

FICHE DE RECUEIL DES FAITS MARQUANTS DEPARTEMENTS/CENTRES

(Renseigner une fiche par fait marquant.

Les départements/centres peuvent choisir de faire la synthèse de plusieurs FM en une seule fiche si pertinent)

Année concernée : 2021 (Publication ou réalisation de 2021)

Fiche envoyée par : MICA/Centre Île de France Jouy-en-Josas - Antony

Priorité attribuée au FM (à renseigner par le CD/PC) :

Titre du fait marquant : BiPSim: un simulateur stochastique flexible et générique des processus cellulaires

Catégorie: Publication <https://doi.org/10.1038/s41598-021-92833-5> (indiquer le DOI)

Contact : Anne Goelzer, Vincent Fromion
Unité : MaIAGE
Département : MICA/MathNum
Centre INRAE : Île de France Jouy-en-Josas - Antony

OS ou OP INRAE 2030 (cf. classification proposée en annexe) :

OS 5.1. Systèmes complexes et évolutifs

Metaprogramme (si adapté) :

Mots-clés (rubrique libre) : bruit cellulaire, logiciel, simulation stochastique

Résumé (10 à 15 lignes max. à rédiger sous une forme exportable dans le Rapport Annuel.)

Simuler le comportement stochastique des systèmes biologiques aux échelles les plus fines en un temps de calcul raisonnable est un enjeu majeur en biologie. Face au nombre colossal d'évènements à simuler, des compromis entre granularité des phénomènes simulés, temps de calcul et généricité du simulateur sont nécessaires. Dans ce contexte, les chercheurs du laboratoire MaIAGE de l'INRAE et de l'équipe MAMBA de l'INRIA ont développé le simulateur stochastique open-source BiPSim, combinant une représentation abstraite efficace des réactions et une implémentation en temps constant de l'algorithme de simulation stochastique de Gillespie pour les réactions. BiPSim permet ainsi de simuler efficacement certains processus cellulaires (comme les processus de polymérisation) de façon détaillée et particulièrement délicats à simuler habituellement compte tenu du nombre d'évènements associés. Différents niveaux de granularité dans la description des processus peuvent être traités et simulés simultanément, offrant ainsi un compromis entre temps de calcul et granularité du processus. Conçu pour être générique et donc aisément adaptable à différents types de cellules, BiPSim rend ainsi accessible à la communauté biologique la simulation stochastique de processus cellulaires et permet de confronter les simulations et des données omiques hétérogènes acquises en cellule unique (données de *single cell*).

(400 à 500 mots/ 2700 à 3400 caractères max. pour l'ensemble des 4 rubriques ci-dessous)

Contexte et enjeux : Le simulateur BiPSim s'inscrit dans un contexte de biologie prédictive, où il s'agit de simuler (de façon efficace) l'ensemble des processus biologiques d'une cellule. La simulation (stochastique) de l'ensemble des évènements moléculaires est actuellement impossible en terme de temps de calcul [1,2]. La modélisation détaillée d'une cellule entière et la simulation associée nécessite ainsi l'intégration de processus cellulaires hétérogènes ayant des formalismes de modélisation différents [3], tout en restant efficace en terme de temps de calcul, et suffisamment générique pour être adapté à une grande diversité de cellules.

Résultats : Dans cet article, nous présentons BiPSim, un simulateur stochastique open-source de processus biologiques de polymérisation tels que la réplication, la transcription et la traduction [4]. BiPSim combine une représentation abstraite efficace des réactions et une implémentation en temps constant de l'algorithme de simulation stochastique (SSA) de Gillespie pour les réactions [5]. BiPSim permet ainsi de simuler efficacement et de manière stochastique des processus de polymérisation à grande échelle. De plus, différents niveaux de granularité dans la description des processus peuvent être traités et simulés simultanément, ce qui permet à l'utilisateur de définir un compromis entre le temps de calcul et la granularité du modèle. Nous avons évalué les performances de BiPSim en simulant l'expression des gènes à l'échelle du génome dans des bactéries comme *Bacillus subtilis* pour plusieurs niveaux de granularité, et en le comparant avec d'autres simulateurs stochastiques sur des benchmarks représentatifs de la complexité des phénomènes biologiques simulés. Enfin, étant donné qu'aucune information spécifique au type de cellule n'est codée en dur dans le simulateur, les modèles peuvent facilement être adaptés à d'autres organismes. BiPSim devrait ainsi ouvrir de nouvelles perspectives pour la simulation à l'échelle du génome de phénomènes stochastiques en biologie.

Perspectives : BiPSim représente la première étape (la partie stochastique) d'un simulateur de cellule bactérienne. BiPSim a été conçu de façon à pouvoir être d'une part interfacé avec d'autres modules de simulation (stochastiques ou déterministes notamment pour les processus métaboliques), et d'autre part adapté facilement à d'autres organismes. Enfin, les simulations sont directement comparables à des données omiques hétérogènes acquises sur cellule unique (données de *single cell*), et permettront d'analyser la variété des phénotypes observés sur cellule unique [3,6].

Valorisation : Elle se fait à travers divers développements, en particulier autour des enjeux portés par les biotechnologies (biologie de synthèse), l'ingénierie métabolique et le design de souches bactériennes, afin d'anticiper *in silico* l'impact de modifications génétiques dans la cellule.

Références bibliographiques :

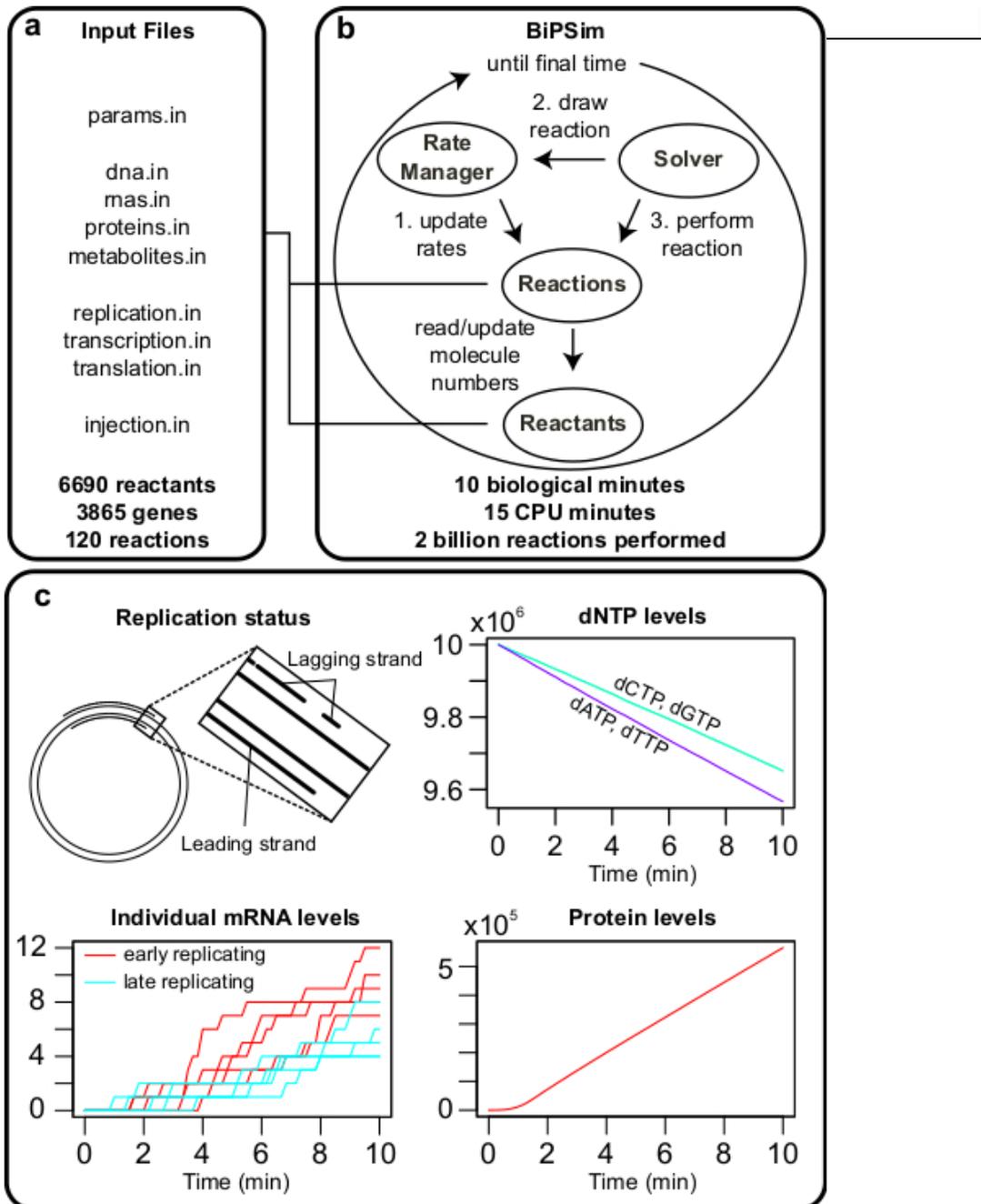
1. Meng, T. C., Somani, S. & Dhar, P. (2004). Modeling and simulation of biological systems with stochasticity. *Silico Biol.* 4, 293–309
2. Köhler, A., Krivine, J. & Vidmar, J (2014). A rule-based model of base excision repair. in International Conference on Computational. *Methods in Systems Biology*, 173–195
3. Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival Jr, B., ... & Covert, M. W. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2), 389-401.
4. Fischer, S., Dinh, M., Henry, V., Robert, P., Goelzer, A., & Fromion, V. (2021). BiPSim: a flexible and generic stochastic simulator for polymerization processes. *Scientific Reports*, 11(1), 1-16.
5. Slepoy, A., Thompson, A. P. & Plimpton, S. J (2008). A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks. *J. Chem. Phys.* 128, 205101.
6. Dessalles, R., Fromion, V. & Robert, P (2020). Models of protein production along the cell cycle: An investigation of possible sources of noise. *Plos one* 15, e0226016.

Illustrations (photos au format jpg, avec légende, auteur de la photo, et copyright s'il y en a un)

Légende : Simulation de l'expression des gènes bactériens à l'échelle du génome. (a) Fichiers d'entrée décrivant des modèles détaillés pour la réplication, la traduction et la transcription. (b) Fonctionnement de BiPSim. Le simulateur instancie les réactants et les réactions déclarés dans les fichiers d'entrée. Le solveur exécute une boucle itérative dans laquelle les taux de réaction sont mis à jour et la réaction suivante est sélectionnée selon l'algorithme de Gillespie. (c) Résultats de la simulation obtenus après 20% du cycle cellulaire simulé. Les axes verticaux indiquent le nombre de molécules. En haut à gauche : état de la réplication et zoom sur une des fourches de réplication. En haut à droite : la consommation de dNTP pour la réplication de l'ADN reflète le biais AT dans la composition des chromosomes de *Bacillus subtilis*. En bas à gauche : niveaux d'ARNm de cinq gènes situés à proximité de l'origine de réplication (rouge) et de cinq gènes qui n'ont pas encore été répliqués (bleu). En bas à droite : production totale de protéines.

Auteur : Stephan Fischer

Copyright : Cette figure est extraite du papier Fischer et al. *Scientific Reports*, 11(1), 1-16, 2021 (<https://doi.org/10.1038/s41598-021-92833-5>).



Champ thématique Mica dans lequel s'inscrit le fait marquant

Inscrivez un X dans la ou les case(s) correspondantes

- | | |
|-------------------------------------|-----------------------------------------------------------------------------|
| <input checked="" type="checkbox"/> | CT1: Biotechnologies industrielles, environnementales et de la santé |
| <input type="checkbox"/> | CT2: Aliments et sûreté des aliments |
| <input type="checkbox"/> | CT3: Santé animale et humaine |

Grands objectifs Scientifique (GOS) et Fronts de sciences (FS) Mica du nouveau SSD dans lequel s'inscrit le fait marquant

Inscrivez un X dans la ou les case(s) correspondantes

- | | |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | GOS1 : Comprendre le fonctionnement et la dynamique des microbiotes de l'aliment pour accompagner la transition agroécologique |
| <input type="checkbox"/> | FS1 : Flux microbiens dans les systèmes alimentaires et impact des changements |
| <input type="checkbox"/> | FS2 : Comprendre et prévoir le fonctionnement des microorganismes et écosystèmes microbiens en réponse aux changements et impacts sur les qualités. Concevoir des solutions adaptées. |
| <input type="checkbox"/> | FS3 : Aliments fermentés pour la durabilité des systèmes et la santé humaine (avec AlimH et TRANSFORM). |
| <input type="checkbox"/> | FS4 : Appréciation des risques et bénéfiques (multicritères) |
| <input type="checkbox"/> | GOS 2 : Etudier et optimiser les systèmes microbiens pour les biotechnologies |
| <input checked="" type="checkbox"/> | FS1 - Approfondir notre connaissance des processus biologiques |
| <input type="checkbox"/> | FS2 - Construire des châssis et des dispositifs performants |
| <input type="checkbox"/> | FS3 – Maitriser les consortia microbiens naturels ou synthétiques (avec TRANSFORM) |
| <input checked="" type="checkbox"/> | FS4 - Co-concevoir des microorganismes, écosystèmes microbiens et procédés (avec TRANSFORM) |
| <input type="checkbox"/> | GOS 3 : Caractérisation fonctionnelle des holobiontes et modulation à des fins de santé |
| <input type="checkbox"/> | FS1 - Comprendre le fonctionnement de l'holobionte) |
| <input type="checkbox"/> | FS2 - Prendre en compte la diversité et la variabilité du microbiote et de l'hôte |
| <input type="checkbox"/> | FS3 - Effet de l'environnement, alimentation et du mode de vie/d'élevage sur l'équilibre de l'holobionte |
| <input type="checkbox"/> | FS4 - Développer des approches préventives et thérapeutiques (avec SA) |
| <input type="checkbox"/> | GOS 4 : Comprendre et maîtriser les pathogènes et la résistance aux antibiotiques dans les systèmes alimentaires pour anticiper leur émergence et réémergence |
| <input type="checkbox"/> | FS1 - Comprendre la dynamique des pathogènes : survie, évolution, adaptation et transition entre leurs différents états |
| <input type="checkbox"/> | FS2 - Comprendre la genèse, le transfert et la dissémination de la résistance aux antimicrobiens |
| <input type="checkbox"/> | FS3 - Identifier de nouvelles stratégies alternatives aux antimicrobiens |

Annexe - Classification des faits marquants –INRAE 2030

- OS 1.1. Changement climatique : intégrer les démarches d'atténuation et d'adaptation
- OS 1.2. Biodiversité : un patrimoine mieux préservé et un levier d'action davantage mobilisé
- OS 1.3. Compréhension et mobilisation des mécanismes d'adaptation du vivant pour la sélection génétique et la préservation de la biodiversité
- OS 1.4. Evaluation et gestion des risques naturels et climatiques
- OS 1.1. Changement climatique : intégrer les démarches d'atténuation et d'adaptation
- OS 1.2. Biodiversité : un patrimoine mieux préservé et un levier d'action davantage mobilisé
- OS 1.3. Compréhension et mobilisation des mécanismes d'adaptation du vivant pour la sélection génétique et la préservation de la biodiversité
- OS 1.4. Evaluation et gestion des risques naturels et climatiques
- OS 2.1. Renforcer la compréhension des processus des transitions et enjeux d'autonomie
- OS 2.2. Progression vers des agricultures sans pesticide de synthèse
- OS 2.3. Transition des élevages
- OS 2.4. Construction des qualités des régimes alimentaires
- OS 2.5. Une alimentation saine et durable accessible et valorisante pour tous
- OS 3.1. Cycles du carbone, de l'azote et du phosphore dans les écosystèmes terrestres
- OS 3.2. Cycle de l'eau, relations entre grand et petit cycles
- OS 3.3. Traitement et usages des biomasses, coproduits, eaux usées et résidus organiques
- OS 3.4. Produits biosourcés : de nouvelles relations marchandes et dynamiques sociales
- OS 4.1. Emergences et re-émergences des maladies transmissibles, au sein et entre les systèmes environnementaux, agricoles et alimentaires
- OS 4.2. Pollutions, contaminants et exposome
- OS 4.3. Une nutrition préventive pour la santé publique et environnementale
- OS 5.1. Systèmes complexes et évolutifs
- OS 5.2. Capteurs et systèmes d'acquisition d'information
- OS 5.3. Des agro-équipements pour la transition agroécologique
- OS 5.4. Technologies de l'information, réseaux et nouveaux pouvoirs

- OP 1.1. Innover par la recherche partenariale en favorisant la co-construction et la co-réalisation
- OP 1.2. Anticiper les grands défis et éclairer les débats sociétaux et les politiques publiques par l'expertise scientifique
- OP 1.3. Ouvrir la science et partager les connaissances
- OP 2.1. Des écosystèmes académiques régionaux aux dispositifs de coordination nationale
- OP 2.2. Une présence et une coopération européennes essentielles
- OP 2.3. Une recherche de référence à l'international
- OP 2.4. Infrastructures de recherche
- OP 3.1. La stratégie RSE porteuse de sens et d'identité
- OP 3.2. INRAE, acteur investi dans la préservation de l'environnement
- OP 3.3. INRAE, employeur engagé
- OP 3.4. INRAE, acteur ouvert et transparent