

## David Parsons

34 rue du Luizet, 69100 Villeurbanne  
06 47 86 87 55 • david.parsons@inria.fr  
www.parsons.eu

# Développeur de logiciels en bio-informatique

Développeur spécialisé dans le domaine de la bio-informatique, j'affectionne particulièrement la recherche du meilleur compromis entre la performance d'une part, et la lisibilité, maintenabilité et généricité du code d'autre part. La Recherche publique constitue pour moi un environnement particulièrement motivant et valorisant.

## COMPÉTENCES

- Langages : Expert C/C++, maîtrise de Java, connaissances en Matlab, Python, R
- Web : HTML, CSS, PHP, Javascript
- Outils de développement : Forge, IDE (NetBeans, C-Lion), VCS (svn, Git), Build (autotools, Cmake), Packaging (deb, rpm), Debug (gdb, ddt, memcheck), Profiling (gprof, massif), Tests (CTest, JUnit, gtest), Intégration Continue, Documentation (UML, doxygen, groff, LaTeX), Refactoring, Programmation parallèle
- Conduite de projets : Méthodes d'inspiration Agile (centralité de l'utilisateur, cycles courts, pair-programming, stand-ups), Encadrement, Formation, Conduite de réunions, Organisation de communautés de développement, Standardisation des pratiques
- Connaissances biologiques : Bonne connaissance générale de la biologie computationnelle, de la biologie des systèmes et dans une moindre mesure, de la génomique. Facilité à communiquer avec les biologistes
- Administration système : Administration de serveurs de calculs et de sauvegardes (Debian, CentOS), Administration réseau

## EXPÉRIENCE EN RECHERCHE ET DÉVELOPPEMENT

**Octobre 2012 – Aujourd'hui**

**INRIA Rhône-Alpes**

**Ingénieur Recherche et Développement**

dans le domaine de la bio-informatique

- Encadrement technique et suivi de 3 ingénieurs juniors et 3 stagiaires
- Accompagnement d'un ingénieur confirmé
- Contribution au développement de plusieurs logiciels dans le domaine de la bio-informatique : Aevol (plateforme d'évolution expérimentale *in silico*), FluobacTracker (segmentation et suivi de cellules dans des images de microscopie time-lapse), Kissplice (détection de polymorphismes dans des données RNA-seq) et SiMuScale (simulation multi-échelles de populations de cellules).
  - Amélioration de la qualité et de la performance des codes existants, ajout de nouvelles fonctionnalités
  - Structuration des équipes de développement autour de la forge et de la communauté d'utilisateurs autour de sites web dédiés
  - Mise en place de bonnes pratiques, de tests, et de processus d'intégration continue
  - Organisation de formations internes, conduite de réunions et de comités de pilotage
  - Standardisation et consolidation des processus de build et d'installation, réalisation et suivi de paquets précompilés
  - Rédaction de documentations à destination des développeurs et des utilisateurs
  - Contribution aux décisions en termes de licences logicielles et de distribution en lien avec le service STIP
  - 3 dépôts APP, 4 publications, 1 communication, organisation de tutoriels dans de grandes conférences internationales
  - Détail des réalisations par projet en annexe
- Conseil, formation et accompagnement des chercheurs dans leurs développements logiciels
- Administration réseau et système, support technique aux usagers du site
- Suivi du chantier d'aménagement des nouveaux bureaux de l'antenne de Lyon la Doua en lien avec les services SIC et SG
- Membre élu au Comité de Gestion Locale de l'AGOS (Comité d'Entreprise)

**Octobre 2008 – Septembre 2012**

**INSA de Lyon**

**Doctorant puis ATER**

Étude de l'Évolution Darwinienne par une approche de modélisation et de simulation

- Développement de la plateforme de simulation Aevol (évolution expérimentale *in silico*)
  - Prise en charge de la gestion du projet (plus de 10 personnes, ~800 jours-hommes)
  - Mise en place de bonnes pratiques de génie logiciel (VCS, tests, build)
  - Refonte complète de la base de code du projet. Performances temporelles  $\times 10$
  - Mise en place de méthodes permettant l'utilisation efficace de centres de calcul
- 8 publications dont 4 en premier auteur, 3 résumés, 2 communications, 3 posters, 2 séminaires invité
- Ouverture de nouvelles perspectives de recherche
- Mise en place de nouvelles collaborations (INSERM)
- Co-encadrement de 4 stagiaires ingénieurs
- Membre élu au Conseil des Études de l'INSA de Lyon, implication forte dans le chantier de réforme de l'organisation de la formation

## EXPÉRIENCE EN ENSEIGNEMENT

<b>2012 – 2015</b>	INSA de Lyon, dpt Bioinformatique et Modélisation Systèmes de Contrôle de Versions (SVN et Git) Méthodes de développement, conception et modélisation UML Programmation en C++	130h EqTD	CM, TD, TP	Vacataire	L3
<b>2011 – 2012</b>	INSA de Lyon, dpt Informatique Programmation en C++ Programmation en Java et persistance avec JPA Conception d'interfaces graphiques Programmation logique avec Prolog Programmation parallèle avec IPC system V	192h EqTD	TD, TP	ATER	L3/M1
<b>2008 – 2011</b>	INSA de Lyon, dpt Premier Cycle Bases de la programmation en Pascal Bases de Données et langage SQL Culture informatique (C2I)	192h EqTD	TD, TP	Moniteur	L1/L2
<b>2009</b>	Université Jean Moulin Interventions lors d'une école d'été de préparation à l'intégration en classes préparatoires organisée par la région Rhône-Alpes	20h EqTD	TP	Vacataire	Bac

## CURSUS UNIVERSITAIRE

2011	INSA de Lyon	<b>Thèse de Doctorat</b> en Informatique. « Sélection Indirecte en Évolution Darwinienne, Mécanismes et Implications »
2008	INSA de Lyon	<b>Master Recherche</b> en Informatique Major dans la spécialité « Intelligence Artificielle et Aide à la Décision »
2007	INSA de Lyon	<b>Ingénieur</b> en Informatique Une année en échange académique à l' <i>Instituto Tecnológico de Buenos Aires</i> (Argentine)

## LANGUES

Français :	Natif
Anglais :	Bilingue (bi-national France – Royaume-Uni)
Espagnol :	Courant (1 an à Buenos Aires)

## CENTRES D'INTÉRÊT

Sauveteur Secouriste du Travail (SST), Plongeur Sauveteur  
Plongée sous-marine, Volley-ball, Sports de montagne  
Voyages (Amérique latine, Asie, Afrique, Europe)

## ANNEXE 1 : PUBLICATIONS

### Journaux

Berenice Batut, [David P. Parsons](#), Stephan Fischer, Guillaume Beslon, Carole Knibbe (2014). In silico experimental evolution: a tool to test evolutionary scenarios. *BMC Bioinformatics*, 14(Suppl 15), S11.

Guillaume Beslon, [David P. Parsons](#), Yolanda Sanchez-Dehesa, Jose Maria Pena, Carole Knibbe (2010). Scaling Laws in Bacterial Genomes: A Side-Effect of Selection of Mutational Robustness. *BioSystems*, 102, 32-40.

Guillaume Beslon, [David P. Parsons](#), Jose Maria Pena, Christophe Rigotti, Yolanda Sanchez-Dehesa (2010). From Digital Genetics to Knowledge Discovery: Perspectives in Genetic Network Understanding. *Intelligent Data Analysis Journal*, 14(2), 173-191.

Yolanda Sanchez-Dehesa, [David P. Parsons](#), Jose Maria Pena, Guillaume Beslon (2008). Modelling Evolution of Regulatory Networks in Artificial Bacteria. *Mathematical Modelling of Natural Phenomena* 3(2), 27-66.

### Conférences

Carole Knibbe, [David P. Parsons](#) (2014). What happened to my genes? Insights on gene family dynamics from digital genetics experiments. *14<sup>th</sup> Intl. Conf. on the Synthesis and Simulation of Living Systems (ALIFE)*. New York City, US-NY.

Guillaume Beslon, Berenice Batut, [David P. Parsons](#), Dominique Schneider, Carole Knibbe (2013). An alife game to teach evolution of antibiotic resistance. *Proceedings of the 12<sup>th</sup> European Conference on the Synthesis and Simulation of Living Systems (ECAL 12)*. Taormina, Italy.

Alice Julien-Laferriere, Gustavo AT Sacomoto, Rayan Chikhi, Erwan Scaon, [David P. Parsons](#), Marie-France Sagot, Pierre Peterlongo, Vincent Miele, Vincent Lacroix (2013). New development in Kissplice: Combining local and global transcriptome assemblers to decipher splicing in RNA-seq data. *Proceedings of the Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM) 2013*. Toulouse, France.

[David P. Parsons](#), Carole Knibbe, Guillaume Beslon (2012). The Paradoxical Effects of Allelic Recombination on Fitness. *Proceedings of the 13<sup>th</sup> Intl. Conf. on the Synthesis and Simulation of Living Systems (ALIFE)*. East Lansing, US-MI.

Dusan Misevic, Antoine Frenoy, [David P. Parsons](#), Francois Taddei (2012). Effects of public good properties on the evolution of cooperation. *Proceedings of the 13<sup>th</sup> Intl. Conf. on the Synthesis and Simulation of Living Systems (ALIFE)*. East Lansing, US-MI.

[David P. Parsons](#), Carole Knibbe, Guillaume Beslon (2011). Homologous and Nonhomologous Rearrangements: Interactions and Effects on Evolvability. *Proceedings of the 11<sup>th</sup> European Conference on the Synthesis and Simulation of Living Systems (ECAL 11)*. Paris, France.

Carole Knibbe, [David P. Parsons](#), Guillaume Beslon (2011). Parsimonious modeling of scaling laws in genomes and transcriptomes. *Proceedings of the 11<sup>th</sup> European Conference on the Synthesis and Simulation of Living Systems (ECAL 11)*. Paris, France.

[David P. Parsons](#), Carole Knibbe, Guillaume Beslon (2010). Importance of the rearrangement rates on the organization of transcription. *Proceedings of the 12<sup>th</sup> Intl. Conf. on the Synthesis and Simulation of Living Systems (ALIFE)*. Odense, Denmark.

[David P. Parsons](#), Carole Knibbe, Guillaume Beslon (2010). Aevol : un modèle individu-centré pour l'étude de la structuration des génomes. *Proceedings of the 8<sup>th</sup> MajecSTIC Conference*. Bordeaux, France.

[David P. Parsons](#), Carole Knibbe, Guillaume Beslon (2010). Influence of the rearrangement rates on the organization of the genome transcription. *Proceedings of the Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM) 2010*. Montpellier, France.

Guillaume Beslon, Yolanda Sanchez-Dehesa, [David P. Parsons](#), Jose Maria Pena, Carole Knibbe (2009). Scaling Laws in Digital Organisms. *Proc. Information Processing in Cells and Tissues IPCAT'09*, Ascona, Switzerland.

### Thèse de Doctorat

[David P. Parsons](#) (2011). *Indirect Selection in Darwinian Evolution: Mechanisms and Implications*. PhD thesis. INSA de Lyon.

## ANNEXE 2 : EXPÉRIENCE EN RECHERCHE ET DÉVELOPPEMENT PAR PROJETS

### Aevol

#### Contexte personnel

Octobre 2008 – Septembre 2011 : Doctorant

Octobre 2011 – Septembre 2012 : ATER

Octobre 2012 – Septembre 2014 : Suivi et interventions ponctuelles

Depuis Octobre 2014 : Développeur à 40 % (dans le cadre d'une Action de Développement Technologique INRIA)

#### Description et contexte du projet

Aevol est une plateforme d'évolution expérimentale *in silico* développée par l'équipe INRIA Beagle (Lyon) et par l'équipe « Biologie systémique évolutive » de l'unité INSERM 1001 (Paris). Aevol permet d'étudier l'évolution de la structure du génome et de sa sélection indirecte sous des contraintes d'évolvabilité et de robustesse.

Équipe de développement bi-localisée d'environ 10 personnes d'horizons divers (biologie, physique, informatique), composée de chercheurs, doctorants, stagiaires et plus récemment ingénieurs.

Code : C++ (45000 lignes de code) outils annexes en Python et R

Outils utilisés : gforge, svn, git, autotools, gdb, ddt, memcheck, massif, gprof, cppunit, gtest, intégration continue

Site web : <http://www.aevol.fr/>

#### Contributions

Pilotage de l'équipe de développement depuis 2008

Encadrement de 3 stagiaires et participation à l'encadrement de 4 autres

Accompagnement d'un Ingénieur Confirmé INRIA depuis septembre 2014

Refonte complète de la base de code ayant conduit au gain d'un facteur 10 dans les performances temporelles du simulateur. Ce gain en performances a notamment permis d'ouvrir de nouvelles perspectives de recherche jusqu'alors rendues impossibles par les temps de calcul prohibitifs du simulateur.

Ajout de nombreuses fonctionnalités

Prototypage d'une version distribuée du simulateur (MPI)

Mise en place d'une collaboration étroite avec l'unité INSERM 1001 (co-développement du logiciel)

Structuration de la communauté de développement autour de la forge INRIA avec notamment la mise en place d'un gestionnaire de version (svn puis Git) et de listes de discussions thématiques dédiées

Mise en place de tests unitaires et d'intégration

Mise en place d'un processus d'intégration continue

Instauration d'un guide de style afin d'uniformiser les pratiques de tous

Packaging : Paquet deb disponible sur les dépôts officiels (29 installations par ce biais connues à ce jour), paquet RPM en cours

Mise en place d'un site web complet à l'attention des utilisateurs

Standardisation et consolidation du processus de build et d'installation

Mise à disposition d'un ensemble de scripts facilitant la mise en production sur des clusters de calcul

Dépôt APP

11 publications, 3 résumés, 3 communications, 3 posters, 2 séminaires invité

Organisation de 2 tutoriels dans les conférences référence du domaine

Participation au projet européen EvoEvo (<http://www.evoevo.eu/>) et à différents projets de moindre ampleur (Finovi, PEPS, PEPiI, ...)

Contribution au développement d'un prototype de démonstrateur sous la forme d'un « serious game »

## FluoBacTracker

### Contexte personnel

Développeur à 40 % pendant 2 ans (dans le cadre d'une Action de Développement Technologique INRIA)

### Description et contexte du projet

FluoBacTracker est un outil de segmentation et de suivi de cellules dans des films de vidéo-microscopie de colonies bactériennes. Il combine au sein d'une interface graphique intuitive une implémentation performante d'algorithmes de traitement d'images de dernière génération et d'une méthode novatrice de segmentation et de tracking en une passe.

Pilotage du projet en collaboration avec les partenaires scientifiques : équipe « traitement d'images », laboratoire MAP5 (Université Paris Descartes) et équipe « réparation de l'ADN, réponse aux stress et vieillissement », unité INSERM 1001.

Équipe de développement constituée d'un Ingénieur Jeune Diplômé (IJD) INRIA et de moi-même.

Code : Java (43000 lignes de code), plugin du logiciel d'analyse d'images ImageJ

Outils utilisés : NetBeans, maven, gforge, git, NetBeans profiler, Hprof, JUnit

Site web : <http://fluobactracker.inrialpes.fr/index.php>

### Contributions

Encadrement d'un Ingénieur Jeune Diplômé (24 mois)

Analyse des besoins et de l'existant

Conception et mise en œuvre d'une interface graphique intuitive

Implémentation des algorithmes de traitement d'images

Mise en œuvre de la méthode de segmentation et de suivi en une passe CellST

Conception et mise en œuvre d'une architecture multi-tâches hiérarchisée permettant de répondre aux besoins de vitesse d'exécution tout en évitant la saturation de la mémoire

Conception et mise en œuvre d'un système d'optimisation et de propagation des apports d'informations de la part de l'utilisateur

Contribution au choix de la licence et au mode de distribution en relation avec le service STIP

Dépôt APP

## Kissplice

### Contexte personnel

Développeur à 40 % pendant environ 1 an (dans le cadre d'une Action de Développement Technologique INRIA)

Suivi et interventions ponctuelles depuis début 2014

### Description et contexte du projet

Kissplice est un outil d'analyse de données RNA-seq qui permet la détection de différents types de polymorphismes dans les données que ce soit avec ou sans génome ou transcriptome de référence. C'est un assembleur local exact de transcriptome qui permet d'identifier des SNPs, des indels, des événements d'épissage alternatif ainsi que d'autres types de polymorphismes. Kissplice peut gérer un nombre arbitraire de conditions biologiques et quantifiera chaque variant dans chaque condition.

Équipe de développement bi-localisée laboratoire LBBE (4 personnes) et antenne INRIA Lyon la Doua (1 personne)

Code : modules C++ indépendants orchestrés avec Python (25000 lignes de code)

Outils utilisés : gforge, svn, CMake, CPack, CTest, intégration continue, gdb, memcheck, massif, gprof

Site web : <http://kissplice.prabi.fr/>

### Contributions

Suivi de 2 Ingénieurs Jeunes Diplômés (12 mois + 12 mois)

Structuration de la communauté de développement autour de la forge INRIA avec notamment la mise en œuvre de branches svn et la mise en place d'un bugtracker et de listes de discussions

Instauration d'un guide de style afin d'uniformiser les pratiques de tous

Refactoring de code ayant permis d'accroître la modularité du code

Documentation (création de pages de man, enrichissement du guide utilisateur et développeur)

Mise en place d'un site web complet à l'attention des utilisateurs

Standardisation et consolidation du processus de build et d'installation (Cmake)

Mise en place d'un processus d'intégration continue avec tests automatiques et suivi des anomalies

Packaging : Paquet deb disponible sur les dépôts officiels debian (28 installations par ce biais connues à ce jour), paquet rpm bientôt disponible dans les dépôts officiels fedora

Dépôt APP

## SiMuScale

### Contexte personnel

Développeur à 40 % depuis Octobre 2014 (dans le cadre d'une Action de Développement Technologique INRIA).

### Description et contexte du projet

SiMuScale est une plateforme de simulation qui modélise à la fois des processus intra- et inter-cellulaires à des échelles de temps et selon des formalismes différents. Son architecture modulaire en permet une extension simple par l'ajout de nouveaux types de modélisation intra-cellulaire. La distribution de SiMuScale est prévue sous la forme d'un SaaS afin d'en faciliter l'adoption par des utilisateurs non informaticiens (biologistes).

Unique développeur du projet à partir d'un prototype de recherche

Code : noyau C++ (5000 lignes de code),

serveur de démo et interface : Ruby + Sinatra, Javascript + d3js, communications par websockets (1000 lignes de code)

Outils utilisés : C-Lion, JSBin, gforge, git, CMake, CPack, CTest, gtest, gdb, outils de développement Chromium

### Contributions

Conception et prototypage d'un portail web utilisable comme un démonstrateur puis pour la distribution SaaS du logiciel. (Serveur en Ruby + Sinatra, Client JavaScript, communication par websockets).

Refactoring complet de la base de code ayant permis de mettre en évidence et de corriger plusieurs bugs critiques et de rendre le code modulaire.

Mise en œuvre de design patterns pour limiter les interdépendances et ainsi rendre le code plus facilement extensible.

Mise en place de tests unitaires (gtest) et d'intégration (CTest).

Documentation technique.

### Activités complémentaires

Depuis ma prise de fonction en tant qu'ingénieur de recherche et développement dans le Service d'Expérimentation et de Développement (SED) d'INRIA Grenoble Rhône-Alpes, j'assume en plus de mes fonctions principales, un rôle de liaison entre les services généraux (SG) et le service informatique du centre (SIC) INRIA d'une part, et les acteurs locaux INSAVALOR, INSA, CISR d'autre part, en particulier pour les problématiques liées au réseau. J'assume également diverses fonctions de support de proximité aux usagers de l'antenne de Lyon la Doua :

- Dimensionnement, achat, installation et administration de serveurs de calculs

- Dimensionnement et administration d'un serveur de sauvegardes

- Implémentation d'une solution de sauvegardes incrémentale ad-hoc

- Installation et suivi des équipements de visioconférences

- Aide à l'établissement de devis pour les équipes

- Support informatique et réseau

- Support d'ordre général

- Suivi de différents dossiers concernant principalement les équipements de l'antenne

En 2014, j'ai pris en charge, en collaboration avec Caroline Lothe, assistante de l'antenne, et en relation avec les différents services et acteurs cités ci-dessus, le suivi du chantier d'aménagement des nouveaux locaux de l'antenne. À ce titre, j'ai été amené à dimensionner l'infrastructure réseau et la salle serveur, à décider de l'emplacement de toutes les prises du plateau, à établir des devis avec divers intervenants extérieurs (storiste, cuisiniste, ...) et à assurer le suivi des prestations correspondantes.

J'ai également piloté une refonte complète du plan d'adressage de l'antenne devenue nécessaire en raison d'une pénurie d'adresses IP, de l'analyse des besoins spécifiques à chaque équipe-projet au suivi de la mise en œuvre de la solution proposée.

Enfin, j'ai récemment été élu au comité de gestion locale de l'AGOS (Comité d'Entreprise) et nommé référent pour les sites distants.