

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 29 日現在

機関番号：82401

研究種目：若手研究(B)

研究期間：2011～2013

課題番号：23700353

研究課題名(和文) 遺伝子機能の環境依存性に着目した代謝ネットワークのロバスト性機構の解明

研究課題名(英文) Analysis of robustness of metabolic networks focusing on environmental dependency of gene knockouts

研究代表者

遠里 由佳子 (Tohsato, Yukako)

独立行政法人理化学研究所・生命システム研究センター・研究員

研究者番号：80346171

交付決定額(研究期間全体)：(直接経費) 2,200,000円、(間接経費) 660,000円

研究成果の概要(和文)：大腸菌の遺伝子機能を解析するためには、その遺伝子が欠失したときの表現型や、遺伝子と表現型がどのようなネットワークにより制御されているかを知ることが重要である。しかし、従来の表現型解析はグルコースやLBなどの限られた培地条件でしかおこなわれていない。そこで(1) Phenotype MicroArray で測定された1920種類の培地環境における大腸菌の野生株と一遺伝子欠失株の細胞増殖の違いを統計的に解析し、(2)大規模な中央代謝の数理モデルの構築することで、培地条件によって遺伝子の必要性が変化する「遺伝子の環境依存性」の解析を試みた。

研究成果の概要(英文)：Systematic studies of genotype-phenotype correlation have revealed that most single-knockout mutants often show no observable phenotype during growth under laboratory conditions. However, such studies have mainly examined phenotypes under a limited number of conditions such as Glucose and LB mediums. Genes showing no phenotypes when absent under one condition may exhibit phenotypic changes under different specific conditions, if the gene function depends on environmental conditions. To elucidate the environmental dependency of genes, we analyzed growth data of wild-type and single-knockout mutant in Escherichia coli in 1920 medium conditions obtained from the Phenotype MicroArray. Furthermore, we develop a kinetic model for the central metabolism of E. coli that includes not only the glucose transport system, but also the glycerol transport system.

研究分野：総合領域

科研費の分科・細目：情報学、生体生命情報学

キーワード：バイオインフォマティクス マイクロアレイ 表現型 数理モデル 統計解析

1. 研究開始当初の背景

ATPやNADHなどのエネルギー分子を生産する基本的な役割を果たす遺伝子の多くは、大腸菌やヒトを含むあらゆる生物で共通している。そして、エネルギー代謝に関わる重要なプロセスを担っている遺伝子を欠失させても、細胞の生存を保つ頑強性を持つ。しかし、遺伝子欠失の影響の検証は、これまでグルコースや LB 培地など限られた培地条件でのみ実施され、栄養源によって遺伝子の必要性が変化する「遺伝子の環境依存性」は十分な検討は行われていなかった。

2. 研究の目的

そこで本研究では (1)Phenotype MicroArray (以下 PM) で測定された 1920 種類の培地環境における大腸菌の 300 の 1 遺伝子欠失株と野生株の細胞増殖過程を統計的に解析し、(2)複数の培地を考慮した大腸菌のエネルギー代謝の数値モデルの構築することで、遺伝子の環境依存性と代謝ネットワークの制御機構の解明を試みた。

3. 研究の方法

(1) 統計解析では、PM データからノイズを除去するために、培地別に観測値の分布を比較する多重解析法を提案する。そして、ノイズ除去した後に得られるベクトルデータに、2006 年に Kemp らによって提案された無限関係モデル (Infinite Relational Model : IRM) とよばれるバイクラスタリングのアルゴリズムを適用することで、欠失株と培地の特徴抽出を行う。IRM のクラスタリングの性能評価は、ソフトウェア BicAT によって公開されている CC と ISA, OPSM, R のライブラリ biclust で公開されている Spectral との比較により行う。評価基準には遺伝子の機能分類 (Gene Ontology: GO) とクラスタ間におけるデータの重複に着目する。

(2) 数値モデルの構築では、Chassagnole らによって開発された大腸菌の解糖系のモデルに、Kadir らにより提案された TCA 回路を含め、さらに、グルコース培地および、グルコースおよびグリセロールの取り込み系を導入することで、中央代謝モデル (解糖系、ペントースリン酸経路、TCA 回路、酢酸代謝、糖新生経路、エントナー・ドッドロフ経路、芳香族アミノ酸経路) を実現する。さらに、実数型遺伝子的アルゴリズムを用いた数値モデルのパラメータ最適化や感度解析を行う。そして、最適化されたパラメータを用いたシミュレーション結果を、Martinez-Gomez らにより測定された大腸菌 (*Escherichia coli* K-12 strain JM101) を好気条件のグルコース培地またはグリセロール培地で pH 7 を維持し回分培養を行ったときの実験結果

で検証する。

4. 研究成果

(1) PM データの統計解析

IRM を PM データに適用した結果、300 種の欠失株は 19 個のクラスタ (遺伝子クラスタ) に分割され、1920 培地条件うちの 1199 培地条件は 25 個のクラスタ (培地クラスタ) に分割された (図 1)。そして図で A から I でラベル付けされた 9 クラスタにおいて、クラスタに含まれる欠失した遺伝子に機能の偏り ($P < 0.05$) が見られた。

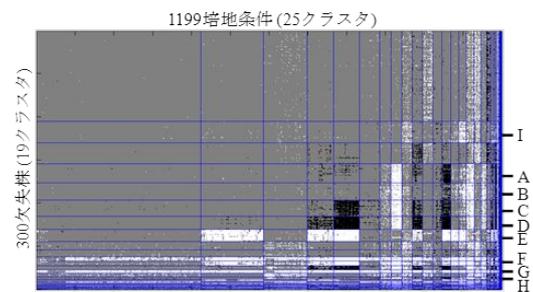


図 1: PM データの IRM によるクラスタリング。縦軸には欠失株を、横軸には 1199 の培地条件を配置し、野生株と比べて欠失株が増加した場合は黒で、減少した場合は白で、それ以外は灰色で示した。

中でもクラスタ I は、25 種の欠失株を含み、アミノ酸を含む培地の一部で NADH 生産量を減少させる傾向が示された。この欠失株クラスタには、ホスファターゼとして化合物に対して脱リン酸化を行う機能を持つタンパク質をコードする欠失遺伝子が集中していた。特に、*AyihX*、*AyfbT*、*AyniC*、*AyqaB*、*AybhA*、*AybjI*、*AyidA*、*AyieH*、*AyjjG* の 9 種の欠失株は、機能未知遺伝子にコードされるタンパク質に対して酵素活性を検証した Kuznetsova らの報告において、いずれも中央代謝経路で触媒される糖類を脱リン酸化する機能が *in vitro* で確かめられた遺伝子の欠失株であった。

IRM のクラスタリング性能を評価した結果、GO の偏りが $P < 0.05$ となる GO が検出された欠失株クラスタの数は IRM が 8 個と最も多く、 $P < 0.001$ では OPSM が最も多かった (図 2)。

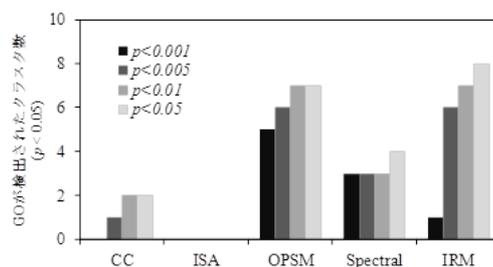


図 2: 得られるクラスタ数の比較

しかし、OPSM は、クラスタ間に重複を許すアルゴリズムであるため、その抽出される遺伝子の重複を Simpson 指数の平均で確認したところ、OPSM は多くのクラスタを検出できるが、機能的には似たものを繰り返し抽出する傾向があることが明らかになった。

(2) 数理モデルの構築

実数型遺伝的アルゴリズムを用いて、パラメータ最適化と感度解析を行うシステムを構築し、モデル中の酵素最大速度 (V_{max}) に着目し、十分な最適化能を確認した。最適化により得られた V_{max} の値を 0 倍から 2 倍まで 0.2 倍ずつ変動させ、変動前後の各代謝物質濃度を観測し、Chassagnole らによる解糖系の数理モデルにおける、大腸菌の増殖において重要な酵素反応と重要でない酵素反応をコードする遺伝子を明らかにした。さらに、影響の大きな酵素反応を補てんできる酵素反応の組み合わせを明らかにした。

で得られた知見を用いて、グルコース培地とグリセロール培地を想定した大規模なモデルを実現した (図 3)。

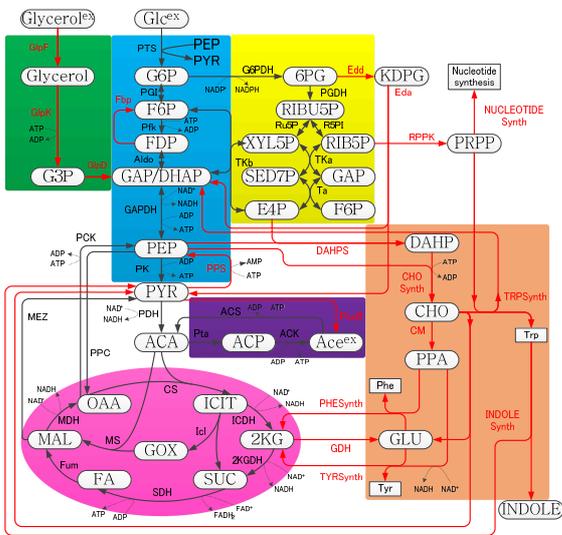


図 3: 構築したグルコースおよびグリセロール代謝モデル。楕円は代謝物質を、実線矢印は酵素反応を示す。Kadir モデルと異なる部分は赤矢印で示す。

グリセロール代謝に着目した数理モデルに Wang らの報告がある。しかし、解糖系および簡略化された TCA 回路のみとその規模は小さく、グリセロールの取り込み系および解糖系や TCA 回路を含む中央代謝全体を含めた大規模なモデル構築の報告はこれまでなかった。

最適化したパラメータを用いたシミュレーション結果は、それぞれの培地条件での炭素源濃度およびグリセロール培地における細胞濃度は実験データと良く一致した (図 4)。

しかし、グリセロール培地におけるインドール濃度の再現に問題があった、これは今回のモデルではインドール生産反応のみを定義しており、インドール濃度が減少しないためと考えられた。

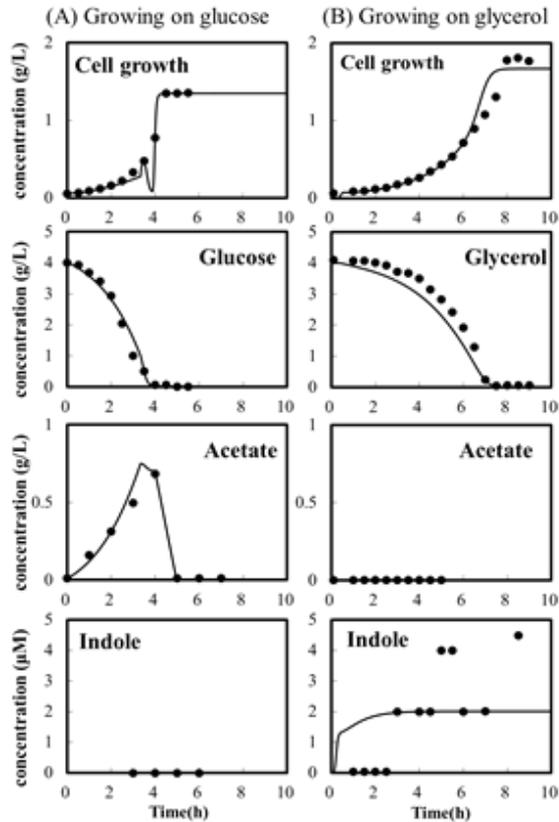


図 4: シミュレーション結果 (実線) と実験結果 (破線)

グルコース培地およびグリセロール培地における中央代謝に含まれる経路の大腸菌増殖への寄与率の時間変化を確認した (図 5)。

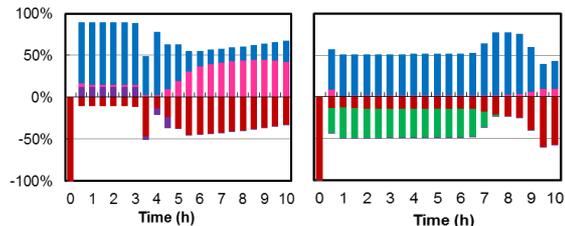


図 5: グルコース培地 (左) とグリセロール培地 (右) における大腸菌増殖への代謝経路の寄与率の時間変化の違い。各色は図 3 の経路に対応している。

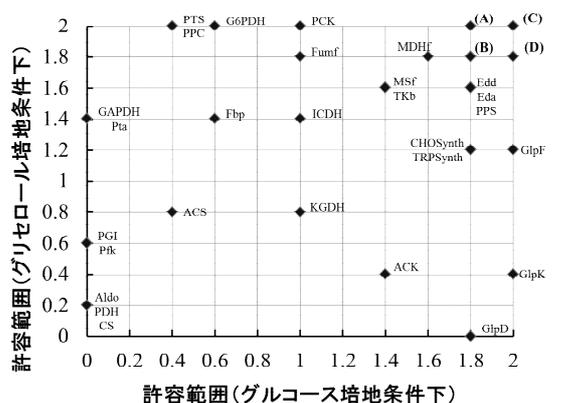
グルコース培地における大腸菌の対数増殖期 (0-4.5h) の細胞増殖には解糖系が最も寄与 (0.5-5.5h) していた。また、TCA 回路は 5h 以降の大腸菌が増殖した後の培養後期に活性化が確認できた。一方、補充反応は負に寄与し、増殖を抑制していた。一方、グリセロール培地においても解糖系が正に寄与していたが、グリセロールのリン酸化反応

(GlpK)は大きく負に影響していた。グリセロールが培地に存在する 0-7h では、GlpK により増殖が抑制されているが、グリセロールが枯渇し、GlpK の反応速度が低下すると大腸菌が解糖系由来のエネルギー分子により増殖すること考えられる。

パラメータ感度解析により V_{max} の影響の違いを各培地条件別に算出し、各 V_{max} のグルコース培地条件とグリセロール培地条件で比較した(図6)。

図6: 培地条件での重要性な酵素反応の違い

その結果、グルコース培地条件下で許容範囲



(A): SDHf SDHr	(B): Ru5P R5PI DAHPS Fumr CM MDHr RPPK	(C): PK PGDH PoxB TA PHESynth Iclf TYRSynth MSr	(D): Tka NUCLEOTIDESynth
-------------------	---	---	-----------------------------

の狭い V_{max} が含まれる反応は、解糖系の PGI、Pfk、Aldo、GAPDH および PDH と ACA からの分岐反応である TCA 回路の CS、酢酸代謝の PTA であった。これらに次いで、グルコース取り込み反応の PTS、PEP から OAA を合成する PPC、酢酸を消費する ACS の影響が大きい。以上からグルコース培地条件下では、グルコースを取り込み、解糖系で代謝した後、PDH を用いた ACA を生産する一連の反応と、ACA を消費する反応の影響が大きいことが明らかになった。しかし解糖系の反応で、PEP からの PYR を生産する PK は他の解糖系の反応と比べ影響が小さい。これはグルコース存在下で、PEP からの PYR の生産が PK よりも PTS で行われるためと考えられる。

5. 主な発表論文等

[雑誌論文](計3件)

Yukako Tohsato, Kunihiko Ikuta, Akitaka Shionoya, Yusaku Mazaki, Masahiro Ito, Parameter optimization and sensitivity analysis for large kinetic models using a real-coded genetic algorithm, *Genes (Suppl. for Genome Informatics 2012)*, 査読有, Vol. 518, No. 1, pp. 84-90, 2013, Doi: 10.1016/j.gene.2012.11.080.
Yoshihiko Matsuta, Masahiro Ito,

Yukako Tohsato, ECOH: An Enzyme Commission number predictor using mutual information and a support vector machine, 査読有, *Bioinformatics*, Vol. 29, No. 3, pp. 365-372, 2012, Doi: 10.1093/bioinformatics/bts700.

杉澤仁, 福田一, 森浩禎, 谷口忠大, 伊藤将弘, 遠里由佳子, 無限関係モデルを用いた大腸菌の表現型マイクロアレイ解析, 査読無, Vol. 71, pp. 49-58, 2012, <http://ci.nii.ac.jp/naid/110009595156>.

[学会発表](計4件)

Akitaka Shionoya, Kunihiko Ikuta, Shuichi Onami, Masahiro Ito, Yukako Tohsato, Large-Scale Kinetic Modeling for the Central Metabolism in *Escherichia coli* Grown on Glucose or Glycerol, *Genome Informatics*, Singapore, Biopolis, 2013/12/16.

Hitoshi Sugisawa, Hajime Fukuda, Hirotada Mori, Masahiro Ito, Tadahiro Taniguchi, Yukako Tohsato, Phenotype microarray analysis of *Escherichia coli* using infinite relational model, *Genome Informatics*, China, Hangzhou, 2012/12/12.

Akitaka Shionoya, Kunihiko Ikuta, Hitomi Dose, Hirotada Mori, Masahiro Ito, Yukako Tohsato, Modeling of central metabolism with regulation of *rpoS* gene in *Escherichia coli*, *Genome Informatics*, China, Hangzhou, 2012/12/12.

Kunihiko Ikuta, Akitaka Shionoya, Masahiro Ito, Yukako Tohsato, Parameter optimization and correlation analysis for a large kinetic model using a real-coded genetic algorithm, *Genome Informatics*, China, Hangzhou, 2012/12/12.

[図書](計1件)

Yukako Tohsato, Natsuko Yamamoto, Toru Nakayshiki, Rikiya Takeuchi, Barry L. Wanner, Hirotada Mori, Daisuke Kihara (Ed.), Protein Function Prediction for Omics Era: 15. Towards elucidation of the *Escherichia coli* K-12 unknownome, pp.289-306, Springer, 2011.

6. 研究組織

(1) 研究代表者

遠里 由佳子 (Yukako Tohsato)

独立行政法人理化学研究所・生命システムセンター・研究員

研究者番号: 80346171