで保管した。

培地

本章において研究に使用した培地組成を以下に示す。全ての培地は121°C、20分の高圧滅菌処理後、終濃度30 µg/L カナマイシンを添加した。

・LB培地

Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	10 g/L

・YESCA培地(Hammar et al., 1996)

Cazamino acid	10 g/L
Yeast Extract	1 g/L

クリスタルバイオレット法によるバイオフィルム量の測定

96ウェルプレートを用いてYESCA培地で28°C、48時間好氣的に静置培養した。培養後、OD₆₀₀で各ウェルを測定した。その後、培養液を捨て、水で3回リンスし、ウェル内が完全に乾くまで風乾した。その後、0.1%クリスタルバイオレット溶液200 µLを入れて、20分静置した。その後、溶液を捨て、水で3回リンスし、溶出液 (エタノール/アセトン混合液) 200 µLでウェルの壁に付着したクリスタルバイオレットを溶出した。20分静置の後、OD₅₉₀で各ウェルを測定した。Biofilm Index(OD₅₉₀/OD₆₀₀)を計算し、WTを1としたときの各株のBiofilm Indexをバイオフィルム量とした。

呈色が顕著に変化した株のリスト

前章までで得られたゲノムワイドな呈色データから、呈色が亢進した上位5% (170 遺伝子) と下位5% (167遺伝子) を選定した。

DAVID使用したtermとcategoryの紹介

DAVIDでは様々なアノテーションデータベースによって各遺伝子に「term」と呼ばれる、生物学的概念を表す用語を紐付けして遺伝子機能を表現している。代表的な用語集にGO(Gene Ontology <http://www.geneontology.org/>)やCOG(Clusters of Orthologous Groups <https://www.ncbi.nlm.nih.gov/COG/>)、KEGG(Kyoto Encyclopedia of Genes and Genomes <http://www.genome.jp/kegg/>)などが挙げられる。注目している遺伝子群をこのようなtermによってアノテーションしていき、どのtermが多いのか解析する手法を enrichment analysisという。DAVIDでは複数のデータベース (DAVIDではCategoryと表現される) で使用されている「term」を一度に扱うことで、注目している遺伝子のアノテーションすることができる。今回、DAVIDで使用したCategoryを以下に示す。

Funcional_Categories	COG_ONTROGY, UP_KEYWORDS, UP_SEQ_FEATURES
Gene_ontrogy	GOTERM_BP_DIRECT, GOTERM_CC_DIRECT, GOTERM_MF_DIRECT
Pathways	KEGG_PATHWAY
Protein_Domains	INTERPRO, PIR_SUPERFAMILY, SMART

DAVIDのクラスタリングの方法 (kappa係数に基づく類似度計算)

DAVIDでは2遺伝子あるいは2アノテーション間の距離をkappa係数に基づいて定義している。kappa係数はある基準で評価された2つの評価結果がどの程度一致しているかを表す指標の一つで、コーエンのkappa係数ともよばれる。具体的な例と計算方法を次に示す。

あるスクリーニング結果で得られた2つの遺伝子にtermをアノテーションしていき、その結果を1（既知）、0（未知）で表した際、次の結果になったとする。

	Cell death	Apoptosis	Ph domain	Sh2 domain	Apoptosis pathway	Membrane
geneA	1	1	0	0	1	0
geneB	1	1	0	1	1	0

geneAとgeneBは次の表のようにまとめることができる。

		geneA		
		1	0	row total
geneB	1	3	1	4
	0	0	2	2
	col total	3	3	6

このとき、2つの遺伝子間のアノテーション結果が一致した数を総数で割った値は $(3+2)/6 = 5/6 = 0.83$

となり、これを「見かけ上の一致率 (observed degree of agreement)」という。この一致率は、「偶然による一致率」を考慮していない。つまり、「あるtermにおけるgeneAとgeneBのアノテーションが偶然1（既知）となる確率」と「あるtermにおけるgeneAとgeneBのアノテーションが偶然0（未知）となる確率」を差し引かなければならない。kappa係数はこの点を考慮した一致率の指標である。

偶然アノテーションが一致した確率は次の式で表される。

$$(3/6 \times 4/6) + (3/6 \times 2/6) = 18/36 = 0.5$$

kappa係数は次の計算式で定義される。

$$K = \text{見かけ上の一致率} - \text{偶然の一致率} / 1 - \text{偶然の一致率}$$

したがってこの例の場合、

$$K = (0.83 - 0.5) / 1 - 0.5 = 0.66$$

となる。

実際のDAVIDではユーザーが任意に指定したkappa係数を閾値にして、複数の遺伝子間のアノテーションの類似度計算を行うことで、遺伝子群の機能別クラスタリングが可能である。

III. 結果

DAVIDによる機能分類

前章で得られた呈色が著しく変化した株についてDAVIDによる機能分類を行った。kappa係数を0.3に設定して機能のクラスタリングを行ったところ、呈色が著しく亢進 (Red)した170株中65株、呈色が著しく低く (White)なった167株中46株が、それぞれ4つと5つのクラスターに分類された(呈色が亢進していたクラスター:図26-29、呈色が低くなっていたクラスター:図30-34)。これら得られたクラスターに含まれる遺伝子について、これまでの知見も含め生物学的な解釈を検討した。Red Cluster1はDNA-bindingやTranscriptionといったtermをもつ遺伝子が集まった。その一部には*csgD*を負に制御する*cpvR*のほか、motilityに必要なflagellaの生合成に関与する*flhC*が含まれていた。その他、レプレッサーとして*nanR*、*nac*、*exuR*がクラスターの一部として入っている。Red Cluster2には*yhbJ*が含まれている。*yhbJ*の欠失株がコンゴレッド染色によって強く呈色することはすでに報告されている。*yhbJ*はsmall regulatory RNAである*glmZ*を不安定化させることが知られている。Red Cluster3にはATP-bindingやkinaseといったtermが目立った。Red Cluster4は多数の遺伝子からなるクラスターだが、とくにmembrane/transmembraneといったtermが目立った。

White Cluster1にはkinaseとしてアノテーションされている遺伝子が複数含まれていた (*fruK*、*thrB*、*rbsK*、*aroK*、*nagK*、*nanK*)。特に、*nagK*、*nanK*はいずれもUDP-GlcNac(Uridine diphosphate N-acetylglucosamine)合成に必要な経路上にマップされる遺伝子である。White Cluster2には*csg*遺伝子群の*csgB*および*csgE*とともに、biofilm形成に関与していることが指摘されている*bhsA*がクラスターを形成していた(Zhang, García-Contreras, & Wood, 2007)。これら比較的機能が良く知られている遺伝子とともに機能未知の*ynfB*が同クラスターを形成していた。White Cluster3にはadhesion domainというtermが目立った。*csg*遺伝子群の*csgA*および*csgB*が含まれ、これらとともに*yebU*と*ybgD*が同クラスターを形成していた。いずれもputative fimbriae proteinとされている。White Cluster4にはTranscriptionやDNA-bindingといったtermが多いことから、バイオフィーム形成に関与する転写因子のクラスターと解釈した。実際、バイオフィーム形成に関わる既知の転写因子が複数含まれていたことから (*rpoS*、*ihfAB*、*csgD*、*mlrA*、*gcvA*、*hipB*)、このクラスターに含まれるほかの転写因子もバイオフィーム形成に関与していることが考えられた。*hipB*は*hipA*とともにtoxin-antitoxin systemを構成している。*hipAB*欠失によって実際にバイオフィームの量が低下するという報告がある(Korch & Hill, 2006)。その他、バイオフィーム形成に関与しているかは問わず、*yieP*を除く他すべてが機能が既知の遺伝子であった。*yieP*はputativeなtranscriptional regulatorと推定されており、現時点で機能未知のいわゆるy-geneであるが、*yieP*以外の他の遺伝子の多くがバイオフィーム形成に関与しているという事実から、バイオフィーム形成に関与する何らかの転写因子ではないかと予想された。White Cluster5にはmembraneというtermをもつ遺伝子が集まった。

Red Cluster 1

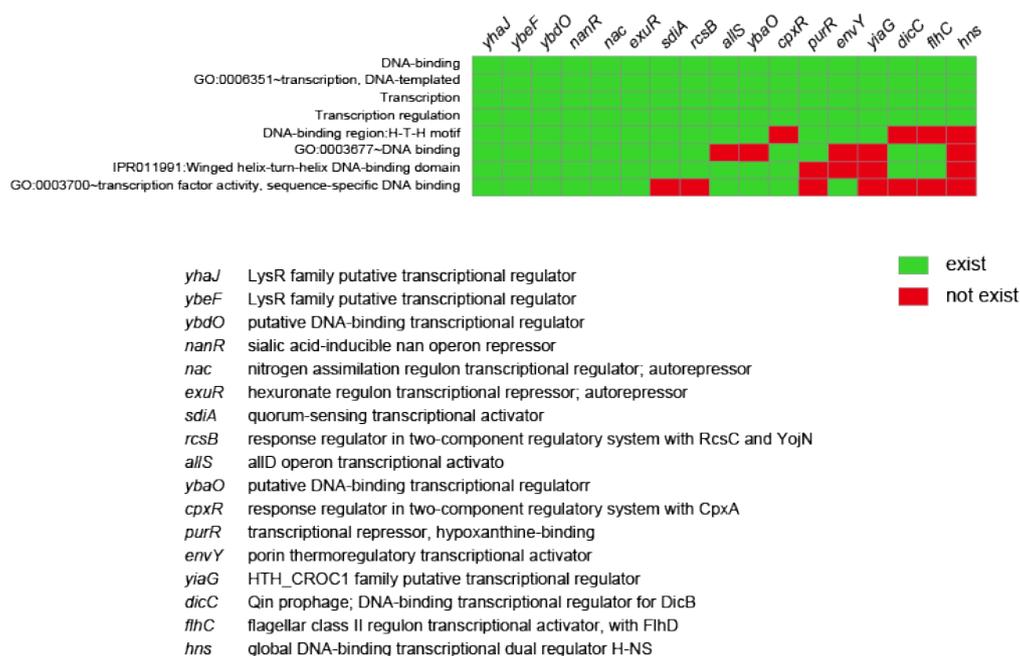
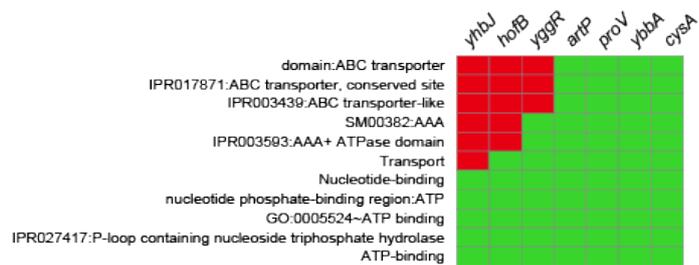


図26 呈色が亢進した遺伝子のクラスター1。DNA-bindingやTranscriptionといったtermをもつ遺伝子が集まった。*PcsgDEFG*に対して通常は負に制御している*cpxR*や*hns*、*rcsB*といった既知遺伝子が含まれている(表6)。また、motilityに必要なflagellaの生合成に関与する*flhC*が含まれていた。

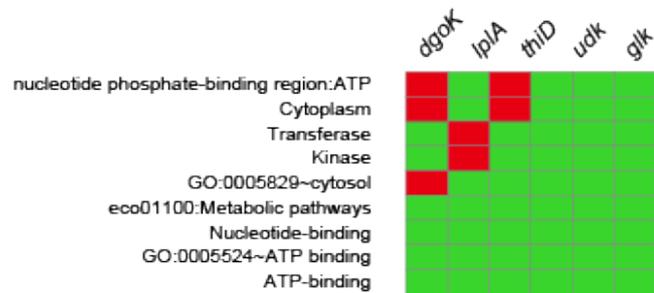
Cluster 2



yhbJ adaptor protein for GlmZ/GlmY sRNA decay, glucosamine-6-phosphate-regulated; NTPase
hofB T2SE secretion family protein; P-loop ATPase superfamily protein
yggR putative PflT family AAA+ ATPase
artP arginine ABC transporter ATPase
proV glycine betaine/proline ABC transporter periplasmic binding protei
ybbA putative ABC transporter ATPasen
cysA sulfate/thiosulfate transporter subunit

図27 呈色が亢進した遺伝子のクラスター2。

Cluster 3



dgoK 2-oxo-3-deoxygalactonate kinase
lplA lipoate-protein ligase A
thiD hydroxy-methylpyrimidine kinase and hydroxy-phosphomethylpyrimidine kinase
udk uridine/cytidine kinase
glk glucokinase

図28 呈色が充進した遺伝子のクラスター3。

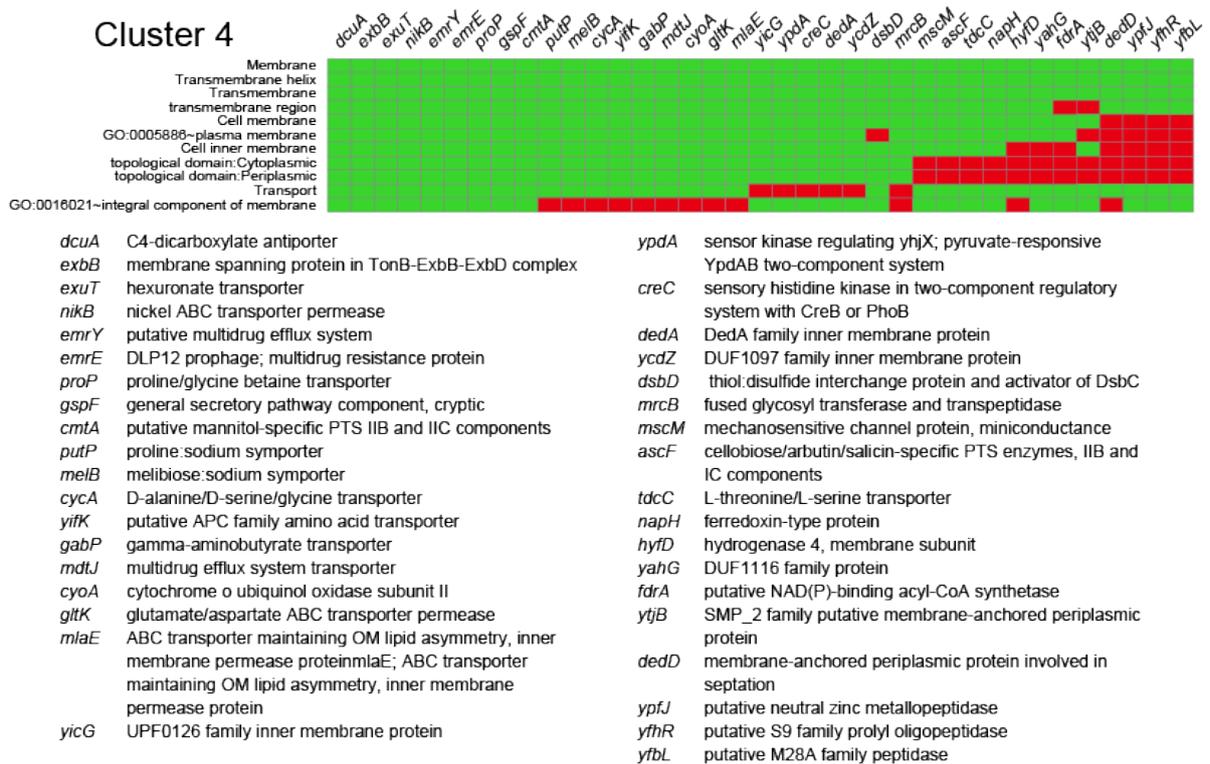
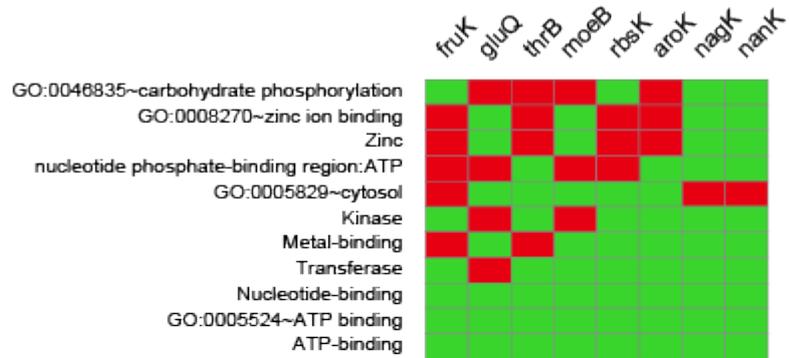


図29 呈色が充進した遺伝子のクラスター4。

White

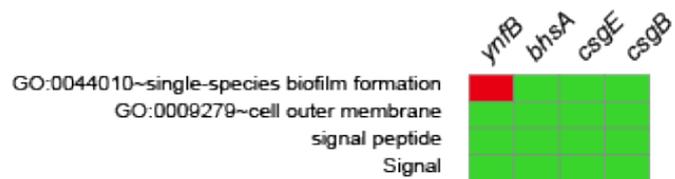
Cluster 1



fruK; fructose-1-phosphate kinase
gluQ; glutamyl-Q tRNA(Asp) synthetase
thrB; homoserine kinase
moeB; molybdopterin synthase sulfurylase
rbsK; ribokinase
aroK; shikimate kinase I
nagK; N-acetyl-D-glucosamine kinase
nanK; N-acetylmannosamine kinase

図30 呈色が下がった遺伝子のクラスター1。kinaseとしてアノテーションされている遺伝子が複数含まれていた (*fruK*、*thrB*、*rbsK*、*aroK*、*nagK*、*nanK*)。

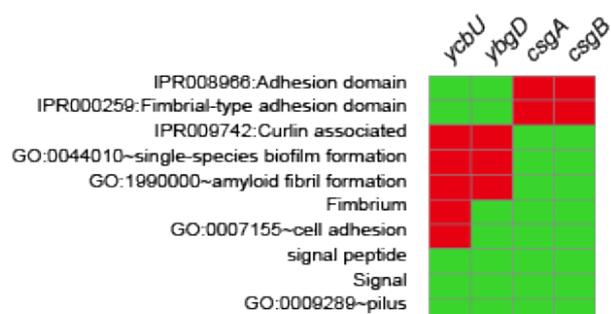
Cluster 2



ynfB UPF0482 family putative periplasmic protein
bhsA biofilm, cell surface and signaling protein
csgE curlin secretion specificity factor
csgB curlin nucleator protein, minor subunit in curli complex

図31 呈色が下がった遺伝子のクラスター2。 *csgBE*、*bhsA*はいずれもバイオフィームまたはCurli線毛合成に関与することが分かっている。*ynfB*は機能未知遺伝子。

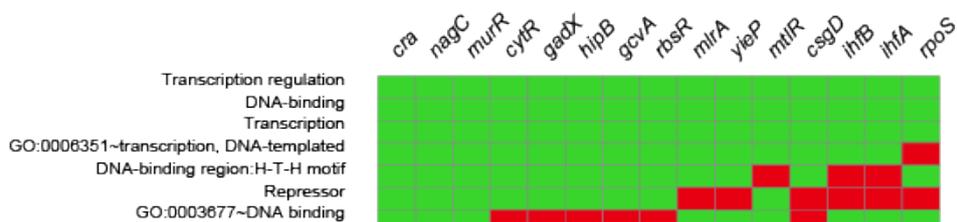
Cluster 3



ycbU putative fimbriae protein
ybgD putative fimbrial-like adhesin protein
csgA curlin subunit, amyloid curli fibers, cryptic
csgB curlin nucleator protein, minor subunit in curli complex

図32 呈色が下がった遺伝子のクラスター3。adhesion domainというtermをもつ遺伝子が集まった。*csgAB*ともにCurli線毛の構成要素。*ycbU*、*ybgD*はいずれもputativeではあるがfimbriae proteinとアノテーションされている。

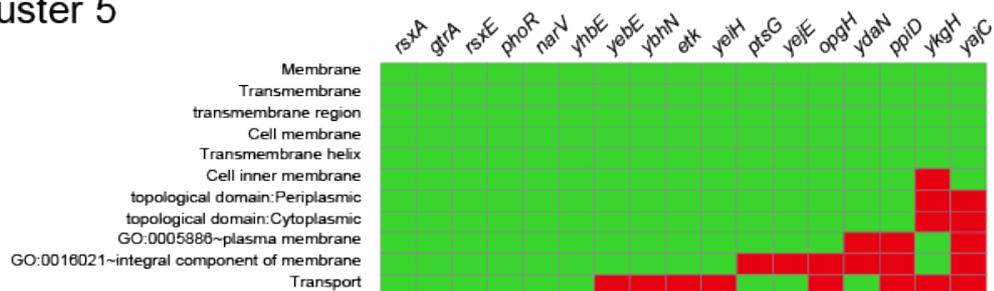
Cluster 4



- cra* transcriptional repressor-activator for carbon metabolism
nagC N-acetylglucosamine-inducible nag divergent operon transcriptional repressor
murR repressor for murPQ, MurNAc 6-P inducible
cytR Anti-activator for CytR-CRP nucleoside utilization regulon
gadX acid resistance regulon transcriptional activator; autoactivator
hipB antitoxin of HipAB toxin-antitoxin system
gcvA glycine cleavage system transcriptional activator; autorepressor
rbsR transcriptional repressor of ribose metabolism
mlrA transcriptional activator of csgD and csgBA
yieP putative transcriptional regulator
mtlR mannitol operon repressor
csgD csgBAC operon transcriptional regulator
ihfB integration host factor (IHF), DNA-binding protein, beta subunit
ihfA integration host factor (IHF), DNA-binding protein, alpha subunit
rpoS RNA polymerase, sigma S (sigma 38) factort

図33 呈色が下がった遺伝子のクラスター4。TranscriptionやDNA-bindingといったtermをもつ遺伝子が集まった。*yieP*以外は機能が判明している遺伝子。

Cluster 5



- rsxA* SoxR iron-sulfur cluster reduction factor component; inner membrane protein of electron transport complex
- gtrA* CPS-53 (KpLE1) prophage; bactoprenol-linked glucose translocase/flippase
- rsxE* SoxR iron-sulfur cluster reduction factor component; electron transport inner membrane NADH-quinone reductase
- phoR* sensory histidine kinase in two-component regulatory system with PhoB
- narV* nitrate reductase 2 (NRZ), gamma subunit
- yhbE* EamA family inner membrane putative transporter
- yebE* DUF533 family inner membrane protein
- ybhN* UPF0104 family inner membrane protein
- etk* tyrosine-protein kinase, role in O-antigen capsule formation
- yeiH* UPF0324 family inner membrane protein
- ptsG* fused glucose-specific PTS enzymes: IIB component/IIC component
- yejE* microcin C ABC transporter permease
- opgH* OPG biosynthetic ACP-dependent transmembrane UDP-glucose beta-1,2 glycosyltransferase; nutrient-dependent cell size regulator, FtsZ assembly antagonist
- ydaN* putative Zn(II) transporter
- ppiD* periplasmic folding chaperone, has an inactive PPIase domain
- ykgH* putative inner membrane protein
- yajC* SecYEG protein translocase auxillary subunit

図34 呈色が下がった遺伝子のクラスター5。membraneというtermをもつ遺伝子が集まった。

クリスタルバイオレット染色法によるバイオフィルム量の測定

呈色が亢進した株および呈色しなかった株の中から、特に呈色の変動している87株（呈色スコアが2以上または-2以下）について、クリスタルバイオレットを用いたバイオフィルム量の測定を行った（図35、36、表7-8）。バイオフィルムの量はBiofilm Index(OD590/OD600)に基づいて評価した。呈色スコアが高かった株の平均Biofilm Indexは0.949であった。中でも、*hns*(BI: 4.923)や*cpxR*(BI: 1.494)は、これまでの報告と同じ結果を示した（表6）。その他、*ycjU*(BI: 1.646)、*sufI*(BI: 1.385)といった株がWTと比してバイオフィルム量が上がっていた。

呈色スコアが低かった株の平均Biofilm Indexは0.641であった。特に*ihfB*(BI: 0.364)、*ihfA*(BI: 0.393)、*pepT*(BI: 0.400)はWTと比べて6割もBiofilm Indexが落ちていた。*ihfAB*は欠失によってバイオフィルムが下がることが既に確認されており、今回の結果と一致していた。また、コンゴレッド法では著しく低い呈色スコアを示す*csg*遺伝子群も軒並み低いBiofilm Indexであった（*csgA*: 0.497、*csgB*: 0.668、*csgD*: 0.567、*csgE*: 0.509、いずれもWTを1としたとき）。一方で、WTよりも高いBiofilm Indexを示す株も含まれており、その中には*csgD*を正に制御する*rpoS*も含まれていた（*rpoS*: 1.261）。

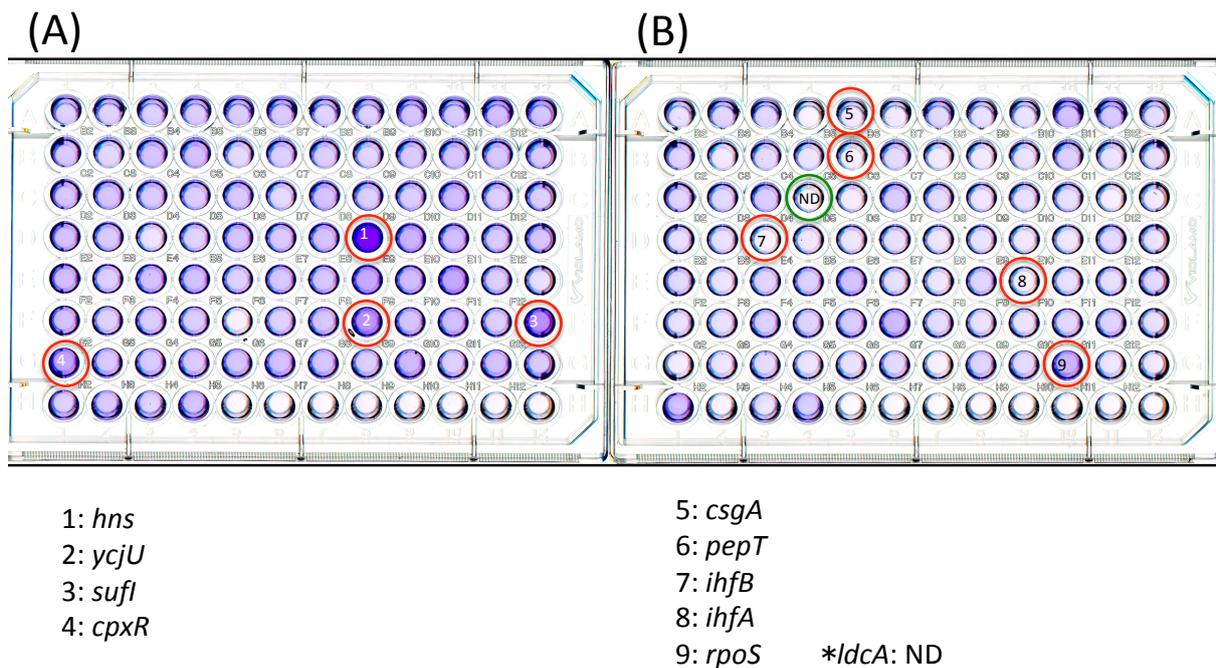


図35 クリスタルバイオレット法によるバイオフィルム検出の結果。(A)呈色スコアが2以上の87株。(B)呈色スコアが-2以下の87株。なお、*ldcA*は本実験で生育出来なかったためNDとした。

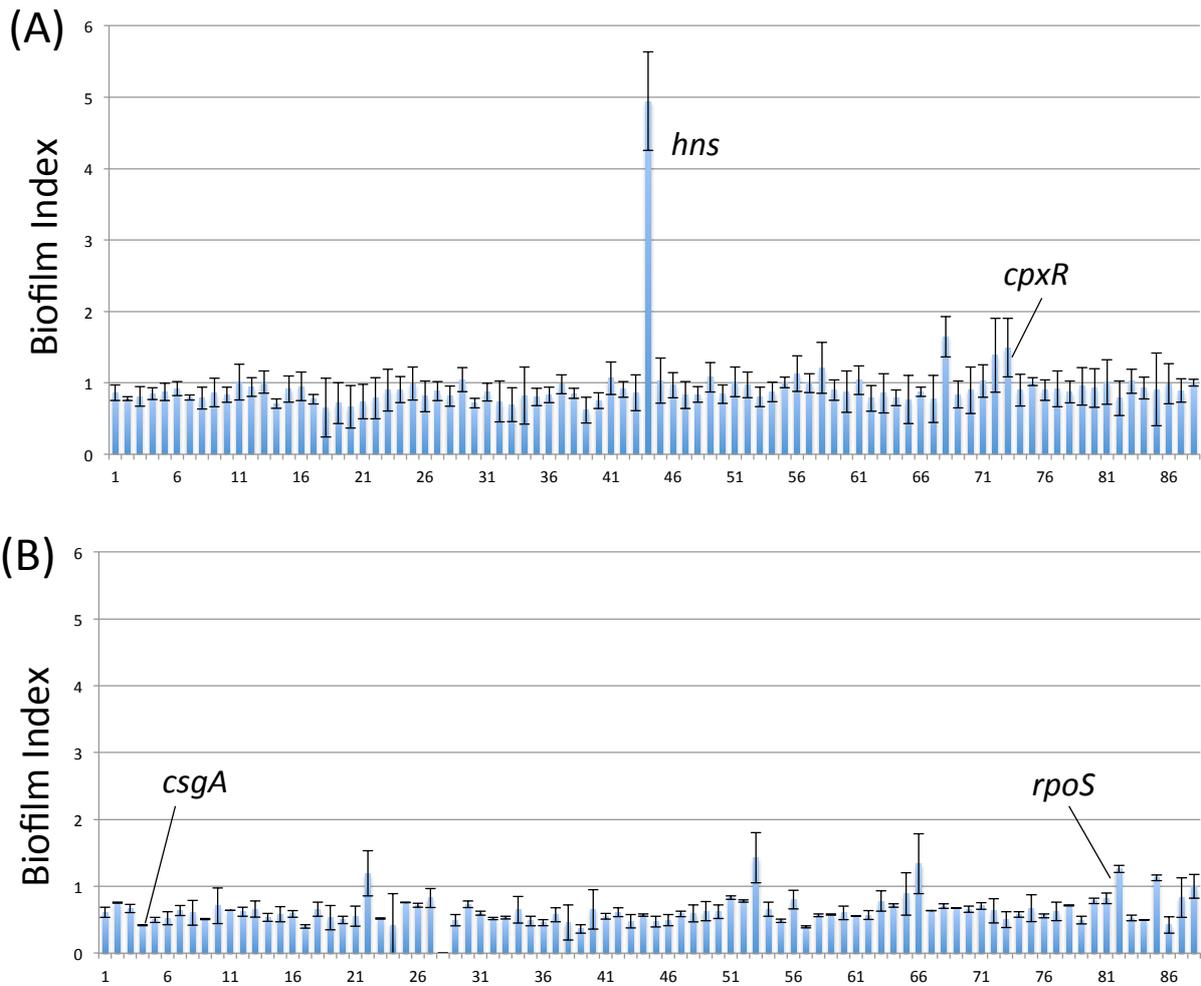


図36 バイオフィルム量をBiofilm Index(OD_{590}/OD_{600})で算出した結果。実験は独立で3回ずつ行い、WT(BW25113株)を1としたときのBiofilm Indexを表示している。エラーバーは各株における測定結果の標準偏差を指す。(A) 呈色スコアが2以上の87株(図35 A)。(B) 呈色スコアが2以下の87株(図35 B)。

pos_num	gene	OD590/OD600	std
1	<i>rcsF</i>	0.863129221	0.109836259
2	<i>idi</i>	0.780922709	0.028355203
3	<i>purR</i>	0.809490497	0.136713774
4	<i>yeaG</i>	0.848252513	0.082854651
5	<i>slt</i>	0.871579312	0.118881383
6	<i>fldB</i>	0.919665927	0.099940177
7	<i>ybeF</i>	0.79499171	0.038247839
8	<i>nac</i>	0.785049312	0.154186003
9	<i>cysG</i>	0.864017132	0.200763656
10	<i>ruvA</i>	0.83088333	0.105246935
11	<i>ruvC</i>	1.008653578	0.248750705
12	<i>proP</i>	0.941780899	0.131839934
13	<i>ypfJ</i>	1.012009793	0.156842272
14	<i>melB</i>	0.709587617	0.066129851
15	<i>cmtA</i>	0.911541215	0.182176841
16	<i>hofB</i>	0.949157852	0.199798336
17	<i>srlB</i>	0.768425309	0.066659851
18	<i>exuT</i>	0.651930773	0.409704328
19	<i>nfo</i>	0.714614266	0.28647593
20	<i>frdB</i>	0.663820067	0.299315298
21	<i>mrcB</i>	0.735739208	0.239309779
22	<i>cdd</i>	0.783207415	0.286988245
23	<i>xapA</i>	0.899306067	0.293638595
24	<i>nth</i>	0.907172651	0.183990155
25	<i>glnK</i>	0.989275922	0.232105159
26	<i>yhbJ</i>	0.810573442	0.215580282
27	<i>gltK</i>	0.883864417	0.135105575
28	<i>viaG</i>	0.814230483	0.142296386
29	<i>ybbS</i>	1.042423235	0.167483031
30	<i>recJ</i>	0.717565588	0.068860156
31	<i>amn</i>	0.868822946	0.122927501
32	<i>gpmI</i>	0.737020737	0.286502305
33	<i>adhE</i>	0.694520735	0.23759677
34	<i>yliK</i>	0.819996955	0.399699278
35	<i>surE</i>	0.801772534	0.122772831
36	<i>dgoK</i>	0.831462938	0.108934532
37	<i>yfaU</i>	0.981231327	0.133896828

38	<i>cysI</i>	0.850862069	0.070845239
39	<i>metL</i>	0.617948507	0.180410108
40	<i>dicC</i>	0.750687437	0.112675082
41	<i>cysJ</i>	1.063243489	0.227023959
42	<i>emrY</i>	0.907934538	0.113062371
43	<i>xerC</i>	0.859003424	0.249426195
44	<i>hns</i>	4.942904026	0.688554604
45	<i>dcuA</i>	1.032944265	0.317644602
46	<i>tfaS</i>	0.964587862	0.175648157
47	<i>cld</i>	0.830202858	0.189268126
48	<i>agp</i>	0.837308007	0.10894813
49	<i>nlpC</i>	1.07624836	0.204541356
50	<i>gcvH</i>	0.841523433	0.131082248
51	<i>smpA</i>	1.012061899	0.208528549
52	<i>ybcS</i>	0.972041847	0.180536777
53	<i>yjdL</i>	0.803645595	0.137026953
54	<i>yihM</i>	0.873146551	0.140359634
55	<i>ypdA</i>	1.004082946	0.075918806
56	<i>dedD</i>	1.130425607	0.250904174
57	<i>nanR</i>	0.995968966	0.131991941
58	<i>ydcS</i>	1.208128206	0.355544974
59	<i>yohF</i>	0.89887802	0.143777891
60	<i>flhE</i>	0.875942988	0.290066399
61	<i>yaaW</i>	1.037180799	0.200360987
62	<i>fliN</i>	0.781849667	0.182957641
63	<i>tdcC</i>	0.85204571	0.276695086
64	<i>yggN</i>	0.789876373	0.112091351
65	<i>fliJ</i>	0.765248082	0.337468198
66	<i>ygjR</i>	0.874715318	0.064810848
67	<i>astB</i>	0.775526147	0.331328303
68	<i>ycjU</i>	1.64556564	0.284419234
69	<i>glvC</i>	0.835864662	0.18683271
70	<i>ygjF</i>	0.897013313	0.327345982
71	<i>rumA</i>	1.024174184	0.22716967
72	<i>sufI</i>	1.385081883	0.517622543
73	<i>cpxR</i>	1.493698315	0.410170387
74	<i>sbp</i>	0.896080637	0.22048718
75	<i>ygiU</i>	1.018301382	0.057684641

76	<i>ppdB</i>	0.89782912	0.144010062
77	<i>glk</i>	0.916697702	0.253555556
78	<i>deaD</i>	0.871540101	0.153727526
79	<i>gatD</i>	0.952266481	0.265059969
80	<i>fumB</i>	0.926000595	0.270798569
81	<i>rcsB</i>	1.010641699	0.310629105
82	<i>xerD</i>	0.784830822	0.242289392
83	<i>malY</i>	1.022860282	0.169702794
84	<i>glpB</i>	0.926471511	0.155237701
85	<i>clpP</i>	0.906864009	0.507695982
86	<i>ybdG</i>	0.989610741	0.282827608
87	<i>trmH</i>	0.891212724	0.161830227
88	WT	1	0.048353847

表 7 呈色スコアが高かった株(図 33A、図 34A)における Biofilm Index。実験は独立で 3 回行い、その平均の Biofilm Index を表示している。

pos_num	gene	OD590/OD600	std
1	<i>fadD</i>	0.609401749	0.076114679
2	<i>trmU</i>	0.754956985	0.005693707
3	<i>rbsR</i>	0.667847645	0.060422669
4	<i>hfq</i>	0.418119066	0.006981576
5	<i>csgA</i>	0.497220048	0.041250261
6	<i>csgG</i>	0.524321637	0.09992205
7	<i>yeeW</i>	0.638200304	0.07496814
8	<i>nagC</i>	0.603620349	0.184267421
9	<i>csgE</i>	0.509403631	0.010575664
10	<i>ydfX</i>	0.709109722	0.268634838
11	<i>aroK</i>	0.640279386	0.000238698
12	<i>rpmF</i>	0.616427571	0.067832878
13	<i>crr</i>	0.66033818	0.120466127
14	<i>yfdG</i>	0.535415528	0.061404974
15	<i>ptsH</i>	0.584799207	0.114808851
16	<i>mutL</i>	0.587207418	0.04900706
17	<i>pepT</i>	0.400760346	0.027005759
18	<i>yheL</i>	0.656430166	0.104211668
19	<i>phoR</i>	0.530796478	0.180698898
20	<i>rbsK</i>	0.497496787	0.055626184
21	<i>mlrA</i>	0.553197719	0.149127122
22	<i>pyrB</i>	1.19226443	0.336495395
23	<i>yicC</i>	0.517712231	0.008028792
24	<i>pgi</i>	0.420403161	0.473167622
25	<i>yhhP</i>	0.756718758	0.003991463
26	<i>gnd</i>	0.71554895	0.028427103
27	<i>gadX</i>	0.825273427	0.141985898
28	<i>ldcA</i>	ND	ND
29	<i>hscB</i>	0.494072594	0.087696848
30	<i>nlpD</i>	0.731264022	0.048751868
31	<i>ybdN</i>	0.599449247	0.032007824
32	<i>rfaP</i>	0.520432173	0.01943494
33	<i>trpD</i>	0.532671856	0.021306909
34	<i>yfbM</i>	0.646973375	0.199114948
35	<i>flgD</i>	0.482064016	0.071920066
36	<i>fruR</i>	0.457224078	0.045842935
37	<i>csgB</i>	0.574480773	0.104281619

38	<i>bipA</i>	0.458541889	0.263179516
39	<i>ihfB</i>	0.363885548	0.063537445
40	<i>yieP</i>	0.656950764	0.295677523
41	<i>rfaF</i>	0.549233233	0.042350323
42	<i>ompC</i>	0.612017907	0.064771783
43	<i>php</i>	0.476445586	0.098557354
44	<i>thrB</i>	0.568658804	0.01830431
45	<i>fruK</i>	0.473839799	0.082914562
46	<i>rdgC</i>	0.494322155	0.082837267
47	<i>ycjS</i>	0.586084686	0.039077807
48	<i>yccT</i>	0.595144913	0.129132614
49	<i>ydaN</i>	0.629570513	0.143930296
50	<i>miaA</i>	0.619581163	0.101431672
51	<i>nusB</i>	0.832359714	0.026931539
52	<i>fabF</i>	0.781993666	0.019954886
53	<i>guaB</i>	1.427947732	0.37441453
54	<i>essQ</i>	0.657799633	0.107634113
55	<i>ddpD</i>	0.482840195	0.027259095
56	<i>hofQ</i>	0.799322034	0.13998605
57	<i>ihfA</i>	0.392635586	0.013955939
58	<i>csgD</i>	0.566508972	0.019962089
59	<i>thrA</i>	0.57929747	0.008906241
60	<i>gcvA</i>	0.602199045	0.101497256
61	<i>sspA</i>	0.559224175	0.005917696
62	<i>yffI</i>	0.570034856	0.065433994
63	<i>ybgD</i>	0.77941177	0.149854061
64	<i>yeiR</i>	0.712970726	0.024447444
65	<i>cytR</i>	0.887091789	0.319513384
66	<i>yejE</i>	1.33714771	0.450261992
67	<i>glcE</i>	0.635087363	0.000940194
68	<i>ydcL</i>	0.702681305	0.032190042
69	<i>trpE</i>	0.675329436	0.006161547
70	<i>ydbH</i>	0.655474604	0.046769077
71	<i>rfaG</i>	0.705308976	0.04811875
72	<i>yehE</i>	0.633427848	0.181068415
73	<i>ycfR</i>	0.502343401	0.121499652
74	<i>kbl</i>	0.578197281	0.039871931
75	<i>ydaF</i>	0.672121254	0.19979079

76	<i>glcD</i>	0.558403472	0.029225386
77	<i>ugd</i>	0.627312049	0.1403041
78	<i>uup</i>	0.714057815	0.006701578
79	<i>rpoZ</i>	0.49550915	0.057503105
80	<i>gmhB</i>	0.782293863	0.04012376
81	<i>yeaD</i>	0.819697173	0.078188582
82	<i>rpoS</i>	1.261597153	0.054259393
83	<i>rsxE</i>	0.523996813	0.043064142
84	<i>ybaJ</i>	0.497698913	0.001480273
85	<i>yccK</i>	1.125498832	0.043493875
86	<i>yccC</i>	0.422612154	0.122765153
87	<i>yajC</i>	0.831661141	0.297444306
88	WT	1	0.179333458

表 8 呈色スコアが低かった株(図 33B、図 34B)における Biofilm Index。実験は独立で 3 回行い、その平均の Biofilm Index を表示している。

IV. 考察

Red Cluster1には*csgD*を負に制御する因子のいくつかが集まった(*cpxR*, *hns*, *rcsB*, 表6)。その他、motilityに必要なflagellaの生合成に関与する*flhC*が含まれていた。FlhCはflagellaのオペロンの転写を調節するマスター転写調節因子として、FlhDと複合体を形成し機能する。一般に、菌はバイオフィルムの状態と水中を浮遊するplanktonicな状態では相反する遺伝子発現制御となっている。菌がバイオフィルムの状態の際、CsgDは*flhCD*の転写を抑制しており、motilityも見られない(Chambers & Sauer, 2013)。*flhC*の欠失によってバイオフィルムの状態がより亢進したことで赤くなった可能性が考えられる。また、Curli合成への阻害機能が知られている*sdiA*が含まれていた。*sdiA*はCurli合成への阻害機能を持つ(Lee, Maeda, Hong, & Wood, 2009)ことから、こちらも先行研究に一致する結果が得られた。Cluster1に属する機能未知遺伝子*yhaJ*, *ybeF*, *ybdO*, *ybaO*に関しても、転写調節因子としてバイオフィルム形成に関わっている可能性が示唆される。これらの欠失株について遺伝子発現解析を行うことで、調節関係を明らかにできる可能性がある。

Red Cluster2に含まれる*yhbJ*はsmall regulatory RNAであるGlmZの不安定化に関与していることが示されている(Kalamorz, Reichenbach, März, Rak, & Görke, 2007)。また、*glmZ*はRNA chaperoneであるhfqと共にglucosamine-6-phosphate synthase (GlmS)の安定化に関与している(Göpel, Papenfort, Reichenbach, Vogel, & Görke, 2013; Kalamorz et al., 2007)。GlcN-6-P(glucosamine-6-phosphate)はバクテリアのcell wallであるpeptidoglycanやlipopolysaccharide(LPS)を合成する経路における重要な物質である。GlmSはfructose-6-phosphateとglutamineをGlcN-6-Pとglutamateに変換する役割を果たす(Kalamorz et al., 2007)。YhbJはRNase Eのadaptor proteinとされ、*glmZ*がRNaseEによって切断されるプロセスを促進する(Göpel et al., 2013)。実際、*yhbJ*の破壊株では細胞内のGlmSの量が増加することが確認されている(Kalamorz et al., 2007)。また、*glmZ*はCsgDも制御していることが分かっている。*glmZ*欠失株では*csgD*のmRNA量が減っていることまでは判明しているが、詳しいメカニズムは依然として不明である(Gruber & Sperandio, 2015)。これらの文献情報から、遺伝子欠失によって呈色が亢進した株の一部に関しては、細胞内のUDP-GlcNacレベルが上昇したために呈色が強くなったと考えられる。一方、GlmUSのプロモーター上流にはNagC(N-acetylglucosamine transcriptional dual regulator)の結合領域が二つ存在しており、GlmUSの転写を促進している(Plumbridge, 2001)。今回の結果においては、*nagC*欠失株は呈色が極端に低い株として検出されており(図33)、文献情報と照らし合わせても矛盾がない。

Red Cluster3, 4に関してはそれぞれ頻出したtermは確認できたが、そこから生物学的な解釈をつけるまでには現時点では至らなかった。ただし、Red Cluster3を構成しているtermであるATP-bindingやkinaseはどちらもタンパク質をリン酸化する意味を持っており、特に*udk*はピリミジンのサルベージ経路、*glk*はグリコーゲン分解経路にそれぞれマップされるkinaseである。これら経路の遮断が原因でCurliの量に変化が生じたと考えられる。Red Cluster4に頻出したtermはmembrane/transmembraneであった。Curliがouter membraneにアンカリングしている構造体であることを考慮すると、遺伝

子欠失によって膜の状態が変化したことでCurliの量が通常よりも上昇したと考えられる。

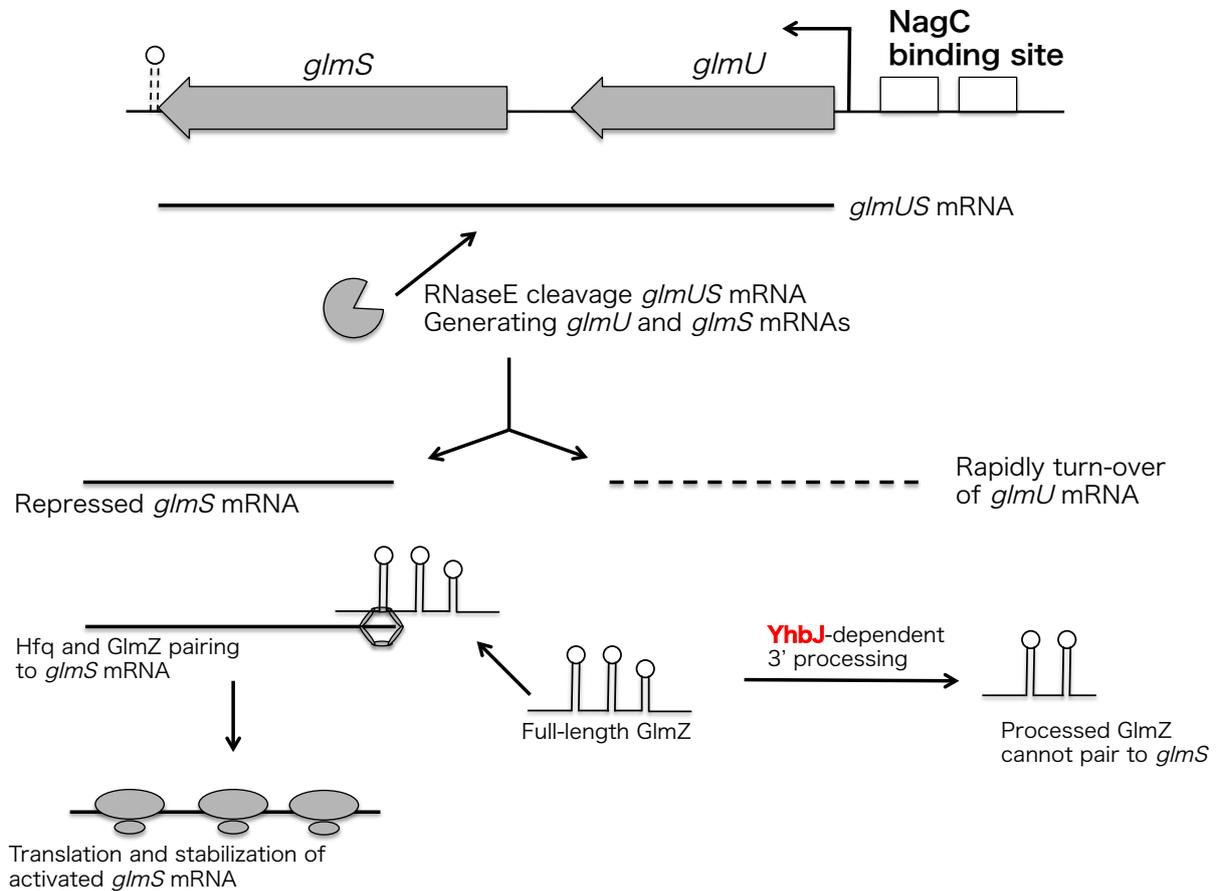


図 37 GlmSU の転写後発現制御に関与する GlmZ および YhbJ の働き。GlmZ は small regulatory RNA であり、RNA chaperone である Hfq と共に GlmS の mRNA の安定化と翻訳の促進に寄与している。YhbJ は RNaseE の adaptor protein であり、GlmZ を切断することで、機能を阻害する。GlmSU のプロモーター上流には NagC 結合領域が存在しており、*glmUS* オペロンの転写に寄与している。

White Cluster1にはkinaseとしてアノテーションされている遺伝子が複数含まれていた (*fruK*, *thrB*, *rbsK*, *aroK*, *nagK*, *nanK*)。このうち、*nagK*、*nanK*はいずれもUDP-GlcNac(Uridine diphosphate N-acetylglucosamine)合成に必要な経路上にマップされる遺伝子である。UDP-GlcNacはpeptidoglycanやlipidAなど、cell wall合成に関わる経路における重要な物質である。*nagK*、*nanK*の欠失はUDP-GlcNacの細胞内レベルが低下することが予想される。Curliが外膜にアンカリングしていることから考えても、UDP-GlcNacの細胞内レベルの低下が、cell wallの不安定化を誘導しCurli合成低下に繋がっているという可能性を想起させる。実際、*nagA*、*nagK*、*nanK*、*nanE*の欠失株でCurliの量が低下していることがすでに報告されている(Smith et al., 2017)。また、*nanATEK-yhcH*(sialic acid catabolic operon)は*nanR*と呼ばれるリプレッサーによって負に制御されている(Kalivoda, Steenbergen, Vimr, & Plumbridge, 2003)。*nanR*欠失株は呈色スコアが高くなる株として検出されており(いくつかのリプレッサーを含むRed cluster1の一部)、リプレッサーの欠失によって*nanK*の発現が上昇し、Curliの量が増加したというモデルが考えられる(図38)。White Cluster2に含まれていた*ynfB*だが、機能未知ではあるがクラスター内の他の3つの遺伝子はいずれもバイオフィームに関係する遺伝子であることから、*ynfB*も同様にバイオフィームに関係していると考えられる。White Cluster4にはバイオフィーム形成に関与することが既知である転写因子が集まった。特に、*cra*、*mlrA*、*ihfA*、*csgD*は通常*PcsgDEFG*を正に制御する因子であることから、欠失によって呈色下がるという結果は妥当である。唯一*yieP*とよばれる機能未知遺伝子はそのクラスターに含まれていた。*yieP*のCsgA量を実際にウェスタンブロットで確認したところ、確かにWTよりも著しく下がっていた(図39)。このことから、少なくとも*yieP*がCurli線毛生合成に関して何らかの機能を持っていることは明らかである。DAVIDの結果から、*yieP*は転写因子の可能性が高い。今後はRNAseqなどの手法を用いて、*yieP*がどのような制御によって最終的にCurli線毛の合成に関与しているのかを探ることが望まれる。

今回、DAVIDによって遺伝子の機能をクラスタリングしたが、その際にDAVIDの解析対象外となってしまった(クラスターに入っていない)株がかなりを占めた(Red: 102株、White: 119株)。これらの株はクラスタリングからはじかれただけであり、Curli線毛の生合成に関与していないとは限らない。今後、これらはじかれた株の情報についても個別で精査する必要がある。

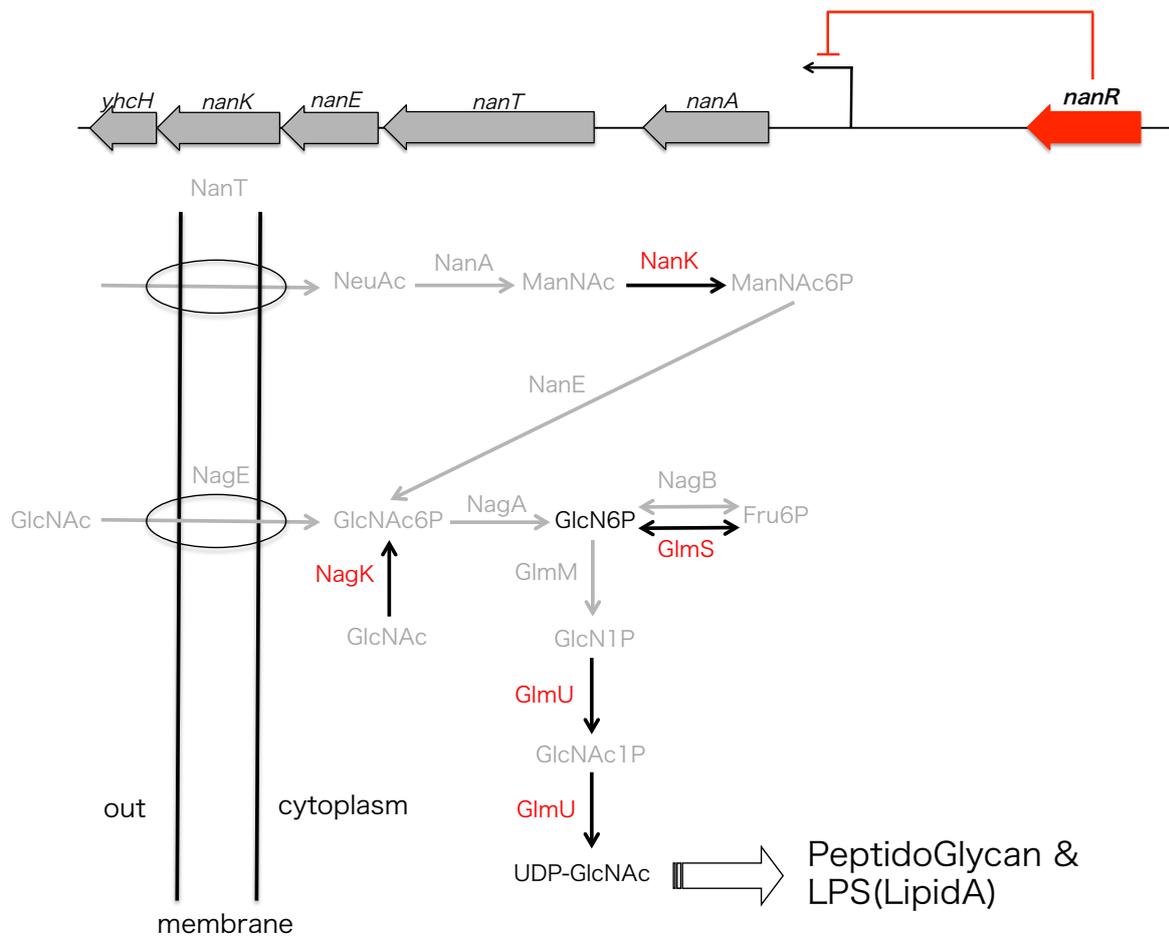


図 38 *nan* オペロンと UDP-GlcNAc 合成経路の概略図。シアル酸代謝に必要な *nan* オペロンは NanR(N-acetyl-neuraminic acid regulator)によって負に制御されている。*nanR* 欠失株は今回の実験において呈色が亢進した株として検出された。赤文字で示された酵素をコードする遺伝子欠失株は全て呈色が著しく低下した株として検出されている。

いくつかの株に関しては実際にバイオフィルム量を測定した。呈色スコアが高かった87株の平均のBiofilm Indexは0.949(WT比)であり、WTとそれほど変わらなかったのに対し、呈色スコアが低かった87株の平均Biofilm Indexは0.641と低い値を示した(図33-34)。このことは、コンゴレッド法でスクリーニングされた呈色が低かった株は平均してバイオフィルム量も低いということを示している。一方で、コンゴレッド法で呈色スコアが高い、あるいは低いからといってバイオフィルム量も必ず一致するというわけではなかった。特に顕著だったのは*rpoS*であった。*rpoS*は本実験ではBiofilm Indexが1.261(WT比)であった。数多くの実験で、RpoSがCsgDを正に制御していることは既に明らかにしているが、他方で*rpoS*欠失株のバイオフィルム量が増加する事実は確かに指摘されている(Corona-Izquierdo & Membrillo-Hernández, 2002; Kim, Lee, & Kim, 2013)。また、今回の実験において*csgA*を欠失させた株は確かにバイオフィルム量が低下していたが、最も低い値ではなかった。このことは、Curli線毛が欠失していても大腸菌はバイオフィルムを形成することが出来ることを意味しており、状況に合わせて巧みにバイオフィルムの構成成分を変化させていると推察される。

本研究ではCurli線毛の量に関して呈色スコアを用いて網羅的に観測することができた。呈色スコアの結果は、概ねクリスタルバイオレット法の結果と一致していたが、一部遺伝子に関しては異なる傾向を観測した。今後は、網羅的な呈色スコアのデータとクリスタルバイオレット法を用いたバイオフィルム量の測定結果の共通点を探し出し、どの遺伝子がバイオフィルム形成にとってコアになっているのか、そしてどの遺伝子がバイオフィルムの構成成分に柔軟性を持たせているのかを明らかにできる可能性と考えられる。

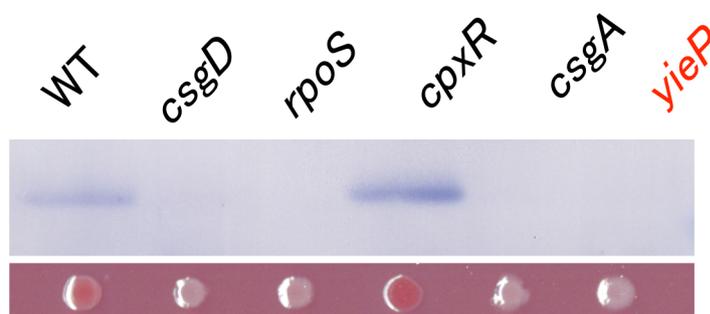


図39 ウェスタンブロットによる*yieP*と他の遺伝子欠失株におけるCsgA量の比較およびCRA上に各株のコロニーの写真。各株の培養液を2 μ lを培地上に滴下し28°Cで32時間培養した。

謝辞

本研究を行うにあたり、適切な助言・指導を頂きました森浩禎教授、武藤愛助教に厚く御礼申し上げます。当研究室にてご自身の研究の合間をぬって、細かくかつ丁寧に助言・指導頂いた任田八恵氏、道瀬ひとみ氏、竹内力矢氏に心から感謝申し上げます。本論文の作成にあたって助言、審査を担当してくださいました、奈良先端科学技術大学、高木博史教授、塚崎智也准教授、佐藤智和准教授、CsgAおよびFtsZ抗体を分けて頂いた慈恵医大・医学部・細菌学講座 杉本真也講師、皆様方に厚く御礼申し上げます。最後になりましたが、長い学生生活において経済的にも精神的にも支えてくださった両親と家族に心から感謝の気持ちを述べさせて頂き、論文の結びとさせて頂きます。

参考文献

- Alper, H., Miyaoku, K., & Stephanopoulos, G. (2005). Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nature Biotechnology*, 23(5), 612–6. <http://doi.org/10.1038/nbt1083>
- Arnqvist, A., Olsén, A., Pfeifer, J., Russell, D. G., & Normark, S. (1992). The Crl protein activates cryptic genes for curli formation and fibronectin binding in *Escherichia coli* HB101. *Molecular Microbiology*, 6(17), 2443–2452. <http://doi.org/10.1111/j.1365-2958.1992.tb01420.x>
- Arnqvist, A., Olsén, A., & Normark, S. (1994). Sigma S-dependent growth-phase induction of the *csgBA* promoter in *Escherichia coli* can be achieved in vivo by sigma 70 in the absence of the nucleoid-associated protein H-NS. *Molecular Microbiology*, 13(6), 1021–1032. <http://doi.org/10.1111/j.1365-2958.1994.tb00493.x>
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., ... Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2, 2006.0008. <http://doi.org/10.1038/msb4100050>
- Barker, C. A., Farha, M. A., & Brown, E. D. (2010). Chemical genomic approaches to study model microbes. *Chemistry & Biology*, 17(6), 624–32. <http://doi.org/10.1016/j.chembiol.2010.05.010>
- Bougdour, A., Lelong, C., & Geiselmann, J. (2004). Crl, a low temperature-induced protein in *Escherichia coli* that binds directly to the stationary phase sigma subunit of RNA polymerase. *The Journal of Biological Chemistry*, 279(19), 19540–50. <http://doi.org/10.1074/jbc.M314145200>
- Brandner, C. J., Maier, R. H., Henderson, D. S., Hintner, H., Bauer, J. W., & Onder, K. (2008). The ORFeome of *Staphylococcus aureus* v 1.1. *BMC Genomics*, 9, 321. <http://doi.org/10.1186/1471-2164-9-321>
- Brettin, T., Altherr, M. R., Du, Y., Mason, R. M., Friedrich, A., Potter, L., ... Arvidson, C. G. (2005). Expression capable library for studies of *Neisseria gonorrhoeae*, version 1.0. *BMC Microbiology*, 5, 50. <http://doi.org/10.1186/1471-2180-5-50>
- Brown, P. K., Dozois, C. M., Nickerson, C. a, Zuppardo, a, Terlunge, J., & Curtiss, R. (2001). MlrA, a novel regulator of curli (AgF) and extracellular matrix synthesis by *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. *Molecular Microbiology*, 41(2), 349–363. <http://doi.org/10.1046/j.1365-2958.2001.02529.x>
- Chambers, J. R., & Sauer, K. (2013). Small RNAs and their role in biofilm formation. *Trends in Microbiology*, 21(1), 39–49. <http://doi.org/10.1016/j.tim.2012.10.008>
- Collins, S. R., Schuldiner, M., Krogan, N. J., & Weissman, J. S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biology*, 7(7), R63. <http://doi.org/10.1186/gb-2006-7-7-r63>
- Corona-Izquierdo, F. P., & Membrillo-Hernández, J. (2002). A mutation in *rpoS* enhances biofilm formation in *Escherichia coli* during exponential phase of growth. *FEMS Microbiology Letters*, 211(1), 105–110. [http://doi.org/10.1016/S0378-1097\(02\)00672-9](http://doi.org/10.1016/S0378-1097(02)00672-9)
- Costerton, J. W., Stewart, P. S., & Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science (New York, N.Y.)*, 284(5418), 1318–22. <http://doi.org/10.1038/nbt1083>
- Donald, R. G. K., Skwish, S., Forsyth, R. A., Anderson, J. W., Zhong, T., Burns, C., ... Phillips, J. W. (2009). A *Staphylococcus aureus* fitness test platform for mechanism-based profiling of antibacterial compounds. *Chemistry & Biology*, 16(8), 826–36. <http://doi.org/10.1016/j.chembiol.2009.07.004>

- Dorel, C., Vidal, O., Prigent-Combaret, C., Vallet, I., & Lejeune, P. (1999). Involvement of the Cpx signal transduction pathway of *E. coli* in biofilm formation. *FEMS Microbiology Letters*, *178*(1), 169–175. [http://doi.org/10.1016/S0378-1097\(99\)00347-X](http://doi.org/10.1016/S0378-1097(99)00347-X)
- Eisert, R., Felau, L., & Brown, L. R. (2006). Methods for enhancing the accuracy and reproducibility of Congo red and thioflavin T assays. *Analytical Biochemistry*, *353*(1), 144–6. <http://doi.org/10.1016/j.ab.2006.03.015>
- Evans, M. L., & Chapman, M. R. (2014). Curli biogenesis: Order out of disorder. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1843*(8), 1551–1558. <http://doi.org/10.1016/j.bbamcr.2013.09.010>
- Fang, X., & Gomelsky, M. (2010). A post-translational, c-di-GMP-dependent mechanism regulating flagellar motility. *Molecular Microbiology*, *76*(5), 1295–305. <http://doi.org/10.1111/j.1365-2958.2010.07179.x>
- Ferrieres, L., & Clarke, D. J. (2003). The RcsC sensor kinase is required for normal biofilm formation in *Escherichiacoli* K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol Microbiol*, *50*(5), 1665–1682. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14651646
- Galperin, M. Y. (2004). Bacterial signal transduction network in a genomic perspective. *Environmental Microbiology*, *6*(6), 552–67. <http://doi.org/10.1111/j.1462-2920.2004.00633.x>
- Gelperin, D. M., White, M. A., Wilkinson, M. L., Kon, Y., Kung, L. A., Wise, K. J., ... Grayhack, E. J. (2005). Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes & Development*, *19*(23), 2816–26. <http://doi.org/10.1101/gad.1362105>
- Gerstel, U., Kolb, A., & Römling, U. (2006). Regulatory components at the *csgD* promoter - Additional roles for OmpR and integration host factor and role of the 5' untranslated region. *FEMS Microbiology Letters*, *261*(1), 109–117. <http://doi.org/10.1111/j.1574-6968.2006.00332.x>
- Gerstel, U., Park, C., & Römling, U. (2003). Complex regulation of *csgD* promoter activity by global regulatory proteins. *Molecular Microbiology*, *49*(3), 639–654. <http://doi.org/10.1046/j.1365-2958.2003.03594.x>
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., ... Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, *418*(6896), 387–91. <http://doi.org/10.1038/nature00935>
- Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., ... Oliver, S. G. (1996). Life with 6000 genes. *Science (New York, N.Y.)*, *274*(5287), 546, 563–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8849441>
- Goller, C., Wang, X., Itoh, Y., & Romeo, T. (2006). The cation-responsive protein NhaR of *Escherichia coli* activates *pgaABCD* transcription, required for production of the biofilm adhesin poly-beta-1,6-N-acetyl-D-glucosamine. *Journal of Bacteriology*, *188*(23), 8022–32. <http://doi.org/10.1128/JB.01106-06>
- Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J., & Görke, B. (2013). Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes and Development*, *27*(5), 552–564. <http://doi.org/10.1101/gad.210112.112>
- Grant, W. D., Sutherland, I. W., & Wilkinson, J. F. (1969). Exopolysaccharide colanic acid and its occurrence in the Enterobacteriaceae. *Journal of Bacteriology*, *100*(3), 1187–93. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4902806>

- Gruber, C. C., & Sperandio, V. (2015). Global analysis of posttranscriptional regulation by GlmY and GlmZ in enterohemorrhagic *Escherichia coli* O157:H7. *Infection and Immunity*, 83(4), 1286–1295. <http://doi.org/10.1128/IAI.02918-14>
- Hammar, M., Arnqvist, A., Bian, Z., Olsén, A., & Normark, S. (1995). Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Molecular Microbiology*, 18(4), 661–70. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8817489>
- Hammar, M., Bian, Z., & Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 93(13), 6562–6566. <http://doi.org/10.1073/pnas.93.13.6562>
- Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nature Reviews. Microbiology*, 7(4), 263–73. <http://doi.org/10.1038/nrmicro2109>
- Hengge, R., Galperin, M. Y., Ghigo, J.-M., Gomelsky, M., Green, J., Hughes, K. T., ... Landini, P. (2015). Systematic Nomenclature for GGDEF and EAL Domain-Containing Cyclic Di-GMP Turnover Proteins of *Escherichia coli*. *Journal of Bacteriology*, 198(1), 7–11. <http://doi.org/10.1128/JB.00424-15>
- Heuermann, K., & Cosgrove, J. (2001). S-Gal: an autoclavable dye for color selection of cloned DNA inserts. *BioTechniques*, 30(5), 1142–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11355350>
- Holmqvist, E., Reimegård, J., Sterk, M., Grantcharova, N., Römling, U., & Wagner, E. G. H. (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *The EMBO Journal*, 29(11), 1840–1850. <http://doi.org/10.1038/emboj.2010.73>
- Hudson, J. R., Dawson, E. P., Rushing, K. L., Jackson, C. H., Lockshon, D., Conover, D., ... Fields, S. (1997). The complete set of predicted genes from *Saccharomyces cerevisiae* in a readily usable form. *Genome Research*, 7(12), 1169–73. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9414322>
- Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., ... Manoil, C. (2003). Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24), 14339–44. <http://doi.org/10.1073/pnas.2036282100>
- Jørgensen, M. G., Thomason, M. K., Havelund, J., Valentin-Hansen, P., & Storz, G. (2013). Dual function of the McaS small RNA in controlling biofilm formation. *Genes & Development*, 27(10), 1132–45. <http://doi.org/10.1101/gad.214734.113>
- Jubelin, G., Vianney, A., Beloin, C., Ghigo, J.-M., Lazzaroni, J.-C., Lejeune, P., & Dorel, C. (2005). CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in *Escherichia coli*. *Journal of Bacteriology*, 187(6), 2038–49. <http://doi.org/10.1128/JB.187.6.2038-2049.2005>
- Kalamorz, F., Reichenbach, B., März, W., Rak, B., & Görke, B. (2007). Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*. *Molecular Microbiology*, 65(6), 1518–1533. <http://doi.org/10.1111/j.1365-2958.2007.05888.x>
- Kalivoda, K. A., Steenbergen, S. M., Vimr, E. R., & Plumbridge, J. (2003). Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*. *Journal of Bacteriology*, 185(16), 4806–4815. <http://doi.org/10.1128/JB.185.16.4806-4815.2003>
- Karow, M., Raina, S., Georgopoulos, C., & Fayet, O. (1991). Complex phenotypes of null mutations in the htr genes, whose products are essential for *Escherichia coli* growth at elevated temperatures. *Research in Microbiology*, 142(2-3), 289–294. [http://doi.org/10.1016/0923-2508\(91\)90043-A](http://doi.org/10.1016/0923-2508(91)90043-A)

- Khurana, R., Uversky, V. N., Nielsen, L., & Fink, A. L. (2001). Is Congo red an amyloid-specific dye? *The Journal of Biological Chemistry*, *276*(25), 22715–21. <http://doi.org/10.1074/jbc.M011499200>
- Kim, T., Lee, J., & Kim, K. S. (2013). Escherichia coli YmdB regulates biofilm formation independently of its role as an RNase III modulator. *BMC Microbiology*, *13*(1). <http://doi.org/10.1186/1471-2180-13-266>
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., & Mori, H. (2005). Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes*, *12*(5), 291–9. <http://doi.org/10.1093/dnares/dsi012>
- Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., ... Ogasawara, N. (2003). Essential Bacillus subtilis genes. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(8), 4678–83. <http://doi.org/10.1073/pnas.0730515100>
- Korch, S. B., & Hill, T. M. (2006). Ectopic overexpression of wild-type and mutant hipA genes in Escherichia coli: Effects on macromolecular synthesis and persister formation. *Journal of Bacteriology*, *188*(11), 3826–3836. <http://doi.org/10.1128/JB.01740-05>
- Labaer, J., Qiu, Q., Anumanthan, A., Mar, W., Zuo, D., Murthy, T. V. S., ... Brizuela, L. (2004). The Pseudomonas aeruginosa PA01 gene collection. *Genome Research*, *14*(10B), 2190–200. <http://doi.org/10.1101/gr.2482804>
- Lease, R. A., & Belfort, M. (2000). Riboregulation by DsrA RNA: trans-actions for global economy. *Molecular Microbiology*, *38*(4), 667–72. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11115103>
- Lee, J., Maeda, T., Hong, S. H., & Wood, T. K. (2009). Reconfiguring the quorum-sensing regulator sdiA of escherichia coli to control biofilm formation via indole and N-acylhomoserine lactones. *Applied and Environmental Microbiology*, *75*(6), 1703–1716. <http://doi.org/10.1128/AEM.02081-08>
- Liu, A., Tran, L., Becket, E., Lee, K., Chinn, L., Park, E., ... Miller, J. H. (2010). Antibiotic sensitivity profiles determined with an Escherichia coli gene knockout collection: generating an antibiotic bar code. *Antimicrobial Agents and Chemotherapy*, *54*(4), 1393–403. <http://doi.org/10.1128/AAC.00906-09>
- Liu, X., & De Wulf, P. (2004). Probing the ArcA-P Modulon of Escherichia coli by Whole Genome Transcriptional Analysis and Sequence Recognition Profiling. *Journal of Biological Chemistry*, *279*(13), 12588–12597. <http://doi.org/10.1074/jbc.M313454200>
- Ma, Q., & Wood, T. K. (2009). OmpA influences Escherichia coli biofilm formation by repressing cellulose production through the CpxRA two-component system. *Environmental Microbiology*, *11*(10), 2735–46. <http://doi.org/10.1111/j.1462-2920.2009.02000.x>
- Macconkey, A. (1905). Lactose-Fermenting Bacteria in Faeces. *The Journal of Hygiene*, *5*(3), 333–79. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20474229>
- Madsen, J. S., Burmølle, M., Hansen, L. H., & Sørensen, S. J. (2012). The interconnection between biofilm formation and horizontal gene transfer. *FEMS Immunology and Medical Microbiology*, *65*(2), 183–95. <http://doi.org/10.1111/j.1574-695X.2012.00960.x>
- Mika, F., Busse, S., Possling, A., Berkholz, J., Tschowri, N., Sommerfeldt, N., ... Hengge, R. (2012). Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in Escherichia coli. *Molecular Microbiology*, *84*(1), 51–65. <http://doi.org/10.1111/j.1365-2958.2012.08002.x>
- Monteiro, C., Papenfort, K., Hentrich, K., Ahmad, I., Le Guyon, S., Reimann, R., ... Römling, U. (2012). Hfq and Hfq-dependent small RNAs are major contributors to

- multicellular development in *Salmonella enterica* serovar Typhimurium. *RNA Biology*, 9(4), 489–502. <http://doi.org/10.4161/rna.19682>
- Niba, E. T. E., Naka, Y., Nagase, M., Mori, H., & Kitakawa, M. (2007). A genome-wide approach to identify the genes involved in biofilm formation in *E. coli*. *DNA Res*, 14(6), 237–246. <http://doi.org/10.1093/dnares/dsm024>
- Nijland, R., Hall, M. J., & Burgess, J. G. (2010). Dispersal of biofilms by secreted, matrix degrading, bacterial DNase. *PloS One*, 5(12), e15668. <http://doi.org/10.1371/journal.pone.0015668>
- Ogasawara, H., Hasegawa, A., Kanda, E., Miki, T., Yamamoto, K., & Ishihama, A. (2007). Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade. *Journal of Bacteriology*, 189(13), 4791–4799. <http://doi.org/10.1128/JB.00319-07>
- Ogasawara, H., Shinohara, S., Yamamoto, K., & Ishihama, A. (2012). Novel regulation targets of the metal-response BasS-BasR two-component system of *Escherichia coli*. *Microbiology (United Kingdom)*, 158(6), 1482–1492. <http://doi.org/10.1099/mic.0.057745-0>
- Ogasawara, H., Yamada, K., Kori, A., Yamamoto, K., & Ishihama, A. (2010). Regulation of the *Escherichia coli* csgD promoter: Interplay between five transcription factors. *Microbiology*, 156(8), 2470–2483. <http://doi.org/10.1099/mic.0.039131-0>
- Ogasawara, H., Yamamoto, K., & Ishihama, A. (2010). Regulatory role of MlrA in transcription activation of csgD, the master regulator of biofilm formation in *Escherichia coli*. *FEMS Microbiology Letters*. <http://doi.org/10.1111/j.1574-6968.2010.02112.x>
- Ogasawara, H., Yamamoto, K., & Ishihama, A. (2011). Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. *Journal of Bacteriology*, 193(10), 2587–2597. <http://doi.org/10.1128/JB.01468-10>
- Okshesky, M., Regina, V. R., & Meyer, R. L. (2015). Extracellular DNA as a target for biofilm control. *Current Opinion in Biotechnology*, 33, 73–80. <http://doi.org/10.1016/j.copbio.2014.12.002>
- Olsén, A., Arnqvist, A., Hammar, M., Sukupolvi, S., & Normark, S. (1993). The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Molecular Microbiology*, 7(4), 523–36. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8459772>
- Paul, K., Nieto, V., Carlquist, W. C., Blair, D. F., & Harshey, R. M. (2010). The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a “backstop brake” mechanism. *Molecular Cell*, 38(1), 128–39. <http://doi.org/10.1016/j.molcel.2010.03.001>
- Plumbridge, J. (2001). Regulation of PTS gene expression by the homologous transcriptional regulators, Mlc and NagC, in *Escherichia coli* (or how two similar repressors can behave differently). *Journal of Molecular Microbiology and Biotechnology*, 3, 371–380.
- Pratt, L. A., & Silhavy, T. J. (1998). Crl stimulates RpoS activity during stationary phase. *Molecular Microbiology*, 29(5), 1225–36. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9767590>
- Reshamwala, S. M. S., & Noronha, S. B. (2011). BioWlm formation in *Escherichia coli* cra mutants is impaired due to down-regulation of curli biosynthesis. *Archives of Microbiology*, 193(10), 711–722. <http://doi.org/10.1007/s00203-011-0708-7>
- Römling, U., & Amikam, D. (2006). Cyclic di-GMP as a second messenger. *Current Opinion in Microbiology*, 9(2), 218–28. <http://doi.org/10.1016/j.mib.2006.02.010>
- Römling, U., Bian, Z., Hammar, M., Sierralta, W. D., & Normark, S. (1998). Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. *Journal of Bacteriology*, 180(3), 722–731.

- Römling, U., Sierralta, W. D., Eriksson, K., & Normark, S. (1998). Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD* promoter. *Molecular Microbiology*, *28*(2), 249–64. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9622351>
- Salama, N. R., Shepherd, B., & Falkow, S. (2004). Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. *Journal of Bacteriology*, *186*(23), 7926–35. <http://doi.org/10.1128/JB.186.23.7926-7935.2004>
- Sanchez-Torres, V., Maeda, T., & Wood, T. K. (2010). Global regulator H-NS and lipoprotein NlpI influence production of extracellular DNA in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, *401*(2), 197–202. <http://doi.org/10.1016/j.bbrc.2010.09.026>
- Shibata, S., Yip, E. S., Quirke, K. P., Ondrey, J. M., & Visick, K. L. (2012). Roles of the structural symbiosis polysaccharide (*syp*) genes in host colonization, biofilm formation, and polysaccharide biosynthesis in *Vibrio fischeri*. *Journal of Bacteriology*, *194*(24), 6736–47. <http://doi.org/10.1128/JB.00707-12>
- Shimada, T., Katayama, Y., Kawakita, S., Ogasawara, H., Nakano, M., Yamamoto, K., & Ishihama, A. (2012). A novel regulator RcdA of the *csgD* gene encoding the master regulator of biofilm formation in *Escherichia coli*. *Microbiologyopen*, *1*(4), 381–394. <http://doi.org/10.1002/mbo3.42>
- Sledjeski, D. D., Gupta, A., & Gottesman, S. (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *The EMBO Journal*, *15*(15), 3993–4000. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8670904>
- Smith, D. R., Price, J. E., Burby, P. E., Blanco, L. P., Chamberlain, J., & Chapman, M. R. (2017). The production of curli amyloid fibers is deeply integrated into the biology of *Escherichia coli*. *Biomolecules*, *7*(4). <http://doi.org/10.3390/biom7040075>
- Stevenson, G., Andrianopoulos, K., Hobbs, M., & Reeves, P. R. (1996). Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *Journal of Bacteriology*, *178*(16), 4885–4893.
- Takeuchi, R., Tamura, T., Nakayashiki, T., Tanaka, Y., Muto, A., Wanner, B. L., & Mori, H. (2014). Colony-live--a high-throughput method for measuring microbial colony growth kinetics--reveals diverse growth effects of gene knockouts in *Escherichia coli*. *BMC Microbiology*, *14*, 171. <http://doi.org/10.1186/1471-2180-14-171>
- Tamae, C., Liu, A., Kim, K., Sitz, D., Hong, J., Becket, E., ... Miller, J. H. (2008). Determination of antibiotic hypersensitivity among 4,000 single-gene-knockout mutants of *Escherichia coli*. *Journal of Bacteriology*, *190*(17), 5981–8. <http://doi.org/10.1128/JB.01982-07>
- Tetz, V. V., & Tetz, G. V. (2010). Effect of extracellular DNA destruction by DNase I on characteristics of forming biofilms. *DNA and Cell Biology*, *29*(8), 399–405. <http://doi.org/10.1089/dna.2009.1011>
- Thomason, M. K., Fontaine, F., De Lay, N., & Storz, G. (2012). A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Molecular Microbiology*, *84*(1), 17–35. <http://doi.org/10.1111/j.1365-2958.2012.07965.x>
- Tong, A. H., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Pagé, N., ... Boone, C. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science (New York, N.Y.)*, *294*(5550), 2364–8. <http://doi.org/10.1126/science.1065810>
- Vianney, A., Jubelin, G., Renault, S., Dorel, C., Lejeune, P., & Lazzaroni, J. C. (2005). *Escherichia coli* *tol* and *rcs* genes participate in the complex network affecting curli synthesis. *Microbiology*, *151*(7), 2487–2497. <http://doi.org/10.1099/mic.0.27913-0>

- Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M., & Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: Involvement of a new *ompR* allele that increases curli expression. *Journal of Bacteriology*, *180*(9), 2442–2449. <http://doi.org/10.1186/gb-2008-9-4-r74>
- Wang, X., Smith, D. R., Jones, J. W., & Chapman, M. R. (2007). In vitro polymerization of a functional *Escherichia coli* amyloid protein. *The Journal of Biological Chemistry*, *282*(6), 3713–9. <http://doi.org/10.1074/jbc.M609228200>
- Wang, X., Zhou, Y., Ren, J.-J., Hammer, N. D., & Chapman, M. R. (2010). Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 163–8. <http://doi.org/10.1073/pnas.0908714107>
- Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., ... Davis, R. W. (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science (New York, N.Y.)*, *285*(5429), 901–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10436161>
- Wozniak, D. J., Wyckoff, T. J. O., Starkey, M., Keyser, R., Azadi, P., O'Toole, G. A., & Parsek, M. R. (2003). Alginate is not a significant component of the extracellular polysaccharide matrix of PA14 and PAO1 *Pseudomonas aeruginosa* biofilms. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 7907–12. <http://doi.org/10.1073/pnas.1231792100>
- Xu, D., Jiang, B., Ketela, T., Lemieux, S., Veillette, K., Martel, N., ... Roemer, T. (2007). Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in *Candida albicans*. *PLoS Pathogens*, *3*(6), e92. <http://doi.org/10.1371/journal.ppat.0030092>
- Yanisch-Perron, C., Vieira, J., & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, *33*(1), 103–19. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2985470>
- Zhang, X. S., Garcia-Contreras, R., & Wood, T. K. (2007). YcfR (BhsA) influences *Escherichia coli* biofilm formation through stress response and surface hydrophobicity. *Journal of Bacteriology*, *189*(8), 3051–3062. <http://doi.org/10.1128/JB.01832-06>
- Zheng, D., Constantinidou, C., Hobman, J. L., & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research*, *32*(19), 5874–5893. <http://doi.org/10.1093/nar/gkh908>