2020/8/20 (revised)

SSBD:database – quantitative data submission guidelines

- SSBD:database primarily shares quantitative data of biological dynamics (i.e. numerical spatiotemporal information on biological dynamics) that can be reused for further research and analysis.
- Spatiotemporal information is normally represented by a set of coordinates to represent how molecules, cells, organelles, individuals, etc., change their positions or shapes that were observed.
- We currently share

  1) spatiotemporal information that can be obtained by image processing or machine learning from microscopy images,

  2) spatiotemporal information that can be obtained by mechanobiological simulation,

  3) omics data (e.g., transcriptome or metabolome data) and their dynamics that are related to spatial information along with microscopy image.

Recommended file formats

- We prefer to receive quantitative data in BDML/BD5 (Biological Dynamics Markup Language) format; the open unified format for spatiotemporal information of biological dynamics.
  - BDML/BD5 format http://ssbd.qbic.riken.jp/bdml/
- We also prefer to receive quantitative data in commonly used file format such as MS Excel, CSV (comma separated value), plain text, etc. SSBD team will convert the provided data into BDML/BD5 format to share in SSBD:database.
- If you are using commercial software (such as Imaris), please provide us the original files, and in addition, please export the quantitative data into commonly used file formats such as MS Excel, CSV, plain text, etc., and provide us with the original files. When the original files are not readable, SSBD team can convert the additional files to BDML/BD5 format.
- In order to help SSBD team to understand your quantitative data, we will appreciate if you can give us a simple explanation of the meaning of the quantitative data in the files, e.g. explanation and meaning of rows and columns, etc.

Name of quantitative data for identification (Local ID)

- Please prepare an ID for each dataset. If the quantitative data contains multiple experimental conditions, please provide an ID for each individual dataset.
- We prefer to use the ID that is normally used in your laboratory or research project that helps to identify to different experiments or conditions. For instance, you can use the mutant name or strain name, date, experiment number and their combination to create a unique ID. If the data is managed and grouped by using folders, we can accept the file path name as ID.

  ex.)

  “RNAi_par-3_180214_03” (combination of experiment name, gene name, date, and experiment number)

  “/data/2018/02/14/par-3_RNAi_embryos/03” (file path name)

Examples of quantitative data in SSBD:database

1. Coordinate information that represents the positions or shapes of molecules/cells/organelles/individuals/etc., when they are moving or changing. These coordinates are measured and obtained by image analysis and they include temporal information.

- 1-dimensional spatial and temporal information: change of length, size, etc. (Example 1)
- 2-dimensional spatial and temporal information: moves or changes of coordinate information on x-y plane. (Example 2)

- 3-dimensional spatial and temporal information: moves or changes of coordinate information in x-y-z space. (Example 3)
- Features associated with spatial information: coordinates in x-y-z space and feature values, i.e. fluorescence intensity of GFP expression in nuclei detected by image analysis. (Example 4)

2. Coordinate information that represents temporal positional changes or shape changes of molecules/cells/organelles/etc., calculated by mechanobiological simulation. (Example 5)

Not recommended for submission to SSBD:database

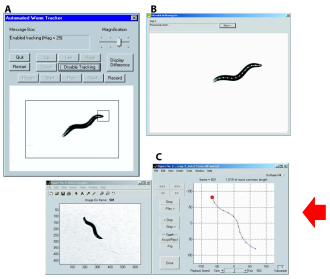
1. Data without temporal information

   However, we will share data containing only spatial information (without temporal information) such as omics data related to pixels of an image even though it is difficult to measure temporal information.

2. Data without spatial information

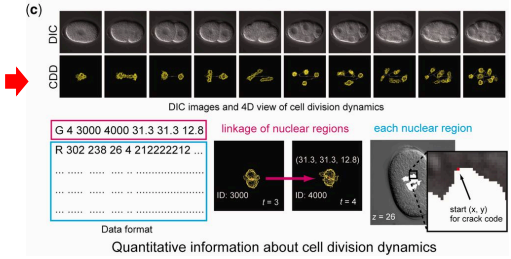
- We do not prefer to receive data contain only temporal dynamics of molecular activity or fluorescence intensity, but lack of spatial information such as positions or shapes of molecules or cells e.g. total sum of luminance value for a fixed area on an image.
- We do not prefer to receive the result of simulation that contains only temporal information without spatial information.

Example 1: single molecule dynamics (position)



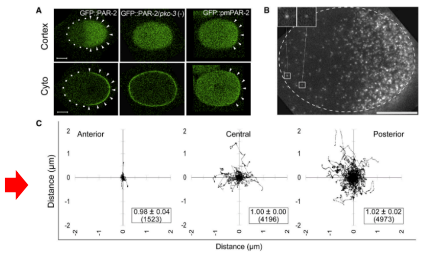
"Figure 2" by Cronin CJ et. al. (2005) BMC Genet, 6: 5 is licensed under CC-BY

Example 2: cell division dynamics (contour)



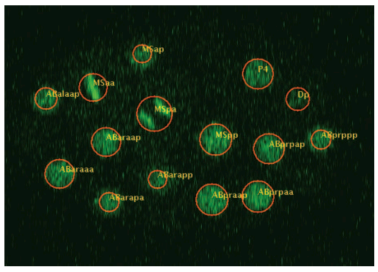
"Figure 1" by Kyoda K et. al. (2013) Nucleic Acids Res, 41(Database issue): D732-7 is licensed under CC-BY-NC

Example 3: individual dynamics (shape)



"Figure 1" by Arata Y et. al. (2016) Cell Rep, 16(8): 2156-2168 is licensed under CC-BY

Example 4: cell division dynamics / additional values (shape / fluorescence intensity)



Nuclear position and size change + GFP expression intensity

"Figure 2" by Bao Z et. al. (2006) Proc Natl Acad Sci U S A, 103(8): 2707-12, "Copyright (2006) National Academy of Sciences."

Example 5: single molecule physical simulation

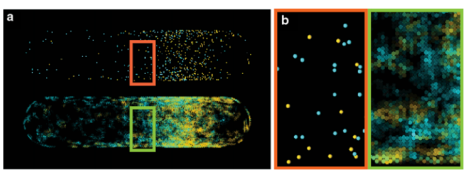


Fig. 5 Comparison of single molecule visualization to simulated microscopy image. **a** Single molecule visualization (*Top*) of MinDE^{m} (*yellow*) and MinE^{m} (*cyan*) on *Escherichia coli* membrane compared to the corresponding simulated fluorescence microscopy image (*Bottom*). The exposure time of the fluorescence image is 500 ms. The proteins were simulated according to reactions (17)–(24). The diffusion coefficients and model parameters are listed in Table 2. **b** The magnified area of the respective boxes shown in (**a**)

"Figure 5" by Arjunan SN & Tomita M (2010) Syst Synth Biol, 4(1): 35-53 is licensed under CC-BY-NC