

# Table des matières

<b>1 L'équipe Statistique et Algorithmique pour la Biologie (SaAB)</b>	<b>1</b>
I – Bilan . . . . .	1
1- Présentation de l'équipe . . . . .	1
2- Produits et activités de la recherche de l'équipe . . . . .	3
II – Projet et stratégie à cinq ans . . . . .	7
1- Analyse SWOT . . . . .	7
2- Structuration, effectifs et orientations scientifiques . . . . .	8
III – Annexes . . . . .	10
<b>2 Annexe 4 - Produits et activités de la recherche de l'équipe Statistique et Algorithmique pour la Biologie (SaAB)</b>	<b>11</b>
I – Production de connaissances et activités concourant au rayonnement et à l'attractivité scientifique de l'équipe . . . . .	11
1- Journaux / Revues . . . . .	11
2- Ouvrages . . . . .	17
3- Production dans des colloques / congrès, séminaires de recherche . . . . .	19
4- Produits et outils informatiques . . . . .	29
5- Productions des personnels recrutés au cours des cinq dernières années réalisées dans leurs anciennes unités . . . . .	29
6- Activités éditoriales . . . . .	30
7- Activités d'évaluation . . . . .	30
8- Post-doctorants et chercheurs accueillis . . . . .	33
9- Indices de reconnaissance . . . . .	33
II – Interaction de l'équipe avec l'environnement non académique, impacts sur l'économie, la société, la culture, la santé . . . . .	34
1- Activités d'expertise scientifique . . . . .	34
2- Produits destinés au grand public . . . . .	35
III – Implication de l'équipe dans la formation par la recherche . . . . .	35
1- Formation . . . . .	35



# Chapitre 1

## L'équipe Statistique et Algorithmique pour la Biologie (SaAB)

### I – BILAN

*On rédigera ce document en complétant les rubriques ci-dessous, suivant le plan indiqué. Les parties du texte en italique bleu seront supprimées dans le document rédigé (via l'option globale 'final')*

#### 1- Présentation de l'équipe

##### Introduction

###### *historique et localisation*

Entièrement hébergée à l'intérieur de l'unité MIAT, l'équipe SaAB (Statistics and Algorithmic for Biology) a pour domaine d'application principal la biologie, avec des approches bioinformatiques et biostatistiques. Initialement focalisée sur des problématiques de génétique (cartographie génétique, recherche de QTL) via des approches de statistique inférentielle, l'équipe a, dès les années 1990, consacré une part croissante de son activité à l'analyse de séquences génomiques (annotation de gènes de protéines et d'ARN en particulier) via des approches informatiques intégrant analyse du texte, optimisation/raisonnement automatique et *machine learning* (en particulier les modèles graphiques discrets, stochastiques ou non). Depuis une dizaine d'années, des niveaux d'échelle supérieurs et inférieurs ont été abordés, d'une part, via une activité croissante sur des problématiques d'intégration de données hétérogènes et d'inférence de réseaux de régulation et d'autre part via une activité en conception algorithmique de protéines.

##### Effectifs et moyens

*En s'appuyant sur les deux fichiers Excel « Données du contrat en cours » (onglets 3, 6) et « Données du prochain contrat » (onglet 2), que l'on aura pris soin de renseigner, on commentera ici l'évolution des effectifs et des moyens financiers de l'unité (équipe / thème). L'équipe SaAB est composée de 8 membres permanents (3 DR dont une à 25% sur la période, 3 CR, 1 IR et 1 IE à 50% depuis cette année). Elle a connu une évolution importantes de ses effectifs ces dernières années avec :*

- le départ d'une IR en bioinformatique en 2017 (mutation sur une autre unité MIA)
- le départ d'un IE en développement en 2018 (mise en disponibilité)
- le départ d'un CR en 2018 (mutation sur une autre unité MIA)
- l'arrivée d'une CR en 2018 (recrutement)
- la promotion d'une CR en DR en 2019

Au niveau des personnels non permanents, sur la période, 15 doctorants ont été rattachés (8 thèses soutenues et une autre prévue fin septembre) à l'équipe ainsi que 6 post-doctorants. Actuellement, l'équipe compte 6 doctorants et 2 post-doctorantes.

## Politique scientifique

*On précisera les missions, objectifs scientifiques, stratégie de l'unité de recherche (équipe / thème) pour le contrat en cours.*

*On détaillera les actions entreprises pour mettre en œuvre les recommandations de la précédente évaluation.*

*On précisera le profil d'activités de l'unité de recherche : comment se répartissent globalement ses activités entre production de connaissance, activités de valorisation et de transfert, appui à la communauté, activités de formation par la recherche (hors service d'enseignement). Ce profil d'activités doit permettre en particulier aux unités fortement orientées vers la recherche finalisée (translationnelle, clinique) de faire apparaître leur spécificité.*

L'équipe a pour objectif de développer et de mettre à disposition des biologistes des méthodes mathématiques, statistiques et informatiques permettant de contribuer à la compréhension du vivant, avec des applications en particulier à la biologie moléculaire et à la biologie de synthèse. D'un point de vue applicatif, l'équipe s'intéresse à la localisation et l'identification d'éléments fonctionnels dans les génomes (assemblage, annotation automatique), à l'intégration de données omiques multi-échelles et à la conception de protéines. Ces recherches s'appuient sur le développement d'outils issus de l'algorithmique du texte, de l'optimisation discrète et du *machine learning*. Pour réaliser ces objectifs, l'équipe s'appuie, pour les parties finalisées, sur ses liens solides avec de nombreux laboratoires de recherche en biologie, dont, pour les plus proches, le Laboratoire des Interactions Plantes/Micro-organismes (LIPM/INRA), celui de Génétique Animale (GenPhySE/INRA) et Toulouse Biotechnology Institute (TBI/INSA). Elle a également de nombreuses collaborations avec des laboratoires de recherche en mathématiques et informatique pour les aspects méthodologiques, notamment localement avec l'Institut de Mathématiques de Toulouse (en particulier équipe projet BioMaths), l'Institut de Recherche en Informatique de Toulouse et le Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS/CNRS). Enfin, elle participe à l'animation de trois réseaux méthodologiques du département MIA/INRA : *Algorithmic Issues for Inference in Graphical Models* (AIGM), *Inférence de réseaux (biologiques)* (NETBIO) et *Optimisation pour les Sciences de la Vie* (Optim, 2015–2017), pour lesquels un séminaire commun a été organisé en 2016 à Toulouse (Cartable 2016).

L'équipe a une activité principalement orientée sur la production de connaissances (environ 60% de son temps d'activité totale) avec une préoccupation marquée pour la valorisation et le transfert (environ 11% de son temps d'activité, production de logiciels et packages R, contrat pré-industriel, ...). Elle investit également de manière importante dans l'appui à la communauté (environ 18% des activités, organisation de conférences, useR!2019, CP2016 et 2019, comités HCERES, CNU 26, bureau de sociétés savantes, comités d'édition, ...). Enfin, elle s'implique nettement dans la formation par la recherche (environ 11% de ses activités, participation à des masters, écoles d'été, encadrements d'étudiants dont 15 thèses sur la période).

Le rapport précédent relevait les faiblesses suivantes :

- **faible nombre de doctorants** : le nombre de doctorants pour le dernier quadriennal était 12 (dont 5 thèses soutenues) et il est, pour ce quadriennal passé à 15 (dont 9 thèses soutenues d'ici fin septembre). Un effort a été réalisé pour accroître la capacité d'encadrement (2 HDR soutenues pour cette période d'évaluation) et pour rechercher des sources de financement de thèse variées (contrat région, institut de convergences #DigitAG, bourse INRA/Inria, ...);
- **manque d'ouverture au niveau local** : le projet de Pôle Informatique et Mathématiques pour les AgroBioSciences (IMABS) a pour objectif d'accroître la visibilité du centre INRA sur la thématique Mathématiques et Informatique : les actions actuelles se sont concentrées sur l'organisation de séminaires transversaux sur les thématiques des maths/info. Ce projet

prendra de l'ampleur au moment de l'ouverture prévue en 2020 d'un nouveau bâtiment destiné à accueillir l'unité, avec, notamment, une volonté affichée de permettre l'accueil de chercheurs ou d'enseignants chercheurs.

L'ouverture au niveau local a été également soutenue par des actions concrètes de l'équipe en direction du tissu local : participation à l'Institut 3IA, nouvelles collaborations matérialisées par des projets de recherche financés avec le LAAS et l'IMT, visibilité croissante de la journée annuelle Bioinfo/Biostat, . . .

- **moins visibilité internationale sur certaines thématiques** : la visibilité internationale des activités de l'équipe est confirmée par l'organisation des conférences internationales CP 2016, CP 2019 (Program Chairs) et useR! 2019. Elle est soutenue par un projet Européen (programme Infect-ERA) sur la thématique de l'analyse de séquences et de l'évolution d'une bactérie pathogène et un autre en cours de montage (programme H2020) sur la thématique de la biologie prédictive (voir partie projet à 5 ans), et plusieurs visites dans des laboratoires étrangers (Université de Monash en Australie, Institute for Protein Design à Seattle et Mayo lab. au CalTech, Institute of Transformative Bio-Molecules (ITbM) et Département of Physics de l'Université de Nagoya, et Institute for Chemical Reaction Design and Discovery (ICReDD) et Integrative Mathematics de l'Université d'Hokkaido). L'équipe a aussi investi une nouvelle thématique (conception de protéines) sur laquelle elle a rapidement acquis une visibilité internationale (voir partie bilan) ;
- **risque de dispersion thématique** : l'équipe développe des recherches méthodologiques dans trois directions méthodologiques principales (algorithmique du texte, *machine learning* et optimisation discrète) et s'est efforcée de créer des liens entre ses diverses thématiques (voir partie projet à 5 ans). En particulier, le recrutement d'une nouvelle CR crée une dynamique de collaborations nouvelles entre bioinformatique du texte et *machine learning* ainsi qu'en optimisation convexe.

## 2- Produits et activités de la recherche de l'équipe

### Bilan scientifique

*L'unité (équipe / thème) appréciera sous forme d'un bilan global, son activité scientifique en s'appuyant sur les données chiffrées de l'annexe 4 ainsi que sur l'onglet 4 du fichier Excel « Données du contrat en cours ».*

**Bioinformatique : assemblage et annotation de génomes** Parmi les données « omiques » se trouvent celles issues du séquençage de l'ADN génomique, dont le traitement continue à poser des problèmes algorithmiques ardues. La taille des données, la complexité des méthodes à appliquer, la difficulté de définir correctement le problème sont autant de verrous techniques. L'équipe a été tout d'abord impliquée sur l'assemblage d'éléments transposables : l'article [47] expose une méthode originale permettant de résoudre ce problème, ainsi qu'un outil, TEDNA. Celui-ci a d'ailleurs été largement utilisé par la communauté, notamment dans [95, 247]. Forte de cette expérience, l'équipe s'est également investie dans le projet SeqOccIn (Région/FEDER), et notamment sur la partie assemblage de génomes.

Par ailleurs, l'équipe a longtemps eu une activité importante dans le développement et la mise en œuvre de méthodes de prédiction de gènes intégratives et automatiques pour les génomes procaryotes et eucaryotes. Ces méthodes avaient été implémentées dans une série de logiciels open source exploitant des variantes discriminatives des modèles de chaînes de Markov cachées de type *Linear Semi Conditional Random Field* et les efforts des dernières années ont porté sur la rationalisation de ces outils en un seul logiciel **EuGene** [40, 292]. Ces outils ont été mis en œuvre sur des génomes très complexes [138, 145].

**Analyse de données omiques** Une partie des activités scientifiques de l'équipe est tournée vers l'analyse de données omiques, dans des collaborations proches avec des biologistes d'autres unités ou d'autres instituts.

Les données transcriptomiques sont parmi les données omiques les plus fréquemment acquises. Nous avons développé une méthode rapide, `mmquant` [193], qui permet de prendre en compte les régions dupliquées et qui a été étendue aux petits ARNs [108] (et package Bioconductor `srnadiff` qui détecte les petits ARNs différentiellement exprimés sans utiliser d'annotation de référence). Des études appliquées sur les petits ARNs ont également été réalisées dans le cadre du projet ANR EpiRNAseIII, où nous avons analysé les données de séquençages sur plusieurs études différentes [86, 153].

L'unité a également étudié le problème de la recherche d'interactions via des approches d'inférence de réseaux, dans le cadre de données provenant d'expériences multiples [46], dans le cas de données multi-échelles [512], en présence d'expériences appariées avec des individus manquants [218], ou en utilisant des informations de co-localisation spatiale des gènes [222].

Enfin, l'intégration de données consiste à pouvoir combiner des données provenant de plusieurs sources différentes et complémentaires (plusieurs types de données omiques, situées à divers niveaux de l'échelle du vivant, plusieurs omiques obtenues dans des conditions différentes, ...). D'un point de vue appliqué, dans le cadre de l'ANR SusOStress, les articles [84, 173] ont abordé la question de l'intégration de données transcriptomiques longitudinales pour une meilleure compréhension des mécanismes de réponse au stress chez le porc. Plus récemment, plusieurs membres de l'unité (et de l'équipe) ont participé au projet FAANG/Fr-Agencode sur l'annotation fonctionnelle des animaux d'élevage (article en cours de révision dans *Genome Biology*). Enfin, une méthode générique d'intégration de données, utilisant une approche multi-noyaux dans un cadre exploratoire, a été publiée [221] et implémentée dans le package CRAN `mixKernel`.

En complément des approches exploratoires, les approches prédictives en biologie permettent de combiner les informations obtenues par le séquençage haut débit pour construire des modèles prédictifs et explicatifs d'un phénomène biologique d'intérêt. Dans [463], une approche par PLS (parcimonieuse) a été utilisée pour extraire des biomarqueurs pertinents du métabolome pour expliquer la réponse à la chaleur chez le porc et dans [254], nous avons développé une méthode basée sur une régression linéaire parcimonieuse pour la quantification automatique de métabolites dans des spectres obtenus par la méthode de résonance magnétique nucléaire  $^1\text{H}$ , méthode qui a été mise à disposition dans le package Bioconductor `ASICS`.

**Apprentissage** Des travaux méthodologiques en apprentissage statistique et fouille de données sont également menés par l'équipe. Ces activités concernent les approches non supervisées (exploration, visualisation, ...) et supervisées (prédiction, ...) et ont donné lieu à deux chapitres de référence sur l'apprentissage connexionniste [289] et sur l'apprentissage de données massives [290] (qui reprend en partie des résultats de l'article [159]).

L'équipe a développé plusieurs approches pour l'analyse exploratoire avec cartes auto-organisatrices (dites « cartes de Kohonen »), qui combinent visualisation et classification non supervisée. En particulier, [73, 74, 168, 439] développent des approches de cartes auto-organisatrices pour des données relationnelles et volumineuses implémentées dans le package CRAN `SOMbrero`.

L'équipe a aussi investi la thématique des données manquantes, essentielle en première analyse des données pour limiter les biais qui y sont associés. Un article de revue sur le sujet du traitement des données manquantes a été publié [271] et a servi de base à la création de la « CRAN Task View » sur les données manquantes.

L'équipe a également développé des méthodes d'apprentissage supervisé pour la prédiction. En particulier, elle porte des compétences en analyse de données fonctionnelles, c'est-à-dire, de données décrites par des fonctions (des courbes) qui sont généralement issues de l'enregistrement continu de phénomènes physiques. En particulier, dans [62, 260] des approches sont proposées par

régression inverse pour la calibration et pour la sélection d'intervalles pertinents pour la prédiction (package CRAN SISIR).

**Réseaux de fonctions de coûts, satisfiabilité et programmation par contraintes** Les Réseaux de Fonctions de Coûts (RFC) offrent un cadre générique, issu du raisonnement automatique en IA, pour exprimer des problèmes d'optimisation combinatoire dans les modèles graphiques stochastiques ou non. Un premier travail sur la période a porté sur l'amélioration des minorants et majorants produits au cours du temps par une recherche arborescente exacte avec ou sans exploitation de la structure du problème [345]. Un second travail dans le cadre d'un projet ANR Ficolofa et suite à une collaboration avec l'Université de Hong-Kong [325] a été l'ajout de fonctions de coûts globales décomposables dans le cadre des RFC, augmentant ainsi leur expressivité [97]. L'efficacité de cette approche reste néanmoins souvent problématique dès que l'on augmente la taille des problèmes. Une proposition de fonction globale dédiée (atmost one) issue de la Recherche Opérationnelle a permis d'améliorer sensiblement les résultats des RFC pour le problème d'enchères combinatoires [429]. Ce travail continue au sein du projet ANR Demograph par une fonction globale d'acyclicité dédiée à l'apprentissage de la structure de réseaux bayésiens. Suite à des expérimentations de cadres de résolution variés pour le problème de conception de protéines [1], une large série de comparaison d'outils de résolution a été menée sur une collection de benchmarks issus de plusieurs communautés scientifiques mettant en valeur le logiciel `toulbar2` développé en majeure partie dans l'équipe [112]. Lorsque le problème devient trop difficile pour une méthode exacte, les méta-heuristiques sont une bonne alternative. Avec l'aide de l'Université de Caen, nous avons proposé une recherche itérative à large voisinage, incluant une version parallèle et fournissant de très bons majorants sur nos benchmarks dont le problème de conception de protéines [444]. Motivés par la conception de protéines et le calcul de fonctions de partition, nous avons proposé une méthode originale de comptage pondéré à approximation garantie [381, 419] qui n'exploite pas la structure du problème mais la concentration des masses de probabilité dans la distribution. Par ailleurs, nous avons amélioré l'état de l'art en programmation par contraintes sur le problème de coloration de graphe [287] (CP 2018 Best Paper Award) et en satisfiabilité sur le problème d'énumération de sous-ensembles incohérents minimaux dans des formules logiques [277] au travers de collaborations avec le LAAS et l'Université de Toronto.

**Conception algorithmique de protéines** Les protéines jouent des rôles centraux dans tous les organismes vivants. L'ingénierie des protéines a pour but d'exploiter la versatilité de ces molécules pour répondre à de nouveaux besoins dans des domaines incluant la santé, les biotechnologies, la chimie verte, *etc.* L'approche rationnelle de la conception de protéines part d'une fonction cible et de la modélisation moléculaire de la protéine visée pour construire une séquence d'acides aminés qui devrait se replier selon ce squelette cible, porteur de la fonction visée [190].

Au cœur de cette problématique, l'optimisation de l'énergie d'une protéine au travers de l'ajustement de sa séquence et de la conformation de ses chaînes latérales définit un problème NP-difficile qui se formule naturellement sous forme d'un réseau de fonctions de coût [1]. L'efficacité de `toulbar2` sur ces problèmes a permis de mettre en évidence les limitations d'une implémentation d'un recuit simulé (celui de `Rosetta`), très rapidement incapable d'explorer les régions de très basse énergie [87]. Nous avons commencé à spécialiser `toulbar2` avec des heuristiques de branchement dédiées [1], des méthodes pour l'énumération de séquences et l'exploration de paysages énergétiques [130, 483], pour l'estimation d'affinité via une intégration discrète [235], pour les problèmes les plus difficiles, via un algorithme de recherche locale à voisinage variable [243]. Enfin, nous avons commencé à relâcher l'hypothèse de squelette rigide [266]. Ces travaux ont profité des collaborations avec nos collègues de TBI, du LAAS, de l'École Polytechnique, de l'Université KU Leuven, de la Duke University et d'un ancien post-doctorant (MdC à l'Université de Toulouse), allant jusqu'à la validation expérimentale [258].

## Faits marquants

### logos/logo\_HCERES . j Fait Marquant SaAB 1 : useR! 2019

L'équipe a porté l'organisation de la conférence internationale useR! qui a eu lieu à Toulouse du 9 au 12 juillet 2019, à l'organisation de laquelle, l'équipe MAD, l'unité InTheReSE, l'unité ORD, l'Institut de Mathématiques de Toulouse et Toulouse School of Economics ont aussi participé. Cette conférence est la conférence annuelle de la fondation R, langage de programmation devenu incontournable en statistique, sciences des données et bioinformatique. Elle a rassemblé plus de 1200 personnes (dont environ 20% de Français), dont la moitié issue du milieu académique et l'autre moitié issue du secteur privé, d'origines thématiques très variées (statistique, développement, *machine learning* & IA, bioinformatique, ...). Au niveau national, elle a été soutenue par la Société Française de Statistique (organisation d'un Datathon notamment) et par quatre GDR (BIM, MaDICS, MASCOT NUM et Statistique & Santé qui ont organisé des prix posters). Voir : <http://user2019.r-project.org>, [https://twitter.com/User2019\\_Conf](https://twitter.com/User2019_Conf) et <https://twitter.com/hashtag/user2019?src=hash>.

### logos/logo\_HCERES . j Fait Marquant SaAB 2 : Interaction 3D et gènes co-exprimés dans le muscle foetal porcin

Dans le secteur de l'élevage porcin, la sélection a été associée à une mortalité plus élevée des nouveau-nés. Le muscle foetal squelettique, de par son implication dans les fonctions motrices et de thermorégulation, est essentiel à la survie du porcelet. Il est par conséquent important de comprendre les processus biologiques liés à la mise en place de la maturité musculaire foetale, pour améliorer la survie à la naissance des porcelets. Une nouvelle méthode statistique a permis d'obtenir un réseau de gènes co-exprimés en prenant en compte l'information de co-localisation des gènes dans le noyau des cellules musculaires. Le réseau obtenu suggère de nouvelles associations entre gènes. Par exemple, nous identifions ainsi des interactions inconnues jusqu'à présent entre des gènes codant un facteur de croissance et une protéine importante du cytosquelette musculaire. Ces associations ont été validées par microscopie. Plus ici : <http://2025.inra.fr/openscience/Billets/FM2018-GA-muscle-foetal-porcine-contribution-a-3Perf-1-et-OpenScience-3>.

### logos/logo\_HCERES . j Fait Marquant SaAB 3 : Conception de nouvelles protéines

En s'appuyant sur l'outil *toulbar2* développé depuis de nombreuses années pour le raisonnement automatique en IA (programmation par contraintes), l'unité a pu mettre à disposition des biologistes des méthodes de conception de protéines originales, permettant en particulier une minimisation de l'énergie avec garantie déterministe à distance finie. Sur ces problèmes NP-difficiles, *toulbar2* a pu empiriquement démontrer les limitations de l'implémentation réputée d'un recuit simulé dédié dans *Rosetta*, puis a été mis en œuvre en collaboration avec une équipe de biochimistes (KUL, Belgique) pour concevoir une protéine symétrique de type  $\beta$ -propeller (appelée *Ika*), capable de s'auto-assembler. Ce travail a, en particulier, fait l'objet d'un article dans le journal « La recherche ». Depuis, *toulbar2* a été adopté comme moteur d'optimisation interne du logiciel de *Protein Design* OSPREY (Duke University, North Carolina, USA) et les équipes développant les logiciels *Explorer* (École Polytechnique) et *Rosetta* (University of Washington, Seattle, USA) ont étendu leurs outils pour les rendre capables de générer des modèles directement résolus par *toulbar2*. Voir : <http://2025.inra.fr/biores/Billets/Le-departement-MIA-contribue-a-1-objectif-BioRes-2>.



## II – PROJET ET STRATÉGIE À CINQ ANS

*L'unité, (l'équipe / le thème) réalisera une analyse SWOT puis présentera la structuration envisagée pour le projet, ses effectifs et ses orientations scientifiques.*

### 1- Analyse SWOT

- En se plaçant dans la perspective de son projet scientifique à cinq ans, l'unité (équipe / thème) dressera la liste :*
- de ses points forts ;
  - de ses points à améliorer ;
  - des possibilités offertes par le contexte / l'environnement dans lequel elle se trouve ;
  - des risques liés à ce contexte / cet environnement.

Points forts	Points à améliorer
<ul style="list-style-type: none"> <li>— L'équipe a une production scientifique importante (publications dans des journaux et conférences sélectifs, voire de premier plan, plusieurs logiciels de notoriété internationale, ...).</li> <li>— L'équipe jouit d'une visibilité internationale en bioinformatique, biostatistique et intelligence artificielle qui se manifeste par l'implication de ses membres dans des comités de programme de plusieurs conférences internationales (PC / Senior PC / Area Chair / Program Chair CP 2019), par l'organisation de conférences nationales et internationales (CP 2016, JFRB 2018, useR! 2019) et d'une école d'été de bioinformatique à Roscoff). Localement, elle assure la responsabilité scientifique des plateformes bioinformatique et biostatistique.</li> <li>— L'équipe a augmenté son nombre de doctorants qu'elle encadre par l'obtention de financements variés (bourses Région, du programme INRA/Inria, de l'institut de convergence #DigitAG, ...) et par l'augmentation de ses capacités d'encadrement (2 HdR soutenues en 2014 et 2016). Au total, avec une promotion DR en son sein en 2019, l'équipe comporte 4 DR ou CR/HdR.</li> <li>— L'équipe a une forte activité de collaboration avec les biologistes des laboratoires environnants (GenPhySE/INRA, Toxalim/INRA, INSERM, LIPM/INRA, TBI/INSA, ...).</li> </ul>	<ul style="list-style-type: none"> <li>— L'équipe n'étant composée que de chercheurs et d'ingénieurs (pas d'enseignant chercheur), elle est loin des étudiants (l'éloignement est structurel et géographique), ce qui pose des difficultés supplémentaires pour recruter des doctorants.</li> <li>— L'équipe n'a qu'une seule ingénieure en biostatistique qui partira à la retraite dans les 4 ans et plus d'ingénieur d'étude en développement.</li> <li>— Les champs d'expertise de l'équipe sont très diversifiés, avec un risque de dispersion scientifique.</li> <li>— Les membres de l'équipe sont impliqués de manière importante dans l'animation (organisation de conférences, activités d'expertise, implication forte dans les diverses instances et dans l'animation de l'INRA, implication dans l'enseignement nécessaire pour le recrutement d'étudiants de thèses ou stage, ...), ce qui présente également un risque concernant une trop grande dispersion des activités.</li> </ul>

Possibilités offertes par le contexte ou l'environnement	Risques liés au contexte ou à l'environnement
<ul style="list-style-type: none"> <li>— La construction d'un nouveau bâtiment pour l'unité, soutenue par la création du pôle IMABS, élargira les possibilités d'accueil de chercheurs et de collaboration pour l'équipe.</li> <li>— La création d'ANITI (Institut Interdisciplinaire d'Intelligence Artificielle, 3IA), dans lequel des membres de l'équipe sont impliqués, est une opportunité à la fois très importante et risquée pour la dynamique des activités et la cohésion de l'équipe.</li> <li>— La création d'un nouveau métaprogramme INRA sur la « biologie prédictive », qui correspond très bien aux activités scientifiques de l'équipe, est une opportunité pour le soutien de ses thématiques au sein de l'INRA.</li> <li>— L'environnement technologique proche (centre de séquençage, plateforme bio-informatique, accélérateur TWB et plateforme européenne IBISBA, ...) facilite les collaborations finalisées en bioinformatique/biostatistique et en biotechnologie (blanche), sur les thématiques d'intérêt de l'équipe.</li> </ul>	<ul style="list-style-type: none"> <li>— Les thématiques de l'analyse des séquences et de l'algorithmique du texte sont désormais à renforcer (avec le départ d'un ingénieur et la baisse d'activité d'un DR sur le sujet) car les besoins locaux sont importants, avec des difficultés spécifiques aux espèces d'intérêt pour l'INRA (régions répétées, organismes polyploïdes et fortement hétérozygotes) ; par ailleurs, ces thématiques sont peu abordées par les méthodologistes sur Toulouse.</li> <li>— Il existe un risque de dépendance financière et scientifique aux projets financés (par l'ANR, l'Europe, ...) entraînant une dispersion scientifique (par mise en place d'une stratégie opportuniste pour décrocher des projets) et donc une difficulté de gestion de la charge. Cette difficulté est renforcée par la faible taille de l'équipe et son isolement géographique des principaux centres de recherche en mathématiques et informatique à Toulouse.</li> <li>— Le pendant du dynamisme des collaborations avec les biologistes localement est une trop grosse sollicitation de ceux-ci sur les compétences de l'unité, qui excède les capacités de réponse.</li> </ul>

## 2- Structuration, effectifs et orientations scientifiques

*Pour l'unité, (l'équipe / le thème), on précisera le mode de structuration et on présentera les effectifs, les orientations scientifiques, les choix stratégiques, les objectifs scientifiques, les moyens à mobiliser pour atteindre ces objectifs, les partenariats, les nouvelles thématiques scientifiques, etc. Pour rédiger cette partie, on prendra soin de commenter les objectifs et réalisations attendues contenues dans l'onglet 3 du fichier Excel « Données du prochain contrat ».*

Le projet de l'équipe à cinq ans s'appuie sur ses trois axes méthodologiques principaux : algorithmique du texte, *machine learning* pour la biologie prédictive et réseaux de fonctions de coûts. Ces trois axes méthodologiques ont pour vocation à répondre à des verrous importants pour faire progresser la biologie moléculaire et de synthèse. Ils sont soutenus par des projets d'importance, en démarrage ou à venir.

**Algorithmique du texte** Les laboratoires de l'Institut, et notamment dans l'entourage de MIAT, utilisent de plus en plus les méthodes de séquençage pour leurs analyses. De nouvelles techniques et protocoles innovants sont régulièrement mis en place afin de résoudre des verrous scientifiques liés aux thèmes abordés à l'INRA. Chacune de ces techniques génère ensuite des masses très importantes de données, qu'il faut analyser informatiquement par des méthodes *ad hoc* efficaces.

MIAT est proche d'une plateforme de séquençage (GeT-PlaGe) et d'une plateforme de bioinformatique analysant les données de séquences et valorisant les *pipelines*. L'unité se trouve donc

idéalement placée pour concevoir ces méthodes bioinformatiques.

Consciente de ces enjeux, l'équipe a déjà conçus plusieurs outils utilisés par la communauté dont *EuGene*, *mmquant*, *srnadiff*, *TEDNA*, ...

À titre d'exemple, le projet FEDER SeqOccIn va contribuer, entre autres, à concevoir des méthodes optimisées pour l'assemblage de génomes complets d'organismes d'intérêts agronomiques (animaux et plantes) à destination des laboratoires de recherche et des partenaires privés. Les travaux en cours, menés avec un post-doctorant recruté sur ce projet, cherchent non seulement à proposer des algorithmes optimisés, déployés sur le cluster de calcul de la plateforme bioinformatique, mais également, à adapter les protocoles utilisés par la plateforme de séquençage, comme le Hi-C ou le 10X.

Afin de consolider cette thématique, nous demandons l'ouverture d'un poste de CR. Nous aborderons avec lui les questions clés de l'assemblage de génomes pour les nombreuses espèces d'intérêts agronomiques pour lesquelles ce problème est encore ouvert en tentant de profiter des avancées variées dans le domaine du séquençage et en combinant divers types de données et diverses technologies. Cette thématique sera notamment abordée dans le cadre de notre participation au projet SeqOccIn. L'équipe soutient de manière importante l'ouverture de ce poste pour aider à la réalisation de ce projet méthodologique ambitieux autour de l'assemblage de génomes.

**Machine learning** Les compétences de l'équipe se sont renforcées dans la thématique du *machine learning* avec une ouverture croissante vers l'optimisation convexe. Cette thématique trouve un ancrage naturel dans le nouveau métaprogramme INRA sur la « biologie prédictive » avec pour spécificité de l'équipe la question de l'intégration de données omiques multi-échelles.

Le futur de cette thématique se matérialise par le montage en cours d'un projet H2020 (programme BIOTECH-07-2020 « Multi-omics integration for genotype to phenotype association ») dans lequel l'équipe joue le rôle de porteur, soutenu par la filiale INRA Transfert, et sera déposé en décembre 2019. Le projet se propose de développer des approches nouvelles pour une meilleure exploitation des données issues des techniques de séquençage, notamment pour prédire, ou optimiser un phénotype à partir de données génotypiques en intégrant de l'information additionnelle apportée par d'autres omiques. L'objectif est d'aller au delà de l'état de l'art d'un point de vue méthodologique et de comparer des approches « usuelles » (principalement modèles linéaires ou modèles mixtes) à des approches issues de la statistique non paramétrique (forêts aléatoires par exemple), du *machine learning* (dont méthodes à noyau) ou de l'intelligence artificielle (réseaux de neurones). Le projet sera réalisé en étroite collaboration avec des industriels des secteurs de l'agriculture et de l'industrie alimentaire, qui apporteront des cas d'études concrets sur des espèces très variées (animaux, plantes, micro-organismes). Le consortium actuel comporte une quinzaine de partenaires issus de France, d'Espagne, de Belgique, d'Allemagne, des Pays-Bas et du Danemark.

En outre, l'équipe souhaite dynamiser les interactions entre *machine learning* et algorithmique du texte, ce qui se matérialisera par un nouveau dépôt du projet ANR SAMBA en 2020, en collaboration avec une équipe Inria. Son objectif est d'intégrer les données de séquences et de structures pour mieux annoter/classifier les ARNs non codants et identifier ceux qui jouent un rôle majeur dans les interactions intra/inter-organismes.

**Réseaux de fonctions de coûts et conception algorithmique de protéines** La thématique optimisation exacte sur les réseaux de fonctions de coûts sera maintenue comme un axe fort de l'équipe, avec une ouverture croissante vers l'optimisation continue, en relation avec la thématique précédente. Cette thématique devrait également se déployer au travers d'applications dans le domaine de la conception de protéines. Elle est soutenue par une chaire du projet 3IA ANITI (2 financements de thèse et un post-doc sur 4 ans). Elle sera principalement ciblée sur des aspects méthodologiques génériques mais potentiellement utiles pour la conception de protéines. En pratique, du côté de la conception de protéines, le projet ANR SpaceHex accepté pour 2020

dans le domaine des nanotechnologies devrait nous permettre d'aborder une problématique de design complexe, dont une formulation naturelle est située au deuxième niveau de la hiérarchie polynomiale. La finalité du projet est de concevoir des micro-réacteurs enzymatiques exploitant les unités asymétriques d'un micro-compartiment bactérien en charge de la fixation du carbone atmosphérique. Ce projet complète des projets finalisés en cours ou à venir avec le Centre de Recherche de Cancérologie de Toulouse, le Japon et un collègue Belge (projet PRHC Tournesol soumis).

**Évolution des effectifs** : Les évolutions prévues du personnel dans les prochaines années devront s'adapter au départ prévu de plusieurs personnes : une IE en biostatistique, qui a déjà partiellement ré-orienté ses activités sur la plateforme biostatistique de Toulouse, devrait partir à la retraite durant la prochaine période d'évaluation. Une DR en bioinformatique, déjà à temps partiel sur l'unité, devrait augmenter sa part d'activités consacrées à la gestion du département. Pour renouveler ces compétences et renforcer les besoins en biostatistique / soutien au développement, un poste d'IE a été demandé à la mutation et sera publié en septembre. En complément, l'équipe soutient depuis plusieurs années la demande d'un poste de CR sur la thématique de l'algorithmique du texte.

### III – ANNEXES

#### Annexe 4 : Produits et activités de la recherche

Voir page 11

## Chapitre 2

# Annexe 4 - Produits et activités de la recherche de l'équipe Statistique et Algorithmique pour la Biologie (SaAB)

En cohérence avec les données chiffrées de l'onglet 4 du fichier Excel « Données du contrat en cours », on remplira ce document destiné à l'évaluation du critère 1 du référentiel de l'évaluation « Produits et activités de la recherche », pour l'ensemble de l'unité et pour chaque équipe / thème.

---

### CAMPAGNE D'EVALUATION 2019-2020 VAGUE A

**Nom de l'équipe :** Statistique et Algorithmique pour la Biologie  
**Acronyme :** SaAB  
**Responsable pour le contrat en cours :** Simon de Givry  
**Responsable pour le contrat à venir :** Simon de Givry

*On rédigera cette annexe en complétant les rubriques ci-dessous, suivant le plan indiqué. Les parties du texte en italique et bleu seront supprimées dans le document rédigé.*

*On dressera la liste des produits de la recherche et des activités de recherche en respectant les indications de nombre et de présentation données dans le dossier d'autoévaluation et après avoir consulté le(s) Guide(s) des produits de la recherche et des activités de recherche correspondant au(x) sous-domaine(s) disciplinaire(s) concerné(s). Ces Guides sont disponibles sur le site du Hcéres.*

## I – PRODUCTION DE CONNAISSANCES ET ACTIVITÉS CONCOURANT AU RAYONNEMENT ET À L'ATTRACTIVITÉ SCIENTIFIQUE DE L'ÉQUIPE

*Pour les personnels recrutés au cours des cinq dernières années, les produits et les activités réalisés dans leur unité de recherche d'origine seront mentionnés dans une liste séparée.*

### 1- Journaux / Revues

*Pour chacune des rubriques ci-dessous, on dressera, s'il y a lieu, la liste des produits, qu'on numérottera de 1 à x, dans l'ordre chronologique, en distinguant d'abord les 20 % jugés les plus significatifs, puis les 80 % restants.*

(Dans le cas de publications cosignées par plusieurs personnes, la liste de tous les auteurs sera établie, excepté lorsque le nombre de co-auteur(e)s est supérieur à la quinzaine (dans ce cas, tous les noms des membres de l'unité ou de l'équipe / thème doivent apparaître à leur position dans la liste, avec des pointillés pour remplacer les noms des auteurs extérieurs). Les noms des membres de l'unité de recherche, de l'équipe interne ou du thème seront soulignés dans la liste des co-auteur(e)s.

### Articles scientifiques les plus significatifs (soit 19/ 74)

- [1] D. ALLOUCHE, I. ANDRE, S. BARBE, J. DAVIES, S. DE GIVRY, G. KATSIRELOS, B. O'SULLIVAN, S. PRESTWICH, T. SCHIEX et S. TRAORE. « Computational protein design as an optimization problem ». In : *Artificial Intelligence* 212 (2014), p. 59–79.
- [24] D. LEROUX, A. RAHMANI, S. JASSON, M. VENTELON, F. LOUIS, L. MOREAU et B. MANGIN. « Clusthaplo : a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. » In : *Theoretical and Applied Genetics* 127.4 (2014), p. 921–933.
- [40] E. SALLET, J. GOUZY et T. SCHIEX. « EuGene-PP : a next generation automated annotation pipeline for prokaryotic genomes. » In : *Bioinformatics* 30.18 (2014), p. 2659–2661.
- [47] M. ZYTNICKI, E. AKHUNOV et H. QUESNEVILLE. « Tedna : a transposable element de novo assembler ». In : *Bioinformatics* (on-line) (2014), p. 1–3.
- [63] S. HIGASHI, C. FOURNIER, C. GAUTIER, C. GASPIN et M.-F. SAGOT. « Mirinho : An efficient and general plant and animal pre-miRNA predictor for genomic and deep sequencing data ». In : *BMC Bioinformatics* 16.1 (2015), p. 179.
- [72] E. MONTASTIER, N. VIALANEIX, S. CASPAR-BAUGUIL, P. HLA VATY, E. TVRZICKA, I. GONZALEZ, W. H. M. SARIS, D. LANGIN, M. KUNESOVA et N. VIGUERIE. « System model network for adipose tissue signatures related to weight changes in response to calorie restriction and subsequent weight maintenance ». In : *Plos Computational Biology* 11.1 (2015), p. 22.
- [86] N. SHAMANDI, M. ZYTNICKI, C. CHARBONNEL, E. ELVIRA-MATELOT, A. BOCHNAKIAN, P. COMELLA, A. MALLORY, G. LEPERE, J. SAEZ-VASQUEZ et H. VAUCHERET. « Plants Encode a General siRNA Suppressor That Is Induced and Suppressed by Viruses. » In : *Plos Biology* 13.12 (2015), e1002326.
- [87] D. SIMONCINI, D. ALLOUCHE, S. DE GIVRY, C. DELMAS, S. BARBE et T. SCHIEX. « Guaranteed Discrete Energy Optimization on Large Protein Design Problems ». In : *Journal of Chemical Theory and Computation* 11.12 (2015), p. 5980–5989.
- [95] E.-M. WILLING, V. RAWAT, T. MANDÁKOVÁ, F. MAUMUS, G. V. JAMES, K. J. NORDSTRÖM, C. BECKER, N. WARTHMAN, C. CHICA, B. SZARZYNSKA, M. ZYTNICKI, M. C. ALBANI, C. KIEFER, S. BERGONZI, L. CASTAINGS, J. L. MATEOS, M. C. BERNIS, N. BUJDOSO, T. PIOFCZYK, L. DE LORENZO, C. BARRERO-SICILIA, I. MATEOS, M. PIEDNOËL, J. HAGMANN, R. CHEN-MIN-TAO, R. IGLESIAS-FERNÁNDEZ, S. C. SCHUSTER, C. ALONSO-BLANCO, R. FRANÇOIS, P. CARBONERO et al. « Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation ». In : *Nature Plants* 1.2 (2015), p. 14023–14028.
- [97] D. ALLOUCHE, C. BESSIERE, P. BOIZUMAULT, S. DE GIVRY, P. GUTIERREZ, J. H. M. LEE, K. L. LEUNG, S. LOUDNI, J.-P. METIVIER, T. SCHIEX et Y. WU. « Tractability-preserving transformations of global cost functions ». In : *Artificial Intelligence* 238 (2016), p. 166–189.

- [105] C. CHAINTREUIL, R. RIVALLAN, D. J. BERTIOLI, C. KLOPP, J. GOUZY, B. COURTOIS, P. LELEUX, G. MARTIN, J.-F. RAMI, D. GULLY, H. PARRINELLO, D. SEVERAC, D. PATREL, J. FARDOUX, W. RIBIERE, M. BOURSOT, F. CARTIEAUX, P. CZERNIC, P. RATET, P. MOURNET, E. GIRAUD et J.-F. ARRIGHI. « A gene-based map of the Nod factor-independent *Aeschynomene evenia* sheds new light on the evolution of nodulation and legume genomes ». In : *DNA Research* 23.4 (2016), p. 365–376.
- [112] B. HURLEY, B. O’SULLIVAN, D. ALLOUCHE, G. KATSIRELOS, T. SCHIEX, M. ZYTNICKI et S. DE GIVRY. « Multi-language evaluation of exact solvers in graphical model discrete optimization ». In : *Constraints* 21.3 (2016), p. 413–434.
- [138] H. BADOVIN, J. GOUZY, C. GRASSA, F. MURAT, S. E. STATON, L. COTTRET, C. LELANDAI-BRIERE, G. OWENS, S. CARRERE, B. MAYJONADE, L. LEGRAND, N. GILL, N. C. KANE, J. E. BOWERS, S. HUBNER, A. BELLEC, A. BERARD, H. BERGES, N. BLANCHET, M.-C. BONIFACE, D. BRUNEL, O. CATRICE, N. CHAIDIR, C. CLAUDEL, C. DONNADIEU, T. FARAUT, G. FIEVET, N. HELMSTETTER, M. KING, S. J. KNAPP et al. « The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution ». In : *Nature* 546.7656 (2017), p. 148–152.
- [152] F. CERUTTI, L. MALLET, A. PAINSET, C. HOEDE, A. MOISAN, C. BECAVIN, M. DUVAL, O. DUSSURGET, P. COSSART, C. GASPIN et H. CHIAPELLO. « Unraveling the evolution and coevolution of small regulatory RNAs and coding genes in *Listeria* ». In : *BMC Genomics* 18.882 (2017).
- [221] J. MARIETTE et N. VIALANEIX. « Unsupervised multiple kernel learning for heterogeneous data integration ». In : *Bioinformatics* (2018).
- [235] C. VIRICEL, S. DE GIVRY, T. SCHIEX et S. BARBE. « Cost function network-based design of protein-protein interactions : predicting changes in binding affinity ». In : *Bioinformatics* 34.15 (2018), p. 2581–2589.
- [254] G. LEFORT, L. LIAUBET, C. CANLET, P. TARDIVEL, M.-C. PERE, H. QUESNEL, A. PARIS, N. IANNUCELLI, N. VIALANEIX et R. SERVIEN. « ASICS : an R package for a whole analysis workflow of 1D 1H NMR spectra ». In : *Bioinformatics* (2019).
- [258] H. NOGUCHI, C. ADDY, D. SIMONCINI, S. WOUTERS, B. MYLEMANS, Van MEERVELT, T. SCHIEX, K. Y. J. ZHANG, J. R. H. TAME et A. R. D. VOET. « Computational design of symmetrical eight-bladed ?-propeller proteins ». In : *IUCrJ* 6.1 (2019), p. 46–55.
- [260] V. PICHENY, R. SERVIEN et N. VIALANEIX. « Interpretable sparse SIR for functional data ». In : *Statistics and Computing* 29.2 (2019), p. 255–267.

### Autres articles scientifiques

- [2] J.-M. AZAÏS, C. DELMAS et C.-E. RABIER. « Likelihood ratio test process for quantitative trait locus detection ». In : *Statistics* 48.4 (2014), p. 787–801.
- [13] M. CHAMPION, C. CIERCO-AYROLLES, S. GADAT et M. VIGNES. « Sparse regression and support recovery with L2-Boosting algorithms ». In : *Journal of Statistical Planning and Inference* 155 (2014), p. 19–41.
- [19] J. DAVIES, G. KATSIRELOS, N. NARODYTSKA, T. WALSH et L. XIA. « Complexity of and algorithms for the manipulation of Borda, Nanson’s and Baldwin’s voting rules ». In : *Artificial Intelligence* 217 (2014), p. 20–42.

- [29] G. MARCHAND, V. A. HUYNH-THU, N. C. KANE, S. ARRIBAT, D. VARES, D. RENGEL, S. AUBOURG-BALZERGUE, L. H. RIESEBERG, P. VINCOURT, P. GEURTS, M. VIGNES et N. LANGLADE. « Bridging physiological and evolutionary time-scales in a gene regulatory network ». In : *New Phytologist* 203.2 (2014), p. 685–696.
- [31] V. MÉROT-L'ANTHOËNE, B. MANGIN, F. LEFEBVRE-PAUTIGNY, S. JASSON, M. RIGOREAU, J. HUSSON, C. LAMBOT et D. CROUZILLAT. « Comparison of three QTL detection models on biochemical, sensory, and yield characters in *Coffea canephora* ». In : *Tree Genetics and Genomes* 10.6 (2014), p. 1541–1553.
- [33] J.-C. OGIER, S. PAGES, G. BISCH, H. CHIAPELLO, C. MÉDIGUE, Z. ROUY, C. TEYSSIER, S. VINCENT, P. TAILLIEZ, A. GIVAUDAN et S. GAUDRIAULT. « Attenuated virulence and genomic reductive evolution in the entomopathogenic bacterial symbiont species, *Xenorhabdus poinarii* ». In : *Genome Biology and Evolution* 6 (2014), p. 1495–1513.
- [36] C.-E. RABIER. « On empirical processes for quantitative trait locus mapping under the presence of a selective genotyping and an interference phenomenon ». In : *Journal of Statistical Planning and Inference* 153 (2014), p. 42–55.
- [37] C.-E. RABIER. « On statistical inference for selective genotyping ». In : *Journal of Statistical Planning and Inference* 147 (2014), p. 24–52.
- [38] C.-E. RABIER. « On the asymptotic robustness of the likelihood ratio test in quantitative trait locus detection ». In : *Electronic Journal of Statistics* 8 (2014), p. 2138–2157.
- [41] M. SAN CRISTOBAL, M. P. SANCHEZ, M. MERCAT, F. ROHART, L. LIAUBET, T. TRIBOUT, C. CANLET, N. MULLER, J. MOLINA, N. IANNUCELLI, B. LAURENT, N. VIALANEIX, A. PARIS et D. MILAN. « Le métabolome, un moyen pour trouver de nouveaux biomarqueurs? » In : *Viandes et Produits Carnés* Février (2014), p. 1–5.
- [46] N. VIALANEIX, M. VIGNES, N. VIGUERIE et M. SAN CRISTOBAL. « Inferring networks from multiple samples with consensus LASSO ». In : *Quality Technology and Quantitative Management* 11.1 (2014), p. 39–60.
- [48] A. BAR-HEN, H. JAVAUX et N. VIALANEIX. « Statistical Analysis of the profiles and of the activity of participants of MOOC ». In : *Revue Internationale des Technologies en Pédagogie Universitaire* 12.1-2 (2015), p. 11–22.
- [58] H. CHIAPELLO, L. MALLET, C. GUERIN, G. AGUILETA, J. AMSELEM, T. KROJ, E. ORTEGA-ABBOUD, M.-H. LEBRUN, B. HENRISSAT, A. GENDRAULT, F. RODOLPHE, D. THARREAU et E. FOURNIER. « Deciphering Genome Content and Evolutionary Relationships of Isolates from the Fungus *Magnaporthe oryzae* Attacking Different Host Plants ». In : *Genome Biology and Evolution* 7.10 (2015), p. 2896–2912.
- [60] T. ESQUERRE, A. MOISAN, H. CHIAPELLO, L. ARIKE, R. VILU, C. GASPIN, M. BOUSQUET et L. GIRBAL. « Genome-wide investigation of mRNA lifetime determinants in *Escherichia coli* cells cultured at different growth rates ». In : *BMC Genomics* 16 (2015).
- [62] N. HERNANDEZ, R. J. BISCAY, N. VIALANEIX et I. TALAVERA. « A non parametric approach for calibration with functional data ». In : *Statistica Sinica* 25.4 (2015), p. 1547–1566.
- [65] C. JESTIN, N. BARDOL, M. LODE, P. DUFFE, C. DOMIN, P. VALLÉE, B. MANGIN, M. MANZANARES-DAULEUX et R. DELOURME. « Connected populations for detecting quantitative resistance factors to phoma stem canker in oilseed rape (*Brassica napus* L.) » In : *Molecular Breeding* 35.8 (2015).



- [69] B. MANGIN, F. SANDRON, K. HENRY, B. DEVAUX, G. WILLEMS, P. DEVAUX et E. GOUDMAND. « Breeding patterns and cultivated beets origins by genetic diversity and linkage disequilibrium analyses ». In : *Theoretical and Applied Genetics* 128.11 (2015), p. 2255–2271.
- [73] M. OLTEANU et N. VIALANEIX. « On-line relational and multiple relational SOM ». In : *Neurocomputing* 147 (2015), p. 15–30.
- [74] M. OLTEANU et N. VIALANEIX. « Using SOMbrero for clustering and visualizing graphs ». In : *Journal de la Société Française de Statistique* 156.3 (2015), p. 95–119.
- [78] C.-E. RABIER. « On stochastic processes for quantitative trait locus mapping under selective genotyping ». In : *Statistics* 49.1 (2015), p. 19–34.
- [84] V. SAUTRON, E. MORMEDE, L. GRESS, Y. LIPPI, Y. BILLON, C. LARZUL, L. LIAUBET, N. VIALANEIX et P. MORMEDE. « Time course of the response to ACTH in pig : biological and transcriptomic study ». In : *BMC Genomics* 16 (2015).
- [92] R. URICARU, C. MICHOTEY, H. CHIAPELLO et E. RIVALS. « YOC, A new strategy for pairwise alignment of collinear genomes ». In : *BMC Bioinformatics* 16.1 (2015), p. 111.
- [93] N. VIALANEIX et A. RUIZ-GAZEN. « Beyond model visualization for multidimensional data ». In : *Statistical Analysis and Data Mining* 8.4 (2015), p. 232–239.
- [99] A. BAR-HEN, H. JAVAUX et N. VIALANEIX. « Analyse statistique des profils et de l'activité des participants d'un MOOC ». In : *Revue Internationale des Technologies en Pédagogie Universitaire* 12.1-2 (2016), p. 11–22.
- [108] C. GASPIN, O. RUÉ et M. ZYTNIKI. « Ingredients for in silico miRNA identification and annotation ». In : *JSM Biotechnology and Biomedical Engineering* 3.5 (2016), p. 1–5.
- [110] S. HAMMOUMI, T. VALLAEYS, A. SANTIKA, P. LELEUX, E. BORZYM, C. KLOPP et J.-C. AVARRE. « Targeted genomic enrichment and sequencing of CyHV-3 from carp tissues confirms low nucleotide diversity and mixed genotype infections ». In : *PeerJ* 4 (2016).
- [124] C.-E. RABIER, P. BARRE, T. ASP, G. CHARMET et B. MANGIN. « On the accuracy of genomic selection ». In : *Plos One* 11.6 (2016), p. 1–23.
- [130] S. TRAORE, K. E. ROBERTS, D. ALLOUCHE, B. R. DONALD, I. ANDRE, T. SCHIEX et S. BARBE. « Fast search algorithms for computational protein design ». In : *Journal of Computational Chemistry* 37.12 (2016), p. 1048–1058.
- [139] C. BÉCAVIN, M. KOUTERO, N. TCHITCHEK, F. CERUTTI, P. LECHAT, N. MAILLET, C. HOEDE, H. CHIAPELLO, C. GASPIN et P. COSSART. « Listeriomics : an Interactive Web Platform for Systems Biology of Listeria ». In : *mSystems* 2.2 (2017).
- [145] R. BLANC-MATHIEU, L. ZURLETTO, J.-M. AURY, M. DA ROCHA, J. GOUZY, E. SALLET, C. MARTIN JIMENEZ, M. BAILLY-BECHET, P. CASTAGNONE, J.-F. FLOT, D. KOZLOWSKI, J. CAZARETH, A. COULOUX, C. DA SILVA, GUY, Y. J. KIM, C. RANCUREL, T. SCHIEX, P. ABAD, P. WINCKER et E. DANCHIN. « Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes ». In : *Plos Genetics* 13.6 (2017), e1006777.
- [146] J. BOLTON, E. MONTASTIER, J. CARAYOL, BONNEL, MIR, MARQUES, IACOVONI, N. VIALANEIX, A. ASTRUP, W. SARIS, VALSESIA, LANGIN et VIGUERIE. « Molecular biomarkers for weight control in obese individuals subjected to a multiphase dietary intervention ». In : *Journal of Clinical Endocrinology and Metabolism* 102.8 (2017), p. 2751–2761.

- [153] C. CHARBONNEL, NIAZI, E. ELVIRA-MATELOT, E. NOWAK, M. ZYTNIKI, De BURES, JOBET, A. OPSOMER, N. SHAMANDI, NOWOTNY, CARAPITO, REICHHELD, H. VAUCHERET et SÁEZ-VÁSQUEZ. « The siRNA suppressor RTL1 is redox-regulated through glutathionylation of a conserved cysteine in the double-stranded-RNA-binding domain ». In : *Nucleic Acids Research* 45.20 (2017), p. 11891–11907.
- [157] S. DOU, N. VIALANEIX, L. LIAUBET, Y. BILLON, M. GIORGI, H. GILBERT, J.-L. GOURDINE, J. RIQUET et D. RENAUDEAU. « (NMR)-N-1H-Based metabolomic profiling method to develop plasma biomarkers for sensitivity to chronic heat stress in growing pigs ». In : *Plos One* 12.11 (2017), Non paginé.
- [159] R. GENUER, J.-M. POGGI, C. TULEAU-MALOT et N. VIALANEIX. « Random Forests for Big Data ». In : *Big Data Research* 9 (2017), p. 28–46.
- [166] F. LEVILLAIN, Y. POQUET, L. MALLET, S. MAZERES, M. MARCEAU, R. BROSCH, F.-C. BANGE, P. SUPPLY, A. MAGALON et O. NEYROLLES. « Horizontal acquisition of a hypoxia-responsive molybdenum cofactor biosynthesis pathway contributed to Mycobacterium tuberculosis pathoadaptation ». In : *Plos Pathogens* 13.11 (2017).
- [167] L. MALLET, T. BITARD-FEILDEL, F. CERUTTI et H. CHIAPELLO. « PhylOligo : a package to identify contaminant or untargeted organism sequences in genome assemblies ». In : *Bioinformatics* 33.20 (2017), p. 3283–3285.
- [168] J. MARIETTE, M. OLTEANU et N. VIALANEIX. « Efficient interpretable variants of online SOM for large dissimilarity data ». In : *Neurocomputing* 225 (2017), p. 31–48.
- [173] E. MORMEDE, V. SAUTRON, C. YDIER, D. BAZOVKINA, A. SEVIN, L. GRESS, Y. LIPPI, C. NAYLIES, Y. BILLON, L. LIAUBET, P. MORMÈDE et N. VIALANEIX. « Time course study of the response to LPS targeting the pig immune gene networks ». In : *BMC Genomics* 18 (2017).
- [176] T. H. H. NGUYÊN, C. BESSIERE, S. DE GIVRY et T. SCHIEX. « Triangle-based consistencies for cost function networks ». In : *Constraints* 22.2 (2017), p. 230–264.
- [190] D. SIMONCINI, T. SCHIEX et K. Y. ZHANG. « Balancing exploration and exploitation in population-based sampling improves fragment-based de novo protein structure prediction ». In : *Proteins* 85.5 (2017), p. 852–858.
- [193] M. ZYTNIKI. « mmquant : how to count multi-mapping reads ? » In : *BMC Bioinformatics* 18.1 (2017), p. 411.
- [206] M. CHAMPION, V. PICHENY et M. VIGNES. « Inferring large graphs using l1 -penalized likelihood ». In : *Statistics and Computing* (2018), p. 905–921.
- [209] M. COTTRELL, M. OLTEANU, F. ROSSI et N. VIALANEIX. « Self-organizing maps, theory and applications ». In : *Revista Investigacion Operacional* 39.1 (2018), p. 1–22.
- [211] S. DANG et N. VIALANEIX. « Cutting edge bioinformatics and biostatistics approaches are bringing precision medicine and nutrition to a new era ». In : *Lifestyle Genomics* 11.2 (2018), p. 73–76.
- [218] A. IMBERT, A. VALSESIA, C. LE GALL, C. ARMENISE, LEFEBVRE, GOURRAUD, N. VIGUERIE et N. VIALANEIX. « Multiple hot-deck imputation for network inference from RNA sequencing data ». In : *Bioinformatics* 34.10 (2018), p. 1726–1732.
- [222] M. MARTI-MARIMON, N. VIALANEIX, V. VOILLET, M. BOUISSOU-MATET, Y. LAHBIB-MANSAIS et L. LIAUBET. « A new approach of gene co-expression network inference reveals significant biological processes involved in porcine muscle development in late gestation ». In : *Scientific Reports* 8 (2018).

- [230] T. SCHIEX. « Contraintes, incertitudes et modèles graphiques ». In : *Bulletin de l'AFIA* 100 (2018), p. 40–42.
- [241] M. CHABOD, L. KHAJAVI-ETTOUATI et R. S. LIBLAU. « An autoimmune basis for narcolepsy type 1 ? » In : *M/S Médecine/Sciences* 35.3 (2019), p. 204–207.
- [243] A. CHARPENTIER, D. MIGNON, S. BARBE, J. CORTES, T. SCHIEX, T. SIMONSON et D. ALLOUCHE. « Variable Neighborhood Search with Cost Function Networks To Solve Large Computational Protein Design Problems ». In : *Journal of Chemical Information and Modeling* 59.1 (2019), p. 127–136.
- [244] S. CHASTANT-MAILLARD, C. SABY, W. ZHANG, R. FOURNIER, R. SERVIEN et N. VIALANEIX. « Reprise atypique de la cyclicité ovarienne chez la vache laitière : une forte association avec la cétose ». In : *Point Vétérinaire* 393 (2019).
- [247] S. DENNENMOSER, F. J. SEDLAZECK, M. C. SCHATZ, J. ALTMÜLLER, M. ZYTNICKI et NOLTE. « Genome-wide patterns of transposon proliferation in an evolutionary young hybrid fish ». In : *Molecular Ecology* 28.6 (2019), p. 1491–1505.
- [261] V. POLONAI, S. NIEHUS, I. WAWRZYNIAK, A. FRANCHET, C. GASPIN, A. BELKORCHIA, M. REICHSTADT, C. BELSER, K. LABADIE, A. COULOUX, F. DELBAC, E. PEYRETAILLADE et D. FERRANDON. « Draft Genome Sequence of *Tubulinosema ratisbonensis*, a Microsporidian Species Infecting the Model Organism *Drosophila melanogaster* ». In : *Microbiology Resource Announcements* 8.31 (2019).
- [262] C.-E. RABIER, J.-M. AZAÏS, J. M. ELSSEN et C. DELMAS. « Chi-square processes for gene mapping in a population with family structure ». In : *Statistical Papers* 60.1 (2019), p. 239–271.
- [263] C.-E. RABIER, B. MANGIN et S. GRUSEA. « On the accuracy in high-dimensional linear models and its application to genomic selection ». In : *Scandinavian Journal of Statistics* 46.1 (2019), p. 289–313.
- [266] D. SIMONCINI, K. Y. J. ZHANG, T. SCHIEX et S. BARBE. « A structural homology approach for computational protein design with flexible backbone ». In : *Bioinformatics* (2019).

### Articles de synthèse / revues bibliographiques (nombre total)

- [270] P. THÉBAULT, R. BOURQUI, W. BENCHIMOL, C. GASPIN, P. SIRAND-PUGNET, R. URICARU et I. DUTOUR. « Advantages of mixing bioinformatics and visualization approaches for analyzing sRNA-mediated regulatory bacterial networks. » In : *Briefings in Bioinformatics* 16.5 (2015), p. 795–805.
- [271] A. IMBERT et N. VIALANEIX. « Décrire, prendre en compte, imputer et évaluer les valeurs manquantes dans les études statistiques : une revue des approches existantes ». In : *Journal de la Société Française de Statistique* 159.2 (2018), p. 1–55.
- [272] N. PEYRARD, M.-J. CROS, S. DE GIVRY, A. FRANC, S. ROBIN, R. SABBADIN, T. SCHIEX et M. VIGNES. « Exact or approximate inference in graphical models : why the choice is dictated by the treewidth, and how variable elimination can be exploited ». In : *Australian and New Zealand Journal of Statistics* 61.2 (2019), p. 89–133.

## 2- Ouvrages

*Pour chacune des rubriques ci-dessous, on dressera, s'il y a lieu, la liste de ces produits, qu'on numérottera de 1 à x, dans l'ordre chronologique, en distinguant d'abord les 20 % jugés les plus significatifs, puis les 80 % restants.*

### Chapitres d'ouvrage (nombre total)

- [273] F. COSTE, C. NÉDELLEC, T. SCHIEX et J.-P. VERT. « Bioinformatique ». In : *Panorama de l'Intelligence Artificielle*. Toulouse FRA : Cepadues Editions, 2014.
- [276] T. SCHIEX, S. DE GIVRY et M. COOPER. « Réseaux de contraintes valués ». In : *Panorama de l'intelligence artificielle*. FRA : Cepadues Editions, 2014.
- [277] F. BACCHUS et G. KATSIRELOS. « Using Minimal Correction Sets to More Efficiently Compute Minimal Unsatisfiable Sets ». In : *Computer Aided Verification*. Cham CHE : Springer International Publishing, 2015, p. 70–86.
- [280] C. BESSIERE, A. DAOUDI, E. HEBRARD, G. KATSIRELOS, LAZAAR, Y. MECHQRANE, N. NARODYTSKA, C.-G. QUIMPER et T. WALSH. « New Approaches to Constraint Acquisition ». In : *Data Mining and Constraint Programming*. Cham CHE : Springer International Publishing AG, 2016, p. 51–76.
- [282] N. VIALANEIX, L. LIAUBET et M. SAN CRISTOBAL. « Depicting gene co-expression networks underlying eQTLs ». In : *Systems Biology in Animal Production and Health*. Sous la dir. d'Haja N. KADARMIDEEN. Switzerland : Springer International Publishing, 2016, p. 1–31.
- [285] S. TRAORE, D. ALLOUCHE, I. ANDRE, T. SCHIEX et S. BARBE. « Deterministic search methods for computational protein design ». In : *Computational Protein Design*. New-York USA : Editions Springer, 2017, p. 107–123.
- [286] A. R. D. VOET, D. SIMONCINI, J. R. H. TAME et K. Y. J. ZHANG. « Evolution-Inspired Computational Design of Symmetric Proteins ». In : *Computational Protein Design*. Sous la dir. d'I SAMISH. NEW YORK : SPRINGER, 2017, p. 309–322.
- [287] E. HEBRARD et G. KATSIRELOS. « Clause learning and new bounds for graph coloring ». In : *Principles and Practice of Constraint Programming*. Berlin DEU : Editions Springer, 2018, p. 179–194.
- [288] S. LAGUERRE, I. GONZALEZ, S. NOUAILLE, A. MOISAN, N. VIALANEIX, C. GASPIN, M. BOUVIER, A. J. CARPOUSIS, M. BOUSQUET et L. GIRBAL. « Large-Scale Measurement of mRNA Degradation in Escherichia coli : To Delay or Not to Delay ». In : *Methods in Enzymology*. Sous la dir. d'Aj CARPOUSIS. SAN DIEGO USA : ELSEVIER ACADEMIC PRESS INC, 2018, p. 47–66.
- [289] N. VIALANEIX et S. CANU. « Apprentissage connexionniste ». In : *Apprentissage Statistique et Données Massives*. Paris FRA : Editions Technip, 2018, p. 271–294.
- [290] N. VIALANEIX et F. ROSSI. « Méthodes pour l'apprentissage de données massives ». In : *Apprentissage Statistique et Données Massives*. Paris FRA : Editions Technip, 2018, p. 417–438.
- [292] E. SALLET, T. SCHIEX et J. GOUZY. « EuGene : An Automated Integrative Gene Finder for Eukaryotes and Prokaryotes ». In : *Gene Prediction*. Sous la dir. de Martin KOLLMAR. 233 Spring Street, New York, NY 10013-1578 USA : Humana Press INC., Springer Science, 2019, p. 97–120.

### Thèses éditées

- [293] E. CADIC. « Recherche de facteurs génétiques impliqués dans l'élaboration du rendement sous contrainte hydrique chez le tournesol *Helianthus annuus* par génétique d'association et analyse de liaison dans une population recombinante ». Thèse de doct. Université Paul Sabatier (Toulouse 3), 2014.

- [294] M. CHAMPION. « Contribution à la modélisation et l'inférence de réseau de régulation de gènes ». Thèse de doct. Université Paul Sabatier (Toulouse 3), 2014.
- [295] T. H. H. NGUYÊN. « Strong consistencies for weighted constraint satisfaction problems ». Thèse de doct. Université Paul Sabatier (Toulouse 3), 2015.
- [298] V. SAUTRON. « Biologie intégrative des réponses au stress et robustesse chez le porc ». Thèse de doct. Institut National Polytechnique de Toulouse, 2016.
- [300] J. MARIETTE. « Apprentissage statistique pour l'intégration de données omiques ». Thèse de doct. Université Paul Sabatier (Toulouse 3), 2017.
- [302] C. VIRICEL. « Contributions au développement d'outils computationnels de design de protéines : méthodes et algorithmes de comptage avec garantie ». Thèse de doct. Institut National des Sciences Appliquées de Toulouse, 2017.
- [304] F. CERUTTI. « Analyse de la dynamique évolutive de petits ARNs non-codants bactériens : mise au point d'une stratégie et application aux firmicutes ». Thèse de doct. Université Paul Sabatier (Toulouse 3), 2018.
- [305] A. IMBERT. « Intégration de données complexes et hétérogènes à partir de tableaux de tailles différentes ». Thèse de doct. Université Toulouse 1 Capitole, 2018.

### 3- Production dans des colloques / congrès, séminaires de recherche

*Pour chacune des rubriques ci-dessous, on dressera, s'il y a lieu, la liste de ces produits qu'on numérotera de 1 à x, dans l'ordre chronologique, en distinguant d'abord les 20 % jugés les plus significatifs, puis les 80 % restants.*

#### Éditions d'actes de colloques / congrès

- [311] S. DE GIVRY et N. PEYRARD, éd. *Actes des Neuvièmes Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes 2018*. 9e Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes 2018 (JFRB 2018). 2018, 114 p.

#### Articles publiés dans des actes de colloques / congrès

- [313] D. ALLOUCHE, S. DE GIVRY, B. HURLEY, G. KATSIRELOS, B. O'SULLIVAN et T. SCHIEX. « Une comparaison de logiciels d'optimisation sur une large collection de modèles graphiques ». In : JFPC 2014 - Dixièmes Journées Francophones de Programmation par Contraintes. 2014, p. 301–310.
- [315] F. BACCHUS, J. DAVIES, M. TSIMPOUKELLI et G. KATSIRELOS. « Relaxation search : a simple way of managing optional clauses ». In : AAAI 2014 - 28th AAAI Conference. 2014.
- [316] C. BESSIERE, E. HEBRARD, G. KATSIRELOS, Z. KIZILTAN, N. NARODYTSKA et T. WALSH. « Reasoning about Constraint Models ». In : *Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence (LNAI)*. 13th Pacific Rim International Conference on Artificial Intelligence. 2014, p. 795–808.
- [317] C. BESSIERE, E. HEBRARD, G. KATSIRELOS, Z. KIZILTAN, E. PICARD-CANTIN, C.-G. QUIMPER et T. WALSH. « The balance constraint family ». In : *Lecture Notes in Computer Science. CP 2014 - 20th International Conference on Principles et Practice of Constraint Programming*. 2014, p. 174–189.
- [318] J. BOELAERT, L. BENDHAIBA, M. OLTEANU et N. VIALANEIX. « SOMbrero : an R package for numeric and non-numeric self-organizing maps ». In : *Advances in Intelligent Systems and Computing*. WSOM 2014. 2014, p. 219–228.

- [324] S. DE GIVRY, B. HURLEY, D. ALLOUCHE, G. KATSIRELOS, B. O'SULLIVAN et T. SCHIEX. « An experimental evaluation of CP/AI/OR solvers for optimization in graphical models ». In : Congrès ROADEF'2014. 2014.
- [325] S. DE GIVRY, J. H. LEE, K. LEUNG et Y. SHUM. « Solving a judge assignment problem using conjunctions of global cost functions ». In : 20. International Conference on Principles et Practice of Constraint Programming - CP2014. 2014, p. 797–812.
- [331] J. MARIETTE, M. OLTEANU, J. BOELAERT et N. VIALANEIX. « Bagged kernel SOM ». In : *Advances in Intelligent Systems and Computing*. 10. International Workshop on Self-Organizing Maps (WSOM 2014). 2014, p. 45–54.
- [332] E. MERLOT, A. PRUNIER, M. DAMON, F. VIGNOLES, N. VIALANEIX, P. MORMÈDE et E. MORMÈDE. « Blood transcriptome response to LPS in pigs ». In : *Annual Meeting of the European Association for Animal Production*. 65. Annual Meeting of the European Federation of Animal Science (EAAP). 2014, p. 274.
- [333] T. H. H. NGUYÊN, S. DE GIVRY, T. SCHIEX et C. BESSIÈRE. « Maintaining Virtual Arc Consistency Dynamically During Search ». In : 26th IEEE International Conference on Tools with Artificial Intelligence - ICTAI 2014. 2014.
- [334] M. OLTEANU et N. VIALANEIX. « Self-organizing maps for clustering and visualization of bipartite graphs ». In : 46e Journées de la Société Française de Statistique. 2014.
- [337] V. PICHENY, J. VANDEL, M. VIGNES et N. VIALANEIX. « Reconstruction quality of a biological network when its constituting elements are partially observed ». In : 17th International Conference on Artificial Intelligence et Statistics - AI & Statistics 2014. 2014.
- [342] N. VIALANEIX. « J'ai testé pour vous... un MOOC ». In : 46e Journées de la Société Française de Statistique. 2014.
- [345] D. ALLOUCHE, S. DE GIVRY, G. KATSIRELOS, T. SCHIEX et M. ZYTNICKI. « Anytime hybrid best-first search with tree decomposition for weighted CSP ». In : CP 2015 - 21st International Conference on Principles et Practice of Constraint Programming. 2015, p. 12–29.
- [352] P. BESSE, N. VIALANEIX et A. RUIZ-GAZEN. « Enseigner la statistique pour l'analyse de mégadonnées ». In : 47ème Journées de Statistique de la SFdS. 2015.
- [353] C. BESSIÈRE, E. HEBRARD, G. KATSIRELOS et T. WALSH. « Reasoning about Connectivity Constraints ». In : 24th International Conference on Artificial Intelligence - IJCAI 2015. 2015, p. 2568–2574.
- [364] R. GENUER, J.-M. POGGI, C. TULEAU-MALOT et N. VIALANEIX. « Random forests and big data ». In : 47ème Journées de Statistique de la SFdS. 2015.
- [371] V. LOUX, M. MARIADASSOU, S. ALMEIDA DA SILVA, H. CHIAPELLO, A. HAMMAMI, J. BURATTI, A. GENDRAULT, V. BARBE, J. M. AURY, S.-M. DEUTSCH, A. NICOLAS, S. PARAYRE-BRETON, M.-N. MADEC, V. CHUAT, G. JAN, P. PETERLONGO, Y. LE LOIR et H. FALENTIN. « Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii* ». In : JOBIM 16. Journées Ouvertes de Biologie, Informatique et Mathématiques. 2015.
- [379] M. SACCAREAU, C. MORENO-ROMIEUX, G. CIAPPESONI, B. GOFFINET, R. FAIVRE, F. BRUN et S. BISHOP. « Modélisation génétique et épidémiologique du parasite gastro-intestinal chez le mouton ». In : *Rencontres autour des Recherches sur les Ruminants*. 22. Rencontres Recherches Ruminants. 2015, p. 36–36.

- [380] C. SIBERTIN-BLANC et N. VIALANEIX. « Data analysis of social simulations outputs - Interpreting the dispersion of variables ». In : *Lecture Notes in Computer Science*. 15th International Workshop on Multi-Agent-Based Simulation (MABS). 2015, p. 133–150.
- [381] C. VIRICEL, D. SIMONCINI, D. ALLOUCHE, S. DE GIVRY, S. BARBE et T. SCHIEX. « Approximate counting with deterministic guarantees for affinity computation ». In : *Advances in Intelligent Systems and Computing*. 3. International Conference on Modelling, Computation, Optimization in Information Systems et Management Sciences - MCO 2015. 2015, p. 165–176.
- [383] F. BACCHUS et G. KATSIRELOS. « Finding a Collection of MUSes Incrementally ». In : *Lecture Notes in Computer Science*. 13th International Conference on Integration of Artificial Intelligence et Operations Research Techniques in Constraint Programming (CPAIOR). 2016, p. 35–44.
- [385] C. BESSIÈRE, E. HEBRARD, G. KATSIRELOS, Z. KIZILTAN et T. WALSH. « Ranking Constraints ». In : 25th International Joint Conference on Artificial Intelligence IJCAI-16. 2016, p. 705–711.
- [389] M. COTTRELL, M. OLTEANU, F. ROSSI et N. VIALANEIX. « Theoretical and applied aspects of the self-organizing maps ». In : *Advances in Intelligent Systems and Computing*. 11th International Workshop WSOM 2016 ; Houston (Etats-Unis) - (2016-01-06 - 2016-01-08) / Workshop. 2016, p. 3–26.
- [398] A. IMBERT et N. VIALANEIX. « Outils pour l'analyse et la simulation de données RNA-seq ». In : Cinquièmes Rencontres R 2016. 2016.
- [399] S. JASSON et D. LEROUX. « A Bayesian approach for the imputation of genotypes on observed markers in complex pedigrees ». In : The Allied Genetics Conference 2016 (TAGC). 2016.
- [403] Y. LAHBIB MANSAIS, M. MARTI-MARIMON, V. VOILLET, H. BARASC, F. MOMPART, J. RIQUET, S. FOISSAC, D. ROBELIN, H. ACLOQUE, Y. BILLON, N. VIALANEIX, L. LIAUBET et M. BOUISSOU-MATET YERLE. « 3D nuclear positioning of IGF2 alleles and trans interactions with imprinted genes ». In : *Journal of Animal Science*. 35. Conference of the International Society for Animal Genetics (ISAG). 2016, p. 181–181.
- [404] Y. LAHBIB MANSAIS, M. MARTI-MARIMON, V. VOILLET, H. BARASC, F. MOMPART, J. RIQUET, S. FOISSAC, D. ROBELIN, H. ACLOQUE, Y. BILLON, N. VIALANEIX, L. LIAUBET et M. BOUISSOU-MATET YERLE. « 3D nuclear positioning of IGF2 alleles and trans interactions with imprinted genes in pig fetal cells ». In : *Chromosome Research*. 22. International Colloquium on Animal Cytogenetics et Genomics (ICACG). 2016, p. 89–89.
- [405] S. MAQROT, S. DE GIVRY, G. QUESNEL et M. TCHAMITCHIAN. « Designing mixed fruit-vegetable cropping systems by integer quadratic programming ». In : *null*. 8th International Congress on Environmental Modelling et Software (iEMSs). 2016, p. 462–470.
- [406] J. MARIETTE et N. VIALANEIX. « Aggregating self-organizing maps with topology preservation ». In : *Advances in Intelligent Systems and Computing*. 11th International Workshop WSOM 2016. 2016, p. 27–37.
- [407] M. MARTI-MARIMON, H. ACLOQUE, D. ROBELIN, M. ZYTNICKI, S. DJEBALI QUELEN, Y. LAHBIB MANSAIS, D. ESQUERRE, F. MOMPART, M. BOUISSOU-MATET YERLE et S. FOISSAC. « Construction of whole genome interaction maps and study of correlation with gene expression level in fetal muscle cells ». In : *Chromosome Research*. 22. International Colloquium on Animal Cytogenetics et Genomics (ICACG). 2016, p. 104–104.

- [409] M. OLTEANU et N. VIALANEIX. « Sparse online self-organizing maps for large relational data ». In : *Advances in Intelligent Systems and Computing*. 10th International Workshop WSOM 2014. 2016, p. 73–82.
- [415] V. SAUTRON, E. MORMEDE, L. GRESS, Y. LIPPI, Y. BILLON, N. VIALANEIX et P. MORMÈDE. « Time course of the response to ACTH in pig : biological and transcriptomic study ». In : *Journal of Animal Science*. 35. Conference of the International Society for Animal Genetics (ISAG). 2016, S54–S55.
- [417] I. TSARMOPOULOS, C. JOLLARD, G. GOURGUES, A. BLANCHARD, I. DUTOUR, R. BOURQUI, A. MOISAN, H. CHIAPELLO, P. THÉBAULT, C. LARTIGUE-PRAT et P. SIRAND-PUGNET. « Investigating mycoplasma small non-coding RNAs using synthetic biology tools ». In : *null*. Colloque Biologie synthétique et systémique, BioSynSys. 2016, p. 16.
- [418] N. VIALANEIX, C. BONTEMPS et S. DEJEAN. « Outils pour chercher de l'information sur R et se former ». In : Cinquièmes Rencontres R 2016. 2016.
- [419] C. VIRICEL, D. SIMONCINI, S. BARBE et T. SCHIEX. « Guaranteed Weighted Counting for Affinity Computation : Beyond Determinism and Structure ». In : *Lecture Notes in Computer Science*. CP16 - 22nd International Conference on the Principles et Practice of Constraint Programming. 2016, p. 733–750.
- [420] D. ALLOUCHE. « Euclidean Variable Neighborhood Search : A method for large computation protein design ». In : Deciphering complex energy landscape et kinetic network from single molecules to cells : a new challenge to make theories meet experiments. 2017.
- [429] S. DE GIVRY et G. KATSIRELOS. « Clique Cuts in Weighted Constraint Satisfaction ». In : *Lecture Notes in Computer Science*. The 23rd International Conference on Principles et Practice of Constraint Programming - CP17. 2017, p. 97–113.
- [431] S. DJEBALI QUELEN, K. MUNYARD, N. VIALANEIX, C. CABAU, A. RAU, E. CRISCI, T. DERRIEN, C. KLOPP, M. ZYTNICKI, S. LAGARRIGUE, H. ACLOQUE, S. FOISSAC et E. GIUFFRA. « Integrative and differential analysis of transcriptomes and chromatin accessibility regions reveals regulatory mechanisms involved in pig immune and metabolic functions ». In : 36th International Society for Animal Genetics Conference. 2017, p. 42.
- [432] S. FOISSAC, S. DJEBALI QUELEN, H. ACLOQUE, P. BARDOU, F. BLANC, C. CABAU, T. DERRIEN, F. DROUET, D. ESQUERRE, S. FABRE, C. GASPIN, I. GONZALEZ, A. GOUBIL, C. KLOPP, F. LAURENT, S. MARTHEY, M. MARTI-MARIMON, F. MOMPART, K. MUNYARD, K. MURET, S. POLLET, P. QUÉRÉ, A. RAU, D. ROBELIN, M. SAN CRISTOBAL, M. TIXIER-BOICHARD, G. TOSSER-KLOPP, N. VIALANEIX, S. VINCENT-NAULLEAU, M. ZYTNICKI et al. « Profiling the landscape of transcription, chromatin accessibility and chromosome conformation of cattle, pig, chicken and goat genomes [FAANG pilot project] ». In : 36th International Society for Animal Genetics Conference. 2017, p. 5–6.
- [434] S. JASSON et D. LEROUX. « Spell-QTL, a New Tool for QTL Analysis on Modern Datasets ». In : PAG XXV - Plant et Animal Genome Conference. 2017.
- [437] L. MALLET, C. HOEDE, F. CERUTTI, A. MOISAN, C. GASPIN, I. VIRLOGEUX-PAYANT, I. SHOMER, O. GAL-MOR, T. SCHIEX et H. CHIAPELLO. « Host tropism and host-pathogen interplay of typhoidal *Salmonella enterica* ». In : JOBIM 2017 - Journées Ouvertes en Biologie, Informatique et Mathématiques. 2017, p.192.
- [438] S. MAQROT, S. DE GIVRY, G. QUESNEL et M. TCHAMITCHIAN. « A mixed integer programming reformulation of the mixed fruit-vegetable crop allocation problem ». In : *Lecture Notes in Computer Science*. 30th International Conference on Industrial, Engineering et Other Applications of Applied Intelligent Systems, IEA/AIE 2017. 2017, p. 237–250.



- [439] J. MARIETTE, F. ROSSI, M. OLTEANU et N. VIALANEIX. « Accelerating stochastic kernel SOM ». In : XXVth European Symposium on Artificial Neural Networks, Computational Intelligence et Machine Learning. 2017, p. 269–274.
- [440] M. MARTI-MARIMON, H. ACLOQUE, M. ZYTNIKI, S. DJEBALI QUELEN, N. VIALANEIX, O. MADSEN, Y. LAHBIB MANS AIS, D. ESQUERRE, F. MOMPART, L. LIAUBET, M. GROENEN, M. BOUISSOU-MATET YERLE et S. FOISSAC. « Characterization of 3D genomic interactions in fetal pig muscle ». In : 36th International Society for Animal Genetics Conference. 2017, p. 43.
- [444] A. OUALI, D. ALLOUCHE, S. DE GIVRY, S. LOUDNI, Y. LEBBAH, F. RIBEIRO ECKHARDT et L. LOUKIL. « Iterative Decomposition Guided Variable Neighborhood Search for Graphical Model Energy Minimization ». In : Conference on Uncertainty in Artificial Intelligence (UAI'17). 2017, p. 550–559.
- [446] C. SABY, W. ZHANG, R. FOURNIER, R. SERVIEN, N. VIALANEIX, F. CORBIERE et S. CHASTANT-MAILLARD. « Reprise de cyclicité post partum et performances de reproduction chez la vache laitière Prim'Holstein en France ». In : Journées Nationales des GTV 2017. 2017, p. 709–716.
- [449] N. VIALANEIX. « Stochastic Self-Organizing Map variants with the R package SOMbrero ». In : 12th International Workshop on Self-Organizing Maps, Learning Vector Quantization, Clustering et Data Visualization (WSOM). 2017.
- [453] D. ALLOUCHE. « Pourquoi les MATH et l'intelligence artificielle sont important pour comprendre et améliorer le fonctionnement de protéine? » In : Fête de la science. 2018.
- [454] D. ALLOUCHE, A. CHARPENTIER, T. SCHIEX et T. SIMONSON. « Variable Neighborhood Search with Cost Function Networks to Solve Large Computational Protein Design Problems ». In : Rencontre des Chimistes Théoriciens Francophones. 2018.
- [463] S. DOU, N. VIALANEIX, L. LIAUBET, H. GILBERT, J.-L. GOURDINE, J. RIQUET et D. RENAUDEAU. « Développement de biomarqueurs sanguins pour évaluer la sensibilité des porcs à la chaleur ». In : Journées d'Animation Scientifique du département Phase (JAS Phase 2018). 2018, p. 66.
- [468] S. JASSON et D. LEROUX. « Method for Computing Map Likelihood Given Marker Observations and an Arbitrary Population Structure ». In : PAG XXVI - Plant et Animal Genome Conference. 2018.
- [470] S. MAQROT, S. DE GIVRY, G. QUESNEL et M. TCHAMITCHIAN. « Improving Wedelin's Heuristic with Sensitivity Analysis ». In : 19. congrès annuel de la Société française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF). 2018.
- [472] M. MARTI-MARIMON, N. VIALANEIX, V. VOILLET, M. YERLE-BOUISSOU, L. LIAUBET et Y. LAHBIB-MANS AIS. « An innovative method of gene co expression network inference reveals significant biological processes involved in fetal porcine muscle development ». In : *Comparative Cytogenetics*. 23rd International Colloquium on Animal Cytogenetics et Genomics. 2018, p. 310–311.
- [476] A. OUALI, D. ALLOUCHE, S. DE GIVRY, S. LOUDNI, Y. LEBBAH, F. RIBEIRO ECKHARDT et L. LOUKIL. « Recherche complète à voisinages variables guidée par la décomposition arborescente pour la minimisation d'énergie dans les modèles graphiques ». In : 9e Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes 2018 (JFRB 2018). 2018, p. 107–113.

- [480] L. POMIÈS, L. GODY, C. PENOUILH-SUZETTE, N. LANGLADE, B. MANGIN et S. DE GIVRY. « Building artificial genetical genomic datasets to optimize the choice of gene regulatory network inference methods ». In : Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM 2018). 2018, p. 338–341.
- [481] L. POMIÈS, L. GODY, C. PENOUILH-SUZETTE, N. LANGLADE, B. MANGIN et S. DE GIVRY. « Building artificial genetical genomic datasets to optimize the choice of gene regulatory network inference methods ». In : 17th European Conference on Computational Biology (ECCB 2018). 2018.
- [483] T. SCHIEX, S. VEREL, D. SIMONCINI et S. BARBE. « Fitness Landscape Analysis around the Optimum in Computational Protein Design ». In : GECCO 2018 @ Tokyo The Genetic et Evolutionary Computation Conference. 2018, p. 355–362.
- [488] D. ALLOUCHE. « Cost function network for molecular design an IA method for discret search space exploration ». In : Symposium on Molecular Design et Bioinformatics 2019 (SEADIM). 2019.
- [491] G. LEFORT, L. LIAUBET, H. QUESNEL, C. CANLET, N. VIALANEIX et R. SERVIEN. « ASICS : identifier et quantifier des métabolites à partir d'un spectre RMN 1H ». In : 51èmes Journées de Statistique de la SFdS. 2019.
- [495] N. RANDRIAMIHAMISON, P. NEUVIAL et N. VIALANEIX. « Classification ascendante hiérarchique, contrainte d'ordre : conditions d'applicabilité, interprétabilité des dendrogrammes ». In : 51èmes Journées de Statistique de la SFdS. 2019.

#### Autres produits présentés dans des colloques / congrès et des séminaires de recherche

- [497] J. AMSELEM, M. ALAUX, N. CHOISNE, O. INIZAN, N. LAPALU, N. MOHELLIBI, F. ALFAMA-DEPAUW, A. KELIET, E. KIMMEL, J. KREPLAK, C. GUERCHE, M. LOAEC, C. POMMIER, M. ZYTNIKI, M.-H. LEBRUN, D. STEINBACH et H. QUESNEVILLE. *URGI plant and fungi platform : distributed resources through GMOD tools*. Plant and Animal Genome conference XXII, GMOD session. 2014.
- [499] N. BRIOT, A. CHATEAU, R. COLETTA, S. DE GIVRY, P. LELEUX et T. SCHIEX. *An integer linear programming approach for genome scaffolding*. 10th Workshop on Constraint-Based Methods for Bioinformatics (WCB), 2014. 2014.
- [512] E. MONTASTIER, N. VIALANEIX, I. GONZALEZ, S. CASPAR-BAUGUIL, W. H. SARIS, D. LANGIN, M. KUNESOVA et N. VIGUERIE. *Adipose tissue signatures related to weight changes in response to calorie restriction and subsequent weight maintenance using lipidome and gene profiling network analysis*. Bioinformatics/Biostatistics regional workshop. 2014.
- [515] C.-E. RABIER. *Gaussian and chi square processes for quantitative trait locus mapping under selective genotyping*. 2014 International Indian Statistical Association Conference (IISA). 2014.
- [520] V. SAUTRON, E. MORMEDE, P. MORMÈDE et N. VIALANEIX. *Genetics systems of stress responses in pigs*. Bioinformatics/Biostatistics regional workshop. 2014.
- [521] N. VIALANEIX. *Contributions à l'analyse de données non vectorielles*. 2014.
- [522] N. VIALANEIX, M. VIGNES, N. VIGUERIE et M. SAN CRISTOBAL. *Inferring networks from multiple samples with consensus LASSO*. ENBIS Spring Meeting. 2014.
- [535] L. MALLET, C. GUERIN, J. AMSELEM, E. FOURNIER et H. CHIAPELLO. *The GEMO project : Hitchhiking DNA in Magnaporthe oryzae*. JOBIM 2015. Meeting of working Group Medicago sativa. 2015.

- [537] M. OLTEANU et N. VIALANEIX. *Classification et visualisation de graphes avec SOMbrero*. Quatrièmes Rencontres R. 2015.
- [544] M. SACCAREAU, C. MORENO-ROMIEUX, G. CIAPPESONI, B. GOFFINET, R. FAIVRE, F. BRUN et S. C. BISHOP. *Genetic-epidemiological modelling of gastrointestinal parasitism in sheep*. 14. Conference of the International Society for Veterinary Epidemiology and Economics (ISVEE 14). 2015.
- [545] T. SCHIEX. *Computational Design as an Optimization Problem*. Séminaire INRIA-IRISA. 2015.
- [546] T. SCHIEX. *Computational protein design as an optimization problem*. BAI 2015 : Advance in Bioinformatics and Artificial Intelligence. 2015.
- [547] T. SCHIEX. *Optimization in Graphical Models*. 34e Journée Francilienne de Recherche Opérationnelle (JFRO). 2015.
- [548] T. SCHIEX. *Optimization in Graphical Models*. 27th IEEE International Conference on Tools with Artificial Intelligence. 2015.
- [549] T. SCHIEX, S. TRAORÉ, D. SIMONCINI, S. BARBE, S. DE GIVRY, G. KATSIRELOS, D. ALLOUCHE et I. ANDRE. *Computational Enzyme Design through deterministic optimization and counting*. 3DSig Structural Bioinformatics and Computational Biophysics 2015. 2015.
- [551] N. VIALANEIX. *What is a MOOC ?* 15th Annual Conference of ENBIS. 2015.
- [553] N. VIALANEIX et M. OLTEANU. *Multiple dissimilarity SOM for clustering and visualizing graphs with node and edge attributes*. International Conference on Machine Learning, Workshop FEAST. 2015.
- [560] S. FOISSAC, H. ACLOQUE, F. MOMPART, D. ESQUERRE, D. ROBELIN, M. SAN CRISTOBAL, S. DJEBALI QUELEN, M. ZYTNICKI et E. GIUFFRA. *Investigating the nuclear spatial conformation of livestock genomes using Hi-C*. NGS'16 Genome Annotation. 2016.
- [561] S. FOISSAC, H. ACLOQUE, F. MOMPART, D. ESQUERRE, D. ROBELIN, M. SAN CRISTOBAL, M. ZYTNICKI, E. GIUFFRA et S. DJEBALI QUELEN. *Profiling the Nuclear Spatial Conformation of Livestock Genomes in Cell Lines and Liver Cells*. PAG XXIV - Plant and Animal Genome Conference. 2016.
- [563] C. GASPIN, H. GROSJEAN, P. SIRAND-PUGNET, C. MARCK, A. BLANCHARD, S. YOSHIKAWA et D. FOURMY. *Translation machinery of mollicutes : economy is possible for proteins, not much for RNAs*. Ribosome Structure and Function 2016. 2016.
- [565] A. IMBERT, C. LE GAL, C. ARMENISE, G. LEFEBVRE, J. HAGER, A. VALSESIA, P.-A. GOURRAUD, N. VIGUERIE et N. VIALANEIX. *Imputation de données manquantes pour l'inférence de réseau à partir de données RNA-seq*. 48. Journées de Statistique de la SFdS. 2016.
- [566] G. KATSIRELOS et F. BACCHUS. *MUSes and MCSes*. AAI 2016 Workshop on Beyond NP. 2016.
- [567] Y. LAHBIB MANSAIS, M. MARTI-MARIMON, V. VOILLET, H. BARASC, F. MOMPART, J. RIQUET, S. FOISSAC, D. ROBELIN, H. ACLOQUE, Y. BILLON, N. VIALANEIX, L. LIAUBET et M. BOUSSOU-MATET. *3D nuclear positioning of IGF2 alleles and trans interactions with imprinted genes in pig fetal cells*. Conference on Genome Architecture in Space and Time. 2016.
- [568] J. MARIETTE, H. CHIAPELLO et N. VIALANEIX. *Integrating Tara oceans data sets using multiple kernels*. 15th European Conference on Computational Biology (ECCB 2016), Workshop "Recent Computational Advances in Metagenomics (RCAM)". 2016.

- [571] M. OLTEANU et N. VIALANEIX. *Using SOMbrero for clustering and visualizing complex data*. 9th International Conference of the ERCIM WG on Computational and Methodological Statistics. 2016.
- [572] L. PANIAGUA, A. LEIP, N. VIALANEIX et W. DE VRIES. *Estimating nitrous oxide fluxes from agricultural soils at European scale using a crop generic meta-model*. 19th Nitrogen Workshop. 2016.
- [575] V. PICHENY, R. SERVIEN et N. VIALANEIX. *Interpretable sparse sliced inverse regression for functional data*. Workshop Learning with Functional Data. 2016.
- [576] V. PICHENY, R. SERVIEN et N. VIALANEIX. *Parcimonie par intervalle pour la régression inverse par tranche fonctionnelle*. 48. Journées de Statistique de la SFdS. 2016.
- [579] V. SAUTRON, M. CHAVENT, N. VIGUERIE et N. VIALANEIX. *Multiway SIR for biological data integration*. Statistical Methods for Post Genomic Data. 2016.
- [580] V. SAUTRON, M. CHAVENT, N. VIGUERIE et N. VIALANEIX. *Multiway-SIR for longitudinal multi-table data integration*. 22nd International Conference on Computational Statistics (COMPSTAT). 2016.
- [581] T. SCHIEX. *Guaranteed energetic optimization, exploration and counting of sequence-conformation in Computational Protein Design*. 4th Prague Protein Spring Meeting 2016. 2016.
- [582] T. SCHIEX. *Guaranteed Optimization on Large Protein Design Problems*. Séminaire au Baker Lab (Institute for Protein Design), Séminaire au Caltech The Mayo Laboratory. 2016.
- [583] T. SCHIEX. *Optimization in Graphical Models - Connecting NP-complete frameworks*. Séminaire Université Catholique de Louvain. 2016.
- [584] T. SCHIEX. *Optimization in Graphical Models - NP-complete optimization and its applications*. Assemblée générale de l'Association Française d'Intelligence Artificielle. 2016.
- [585] T. SCHIEX. *Removing Dead Ends in (re)search - Emeritatus of Marc de Maeyer*. Eméritatus de Marc de Maeyer. 2016.
- [586] R. SERVIEN, V. PICHENY et N. VIALANEIX. *Interval sparsity for functional inverse regression*. 2016 CRoNoS Summer Course, 22nd International Conference on Computational Statistics (COMPSTAT). 2016.
- [587] N. VIALANEIX, R. GENUER, J.-M. POGGI et C. TULEAU-MALOT. *Random forests for big data*. Journées de Statistique de Rennes (JSTAR). 2016.
- [589] M. ZYTNIKI. *De quelle épigénétique parle-t-on ?* Journées bioinformatiques de l'INRA. 2016.
- [597] A. CHARCOSSET, A. ZANETTO, B. GOUESNARD, C. WELCKER, C. BAULAND, L. MOREAU, C. PALAFFRE, Y. BARRIÈRE, M. REYMOND, A. MURIGNEUX, S. MEZMOUK, G. GUÉRIN, S. MELKIOR, P. CAROLO, L. BOSSAY, B. REVOL, P. SEHABIAGUE, N. RANC, A. LAMBERT et J. BEIGBEDER. *French National Institute for Agricultural Research, France INRA-PROMAIS : 35 years of public-private cooperation in France on the conservation of maize genetic resources - study of their diversity and methodology of use in breeding*. EU-CARPIA Genetic Resources 2017 - Crop diversification in a changing world : Mobilizing the green gold of plant genetic resources. 2017.
- [603] A. DE LA FOYE, A.-F. YAO, M.-C. NOT, A. IMBERT et C. SAMIR. *Application d'une régression PLS fonctionnelle : recherche de liens entre deux ensembles de variables omiques*. 49. Journées de Statistique. JDS 2017. 2017.

- [606] R. GENUER, J.-M. POGGI, C. TULEAU-MALOT et N. VIALANEIX. *Random forests for big data*. 10th International Conference of the ERCIM WG on Computational and Methodological Statistics. 2017.
- [608] G. KATSIRELOS. *Reparameterization : a Universal Tool for Optimization and Counting*. Beyond NP meeting 2017. 2017.
- [611] S. MAQROT, S. DE GIVRY, G. QUESNEL et M. TCHAMITCHIAN. *Mixed Fruit-Vegetable Crop Allocation Problem using MIP*. 18e congrès annuel de la Société française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF). 2017.
- [613] J. MARIETTE et N. VIALANEIX. *Unsupervised multiple kernel learning to integrate various metagenomic sources*. 4ème colloque de Génomique Environnementale. 2017.
- [621] S. POLLET, A. BOLAND-AUGE, R. OLASO, C. HOEDE, C. GASPIN, N. MACH CASELLAS, J.-F. DELEUZE et E. BARREY. *Hippocode : A project to improve the annotation of equine non-coding RNAs involved in the Regulation of mitochondrial functions*. Targeting Mitochondria world congress. 2017.
- [626] R. SABBADIN, N. PEYRARD, M.-J. CROS, S. DE GIVRY, A. FRANC, S. ROBIN, T. SCHIEX et M. VIGNES. *Exact and approximate inference in graphical models*. 10th International Conference on Scalable Uncertainty Management (SUM). 2017.
- [627] C. SABY, W. ZHANG, R. FOURNIER, R. SERVIEN, N. VIALANEIX, F. CORBIERE et S. CHASTANT-MAILLARD. *Progesterone and betahydroxybutyrate in line measurements for a better description and understanding of Holstein cows fertility in field conditions*. 8th European Conference on Precision Livestock Farming (EC-PLF). 2017.
- [635] D. ALLOUCHE, S. DE GIVRY, G. KATSIRELOS, T. SCHIEX, M. ZYTNICKI, A. OUALI et S. LOUDNI. *Recent algorithmic advances for combinatorial optimization in graphical models*. 23rd International Symposium on Mathematical Programming (ISMP-18). 2018.
- [636] C. AMBROISE, A. DEHMAN, M. KOSKAS, P. NEUVIAL, G. RIGAILL et N. VIALANEIX. *Adjacency-constrained hierarchical clustering of a band similarity matrix with application to genomics*. Journée Régionale de Bioinformatique et Biostatistique, Génopole Toulouse. 2018.
- [646] R. GENUER, J.-M. POGGI, C. TULEAU-MALOT et N. VIALANEIX. *Random forests for big data*. Data Science, Statistics & Visualization (DSSV) 2018. 2018.
- [647] A. IMBERT, A. VALSESIA, C. ARMENISE, G. LEFEBVRE, P.-A. GOURRAUD, N. VIGUERIE et N. VIALANEIX. *Multiple hot-deck imputation for network inference from RNA sequencing data*. European Conference on Computational Biology (ECCB 2018). 2018.
- [648] A. IMBERT et N. VIALANEIX. *RNAseqNet : un package pour l'inférence de réseaux à partir de données RNA-seq*. 7ème Rencontres R. 2018.
- [649] G. LEFORT, L. LIAUBET, C. CANLET, N. VIALANEIX et R. SERVIEN. *ASICS : a new R package for identification and quantification of metabolites in complex 1D 1H NMR spectra*. European Conference on Computational Biology (ECCB 2018). 2018.
- [650] G. LEFORT, L. LIAUBET, C. CANLET, N. VIALANEIX et R. SERVIEN. *ASICS : un package R pour l'identification et la quantification de métabolites dans un spectre RMN 1H*. 7ème Rencontres R. 2018.
- [651] G. LEFORT, L. LIAUBET, P. TARDIVEL, C. CANLET, M. TREMBLAY FRANCO, L. DEBRAUWER, N. VIALANEIX et R. SERVIEN. *Using ASICS to quantify metabolites in 1D 1H NMR spectra : an application to perinatal survival in pigs*. Journée des Plateformes BioInfo et BioStat du Génotoul. 2018.

- [654] C. MESTRE, S. DJEBALI QUELEN, T. FARAUT, N. VIALANEIX, A. RAU, C. CABAU, M. ZYTNICKI, T. DERRIEN, S. LAGARRIGUE, E. GIUFFRA et S. FOISSAC. *Integrative analyses of chromosome conformation, chromatin accessibility and gene expression in human and livestock genomes*. European Conference on Computational Biology (ECCB 2018). 2018.
- [665] T. SCHIEX. *Designing molecules with cost function networks - Bridging symbolic and numerical AI*. Journées plénières du GDR IA du CNRS. 2018.
- [666] T. SCHIEX. *How Computers Break (Serious) Puzzles with logic and (a different breed of) learning*. French Académie des sciences Symposium on « Machine Learning for Artificial Intelligence ». 2018.
- [667] T. SCHIEX, S. BARBE, D. SIMONCINI, J. VUCINIC, M. RUFFINI et D. ALLOUCHE. *Constraint Programming and Graphical models - Pushing data into your models, The protein design case*. 23rd International Symposium on Mathematical Programming (ISMP-18). 2018.
- [671] N. VIALANEIX. *Learning from (dis)similarity data*. European R Users Meeting (eRum 2018). 2018.
- [675] C. AMBROISE, A. DEHMAN, P. NEUVIAL, G. RIGAILL et N. VIALANEIX. *Adjacency-constrained hierarchical clustering of a band similarity matrix with application to genomics*. Statistical Methods for Post Genomic Data (SMPGD 2019). 2019.
- [680] S. DJEBALI QUELEN, S. FOISSAC, N. VIALANEIX, K. MUNYARD, A. RAU, T. FARAUT, S. LAGARRIGUE, H. ACLOQUE et E. GIUFFRA. *Chromatin accessibility conservation across four livestock species*. Conference of the International Society for Animal Genetics (ISAG 2019). 2019.
- [681] S. FOISSAC, S. DJEBALI QUELEN, N. VIALANEIX, M. ZYTNICKI, A. RAU, S. LAGARRIGUE, H. ACLOQUE et E. GIUFFRA. *Multi-level conservation of chromosome conformation across livestock species reveals evolutionary links between genome structure and function*. Conference of the International Society for Animal Genetics (ISAG 2019). 2019.
- [682] J. FOURQUET, A. CHAUBET, H. CHIAPELLO, C. GASPIN, M. HAENNI, C. KLOPP, A. LUPO, J. MAINGUY, C. NOIROT, T. ROCHEGUE, M. ZYTNICKI, T. FERRY et C. HOEDE. *Whole metagenome analysis with metaWGS*. JOBIM 2019. 2019.
- [687] Y. LAHBIB MANSAIS, M. MARTI-MARIMON, N. VIALANEIX, S. FOISSAC, M. BOUISSOU-MATET et L. LIAUBET. *Organisation nucléaire et expression génique lors du développement chez le porc*. Réunion d'Animation épiPHASE. 2019.
- [688] G. LEFORT, L. LIAUBET, C. CANLET, H. QUESNEL, N. VIALANEIX et R. SERVIEN. *ASICS : a new R package for identification and quantification of metabolites in complex 1H NMR spectra*. useR! 2019. 2019.
- [689] G. LEFORT, L. LIAUBET, H. QUESNEL, C. CANLET, N. VIALANEIX et R. SERVIEN. *ASICS : identification and quantification of metabolites in complex 1D 1H NMR spectra*. 15th Annual Conference of the Metabolomics Society (Metabolomics 2019). 2019.
- [690] G. LEFORT, N. VIALANEIX, H. QUESNEL, M.-C. PERE, N. IANNUCELLI, C. CANLET, A. PARIS, R. SERVIEN et L. LIAUBET. *Study of fetal pig maturity in relation with neonatal survival using a multi-fluids metabolomic approach*. 15th Annual Conference of the Metabolomics Society (Metabolomics 2019). 2019.
- [692] J. MAINGUY, O. BOUCHEZ, A. CASTINEL, S. COMBES, C. GASPIN, D. MILAN, C. DONNADIEU, C. IAMPIETRO, C. HOEDE et G. PASCAL. *Identification of genomic regions for high-resolution taxonomic profiling using long-read sequencing technology*. JOBIM : Journées Ouvertes Biologie Informatique Mathématiques. 2019.

- [696] N. RANDRIAMIHAMISON, P. NEUVIAL et N. VIALANEIX. *Classification ascendante hiérarchique, contrainte d'ordre : conditions d'applicabilité, interprétabilité des dendrogrammes*. Conférence sur l'Apprentissage Automatique (CAp 2019). 2019.
- [698] T. SCHIEX. *Intelligence Artificielle : entre logique et intuition*. Journée Science pour tous. 2019.

#### 4- Produits et outils informatiques

*La liste de ces produits de la recherche n'a pas à être exhaustive. On mettra en évidence les réalisations les plus significatives.*

##### Logiciels

- [526] M. ZYTNIKI. *Tedna*. Application. 2014.
- [552] N. VIALANEIX, J. MARIETTE, M. OLTEANU, F. ROSSI, L. BENDHAIBA et J. BOELAERT. *SOMbrero : SOM Bound to Realize Euclidean and Relational Outputs (R package, available on CRAN)*. Librairie. 2015.
- [588] N. VIALANEIX, V. PICHENY et R. SERVIEN. *SISIR : Sparse Interval Sliced Inverse Regression (R package, available on CRAN)*. Librairie. 2016.
- [590] M. ZYTNIKI. *mmquant*. Application. 2016.
- [610] G. LEFORT, R. SERVIEN, P. TARDIVEL et N. VIALANEIX. *ASICS : Automatic Statistical Identification in Complex Spectra (R package, available on Bioconductor)*. Librairie. 2017.
- [612] J. MARIETTE et N. VIALANEIX. *mixKernel : Omics Data Integration Using Kernel Methods*. Librairie. 2017.
- [614] P. NEUVIAL, C. AMBROISE, S. CHATURVEDI, A. DEHMAN, M. KOSKAS, G. RIGAILL et N. VIALANEIX. *adjclust : Adjacency-Constrained Clustering of a Block-Diagonal Similarity Matrix (R package, available on CRAN)*. Librairie. 2017.
- [672] N. VIALANEIX et A. IMBERT. *RNAseqNet : Log-Linear Poisson Graphical Model with Hot-Deck Multiple Imputation (R package, available on CRAN)*. Librairie. 2018.
- [674] M. ZYTNIKI. *Rmmquant*. Application. 2018.
- [694] G. QUESNEL, S. DE GIVRY et S. MAQROT. *Baryonyx*. Application. 2019.
- [695] G. QUESNEL, S. DE GIVRY et S. MAQROT. *rbaryonyx*. Plugin. 2019.
- [701] M. ZYTNIKI et I. GONZALEZ. *srnadiff*. Application. 2019.

#### 5- Productions des personnels recrutés au cours des cinq dernières années réalisées dans leurs anciennes unités

- [706] C. BROUARD, H. SHEN, K. DÜHRKOP, F. D'ALCHÉ-BUC, S. BÖCKER et J. ROUSU. « Fast metabolite identification with Input Output Kernel Regression ». In : *Bioinformatics* 32.12 (2016), p. i28–i36.
- [707] C. BROUARD, M. SZAFRANSKI et F. D'ALCHE-BUC. « Input output Kernel regression : supervised and semi-supervised structured output prediction with operator-valued kernels ». In : *Journal of Machine Learning Research* 17 (2016).
- [708] J. KLUDAS, M. ARVAS, S. CASTILLO, T. PAKULA, M. OJA, C. BROUARD, J. JÄNTTI, M. PENTTILA et J. ROUSU. « Machine Learning of Protein Interactions in Fungal Secretory Pathways ». In : *Plos One* 11 (2016), p. 1–20.

- [710] E. L. SCHYMANSKI, C. RUTTKIES, M. KRAUSS, C. BROUARD, T. KIND, K. DÜHRKOP, F. ALLEN, A. VANIYA, D. VERDEGEM, S. BÖCKER, J. ROUSU, H. SHEN, H. TSUGAWA, T. SAJED, O. FIEHN, B. GHESQUIÈRE et S. NEUMANN. « Critical Assessment of Small Molecule Identification 2016 : automated methods ». In : *Journal of Cheminformatics* 9.1 (2017).
- [711] E. BACH, S. SZEDMAK, C. BROUARD, S. BÖCKER et J. ROUSU. « Liquid-chromatography retention order prediction for metabolite identification ». In : *Bioinformatics* 34.17 (2018), p. i875–i883.
- [720] H. SHEN, S. SZEDMAK, C. BROUARD et J. ROUSU. « Soft Kernel Target Alignment for Two-Stage Multiple Kernel Learning ». In : *Lecture Notes in Computer Science*. 19th International Conference on Discovery Science. 2016, p. 427–441.
- [723] C. BROUARD, E. BACH, S. BÖCKER et J. ROUSU. « Magnitude-Preserving Ranking for Structured Outputs ». In : Asian conference on machine learning (ACML 2017). 2017, p. 407–422.

### Outils présentés dans le cadre de compétitions de solveurs (ST uniquement)

- Logiciel d'optimisation dans les modèles graphiques Proteus/toulbar2 (<https://github.com/9thbit/uai-proteus>) à la compétition UAI 2014 (<http://auai.org/uai2014/competition.shtml>).

## 6- Activités éditoriales

*La liste de ces produits de la recherche n'a pas à être exhaustive. On mettra en évidence les réalisations les plus significatives.*

### Participation à des comités éditoriaux (journaux scientifiques, revues, collections, etc.)

Nous sommes éditeurs associés pour les revues :

- Artificial Intelligence Journal,
- Constraints,
- Journal of Artificial Intelligence Research,
- Journal of Statistical Software,
- Lifestyle Genomics.

## 7- Activités d'évaluation

*La liste de ces produits de la recherche n'a pas à être exhaustive. On mettra en évidence les réalisations les plus significatives.*

### Évaluation d'articles et d'ouvrages scientifiques (relecture d'articles / reviewing)

Nombreuses relectures d'articles pour :

- des revues internationales en statistique et en IA : Artificial Intelligence Journal, Bernouilli, Computational Optimization and Application, Computers & Operations Research, Communication in Statistics, Computational Statistics, Computational Statistics and Data Analysis, Constraints, Electronic Journal of Applied Statistical Analysis, Environmental Modelling and Software, Fuzziness and Knowledge-Based Systems IEEE Transactions on Neural



- Networks and Learning Systems, Information Processing Letters, Information Retrieval, Information Sciences, INFORMS Journal on Computing, International Journal of Reasoning-based Intelligent Systems, International Journal of Uncertainty, Journal of Artificial Intelligence Research, Journal of Machine Learning Research, Neural Networks, Neurocomputing, Pattern Recognition Letters et Theoretical Computer Science ;
- des revues internationales appliquées en biologie et en écologie : Biosystems, BMC Genomics, BMC Bioinformatics, Bioinformatics, Genetics and Molecular Research, Genomics, IEEE Journal of Biomedical and Health Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Molecular Ecology, Nucleic Acids Research, PLoS ONE, RNA, Statistical Applications in Genetics and Molecular Biology, Trends in Plant Science et World Rabbit Science ;
  - des conférences nationales : Journées Francophones de Programmation par Contraintes (JFPC 2014–2016, 2018–2019), Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes (JFRB 2014, 2016, 2018), Journées Ouvertes Biologie, Informatique et Mathématiques (JOBIM 2015–2017, 2019), Journées de Statistique de la SFdS 2016, Rencontres R 2015 et Rencontres de la Société Francophone de Classification 2014 ;
  - et des conférences internationales : AAAI Conference on Artificial Intelligence (AAAI 2014–2018), International Conference on Algorithms for Computational Biology (AlcoB 2014), Conference on Neural Information Processing Systems (NIPS 2019), Conference on Uncertainty in Artificial Intelligence (UAI 2014), European Conference on Artificial Intelligence (ECAI 2014), European Conference on Computational Biology (ECCB 2014, 2016, 2018), European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2014–2019), International Conference on Computational Statistics (COMPSTAT 2016), International Conference on Principles and Practice of Constraint Programming (CP 2014–2019), International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR 2014–2016), International Conference on Theory and Applications of Satisfiability Testing (SAT 2018), International Conference on Tools with Artificial Intelligence (ICTAI 2014–2016, 2018), International Joint Conference on Artificial Intelligence (IJCAI 2015–2019), International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM 2015, 2017, 2019).

### Évaluation de projets de recherche

Rapports d'expertise ANR, pour la région Auvergne Rhône Alpes (2019), pour la région Pays de la Loire, pour des métaprogrammes INRA. Evaluation d'un projet de recherche canadien (2016) et d'un ERC (remote referee).

### Évaluation de laboratoires

- 2014 : MICALIS, Jouy-en-Josas,
- 2018 : URGI, Versailles,
- 2019 : CRIL, Lens.

### Responsabilités au sein d'instances d'évaluation

- Membre de la CNU 26 (2011–2018),
- Membre de la CSS MBIA (INRA, 2011–2020), puis du bureau de la CSS MBIA (2015–2020),
- Membre de la Commission d'évaluation de l'Inria (2015–2018),

— Membre de la Commission d'évaluation des Ingénieurs INRA (CEI) MPR1&6 (2016) MPR3&6 (2017).

### Autres contrats européens en tant que partenaire

Contrats gérés dans l'unité MIAT				
Nom	Début	Fin	Financier	Montant
SeqOccIn	01/01/2019	31/12/2019	FEDER	4 800 000,00€

### Contrats nationaux (ANR, PHRC, FUI, INCA, etc.) en tant que partenaire

Contrats gérés dans l'unité MIAT				
Nom	Début	Fin	Financier	Montant
memRNase	01/01/2014	31/12/2018	ANR	86 720,00€
SAL HOST TROP	01/12/2015	31/05/2020	ANR	169 201,00€
CASPAR	01/10/2016	02/10/2020	ANR	65 520,00€
DE-MO-GRAPH	16/01/2017	15/01/2022	ANR	133 829,00€
METAhCOL	01/10/2017	30/09/2020	ANR	29 448,00€
ANTI-SELFISH	01/01/2019	31/03/2020	LABEX ECOFECT	40 500,00€
FICOLOFO	01/12/2010	31/11/2014	ANR	56 602,99€
BACNET	01/11/2011	31/10/2016	ANR	113 530,79€

### Contrats avec les collectivités territoriales en tant que porteur

Contrats gérés dans l'unité MIAT				
Nom	Début	Fin	Financier	Montant
Doctorant (Clément Viricel)	15/05/2014	01/01/2018	Région	47 672,00€
Doctorant (Franck Cerutti)	15/05/2015	01/01/2018	Région	50 691,00€
Doctorant (Alyssa Imbert)	04/06/2015	01/01/2019	Région	51 700,00€
Doctorant (Walid Ben Saoud Benjerri) <sup>1</sup>	27/05/2016	01/09/2018	Région	51 700,00€
Doctorante (Jelena Vucinic)	01/10/2017	01/01/2021	Région	55 320,00€

### Contrats financés dans le cadre du PIA en tant que partenaire

Contrats gérés dans l'unité MIAT				
Nom	Début	Fin	Financier	Montant
BACNET	01/11/2011	31/10/2016	PIA BIOINFOR- MATIQUE	137 280,00€

1. Théoriquement du 27/05/2016 au 01/01/2020.

<b>AMAIZING</b> <sup>2</sup>	01/10/2011	30/09/2019	PIA	1 075,00€
<b>SUNRISE</b>	01/09/2012	30/06/2020	PIA	115 406,00€
<b>AKER</b> <sup>2</sup>	01/01/2013	31/12/2019	PIA BIOTECHNO- LOGIES ET BIORESSOURCES	145 600,00€

---

#### Contrats gérés hors de l'unité MIAT

Nom	Début	Fin	Financier	Montant
<b>AMAIZING</b> <sup>2</sup>	01/10/2011	30/09/2019	PIA	1 075,00€
<b>AKER</b> <sup>2</sup>	01/01/2013	31/12/2019	PIA BIOTECHNO- LOGIES ET BIORESSOURCES	145 600,00€

---

## 8- Post-doctorants et chercheurs accueillis

*On précisera ici les faits observables pertinents pour l'évaluation : nombre, financements, durée etc. On pourra aussi donner une sélection des publications les plus significatives dans ce domaine.*

**Post-doctorants (nombre total) : 6**

**Post-doctorants étrangers : 1**

Obtention de la première place à la compétition internationale d'optimisation sur les modèles graphiques à UAI 2014 (logiciel Proteus/toulbar2) [112].

**Chercheurs accueillis (nombre total) : 3**

**Chercheurs étrangers accueillis : 2**

## 9- Indices de reconnaissance

*On s'en tiendra aux indices de reconnaissance les plus significatifs pour le rayonnement et l'attractivité de l'unité de recherche (équipe / thème).*

### Prix et/ou Distinctions

- 2015 : Keynote Speaker Award (ICTAI 2015) ;
- 2016 : Fellow of the European Association for Artificial Intelligence ; Distinguished Service Award of the Association for Constraint Programming (CP 2016) ;
- 2018 : Distinguished Senior PC member (IJCAI 2018) ; Best Paper Award (CP 2018).

### Responsabilités dans des sociétés savantes

- Membre du bureau du GDR MASCOT NUM (2011-2013) ;
- Membre du bureau (depuis 2014) et vice-présidente (depuis 2017) de la Société Française de BioInformatique (SFBI) ;
- Trésorier de l'Association for Constraint Programming (2008-2015) ;
- Webmestre (2011-2016), membre du conseil (2013-2017) et du bureau (2014-2017) de la Société Française de Statistique (SFdS).

---

2. Transféré au LIPM en 2015.

**Organisations de colloques / congrès à l'étranger**

- 2016 : Local chair of the 22th International Conference on Principles and Practice of Constraint Programming (CP 2016, Toulouse, France) ;
- 2019 : Program co-chairs of the 25th International Conference on Principles and Practice of Constraint Programming (CP 2019, Stamford, USA) ;
- 2019 : Local chair of the 15th R User Conference (useR! 2019, Toulouse, France).

**Invitations à des colloques / congrès à l'étranger**

- 2014 : Workshop on Discrete Graphical Models, Heidelberg, Germany ;
- 2016 : 9th International Conference of the ERCIM WG on Computational and Methodological Statistics, Seville, Spain ; conférence invitée pour l'éméritat de Marc De Maeyer, KUL, Leuven, Belgique ; Prague Protein Spring, Pragma, Tchèque ;
- 2017 : Formateur invité dans l'école internationale d'hiver *Algorithms in Structural Bioinformatics* ; workshop international *Deciphering complex energy landscape and kinetic network from single molecules to cells*, Dijon, France ;
- 2018 : European R Users Meeting, Budapest, Hungary ; Académie des sciences *Symposium on Artificial Intelligence*, Paris, France ; 23rd International Symposium on Mathematical Programming, Bordeaux, France ;
- 2019 : 16th International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research, Thessaloniki, Greece ; 7th International Chemistry Symposium SIQ - SEADIM, Toulouse, France ; 12th Molecular Design and Bioinformatics Symposium, Cayo Santa Maria, Cuba.

**Séjours dans des laboratoires étrangers**

- 2016 : Institute for Computational Protein Design, Seattle, USA ; CalTech, California, USA ;
- 2017 : Dept. Physics, University of Nagoya, Japan ; Riken, Kobe, Japan ; Research Institute for Electronic Science, Hokkaido University, Japan ; Universitat Lleida, Spain ;
- 2018 : University of Toronto, Canada ; Monash University, Melbourne, Australia.

**II – INTERACTION DE L'ÉQUIPE AVEC L'ENVIRONNEMENT NON ACADÉMIQUE, IMPACTS SUR L'ÉCONOMIE, LA SOCIÉTÉ, LA CULTURE, LA SANTÉ**

**Contrats de R&D avec des industriels**

<b>Contrats gérés dans l'unité MIAT</b>				
<b>Nom</b>	<b>Début</b>	<b>Fin</b>	<b>Financeur</b>	<b>Montant</b>
<b>Doctorant (Alyssa Imbert)</b>	01/11/2015	30/10/2018	METHODOMICS	47 300,00€
<b>FSOV Criblage</b>	01/10/2012	31/03/2016	FLORIMOND DESPREZ/GNIS	11 550,95€

**1- Activités d'expertise scientifique**

## Activités de consultant

- 2017 : Membre du sous-groupe recherche chargé de proposer des actions pour la Stratégie en Intelligence Artificielle pour la France initié par Thierry Mandon (secrétaire d'Etat en charge de l'Enseignement Supérieur et de la Recherche) et Axelle Lemaire (secrétaire d'Etat en charge du Numérique et de l'Innovation) ;
- 2018 : Expert interrogé par le comité Cédric Villani sur l'Intelligence Artificielle.

## 2- Produits destinés au grand public

### Produits de vulgarisation : articles, interviews, éditions, vidéos, produits de médiation scientifique, débats science et société, etc.

- 2014 : Participation à la fête des maths, lycée Jules Fil, Carcassonne ; article pour des professionnels *Viandes Et Produits Carnés* ; participation à l'ouvrage *Brèves de Maths* ;
- 2015 : Coordination pour la SFdS du zoom ONISEP sur les métiers des mathématiques et de l'informatique (voir <http://www.onisep.fr/Publications/Zoom/Les-metiers-des-mathematiques-et-de-l-informatique>) ;
- 2018 : Participation à Sciences en fête (organisation d'ateliers en Intelligence Artificielle et conception de protéines), avec une présentation *How computers break (serious) puzzles* ;
- 2019 : Article dans la revue *La recherche* décrivant nos résultats en conception algorithmique de protéines (avec la protéine *Ika*, qui s'auto-assemble) ; article pour des professionnels *Le Point Vétérinaire.Fr*.

## III – IMPLICATION DE L'ÉQUIPE DANS LA FORMATION PAR LA RECHERCHE

*La liste de ces produits et activités n'a pas à être exhaustive. On mettra en évidence les réalisations les plus significatives.*

### Productions scientifiques (articles, ouvrages, etc.) issus des thèses

Nombre moyen d'articles par doctorant : 3

#### 1- Formation

Nombre de personnes Habilitées à Diriger des Recherches (HDR) : 4

Nombre d'HDR soutenues : 2

Doctorants (nombre total) : 16

Doctorants bénéficiant d'un contrat spécifique au doctorat : 3 (moniteur)

Nombre de thèses soutenues : 10

Durée moyenne des thèses : 3 ans et 3 mois

Stagiaires accueillis (M1, M2) : 31