慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	E-Cellプロジェクト20周年
Sub Title	20th anniversary of the E-Cell project
Author	高橋, 恒一(Takahashi, Koichi) 内藤, 泰宏(Naito, Yasuhiro) 佐野, ひとみ(Sano, Hitomi)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2015
Jtitle	Keio SFC journal Vol.15, No.1 (2015.) ,p.40- 62
JaLC DOI	10.14991/003.00150001-0040
Abstract	E-Cellプロジェクトは, 1995年の発足以来一貫して, その究極のゴールを全細胞シミュレーションに定めている。本稿では, E-Cellプロジェクトの発足から, 汎用的な細胞シミュレーション環境の実現を目指したE-Cell Systemの開発, 仮想的な自己維持細胞を始めとする様々な細胞のモデル構築, そして, 代謝, 遺伝子発現, 信号伝達の統合シミュレーションの実現を目指したE-Cell System Version 4の開発までの軌跡を辿る。
Notes	特集 世界を救え : SFCバイオの挑戦#招待論文#第1章 SFCが先導した生命科学
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=0402-1501-0040

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

［招待論文］

E-Cell プロジェクト 20 周年

20th Anniversary of the E-Cell Project

高橋 恒一
理化学研究所生命システム研究センター生化学シミュレーション研究チーム
チームリーダー

Koichi Takahashi
Team Leader, Laboratory for Biochemical Simulation, RIKEN Quantitative Biology Center

内藤 泰宏
慶應義塾大学環境情報学部准教授

Yasuhiro Naito
Associate Professor, Faculty of Environment and Information Studies, Keio University

佐野 ひとみ
慶應義塾大学環境情報学部専任講師

Hitomi I. Sano
Assistant Professor, Faculty of Environment and Information Studies, Keio University

Abstract:　E-Cell プロジェクトは、1995 年の発足以来一貫して、その究極のゴールを全細胞シミュレーションに定めている。本稿では、E-Cell プロジェクトの発足から、汎用的な細胞シミュレーション環境の実現を目指した E-Cell System の開発、仮想的な自己維持細胞を始めとする様々な細胞のモデル構築、そして、代謝、遺伝子発現、信号伝達の統合シミュレーションの実現を目指した E-Cell System Version 4 の開発までの軌跡を辿る。

The ultimate goal of the E-Cell Project, founded in 1995, is to construct a whole cell model. In this review, we describe how we launched the E-Cell Project, developed E-Cell System Version 1 to 3, constructed various cell models including the hypothetical self-sustaining cell on the E-Cell System, and developed E-Cell System Version 4 to enable integrated simulation of various cellular processes including metabolism, gene expression, and signal transductions.

Keywords: マイコプラズマ、自活細胞モデル、マルチアルゴリズム・マルチタイムスケールシミュレーション、1 分子粒度シミュレーション

mycoplasma genitalium, self-sustaining cell model, multi-algorithm and multi-time scale simulation

1　ことの起こり

E-Cell プロジェクトの発足は 1995 年である。この年はいくつかの細菌とともに、自立して細胞を構成する生き物として当時最小のゲノムを持つと考えられた Mycoplasma genitalium（マイコプラズマ）の全ゲノム情報[1]が公開された年であった。当時慶應義塾大学環境情報学部助教授だった冨田勝（以降冨田研究室の伝統にちなみ「冨田さん」とする）は、一個の生命を構成するこのマイコプラズマのゲノム情報が高々 580K base pair (bp) 、つまり 58 万塩基対[注1]でしかないことに驚き、この全ゲノム情報をデジタルコンピューター上で展開して細胞まるごとシミュレーションするという E-Cell プロジェクトの原型となる着想を得た。

1.1 「マイコプラズマプロジェクト」

1996 年 4 月からの春学期からは、早速このマイコプラズマの全ゲノムシミュレーションに向け冨田研究室の学生が手分けしてこのゲノムがコードする 480 程度の全ての遺伝子の既知機能を調査、整理してデータベース化するプロジェクトが始動した（通称「マイコプラズマプロジェクト」と呼ばれていた；なお現在ではこのマイコプラズマゲノムが持つ全ての遺伝子は 525 個とされている）。本稿執筆者の高橋も 1996 年 4 月から冨田研究室に参加した学部 3 年生のメンバーの一人である。

1996 年 10 月からの秋学期では、このゲノム情報を具体的なコンピュータープログラムとして組み上げる方法の模索が始まり、何名かの学生が各々違ったアプローチでプロトタイプの作成に取り組んだ。このプロトタイピングでは細胞の増殖速度に注目したものや遺伝子の発現レベルに注目したものなど様々な可能性が試された。高橋も冨田さんから「高橋君、君ならどうする？」とけしかけられたうちの一人であった。ひとしきり分子遺伝学、細胞生物学、酵素反応速度論などについての文献を読み込んだ後、高橋が 11 月 15 日にプロジェクト内向けに行ったプレゼンテーションでは、このプロジェクトの方向性として (1) 代謝系酵素の反応速度論（キネティクス）とそれら酵素をコードする遺伝子の発現制御に絞ったモデリングを出発点とすること、また (2) 当時注目を浴びていたオブジェクト指向の思想と親和性の高い C++ 言語

で汎用の細胞シミュレーションソフトウエアシステムを開発する事、の 2 点を提案した。結果的に、この提案が初期のプロジェクト設計を方向づける事となった。

1.2　E-Cell プロジェクト創成期

高橋は文字通り寝食を忘れてこのソフトウエアの開発に取り組み、12 月 29 日にはアルファ版 (version 0.0-pre；アルファ版とは関係者向けに限定的に公開されるプロトタイプの事) をプロジェクト内部向けに公開した。当初のプログラムは C++ 言語で 3,500 行ほど、現在よりもかなりコンパクトなものであったが、細胞内の酵素一つひとつの反応速度論に基づいて、細胞内の代謝反応ネットワークの挙動をオブジェクト指向でモデリングし、シミュレーション計算をするために必要な最小限の機能を備えていた。このソフトウエアは当初 Electronic Cell Laboratory と呼ばれていたが、後に橋本健太 (当時学部 4 年生；その後クックパッド社の設立に関わり、最高技術責任者などを歴任) の発案で現在のように Electronic Cell System (E-Cell System) と呼ばれる事になった。この E-Cell System の「アルファ版」公開前からの世界で最初のユーザーは高橋と同じく当時学部 3 年生だった松嶋亮であった。

初期の E-Cell System のオブジェクトモデルは物質の量を表す Substance オブジェクトと物質間の反応を表す Reactor オブジェクトの 2 つのオブジェクトクラスの組み合わせで反応ネットワークを表現するもので、Substance-Reactor モデルと呼ばれた。1997 年の 3 月までには、高橋、橋本、さらに当時修士課程 1 年生であった清水友益 (ケンブリッジ大学、ハーバード大学を経て現在オランダ原子分子物理研究所 AMOLF [注 2] で研究室を主宰) の間で行われた議論に基づき、細胞内の物理的あるいは機能的なコンパートメントを表現する System クラスを加え、プリミティブと呼ばれる 3 つのオブジェクトクラスの組み合わせで細胞内のあらゆる機能を表現する方針を固めた。このオブジェクトモデルは Structured Substance-Reactor モデルと呼ばれた。その後、E-Cell System へは橋本を中心として化学平衡を計算する Reactor クラス、一般的な遺伝子発現を表現するモデルモジュールが追加さ

図 1　E-Cell プロジェクト 2000 年頃

れた。我々のプロジェクトが E-Cell プロジェクトと呼ばれ始めたのはこのころである。さらに 1997 年夏までには当時学部 3 年生だった柚木克之（現在東京大学大学院理学系研究科）らが全ゲノム規模での遺伝子発現状況（主に各遺伝子に関連する mRNA: messenger RNA の量）を一望出来る可視化モジュールを開発し、これは後のデモンストレーションで好評を博することになる。このように、E-Cell System の開発は冨田研究室の多くの仲間を巻き込んで飛躍的な速さで進み、冨田さんをして「これなら行ける」と思わせたのか否か、マイコプラズマゲノムを解読したアメリカ TIGR [注3] 研究所（現在は J. C. Ventor 研究所）まで押しかけ合宿で細胞シミュレーターを完成するという思い切った動きに発展する。

1.3　アメリカ合宿と自活細胞モデルの完成

1997 年の 8 月から 9 月にかけての夏休み、E-Cell プロジェクトの主要なメンバーはアメリカワシントン D.C. 郊外、バルチモアにある TIGR 近くのアパートを何室か借り、6 週間に渡る合宿を行った。前述の通り TIGR はショットガンシーケンス法を武器として当時いくつかの生物種の全ゲノム解読で独走していた。我々は TIGR のマイコプラズマ菌の専門家であるクライド＝ハッチソン博士とも議論し、当時まだアノテーションが十分整備されていなかったマイコプラズマの 480 遺伝子程度のゲノムから 120 個の遺伝子を選び、さらに 7 個を大腸菌などの他のモデル生物から拝借して、127 遺伝子を持つ仮想細胞を構築する事にした。この仮想モデル生物は、グルコース

の取り込みから連なり細胞内のエネルギー通貨である ATP 分子を生産する解糖系、酵素などのタンパク質の材料となるアミノ酸生合成系、細胞膜を構成するリン脂質の生合成系、さらに遺伝子の転写と翻訳に必要となるポリメラーゼやリボソーム、転移 RNA などをコードする遺伝子を持ち、細胞周期を持たず増殖も環境変化への応答を行うシグナル伝達系も持たないが、自分自身の生命活動を維持するのに必要な最小の遺伝子セットを持つものであった[2]。この合宿に参加したのは、高橋、橋本、清水、柚木に加え、松崎由理(現在東京工業大学)、三由文彦、斎藤佳奈子、谷田さくらの 8 名の学生と冨田さんの計 9 名であった。この成果はさらに磨き上げられ、TIGR のディレクター J. クレイグ＝ベンターを最終著者とし、またハッチソン博士も著者に加え、1999 年に Bioinformatics 誌に掲載された[3]。この論文の 2015 年 2 月時点までの引用回数は約 800 回である。また、同年には Science 誌からは“Building Working Cells ‘in silico’”[4] とするコメンタリーで、また Nature 誌から“Going for Grand Challenges”[5] として 21 世紀へ向けた挑戦への先駆けとして、それぞれ注目を浴びた。

2　シミュレーション環境の開発：E-Cell System Version 3 まで

2.1　E-Cell System Version 1.0 と SBML (Systems Biology Mark-up Language) コミュニティーへの貢献

E-Cell System の開発は冨田研究室の多くの学生やプログラマーにより改良が加えられ、2000 年 3 月 2 日には E-Cell Beta-testing License の下で実行形式のインターネットでの配布を始めた。さらに 2000 年 10 月 13 日には GNU General Public License version 2 の下でオープンソースソフトウエアとなった。その後、Linux で動作する E-Cell System Version 1 (E-Cell 1) の派生版として、石川直太により Windows で動作する E-Cell System Version 2 (E-Cell 2) が作成された[6]。

1999 年の終わりから 2000 年にかけて、北野宏明とジョン C. ドイルの呼びかけにより、当時黎明期にあったシステム生物学におけるコンピューターモデリングプラットフォームの開発者がカリフォルニア工科大学で一同

に会し、E-Cell プロジェクトも招聘された。数度に渡る議論の結果、システム生物学研究における共通モデル交換フォーマットの必要性が合意され、2000 年 11 月に東京で開催された国際システム生物学学会[注4]の初回で行われた議論などをベースに SBML(Systems Biology Mark-up Language) の仕様が策定された[7]。E-Cell 1 の思想の一部も反映された SBML はその後のシステム生物学研究におけるデファクト・スタンダードとなった。

2.2　E-Cell System Version 3.0

主な開発者だった高橋が大学院政策・メディア研究科で修士号を取得する 2000 年 3 月頃には、E-Cell 1 の仕様はほぼ成熟していた。博士課程に進んだ高橋は、細胞生物学におけるシミュレーションプラットフォームの要求仕様について考察し、以下の 7 つの要求を定義した[8]。

(1) マルチアルゴリズムシミュレーション
（細胞内の性質の違う様々な現象を扱うため）
(2) マルチタイムスケールシミュレーション。
（細胞内の時間スケールの違う様々な現象を扱うため）
(3) オブジェクト指向モデル
（複雑な細胞内分子や構造物のオントロジーを表現するため）
(4) オブジェクト指向シミュレーション
（オブジェクト指向構造を保持したままシミュレーションを行う事）
(5) リアルタイムユーザーインタラクション
（複雑な細胞モデル挙動の理解を助けるため、リアルタイムで実行中のシミュレーションモデルを操作出来る事）
(6) 動的なモデル構造
（細胞分裂のように自発的かつ動的な構造変化）
(7) 空間的なモデリングとシミュレーション
（細胞内の分子の局在や拡散運動の表現）

このうち (3), (4), (5) は E-Cell 1 でサポートしていたが、残りの全てに対

応するためには抜本的なソフトウエアの再設計が必要であったため、博士課程での研究として E-Cell 1 と E-Cell 2 で用いられた第一世代カーネルに続く第二世代カーネルを搭載した E-Cell System Version 3 (E-Cell 3) の開発が始まった。

E-Cell 3 の開発では、特に (1) と (2) で提示するマルチアルゴリズム、マルチタイムスケールの対応が焦点となった。例えば、多数の小分子が関与する代謝系では常微分方程式により連続量で表現された濃度変化の計算が行われるが、遺伝子発現系では少数の分子により行われるため離散的かつ確率論的である。その他にも細胞周期やシグナル伝達系、細胞膜や細胞内小器官の構造変化など、多種多様な現象が存在し、それらを統一的に扱える汎用のシミュレーションフレームワークをいかにつくり上げるかが大きな課題であった。

E-Cell 3 では離散事象法を統一の「世界観」とし、様々なシミュレーションアルゴリズムがこの世界観に基づき変数と時間の更新を行う事で統合するのが基本的なアイディアであった。開発の難関は、連続的に変化する連続量と離散的に変化する離散量の更新をいかに矛盾なく行うかであった。この問題にヒントをもたらしたのが、共同研究先の理化学研究所の戎崎俊一博士との議論の中で紹介されたエルミート積分法（重力多体問題などで粒子一つひとつが別々の時間ステップを持ち非同期で運動量を更新する際に用いられる）である。エルミート積分法と離散事象法は通常組み合わせて用いられる事はないが、これらを矛盾なく組み合わせるシミュレーションステップの進行手順を突然ひらめいたある明け方が、高橋にとっての E-Cell 開発のクライマックスの１つであった。E-Cell 3 は 2004 年にリリースされた [9]。E-Cell 1 から E-Cell 3 にかけての開発では、中山洋一、ガボール＝ベレツキー、クリストファー＝ピケット、ビン＝フ (SRA インターナショナル)、石田達也、石田貴士（東京工業大学)、北山朝也、山田洋平、伊東大輝、丸岡和人、高橋健太郎、藤田由起、杉本昌弘（慶應義塾大学先端生命科学研究所)、井元淳（ヒューマンメタボロームテクノロジーズ)、岡山真之、斎藤裕介などをはじめとする多くの学生、研究スタッフが貢献した。

このように開発された E-Cell 3 のマルチアルゴリズム・マルチタイムスケールシミュレーションフレームワークが力を発揮した例としては、大腸

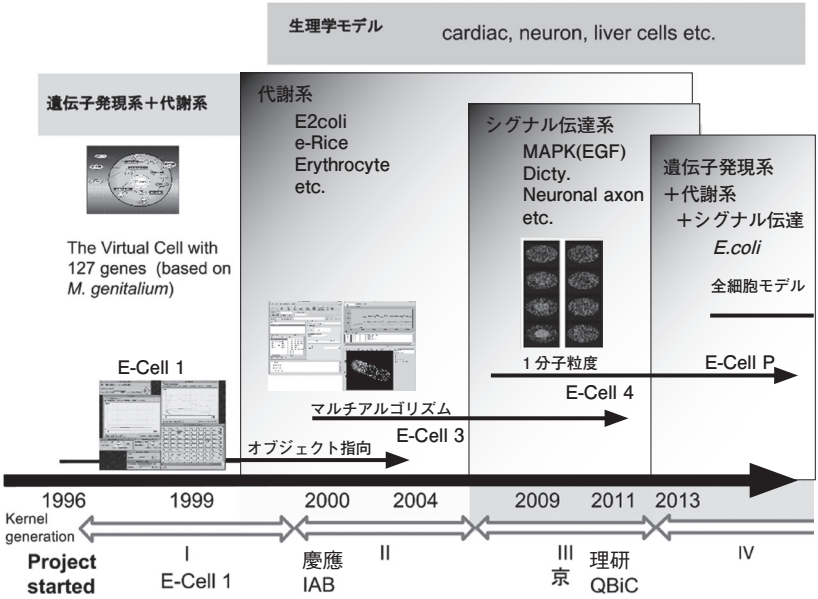

図 2　E-Cell のカーネル世代、それぞれの特徴や主なプロジェクト

菌ヒートショックモデルにおける確率・離散的手法と決定論的・連続的手法との結合、代謝系の動的＝静的ハイブリッドモデリング手法の開発などがあった。

(6) 動的なモデル構造と (7) 空間的なモデリングとシミュレーションは E-Cell System Version 4 (E-Cell 4) 以後の課題として持ち越された。

3　細胞モデルの開発

前述の通り、E-Cell プロジェクトは、そもそも自活細胞の数理モデルとそのシミュレーション環境を同時に開発する研究プログラムとして始まったものであり、その後も、シミュレーション実行環境の開発と細胞モデルの構築を平行して進め、相互にフィードバックしあうスタイルを現在まで維持している。

3.1　赤血球モデル

モデル化対象として最初期に注力した対象は、赤血球だった。赤血球は、細胞シミュレーションの対象に適する特徴を有しており、1970 年代からその代謝経路の数理モデル化とシミュレーションの研究が行われている [10]。ヒトをはじめとする哺乳類の赤血球は、その成熟過程の途上で、細胞核やミトコンドリアといった細胞内小器官の大部分を失っている。これに伴い、タンパク質を合成するシステムが失われるため、成熟赤血球では、新たなタンパク質の合成は起こらない。赤血球は、骨髄から遊出した際に細胞中に含んでいたタンパク質分子によって淡々と代謝を実行する、いわばミクロの試験管であり、細胞シミュレーションのプロトタイプとして格好の対象だった。

赤血球の代謝経路の中でも、生命活動の基幹となるエネルギー合成経路を主たるモデル化対象とした研究は 1970 年代から盛んに行われてきた。これに対し、E-Cell プロジェクトは、木下綾子、西野泰子、下英恵 (ケンブリッジ大学大学院博士課程)、中山らを中心として、プロジェクトの目標である全細胞シミュレーションを見据え、赤血球についても赤血球に含まれる全代謝経路のモデルからなる全赤血球シミュレーションをめざした [11]。

赤血球は、酸素の運搬に特化した細胞であり、酸素と結合し運搬するヘモグロビンを大量に含んでいる。酸素運搬という機能ゆえに、赤血球は絶えず高度な酸化ストレスにさらされている。全赤血球モデルを構築すると、酸化ストレスなどの外界からの刺激、環境の変化が、細胞に対して総和としてどんな影響を及ぼすかを計算し、推定することができる。E-Cell プロジェクトで開発した赤血球モデルを用いて、赤血球が酸化ストレスに対応するメカニズムの推定 [12, 13]、しばしば見られる酵素の遺伝的異常の影響が代謝経路全体に波及する状況の推定 [14] などに成功した。また、実践的な研究として、長期保存した赤血球に起こる変化のシミュレーションを通し、長期保存後の赤血球の状態を新鮮血に近い状態に回復するために必要な条件の解析なども行っている [15, 16]。

3.2　ミトコンドリアモデル

ミトコンドリアは動物や植物などの真核生物の細胞に含まれる細胞内小器

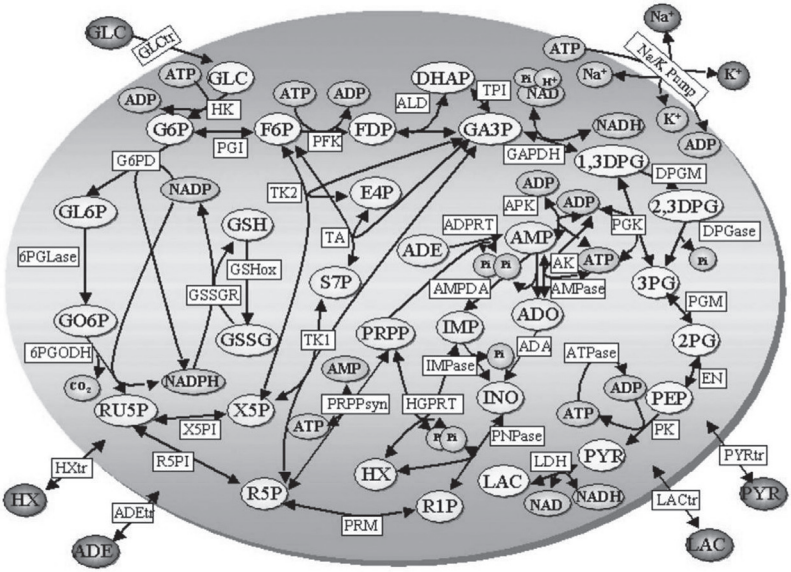

図 3　E-Cell 赤血球モデルの概要

官で、エネルギー代謝経路の中心的な役割を持つ分子が多数集約された「エネルギー産生工場」である。E-Cell プロジェクトでは、柚木、中山らを中心に、ミトコンドリアに含まれる諸要素を個別に解析した先行研究を統合し、58 種の酵素、117 の代謝物質からなるミトコンドリアの全代謝モデルを構築した[17]。このモデルには、ミトコンドリアの主要代謝経路である電子伝達系、TCA 回路、脂肪酸 β 酸化の他、外部環境との物質交換を担うミトコンドリア内膜の輸送系が含まれている。

平行して大規模なモデルの構築手法の開発にも取り組んだ。対象に含まれる全代謝経路をモデル化しようとした時、全ての反応について詳細な情報があることは稀で、多くの場合、情報の不足する反応がある。特に、反応ネットワークの時間変化を予測しようとする動的なシミュレーションをめざす場合、必要な情報が全て揃うことはほとんど期待できない。柚木、中山らを中心としたチームは、一部の酵素について動特性に関する情報が得られない場合にも、近隣の反応を触媒する酵素群の動特性から代謝ネットワーク全体のダイナミクスを

精度良く予測する動的／静的ハイブリッド法を開発した[18]。

3.3　概日リズムモデル

生物には、自律的に１日を計測する「体内時計」が備わっていることが知られており、概日リズムと呼ばれている。概日リズムの分子機構の研究から、一つひとつの細胞に概日リズムを計測する遺伝子回路が備わっていることがわかってくると、回路の出力の明解さ（約１日という周期で振動する振動子）と、それを実現する遺伝子回路が比較的単純であることから、概日リズムは数理モデル化の好適な対象になった。E-Cell プロジェクトでは、三由、小川らが概日リズムを構成するネットワークのモデル化に取り組み、分子間相互作用の様式の予測[19]、複数の数理モデルの比較研究による概日リズムの遺伝子回路の特性の解析[20]などを行った。

3.4　肝小葉モデル

生命科学を実践的に応用する際、その対象には多くの多細胞生物が含まれる。多細胞生物は、単なる細胞の寄せ集めではなく、細胞を要素単位とした多階層の高次構造をもっている。例えば動物では、ミクロからマクロにかけて、細胞～組織～器官（臓器）～器官系～個体といった階層が解剖学的に定義されており、それぞれの階層は、その直下の階層を要素とする秩序だった構造を有している。

細胞シミュレーションを生物個体のシミュレーションへと敷衍する上で、下位階層である細胞のモデルを活用しながら、いかにして上位階層を数理モデル化するかが重要な課題となる。大野浩、内藤泰宏らは、比較的小規模な機能ユニットの繰り返し構造として理解可能な器官として肝臓を選んだ。肝臓は、肝小葉と呼ばれる機能単位が超並列回路を構成して血管などの入出力経路に接続されている器官として記述可能である。大野らは、入力血管である門脈および肝動脈の細枝と、出力血管である中心静脈を結ぶ類洞と呼ばれる多孔性の特殊な毛細血管と、類洞に沿って並ぶ肝細胞を機能単位とする肝小葉の数理モデルを構築し、肝小葉内ではたらく部位特異的な遺伝子発現調節の機能的な意義を解析した[21]。

3.5　心筋細胞モデル

心筋細胞の動的な挙動、主に膜興奮を数理モデルで説明する試みは、半世紀以上前から行われている。1960 年代に Hodgkin-Huxley 式に基づき Na+ チャネルと K+ チャネルが流す電流 2 種類が数式で表されていた心筋細胞の電気生理学モデルは、細胞膜を流れる電流の精密な定量が可能なパッチクランプ法の発達により、単一細胞レベルに留まらず、単一イオンチャネルを流れるイオンの高精度な定量データの蓄積とともに、心臓生理に関連する現象を高精度に再現できるモデルへと発展した。

E-Cell プロジェクトでは、京都大学の野間らが開発を進めていた心筋細胞の包括的心筋細胞モデル [22] の E-Cell System 上での構築に取り組んだ。代謝ネットワークを主たる対象としていた E-Cell プロジェクトにとって、細胞の電気生理現象、筋収縮などの物理現象のモデルの開発は、全細胞モデルに向けても重要な挑戦になった。内藤、橋本、米田元洋、石鍋沙耶花、宮本章子らによって開発が始まり、ほどなく佐野ひとみらが参加した。

佐野らは、心臓が拍動を開始する胎仔期から成体に至るまでの過程をシミュレートすることを目指した。拍動を開始して間もない心臓を構成する細胞の半数以上は、「ペースメーカー」としての機能も有する。「ペースメーカー」と「収縮」の両方の機能を有する心筋細胞は、胎生期の発生により発達した刺激伝導系から伝わる刺激に応答して収縮する受動的な細胞へと変化を遂げる。佐野らは、発生過程による膜興奮の変化がイオンチャネルの量的変化のみで説明できることを示した [23]。また、胎生期の発生過程において、恒常的な収縮の維持に寄与する要素を同定するため、胎生期に量的変化を伴う 9 つの要素を対象に、量的変化の組み合せを網羅的に検討する新たな方法を提案した [24]。

4　シミュレーション環境の開発：E-Cell System Version 3 以降

2004 年の E-Cell 3 のシミュレーションエンジンの完成を受け、2004 年 10 月から 2010 年 3 月までの 5 年半では科学技術振興機構（JST [注 5]）の研究助成（CREST [注 6]）を受け、研究課題「システムバイオロジーのためのモ

デリング・シミュレーション環境の構築」の遂行と通じて、E-Cell System の開発は一層の広がりを見せた。

この時期の E-Cell System 開発の中心となったのは、(1) E-Cell System Version 3 のシミュレーションエンジンをベースに細胞モデリング・シミュレーション統合開発環境の整備を目的とする E-Cell IDE (Integrated Development Environment)、そして (2) E-Cell 3 のさらに先、E-Cell 4 へと向かう先進的なシミュレーション手法とその要素技術の開発の主に 2 つの活動であった。

4.1　E-Cell IDE (Integrated Development Environment)

E-Cell IDE は E-Cell 3 をシミュレーションエンジンとして、モデリング・シミュレーション・解析環境をユーザに提供するためのソフトウェアプラットフォームである。開発プロジェクトリーダーは当時大学院生の櫻田剛史が担当し、三菱スペースソフトウエア株式会社との共同開発であった。E-Cell IDE は、シミュレーション技術の専門知識を持つ研究者だけでなく、生化学・分子生物学の実験ラボの研究者のような一般的な生物学者にとっても幅広く使えるソフトウェアを目指した統合プラットフォームで、モデリング・シミュレーションのプロジェクト管理機能、パスウェイエディタ機能、シミュレーション可視化 ・解析・デバッグ機能、さらに分散コンピューティングによりバックエンドの計算機クラスター行われる並列計算を行う機能を備えていた。E-Cell IDE の開発は現在では終息しているが、その機能開発の成果は現在の E-Cell 3 本体にフィードバックされたほか、最新版の E-Cell 4 にも生かされている。

4.2　カリフォルニア拠点の始動

E-Cell 3 のマルチアルゴリズム手法の開発で大学院政策・メディア研究科の博士学位を取得した後、高橋はドリュー＝エンディ (MIT [注7] 当時) から、当時米国国立保健衛生研究所 (NIH [注8]) の大型助成を受けて酵母のフェロモン応答経路のモデリングプロジェクト (アルファプロジェクト) を率いていたロジャー＝ブレント (現在フレッドハッチソン癌研究センター) を紹

介され、アルファプロジェクトの本拠地であるMSI[注9]（米国カリフォルニア州バークレー市）で客員研究員の身分を得た。MSIは2002年にノーベル賞を受賞したシドニー＝ブレナーが1996年に設立した独立系の非営利研究機関で、後述のルールベース反応経路モデリング技術やレーザー顕微鏡による一細胞観察技術などの技術面だけでなく、所長・最高経営責任者のブレントの下で細胞内反応経路のノイズ解析や細胞間の応答不均一性など、当時から10年を経た2015年現在の水準で考えても先端的な思想の下に研究を行う、システム生物学の発信源の1つであった（アルファプロジェクト収束後、現在MSIはシリコンバレー中心地のミルピタス市に場所を移し、バイオテクノロジーのインキュベーターとして活動）。

SFCの教員も兼務し、日本と米国を行き来していた高橋は、2005年10月からはフランスストラスブール市に本部を置く国際組織ヒューマン・フロンティア・サイエンス・プログラム（HFSP）の学際フェローの第一期生に選ばれ、MSIに本拠を移して研究を開始した。これとともに、MSIはJSTプロジェクトの一拠点として2008年10月までの3年間E-Cellプロジェクトに加わった。

4.3　最後のピース、シグナル伝達系

E-Cell Systemのシミュレーションエンジンの開発において、E-Cell 1の第一世代カーネルでは確率的かつ離散的な遺伝子発現系モデルと決定論的かつ連続的な代謝反応系の結合が、またE-Cell 3の第二世代カーネルでは代謝系モデルの中に混在する、反応速度論に基く動的なモデルと代謝流束モデルに基く静的なモデルの矛盾のない結合がモデル構築グループ側からの大きな要請であった。これらの要求に応えることを通じ、実験的にいかにパラメータを取得しモデルを構築するかは依然問題として残っていたものの、計算技術としては遺伝子発現系と代謝系を扱う方法は概ね確立されていた。E-Cellプロジェクトの最終目的である「細胞まるごとシミュレーション」の実現に向けて、残った最後の主要な生化学経路がシグナル伝達系であった。

ATPやグルコースなど、小さく、速く、大量にある小分子が主役の代謝系や、DNAやポリメラーゼ、リボソームなどの核酸とタンパク質の複合体

を中心として形成される巨大な超分子機械が主役の遺伝子発現系と異なり、シグナル伝達系では比較的大き目の分子であるタンパク質であり、また分子間の反応は速く単純なものが多い。その結果として、シグナル伝達系では分子の拡散運動や局在、細胞内巨大分子混雑などの空間的な要素を考慮しなければシミュレーションの結果が正確さを欠いてしまう場合がある[25, 26]。それまでも、大腸菌走化性経路や神経細胞の長期抑制モデルなど、いくつかの優れた部分的なシグナル伝達経路モデルが E-Cell 1 や 3 を用いて開発されていたが、より一般のシグナル伝達経路の挙動を精密に表現するためには、細胞内空間での分子の拡散運動の表現が必須であった。E-Cell 3 の次、E-Cell 4 に向かっては、もう一度シミュレーターの設計を抜本的に見直すだけでなく、それまでは存在しないいくつかの要素技術の開発から始める必要があった。

4.4　地道な要素技術開発

MSI で E-Cell プロジェクトのために開発された技術のうち大きなものは、1 分子粒度のシミュレーション技術そして反応ネットワークモデルで起きる組み合わせ爆発を抑制するルールベースモデリング技術の 2 つであった。

これまでの電気回路のような非空間的なシミュレーションから飛躍し、粒子を用いて細胞内の分子一つひとつの運動のレベルから細胞内の反応ネットワークの挙動を再現する「1 分子粒度」のシミュレーション技術はシグナル伝達系のように分子の空間的な配置や運動が系全体の性質に影響する際には非常に効果的であるが、それまではアドホックな近似を用い精度が十分に保証されないものや、性能が不十分でごく短い時間スケールの計算にしか用いることが出来ないものが多かった。高橋は、当時ハーバード大学ハワード＝バーグの研究室に籍をおいていた E-Cell プロジェクト創設メンバーの一人、清水友益から入手した一本の未発表論文を読み、オランダ AMOLF のピーター＝レイン＝テンウォルデ教授が画期的な超高性能、高精度の粒子反応拡散系シミュレーション法である Greens Function Reaction Dynamics (GFRD) を開発していることを知った。早速テンウォルデ教授を尋ねた高橋のアムステルダム訪問は、GFRD の弱点を克服し、より高性能かつ高精度な新手法

である The enhanced Greens Function Reaction Dynamics (eGFRD) の共同開発へと進展した。eGFRDのコアはスモルコフスキーの反応拡散方程式の厳密解であるグリーン関数を用いて粒子の位置の更新と反応の処理を行う複雑な離散事象法である。eGFRDのコードはSFCにあってプログラミングスキルで一度も引け目を感じたことがない高橋をして何年にも渡る灰色の戦いを強いるほど難航したものの、2008年頃までには完成した。eGFRD法は2015年現在においても近似を用いない正確な粒子反応拡散手法としては世界最高の性能を誇り、世界で初めてMAPキナーゼの2重リン酸化反応における酵素＝基質再結合が系の挙動を大幅に変える事を示した[27]。これらの仕事はAMOLFと理研の他、スウェーデンのウプサラ大学およびアムステルダム自由大学でそれぞれ発展させられ、2名の博士が誕生するなど、現在でも注目を浴びている。

一方慶應側では、米国留学前の高橋がシンガポールでリクルートし政策・メディア研究科博士課程への進学を薦めた大学院生サティア＝アルジュナンが空間格子を用いた高機能な1分子粒度手法であるSpatiocyteを開発し[28]、2009年に博士学位を取得した。これはその後スパコン「京」で動作する超並列シミュレーターpSpatiocyteの基礎ともなる。

シグナル伝達系を扱うためのもう1つの重要な技術が、ルールベースモデリングである。シグナル伝達系のタンパク質分子はしばしば複雑な複合体を形成し、またリン酸化などの共有化学修飾を受けるため、その組み合わせは膨大な数になる。例えば、酵母アルファ経路に関与するたった20種類程度のタンパク質が構成する複合体の組み合わせは10万を超える。このため、MSIではラリー＝ロック博士を中心に反応の種類を全てあらかじめ数え上げるのではなく少数の分子間の結合関係のルールとして記述出来るMoleculizerソフトウエアを開発した。MSIでE-Cell System開発に取り組んだプログラマーのネイサン＝アディは、この技術をE-Cell 3に移植し、その有効性を示した。また、この時開発された基礎コードは、同時期にMSIに在籍していたスティーブ＝アンドリュースが開発する著名なシミュレーターSmoldynでも利用されている。

4.5 「京」の理化学研究所へ

IAB で E-Cell プロジェクトにも関わったパワン＝ダール博士はその後理化学研究所横浜研究所のゲノム科学研究センター（GSC [注10]）で研究室を主宰しており、3 年間の HFSP フェローシップ後の展開を考えていた高橋をセミナーに招聘した。同じく GSC のメンバーであった泰地真弘人博士はその後の訪米時に高橋を MSI に訪ね、当時計画中であった 10 ペタフロップス級スーパーコンピューター（その後「京」と名付けられる）に向けた計算生物学プロジェクトに誘った。これに応じた高橋は 2008 年 10 月から帰国し横浜理研に研究の場所を移して生化学シミュレーション研究チームを立ち上げ、基幹研究所等を経て、2011 年からは大阪府に新設された理研生命システム研究センター（QBiC [注11]）のメンバーとなった。基幹研究所時代、2009 年にはドイツで並列離散事象法を専門とするロストック大学のアデリンデ＝ウールマッハー教授の研究室から大学院生のマティアス＝エシュケが滞在し eGFRD の分散並列化に取り組んだり、2010 年には AMOLF からトーマス＝ミエデマが同じく eGFRD コード開発に協力するため滞在したりと、国際協力も盛んであった。理研の研究チームには、サティア＝アルジュナン、E-Cell IDE と eGFRD のコード開発にも貢献した小泉守義、SFC 学部時代に E-Cell 3 開発にも参加した海津一成なども合流した。これを受けて、E-Cell プロジェクトも慶應 SFC・IAB と理化学研究所の共同プロジェクトとして発展していくこととなった。

4.6　理研 QBiC と E-Cell System Version 4

QBiC は、計算生物学だけでなく、計測技術や合成生物学に力を入れた、理化学研究所としては初めてシステム生物学をミッションの中心に据えた研究センターである。これまでの要素技術の開発を受けて本格始動した E-Cell 4 の開発は、高橋が基本コンセプトを固めた後、海津が実質的な開発のリーダーとなっている。QBiC が特に得意とするレーザー蛍光顕微鏡を用いた生きた細胞の 1 分子計測と E-Cell 4 の 1 分子粒度シミュレーション手法群と粒度がよくマッチするため、これまでの細胞モデリングの大きな問題であった実験データとシミュレーションの入力データのミスマッチが解消

される可能性がある。素粒子物理分野から合流した渡部匡己は レーザー蛍光顕微鏡シミュレーターを開発し、実験とシミュレーションの融合に取り組んでいる。また、宮内敦と岩本一成は Spatiocyte を発展させ「京」で動作する pSpatiocyte を開発し、ほ乳細胞の細胞まるごと規模での大規模なシグナル伝達経路シミュレーションを実現した。

5　我々はなぜ細胞シミュレーションの実現を目指すのか

ATGC のゲノム配列だけを見て、それがどんな形をした生き物で、どんな環境で生き、何を食べ、どれくらいの寿命を持つのかを知ることは可能だろうか？　ワトソンとクリックの DNA 二重らせん構造の発見から 60 年、分子生物学が成熟した現在においても、我々はこの「遺伝子型から表現型を直接予測する技術」をいまだ手にしていないし、今後 5 年、10 年で可能になるような簡単な問題であるとも考えられない。しかし、もし今から 50 年後、100 年後の将来の生命科学が「遺伝子型から表現型を予測する技術」、そして「表現型を自在に操作するための遺伝子型の設計技術」を基礎技術としない姿をしているとしたら、これも考えられない事ではないだろうか？　では、2015 年の今現在、すべき事、出来ることは何だろうか？　我々が 21 世紀初頭の今成すべきこと、それこそが、細胞シミュレーション技術の構築であり、それがゆえに細胞シミュレーションは 21 世紀最大の科学的挑戦の 1 つと定義されている [29]。

5.1　原点回帰と次の一手

ここまで見て来たように、E-Cell 1 の遺伝子発現系と代謝系の複合シミュレーションに始まり、E-Cell 3 では代謝系のより大規模かつ精密なモデリングや異る時間スケールなどの扱い、さらに E-Cell 4 に向けてはシグナル伝達系を精度よくシミュレーションするための技術基盤が揃った。つまり、細胞内の 3 つの主要な生化学反応経路をカバーするシミュレーション技術が E-Cell プロジェクト開始から 20 年を経て、やっと、そして遂に、全て揃ったのである。言うまでもなく、次にやるべきことはその全てをつなぎあわせて全細胞シミュレーションを実現するという原点回帰に他ならない。

ゲノム配列情報のみからタンパクや核酸分子の構造と機能を予測し、それらの間の相互作用ネットワークを定量的に記述し、細胞や個体の機能と一対一に結び付けるという、「遺伝子型からの表現型の予測」の完全な姿の実現は現状では夢の域を出ない（図 4 上段）。しかし、近年の定量マルチオミクス解析技術の進展や本稿で振り返ってきたようなシミュレーション計算技術の発展を外挿すれば、次の事であれば現状の現実的な延長線上で十分に手が届く距離にある。つまり、「よく調べられたモデル生物の遺伝子配列に加わった少数の変異が特定の表現型をどのように変化させるのかを、定量マルチオミクスが与える網羅的な測定情報に沿ったモデリングとシミュレーションにより予測する」（図 4 下段）。これが、E-Cell プロジェクトが次の大きな目標である。この手始めに、海津らが中心となり、最もよく研究されたモデル原核生物である大腸菌の全ゲノム規模モデルの構築を行っており、理研

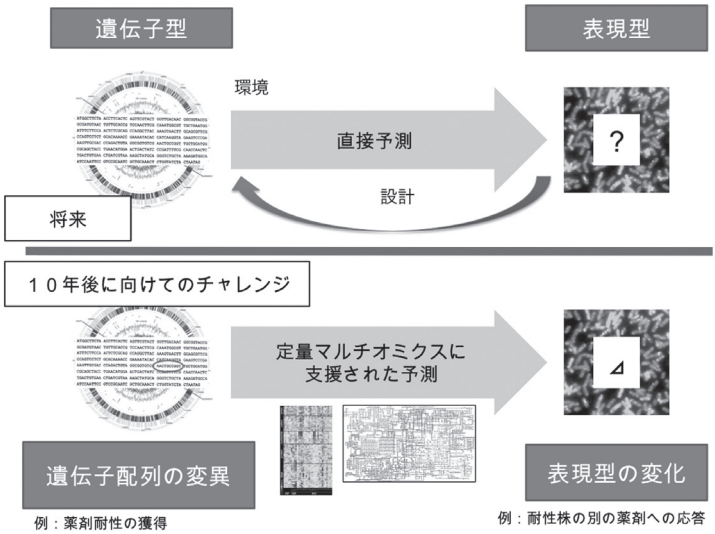

図４　（上段）「遺伝型からの表現型の予測」配列情報のみからタンパクや核酸分子の構造と機能を予測し、それらの間の相互作用ネットワークを定量的に記述し、細胞や個体の機能と一対一に結び付ける。（下段）定量マルチオミクスが与える網羅的な測定情報に沿ったモデリングとシミュレーションにより遺伝子型から表現型を予測する挑戦。

QBiC の古澤力らとも協力し、ある特定の薬剤に耐性を獲得した単剤耐性変異株の他の薬剤に対する増殖速度の変化（交差耐性）などの応答を予測するシミュレーションの実現を目指している。これは我々にとって 2010 年代のうちに実現したい目標である。

5.2　細胞シミュレーションで分子生物学は終わるか

ゲノム配列の遺伝子型だけを見て細胞や個体の表現型を完全に予測出来るようになる事、それは生物学の一つの綺麗な終わらせ方と言える。しかし、それは本当に終わりだろうか？　生物学とは本質的に多様性の学問である。太古の地球に発生した、我々地球上の生物全ての祖先である LUCA（Last Universal Common Ancestor）以来、生態系の豊かさとはそこに生きる生物種の多様性と同義だった。自己複製をかろうじて行うだけの単純な生命から動物と植物が発生し、多細胞生物となり、また海から陸上へと棲息環境を広げてゆく、といった新たな特質の獲得と適応が、現在の我々、ヒトとヒトの文明を取り巻く自然環境へと繋がって来た。生命と「死」の獲得、世代交代と突然変異そして進化という、入力資源を多様性の増大へと向けて出力する生命の特性は、より幅広い環境変化に対する生存可能性（survivability）、またより効率的に生命を営む進化可能性（evolvability）と密接な関係があり。だから、生命が続く限り、立ち現れる生き物のかたちに限りはなく、だから遺伝型と表現型の組み合わせにも限りがない、そのことこそが生物学の本質のように思える。

「遺伝型からの表現型の直接予測を可能にする技術の樹立」は、生命科学の終わりではなく、むしろそこが出発点となり、より本質的な問題、例えば進化や共生、生物多様性などの問題に本格的に取り組む土台となるはずである。であれば、現代の生命科学者にとっての大きな課題は、その出発点にいかに早くたどり着くのかである、と我々は感じている。

最後に、E-Cell プロジェクトが毎年夏に行っているプログラミングイベント E-Cell Sprint の 2014 年の会の集合写真とともに、プロジェクトの今後の新たな展開に思いを新たにしながら本稿の筆を置きたい。

謝辞

プロジェクトの初期メンバーの一人であり現在東京工業大学に在籍する松崎由理氏には本稿のチェックや編集に手助けをいただいた。本稿は SFC ジャーナル向けに書かれたこともあり、E-Cell プロジェクトに貢献したメンバーの名前を出来るだけ多く挙げるよう努力したが、文脈に不自然でない形で全ての人物を網羅することは到底不可能であり、本稿に名前がない人が名前がある人よりもプロジェクトに対する貢献度が低いという事を意味するわけでは全くない事は言うまでもない。

図 5　E-Cell Sprint 2014 の写真

注

1　DNA は A, T, G, C の 4 種類（2 ビット）の塩基が鎖状に連なった巨大分子なので、58 万塩基対の情報量は概ね 150 キロバイトに相当する。
2　The Institute for Atomic and Molecular Physics
3　The Institute for Genomic Research
4　International Conference on Systems Biology
5　Japan Science and Technology Agency
6　Core Research for Evolutional Science and Technology
7　Massachusetts Institute of Technology
8　National Institute for Health
9　The Molecular Sciences Institute
10　Genome Science Center
11　Quantitative Biology Center
12　http://www.e-cell.org/ja/publications/

引用文献 注12

[1]　Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D.,

Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A. 3rd., and Venter, J.C., "The minimal gene complement of Mycoplasma genitalium.." *Science*, 270(5235), 1995, pp.397-403.

[2] Itoh, H., Naito, Y. and Tomita, M., "Simulation of developmental changes in action potentials with ventricular cell models." *Syst. Synth. Biol*., 1(1), 2007, pp.11-23.

[3] Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T.S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J.C. and Hutchison, C.A., 3rd., "E-CELL: software environment for whole-cell simulation." *Bioinformatics*, 15(1), 1999, pp.72-84.

[4] Normile, D., "Building working cells 'in silico'." *Science*, 284(5411), 1999, pp.80-81.

[5] Butler, D., "Computing 2010: from black holes to biology." *Nature*, 402(6761 Suppl), 1999, pp. C67-70.

[6] Takahashi, K. *et al*., "E-Cell 2: multi-platform E-Cell simulation system." *Bioinformatics*, 19(13), 2003, pp.1727-1729.

[7] Hucka, M. *et al*., "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models." *Bioinformatics*, 19(4), 2003, pp.524-531.

[8] Takahashi, K. *et al*., "Computational challenges in cell simulation: a software engineering approach." *IEEE Intelligent Systems*, 17, 2002.

[9] Takahashi, K. *et al*., "A multi-algorithm, multi-timescale method for cell simulation." *Bioinformatics*, 20(4), 2004, pp.538-546.

[10] Rapoport, T.A., Heinrich, R., Jacobasch, G. and Rapoport, S., "A linear steady-state treatment of enzymatic chains. A mathematical model of glycolysis of human erythrocytes." *European Journal of Biochemistry*, 42(1), 1974, pp.107–120.

[11] Yachie-Kinoshita, A., Nishino, T., Shimo, H., Suematsu, M. and Tomita, M., "A metabolic model of human erythrocytes: practical application of the E-Cell Simulation Environment." *J. Biomed. Biotechnol*., 2010, p.642420.

[12] Kinoshita, A., Nakayama, Y., Kitayama, T. and Tomita, M. "Simulation study of methemoglobin reduction in erythrocytes. Differential contributions of two pathways to tolerance to oxidative stress." *FEBS J*., 274(6), 2007, pp.1449-1458.

[13] Shimo, H. *et al*., "Particle simulation of oxidation induced band 3 clustering in human erythrocytes." *PLoS Comput. Biol*., to appear.

[14] Shimo, H., Nishino, T. and Tomita, M., "Predicting the Kinetic Properties Associated with Redox Imbalance after Oxidative Crisis in G6PD-Deficient Erythrocytes: A Simulation Study." *Adv. Hematol*., 2011, p.398945.

[15] Nishino, T., Yachie-Kinoshita, A., Hirayama, A., Soga, T., Suematsu, M. and Tomita, M., "In silico modeling and metabolome analysis of long-stored erythrocytes to improve blood storage methods." *J. Biotechnol*., 144(3), 2009, pp.212-223.

[16] Nishino, T., Yachie-Kinoshita, A., Hirayama, A., Soga, T., Suematsu M. and Tomita, M., "Dynamic simulation and metabolome analysis of long-term erythrocyte storage in adenine-guanosine solution." *PLoS One*, 8(8), 2013, p.e71060.

[17] Yugi, K. and Tomita, M., "A general computational model of mitochondrial metabolism in a whole organelle scale." *Bioinformatics*, 20(11), 2004, pp.1795-1796.

[18] Yugi, K., Nakayama, Y., Kinoshita, A. and Tomita, M., "Hybrid dynamic/static method for large-scale simulation of metabolism." *Theor. Biol. Med. Model*., 2, 2005, p.42.
[19] Miyoshi, F. *et al*., "A mathematical model for the Kai-protein-based chemical oscillator and clock gene expression rhythms in cyanobacteria." *J. Biol. Rhythms*, 22(1), 2007, pp.69-80.
[20] Ogawa, Y. *et al*., "Comparative study of circadian oscillatory network models of Drosophila." *Artif. Life*, 14(1), 2008, pp.29-48.
[21] Ohno, H., Naito, Y., Nakajima, H. and Tomita, M., "Construction of a biological tissue model based on a single-cell model: a computer simulation of metabolic heterogeneity in the liver lobule." *Artif. Life*, 14(1), 2008, pp.3-28.
[22] Matsuoka, S., Sarai, N., Kuratomi, S., Ono, K. and Noma, A., "Role of individual ionic current systems in ventricular cells hypothesized by a model study." *Jpn. J. Physiol*., 53(2), 2003, pp.105-123.
[23] Itoh, H., Naito, Y. and Tomita, M., "Simulation of developmental changes in action potentials with ventricular cell models." *Syst. Synth. Biol*., 1(1), 2007, pp.11-23.
[24] Okubo, C., Sano, H.I., Naito, Y. and Tomita, M., "Contribution of quantitative changes in individual ionic current systems to the embryonic development of ventricular myocytes: a simulation study." *J. Physiol. Sci*., 63(5), 2013, pp.355-367.
[25] Takahashi, K. *et al*., "Space in systems biology of signaling pathways--towards intracellular molecular crowding in silico." *FEBS Lett*., 579(8), 2005, pp.1783-1788.
[26] Mugler, A. *et al*., "Membrane clustering and the role of rebinding in biochemical signaling." *Biophys. J*., 102(5), 2012, pp.1069-1078.
[27] Takahashi, K. *et al*., "Spatio-temporal correlations can drastically change the response of a MAPK pathway." *Proc. Natl. Acad. Sci. U S A*., 107(6), 2010, pp.2473-2478.
[28] Arjunan, S.N. *et al*., "A new multicompartmental reaction-diffusion modeling method links transient membrane attachment of E. coli MinE to E-ring formation." *Syst. Synth. Biol*., 4(1), 2010, pp.35-53.
[29] Tomita, M., "Whole-cell simulation: a grand challenge of the 21st century." *Trends Biotechnol*., 19(6), 2010, pp.205-210.

〔受付日　2015. 2. 26〕