



Annuaire des  
plateformes  
technologiques

## **Plateformes de Criblage à Haut Débit ..... 2**

PlateformeARIADNE_HCS U761 CIIL.....	3
Criblage à haut débit et chimie combinatoire - CEA.....	6
CIBLOT: Ciblage, Interface Biologie-Chimie et Laboratoire Opérationnel de Transfert.....	8
Assay development and chemical library screening team.....	10
Criblage primaire de protéine fluorescente sur puces à petites molécules .....	13
CMBA (Criblage pour des Molécules Bio-Actives).....	15
Antiviral Drug Design Platform-AD2P .....	19
PlateformeARPEGE-Pharmacologie-Criblage-Interactome .....	21
ImPACcell (Imagerie Pour Analyse du Contenu cellulaire) .....	24
PICT - Plateforme Intégrée de Criblage de Toulouse.....	26
Lead Optimization .....	29
TechMed <sup>ILL</sup> .....	31
PCBIS (Plate-forme de Chimie Biologique Intégrative de Strasbourg) .....	33

## **Plateformes de Séquençage Haut Débit, Génomique, Transcriptomique ..... 37**

GenoScreen SAS.....	38
Plateforme Gentyane Génotypage et Séquençage à Haut Débit.....	41
INTEGRAGEN GENOMIC SERVICES.....	46
Transcriptomique et Génomique Marseille-Luminy (TGML).....	56
MGX-MontpellierGenomiX.....	60
Biogenouest Génomique .....	63
Plateforme de Nice Sophia Antipolis .....	64
Plateforme Génomique Imagine & Plateforme de Bioinformatique Paris-Descartes.....	71
Plateforme de génomique et microgénomique : profileXpert .....	78
Plateforme de Génomique - Lille .....	90
Plateforme biopuces et séquençage .....	100
Plateforme Genome et Transcriptome Genotoul .....	109
Plateforme Génome Transcriptome de Bordeaux - PGTB.....	115

**Fiches d'identité:**

# **Plateformes de Criblage à Haut Débit**

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### Plateforme ARIADNE\_HCS U761 CIIL

<b>Responsable</b>	Benoît Déprez
<b>Structure de rattachement</b>	Inserm/Université Lille 2/Institut Pasteur de Lille
<b>Adresse complète</b>	Inserm U761 Institut Pasteur de Lille 1 rue du professeur Calmette 59019 Lille
<b>Site Web</b>	<a href="http://www.deprezlab.fr/">http://www.deprezlab.fr/</a> <a href="http://www.cdithem.fr/">http://www.cdithem.fr/</a>
<b>Contact</b>	Florence Leroux
<b>Date de mise en place de la PTF</b>	2005
<b>Labels / démarche qualité</b>	
IBiSA	oui
Certification ISO 9001	non
<b>Réseaux (national, international)</b>	GDR_3056
<b>Principales réalisations</b>	
nombre de projets traités depuis la création	24
nombre de projets traités par an	1 à 4
domaines de compétence / aires thérapeutiques	En fonction de la nature de la cible thérapeutique Historiquement : Maladies métaboliques & infectieuses.
<b>Equipe (effectifs)</b>	
chercheurs	2
ingénieurs	2
techniciens	3
administratifs	1,5
autres (préciser)	contractuels et étudiants en nombre variable

## Chimiothèque

### Description de la collection :

70 000 molécules chimiques dont 10 000 préparés dans l'équipe U761 et 60 0000 achetés auprès de différents fournisseurs (Asinex, ChemDiv, Enamine ...)

les 3/4 sont des molécules "drug-like", le reste des molécules "lead-like"

Focus "interactions protéine-protéine" pour 1/3 des produits.

Focus "cibles à ligand acide" pour 1/10 des produits.

toctriscreen total: 1120 biologically active compounds

toctriscreen stem cell modulators: 44 compounds

prestwick chemical library: 1280 FDA approved drug

librairie de 960 fragments

### Conditionnement :

Stock poudre en tube "matrix" + 1 copie en solution (DMSO) en plaques 96 puits SBS

1 copie en solution (DMSO) en plaques 384 puits Echo compatible

code à barres sur les plaques et code 2D sur les tubes

### Base de données :

Gestion de données : base ISIS pour les structures et bases access pour le compound management et le stockage des résultats. Les bases sont interrogeables grâce au logiciel Pipeline Pilot (Accelrys) et via le web.

**Membre de la chimiothèque nationale :** non

## Cibles

Cibles destinées au service de criblage Récepteur, enzyme, interaction prot-prot, cellules

## Criblage à haut débit

Nombre approximatif de mesures/ jour 4000 points / jours, variable selon les essais

Tests biologiques proposés :

tests enzymatiques in vitro, interactions protéines/protéines in vitro ou in cellulo, tests cellulaires (gene reporter, sécrétion, analyses d'image ...)

## Technologies / Equipements

multimode readers equipped with stackers (Mithras LB, Victor 3V, Polarstar Omega)

Xmap technology (Magpix)

Thermal shift assay (Light Cycler 480II)

high content screening platforms (1 system en L2 lab , 1 system en P3)

liquid handler (Cybi-well, Zephyr)

acoustic dispenser Echo 550

HPLC-mass spectrometry

## Activité Post-criblage

Hit-2-Lead : Chimie médicinale, Caractérisation physico-chimique, Optimisation ADME-PK

## Valorisation

Mathieu, A.L. et al., Identification of small inhibitory molecules targeting the Bcl-1 anti-apoptotic protein that alleviate resistance to ABT-737, *Journal Biomolecular Screening*, **2014**.

Charton, J. et al., Imidazole-derived 2-[N-carbamoylmethyl-alkylamino]acetic acids, substrate-dependent modulators of insulin-degrading enzyme in amyloid- $\beta$  hydrolysis. *European Journal of Medicinal Chemistry*, **2014**.

Maingot, L. et al., Aggrecanase-2 Inhibitors Based on the Acylthiosemicarbazide Zinc-Binding Group. *Eur J Med Chem*, **2013**.

Gonzalez, S., et al., Rescue of nonsense mutations by amlexanox in human cells. *Orphanet Journal of Rare Diseases*, **2012**.

Flipo, M., et al., Discovery of novel N-phenyl-phenoxyacetamide derivatives as EthR inhibitors and ethionamide boosters by combining High-Throughput Screening and Synthesis. *Journal of Medicinal Chemistry*, **2012**.

Elbakali, J., et al., Novel Selective Inhibitors of Neutral Endopeptidase: Discovery by Screening and Hit-to-Lead Optimisation. *Med.Chem.Comm.*, **2012**.

Maingot, L. et al., New non-hydroxamic ADAMTS-5 inhibitors based on the 1,2,4-triazole-3-thiol scaffold. *Bioorganic & Medicinal Chemistry Letters*, **2010**.

Deprez-Poulain, Rebecca; Charton, Julie; Leroux, Florence; Deprez, Benoit, Gauriot, Marion; Tang, Wei-Jen; Totobenazara, Jane "Ligands of Insulin-Degrading Enzyme and their uses". Mar 2010.

Deprez Benoit, Willand Nicolas, Dirie Bertrand, Toto Patrick "Compounds having a potentiating effect on the activity of ethionamide and uses thereof" WO2008003861. 10th Jan. 2008.

Beghyn Terence [fr]; Deprez Benoit [fr]; Gonzalez-Hilarion Sara Sofia [fr]; Lejeune Fabrice "composé utile pour le traitement de maladies médiées par une mutation non-sens et composition pharmaceutique comprenant ledit composé" wo 2012016930, 2012-02-09

## Accès

Equipes du site	oui
Equipes extérieures académiques	oui

## Formation

Néant. Prévision pour le futur.

Mise à jour

06/10/2014

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### Criblage à haut débit et chimie combinatoire - CEA

<b>Responsable</b>	Jean-Christophe Cintrat
<b>Equipe (effectifs)</b>	2
chercheurs	1
techniciens	1
<b>Structure de rattachement</b>	CEA
<b>Adresse complète</b>	CEA Saclay DSV-Ibitecs-SCBM Bâtiment 547 91191 Gif sur Yvette Cedex
<b>Site Web</b>	<a href="http://www-dsv.cea.fr/ibitecs/scbm/lcb/chimie-criblages">http://www-dsv.cea.fr/ibitecs/scbm/lcb/chimie-criblages</a>
<b>Contact</b>	Jean-Christophe Cintrat
<b>Date de mise en place de la PTF</b>	2004
<b>Labels / démarche qualité</b>	
IBiSA	OUI
Certification ISO 9001	NON
<b>Réseaux (national, international)</b>	Chimiothèque nationale GDR criblage G5C (CEA Grenoble)
<b>Principales réalisations</b>	
nombre de projets traités depuis la création	10 projets de criblage 20 projets d'optimisation de molécules
nombre de projets traités par an	3
domaines de compétence / aires thérapeutiques	Cancer Toxines Résistances bactériennes Transporteurs

## Chimiothèque

Description de la collection	100 000 composés favorisant la diversité chimique, libres de droit.
Passerelles vers d'autres collections	Chimiothèque nationale
Conditionnement	Dans le DMSO, en plaques 96 puits
Base de données (structure, conditions d'accessibilité)	Partenariat avec des structures de criblage in silico
Membre de la chimiothèque nationale	OUI depuis 2010

## Cibles

Cibles destinées au service de criblage	La plateforme peut travailler sur tous les types de cibles
---	--

## Activités

Culture cellulaire

## Criblage à haut débit

Nombre approximatif de mesures/ jour	2000
Tests biologiques proposés	Tests enzymatiques in vitro Tests sur cellules (essais de liaison, survie cellulaire, synthèse protéique....) Tests de compétition Radioactivité

## Technologies / Equipements

Robot de criblage, automate de pipetage et dilution  
Absorbance, luminescence, fluorescence, HTRF, radioactivité (3H, C14, S35, I125)

## Activités post-criblage

Optimisation de hits (amélioration de l'activité biologique et des propriétés physico-chimiques)  
Identification de cible

## Accès

Equipes du site	OUI
Equipes extérieures académiques	OUI
Equipes extérieures privées	OUI

*Mise à jour* 06/10/2014

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### CIBLOT: Ciblage, Interface Biologie-Chimie et Laboratoire Opérationnel de Transfert

<b>Responsable scientifique</b>	Catherine Brenner
<b>Equipe (effectifs)</b>	
Chercheurs	1
Ingénieurs	2
<b>Structure de rattachement</b>	IFR141-IPSIT
<b>Adresse complète</b>	CIBLOT, Faculté de Pharmacie, Université Paris Sud, 5 rue Jean-Baptiste Clément, Tour D1 92290 Châtenay-Malabry
<b>Site Web</b>	<a href="http://www.ciblot.u-psud.fr">www.ciblot.u-psud.fr</a>
<b>Contact</b>	Catherine Brenner
<b>Date de mise en place de la PTF</b>	2008
<b>Labels / démarche qualité</b>	
IBiSA	non
Certification ISO 9001	non
<b>Réseaux (national, international)</b>	GDR CNRS 3056: ChemBioScreen; LabEx LERMIT
<b>Principales réalisations</b>	
Nombre de projets traités depuis la création	4
Nombre de projets traités par an	4
Domaines de compétence / aires thérapeutiques	Maladies cardiovasculaires, hématologie, cancérologie, obésité

## Chimiothèques

Description de la collection	GreenPharma (7200 molécules), BIOCIS (1578 molécules) Prestwick (1120 molécules) Passerelles vers d'autres collections: Chimiothèque Nationale Essentielle (640 molécules), ICSN (4400 molécules)
Conditionnement	Plaques 96 puits, solutions environ 10mM DMSO
Base de données	ISIS base
Membre de la chimiothèque nationale	oui

## Cibles

Cibles destinées au service de criblage :	Toute cible moléculaire protéique et cellulaire, vecteur d'expression, vecteur luciférase, FRET
---	---

## Activités

Clonage et expression de protéines, obtention de lignées cellulaires/bactéries exprimant des protéines recombinantes, culture cellulaire de lignées cancéreuses humaines, mesure d'interaction protéine-protéine

## Criblage à haut débit

Nombre approximatif de mesures/ jour	Jusqu'à 4000 suivant type de mesure
--------------------------------------	-------------------------------------

### Tests biologiques proposés:

Activités enzymatiques en phase homogène ou hétérogène (liposomes) *in vitro* et *in cellulo*, tests sur cellules (viabilité, mort cellulaire) avec/sans prétraitement, interaction protéine-protéine (immunoprécipitation, polarisation de fluorescence, FRET, BRET)

## Technologies / Equipements

Robots diluteurs, FPLC, ultracentrifugeuses, spectrophotomètre (fluorescence, luminescence, absorbance), salles de culture

## Activités post-criblage

Confirmation, Structure Activity Relationship (SAR), Elucidation mode d'action

## Valorisation

Rédaction de protocole, dépôt de brevet, publication

## Accès/tarifs (sur demande)

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

Mise à jour 06/10/2014

## Fiche d'identité – Plateforme CBC (Criblage Biologique et Chimiothèques)

### Assay development and chemical library screening team

<b>Responsable</b>	Hélène Munier-Lehmann
<b>Equipe (effectifs)</b>	
chercheurs	1
techniciens	1
<b>Structure de rattachement</b>	Institut Pasteur, Unité de Chimie et Biocatalyse et CNRS UMR3523
<b>Adresse complète</b>	Institut Pasteur, Unité de Chimie et Biocatalyse 28 Rue du Dr Roux, 75724 Paris cedex 15
<b>Contact</b>	<a href="mailto:hmunier@pasteur.fr">Hélène Munier-Lehmann : hmunier@pasteur.fr</a>
<b>Date de mise en place de la PTF</b>	1 <sup>er</sup> trimestre 2007

#### Labels / démarche qualité

Pas de certification ISO. Mais toutes les plaques sont code-barrées et sont systématiquement scannées dès l'entrée sur la plateforme. Les données de pipettage et de lecture sont ensuite envoyées sur un serveur qui héberge un LIMS développé à façon par la société Modul-Bio et qui permet de relier les codes-barres à des plaques et donc à des composés (qui ont un code noté SID attribué par le LIMS mais pour lesquels on a aussi l'identifiant donné par le fournisseur) dans chaque puits de la plaque, et de relier aussi les plaque-mères aux plaques de dilution et/ou aux plaques expériences. Chaque composé est aussi consultable via son SID sous une forme de fiche listant les informations structurales et de criblage. Il y a ainsi une traçabilité des composés, des plaques et des résultats.

<b>Réseaux (national, international)</b>	GDR 3056 ChemBioScreen (Criblage biologique de petites molécules)
--	--

#### Principales réalisations

nombre de projets traités depuis la création	26
nombre de projets traités par an	en moyenne 4 projets

domaines de compétence / aires thérapeutiques thème principal : infectiologie  
criblages enzymatiques ou cellulaires (readout en absorbance, fluorescence, luminescence et AlphaScreen)

## Chimiothèque

Description de la collection :

Chimiothèques commerciales (Prestwick, ChemDiv et Chem-X-Infinity) : 25104 composés;  
Chimiothèque Nationale (Institut Curie, Faculté de Pharmacie de Strasbourg, CERMN Caen, ICBMS Lyon) : 25220;

Chimiothèque Institut Pasteur : 2480 composés. Total: environ 53000 composés.

prochainement dépositaire d'une copie complète de la Chimiothèque Nationale disponible en plaques soit environ 40000 composés (15000 composés supplémentaires par rapport à aujourd'hui)

Conditionnement : plaques 96 et 384-puits

Base de données :

LIMS développé et mis à jour par la société ModulBio incluant les structures chimiques des composés, et reliant de façon automatique chaque composé à sa position dans une plaque (toutes les plaques sont code-barrées et sont scannées dès leur entrée sur la plateforme). Accès au serveur via internet avec identifiant et mot de passe (possibilité de définir des profils d'utilisateurs).

Membre de la chimiothèque nationale :

L'Institut Pasteur a adhéré au GIS Chimiothèque Nationale le 14 septembre 2011. Notre Unité est membre de la Chimiothèque Nationale et nous avons une chimiothèque en interne.

## Cibles

Cibles destinées au service de criblage :

Nous n'avons pas de test en routine. Nous développons des tests et réalisons les campagnes de criblage en fonction des projets. Nous avons effectué des tests sur des protéines purifiées, sur des cultures cellulaires ou sur des pathogènes (bactéries, champignons, etc.).

## Criblage à haut débit

Nombre approximatif de mesures/ jour :

on est capable de tester jusqu'à 84 plaques 96-puits par jour (soit environ 8000 mesures par jour) au cours d'une campagne. En 2012 nous avons effectué 216 000 mesures au cours de campagnes de criblage.

Tests biologiques proposés :

tests enzymatiques (par exemple, cinétique rapide avec ou non couplage enzymatique avec mesure en absorbance; end-point avec mesure en luminescence);

tests sur pathogènes (mesure de viabilité en absorbance par exemple);

tests cellulaires (suivi de l'interaction d'un récepteur avec son ligand par HTRF par exemple).

## Technologies / Equipements

2 plateformes TECAN avec systèmes de pipettage (tête 96 cônes ou aiguilles, bras 4 ou 8-canaux indépendants), bras robotiques, carroussels, incubateurs;

4 lecteurs dont 3 intégrés (Infinite M1000 Pro incluant la technologie Alphascreen/AlphaLisa, Sapphire2, Infinite M200 Pro et Sunrise);

les 2 plateformes sont situés sous des plafonds soufflants permettant des conditions stériles si nécessaire et le tout est localisé dans un laboratoire de niveau de sécurité P2.

## Activités post-criblage

Vérification de la qualité du criblage (notamment, pour chaque microplaque, calcul du facteur  $Z'$ ), analyse des résultats (calculs selon équation(s) prédéfinie(s) puis requêtes selon un seuil variable ou en totalité, et principalement par chimiothèque),

fourniture des résultats sous forme de fichier excel avec pour chaque ligne correspondant à un composé les colonnes suivantes : le SID, la structure, le cmpd identifier d'origine ou external ID, la masse molaire, le nom de l'expérience, le code-barre de la plaque, le puits dans la plaque, la concentration du composé, la valeur de la donnée analysée, le facteur  $Z'$  ;

échanges avec les chimistes pour étapes de validation et études SAR

## Accès

Pour l'instant, projets uniquement en collaboration avec des équipes de l'Institut Pasteur

*Mise à jour*

*06/10/2014*

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### Criblage primaire de protéine fluorescente sur puces à petites molécules

<b>Responsable</b>	Christian GIRARD
<b>Equipe (effectifs)</b>	2
chercheurs	1
ingénieurs	1
<b>Structure de rattachement</b>	UMR8258 CNRS / U1022 INSERM
<b>Adresse complète</b>	UTCBS - ChimieParistech 11 rue Pierre et Marie Curie 75005 PARIS
<b>Contact</b>	<a href="mailto:christian.girard@chimie-paristech.fr">christian.girard@chimie-paristech.fr</a> 01 44 27 67 48
<b>Date de mise en place de la PTF</b>	01/01/2008

#### Principales réalisations

nombre de projets traités depuis la création 6

nombre de projets traités par an 1 à 2

domaines de compétence / aires thérapeutiques :

Criblage à haut débit de tout type de protéine purifié. Criblage primaire d'interaction Protéine-Petite molécule. Marquage par un agent fluorescent de la protéine

Fonctionnalisation de surface de lame de verre

Utilisation d'une chimiothèque issue des synthèses de notre laboratoire. Optimisation des "hit" par synthèse organique (Robot de synthèse, MW, flux continu)

## Chimiothèque

Description de la collection	1000 molécules issues de la synthèse orientée vers la diversité (thiazolidine, Triazole, Macrocycle, C-glycoside, Thiourée, Sucre...)
Passerelles vers d'autres collections	Chimiothèque nationale, Chimiothèque Institut Curie
Conditionnement	En solution dans le DMSO (10 mg/ml) + en poudre
Base de données	La chimiothèque est décrite dans un fichier word rassemblant la structure, le nom, la formule brute, la masse molaire des petites molécules.

## Cibles

Criblage de tout type de protéine purifiée comme cible, lysats cellulaires, cellules

## Activités

Synthèse de banques, Marquage de protéines

## Criblage à haut débit

Nombre approximatif de mesures/ jour 100/jour

## Technologies / Equipements

Microdépenseur GENETIX / Scanner à fluorescence GENEPIX

## Activités post-criblage

Synthèse en quantité de la (des) touche(s) pour tests complémentaires in vitro / in vivo  
Optimisation des touches identifiées par synthèse organique

## Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## Formation

Cours de troisième année (MR) de l'ENSCP - ChimieParistech

Mise à jour 20/10/2014

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### CMBA (Criblage pour des Molécules Bio-Actives)

<b>Responsable</b>	Dr Marie-Odile Fauvarque
<b>Equipe (effectifs)</b>	
chercheurs	Dr Marie-Odile Fauvarque
ingénieurs	Dr Caroline Barette : responsable opérationnelle criblage haut débit (HTS) Emmanuelle Soleilhac : responsable criblage phénotypique à haut contenu (HCS)
techniciens	Catherine Pillet : mise au point de tests et valorisation de molécules
administratifs	Gestionnaire partagé
autres (préciser)	Support bioinformatique de l'IRTSV (GIPSE): conception et implémentation d'un logiciel dédié TAMIS
<b>Structure de rattachement</b>	Institut de Recherches en Technologies et Sciences pour le Vivant (iRTSV) Laboratoire BGE , UMR_S 1038 INSERM, CEA, Université Grenoble-Alpes
<b>Adresse complète</b>	iRTSV - LBGE - Gen&Chem -CMBA U1038 INSERM/CEA/UJF CEA Grenoble, 17, rue des Martyrs 38054 Grenoble Cedex 09
<b>Site Web</b>	<a href="http://www-dsv.cea.fr/dsv/instituts/institut-de-recherches-en-technologies-et-sciences-pour-le-vivant-irtsv/unites/biologie-a-grande-echelle-bge/equipe-gen-chem/plate-forme-de-criblage-pour-des-molecules-bioactives-cmba">http://www-dsv.cea.fr/dsv/instituts/institut-de-recherches-en-technologies-et-sciences-pour-le-vivant-irtsv/unites/biologie-a-grande-echelle-bge/equipe-gen-chem/plate-forme-de-criblage-pour-des-molecules-bioactives-cmba</a>
<b>Contact</b>	Tel: +33 (0)4 3878 26 37 (office) or - 5238 (lab)
<b>Date de mise en place de la PTF</b>	2008

## Labels / démarche qualité

IBiSA

labellisée IBiSA avec le groupe de Chimie Combinatoire et Criblage à Haut Débit du CEA-Saclay (CBMA biologie, Saclay chimie)

## Réseaux (national, international)

Réseau National de plateformes de criblage national (FR-openscreen)

## Principales réalisations

nombre de projets traités par an  
domaines de compétence / aires thérapeutiques

8 à 10, activités de collaboration et de service diversifiés et adaptables (modification du cytosquelette, cancer, infectieux...)  
expérience maladies rares

## Chimiothèque

Description de la collection

>33,000 molécules chimiques (dont environ 50% chimiothèque nationale et 50% commerciales incluant Prestwick, Infarmatik, Chembridge)

Passerelles vers d'autres collections

sur demande, inclus accès aux autres chimiothèques (Paris, Strasbourg etc...)

Conditionnement

microplaques 96 puits, reconditionnement possible vers microplaques 384 puits

Membre de la chimiothèque nationale

Plate-forme référencée comme plate-forme ouverte de criblage auprès de la Chimiothèque Nationale.

## Cibles

Cibles destinées au service de criblage :

cibles définies et fournies par l'utilisateur pouvant être des cibles cellulaires (criblage sur cellules y compris bactéries et levures), cibles protéiques ou enzymatiques (protéines purifiées par utilisateur ou disponibles commercialement)

## Activités

activité de service et de collaboration pour le criblage de chimiothèques sur des tests cellulaires ou biochimiques (High Throughput Screening, HTS) :

- Analyse du projet et évaluation de la faisabilité
- Conseils et aide à la miniaturisation.
- Accès aux locaux aménagés et aux appareils pour mise au point des tests.
- Accès aux collections de molécules chimiques, dans des conditions cadrées par les accords entre les plates-formes et les fournisseurs de chimiothèques (lorsqu'il s'agit de chimiothèques issues de laboratoires académiques).
- Assistance dans le choix des collections chimiques à cribler, en fonction des objectifs visés
- Programmation robotique et optimisation du test
- Criblage de tout ou partie des collections de molécules chimiques, en microplaques de 96 puits.
- Analyse et sélection des touches (hits)
- Réalisation du Hit-picking et d'un criblage de validation.
- Transfert des résultats (au minimum sous forme de données brutes associées aux structures des composés de la chimiothèque) simultanément au biologiste porteur du projet et aux chimistes détenteurs des chimiothèques (dans le seul cas du criblage de chimiothèques académiques).

activité de recherche et développement pour l'analyse de phénotypes cellulaires par imagerie à haut contenu d'information (High Content Screening, HCS), et l'analyse statistique des résultats

## Criblage à haut débit

### Nombre approximatif de mesures/ jour

variable en fonction du format de criblage (96 ou 384 puits) et de la complexité du test, typiquement de 5 à 30 plaques/jour

### Tests biologiques proposés

tests cellulaires dont phénotypiques, de survie cellulaire, de gènes rapporteurs, d'intégration virale... Tests enzymatiques, blocage de fonctions enzymatiques, interactions protéine-protéine. Tout type de test pouvant être porté sur les instruments composant la plate-forme.

limite: pas de centrifugation, pas de radioactivité

## Technologies / Equipements

plateforme entièrement robotisée avec incubateur à cellules, laveur de microplaques adapté aux cellules et lecteur de microplaques (fluorescence dont HTRF et polarisation, absorbance, luminescence) en 96 et 384 puits

analyses informatiques et statistiques (logiciel maison, TAMIS)

## Activités post-criblage

essais automatisés sur les hits et/ou analogues (cytotoxicité cellulaire, dose/effet, HCS,...)

## Valorisation

Licence logiciel TAMIS pour la gestion des données chimiques, et la gestion et l'analyse statistique des données biologiques

Composé Liminib (Dr Laurence Lafanechère, CNRS), anti-cancéreux (société Cellipse en incubation)

Composés anti-angiogéniques (Dr Jean-Jacques Feige, INSERM)

Composés ciblant la CK2 (Dr Claude Cochet, INSERM)

## Accès

Equipes du site oui

Equipes extérieures académiques oui

Equipes extérieures privées oui

## Formation

Ecole Thématique de Criblage

Activité diverses de formation sur le criblage HTS ou HCS (cours, ateliers)

Formation au logiciel TAMIS (destinée au responsable de Plateforme)

*Mise à jour*

*04/11/2013*

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### Antiviral Drug Design Platform-AD2P

<b>Responsable</b>	Jean-Claude GUILLEMOT
<b>Equipe (effectifs)</b>	9
chercheurs	4
ingénieurs	2
techniciens	1
autres (préciser)	Thésards 2
<b>Structure de rattachement</b>	laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB) - Unité des virus émergents de la Timone (UVE) - Centre de recherche en cancérologie de Marseille (CRCM)
<b>Adresse complète</b>	Campus de Luminy AFMB-Polytech 13288 Marseille cedex 09
<b>Site Web</b>	<a href="http://www.AFMB.univ-mrs.fr">www.AFMB.univ-mrs.fr</a>
<b>Contact</b>	Jean-Claude.Guillemot@univ-amu.fr
<b>Date de mise en place de la PTF</b>	2007
<b>Labels / démarche qualité</b>	
IBiSA	oui
Certification ISO 9001	non
<b>Réseaux (national, international)</b>	Groupement De Recherche (GDR) ChembioScreen, CoreBio-PACA, Eurobiomed, GDR ChemInfo
<b>Principales réalisations</b>	
nombre de projets traités depuis la création	25
nombre de projets traités par an	6 à 10 (depuis 3 ans)
domaines de compétence / aires thérapeutiques	Antiviraux (de l'enzyme à l'isolat clinique) 2P2I

## Chimiothèque

Description de la collection :

Chimiothèque Nationale, Prestwick, Chembridge (30 000), 2P2I3D, >70 000 molécules

Conditionnement :	Plaques 96, DMSO
Base de données :	LIMS spécifique de ModulBio
Membre de la chimiothèque nationale :	Collaborations actives

## Cibles

Cibles protéiques : enzymes virales, complexes viraux (prot/prot)

Cellulaires: Isolats Cliniques, Réplicon

## Activités

Criblage sur enzyme, cellules	Constitution de chimiothèque dédiée
-------------------------------	-------------------------------------

## Criblage à haut débit

Nombre approximatif de mesures/ jour	2 000/semaine
--------------------------------------	---------------

Tests biologiques proposés :

Activité enzymatique: Polymérase, hélicase, méthyltransférase. Chiminformatique

Test avec détection radioactive, Fluorescence, 2P2I. Criblage sur isolats cliniques

## Technologies / Equipements

2 robots, espace ARS, 1 BSL2, 1BSL3, Multiples lecteurs

## Activité Post-criblage

Techniques biophysiques: ITC, Biacore, Thermophorèse, Chimie intégrée

Thermal shift assays, Co-cristallisation/soaking (3D)

## Valorisation

SATT-PACA, 1 brevet déposé, une maturation en cours.

## Accès

Equipes du site	Oui
Equipes extérieures académiques	Oui
Equipes extérieures privées	Oui

## Formation

Ecole de criblage, formation d'ingénieurs (Polytech-Marseille, Centrale Marseille)

Mise à jour 02/12/2013

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### Plateforme ARPEGE -Pharmacologie-Criblage-Interactome

<b>Responsable</b>	Laurent Prézeau, DR2 CNRS
<b>Equipe (effectifs)</b>	
chercheurs	1 DR2, 30% ETP (Responsable Scientifique) 1 Ingénieur de Recherche, 100% ETP
ingénieurs	(Responsable Technique) 2 Ingénieures d'études IE2, 50% ETP chacune
autres (préciser)	1 CDD Assistant Ingénieur, un an
<b>Structure de rattachement</b>	UMS Biocampus
<b>Adresse complète</b>	Institut de Génomique Fonctionnelle (IGF) 141, rue de la Cardonille 34094 Montpellier cedex 5
<b>Site Web</b>	<a href="http://www.arpege.cnrs.fr/">http://www.arpege.cnrs.fr/</a>
<b>Contact</b>	technique@arpege.cnrs.fr direction@arpege.cnrs.fr
<b>Date de mise en place de la PTF</b>	2004
<b>Labels / démarche qualité</b>	
IBiSA	Label Ibis depuis 2008
Certification ISO 9001	Audit Prévu en 2014
<b>Réseaux (national, international)</b>	Cancéropôle Grand Sud-Ouest ; GDR Criblage

## Principales réalisations

**nombre de projets traités depuis la création** 10 projets de criblage

**nombre de projets traités par an** 1 projet de criblage /an

### **domaines de compétence / aires thérapeutiques :**

Compétences en analyse pharmacologique ; récepteurs couplés aux protéines G ; récepteurs membranaires ; signalisation protéines G ; signalisation arrestine ; phosphorylation de protéines de signalisation.

Compétences techniques : senseurs fluorescents et tests basés sur le principe de transfert d'énergie (BRET, FRET, TR-FRET, HTRF) ; tests cellulaires (prolifération, migration, mort cellulaire, morphologie cellulaire) en technique Labelfree (impédance, redistribution de masse). Applications en de nombreux domaines physio-pathologiques et thérapeutiques.

## Chimiothèque

Description de la collection Chimiothèques fournies par les clients (commerciales, propres, Chimiothèque Nationale)

Membre de la chimiothèque nationale Partenaire de la Chimiothèque Nationale au sein des réseaux de plateformes

## Cibles

### **Cibles destinées au service de criblage :**

Cibles protéiques : protéines membranaires, récepteurs membranaires (RCPGs, Récepteurs canaux, Récepteurs enzymes, etc), protéines intracellulaires, etc.

## Activités

Mise au point des tests de criblage faible à moyen débit (criblage de petites banques de quelques centaines de molécules)

## Criblage à haut débit

Nombre approximatif de mesures/ jour 10 plaques 96 puits/jour ; 5-10 plaques 384 puits/jour, en fonction des tests utilisés

### Tests biologiques proposés :

Analyse Pharmacologique : test sur cellules de liaison de ligands fluorescents, analyse fonctionnelle (recrutement de protéines de signalisation, activation de protéines effectrices, production de seconds messagers intracellulaires, etc), tests senseurs de signal (calciques, etc)

## Technologies / Equipements

7 Lecteurs de plaques Spectrofluorimètres et Luminomètres pour les techniques basées sur la fluorescence et les transferts d'énergie par résonance : BRET, FRET, TR-FRET, HTRF ;  
2 Lecteurs de plaques Labelfree (mesure d'impédance ou de redistribution de masse) pour les techniques de mesures de paramètres cellulaires (prolifération, migration, etc)

## Activités post-criblage

Caractérisation pharmacologique : de la liaison des ligands aux comportements cellulaires, en passant par les aspects fonctionnels des cibles.

Détermination de la formation de complexes protéiques.

## Valorisation

Au cas par cas

### Accès

Equipes du site	70%
Equipes extérieures académiques	25%
Equipes extérieures privées	5%

## Formation

TP Masters Université Montpellier 2  
Ateliers annuels technologiques Biocampus

*Mise à jour*

20/10/2014

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### ImPACcell (Imagerie Pour Analyse du Contenu cellulaire)

<b>Responsable</b>	Rémy Le Guével
<b>Equipe (effectifs)</b>	
chercheurs	1
ingénieurs	2
<b>Structure de rattachement</b>	Structure Fédérative de Recherche BIOSIT
<b>Adresse complète</b>	Université de Rennes 1, Bat 8 Campus de Villejean 2 Avenue du Pr. Leon Bernard 35043 Rennes cedex
<b>Site Web</b>	<a href="http://imagerie-puces-a-cellules.univ-rennes1.fr/">http://imagerie-puces-a-cellules.univ-rennes1.fr/</a>
<b>Contact</b>	Rémy Le Guével
<b>Date de mise en place de la PTF</b>	2005
<b>Labels / démarche qualité</b>	
IBiSA Certification ISO 9001	2010 en projet
<b>Réseaux (national, international)</b>	Biogenouest, Canceropole grand Ouest
<b>Principales réalisations</b>	
nombre de projets traités par an	10
domaines de compétence / aires thérapeutiques :	Cancers, Toxicité environnementale, hepatotoxicité

## Cibles

Cibles destinées au service de criblage :

Cellules tumorales humaines, lignée Hépatocytaire HepaRG

Cytosquelette, cycle cellulaire, signalisation intracellulaire, mort cellulaire, motilité et mobilité cellulaire, genotoxicité

## Activités

Culture cellulaire, criblage molécules d'intérêt pharmacologique

Détermination du mode d'action de molécules d'intérêt

## Criblage à haut débit

Tests biologiques proposés

Test sur cellules en immunodétection

Analyses d'images

## Technologies / Equipements

Robot d'imagerie HCS Cellomics; robots de pipettage Hamilton

Microscope à fluorescence, vidéo microscope

*Mise à jour*

*05/09/2013*

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### PICT - Plateforme Intégrée de Criblage de Toulouse

<b>Responsable</b>	Pr. Laurent Maveyraud
<b>Equipe (effectifs)</b>	
chercheurs	19
ingénieurs	11
techniciens	5
administratifs	1
autres (préciser)	1 post doc, 2 doc
<b>Adresse complète</b>	3 sites toulousains
<b>Site principal (Pr Laurent Maveyraud)</b>	Institut de Pharmacologie et Biologie Structurale UMR 5089 - BP64182 205, route de Narbonne 31077 Toulouse Cedex 04
<b>Site criblage (Dr Frédéric Ausseil)</b>	USR CNRS-Pierre Fabre 3388 CRDPF - 3, avenue Hubert Curien - BP 13562 31035 Toulouse Cedex 1
<b>Site chimie (Pr Michel Baltas)</b>	LSPCMIB UMR 5068 Université Paul Sabatier. 118, route de Narbonne, 31062 TOULOUSE Cedex 09
<b>Site Web</b>	<a href="http://cribligand.ipbs.fr/">http://cribligand.ipbs.fr/</a>
<b>Labels / démarche qualité</b>	
IBiSA	oui
Certification ISO 9001	oui
Autre	CNOC
<b>Réseaux (national, international)</b>	GiS Chimiothèque Nationale Réseau National de Criblage GDR Chembioscreen ELRIGfr Génotoul AFSP GDR assemblages supramoléculaire et mbs bios Réseau RMN Midi Pyrénées Fédération Recherche Chimie Midi Pyrénées

## Principales réalisations

nombre de projets traités depuis la création >100

nombre de projets traités par an de 5 à 10

domaines de compétence / aires thérapeutiques :

Conception de chimiothèques et synthèse de ligands

Analyse et purification par HPLC, SFC

Criblage robotisé à haut débit

Détermination de structures 3D des cibles par RMN et radiocristallographie des rayons X

Modélisation moléculaire et criblage virtuel

Criblage de ligands par des méthodes biophysiques, de RMN et la cristallographie des rayons X

Étude biophysique et thermodynamique des interactions cibles-ligands (RMN, fluorescence, calorimétrie)

Ingénierie enzymatique : approche rationnelle ou évolution aléatoire ou semi-rationnelle de protéines

Criblage fonctionnel haut-débit d'activités enzymatiques : banques d'évolution dirigée, banques métagenomiques

## Chimiothèque

Description de la collection :

Banque de 1200 fragments pour criblage biophysique

Banque de 3000 molécules correspondant à des archétypes structuraux maîtrisés à PICT (composante chimie)

Banque Prestwick

Passerelles vers d'autres collections :

Chimiothèque Nationale

Chimiothèque Pierre Fabre (sous conditions contractuelles)

Base de données

Activity Base

Membre de la chimiothèque nationale

oui

## Cibles

Cibles destinées au service de criblage :

Enzymes, canaux, récepteurs, voies de transduction, fonctions cellulaires, phénotypes cellulaires, zebrafish

## Activités

Entretien lignées, culture cellulaire, scale-up culture cellulaire

## Technologies / Equipements

2 stations HTS full automation (Beckman FX/BRT)

2 labos, miniaturisation/automatisation des essais

2 labos de culture cellulaire

1 station HCS (In cell ->Arrayscan cellomics)

## Accès

Equipes du site

Oui

Equipes extérieures académiques

Oui (sous conditions)

Equipes extérieures privées

Non

*Mise à jour*

*10/01/2014*

## Fiche d'identité – Plateforme ADME

### Lead Optimization

<b>Responsable</b>	Benoît Déprez, Florence Leroux
<b>Structure de rattachement</b>	Inserm, Institut Pasteur de Lille, Université Lille2
<b>Adresse complète</b>	Faculté de Pharmacie de Lille 3 Rue du Professeur Laguesse BP83 F-59006 Lille FRANCE
<b>Site Web</b>	<a href="http://www.drugdiscoverylille.org/">http://www.drugdiscoverylille.org/</a> <a href="http://www.deprezlab.fr/">http://www.deprezlab.fr/</a>
<b>Contact</b>	Catherine Piveteau
<b>Date de mise en place de la PTF</b>	2008
<b>Labels / démarche qualité</b>	
IBiSA	oui
Certification ISO 9001	non
<b>Principales réalisations</b>	
nombre de projets traités depuis la création	147 Etudes
nombre de projets traités par an	de 9 à 41
domaines de compétence / aires thérapeutiques	Transversale
<b>Equipe (effectifs)</b>	
chercheurs	2
ingénieurs	2
techniciens	1
administratifs	1,5

## Etude des propriétés des hits

Profil physicochimique

Solubilité, LogD, pKa..

Propriétés d'ADME-toxicologie

Stabilité microsomale, stabilité plasmatique, profil métabolique, caractérisation CYP et estérases...

Comportement pharmacocinétique

Distribution, Cmax, AUC, F%, Interaction PK, PK-PD...

## Technologies / Equipements

LC-MS / Triple Quadripole Waters (Acquity I Class - Xevo TQD)

LC-MS (Analytique) / Simple Quadripole Waters (ZQ2000)

LC-MS (Préparative et Analytique) / Simple Quadripole Waters (3100)

LC-MS Temps de Vol (LCT Premier XE)

Préparateur d'échantillon Gilson (GX271 ASPEC)

## Valorisation

Villemagne, B. et al., Ligand efficiency driven design of new inhibitors of Mycobacterium tuberculosis transcriptional repressor EthR using fragment growing, merging and linking approaches. *Journal of Medicinal Chemistry*, **2014**

Deprez-Poulain R. et al. Structure-Activity Relationships and Blood Distribution of Antiplasmodial Aminopeptidase-1 Inhibitors. *Journal of Medicinal Chemistry*, **2012**

Flipo M. et al. Discovery of novel N-phenyl-phenoxyacetamide derivatives as EthR inhibitors and ethionamide boosters by combining High-Throughput Screening and Synthesis. *Journal of Medicinal Chemistry*, **2012**

Flipo, M. et al., Ethionamide Boosters. 2. Combining Bioisosteric Replacement and Structure-Based Drug Design To Solve Pharmacokinetic Issues in a Series of Potent 1,2,4-Oxadiazole EthR Inhibitors. *Journal of Medicinal Chemistry*, **2012**

Flipo, M. et al., Ethionamide Boosters: Synthesis, Biological Activity, and Structure-Activity Relationships of a Series of 1,2,4-Oxadiazole EthR Inhibitors. *Journal of Medicinal Chemistry*, **2011**

Flipo, M. et al., Hydroxamates: Relationships between structure and plasma-stability. *J Med Chem*, **2009**

Willand, N. et al., Synthetic EthR inhibitors boost anti-tuberculous activity of ethionamide. *Nature Medicine*, **2009**

## Accès

Equipes du site

U761

Equipes extérieures académiques

Equipes extérieures privées

Pas d'accès libre

## Formation

Pas pour le moment, mais envisageable

Mise à jour

06/10/2014

## Fiche d'identité – Plateforme ADME

### TechMed<sup>ILL</sup>

<b>Responsable</b>	Patrick Gizzi
<b>Structure de rattachement</b>	UMR 7242 (directeur Jean-Luc Galzi)
<b>Adresse complète</b>	Plate-forme TechMed <sup>ILL</sup> ESBS, 300 boulevard Sébastien Brant 67412 Illkirch
<b>Site Web</b>	<a href="http://www.pcbis.fr">www.pcbis.fr</a>
<b>Contact</b>	Patrick GIZZI
<b>Date de mise en place de la PTF</b>	juin-09
<b>Labels / démarche qualité</b>	
IBiSA	Oui
Certification ISO 9001	Oui
<b>Réseaux (national, international)</b>	EU-OPENSREEN
<b>Principales réalisations</b>	
nombre de projets traités depuis la création	215
nombre de projets traités par an	40
domaines de compétence / aires thérapeutiques	Prestations ADME
<b>Equipe (effectifs)</b>	
ingénieur	1

## Etude des propriétés des hits

Profil physicochimique

Solubilité, lipophilie (log P)

Constante d'acidité (pKa)

Stabilité chimique

Propriétés d'ADME-toxicologie

Perméabilités (PAMPA, Caco-2)

Liaison aux protéines plasmatiques, stabilité plasmatique

Métabolisme hépatique *in vitro* (stabilité, inhibition, induction des CYP)

Cytotoxicité

Comportement pharmacocinétique

PK chez la souris, biodisponibilité, biodistribution

## Technologies / Equipements

HPLCs (détection UV-vis)

LC-MS (triple quadripôle)

Électrophorèse capillaire

## Accès

Equipes du site

Oui

Equipes extérieures académiques

Oui

Equipes extérieures privées

Oui

*Mise à jour*

*06/10/2014*

## Fiche d'identité – Plateforme de Criblage à Haut Débit

### PCBIS (Plate-forme de Chimie Biologique Intégrative de Strasbourg)

<b>Responsable</b>	Pascal Villa
<b>Equipe (effectifs)</b>	13
ingénieurs	7
techniciens	2
administratifs	2
post-doc	2
<b>Structure de rattachement</b>	UMS 3286 CNRS-Unistra
<b>Adresse complète</b>	UMS 3286 PCBIS ESBS et Faculté de Pharmacie 300 boulevard Sébastien Brant 67412 ILLKIRCH
<b>Site Web</b>	<a href="http://www.pcbis.fr">www.pcbis.fr</a>
<b>Contact</b>	Pascal Villa
<b>Date de mise en place de la PTF</b>	1999
<b>Labels / démarche qualité</b>	
IBiSA	oui (2008)
Certification ISO 9001	oui (janvier 2007)
Autre	Labex Medalis
<b>Réseaux (national, international)</b>	GDR Chembioscreen Eu-OPENSREEN

## Principales réalisations

nombre de projets traités depuis la création	> 70
nombre de projets traités par an	4 à 10
domaines de compétence / aires thérapeutiques	maladies rares; cancer; inflammation; développement et validation d'essais biologiques miniaturisation et robotisation d'essais criblage de chimiothèques (molécules de synthèse ou extraits naturels) études de relation structure/activité tests d'évaluation des propriétés ADME-Tox précoces des molécules GPCR tests de liaisons

## Chimiothèque

Description de la collection	45 000 molécules de synthèse libres de droit (méthodes de synthèses connues) 15 450 substances naturelles
Passerelles vers d'autres collections	Chimiothèque Nationale; chimiothèque européenne (à venir)
Conditionnement	poudres en tubes; plaques 96 puits; plaques 384 puits
Base de données	ISIS (accessibilité par le biais de la chimiothèque Nationale) et activity base
Membre de la chimiothèque nationale	dépositaire de la collection entière de la chimiothèque Nationale (produits en plaques)

## Cibles

### Cibles destinées au service de criblage (liste non exhaustive)

protéines purifiées solubles: tout type de protéine (kinases, phosphatases, phosphodiesterases, calmoduline, protéines nucléaires, etc)

Protéines membranaires: expression dans des cellules eucaryotes (GPCRs)

Modèles cellulaires: lignées cellulaires exprimant des protéines d'intérêt (genes rapporteurs)

lignées de cellules souches cancéreuses

Développement et optimisation de nouvelles cibles avec les laboratoires de recherche

## Activités

clonage

expression et purification de protéines

développement de lignées cellulaires

gestion et stockage des lignées cellulaires

transfections et transformation de lignées

service de culture cellulaire

développement d'essais biologiques avec mesures de fluorescence, luminescence, FRET, BRET, HTRF, TR-FRET, Polarisation de fluorescence

## Criblage à haut débit

Nombre approximatif de mesures/ jour 10 000

Tests biologiques proposés (liste non exhaustive (contacts sur [www.pcbis.fr](http://www.pcbis.fr))) :

tests enzymatiques (mesures en fluorescence, luminescence, HTRF)

essais de liaisons sur protéines solubles (technologie DMR)

essais de liaison de molécules sur protéines solubles (technologie de polarisation de fluorescence)

mesures de survie cellulaire, cytotoxicité, apoptose (essais en fluorescence, mesures de confluence cellulaire)

Essais de blessure cellulaire/cicatrisation ("wound healing") par analyse d'images

Mesure des effets anti-inflammatoires (recherche sur l'inhibition de la sécrétion de cytokines;

recherche de l'activité d'inhibition de la nitrite oxydase)

Tests d'anti-microbiens (bactéries, levures, etc)

## Technologies / Equipements

stations automatisées avec robots de pipetage, carrousel de stockage, incubateur à cellules, lecteurs de microplaques, distributeur de liquides et laveur de plaques). Conditions stériles permettant le travail avec des cellules.

lecteurs de microplaques permettant les mesures avec les technologies:

absorbance, fluorescence, HTRF, TR-FRET, polarisation de fluorescence, BRET, FRET, luminescence, DLS

Salle de culture; transfection de lignées (système 96 puits Amaxa-Lonza)

Mesures pKa 96 puits

HPLC, LCMSMS

## Activités post-criblage

aide à l'optimisation des molécules actives

recherche des effets cytotoxiques des molécules

études de relation structure/activité

tests secondaires sur molécules actives

Etudes ADME-Tox des molécules actives

## Valorisation

une molécule en phase 2 (minozac: Alzheimer)

un outil de recherche commercialisé par SIGMA (chalcone-4)

une molécules anti-cancéreuse (sur cellules souches cancéreuses) brevetée, en évaluation

plusieurs molécules actives sur des chimiokines (brevets en cours)

plusieurs molécules en évaluation post-criblage (sur modèles animaux ou autres)

## Accès

Equipes du site

27 équipes du site d'Illkirch

Equipes extérieures académiques

16 équipes extérieures

Equipes extérieures privées

10 équipes privées

## Formation

Formation continue Université de Strasbourg:

techniques de fluorescence et techniques sans marquage pour le criblage à haut débit (annuel: janvier)

Ecole thématique chemical biology (bi-annuel, multi-plates-formes)

*Mise à jour*

*06/10/2014*

**Fiches d'identité:**

**Plateformes de Séquençage Haut  
Débit, Génomique,  
Transcriptomique**

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### GenoScreen SAS

<b>Responsables</b>	Mr André Tordeux: Président Mr Guillaume de Saint Martin: Responsable Commercial & Marketing Mme Hélène Blanquart: Responsable NGS Dr Stéphanie Ferreira: Responsable R&D et Innovation Santé humaine Dr Cyril Gaudin: Responsable R&D et Innovation Microbiologie Moléculaire
<b>Adresse complète</b>	1 rue Du Professeur Calmette Campus de l'Institut Pasteur de Lille 59000 Lille France
<b>Site Web</b>	<a href="http://www.genoscreen.fr">www.genoscreen.fr</a>
<b>Contact</b>	<a href="mailto:commercial.dept@genoscreen.fr">commercial.dept@genoscreen.fr</a>
<b>Date de mise en place de la Plateforme</b>	
activités de séquençage à haut débit	2001
activités de génomique - transcriptomique	
<b>Labels / démarche qualité</b>	
IBiSA	Oui
Certification ISO 9001	en cours
Autre	Non

## 1. Séquençage à haut débit

### Domaines de réalisation

**Génétique humaine** Oui

**Génétique animale et végétale** Oui

**Microorganismes** Oui

### Services proposés (domaines d'applications / expertise)

**Séquençage *de novo*** Oui sur GS-FLX (454)  
Illumina HiSeq2500 et MiSeq

**Métagénomique** Oui sur GS-FLX (454) et Illumina MiSeq

#### Reséquençage

génomome entier	Oui sur Illumina HiSeq2500 et MiSeq
exome	Oui sur Illumina HiSeq2500
région ciblée	Oui sur Illumina HiSeq2500 et Miseq
RNAseq	Oui sur Illumina HiSeq2500 et Miseq
smallRNAseq	Oui sur Illumina HiSeq2500 et Miseq
ChIPseq	Oui sur Illumina HiSeq2500 et Miseq

### Equipe (effectifs)

32 personnes au total

### Equipements (séquenceur, robot, informatique)

2 Hi Seq 2500

2 MiSeq

1 GS junior

1 GSFlex+

1 GSFlex

3 Séquenceur ABI 3730 XL

1 ABI 7900 HT (PCR quantitative)

15 Thermocycleurs

2 automates de pipettage Tecan (tête 96 et bras 8 aiguilles)

Blade Dell composé de 3 lames de chacune 32 cœurs et 22Go de mémoire + 10 To de stockage

Serveur LDLC avec 24 cœurs et 48 Go de mémoire + 4To de stockage

Serveur Dell avec 8 cœurs et 16Go de mémoire + 1To de stockage

Accès à un serveur avec 64 cœurs et 10To de stockage

## Technologies utilisées

### Kits de capture

TruSeq, SureSelect 50 et 71 Mb, Haloplex, Raindance, Multiplicom

### Format de séquençage

Single and paired-end: 50, 100, 150, 200, 250, 300 bp

### Temps d'un run

48h pour GS, MiSeq et HiSeq2500 Rapid Run  
sinon 2 semaines pour HiSeq 2500 High Output run

### Multiplexage des échantillons

nombre d'échantillons/ligne

Jusqu'à 384

## Traitement, analyse et stockage des données

Pipeline d'analyse Bioinformatique spécifique à chaque type de service : nous consulter

## Développement d'outils de traitement des données

Nous consulter

## Rendu actuel des résultats

### Qualité des données

couverture

profondeur moyenne de lecture (min-max)

projets 'standard'

projets 'deep sequencing'

A définir avec le client en fonction du projet et de ses objectifs

La plateforme dispose-t-elle d'une base de données de variants?

Non

## Délai de rendu des résultats / Accès

Equipes du site

Equipes extérieures académiques

Equipes extérieures privées

De 2 à 12 semaines en fonction du projet et de la technologie utilisée

## Formation

Formation thématique ou sur mesure - Tous niveaux (débutant/intermédiaire/expérimenté) - En individuelle (sur 2 à 3 jours) ou en groupe (8 à 12 personnes sur 2 à 3 jours).

Exemples de thématiques: Assemblage, annotation et analyse de génomes / Blast - Technologies de séquençage / Lignes de commandes en environnement UNIX/LINUX...

Mise à jour

06/10/2014

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme Gentyane Génotypage et Séquençage à Haut Débit

<b>Responsable</b>	Charles PONCET
<b>Equipe (effectifs)</b>	
ingénieurs	2
techniciens	4
bioinformaticiens	1
administratifs	bénéficie des personnels administratifs de l'unité
<b>Structure de rattachement</b>	INRA UMR 1095 Génétique Diversité et
<b>Adresse complète</b>	5 chemin de Beaulieu, 63039 Clermont-Ferrand cedex 02
<b>Site Web</b>	en cours de création (Overscan)
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2012
2. des activités de génomique - transcriptomique	2007
<b>Labels / démarche qualité</b>	Plateforme Stratégique INRA
IBiSA	IBiSA 2009
Certification ISO 9001	2012

## 1. Séquençage à haut débit

### Principales réalisations

<b>Nombre de projets traités depuis la création</b> <i>préciser</i>	19 Séquençage de BAC et d'amplicon sur blé, bactérie, tique, campagnols, souris, coquelicot, etc ...
<b>Nombre d'échantillons traités depuis la création</b> <i>préciser</i>	plusieurs milliers Séquençage de BAC en mélange, plusieurs centaines d'amplicons en mélange
<b>Nombre de projets traités par an</b> <i>préciser</i>	15 en moyenne Clients INRA, CNRS, INSERM, Limagrain, CHRU, Université
<b>Nombre d'échantillons traités par an</b>	Plusieurs centaines

### Domaines de réalisation

<b>Génétique animale et végétale</b>	souris, coquelicot, blé, etc...
<b>Microorganismes</b>	bactéries

### Services proposés (domaines d'applications / expertise)

<b>Métagénomique</b>	oui
<b>Séquençage</b> de régions ciblées	oui
<b>RNAseq</b>	en cours

### Perspectives de développement de services

Rad Seq (réduction de complexité par digestion enzymatique), automatisation de préparation de bibliothèques NGS

### Equipements (séquenceur(s), robot(s), informatique)

Séquenceurs MiSeq Illumina et Roche 454 Junior, génotypage sur Fluidigm Biomark, LC 480, Affymetrix Genetitan, Illumina BeadXpress  
7 robots de pipetage tête 96 et 384 et 8 canaux (Hamilton et Beckman)

## Technologies utilisées

<b>Format de séquençage</b>	2*300 pb sur MiSeq et 500pb en moyenne sur 454 Junior
<b>Temps d'un run</b>	variable selon programme et kit utilisé
<b>Multiplexage des échantillons</b> nombre d'échantillons/ligne	jusqu'à 10 BAC tagués et 1200 amplicons

## Traitement, analyse et stockage des données

Traitement de premier niveau. Le traitement des séquences est assuré par le client. Logiciel NewBler Roche et suite Illumina

## Rendu actuel des résultats

<b>Qualité des données</b>	
couverture	variable selon projets
<b>Format de rendu des résultats de l'analyse bioinfo</b>	
reads	oui
alignements	oui
La Plateforme dispose-t-elle d'une base de données interne de variants?	non

## Accès et délai de rendu des résultats

Equipes du site	une semaine
Equipes extérieures académiques	selon planning
Equipes extérieures privées	selon planning

## Formation

Formation interne unité

## 2. Génomique, transcriptomique

### Principales réalisations

**Nombre de projets traités depuis la création** 230

**Nombre de projets traités par an** 40

**Domaines de réalisation** Développement de marqueurs, contrôle qualité, SAM, étude de population, etc...

### Services proposés (domaines d'applications / expertise)

Développement et génotypage de marqueurs moléculaires (SSR, SNP, autres) à très haut débit. De quelques marqueurs à plusieurs centaines de milliers de marqueurs.

### Perspectives de développement de services

Baisse des coûts de génotypage, automatisation poussée

### Equipe

Ingénieurs	2
Techniciens	4
Bioinformaticiens	1

### Equipements

5 robots de pipetage 8 à 384 canaux (Hamilton et Beckman), Séquenceur ABI 3730XL, Fluidigm Biomark, LC480 Roche, Affymétrie Genetitan, Illumina BeadXpress.

### Technologies utilisées

Technologie Illumina BeadXpress, Technologie Fluidigm Biomark Chimie Kaspar, Technologie hybridation sur Puce Affymetrix GeneTitan, PCR Temps réel LC 480 Roche.

### Traitement, analyse et stockage des données

Nous proposons le traitement des données mais aussi la formation des utilisateurs, nous stockons les données en standard 2 mois (Stockage sécurisé).

### Développement d'outils de traitement des données

Utilisation des logiciels associés aux différentes technologies

## Rendu actuel des résultats

Soit données brutes soit données analysées

## Délai de rendu des résultats

De quelques jours à quelques mois selon la taille du projet

## Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

*Mise à jour* 06/09/2013

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### INTEGRAGEN GENOMIC SERVICES

<b>Responsable</b>	Directeur des Services (commercial) : Emmanuel Martin Directeur laboratoire : Francis Rousseau Responsable Plateforme HT-Seq : Jean-Paul Saraiva Directeur Bioinformatique : Bérengère Génin
<b>Equipe (effectifs)</b>	
ingénieurs	6
techniciens	
bioinformaticiens	4 bio-informaticiens + 2 informaticiens
administratifs	2
autres (préciser)	6 autres techniciens et informaticiens non en charge du séquençage
<b>Adresse complète</b>	Genopole Campus 1- Genavenir 8 5, rue henri Desbruères 91030 EVRY Cedex
<b>Site Web</b>	<a href="http://www.integragen.com">www.integragen.com</a>
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2008
2. des activités de génomique - transcriptomique	2006 en service

## 1. Séquençage à haut débit

### Principales réalisations

<b>Nombre de projets traités depuis la création</b>	500 depuis 2009 480 en HiSeq ou GAIIX, 20 en MiSeq 296 projets d'exomes
<b>Nombre d'échantillons traités depuis la création</b>	4000 exomes fin 2013 (le reste confidentiel)
<b>Nombre de projets traités par an</b>	> 200 en 2012

### Domaines de réalisation

<b>Génétique humaine</b>	X
<i>préciser les spécialités</i>	Maladies génétiques, Génomique du Cancer, Epidémiologie, Pharmacologie
<b>Génétique animale et végétale</b>	X

### Services proposés (domaines d'applications / expertise)

<b>Séquençage <i>de novo</i></b>	oui
<b>Métagénomique</b>	dépend du projet
<b>Séquençage</b>	
de génomes entiers (WGS)	oui
d'exomes entiers (WES)	oui
de régions ciblées	oui
<b>RNAseq</b>	oui
<b>SmallRNAseq</b>	oui, y compris smRNA plasmatiques
<b>ChIPseq</b>	oui
<b>MeDIPseq</b>	oui
<b>Autres</b>	RRBS (reduced representation bisulfite sequencing)

## Perspectives de développement de services

- \* séquençage à partir de très faibles quantités d'ADN, de produits WGA, ...
- \* développement d'un pipeline d'analyse oncologie en collaboration avec des équipes académiques et hospitalières

## Equipements (séquenceur(s), robot(s), informatique)

HiSeq2000, HiSeq2500  
NextSeq500, MiSeq

## Technologies utilisées

<b>Kits de capture</b>	Agilent SureSelect all exon V5 50Mb et V5UTR 70Mb Agilent SureSelect custom
<b>Format de séquençage</b>	dépendant du projet: paired-end 2x75 (exome), paired-end 2x100 (RNA-Seq, de novo, methylome...), 1x36 (smallRNA)
<b>Temps d'un run</b>	dépendant du type de run
<b>Multiplexage des échantillons</b>	oui, dépendant du projet (multiplexage possible jusqu'à 96 banques par canal sur HiSeq)

## Traitement, analyse et stockage des données

Analyse des données d'exome, whole genome, RNA-Seq, smRNA, Methyl-Seq et RRBS.  
Stockage des données brutes (Fastq et BAM) pendant 15 jours à compter de la livraison sur disque dur externe.

## Développement d'outils de traitement des données

ERIS Exome Resequencing Intelligent Server, outil de filtre et d'analyse des variants permettant des comparaisons intra- et inter-famille(s) ou groupe(s) d'individus.  
Un module Oncologie permet une analyse spécifique des variants somatiques.

## EXOME ou reséquençage ciblé par capture

Quantité de matériel requis: 1µg d'ADN minimum

Un protocole pour de l'ADN FFPE est également disponible. Nous consulter.

### Rendu actuel des résultats

#### Caractéristiques des données produites

Garanties techniques:

\*Validation de la concordance d'un panel de marqueurs génotypés sur chaque ADN réceptionné avec les résultats d'exome correspondant.

\*Quantité de données minimum garantie par échantillon (dépendant du kit de capture utilisé), dépendante de l'objectif du projet

#### Qualité des données

couverture	>90% à 10X; environ 85% à 25X
profondeur moyenne de lecture (min-max)	en moyenne 75X après mapping et filtre des séquences
projets 'standard'	Adaptable en fonction de l'objectif du projet.
projets 'deep sequencing'	dépendant de l'objectif du projet et de la fréquence attendue des mutations recherchées

#### Format de rendu des résultats de l'analyse bioinfo

reads	fichiers Fastq (et contrôles qualités du run associés)
alignements	BAM
détection des variants	fichier tabulaire et vcf: comptage des bases, génotype observé et scores qualité associés

Annotation des variants : fichier tabulaire et vcf: pipeline d'annotation développé par IntegraGen

Pour chaque transcrit NM# (RefSeq) :

- position sur le génome, sur le gène, sur le transcrit, sur la protéine, sur les miARN, distance à l'intron
- type de mutation (intron, synonyme, faux sens, non sens, 3'UTR, 5'UTR), changement d'acide aminé, ou de phase de lecture .
- annotation des variants présents dans les principales bases de données publiques
- annotation des variants présents dans la base de données Exome d'IntegraGen

## Outils de prédiction de la pathogénicité des variants

outils existants

Polyphen

## Autres

Analyses CNV

## Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement

ERIS Exome Resequencing Intelligent Server, interface d'analyse des données d'exome en ligne (inclut une passerelle vers Alamut).

Accompagnement personnalisé à la prise en main des données pour tout nouveau client.

La Plateforme dispose-t-elle d'une base de données interne de variants? oui

## Délai de rendu des résultats

6-10 semaines

## Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## Reséquençage ciblé par PCR haut-débit

Service incluant le design des PCR à partir d'une liste de gènes, la production des PCR haut-débit à l'aide de l'Access Array Fluidigm, le séquençage haut-débit sur MiSeq en 2x150 bases.

### Rendu actuel des résultats

#### Caractéristiques des données produites

Quantité minimum produite dépendante de la taille cumulée de la cible et du nombre d'échantillon.

#### Qualité des données

couverture	en moyenne >95% de la cible couverte
profondeur moyenne de lecture (min-max) projets 'standard'	200X minimum
projets 'deep sequencing'	dépendant de l'objectif du projet en fonction de la fréquence attendue des mutations recherchées.

#### Format de rendu des résultats de l'analyse bioinfo

reads	Fastq
alignements	BAM
détection des variants	fichier tabulaire et vcf
annotation des variants	fichier tabulaire et vcf:idem exome
outils de prédiction de la pathogénicité des variants	
outils existants	Polyphen
développement d'outils propres	non

#### Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement

ERIS Exome Resequencing Intelligent Server, interface d'analyse des données d'exome en ligne.

La Plateforme dispose-t-elle d'une base de données interne de variants? oui

### Délai de rendu des résultats

6-10 semaines

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## RNA-Seq

de 250ng à 4µg d'ARN total, RIN >7

### Rendu actuel des résultats

#### Caractéristiques des données produites

PE 2x75 ou 2x100 bases, quantité de séquences dépendante de l'objectif du projet

#### Format de rendu des résultats de l'analyse bioinfo

- QC extensif des données de séquences,
- Demultiplexage des séquences par échantillon
- Mapping des paires de séquence sur le génome de référence avec le logiciel TopHat
- Détection des jonctions de splice, assemblage et quantification des gènes et transcrits (FPKM) avec Cufflinks
- Annotation des gènes/transcrits basés sur RefSeq, mise en évidence des transcrits connus et des nouvelles isoformes avec CuffCompare
- Détection et annotation des transcrits de fusion avec TopHat-Fusion
- Détection et annotation des variants SNP
- Détection et annotation des long non-coding RNA (nous consulter)

Les résultats livrés: Fichiers Fastq et contrôles qualités associés, alignement BAM, liste des transcrits annotés avec niveau de quantification sous format tabulaire, liste des SNP annotés sous format tabulaire et vcf.

### Délai de rendu des résultats

6 à 10 semaines

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## SmallRNA-Seq

Protocole standard: à partir de 1 à 10µg d'ARN total, RIN >7

Protocole smallRNA circulants: nous consulter

### Rendu actuel des résultats

#### Caractéristiques des données produites

15 à 20 millions de reads de 35 bases par banque

#### Format de rendu des résultats de l'analyse bioinfo

Contrôle qualité du run

Séquences Fastq

Séquences trimmés des adaptateurs (fasta)

Fichiers tabulaires contenant les séquences fasta des miRNAs séquencés et comptage associé

Alignement sur un outil opensource type MirAnalyser pour un comptage des miRNA connus ou putatifs présents, et statistiques associés.

### Délai de rendu des résultats

6-10 semaines

### Accès

Equipes du site

oui

Equipes extérieures académiques

oui

Equipes extérieures privées

oui

## RRBS (Reduced Representation Bisulfite Sequencing)

Quantité de matériel demandé: min. 50 ng d'ADN extrait

### Rendu actuel des résultats

**Caractéristiques des données produites** 20 à 30 millions de paires de 75 bases par banque RRBS

### Format de rendu des résultats de l'analyse bioinfo

- QC extensif des données de séquences.
- Trimming des adaptateurs.
- Conversion bisulfite des séquences.
- Alignement des paires sur le génome "converti bisulfite" en mode RRBS ou WGBS suivant le type de banques.
- Détermination de l'état de méthylation des cytosines; et du taux de méthylation.
- Livraison sur disque dur externe:
- QC extensif des données de séquences;
- Demultiplex des séquences par banque (sur la base des index fournis par le client si IntegraGen reçoit les banques prêtes à séquencer);
- Livraison sur disque dur externe:
  - des contrôles qualité du run pour chaque échantillon: # clusters raw, # clusters PF, % of  $\geq$  Q30 Bases (PF), Mean Quality Score (PF);
  - les fichiers fastq correspondants à chaque banque;
  - rapports qualité FastQC associés.
  - le rapport d'alignement ((% conversion, efficacité du mapping, nombre de cytosines détectées dans les contextes CpG, CHG, et CHH, ...)
  - le fichier d'alignement au format SAM
  - le fichiers d'extraction des cytosines (incluant le taux de méthylation) dans les 3 contextes CpG, CHG, et CHH en mode RRBS; uniquement dans le contexte CpG en mode WGBS.

### Délai de rendu des résultats

8-12 semaines

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## 2. Génomique, transcriptomique

### Principales réalisations

**Domaines de réalisation** génétique humaine; génétique animale et végétale

### Services proposés (domaines d'applications / expertise)

Génotypage SNP  
GWAS  
Pharmacogénétique  
Analyses de méthylation sur puce  
Gene expression (Nanostring)

### Equipe

**Ingénieurs**  
**Techniciens** 6  
**Bioinformaticiens** 1 bioinformaticien et 1 biostatisticien

### Equipements

Illumina iScan  
Fluidigm Biomark  
ABI 7900HT  
Plateforme nanoString

### Technologies utilisées

Illumina Infinium et GoldenGate  
TaqMan  
AS-PCR

### Traitement, analyse et stockage des données

Analyse statistique des résultats de GWAS possible sur cahier des charges de l'étude.

*Mise à jour*

*20/10/2014*

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Transcriptomique et Génomique Marseille-Luminy (TGML)

<b>Responsable</b>	Jean Imbert
<b>Equipe (effectifs)</b>	
ingénieurs	3
bioinformaticiens	3
<b>Structure de rattachement</b>	TAGC UMR1090 Inserm-Université d'Aix-Marseille
<b>Adresse complète</b>	TAGC UMR_S 1090 Inserm – Aix-Marseille Université 163 Avenue de Luminy 13009 Marseille - France
<b>Site Web</b>	<a href="http://tagc.univ-mrs.fr/tagc/index.php/tgml-facility">http://tagc.univ-mrs.fr/tagc/index.php/tgml-facility</a>
<b>Contact</b>	<a href="mailto:jean.imbert@inserm.fr">jean.imbert@inserm.fr</a>
<b>Date de mise en place de la Plateforme</b>	
des activités de séquençage à haut débit	01/09/2009
des activités de génomique	01/09/2009
des activités de transcriptomique	01/01/2002
<b>Labels / démarche qualité</b>	
IBiSA	oui
Certification ISO 9001	certification en cours
Autre	Membre de l'infrastructure nationale distribuée France-Génomique (ANR-10-INBS-09-10)

## Principales réalisations

**Nombre de projets traités depuis la création** plusieurs centaines

**Nombre d'échantillons traités depuis la création** plusieurs milliers

**Nombre de projets traités par an** plusieurs dizaines

**Nombre d'échantillons traités par an** plusieurs centaines

## Domaines de réalisation

**Génétique humaine** maladies rares, oncogénétique, etc.

## Services proposés (domaines d'applications / expertise)

**Séquençage *de novo*** en collaboration uniquement

**Métagénomique** en collaboration uniquement

### Séquençage

de génomes entiers (WGS) oui

d'exomes entiers (WES) oui

de régions ciblées oui

**RNaseq** oui

**SmallRNaseq** oui

**ChIPseq** oui

**Autres** FAIRE-seq, Mnase-seq

## Perspectives de développement de services

Whole genome, single cell sequencing

## Equipements (séquenceur(s), robot(s), informatique)

**Séquençage :** Illumina Nextseq 500, Ion Torrent PGM, OneTouch System, TECAN EVO150 -  
**Informatique :** Illumina BaseSpace OnSite System, Cluster #1 : 3 serveurs Dell PE1950 (2 CPU E5420@2.5GHz, 24 cœurs) ; Cluster # 2 : 4 serveurs Dell R710 (2 CPU E5540@2.53GHz, 32 cœurs) ; Cluster #3 : 16 serveurs Dell C6220 (2 CPU E5-2640@2.5GHz, 192 cœurs). Stockage : 4 volumes (baies Dell MD1000, MD1200 et MD3260) de 140 à 220 Tb.

## Technologies utilisées

<b>Kits de capture</b>	Agilent SureSelect tous formats
<b>Format de séquençage</b>	Illumina NextSeq 500 : 20-120 Gb (fragments 75b et paired-end 2x150b ; Ion Torrent PGM : 30Mb-2Gb (fragments 200-400b)
<b>Temps d'un run</b>	NextSeq 500 : 12-30 heures Ion Torrent PGM : 2-4h

## Multiplexage des échantillons

Nombre d'échantillons/ligne : sur demande car cela dépend de la profondeur de séquençage voulu, de l'espèce et du type de séquençage

## Traitement, analyse et stockage des données

La plateforme offre un service complet d'analyses modulable à la demande

## Développement d'outils de traitement des données

Variants génomiques (GeVaRa), Transcriptome (RNA-seq et smallRNA-seq), Epigénomique (ChIP-seq, FAIRE-seq, Mnase-seq, STAR-seq), suite logicielle RSAT : <http://139.124.66.4/rsat/>

## Rendu actuel des résultats

### Caractéristiques des données produites

Sur demande car cela dépend de la profondeur de séquençage voulu, de l'espèce et du type de séquençage.

La Plateforme dispose-t-elle d'une base de données interne de variants? Non

### Accès

Equipes du site Oui  
Equipes extérieures académiques Oui  
Equipes extérieures privées Oui

### Formation

Participation régulière aux ateliers de formation et aux formations permanentes Inserm et AVIESAN, possibilité de formation individualisée

Liste des expertises pouvant être transmises Toutes celles maîtrisées par la plateforme et les chercheurs et ingénieurs du site

*Mise à jour* 20/10/2014

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### MGX-Montpellier GenomiX



<b>Responsable</b>	Laurent Journot
<b>Equipe (effectifs)</b>	
ingénieurs	3
bioinformaticiens	4
administratifs	1
<b>Structure de rattachement</b>	BioCampus Montpellier - Unité Mixte de Service 3426 (CNRS-INSERM-UM1-UM2)
<b>Adresse complète</b>	c/o Institut de Génomique Fonctionnelle 141, rue de la cardonille 34094 Montpellier Cedex 5
<b>Site Web</b>	<a href="http://www.mgx.cnrs.fr">www.mgx.cnrs.fr</a>
<b>Contact</b>	<a href="mailto:contact@mgx.cnrs.fr">contact@mgx.cnrs.fr</a>
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2008
2. des activités de génomique - transcriptomique	1999
<b>Labels / démarche qualité</b>	
IBiSA	oui (2006)
Certification ISO 9001	oui (2010)
Autre	Infrastructure nationale "France Génomique"

## Principales réalisations

Nombre de projets traités depuis la création >250

Nombre d'échantillons traités depuis la création >2500

Nombre d'échantillons traités par an ~700

## Domaines de réalisation

**Génétique humaine** oui  
*préciser les spécialités* séquençage de génomes et d'exomes, transcriptomique

**Génétique animale et végétale** oui

**Microorganismes** peu

**Autres** biodiversité, agronomie

## Services proposés (domaines d'applications / expertise)

**Séquençage *de novo*** oui

**Séquençage**  
de génomes entiers (WGS) oui  
d'exomes entiers (WES) oui  
de régions ciblées oui

**RNaseq** oui

**SmallRNaseq** oui

**ChIPseq** oui

**Autres** RADseq, HRSseq

## Perspectives de développement de services

Mise en place de la technologie "Moleculo" (Illumina)

## Equipements (séquenceur(s), robot(s), informatique)

HiSeq2000, HiSeq1500 (2014), robots Tecan Genesis + Evo, serveurs Dell

## Technologies utilisées

<b>Kits de capture</b>	Agilent SureSelect, Illumina
<b>Format de séquençage</b>	SR 50, PE 2x100
<b>Temps d'un run</b>	SR 50: 3 jours; PE 2x100 : 14 jours
<b>Multiplexage des échantillons</b> nombre d'échantillons/ligne	oui jusqu'à 96 (RADseq)

## Traitement, analyse et stockage des données

oui

## Rendu actuel des résultats

<b>Caractéristiques des données produites</b>	fixées par le demandeur
---	-------------------------

### Qualité des données

couverture	fixée par le demandeur
profondeur moyenne de lecture (min-max)	50-2x100

### Format de rendu des résultats de l'analyse bioinfo

reads	oui
alignements	oui
détection des variants	oui
annotation des variants	oui

La Plateforme dispose-t-elle d'une base de données interne de variants?	non
---	-----

## Accès

Equipes du site	2-5 mois selon le nombre d'échantillons
Equipes extérieures académiques	2-5 mois selon le nombre d'échantillons
Equipes extérieures privées	2-5 mois selon le nombre d'échantillons

## Formation

Ateliers Technologiques BioCampus  
RNAseq  
Statistiques

*Mise à jour*

25/11/2013

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Biogenouest Génomique

<b>Responsable</b>	Richard Redon
<b>Equipe (effectifs)</b>	
ingénieurs	8
techniciens	5
bioinformaticiens	4
administratifs	0
autres (préciser)	0
<b>Structure de rattachement</b>	Plateforme multisite : SFR Santé de Nantes, SFR Biosit de Rennes, SFR Osur de Rennes, Station biologique de Roscoff
<b>Adresse complète</b>	L'institut du thorax, IRS-UN, 8 quai Moncoussu, BP70721, 44007 Nantes cedex 1
<b>Site Web</b>	<a href="http://www.biogenouest.org/contenu/plateformes/genomique/biogenouest-genomique">http://www.biogenouest.org/contenu/plateformes/genomique/biogenouest-genomique</a>
<b>Contact</b>	<a href="mailto:richard.redon@univ-nantes.fr">richard.redon@univ-nantes.fr</a>
<b>Date de mise en place de la Plateforme</b>	
des activités de séquençage à haut débit	2012
des activités de génomique - transcriptomique	2004
<b>Labels / démarche qualité</b>	
IBiSA	2011
Certification ISO 9001	2012
<i>Mise à jour</i>	<i>06/09/2013</i>

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme de Nice Sophia Antipolis

<b>Responsable</b>	Pascal Barbry
<b>Equipe (effectifs)</b>	
ingénieurs	4
bioinformaticiens	4
administratifs	1
<b>Structure de rattachement</b>	IPMC UMR7275 CNRS/UNS
<b>Adresse complète</b>	IPMC UMR7275 CNRS/UNS 660, route des lucioles 06560 Sophia Antipolis
<b>Site Web</b>	<a href="http://www.genomique.info">http://www.genomique.info</a>
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2008
2. des activités de génomique - transcriptomique	1999
<b>Labels / démarche qualité</b>	
IBiSA	oui (partenaire de l'infrastructure nationale France-Génomique)
Certification ISO 9001	depuis 2006

## 1. Séquençage à haut débit

### Principales réalisations

<b>RNAseq</b>	82 projets, 543 banques
<b>ChIPseq</b>	9 projets, 155 banques
<b>SmallRNAseq</b>	94 projets, 927 banques
<b>SAGE</b>	5 projets, 26 banques
<b>De novo sequencing</b>	3 projets, 3 banques
<b>CLIPseq</b>	10 projets, 47 banques

### Services proposés (domaines d'applications / expertise)

<b>Séquençage <i>de novo</i></b>	oui (sur petits génomes)
<b>RNAseq</b>	oui
<b>SmallRNAseq</b>	oui
<b>ChIPseq</b>	oui

### Technologies utilisées

#### Équipement

SOLiD™ WildFire , PGM, Ion Proton (Life Technologies)

Agilent scanner

tout l'équipement de base pour la réalisation des analyses (bioanalyzer, Qbit, Covaris,...)

C1, Biomark (Fluidigm)

#### Format de séquençage

Small RNAseq (50 bases); RNAseq (75 bases); MATE-pair (2 x 50 bases); paired-end (75 bases + 50 bases); DNA (2 x 50 bases)

#### Temps d'un run

SOLiD: de 4 jours (smallRNAseq, 50 bases) à ~15 jours (mate-pair, 2 x 50 bases).

PGM/Ion Proton: <24 heures

#### Multiplexage des échantillons

nombre d'échantillons/ligne

dépend de la nature du projet :

Small RNAseq = 60 échantillons/ligne

RNAseq = 6 échantillons/ligne

MATE-pair = 1 échantillon/ligne

paired-end = 2-4 échantillons/ligne

## Rendu actuel des résultats

<b>Caractéristiques des données produites</b>	Dépend de la nature du projet. smallRNAseq = 10 millions de séquences. RNAseq = 100 millions de séquences.
---	--

## Format de rendu des résultats de l'analyse bioinfo

reads	oui
alignements	oui (fichiers bam)
détection des variants	oui
annotation des variants	oui

## Outils de prédiction de la pathogénicité des variants

Oncotator, Polyphen, Variant Effect Predictor, SIFT

La Plateforme dispose-t-elle d'une base de données interne de variants?	utilisation des données de dbSNP, 1000GP, COSMIC
---	--

## Accès

Trois tarifications différentes, validées par l'organisme gestionnaire (CNRS)

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## 2. Génomique, transcriptomique

### Principales réalisations

<b>Nombre de projets traités depuis la création</b>	1058 projets depuis 2004 (13046 échantillons)
<b>Nombre de projets traités par an</b>	100 projets par an, 1300 analyses par an
<b>Domaines de réalisation</b>	Transcriptomique (mRNA, smallRNA, ChIPseq, de novo)

### Services proposés (domaines d'applications / expertise)

The platform has analyzed so far 1713 distinct libraries. A large, but not exclusive, fraction corresponds to RNA libraries (RNAseq: 543 samples; small RNAseq: 927 samples). Several pipelines are now fully validated: (1) DNA Fragment Libraries, appropriate for resequencing, also suitable for ChIP-Seq experiments. We have optimized the preparation and analysis of mate-pair libraries, which have been successfully used in a *de novo* sequencing of a fungal genome. A PGM and an Ion Proton are also available for multiple purposes. (2) directional RNA and small RNA libraries. Protocols for RNAseq have been optimized to work with small RNAs (miRNAs, piRNAs, ...), total RNA depleted of ribosomal RNA, mRNA (polyA+ purifications), and also minute amounts of total RNA (1-10ng). These protocols allow comprehensive profiling and mutation discovery at the level of a full genome. A SAGE approach was also validated (Chatonnet et al, 2013). Several in-house adaptations, such as the use of barcodes, have transformed the high throughput sequencing into a cost-effective technology in our hands. Our specific expertise in different types of HTS/microarray data analyses has been acknowledged in more than 80 publications since 2003, including several clinically-oriented gene expression studies ([http://www.genomique.info/joomla\\_2.5.9/index.php/homepage/related-platform-publications](http://www.genomique.info/joomla_2.5.9/index.php/homepage/related-platform-publications)).

### Perspectives de développement de services

Tout domaine d'analyse des ARN  
Extension en cours à l'analyse du contexte chromatinien

### Equipe

Ingénieurs	4
Bioinformaticiens	4
Administratifs	1
Autres	lien fonctionnel important avec l'équipe Physiologie Génomique des Eucaryotes de l'IPMC

## Equipements

SOLiD™ WildFire (Life Technologies), PGM, Ion Proton

## Traitement, analyse et stockage des données

A strong emphasis has been made since the creation of the platform regarding data storage, and a specific information system, called Mediante, has been developed for more than 10 years. All experimental data, including runs of high throughput sequencing, are accessible today through the Mediante server, which functions as a J2EE three-tier application (Le Brigand and Barbry, 2007).

At the end of every analysis, results are backed-up according to a standardized procedure (ISO 9001 certified in June 2010) in the Mediante information system (<http://www.microarray.fr:8080/merge/index>). Mediante is a MIAME-compliant gene expression data manager that links together annotations and experimental data. Developed as a J2EE three-tier application, Mediante integrates a management system for production interface and a user interface dedicated to the management of microarrays/sequencing projects. Several tools allow quality control of hybridizations and submission of validated data to public repositories. The backup format is based on the latest standards of the domain (HTS MINSEQE, <http://www.mged.org/minseqe/>). Mediante data are replicated once per week on external backup hard drives.

## Développement d'outils de traitement des données

<http://www.microarray.fr:8080>

Les différents pipelines sont proposés aux utilisateurs dans le cadre du système d'information Mediante, mis en place sur la plate-forme depuis 2003.

## **RNAseq:**

Whole transcriptome pipeline in Lifescape v2.5.1 (quality analysis of the reads via FastQC, filtering against a non specific sequence database (repeated sequences, barcodes, adapters, ribosomal RNA, ...), mapping on the genome reference with mapreads (corona), quantification and counting of transcripts based on \*.gff files. In early July 2013, the paired-end option has been installed on our SOLiD sequencer, in order to facilitate the identifications of translocations in our RNA samples.

RNAseq and smallRNA-seq data are analyzed for differential gene expression using a in-house Bioconductor pipeline, developed with libraries such as DEseq. This pipeline allows the generation of a standardized PDF report that includes different quality controls as well as information about the quality of the samples.

## **SmallRNA-seq:**

The smallRNAAnalyzer pipeline was developed in-house for processing smallRNA-seq libraries. It integrates the smallRNAseq pipeline from Lifescape2.5.1 that filters the reads against a database of sequences (adapters, barcodes, rRNAs, homopolymers) and maps them against the ad hoc genome. The creation of a \*.bam file allows detection and counting of all small RNAs present in the samples. Java classes (derived from Picard libraries) have been developed for counting. Usual reference databases correspond to miRBase v18, tRNAs from UCSC, YRNAs, non-coding RNA Ensembl ncRNA r69, fRNAdb (piRNA). When necessary, repeated sequences from RepeatMasker are also incorporated into the analysis. The mapping of expressed yet unannotated regions with MACS allows the detection of potential new small RNA loci.

RNAseq and smallRNA-seq data are analyzed for differential gene expression using a in-house Bioconductor pipeline, developed with libraries such as DEseq. This pipeline allows the generation of a standardized PDF report that includes different quality controls as well as information about the quality of the samples.

## **De novo RNAseq:**

A *de novo* RNA-seq pipeline for the treatment of SOLiD data was developed: it includes error correction reads with SAET, filtering low-quality reads with the SOLiD preprocessor suite, standardization with the digital-CTB SOFTWARE Khmer in order to decrease the amount of reads in assembler input (i.e. it normalizes the covers for highly expressed transcripts). The Velvet / Oases software suite is used to assemble at different values of k-mer. The conversion to base space is done with denovo2. Functional annotation is made by BLASTX against Uniprot and / or against sequences of related species. Count and differential analyses are done using Bowtie.

### **De novo genomic DNA:**

*De novo* genome assembly requires a important filtering of sequencing errors. SOLiD preprocessor directly filters the reads based on quality parameters obtained during sequencing. Good quality mate-paired reads are then analyzed with denovo2 (suitable for genomes of no more than ~ 50Mb). The pipeline integrates SAET for correcting specific sequencing errors, Velvet (version 1.0.15) for de novo reconstruction, and Assembly Assistant for SOLiD™ (ASID v.1.0) for scaffolding and conversion to base space.

### **ChIP-seq:**

Mapping is done with the Lifescope2.5.1 suite. Peaks are detected with MACS. Peaks are annotated with PeakAnalyser. Motifs are detected with MEME.

### **SNP calling, SNV detection:**

Performed using the Ion Torrent server, GATK dbSNP, Annovar.

### **Rendu actuel des résultats**

Document pdf + lien vers différents fichiers accessibles via l'interface Mediante

### **Délai de rendu des résultats**

Dépend du type de projets

### **Accès**

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

*Mise à jour* 06/10/2014

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme Génomique Imagine & Plateforme de Bioinformatique Paris-Descartes

Les données produites par la plateforme de Génomique sont analysées par la plateforme de Bioinformatique.

La plateforme de Bioinformatique analyse également des données issues d'autres plateformes pour des utilisateurs de l'Université Paris V.

#### Plateforme Génomique Imagine

##### Responsables Plateforme Génomique

Christine Bole-Feysot

Responsable scientifique : Laurence Colleaux

##### Equipe (effectifs)

ingénieurs

2

techniciens

3

administratifs

0 (intervention d'agents administratifs extérieurs : IFR94 & Imagine)

##### Structure de rattachement

Imagine, Institut des maladies génétiques & ISFR-Necker

##### Adresse complète

Imagine - Institut des maladies génétiques  
Plateforme Génomique, 3e étage, Bureau 333  
24 Boulevard du Montparnasse - 75015 Paris

##### Site Web

<http://www.institutimagine.org/fr/la-recherche/8-plateformes-technologiques/31-genomic-platform.html>

##### Contact

Christine Bole-Feysot : christine.bole@inserm.fr

##### Date de mise en place de la Plateforme

1. des activités de séquençage à haut débit avr-10
2. des activités de génomique - transcriptomique mai-08

## Plateforme de Bioinformatique Paris-Descartes

<b>Responsables Plateforme de Bioinformatique</b>	Patrick Nitschké Responsable scientifique : Jean-Philippe JAÏS
<b>Equipe (effectifs)</b> bioinformaticiens	7
<b>Structure de rattachement</b>	Université Paris Descartes
<b>Adresse complète</b>	Plateforme de Bioinformatique Paris-Descartes Hopital Necker, Batiment Lavoisier, 7 eme étage 75730 PARIS Cedex 15
<b>Site Web</b>	<a href="http://www.institutimagine.org/fr/la-recherche/8-plateformes-technologiques/32-bioinformatic.html">http://www.institutimagine.org/fr/la-recherche/8-plateformes-technologiques/32-bioinformatic.html</a>
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2010
2. des activités de génomique - transcriptomique	2007

## 1. Séquençage à haut débit

### Principales réalisations

#### Nombre d'échantillons et de projets totaux traités en NGS par les Plateformes Génomique Imagine et Bioinformatique Paris-Descartes

Entre 2010 et mi-novembre 2014 :

Nombre d'échantillons traités en Whole Exome Sequencing (WES) :

Séquençage : 2560 exomes en tout (253 projets),

Analyse bioinformatique: 4740 exomes et 402 projets

Nombre d'échantillons de reséquençage ciblé hors WES ou autres applications :

Séquençage : 3237 échantillons en tout (162 projets)

Analyse bioinformatique : 3500 et 130 projets

#### Nombre d'échantillons et de projets traités en 2014 par la Plateforme Génomique Imagine

De janvier à mi-novembre 2014 : 2501 échantillons traités en NGS (188 projets) dont 1036 exomes (122 projets)

#### Nombre d'échantillons et de projets traités en 2014 par la Bioinformatique Paris-Descartes

De janvier à mi-novembre 2014 : 1882 échantillons + 2113 gènes ciblés

### Domaines de réalisation

#### Génétique humaine

maladies génétiques rares essentiellement

#### Microorganismes

ponctuellement séquençage de génomes de bactéries et de virus

### Services proposés (domaines d'applications / expertise)

#### Séquençage

de génomes entiers (WGS)

WGS de petits génomes uniquement sous certaines conditions

d'exomes entiers (WES)

oui, principale activité

de régions ciblées

oui, sous certaines conditions

#### RNAseq

en cours de mise en place

#### ChIPseq

en cours de mise en place

#### Autres

reséquençage ciblé à la carte ou autres types de projets utilisant le séquençage de nouvelle génération (sous certaines conditions)

## Perspectives de développement de services

Whole Genome Sequencing + projets de reséquençage ciblé à façon (enrichissement obtenu par différentes méthodes : PCR Multiplex, capture par hybridation, autre)

## Equipements (séquenceur(s), robot(s), informatique)

2 HiSeq2500, Illumina (up-grade HiSeq2500, chimie V4 en novembre 2014)  
1 Ion Torrent PGM  
1 MiSeq  
1 Cluster : 21 nodes totalizing 252 CPU cores, 1 To RAM and 30TB of storage disk  
1 Application Server : 8 cores, 96 GB RAM, 1 TB of storage disk  
1 Database Server : 8 cores, 96 GB RAM, 40TB of storage disk  
1 Clustered Network Attached Storage : 120 TB of storage disk  
1 Clustered Network Attached Storage : 500 TB of storage disk

## Technologies utilisées

### Kits de capture

WES : Agilent SureSelect All exon V5, 51Mb (Agilent)

autres méthodes d'enrichissement : capture par hybridation SureSelect Custom (Agilent) : différents types de panels utilisés

PCR Multiplex (Multiplicom), AmpliSeq (Life Technologies), Long Range PCR, séquençage de produits de RT-PCR...

### Format de séquençage

dépend des projets et du séquenceur utilisé  
WES: HiSeq2500, mode HighOutput, 2x100 bases (chimie V3) (prochainement 2x125 bases, chimie V4)

### Temps d'un run

dépend des projets, du séquenceur utilisé, du nombre de cycles de séquençage - WES : run HiSeq2500, mode HighOutput, 2x100bases : ~10 jours (amplification clonale + séquençage)

### Multiplexage des échantillons

nombre d'échantillons/ligne

dépend des projets, WES : capture simplex (1 capture par échantillon) et séquençage sur HiSeq2500, mode HighOutput multiplex : 3 exomes par ligne (chimie V3)

## **Whole Exome Sequencing (WES)**

Capture SureSelect All exon V5, 51Mb (Agilent) , séquençage sur HiSeq2500 (Illumina)

### **Rendu actuel des résultats**

#### **Caractéristiques des données produites**

séquençage Paired-End 100+100, équivalent de 3 exomes par ligne de FlowCell (mode High Output, chimie V3)

#### **Qualité des données**

couverture

en moyenne : 99% à 5X, >97% à 15X, >90% à 30X

profondeur moyenne de lecture (min-max)

>100X

#### **Format de rendu des résultats de l'analyse bioinfo**

Interface web d'interrogation développée in-house (polyweb)

reads

sur demande

alignements

sur demande

détection des variants

sur demande

annotation des variants

sur demande

#### **Outils de prédiction de la pathogénicité des variants**

outils existants

polyphen sift

### **Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement**

Plateforme de Bioinformatique Paris-Descartes

<http://www.institutimagine.org/Larecherche/Lesplateformestechiques/Lesfichesplateformes/Plateformebioinformatique.aspx>

La Plateforme dispose-t-elle d'une base de données interne de variants?

oui

### **Délai de rendu des résultats**

~3 mois à réception des ADN

#### **Accès**

Equipes du site

oui

Equipes extérieures académiques

oui, sous certaines conditions

Equipes extérieures privées

non

#### **Formation**

Formation à notre outil polyweb de visualisation de résultat de WES (1 fois par mois).

## 2. Génomique, transcriptomique (autre que séquençage de nouvelle génération)

### Principales réalisations

<b>Nombre de projets traités depuis la création</b>	Puces + Analyse bioinformatique : 2179 échantillons traités entre mai 2008 et mi-novembre 2014 (252 projets) Analyse bioinformatique uniquement : 2600 échantillons analysés (puces faites par une autre plateforme que la Plateforme Génomique Imagine)
<b>Nombre de projets traités par an</b>	Puces + Analyse bioinformatique : entre 200 et 500 échantillons analysés par an Analyse bioinformatique uniquement : entre 500 et 750 échantillons analysés par an (puces faites par une autre plateforme que la Plateforme Génomique Imagine)
<b>Domaines de réalisation</b>	Maladies génétiques, immunologie, signalisation, autres...

### Services proposés (domaines d'applications / expertise)

Analyses sur puces Affymetrix : étude de liaison sur puce de génotypage, transcriptomique sur divers types de puces expression, études des Copy Number Variations et des pertes d'hétérozygocité (puces CytoScan HD)

### Perspectives de développement de services

Mise en place de nouveaux types de puces Affymetrix suivant la demande

### Equipements

1 plateforme de microarrays Affymetrix

### Technologies utilisées

Microarrays Affymetrix, kits Affymetrix et kits Nugen

## Traitement, analyse et stockage des données

Utilisation des outils d'analyse primaire et secondaire d'Affymetrix : Commande Console, Genotyping Console, Expression Console, Chromosome Analysis Suite  
Etude de liaison : merlin, transcriptomique : bioconductor, R etc...  
Ingenuity

Analyses sur puces Affymetrix : nous contacter pour plus de détails  
1-étude de liaison sur puce de génotypage,  
2-transcriptomique sur diverses puces expression  
3-études des Copy Number Variations et des pertes d'hétérozygoté

## Accès

Equipes du site	oui
Equipes extérieures académiques	sous certaines conditions
Equipes extérieures privées	non

*Mise à jour* 19/11/2014

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme de génomique et microgénomique : profileXpert

<b>Responsable</b>	Pr Joel LACHUER
<b>Equipe (effectifs)</b>	<b>14</b>
ingénieurs	4
techniciens	2
bioinformaticiens	2
administratifs	1
autres (préciser)	professeurs et maîtres de conférences 5
<b>Structure de rattachement</b>	Université Lyon1; SFR santé Lyon Est , UCBL UMS 3453 CNRS – US7 INSERM
<b>Adresse complète</b>	Faculté de Médecine et Pharmacie Rockefeller 2ème étage aile C et 4ème étage aile B 8 avenue Rockefeller 69008 Lyon
<b>Site Web</b>	<a href="http://www.profilexpert.fr">www.profilexpert.fr</a>
<b>Contact</b>	<a href="mailto:Contact@profilexpert.fr">Contact@profilexpert.fr</a>
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2011
2. des activités de génomique - transcriptomique	2004
<b>Labels / démarche qualité</b>	
IBiSA	depuis 2009
Certification ISO 9001	audit LRQA : 21 et 22 octobre 2013

## 1. Séquençage à haut débit

### Principales réalisations

<b>Nombre de projets traités depuis la création</b>	75 : 9 (2011) 34 (2012) 32 (juillet 2013)
<b>Nombre d'échantillons traités depuis la création</b>	883 : 58 (2011) 344 (2012) 481 (juillet 2013)
<b>Nombre de projets traités par an</b>	37
<b>Nombre d'échantillons traités par an</b>	441

### Domaines de réalisation

<b>Génétique humaine</b> <i>préciser les spécialités</i>	8 projets exome et haloplex
<b>Génétique animale et végétale</b>	6 projets (exome et DNA seq de novo ou reséquençage)
<b>Microorganismes</b>	23 projets (DNAseq bactérien et virus discovery)
<b>Autre (préciser)</b>	38 projets (RNA seq, Chip-seq, RAD-seq, RRBS)

### Services proposés (domaines d'applications / expertise)

<b>Séquençage <i>de novo</i></b>	oui : (RNA seq) et (DNA seq) : évolution, écologie, toxicologie, infectiologie
<b>Métagénomique</b>	oui : rad seq : écologie, infectiologie
<b>Séquençage</b> de génomes entiers (WGS)	oui (homme, souris, espèces modèles) : cancérologie, génétique humaine, infectiologie
d' exomes entiers (WES)	oui : cancérologie, neurosciences, génétique humaine
de régions ciblées	Haloplex, amplicon classique, enrichissement en phase liquide custom (voir kit enrichissement) : cancérologie; génétique humaine

<b>RNAseq</b>	oui : tous domaines
<b>SmallRNAseq</b>	oui : tous domaines
<b>ChIPseq</b>	oui : tous domaines
<b>MeDIPseq</b>	oui : oncologie, neurosciences
<b>Autres</b>	Tnseq et INSseq : recherche de sites d'insertion de génome viraux, de transposons RRBS (recherche de site de méthylation sur génome de complexité réduite) Bar-seq Génome entier bisulfite BS-seq (étude de méthylation)

### Perspectives de développement de services

Développement de la génomique sur cellule isolée: acquisition C1 de fluidigm ou deeparray exyclone. En cours de discussion, financement acquis  
NGS sur échantillons FFPE  
Développement de système d'enrichissement custom

### Equipements (séquenceur(s), robot(s), informatique)

Hiseq 2500 (Illumina) et MiSeq (Illumina)  
Cluster informatique HP (48 To memory, 8 cores)  
Access array (fluidigm)

### Technologies utilisées

<b>Kits de capture</b>	Truseq exome enrichment human (Illumina) SureSelect XT mouse all exon kit (Agilent) Nimblegen SeqCap EZ library SR (Nimblegen) : custom virus
<b>Format de séquençage</b>	Hiseq2000/2500 (single read 50 et PE 2x100) Miseq (SR 250/500 ou PE 2x250bp) Autres sur demande
<b>Temps d'un run</b>	variable
<b>Multiplexage des échantillons</b>	variable

### Traitement, analyse et stockage des données

Traitement des données par des logiciels commerciaux.  
Stockage des données sur bandes magnétiques.

## RNAseq

### Rendu actuel des résultats

#### Caractéristiques des données produites

HiSeq 2000/2500 (single read 50 ou 100 bp; paired end 2x50 ou 2x100bp) Déplétion ribosomique ou DSN optionnel, RNAseq directionnel

Quantité minimum de séquences produites par échantillon : 25 millions reads (homme, rat , souris)

#### Qualité des données

couverture	variable
profondeur moyenne de lecture (min-max)	variable
projets 'standard'	RNAseq (homme) : 6 échantillons par ligne/25 millions de reads/échantillon (comptage classique) SR 50 bp
projets 'deep sequencing'	RNAseq (homme) : 1 échantillon par ligne/150 millions de reads/échantillon (Etude polymorphismes, fusion) PE 50 bp

#### Format de rendu des résultats de l'analyse bioinfo

Chaque étape est indépendante et peut faire l'objet de service à la carte

reads	demultiplexing (CASAVA 1.8.2) : création de fichiers FASTQ
alignements	alignement/genome de référence hg19 (TopHatBowtie 2.0.2): BAM files visualisation par IGV

#### Autres

Rapport qualité  
Analyse statistique (tri multicritères)  
Clusterisation, ACP  
Voies de signalisation (ingenuity) Gene fusion, polymorphismes, variants (Partek GS 6.6 software)

### Délai de rendu des résultats

1 mois

#### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

#### Formation

via la société Fc3Bio

## WES

### Truseq enrichment kit (or SureSelect Agilent kit) & paired-end (2x100) Séquençage sur hiseq2000 (homme, rat, souris)

#### Rendu actuel des résultats

**Caractéristiques des données produites** 6 échantillons par ligne/25-30 millions de reads/échantillon

#### Qualité des données

Couverture 99,5% avec au moins 1 read  
Profondeur moyenne de lecture (min-max) 48-52%  
projets 'standard' 6 échantillons par ligne  
projets 'deep sequencing' 3 échantillons par ligne

#### Format de rendu des résultats de l'analyse bioinfo

Chaque étape est indépendante et peut faire l'objet de service à la carte

reads demultiplexing (CASAVA 1.8.2) : création de fichiers FASTQ  
alignements alignement/génomome de référence hg19 : BAM  
files visualisation par IGV  
détection des variants CNV (FREEC), SNP et InDel (Samtools)  
annotation des variants snpeff  
outils de prédiction de la pathogénicité des variants  
outils existants snpEff 3.2.a (SIFT score, Polyphen score, dbsnp, 1000 exomes db, 1000 genomes db)

La Plateforme dispose-t-elle d'une base de données interne de variants? oui

#### Accès

Equipes du site oui  
Equipes extérieures académiques oui  
Equipes extérieures privées oui

#### Formation

via la société Fc3Bio

## WGS

### Paired-end (2X100) & séquençage sur Hiseq2000 (homme, rat, souris)

#### Rendu actuel des résultats

<b>Caractéristiques des données produites</b>	1 échantillon pour 3 lignes/450 millions de reads/échantillon
---	---

#### Qualité des données

Couverture	93,6%
Profondeur moyenne de lecture (min-max) projets 'standard'	30,1X 1 échantillon pour 3 lignes (30X)

#### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

#### Formation

Via la société Fc3Bio

## Small RNA seq (homme)

### Single read (50) sur Hiseq 2500 (Illumina)

#### Rendu actuel des résultats

<b>Caractéristiques des données produites</b>	12 échantillons par lignes/15 millions reads/échantillon
---	--

#### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

#### Formation

Via la société Fc3Bio

## ChIP-seq

Single read (50) sur Hiseq 2500 (Illumina)

### Rendu actuel des résultats

<b>Caractéristiques des données produites</b>	3 échantillons par lignes/50 millions reads/échantillon
---	---

### Qualité des données

couverture	ND
profondeur moyenne de lecture (min-max)	30 à 100x
projets 'standard'	3 échantillons par lignes/50 millions reads/échantillon

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

### Formation

Via la société Fc3Bio

## Haloplex

Sur Miseq (Illumina) : exemple homme: 48 échantillons, 16 gènes, 19 régions, 3000 amplicons

### Rendu actuel des résultats

**Caractéristiques des données produites** 48 échantillons par ligne : 300 000-400 000 reads/amplicon de 100bp (moyenne)

### Qualité des données

couverture	98,9 à 99,1% des cibles
profondeur moyenne de lecture (min-max)	300 à 400 fois
efficacité d'enrichissement	93%

### Format de rendu des résultats de l'analyse bioinfo

Chaque étape est indépendante et peut faire l'objet de service

reads	demultiplexing (CASAVA 1.8.2) : création de fichiers FASTQ
alignements	alignement/génome de référence hg19 (Bowtie 2.0.2) : BAM files visualisation par IGV
détection des variants	variants calling (SNP/InDel) : Samtools 0.1.16
annotation des variants	variants annotation : snpEff 3.2.a
outils de prédiction de la pathogénicité des variants	
outils existants	snpEff 3.2.a (SIFT score, Polyphen score, dbsnp, 1000 exomes db, 1000 genomes db)
développement d'outils propres	non

La Plateforme dispose-t-elle d'une base de données interne de variants? oui

### Délai de rendu des résultats

1 mois (en fonction du nombre d'échantillon à analyser et du nombre d'amplicons)

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

### Formation

Via la société Fc3Bio

## 2. Génomique, transcriptomique

### Principales réalisations

Nombre de projets traités depuis la création 564

Nombre de projets traités par an 84

### Domaines de réalisation

#### Transcriptome, Génotypage, Méthylome

### Services proposés (domaines d'applications / expertise)

#### QUANTITE STANDARD

RNA chip : puces à ADN (Affymetrix, Illumina, Agilent, custom)

miRNA-chip : Agilent, Affymetrix, TLDA

ChIP-chip : affymetrix

CNV : Puces à ADN (agilent CGH, affymetrix SNP 6.0)

SNP : Affymetrix 6.0

SNP, CNV, translocation et gene fusion : NGS

Expression classique gènes dédiés : PCR (light cycler, LC480)

#### MICROQUANTITE (100 à 200 cellules)

##### (eucaryotes et procaryotes)

RNA chip : puces à ADN (Affymetrix)

RNA seq : NGS

miRNA (TLDA, Applied)

CNV et SNP : puces à ADN (agilent CGH, affymetrix SNP 6.0)

SNP : affymetrix 6.0

Microdissection Laser

Amplification ARN ou ADN

#### Echantillon paraffinés ou difficiles

puces à ADN et NGS

### Perspectives de développement de services

Transcriptomique et génomique sur cellule unique (1 cellule)

Développement de protocole pour l'étude d'échantillons difficiles

## Equipe

ingénieurs	4
techniciens	2
bioinformaticiens	2
administratifs	1
autres (préciser)	professeurs et maîtres de conférences : 5

## Equipements

Plateforme Affymetrix (puces à ADN)  
Plateforme Agilent (puces à ADN)  
Plateforme Iscan Illumina (beads arrays)  
Scanner Axon 2000B  
Light cyclers, LC480 (Roche) et  
HiSeq 2500, MiSeq (séquenceurs), Bioruptor (diagenode)  
Microdissecteur laser (Pixcell arturus) : microgénomique  
Bioanalyseur 2100 (agilent), Robot Qiacube (extraction ADN, ARN), nanodrop, Cubit

## Technologies utilisées

microarrays, NGS, PCR, microdissection laser

## Traitement, analyse et stockage des données

Partek software  
Ingenuity software

## RNA-chip

### Quantité standard ou microquantités

#### Rendu actuel des résultats

Rapport qualité (contrôle qualité)  
Signaux normalisés (médiane) des puces format excel  
Analyse statistique des données (tri multicritères)  
Clusterisation, ACP  
Voies de signalisation (ingenuity)

#### Délai de rendu des résultats

15 jours à partir réception échantillons et consommables

#### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## miRNA-chip

#### Rendu actuel des résultats

Rapport qualité (contrôle qualité)  
Signaux normalisés (méthode globale) des puces  
Analyse statistique des données (tri multicritères)  
Clusterisation, ACP  
Voies de signalisation (ingenuity)  
Recherche des ARNm cibles

#### Délai de rendu des résultats

15 jours à partir réception échantillons et consommables

#### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## Méthylation

### Rendu actuel des résultats

Rapport qualité (contrôle qualité)  
Signaux normalisés des puces  
Recherche des sites méthylés (enrichissement)

### Délai de rendu des résultats

15 jours à partir réception échantillons et consommables

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

## CNV & SNP

### Rendu actuel des résultats

rapport qualité (contrôle qualité)  
signaux normalisés des puces  
genotype

### Délai de rendu des résultats

15 jours à partir réception échantillons et consommables

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

*Mise à jour*

*06/09/2013*

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme de Génomique - Lille

#### Responsable

Amélie Bonnefond (Recherche) Véronique Dhennin (Prestations) sous la responsabilité de Jérôme Delplanque et Philippe Froguel

#### Structure de rattachement

CNRS

#### Adresse complète

Génomique et Maladies Métaboliques-UMR 8199  
1 rue du Professeur Calmette 59000 Lille

#### Site Web

<http://www-good.ibl.fr/http://www-good.ibl.fr/index.php/services-prestations>

#### Contact

contact@good.ibl.fr  
veronique.dhennin@good.ibl.fr

#### Date de mise en place de la Plateforme

1. activités de séquençage à haut débit 2009
2. activités de génomique - transcriptomique 2008

#### Labels / démarche qualité

IBiSA 2009  
Certification ISO 9001 non  
Autres :

mise en place de cahiers d'utilisation auprès de certains équipements (fin 2011), mise en place de calendriers de réservation via intranet (fin 2011), mise en place d'un système de surveillance des frigos (début 2012). Pour les prestations relatives aux extractions d'ADN : mise en place d'un LIMS (2010). Tableau excel pour le suivi des consommables (2012), fiches et procédures d'utilisation des équipements, mise en place d'un système de réception des échantillons (2013).

## 1. Séquençage à haut débit

### Principales réalisations

#### Nombre de projets traités depuis la création

Environ 120 projets. Ce total ne tient pas compte du volume d'échantillons / projet.

#### Nombre d'échantillons traités depuis la création

les mises au point ou test des différentes librairies sont reprises dans le décompte ci après

120 exomes séquencés / prestations & collaborations

735 exomes séquencés / recherche (capture Agilent 38Mb, 50Mb, 71Mb, 50Mb\_diab,

230 échantillons en séquençage ciblé / recherche

#### Nombre de projets traités par an

en moyenne : 10 projets collaboration & recherche / an. Cette moyenne ne tient pas compte du volume d'échantillons / projet.

projets prestation : très variable d'une année à l'autre.

Décomptes annuels :

2009	10 projets collaboration & recherche
2010	2 projets prestation GAIIX : exomes
2011	10 projets collaboration & recherche
2012	Un total de 10 projets de type prestation Hiseq : 5
2013	17 projets réalisés en prestation incluant la préparation de librairies d'Exome, puis séquençage avec ou sans les analyses bioinformatiques, et les projets ready to load 10 projets collaboration & recherche
2014	28 projets réalisés en prestation incluant la préparation de librairies d'Exome, puis séquençage avec ou sans les analyses bioinformatiques, et les projets ready to load 10 projets collaboration & recherche

#### Nombre d'échantillons traités par an

la capacité des équipements utilisés en 2009, 2010 et jusqu'à août 2011 était liée à la capacité du GAIIX. Août 2011 : acquisition Hiseq1, décembre 2012 : installation second Hiseq : capacités beaucoup plus importantes. Acquisition de deux Miseq

2010 : estimation : 50 exomes

2011 : estimation : 80 exomes

2012 : estimation : 280 exomes - 80 librairies raindance séquencées via Hiseq

2013 : 270 exomes - 50 librairies raindance séquencées via Miseq

2014 : 175 exomes - 100 librairies raindance et haloplex séquencées via Miseq

## Domaines de réalisation

### Génétique humaine

séquençage d'exomes et séquençage ciblé via les technologies raindance et haloplex

### Génétique animale et végétale

en prestations : séquençage génomique de cione, plasmodium, loup de mer. RNAseq murin et sur diatomées (en cours). Nous acceptons les librairies prêtes à être séquencées. Les clients doivent aller jusqu'à l'étape de multiplexage des librairies et nous fournir les pools à 2nM.

### Micro-organismes

notre partenaire Genoscreen a développé le nextera dual index sequencing bactérien

### Autre

RADseq ready to load

## Services proposés (domaines d'applications / expertise)

### Reséquençage

- génome entier via Spriworks
- exome via kits SureSelect
- région ciblée via technologie Raindance, Haloplex

### RNAseq

chaque projet est spécifique, et de nombreux kits sont proposés selon les application. Au cas par cas

### SmallRNAseq

en ready to load (librairies préparées par le client)

### ChIPseq

en ready to load (librairies préparées par le client)

### Methyl Seq

oui - capture Agilent 82Mb

### MeDIPseq

oui

### Haloplex

oui

### Autres

radSEQ avec librairies préparées par le client

## Perspectives de développement de services

à moyen terme	miRNA, whole genome PCR free, captures Roche 92Mb ?
à long terme	selon technologies développées par les fournisseurs

## Equipe NGS (effectifs)

ingénieurs	4
assistant ingénieur	1
bioinformaticiens	3
postdoc	1
doctorant	1
labmanager	1

## Equipements (séquenceur, robot, informatique)

1 fragmenteur Covaris E220, 1 fragmenteur Bioruptor, 1 T-Storm : automate de préparation de bibliothèques pour le séquençage ciblé, 1 Spriworks : automate pour la préparation de bibliothèques génomes ou finalisation de bibliothèques cDNA, 1 Bravo : automate pour enrichissement des bibliothèques Sureselect, 1 automate Hamilton pour la préparation de bibliothèques, 2 séquenceurs HiSeq2500, 2 séquenceurs de paillasse MiSeq, 1 séquenceur de paillasse GS junior

## Technologies utilisées

**Fragmentation** Bioruptor, Covaris E220

**Kits de capture** Agilent SureSelect All exon V4, et custom 50Mb, 71Mb  
Spriworks pour la finalisation de bibliothèques à séquencer sur les équipements Illumina et Roche  
Raindance T-Storm - Haloplex

## Format de séquençage

via HiSeq mode high output : PE 2x75, PE 2x100, quelque soit la taille du projet, SR 50, SR 100 pour des flow cell complètes - Pour le mode rapid run et le séquençage via MiSeq : flow cell complétée plus facilement, donc selon les besoins du client / collaborateur.

## Temps d'un run

Le temps de run est proportionnel au type de run choisi et varie selon qu'on est en mode high output ou rapid run (p.ex. : 11 jours en mode high output PE 2x100, 27H en mode rapid run).

## **Multiplexage des échantillons**

Variable selon la couverture choisie. De façon générale compter 30-35Gb / lane en mode high output en PE 2x100, et 25-30Gb / lane en mode rapid run.

nombre d'échantillons/ ligne

exome 50Mb : 24-27 / flow cell pour une couverture de 100x de moyenne

exome 71Mb : 17-19 / flow cell pour une profondeur de 100x de moyenne

## **Traitement, analyse et stockage des données**

1 serveur dédié par Hiseq

Analyse d'exome : en routine via outil d'analyse CASAVA

Analyse séquençage ciblé raindance et haloplex : en routine

## **Développement d'outils de traitement des données**

Pipeline Exome seq développé

Pipeline RNA seq développé

Pipeline pour le Methyl seq développé

Pipeline pour le miRNA seq développé

Pipeline pour le CNV à partir des données exomes prévu

Pipeline pour le CHIP seq prévu

## Exome

### Rendu actuel des résultats

**Caractéristiques des données produites** Données FastQ et analysées selon différents niveaux

#### Qualité des données

couverture	80x-100x
profondeur moyenne de lecture (min-max)	min 8 - max : variable, pas de limite supérieure selon demande client - nos projets standard étant réalisés en "deep sequencing"
projets 'standard'	
projets 'deep sequencing'	pour de l'exome : minimum de 100x de couverture

#### Format de rendu des résultats de l'analyse bioinfo

*Possibilité de récupérer les fichiers .bcl, pour des flow cell complètes, pour les clients autonomes au niveau analyse des data*

##### *Analyse exome niveau 1*

FastQ; ListeSNP brut avec un filtre 8 profondeur; fichier BAM; statistiques; compte rendu

##### *Analyse exome niveau 2*

Fichiers FastQ, BAM et VCF, liste des variants annotés (gene symbol, identifiants Ensembl gène et transcrit, changement d'acide aminé, conséquence de la mutation, identifiant Uniprot, description du gène, identifiant dbSNP, fréquences alléliques dans projets 1000génomés/ESP/Complete Genomics, prédictions fonctionnelles, conservation, expression, interactions, pathways, MIM), statistiques, compte rendu

##### *Analyse exome niveau 3*

SUR DEMANDE (ex: étude comparative entre individus; comparaison avec données de génotypage; études de ségrégation...)

##### *Analyse séquençage ciblé via libraires raindance*

idem analyse exome

reads	FastQ
détection des variants	sous forme de tableau excel
annotation des variants	sous forme de tableau excel

## **Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement**

Selon les volumes de données : mise à disposition des données via téléchargement ftp, ou envoi des data sur disque dur externe. Le client est prévenu de la mise à disposition ou de l'envoi des données par mail

La Plateforme dispose-t-elle d'une base de données de variants? oui, liée aux thématiques de recherche

### **Délai de rendu des résultats**

5 à 6 semaines pour la préparation des librairies, 5 à 6 semaines pour le run en mode high output 2x100. En mode rapid run : 5 à 6 semaines pour la préparation des librairies et 2 à 3 semaines pour le run - pour des flow cell complètes

## Raindance

### Rendu actuel des résultats

#### Caractéristiques des données produites

dépend de la couverture choisie et de la taille de la librairie (pe: run Miseq PE 2x100, couverture 300x librairie 310kb 1400000 reads PF)

#### Qualité des données

couverture	300 - 400x
profondeur moyenne de lecture (min-max)	min 8 - max : variable, pas de limite supérieure
projets 'standard'	selon demande client - nos projets standard étant réalisés en "deep sequencing"
projets 'deep sequencing'	300x

#### Format de rendu des résultats de l'analyse bioinfo

*Possibilité de récupérer les fichiers .bcl, pour des flow cell complètes, pour les clients autonomes au niveau analyse des data*

##### *Analyse exome niveau 1*

FastQ; liste SNP brut avec un filtre 8 profondeur; fichier BAM; statistiques; compte rendu

##### *Analyse raindance niveau 2*

Fichiers FastQ, BAM et VCF, liste des variants annotés (gene symbol, identifiants Ensembl gène et transcrit, changement d'acide aminé, conséquence de la mutation, identifiant Uniprot, description du gène, identifiant dbSNP, fréquences alléliques dans projets 1000génomés/ESP/Complete Genomics, prédictions fonctionnelles, conservation, expression, interactions, pathways, MIM), statistiques, compte rendu

##### *Analyse raindance niveau 3*

SUR DEMANDE (ex: étude comparative entre individus; comparaison avec données de génotypage; études de ségrégation...)

##### *Analyse séquençage ciblé via libraires raindance*

idem analyse exome

reads	FastQ
détection des variants	sous forme de tableau excel
annotation des variants	sous forme de tableau excel

## **Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement**

Selon les volumes de données : mise à disposition des données via téléchargement ftp, ou envoi des data sur disque dur externe. Le client est prévenu de la mise à disposition ou de l'envoi des données par mail.

La Plateforme dispose-t-elle d'une base de données de variants? oui, liée aux thématiques de recherche

## **Délai de rendu des résultats**

5 à 6 semaines pour la préparation des librairies - séquençage via Miseq 2 à 3 semaines pour le run

## **Formation**

possibilité de formation aux préparations de capture Sureselect

## 2. Génomique, transcriptomique

### Principales réalisations

extractions d'ADN génomique / génotypage / puces d'expression

### Services proposés (domaines d'applications / expertise)

extractions d'ADN génomique / génotypage / puces d'expression

### Equipements

Autopure LS, Qiacube

nanosting (formation en cours)

Iscan Biomark

Light cycler 1536 Access array

Viiia7

### Développement d'outils de traitement des données

Analyse biostatistique, selon les projets

*Mise à jour*

*06/10/2014*

## Plateforme biopuces et séquençage

<b>Responsable</b>	Christelle Thibault
<b>Equipe (effectifs)</b>	13
ingénieurs biologistes	4
techniciens biologistes	3
ingénieurs bioinformaticiens	6
<b>Structure de rattachement</b>	IGBMC Strasbourg (UMR 7104 CNRS-Université de Strasbourg-INSERM U964)
<b>Adresse complète</b>	1, rue Laurent Fries 67404 Illkirch Cedex France
<b>Site Web</b>	<a href="http://www-microarrays.u-strasbg.fr/">http://www-microarrays.u-strasbg.fr/</a>
<b>Contact</b>	Bernard Jost (jost@igbmc.fr)
<b>Date de mise en place de la Plateforme</b>	
1. des activités de séquençage à haut débit	2008
2. des activités de génomique - transcriptomique	2000
<b>Labels / démarche qualité</b>	
IBiSA	oui depuis 2006 (rio depuis 2003 et ibisa depuis 2008)
Certification ISO 9001	oui depuis 2007
Autre	Membre du consortium France Génomique, Prestataire pour la Ligue contre le Cancer

## Equipements (séquenceur(s), robot(s), informatique)

<b>Contrôle qualité</b>	Bioanalyzer, Nanodrop, Qubit, Varioskan
<b>Fragmentation</b>	Covaris E210
<b>Automatisation de la préparation des librairies</b>	Biomek FX (Beckman), SPRI-TE (Beckman), Biomek 3000 (Beckman); Mondrian (Nugen)
<b>Séquençage</b>	Illumina Hiseq2500
<b>Puces à ADN</b>	Affymetrix GeneChip technology
<b>Acquisition récente</b>	Biomark HD MX/HX et C1 single-cell auto prep system (Fluidigm)
<b>Informatique</b>	Stockage (RAID) : 97 To utiles, calcul : 192 cœurs

## 1. Séquençage à haut débit

### Principales réalisations

**Nombre de projets traités depuis la création** environ 800

**Nombre d'échantillons traités depuis la création** environ 4000

**Nombre de projets traités par an** > 150

**Nombre d'échantillons traités par an** > 1500

### Domaines de réalisation

**Génétique humaine** oui  
*préciser les spécialités* Neurogénétique, maladies neuromusculaires, maladies rares (BBS, Friedrich...)

**Génétique animale et végétale** oui

**Micro-organismes** non

**Autre (préciser)** Génomique fonctionnelle, biologie du développement, cellules souches, cancer

### Services proposés (domaines d'applications / expertise)

**Séquençage *de novo*** oui

**Métagénomique** non

**Séquençage**  
de génomes entiers (WGS) oui (occasionnellement)  
d'exomes entiers (WES) oui (prestation standard fréquemment réalisée)  
de régions ciblées oui (prestation standard fréquemment réalisée)

**RNAseq** oui (prestation standard fréquemment réalisée)

**SmallRNAseq** oui (prestation standard fréquemment réalisée)

**ChIPseq** oui (prestation standard fréquemment réalisée)

### Perspectives de développement de services

Methyl-seq, single cell transcriptomics, single-cell DNA-seq

## Technologies utilisées

### Kits de capture

Agilent SureSelectXT Human All exon 50Mb, Agilent SureSelectXT Custom (régions ciblées)

### Format de séquençage

single read : de 1x50 à 1x100 en mode standard, jusqu'à 1x150 en mode rapide  
paired-end : de 2x50 à 2x100 en mode standard, jusqu'à 2x150 en mode rapide

### Temps d'un run

De 3 à 12 jours en mode standard, de 24 à 48h en mode rapide

### Multiplexage des échantillons

oui

nombre d'échantillons/ligne

Jusqu'à 24 en routine, possibilité d'aller jusque 96

## Traitement, analyse et stockage des données

Pour toutes les applications :

Démultiplexage, suppression des adaptateurs, contrôle qualité des séquences, mise à disposition des séquences et d'un rapport de contrôle qualité sur le serveur ftp de la plateforme sécurisé par login et mot de passe.

## Développement d'outils de traitement des données

Un pipeline d'analyse a été développé par la plateforme pour chaque application (DNaseq, RNAseq, smallRNAseq et CHIPseq), ainsi que le logiciel SeqMINER (Comparaison de jeux de données de CHIPseq)

## DNaseq

### Rendu actuel des résultats

#### Caractéristiques des données produites

>75 millions de séquences minimum et en moyenne 150 millions de séquences par ligne (le nombre de séquences par échantillon dépend du nombre d'échantillons que le porteur de projet souhaite multiplexer par ligne).

#### Qualité des données

couverture	de 50X (all exon avec 4 échantillons par piste) à 100X (all exon avec 2 échantillons par piste)
moyenne des scores de qualité (Phred score)	de 50X à 1000X (régions ciblées) ≥ 30 pour 95% des positions

#### Format de rendu des résultats de l'analyse bioinfo

reads	Format fastq
alignements	Format BAM
détection des variants	Format VCF
annotation des variants	Fichier tsv

#### Outils de prédiction de la pathogénicité des variants

outils existants	Polyphen2, SIFT
développement d'outils propres	Varrank

#### Comparaison de variants entre plusieurs individus

En collaboration avec la plateforme

#### Détection des remaniements chromosomiques

En collaboration avec la plateforme

### Délai de rendu des résultats

2 mois pour les données brutes + 1 mois supplémentaire pour l'analyse des données (ces délais peuvent varier en fonction de la taille et de la nature du projet)

#### Accès

Equipes du site	Oui
Equipes extérieures académiques	Oui
Equipes extérieures privées	Oui

#### Formation

Formation individuelle des porteurs de projets et formation dans le cadre du DU "DU Séquençage Haut Débit et maladies génétiques" (Dijon)

Visualisation des données, priorisation des variants, interprétation des résultats, comparaison d'échantillons, détection des remaniements chromosomiques.

## ChIPseq

### Rendu actuel des résultats

**Caractéristiques des données produites** >25 millions de séquences

### Qualité des données

Moyenne des scores de qualité (Phred score)  $\geq 30$  pour 95% des positions

### Format de rendu des résultats de l'analyse bioinfo

reads Format fastq  
alignements Formats BAM et BED

En collaboration avec la plateforme :

recherche de pics, localisation des pics par rapport à des annotations fonctionnelles du génome de référence, recherche de motifs significativement retrouvés dans les pics, analyse de colocation des pics entre plusieurs jeux de données (clustering)

### Délai de rendu des résultats

2 mois pour les données brutes + 1 mois supplémentaire pour l'analyse des données (ces délais peuvent varier en fonction de la taille et de la nature du projet)

### Accès

Equipes du site Oui  
Equipes extérieures académiques Oui  
Equipes extérieures privées Oui

### Formation

Formation individuelle des porteurs de projets et formation CNRS "Introduction à l'analyse bioinformatique de données de séquençage haut débit" (2 sessions / an)

Visualisation et interprétation des données, recherche et annotations des pics, recherche de motifs, comparaison de plusieurs jeux de données.

## **RNAseq**

### **Rendu actuel des résultats**

**Caractéristiques des données produites** >25 millions de séquences

### **Qualité des données**

Moyenne des scores de qualité (Phred score)  $\geq 30$  pour 95% des positions

### **Format de rendu des résultats de l'analyse bioinfo**

reads Format fastq  
alignements Formats BAM

En collaboration avec la plateforme :

alignement, quantification, normalisation, analyses statistiques, analyses fonctionnelles

### **Délai de rendu des résultats**

2 mois pour les données brutes + 1 mois supplémentaire pour l'analyse des données (ces délais peuvent varier en fonction de la taille et de la nature du projet)

### **Accès**

Equipes du site Oui  
Equipes extérieures académiques Oui  
Equipes extérieures privées Oui

### **Formation**

Formation individuelle des porteurs de projets et formation CNRS "Introduction à l'analyse bioinformatique de données de séquençage haut débit" (2 sessions / an)

Visualisation et interprétation des données, analyses statistiques, analyses fonctionnelles, comparaison de différents jeux de données

## smallRNAseq

### Rendu actuel des résultats

**Caractéristiques des données produites** >5 millions de séquences

### Qualité des données

moyenne des scores de qualité (Phred score)  $\geq 30$  pour 95% des positions

### Format de rendu des résultats de l'analyse bioinfo

reads Format fastq  
alignements Formats BAM

En collaboration avec la plateforme : suppression des adaptateurs, alignement, quantification des petits ARN connus, normalisation, analyses statistiques, prédiction de nouveaux miRNA.

### Délai de rendu des résultats

2 mois pour les données brutes + 1 mois supplémentaire pour l'analyse des données (ces délais peuvent varier en fonction de la taille et de la nature du projet)

### Accès

Equipes du site	Oui
Equipes extérieures académiques	Oui
Equipes extérieures privées	Oui

### Formation

Formation individuelle des porteurs de projets

Visualisation et interprétation des données, prédiction de nouveaux miRNA, comparaison de différents jeux de données, analyses statistiques

## 2. Génomique, transcriptomique (puces à ADN uniquement)

### Principales réalisations

Nombre de projets traités depuis la création	Plus de 15000 échantillons traités
Nombre de projets traités par an	Environ 100
Domaines de réalisation	Transcriptomique, génotypage de SNP et CGH

### Services proposés (domaines d'applications / expertise)

Analyse du transcriptome sur puces Affymetrix et Agilent  
Génotypage de SNP sur puces Affymetrix

### Perspectives de développement de services

Technologie Fluidigm

### Technologies utilisées

Affymetrix

### Traitement, analyse et stockage des données

Contrôles qualité, normalisation, analyses statistiques des données, analyses fonctionnelles

### Développement d'outils de traitement des données

Flora (clustering flou), Zoe (outil statistique de sélection de gènes différentiellement exprimés)

### Rendu actuel des résultats

Données brutes et fichiers issus de l'analyse des données

### Délai de rendu des résultats

4 semaines à partir de la validation des échantillons de départ

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

*Mise à jour* 06/10/2014

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme Genome et Transcriptome Genotoul

<b>Responsable</b>	Denis Milan
<b>Equipe (effectifs)</b>	19
bioinformaticiens	4
administratifs	1
autres (préciser)	3 Chercheurs - 1 Directeur de recherche - Biostatisticien
<b>Structure de rattachement</b>	INRA
<b>Adresse complète</b>	Chemin de Borde rouge 31326 Castanet tolosan
<b>Site Web</b>	<a href="http://get.genotoul.fr/">http://get.genotoul.fr/</a>
<b>Contact</b>	get@gentoul.fr
<b>Date de mise en place de la Plateforme</b>	2001
1. des activités de séquençage à haut débit	2008
2. des activités de génomique - transcriptomique	2001
<b>Labels / démarche qualité</b>	
IBiSA	depuis 2008
Certification ISO 9001	depuis 2008
Autre	Investissement d'avenir - Programme France Génomique

## 1. Séquençage à haut débit

### Principales réalisations

<b>Nombre de projets traités depuis la création</b>	168 projets depuis 2008
<b>Nombre d'échantillons traités depuis la création</b>	plusieurs milliers d'échantillons
<b>Nombre de projets traités par an</b>	44 projets en 2012
<b>Nombre d'échantillons traités par an</b>	682 libraires analysées en 2012 plrs échantillons traités dans 1 librairie

### Domaines de réalisation

<b>Génétique humaine</b> <i>préciser les spécialités</i>	Environ 30% des projets sont du domaine de la santé maladies métaboliques, cancerologie
<b>Génétique animale et végétale</b>	Environ 50% des projets sont du domaine agricole
<b>Microorganismes</b>	Environ 20% des projets sont sur des micro-organismes ou autres

### Services proposés (domaines d'applications / expertise)

<b>Séquençage <i>de novo</i></b>	oui
<b>Métagénomique</b>	oui
<b>Séquençage</b>	
de génomes entiers (WGS)	oui
d'exomes entiers (WES)	oui
de régions ciblées	oui
<b>RNAseq</b>	oui
<b>SmallRNAseq</b>	oui
<b>MeDIPseq</b>	oui
<b>Autre (préciser)</b>	RADseq, Bisulfite, RNAseq et smallRNAseq directionnel ...

## Perspectives de développement de services

Séquençage NGS 3G, droplet PCR (raindance), Pyroséquençage

## Equipements (séquenceur(s), robot(s), informatique)

HISEQ 2500 n°1  
HISEQ 2000 n°2  
MISEQ (Illumina)  
Ion-Proton  
Robot genesis 150/8 (TECAN)  
Robot genesis 200/8 (TECAN)  
Robot Evo200 (TECAN)

La plateforme a établi un partenariat étroit avec la plateforme Bioinformatique de Toulouse. Le stockage, et la sauvegarde des données sont assurés par celle-ci. Elle est équipée de 2000 cpu, 170tb d'espace d'analyse, 200Tb d'espace de stockage.

## Technologies utilisées

<b>Format de séquençage</b>	2*100pb, 2*250pb, 1*500pb, 1*150pb
<b>Temps d'un run</b>	dépendant des machines - 3j pour un run HISEQ2500, 24h pour un run Ion Proton
<b>Multiplexage des échantillons</b>	oui
nombre d'échantillons/ligne	dépend des projets de recherche

## Traitement, analyse et stockage des données

Livraison des séquences brutes et des qualités associées (accessibles via ng6, pendant 2 ans minimum sur un espace sauvegardé), rapport de séquençage et contrôle qualité, recherche de contaminants, alignement des lectures contre une séquence de référence.

## Développement d'outils de traitement des données

Développement d'une interface de visualisation et de sa base de données (NG6) ainsi que d'un gestionnaire de pipeline d'analyse basé sur make flow.

En prestation de service, GeT peut aussi réaliser vos analyses de données de séquençage (qualité, mapping des séquences sur génome de référence, de reconstruction de transcrits, quantification de l'expression des gènes ou transcrits, recherche de variants). Pour ce type de prestation, contacter Delphine Laboudette de GeT-Biopuces ([labourdette@insa-toulouse.fr](mailto:labourdette@insa-toulouse.fr)).

## Rendu actuel des résultats

### Caractéristiques des données produites

350 millions de séquences produites par lane sur HISEQ 2000, 80 millions de séquences produites par puce PI sur Ion Proton.

### Qualité des données

couverture dépendante du génome, multiplexage...  
profondeur moyenne de lecture (min-max) dépendante du génome, multiplexage...

### Format de rendu des résultats de l'analyse bioinfo

reads fastQ  
alignements BAM  
détection des variants vcf  
annotation des variants non  
outils de prédiction de la pathogénicité des variants non

### Outils/portail d'analyse hébergés par la plateforme et mis à disposition du porteur via une interface web, accompagnement

voir avec la plateforme Bioinformatique de Gentoul : <http://bioinfo.genotoul.fr/>

La plateforme dispose-t-elle d'une base de données interne de variants? non

## Délai de rendu des résultats

Dépendant du projet

### Accès

Equipes du site oui  
Equipes extérieures académiques oui  
Equipes extérieures privées oui

### Formation

#### Existence d'un programme de formation (quel format ?)

Les équipes utilisatrices sont invitées à participer aux premières manipulations, elles sont formées à la préparation des librairies. Elles peuvent par la suite accéder en autonomie à certaines machines NGS. Formation à l'analyse de données RNAseq en collaboration avec la plateforme Bioinformatique de la Génopole de Toulouse (alignement épissé, découverte de nouveaux transcrits, quantification de l'expression des gènes et des transcrits).

## 2. Génomique, transcriptomique

### Principales réalisations

#### Nombre de projets traités depuis la création

Plus d'un millier de projets ont été traités depuis la création de la plateforme

#### Nombre de projets traités par an

460 projets en 2012

#### Domaines de réalisation

30% des projets sont du domaine de la santé  
50% des projets sont du domaine agronomique  
20% des projets sont sur des micro-organismes ou autres

### Services proposés (domaines d'applications / expertise)

Séquençage sanger

Analyse de polymorphisme (puces ou autres)

Analyse sur cellule unique

Analyse transcriptomique sur microarrays ou PCR quantitative

### Perspectives de développement de services

Analyse de méthylation

Quantification allèle spécifique

### Equipements

BeadXpress

BioMark

MISEQ (Illumina) - Ion-Proton

Séquenceur 4 Capillaires (ABI) EDB

Séquenceur 48 Capillaires (ABI) 3730

7900 HT (ABI)

LC480 (Roche)

C1

Séquenceur 16 Capillaires (ABI)

2 Stepone

ABI 7500

ViiA 7

Experion

Nanodrop 2000

PCR Veriti

Robot de dépôt Q Array Mini

Chambre d'hybridation automatique

SPRi Plex

Station Affymetrix

Innoscan 900

Scanner MS200

Scanner de microarrays G2505C

Four d'hybridation

Automate de distribution Bravo

Appareil qPCR ViiA7

Multi NA

BioSpec Nano

## Technologies utilisées

Genotypage, séquençage, PCR quantitative nanovolume, microarrays, isolation de cellules, ...

## Traitement, analyse et stockage des données

Livraison des données brutes (accessibles via esitoul, pendant 2 ans minimum sur un espace sauvegardé)

## Développement d'outils de traitement des données

Développement d'un système d'information de la gestion et le stockage des échantillons et des données : esitoul

Pour les microarrays: analyse des données avec le logiciel R ainsi que les packages Bioconductor appropriés à chaque type de puces.

## Rendu actuel des résultats

En fonction du type de projet données brutes ou analysées. Compte rendu de résultats.

## Délai de rendu des résultats

Dépendant du projet

## Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

*Mise à jour* 06/09/2013

## Fiche d'identité – Plateforme de Séquençage Haut Débit, Génomique, Transcriptomique

### Plateforme Génome Transcriptome de Bordeaux - PGTB

#### Responsable

Co-Direction : Patricia FERGELOT / Pascal SIRAND-PUGNET

#### Equipe (effectifs)

ingénieurs	4
techniciens	5
direction	2
administratifs	1

#### Structure de rattachement

Université de Bordeaux / INRA - (Plateforme intégrée à la SFR infrastructure CGFB\*)  
\*Centre de Génomique Fonctionnelle de Bordeaux

#### Adresse complète

Plateforme Génome Transcriptome de Bordeaux  
Centre de Génomique Fonctionnelle -  
Université de Bordeaux  
146 rue Léo Saignat - 33076 Bordeaux Cedex

#### Site Web

<http://www.pgtb.u-bordeaux2.fr>

#### Contact

[christophe.hubert@u-bordeaux.fr](mailto:christophe.hubert@u-bordeaux.fr),  
[PGTP@pierroton.inra.fr](mailto:PGTP@pierroton.inra.fr)

#### Date de mise en place de la Plateforme

1. des activités de séquençage à haut débit	2009
2. des activités de génomique - transcriptomique	2002

#### Labels / démarche qualité

IBiSA	Plateforme Stratégique IBISA
Certification ISO 9001	Certification ISO 9001: 2008
Autre	CNOC (INRA) : identifiée comme Plateforme Stratégique Régionale

## 1. Séquençage à haut débit

### Principales réalisations

#### Nombre de projets traités depuis la création >100

La plateforme offre des prestations de service et d'accompagnement scientifique pour des projets principalement régionaux et académiques, mais aussi nationaux, internationaux et privés. Elle participe ainsi aux projets fondamentaux et appliqués d'une large communauté scientifique travaillant dans les domaines médicaux et agro-industriel. Elle est reconnue comme plateforme de référence en génomique dans le cadre du projet européen d'excellence Evoltree et dans la partie transnational accès du projet européen T4F. Elle met à disposition des utilisateurs les techniques de séquençage haut débit disponibles sur les séquenceurs:

- Séquençage de novo
- Transcriptome
- Séquençage d'exome humain
- Séquençage ciblé (ADN, ARN) notamment par capture sélective de régions génomiques (technique SureSelect targetting enrichment).
- Capture sélective d'ADN mitochondrial humain
- Métagénomique 16S humaine et végétale
- Séquençage d'amplicons humains et non humains

#### Nombre d'échantillons traités depuis la création >1000

Depuis sa création, notre structure a analysé des échantillons émanant de tous types d'organismes : bactéries, virus, levures, champignons, eucaryotes animal ou végétal. Cependant la plus forte augmentation ces deux dernières années s'est faite dans le domaine de l'humain avec l'explosion des approches de capture, amplicon et exome. Dans le domaine de l'humain, la plateforme a réalisé plusieurs centaines de bibliothèques, notamment dans le cadre de projets de :

- Capture sélective d'ADN mitochondrial
- Capture exonique
- Métagénomique 16S
- Capture sélective de régions génomiques
- Séquençage génome entier sur prélèvement tumoral
- Séquençage d'amplicons

#### Nombre de projets traités par an 30-60

Le nombre de projets accepté varie en fonction de la taille des projets. Certains ne comportent qu'un run machine alors que d'autres peuvent monopoliser nos séquenceurs pendant plusieurs semaines

#### Nombre d'échantillons traités par an >1000

Les échantillons traités sont d'origines assez variables. On peut cependant observer depuis 2011/2012 une proportion croissante d'échantillons d'origine humaine. Cette proportion devrait encore s'accroître à l'avenir.

## Domaines de réalisation

### Génétique humaine

Maladies rares (Laboratoire MRGM) : séquençage + analyse en NGS  
Oncogénétique (Institut Bergonié) : séquençage NGS de région spécifique  
Pharmaco-génétique  
Approche Ampliseq via des cancer panel ou custom panel  
Sequençage d'ADN circulant ccfDNA  
Sequençage ciblé de genes d'intérêts (capture ou ampliseq)

### Génétique animale et végétale

Analyse de populations animales , RNA-Seq végétaux  
Séquençage d'Amplicon pour identification d'espèces, approches de GBS par capture ou amplicons ou ampliseq

### Microorganismes

séquençage de novo de bactéries et levures  
séquençage de population de virus  
métagénomique humaine / Végétale

### Autres

Paléogénétique humaine et méta-paléogénétique.

## Services proposés (domaines d'applications / expertise)

### Séquençage *de novo*

oui (selon taille génome)

### Métagénomique

oui 16S,18S,marqueurs spécifiques comme Rbcl,ITS,Cox ...

### Séquençage

de génomes entiers (WGS)  
d'exomes entiers (WES)  
de régions ciblées

oui (selon taille génome)  
oui (ampliseq)  
par capture ou ampliseq

### RNAseq

oui

### SmallRNAseq

possibilité

## Perspectives de développement de services

Développement "long read" moleculo d'illumina, étude du methylome et de sites de methylation de l'ADN

## Equipements (séquenceur(s), robot(s), informatique)

1 MiSeq (Illumina)	séquenceur NGS
1 Proton de Life technologies	séquenceur NGS
1 PGM de life Technologies	séquenceur NGS
1 GS Junior de Roche	séquenceur NGS
1 serveur Galaxy dédié a l'analyse des données	informatique
suite de logiciels commerciaux (CLC...)	informatique
stockage de 80 teraOctets	informatique
2 robots Hamilton	robotique

## Technologies utilisées

### Kits de capture

Kit Agilent SureSelect V5 avec et sans UTR,  
Kit Agilent SureSelect Custom, mybaits ou  
autre type de capture

### Format de séquençage

**MiSeq:** Paired end 2 x75b, paired end 2 x  
100b, Paired end 2 x 150b et 2 x 250b, 2 x  
300pb

**PGM:** single read en 200 ou 400 bp (6  
millions de reads/run)

**PROTON:** single reads 200 bp (85 000 000  
de reads/run)

### Temps d'un run

Variable selon longueur de lecture et du type  
de densité de lecture.

**MiSeq:** 12 à 36h

**PGM:** 8h max

**PROTON:** 2h30

### Multiplexage des échantillons

nombre d'échantillons/ligne

oui

Jusqu'à 96 par manipulation

## Traitement, analyse et stockage des données

La plateforme possède ses propres structures de stockage et d'analyse de données. Selon les projets, la plateforme travaille en association avec le CBIB (plateforme de bioinformatique du CGFB) qui gère alors le traitement et l'analyse des données. Leur stockage n'est géré que le temps de l'analyse et de la récupération par l'utilisateur.

## Rendu actuel des résultats

### Caractéristiques des données produites

Toutes les caractéristiques des données produites sont accessibles dans notre catalogue de prestation a cette adresse :

[http://www.pgtb.ubordeaux2.fr/content/download/3212/34244/version/1/file/CataloguePrestationsServices\\_v27062014.pdf](http://www.pgtb.ubordeaux2.fr/content/download/3212/34244/version/1/file/CataloguePrestationsServices_v27062014.pdf).

L'hétérogénéité des projets traités par la plateforme ne permet pas de détailler ici le nombre de reads par échantillon. Exemple pour le MiSeq: pourcentage de Q30 doit être supérieur à 80% pour des reads de 100pb, rendu de 1 à 25 millions de reads par run) avec une quantité de base de 0,3Gb à 15Gb par séquençage, données rendues sur serveur FTP sécurisé ou par disque dur externe transmit par transporteur, rapport de qualité établit à la génération des séquences résumant le nombre de reads, le Q30, la répartition d'index (dans le cadre d'un multiplexage), le % d'alignement, le % GC, le phasing/préPhasing. L'ensemble est validé par un standard (PhiX) contenu dans la flowcell (% d'alignement et taux de clusterisation standard).

### Qualité des données

couverture	Au choix de l'utilisateur
profondeur moyenne de lecture (min-max)	1x à 10000x
projets 'standard'	30x à 100x
projets 'deep sequencing'	10000x

### Format de rendu des résultats de l'analyse bioinfo

reads	fastq, fasta, fastq.gz
alignements	.bam - .sam
détection des variants	.vcf
annotation des variants	en fonctions des logiciels utilisés (bed,gff)

La Plateforme dispose-t-elle d'une base de données interne de variants? non

## Délai de rendu des résultats

Entre 2 semaines et 4 mois en fonction de la taille des projets

### Accès

Equipes du site	oui
Equipes extérieures académiques	oui
Equipes extérieures privées	oui

### Formation

Possibilité de formation théorique et pratique sur les technologies NGS, dans le cadre de projets ou de transferts de technologie

## 2. Génomique, transcriptomique

### Principales réalisations

**Nombre de projets traités depuis la création** >1000

**Nombre de projets traités par an** >50

**Domaines de réalisation** genotypage SSR et SNP par différentes approches technologiques

### Services proposés (domaines d'applications / expertise)

genotypage en multiplex de SNP par sequenom

genotypage de Microsatellites

génotypage par GBS

### Equipements

sequenom (40 SNP/384 echantillons)

sequenceurs NGS

genotypage SNP

genotypage GBS

### Technologies utilisées

technologie sequenom

technologie GBS

technologie SSR

massArray

capture , amplicon, ampliseq

marquage fluo des microsatellites

### Traitement, analyse et stockage des données

La plateforme met à disposition tous les logiciels d'analyses et de traitements et met à disposition ses capacités de stockage

### Rendu actuel des résultats

Toutes les caractéristiques des données produites sont accessibles dans notre catalogue de prestation à cette adresse :

[http://www.pgtb.ubordeaux2.fr/content/download/3212/34244/version/1/file/CataloguePrestationsServices\\_v27062014.pdf](http://www.pgtb.ubordeaux2.fr/content/download/3212/34244/version/1/file/CataloguePrestationsServices_v27062014.pdf)

### Délai de rendu des résultats

entre 2 semaines et deux mois selon la taille des projets

### Accès

Equipes du site oui

Equipes extérieures académiques oui

Equipes extérieures privées oui

*Mise à jour* 26/11/2014