# SSBD: An Integrated Database of Spatiotemporal Dynamics and Microscopy Data

Yukako Tohsato[1], Koji Kyoda[1], Kenneth Ho[1], Shuichi Onami[1,2]
1) Lab. for Dev. Dyn., RIKEN QBiC, Japan, 2) NBDC, JST, Japan

## Overview

SSBD aims to provide a computer platform to support and drive forward a modeling and experimental approach for understanding molecular/cellular/developmental systems.

We have developed a Systems Science of Biological Dynamics (SSBD) database. It is a repository of quantitative data of biological dynamics. A new unified data format Biological Dynamics Markup Language (BDML) is being used by SSBD. Quantitative data can be exchanged easily using BDML/SSBD (Fig. 1). New software tools for analysis and visualization can be built utilizing SSBD and BDML.

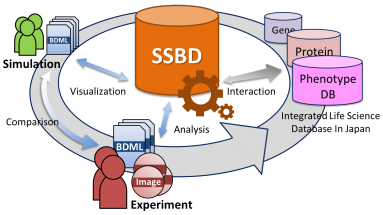


Fig. 1: Project aim

## Biological Dynamics Markup Language (BDML)

We propose a new BDML schema (ver. 0.15) to represent different types of biological dynamical data. BDML allows us to use a unified format to analyze biological dynamics (Fig. 2).

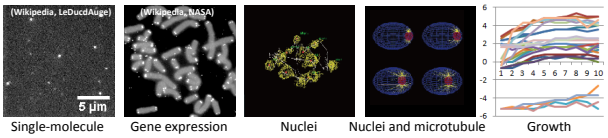


Fig. 2: Heterogeneous biological dynamics by representing BDML format

The schema consists of five top-levels. Data are represented in four different types of features, point, line, circle, and sphere (Fig. 3).

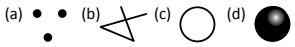


Fig. 3: Types of features.

1. Info: Information of the BDML file, e.g. license, release date, etc.
2. Summary: Short summary of the numerical information of biological dynamics
3. Contact: Contact information of the author/owner/maintainer of the BDML files
4. Methods: Sources files and extraction protocols
5. Data: Description of numeric data

GFP-histone-labeled nuclei (Sarov et al. 2012) can be described in BDML (Figs. 4-6).

```
<data>
  <scaleUnit>
    <xScale>0.09</xScale>
    <yScale>0.09</yScale>
    <zScale>1.0</zScale>
    <xyzUnit>micrometer</xyzUnit>      } xyz,t units
    <tScale>1.0</tScale>
    <tUnit>minute</tUnit>
  </scaleUnit>
  <object>
    <objectName>nucleus</objectName>   ← Summary of dynamics
  </object>
  <feature>  ← Summary of property information (GFP)
    <featureName>totalGFPsignal</featureName>
    <featureScale>totalGFPsignal</featureScale>
    <featureUnit>intensity</featureUnit>
  </feature>
  <component>  ← Component name
    <componentID>1001</componentID>
    <componentName>polar1</componentName>
    <time>1</time>  ← Time point
    <measurement>
    <objectRef>nucleus</objectRef>  ← Kind of dynamics
    <sphere>
      <coords>163,343,6.1</coords>  ← x,y,z
      <radius>1.62</radius>  ← Radius
      <property>
        <featureRef>totalGFPsignal</featureRef>
        <featureValue>103162</featureValue>
      </property>
    </sphere>
  </measurement>  ↑ Property information (GFP)
  </component>
  <component>
    <componentID>1002</componentID>
    <componentName>EMS</componentName>
    <time>1</time>
    ...
```

Fig. 5: Example of data section

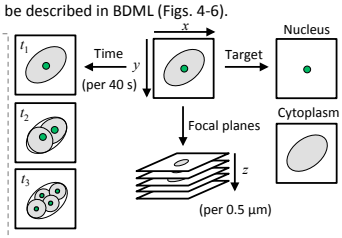


Fig. 4: Cell dynamics with GFP information

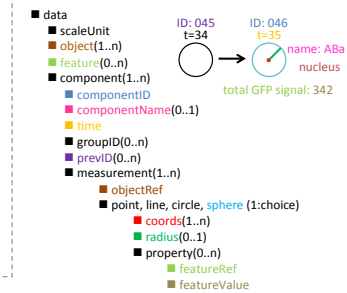


Fig. 6: Structure of the data section – e. g. representing a sphere

## Data analysis by BDML format

Phenochar (ver. 1.0) calculates phenotypic characters from BDML data (Fig . 7).

Results of phenotype analysis for *D. melanogaster* and *C. elegans* can be found in Fig. 8.

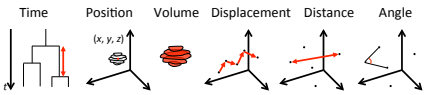


Fig 7: Calculating phenotypic characters among components

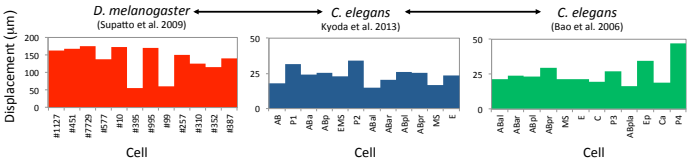


Fig. 8: Examples of phenotype analysis

## SSBD database http://ssbd.qbic.riken.jp/

Systems Science of Biological Dynamics (SSBD) database manages both raw image data and quantitative numerical data (Fig. 9).

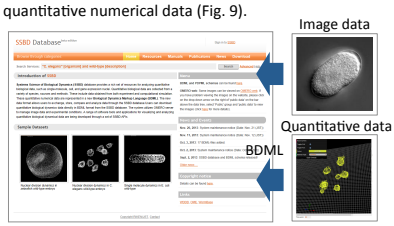


Fig. 9: SSBD database

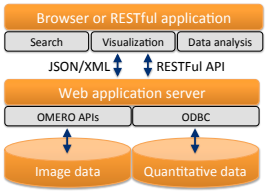


Fig. 10: Overview of our SSBD database

OMERO is integrated into SSBD allowing image data, quantitative data and experimental conditions to be organized (Fig. 10). More than 2 million DIC images can be accessed via SSBD (Fig. 11).

Fig. 11: Browsing images remotely on browser via OMERO server.

SSBD provides a wide variety of quantitative data (Fig. 12 and Table 1). JavaScript viewer is used to visualize quantitative data through a web browser.

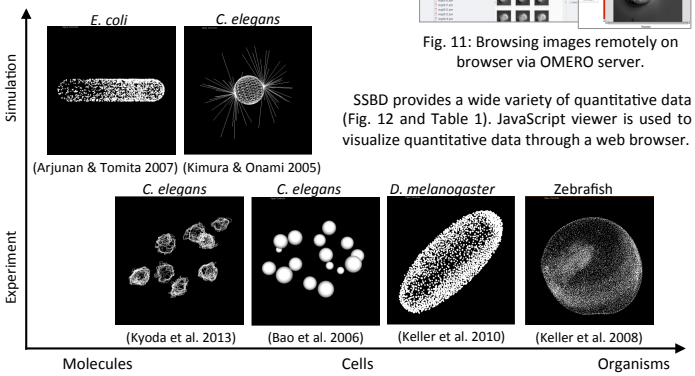


Fig. 12: Visualized quantitative data

Table 1: List of available resources in SSBD: ■planned release, □currently available

| Organism | Dynamics | Type of source | Paper | #BDML | Type | #Entries | #Image | Type |
|---|---|---|---|---|---|---|---|---|
| Mouse | nuclear | experiment | Kurotaki et al. 2007 | 1 | point | 12096 | 80 | liff |
| Mouse | nuclear | experiment | Bashar et al. 2012 | 1 | point | | | |
| Zebrafish | nuclear | experiment | Keller et al. 2008 | 7 | point | 9613526 | | |
| *D. melanogaster* | nuclear | experiment | Keller et al. 2010 | 2 | sphere | 3624857 | | |
| *D. melanogaster* | nuclear | experiment | Supatto et al. 2009 | 1 | point | 40534 | | |
| *C. elegans* | nuclear | experiment | Kyoda et al. 2013 | 186 | line | 262 | 33480 | ipm |
| *C. elegans* | nuclear | experiment | Bao et al. 2005 | 2 | sphere | 23980 | | |
| *C. elegans* | nuclear | simulation | Kimura & Onami 2005 | 100 | sphere + line | 24001 | | |
| *C. elegans* | nuclear | experiment | Kyoda et al. | 261 | line | 597 | 256 | lif |
| *C. elegans* | nuclear | experiment | Sönnichsen et al. 2005 | 979 | line | 1197 | | |
| *C. elegans* | nuclear | experiment | Gonczy et al. 2000 | 1 | line | 2988 | | |
| *C. elegans* | nuclear | experiment | Takayama et al. | 12 | line | 1628 | 4899 | tiff |
| *C. elegans* | expression | experiment | Sarov et al. 2012 | 273 | sphere | 17456 | | |
| *C. elegans* | movement | experiment | Cronin et al. 2005 | 11 | line | 1450 | | |
| *D. discoideum* | single- molecule | experiment | Jin et al. | 1 | point | 987 | 1 | avi |
| *E. coli* | single- molecule | simulation | Arjunan & Tomita 2010 | 1 | point | 721 | | |

## ImageJ coordination

We developed an ImageJ plugin named BDML4DViewer for visualization (Fig. 13). We are also planning to develop software applications to work with BDML content, e.g. programming libraries to create/edit/translate BDML files

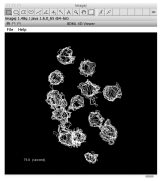

Fig. 13: A snapshot of BDML4DViewer

## Future plans

### RESTful API

The RESTful API allows third party applications to access SSBD using various programming languages, tools and development frameworks (Fig. 14).

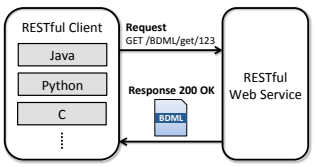


Fig. 14: RESTful API service

### Experiment versus simulation

Developing an innovative environment for understanding molecular, cellular, developmental systems by comparative analysis of data obtained from experiment and simulation (Fig. 15).

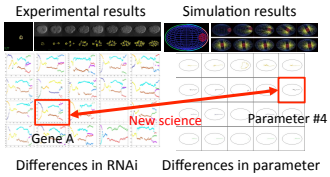


Fig. 15: Comparative analysis