



Title	微生物育種に向けた多重遺伝子破壊の探索手法の開発
Author(s)	大野, 聡
Citation	大阪大学, 2015, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/52028
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

微生物育種に向けた多重遺伝子破壊の
探索手法の開発

提出先 大阪大学大学院情報科学研究科

提出年月 2015 年 1 月

大野 聡

学位取得に関わる発表論文

学術雑誌

- [1] Satoshi Ohno, Hiroshi Shimizu, Chikara Furusawa. “FastPros: Screening of reaction knockout strategies for metabolic engineering”. *Bioinformatics*, **30**(7), 981-987. 2014. (博士論文第 3 章)
- [2] Satoshi Ohno, Chikara Furusawa, Hiroshi Shimizu. ”*In silico* screening of triple reaction knockout *Escherichia coli* strains for overproduction of useful metabolites”. *Journal of Bioscience and Bioengineering*, **115**(2), 221-228. 2013. (博士論文第 2 章)

国際会議

- [1] Satoshi Ohno, Chikara Furusawa, Hiroshi Shimizu. “FastPros: Screening method of multiple gene knockout for microbial production using genome-scale metabolic model”. Metabolic Engineering IX, poster #66, Le Bellevue Congress and Exhibition Centre in Biarritz, Biarritz, France. June 3-8, 2012. (博士論文第 3 章)

内容梗概

微生物による有用化合物生産は、現代社会において欠かせないものである。古くからのアルコール飲料、アミノ酸、抗生物質などに加えて、近年ではバイオ燃料やプラスチック原料などの生産にも、微生物を利用した生産プロセスが注目を集めている。そのため、ターゲットとする有用化合物を高効率に生産可能な微生物を、合理的かつ効率的に育種する手法の開発が至上命題となっている。従来より、微生物の代謝をターゲット生産に向けて最適化するため、宿主微生物への異種生物遺伝子の導入によるターゲット生合成経路の強化や、遺伝子破壊による不要な代謝経路の不活性化により、代謝フラックス分布の改変が行われてきた。遺伝子破壊に関しては、近年発展した代謝シミュレーション手法を利用し、ターゲット生産性向上に有望な破壊候補遺伝子を計算機上で探索する手法が開発されてきた。しかし、同時に破壊する遺伝子数が増加すると、その組み合わせが指数関数的に増加することが課題となっている。本研究では、ゲノムスケールの代謝モデルについて、多重破壊の大域探索手法および局所探索手法の開発を目指した。

学位論文は、第1章から第4章で構成される。以下に、各章の内容を示す。第1章では、研究の背景および目的について記述した。代謝シミュレーション手法としては、**Flux balance analysis (FBA)** と呼ばれる手法を使用した。FBAは、微生物のゲノム情報や栄養条件を入力とし、各代謝反応の反応速度（代謝フラックス）を出力とするシミュレーション手法である。これは、代謝の定常を満たす条件で、細胞の増殖速度を最大化する線形最適化問題として設定できる。代謝モデルに含まれる1000以上の代謝反応について代謝フラックスのシミュレーションが可能であり、広く利用されている。

第2章では、ターゲット化合物を高生産可能な多重遺伝子破壊を大域的に探索する手法を開発した。解析したい栄養条件に応じて、縮約した代謝モデルを自動的に構築する方法を提案し、破壊対象の遺伝子の数を削減することで探索を高速化した。例えば、1260の遺伝子からなる大腸菌の代謝モデルでは、グルコース炭素源好気条件下での破壊対象の遺伝子は393まで減少し、3重破壊の探索時間は 10^7 秒から 10^2 秒に短縮された。開発手法を、大腸菌による1-ブタノールや1,3-プロパンジオールなどの生産に対して適用し、高い生産収率が予測される3重破壊を新たに複数見つけ出すことに成功した。

第3章では、さらなる多重遺伝子破壊を局所的に探索する手法 **FastPros (Fast Algorithm of knockout Screening based on Target PROduction using Shadow price analysis)** を開発した。ターゲット生産速度と増殖速度との関連性を、線形計画問題におけるシャドウプライスを用いて評価し、これを指標とした局所的探索手法 **FastPros** を開発した。FastProsの性能を評価するため、大腸菌細胞内の625代謝物質それぞれの生産に有望な破壊候補遺伝子を、FastProsにより探索した。その結果、生産収率を指標とした従来の局所的探索手法より多い472の代謝物質について、破壊候補遺伝子が同定され、大域的探索では適用が難しい6遺伝子以上の同時破壊も予測が可能となった。

第 4 章では、まとめと展望を述べた。開発した探索手法は、ゲノムの解読された微生物種に対して広く利用可能であり、微生物育種において破壊候補遺伝子を探索する汎用的な手法となることが期待される。

目次

第1章	序論	1
1.1	微生物を用いた有用化合物生産	1
1.2	ゲノムスケール代謝モデルを用いた代謝シミュレーション	2
1.3	微生物育種に向けた遺伝子破壊の <i>in silico</i> 探索と課題	4
1.4	本研究の目的および論文の構成	6
第2章	育種に向けた多重遺伝子破壊の大域的探索	9
2.1	緒言	9
2.2	手法	11
2.2.1	代謝モデル	11
2.2.2	FBA	14
2.2.3	代謝モデルの縮約	15
2.2.4	破壊により代謝が変動しない遺伝子の特定	18
2.2.5	大域的探索手法	18
2.3	結果	20
2.3.1	計算時間	20
2.3.2	栄養条件に応じた縮約モデルの違い	22
2.3.3	大腸菌による 1-ブタノール生産への本手法の適用	27
2.3.4	大腸菌による 1-プロパノール生産への本手法の適用	32
2.3.5	大腸菌による 1,3-プロパンジオール (13PD) 生産への本手法の適用	37
2.4	考察	41
2.5	結言	43

第3章	育種に向けた多重遺伝子破壊の局所的探索 (FastPros)	44
3.1	緒言	44
3.2	手法	46
3.2.1	代謝モデル	46
3.2.2	u_{TARGET}	47
3.2.3	u_{TARGET} の計算	49
3.2.4	FastPros	51
3.2.5	FastPros の結果に基づく OptKnock	55
3.2.6	破壊候補遺伝子の類似度に基づくクラスタリング	56
3.3	結果	57
3.3.1	FastPros によるターゲット化合物生産に向けた遺伝子破壊の探索	57
3.3.2	既存の探索手法との比較	60
3.3.3	u_{TARGET} を利用した探索における適応度地形	63
3.3.4	FastPros の結果に基づく OptKnock によるターゲット生産性の向上	65
3.3.5	工学的に有用な代謝物質生産に向けて見つけ出された多重遺伝子破壊	67
3.3.6	複数の代謝物質生産に共通して予測された破壊候補遺伝子の解析	72
3.4	考察	74
3.5	結言	75
第4章	結論	76
4.1	結果のまとめ	76
4.2	得られた知見の寄与するところ	78
4.3	今後の展望	79
参考文献		82
付録		91
謝辞		96

第1章 序論

1.1 微生物を用いた有用化合物生産

微生物を用いた有用化合物生産は、人類社会に欠かせないものである。古くより、アルコール・発酵食品・アミノ酸・抗生物質などが微生物を用いて生産されてきた (Kourkoutas *et al.*, 2004; Motoki and Seguro, 1998)。例えば、抗生物質であるペニシリンは、20 世紀前半にアオカビより産生されることが発見され、第二次世界大戦以降の感染症の治療に大きく貢献した (Manring *et al.*, 2009)。また、うま味の成分であるグルタミン酸は、もともとは大豆や小麦などを加水分解する方法で生産されていたが、現在では微生物発酵を用いる方法が主流となっている (Sano, 2009)。近年では、植物由来のバイオマスから燃料や化学品を生産するプロセスが注目を集めており、カーボンニュートラルな社会の実現が目指されている。ブラジルでは、バイオエタノールを燃料として利用可能な自動車が、国産自動車販売の半分以上を占めている (株式会社資源・食糧問題研究所、2012 年)。米国でも、DuPont Tate & Lyle BioProducts 社により、微生物を用いた 1,3-プロパンジオールの商業生産が、2007 年より行われている (U.S.A.: World's first propanediol production from corn sugar opened, Aug. 2007, Bio-based-News)。

微生物に有用物質を生産させるプロセスで重要なことの一つに、ターゲット化合物の生産能や生産物ストレスなどに対する耐性の高い菌株の育種が挙げられる。有用物質を生産させる際には、冷凍保存した菌株の一部を生育させ、そして有益な形質が安定して観察される範囲で培養することが一般的である。産業的な生産プロセスに利用可能な微生物の育種では、古くより、紫外線や変異誘導物質でゲノムにランダムな変異を導入し、得られた多数の変異株より優れた形質を持つもののスクリーニングが行われてきた (Rowlands, 1984)。このランダムな変異とスクリーニングによる育種は、生物学的知見が乏しい微生物に対しても適用可能である。現在までに、エタノール耐性株 (Ogawa *et al.*, 2000) やアミノ酸生産株 (Demain, 2000) など、数多くの有用微生物株が育種されてきた。しかし、この方法では、細胞にとって有害な変異も蓄積し (Ohnishi *et al.*, 2005)、また目的とする形質に影響した変異の特定が難しいため、新たな育種に活かし難いといった弱みがあった。遺伝子組換え技術 (Cohen *et al.*, 1973) などの分子生物学的な実験技術が発達すると、目的の遺伝子を人為的に操作し、有用な表現型を示す株を獲得することが可能となった。1990 年代には、「組換え DNA 技術を用いた細胞内の特定の反応の改変や新しい反応の導入により、指定された目的物質生産や菌体の特性の改変を目指す」学問である、代謝工学が発展した (Stephanopoulos *et al.*, 1998)。炭素の安定同位体を用いた代謝フラックス解析手法が登場し、遺伝子操作が細胞内代謝に及ぼす影響を、各代謝反応の単位菌体量あたりの反応速度 (代謝フラックス) から解析することが可能になった (Zamboni *et al.*, 2009)。

代謝工学的に微生物を育種する場合には、細胞内の代謝フラックス分布を、ターゲット

とする有用化合物生産に向けて最適化することが重要である。そのために、ターゲット化合物の生合成経路を活性化し、不要な代謝経路を不活性化させることが一般に行われる (Zaldivar *et al.*, 2001; Becker *et al.*, 2011; Fisher *et al.*, 2014)。生合成経路の強化では、異種生物種由来の酵素の発現 (Ro *et al.*, 2006)、アミノ酸置換による酵素の改良 (Leonard *et al.*, 2010)、フィードバック阻害の回避 (Krämer *et al.*, 2003) などが行われている。不要な代謝経路の不活性化では、対応する遺伝子を欠損させ、副生産物の生産を阻害し、細胞内の補酵素のバランスが調整されている (Atsumi *et al.*, 2008; Jain *et al.*, 2012)。遺伝子操作により代謝を改変した株が獲得されると、その株の培養結果から、ターゲット化合物の生産収率、培養液中の濃度、生産速度等が評価される。また、代謝フラックス解析により培養中の細胞の代謝状態が推定され、得られた知見より代謝改変戦略を立てて、生産性向上に向けた遺伝子操作等が繰り返される (Becker *et al.*, 2011)。しかし、これらの遺伝子操作は、蓄積されてきた知見に基づいてはいるものの、経験や勘で行う場合も多く、戦略通りの表現型を持つ株の獲得には多くの試行錯誤を要していた。

近年、計算機を利用し、代謝改変に必要な遺伝子操作を合理的に検討する技術が発展してきた。例えば、異種生物由来の酵素を利用して生合成経路を強化する際に、どの生物のどの酵素を発現させればよいかを、データベースより探索する手法が開発されている (Tomar and De, 2013; Zomorodi *et al.*, 2012)。また、酵素と基質のドッキングシミュレーションにより、基質を特異的に認識するアミノ酸残基が同定され、新規化合物生産に有益な知見が提供されている (Korman *et al.*, 2004; Torres *et al.*, 2013)。不要な代謝経路の不活性化では、後述するような代謝シミュレーション手法により、破壊候補遺伝子を計算機上で探索する手法が開発されてきた。微生物が持つ数千の遺伝子の中から、望まれる代謝フラックス分布を実現する可能性の高い破壊候補遺伝子が同定されている。本学位論文では、このターゲット生産を目指した破壊候補遺伝子の探索に着目し、研究を遂行した。

1.2 ゲノムスケール代謝モデルを用いた代謝シミュレーション

DNA の塩基配列のシーケンシング技術とアノテーション技術の発展により、生物の代謝反応の化学量論式をゲノムワイドに記述した数学的なモデル (ゲノムスケール代謝モデル) が、構築可能になった (Feist and Palsson, 2008; Oberhardt *et al.*, 2009)。このモデルと、フラックスバランス解析 (FBA) と呼ばれるシミュレーション手法により、細胞内のそれぞれの代謝反応がどの程度のフラックスをとるか (代謝フラックス分布) をシミュレートすることができる (Orth *et al.*, 2010; Varma and Palsson, 1994)。FBA による代謝シミュレーション手法の最大の特徴は、他の速度論的なシミュレーション手法と比較し、代謝中間体濃度、酵素量、速度論的パラメータなどを必要とせず、化学量論とフラックスにのみ着目している点である。KEGG (Kyoto Encyclopedia of Genes and Genomes) や MetaCyc (Encyclopedia of

Metabolic Pathways) などのデータベースより、解析したい生物の持つ代謝反応の化学量論式を取得し、文献や実験より細胞を構成するアミノ酸、ヌクレオチドなどの割合を獲得すれば、代謝モデルの構築が可能となる。1999年に *Haemophilus influenzae* の代謝モデルが構築されて以来 (Edwards, 1999)、大腸菌 (Reed *et al.*, 2003; Orth *et al.*, 2011) や酵母 (Duarte *et al.*, 2004) を始めとする様々な生物種について、ゲノムスケール代謝モデルが構築されてきた (Monk *et al.*, 2014)。近年では、モデル構築の自動化が進展し (Henry *et al.*, 2010)、また代謝モデルを容易に解析するための Toolbox も公開され (Becker *et al.*, 2007; Schellenberger *et al.*, 2011)、今後もさらに多くの生物種について代謝モデルが構築されることが予想される。構築された代謝モデルは、代謝工学、表現型の予測、代謝ネットワーク解析など、幅広く利用されている (Feist and Palsson, 2008; McCloskey *et al.*, 2013)。

FBA は、化学量論式が記述された代謝モデルより、代謝フラックス分布や細胞の増殖速度等が計算されるシミュレーション手法である (Fig. 1.1)。この手法は、代謝モデルに含まれる反応のフラックスを変数とし、ある制約条件下で目的関数を最大化するという線形計画問題として記述される。

FBA における主な制約条件は、代謝の定常の仮定である。定常とは、細胞内の代謝フラックス分布や代謝物質濃度等が時間変化しないことを意味している。この仮定より、それぞれの代謝物質について生産フラックスの総和と消費フラックスの総和が等しいという制約式が立てられる。これは、代謝フラックスの計算を簡略化するために仮定されたものであり、回分培養における対数増殖期や、連続培養での細胞の状態に相当する。

目的関数としては、一般に細胞の増殖速度が採用されており、微生物は与えられた栄養条件において、最大の増殖速度をとるという仮定に基づくものである。代謝モデル中での細胞の増殖は、「アミノ酸やヌクレオチドなどの菌体構成成分が一定の比率で消費され、新しい細胞 1 グラムが合成される」という擬似的な化学反応で記述されている。この擬似的な菌体合成反応のフラックスが、細胞の比増殖速度を表現している。

このように FBA では、代謝の定常を制約条件として、細胞の比増殖速度 (菌体合成反応のフラックス) を最大化する線形計画問題として、代謝フラックス分布が計算される。代謝モデルには、栄養条件や酸素、発酵産物の膜輸送も加味されているので、炭素源の違い、通気条件、代謝反応の有無による代謝フラックス分布や増殖速度の違いを解析することができる。例えば、グルタミン酸生産菌として知られる *Corynebacterium glutamicum* については、通気量の減少に伴い乳酸やコハク酸などの発酵産物の生産が増加する傾向が、実験と FBA による予測で一致することが報告されている (Shinfuku *et al.*, 2009)。また、大腸菌の遺伝子欠損と増殖能の関係性は、グルコースを単一炭素源とする最小培地において、実験値と予測値で 90% 以上一致している (Feist *et al.*, 2007)。

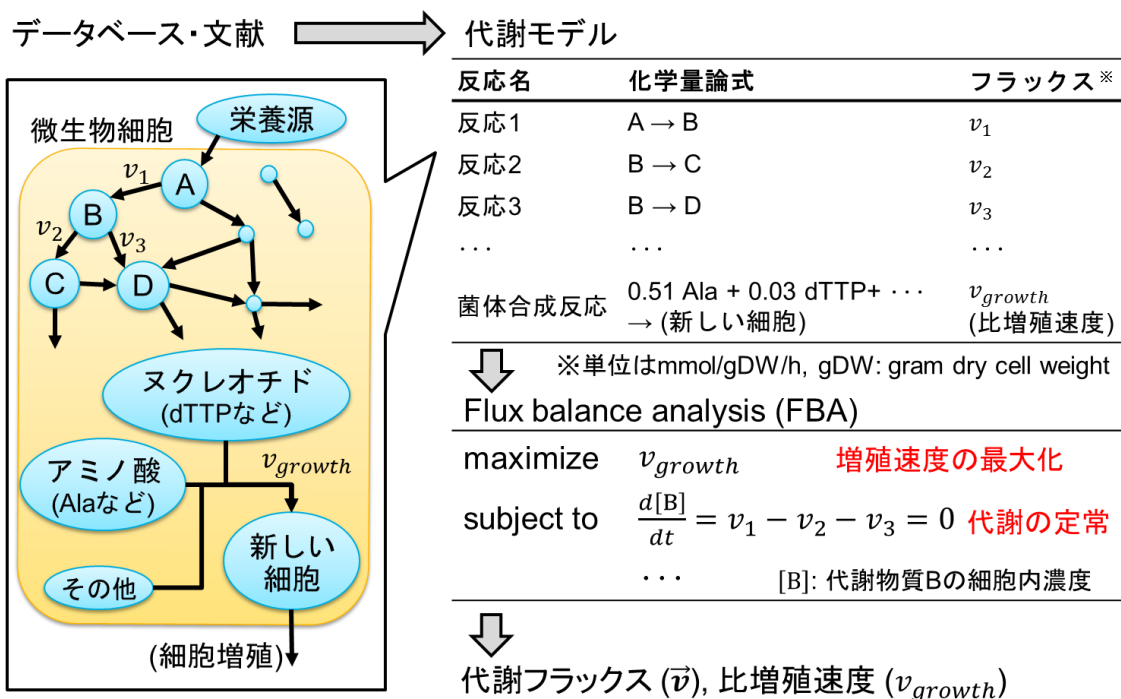


Fig. 1.1 ゲノムスケール代謝モデルと FBA による代謝シミュレーション

1.3 微生物育種に向けた遺伝子破壊の *in silico* 探索と課題

現在までに、ゲノムスケール代謝モデルと FBA を利用し、微生物育種に向けた遺伝子破壊を計算機上で (*in silico*) 探索する手法が提案されており、実験的に検証した例も多く報告されている (Lee *et al.*, 2005; Yim *et al.*, 2011; Santala *et al.*, 2011; Fong *et al.*, 2005; Ng *et al.*, 2012; Xu *et al.*, 2013)。これらの手法は、大きく分けて以下の2種類に分類できる。

一つ目の探索手法は、一定数以下の遺伝子破壊の全ての組み合わせの中で、大域的に最適な解を探索する手法である。この探索手法を、ここでは大域的探索手法と呼ぶ。例えば OptKnock という手法は、増殖速度を最大化する条件下で、ターゲット化合物の生産速度が最大となるような反応破壊 (実験的には対応する遺伝子の破壊) の組み合わせを大域的に探索する手法である (Burgard *et al.*, 2003)。OptKnock では、代謝フラックスを連続変数、ある反応を破壊するか否かを 0 か 1 しか取りえない 2 値変数で表現し、ターゲット化合物生産速度を目的関数とした最適化問題として記述される。OptKnock は広く利用されている探索手法であり、反応破壊だけでなく、遺伝子破壊に拡張したもの (Kim and Reed, 2010)、反応の活性化に拡張したもの (Pharkya and Maranas, 2006)、異種生物由来の反応経路の追加に拡張したもの (Pharkya *et al.*, 2004) などが提案されている。予測結果の実験的な検証もなされており、大腸菌に 1,4-ブタンジオールを生産させる際に有望な破壊候補遺伝子を OptKnock で予測し、それを実際の大腸菌株に対して適用した結果、1,4-ブタンジオールの

生産濃度が7倍に上昇したことが報告されている (Yim *et al.*, 2011)。

破壊候補遺伝子の大域的探索手法は、可能な遺伝子破壊の組み合わせについて、計算機上で取りこぼしなく評価できる利点を持つ。複数の遺伝子の同時破壊が化合物生産に影響する場合でも、大域的探索ではその破壊の組み合わせが評価される。しかし、大域的探索手法では、同時に破壊する遺伝子の数の増加に伴い、その組み合わせが指数関数的に増加するので、現実的な計算時間以内で探索可能な破壊数に限界があった。例えば、第2章で示すように、大腸菌のゲノムスケール代謝モデルに含まれる1260の破壊候補の中からは、一般的な計算機で探索可能な破壊数の上限は2遺伝子破壊だった。育種のための代謝フラックス分布の最適化では、しばしばそれ以上の遺伝子破壊が必要になるため (Jantama *et al.*, 2008; Lee *et al.*, 2007; Zhang *et al.*, 2011)、より高速で探索可能な手法の開発が求められている。

もう一方の探索手法は、局所的に最適な多重遺伝子破壊を探索する手法である。大域的探索と異なり、現実的な時間以内で出来る限り良質な多重遺伝子破壊を探索する手法であり、このような手法をここでは局所的探索手法と呼ぶ。例えば Alper らは、大腸菌によるリコペン生産性向上を目指し、まず大腸菌の代謝モデルに対して、1遺伝子を破壊した時の代謝フラックス分布を網羅的に計算した (Alper *et al.*, 2005)。最大のリコペン生産収率が予測される遺伝子破壊を特定したのちに、その遺伝子破壊に加えて、新たに1遺伝子を破壊し、リコペン収率が最大となる2重遺伝子破壊を予測した。同様に、3重遺伝子破壊も予測した。この単純な局所的探索により予測された3重遺伝子破壊を実際の大腸菌に対して適用したところ、リコペンの生産濃度を37%上昇させることに成功した。この探索は、同時に破壊する遺伝子数の増加に対して、計算時間は線形的に増加するため、計算時間が指数的に増加する大域的探索と比較して非常に高速である。また、Lun らは Global Design through Local Search (GDLS) と呼ばれる局所的探索手法を開発し、より広い範囲の破壊候補遺伝子の組み合わせの探索を可能にした (Lun *et al.*, 2009)。GDLS では、もとの代謝モデルに対して2反応程度を破壊できる条件で OptKnock を実行し、ターゲット生産収率の局所的に最大となる反応破壊が獲得される。そして獲得された遺伝子破壊のそれぞれに加えて、新たに1-3反応破壊できる条件で OptKnock を実行し、さらに生産速度の高い遺伝子破壊の組み合わせを獲得することを繰り返す探索手法である。他に、進化的アルゴリズムを利用した探索手法も開発され (Patil *et al.*, 2005)、予測された遺伝子破壊により、大腸菌細胞内で異種生物由来の NADPH 消費反応を活性化させることに成功した例も報告されている (Chemler *et al.*, 2010)。

従来の局所的探索手法では、ターゲット化合物の生産収率が探索における指標として、用いられてきた。しかし、この指標を用いた場合には、単独の遺伝子破壊ではターゲット生産に影響せず、他の遺伝子と同時に破壊することによって初めて、ターゲット生産の上昇が予測される場合に、破壊候補遺伝子の同定が困難となる。例えば、5遺伝子以上の同時破壊によって、ターゲット生産が初めて予測される場合には、4遺伝子以下の破壊ではターゲット生産収率が変化しないため、探索中に有望な遺伝子破壊を選抜することが難しく

なる。このように、ターゲット化合物の生産収率を指標とした局所的探索手法では、ターゲット生産に必要な遺伝子が比較的多い場合における破壊候補遺伝子の同定に課題があった。

以上のように、育種に向けた遺伝子破壊を FBA を利用して同定する手法が開発されてきた。そして、増殖速度最大化の条件下で高いターゲット生産収率が予測される遺伝子破壊は、実際の微生物株においても、生産性の向上に寄与することが、多くの文献で確かめられてきた。しかしながら、多くの遺伝子の同時破壊がターゲット生産及ぼす影響を加味して破壊候補遺伝子を効率よく探索することには課題あり、育種における破壊候補遺伝子の同定に、代謝モデルと FBA を利用したシミュレーションの適用可能な範囲を拡大するには、上記した課題を解決する必要があった。

1.4 本研究の目的および論文の構成

微生物を用いた有用化合物生産は、近年の地球環境問題に対する関心の高まりを受けて、これまでも増して注目されており、産業利用可能な菌株を効率的に育種するための手法の開発が求められている。近年では、代謝の最適化に必要な遺伝子操作を計算機上で検討する技術が進歩し、経験や勘に頼るのではなく、より合理的で効率的な育種法が開発が目指されてきた。遺伝子破壊により不要な代謝経路を不活性化することで、細胞内の代謝を改変する場合に、ゲノムスケール代謝モデルと代謝シミュレーション手法である FBA を利用し、破壊候補遺伝子を探索することが広く行われている。増殖速度最大化の条件下でターゲット化合物の高い生産収率が予測される遺伝子破壊は、これまでの多くの先行研究で、実際の微生物での生産性を向上させたことが報告されている。破壊候補遺伝子の探索に関して、理想的には大域的に探索し、可能な全ての破壊の組み合わせを取りこぼしなく評価することが求められている。しかし、同時に破壊する遺伝子数が多く、現実的な計算時間以内で大域的に最適な多重遺伝子破壊を同定することが困難な場合には、取りこぼしの少ない局所的探索手法を適用し、現実的な計算時間で優良な破壊候補遺伝子を同定するアプローチが有効だと考えられる。したがって、破壊候補遺伝子の探索に対して FBA による代謝シミュレーションを利用する場合には、大域的探索手法において同時に破壊可能な遺伝子数の上限を向上させるとともに、取りこぼしの少ない局所的探索手法を開発する必要があった。そこで本研究では、微生物育種に向けて、増殖速度最大化の条件下で高いターゲット生産収率が予測される多重遺伝子破壊の、大域的探索手法および局所的探索手法の開発を目指した。

本学位論文では以下の第 1 章から第 4 章で構成される。概要を Fig. 1.2 に示す。

第 1 章では、研究の背景として、代謝モデルを利用した微生物育種に関する研究例と課題、および本論文の目的と構成を述べた。

第 2 章では、ターゲット化合物生産に有望な多重遺伝子破壊の大域的探索手法を開発した。破壊対象となる遺伝子の数の削減により、探索を高速化するため、宿主微生物のゲノムスケール代謝モデルの縮約を栄養条件に応じて自動的に行う方法を提案した。代謝モデルの解析により、計算機上で解析したい栄養条件下で使用される可能性のない反応をコードする遺伝子や、細胞増殖に必須の遺伝子を特定し、破壊対象の遺伝子から除外した。このような方法により、例えば 1260 遺伝子から構成される大腸菌の代謝モデルの場合には、グルコース炭素源好気条件下での破壊対象の遺伝子は 393 まで減少し、3 重破壊の探索時間は 10^7 秒から 10^2 秒に短縮された。一般的な計算機を用いた場合に探索可能は遺伝子破壊数の上限については、もとの代謝モデルでは 2 遺伝子破壊だったのに対して、開発手法であれば 5 遺伝子破壊まで可能であることが推測され、代謝モデルを用いた破壊候補遺伝子の探索における破壊数の上限を向上させた。通気条件・炭素源・窒素源が異なる栄養条件下で構築された縮約モデルについては、その差異が解析され、研究者が専門的知見に基づき破壊対象より除外する遺伝子を、提案する縮約方法で栄養条件に応じて自動的に特定できていることも確認された。開発手法の有効性を評価するため、宿主微生物として、微生物育種に広く利用されている大腸菌を選択し、生産ターゲットとして、燃料やプラスチック原料として有益な 1-ブタノール、1-プロパノール、1,3-プロパンジオールの 3 種類を採用した。そして、それぞれの生産性向上に有望な破壊候補遺伝子の探索を行い、3 重破壊の場合について詳細に解析した。その結果、例えば 1,3-プロパンジオールの生産収率が最大となる時の 3 重遺伝子破壊として、*tpiA*、*ptsG*、*pykF* の同時破壊が見つげ出され、解糖系および TCA サイクルの補充経路のフラックスを変化させる代謝改変戦略が予測された。1,3-プロパンジオールは 1 遺伝子破壊では生産が予測されなかったため、従来の単純な局所的探索のように、1 遺伝子破壊と高い生産収率が予測されたものの選出を繰り返す方法では見つけ出せない破壊候補遺伝子も、開発した探索手法により同定することが可能となった。

第 3 章では、大域的探索の適用が難しい多重遺伝子破壊を局所的に探索する手法 **FastPros** (Fast Algorithm of knockout Screening for Target PROduction based on Shadow price analysis) を開発した。ターゲット化合物の生産収率を指標とした従来の局所的探索手法では、単独の遺伝子破壊では生産収率が変化せず、複数の遺伝子の同時破壊により初めて生産収率が上昇する場合には探索が困難だった。そこで、ターゲット生産速度と比増殖速度との関連性を、線形計画問題におけるシャドウプライスを用いて評価することを提案した。そして、これを指標として繰り返し遺伝子を破壊することで、増殖速度最大化の条件下でターゲット生産が予測される多重遺伝子破壊の組み合わせを同定する局所的探索手法 **FastPros** を開発した。**FastPros** の性能を評価するため、大腸菌の細胞内の 625 代謝物質それぞれの生産に有望な破壊候補遺伝子を **FastPros** により探索した。その結果、472 の代謝物質について、増殖速度最大化の条件下で生産が予測される破壊候補遺伝子が同定され、大域的探索または従来の局所的探索では適用が難しい 6 遺伝子以上の同時破壊も予測が可能となった。3 重破壊の大域探索により破壊候補遺伝子が同定される化合物は 152 のみだったことから、

FastPros による探索は、より多くのターゲット化合物に対して適用可能であることが示された。

第 4 章では、本研究で得られた知見をまとめ、微生物育種にどのように貢献するのかについて論じ、今後の展望について述べた。

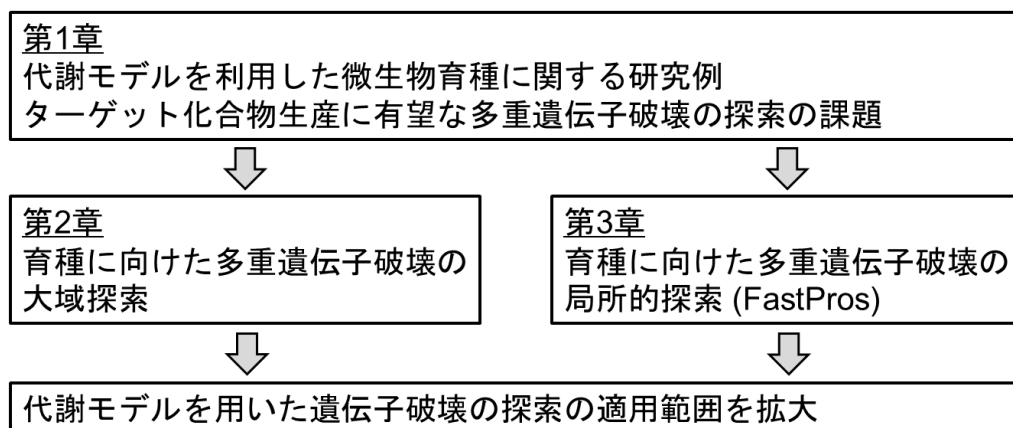


Fig. 1.2 本論文の構成

第2章 育種に向けた多重遺伝子破壊の大域的探索

2.1 緒言

バイオ燃料やプラスチック原料を高効率に生産可能な微生物の育種では、細胞内の代謝フラックス分布をターゲット化合物生産に向けて最適化することが、一般的に行われている。そのためには、異種生物遺伝子の導入などでターゲット化合物の生合成経路を活性化し、複数の遺伝子の破壊により不要な代謝経路の不活性化する必要がある。ターゲット生産性の向上に有望な遺伝子操作の候補を、計算機上で探索する手法が発展し、実験的にも予測された遺伝子操作の影響が評価されてきた。遺伝子破壊の探索に関しては、代謝反応の化学量論式より構成される代謝モデルと、Flux balance analysis (FBA) と呼ばれる代謝シミュレーション手法を活用し、細胞内の数千の遺伝子より破壊候補遺伝子を探索することが広く行われている。

計算機上でターゲット生産性の向上に有望な破壊候補遺伝子を比較的に簡単に予測する方法として、宿主微生物の代謝モデルに対して、1 遺伝子を破壊 (計算上は対応する反応を破壊) した時の代謝フラックス分布を網羅的にシミュレートし、その中でターゲット化合物生産収率の高いものを見つけ出すといった、単純な局所的探索手法が挙げられる。より高い生産収率に必要な破壊候補遺伝子を同定する場合には、高い生産収率が達成された遺伝子破壊に対して、新たに 1 遺伝子を破壊した場合の代謝フラックス分布を網羅的に計算し、結果として生産収率が最大となった遺伝子破壊の組み合わせを見つけ出される。この単純な局所的探索手法で、例えば Alper らは、大腸菌によるリコペン生産に有望な 3 重遺伝子破壊を予測し、その遺伝子破壊によって実際の大腸菌によるリコペン生産濃度を 37% 上昇させることに成功した (Alper *et al.*, 2005)。また、Park らは、大腸菌による L-バリン生産に対して同様の手法を適用し、予測に基づく遺伝子破壊により L-バリン生産濃度を 2.3 倍に上昇させた (Park *et al.*, 2007)。このように、増殖速度最大化の条件下で高いターゲット生産収率が予測される遺伝子は、実際の微生物においても生産性向上に寄与することが確かめられてきたため、様々な生物種・栄養条件・生産ターゲットに対して、FBA に基づき破壊候補遺伝子を探索することが一般的になりつつある。しかしながらこの方法では、1 遺伝子破壊ではターゲット生産収率が変化せず、2 遺伝子以上の同時破壊によって初めて生産収率の増加が予測される場合に、破壊候補遺伝子の同定ができない、という問題があった。

そこで本章では、ゲノム中の 10^3 程度の代謝遺伝子を破壊の対象とした場合に、ターゲット化合物生産に有望な破壊候補遺伝子の組み合わせを大域的に探索する手法を開発した。探索を高速化するため、栄養条件に応じて縮約したモデルを自動的に構築し、評価の必要のない遺伝子を破壊対象から除くことで、破壊対象の遺伝子数を削減する方法を提案した。そして、縮約した代謝モデルを用いて、遺伝子破壊に伴う代謝変動を FBA でシミュレートし、ターゲット化合物の生産収率の上昇が大きい遺伝子破壊を抽出することで、ターゲッ

ト生産性向上に有望な破壊候補遺伝子を同定した。

開発した手法の有効性を示すため、大腸菌に工学的に有用なターゲット化合物を生産させる場合に対して適用した。大腸菌は、モデル生物として知見が多く、また様々なターゲット生産における宿主微生物としても利用されている。さらに、ゲノムスケール代謝モデルも構築されており、育種に向けた遺伝子操作等の探索に利用されてきたため、本章での宿主微生物として採用した (Feist *et al.*, 2007)。ターゲット化合物としては、バイオ燃料やプラスチック原料として注目を集めている 1-ブタノール、1-プロパノール、1,3-プロパンジオール (13PD) の3種類を選択し、それぞれのターゲット生産に有望な破壊候補遺伝子を、開発手法により同定した。

2.2 提案手法

2.2.1 代謝モデル

本章では、大腸菌 K-12 MG1655 株のゲノムスケール代謝モデル iAF1260 (Feist *et al.*, 2007) を使用した。この代謝モデルは、1260 遺伝子、2077 反応、1039 代謝物質より構成される。このモデルは、同じ研究グループが 2003 年に発表した大腸菌の代謝モデル (Reed *et al.*, 2003) を拡張したものであり、データベース EcoCyc (EcoCyc: Encyclopedia of *E. coli* Genes and Metabolic Pathways) および MetaCyc (MetaCyc: Encyclopedia of Metabolic Pathways) との比較により予測精度と網羅性が向上された。この代謝モデルを用いて、FBA を実行した場合、グルコース単一炭素源下での培養における遺伝子の増殖必須性が 92% 一致し、広く利用されていることから、信頼できる代謝モデルとして本研究で使用した。

代謝モデル iAF1260 を使用する際に、文献 (Feist *et al.*, 2007) にならい FBA の条件を設定した。細胞増殖を表現する反応としては、代謝モデル中の **core biomass objective function** を採用し、FBA における目的関数として使用した。**core biomass objective function** は、観測された菌体構成成分がすべて含まれている菌体合成反応を、遺伝子の増殖必須性が実験値と一致するように修正したものである (Feist *et al.*, 2007)。計算機上での最小培地として、Table 2.1 に示す栄養源は細胞が自由に消費できるものとした。また、破壊候補遺伝子を探索する際に、ターゲット化合物生産における原料となる炭素源としては、グルコースあるいはグリセロールを選択した。グルコースは、発酵生産において原料として広く利用されている廃糖蜜の成分であり、グリセロールは、バイオディーゼル生産における副産物として、近年価格が低下している化合物である (Ciriminna *et al.*, 2014)。培養条件としては、増殖連動での高いターゲット化合物生産性を目指すために、微好気条件を想定した。好気条件下では、消費された炭素が菌体合成に使用される割合が大きいためである。また、Catalase、Superoxide dismutase、Superoxide dismutase in periplasm の 3 反応については、酸化ストレスに対する応答であり、本研究で想定する培養条件とは異なるため、あらかじめ代謝モデルより取り除いた。水素発生に関わる Formate-hydrogen lyase についても、pH が 7 付近の通常の培養条件下では活性を持たないため、モデルより除外した。

開発手法の性能を調べるために本章で採用したターゲット化合物は、1-ブタノール、1-プロパノール、13PD であり、大腸菌の野生型株はいずれの化合物の生合成経路も所持しない。そこで、Table 2.2-4 に示す異種生物種由来の生合成経路およびトランスポートをそれぞれ代謝モデルに追加した後に、破壊候補遺伝子の探索を行った。

Table 2.1 計算機上の最小培地において自由に消費できる栄養源

栄養源	分類	栄養源	分類
NH ₄ ⁺	窒素源	K ⁺	無機イオン
phosphate	リン源	Mg ²⁺	無機イオン
SO ₄ ²⁻	硫黄源	Mn ²⁺	無機イオン
Ca ²⁺	無機イオン	molybdate	無機イオン
Cl ⁻	無機イオン	Na ⁺	無機イオン
Co ²⁺	無機イオン	Zn ²⁺	無機イオン
Cu ²⁺	無機イオン	CO ₂	その他
Fe ²⁺	無機イオン	H ⁺	その他
Fe ³⁺	無機イオン	H ₂ O	その他

代謝モデル iAF1260 の文献中で記述されているものと同じである (Feist *et al.*, 2007)。炭素源としてグルコースあるいはグリセロールを消費可能とする場合は、消費速度の上限をそれぞれ 10、20 mmol/gDW/h (gDW: gram dry cell weight) とした。酸素に関しては、微好気条件を想定する場合は、その消費速度の上限を 5 mmol/gDW/h とした。

Table 2.2 代謝モデルに追加した *Clostridium acetobutylicum* 由来の 1-ブタノール生合成経路

追加した反応	EC 番号	反応の化学量論式
Butyryl-CoA dehydrogenase	1.3.8.1	crotonoyl-CoA + NADH + H ⁺ → butyryl-CoA + NAD ⁺
Butanal dehydrogenase	1.2.1.10	butyryl-CoA + NADH + H ⁺ → butanal + NAD ⁺
Butanol dehydrogenase	1.1.1.-	butanal + NADH + H ⁺ → 1-butanol + NAD ⁺
1-Butanol transport (cytosol → periplasm)	-	1-butanol + H ⁺ → 1-butanol[p] + H ⁺ [p]
1-Butanol transport (periplasm → extracellular)	-	1-butanol[p] → 1-butanol[e]
1-Butanol exchange	-	1-butanol[e] →

[p]、[e]はそれぞれ、その代謝物質がペリプラズム空間あるいは細胞外空間にあることを示し、他の代謝物質は細胞質にあることを意味している。*C. acetobutylicum* の 1-ブタノールトランスポーターは同定されていないので、代謝モデルに追加する 1-ブタノールのトランスポーターは、代謝モデル中のエタノールのトランスポーターと同様に、プロトンシンポートとした。また、アセチル CoA からのクロトノイル CoA 合成を可能にするため、代謝モデルに含まれる Acetyl-CoA transferase 反応は不可逆反応から可逆反応へと修正された。

Table 2.3 代謝モデルに追加した *Klebsiella pneumoniae* 由来の 1-プロパノール生合成経路

追加した反応	EC 番号	反応の化学量論式
Glycerol dehydratase	4.2.1.30	glycerol → lactaldehyde + H ₂ O
Propanediol dehydratase	4.2.1.28	1,2-propanediol → propanal + H ₂ O
1,3-Propanediol dehydrogenase	1.1.1.202	propanal + NADH + H ⁺ → 1-propanol + NAD ⁺
1-Propanol transport (cytosol → periplasm)	-	1-propanol + H ⁺ → 1-propanol[p] + H ⁺ [p]
1-Propanol transport (periplasm → extracellular)	-	1-propanol[p] → 1-propanol[e]
1-Propanol exchange	-	1-propanol[e] →

K. pneumoniae の 1-プロパノールトランスポーターは同定されていないため、モデルに追加する 1-プロパノールのトランスポーターは、使用する代謝モデル中のエタノールのトランスポーターと同様に、プロトンシンポートとした。

Table 2.4 代謝モデルに追加した *Klebsiella pneumoniae* 由来の 1,3-プロパンジオール (13PD) 生合成経路

追加した反応	EC 番号	反応の化学量論式
Glycerol dehydratase	4.2.1.30	glycerol → 3HPA + H ₂ O
13PD dehydrogenase	1.1.1.202	3HPA + NADH + H ⁺ → 13PD + NAD ⁺
13PD transport (cytosol → periplasm)	-	13PD → 13PD[p]
13PD transport (periplasm → extracellular)	-	13PD[p] → 13PD[e]
13PD exchange	-	13PD[e] →

3HPA は 3-ヒドロキシプロピオンアルデヒドの略称である。*K. pneumoniae* の 13PD トランスポーターは同定されていないため、モデルに追加する 13PD のトランスポーターは、代謝モデル中の 1,2-プロパンジオールのトランスポーターと同様に、単純拡散とした。

2.2.2 FBA

本章では、FBA を利用して、代謝フラックス分布および細胞の増殖速度をシミュレートした。FBA は、定常を満たすフラックス分布の中で、細胞の比増殖速度が最大となるものが予測されるシミュレーション手法である。本章での FBA は、次の線形計画問題として記述される。

$$\begin{aligned}
 & \max \quad v_{growth} + \alpha \cdot v_{target} \\
 & (\mathbf{v}) \\
 \text{s. t.} \quad & \sum_{j \in R} S_{i,j} \cdot v_j = 0 \quad (\forall i \in M) \\
 & v_{carbon_uptake} \geq -CUR \\
 & v_{o2_uptake} \geq -OUR \\
 & v_{atp_main} \geq NGAM \\
 & v_{growth} \geq v_{growth}^{min} \\
 & v_j^{min} \leq v_j \leq v_j^{max} \quad (\forall j \in R) \\
 & 0 \leq T_{k,j} \cdot v_j \leq 0 \quad (\forall j \in R, \forall k \in \mathbf{G}_{knockout})
 \end{aligned}$$

ただし、

v_j	: 反応 j のフラックス (mmol/gDW/h) (gDW: gram dry cell weight)
v_{carbon_uptake}	: 炭素源消費反応のフラックス
v_{o2_uptake}	: 酸素消費反応のフラックス
v_{atp_main}	: ATP maintenance 反応のフラックス
v_{growth}	: 菌体合成反応のフラックス (細胞の比増殖速度)
v_{target}	: ターゲット化合物の生産速度
CUR	: 炭素源の消費速度の上限値 (グルコース: 10 mmol/gDW/h、グリセロール: 20 mmol/gDW/h)
OUR	: 酸素消費速度の上限値 (5 mmol/gDW/h)
NGAM	: 増殖に関係なく細胞維持に必要な ATP 量 (8.39 mmol/gDW/h) (Feist <i>et al.</i> , 2007)
v_{growth}^{min}	: 細胞の比増殖速度の下限值 (0.05 mmol/gDW/h)
v_j^{max}	: 反応 j のフラックスの上限値 (10^6 mmol/gDW/h)
v_j^{min}	: 反応 j のフラックスの下限值 (可逆反応: -10^6 mmol/gDW/h、不可逆反応: 0 mmol/gDW/h)
$S_{i,j}$: 反応 j における代謝物質 i にかかる化学量論係数
$T_{k,j}$: 遺伝子 k が反応 j をコードするとき 1、それ以外の時 0 となる 行列
M	: 代謝物質の集合
R	: 代謝反応の集合
$\mathbf{G}_{knockout}$: 破壊対象の遺伝子の集合

α : 増殖速度最大化の条件下でのターゲット化合物生産速度を一意に定まめるため係数(10^{-6})

である。

菌体合成反応のフラックスの下限值 (v_{growth}^{min}) は 0.05 mmol/gDW/h とし、遺伝子破壊によりこれ以上の増殖速度を細胞がとれない場合には、この遺伝子破壊株を実験的に構築するのは難しいとみなされ、探索結果から自動的に除外された。遺伝子破壊に伴う反応の不活性化は、その代謝フラックスを 0 に制限することで表現した。これは、代謝モデルよりその反応を除外することと本質的には変わらない。そこで、本論文では以降、「遺伝子破壊により対応する反応のフラックスを 0 と制限して FBA を実行する」場合には、「破壊後の代謝モデルに対して FBA を実行する」と記述する。

2.2.3 代謝モデルの縮約

一般に、ある遺伝子破壊がターゲット化合物生産にどの程度有望であるかを FBA を利用して評価する際には、遺伝子破壊の前後でのターゲット化合物生産収率の差が FBA で評価される。したがって、破壊対象となる遺伝子数を g 、破壊数を k 、代謝モデルに対して FBA を実行する際に必要な計算時間の平均を t と置くと、破壊遺伝子の探索に必要な計算時間は $gC_k \times t$ と記述できる。そこで本章では、解析したい栄養条件に応じて縮約した代謝モデルを自動的に構築する方法を提案し、破壊対象遺伝子数 g を減少させることで、破壊候補遺伝子の探索を高速化した。

提案する代謝モデルの縮約は、以下の 4 つのステップより構成される (Fig. 2.1)。ステップ 1 では、解析したい栄養条件下で使用される可能性のない反応およびそれをコードする遺伝子を、代謝モデルより除外した。これまで、準最適な代謝状態 (Reed and Palsson, 2004) や代謝ネットワークの冗長性 (Thiele *et al.*, 2010) を解析するために利用されてきた Flux variability analysis (Mahadevan and Schilling, 2003) と呼ばれる手法を、破壊前の代謝モデルに対して適用し、各代謝フラックスの上限値と下限値を計算した。具体的には、2.2.3 で定義された制約条件に対して、各代謝フラックス v_j を最大化あるいは最小化する線形計画問題を解いた。各代謝フラックスでその結果、上限値と下限値がともに 0 だと計算された反応およびそれをコードする遺伝子については、以降のシミュレーション結果に影響しないため、代謝モデルより除外した。なお、除外された反応は主に、エキステンジ反応、トランスポート反応、炭素源の異化反応であり、これらのカテゴリーに属する反応は、除外された反応の約 70% を占めた。

ステップ 2 では、枝分かれのない連続した反応とそれをコードする遺伝子を連結した。代謝モデル中のある代謝物質について、それに関わる反応が 2 反応のみである場合には、定常を満たすフラックス分布においては、一方の代謝フラックスが定まれば、もう一方の代謝フラックスも一意に定まる。そこで、このような 2 反応を連結し 1 反応とみなし、関

連する遺伝子もまとめて 1 遺伝子とみなすことで、代謝モデルの持つ反応数および遺伝子数を削減した。この反応および遺伝子の連結は、代謝モデル中に上記した代謝物質が存在しなくなるまで、繰り返し行われた。連結後の反応の可逆性は、連結前の反応の可逆性に従って定めた。

ステップ 3 では、計算機上で増殖に必須の遺伝子を破壊対象から除いた。1 遺伝子が破壊された代謝モデルに対して FBA を実行し、計算される増殖速度が 0 となった遺伝子については、破壊対象から除いた。なぜなら、これらの遺伝子の破壊は現実の細胞にとっても、細胞増殖を著しく減少させ、破壊株の構築が困難であるからである。

ステップ 4 では、その他の生物学的な理由から一部の遺伝子を破壊対象から除いた。*gapA*、*pgk*、*eno*、*gpmA* は、実際の大腸菌の増殖に必須の遺伝子であることが報告されているので (Baba *et al.*, 2006; Foster *et al.*, 2010)、破壊対象から除いた。 H_2O のトランスポートに関わる外膜タンパクをコードする遺伝子については、他の化合物の膜輸送にも関わり、 H_2O の膜輸送のみをこの遺伝子の破壊により阻害することは困難であるため、破壊対象から除外した。電子伝達系の ATP synthase をコードする遺伝子については、破壊後の代謝モデルについて、増殖速度最大化の条件下でそれぞれの代謝フラックスがとりうる値の範囲が比較的に大きく、現実的なシミュレーションがなされていないため、破壊対象から除外した。

以上の 4 つのステップにより、縮約した代謝モデルを構築する方法を提案した。グルコース炭素源好気条件下での縮約において、それぞれのステップでのモデル中の代謝物質数・代謝反応数・破壊対象の遺伝子数を Table 2.5 に示す。最終的には、破壊対象の遺伝子数は、縮約前の 31%まで削減された。提案した縮約方法は、代謝反応の化学量論式から構成される代謝モデルそのものに基づくものであり、ステップ 4 を除き、生物種や栄養条件に特有の知見は利用されていない。そのため、提案した代謝モデルの縮約方法は、化学量論式からなる代謝モデルと FBA を利用して破壊候補遺伝子を探索する場合に、汎用的に利用可能である。また、FBA による代謝シミュレーション結果が全く変化しないままに、代謝モデルの中で不要な遺伝子や反応を除外している点も、提案する縮約方法の特徴である。1 遺伝子が破壊された代謝モデルに対して FBA を実行し、計算されたフラックス分布がモデルの縮約前後で完全に一致することから、縮約後の代謝モデルの妥当性が確認された。また、育種に向けた探索により同定された破壊候補遺伝子についても、その遺伝子破壊を適用したのちに FBA でシミュレートされる代謝フラックス分布が、モデルの縮約前後で完全に一致することを確認した。以上のように、モデルの縮約前後で FBA の結果が変わらないようにしつつも、代謝モデルそのものの情報を利用し、解析したい栄養条件に応じて異なる縮約モデルを自動的に構築する方法を提案した。

以下、探索の結果として示す破壊候補遺伝子は、代謝モデルの縮約によりまとめられた遺伝子のうち、代表的なもののみを記述するものとする。例えば、アセチル CoA から酢酸の生合成に関わる反応は、Phosphotransacetylase と Acetate kinase の 2 反応であり、それぞれ (*eutD* or *pta*) と (*tdcD* or *ackA* or *purT*) によってコードされている。遺伝子破壊によりこれら

の反応を不活性化する場合には、煩雑さを避けるため、*pta* 遺伝子の破壊として記述する。各遺伝子がコードする反応については、Table A1 に示す。

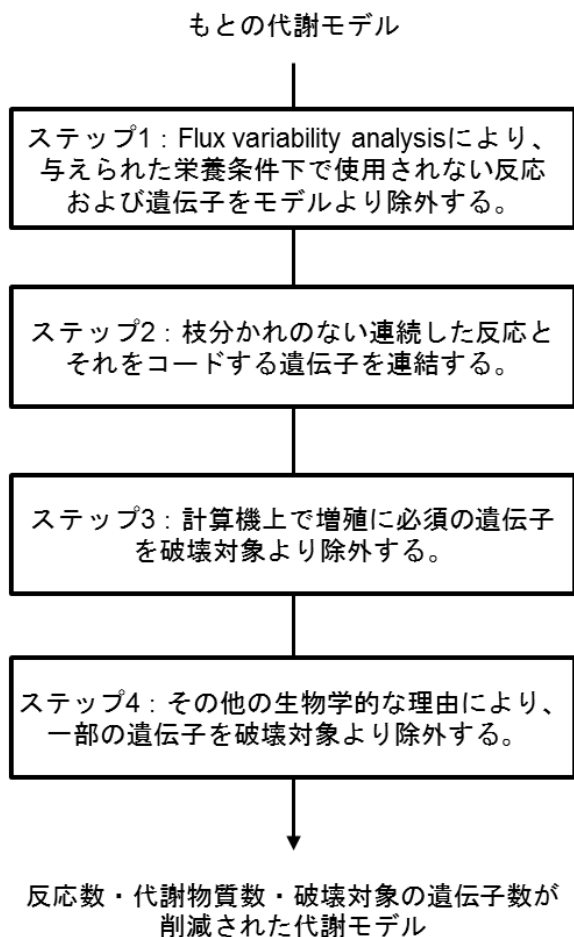


Fig. 2.1 代謝モデルの縮約の概要

Table 2.5 大腸菌代謝モデルの縮約結果

	縮約前	ステップ 1	ステップ 2	ステップ 3	ステップ 4 (最終)
代謝物質数	1668	1032	505	505	505
代謝反応数	2382	1531	1010	1010	1010
破壊の対象の 遺伝子数	1260	661	518	399	393

グルコース炭素源好気条件での縮約結果である。他の栄養条件の場合でも同様の結果が得られた。ここでの代謝反応はエキスチェンジ反応を含み、代謝物質は、細胞質・ペリプラズム空間・細胞外空間で区別している。

2.2.4 破壊により代謝が変動しない遺伝子の特定

ここでは、FBA によって代謝シミュレーションの必要のない遺伝子破壊後の代謝モデルについて議論する。ある代謝モデルのフラックス分布を FBA でシミュレートすると、代謝フラックスが 0 と計算された反応が少なからず存在する。このような反応のみをコードする遺伝子のみを新たに破壊した場合には、FBA で計算される代謝フラックス分布は、遺伝子破壊の前後で変動しない。そこで、探索時に FBA で計算される全ての代謝フラックス分布を用いて、新たな破壊により代謝が変動しない遺伝子を特定し、その遺伝子破壊に伴う代謝変動の計算を省略した。このような遺伝子は、縮約後の代謝モデルにおける破壊対象遺伝子のうち、およそ 75% 存在した。なお、ランダムに 1000 程度選んだ 3 重破壊について、それらの破壊に伴う代謝変動が、シミュレーションを省略しない場合の代謝変動を完全に一致することは確認されている。ここでも、探索結果として予測される破壊候補遺伝子に影響が現れないように、探索を高速化した。

2.2.5 大域的探索手法

以上の手法により、育種に向けた大域的探索手法を開発した (Fig. 2.2)。具体的には、もとの代謝モデルを縮約した後に、FBA による代謝シミュレーションを実行し、遺伝子破壊に伴う代謝変動にシミュレートした。この際に、破壊に伴い代謝が変動しない遺伝子を特定し、その遺伝子破壊に伴う代謝変動のシミュレーションは省略した。そして、この FBA による代謝シミュレーションと、次の遺伝子破壊におけるシミュレーションの省略を、破壊数を増加しながら順に行った。最後に、破壊によりターゲット化合物の生産収率の増加が大きい遺伝子を、そのターゲット生産に有望な破壊候補遺伝子として抽出した。この手法では、破壊候補遺伝子の複数の組み合わせを同時に獲得することが可能である。

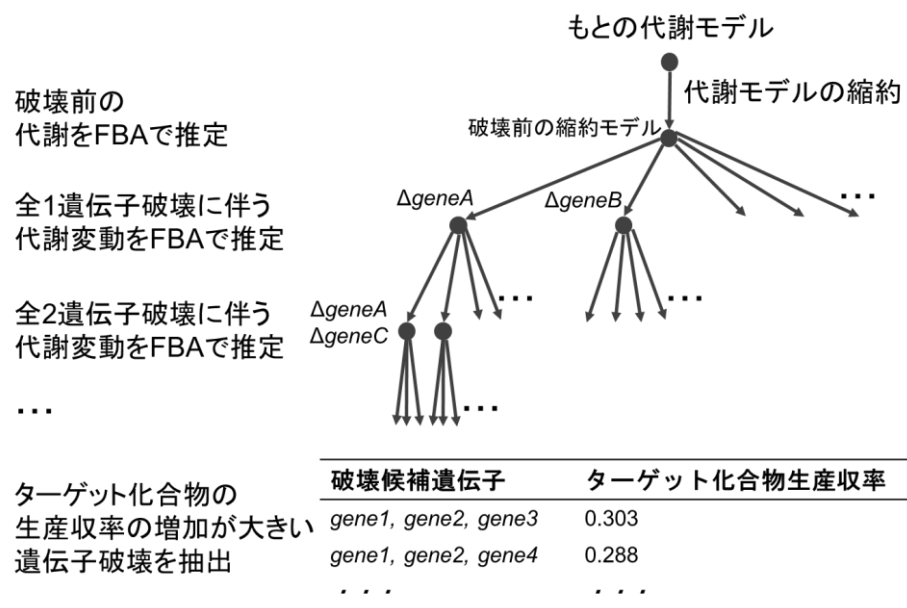


Fig. 2.2 開発した多重遺伝子破壊の大域的探索手法の概要

2.3 結果

2.3.1 計算時間

本章では、微生物育種に向けた多重遺伝子破壊の大域的探索手法を開発した。栄養条件に応じて縮約した代謝モデルを自動的に構築する方法を提案し、破壊対象の遺伝子数を削減することで探索を高速化した。開発手法の性能を調べるため、大腸菌に工学的に有用な化合物を生産させる場合に対して適用した。大腸菌はモデル生物として多くの知見が蓄積されており、発酵生産に広く利用され、またゲノムスケールの代謝モデルが精密に構築されているため、宿主微生物として選択した。生産のターゲットとしては、燃料やプラスチック原料として有用な 1-ブタノール、1-プロパノール、13PD を選択した。そして、グルコースあるいはグリセロールを炭素源とし、微好気条件を想定した場合について詳細に解析した。また、すべての計算は MATLAB (Mathworks Inc) を Windows 上で使用して行った。FBA などの最適化問題を解く場合には、ソルバーとして GLPK (GNU linear programming kit) の関数を MATLAB 上で使用した。CPU には、Intel Xeon CPU (2.93 GHz, 12 core) を使用した。

大腸菌のゲノムスケール代謝モデルを用いて、グルコース炭素源下での 1-ブタノールの生産に向けた多重遺伝子破壊を探索し、計算時間を測定した (Fig. 2.3)。その結果、探索時間が大幅に短縮され、例えば 3 重破壊に関して縮約前のモデルを用いた従来の探索手法の場合、 1.4×10^7 秒 (160 日) 必要だと推測されたが、開発手法により 2.3×10^2 秒 (3.8 分) に短縮された。探索時間が短縮された主な要因は、FBA の実行回数が、 10^9 から 10^4 のオーダーまで大幅に削減されたことであり、これは主に、代謝モデルの縮約による破壊対象数の削減によるものだった。代謝モデルの縮約には 4.1×10^1 秒を要したため、同時に 2 遺伝子以上を破壊する場合には、従来手法より開発手法の方が短時間で探索可能であることが示された (Fig. 2.3)。本研究の計算環境において、現実的に探索可能な破壊数については、従来手法では 2 重破壊 (5 時間) 程度が限界だったが、開発手法では 5 重破壊 (13 時間) まで可能だと推測される。なお、ターゲット化合物として 1-プロパノールあるいは 13PD を選択した場合や、炭素源をグリセロールに変更した場合にも、これらの結果はおおよそ同じだった。このことから、開発した探索手法は、ターゲットや炭素源によらず汎用的に利用可能であることが示された。

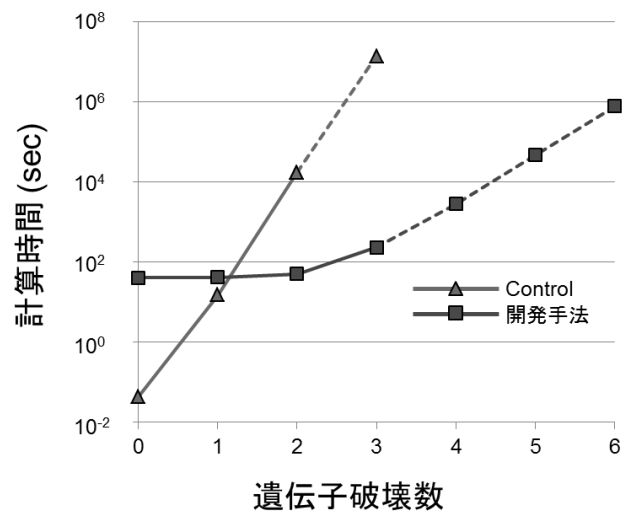


Fig. 2.3 開発した多重遺伝子破壊の大域的探索手法の計算時間

実線は実測値、点線は推定値である。Control は、モデル縮約前の代謝モデルを用いて、考え得る全ての破壊対象について FBA を実行した時の計算時間である。開発手法の計算時間は、モデル縮約の計算時間 4.1×10^1 秒を加味している。

2.3.2 栄養条件に応じた縮約モデルの違い

探索時間を高速化するために提案した代謝モデルの縮約は、解析したい栄養条件に応じて異なる縮約モデルが自動的に構築される。栄養条件に対して縮約モデル中の破壊対象の遺伝子が異なれば、解析したい栄養条件に応じて縮約した代謝モデルを構築し、破壊候補遺伝子を探索する必要があるといえる。そこで、栄養条件に応じた縮約モデルの違いに関して解析した。

まず、通気条件について、好気条件と嫌気条件の 2 種類の場合で構築された縮約モデルの違いを解析した。好気条件とは、Table 2.1 に示す栄養とグルコースを消費できるものとし、さらに酸素も自由に消費できる条件である。嫌気条件とは、好気条件から酸素が全く消費できないように変更したものである。もとの大腸菌代謝モデルより、好気条件と嫌気条件では、それぞれ 399 と 369 の破壊対象の遺伝子が含まれる縮約モデルが構築された。好気条件と嫌気条件の縮約モデル中の破壊対象遺伝子の違いを、Table 2.6 に示す。好気条件の縮約モデルでのみ破壊対象となる 31 遺伝子は、酸素またはユビキノンに関わる反応に関わるものであり、電子伝達系、炭素源の異化経路、補酵素合成、アミノ酸代謝などの代謝全体に散在していた。嫌気条件では、電子伝達系において最終的に電子を受容する酸素が存在しないので、その前段階としてのユビキノンが電子を受容することができない。そのため、ユビキノンの酸化還元が関与する反応の遺伝子が、嫌気条件下での代謝モデルでは欠失していた (Fig. 2.4)。なお、同じ電子伝達物質である NAD^+ 、メナキノン、ジメチルメナキノンについては、エタノール、乳酸、コハク酸の生合成等で他の化合物に電子を受け渡すことが可能であるため、これらの電子伝達物質に関わる反応の遺伝子は嫌気条件下での縮約モデル中にも存在していた。このように、研究者が専門的知見に基づき破壊対象から除外していた遺伝子を、本章で提案する縮約方法により自動的に特定することで、微生物育種における破壊候補遺伝子の探索を高速化することに成功した。

次に、異なる炭素源に関して、縮約モデルの違いを解析した。好気条件グルコース炭素源の場合と、炭素源をグリセロールあるいはキシロースに変化させた場合について、縮約モデルで破壊対象となる遺伝子を比較した (Table 2.7, 8)。グルコース炭素源の場合でのみ破壊対象となる 9 遺伝子は、グルコースからイドン酸への変換反応および中間生成物のトランスポート反応をコードしており (Fig. 2.5)、比較対象としてグリセロールあるいはキシロースを炭素源として選択した場合で共通していた。グリセロール炭素源でのみ破壊対象となる遺伝子は存在しなかったが、キシロース炭素源でのみ破壊対象となる遺伝子は、キシロースのトランスポート反応およびペントースリン酸経路中のキシロース 5 リン酸への変換反応をコードしていた。

また、異なる窒素源に関しても、縮約モデルの違いを解析した。好気条件グルコース炭素源において、窒素源としてアンモニアあるいは硝酸を選択した場合の縮約モデルに関して、縮約モデルで破壊対象となる遺伝子を比較した (Table 2.9)。硝酸を窒素源とした場合に

のみ破壊対象となる 7 遺伝子は、硝酸を還元してアンモニアに変換する反応に関わるものだった (Fig. 2.6)。このように、炭素源あるいは窒素源を変化させた場合にも、縮約モデル中で破壊対象となる遺伝子に差異が現れた。

以上より、解析したい栄養条件に応じて、自動的に構築された縮約モデルに違いが生じ、破壊対象の遺伝子に差異が現れることが示された。この差異は、代謝全体に生じている場合もあり、単純に栄養源の取り込みやそれに直接関わる代謝経路のみを考慮するだけでは不十分である。したがって、本章で開発した手法により縮約した代謝モデルを自動的に構築することは、微生物育種や、そのための破壊候補遺伝子の探索に有益だといえる。

Table 2.6 好気条件あるいは嫌気条件での縮約モデルに含まれる破壊対象遺伝子の違い

好気条件の縮約モデルでのみ破壊対象となる遺伝子			嫌気条件の縮約モデルでのみ破壊対象となる遺伝子
<i>allA</i>	<i>allB</i>	<i>allD</i>	<i>mtn</i>
<i>allP</i>	<i>bcd</i>	<i>cbdA</i>	<i>pyrC</i>
<i>cueO</i>	<i>cydA</i>	<i>cyoA</i>	
<i>dld</i>	<i>gcd</i>	<i>glpD</i>	
<i>gntT</i>	<i>hemE</i>	<i>hemF</i>	
<i>hemG</i>	<i>hemH</i>	<i>hemN</i>	
<i>idnK</i>	<i>idnO</i>	<i>idnT</i>	
<i>msrA</i>	<i>msrB</i>	<i>pdxH</i>	
<i>poxB</i>	<i>puuB</i>	<i>pyrB</i>	
<i>pyrD</i>	<i>pyrE</i>	<i>sdhC</i>	
<i>ubiB</i>			

代謝モデルの縮約によりまとめられた遺伝子のうち、代表的なもののみを記述している。

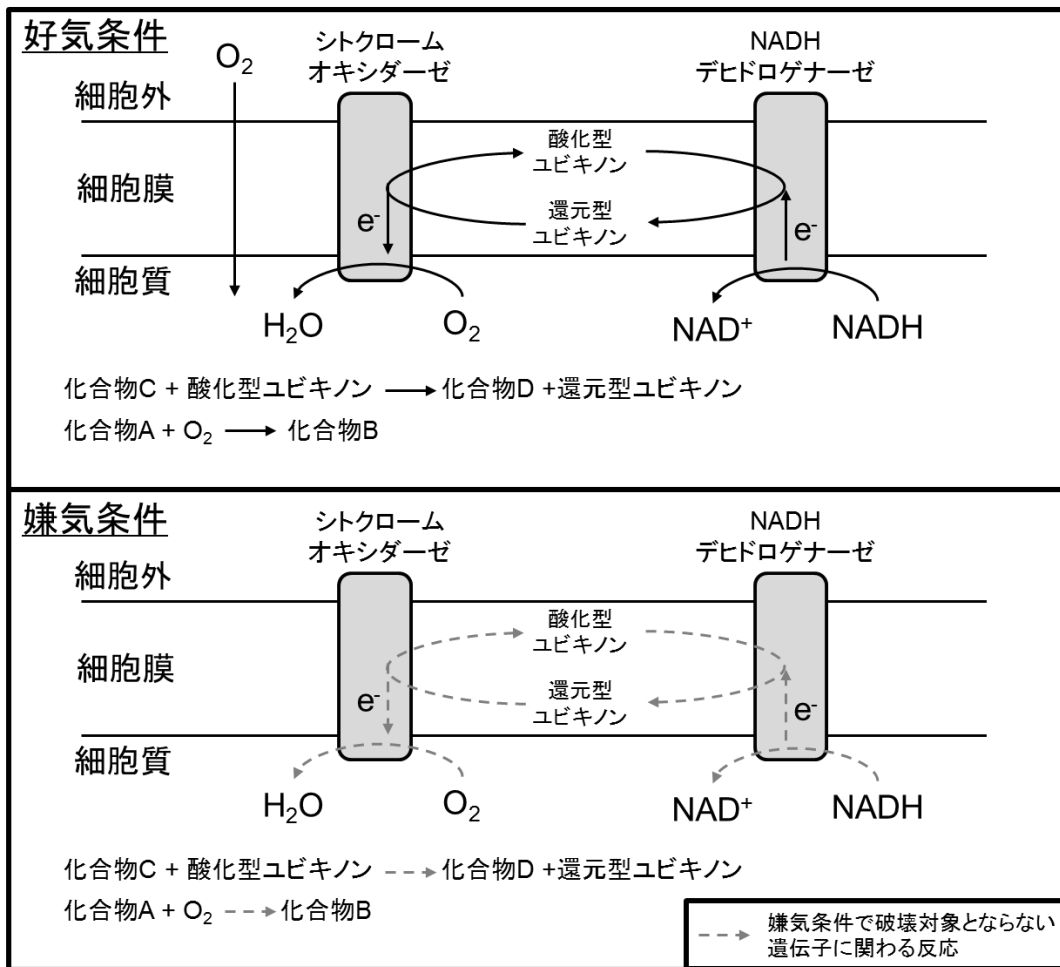


Fig. 2.4 通気条件による破壊対象遺伝子に関わる反応の違い

嫌気条件下では、酸素またはユビキノンの関わる反応に関わる遺伝子が、破壊対象より除外された。

Table 2.9 アンモニア窒素源と硝酸窒素源の違いによる縮約モデルの代謝反応の違い

アンモニア窒素源の縮約モデルで 硝酸窒素源の縮約モデルでのみ破壊対象となる遺伝子のみ破壊対象となる遺伝子

(None)	<i>napC</i>	<i>napG</i>	<i>narG</i>
	<i>narU</i>	<i>nirB</i>	<i>nirC</i>
	<i>nrfA</i>		

代謝モデルの縮約によりまとめられた遺伝子のうち、代表的なもののみを記述している。

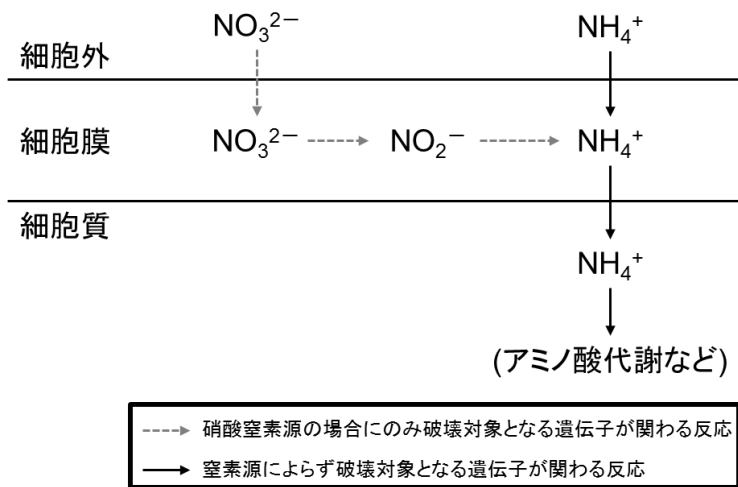


Fig. 2.6 窒素源による破壊対象遺伝子に関わる反応の違い

2.3.3 大腸菌による 1-ブタノール生産への本手法の適用

ここからは、開発した 3 重遺伝子破壊の探索手法を、工学的に有用な化合物生産を目指して適用した結果を示す。宿主微生物としては、発酵生産に広く利用され、また代謝モデルが精密に構築されている大腸菌を選択した。生産のターゲットとしては、燃料やプラスチック原料として有用な 1-ブタノール、1-プロパノール、13PD を選択した。

大腸菌による 1-ブタノール生産の場合には、まず、大腸菌の代謝モデル iAF1260 (Feist *et al.*, 2007) に対して、Table 2.2 に示す代謝反応およびトランスポート反応を追加した。この反応の追加により、NADH の酸化を伴う反応経路によりアセチル CoA から 1-ブタノールを生合成することが可能となる。次に、グルコースあるいはグリセロールを単一炭素源とし、微好気条件を想定して代謝モデルの縮約を行い、開発した手法により 1-ブタノールの高い生産収率が予測される遺伝子破壊を探索した。その結果、Table 2.6 に示す破壊候補遺伝子が同定された。なお、野生型の代謝モデルに 1-ブタノール生合成経路を導入しただけでは、グルコースとグリセロールのいずれの炭素源の場合でも、細胞の増殖速度を最大化する条件下では 1-ブタノールは生産されなかった。

グルコースを単一炭素源とした場合には、高い 1-ブタノール生産収率が予測された破壊候補遺伝子の組み合わせの全てに、*adhE* 遺伝子と *pta* 遺伝子が含まれていた (Table 2.10)。*adhE* と *pta* はそれぞれ、アセチル CoA からのエタノールあるいは酢酸の生合成経路をコードする遺伝子であり、これらの遺伝子の 2 重破壊により、0.271 g/g-glucose の収率での 1-ブタノール生産が予測された。したがって、高い 1-ブタノール収率を達成するには、これらの 2 重破壊が重要であることが示唆された。*adhE* 遺伝子を破壊すると、微好気条件下でエタノール生合成に伴う NADH の酸化が阻害されるため、同じく NADH の酸化を伴う 1-ブタノール生合成が促進された。大腸菌野生型株では、NADH の酸化を伴う代表的な反応として、エタノール生合成反応の他に乳酸生合成反応が挙げられる。しかし、乳酸は 1-ブタノールよりも生合成における NADH 酸化の効率が低いため、破壊候補遺伝子として予測はされなかった (1 mol のピルビン酸からの生合成で酸化される NADH 量は、乳酸合成では 1 mol、1-ブタノール合成では 2 mol である)。一方で、*pta* 遺伝子は ATP の合成を伴う酢酸生合成経路をコードしている。そのため、*pta* 遺伝子を破壊すると、菌体合成に必要な ATP を補うため ATP 合成を伴う TCA サイクルのフラックスが上昇し、これにより生じた NADH を酸化するために 1-ブタノール生合成が促進された。

adhE 遺伝子と *pta* 遺伝子の 2 重破壊に加えて、1-ブタノール生産収率向上に有望な破壊候補遺伝子としては、*tpiA* 遺伝子や *pgi* 遺伝子などが予測された。*tpiA* と *pgi* はそれぞれ解糖系上流の Triose phosphate isomerase 反応と Phosphoglucose isomerase 反応をコードしている。*adhE* と *pta* の 2 重破壊に加えて、*tpiA* 遺伝子、あるいは *pgi* 遺伝子を破壊すると、酸化的ペントースリン酸経路のフラックスが上昇し、それに伴い NADPH の生産が増加する。この NADPH は Transhydrogenase 反応により NADH に変換され、結果として NADH の酸化を伴う 1-ブタノール生合成が促進された (Fig. 2.7)。この結果から、1-ブタノールの生産性向上

につながる戦略として、NADH の酸化を伴う反応経路を阻害することに加えて、NADPH を Transhydrogenase 反応を通じて NADH に変換する代謝改変戦略が提案された。

先行研究において、1-ブタノールを高生産する大腸菌の育種が行われている (Atsumi *et al.*, 2008)。この文献では、*Clostridium* 属由来の 1-ブタノール生合成遺伝子を発現させた大腸菌に対して、*adhE*、*ldhA*、*frd*、*fnr*、*pta* 遺伝子を同時に破壊することで、1-ブタノールの生産収率が 2.8 倍に上昇した。このうち、*adhE* 遺伝子と *pta* 遺伝子の破壊については、本章での同定された破壊候補遺伝子と一致しており、*ldhA*、*frdA* 遺伝子がそれぞれ関わる乳酸生合成とコハク酸生合成については、いずれもそのフラックスが 0 だと予測されている。したがって、本章での破壊候補遺伝子の探索結果は、先行研究の実験結果と一致しているといえる。また、*adhE* と *pta* は、高い生産収率が予測される上位 10 組の破壊候補遺伝子に共通している遺伝子あることから、これらの遺伝子破壊による代謝改変戦略が 1-ブタノールの生産性向上に対して適切なアプローチであることが、代謝モデルを用いた解析からも裏付けされた。さらに *tpiA* や *nuoN* は先行研究での報告がない新規破壊候補遺伝子であり、本章で提案する手法により初めて見つけ出された。なお、*fnr* 遺伝子は転写因子をコードし、代謝モデルには含まれていない遺伝子であるため、本章の手法により破壊候補遺伝子として同定されることはない遺伝子である。

グリセロールを単一炭素源とした場合では、すべての破壊候補遺伝子に *adhE* 遺伝子が含まれていた (Table 2.10)。グルコース炭素源の場合と同様に、*adhE* 遺伝子が破壊されると、エタノールの生合成による NADH の酸化が阻害され、解糖で生じた NADH を酸化するため、NADH の酸化を伴う 1-ブタノール生産が促進されると予測された。また、*pflB* 遺伝子あるいは *focA* 遺伝子のどちら一方も、*adhE* と同様にすべての破壊候補遺伝子に含まれていた。*pflB* は Pyruvate formate lyase 反応をコードし、*focA* はギ酸のトランスポーターをコードする。*pflB* あるいは *focA* を破壊すると、ピルビン酸からアセチル CoA へは Pyruvate dehydrogenase 反応により変換される。この反応は NAD⁺ の還元を伴うため、生じた NADH を酸化するために 1-ブタノール生産が促進された。*adhE* 遺伝子と、*focA* あるいは *pflB* 遺伝子に加え、電子伝達系の NADH dehydrogenase 反応をコードする *nuoN* 遺伝子が、1-ブタノール生産性向上に寄与する破壊候補遺伝子として同定された。*adhE*、*pflB*、*nuoN* の 3 重破壊に伴う代謝フラックス分布の変動を Fig. 2.8 に示す。*nuoN* の破壊により、*ndh* 遺伝子がコードするもう一つの NADH dehydrogenase 反応が活性化するが、この反応は NADH の酸化に伴うプロトンの排出量が小さい。そのため、細胞の増殖速度が低下し、アミノ酸合成に必要な還元力が減少する。したがって、*nuoN* 遺伝子の破壊により細胞の増殖速度を減少させ、余剰の還元力を利用して 1-ブタノール生産を活性化させる戦略が予測された。

2.2 で述べたように、本研究のモデル縮約および探索手法は、探索時間のみを短縮するものであり、もとのモデルを用いて探索したとしても、時間をかければ同じ破壊候補遺伝子が同定される。そのため、以下の Table 2.10-12 で示すターゲットの生産収率以上が予測される破壊候補遺伝子は、同じ代謝モデルや栄養条件等を使用する限り、3 遺伝子以下の破壊

では存在しない。その点で、局所的探索における破壊候補遺伝子を取りこぼす可能性がないことが本章の大域的探索の特徴である。

Table 2.10 1-ブタノール生産収率の向上が予測された上位 10 組の破壊候補遺伝子

炭素源	破壊候補遺伝子			収率 (g/g-substrate) (理論最大に対する%)	比増殖速度 (1/h) (野生型に対する%)
グルコース	<i>adhE</i>	<i>pta</i>	<i>tpiA</i>	0.303 (73.5)	0.246 (61.7)
	<i>adhE</i>	<i>pta</i>	<i>cyoA</i>	0.288 (70.0)	0.294 (73.7)
	<i>adhE</i>	<i>pta</i>	<i>nuoN</i>	0.286 (69.6)	0.299 (75.0)
	<i>adhE</i>	<i>pta</i>	<i>gdhA</i>	0.280 (68.2)	0.319 (79.9)
	<i>adhE</i>	<i>pta</i>	<i>ptsG</i>	0.280 (68.0)	0.321 (80.5)
	<i>adhE</i>	<i>pta</i>	<i>pgi</i>	0.278 (67.6)	0.326 (81.9)
	<i>adhE</i>	<i>pta</i>	<i>serC</i>	0.277 (67.4)	0.329 (82.4)
	<i>adhE</i>	<i>pta</i>	<i>glyA</i>	0.275 (66.9)	0.335 (84.1)
	<i>adhE</i>	<i>pta</i>	<i>ppk</i>	0.275 (66.8)	0.337 (84.4)
	<i>adhE</i>	<i>pta</i>	<i>fold</i>	0.275 (66.8)	0.337 (84.6)
	(None)		0.000 (0.0)	0.399 (100.0)	
グリセロール	<i>adhE</i>	<i>focA</i>	<i>nuoN</i>	0.379 (80.8)	0.196 (56.2)
	<i>adhE</i>	<i>focA</i>	<i>glpK</i>	0.372 (79.1)	0.223 (63.7)
	<i>adhE</i>	<i>focA</i>	<i>gpsA</i>	0.372 (79.1)	0.223 (63.7)
	<i>adhE</i>	<i>focA</i>	<i>cyoA</i>	0.371 (79.1)	0.224 (63.9)
	<i>adhE</i>	<i>pflB</i>	<i>ppk</i>	0.369 (78.7)	0.230 (65.8)
	<i>adhE</i>	<i>focA</i>	<i>gdhA</i>	0.366 (78.0)	0.241 (68.8)
	<i>adhE</i>	<i>focA</i>	<i>ppk</i>	0.358 (76.3)	0.270 (77.1)
	<i>adhE</i>	<i>pflB</i>	<i>nuoN</i>	0.358 (76.2)	0.167 (47.7)
	<i>adhE</i>	<i>focA</i>	<i>rpe</i>	0.357 (76.0)	0.273 (78.1)
	<i>adhE</i>	<i>focA</i>	<i>ybcF</i>	0.356 (75.9)	0.275 (78.7)
	(None)		0.000 (0.0)	0.349 (100.0)	

遺伝子がコードする反応は、付録の Table A1 参照。

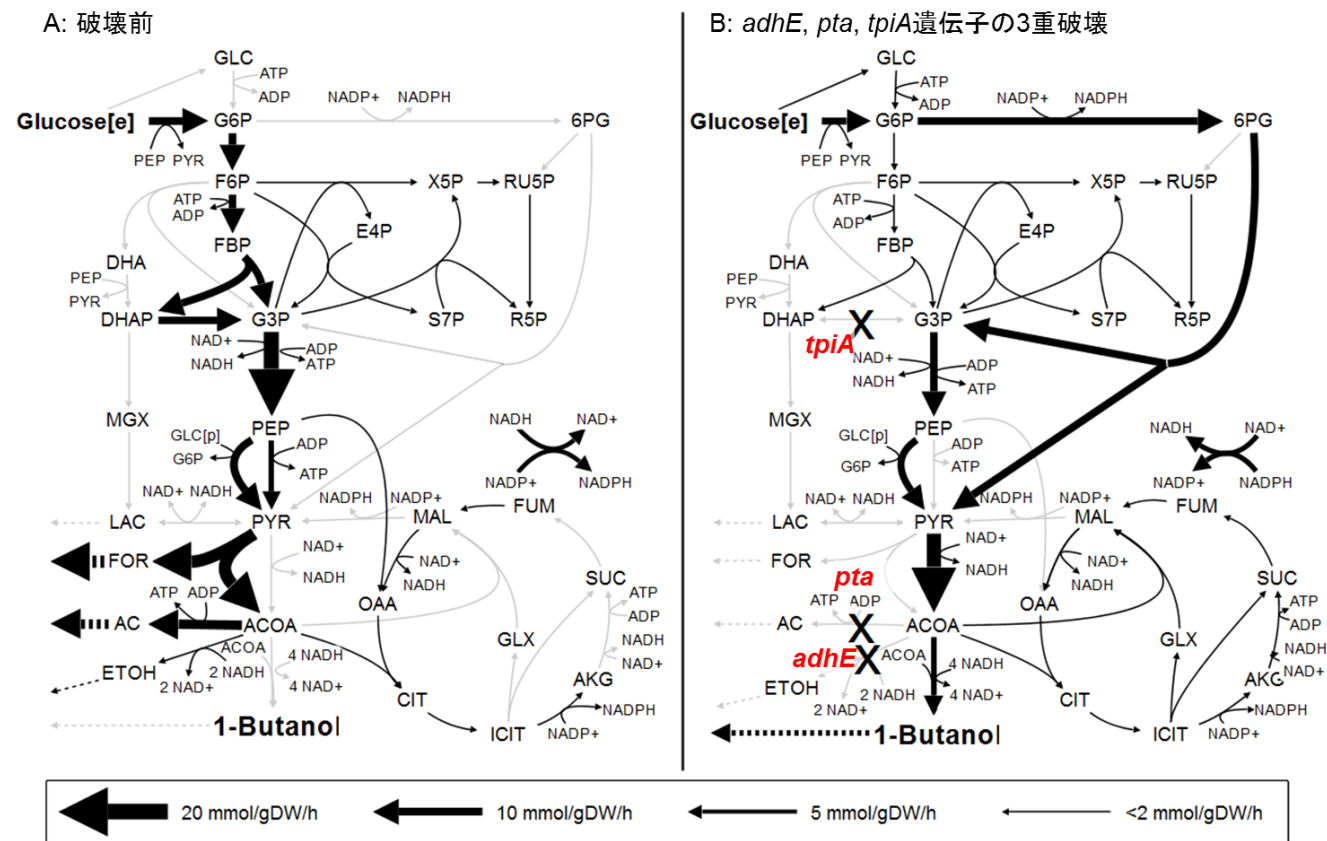


Fig. 2.7 グルコース炭素源下での野生型 (A) および 1-ブタノール生産に有望な 3 重遺伝子破壊 (B) に対応する代謝フラックス分布

矢印の太さはフラックスの大きさを示し、灰色の矢印はフラックスが 0 であることを意味している。実線の矢印は代謝反応を意味し、破線の矢印はトランスポート反応を意味している。×印は遺伝子破壊による反応の不活性化を意味している。代謝物質の略称については、付録の Table A2 参照。*adhE*, *pta*, *tpiA* 遺伝子の 3 重破壊により、消費されたグルコースは酸化的ペントースリン酸経路と ED 経路を通じて異化され、生じた還元力を用いて 1-ブタノール生産が促進することが予測された。

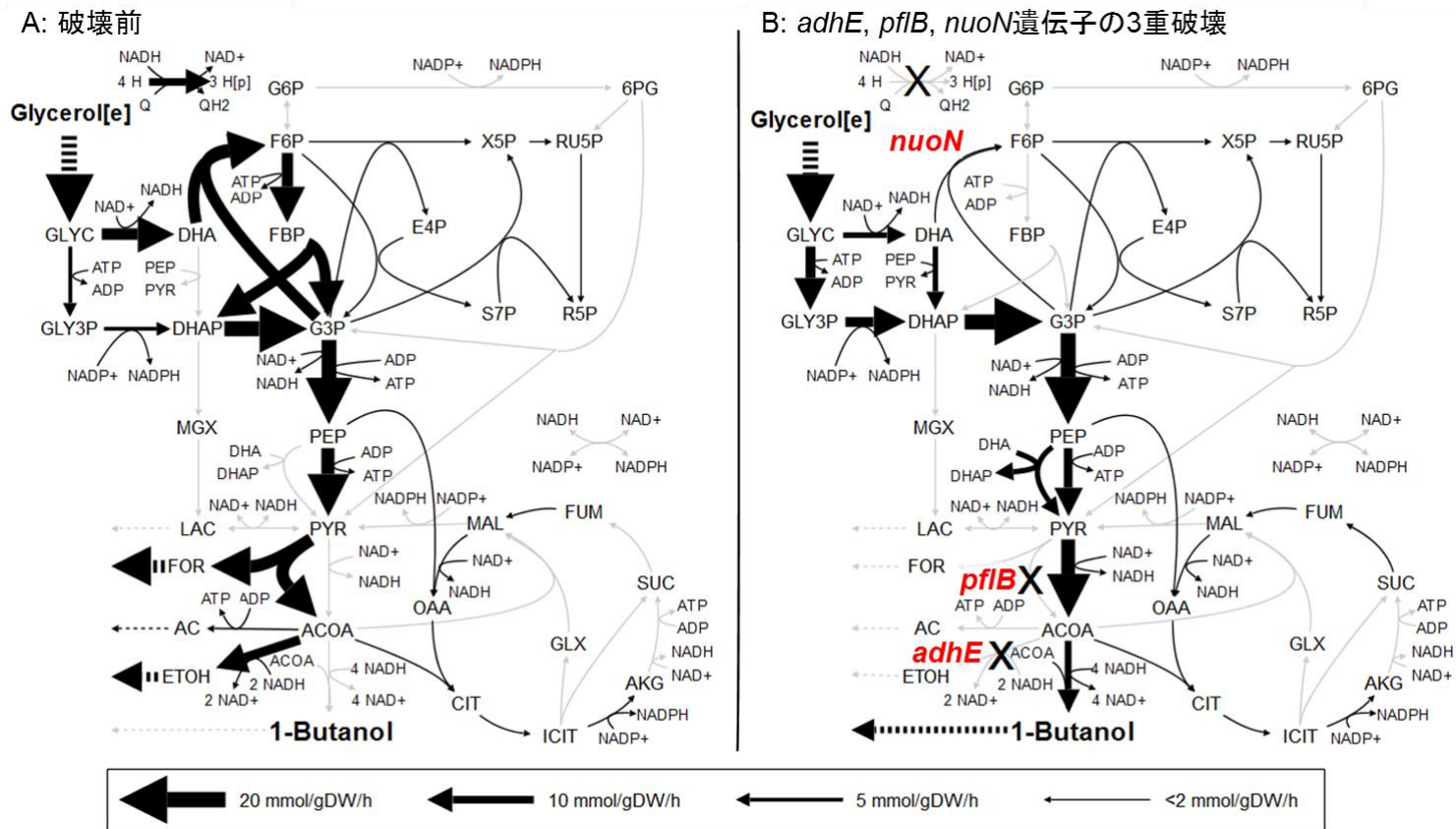


Fig. 2.8 グリセロール炭素源下での野生型 (A) および 1-ブタノール生産に有望な 3 重遺伝子破壊 (B) に対応する代謝フラックス分布

凡例は Fig. 2.7 と同様である。*adhE*, *pflB*, *nuoN* 遺伝子の 3 重破壊により、グリセロール異化によって生じた NADH を酸化するために、NADH の酸化を伴う 1-ブタノール生産が促進されると予測された。

2.3.4 大腸菌による 1-プロパノール生産への本手法の適用

次に、大腸菌による 1-プロパノール生産に対して開発手法を適用し、生産収率向上に有望な破壊候補遺伝子を探索した。1-ブタノール生産の場合と同様に、大腸菌のゲノムスケール代謝モデルに *Klebsiella pneumoniae* 由来の反応 (Table 2.3) を追加し、グリセロールから 1-プロパノールへの変換経路を付加した。1-プロパノールのトランスポート反応も同様に代謝モデルに追加した。グルコース炭素源下では、代謝モデルに 1-プロパノール生合成経路を追加しただけでは、細胞の増殖速度を最大化する条件下で 1-プロパノールは生産されなかった。一方で、グリセロール炭素源下では 1-プロパノール生合成経路を追加しただけで、0.200 g/g-glycerol の収率での 1-プロパノール生産が予測された。グリセロール炭素源微好気条件下での野生型の代謝モデルのフラックス分布では、解糖で生じた NADH を再酸化するために、エタノール合成が促進される。ここで、プロパノール生合成経路を代謝モデルに追加すると、1-プロパノール合成の方がエタノール合成よりも NADH の酸化効率が高いため、一部のグリセロールより 1-プロパノールの生産が予測された (Fig. 2.10 A)。

グルコースを炭素源として、1-プロパノールの高い生産収率が予測される 3 重遺伝子破壊を探索した結果、予測された 10 組の破壊候補遺伝子の組み合わせ全てに、*tpiA* 遺伝子が含まれていた (Table 2.11)。*tpiA* 遺伝子は、解糖系中のジヒドロキシアセトンリン酸 (DHAP) と グリセルアルデヒド 3 リン酸 (G3P) の変換反応をコードしている。この反応が不活性化すると、DHAP は NADH の酸化を伴いながら、5 反応を経て 1-プロパノールへ変換され、G3P は ATP や菌体構成成分 (アミノ酸など) を合成するためにピルビン酸等に異化された。DHAP から 1-プロパノールへの変換は、G3P 異化で生じる NADH の酸化還元バランスを維持するために促進された。*tpiA* 遺伝子に加えて、ギ酸の生合成反応をコードする *pfB* 遺伝子、あるいはギ酸のトランスポーターをコードする *focA* 遺伝子が、1-プロパノール生産に向けた破壊候補遺伝子として見つけ出された (生産収率の高い 10 組の破壊候補遺伝子のうち 6 組に含まれていた)。ギ酸生産に関わるこれらの遺伝子破壊により、ピルビン酸からアセチル CoA へ変換するために、NADH 生産を伴う Pyruvate dehydrogenase 反応が活性化される。その結果、NADH の酸化を伴う 1-プロパノール生産が増加した。

1-プロパノールの最大生産収率は、上記した 2 つの遺伝子破壊に加え、Phosphoenolpyruvate (PEP) carboxylase をコードする *ppc* 遺伝子の破壊により達成された (Fig. 2.9)。この反応は解糖系から TCA サイクルへの補充反応である。*ppc* 遺伝子が破壊されると、グリオキシル酸回路が活性化しアミノ酸合成に必要なオキサロ酢酸が供給される。このオキサロ酢酸は、NADH 生産を伴う Malate dehydrogenase 反応を通じて合成されるため、*ppc* 遺伝子破壊により 1-プロパノール生産が促進された。また、*zwf*、*adhE*、*cyoA* 遺伝子も、1-プロパノール生産収率向上に有望な破壊候補遺伝子として同定された。*zwf* 遺伝子が破壊されると、グルコース 6 リン酸から酸化的ペントースリン酸経路へのフラックスが阻害され、解糖系のフラックスが増加する。そのため、解糖系での生産が上昇した NADH を酸化するため、1-プロパノール合成が促進された。

グリセロールを単一炭素源とした場合、*tpiA* 遺伝子はグルコース炭素源の場合と同様に、生産収率の高い破壊候補遺伝子の全ての組み合わせに含まれていた (Table 2.11)。*tpiA* 遺伝子が破壊されると、グリセロール由来の炭素は解糖系の下流に直接流れることができなくなり、代わりに酸化のペントースリン酸経路と Entner-Doudoroff (ED) 経路を通じてピルビン酸等が合成される。酸化のペントースリン酸経路のフラックスが増加すると、NADPH 生産が上昇する。その結果、NADPH を Transhydrogenase 反応で NADH に変換し、NADH を酸化するために 1-プロパノール生産が増加するような代謝フラックス分布が予測された。

tpiA 遺伝子の破壊に加えて、酢酸合成に関わる *pta* 遺伝子の破壊が、破壊候補遺伝子として見つけ出された (生産収率の高い破壊候補遺伝子 10 組のうち 5 組に含まれていた)。酢酸合成では ATP 生産を伴うため、この合成反応を阻害すると TCA サイクル中の Succinyl-CoA synthase による ATP 合成が促進される。その結果、TCA サイクルにおける NADH 生産が上昇し、NADH 酸化のために 1-プロパノール生産が上昇した。

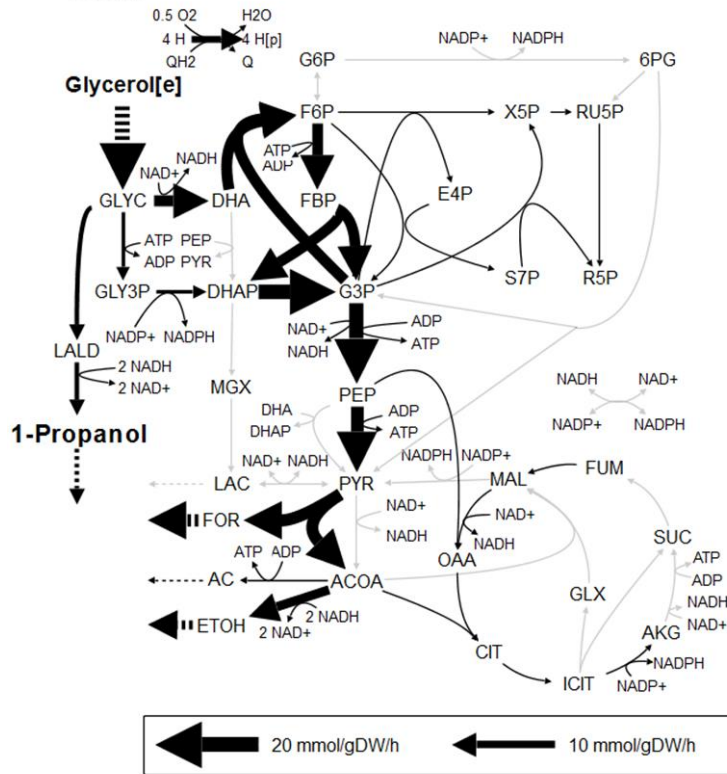
tpiA と *pta* の破壊に加えて、*cyoA* 遺伝子を破壊した場合に、最大の 1-プロパノール生産収率が予測された。*cyoA* 遺伝子は電子伝達系の Cytochrome oxidase 反応をコードする遺伝子である。*cyoA* 遺伝子を破壊すると、*cydA* 遺伝子がコードするもう一つの Cytochrome oxidase 反応が活性化される。*cydA* がコードするこの反応も、還元型キノンを酸化する際に細胞質のプロトンのトランスポートを伴うが、その効率は *cyoA* がコードする反応より低い。そのため、*cyoA* 遺伝子の破壊により、プロトン勾配を利用した ATP 合成が減少し、Succinyl-CoA synthase を介した ATP 合成がより促進される。その結果、*pta* 破壊と同様に TCA サイクルにおける NADH 生産が上昇し、NADH 酸化のために 1-プロパノール生産が上昇した (Fig. 2.10)。

Table 2.11 1-プロパノール生産収率の向上が予測された上位 10 組の破壊候補遺伝子

炭素源	破壊候補遺伝子			収率 (g/g-substrate) (理論最大に対する%)	比増殖速度 (1/h) (野生型に対する%)
グルコース	<i>tpiA</i>	<i>pflB</i>	<i>ppc</i>	0.173 (41.1)	0.287 (72.0)
	<i>tpiA</i>	<i>focA</i>	<i>ppc</i>	0.173 (41.0)	0.287 (72.1)
	<i>tpiA</i>	<i>pflB</i>	<i>adhE</i>	0.161 (38.1)	0.229 (57.5)
	<i>tpiA</i>	<i>focA</i>	<i>adhE</i>	0.160 (38.0)	0.230 (57.6)
	<i>tpiA</i>	<i>ptsG</i>	<i>zwf</i>	0.159 (37.7)	0.237 (59.5)
	<i>tpiA</i>	<i>pgi</i>	<i>zwf</i>	0.159 (37.7)	0.237 (59.5)
	<i>tpiA</i>	<i>ptsG</i>	<i>pykF</i>	0.158 (37.5)	0.244 (61.3)
	<i>tpiA</i>	<i>pflB</i>	<i>cyoA</i>	0.157 (37.2)	0.251 (63.0)
	<i>tpiA</i>	<i>fsaA</i>	<i>zwf</i>	0.157 (37.2)	0.249 (62.6)
	<i>tpiA</i>	<i>focA</i>	<i>cyoA</i>	0.156 (37.1)	0.251 (63.1)
	(None)		0.000 (0.0)	0.399 (100.0)	
グリセロール	<i>tpiA</i>	<i>pta</i>	<i>cyoA</i>	0.440 (87.6)	0.097 (27.8)
	<i>tpiA</i>	<i>aldA</i>	<i>edd</i>	0.434 (86.6)	0.117 (33.4)
	<i>tpiA</i>	<i>aceF</i>	<i>edd</i>	0.434 (86.5)	0.119 (33.9)
	<i>tpiA</i>	<i>pta</i>	<i>edd</i>	0.432 (86.1)	0.122 (34.8)
	<i>tpiA</i>	<i>pta</i>	<i>adhE</i>	0.431 (86.0)	0.124 (35.3)
	<i>tpiA</i>	<i>aceF</i>	<i>pflB</i>	0.431 (85.9)	0.127 (36.4)
	<i>tpiA</i>	<i>aceF</i>	<i>focA</i>	0.431 (85.9)	0.127 (36.4)
	<i>tpiA</i>	<i>pta</i>	<i>gdhA</i>	0.427 (85.0)	0.138 (39.5)
	<i>tpiA</i>	<i>pta</i>	<i>ppk</i>	0.424 (84.5)	0.146 (41.7)
	<i>tpiA</i>	<i>pta</i>	<i>aceA</i>	0.424 (84.5)	0.146 (41.8)
	(None)		0.200 (39.9)	0.350 (100.3)	

遺伝子がコードする反応は、付録の Table A1 参照

A: 破壊前



B: *tpiA*, *pta*, *cyoA* 遺伝子の3重破壊

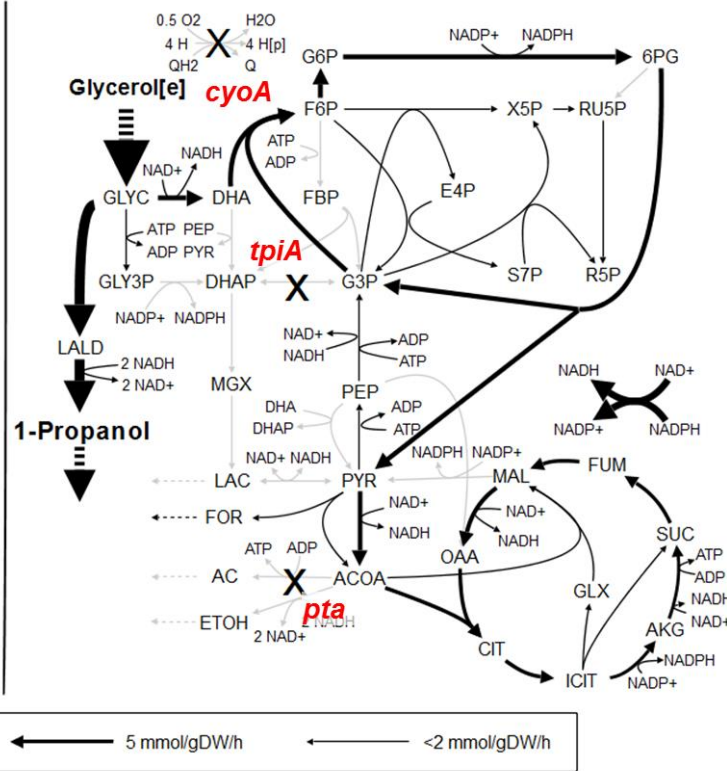


Fig. 2.10 グリセロール炭素源下での野生型 (A) および 1-プロパノール生産に有望な 3 重遺伝子破壊 (B) に対応する代謝フラックス分布
 凡例は、Fig. 2.7 と同様である。 *tpiA*, *pta*, *cyoA* 遺伝子の破壊により、取り込まれたグリセロールは酸化的ペントースリン酸経路と ED 経路を通じて異化された。また TCA サイクルのフラックスが上昇した。ペントースリン酸経路と TCA サイクルで生じた還元力を利用して、1-プロパノール生産が促進された。

2.3.5 大腸菌による 1,3-プロパンジオール (13PD) 生産への本手法の適用

最後に、大腸菌による 13PD 生産に対して開発手法を適用し、生産収率向上に有望な破壊候補遺伝子を探索した。1-ブタノール、1-プロパノールの場合と同様に、大腸菌のゲノムスケール代謝モデルに *K. pneumoniae* 由来の 13PD 生合成反応および 13PD トランスポート反応 (Table 2.4) を追加した。1-ブタノールの場合と同様に、グルコースとグリセロールのいずれの炭素源下でも、代謝モデルにこれらの反応を追加しただけでは、増殖速度最大化の条件下で 13PD が生産される代謝フラックス分布は計算されなかった。次に、13PD 生産に向けた 3 重遺伝子破壊を探索した。探索結果を Table 2.12 に示す。

グルコース炭素源の場合は、13PD 高い生産収率が予測された 10 組の 3 重遺伝子破壊の全てに、*tpiA* 遺伝子の破壊が共通して含まれていた。1-プロパノールの場合と同様に、*tpiA* 遺伝子が破壊されると DHAP から GAP への変換が阻害される。結果として DHAP から解糖系下流への代謝フラックスが減少し、生じた DHAP のほぼすべてが 13PD に変換された。

tpiA 遺伝子の破壊に加えて、*pykF* 遺伝子および *ptsG* 遺伝子の破壊により、最大の 13PD 生産収率が予測された (Fig. 2.11)。*pykF*、*ptsG* はいずれも PEP からピルビン酸への変換に関わる。これらの遺伝子を破壊すると、PEP からピルビン酸への変換は、TCA サイクルの補充経路を通じて、オキサロ酢酸とリンゴ酸を経由して変換された。このピルビン酸への変換経路は NADPH の生産を伴うので、NADPH の酸化を伴う 13PD 生産が促進された。また、先行研究より、*tpiA* の破壊と、*ptsG* 遺伝子がコードする Phosphotransferase system の発現抑制が、実際の大腸菌においても 13PD 生産に有効であることが報告されている (Nakamura and Whited, 2003)。残りの破壊候補遺伝子については、1-ブタノールや 1-プロパノール生産の場合とよく似ていた。

グルコース炭素源での 13PD 生産では、可能な全ての 1 遺伝子破壊の中で、13PD 生産収率が向上するような破壊は予測されなかった。また、2 重遺伝子破壊では *tpiA*、*adhE* が予測されたが、その結果から新たに 1 遺伝子を破壊しても、最大の 13PD 生産収率が推測された *tpiA*、*pykF*、*ptsG* の 3 重遺伝子破壊は導かれなかった。したがって、グルコースからの 13PD 生産に有望な破壊候補遺伝子を探索する際には、少なくとも同時に 2 遺伝子を破壊した際の代謝状態を FBA でシミュレートする必要がある。したがって、先行研究 (Alper *et al.*, 2005) で実施された単純な局所的探索手法では、本章で見つけ出された *tpiA*、*pykF*、*ptsG* の 3 重破壊を見つけ出すことはできず、大域的に探索することで初めて見つけ出すことができた。

グリセロール炭素源の場合は、*tpiA* 遺伝子と *edd* 遺伝子が、13PD 生産収率の高い破壊候補遺伝子の組み合わせ全てに含まれていた。*tpiA* 遺伝子は DHAP と G3P との変換反応をコードし、*edd* 遺伝子は ED 経路中の 6-Phosphogluconate dehydratase をコードする。*tpiA* と *edd* 遺伝子の 2 重破壊により、予測される 13PD 生産収率は 0.633 g/g-glycerol であり、3 重破壊での最大収率 (0.636 g/g-glycerol) とほぼ変わらなかった。*tpiA* 遺伝子破壊のみでは、消費されたグリセロールの一部はジヒドロキシアセトンに変換され、その後に ED 経路を通じて解

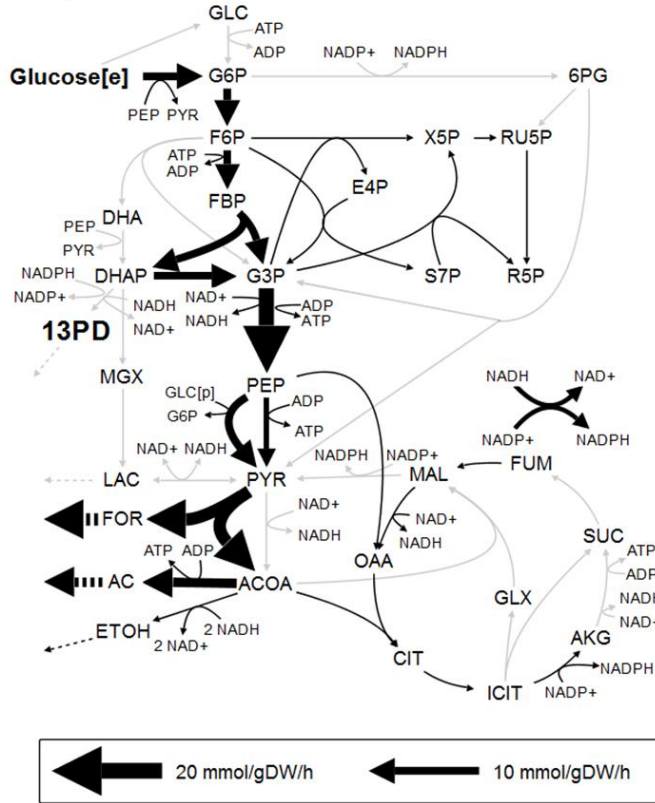
糖系下流の代謝物質に変換された。ここで、*edd* 破壊を加えると、一定量のグリセロールはメチルグリオキサール経路を通じてピルビン酸に変換され、そして TCA サイクルやアセチル CoA 由来の菌体構成成分が合成され、余剰のグリセロールが 13PD とペントースリン酸経路中の代謝物質に変換された (Fig. 2.12)。酸化のペントースリン酸経路で生じる NADPH は、Transhydrogenase 反応により NADH に変換され、13PD 生合成に伴う NADH の酸化とバランスがとられていた。なお、*tpiA* と *edd* の 2 重破壊は、グルコース炭素源における *ptsG*、*pykF* の 2 重破壊と同様に、同時に破壊することが 13PD 生産収率の向上に必要だった。

Table 2.12 1,3-プロパンジオール生産収率の向上が予測された上位 10 組の破壊候補遺伝子

炭素源	破壊候補遺伝子			収率 (g/g-substrate) (理論最大に対する%)	比増殖速度 (1/h) (野生型に対する%)
グルコース	<i>tpiA</i>	<i>ptsG</i>	<i>pykF</i>	0.213 (36.6)	0.217 (54.6)
	<i>tpiA</i>	<i>adhE</i>	<i>pflB</i>	0.123 (21.2)	0.268 (67.1)
	<i>tpiA</i>	<i>adhE</i>	<i>focA</i>	0.123 (21.1)	0.268 (67.2)
	<i>tpiA</i>	<i>adhE</i>	<i>ppc</i>	0.115 (19.8)	0.286 (71.8)
	<i>tpiA</i>	<i>adhE</i>	<i>nuoN</i>	0.091 (15.6)	0.232 (58.3)
	<i>tpiA</i>	<i>adhE</i>	<i>cyoA</i>	0.086 (14.8)	0.255 (63.9)
	<i>tpiA</i>	<i>adhE</i>	<i>rpe</i>	0.083 (14.4)	0.305 (76.4)
	<i>tpiA</i>	<i>adhE</i>	<i>gdhA</i>	0.080 (13.7)	0.283 (71.0)
	<i>tpiA</i>	<i>adhE</i>	<i>talA</i>	0.079 (13.6)	0.307 (77.0)
	<i>tpiA</i>	<i>adhE</i>	<i>ptsG</i>	0.079 (13.6)	0.287 (72.1)
	(None)			0.000 (0.0)	0.399 (100.0)
グリセロール	<i>tpiA</i>	<i>edd</i>	<i>gdhA</i>	0.636 (89.8)	0.079 (22.7)
	<i>tpiA</i>	<i>edd</i>	<i>mgsA</i>	0.636 (89.7)	0.080 (22.8)
	<i>tpiA</i>	<i>edd</i>	<i>ldhA</i>	0.635 (89.6)	0.082 (23.3)
	<i>tpiA</i>	<i>edd</i>	<i>ppk</i>	0.634 (89.5)	0.083 (23.8)
	<i>tpiA</i>	<i>edd</i>	<i>serC</i>	0.634 (89.5)	0.083 (23.8)
	<i>tpiA</i>	<i>edd</i>	<i>ppsA</i>	0.634 (89.5)	0.083 (23.8)
	<i>tpiA</i>	<i>edd</i>	<i>glyA</i>	0.634 (89.5)	0.084 (24.0)
	<i>tpiA</i>	<i>edd</i>	<i>folD</i>	0.634 (89.5)	0.084 (24.1)
	<i>tpiA</i>	<i>edd</i>	<i>mdh</i>	0.634 (89.4)	0.084 (24.1)
	<i>tpiA</i>	<i>edd</i>	<i>ybcF</i>	0.634 (89.4)	0.085 (24.2)
	(None)			0.000 (0.0)	0.349 (100.0)

遺伝子がコードする反応は、付録の Table A1 参照

A: 破壊前



B: *tpiA*, *ptsG*, *pykF* 遺伝子の3重破壊

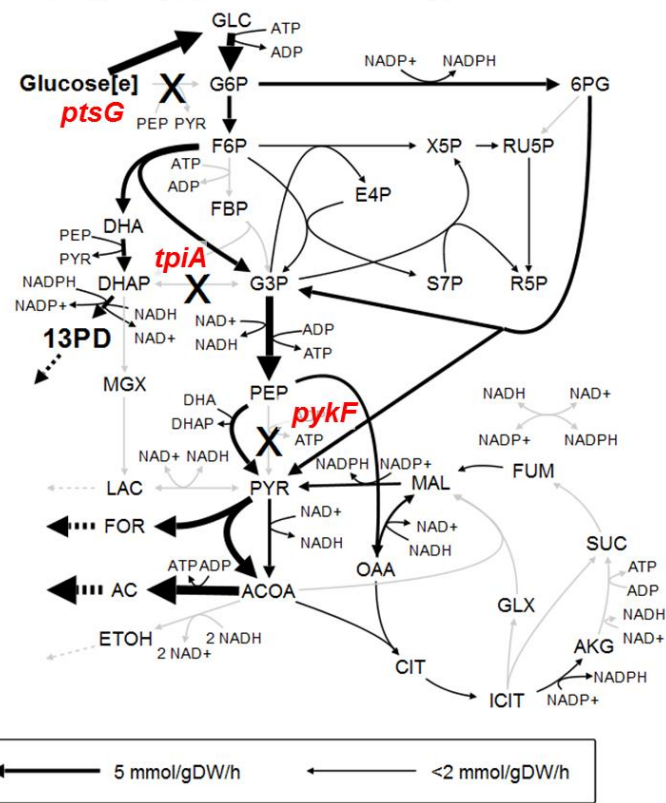


Fig. 2.11 グルコース炭素源下での野生型 (A) および 1,3-プロパンジオール (13PD) 生産に有望な 3 重遺伝子破壊 (B) に対応する代謝フラックス分布

凡例は Fig. 2.7 と同様である。 *tpiA*、*ptsG*、*pykF* 遺伝子の破壊により、グルコース異化により生じる DHAP と G3P は分断され、PEP からピルビン酸への変換を TCA サイクルとその補充経路により行われ、生じた NADPH を利用して 13PD 生産が促進することが予測された。

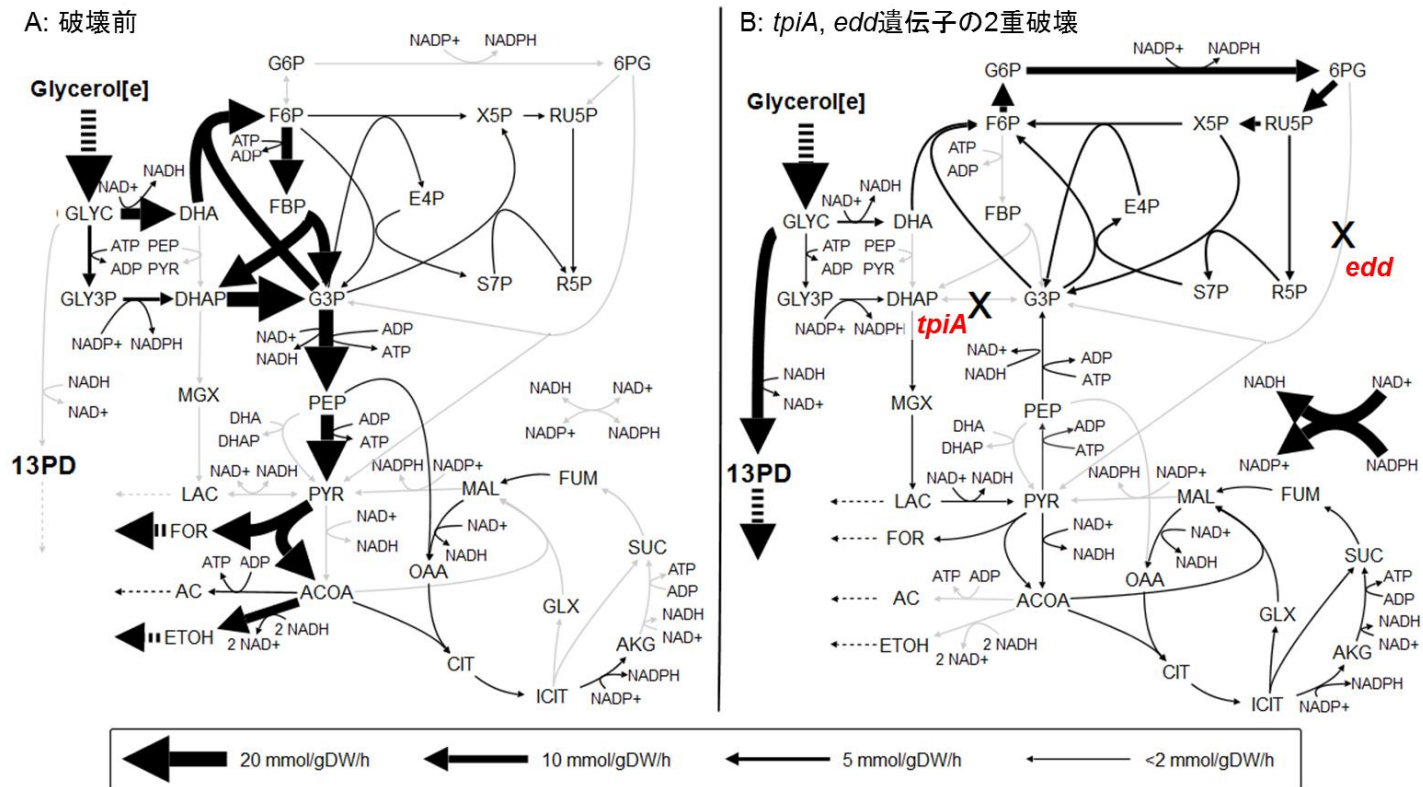


Fig. 2.12 グリセロール炭素源下での野生型 (A) および 1,3-プロパンジオール (13PD) 生産に有望な 3 重遺伝子破壊 (B) に対応する代謝フラックス分布

凡例は Fig. 2.7 と同様である。*tpiA*, *edd* 遺伝子の破壊により、消費されたグリセロールの一部は酸化的ペントースリン酸経路に流れ、その時に生じた NADPH は Transhydrogenase 反応により NADH に変換され、結果として NADH の酸化を伴う 13PD 生産が促進することが予測された。

2.4 考察

本章では、栄養条件に応じて縮約した代謝モデルを自動的に構築し、破壊対象の遺伝子数を削減することで、微生物育種に有望な多重遺伝子破壊を高速かつ大域的に探索する手法を開発した。代謝モデルの縮約では、Flux variability analysis と呼ばれる手法を適用し、与えられた栄養条件下で使用される可能性のない反応およびそれをコードする遺伝子を代謝モデルより除外したり、また計算機上で細胞増殖に必須の遺伝子を破壊対象より除外することで、破壊候補遺伝子が削減された。縮約された代謝モデルの違いを解析したところ、特に通気条件の違いにより、破壊候補遺伝子に 8% 程度の差異が現れることが示され、研究者が専門的知見により破壊対象から除外する遺伝子を、提案する縮約方法で自動的に除外することが可能となった。

従来の単純な局所的探索手法として、1 遺伝子破壊に伴う代謝変動を網羅的に解析し、優れた形質が予測されたものを選抜して、新たに 1 遺伝子を破壊した時の代謝変動を解析する、といった手法が開発されていた。しかし、そのような局所的探索手法では、複数遺伝子の同時破壊がターゲット生産に影響する場合に、探索が困難となる。対して、本手法は、多重遺伝子破壊がターゲット生産に及ぼす影響を網羅的に解析することが可能である。例えば、本章のグルコース炭素源における 13PD 生産では、可能な全ての 1 遺伝子破壊の中で、13PD 生産収率が向上するものは予測されなかった。また、2 重遺伝子破壊では *tpiA*、*adhE* が予測されたが、その結果から新たに 1 遺伝子を破壊しても、最大の 13PD 生産収率が推測された *tpiA*、*pykF*、*ptsG* の 3 重遺伝子破壊は導かれなかった。したがって、グルコースからの 13PD 生産に有望な破壊候補遺伝子を探索する際には、少なくとも同時に 2 遺伝子を破壊した際の代謝状態をシミュレートする必要がある。本章で開発した探索手法では、そのような破壊候補遺伝子の予測にも成功している。

本章では、縮約した代謝モデルを自動的に構築し、破壊対象となる遺伝子数を栄養条件に応じて減少させ、破壊候補遺伝子の探索を高速化する手法を提案した。探索の高速化を目指す場合には、代謝モデルをいくつかのモジュールに分割し、部分問題として取り扱う方法も考えられる。しかし、代謝モデル中の代謝物質をノード、反応をエッジとした代謝ネットワークは、スケールフリー性を持つことが報告されており (Jeong *et al.*, 2000)、一部の代謝物質が多くの代謝反応に関与していることが知られている。例えば、本章で使用した大腸菌の代謝モデルでは、全 2382 反応中、ATP の関与する反応は 339 反応、NADH の関与する反応は 109 反応存在し、解糖系・TCA サイクル・アミノ酸合成経路など、様々な代謝経路に関わっている。これらの物質は、FBA で予測される代謝フラックス分布に大きく影響するため、現状では代謝モデルを適切にモジュール分割することは難しい。本章で提案した手法は、そのような代謝ネットワークを適切に縮約することで、破壊候補遺伝子の探索を高速化している。

また、本研究の手法では破壊候補遺伝子の複数の組み合わせを同時に示すことができる。そのため、予測された遺伝子破壊による生産性向上の実現性を議論することが可能である。

例えば、大腸菌によるグリセロール炭素源での 13PD 生産では *tpiA* 遺伝子と *edd* 遺伝子が、高い生産収率に必要な破壊候補遺伝子として共通して予測された。したがって、これらの遺伝子破壊は、ターゲットとする 13PD の生産性向上に特に有望な遺伝子破壊であり、実際の細胞内でのその遺伝子破壊の効果を優先的に検証されるべきであることが示唆されている。

本章の大域的探索手法では、5 遺伝子程度までは現実的な時間以内で破壊候補遺伝子を探索することが可能だと推測され、それ以上の同時破壊を探索するのは現実的でないと思われる。しかし、5 遺伝子以下の破壊によりターゲット化合物の生産収率向上が見込まれる破壊候補遺伝子が予測されれば、代謝モデルに対してその遺伝子破壊を実行し、追加する遺伝子破壊の探索を行うことが有効である。この場合には、有望な多重遺伝子破壊を取りこぼす可能性があるが、育種に向けた遺伝子破壊の探索における、破壊遺伝子数の上限を上昇させることは可能である。従来の局所的探索手法と比較し、1 度の探索における破壊遺伝子数は大きいため、より広い範囲の探索空間から、効率よく破壊候補遺伝子が見つげ出されることが期待される。

本章の手法は、化学量論式からなる代謝モデルそのものの情報に基づいてモデルを縮約し、探索を高速化した。そのため、本章の手法は異なる生物種、ターゲット化合物、探索指標を用いる場合にも汎用的に使用することが可能である。これまでに様々な生物種について代謝モデルが構築されており (Monk *et al.*, 2014)、それらのモデルをベースとして探索を行えば、他の生物種での有用化合物生産に向けた破壊候補遺伝子を探索できる。また、ターゲット化合物の生合成経路が同定されていれば、その生合成経路をもとの代謝モデルに追加 (実験的には生合成経路に関連する遺伝子を発現) することで、本章と異なるターゲット化合物の生産に向けて探索を行える。また、生産収率とは異なる探索指標として、細胞の増殖速度と生産収率の積を使用することも提唱されており (Patil *et al.*, 2005; Rocha *et al.*, 2008)、本章の手法はそのような探索指標を利用することもできる。このように、様々な事例に対して拡張できる点は、本章の探索指標の特徴であり、既存の OptKnock (Burgard *et al.*, 2003) などの大域的探索手法では容易にはできないことである。

代謝シミュレーション手法である FBA では、代謝状態の定常と目的関数の最大化により代謝フラックス分布が計算され、代謝物質濃度、酵素の速度論的パラメータ、遺伝子発現の制御等の情報は使用されない。そのため FBA は、速度論的なシミュレーションと比較し、使用するパラメータが少なく、計算が高速であるという利点がある。一方で FBA は、遺伝子発現制御や、代謝の中間物質の蓄積の影響を考慮していないという限界があり、予測された代謝変動が実際の細胞内で実現されない可能性もある。例えば、本章では *adhE*、*pta*、*tpiA* の 3 重遺伝子破壊の結果、ED 経路が活性化することが予測されたが (Fig. 2.7)、ED 経路の酵素の発現は、転写因子 Cra によって抑制されていることが知られている (Ramseier, 1996)。本章の代謝シミュレーションでは、そのような転写因子の影響を考慮していないため、予測された 3 重破壊により実際の細胞内で ED 経路が本当に活性化するかを断言するこ

とはできない。このような場合には、遺伝子破壊の探索を本章の手法で行った後に、遺伝子発現制御の情報を文献やデータベースより取得し、また代謝物質の蓄積を実験的に定量し、それらの影響を回避するように遺伝子操作を新たに加えることで、工学的に有用な株を獲得することが期待される。

2.5 結言

育種に向けた多重遺伝子破壊の大域的探索手法を開発した。栄養条件に応じた代謝モデルの縮約により、探索の高速化に成功し、例えば大腸菌のゲノムスケール代謝モデルを使用した時の3重破壊の探索時間に関しては、 1.4×10^7 秒から 2.3×10^2 秒に短縮することに成功した。また、開発した探索手法を、大腸菌による 1-ブタノール、1-プロパノール、13PD の生産に対して適用し、生産収率の大きな上昇が見込まれる複数の3重遺伝子破壊を探索した。結果、同時に2遺伝子以上の同時破壊により、高い生産収率が達成されるような破壊候補遺伝子の組み合わせについても、本手法で同定することに成功した。これは、単純な局所的探索手法では、破壊候補遺伝子の同定が難しいものである。予測された破壊候補遺伝子には、既に実験的な検証がなされているものあり、実際の代謝改変戦略が妥当であることを、代謝モデルを用いた解析より裏付けすることが可能となった。先行研究のない新規破壊候補遺伝子も見つけ出され、実験的な検証が望まれる。さらに、生産収率の上昇が大きい破壊候補遺伝子の組み合わせをリストアップし、リスト中に共通する遺伝子破壊および予測される代謝変動を示すことで、ターゲット化合物生産に向けた代謝改変戦略について議論された。

開発した探索手法では、化学量論式から構成される代謝モデルそのものの情報に基づいてモデルの縮約を行い、探索を高速化したため、様々な宿主微生物、ターゲット化合物等について汎用的に利用することが可能である。今後の育種における、遺伝子破壊による代謝改変戦略を予測する手法として、広く利用されることが期待される。

第3章 育種に向けた多重遺伝子破壊の局所的探索 (FastPros)

3.1 緒言

有用化合物生産を目指した微生物育種において、ターゲットとする化合物生産に最適な代謝フラックス分布を細胞内で実現させることは重要である。その際に、遺伝子破壊より不要な代謝経路を不活性化することが一般に行われる。近年では、ゲノムスケール代謝モデルと FBA による代謝シミュレーションを利用し、増殖速度最大化の条件下で高いターゲット生産収率が予測される破壊候補遺伝子を探索する手法が開発されてきた。そして、予測結果が実験的に確かめられてきた。しかし、同時に破壊する遺伝子の数の増加にしたがって、その組み合わせが指数関数的に増加し、探索における最大の破壊数に限界があることが問題となっている。第 2 章で開発した破壊候補遺伝子の探索手法や、OptKnock などの他の大域的探索手法においても、1000 程度の遺伝子より 6 遺伝子以上を同時に破壊する場合には、現実的な計算時間以内で破壊候補遺伝子を見つけ出すのは極めて困難である。

現実的な計算時間以内で、ターゲット生産に寄与する多重遺伝子破壊を効率よく見つけ出すために、局所的探索手法がいくつか提案されている。例えば、1 遺伝子の破壊に対応する代謝フラックス分布の変動を網羅的にシミュレートした後に、ターゲット化合物の生産収率の高い遺伝子破壊の選択し、選択された遺伝子破壊に加えて新たに破壊を行うことを繰り返す探索手法が挙げられる (Alper *et al.*, 2005; Park *et al.*, 2007)。この単純な局所的探索手法では、同時に破壊する遺伝子数の増加にしたがって、計算時間は線形的に増加するため、計算時間が指数的に増加する大域的探索と比較して非常に高速である。しかし、ターゲット化合物生産に 2 遺伝子以上の同時破壊が必要な場合には、その破壊候補遺伝子の組み合わせを取りこぼすという弱点がある。可能な破壊の組み合わせをより広く探索するために、複数のサーチパスを用いた局所的探索手法である Global Design through Local Search (GDLS) (Lun *et al.*, 2009)、遺伝的アルゴリズムを利用した OptGene (Patil *et al.*, 2005)、焼きなまし法を利用した探索手法 (Rocha *et al.*, 2008) などが開発され、実験的にも検証されてきた (Chemler *et al.*, 2010)。探索指標としてはターゲット化合物の生産速度、生産収率、あるいは生産収率と細胞の増殖速度との積が利用されている。しかしながら、これらの指標を用いた探索では、計算機上で多くの遺伝子の同時破壊がターゲット化合物生産に必要であり、単独の遺伝子破壊ではターゲット生産が予測されない場合に、そのような多重破壊を見つけ出すことが非常に困難になる。この問題を解決する方法としては、増殖速度最大化の条件下で高い生産収率が予測される破壊候補遺伝子を探索する際に、単純な生産収率や生産速度とは異なる探索指標を利用し、たとえ単独の遺伝子破壊ではターゲット生産が予測されなくても、他の遺伝子と同時に破壊することで化合物生産収率を上昇させるような遺伝子を評価することが考えられる。

そこで本章では、大域的探索の適用が困難な破壊数の場合に、微生物育種に有望な多重遺伝子破壊を見つけ出すための局所的探索手法FastPros (Fast Algorithm of knockout Screening for Target PROduction based on Shadow price analysis) を開発した。他の遺伝子との同時破壊によりターゲット化合物の生産収率を上昇させる遺伝子を評価するため、ターゲット化合物の生産速度と細胞の増殖速度を関連付ける指標 u_{TARGET} を新たに提案し、これを指標とした多重遺伝子破壊の局所的探索手法として FastPros を開発した。これは、FBA のように細胞の増殖速度を最大化する条件下で、高いターゲット生産収率の実現に必要な多重遺伝子破壊を探索する手法である。 u_{TARGET} は、ターゲット化合物生産速度の微小変化に対する、細胞の増殖速度の微小変化の割合として定義され、線形計画問題におけるシャドウプライスとして計算される。本章では、大腸菌細胞内の 625 の代謝物質をそれぞれ生産ターゲットとして FastPros を適用することで、 u_{TARGET} を利用した破壊候補遺伝子の探索が有効であることを示す。また、FastPros を既存の OptKnock と組み合わせることで、ターゲット化合物の生産収率をより上昇させるような破壊候補遺伝子の組み合わせを予測した。また、予測された破壊候補遺伝子の類似性に基づき、ターゲット化合物のクラスタリングを行うことで、複数の化合物生産において共通性の高い遺伝子破壊を同定し、様々な有用化合物生産に共通して利用可能な親株のデザインを提案した。

3.2 提案手法

3.2.1 代謝モデル

第2章と同様に、宿主微生物として大腸菌 K-12 MG1655 株を選択し、そのゲノムスケール代謝モデル iAF1260 (Feist *et al.*, 2007) を使用した。大腸菌は、モデル生物として知見が多く、また様々なターゲット生産における宿主微生物としても利用されてきたため、本章でも宿主微生物として採用した。この代謝モデルは、1260 遺伝子、2077 反応、1039 代謝物質より構成される。培養条件としては、グルコース単一炭素源下における微好気的な培養を想定し、グルコース消費速度の上限を 10 mmol/gDW/h、酸素消費速度の上限を 5 mmol/gDW/h とした。これは、微好気条件下での培養を想定している。好気条件下では、消費された炭素が菌体合成に使用される割合が大きいため、高いターゲット化合物生産速度を達成する場合には、微好気条件下での培養を想定するのが適切だと考えられる。細胞増殖を表現する反応としては、代謝モデル中の **core biomass objective function** を採用し、FBA や u_{TARGET} の計算時の線形計画問題における目的関数として使用した。計算機上での最小培地として、第2章 Table 2.1 に示す栄養源は細胞が自由に消費できるとした。また、Catalase、Superoxide dismutase、Superoxide dismutase in periplasm、Formate-hydrogen lyase の4反応については、本章で想定する培養条件下では活性を持たないとみなし、あらかじめモデルより取り除いた (Feist *et al.*, 2007)。

本章におけるターゲット化合物として、大腸菌細胞内に存在し得る 625 の代謝物質を選択した。これらの代謝物質は、i. 大腸菌のゲノムスケール代謝モデル iAF1260 に含まれる、ii. 構成原子に炭素を含む、iii. 上記したグルコース単一炭素源下での栄養条件で生合成可能である、iv. 遺伝子破壊前ではその代謝物質の生産が予測されない、という4つの条件を満たすものである。iv について、遺伝子破壊前でその代謝物質の生産が予測される場合には、生産収率を指標とした従来の局所的探索手法で有望な破壊候補遺伝子を比較的容易に同定することが可能であるため、本章の対象とはしなかった。ターゲット化合物のトランスポート反応に関して、代謝モデルに含まれている場合はそれを採用し、含まれていない場合は単純拡散を仮定したトランスポート反応を、新たに代謝モデルに追加した。

また、第2章と同様に、計算時間の短縮のために代謝モデルの縮約を行った。ただし、生物学的な理由から一部の遺伝子を破壊対象から除くステップを少し修正した。*gapA*、*pgk*、*eno*、*gpmA* は、第2章で述べたように、現実の大腸菌の増殖に必要であることが報告されているので (Baba *et al.*, 2006; Foster *et al.*, 2010)、破壊対象より除外した。そして、乳酸トランスporterをコードする *lldP* 遺伝子は、乳酸生合成反応をコードする *ldhA* 遺伝子の破壊とほぼ等価であるため、新たに破壊対象より除外した。同様に、ギ酸のトランスporterをコードする *focA* 遺伝子も、ギ酸の生合成反応をコードする *pflB* 遺伝子の破壊とほぼ等価であるため、破壊対象より除外した。

3.2.2 u_{TARGET}

FBA による代謝シミュレーションにおいて、細胞の増殖速度を最大化する条件下で、ターゲット化合物の生産速度が 0 となる場合を考える。このとき、ターゲット生産速度が 0 から少し増加すると、細胞の増殖速度は少し減少する。対して、増殖速度を最大化する条件下で、ターゲット生産速度が正となると、その生産速度が 0 から少し増加すると、細胞の増殖速度は少し増加する。そのため、ターゲット化合物生産速度の変化量に対する細胞の増殖速度の変化量の割合は、FBA でシミュレートされる代謝フラックス分布がターゲット化合物を生産する代謝状態にどの程度近いかを評価する指標として、利用可能だと考えられる。そこで本章では、それを評価する指標として、 u_{TARGET} を次のように定義した。

$$u_{TARGET} = \frac{\Delta v_{growth}}{\Delta v_{target}}$$

ここで、 Δv_{growth} は、ターゲット化合物生産速度が 0 から $\Delta v_{target} (> 0)$ だけ変化した際の細胞の増殖速度の変化量である (Fig. 3.1)。 u_{TARGET} が正の時は、細胞の増殖速度を最大化する条件下で、ターゲット化合物は生産される (Fig. 3.2 A)。対して u_{TARGET} が負の時は、この条件下でターゲット化合物は生産されない (Fig. 3.2 B)。ここで、 u_{TARGET} の絶対値は、この正負が変わることの難しさを表現している可能性がある。すなわち、 u_{TARGET} が 0 に近い負の値をとれば、新たな遺伝子破壊によって正の値に変わりやすく、逆に u_{TARGET} が大きな負の値をとれば、新たな遺伝子破壊によって正の値に変わる可能性が小さいことが期待できる。

開発した FastPros で最も重要な点は、ターゲット生産速度と増殖速度を関連付けている u_{TARGET} を、探索における指標として利用している点である。野生型の代謝モデルにおいて、 u_{TARGET} が大きな負の値であっても、 u_{TARGET} が増加するような 1 遺伝子破壊を繰り返し行えば、最終的に u_{TARGET} が正の値になり、細胞の増殖速度を最大化する条件下でターゲット化合物が生産される代謝フラックス分布および必要な破壊遺伝子を予測できることが期待される。

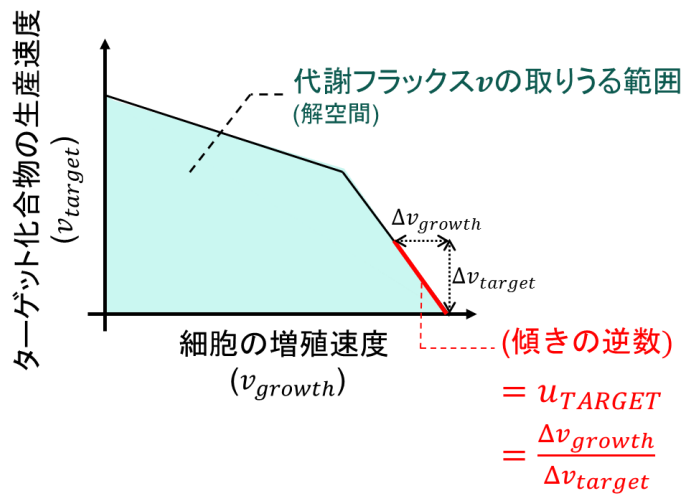


Fig. 3.1 u_{TARGET} の定義

u_{TARGET} は、増殖速度を最大化する条件下における、ターゲット化合物生産速度の変化量 (Δv_{target}) に対する細胞の増殖速度の変化量 (Δv_{growth}) として定義される。

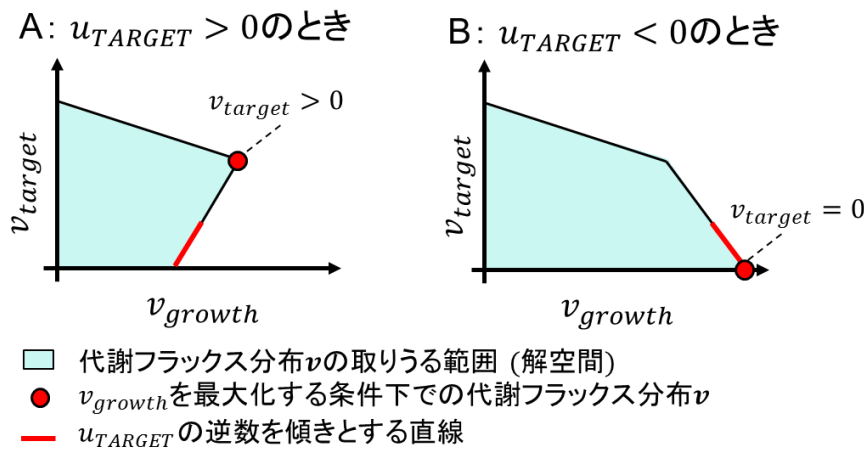


Fig. 3.2 u_{TARGET} が正のとき (A) と負のとき (B) の解空間の形状

$u_{TARGET} > 0$ のとき、細胞の増殖速度 (v_{growth}) を最大化する条件下で、ターゲット化合物生産速度 (v_{target}) が 0 より大きくなる。一方で、 $u_{TARGET} < 0$ のときは、細胞の増殖速度を最大化する条件下で、ターゲット化合物生産速度が 0 となる。

3.2.3 u_{TARGET} の計算

u_{TARGET} は、増殖速度を目的関数とした線形計画問題において、ターゲット化合物生産速度の制約式に対応するシャドウプライスに相当する。シャドウプライスは、特定の制約条件を強化あるいは緩和した時の最適値の変化量として定義される。したがって、 u_{TARGET} は次の線形計画化問題を利用して計算される。

$$\begin{aligned}
 & \max v_{growth} \\
 & (\mathbf{v}) \\
 & \text{s. t.} \quad \sum_{j \in R} S_{i,j} \cdot v_j = 0 \quad (\forall i \in M) \\
 & \quad v_{glc_uptake} \geq -GUR \\
 & \quad v_{o2_uptake} \geq -OUR \\
 & \quad v_{atp_main} \geq NGAM \\
 & \quad v_{growth} \geq v_{growth}^{min} \\
 & \quad v_j^{min} \leq v_j \leq v_j^{max} \quad (\forall j \in R) \\
 & \quad 0 \leq T_{k,j} \cdot v_j \leq 0 \quad (\forall j \in R, \forall k \in \mathbf{G}_{knockout}) \\
 & \quad v_{target} = b_{target}
 \end{aligned}$$

ただし、

v_j	: 反応 j のフラックス (mmol/gDW/h)
v_{glc_uptake}	: 炭素源消費反応のフラックス
v_{o2_uptake}	: 酸素消費反応のフラックス
v_{atp_main}	: ATP maintenance 反応のフラックス
v_{growth}	: 菌体合成反応のフラックス (細胞の比増殖速度)
v_{target}	: ターゲット化合物の生産速度
GUR	: グルコース消費速度の上限値 (10 mmol/gDW/h)
OUR	: 酸素消費速度の上限値 (5 mmol/gDW/h)
NGAM	: 増殖に関係なく細胞維持に必要な ATP 量 (8.39 mmol/gDW/h)
v_{growth}^{min}	: 細胞の比増殖速度の下限值
v_j^{max}	: 反応 j のフラックスの上限値 (10^5 mmol/gDW/h)
v_j^{min}	: 反応 j のフラックスの下限值 (可逆反応: -10^5 mmol/gDW/h、不可逆反応: 0 mmol/gDW/h)
b_{target}	: 最適解における u_{TARGET} を一意に定めるための値 (10^{-5})
$S_{i,j}$: 反応 j における代謝物質 i にかかる化学量論係数
$T_{k,j}$: 遺伝子 k が反応 j をコードするとき 1 となり、それ以外するとき 0 となる行列
M	: 代謝物質の集合

\mathbf{R} : 代謝反応の集合
 $\mathbf{G}_{knockout}$: 破壊対象の遺伝子の集合

である。

菌体合成反応のフラックスの下限値 (v_{growth}^{min}) は 0.05 mmol/gDW/h とし、遺伝子破壊によりこれ以上の増殖速度を細胞がとれない場合には、この遺伝子破壊株を実験的に構築するのは難しいとみなされ、探索結果から除外された。この線形最適化問題において、 u_{TARGET} はターゲット化合物生産速度に関する制約式 ($v_{target} = b_{target}$) に対応するシャドウプライスとして計算される。

線形計画問題におけるシャドウプライスは、その双対問題の最適解と等価である。前ページの線形計画問題の双対問題は、次のように記述できる。

$$\begin{aligned}
 \min \quad & \text{GUR} \cdot u_{GU} + \text{OUR} \cdot u_{OU} - \text{NGAM} \cdot u_{ATPM} - v_{growth}^{min} \cdot u_{GR} \\
 & + b_{target} \cdot u_{TARGET} \\
 (\boldsymbol{\lambda}, \mathbf{u}) \\
 \text{s. t.} \quad & \sum_{i \in M} S_{i,growth} \cdot \lambda_i - u_{GR} \geq 1 \\
 & \sum_{i \in M} S_{i,glc_uptake} \cdot \lambda_i - u_{GU} = 0 \\
 & \sum_{i \in M} S_{i,o2_uptake} \cdot \lambda_i - u_{OU} = 0 \\
 & \sum_{i \in M} S_{i,atp_main} \cdot \lambda_i - u_{ATPM} \geq 0 \\
 & \sum_{i \in M} S_{i,target} \cdot \lambda_i - u_{TARGET} \geq 0 \\
 & 0 \leq (1 - T_{k,j}) \sum_{i \in M} S_{i,j} \cdot \lambda_i \leq b_j^{max} \quad (\forall j \in \mathbf{R}) \\
 & \lambda_i \in \mathbb{R} \quad (\forall i \in \mathbf{M}) \\
 & u_{GU}, u_{OU}, u_{ATPM}, u_{GR} \geq 0 \\
 & u_{TARGET} \in \mathbb{R}
 \end{aligned}$$

ただし、

λ_i : 代謝物質 i の細胞内濃度の定常を表現する制約式に対応する双対変数
 u_{GU} : v_{glc_uptake} の制約式に対応する双対変数
 u_{OU} : v_{o2_uptake} の制約式に対応する双対変数
 u_{ATPM} : v_{atp_main} の制約式に対応する双対変数
 u_{GR} : v_{growth} の制約式に対応する双対変数
 u_{TARGET} : v_{target} の制約式に対応する双対変数
 b_j^{max} : $\sum_{i \in M} S_{i,j} \cdot \lambda_i$ の上限値
(反応 j が可逆反応の時は 0、不可逆反応の時は 10^5)

である。

線形計画問題における双対定理 (坂和、1999) より、もとの問題の最適値と双対問題の最適値が一致する。したがって、もとの問題の最適解を \mathbf{v}' 、双対問題の最適解を $\boldsymbol{\lambda}'$ 、 \mathbf{u}' と

すると、次の式が成り立つ。

$$v'_{growth} = GUR \cdot u'_{GU} + OUR \cdot u'_{OU} - NGAM \cdot u'_{ATPM} - v_{growth}^{min} \cdot u'_{GR} + b_{target} \cdot u'_{TARGET}$$

したがって、 b_{target} が Δb_{target} だけ変化すると、 v_{growth} は $\Delta b_{target} \cdot u'_{TARGET}$ だけ変化するため、

$$\frac{\Delta v_{growth}}{\Delta b_{target}} = \frac{\Delta b_{target} \cdot u'_{TARGET}}{\Delta b_{target}} = u'_{TARGET}$$

となる。以上より、 u_{TARGET} は、ターゲット化合物生産速度の制約式に対応する双対問題の最適解 (シャドウプライス) として計算される。

3.2.4 FastPros

FastPros は、 u_{TARGET} が上昇するような遺伝子破壊を繰り返し行い、ターゲット化合物生産に必要な多重遺伝子破壊を局所的に探索する手法である。FastPros の概要を Fig. 3.3 に示す。破壊対象の遺伝子数が g の代謝モデルを用いて、FastPros ではまず、可能な全ての 2 重遺伝子破壊に対して u_{TARGET} を計算する。次に、 u_{TARGET} が大きい上位 p 組の遺伝子破壊の組み合わせを選択し、母集団となる遺伝子破壊を用意する。それぞれの世代において、母集団の遺伝子破壊それぞれに対して、可能な全ての 1 遺伝子破壊を追加し ($p \times g$ の遺伝子破壊の組み合わせが作成される)、 u_{TARGET} を計算する。そして、遺伝子破壊によって u_{TARGET} が上昇した遺伝子破壊のうち、 u_{TARGET} が大きい上位 p 組の遺伝子破壊の組み合わせを選択し、次の世代での母集団とする。ある遺伝子破壊の組み合わせで u_{TARGET} が 0 以上となった場合には、この遺伝子破壊は探索から取り除かれ、残りの遺伝子破壊については探索が続けられる。この遺伝子破壊と選択は、指定する最大の破壊数に到達するまで繰り返される。本章では最大の遺伝子破壊数は 25 とした。最後に、 u_{TARGET} が 0 以上となった遺伝子破壊については、もとの代謝モデルに対してその遺伝子破壊を適用し、第 2 章と同様に、次の線形計画問題を解くことで代謝フラックス分布を計算する。

$$\begin{aligned}
& \max v_{growth} + \alpha \cdot v_{target} \\
& (v) \\
& \text{s. t.} \quad \sum_{j \in R} S_{i,j} \cdot v_j = 0 \quad (\forall i \in M) \\
& \quad v_{glc_uptake} \geq -GUR \\
& \quad v_{o2_uptake} \geq -OUR \\
& \quad v_{atp_main} \geq NGAM \\
& \quad v_{growth} \geq v_{growth}^{min} \\
& \quad v_j^{min} \leq v_j \leq v_j^{max} \quad (\forall j \in R) \\
& \quad 0 \leq T_{k,j} \cdot v_j \leq 0 \quad (\forall j \in R, \forall k \in G_{knockout})
\end{aligned}$$

(変数等については page 49-50 参照)

ただし、 α は増殖速度最大化の条件下での v_{target} を一意に定めるための係数であり、 10^{-6} とした。以上より、FastPros では u_{TARGET} が 0 以上となる、すなわち増殖速度最大化の条件下でターゲット化合物生産が予測されるような多重遺伝子破壊が探索される。

母集団となる遺伝子破壊の組み合わせ数 p は、探索範囲に影響するパラメータであり、ヒューリスティックに決められた。 p を $0.5 \times g$, $1 \times g$, $1.5 \times g$ に設定し、大腸菌代謝モデルよりランダムに選択された 30 の代謝物質生産に向けて FastPros を実行した。その結果、 $p=1 \times g$ の場合で遺伝子破壊による生産が予測された代謝物質のうち 18%については、 $p=0.5 \times g$ 場合には生産が予測されなかった。また、 $p=1.5 \times g$ の場合には $p=1 \times g$ の場合とほぼ同じ代謝物質が、遺伝子破壊による生産が予測された。そこで、本章での大腸菌代謝モデルを用いた探索では、 $p=1 \times g$ とするのが適切であると判断した。

FastPros による遺伝子破壊の探索において、遺伝子破壊による解空間の変化の概要を Fig.3.4 に示す。遺伝子破壊を繰り返すごとに解空間の形状が変化し、 u_{TARGET} が正の値をとった時に、増殖速度 (v_{growth}) が最大となる点でのターゲット化合物生産速度 (v_{target}) が正の値をとる。ターゲット化合物生産に複数の遺伝子の同時破壊が必要な場合、遺伝子破壊の一部を実行しても、増殖速度が最大となる時の v_{target} は 0 から変化しない。対して u_{TARGET} は、遺伝子破壊によって細胞増殖最大化で v_{target} が正になるような解空間にどの程度近づいたかを評価できる。したがって、 v_{target} を指標とした探索と比較して、ターゲット化合物生産に有望な多重遺伝子破壊を効率よく見つけ出せることが予想される。

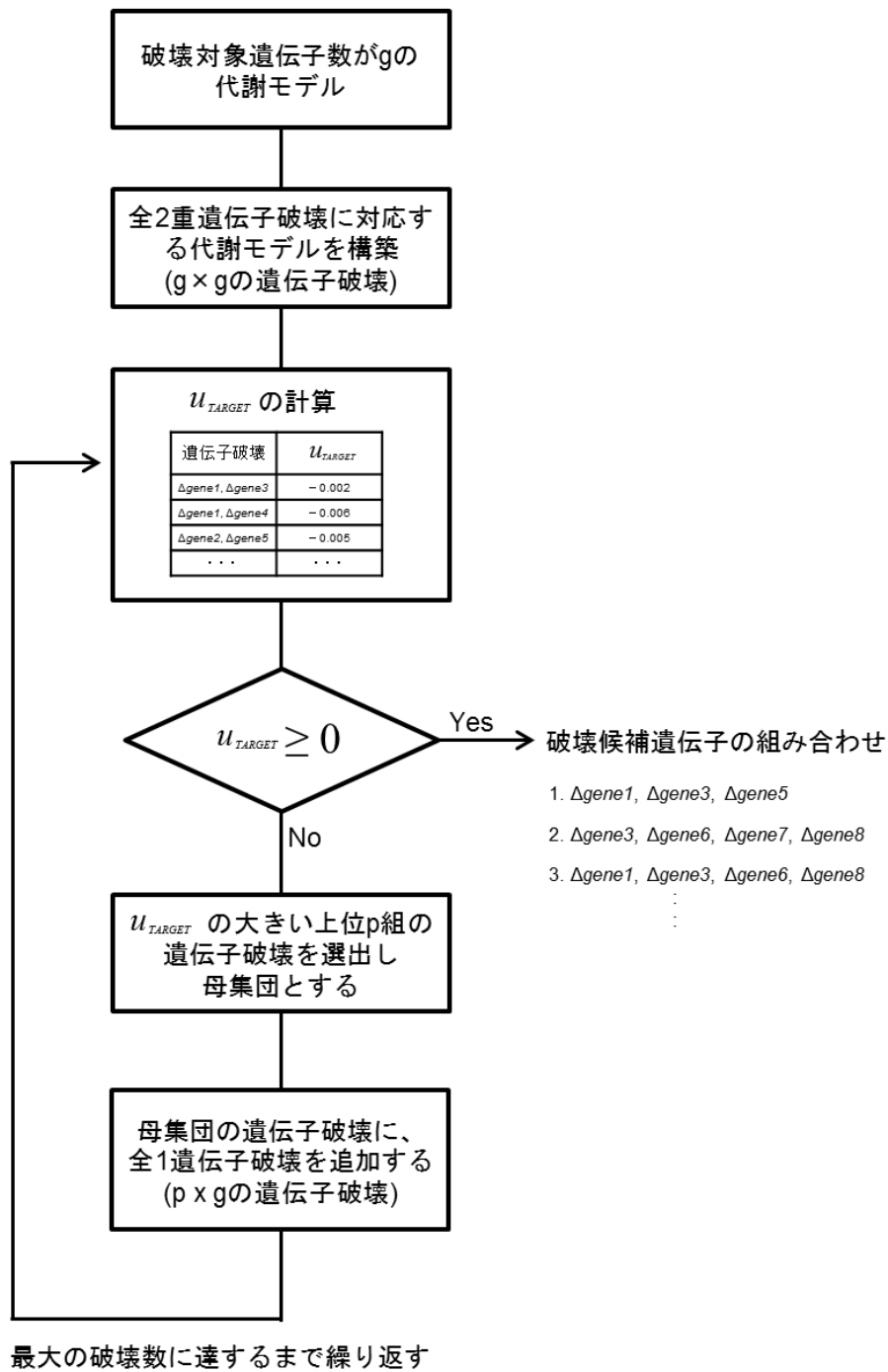


Fig. 3.3 FastPros の概要

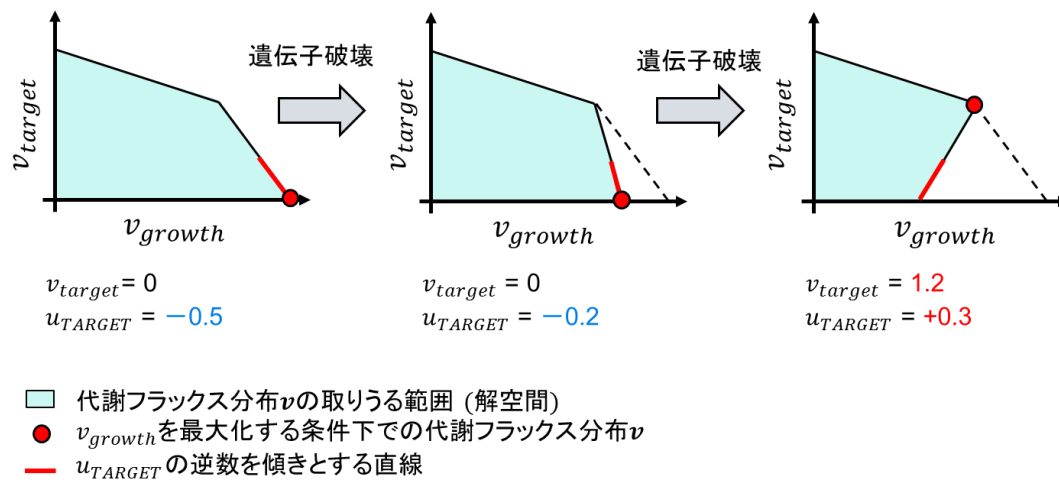


Fig. 3.4 FastPros での探索における、遺伝子破壊による解空間の変化の例

遺伝子破壊を繰り返すごとに解空間の形状が変化し、 u_{TARGET} が正の値となった時に、細胞の増殖速度 (v_{growth}) が最大となる点でのターゲット化合物生産速度 (v_{target}) が正の値をとる。 u_{TARGET} が上昇する遺伝子破壊を行うとき、図中の赤色の直線の傾きが変化し、解空間の形状が v_{growth} 最大化で v_{target} が正となる場合の形状に近づく。

3.2.5 FastPros の結果に基づく OptKnock

FastPros ではターゲット化合物生産を導く遺伝子破壊が同定されるが、同定された遺伝子破壊は、必ずしも高い生産収率を導くものではない。 u_{TARGET} の正負は、増殖速度最大化の条件下でターゲットを生産が予測されるか否かを表現する値であり、その絶対値は生産性に対応していない。そこで、開発した FastPros と、先行研究で提案された大域的探索手法である OptKnock を組み合わせることで、高い生産収率が予測される破壊候補遺伝子を同定する方法を提案した。FastPros で見つけ出された遺伝子破壊は、代謝状態をターゲット化合物生産に適した状態に改変するうえで、比較的大きな影響を持っていると予想される。そこで、FastPros の結果得られた破壊候補遺伝子を、可能な破壊対象として、OptKnock を実行した。この OptKnock は次の 2 段階の最適化問題として記述される。

$$\begin{array}{ll}
 \max & v_{target} - \beta \cdot \sum_{k \in G} (1 - y_k) \\
 (\mathbf{y}) & \\
 \text{s. t.} & \left(\begin{array}{ll}
 \max & v_{growth} \\
 (\mathbf{v}) & \\
 \text{s. t.} & \sum_{j \in R} S_{i,j} \cdot v_j = 0 \quad (\forall i \in M) \\
 & v_{glc_uptake} \geq -GUR \\
 & v_{o2_uptake} \geq -OUR \\
 & v_{atp_main} \geq NGAM \\
 & v_{growth} \geq v_{growth}^{min} \\
 & y_k \cdot v_j^{min} \leq v_j \leq y_k \cdot v_j^{max} \\
 & \quad (\forall j \in R_{ko}, \text{reaction } j \text{ is encoded by gene } k, \forall k \in G_{knockout}) \\
 & v_j^{min} \leq v_j \leq v_j^{max} \quad (\forall j \in R_{nko}) \\
 & \sum_{k \in G_{ko}} (1 - y_k) \leq K^{max} \\
 & y_k \in \{0, 1\} \quad (\forall k \in G_{knockout})
 \end{array} \right)
 \end{array}$$

(変数等については page 49-50、 または以下の文を参照)

y_k は、遺伝子が破壊されているかを示す 2 値変数であり、0 の時は遺伝子 k が破壊され、1 の時は破壊されていないこと意味する。この 2 段階の最適化問題において、内側では代謝フラックス \mathbf{v} を変数とし、遺伝子破壊後の代謝フラックス分布を FBA でシミュレートしている。外側では遺伝子の状態 (破壊されているか否か) を示す \mathbf{y} を変数とし、ターゲット化合物の生産速度が最大化される。実際にこの 2 段階の最適化問題と解くときには、内側の問題とその双対問題の最適値が一致すること (強双対定理) を利用し、1 段階の 0-1 混合整数線形計画問題として解かれた (Burgard *et al.*, 2003)。FastPros での探索で u_{TARGET} が 0 以上となった様々な破壊候補遺伝子の中で、高い生産収率が予測された 30 の破壊候補遺伝子を選択し、OptKnock における破壊対象の遺伝子 ($G_{knockout}$) とした。 R_{ko} は、 $G_{knockout}$

に含まれる遺伝子によりコードされる反応であり、 R_{nko} はそれ以外の反応である。最大の破壊数 K^{max} は 3、5、10 の 3 つの条件下で OptKnock は実行された。 β は遺伝子破壊のペナルティーを表現するパラメータであり、 10^{-6} とした。これにより、最大化された生産速度が同じである場合には、破壊遺伝子数が少ないものが最適解として選ばれるようにした。

3.2.6 破壊候補遺伝子の類似度に基づくクラスタリング

複数の代謝物質生産に共通する破壊候補遺伝子の同定を目指し、FastPros の結果に基づく OptKnock の結果として得られた破壊候補遺伝子を用いて、階層的クラスタリングを行った。ターゲット化合物間の破壊候補遺伝子の類似度は、Jaccard 距離 (Levanodowsky and Winter, 1971) で表現し、クラスタリングには群平均法を用いた。

ターゲット化合物 A の破壊候補遺伝子 G^A と、ターゲット化合物 B の破壊候補遺伝子 G^B の類似度を示す Jaccard 距離 J_{AB} は、次の式によって定義される。

$$J_{AB} = 1 - \frac{|G^A \cap G^B|}{|G^A \cup G^B|}$$

Jaccard 距離は、0 から 1 までの値をとり、0 に近いほど破壊候補遺伝子の類似度が大きいことを意味している。

本研究の階層的クラスタリングでは、群平均法を用いて、破壊候補遺伝子が似ている化合物を段階的にグルーピングした。群平均法では、グループ間の距離は、2 つグループ中の化合物の距離の平均値として定義され、グループ間の距離が小さいものから順にグルーピングされる。グループ c_i とグループ c_j の間の距離は、次の式で記述される。

$$d_{ij} = \frac{1}{n_i n_j} \sum_{k=1}^{n_i} \sum_{s=1}^{n_j} J_{ks}$$

ここで、 n_i 、 n_j はグループ c_i 、 c_j に含まれる化合物数を、それぞれ意味する。

3.3 結果

3.3.1 FastPros によるターゲット化合物生産に向けた遺伝子破壊の探索

本章では、ターゲットの生産速度と細胞の増殖速度を関連付ける指標として u_{TARGET} を新たに提案した。そして、 u_{TARGET} を探索指標として、育種に向けた多重遺伝子破壊の局所的探索手法 FastPros を開発した。 u_{TARGET} が負の時は、増殖速度が最大となる条件下でターゲット生産速度が 0 の場合に、 u_{TARGET} が正の時は、ターゲット生産速度が正の場合に対応する。そこで FastPros では、野生型の代謝モデルで u_{TARGET} が負の時に、 u_{TARGET} が上昇する遺伝子破壊が繰り返し行われる。最終的に u_{TARGET} が 0 以上となれば、細胞の増殖速度を最大化する条件下でターゲット生産が予測される代謝フラックス分布と、必要な破壊候補遺伝子が同定される。

FastPros の性能を調べるため、大腸菌のゲノムスケール代謝モデル iAF1260 に含まれる 625 の代謝物質を生産ターゲットとして選択し、増殖速度最大化の条件下でのターゲット化合物の生産に必要な遺伝子破壊を探索した。この時、最大の遺伝子破壊数は 25 とした。結果、75%にあたる 472 の代謝物質について、 u_{TARGET} が 0 以上となる遺伝子破壊がを見つけ出された。これは、を見つけ出された遺伝子破壊をもとの代謝モデルに適用すると、増殖速度最大化の条件下で対応するターゲット化合物が生産されることを意味する。生産が予測された代謝物質には、アミノ酸、核酸、脂質、補酵素などが含まれており、幅広い範囲の代謝物質の生産に対して、FastPros による破壊候補遺伝子の探索が可能であることが示された。Fig. 3.5 に、ゲラニルピロリン酸 (GPP) を生産ターゲットとした FastPros での探索で、遺伝子破壊数の増加に伴う u_{TARGET} 、GPP 生産速度、細胞の増殖速度と GPP 生産速度の取りうる範囲の変化を示している。野生型の代謝モデルでは、 u_{TARGET} は負の値 (-0.082) であり、GPP 生産は細胞の増殖にとってデメリットであることを意味している。この代謝モデルに対し、*adhE*、*ldhA*、*ppc*、*mgsA* 遺伝子を順に破壊すると、 u_{TARGET} は徐々に増加する (Fig. 3.5 A)。これは、GPP 生産に伴う細胞の増殖速度の減少量が、遺伝子破壊を行うごとに逡減していることを意味している。さらに、上記した 4 重遺伝子破壊に加えて *pntA* 遺伝子を破壊すると、 u_{TARGET} は正の値 (0.017) をとり、増殖速度が最大となる時、 1.0 mmol/gDW/h の生産速度 (0.19 g/g-glucose の生産収率) で GPP が生産されることを意味する。一方で、一般的な FBA では、代謝フラックス分布は増殖速度を最大化する条件下で計算されるため、 u_{TARGET} が負の間は、FBA で計算される GPP 生産速度は常に 0 となる (Fig. 3.5 B)。遺伝子破壊に伴って、定常を満たす代謝フラックス分布がとりうる範囲 (解空間) が変化するが、 u_{TARGET} の変化はこの解空間の変化を端的に表現している (Fig. 3.5 C)。

3.3.5 で詳しく述べるように、GPP の生合成では、NADH などの還元力が必要とされる。そのため、遺伝子破壊により、還元力を消費するエタノール、乳酸、コハク酸、1,2-プロパンジオール、L-バリンの生合成が阻害されると、解糖で生じた還元力を消費するため、GPP が生産されることが予測された。なお FastPros において、 u_{TARGET} が上昇するように遺伝子

破壊を繰り返し行うことは、生物学的には、ATP の生産量が多くかつ還元力の消費量が多い生合成経路から順に、阻害したことに相当する (Table 3.1)。

なお、すべての計算は MATLAB (Mathworks, Inc) を Windows 上で使用して行った。FBA などの最適化問題を解く場合には、ソルバーとして GLPK (GNU linear programming kit) の関数を MATLAB 上で使用した。このソルバーでは、線形計画問題を解くときに、その問題のシャドウプライスも同時に計算される。CPU には、Intel Xeon CPU (2.93 GHz, 12 core) を使用した。この計算環境下で一つの代謝物質を生産ターゲットとして FastPros を実行した時、計算時間は、 2.9×10^3 秒 (48 分) だった。

Table 3.1 1 mol のグリセルアルデヒド 3 リン酸 (G3P) からのエタノール、乳酸、コハク酸、1,2-プロパンジオール、ゲラニルピロリン酸 (GPP) 生合成に消費する還元力および ATP 生産量

1 mol の G3P から生合成される代謝物質		生産される ATP 量 (mol)	消費される還元力 (mol-e ⁻)
エタノール	(1 mol)	2	1
乳酸	(1mol)	2	0
コハク酸	(1 mol*)	1	1
1,2-プロパンジオール	(1mol)	0	2
L-バリン	(0.5 mol)	2**	0
GPP	(0.25 mol)	0***	1

*Phosphoenolpyruvate carboxylase 反応により、1 mol の CO₂ より炭素が固定される

**ただし、Transhydrogenase 反応により、プロトン勾配を利用して、NADH から NADP⁺への電子の移動が必要である。これは電子伝達系での ATP 合成に必要なプロトン勾配を減少させるため、電子伝達系での ATP 合成量は減少する。

***ATP の生産量と消費量は釣り合うが、0.5 mol の CTP の脱リン酸を伴う。

FastPros で順に破壊された *adhE*, *ldhA*, *ppc*, *mgsA* は、それぞれエタノール、乳酸、コハク酸、1,2-プロパンジオールの生合成に関わる遺伝子である。FastPros において、 u_{TARGET} が上昇するように遺伝子破壊を繰り返し行うことは、生物学的には、ATP の生産量が多くかつ還元力の消費量の多い生合成経路から順に、阻害したことに相当する。

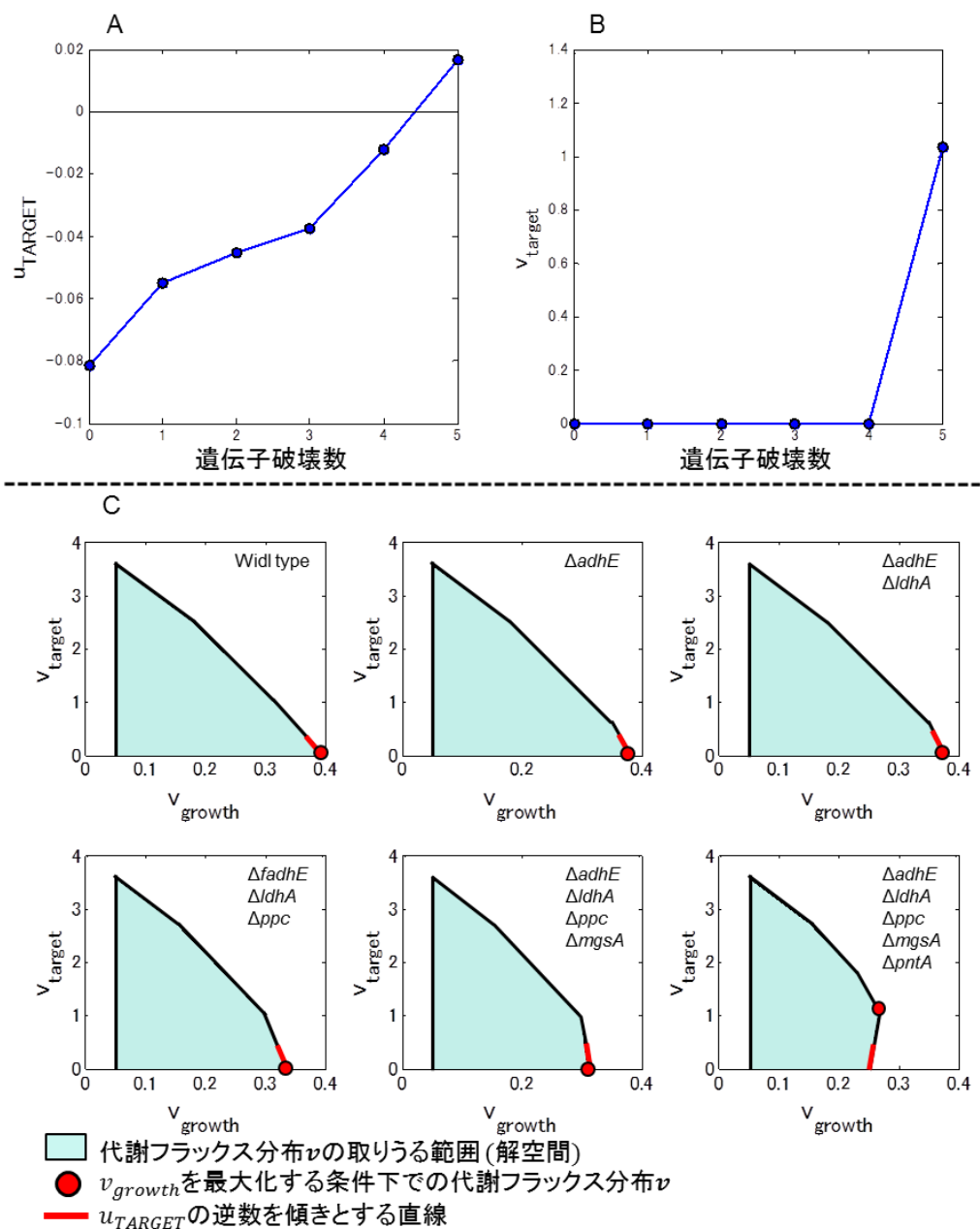


Fig. 3.5 グラニルピロリン酸 (GPP) を生産ターゲットとしたときの、FastPros による多重遺伝子破壊の探索

A: 遺伝子破壊に伴う u_{TARGET} の変化、B: FBA で予測される GPP 生産速度 (v_{target}) の、遺伝子破壊に伴う変化、C: 遺伝子破壊に伴う GPP 生産速度と細胞の増殖速度 (v_{growth}) の取りうる範囲 (解空間) の変化。遺伝子破壊により u_{TARGET} が増加し、5 遺伝子の破壊により、0.182 g/g-glucose の収率での GPP 生産が予測された。解空間については、3.2.4 で記述された FBA の線形計画問題において、 v_{target} を変化させたときの最大の v_{growth} の値より描かれた。 v_{growth} が 0.05 以下の領域が存在しないのは、 v_{growth} の下限値を 0.05 としているからである。

3.3.2 既存の探索手法との比較

FastPros で破壊候補遺伝子が同定された代謝物質について、既存の破壊候補遺伝子の大域的探索手法である OptKnock と比較し、探索の正確さを評価した。OptKnock を用いると、3 遺伝子以下の破壊により、全 625 の代謝物質のうち 152 についてそれぞれの生産が予測された。この 152 の代謝物質のうち 136 については、FastPros でも 3 遺伝子以下の破壊により生産された (Fig. 3.6 A)。残りの 16 の代謝物質についても、FastPros では 4 遺伝子以上の遺伝子破壊により生産が予測された (Fig. 3.6 B)。これは、FastPros は解の取りこぼしが少ない探索手法であることを意味している。なお、OptKnock での計算時間は破壊遺伝子数に対して指数関数的に増加するため、破壊可能な遺伝子数には限界がある。一方で、FastPros での計算時間は破壊遺伝子数に対して線形で増加するため、小さい計算コストでターゲット化合物生産に向けた多重遺伝子破壊を探索可能である。Fig. 3.7 A に、FastPros の結果示された、ターゲット生産に必要な遺伝子破壊数を示している。FastPros ではターゲット化合物生産に 10 遺伝子以上が必要な場合でも、遺伝子破壊を見つけ出すことに成功している

また、他の破壊候補遺伝子の探索手法である GDLS と OptGene についても、FastPros の探索結果と比較した (Fig. 3.6 C, D)。GDLS はターゲット生産速度を探索の指標とした局所的探索手法であり、OptGene は遺伝的アルゴリズムを利用した探索手法である。探索でのパラメータについては先行研究のものを使用した (付録の Table A3, 4 参照)。その結果、625 の代謝物質のうち、GDLS では 55 の代謝物質、OptGene では 247 の代謝物質について、最大 25 遺伝子破壊まででの生産が予測された。一方、FastPros で破壊候補遺伝子が同定された代謝物質は、GDLS で生産された 55 のうち全て、OptGene で生産された 247 のうち 244 だった。GDLS と OptGene では、比較的破壊遺伝子数が少ない場合には、生産に有望な破壊候補遺伝子を同定することに成功したが、生産に必要な破壊数が多い場合には同定に失敗していた (Fig. 3.7 B, C)。これは、GDLS と OptGene では探索の指標としてターゲット生産速度あるいは収率が使用されており、破壊候補遺伝子を同定するには、この生産速度が増加する必要があるからである。野生型あるいは遺伝子破壊数が少ない条件でターゲット化合物が生産される場合であれば、従来手法のように生産速度あるいは収率を探索指標として採用することは有効である。しかし、ターゲット生産に多数の遺伝子破壊が必要な場合には、生産されるまでに莫大な遺伝子破壊の組み合わせを探索する必要があるため、生産速度あるいは収率を指標とした探索は非常に困難になる。例えば、遺伝的アルゴリズムである OptGene において、探索における最初の集団に含まれる遺伝子破壊は、ランダムに決定される。もし、その集団でターゲット生産速度が正となるような遺伝子破壊が見つからなければ、有望な遺伝子破壊を選抜して次世代に残すことはできない。対して FastPros では、ターゲット生産が集団中に見られない場合においても、 u_{TARGET} を探索指標として利用することで、ターゲット化合物生産に潜在的に貢献する遺伝子破壊を見つけ出すことが可能である。

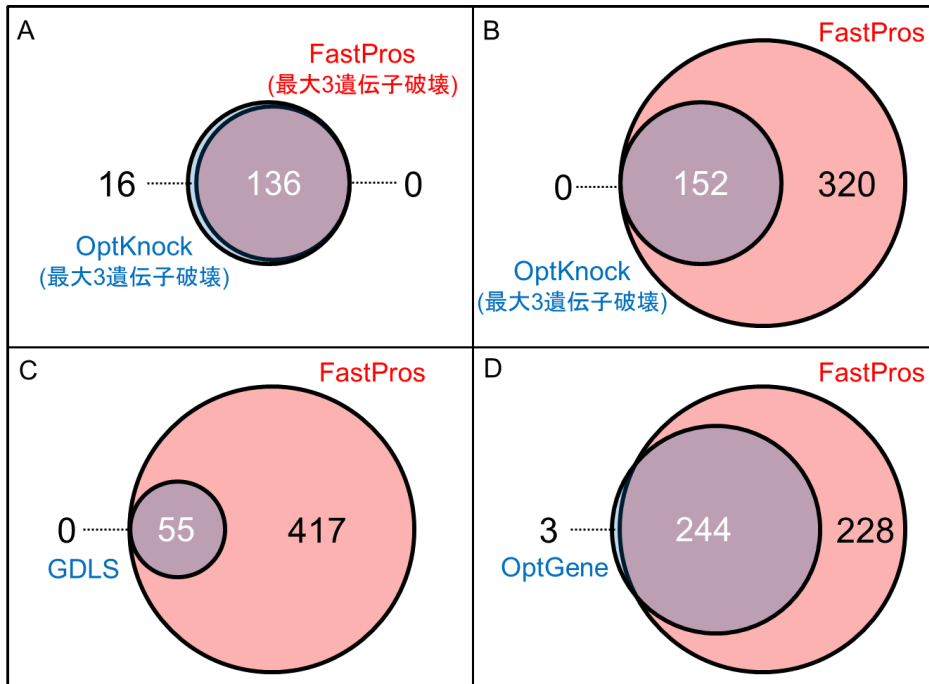


Fig. 3.6 破壊遺伝子の探索により生産された代謝物質のベン図

A: OptKnock (最大3 遺伝子破壊) vs FastPros (最大3 遺伝子破壊), B: OptKnock (最大3 遺伝子破壊) vs FastPros, C: GDLS vs FastPros, D: OptGene vs FastPros。最大の遺伝子破壊数は、断りがない限り 25 としている。大域的探索手法である OptKnock に対して、FastPros では生産された代謝物質の取りこぼしの少なかった。また、GDLS と OptGene で生産された代謝物質についても、FastPros ではそのほぼ全てが生産された。GDLS と OptGene のパラメータについては付録 Table A3, 4 参照。

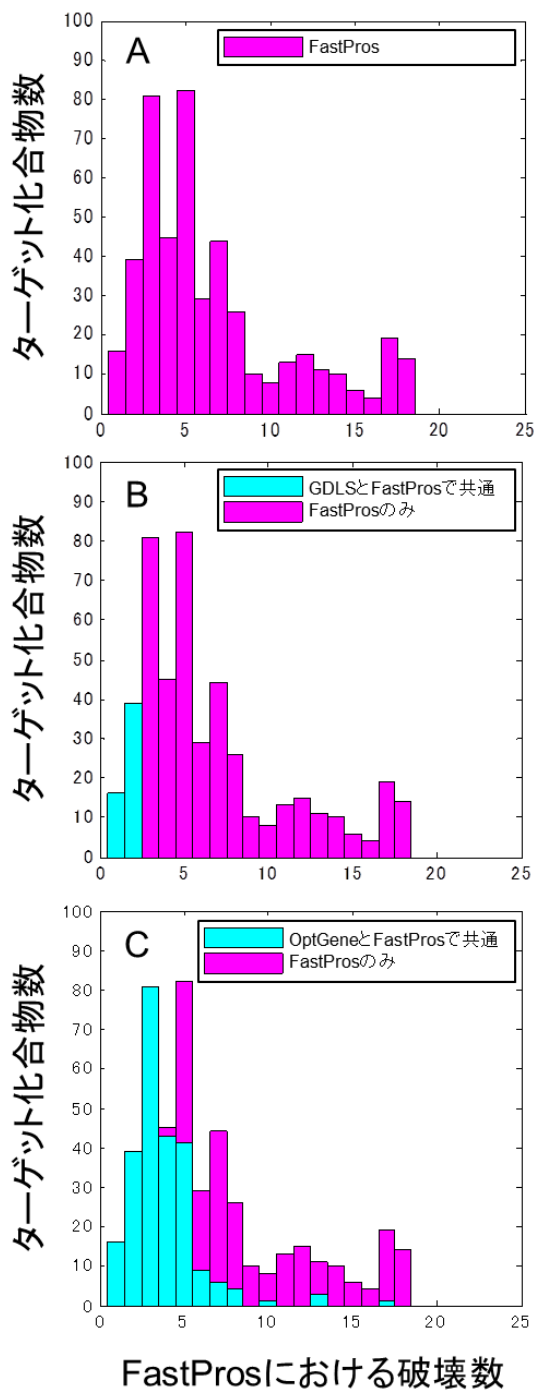


Fig. 3.7 ターゲット化合物生産に必要な破壊遺伝子数の分布

A: FastPros、B: FastPros と GDLs の比較、C: FastPros と OptGene の比較。FastPros は他の手法と比較して、生産に必要な遺伝子破壊数が多いターゲット化合物についても、その遺伝子破壊を見つけ出すことに成功した。

3.3.3 u_{TARGET} を利用した探索における適応度地形

FastPros は、新規探索指標 u_{TARGET} を利用した探索手法である。ターゲット生産速度を探索指標とした従来手法と比較して、多くの代謝物質生産に必要な多重遺伝子破壊が同定された。そこで、破壊候補遺伝子の探索を、「遺伝子型が形成する空間における適応度地形上の山登り」と考えた時に、 u_{TARGET} を探索指標とした場合の適応度地形はどのような特徴を持つのかを解析した。ここで「適応度」は、探索指標である u_{TARGET} の大きさを意味する。FastPros での破壊候補遺伝子の探索は、適応度が増加するように遺伝子型の空間を探索すると考えることができる。その際に、適応度地形の形状は、探索の安定性や効率性を議論するうえで重要な概念となる (Tavares *et al.*, 2008)。

FastPros で生産が予測されたターゲット化合物および同定された破壊候補遺伝子に対して、破壊の順番に対する u_{TARGET} の変化を解析した。破壊の順番に関わらず u_{TARGET} が単調増加する場合には、適応度地形は凹凸のない形状をしているといえる。この場合、増殖速度最大化の条件下でターゲットが生産される遺伝子破壊が存在すれば、 u_{TARGET} が単調増加するような遺伝子破壊を繰り返し行うことで、必ずその遺伝子破壊を見つけ出すことができる。一方で、破壊の順番によって u_{TARGET} が減少するときもある場合には、適応度地形は凹凸のある形状をしているといえる。この場合には、増殖速度最大化の条件下でターゲットが生産されるような遺伝子破壊が存在していても、FastPros での探索ではその破壊候補遺伝子が同定されなかった可能性がある。

FastPros において、10 遺伝子以下の破壊での生産が予測された 380 のターゲット化合物およびその時の破壊候補遺伝子について、破壊の順序に伴う u_{TARGET} の変化を解析し、すべての破壊の順序で u_{TARGET} が単調増加するターゲット化合物を数えた。例えば、GPP に関しては、5 重破壊での生産が FastPros で予測されたので、 $5! = 120$ 通りの破壊の順序に対する u_{TARGET} の変化を計算し、すべての破壊の順序で u_{TARGET} が単調増加するかを調べた。ターゲットによって、生産が予測される遺伝子破壊の組み合わせが複数見つけ出されている場合には、破壊数が最小であるものの中で生産収率が最大となる遺伝子破壊の組み合わせ 1 つを選択した。その結果、380 のターゲット化合物のうち、すべての破壊の順序で u_{TARGET} が単調増加したのは、153 のターゲットだった。これらのターゲット化合物における適応度地形は凹凸のない形状をしており、 u_{TARGET} を探索指標とすることで、効率的に破壊候補遺伝子を同定できると考えられる。そのターゲット生産に必要な破壊数の分布を、Fig. 3.8 に示す。 u_{TARGET} の変化より推測される適応度地形の凹凸度は、ターゲットに依存するところが多いが、9 遺伝子破壊におけるその $9! = 3.6 \times 10^5$ 通りの破壊の順序すべてで u_{TARGET} の単調増加がみられるターゲットも存在した。 u_{TARGET} の変化より推測され 3.3.2 では、従来手法よりも多くのターゲットに対して FastPros による探索の有効性が示されたが、増殖速度最大化の条件下でターゲット生産が予測される全ての破壊候補遺伝子を取りこぼしている可能性があると言えた。

以上の解析より、FastPros では幅広い範囲のターゲットについて、細胞増殖最大化の条件

下でターゲット生産が予測される破壊候補遺伝子が同定されたものの、必ずしもすべての破壊候補遺伝子を同定できるわけではないことが示唆された。さらに、破壊候補遺伝子が同定可能か否かは、ターゲットに依存することも示された。一方で、3.2.4 で述べたように、本章では遺伝子破壊が適用される母集団の数を (p) を、破壊対象の遺伝子数 (g) として探索を行ったが、 p を $1.5 \times g$ に上昇し、破壊候補遺伝子が同定されたターゲット数に変化はなかった。そのため、単純に p を増加させて探索範囲を広げただけでは、破壊候補遺伝子が同定されるターゲット数に大きな違いはないと考えられる。

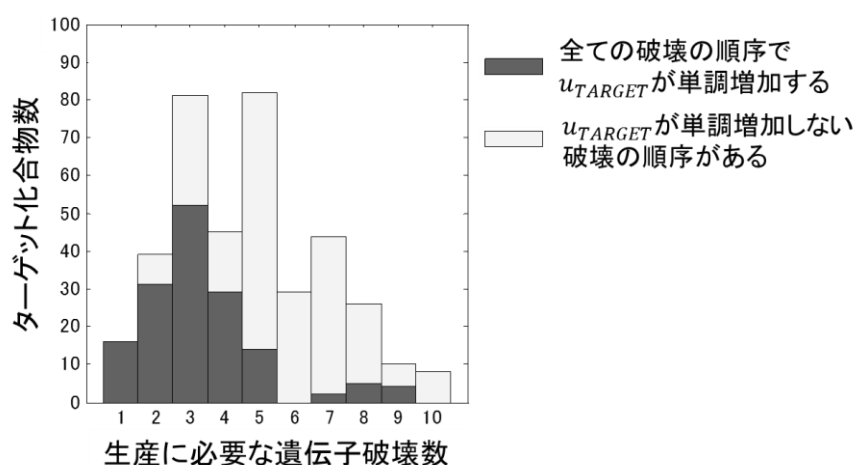


Fig. 3.8 破壊の順序による u_{TARGET} の単調増加性に対する、ターゲット生産に必要な破壊数の分布

FastPros において、10 遺伝子以下の破壊での生産が予測された 380 のターゲット化合物のうち、153 のターゲットは全ての破壊の順序で u_{TARGET} の単調増加が観察された。これらのターゲット化合物の適応度地形は凹凸のない形状をしており、 u_{TARGET} を探索指標とすることで、効率的に破壊候補遺伝子を同定できると考えられる。

3.3.4 FastPros の結果に基づく OptKnock によるターゲット生産性の向上

FastPros では、 u_{TARGET} が上昇するように遺伝子破壊を繰り返し行うことで、増殖速度を最大化する条件下でターゲット生産が予測される遺伝子破壊がを見つけ出される。しかし、見つけ出された遺伝子破壊により、必ずしも高い生産収率を伴う代謝フラックス分布がシミュレーションの結果として得られるわけではない。増殖速度とターゲット生産速度を関連付けた u_{TARGET} の正負は、増殖速度最大化の条件下でターゲットが生産されるか否かを表現する値であり、その絶対値は生産収率に対応するわけではない。したがって、高い生産収率に必要な遺伝子破壊を見つけ出すためには、 u_{TARGET} を利用した探索だけでは不十分である。

そこで、高い生産収率が予測される遺伝子破壊を同定するため、FastPros と OptKnock を組み合わせた探索方法を提案した。前述したように、FastPros では u_{TARGET} が正になるような遺伝子破壊がを見つけ出される。そのため、見つけ出された遺伝子破壊は、代謝状態をターゲット化合物生産に適したものに改変するうえで、比較的に大きな影響を持っていると予想される。したがって、FastPros によって同定された破壊候補遺伝子より、生産収率が高い破壊の組み合わせを見つけ出すのは妥当だと推測される。そこで、FastPros で同定された破壊候補遺伝子を可能な破壊対象として、大域的探索手法である OptKnock を実行した。 u_{TARGET} が 0 以上となった様々な破壊候補遺伝子の中で、高い生産収率が予測された 30 の破壊候補遺伝子を選択し、その中から最大で 3、5、10 遺伝子を破壊できる条件下で OptKnock を実行した。生産ターゲットとしては、FastPros で 10 遺伝子以下の破壊により生産された 380 の代謝物質を選択した。その結果、FastPros のみでの探索結果と比較して、OptKnock を組み合わせることで、106 (28%) のターゲット化合物の生産収率が、理論最大値の 5% 以上増加した。例えば、コハク酸とセドヘプツロース 7 リン酸の生産収率は、FastPros での探索ではそれぞれ理論最大値の 22% と 55% だったが、FastPros の結果に基づく OptKnock では、それぞれ 74% と 79% に上昇した (Fig. 3.9 A, B)。また、グリセロールを生産ターゲットとした場合は、3 遺伝子破壊以下の条件で OptKnock のみを実行した場合には、増殖速度最大化の条件下で生産が予測される遺伝子破壊は見つけ出されなかったが、FastPros の結果に基づく OptKnock では、理論最大値の 68% という高い生産収率が予測された (Fig. 3.9 C)。FastPros の結果に基づく OptKnock で予測される生産収率は、GDLS あるいは OptGene の予測結果とも比較された。結果、提案した手法は既存の手法と同等以上の生産収率が予測された (Fig. 3.10)。

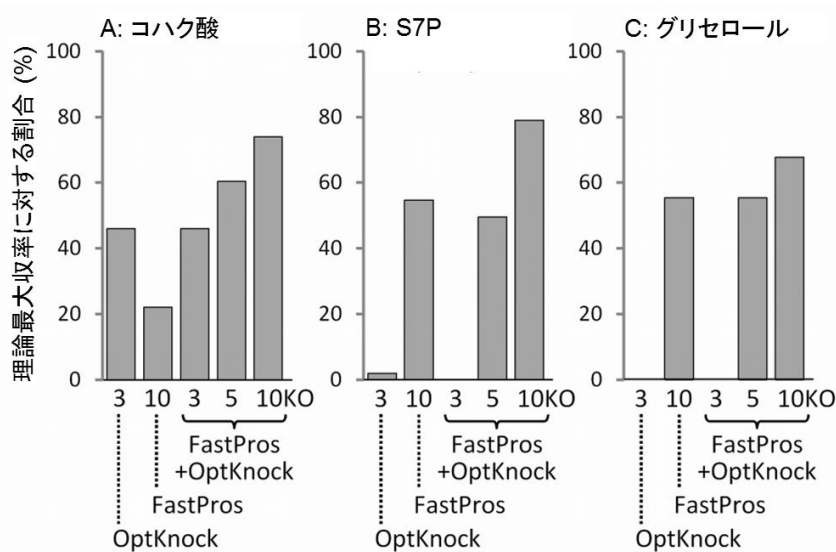


Fig. 3.9 コハク酸 (A)、セドヘプツロース7リン酸 (S7P) (B)、グリセロール (C) の予測された生産収率

最大の遺伝子破壊数は、OptKnock では3、FastPros では10、FastPros の結果に基づく OptKnock (FastPros+OptKnock)では 3、5、10 とした。FastPros で得られた破壊候補遺伝子を、可能な破壊対象として OptKnock を実行することで、高い生産収率を目指して遺伝子破壊の組み合わせが最適化された。

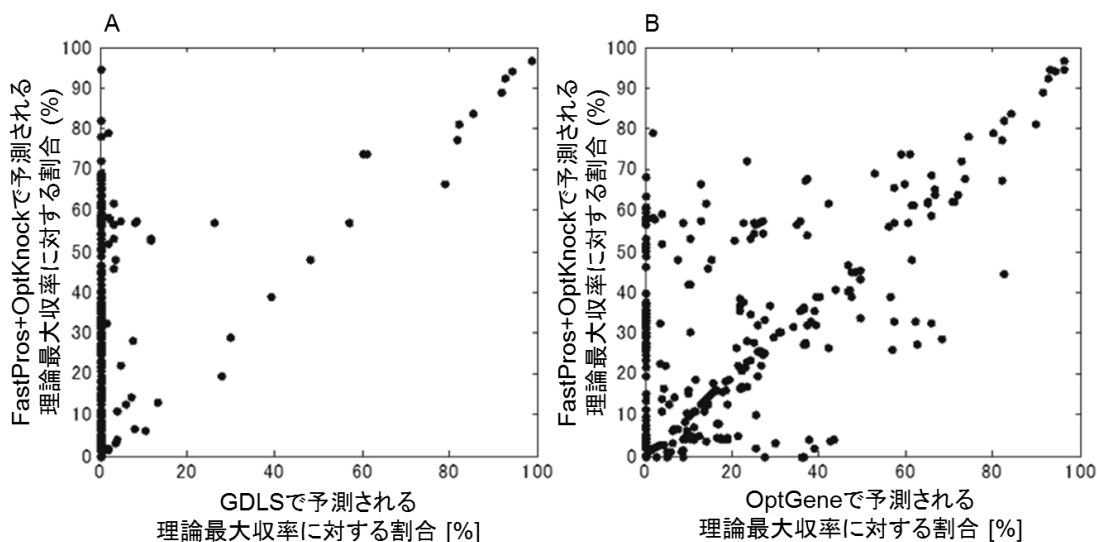


Fig. 3.10 FastPros の結果に基づく OptKnock で予測される生産収率と、GDLS (A) あるいは OptGene (B) での生産収率の比較

380 の代謝物質について、予測される生産収率をプロットしている。FastPros の結果に基づく OptKnock は、既存の探索手法である GDLS または OptGene と同等以上の生産収率が予測された。

3.3.5 工学的に有用な代謝物質生産に向けて見つけ出された多重遺伝子破壊

FastPros とその結果に基づく OptKnock により、破壊候補遺伝子が同定された代謝物質の中には、プラスチック原料・アミノ酸・薬剤原料など、工学的に有用なものも含まれていた。Table 3.2 にその例を示す。ここでは、GPP と L-フェニルアラニン (L-Phe) について、見つけ出された遺伝子破壊が代謝に及ぼす影響について詳細に議論する。

Table 3.1 FastPros の結果に基づく OptKnock により、破壊候補遺伝子が予測された有用化合物の例

化合物名	用途	遺伝子破壊の数	理論最大収率に対する割合 [%]
乳酸	生分解性プラスチックであるポリ乳酸の原料	10	97
1,2-プロパンジオール	保湿剤・溶媒・プラスチックの中間原料	10	74
ゲラニルピロリン酸 (GPP)	テルペノイド合成時の中間体	7	53
L-フェニルアラニン (L-Phe)	甘味料アスパルテームの原料	10	58
L-バリン	化粧品・医薬品・飼料の原料	6	93
シキミ酸	抗インフルエンザ薬タミフルの原料	10	57
アロース-6-リン酸	抗酸化作用・ガン細胞増殖抑制作用を持つ希少糖アロースの前駆体	10	57

3.3.5.1 ゲラニルピロリン酸 (GPP)

GPP は、香水等に広く使われているゲラニオールの前駆体である。また、GPP はテルペノイド生合成経路 (非メバロン酸経路) の中間体でもある。したがって、GPP の生合成フラックスを上昇させる遺伝子破壊は、大腸菌によるゲラニオールやテルペノイドの生産性向上に有望だと期待される。

GPP を生産ターゲットとして FastPros を実行した結果、25 の破壊候補遺伝子からなる 88 の組み合わせが同定された。それぞれの破壊候補遺伝子の組み合わせは、増殖速度を最大化する条件下で GPP が生産されることを意味している。GPP の高い生産性を目指し、これら 25 遺伝子を破壊の対象として OptKnock を実行したところ、*adhE*、*ldhA*、*ppc*、*mgsA*、*pntA*、*edd*、*sdhA* の 7 遺伝子の同時破壊により、0.34 g/g-glucose (理論最大値の 53%) での GPP 生産が予測された。Fig. 3.11 に、野生型および GPP 生産に向けた 7 重遺伝子破壊後の代謝モデルで計算されるフラックス分布を示す。1 mol の GPP は、2 mol のグリセルアルデヒド 3 リン酸 (G3P) と 2 mol のピルビン酸より合成され、4 mol の NADH と 2 mol の NADPH の酸化を伴う。したがって、GPP の生産には、NADH または NADPH の酸化に関して競合する反応経路が阻害された。具体的には、*adhE*、*ldhA*、*ppc* の破壊により、それぞれエタノール、乳酸、コハク酸の生合成経路が阻害された。メチルグリオキサール経路中の反応をコードする *mgsA* もまた、NADH の酸化を伴う 1,2-プロパンジオール生合成を阻害するために破壊された。Transhydrogenase 反応をコードする *pntA* 遺伝子については、その遺伝子の破壊により NADH から NADP⁺ への電子の伝達を不活性化し、NADPH の消費を伴う L-バリン生産が阻害された。最後に、*edd* 遺伝子の破壊により、Entner-Doudroff (ED) 経路を通じた解糖系下流へのフラックスが阻害され、GPP 生産に必要な G3P が供給された。上記した 6 遺伝子の同時破壊は、この 7 遺伝子破壊の中で GPP 生産に必須の遺伝子破壊であり、いずれかの破壊が欠けていた場合には GPP は生産されなかった。これより、第 2 章での大域的探索や従来の局所的探索では、この 6 重遺伝子破壊を見つけ出すには大きな計算コストを必要とすることが分かる。また、この 6 重遺伝子破壊は GPP 生産収率の大部分 (0.31 g/g-glucose) に寄与し、*sdhA* 遺伝子の破壊は増殖速度をわずかに減少させ、余剰の炭素により GPP へのフラックスが増加した。

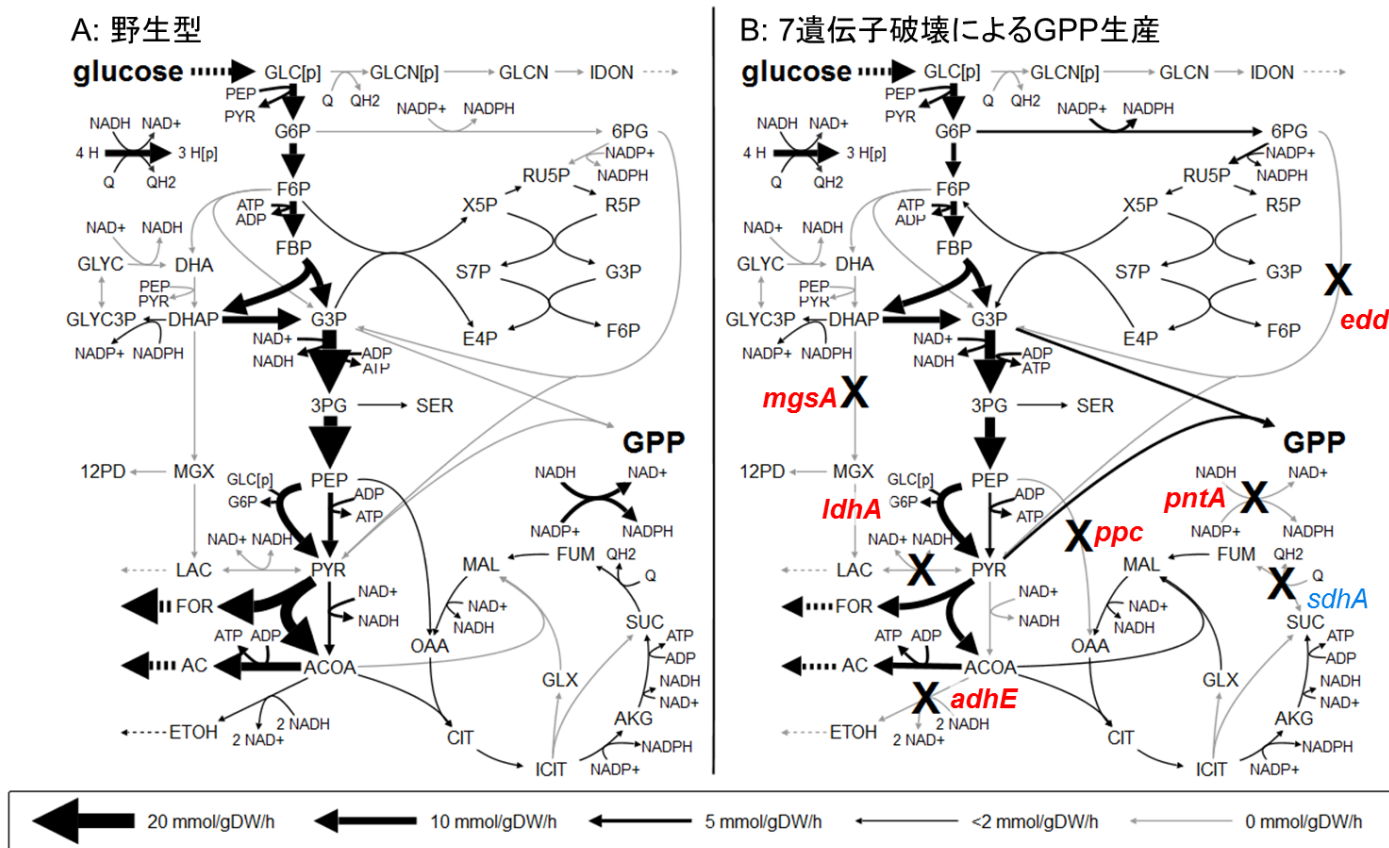


Fig. 3.11 野生型 (A) およびゲラニルピロリン酸 (GPP) 生産に有望な7 遺伝子破壊 (B) に対応する代謝フラックス分布

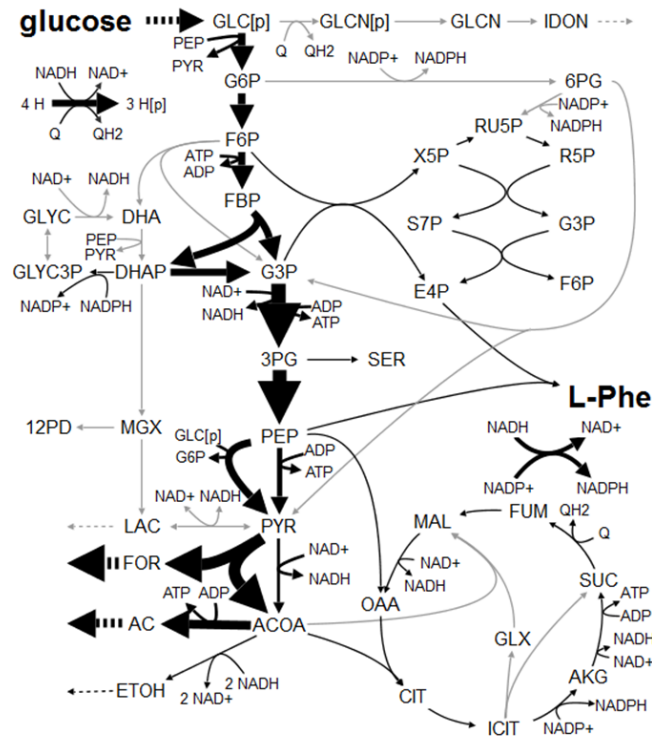
矢印の太さはフラックスの大きさを示し、灰色の矢印はフラックスが0であることを意味する。実線の矢印は代謝反応を意味し、破線の矢印はトランスポート反応を意味する。×印は、赤色あるいは青色の遺伝子の破壊による反応の不活性化を意味する。赤色の6 遺伝子の破壊は、この7 遺伝子の破壊の中でGPP 生産に必須である。代謝物質の略称については、付録のTable A2 参照。遺伝子破壊により、NADH またはNADPH の酸化を伴う反応経路が阻害され、余剰の還元力を利用してGPP 生産が促進した。

3.3.5.2 L-フェニルアラニン (L-Phe)

芳香族アミノ酸である L-Phe は、甘味料アスパルテームの原料であり、食品や製薬の産業で広く利用されている。化学的にはラセミ体として Phe が合成されるため、純粋な L-Phe の生物工学的な生産プロセスは重要である。大腸菌は元来 L-Phe の生合成経路を所持し、商業生産に利用されてきた。

L-Phe 生産に向けて FastPros を実行し、44 の遺伝子の 57 の組み合わせが生産性向上に有望な破壊候補遺伝子として見つけ出された。これらの遺伝子より、高い L-Phe 生産収率に寄与した上位 30 の遺伝子を選択し、OptKnock を実行した。結果、10 遺伝子の破壊により 0.27 g/g-glucose (理論最大値の 58%) で L-Phe 生産が予測された。破壊された遺伝子は、*pykF*、*fsaA*、*ppc*、*zwf*、*serC*、*gldA*、*pta*、*adhE*、*nuoN* および *gntT* である。Fig. 3.12 に、野生型および L-Phe 生産に向けた 10 遺伝子が破壊された代謝モデルでシミュレートされるフラックス分布を示す。1 mol の L-Phe は、2 mol のホスホエノールピルビン酸 (PEP) と、ペントースリン酸経路中の 1 mol のエリスロース 4 リン酸 (E4P) から合成され、2 mol の NADPH の酸化を伴う。そのため、*pykF*、*fsaA*、*ppc* 遺伝子の破壊により、PEP を直接消費する反応が阻害され、L-Phe 合成の前駆体となる PEP が供給された。*zwf* の破壊は ED 経路を通じたピルビン酸への炭素供給を阻害し、*serC* の破壊は Serine deaminase 反応を通じたピルビン酸の供給を阻害した。この時、細胞増殖に必要な L-セリンはグリシンから合成され、グリシンは L-スレオニンの分解によって合成された。*gldA* 遺伝子の破壊は、Dihydroxyacetone phosphotransferase 反応による PEP からピルビン酸への供給を阻害した。*pta* 遺伝子は ATP 合成を伴う酢酸の生合成経路の反応をコードし、この破壊により解糖系中の Phosphoglycerate kinase 反応での ATP 合成が促進され、L-Phe 合成に必要な PEP が供給された。*adhE* 遺伝子は NADH 合成を伴うエタノール生合成経路の遺伝子であり、この破壊により Transhydrogenase 反応を通じた NADH から NADP⁺への電子伝達が促進し、L-Phe 合成に必要な NADPH が供給された。これらの 8 遺伝子の同時破壊は、この 10 遺伝子破壊の中で L-Phe 生産に必須の遺伝子破壊であり、またおおよそ全ての L-Phe 生産収率 (0.26 g/g-glucose) に寄与した。*nuoN* と *gntT* 遺伝子の破壊は、細胞の増殖速度をわずかに減少させ、余剰の炭素により L-Phe 生産が促進した。*nuoN* 遺伝子破壊のみでは、*gntT* がコードするグルコン酸トランスポーターを通じてイドン酸合成が活性化するため、L-Phe 生産は上昇しなかった。

A: 野生型



B: 10遺伝子破壊によるL-Phe生産

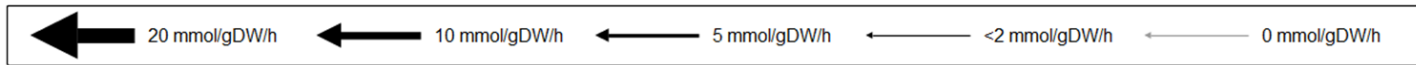
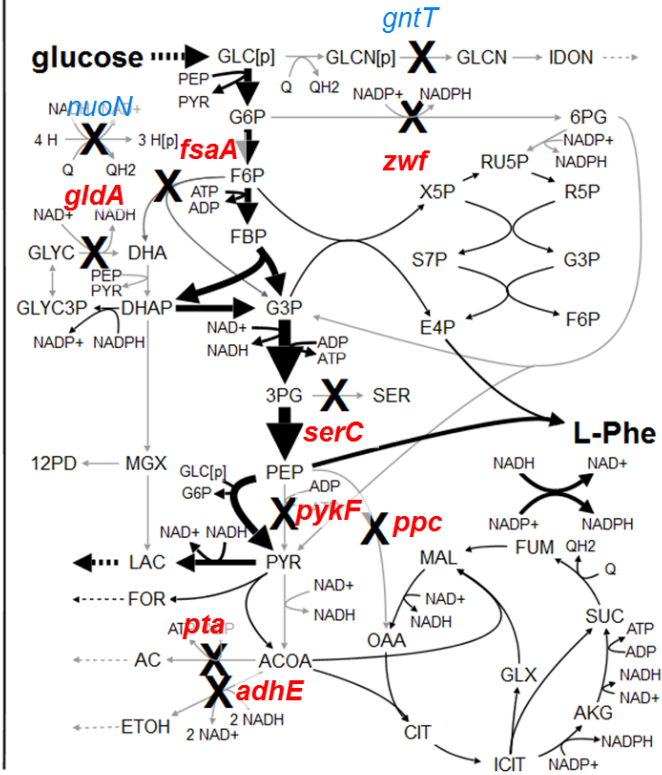


Fig. 3.12 野生型 (A) およびL-フェニルアラニン (L-Phe) 生産に有望な 10 遺伝子破壊 (B) に対応する代謝フラックス分布

凡例は Fig. 3.11 と同様である。赤色の 8 遺伝子の破壊は、この 10 遺伝子の破壊の中で L-Phe 生産に必須である。代謝物質の略称については、付録の Table A2 参照。遺伝子破壊により、PEP または解糖系上流の代謝物からピルビン酸への変換が阻害され、余剰の炭素を用いて PEP と E4P からの L-Phe 生産が促進した。

3.3.6 複数の代謝物質生産に共通して予測された破壊候補遺伝子の解析

FastPros による多重遺伝子破壊の探索を、細胞内の様々な代謝物質に対して適用することで、従来の手法よりも多くの代謝物質について有望な破壊候補遺伝子が同定された。そこで、代謝物質間で共通する破壊候補遺伝子について解析した。複数の代謝物質生産に共通して有望な破壊候補遺伝子が同定されれば、様々な有用化合物生産に向けた育種において、共通する親株をデザインすることが可能になる。この親株に対して、それぞれのターゲット化合物生産に特有の遺伝子破壊を適用することで、生産株を効率よく構築することが可能になると期待される。

そこで、FastPros の結果に基づく OptKnock で同定された破壊候補遺伝子について、その類似性を Jaccard 距離で評価し、それに基づく代謝物質の階層的クラスタリングを行った。この計算には、MATLAB の Statistics Toolbox 中の関数を使用した。クラスタリングの結果として得られた樹形図を Fig. 3.13 に示す。いくつかの代謝物質のクラスターが観察され、同じクラスター内の代謝物質は共通する特徴が見られた。例えば、Fig. 3.13 の Cluster 1 は、103 の代謝物質から構成され、それらのほとんどは脂質やその派生物 (デカン酸、ホスファチジルグリセロール、GPP など) だった。このクラスターでは、*adhE*、*ldhA*、*mgsA*、*ppc* 遺伝子の破壊が、90%以上の代謝物質間で共通しており、これらの遺伝子破壊はこのクラスター内の代謝物質生産における基本的な代謝改変戦略となりうる。脂質の生合成経路における炭素鎖伸長では、NADPH や NADH などの還元力が要求される。したがって、GPP 生産の場合で議論したように、細胞内の NADH または NADPH の酸化反応に関わる遺伝子が破壊されることで、脂質合成が促進された。Cluster 2 は芳香族アミノ酸やその生合成経路中の中間体である 15 の代謝物質 (3-デヒドロシキミ酸、L-Phe など) から構成された。共通する遺伝子破壊は、*ppc*、*pykF*、*mgsA*、*serC* だった。L-Phe 生産の場合で議論したように、芳香族アミノ酸合成において PEP は重要な前駆体である。そのため、PEP を消費する反応の遺伝子が破壊され、余剰の PEP を利用して芳香族アミノ酸生合成経路のフラックスを上昇させる代謝改変戦略が提案された。先行研究より、実験的に大腸菌に L-Phe などの芳香族アミノ酸を生産させる場合には、*ppc* 遺伝子や *pykF* 遺伝子が一般的に破壊され (Ikeda, 2006; Bongaerts *et al.*, 2001)、本章の結果と一致する。Cluster 3 では、糖と糖リン酸から主に構成される 29 の代謝物質からなり、*ppc*、*pykF*、*serC*、*zwf* の破壊が共通していた。これらの遺伝子破壊により、解糖系やペントースリン酸経路中の代謝物質からピルビン酸への変換が妨げられ、解糖系上流からの糖あるいは糖リン酸の合成が増加した。

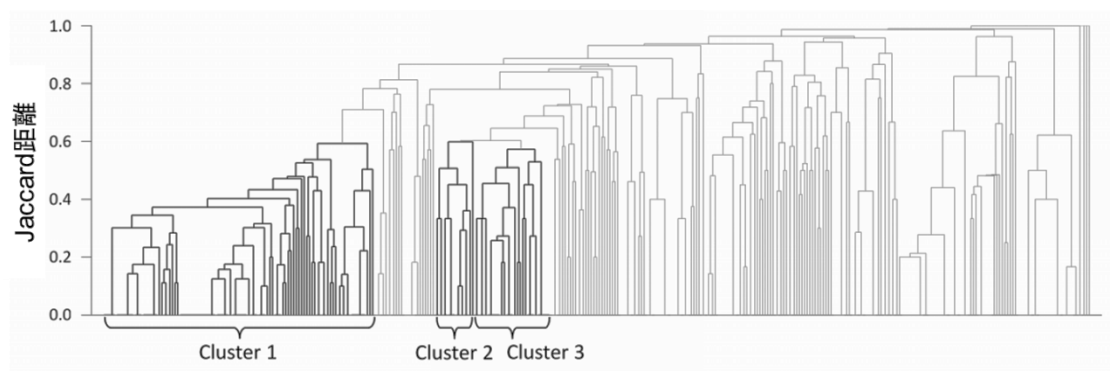


Fig. 3.13 同定された破壊候補遺伝子の類似性に基づく代謝物質の樹形図

それぞれのクラスターの特徴は、Table 3.2 参照。

Table 3.2 Fig. 3.13 のクラスターを構成する代謝物質と共通する破壊遺伝子

	Cluster 1	Cluster 2	Cluster 3
主要な代謝物質	脂質とその派生物 (デカン酸、GPP、ホスファチジルグリセロール、etc.)	芳香族アミノ酸とその前駆体 (3-デヒドロシキミ酸、4-ヒドロキシ安息香酸、L-Phe、etc.)	糖と糖リン酸 (リボース 1 リン酸、アロース 6 リン酸、トレハロース、etc.)
代謝物質の用途	燃料、テルペノイド、薬剤合成における中間体	飼料や食品の添加物薬剤合成における中間体	希少糖、食品添加物
代謝物質数	103	15	29
共通する破壊遺伝子	<i>adhE</i> 、 <i>ldhA</i> 、 <i>mgsA</i> 、 <i>ppc</i>	<i>ppc</i> 、 <i>pykFA</i> 、 <i>mgsA</i> 、 <i>serC</i>	<i>ppc</i> 、 <i>pykFA</i> 、 <i>serC</i> 、 <i>zwf</i>

3.4 考察

本章では、育種に向けた多重遺伝子破壊を探索する手法 FastPros を開発し、大腸菌細胞内の様々な代謝物質の生産に対して、生産性向上に有望な破壊候補遺伝子が同定された。これを可能にしたのは、増殖速度とターゲット生産速度を関連付ける u_{TARGET} を新たに提案し、探索における指標として採用したからである。 u_{TARGET} の正負は、増殖速度最大化の条件下でターゲットが生産されるか否かに対応し、 u_{TARGET} が負の場合でも 0 に近いときは、新たに遺伝子破壊を加えることでターゲット生産が予測されやすいとみなした。これにより、大域的探索の適用が難しい、6 遺伝子以上の同時破壊により初めてターゲット生産が予測される場合でも、破壊候補遺伝子の探索を可能にした。

FastPros による破壊候補遺伝子の探索は、代謝工学における他の *in silico* の手法と共に利用することが可能である。例えば、本章では大腸菌が生産し得る代謝物質をターゲットとしたが、異種生物種由来の生合成経路による新規化合物生産に向けても、FastPros を適用できる。近年、ターゲット化合物の生合成経路を探索する手法が開発されており (Chatsurachai *et al.*, 2012; Cho *et al.*, 2010; Pharkya *et al.*, 2004)、これらの手法で同定された生合成経路を代謝モデルに導入することで、破壊候補遺伝子の探索が可能である。別の例として、FastPros によって予測された破壊候補遺伝子や代謝フラックス分布は、他の破壊候補遺伝子の探索手法 (OptKnock や OptStrain (Pharkya *et al.*, 2004)) における初期状態として使用し、探索を高速化することが可能である。高いターゲット生産性を導く遺伝子破壊の組み合わせをあらかじめ用意しておけば、効果的に準最適解を切り取ることができ、結果として大域最適解を見つけ出すのに必要な計算時間が減少する (Kim *et al.*, 2011)。さらに、先行研究においてターゲット生産性に寄与する遺伝子発現増強を同定する戦略も提案されている (Pharkya and Maranas, 2006; Choi *et al.*, 2010)。この戦略を FastPros と組み合わせれば、遺伝子の破壊と発現増強のどちらをも含む遺伝子操作を探索し、高い生産収率に有望な遺伝子操作を予測することが可能となることが期待される。

FastPros では、細胞の増殖速度を最大化する、すなわち細胞は自身の増殖にとって最適な代謝フラックス分布とることが仮定されている。もちろん、この仮定は実際の細胞内で必ず成り立つわけではない。予測された遺伝子破壊により細胞の増殖能が著しく低下し、破壊株を実験的に獲得することが難しい場合には、そのような遺伝子破壊が選ばれない条件で、再び探索を行うべきである。合成致死に関するデータ (Typas *et al.*, 2008) を FastPros に統合すれば、予測と実験の増殖能の差を解消することが可能になると推測される。また、FBA を実行する際に細胞の増殖速度の下限値 (page 49 の v_{growth}^{min}) を大きく設定すれば、増殖速度の大きな減少を引き起こす遺伝子破壊が同定されるのを避けることができると考えられる。

3.5 結言

本章では、育種に向けた多重遺伝子破壊を探索する手法 **FastPros** を開発した。**FastPros** では、探索における指標として u_{TARGET} を新たに提案した。これは、線形計画問題におけるシャドウプライスとして計算される値であり、細胞の増殖速度とターゲットの生産速度を関連付けたものである。**FastPros** は、従来の大域的探索手法 (**OptKnock** など) および局所的探索手法 (**GDLS** など) と比較し、特に化合物生産に必要な破壊数が比較的が多い (5 遺伝子以上) 場合に、破壊候補遺伝子を同定できることが示された。遺伝子破壊に対応する代謝フラックス分布のシミュレーションは、従来手法と **FastPros** で共通するため、**FastPros** による予測は従来手法と同程度の現実性を保ちながら、破壊候補遺伝子を探索可能な範囲を拡大できたといえる。また、複数のターゲット生産に共通して予測された遺伝子破壊が同定され、様々な有用化合物生産に向けた育種において、共通する親株をデザインすることが可能になることが示唆された。開発した探索手法と u_{TARGET} の概念により、代謝工学的な技術が発展し、幅広い有用化合物の生物学的な生産が促進することが期待される。

第4章 結論

4.1 結果のまとめ

微生物を用いた有用化合物生産は、近年の地球環境問題に対する関心の高まりを受けて、これまでも増して注目されており、産業利用可能な菌株を効率的かつ合理的に育種するための手法の開発が求められている。遺伝子破壊で不要な代謝経路を削除し、ターゲット化合物生産に向けて細胞内の代謝フラックス分布を最適化するにあたり、計算機上での Flux balance analysis (FBA) などの代謝シミュレーションを利用し、破壊候補遺伝子を絞り込む手法が開発されてきた。FBA で想定されている増殖速度最大化の条件下で高い生産収率が予測される遺伝子破壊は、これまでの多くの先行研究で、実際の微生物での生産性を向上させたことが報告されてきた。しかし、同時に破壊する遺伝子数が増加すると、その組み合わせが指数関数的に増加し、現実的な計算時間以内での探索が困難になることが課題となっている。破壊候補遺伝子の探索に対して FBA による代謝シミュレーションを適用する場合には、大域的探索手法で同時に破壊可能な遺伝子数の上限を向上させるとともに、取りこぼしの少ない局所的探索手法を開発する必要がある。そこで本研究では、ゲノムスケールの代謝モデルと FBA を利用し、増殖速度最大化の条件下で高いターゲット生産収率が予測される多重遺伝子破壊の、大域的探索手法および局所的探索手法の開発を目指した。以下に、本研究で得られた知見を簡単にまとめる。

第 2 章では、ターゲット化合物生産に有望な多重遺伝子破壊を大域的探索を高速化する手法を開発した。破壊対象となる遺伝子の数の削減により探索の高速化を実現するため、栄養条件に応じて宿主微生物のゲノムスケール代謝モデルを自動的に縮約する方法を提案した。具体的には、解析したい栄養条件下で使用される可能性のない反応に関わる遺伝子や、細胞増殖に必須の遺伝子を計算機上で自動的に特定し、それらを破壊対象より除くといったことを行った。結果、例えば 1260 遺伝子からなる大腸菌の代謝モデルでは、グルコース炭素源好気条件下での破壊対象の遺伝子数は 393 まで減少し、3 重破壊の探索時間は、 10^7 秒から 10^2 秒に短縮された。現実的な時間以内に計算が可能だった遺伝子破壊数の上限について、もとの代謝モデルでは 2 遺伝子破壊だったのに対して、開発手法により 5 遺伝子破壊程度まで可能であることが推測された。

開発手法の有効性を評価するため、宿主微生物として広く利用されている大腸菌を選択し、ターゲット化合物として、燃料やプラスチック原料として有用な 1-ブタノール、1-プロパノール、1,3-プロパンジオールの 3 種類を採用した。そして、破壊候補遺伝子の探索を行い、3 重破壊の場合について詳細に解析した。結果、例えばグルコース炭素源好気条件下で 1,3-プロパンジオールの生産収率が最大となる 3 重遺伝子破壊として、*tpiA*、*ptsG*、*pykF* の同時破壊が予測され、解糖系および TCA サイクルの補充経路のフラックスを変化させる代謝改変戦略が予測された。1,3-プロパンジオールは 1 遺伝子破壊では生産が予測されなか

ったため、従来の単純な局所的探索のように、1 遺伝子破壊で高い生産収率が予測されたものの選出を繰り返す方法では、この 3 重遺伝子破壊を見つけ出すことは難しいといえる。同定された破壊候補遺伝子には、実験的にターゲット生産性向上に寄与することが報告されているものも含まれており、先行研究での代謝改変戦略が適切であることが、代謝モデルを用いた解析から裏付けされた。

第 3 章では、大域的探索の適用が難しい多重遺伝子破壊を局所的に探索する手法 **FastPros** を開発した。従来の局所的探索手法では、探索指標としてターゲット化合物の生産収率が採用されており、複数の遺伝子の同時破壊により初めてターゲット生産が上昇する場合には探索が困難だった。そこで、本研究では、ターゲット生産速度と増殖速度を関連付ける新規探索指標として u_{TARGET} を提案し、ターゲット生産性向上に寄与する可能性の高い破壊候補遺伝子を同定した。そして **FastPros** では、この新規探索指標が増加するように繰り返し遺伝子を破壊することで、ターゲット生産に有望な破壊候補遺伝子の同定を行った。**FastPros** の性能を評価するため、大腸菌の細胞内の 625 の代謝物質を生産ターゲットとして、**FastPros** を実行した。その結果、472 の代謝物質について、増殖速度最大化の条件下で高い生産収率が予測される破壊候補遺伝子が同定された。3 重破壊の大域探索により破壊候補遺伝子が同定される化合物は 152 のみだったことから、**FastPros** による探索は、より多くのターゲット化合物に対して適用可能であることが示された。 u_{TARGET} を探索指標とした場合の適応度地形についても解析され、**FastPros** での探索で破壊候補遺伝子を同定するには、ある程度の探索範囲を広くとるが必要であり、また、必ずしもすべての破壊候補遺伝子を同定できるわけではないことが示された。

さらに、**FastPros** と既存の探索手法 **OptKnock** を組み合わせて用いることで、より高い生産収率が予測される多重遺伝子破壊を見つけ出すことに成功し、ガラニルピロリン酸と L-フェニルアラニンを生産ターゲットとした場合について、遺伝子破壊に伴う代謝変動に関して具体的に議論した。最後に、複数の代謝物質生産に共通して予測された破壊候補遺伝子が同定され、様々な有用化合物生産に向けた育種において、共通して利用可能な親株のデザインが提案された。

4.2 得られた知見の寄与するところ

本学位論文では、微生物育種に向けた多重遺伝子破壊の探索手法を開発した。以下に、得られた知見が寄与するところについて述べる。

一般に、工学的に求められる化合物を微生物に生産させることは、その微生物にとって都合が悪いことが多い。細胞増殖に必要な栄養源を、別の化合物として細胞外に排出することになるため、生産性の高い細胞は増殖速度が低下する傾向にある。したがって、生産能の高い変異株の培養中に、変異や遺伝子発現量の摂動により生産能の低下と増殖能の回復が生じて、その細胞が集団の多数を占めてしまう可能性が存在する (Summers, 1991; Siegel and Ryu, 1985)。しかし、FBA を用いた破壊候補遺伝子の探索では、増殖速度最大化の条件下で高いターゲット生産収率が予測される代謝状態が設計されるため、ターゲット化合物生産が細胞増殖にとっても都合のよいように、代謝を設計することが可能である。したがって、本研究で開発された手法により遺伝子破壊を探索し、化合物生産と細胞増殖のどちらにも適した代謝を設計すれば、増殖連動での生産プロセスにおいて、生産性の安定した株を構築できると期待される。

本研究では、遺伝子破壊を評価するためにターゲット化合物の生産収率を採用したが、微生物を用いた有用化合物生産では、生産速度や生産濃度なども重要な評価指標である (バイオ燃料技術革新協議会、2008; DOE, USA, 2004)。生産速度に関しては、予測された生産収率 (消費した栄養源量あたりの生産量) に対して、栄養源の消費速度と菌体濃度より算出することが可能である。例えば、大腸菌のグルコース消費速度を 10.57 mmol/gDW/h (Holms, 1996) とすると、第 2 章の Table 2.10 で示されている最大の 1-ブタノール生産収率 ($0.303 \text{ g/g-glucose}$) とグルコースの分子量を用いて、比生産速度 (単位菌体量あたりの生産速度) は 0.57 g/gDW/h と算出できる。そして、菌体濃度を 3 g/L (Shen *et al.*, 2011) とすると、算出された比生産速度と菌体濃度の積を計算し、生産速度を 1.71 g/L/h と計算可能である。ただし、一般に FBA では、グルコースなどの栄養源の消費速度はシミュレーションにおける入力値として使用されるものであり、破壊後の株のグルコース消費速度の予測は難しいのが現状である。また、生産濃度や菌体濃度に関しても、生産濃度に応じた増殖阻害等が影響し、濃度情報を捨象した本研究のシミュレーション手法では考慮することができない。したがって、生産速度や生産濃度も予測したい場合には、FBA よりも計算コストがかかるが酵素の速度論等を加味している代謝シミュレーション手法を適用し、大規模な代謝改変により生産収率向上を目指す場合には、本研究で開発したような、FBA を利用した多重遺伝子破壊の探索手法を活用することが有効だと考えられる。

産業利用可能な微生物を育種する際には、遺伝子操作による株の構築と、培養・分析による株の評価が繰り返し行われる。この際に、単独の遺伝子の操作では表現型に影響しないが、そのような遺伝子を複数操作した場合には表現型に影響するような場合が十分に考えられる。例えば、大腸菌の *ompA* 破壊株と *pal* 破壊株の増殖挙動は、野生型株とほぼ同じ

であるが、*ompA* と *pal* の 2 重破壊株は、誘導期が長期化することが報告されている (Typas *et al.*, 2008)。したがって、単独の遺伝子操作が表現型に及ぼす影響を評価するだけでなく、理想的には複数の遺伝子操作が表現型に及ぼす影響を評価することが重要だと考えられる。本研究で開発した探索手法は、計算機上で多数の遺伝子の同時破壊の影響を評価することが可能であり、従来の探索手法では同定できない多重遺伝子破壊を、育種の研究者に提案することができる。

本研究では、育種に向けた遺伝子破壊の探索手法として大域的探索手法と局所的探索手法の 2 つを開発した。大域的な探索手法では、代謝モデル上で可能な遺伝子破壊が化合物生産に及ぼす影響を網羅的に評価されるが、探索可能な遺伝子破壊数に限界がある。一方で、局所的な探索手法では、多数の遺伝子の同時破壊の影響を評価できる一方で、より優れた遺伝子破壊を取りこぼす可能性が常に存在する。FastPros についても、 u_{TARGET} を探索指標とした場合の適応度地形に凹凸が存在したことから、破壊候補遺伝子を取りこぼしている可能性が存在することが示唆された。また、実際に、FastPros の結果に基づく OptKnock と、先行研究で報告された探索手法を比較した場合に、先行研究の方が生産収率の高い遺伝子破壊の組み合わせを同定できている化合物も多く存在した (Fig. 3.10)。したがって、本研究で使用したものと同等程度の計算機を利用する場合には、5 遺伝子破壊程度までは大域的探索により可能な破壊の組み合わせを網羅的に評価し、それ以上の破壊の場合は FastPros を利用することで有望な破壊候補遺伝子を探索することが適当だと思われる。

4.3 今後の展望

本論文では、代謝反応の化学量論式から構築される代謝モデルと、代謝の定常と増殖速度の最大化により代謝フラックスをシミュレートする FBA を利用し、育種に向けた多重遺伝子破壊の探索手法を開発した。この代謝モデルと FBA による代謝シミュレーション手法では、細胞内での代謝物質濃度、酵素の速度論、遺伝子発現や酵素活性の制御等の情報を捨象し、代謝の定常における物質収支と増殖速度の最適化に着目することで、線形計画法により高速で代謝のシミュレーションが可能となった。これにより、育種に向けた破壊候補遺伝子を、ゲノム全体から探し出せるようになり、細胞全体の代謝フラックス分布の最適化が可能になりつつある。しかし実際には、中間代謝物質が蓄積したり酵素の性能の関係で、現実の細胞内の代謝フラックス分布が予測結果と異なる場合も十分に考えられる。

予測と実験の齟齬を解消する戦略の一つは、代謝モデルやシミュレーション手法を改良し、実際の微生物の代謝をより正確に予測できるようにすることである。本論文では一般的なシミュレーション手法として FBA を使用したが、最適化問題を利用して代謝をより正確に計算機上で表現する手法が他にも研究されてきた (Lewis *et al.*, 2012)。例えば、細胞内の混雑さや細胞膜中のタンパクの占有に着目し、代謝フラックスや電子伝達系の反応速度

に現実的な制限を課すことで、予測精度が向上することが、先行研究により報告されている (Vazquez *et al.*, 2008; Zhuang *et al.*, 2011)。遺伝子発現の制御やシグナル伝達を制約条件として加える手法も提案されている (Covert *et al.*, 2004, 2008)。目的関数についての研究も進んでおり、野生型株の代謝については増殖速度最大化の他に、ATP 生産の最大化やフラックスの総和の最小化が (Schuetz *et al.*, 2007, 2012)、変異株の代謝については、変異に伴う代謝変動の最小化などが提唱されている (Segrè *et al.*, 2002; Shlomi *et al.*, 2005; Kim and Reed, 2012)。工学的には、特に遺伝子操作に伴う代謝フラックス分布の変化を正確に予測できることが重要であるが、多数の遺伝子操作の組み合わせが細胞の表現型に与える影響については、不明な点が多かった。しかし、近年では遺伝子操作や表現型の評価をハイスループットで行う技術が開発されつつある (Wang *et al.*, 2009)。予測と実験との齟齬が明らかになり、その情報をもとに細胞の代謝をより正確にシミュレートできるようになれば、ターゲット生産に有望な遺伝子操作を計算機上で探索する場合に、信頼性の高い現実的な予測が可能になると推測される。

予測と実験の不一致を解消する別の戦略は、計算機上で予測される代謝フラックス分布を目標として、細胞内の代謝を実験的に改変するものである。例えば、実験的な代謝フラックス解析手法で推定される代謝フラックス分布と、FBA で予測される代謝フラックス分布とを比較し、差の大きい反応の活性化あるいは不可性化することで、予測される代謝フラックス分布を細胞内で実現する手法が考えられる。実際に、大腸菌による脂肪酸生産に対して、目標と実際の代謝フラックス分布が比較され、差の解消に向けて同定された *fabZ* の発現増強を行った結果、脂肪酸の生産性が実験的に向上されたことが報告されている (Ranganathan *et al.*, 2012)。また、細胞内の代謝物質濃度を網羅的に定量する技術 (メタボローム分析) も発展し (Bennett *et al.*, 2009)、理想とする代謝フラックス分布を達成するのに律速となっている反応を、より合理的に同定することも可能になりつつある。別の方法として、人工進化実験を利用した代謝改変も提案されている。先行研究で、微生物株に対して長期間の継代培養を行い、その培養環境下で増殖速度の高い細胞が選択されるように菌株を人工的に進化させると、細胞内の代謝状態は計算機上で予測された最適な代謝状態に近づくことが報告されている (Fong *et al.*, 2003; Lewis *et al.*, 2010)。また Fong らは、増殖速度最大化の条件下で乳酸生産性の高い大腸菌の破壊株を予測し、その株の人工進化実験を行うことで、予測されたように乳酸生産性を向上させることにも成功した (Fong *et al.*, 2005)。もちろん、継代培養が予測と実験の不一致を解消するのに有効であるためには、線形の制約式で定義される代謝の取りうる範囲 (Fig. 3.1) において、実際の細胞が安定した代謝状態をとりながら、局所的に増殖速度が最大となる領域に陥らずに、破壊株の代謝状態が FBA で予測される代謝状態まで推移する必要がある。現在の知見では、代謝の安定性や実験室進化の有効性については明らかではないが、メタボローム分析や代謝フラックス解析、酵素の速度論を考慮した代謝モデル等を用いて、実験室進化に伴う細胞内の代謝状態の推移を詳細に解析すれば、代謝の安定性や実験室進化の有効性について議論可能になると期待

される。以上のように、FBA で予測される代謝状態を目標として株を構築し、代謝改変後の株を解析することで、予測と実験の差を解消できる。

本論文では、育種に向けた多重遺伝子破壊の探索手法を開発し、大腸菌を宿主微生物として利用した場合に探索可能な破壊遺伝子数や破壊候補遺伝子を同定できる代謝物質に関して議論した。開発した手法は、化学量論式からなる代謝モデルと FBA に基づき、増殖速度最大化の条件下でターゲット化合物の高い生産収率が予測される破壊候補遺伝子を探索することができる。代謝モデルを別の生物種のものに置き換えたり、栄養源やターゲット化合物を別のもので変更することで、異なる生物種・栄養条件・ターゲット化合物に対して、開発手法を適用可能である。今後の微生物育種での代謝改変において、有望な破壊候補遺伝子を同定する手法として汎用的に利用され、産業的に有益な表現型を持つ株が合理的かつ効率的に獲得可能になることが期待される。

参考文献

- Alper, H., Jin, Y.-S., Moxley, J. F., & Stephanopoulos, G. N. (2005). Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.*, **7**(3), 155–164.
- Atsumi, S., Cann, A. F., Connor, M. R., Shen, C. R., Smith, K. M., Brynildsen, M. P., Chou, K. J. Y., Hanai, T., & Liao, J. C. (2008). Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.*, **10**(6), 305–311.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
- Becker, J., Zelder, O., Häfner, S., Schröder, H., & Wittmann, C. (2011). From zero to hero--design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metab. Eng.*, **13**(2), 159–168.
- Becker, S. a, Feist, A. M., Mo, M. L., Hannum, G., Palsson, B. Ø., & Herrgard, M. J. (2007). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.*, **2**(3), 727–738.
- Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., & Rabinowitz, J. D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.*, **5**(8), 593–599.
- Bongaerts, J., Krämer, M., Müller, U., Raeven, L., & Wubbolts, M. (2001). Metabolic engineering for microbial production of aromatic amino acids and derived compounds. *Metab. Eng.*, **3**(4), 289–300.
- Burgard, A. P., Pharkya, P., & Maranas, C. D. (2003). OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.*, **84**(6), 647–657.
- Chaturachai, S., Furusawa, C., & Shimizu, H. (2012). An *in silico* platform for the design of heterologous pathways in nonnative metabolite production. *BMC Bioinformatics*, **13**(1), 93.
- Chemler, J. A., Fowler, Z. L., McHugh, K. P., & Koffas, M. A. G. (2010). Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering. *Metab. Eng.*, **12**(2), 96–104.

- Cho, A., Yun, H., Park, J. H., Lee, S. Y., & Park, S. (2010). Prediction of novel synthetic pathways for the production of desired chemicals. *BMC Syst. Biol.*, **4**(1), 35.
- Choi, H. S., Lee, S. Y., Kim, T. Y., & Woo, H. M. (2010). *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.*, **76**(10), 3097–3105.
- Ciriminna, R., Pina, C. Della, Rossi, M., & Pagliaro, M. (2014). Understanding the glycerol market. *Eur. J. Lipid Sci. Technol.*, **116**(10), 1432–1439.
- Cohen, S. N., Chang, A. C., Boyer, H. W., & Helling, R. B. (1973). Construction of biologically functional bacterial plasmids *in vitro*. *Proc. Natl. Acad. Sci. U. S. A.*, **70**(11), 3240–3244.
- Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J., & Palsson, B. Ø. (2004). Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, **429**(6987), 92–96.
- Covert, M. W., Xiao, N., Chen, T. J., & Karr, J. R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics*, **24**(18), 2044–2050.
- Demain, A. L. (2000). Microbial biotechnology. *Trends Biotechnol.*, **18**(1), 26–31.
- Department of Energy, U.S.A. (2004) Top Value Added Chemicals from Biomass
- Duarte, N. C., Herrgard, M. J., & Palsson, B. Ø. (2004). Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.*, **14**(7), 1298–1309.
- EcoCyc: Encyclopedia of *E. coli* Genes and Metabolic Pathways. <http://ecocyc.org/>
- Edwards, J. S. (1999). Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.*, **274**(25), 17410–17416.
- Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., & Palsson, B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.*, **3**(121), 121.
- Feist, A. M., & Palsson, B. Ø. (2008). The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nat. Biotechnol.*, **26**(6), 659–667.
- Fisher, A. K., Freedman, B. G., Bevan, D. R., & Senger, R. S. (2014). A review of metabolic and

- enzymatic engineering strategies for designing and optimizing performance of microbial cell factories. *Comput. Struct. Biotechnol. J.*, **11**(18), 91–99.
- Fong, S. S., Burgard, A. P., Herring, C. D., Knight, E. M., Blattner, F. R., Maranas, C. D., & Palsson, B. Ø. (2005). *In silico* design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.*, **91**(5), 643–648.
- Fong, S. S., Marciniak, J. Y., & Palsson, B. O. (2003). Description and interpretation of adaptive evolution of *Escherichia coli* K-12 MG1655 by using a genome-scale *in silico* metabolic model. *J. Bacteriol.*, **185**(21), 6400–6408.
- Foster, J. M., Davis, P. J., Raverdy, S., Sibley, M. H., Raleigh, E. a, Kumar, S., & Carlow, C. K. S. (2010). Evolution of bacterial phosphoglycerate mutases: non-homologous isofunctional enzymes undergoing gene losses, gains and lateral transfers. *PLoS One*, **5**(10), e13576.
- Henry, C. S., DeJongh, M., Best, A. a, Frybarger, P. M., Linsay, B., & Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.*, **28**(9), 977–982.
- Holms, H. (1996). Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol. Rev.*, **19**, 85–116.
- Ikeda, M. (2006). Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering. *Appl. Microbiol. Biotechnol.*, **69**(6), 615–626.
- Jain, V. K., Divol, B., Prior, B. A., & Bauer, F. F. (2012). Effect of alternative NAD⁺-regenerating pathways on the formation of primary and secondary aroma compounds in a *Saccharomyces cerevisiae* glycerol-defective mutant. *Appl. Microbiol. Biotechnol.*, **93**(1), 131–141.
- Jantama, K., Haupt, M. J., Svoronos, S. A., Zhang, X., Moore, J. C., Shanmugam, K. T., & Ingram, L. O. (2008). Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. *Biotechnol. Bioeng.*, **99**(5), 1140–1153.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. *Nature*, **407**(6804), 651–654.
- KEGG: Kyoto Encyclopedia of Genes and Genomes. <http://www.genome.jp/kegg/>
- Kim, J., & Reed, J. L. (2010). OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst. Biol.*, **4**(1), 53.

- Kim, J., & Reed, J. L. (2012). RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. *Genome Biol.*, **13**(9), R78.
- Kim, J., Reed, J. L., & Maravelias, C. T. (2011). Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. *PloS One*, **6**(9), e24162.
- Korman, T. P., Hill, J. A., Vu, T. N., & Tsai, S. (2004). Structural analysis of actinorhodin polyketide ketoreductase: cofactor binding and substrate specificity. *Biochemistry*, **43**(46), 14529–14538.
- Kourkoutas, Y., Bekatorou, A., Banat, I. M., Marchant, R., & Koutinas, A. A. (2004). Immobilization technologies and support materials suitable in alcohol beverages production: a review. *Food Microbiol.*, **21**(4), 377–397.
- Krämer, M., Bongaerts, J., Bovenberg, R., Kremer, S., Müller, U., Orf, S., Wubbolts, M., & Raeven, L. (2003). Metabolic engineering for microbial production of shikimic acid. *Metab. Eng.*, **5**(4), 277–283.
- Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., & Lee, S. Y. (2007). Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.*, **3**(149), 149.
- Lee, S. J., Lee, D.-Y., Kim, T. Y., Kim, B. H., Lee, J., & Lee, S. Y. (2005). Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and *in silico* gene knockout simulation. *Appl. Environ. Microbiol.*, **71**(12), 7880–7887.
- Leonard, E., Ajikumar, P. K., Thayer, K., Xiao, W.-H., Mo, J. D., Tidor, B., Stephanopoulos, G., & Prather, K. L. J. (2010). Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U. S. A.*, **107**(31), 13654–13659.
- Levanodowsky, M., & Winter, D. (1971). Distance between sets. *Nature*, **234**(5), 34–35.
- Lewis, N. E., Hixson, K. K., Conrad, T. M., Lerman, J. A., Charusanti, P., Polpitiya, A. D., Adkins, J. N., Schramm, G., Purvine, S. O., Lopez-Ferrer, D., Weitz, K. W., Eils, R., Smith, R. D., & Palsson, B. Ø. (2010). Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.*, **6**, 390.
- Lewis, N. E., Nagarajan, H., & Palsson, B. Ø. (2012). Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods. *Nat. Rev. Microbiol.*, **10**(4), 291–305.
- Lun, D. S., Rockwell, G., Guido, N. J., Baym, M., Kelner, J. A., Berger, B., Galagan, J. E., & Church, G. M. (2009). Large-scale identification of genetic design strategies using local search. *Mol.*

Syst. Biol., **5**, 296.

- Mahadevan, R., & Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.*, **5**(4), 264–276.
- Nakamura, C. E. & Whited, G. M. (2003). Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.*, **14**, 454–459.
- Manring, M. M., Hawk, A., Calhoun, J. H., & Andersen, R. C. (2009). Treatment of war wounds: a historical review. *Clin. Orthop. Relat. Res.*, **467**(8), 2168–2191.
- McCloskey, D., Palsson, B. Ø., & Feist, A. M. (2013). Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol. Syst. Biol.*, **9**(661), 661.
- MetaCyc Encyclopedia of Metabolic Pathways. <http://metacyc.org/>
- Monk, J., Nogales, J., & Palsson, B. O. (2014). Optimizing genome-scale network reconstructions. *Nat. Biotechnol.*, **32**(5), 447–52.
- Motoki, M., & Seguro, K. (1998). Transglutaminase and its use for food processing. *Trends Food Sci. & Technol.*, **9**(5), 204–210.
- Ng, C. Y., Jung, M.-Y., Lee, J., & Oh, M.-K. (2012). Production of 2,3-butanediol in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering. *Microb. Cell Fact.*, **11**(1), 68.
- Oberhardt, M. A., Palsson, B. Ø., & Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.*, **5**, 320.
- Ogawa, Y., Nitta, A., Uchiyama, H., Imamura, T., Shimoi, H., & Ito, K. (2000). Tolerance mechanism of the ethanol-tolerant mutant of sake yeast. *J. Biosci. Bioeng.*, **90**(3), 313–320.
- Ohnishi, J., Katahira, R., Mitsuhashi, S., Kakita, S., & Ikeda, M. (2005). A novel *gnd* mutation leading to increased L-lysine production in *Corynebacterium glutamicum*. *FEMS Microb. Lett.*, **242**(2), 265–274.
- Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Mol. Syst. Biol.*, **7**, 535.
- Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nat. Biotechnol.*, **28**(3), 245–248.
- Park, J. H., Lee, K. H., Kim, T. Y., & Lee, S. Y. (2007). Metabolic engineering of *Escherichia coli*

- for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc. Natl. Acad. Sci. U. S. A.*, **104**(19), 7797–7802.
- Patil, K. R., Rocha, I., Förster, J., & Nielsen, J. (2005). Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinformatics*, **6**(1), 308.
- Pharkya, P., Burgard, A. P., & Maranas, C. D. (2004). OptStrain: a computational framework for redesign of microbial production systems. *Genome Res.*, **14**(11), 2367–2376.
- Pharkya, P., & Maranas, C. D. (2006). An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab. Eng.*, **8**(1), 1–13.
- Ramseier, T. M. (1996). Cra and the control of carbon flux via metabolic pathways. *Res. Microbiol.*, **147**(6-7), 489–493.
- Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., & Maranas, C. D. (2012). An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metab. Eng.*, **14**(6), 687–704.
- Reed, J. L., & Palsson, B. Ø. (2004). Genome-scale *in silico* models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res.*, **14**(9), 1797–1805.
- Reed, J. L., Vo, T. D., Schilling, C. H., & Palsson, B. Ø. (2003). An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.*, **4**(9), R54.
- Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., & Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**(7086), 940–943.
- Rocha, M., Maia, P., Mendes, R., Pinto, J. P., Ferreira, E. C., Nielsen, J., Patil, K. R., & Rocha, I. (2008). Natural computation meta-heuristics for the *in silico* optimization of microbial strains. *BMC Bioinformatics*, **9**(1), 499.
- Rowlands, R. T. (1984). Industrial strain improvement: mutagenesis and random screening procedures. *Enzyme Microb. Technol.*, **6**(1), 3–10.
- Sano, C. (2009). History of glutamate production. *Am. J. Clin. Nutr.*, **90**(3), 728S–732S.
- Santala, S., Efimova, E., Kivinen, V., Larjo, A., Aho, T., Karp, M., & Santala, V. (2011). Improved

- triacylglycerol production in *Acinetobacter baylyi* ADP1 by metabolic engineering. *Microb. Cell Fact.*, **10**(1), 36.
- Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., Kang, J., Hyduke, D. R., & Palsson, B. Ø. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.*, **6**(9), 1290–1307.
- Schuetz, R., Kuepfer, L., & Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.*, **3**(119), 119.
- Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., & Sauer, U. (2012). Multidimensional Optimality of Microbial Metabolism. *Science*, **336**(6081), 601–604.
- Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., & Liao, J. C. (2011). Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.*, **77**, 2905–2915.
- Shinfuku, Y., Sorpitiporn, N., Sono, M., Furusawa, C., Hirasawa, T., & Shimizu, H. (2009). Development and experimental verification of a genome-scale metabolic model for *Corynebacterium glutamicum*. *Microb. Cell Fact.*, **8**(1), 43.
- Siegel, R., & Ryu, D. D. (1985). Kinetic study of instability of recombinant plasmid pPLc23trpAl in *E. coli* using two-stage continuous culture system. *Biotechnol. Bioeng.*, **27**(1), 28–33.
- Stephanopoulos, G. N., Aristidou, A. A., & Nielsen, J. (1998). *Metabolic Engineering: Principles and Methodologies*. Academic Press, Orlando, FL.
- Summers, D. K. (1991). The kinetics of plasmid loss. *Trends Biotechnol.*, **9**(8), 273–278.
- Tavares, J., Pereira, F. B., & Costa, E. (2008). Multidimensional knapsack problem: a fitness landscape analysis. *IEEE Trans. Syst. Man. Cybern. B*, **38**(3), 604–616.
- Thiele, I., Fleming, R. M. T., Bordbar, A., Schellenberger, J., & Palsson, B. Ø. (2010). Functional characterization of alternate optimal solutions of *Escherichia coli*'s transcriptional and translational machinery. *Biophys. J.*, **98**(10), 2072–2081.
- Tomar, N., & De, R. K. (2013). Comparing methods for metabolic network analysis and an application to metabolic engineering. *Gene*, **521**(1), 1–14.
- Torres, L. L., Cantero, A., del Valle, M., Marina, A., López-Gallego, F., Guisán, J. M., Berenguer, J., & Hidalgo, A. (2013). Engineering the substrate specificity of a thermophilic penicillin acylase

from *Thermus thermophilus*. *Appl. Environ. Microb.*, **79**(5), 1555–1562.

Typas, A., Nichols, R. J., Siegele, D. A., Shales, M., Collins, S. R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B. L., Mori, H., Weissman, J. S., Krogan, N. J., & Gross, C. A. (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nat. Methods*, **5**(9), 781–787.

U.S.A.: World's first propanediol production from corn sugar opened, Aug. 2007, Bio-based-News, <http://bio-based.eu/news/u-s-a-worlds-first-propanediol-production-from-corn-sugar-opened/>

Varma, A., & Palsson, B. Ø. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.*, **60**(10), 3724–3731.

Vazquez, A., Beg, Q. K., DeMenezes, M. A., Ernst, J., Bar-Joseph, Z., Barabási, A.-L., Boros, L. G., & Oltvai, Z. N. (2008). Impact of the solvent capacity constraint on *E. coli* metabolism. *BMC Syst. Biol.*, **2**(1), 7.

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., & Church, G. M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, **460**(7257), 894–898.

Xu, C., Liu, L., Zhang, Z., Jin, D., Qiu, J., & Chen, M. (2013). Genome-scale metabolic model in guiding metabolic engineering of microbial improvement. *Appl. Microb. Biotechnol.*, **97**(2), 519–539.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A. P., Boldt, J., Khandurina, J., Trawick, J. D., Osterhout, R. E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H. B., Andrae, S., Yang, T. H., Lee, S. Y., Burk, M. J., & Van Dien, S. (2011). Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.*, **7**(7), 445–452.

Zaldivar, J., Nielsen, J., & Olsson, L. (2001). Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Appl. Microbiol. Biotechnol.*, **56**(1-2), 17–34.

Zamboni, N., Fendt, S.-M., Rühl, M., & Sauer, U. (2009). ¹³C-based metabolic flux analysis. *Nat. Protoc.*, **4**(6), 878–892.

Zhang, X., Wang, X., Shanmugam, K. T., & Ingram, L. O. (2011). L-malate production by metabolically engineered *Escherichia coli*. *Appl. Environ. Microbiol.*, **77**(2), 427–434.

Zhuang, K., Vemuri, G. N., & Mahadevan, R. (2011). Economics of membrane occupancy and

respiro-fermentation. *Mol. Syst. Biol.*, **7**, 500.

Zomorodi, A. R., Suthers, P. F., Ranganathan, S., & Maranas, C. D. (2012). Mathematical optimization applications in metabolic networks. *Metab. Eng.*, **14**(6), 672–686.

株式会社資源・食糧問題研究所、2012年、ブラジルのバイオエタノールをめぐる動向
http://www.alic.go.jp/joho-s/joho07_000557.html

バイオ燃料技術革新評議会、2008年、バイオ燃料技術革新計画

坂和 正敏、1999年、数理計画法の基礎、森北出版株式会社

付録

Table A1 本論文における遺伝子名と反応名の関係

遺伝子名	反応名
<i>aceA</i>	Isocitratelase
<i>aceF</i>	Pyruvate dehydrogenase
<i>adhE</i>	Alcohol dehydrogenase
<i>adhE</i>	NADH dehydrogenase
<i>aldA</i>	Lactaldehyde dehydrogenase
<i>allA</i>	Ureidoglycolate hydrolase
<i>allB</i>	Allantoinase
<i>allD</i>	Ureidoglycolate dehydrogenase
<i>allP</i>	Allantoin transport (periplasm)
<i>bcd</i>	Nitrite reductase (ubiquinone) (periplasm)
<i>cbdA</i>	Cytochrome oxidase <i>bd</i> (menaquinone)
<i>cueO</i>	Ferroxidase oxidase
<i>cydA</i>	Cytochrome oxidase <i>bd</i> (ubiquinone)
<i>cyoA</i>	Cytochrome oxidase <i>bo</i>
<i>dld</i>	D-Lactate dehydrogenase (ubiquinone)
<i>edd</i>	6-Phosphogluconate dehydratase
<i>focA</i>	Formate transport via diffusion
<i>folD</i>	Methenyltetrahydrofolatecyclohydrolase
<i>fsaA</i>	Fructose 6-phosphate aldolase
<i>galP</i>	Glucose transport in via proton symport
<i>gcd</i>	Glucose dehydrogenase
<i>gdhA</i>	Glutamate dehydrogenase
<i>gldA</i>	Lactaldehyde oxidoreductase
<i>glpD</i>	Glycerol 3-phosphate dehydrogenase (ubiquinone)
<i>glpK</i>	Glycerol kinase
<i>glyA</i>	Glycine hydroxymethyltransferase
<i>gntT</i>	Gluconate transporter
<i>gpsA</i>	Glycerol 3-phosphate dehydrogenase (NADPH)
<i>hemE</i>	Uroporphyrinogen decarboxylase (uroporphyrinogen III)
<i>hemF</i>	Coproporphyrinogen oxidase (O ₂ required)
<i>hemG</i>	Protoporphyrinogen oxidase

(次ページに続く)

遺伝子名	反応名
<i>hemH</i>	Ferrochelatase
<i>hemN</i>	Oxygen Independent coproporphyrinogen-III oxidase
<i>idnK</i>	Gluconokinase
<i>idnO</i>	5-Dehydro-gluconate reductase
<i>idnT</i>	Idonate transport
<i>lamB</i>	Glucose transport (extracellular to periplasm)
<i>ldhA</i>	D-Lactate dehydrogenase
<i>mdh</i>	Malate dehydrogenase
<i>mgIA</i>	Glucose transport via ABC system
<i>mgsA</i>	Methylglyoxal synthase
<i>msrA</i>	Methionine-S -sulfoxide reductase
<i>msrB</i>	Methionine-R-sulfoxide reductase
<i>mtn</i>	Methylthioadenosine nucleosidase
<i>napC</i>	Nitrate reductase (menaquinone)
<i>napG</i>	Nitrate reductase (ubiquinone)
<i>narG</i>	Nitrate reductase (ubiquinone) (periplasm)
<i>narU</i>	Nitrate transport in via nitrite antiport
<i>nirB</i>	Nitrite reductase (NADH)
<i>nirC</i>	Nitrite transport in via proton symport
<i>nrfA</i>	Nitrite reductase (ubiquinone)
<i>pdxH</i>	Pyridoxamine 5-phosphate oxidase
<i>pflB</i>	Pyruvate formatelyase
<i>pgi</i>	Glucose6-phosphate isomerase
<i>poxB</i>	Pyruvate oxidase
<i>ppc</i>	Phosphoenolpyruvate carboxylase
<i>ppk</i>	Polyphosphate kinase
<i>ppsA</i>	Phosphoenolpyruvate synthase
<i>pta</i>	Phosphotransacetylase
<i>ptsG</i>	Glucose transport via phosphotransferase system
<i>puuB</i>	Gamma glutamyl putrescine oxidase
<i>pykF</i>	Pyruvate kinase
<i>pyrB</i>	Aspartate carbamoyltransferase
<i>pyrC</i>	Dihydroorotase
<i>pyrD</i>	Dihydroorotic acid dehydrogenase (ubiquinone)

(次ページに続く)

遺伝子名	反応名
<i>pyrE</i>	Orotate phosphoribosyltransferase
<i>rpe</i>	Ribulose 5-phosphate 3-epimerase
<i>sdhC</i>	Succinate dehydrogenase
<i>serC</i>	Phosphoserine transaminase
<i>talA</i>	Transaldolase
<i>tpiA</i>	Triose-phosphate isomerase
<i>ubiB</i>	2-Octaprenylphenol hydroxylase
<i>xylA</i>	Xylose isomerase
<i>xylE</i>	Xylose transport in via proton symport
<i>xylF</i>	Xylose transport via ABC system
<i>ybcF</i>	Carbamate kinase
<i>zwf</i>	Glucose 6-phosphate dehydrogenase

Table A2 代謝物質の略称

代謝物質の略称	代謝物質名
13PD	1,3-propanediol
6PG	6-phospho- gluconate
AC	acetate
ACOA	acetyl-CoA
AKG	2-oxoglutarate
CIT	citrate
DHA	dihydroxyacetone
DHAP	dihydroxyacetone phosphate
E4P	erythrose 4-phosphate
ETOH	ethanol
F6P	fructose 6-phosphate
FBP	fructose 1,6-bisphosphate
FOR	formate
FUM	fumarate
G3P	glyceraldehyde 3-phosphate
G6P	glucose 6-phosphate
GLC	glucose
Glucose[e]	glucose (細胞外)
GLX	glyoxylate
GLY3P	glycerol 3-phosphate
GLYC	glycerol
Glycerol[e]	glycerol (細胞外)
GPP	geranyl diphosphate
ICIT	isocitrate
LAC	D-lactate
MAL	L-malate
MGX	methylglyoxal
OAA	oxaloacetate
PEP	phosphoenolpyruvate
L-Phe	L-phenylalanine
PYR	pyruvate
R5P	ribose 5-phosphate
RU5P	ribulose 5-phosphate

(次ページに続く)

代謝物質の略称	代謝物質名
S7P	sedoheptulose 7-phosphate
SUC	succinate
X5P	xylulose 5-phosphate

Table A3 GDLS のパラメータ

探索のパスの数	2
近傍の数	2
最大の遺伝子破壊数	25

パラメータは、先行研究 (Lun *et al.*, 2009) に従った。

Table A4 OptGene のパラメータ

集団のサイズ	125
最大の世代数	5000
最大の遺伝子破壊数	25
変異率	1 / (破壊対象の遺伝子数)

パラメータは、先行研究 (Patil *et al.*, 2005) に従った。

謝辞

本研究を遂行するにあたり、大阪大学工学部応用自然科学科の4年目と、大阪大学大学院情報科学研究科バイオ情報工学専攻の博士前期課程と博士後期課程の合計7年間、研究の場を与えてくださり、終始暖かく辛抱強く、御指導と御助言を頂きました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座の清水浩教授に、心より御礼申し上げます。論文審査委員として、有益なご助言、御指導を頂きました、大阪大学大学院情報工学研究科バイオ情報工学専攻の松田秀雄教授、若宮直紀教授、四方哲也教授、前田太郎教授、および古澤力招へい教授に熱く御礼申し上げます。アドバイザー委員として、研究に関して深く議論し、有益なご助言をくださった、大阪大学薬学研究科附属薬用植物園の原田和生講師に、厚く御礼申し上げます。研究の遂行にあたり、方針や結果に関して議論を行い御指導を頂きました、代謝情報工学講座の古澤力招へい教授に心より感謝申し上げます。研究に関して、非常に客観的かつ論理的な御助言を頂きました代謝情報工学講座の松田史生准教授に厚く御礼申し上げます。生物学的な視点から御助言を頂きました、代謝情報工学講座の平沢敬招へい准教授に、心より御礼申し上げます。数理モデルやその解析に関して、御助言を頂きました奈良先端科学技術大学院大学の小野直亮助教(元代謝情報工学講座特任准教授)に心より感謝申し上げます。研究室生活や研究方法について御指導頂きました代謝情報工学講座の吉川勝徳助教に厚く感謝申し上げます。代謝フラックス解析や数理モデルを用いたシミュレーションに関して、親身にご助言を頂きました代謝情報工学講座の戸谷吉博特任助教に、厚く御礼申し上げます。探索プログラムのテストをして頂きました、代謝情報工学講座の吉川勝徳助教、戸谷吉博特任助教、徳山健斗氏および Mitr Phol Sugar 社の Sunisa Chaturachai 氏(代謝情報工学講座卒業生)に厚く御礼申し上げます。同級生として共に研究活動に励み、様々な面でお世話になりました代謝情報工学講座の梶畠秀一氏および仲嶋翼氏に心より感謝申し上げます。研究室生活の面で、数々の面でお世話になりました代謝情報工学講座の皆様に、厚く御礼申し上げます。

最後に、長年にわたる大学生生活を、精神的、経済的に支えてくださった家族に、心より感謝申し上げます。