

Concours CNRS
Directeur de Recherche 2^{ème} classe
CID 51

Franck Picard

LBBE, Univ. Lyon 1, F-69622 Villeurbanne, France

24 Avril 2015

Outline

- 1 **Présentation**
- 2 Positionnement de la recherche
- 3 Exemples de résultats obtenus
- 4 **Projet: Modèles spatiaux pour les données de séquençage à haut débit**

Éléments de Curriculum Vitæ

Parcours professionnel

- | | |
|-------|-----------------------------------------------------------------------------------------------------------|
| 2014 | Habilitation à diriger des recherches , Univ. Lyon 1 5 thèses en co-direction |
| 2007- | Chargé de recherche CNRS , Sections 41-51, Laboratoire de Biométrie et Biologie Évolutive, Lyon |

Principales responsabilités

- | | |
|-----------|------------------------------------------------------------------------------------------------------|
| 2014- | Responsable d'équipe (2 CR, 2 Doct., 2 postdocs) |
| 2012-2016 | ABS4NGS, Projet investissement d'avenir Responsable du WP "Epigenomics" |
| 2011-2013 | Porteur d'un Projet PEP2 CNRS |
| 2013- | Master 2 Mathématique en Action, Univ. Lyon 1 ▷ Statistique en grande dimension pour la génomique |
| 2012- | License 3 Biologie, ENS Lyon ▷ Module de Biostatistique |

Outline

- 1 Présentation
- 2 Positionnement de la recherche**
- 3 Exemples de résultats obtenus
- 4 Projet: Modèles spatiaux pour les données de séquençage à haut débit

Biologie

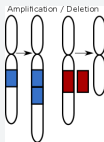
Statistique

Logiciels

Biologie

Variations du nombre de copies

1 review



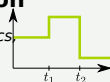
Statistique

Segmentation

5 articles (*Biometrics*,
Biostatistics, *CSDA*)



thèse M. Giacofci

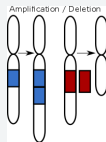


Logiciels

Biologie

Variations du nombre de copies

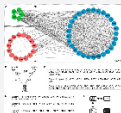
1 review



Réseaux biologiques

2 articles

(*Mol. Sys. Biol.*,
Meth. Ecol. Evol.)



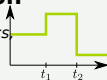
Statistique

Segmentation

5 articles (*Biometrics*,
Biostatistics, *CSDA*)



thèse M. Giacofci



Modèles de graphes aléatoires

5 articles (*AOAS*,
Stat& Comp,
Bioinformatics)
ANR Nemo

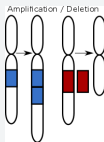


Logiciels

Biologie

Variations du nombre de copies

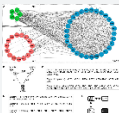
1 review



Réseaux biologiques

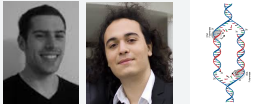
2 articles

(*Mol. Sys. Biol.*,
Meth. Ecol. Evol.)



Réplication de l'ADN

1 article (*PLoS Genetics*)



Thèses: G. Durif, S. Ivanoff

Statistique

Segmentation

5 articles (*Biometrics*,
Biostatistics, *CSDA*)

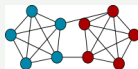


thèse M. Giacofci



Modèles de graphes aléatoires

5 articles (*AOAS*,
Stat& Comp,
Bioinformatics)
ANR Nemo



ANR

Modélisation NGS

- détection de pics
- SNP liés au sexe

ANR

Institut des sciences biologiques
Centre national de la recherche scientifique

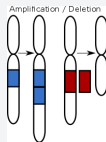


Logiciels

Biologie

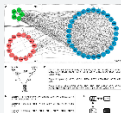
Variations du nombre de copies

1 review



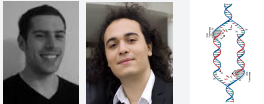
Réseaux biologiques

2 articles
(*Mol. Sys. Biol.*,
Meth. Ecol. Evol.)



Réplication de l'ADN

1 article (*PLoS Genetics*)



Thèses: G. Durif, S. Ivanoff

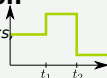
Statistique

Segmentation

5 articles (*Biometrics*,
Biostatistics, *CSDA*)



thèse M. Giacofci



Modèles de graphes aléatoires

5 articles (*AOAS*,
Stat& Comp,
Bioinformatics)
ANR Nemo



Modélisation NGS

- détection de pics
- SNP liés au sexe



Institut des sciences biologiques
Centre national de la recherche scientifique



Logiciels

Packages R



Calcul parallèle, C++



grandes masses de données



Questions Biologiques

Expression

T. Burger, J. Mold, C. Vaissé



University of California
San Francisco



Karolinska
Institutet

Déterminisme du sexe

G. Marais



Résistances bact.

J.-B. Veyrieras, P. Mahé



BIOMÉRIEUX
INDUSTRY

Réplication

M. N. Prioleau

Statistique

Statistique en grande dimension pour la génomique

2 CR, 2 thèses, 2 postdocs



LBBE
BIOMETRIE ET BIOLOGIE EVOLUTIVE

Bioinformatique

Questions Biologiques

Expression

T. Burger, J. Mold, C. Vaisse



University of California
San Francisco



Karolinska
Institutet

Déterminisme du sexe

G. Marais



Résistances bact.

J.-B. Veyrieras, P. Mahé



BIOMÉRIEUX
INDUSTRY

Réplication

M. N. Prioleau

Statistique en grande dimension pour la génomique

2 CR, 2 thèses, 2 postdocs



LBBE
BIOMETRIE ET BIOLOGIE EVOLUTIVE

Statistique

Réduction de dimension

S. Lambert-Lacroix, J. Mairal, V. Rivoirard

Apprentissage statistique

J. Mairal, J.-P. Vert



Processus ponctuels

P. Reynaud-Bouret



Graphes

V. Lacroix

Bioinformatique

Questions Biologiques

Expression

T. Burger, J. Mold, C. Vaissé



University of California
San Francisco



Karolinska
Institutet

Déterminisme du sexe

G. Marais



Résistances bact.

J.-B. Veyrieras, P. Mahé



BIOMÉRIEUX
INDUSTRY

Réplication

M. N. Prioleau

Statistique

Réduction de dimension

S. Lambert-Lacroix, J. Mairal, V. Rivoirard

Apprentissage statistique

J. Mairal, J.-P. Vert



Processus ponctuels

P. Reynaud-Bouret



Graphes

V. Lacroix

Statistique en grande dimension pour la génomique

2 CR, 2 thèses, 2 postdocs



LBBE
BIOMÉTRIE ET BIOLOGIE ÉVOLUTIVE

Projet ABS4NGS

Curie, AgroParisTech, MinesParisTech,
Marne-la Vallée, LyonSynergie, Genostar

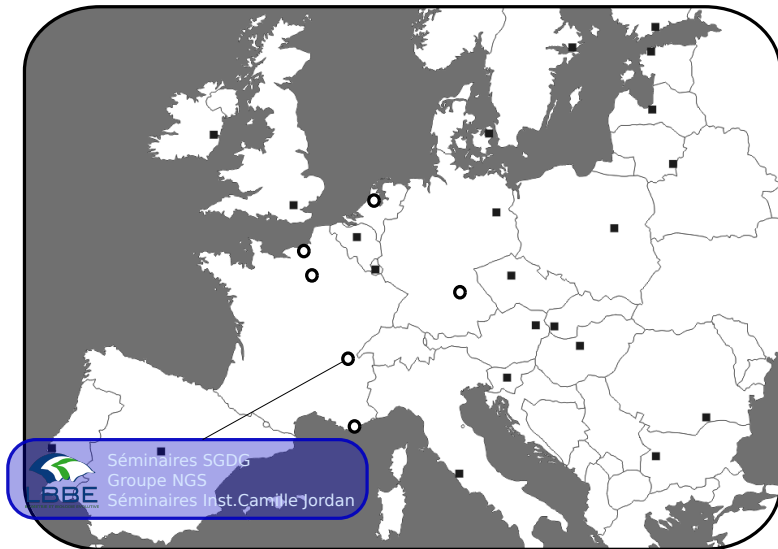


Plateforme PRABI

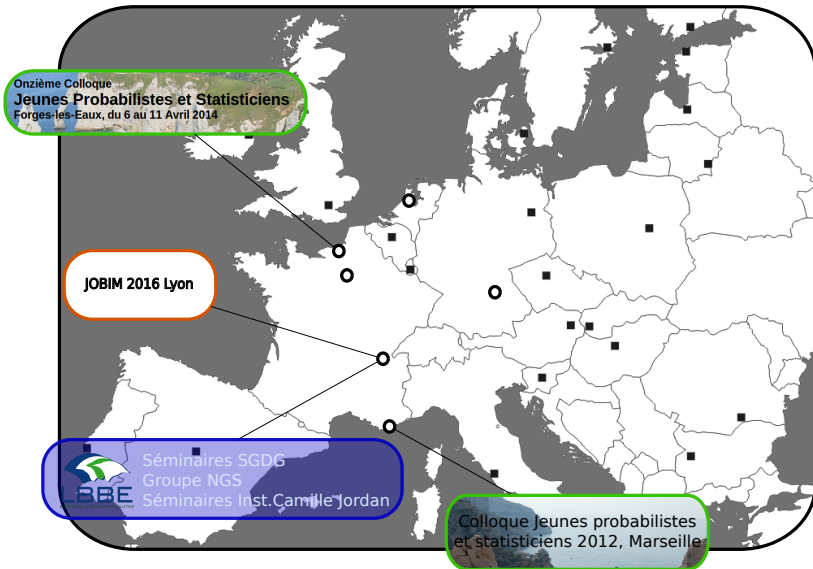


Bioinformatique

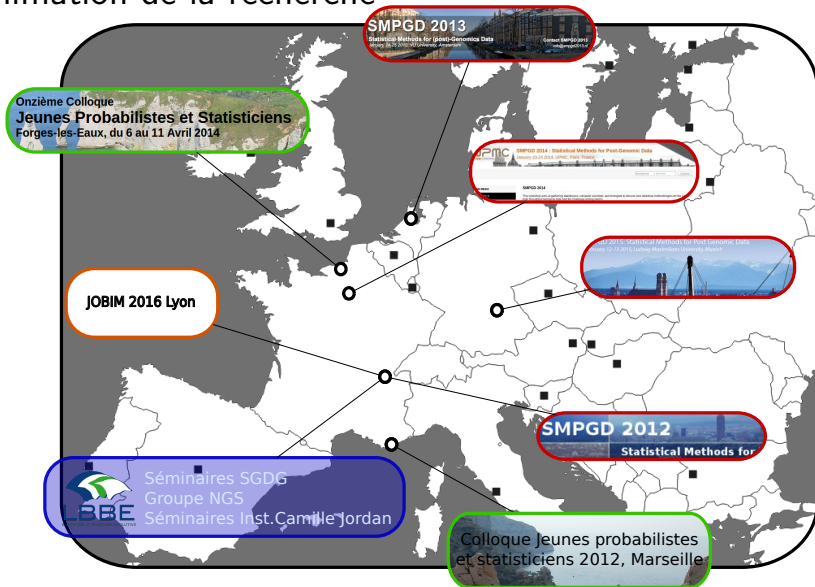
Animation de la recherche



Animation de la recherche



Animation de la recherche

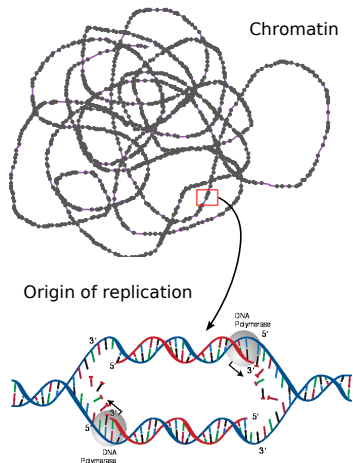


Outline

- ① Présentation
- ② Positionnement de la recherche
- ③ Exemples de résultats obtenus
- ④ Projet: Modèles spatiaux pour les données de séquençage à haut débit

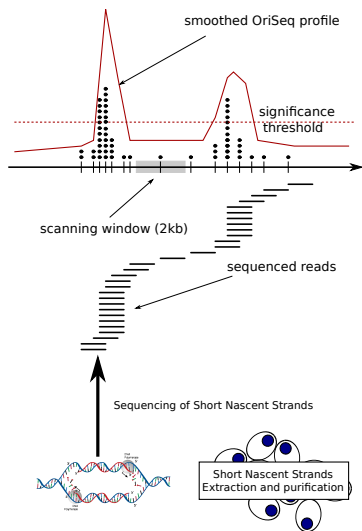
Cartographie des origines de réplication du génome humain

- Réplication du génome humain sous contrôle:
 - spatial (choix de l'origine)
 - temporel (cycle cellulaire)
- Quels sont les déterminants (épi)génétiques de ces contrôles ?
- Collaboration avec M.-N.Prioleau (IJM), L. Duret (LBBE)
- *PloS Genetics* (2014)



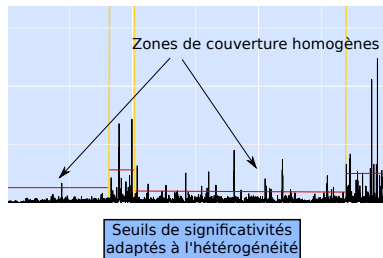
Développements méthodologiques

- Détection d'accumulations exceptionnelles de reads



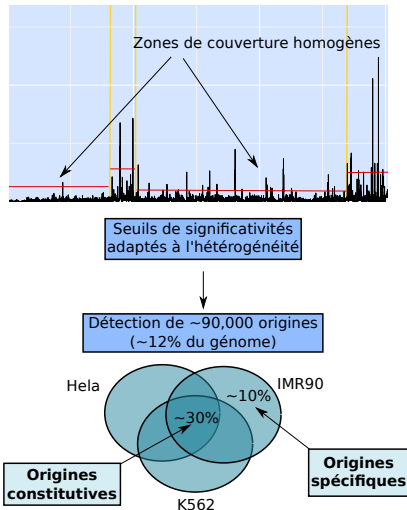
Développements méthodologiques

- Détection d'accumulations exceptionnelles de reads
- Prise en compte de l'hétérogénéité de couverture
- Contrôle du niveau du test (tests multiples)



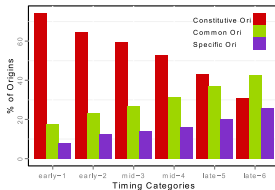
Développements méthodologiques

- Détection d'accumulations exceptionnelles de reads
- Prise en compte de l'hétérogénéité de couverture
- Contrôle du niveau du test (tests multiples)
- Validation expérimentale et bioinformatique
- Détection d'un ensemble d'origines de réplication sur plusieurs lignées cellulaires



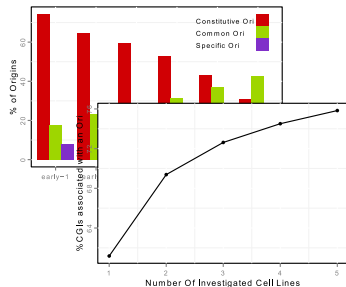
Caractérisation génétique et épi-génétique des origines de réplication

- Les origines constitutives précoces associées aux îlots CpG constituent un pool d'origines "drivers"



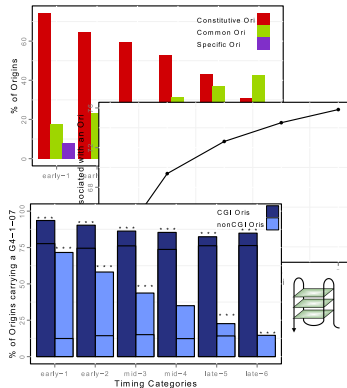
Caractérisation génétique et épi-génétique des origines de réplication

- Les origines constitutives précoces associées aux îlots CpG constituent un pool d'origines "drivers"
- 75% des îlots CpG sont associés à au moins une origine



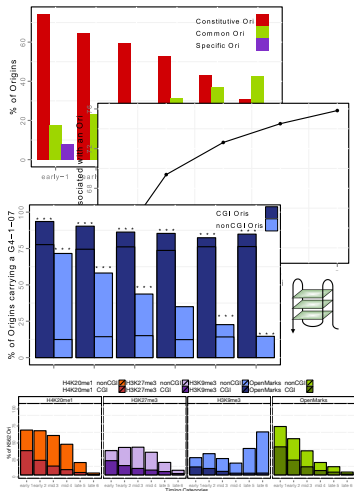
Caractérisation génétique et épi-génétique des origines de réplication

- Les origines constitutives précoces associées aux îlots CpG constituent un pool d'origines "drivers"
- 75% des îlots CpG sont associés à au moins une origine
- Les origines sont fortement associées aux G-quadruplex



Caractérisation génétique et épi-génétique des origines de réplication

- Les origines constitutives précoces associées aux îlots CpG constituent un pool d'origines "drivers"
- 75% des îlots CpG sont associés à au moins une origine
- Les origines sont fortement associées aux G-quadruplex
- L'association des origines avec certaines marques chromatinienne est dynamique

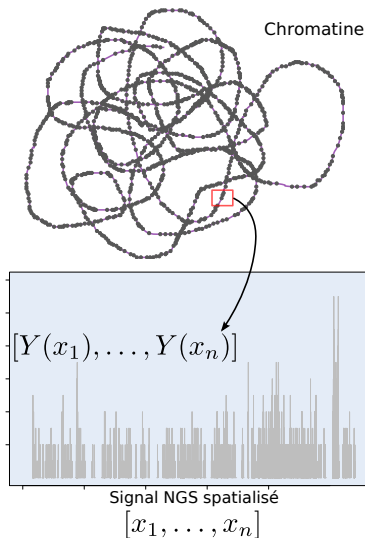


Outline

- ① Présentation
- ② Positionnement de la recherche
- ③ Exemples de résultats obtenus
- ④ **Projet: Modèles spatiaux pour les données de séquençage à haut débit**

Des processus biologiques spatialisés

- Modifications de la chromatine, recombinaison, **réplication**
- Données NGS spatialisées le long des génomes
- Quel cadre statistique ?
- Approches existantes: fenêtres, HMM
- Intégration de données ?
Régression ? Comparaison ?



Modélisation fonctionnelle pour l'analyse des données NGS

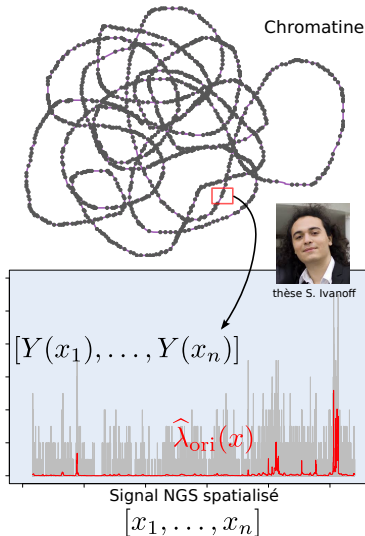
- Modèle de Poisson fonctionnel:

$$Y(x_t) \sim \mathcal{P}(\lambda(x_t))$$

- $\lambda(x)$ une fonction d'intensité décomposée sur une base de fonctions

$$\log \lambda(x) = \sum_{j=1}^p \beta_j \phi_j(x)$$

- Défi: sélection des β_j
- Cadre pour l'intégration de données



Modélisation fonctionnelle pour l'analyse des données NGS

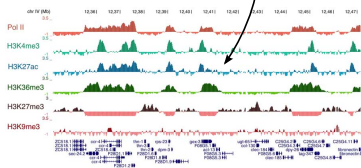
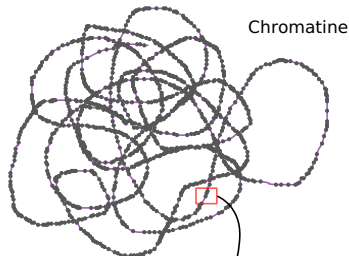
- Modèle de Poisson fonctionnel:

$$Y(x_t) \sim \mathcal{P}(\lambda(x_t))$$

- $\lambda(x)$ une fonction d'intensité décomposée sur une base de fonctions

$$\log \lambda(x) = \sum_{j=1}^p \beta_j \phi_j(x)$$

- Défi: sélection des β_j
- Cadre pour l'intégration de données



$$Y_{\text{ori}}(x) = f(Y_{\text{H3H4me3}}(x), \dots, Y_{\text{H3H27me3}}(x))$$

$$\lambda_{\text{ori}}(x) = \sum_{k=1}^K \alpha_k(x) \lambda_k(x)$$

Modélisation fonctionnelle pour l'analyse des données NGS

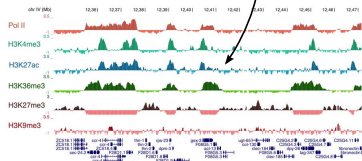
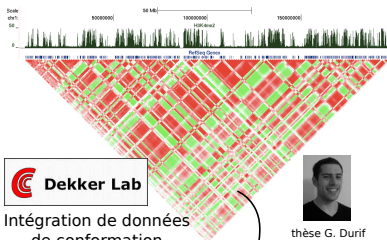
- Modèle de Poisson fonctionnel:

$$Y(x_t) \sim \mathcal{P}(\lambda(x_t))$$

- $\lambda(x)$ une fonction d'intensité décomposée sur une base de fonctions

$$\log \lambda(x) = \sum_{j=1}^p \beta_j \phi_j(x)$$

- Défi: sélection des β_j
- Cadre pour l'intégration de données



$$\lambda_{\text{ori}}(x) = \sum_{k=1}^K \alpha_k(x) \lambda_k(x)$$