Form 3

Synopsis of Thesis

Title: Development of computational method for heterologous pathways design
(異種代謝経路の計算機デザイン手法の開発に関する研究)

Name of Applicant Sunisa Chatsurachai

Abstract

Genome-scale metabolic models of industrial host cells are conventionally exploited as tools for finding engineering strategies to produce native metabolites at high yields. However, it is impossible to produce nonnative metabolites (non-existent metabolites in these hosts), which are value-added products today. To establish a computational method/*in silico* platform for the design of host-specific heterologous pathways/genes for production of nonnative metabolites, an algorithm was developed to search for feasible pathways and to suggest candidate heterologous genes by CAI scores. This *in silico* platform was named as "ArtPathDesign", and effectively designed the heterologous pathways for the production of nonnative metabolites including industrial chemicals such as 1,3-propanediol and 2,3-butanediol. Furthermore, the ArtPathDesign system could automatically identify complex heterologous pathways, e.g. heterologous pathway of sophorol (plant secondary metabolite) by adding 10 heterologous enzymes. These findings could provide options of metabolic engineering strategies to make fine chemicals using microorganisms.

SUMMARY OF THE WORK

Chapter 1. General introduction

Currently, the demands of fuels and chemical feedstocks are largely increased, while petroleum resources are limited and unsustainable. Petroleum-based process for fuels and chemicals generates toxic wastes and CO_2-emissions, and needs high costs for facilities and maintenance. One of the solutions for these problems is utilization of bio-based process to produce industrial chemicals by microorganisms. Natural organisms are not easy to cultivate and produce high levels of bio-product, as well as, metabolic information and genetic manipulation tools of these organisms are unavailable and not being well-developed. Genome-scale metabolic models are metabolic engineering tools, which are widely used for finding strategies to modify/redirect existing metabolic pathways for producing native metabolites. To produce nonnative metabolites, the design of heterologous pathways is required to incorporate to the genome-scale metabolic models of well-characterized hosts, including *E. coli*, *Saccharomyces cerevisiae*, *Bacillus subtilis*, etc. Still, it is difficult to answer the following questions; how do we obtain the feasible heterologous pathways/genes for the production of industrial chemicals from huge number of possible pathways/genes? Which host is suitable for a target production? Several computational methods were reported in order to answer some of those questions[1–3]. Among those methods, host-specific heterologous pathways/genes are not developed yet. Therefore, this study aimed to develop a computational method to overcome these problems and to provide options of metabolic engineering strategies for industrial chemicals by microorganisms.

Chapter 2. Development of an algorithm to design heterologous pathway

Information of metabolic reactions were retrieved from KEGG database and were parsed into a constructed in-house database. This in-house database contained 7,769 known metabolic reactions, 6,635 compounds and necessary information of genes, enzymes, pathways and organisms. All metabolic reactions were considered as candidate heterologous reactions that could be added to the host's metabolic network. To design heterologous pathways, an iterative algorithm was developed to search for a set of heterologous reactions used to connect nonnative metabolites with the host's metabolic network. This algorithm was developed to sequentially adding heterologous reactions to the host's metabolic network. At 1st iteration, a set of nonnative metabolites

which can connect to host's metabolic network by adding single heterologous reaction was obtained. Next, the host's metabolic network was expanded by including a set of heterologous reactions identified from the 1st iteration. The process to find connectable nonnative metabolites obtained by adding heterologous reactions were iterated, and the host's metabolic network was expanded. This iterative process was finished when no further nonnative metabolites can connect to the expanded host's metabolic network. As results, there are 3,244, 3,154, 3,112 connectable nonnative metabolites that are found to connect to *E. coli*, *S. cerevisiae*, and *C. glutamicum* metabolic networks, respectively. The heterologous pathways of those nonnative metabolites were automatically generated by using the developed algorithm, including the identical pathways previously reported for 1,3-propanediol, isoprene, cadaverine, etc[4–6]. With the developed algorithm, the *in silico* platform for the design of heterologous pathways to produce nonnative metabolites was developed and successfully identified heterologous pathways of useful bio-products.

Chapter 3. Selection of heterologous genes using CAI score

The Michaelis-Menten constant, Km, is equal to the substrate concentration at which the reaction rate is half its maximum value. Thus, the smaller the value of Km of the enzyme is, the more efficiently enzyme reaction occurs. Although Km-value is a useful score for selection of candidate heterologous genes, the information of Km is depending on experimental data and unavailable for several enzymes. The inefficient target production by introduced heterologous genes may cause by a low or lack of expression of heterologous enzymes in the host cell. For these reasons, the new score, Codon Adaptation Index (CAI), was applied to select host-specific heterologous genes. CAI score is one of the most measures to estimate the extent of codon usage bias in genes. To evaluate CAI score, the relationship between CAI of genes and protein abundance was investigated, and positive correlation was found in both *E. coli* and *S. cerevisiae*. Therefore, the *in silico* platform developed in chapter 2 was extended by including the CAI score for selecting candidate heterologous genes. This computational platform named as ArtPathDesign was improved to overcome the problems about host-specific heterologous pathways and genes. Besides, the complex heterologous pathway by adding about 10 heterologous enzymes to the host *E. coli* was automatically generated. The host-specific catalog of nonnative metabolites was developed to get information of generated pathways as a good format of an html file easily opened by web browsers such as Google Chrome and Mozilla Firefox.

Chapter 4. General conclusion and future perspective

The ArtPathDesign system was developed and successfully used as a metabolic engineering tool for the design of host-specific heterologous pathways/genes and for providing the host-specific catalog of compounds. The ArtPathDesign will become a useful tool to enable scientists to engineer host microorganisms for industrial chemicals. Since the current version of ArtPathDesign has been developed, the obtained results are depending on metabolic reaction data available on public databases. Generating new candidates of metabolic reactions (even non-existent reactions in databases) is desirable features of pathway design methods. To improve the ArtPathDesign system capabilities, a computational method to design new metabolic reactions will be further developed in the future.

様式７

論文審査の結果の要旨及び担当者

氏　名　（　Sunisa Chatsurachai　）			
		(職)	氏　　名
論文審査担当者	主 査	教授	清水　浩
	副 査	教授	大竹久夫
	副 査	教授	村中俊哉
	副 査	教授	福崎英一郎
	副 査	教授	原島　俊
	副 査	教授	渡邉　肇
	副 査	教授	福井希一
	副 査	教授	紀ノ岡正博
	副 査	教授	金谷茂則
	副 査	教授	永井健治
	副 査	教授	仁平卓也
	副 査	教授	藤山和仁

論文審査の結果の要旨

本論文は、異なる生物の経路を含めた代謝経路のコンピュータ援用設計法の開発に関する研究に関するものである。微生物を用いたバイオ燃料やバイオ化成品の生産は、再生可能な社会の構築のため重要な技術となっている。大腸菌や出芽酵母をはじめとする産業有用微生物に元来は生成しない代謝物質を生産させるためには、その代謝経路や遺伝子を他の生物から取得し、導入する必要があるが、計算機を援用してそのような代謝設計を行うシステムの開発は、十分になされてはいないのが現状である。近年、ゲノム解読による遺伝子、代謝反応のデータは加速的に蓄積されている。蓄積された情報の中から宿主生物に標的代謝物質を生産させるための代謝設計、遺伝子の取得を合理的に行う方法の開発が必須の課題となっている。本論文では、異種の遺伝子や代謝経路を宿主微生物に導入し、代謝経路を設計するための計算機援用システムの開発に関する研究である。本論文は、序章と結論を含め４章からなっている。

１章においては緒言として、代謝工学による微生物の燃料、化成品生産に関する研究背景、ゲノム解読に伴う遺伝子情報や代謝情報の蓄積と利用、計算機援用代謝設計法や代謝予測法の開発の研究背景と意義、および既存の研究成果について述べ、本研究の目的および構成について詳述している。

２章においては、石油化学由来の化学物質、燃料を微生物で生産するための計算機援用代謝設計法の開発について述べている。宿主微生物に対して、宿主微生物が元来生成する経路を有しない物質の生産を目的として、異種生物の代謝反応を取得し、宿主微生物へ導入して物質生産を可能とする経路の合成法を開発している。遺伝子を取得する際に、宿主に対して最も短い経路で標的の化合物を生成する代謝経路を抽出するアルゴリズムを開発している。また、候補遺伝子の中で、酵素が前駆体と強い親和力で結合する能力を有する遺伝子をデータベースから選択し、ランキングを生成する方法を開発している。本方法による既に熟練研究者によって生産されている化合物の代謝経路を自動的に設計することが可能となり、本方法の有効性を確認することが可能となった。

３章においては、コドン適応インデックス（Codon Adaptation Index (CAI)）という指標を用いた遺伝子配列評価に基づく遺伝子の選択に関する方法の開発について述べている。宿主における外来遺伝子の発現量はコドンの使用頻

度によって影響を受ける場合が多い。この影響を考慮して効率的に遺伝子を選択することが経路設計において重要である。本章では、コドンの使用頻度に基づく CAI という評価基準を用いて、より適切な遺伝子の選択法の開発を行っている。コドン使用頻度とタンパク質の発現量は相関があり、本指標を利用することにより、より適切な遺伝子選択が可能となると考えられる。

４章では、結言として本研究で得られた知見をまとめ、計算機援用代謝設計法の開発に関して今後の展望について述べている。

以上のように、本論文は、産業有用微生物の計算機援用代謝設計に関する開発研究を行っている。多くの生物におけるゲノムが飛躍的なスピードで解読され、遺伝子、代謝に関するデータベースの蓄積は非常に充実している。本論文は、合理的な代謝経路の計算機援用設計法の開発を行ったものであり、今後、多くの産業湯曜日生物に対する多様な物質の生産の代謝経路改良設計に大きく貢献することが期待される。

よって本論文は博士論文として価値あるものと認める。