

分子ネットワークのコンピュータ支援設計法の開発とグルコース同化システムへの応用

●倉田 博之

九州工業大学大学院情報工学研究院

<研究の目的と進め方>

システム生物学の最終目標は、環境ストレス応答や遺伝的变化の中でどのようにして生化学ネットワークが特定の細胞機能を生み出すのかを理解し、分子プロセスの合理的設計をおこなうことである。多様な環境ストレス、遺伝子変化、確率的なゆらぎに対して恒常性を維持するための細胞固有の性質であるロバストネスは、様々な細胞の機能を生み出す分子ネットワーク構造を解析し、設計する上で大切なキーワードである。

分子相互作用レベルで生化学システムを合理的に設計するためには、生化学ネットワークマップを同定すること、システムの動的モデルを構築し、システム解析を行うことが、必要不可欠である。それらによって分子ネットワーク構造と細胞機能を関連付けることができる。現在、Computer-Aided Design (CAD) は、数学モデルを作り、遺伝子組換え細胞の分子ネットワークの数値シミュレーションを行う上で重要なコンセプトとなっている。摂動解析はシステムの性能に影響を及ぼす決定的パラメータの同定に役に立つ。

制御工学との類似性に基づいて、機能モジュールとフラックスモジュールからなる階層的モジュール構造へ生化学ネットワークマップを分解するような、モジュールを基にした解析が提案された。生化学ネットワークの目的が与えられたならば、機能モジュールはすぐに特定され、それに工学的機能が割り当てることができる。それによって、望ましい成果を達成するために必要な遺伝子組換え操作の合理的設計が進展し、生化学ネットワークをどのように改良し、修正するのかを直感的に理解できるようになる。

微生物による有用物質の工業的生産では、グルコースは最も重要な基質のひとつであり、物質生産や細胞の増殖に必要な不可欠である。本研究では、摂動解析、すなわち遺伝子のコピー数の変化の影響を解析するために、タンパク質の転写調節を含む、大腸菌のグルコースリン酸転移システム (phosphotransferasesystem (PTS)) の大規模な数学モデルを開発する。最終目標は、PTS のグルコース取り込み効率を遺伝的レベルから改良することによって、発酵生産性を向上させることである。

CADLIVE システムを用いて *ptsG*, *ptsH*, *ptsI*, *crp*, *cyaA*, *crp*, *mlc* の遺伝子から成るグルコース PTS の分子ネットワーク構造を設計する。具体的には、グルコース取り込み速度を向上させるため、どのように PTS を遺伝子組換え操作をするのか、もしくは、どの遺伝子を削除し、過剰発現すべきなのか、を決定する合理的計画を作るために、制御工学との類似性に基づいて、PTS を機能モジュールとフラックスモジュールからなる階層的モジュールに分解する。そしてグルコース取り込み速度に与える遺伝子発現量の変化の影響をシミュレーションして、どの遺伝子が取り込み速度を増加させるのに有用なのかを予測する。設計した細胞の挙動の予測を生物実験を行って検証する。

<研究の目的と進め方>

モデリング、シミュレーション、システム解析に基づいて、

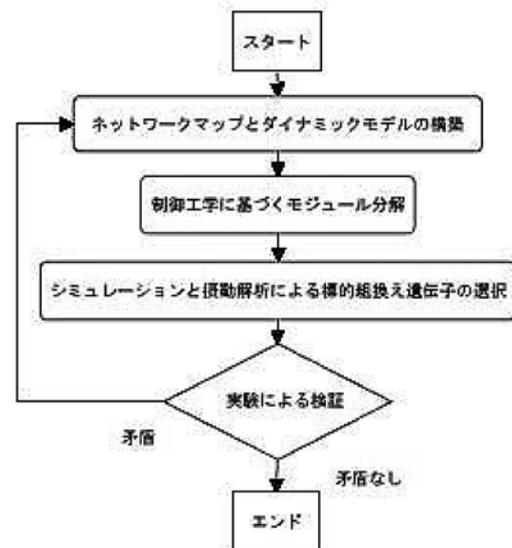


図1 細胞の合理的設計戦略

Glucose PTS を分子レベルから設計するための合理的戦略を提案し、その予測結果を実験によって証明する。具体的な戦略は図1 (次ページ) のようになる。設計戦略は、(1) 分子ネットワーク同定とダイナミックモデルの構築、(2) 工学のアナロジーによるネットワーク分解とシステム解析、(3) 摂動解析による標的遺伝子群の探索、(4) 生物実験による設計戦略の証明、である。

モジュール分解

人工物回路と分子ネットワークの構造が類似であることは、過去の研究によって示されているので (Kurata H et al., PLoSComput Biol 2: e59, 2006), 制御工学のアナロジーに基づいて、システムを階層的モジュールに分解し、制御メカニズムと細胞機能の関係を容易に理解することができる。分子ネットワークを機能モジュールと機能モジュール間を結ぶ FLUX モジュールに分解する。分子や信号の流れを示す FLUX モジュールは、細胞機能を発現させる本質的制御構造である。

図2に大腸菌の glucose PTS の分子ネットワークとモジュール分解図を示す。リン酸転移カスケード (PTS タンパク質, IICB, IIA, Hpr, EI) を通してグルコース取り込み行う「プラント」モジュールを定義する。「プラント」を駆動する PTS タンパク質は「アクセルアクチュエータ」(CRP:cAMP) によって発現が促進される。プラント中のグルコース同化の状態 (IIA-P) は、「センサー」によって検知されて、その信号は「コンピュータ」(CRP, CYA) に送られる。「コンピュータ」は「アクセルアクチュエータ」や「ブレーキアクチュエータ」(Mlc) の操作変数を適切に決定する。2つの「アクチュエータ」の働きによって、「プラント」におけるグルコース同化速度が制御される。これらの機能モジュール間を結合する FLUX モジュールは、PTS 遺伝子を発現させる

ことによってグルコース取り込み速度を制御する。ここでは、大きくわけて、「アクセル FLUX モジュール」と「ブレーキ FLUX モジュール」があり、両者の適切な組み合わせによって、グルコース同化速度が制御される。

システム解析

CADLIVE を用いて Glucose PTS のダイナミックモデルを作成し、数値シミュレーションを行う。CADLIVE は分子ネットワークマップ構築、ダイナミックモデリング、システム解析を支援するシステムである。ダイナミックモデルを検証するために、培地中のグルコース濃度変化に対する PTS タンパク質の濃度変化を測定し、シミュレーション結果と比較する。

次に、ダイナミックモデルのロバストネスを検証する。一般にフィードバック制御系をもつ生命システムは遺伝子発現量の変化に対して、大きな変化を示さない、すなわちロバストネスを示すことが知られている。そこで、遺伝子発現量の変化に対するグルコース同化速度の感度を計算する。感度値が極端に高い遺伝子発現変化があれば、注意が必要であろう。

モデルの構築に用いなかった実験データを用いて、ダイナミックモデルの正当性を検証する。具体的には、培地中のグルコース濃度を低下させて細胞内の cAMP や IIA-P 濃度を調べる。グルコース濃度低下によって cAMP 濃度が上昇するので、PTS タンパク質の発現が変化することが期待される。また、Mlc ノックアウト株における PTS タンパク質制御について、シミュレーションと生物実験を比較する。

摂動解析

グルコース同化速度を向上させる標的遺伝子の組み合わせを特定するために、遺伝子発現の摂動解析をコンピュータ上で行う。個々の遺伝子、あるいは複数の遺伝子についてノックアウトや過剰発現を行ったダイナミックモデルを用いて、グルコース同化速度の変化をシミュレーションする。モジュール分解図に基づいて考えると、PTS タンパク質の過剰発現、ブレーキアクチュエータの抑制 (Mlc ノックアウト)、アクセルアクチュエータ (CRP:cAMP) の促進によって、グルコース同化速度が向上すると考えられる。それらの遺伝子組換え操作がダイナミックモデルに与える影響をシミュレーションする。

細胞設計結果の実験による検証

摂動解析を通して、グルコース同化速度の向上につながる遺伝子群を選択し、遺伝子組換え戦略を構築する。遺伝子組換え戦略の正当性は、遺伝子ノックアウトや過剰発現株の作成実験を行うことによって検証する。生物実験とシミュレーションの結果を比較するために、培養環境を正確に制御できる連続培養装置を用いた実験を行う。

予測通りに細胞が設計できるかどうかを検証する。ダイナミックモデルは、グルコース同化システムに焦点をあてているが、その周辺の未知の信号伝達経路や増殖速度の変化などは考慮していない。もしそれらの因子がシステムに影響を与えるならば、モデルと実験結果は必ずしも一致しないであろう。そのような場合、モデルと実験が不一致となる原因を探索することによって、新しい分子メカニズムの発見が期待できる。その発見に基づいて細胞を再設計することによって、より現実的な細胞が作れるであろう。ネットワーク再設計や生物実験による検証を繰り返すことによって、生命システムの理解が進み、精密な設計が実現する。

他に、PTS リン酸リレイの律速段階の同定、グルコース同化システムと他の糖同化システムとの連携についての解明を行う。

多様な糖の取り込みに優先順位をつけるメカニズムの解明が期待できる。

E. coli MG1655 を L-broth, M9 培地 (20mM グルコース) を用いて好気的条件下で培養する。遺伝子ノックアウトは DatsenkoWanner の方法を用い、遺伝子過剰発現は、高コピープラスミド (p UC118) に目的の遺伝子を導入する。発現プロモータとして *Morganella morganii* の *phoC* (フォスファターゼ) のプロモータを用いる。グルコース同化速度、増殖、細胞外の cAMP 濃度を測定する。

<研究期間の成果>

グルコースPTS のモジュール構造と数学モデリング

CADLIVE を用いて、大腸菌グルコース PTS の生化学ネットワークマップ (図 2) を記述して、動的モデルを開発した。図の表記法の詳細は他で述べる。一般的に、物質生産増強のために、糖質の取り込み速度を増加させることは、非常に重要である。ここでは、グルコース取り込み速度を増加させることを目的とする。取り込み速度増強の観点から、各々の遺伝子の機能を特定するために、制御工学ブロックダイヤグラムとの類似性に基づいて、グルコース PTS の生化学マップを機能モジュール (図 2A) とフラックスモジュール (図 2B) に分解する。制御対象であるプラントとして、ptsG, ptsH, ptsI, crr のリン酸転移酵素カスケードを定義すると、プラントを操作する変数は、活性化アクチュエータからのシグナルになる。IIA-P からのリン酸化シグナルは、フィードバックセンサーモジュールが感知して、アデニル酸シクラーゼ、cAMP、cAMP 受容体タンパク質 (CRP) から成るコンピュータモジュールに送られる。一方、他の糖輸送系を増強するシグナル伝達経路にも送られる。フィードバックセンサーモジュールの中で生成する IIA-P とアデニル酸シクラーゼの複合体はコンピュータモジュール中で、cAMP シグナルを生成する。その結果、cAMP 受容体タンパク質と cAMP 複合体が生成する。cAMP 受容体タンパク質と cAMP の複合体は、コンピュータモジュールからの出力であり、活性化と抑制のアクチュエータモジュールを調節する。活性化アクチュエータモジュールは、HICB, HPr, EI タンパク質の合成プロセスから構成され、これらのリン酸転移酵素タンパク質をプラントモジュールに供給する。一方、Mlc タンパク質からなる抑制アクチュエータモジュールはリン酸転移酵素タンパク質の合成を抑制する。

これらの機能モジュールを重ねあわせることによって、活性化

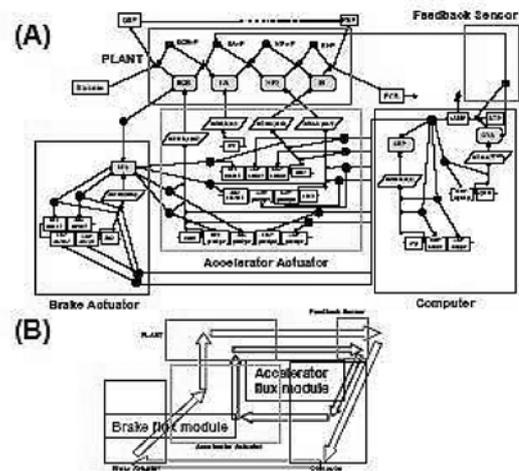


図2 グルコース PTS の生化学マップ
A:機能モジュール, B フラックスモジュール

フラックスモジュールと抑制フラックスモジュールという二つの主要なフラックスモジュールが同定される。これらのフラックスモジュールに関する記述は定性的なものであるが、フラックスモジュールの成分は簡単に明らかにできる (図2A)。まず、IIA-P のシグナルが、フィードバックセンサー、コンピュータ、活性化アクチュエータ、プラントモジュールを通して IICB の合成を引き起こす活性化フラックスモジュールを明らかにした。次に、IIA-P のシグナルが、フィードバックセンサー、コンピュータ、抑制アクチュエータ、プラントモジュールを通して、活性化アクチュエータを抑制することによって、リン酸転移酵素の合成を調節する、これは、抑制フラックスモジュールである。

動的シミュレーション

数学モデルが、グルコース濃度変化に対するグルコース PTS の動的挙動を再現することを証明するため、数学モデルをシミュレーションした。シミュレーションによる動的挙動は、実験のデータと一致した。グルコースの枯渇後、IIA 濃度が急速に減少すると同時に、IIA-P の濃度は増加した。詳細は文献を参照。

グルコース取り込み速度を向上させるために重大な役割を担う遺伝子の抽出

PTS モデルの機能モジュールとフラックスモジュールへの分解は、グルコース取り込み速度を向上させるための合理的戦略を容易に供給する。転写調節因子である *mlc* 遺伝子と *crp* 遺伝子は、複数の遺伝子の発現に影響を及ぼす。*mlc* 遺伝子は抑制アクチュエータを構成するので、*mlc* 遺伝子欠損は PTS のタンパク質の合成を促進することが期待される。*crp* 遺伝子は活性化アクチュエータを制御するコンピュータモジュールに属しているため、*crp* 遺伝子の過剰発現は PTS のタンパク質の合成を強める可能性がある。*ptsG*, *ptsH*, *ptsI*, *err* 遺伝子はプラントモジュールのリン酸化カスケードの触媒となるので、グルコース取り込み速度を上げるため、それらの遺伝子の発現を促進することは合理的である。グルコース取り込み速度を上げるためには、プラントモジュールの PTS タンパク質を増やすこと (*ptsG*, *ptsH*, *ptsI*, *err* 遺伝子)、抑制アクチュエータモジュール (*mlc* 遺伝子) を削除すること、コンピュータモジュール (*crp* 遺伝子) を通して活性化アクチュエータモジュールを強めることが、重要な意味を持つ。*crp* 遺伝子の操作に関しては、CRP:cAMP 複合体が活性化アクチュエータモジュールだけでなく、抑制アクチュエータモジュールも強めるため、活性化フラックスモジュールと抑制フラックスモジュールのバランスを考慮することが必要である。

設計したPTSのシミュレーションと摂動解析

グルコース取り込み速度に関係のある目標遺伝子をさらに正確に調査するため、*crp*, *mlc*, *ptsG*, *ptsH*, *ptsI*, *err* 遺伝子発現に関して摂動解析を行った。これらの遺伝子のコピー数を 10 倍増やして、比グルコース取り込み速度をシミュレーションした。目標遺伝子の濃度が初期状態 (野生型) から 10 倍に変化したときの、変異株と野生型の比グルコース取り込み速度の比をシミュレーションした。比グルコース取り込み速度に与える *mlc* 遺伝子破壊と *crp* 遺伝子の増幅の効果を算出した。*ptsI* 遺伝子の増幅は、野生型、*mlc* 遺伝子欠損株、*crp* 遺伝子過剰発現変異株のすべてにおいて、最も高いグルコース取り込み速度を示した。*ptsI* 遺伝子と *crp* 遺伝子を過剰発現する *mlc* 遺伝子欠損変異株は、比グルコース取り込み速度が大幅に向上することが予想された。*ptsI* の増幅が比グルコース取り込み速度に与える影響をシミュレーションし、*ptsI* の増幅がグルコース取り込みに常に効果的であること

を確認した。

ptsI 遺伝子の増幅が、比グルコース取り込み速度を増加させる仕組みを解析するために、*ptsI* 遺伝子発現が変化させながら、PTS の構成要素濃度をシミュレーションした。比グルコース取り込み速度が、細胞あたりのリン酸化 IICBGlc によるグルコースへのリン酸基転移反応、つまりグルコース 6 リン酸の産生速度であると定義するので、比グルコース取り込み速度は、IICBGlc タンパク質のリン酸化速度と密接に関係する。リン酸化タンパク質の濃度は、*ptsI* 遺伝子の濃度の増加と共に増加した。このシミュレーション結果は、PTS 流束の増加は EI タンパク質濃度の増加が原因であることを示している。シミュレーションでは、*ptsI* 遺伝子発現量が初期の状態から 10 倍になると、リン酸化 IIAGlc タンパク質濃度は増加する。IIAGlc タンパク質のリン酸化状態は細胞外の糖質に依存することが知られており、グルコースが使われる時、IIAGlc タンパク質は脱リン酸化される。

ptsG, *ptsH*, *err*, *ptsI* 遺伝子の中から、比グルコース取り込み速度を最も増大させる *ptsI* 遺伝子を詳細な解析のために選択する。*ptsI* 遺伝子は PEP から EI タンパク質へのリン酸基転移反応の触媒となる EI タンパク質をコードする。この反応は、PTS のグルコース取り込みの律速段階になることが示されており、EI タンパク質濃度の増加はグルコース PTS 流束の増加を導くことが予想される。*ptsH* や *ptsI* の転写減衰が起こり、*ptsI* の転写量は、*ptsH* の転写量の 10 分の 1 になることが示されている。おそらく、*ptsI* 遺伝子の転写量の減衰は、グルコース PTS 流束を調節する本質の一つである。それゆえに、プラントモジュール中のリン酸化カスケード中から、*ptsI* 遺伝子を選択することは合理的である。

合理的に設計した細胞の実験的検証

mlc 遺伝子欠損株での *crp* 遺伝子と *ptsI* 遺伝子の過剰発現は、比グルコース取り込み速度を最も増大させることを予想した。この戦略の妥当性を調は、生物学実験を行って証明した。詳細は文献を参照。

コンピュータを用いた合理的設計の力

工学的目的に対して、生化学ネットワークを合理的に設計する一般的な戦略を提示した。その戦略は、生化学ネットワーク、動的シミュレーション、モジュール解析、摂動解析、実験による証明から構成される (図1)。CAD は生化学ネットワークの構築、動的モデルの設計、システムの解析を支援する。モジュールに基づく解析は、生化学ネットワークマップを機能モジュールとフラックスモジュールから成る階層的モジュールという制御工学構造に似た様式に分解する。もし工学的目的が与えられたならば、機能モジュールは直ちに同定されて、工学的機能が各モジュールに割り当てられる。このことは、目的達成するために必要な遺伝子操作の合理的設計や、生化学ネットワークをどのように改良し、修正できるのかを直感的に理解することを助ける。さらに、摂動解析は、目的を達成するために決定的に重要な遺伝子を探索することに役立つ。

CADLIVE を使って、グルコース取り込み速度の向上を目指して、どのようにグルコース PTS を遺伝子レベルから設計する戦略を立てるために、数学モデリング、シミュレーション、システム解析を行った。比グルコース取り込み速度を高めるために重要な因子、例えば *ptsI* 遺伝子の増幅や *mlc* 遺伝子欠損を見つけた。数学シミュレーション、それに続く摂動解析は、*ptsI* 遺伝子増幅や *mlc* 遺伝子欠損の組み合わせが比グルコース取り込み速度を高めることを予測し、最終的に、実験データを用いてこの予測を立

証した。

完全な設計に向けて

シミュレーション結果は必ずしも実験結果を説明するとは限らない。今回、詳しく述べないが、数学モデルと実験モデルの間で、cAMP 濃度の変化には矛盾が生じた。ptsI 遺伝子増幅と mlc 遺伝子欠損の組み合わせは、比グルコース取り込み速度を高める適切な戦略になると予測されたが、それは増殖阻害を引き起こした。数学モデルから、比グルコース取り込みは、ptsI 遺伝子の過剰発現株は元株と比べ 3.87 : 1 の比を持つと予測されたが、実験モデルでは最高でも 1.2 : 1 までしか観測されなかった。同様に mlc 遺伝子欠損株では 5.7 : 1 の比を持つと予測されたが、実験モデルでは 1.7 : 1 の比であった。これらの相違はモデルのサイズが不十分であるように、あるいは、実際の細胞の遺伝子組み換えにロバストネスを与える遺伝子発現や酵素のフィードバック調節の記述に不足があるように見える。加えて、グルコース浸透に関わる、マンノース PTS、PTS 非輸送体、GalP などの他のタンパク質が、実験結果に影響を及ぼした可能性がある。

増殖は様々な分子の複雑な相互作用の出力なので、分子相互作用レベルで細胞増殖を正確に予測することはまだ難しい。解糖系、ペントースリン酸経路、TCA サイクル、嫌氣的経路などの様々な酵素阻害や遺伝子発現を含む大規模な複雑なモデルが必要となるだろう。拡張モデルは実験結果を明確に表現することが期待される。もし数学モデルと実験モデルの間に矛盾があれば、この情報をモデルの改良にフィードバックする必要がある。理想的には予測が実験結果と完全に一致するまで、システム解析と実験による検証を繰り返し行うべきである。このサイクルは生化学システムの合理的設計の鍵となる。

<国内外での成果の位置づけ>

モデリング、シミュレーション、システム解析技術の開発を得意分野とし、生命システムに特有の問題点について解明した。工学における設計は、すべての部品の数値的特性は高精度で与えられているので、それらを組み合わせたシステムは予測通りのパフォーマンスを発揮する。一方、生物は、個々の分子の動力学的パラメータはほとんどわからないという問題点がある。パラメータ値が未知の状態では、システム解析し、設計することが求められる。その問題の解決策として、潜在的にとりうるパラメータ値をすべて考慮し、パラメータ値の変化が細胞性能に与える影響を理解しながら、システムを設計しなければならない。すなわち、モデリングに高いレベルの最適化技術が必要とされる。本研究は、そのような最適化技術やそれに基づくシステム解析技術に十分な経験と知識を習得しているため、モデリングと実験を組み合わせた研究を行って、生命設計学を創成し発展させることができた。

究極の目的は、多様な生物情報から、生命の設計原理を探索し、それに基づいて分子ネットワーク(細胞)を合理的に設計する学問を創造することである。膨大なスクリーニングによる従来の方法論ではなく、遺伝子や分子の基本的な構成因子から設計原理に基づいて合理的に細胞を設計することは革新性がある。「ものづくり」という合成の観点から、生命科学、情報科学、システム工学を収斂することによって、理論、シミュレーション、実験に基づく生命設計学という革新的領域を拓いた。

独創的な点は、研究計画で詳述するが、分子ネットワークを工学のアナロジーに基づいてモジュール分解することによって、システムの目的に対する各遺伝子機能の役割を決める技術、摂動解析を用いてもっとも効果的なネットワークの標的部位を予測する技術の開発である。これらの一連の研究戦略は普遍的方法論とし

て提案した。

グルコース等の基質の取り込み速度の向上が、物質生産性や増殖に本質的影響を与えることは容易に予測できる。既往の研究では、glucose PTS タンパク質のシグナルトランスダクションのモデリングが中心であり、タンパク質の発現制御はほとんど考慮されていない。一方、本研究は遺伝子制御ネットワークを含む大規模モデルを構築するので、遺伝子組換えによる遺伝子発現制御の動的変化を予測することができる。

有用物質生産性に本質的に関わるグルコース同化システムの性能を向上させることは、バイオテクノロジーの観点から大きな意義がある。合成された有用株はそのまま生産現場で利用することができる。

<達成できなかったこと、予想外の困難、その理由>

研究は順調に進行した。

<今後の課題、展望>

遺伝子制御ネットワークを設計するための合理的戦略を構築するために、グルコース取り込み速度に与える遺伝子発現変化の影響をシミュレーションした。そのような設計や解析を通して、ptsI 遺伝子の過剰発現、mlc 遺伝子欠損がグルコース取り込みを強化することを予測できたので、生物実験を用いて、合理的設計戦略やその予測の正当性を確認し、グルコース取り込みを増強して、高機能の細胞を合成して、バイオテクノロジー分野へ貢献する。

参考資料

Yousuke Nishio, Yoshihiro Usuda, Kazuhiko Matsui, Hiroyuki Kurata, Computer-aided rational design of the phosphotransferase system for enhanced glucose uptake in *Escherichia coli*, *Mol Syst Biol*, 4:160, 2008

<研究期間の全成果公表リスト>

- 0901061748
Kazuhiro Maeda, Hiroyuki Kurata, Two-phase search (TPS) method: Nonbiased and high-speed parameter search for dynamic models of biochemical networks, *IPSI Transactions on Bioinformatics*, 2:1-13, 2009
- 0911230919
Quanyu Zhao, Hiroyuki Kurata, Genetic modification of flux for flux prediction of mutants, *Bioinformatics* 25: 1702-1708, 2009
- 0911230925
Quanyu Zhao, Hiroyuki Kurata, Maximum entropy decomposition of flux distribution at steady state to elementary modes. *J Biosci Bioeng*, 107: 84-89, 2009