

生命動態研究をデータベースで 駆動する

大浪修一

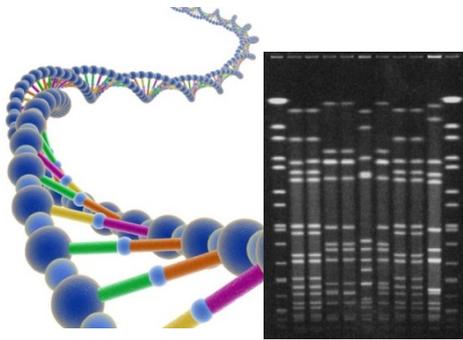
理化学研究所生命システム研究センター



©2013 大浪 修一 (理化学研究所) licensed under CC表示2.1日本

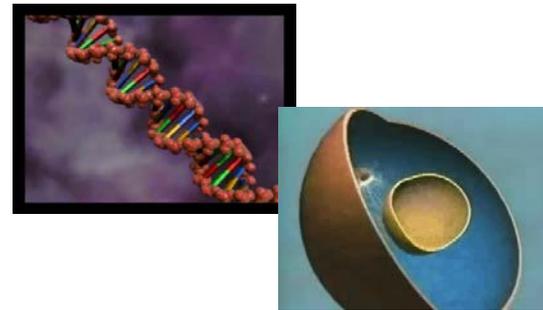
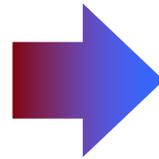
生命動態システム科学

- 生命を動的システムとして理解し、操作する生命科学
- 新たな生命科学の潮流として21世紀の科学全体への大きな影響が期待される



分子生物学

遺伝子の制御



生命動態システム科学

細胞の制御

複雑な生命現象の動態を時・空間を有する先端定量計測と高精密モデリングをもとに、*in silico*と*in vitro*で再構成

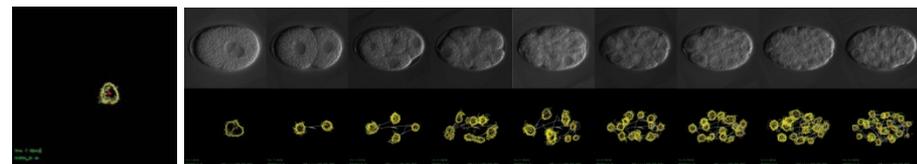
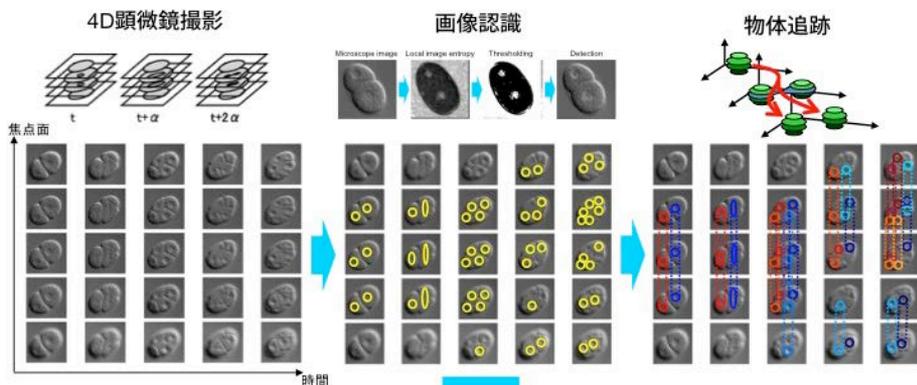
生命動態システム科学のデータ

時空間情報を数値として含む新しい様式の生命科学の研究データ

具体例

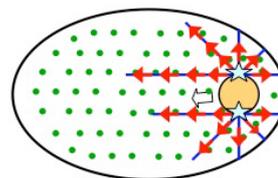
- 様々な摂動条件下の生命動態の時空間定量計測データ
- 様々なパラメータ下の生命現象の時空間動態シミュレーション結果

線虫胚の細胞核分裂動態計測データ



細胞核動態のシミュレーション結果

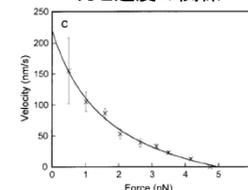
微小管の長さ依存の引力



$$f_{PULL} = F_{motor} \cdot D \cdot L$$

F_{motor} : force per motor
 D : density of motors
 L : length of the microtubule

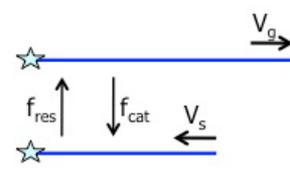
分子モーターの力と速度の関係



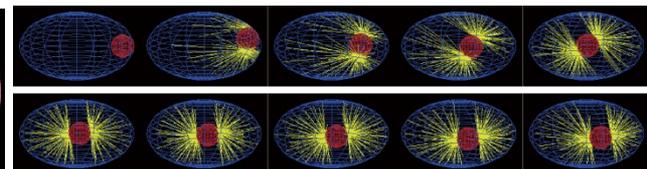
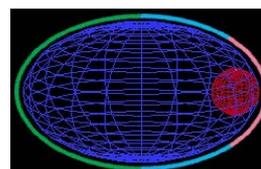
$$F_{motor} = F_{stall} \cdot \{1 - (v/V_{max})\}$$

F_{stall} : stall force of motor
 v : velocity of motor
 V_{max} : maximum velocity of motor

微小管の動的不安定性

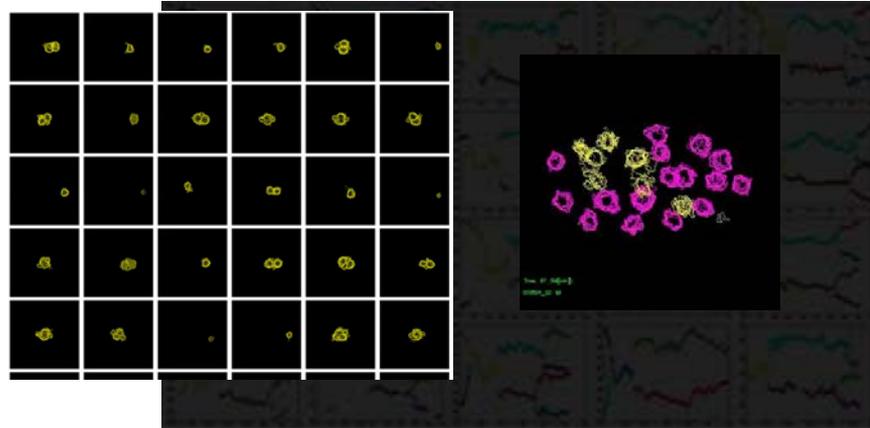


V_g : Growth velocity
 V_s : Shrinkage velocity
 f_{cat} : Catastrophe frequency
 f_{res} : Rescue frequency

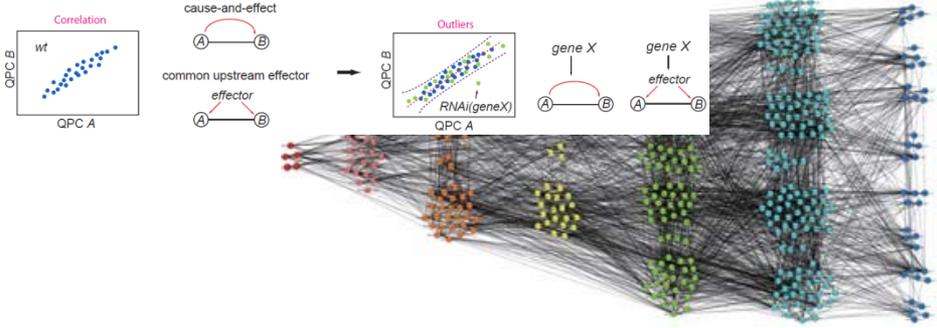


生命動態システム科学計測データの活用

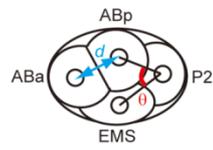
新しいバイオインフォマティクス領域を創成



可視化



情報・数理解析

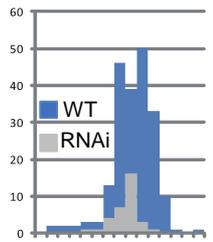


NucDis_ABa-ABp

$$d = \sqrt{(\epsilon_{ABa} - x_{ABp})^2 + (\epsilon_{ABa} - y_{ABp})^2 + (\epsilon_{ABa} - z_{ABp})^2}$$

NucAng_ABp-P2-EMS

$$\theta = \arccos \left(\frac{(\epsilon_{ABp} - x_{P2})(\epsilon_{ABp} - x_{EMS}) + (\epsilon_{ABp} - y_{P2})(\epsilon_{ABp} - y_{EMS}) + (\epsilon_{ABp} - z_{P2})(\epsilon_{ABp} - z_{EMS})}{\sqrt{(\epsilon_{ABp} - x_{P2})^2 + (\epsilon_{ABp} - y_{P2})^2 + (\epsilon_{ABp} - z_{P2})^2} \sqrt{(\epsilon_{ABp} - x_{EMS})^2 + (\epsilon_{ABp} - y_{EMS})^2 + (\epsilon_{ABp} - z_{EMS})^2}} \right)$$



計算表現型解析

生命動態システム科学のDB統合化

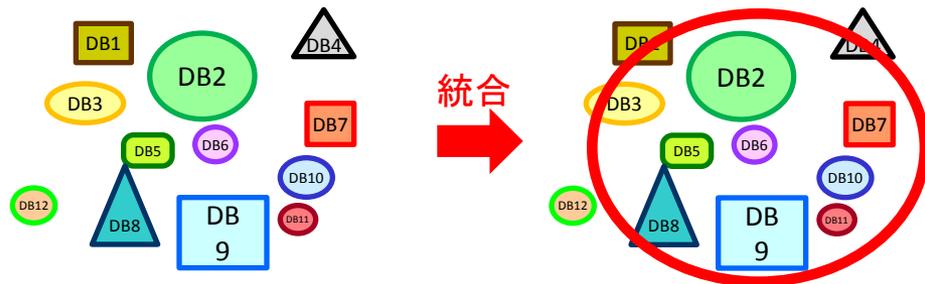
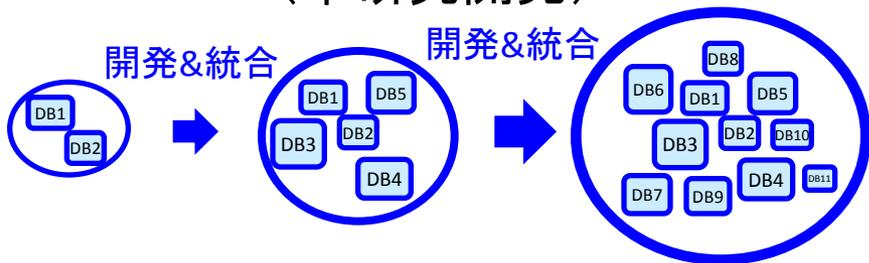
- 現在公開・運用されているデータベースは少ない
- 分野の成長に伴い、今後の急速な増加が見込まれる



今なら、データベースの乱立に先行して統合化のスキームを構築し、当分野の全てのデータベースが統合化スキームの下で開発される体制の構築が可能

生命動態システム科学のDB統合化
(本研究開発)

通常のDB統合化



→ 統合化スキームの先行構築によりDB統合化の諸困難を解決!!

平成24年度 ライフサイエンスデータベース統合推進事業
統合化推進プログラム研究開発課題

生命動態システム科学のデータベースの統合化

理化学研究所生命システム研究センター

大浪 修一

本研究開発の目標

- 我が国の生命動態システム科学分野の全てのデータベースを統合する体制と仕組みの構築
- 研究開発終了時における、我が国の生命動態システム科学分野の全てのデータベースの統合
- 生命動態システム科学分野の全世界のデータベースの統合

データベース統合の体制と仕組みの構築

生命動態システム科学のコミュニティの合意の上で、当分野のデータベース統合の体制、仕組み、スケジュール等を決定する。

プロジェクトグループ

- 大規模なデータ収集が見込まれる研究者、DB統合に関心を持つ研究者等
- プロジェクトの進捗の確認、助言等
- 年1回ペースで打合せ



ワーキングチーム

- プロジェクトチームから選抜(5名程度)
- プロジェクトの具体的な決定等
- 年2回ペースで打合せ

生命動態システム科学 All Japanコミュニティ

- 理研QBiC/大阪大
- さきがけ/CREST
- 新学術研究領域

など



データベースの統合

コミュニティで合意した体制、仕組み、スケジュールの下で、生命動態システム科学のわが国の全てのDBを開発と平行して統合する

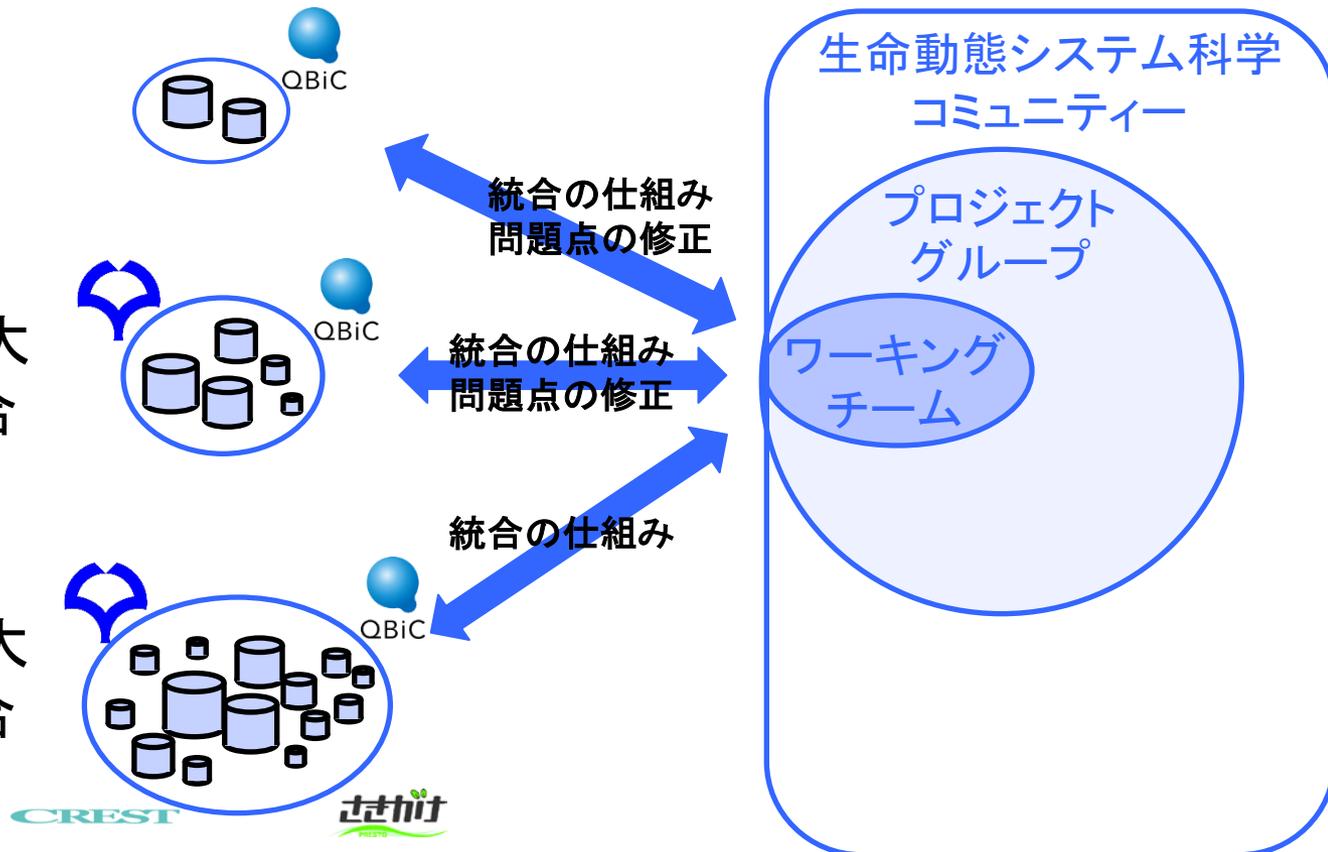
1. 大浪研究室のDBの統合



2. 理研QBiC/大阪大の他のDBの統合



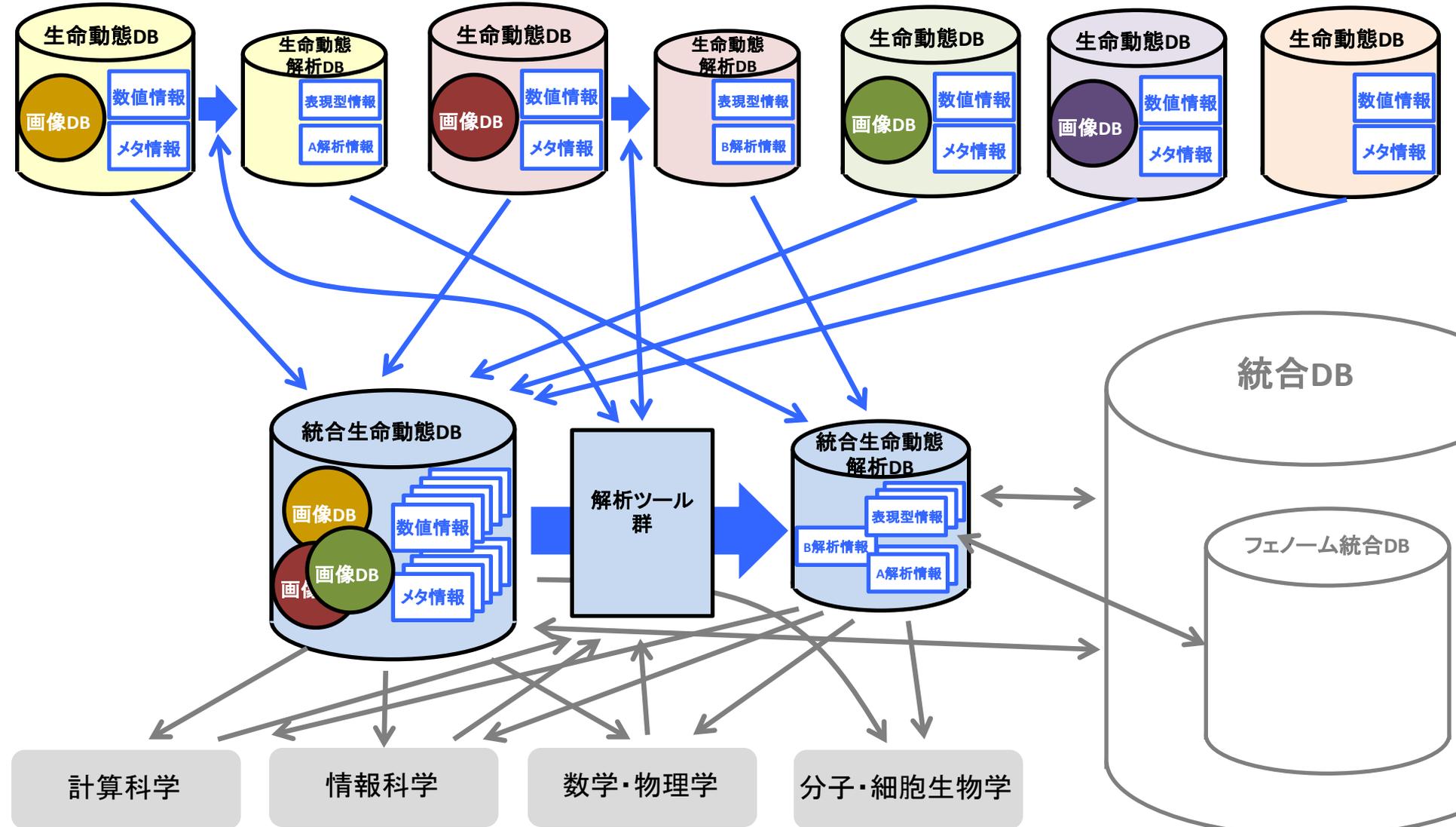
3. 理研QBiC/大阪大以外のDBの統合



データベース統合化の具体的なイメージ

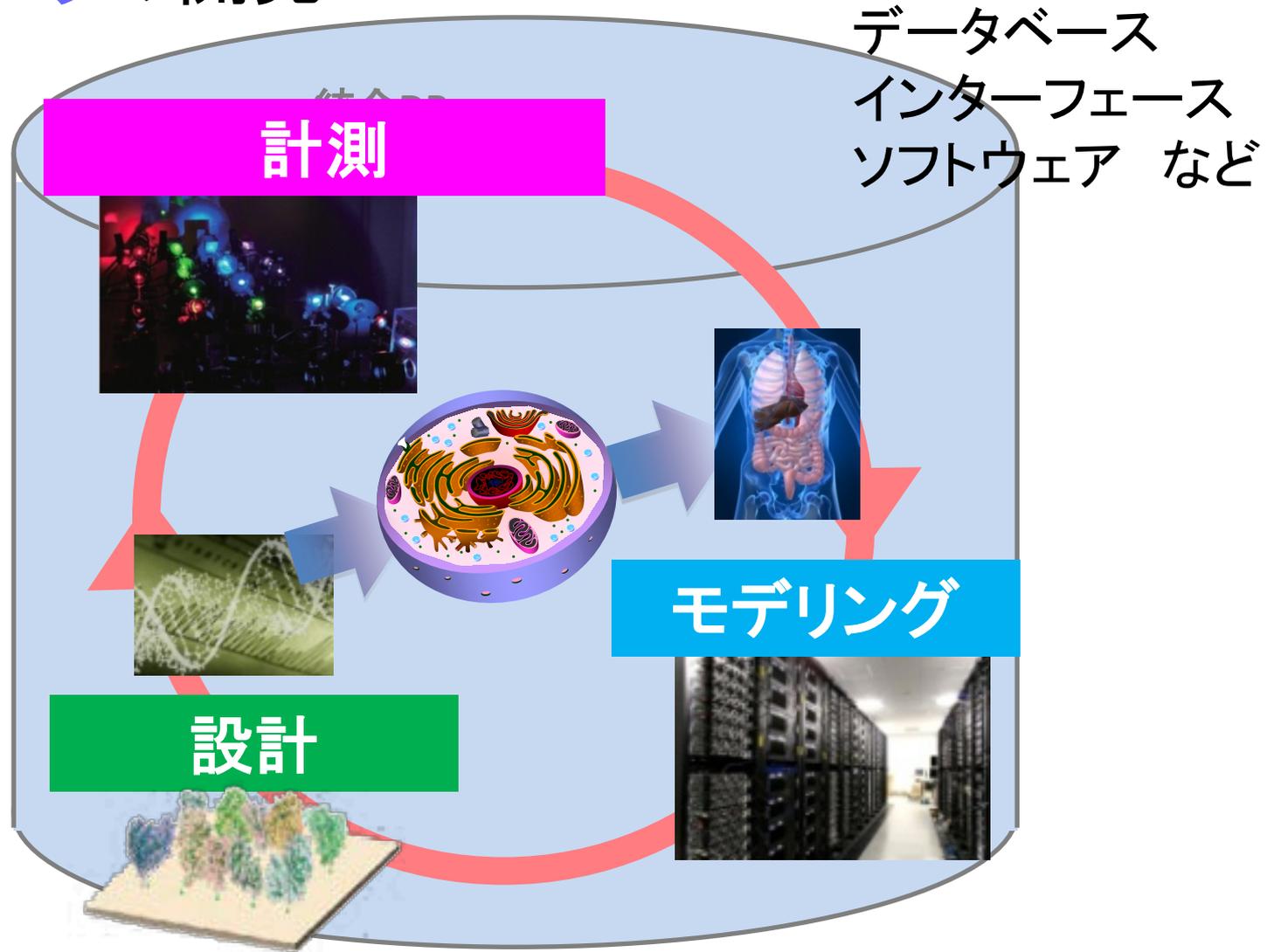
データのフォーマットを統一し統合化

→解析ツールを共有できる！

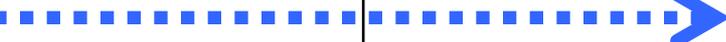


チャレンジ

計測、モデリング、設計の有機的な連携を加速する
情報インフラの開発



研究開発のスケジュール

研究項目	H24年度	H25年度	H26年度
1. DB統合の体制と仕組みの構築			
2. 理研QBiC/大阪大内のDB統合			
3. 理研QBiC/大阪大外のDB統合			

H24年度の進捗状況

- 研究項目1 DB統合の体制と仕組みの構築
 - プロジェクトグループの構築(達成)
 - 統合データフォーマットの構築(BDD-XML)
 - 統合データベースサーバの構築(設計まで)
- 研究項目2 理研QBiC/大阪大内のDB統合
 - 大津研究室が保有する2つの大規模データの統合(達成)

H25年度の目標、当初計画

- 研究項目1 DB統合の体制と仕組みの構築
 - － 理研QBiC内のデータの調査
 - － 統合データフォーマットの構築
 - － 統合データベースの構築
- 研究項目2 理研QBiC/大阪大内のDB統合
 - － 理研QBiC内のデータベースの統合
- 研究項目3 理研QBiC/大阪大外のDB統合
 - － 理研QBiC外のデータベースの統合の開始

H25年度の進捗状況

- 研究項目1 DB統合の体制と仕組みの構築
 - 理研QBiC内のデータの調査
 - 統合データフォーマットの構築
 - 統合データベースの構築
- 研究項目2 理研QBiC/大阪大内のDB統合
 - 理研QBiC内のデータベースの統合
- 研究項目3 理研QBiC/大阪大外のDB統合
 - 理研QBiC外のデータベースの統合の開始

理研QBic内のデータの調査

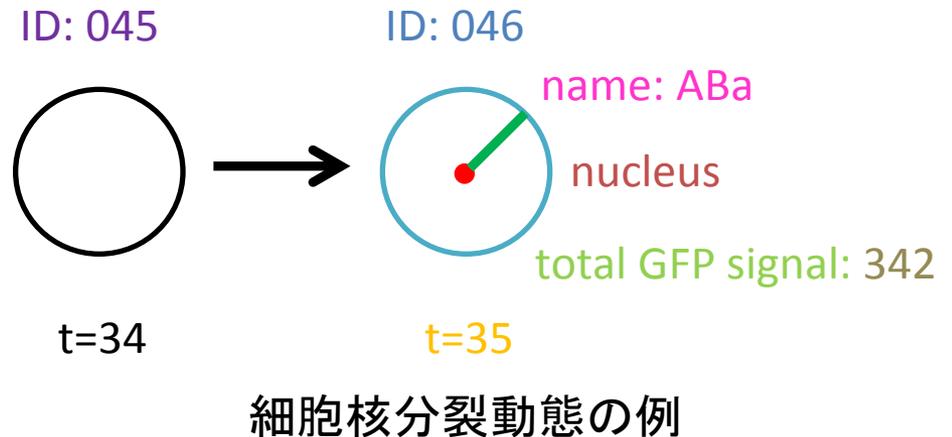
生物種	対象	方法	研究室
Hela細胞	一分子	計測	上田昌宏
Hela細胞	FCS	計測	神 隆
Hela細胞	一分子	モデル	高橋恒一
肝臓細胞	MS	計測	升島 勉
マウス	一分子	計測	岡田康志
マウス	発現	計測	上田泰己
ゼブラフィッシュ	核	計測	岡田康志
線虫	核	計測	大浪修一
線虫	核	計測	大浪修一
線虫	核	モデル	大浪修一
粘菌	一分子	計測	神 隆
出芽酵母	発現	計測	谷口雄一
大腸菌	生育	計測	古澤 力
大腸菌	一分子	モデル	高橋恒一

理研QBic内のデータの調査

生物種	対象	方法	研究室
Hela細胞	一分子	計測	上田昌宏
Hela細胞	FCS	計測	神 隆
Hela細胞	一分子	モデル	高橋恒一
肝臓細胞	MS	計測	升島 勉
マウス	一分子	計測	岡田康志
マウス	発現	計測	上田泰己
ゼブラフィッシュ	核	計測	岡田康志
線虫	核	計測	大浪修一
線虫	核	計測	大浪修一
線虫	核	モデル	大浪修一
粘菌	一分子	計測	神 隆
出芽酵母	発現	計測	谷口雄一
大腸菌	生育	計測	古澤 力
大腸菌	一分子	モデル	高橋恒一

BDML (ver. 0.15) データ部の構造

- data(1)
 - scaleUnit(1)
 - object(1..n)
 - feature(0..n)
 - component(1..n)
 - componentID(1)
 - componentName(0..1)
 - time(1)
 - groupID(0..n)
 - prevID(0..n)
 - measurement(1..n)
 - objectRef(1)
 - point, line, face, circle, sphere(1:choice)
 - coords(1..n), radius(0..1)
 - property(0..n)
 - featureRef(1), featureValue(1)



統合データベースSSDBを公開 (2013年9月2日)

SSBD Database beta edition

Sign in to [SSBD](#)

Browse through categories:

[Home](#)

[Resources](#)

[Manuals](#)

[Publicatons](#)

[News](#)

[Download](#)

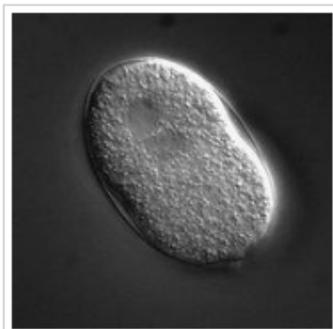
Search Services:

[Advanced](#) [Help](#)

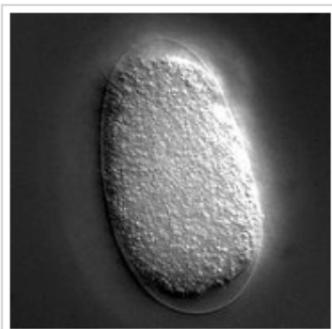
Introduction of SSBD

Systems Science of Biological Dynamics (SSBD) database provides a rich set of resources for analyzing quantitative biological data, such as single-molecule, cell, and gene expression nuclei. Quantitative biological data are collected from a variety of species, sources and methods. These include data obtained from both experiment and computational simulation. These quantitative numerical data are represented in a new **Biological Dynamics Markup Language (BDML)**. The new data format allows users to exchange, store, compare and analyze data through the SSBD database. Users can download quantitative biological dynamics data directly in BDML format from the SSBD database. The system utilizes OMERO server to manage image data and experimental conditions. A range of software tools and applications for visualizing and analyzing quantitative biological dynamical data are being developed through a set of SSBD APIs.

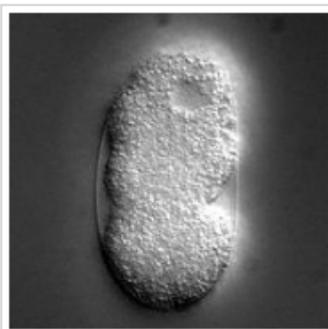
Sample Datasets



Nuclear division dynamics in *C. elegans* wild-type embryo, IDENTIFIER: wt(N2)030116_02



Nuclear division dynamics in *C. elegans* RNAi embryo, IDENTIFIER: F58B6.3(gPCR)080610_02



Nuclear division dynamics in *C. elegans* RNAi embryo, IDENTIFIER: C02F5.1(gPCR)061219_02

Menu

BDML and PDPML schemas can be found [here](#).

OMERO web: Some images can be viewed on [OMERO web](#). If you have problem viewing the images on the website, please click on the drop-down arrow on the right of 'public data' on the bar above the data tree, select 'Public' group and 'public data' to view the images (click [here](#) for more details).

News and Events

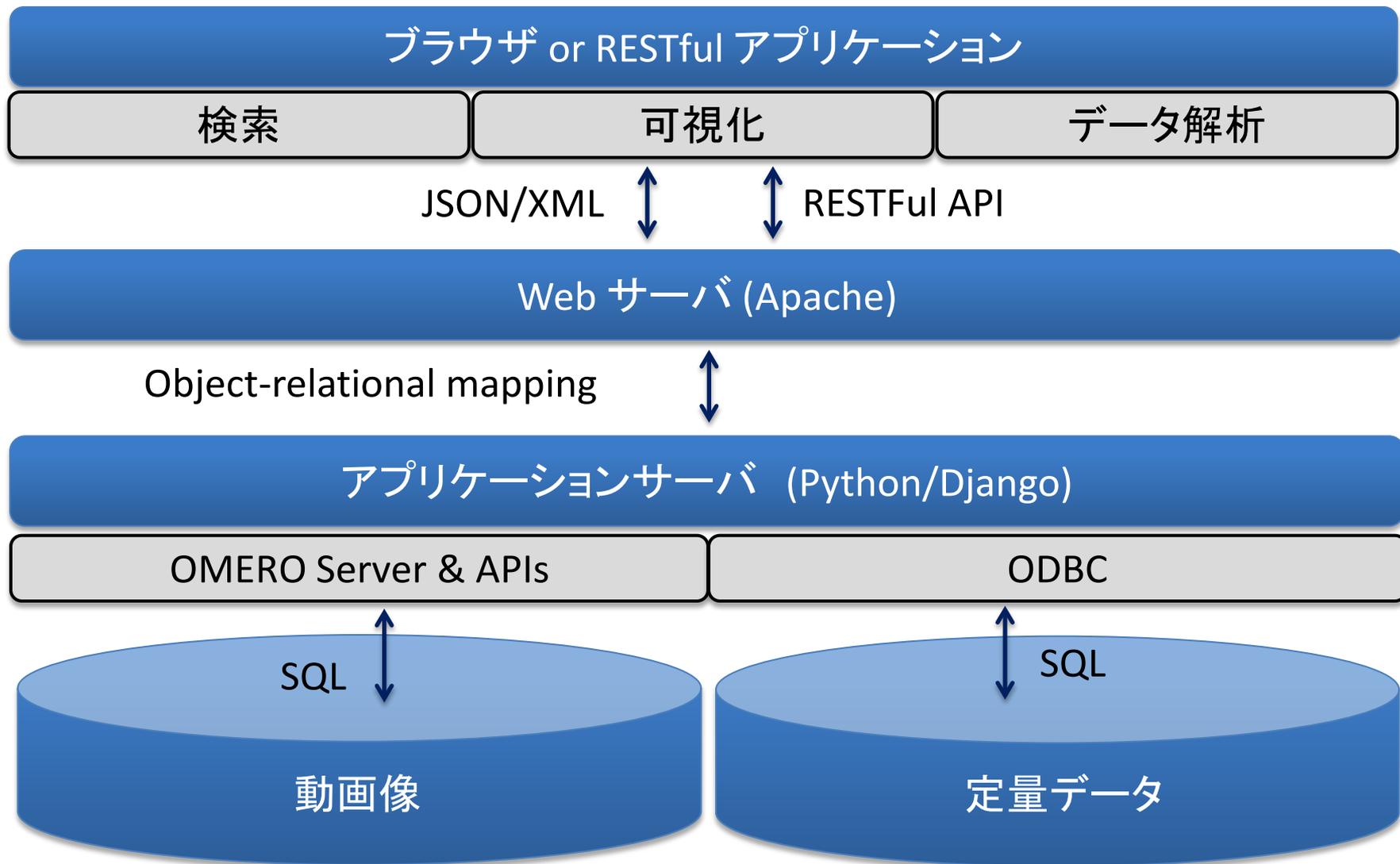
Sept. 2, 2013: SSBD database is now open for public access.

[Older news ...](#)

Links

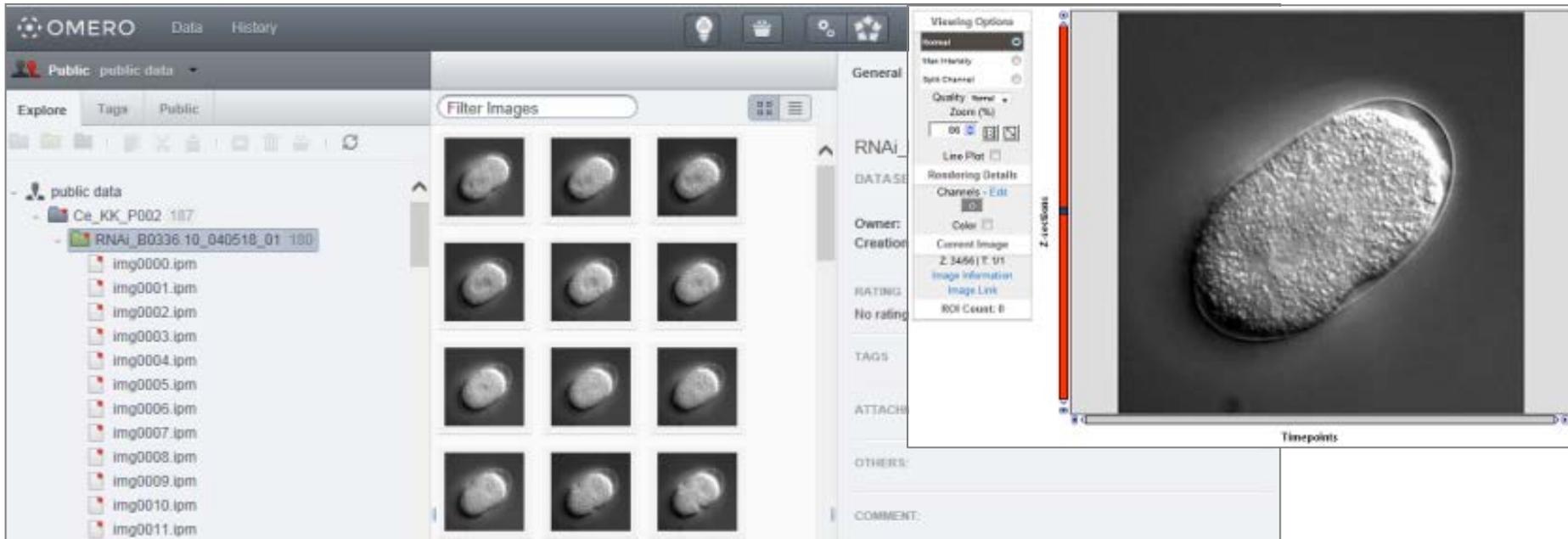
[WDDD](#), [OME](#)

データベースのシステム構成

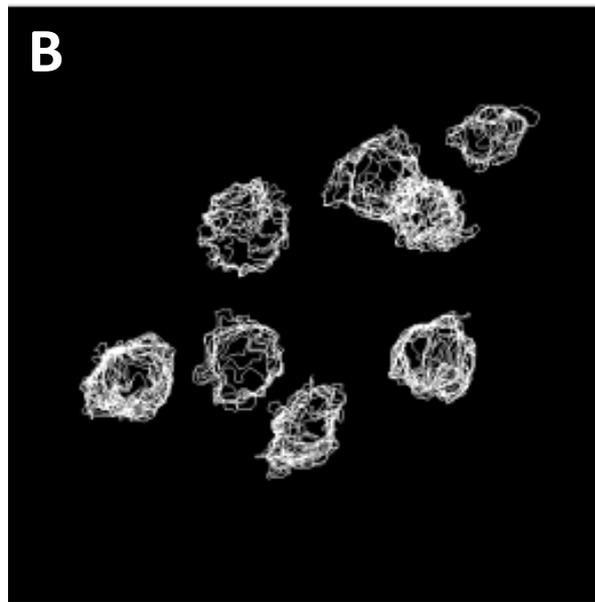
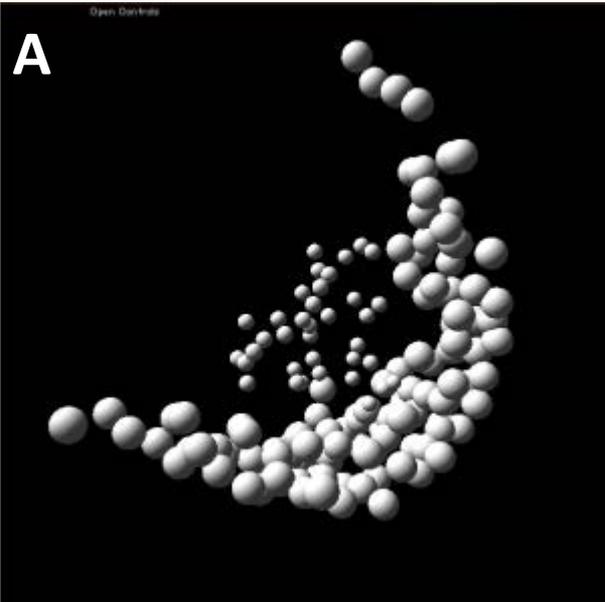


OMEROシステムによる動画像の管理

- 顕微鏡から得られるテラバイト単位の動画像の管理を目的に開発されたフリーのソフトウェア
 - 127種類の動画像のフォーマットに対応
 - ImageJ, MATLAB, Java, C++, Python によるクライアント・アプリケーションの開発が可能



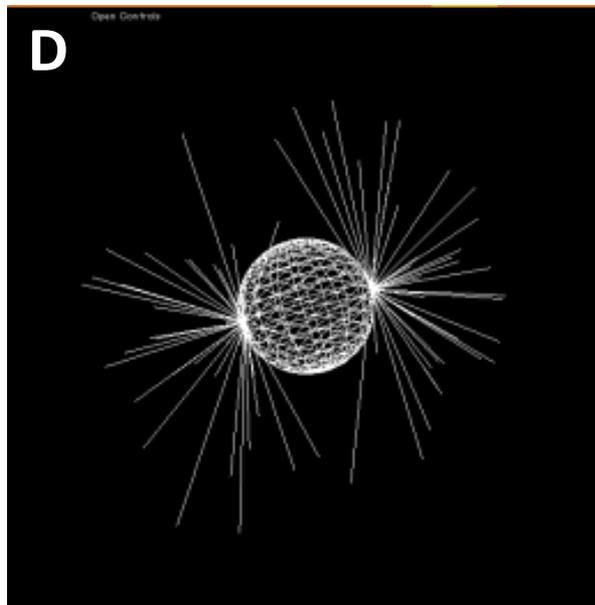
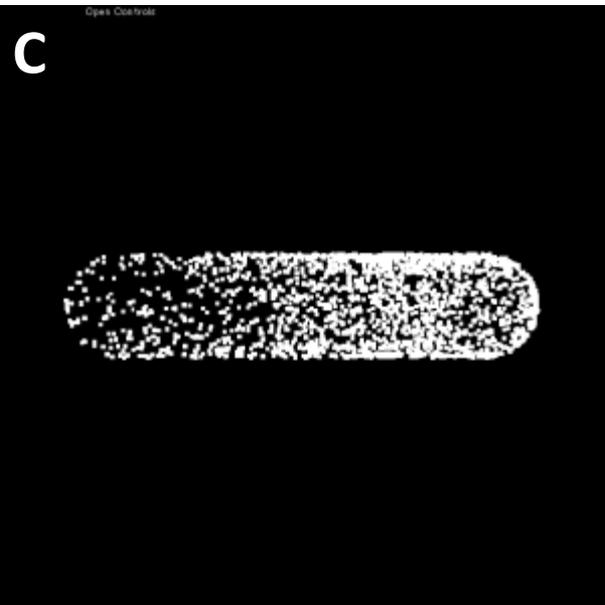
BDMLデータの可視化



(A) ショウジョウバエの
核分裂動態
(Supatto et al. 2009)

(B) 線虫の核分裂動態
(Kyoda et al. 2013)

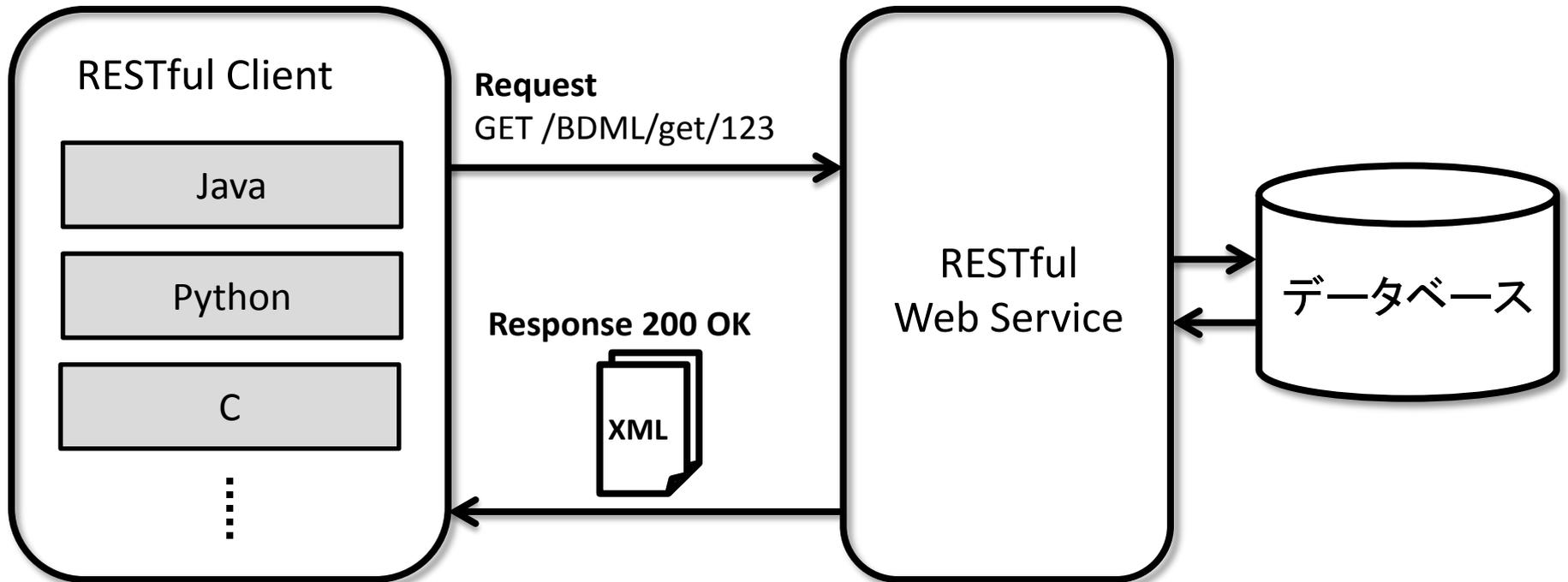
(C) 大腸菌の一分子
シミュレーション結果
(Arjunan & Tomita 2010)



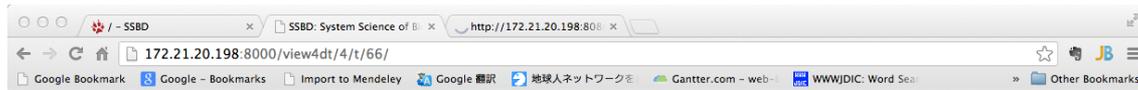
(D) 線虫の核と紡錘体のシ
ミュレーション結果
(Kimura & Onami 2005)

RESTful API を構築

- Representational State Transfer
- HTTPのURI (Uniform Resource Identifier) でデータの読み取り・作成・更新・削除 ができる仕組み
 - 第三者がクライアントのアプリケーションをプログラミング言語やプラットフォームに制限されることなく開発することができる



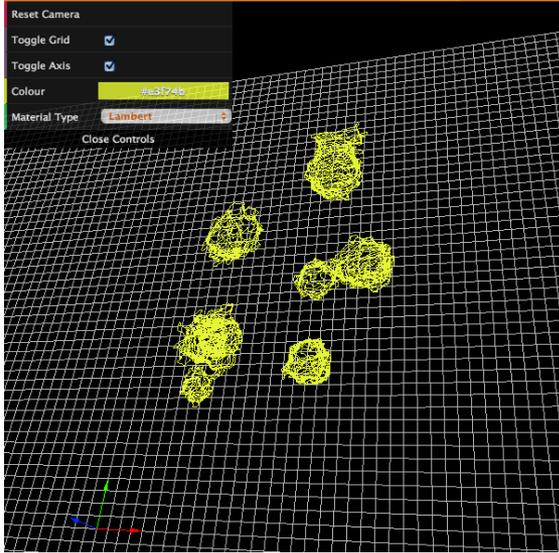
ブラウザ上でBDMLデータを可視化



SSBD Database beta edition

Sign in to SSBD

Browse through categories: Home Resources Manuals Publications News Download



BDML ID: 0ac550a0-90ec-4779-876f-6f1992a4f870
Title BDML file for quantitative information about nuclear division dynamics of wild-type embryo
Version Version 1, updated on May 18, 2012, CC BY-NC-SA license
Description quantitative information about nuclear division dynamics in wild-type embryo
Organism: C. elegans
Datatype nuclear division dynam
Identifier wt(N2)030116_02
Based on Measurement
Contributor Kyoda, K., Furukawa, M
PubMed ID [23172286](#)
Contact Information Shuichi Onami, RIKEN, Center, Laboratory for C
Summary of methods See details in Kyoda et Res 41, D732-D737.
Data Scale (XYZ, T) 0.105 x 0.105 x 0.5 (mic

Time point: 66

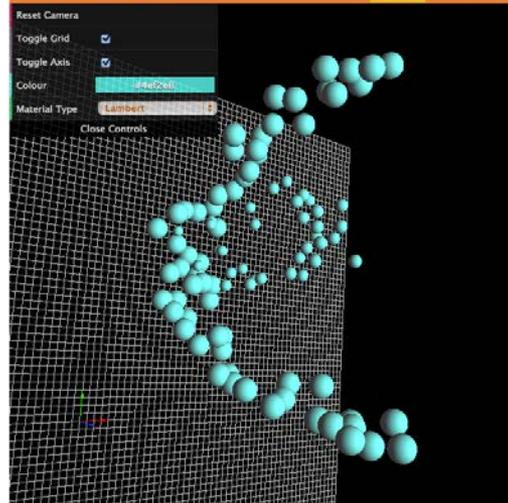
Copyright SSBD Project, Contact



SSBD Database beta edition

Sign in to SSBD

Browse through categories: Home Resources Manuals Publications News Download



BDML ID: 649141f0-168e-11e3-8512-1a9bbf263d8a
Title BDML file for quantitative information about cell dynamics of D. melanogaster wild-type embryo
Version Version 1, updated on Sept. 6, 2013, CC BY-NC-SA license
Description quantitative information about cell dynamics in D. melanogaster wild-type embryo
Organism: D. melanogaster
Datatype cell dynamics
Identifier Dme_Cell_SW
Based on Measurement
Contributor Supatto, W., MacMahon, A., Fraser, S. E. and Stathopoulos, A.
PubMed ID [19745822](#)
Contact Information Willy Supatto, CNRS UMR 7592, Universite Paris 7, Institut Jacques Monod, Laboratory for Optics and Biosciences
Summary of methods See details in Supatto et al. (2009) Nature Protocols 4(10), 1397-1412.
Data Scale (XYZ, T) 0.5 x 0.5 x 1.0 (micrometer), 50.0 (second)

Time point: 0

Copyright SSBD Project, Contact

H25年度の進捗状況

- 研究項目1 DB統合の体制と仕組みの構築
 - 理研QBiC内のデータの調査
 - 統合データフォーマットの構築
 - 統合データベースの構築
- 研究項目2 理研QBiC/大阪大内のDB統合
 - 理研QBiC内のデータベースの統合
- 研究項目3 理研QBiC/大阪大外のDB統合
 - 理研QBiC外のデータベースの統合の開始

理研QBiC内のデータの格納

生物種	対象	方法	研究室	BDML	画像
Hela細胞	一分子	計測	上田昌宏		
Hela細胞	一分子	モデル	高橋恒一		
マウス	一分子	計測	岡田康志		
マウス	発現	計測	上田泰己		
ゼブラフィッシュ	核	計測	岡田康志		
線虫	核	計測	大浪修一	186	33480
線虫	核	計測	大浪修一	256	256
線虫	核	モデル	大浪修一	100	
粘菌	一分子	計測	神 隆	1	
出芽酵母	発現	計測	谷口雄一		
大腸菌	生育	計測	古澤 力		
大腸菌	一分子	モデル	高橋恒一	1	

H25年度の進捗状況

- 研究項目1 DB統合の体制と仕組みの構築
 - 理研QBiC内のデータの調査
 - 統合データフォーマットの構築
 - 統合データベースの構築
- 研究項目2 理研QBiC/大阪大内のDB統合
 - 理研QBiC内のデータベースの統合
- 研究項目3 理研QBiC/大阪大外のDB統合
 - 理研QBiC外のデータベースの統合の開始

理研QBic外のデータの格納

生物種	対象	方法	論文	BDML	画像
Hela細胞	細胞	計測	Neumann et al. 2010		
マウス	核	計測	Kurotaki et al. 2007	1	80
マウス	核	計測	Bashar et al. 2012	1	
ゼブラフィッシュ	核	計測	Keller et al. 2008	7	
ショウジョウバエ	核	計測	Keller et al. 2010	2	
ショウジョウバエ	細胞	計測	Supatto et al. 2009	1	
線虫	核	計測	Bao et al. 2005	2	
線虫	核	計測	Sönnichsen et al. 2005	979	
線虫	核	計測	Gonczy et al. 2000	1	
線虫	発現	計測	Sarov et al. 2012	273	
線虫	行動	計測	Cronin et al. 2005	11	

SSBDから公開したBDMLデータの一覧

理研QBiC内

生物種	対象	方法	論文	BDML	画像
線虫	核	計測	Kyoda et al. 2013	186	33480
線虫	核	モデル	Kimura & Onami 2005	100	
大腸菌	一分子	モデル	Arjunan & Tomita 2010	1	

理研QBiC外

生物種	対象	方法	論文	BDML	画像
マウス	核	計測	Kurotaki et al. 2007	1	80
ゼブラフィッシュ	核	計測	Keller et al. 2008	7	
ショウジョウバエ	核	計測	Keller et al. 2010	2	
ショウジョウバエ	細胞	計測	Supatto et al. 2009	1	
線虫	核	計測	Bao et al. 2005	2	
線虫	行動	計測	Cronin et al. 2005	11	

H25年度の進捗状況のまとめ

- 研究項目1 DB統合の体制と仕組みの構築
 - 理研QBiC内で産出されるデータを調査した
 - 理研QBiC内のデータに対応した統合データフォーマットを開発した(BDML ver. 0.15)
 - 生命動態システム科学の統合DBを公開した(SSBD DB)
 - OMEROによる統括的な画像の管理の実現
 - データ化可視化ソフトの導入
 - RESTful APIの構築
- 研究項目2 理研QBiC/大阪大内のDB統合
 - 理研QBiC内のデータをSSBD DBへ格納した
- 研究項目3 理研QBiC/大阪大外のDB統合
 - 理研QBiC外のデータのSSBD DBへの格納を開始した

研究開発のスケジュール

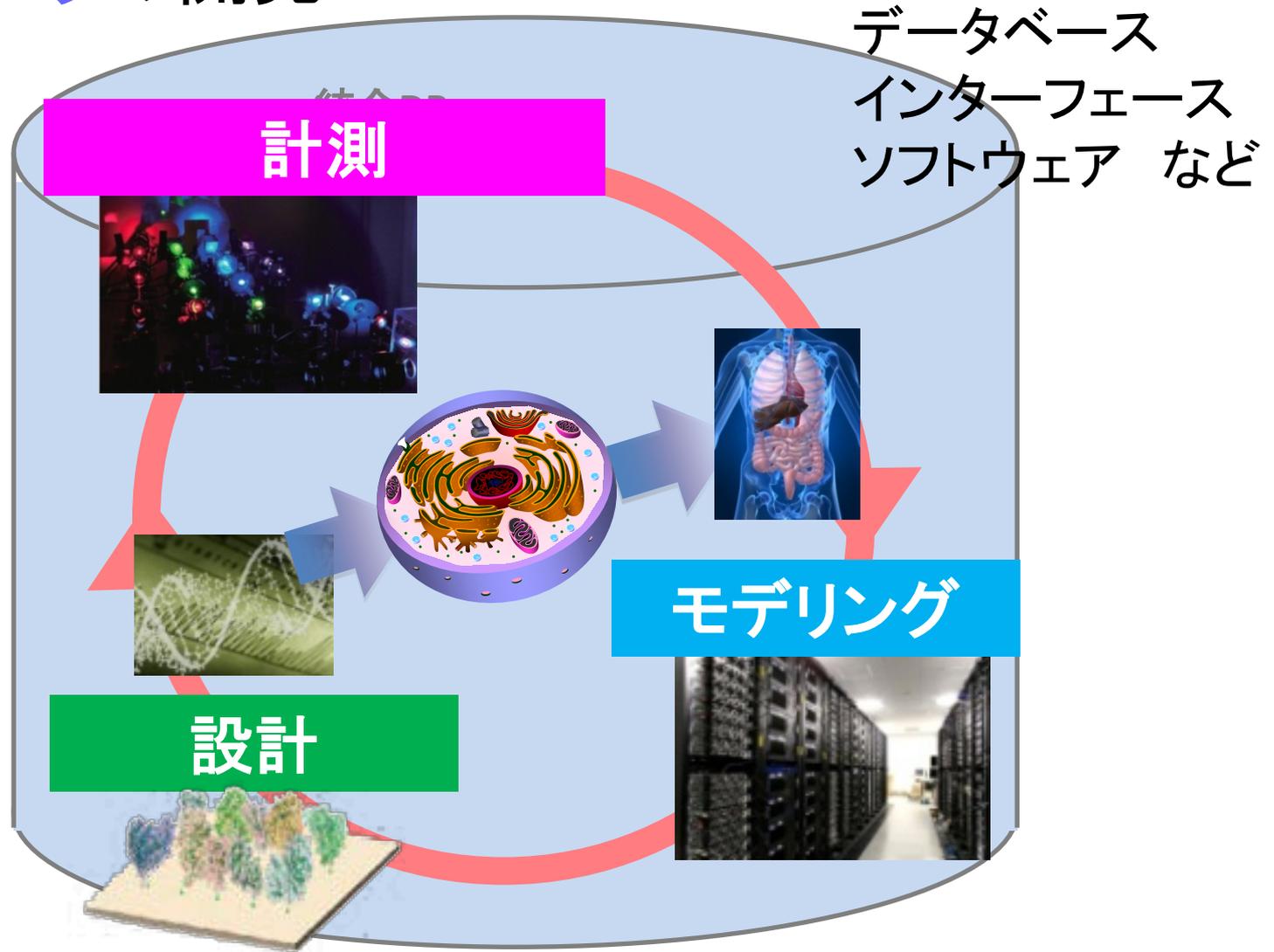
研究項目	H24年度	H25年度	H26年度
1. DB統合の体制と仕組みの構築			
2. 理研QBiC/大阪大内のDB統合			
3. 理研QBiC/大阪大外のDB統合			

H25年度後半以降の計画

- 研究項目1 DB統合の体制と仕組みの構築
 - SSBD DBに未格納のデータを統合DBに格納する仕組みの開発
 - 理研QBiC内で産出されるデータ
 - 国内外のデータ
 - SSBD DBのデータ可視化ソフトの改良
 - 時間軸方向の表示更新の円滑化
 - 追跡情報の選択的な表示を可能に
 - APIの改良、充実
- 研究項目2 理研QBiC/大阪大内のDB統合
 - SSBD DBに未格納の理研QBiC内のデータを統合DBに格納
- 研究項目3 理研QBiC/大阪大外のDB統合
 - 国内外のデータを SSBD DBへ登録

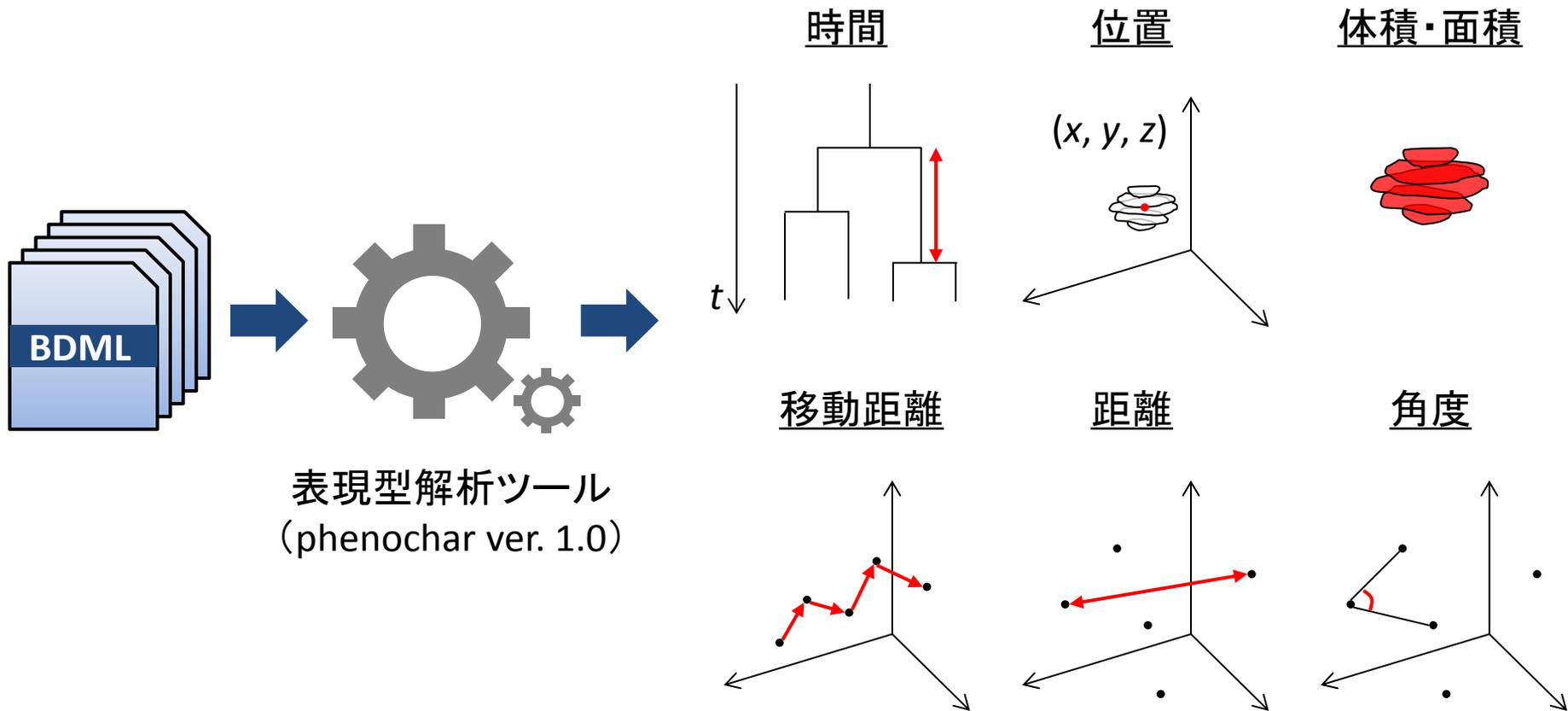
チャレンジ

計測、モデリング、設計の有機的な連携を加速する
情報インフラの開発



表現型解析ツールの開発

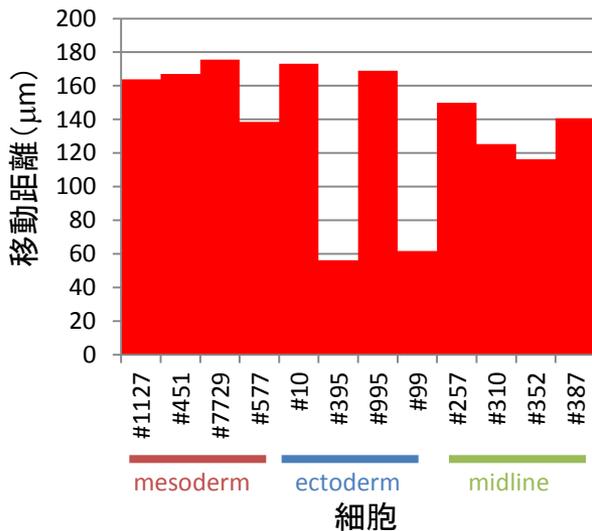
- BDML形式のデータから表現型特徴を計算



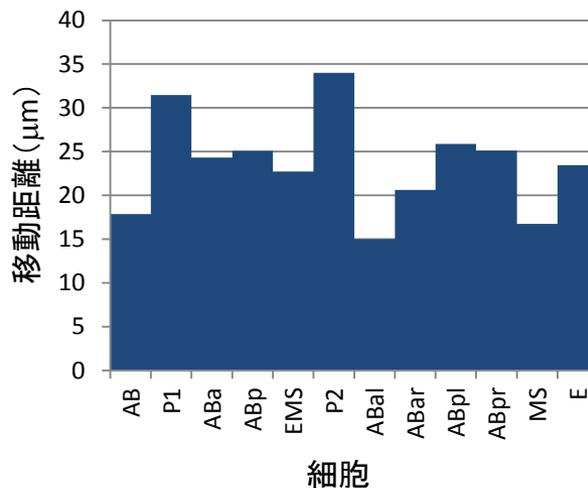
表現型解析ツールの使用例

- 細胞(核)の移動距離について

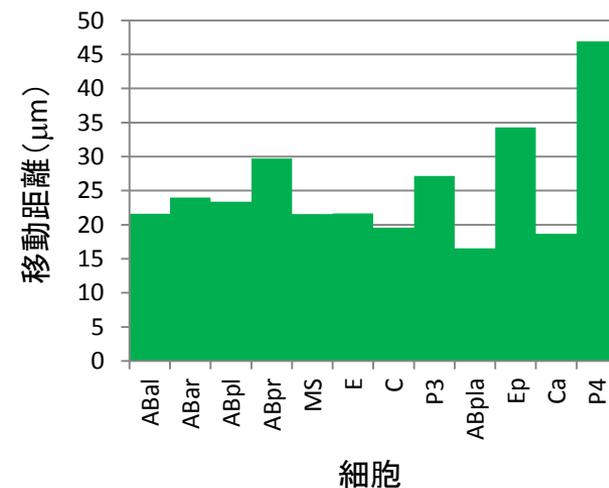
Supatto et al. 2009
ショウジョウバエ



Kyoda et al. 2013
線虫



Bao et al. 2006
線虫

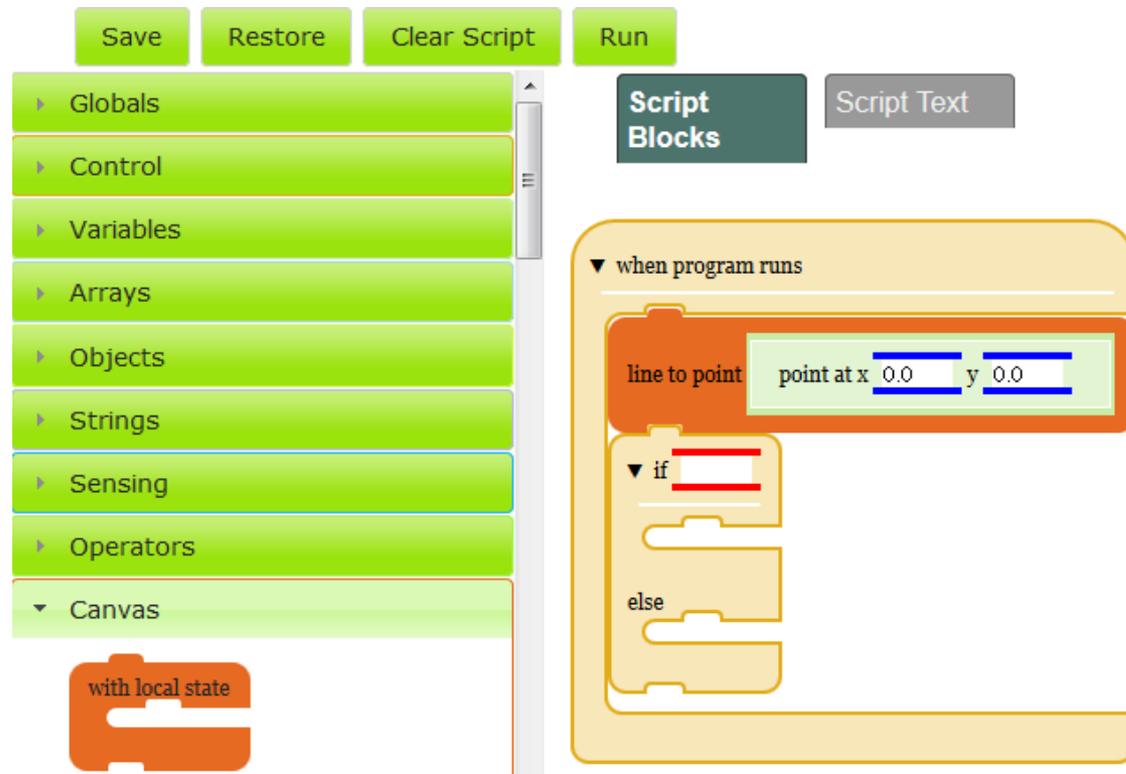


BDML形式で記述された
様々な定量データの解析が可能

BDML形式で記述された
異なる研究室のデータの比較解析が可能

ビジュアルプログラミングによる ウェブブラウザ上での表現型解析

- 実験生物学者が視覚的操作により、定量データを解析できるような仕組みを構築



動画像からBDMLデータを生成するインフラの構築

- OMERO と ImageJ + MATLAB の連携
 - SSBDのOMEROシステムから画像を読み込み、画像処理したデータをBDML形式で出力するAPIを作成する



RDFによるデータベース連携

- SSBDのリソース (URI) をRDFを介し、**オントロジーDB** に接続し、異なる生物間の動態を比較する

RNAi 線虫の動態

KOマウスの動態

RDF/フェノーム統合DB

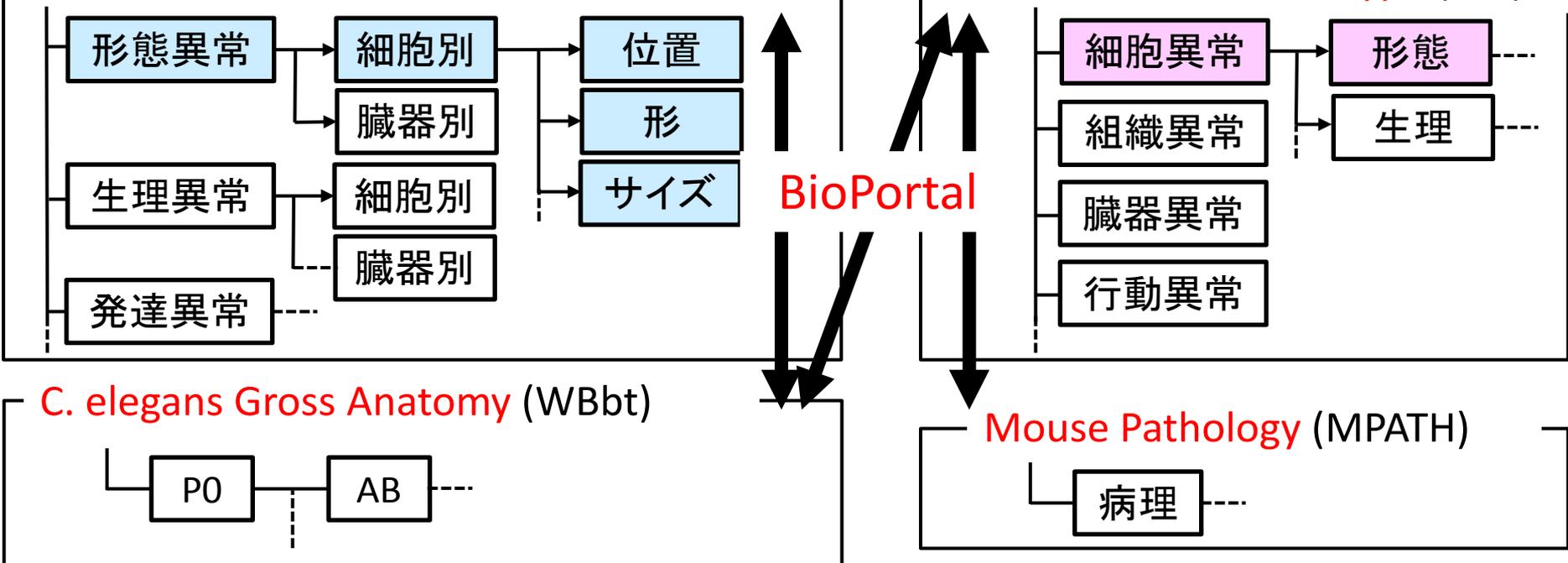
C. elegans Phenotype (WBPhenotype)

Mammalian Phenotype (MP)

BioPortal

C. elegans Gross Anatomy (WBbt)

Mouse Pathology (MPATH)



計測とシミュレーションデータによる予測

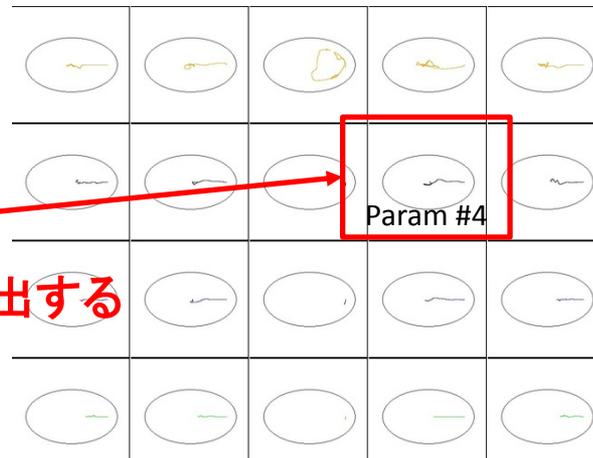
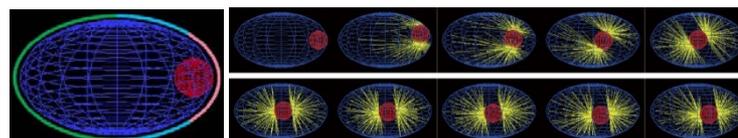
- 動態の比較から遺伝子の機能を予測する

計測データ



遺伝子機能の阻害時の定量データ

シミュレーションの結果



パラメータの変動時のシミュレーション結果

類似した動態を検出する



遺伝子の細胞内における物理的な機能が予測できる

シミュレーション研究を支援する情報 インフラの構築

- パラメータ・データベース
 - パラメータの格納
 - 元データの格納
 - ユーザー評価システム
 - いいね！など

Table 1. Simulation Parameters and Experimentally Obtained Parameters

Parameter Description	Values Used in Simulation		Experimentally Obtained Values	References Obtained
	Standard Condition	Range of Fluctuation		
Fixed Parameters				
Long axis of egg [$\times 10^{-6}$ m]	50			
Short axes of egg [$\times 10^{-6}$ m]	30			
Radius of pronucleus [$\times 10^{-6}$ m]	5			
Varied Parameters				
Dynamic Instability and Number of Microtubules (MTs)				
Growth velocity of MT (V_g) [$\times 10^{-6}$ m/s]	0.12	0.118–0.328	0.118–0.328	Cassime 1990; V and W
Shrinkage velocity of MT (V_s) [$\times 10^{-6}$ m/s]	0.288	0.157–0.537	0.157–0.537	See above
Catastrophe frequency of MT (f_{cat}) [1/s]	0.014	0.0115–0.046	0.0115–0.046	See above
Rescue frequency of MT (f_{res}) [1/s]	0.044	0.0113–0.133	0.0113–0.133	See above
Number of MTs (N)	98	30–550	>30	This study
Pushing Forces				
Rigidity of MT (κ) [$\times 10^{-24}$ Nm ²]	10	2–50	4.6–41	Venier et al. 1995; Howard, 2004
Elongation-rate parameter (A) ^b [$\times 10^{-6}$ m/s]	V_g	$V_g-1.25 \times V_g$	$\sim V_g$	Dogteron 2001; Tran et al.
Force-dependency parameter (B) ^b [$\times 10^{10}$ /N]	3.2	3.2–53	3.2–53	
Pulling Forces				
Stall force of motor (F_{stall}) [$\times 10^{-12}$ N]	1.1	0.5–1.5	0.78–1.1	Gittes et al. 1998; Mallik and Howard, 2002
Maximum velocity of motor (V_{max}) [$\times 10^{-6}$ m/s]	2.0	0.5–5.0	1.1–2.0	
Density of motor on MT (D) [$\times 10^3$ /m]	100	20–500	N.A. ^c	
Drag Force of Pronucleus				
Viscosity of cytosol (η) [Ns/m ²]	1	0.001–10	0.0018–3	Hiramoto et al. 2001
Stokes' radius of pronucleus (r) [$\times 10^{-6}$ m]	10	5–15	5–15 ^d	
Model-Specific Parameter				
Time step [s]	0.05	0.01–0.1		Nédélec, 2001

^a Only the references for values in the Table are given.

^b MT force-velocity relationship for MT growth velocity (v) encountering an opposing force (F) modeled as $v = V_g - AV_g/F$. V_g is the growth velocity for freely growing MTs, and A and B are two parameters defining the MT force-velocity relationship (see Supplemental Data).

^c N.A. indicates data not available. See text for choice of the parameters.

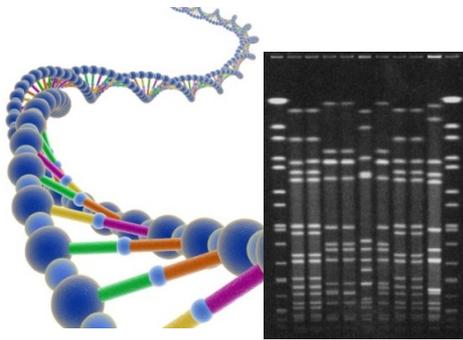
^d Stokes' radius of the pronucleus was estimated such that its value was between the radius of the pronucleus and the radius of the egg because MTs associated with the pronucleus increase the drag force on the pronucleus (Reinsch and Golev, 2004).

本研究開発の目標

- 我が国の生命動態システム科学分野の全てのデータベースを統合する体制と仕組みの構築
- 研究開発終了時における、我が国の生命動態システム科学分野の全てのデータベースの統合
- 生命動態システム科学分野の全世界のデータベースの統合
- 計測、モデリング、設計の有機的な連携を加速する情報インフラの開発

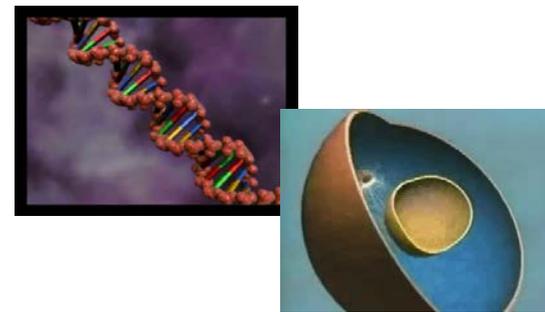
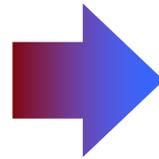
生命動態システム科学

- 生命を動的システムとして理解し、操作する生命科学
- 新たな生命科学の潮流として21世紀の科学全体への大きな影響が期待される



分子生物学

遺伝子の制御



生命動態システム科学

細胞の制御

複雑な生命現象の動態を時・空間を有する先端定量計測と高精密モデリングをもとに、*in silico*と*in vitro*で再構成