



INRAE

Matinée des utilisateurs

Marie Gislard

Le 2 Octobre 2020

<http://get.genotoul.fr>

get@genotoul.fr

 [@get_genotoul](https://twitter.com/get_genotoul)



Matinée des utilisateurs

Retour de Sondage



Merci à tous pour votre participation au sondage, voici quelques résultats :

95% satisfaits de l'accueil sur la plateforme

91% satisfaits de l'équipe GeT-PlaGe qui a su répondre à vos demandes

50% souhaitent avoir plus d'interactions avec l'équipe GeT-PlaGe

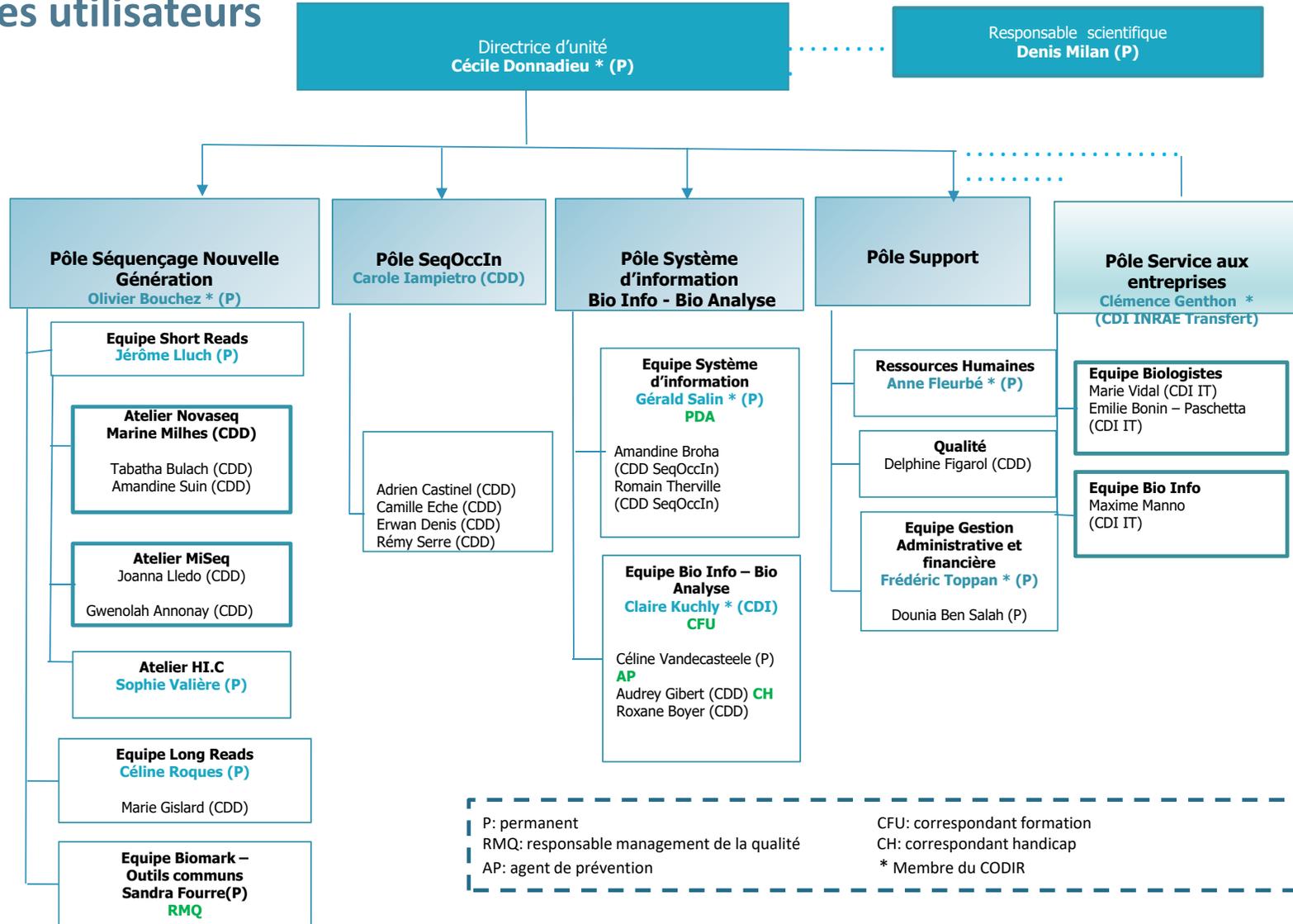
90% satisfaits de la formation proposée par l'équipe GeT-PlaGe

75% satisfaits de l'infrastructure et des outils proposés

Ce sondage est pour nous l'occasion de jauger votre niveau de satisfaction générale, mais surtout d'évaluer vos besoins. Nous tachons de prendre en compte toutes vos remarques et d'y répondre si possible.

Matinée des utilisateurs

Organigramme



P: permanent
 RMQ: responsable management de la qualité
 AP: agent de prévention
 CFU: correspondant formation
 CH: correspondant handicap
 * Membre du CODIR



Matinée des utilisateurs

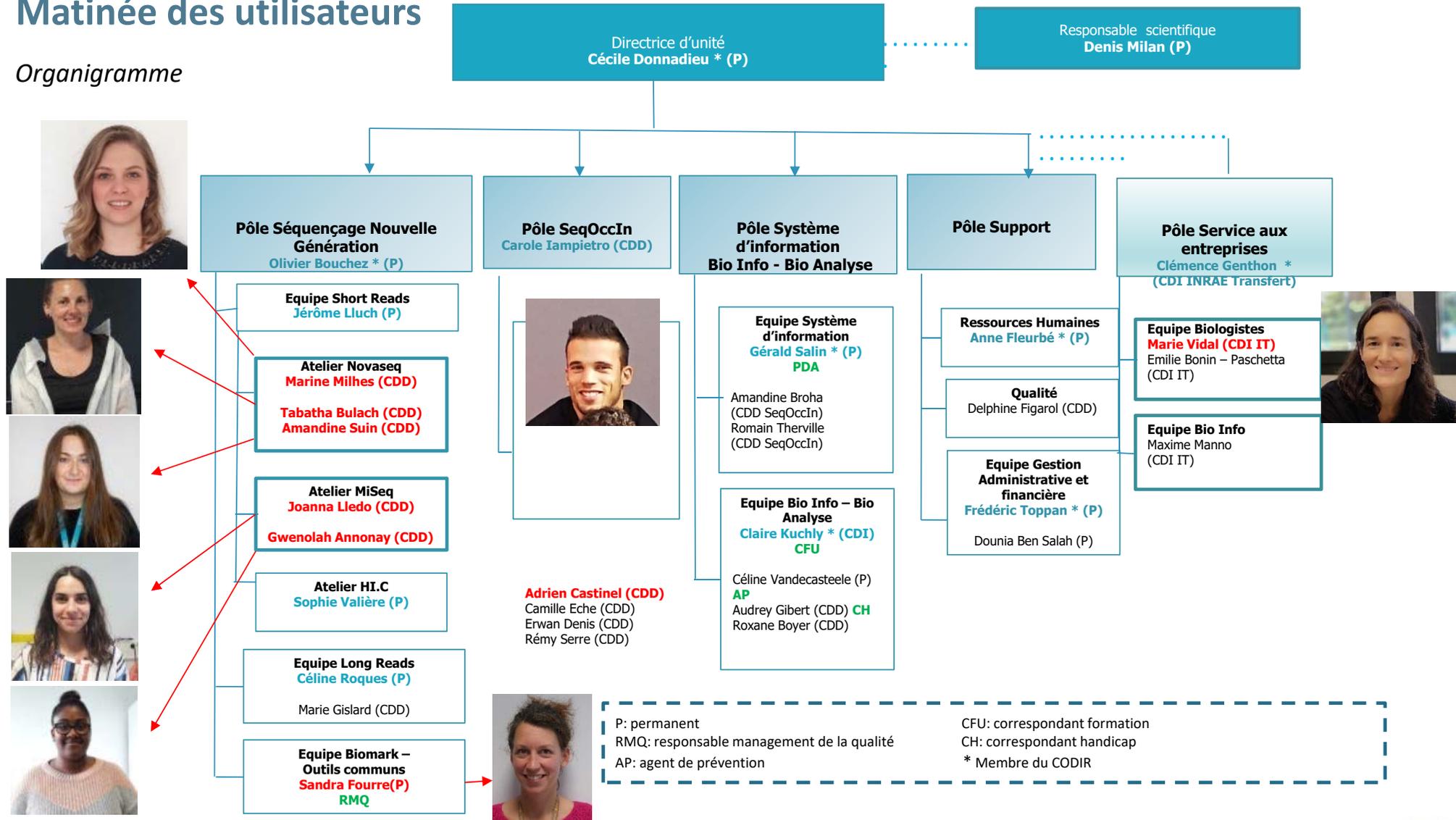
Les Responsables Techniques

Noms des Machines	Noms et mails des Responsables	Noms et mails des Seconds Responsables
ABI3730*	Sandra Fourre sandra.fourre@inrae.fr	
Biomark*	Sandra Fourre sandra.fourre@inrae.fr	
QuantStudio6	Sandra Fourre sandra.fourre@inrae.fr	Amandine Suin amandine.suin@inrae.fr
Nanodrop8000	Sandra Fourre sandra.fourre@inrae.fr	Amandine Suin amandine.suin@inrae.fr
Tecan EVO	Marie Vidal marie.vidal@inrae.fr	Adrien Castinel adrien.castinel@inrae.fr
Ready-To-Load MiSeq	Joanna Lledo joanna.lledo@inrae.fr	Gwenolah Annonay gwenolah.annonay@inrae.fr
Préparation librairies Illumina et séquençage sur MiSeq	Joanna Lledo joanna.lledo@inrae.fr	Gwenolah Annonay gwenolah.annonay@inrae.fr
Ready-To-Load NovaSeq	Tabatha Bulach tabatha.bulach@inrae.fr	Marine Milhes marine.milhes@inrae.fr

Pour tous les Responsables Techniques qui souhaitent des rappels ou compléments de formations, **nous sommes disponibles au cas par cas**.
N'hésitez pas à nous envoyer un mail, au responsable de la machine concernée ou à get-plage.contact@genotoul.fr pour convenir d'un créneau.

Matinée des utilisateurs

Organigramme



Matinée des utilisateurs

Les petits nouveaux

- **Bienvenue au PIXUL à l'essai sur GeT-PlaGe en octobre pour remplacer le Covaris**

Le PIXUL de Active Motif est le premier et le seul sonicateur multi-échantillons jusqu'à 96 échantillons traités en parallèle

Compatible avec la chromatine, l'ADN, l'ARN et les protéines provenant d'échantillons, de cellules ou de tissus



- **L'arrivée du Sequel II de Pacific Biosciences**

Remplace le PacBio RS II

Formation méthodologique et technique faite par le fournisseur en Mars 2020

Machine validée en juillet 2020



Matinée des utilisateurs

Les nouveautés de la plateforme

- Les gants

ecoSHIELD* Eco Nitrile PF 250



- Category III PPE glove (PPE Regulation (EU) 2016/425)
- Complex Design - For mortal and irreversible risks
- Class 1 MDD glove (Council Directive 93/42/EEC)
- Powder-free green/white nitrile glove
- twinSHIELD* double-walled protection
- Ambidextrous
- 250 mm / 0.10 mm (XL/10: 260 mm) (EN 420:2003+A1:2009)
- Biological risk (ISO 374-5:2016 VIRUS)
- AQL 0.25 (EN 374-2:2014 Level 3)
- Viral penetration test (ISO 16604:2004 Procedure B)
- Chemical risk (ISO 374-1:2016+A1:2018 - Type B JKPT)
- Waterproof and for low chemical protection
- Tested for chemical permeation (EN 16523-1:2015+A1:2018)

Des gants « verts » : équivalents des bleus existants sur la plateforme :

- Non stériles
- Non DNaseFree
- Non RNaseFree

SHIELDskin* ORANGE NITRILE* 300 Sterile



- Category III PPE glove (PPE Regulation (EU) 2016/425)
- Complex Design - For mortal and irreversible risks
- Sterile powder-free orange/white nitrile/polychloroprene glove
- twinSHIELD* double-walled protection
- Ambidextrous
- 300 mm / 0.14 mm (EN 420:2003+A1:2009)
- Biological risk (ISO 374-1:2016 VIRUS)
- AQL 0.25 (EN 374-2:2014 Level 3)
- Viral penetration test (ISO 16604:2004 Procedure B)
- Chemical risk (ISO 374-1:2016+A1:2018 - Type B JKPT)
- Waterproof and for low chemical protection
- Tested for chemical permeation (EN 16523-1:2015+A1:2018)
- RNase and DNase free

Des gants « orange » :

- Stériles
- DNaseFree
- RNaseFree

Matinée des utilisateurs

- La boîte à suggestion

Afin de conserver cette proximité tout au long de l'année, nous vous proposons « **une boîte à suggestion** » qui sera disponible à l'entrée de la plateforme, vous pourrez si besoin nous faire part, anonymement ou pas, de vos remarques, questions...



- Les blouses

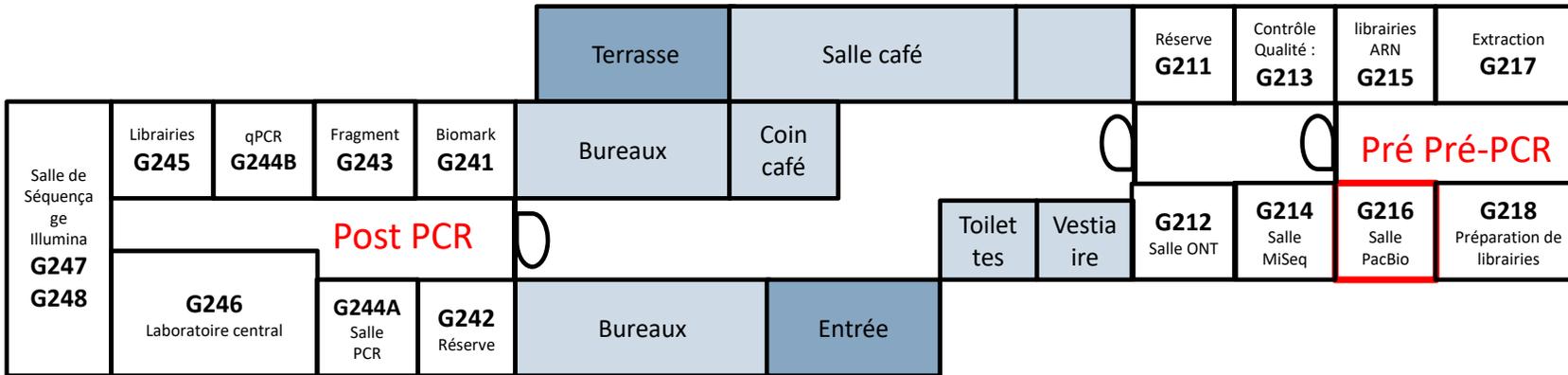
Les utilisateurs peuvent amener leur blouse et la laisser dans le casier

Si l'activité sur la plateforme est de plusieurs jours, il est possible d'emprunter une blouse en tissu en prenant soin de l'identifier par un badge puis de la mettre dans la pаниère à linge sale à la fin de la semaine

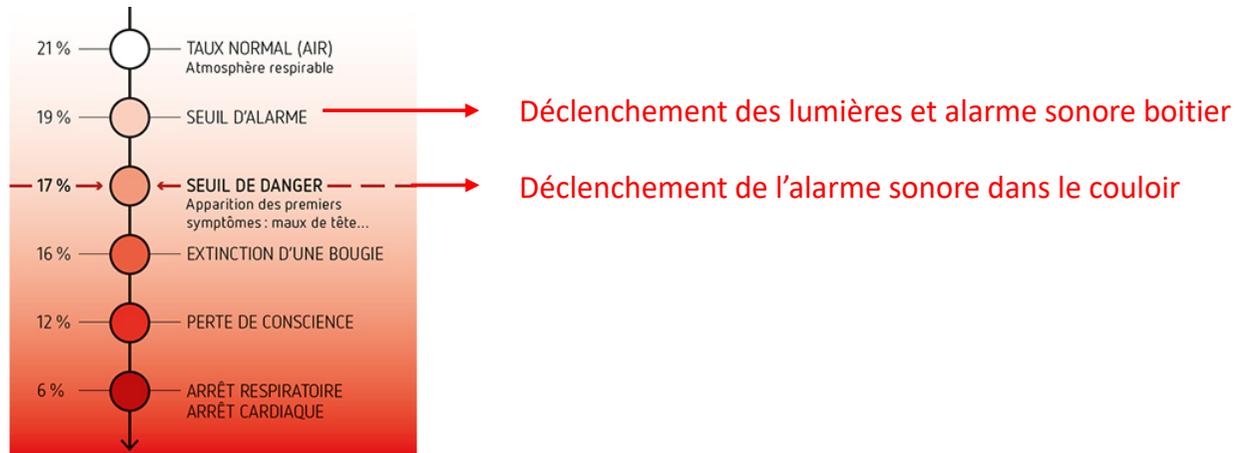


Matinée des utilisateurs

- Le système de sécurité circuit Azote et détection en salle PacBio G216



La porte de cette salle PacBio G216 doit toujours être tenue fermée sauf si un personnel GeT-PlaGe manipule à l'intérieur.
Un système d'alarme se déclenche suivant le taux d'oxygène de la pièce :



Matinée des utilisateurs

- Ménage du Y

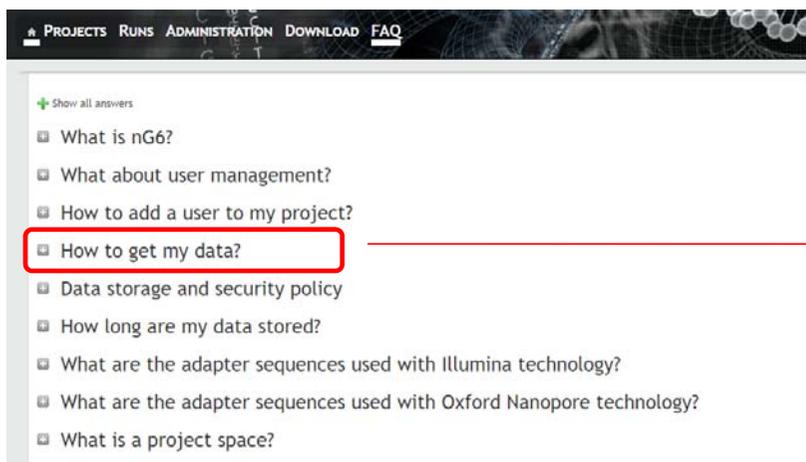
N'oubliez pas de récupérer et effacer vos fichiers temporaires qui sont stockés sur le Y.

Attention, nous faisons le ménage régulièrement...

- Récupération des données sur NG6

FAQ NG6 : Les bioinformaticiens ont mis en place une FAQ et des vidéos explicatives sur le site NG6 pour vous aider à récupérer vos données.

<https://ng6.toulouse.inra.fr/index.php?id=57>



PROJECTS RUNS ADMINISTRATION DOWNLOAD FAQ

Show all answers

- What is nG6?
- What about user management?
- How to add a user to my project?
- How to get my data?**
- Data storage and security policy
- How long are my data stored?
- What are the adapter sequences used with Illumina technology?
- What are the adapter sequences used with Oxford Nanopore technology?
- What is a project space?

A tutorial video for download by URL on Linux system is available here

- On a Windows workstation (Command lines), copy/paste this list in a file "listOfFiles" (Be aware that if the same filename is found in different url, the local file will be overridden). Then you can download curl for Windows (<https://curl.haxx.se/download.html>), extract the archive, and run this following command in a windows prompt

```
for /f %d in (C:\Path\to\listOfFiles.txt) do C:\Path\to\bin\curl.exe --output %~nd.gz %d
```

Symbolic Link :

A tutorial video is also available here

a dialog box asks you

- a login/password couple existing in the Bioinformatics Genotoul platform
- a repository where the symbolic links will be created on the genotoul server. It should be :
 - on the /work space
 - on a folder where YOU can access in writing mode, eg. /work/user/projectX/

- Site Web

Le site web est régulièrement actualisé, **les différentes applications** de GeT sont disponibles sur le site :

<https://get.genotoul.fr/la-plateforme/get-plage>

Matinée des utilisateurs

SERVCOM QC Fragment Analyzer 50Kb

75% satisfaits de l'infrastructure et des outils proposés mais une forte demande de QC Fragment Analyzer pour des échantillons hors projet de séquençage

Mise en place d'un service SERVCOM Fragment pour le contrôle qualité des échantillons hors projet de séquençage (pour l'instant sur l'application 50Kb) en test sur 2 équipes -> si ok, ouverture !

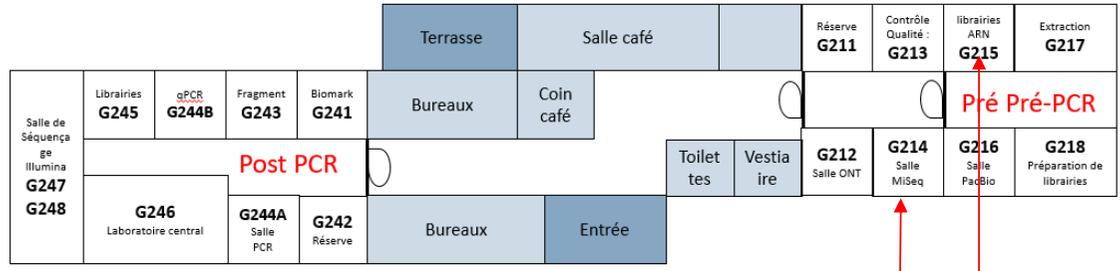
Modalités définies :

- Formation au dépôt des échantillons sur une plaque (1 à 2 personnes par laboratoire)
- Mise en place d'un fichier de suivi pour la réservation de place sur 1 ligne FA
- Lancement des runs le vendredi après-midi
- Facturation à l'échantillon

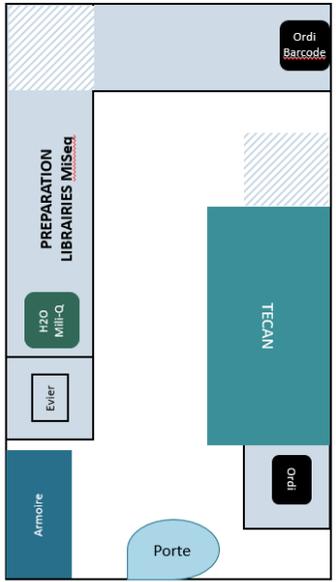


Matinée des utilisateurs

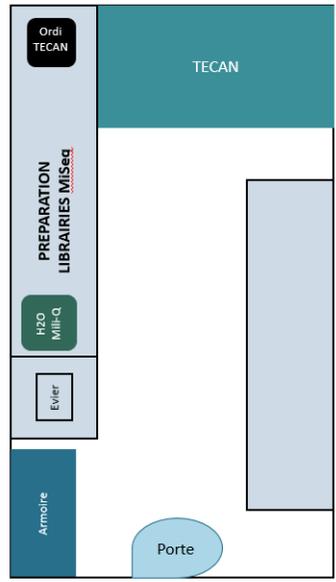
- Réorganisation de la plateforme



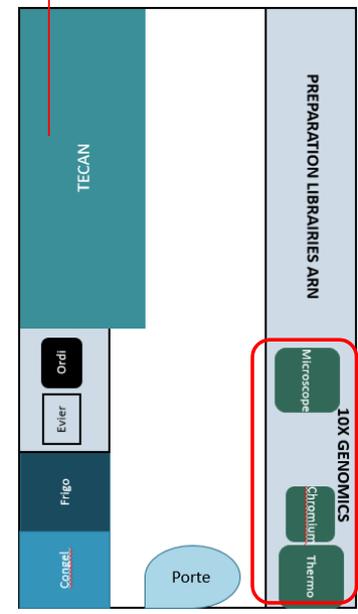
AVANT



APRES



+ 2 places



+ 1 place
après le
départ du
10X



Matinée des utilisateurs

COVID-19

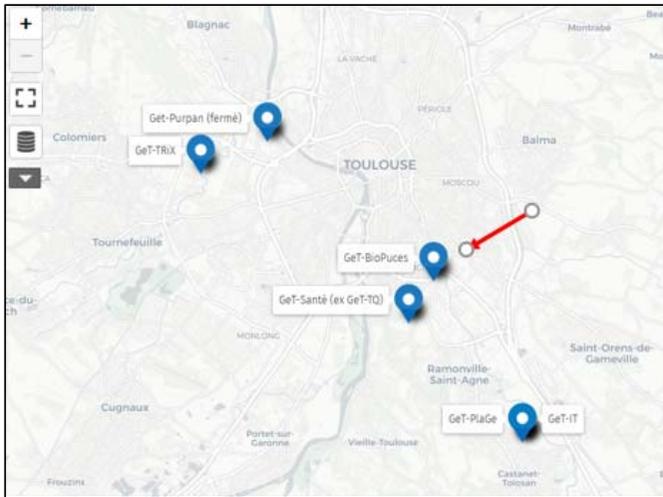
- ⌘ Le nombre d'extérieurs est limité à 3 par jour sur la plateforme
- ⌘ Chaque utilisateur doit planifier sa venue avec le Responsable de la technologie qui l'intéresse
- ⌘ Une formation Hygiène & Sécurité est dispensée pour la première venue des utilisateurs sur site
- ⌘ Les utilisateurs peuvent amener leur blouse et la laisser dans le casier
- ⌘ Pour le QS6 et Nanodrop, une seule personne par unité a été définie au préalable pour venir faire les manips

Les protocoles ci-dessus peuvent changer à tout moment suivant l'évolution des conditions sanitaires



Matinée des utilisateurs

Le transfert de l'activité amplicon vers Biopuce pour le début d'année 2021



Lidwine TROUILH

Actuellement en formation
sur GeT-PlaGe



- Aucun changement pour les collaborateurs qui souhaitent programmer un projet avec construction de bibliothèques amplicons et séquençage sur Miseq en terme de prix et communication : get-plage.contact@genotoul.fr
- 1 MiSeq sur Biopuce
- 3 MiSeq restent sur PlaGe + probablement 1 achat pour 2021
- **Formation des utilisateurs MiSeq** (constructions des bibliothèques et lancement MiSeq) → toujours dispensée sur GeT-PlaGe
- **Pour les utilisateurs en autonomie MiSeq** → Aucun changement
- **Ready-To-Load** → Aucun changement

Matinée des utilisateurs

Baisse des prix du NovaSeq pour les nouveaux projets

Kit de séquençage	Nombre d'échantillons	Données	Prix INRA HT 2020	Prix INRA HT 2021	Baisse
SP 2x150	48 ARN	325M de reads	7470€	7120€	4,7%
S4 2x150	48 ARN	2000M de reads	14840€	11090€	25,3%
SP 2x150	48 ADN	100 Gb	6510€	6160€	5,4%
S4 2x150	48 ADN	600 Gb	13880€	10130€	27%

Ces prix seront bientôt d'actualité sur la plateforme, nous attendons encore la signature du marché



PacBio : Applications & Premiers résultats !

Matinée Utilisateurs



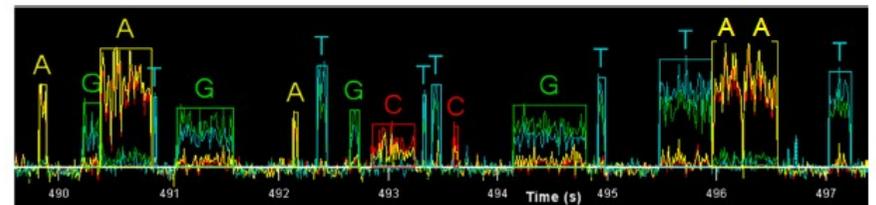
<http://get.genotoul.fr>
get@genotoul.fr
 [@get_genotoul](https://twitter.com/get_genotoul)



Il y a quelques années...

PacBio RSII

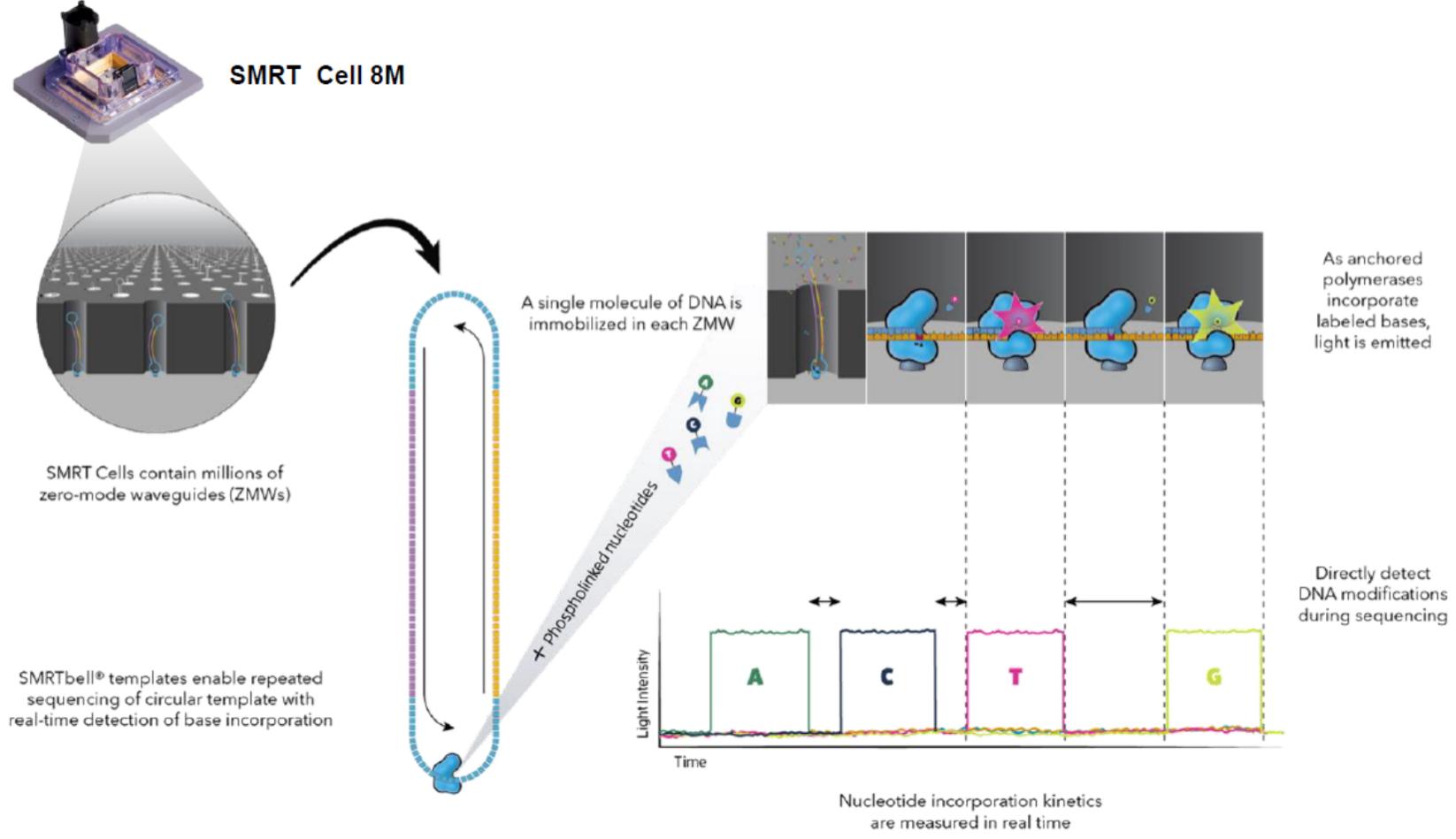
- 500Mb à 1Gb par run
- N50 polymerase reads > 14kb
- Error rate : ~15%
- Durée : 4h
- 8 à 16 SMRTcell max/semaine = ~16 Gb données



Notre PacBio RSII était le seul RSII disponible en France (2015)

PacBio Sequel II : principe de séquençage

SINGLE MOLECULE, REAL-TIME (SMRT) SEQUENCING



PacBio Sequel II

CARACTÉRISTIQUES

Longues lectures : de 500pb jusqu'à 100kb

8 millions de puits sur 1 SMRTcell

8 SMRTcells

Chimie 2.0 :

Mode HIFI (CCS) : taille d'insert entre 10 et 20kb - 15Gb de données CCS

Mode CLR : taille d'insert > 25 kb - 80Gb de données

Multiplexage

Couverture uniforme, pas d'amplification et moins de biais liés au pourcentage de GC

APPLICATIONS

Séquençage ADN *de novo*, détection de variants

Accès à l'information des marques épigénétiques portées par l'ADN

Annotation de génome

Transcriptomique

Métagénomique (16S, opéron...)

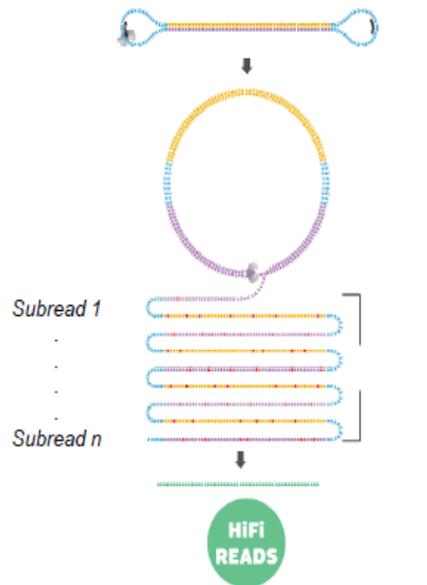


PacBio Sequel II : Mise en place d'un nouveau séquenceur sur la plateforme

TWO MODES OF SMRT SEQUENCING

Circular Consensus Sequencing (CCS) Mode

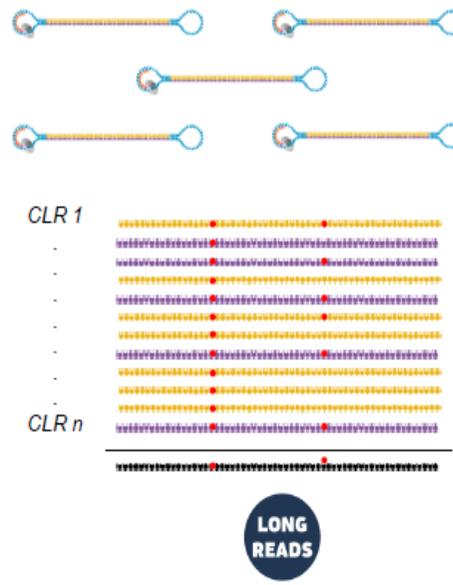
Inserts 10-20 kb



Single-molecule consensus sequence

Continuous Long Read (CLR) Sequencing Mode

Inserts >25 kb, up to 175 kb



Multi-molecule consensus sequence

Spécifications PacBio :
 Mode CLR : 80Gb par SMRTcell
 Mode CSS : 15Gb parSMRTcell

- MODE CCS (HIFI) : 31 SMRTcells passées
 → min 7Gb - max 37 Gb CCS (N50=16kb)
- MODE CLR => 6 SMRTcells passées
 → min 91Gb - max 166 Gb (N50 35kb)

Importance de la qualité des échantillons & pré requis

- Pureté



A260/280 = 1.8-2.0

A260/230 = 2.0-2.2

Nanodrop

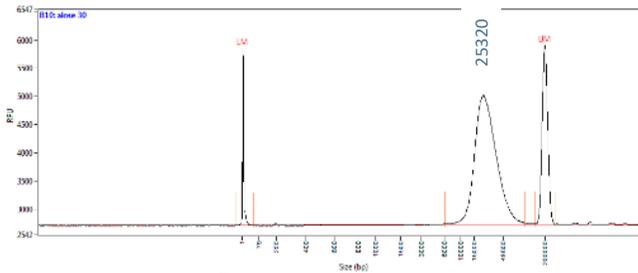
- Quantité



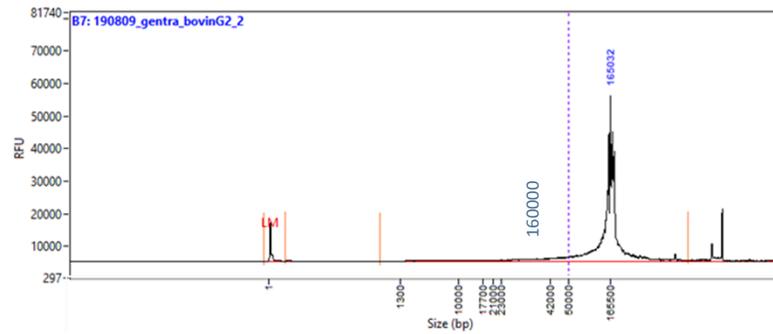
Qubit

Pré-requis ADN PACBIO	
Qté min. de matrice (Qubit) :	15µg (CLR) 20µg (CCS)
concentration min-max :	30 - 50 ng/µl
Pureté 260/280 :	1,8-2
Pureté 260/230 :	2-2,2
Taille ADN minimum :	>20kb

- Qualité



Fragment analyzer
DNA < 40 kb



FEMTO
DNA > 40 kb (up to 165 kb)



La qualité de l'échantillon peut fortement impacter le séquençage



Mode CCS : impact de la qualité de l'échantillon

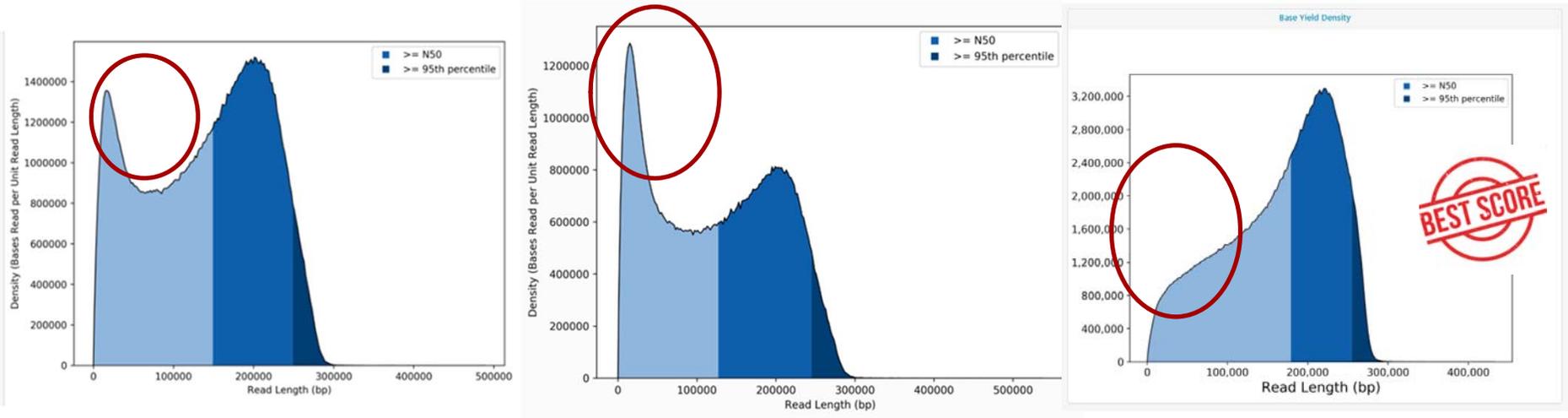
PacBio Sequel II : Premiers runs sur le bovin en mode CCS 20kb



	Extraction 1	Extraction 1 vieillie	Extraction 2
Gb produits	293 Gb	184 Gb	493 Gb
Gb CCS produits	17 Gb	8,7 Gb	36,8Gb



Pic de reads à 0kb : la polymérase se décroche → ADN abimé → run faible en quantité et qualité



Bovin : 4 SMRTcell HIFI → 48Gb → 15X de reads CCS
 1 SMRTcell HIFI → 36,8 Gb → 12X de reads CCS

PacBio Sequel II : Continuous Long Reads (CLR)



Bovin : 3Gb

1 SMRTcell → 42X

Raw Data Report

Value	Analysis Metric
127,528,796,131	Polymerase Read Bases
6,474,789	Polymerase Reads
19,696	Polymerase Read Length (mean)
31,158	Polymerase Read N50
17,914	Subread Length (mean)
28,182	Subread N50
18,407	Insert Length (mean)
28,701	Insert N50
119,138,951,168	Unique Molecular Yield



Mais : 2,4 Gb

1 SMRTcell → 70X

Raw Data Report

Value	Analysis Metric
166,888,040,763	Polymerase Read Bases
6,481,912	Polymerase Reads
25,715	Polymerase Read Length (mean)
41,668	Polymerase Read N50
20,721	Subread Length (mean)
31,117	Subread N50
21,751	Insert Length (mean)
32,673	Insert N50
140,879,921,152	Unique Molecular Yield

Spécifications PacBio :
CLR : 80Gb par SMRTcell

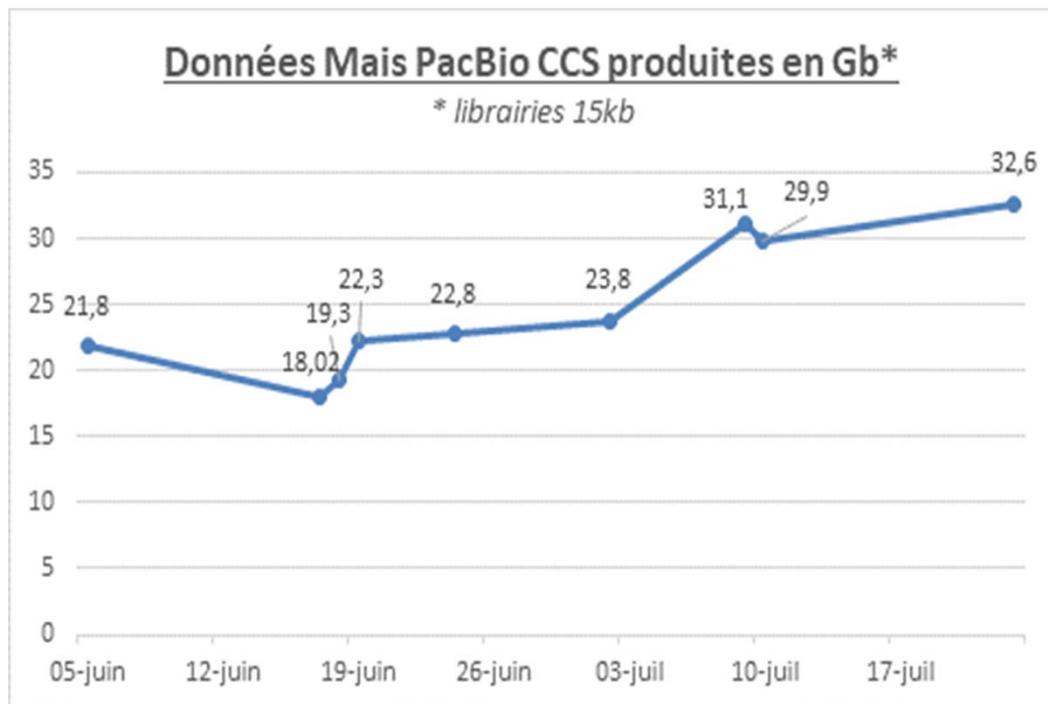
Variabilité des runs en fonction des échantillons et/ou espèces

Modes CCS et CLR : exemple de productibilité avec le Maïs

MODE	Infos	Maïs 1	Maïs 2	Maïs 3	Maïs 4
CLR	Nombre de runs CLR	1	1	1	1
	Total Gb CLR (Gb)	161	166	131	145
	Couverture (X) CLR	67	69	55	60
CCS	Nombre de runs CCS	3	3	3	3
	Total Gb CCS (Gb)	1062	1356	1224	1331
	HIFI (48 Gb attendu)	66,9	85,7	75,5	79,9
	Couverture (X) CCS	28	36	31	33

Spécifications PacBio :
 CLR : 80Gb par smrtcell
 CSS : 15Gb de données HIFI

Evolution de la production de données CCS (HIFI)



Mode CCS : Métriques d'assemblage de contigs sur le Maïs

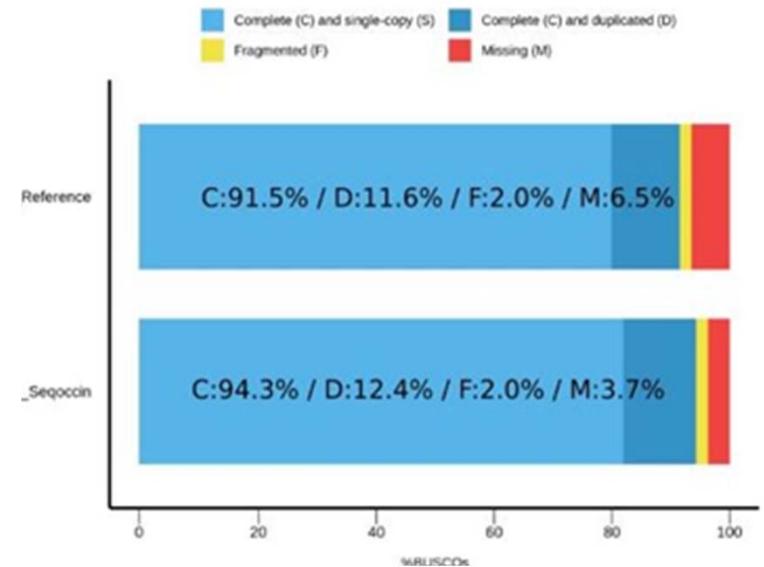
Métrique des Contigs produits par reads PacBio Hifi (CCS) sur le Sequel II de GeT-PlaGe

Couverture	33 X
Nb contigs	2 587
Somme des Tailles des contigs	2 252 981 446
Plus long contig	200 256 324
Longueur N50	51 617 720
Longueur N90	12 650 985
Nombre L50	14
Nombre L90	44

90 % du génome est assemblé grâce à 44 contigs !

→ Comparaison avec l'assemblage du génome de référence réalisé par NRgene

Le % de gènes Busco est supérieur pour SeqOcln :
→ on retrouve plus de gènes !



Combinaison de technologies en fonction de la question biologique



Génome de référence hétérozygote peu répété Ex : BOVIN	Individu concerné	Technologie	Mode	Profondeur
Approche Trio	Père	Illumina	Short Read	15X
	Mère	Illumina	Short Read	15X
Option 1 : pour un bon assemblage	Descendant	PacBio	CLR	60X
Option 2 : pour un très bon assemblage	Descendant	PacBio	HIFI	30X
Option complémentaire : A faire si, pas de génome de référence proche, ou pour voir des éventuelles réorganisations	Espèce concernée	Illumina	Hi-C	20X

Génome de référence homozygote répété Ex : MAIS	Technologie	Mode	Profondeur
Pour un bel assemblage	PacBio	HIFI	20X
Option complémentaire : A faire si pas de génome de référence proche ou si nouvelle lignée	Illumina	HI-C	20X

Questions biologiques

		Technologie	Mode	Profondeur	
Détection de variants					
SNP		PacBio	CLR	60X	
Variants de structures OK et SNVs 77% détectés, besoin d'imputation pour retrouver le reste, besoin d'un génome de référence		PacBio	CLR	15X	
Transcriptomique (IsoSeq)					
Annontation de génome / 1 par lignée suffirait		8 ech	PacBio	IsoSeq	0,125 M
Deep Transcriptomique		1 ech	PacBio	IsoSeq	1M
Sparse sequencing					
12 individus 16X (1ligne S4)			Illumina		16X
1000 individus 0,8X (4 lignes S4)			Illumina		0,8X
		Technologie	Mode	Profondeur	
Metagénomique ciblée					
30 échantillons sur 1 SMRTcell Pacbio en CCS Hifi		PacBio	Hifi	50 000 reads/ech	
Whole metagénome					
3 échantillons sur 1 SMRTcell Pacbio en CCS Hifi		PacBio	Hifi	800 000 Hifi reads/ech	



INRAE

Hi-C

Carole Lampietro

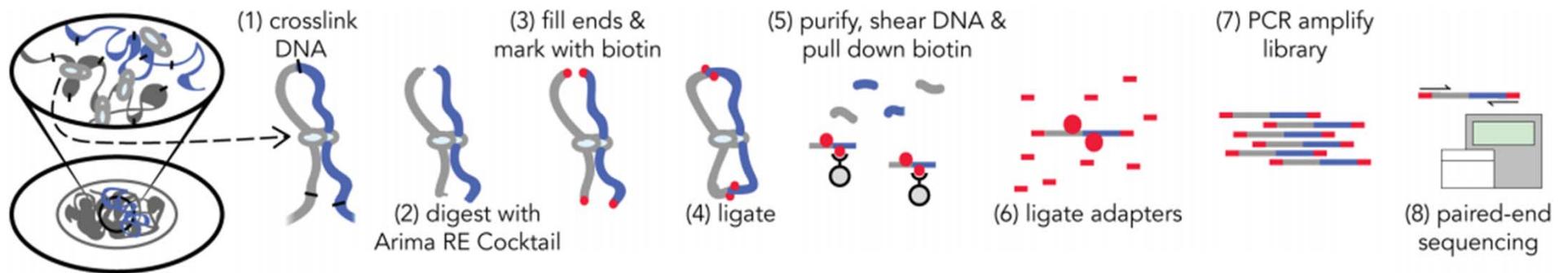
Matinée Utilisateurs

<http://get.genotoul.fr>

get@genotoul.fr

 [@get_genotoul](https://twitter.com/get_genotoul)





§ Protocole sang animal:

- Prérequis: Sang EDTA congelé (10-100 ul sang nucléé, 250ul sang mammifère) , profondeur de séquençage minimum 20X (scaffolding).

§ Protocole plante en cours de validation:

- Prérequis: 1 gramme de jeune feuilles congelées, profondeur de séquençage minimum 20X (scaffolding).

Matinée des utilisateurs



Merci à tous et prenez soin de vous