How does the Nanopore technology meet the needs of our users ?

ONT – 6 Juin 2017

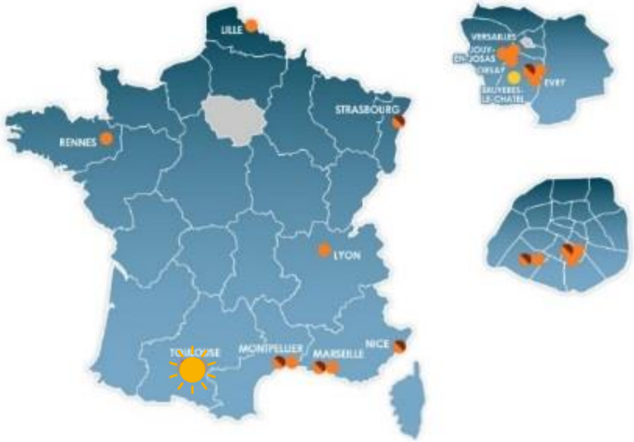
Catherine Zanchetta

http://get.genotoul.fr
get@genotoul.fr
@get_genotoul

GeT

Core Facility

- Genomics and Transcriptomics (GeT) Platform of Genotoul hosted by INRA



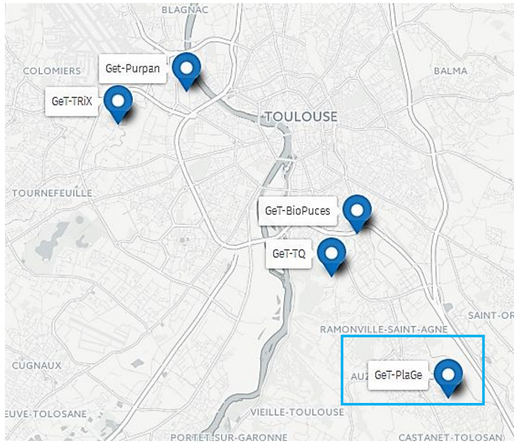
- A strong partnership with a bioinformatics core facility Genotoul Bioinfo
- A node of the National Distributed Infrastructure « France Génomique »
- Quality certifications ISO9001 & NFX 50 900, Propel

GeT

Team and expertise

35 people on 5 sites

- **Experts in Agronomy, Environment, Microbiology, Health**
- **Competence in biology, bioinformatics, biostatistics**

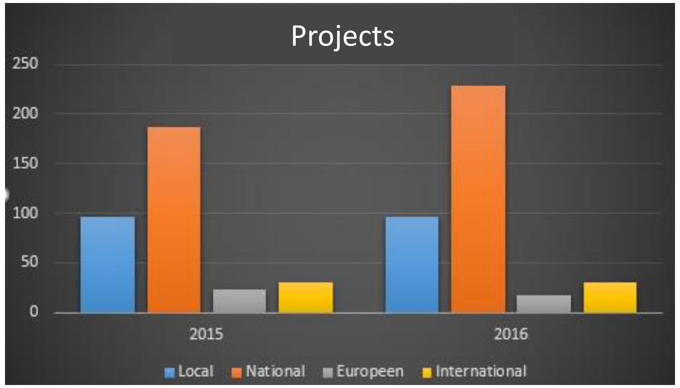


GeT-PlaGe

Activity

More than 140 laboratories in 2016 **(INRA, CNRS, INSA, INSERM, CHU, CIRAD …)**

- **More than 240 research teams**
- **More than 360 projects**
- **>2M€ of activity**



10 R&D projects

- **Chromium, MinION, Chipseq, methylation, HIC…**

19 Research projects **(ANR, INCA, H2020 …)**

91 Publications as co-author **since 2010**


nature COMMUNICATIONS

ARTICLE

Reinforcement of ST
human embryonic st
pluripotency

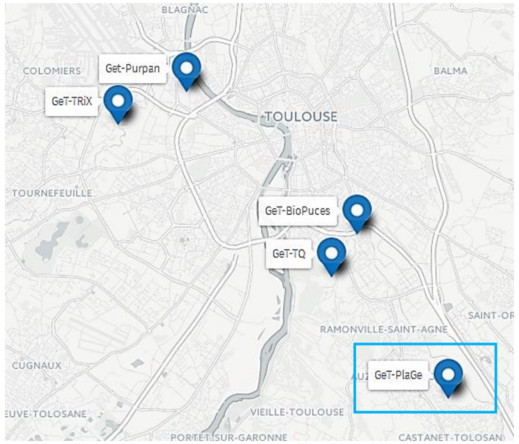
nature genetics

ARTICLES

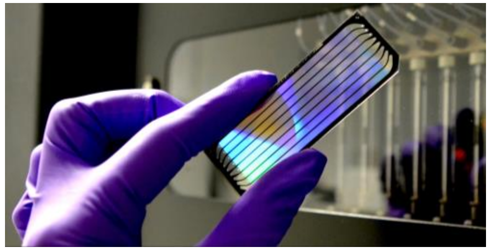
Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle

GeT-PlaGe

NGS technologies



Short reads

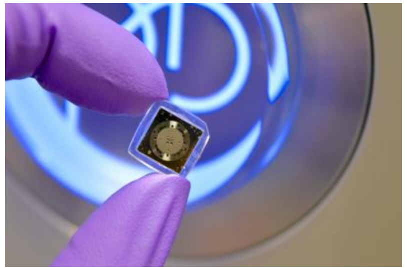


Illumina : MiSeq + HiSeq 3000

Synthetic long reads

10x genomics : Chromium

Long reads



Pacbio : RSII

ONT : MinION

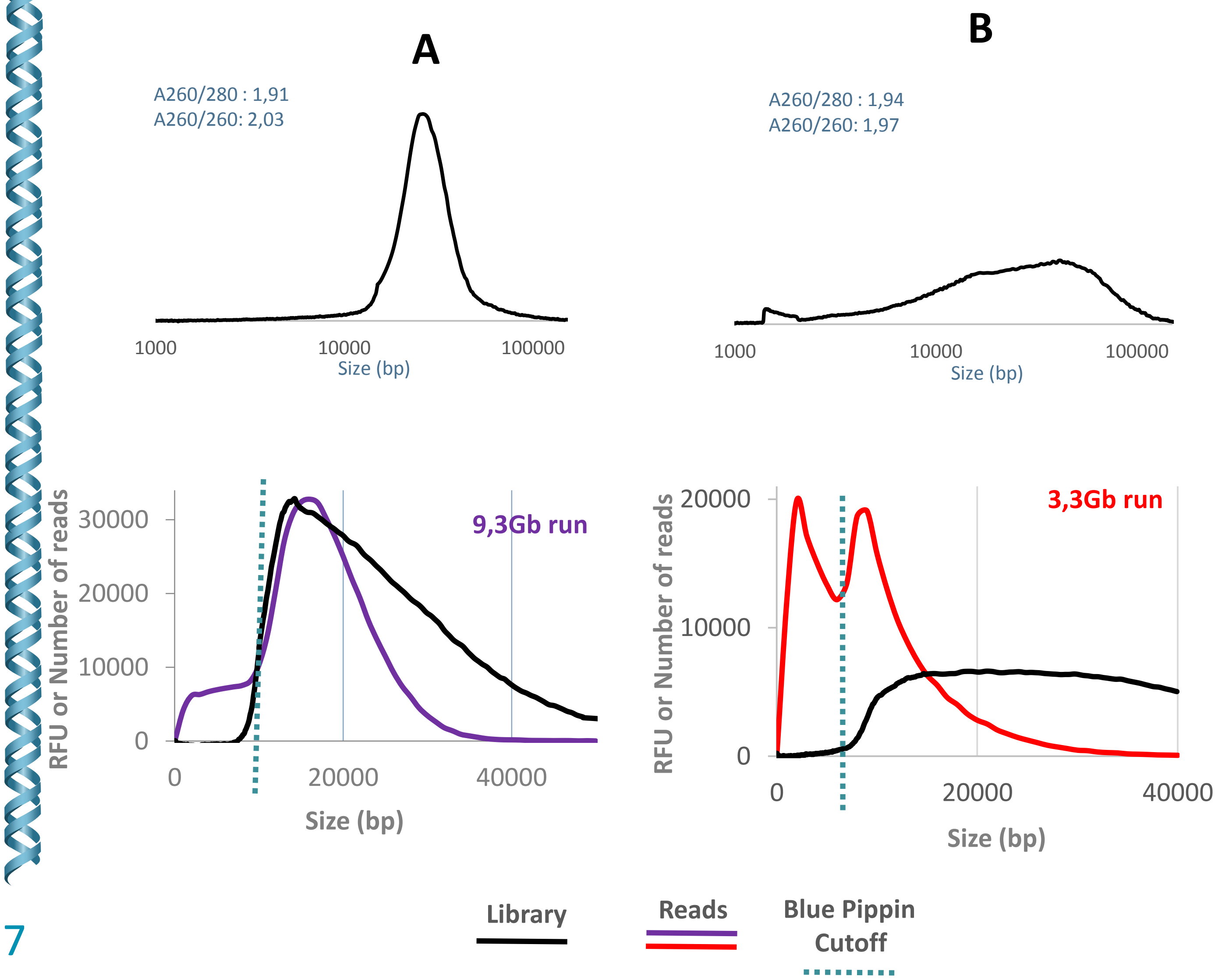
What am I going to talk about for the next 20 minutes ?

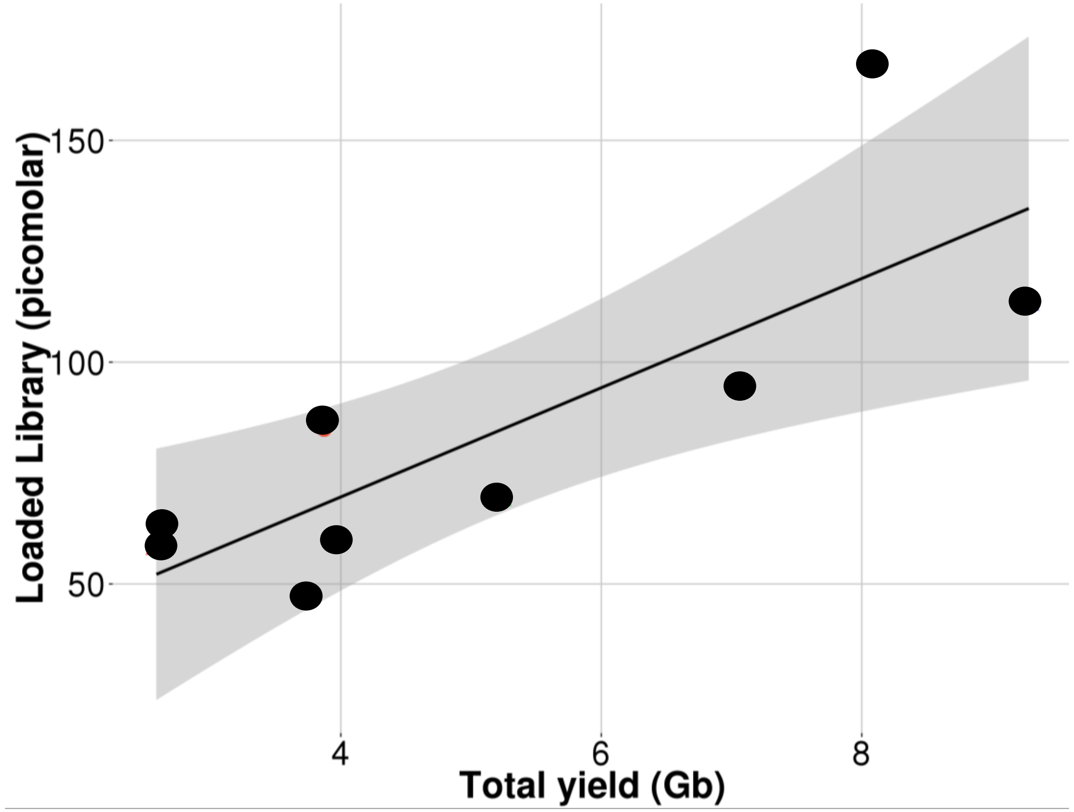
- **Assesment of the technology**

- **4 biological questions (gDNA)**
- **Production of sequencing data**
- **Veterinary diagnosis**
- **Agronomics reasearch**
- **Complexe genome, contaminant molecule**

Assesment of the technology

What can I expect with the MinION ?

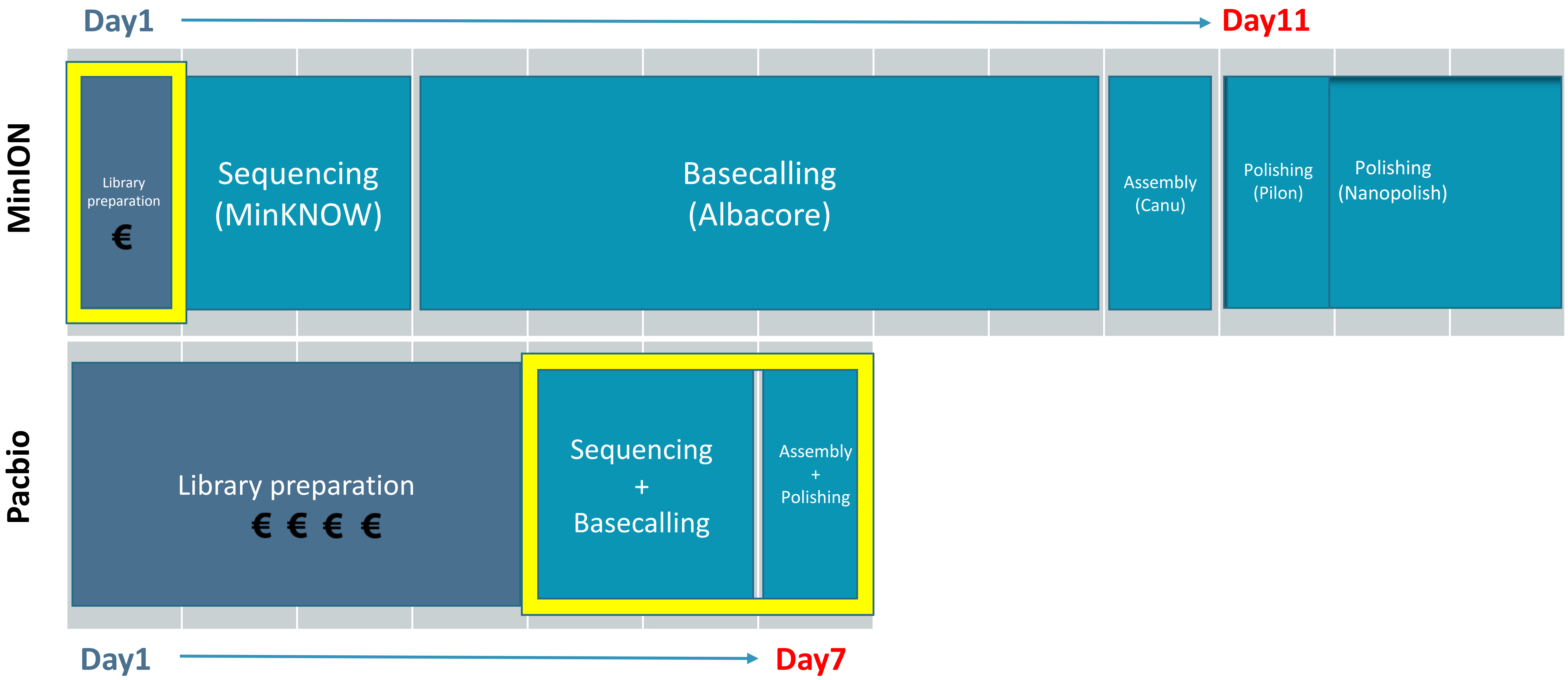




Assesment of the rechnology

When will I get my results ?

Example of a bacterial genome assembly :
Same quantity of data / Same informatics resources



Day1 → Day11

MinION

Library preparation
€

Sequencing
(MinKNOW)

Basecalling
(Albacore)

Assembly
(Canu)

Polishing
(Pilon)

Polishing
(Nanopolish)

Pacbio

Library preparation
€ € € €

Sequencing
+
Basecalling

Assembly
+
Polishing

Day1 → Day7

Assesment of the technology

Accuracy on Raw Data and Qualitative analysis of an assembly what is the best option ?

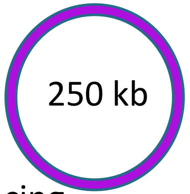
Example of a bacterial genome assembly

Technologie	Accuracy
Illumina	99.7 %
Pacbio	**81.3 %**
MinION Albacore V0	81.0 %
MinION Albacore V1	**86.3 %**

Assembly methods	% Complete genes	% Fragmented genes	% Missing genes
MinION	12.2	27	60.8
MinION-nanopolish	71.6	15.5	12.9
MinION-pilon	**95.3**	**0**	**4.7**
Pacbio-hgap3	**95.3**	**0**	**4.7**

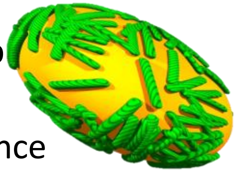
MinION and Users needs

Biological questions, technical answers



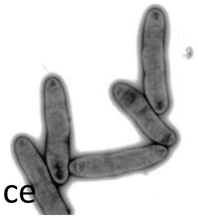
Production
BAC clones sequencing
Same or better assemblies compared to the Pacbio

300 kb

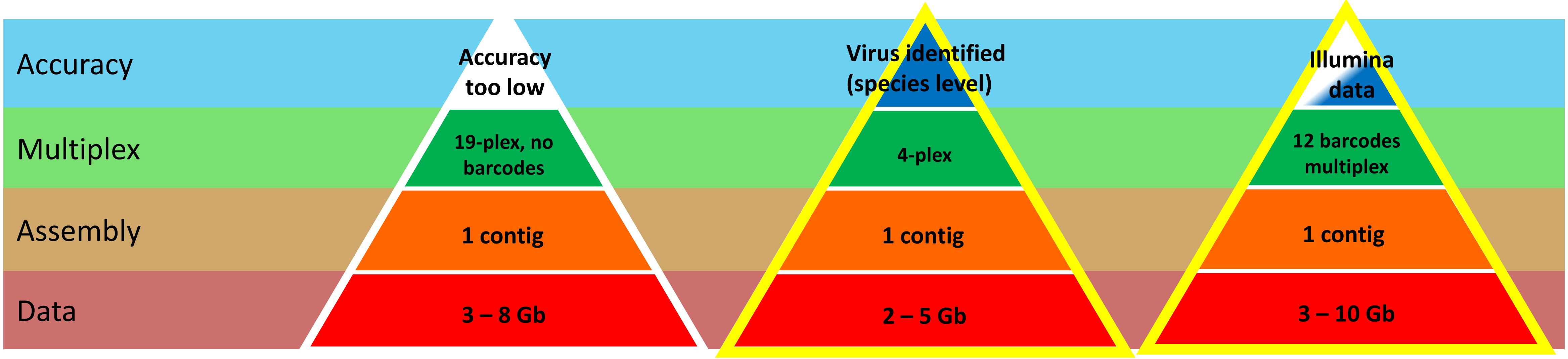


Veterinary surveillance
Support tool
Virus identification in a biological sample

5 Mb



Plant-pathogen arm race
Link *Xcc* genotypes to *A. thaliana* phenotypes
Identification of SNPs



	Production	Veterinary surveillance	Plant-pathogen arm race
Accuracy	**Accuracy too low**	**Virus identified (species level)**	**Illumina data**
Multiplex	**19-plex, no barcodes**	**4-plex**	**12 barcodes multiplex**
Assembly	**1 contig**	**1 contig**	**1 contig**
Data	**3 – 8 Gb**	**2 – 5 Gb**	**3 – 10 Gb**

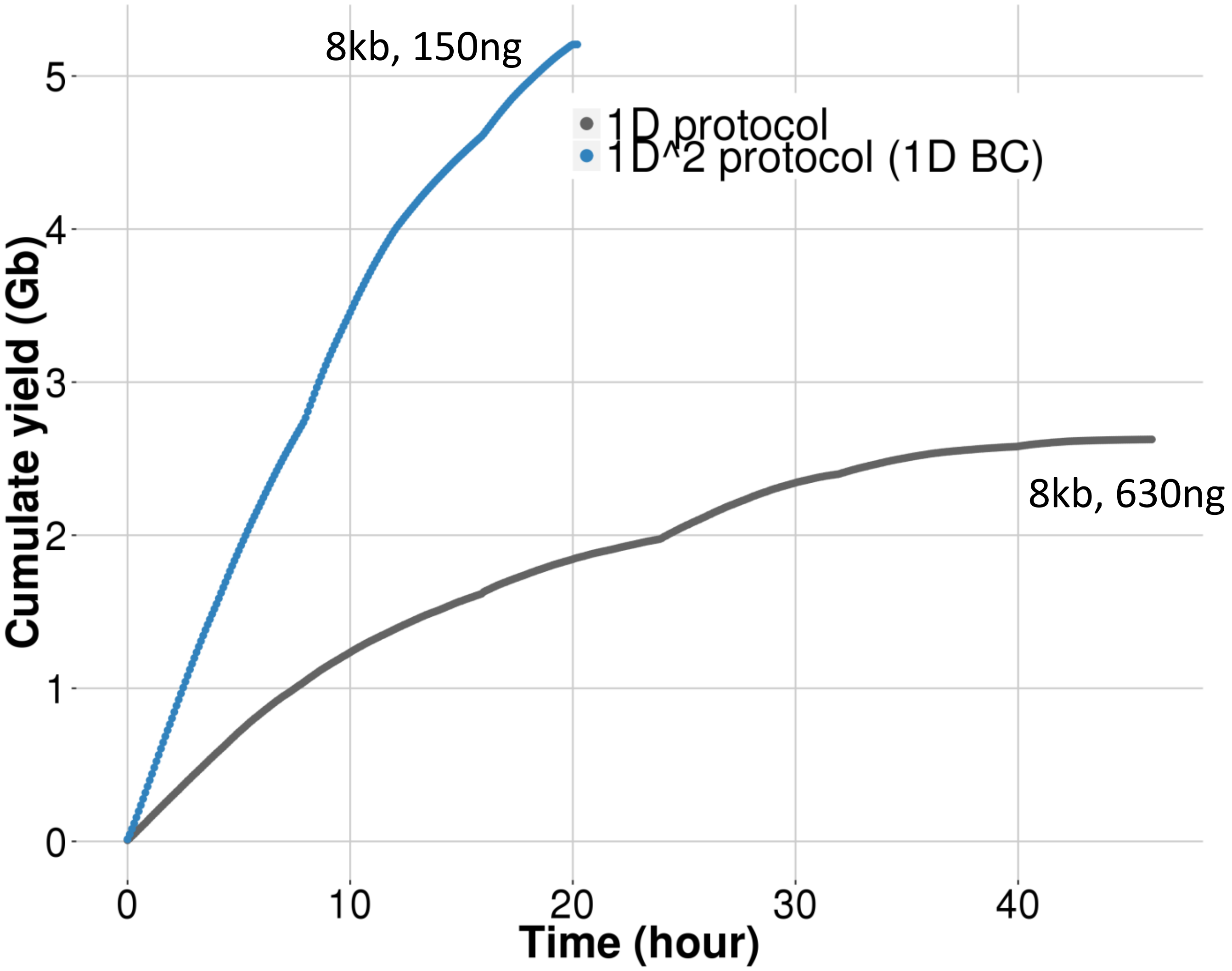
Caroline Callot, Stéphane Cauet, Hélène Berges

Guillaume Croville, Jean-luc Guerin

Baptiste Mayjonade, Jérôme Gouzy, Fabrice Roux

MinION vs Users needs

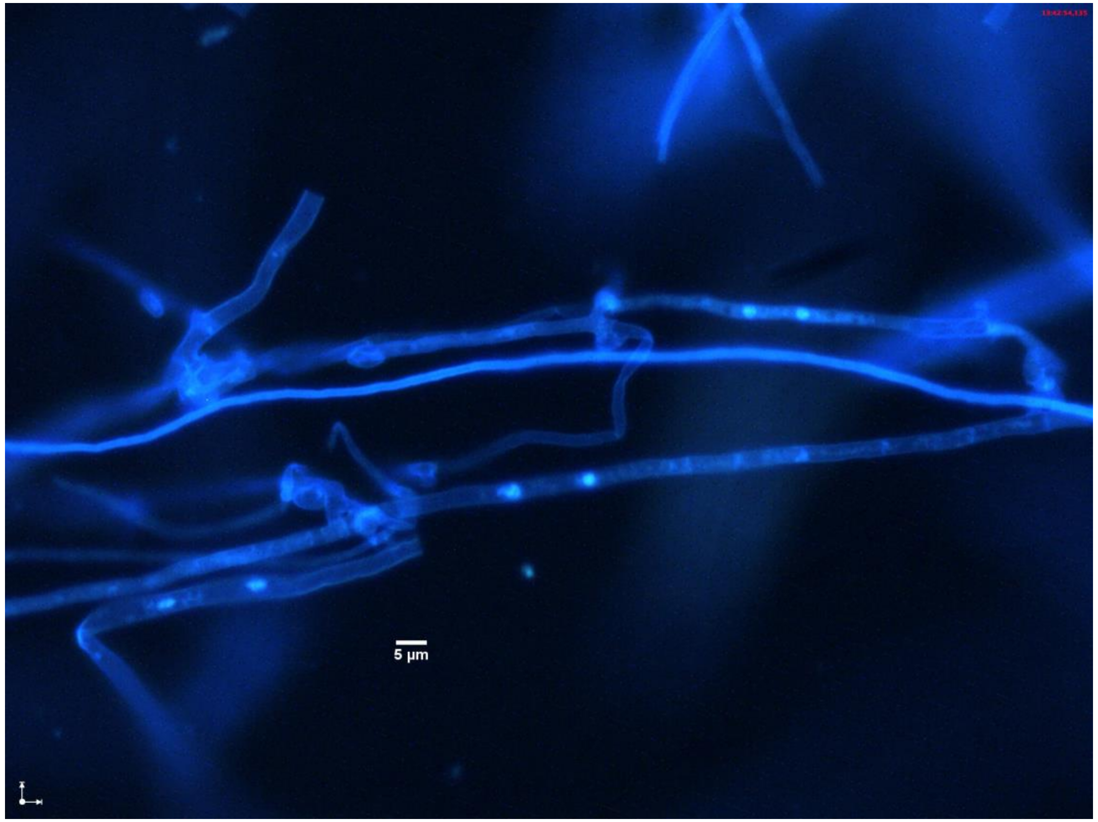
New chemistry : 1D² / R9.5



1D² : R9.5 / MinKNOW 1.6.11 / Albacore 1.1.0
1D : R9.4 / MinKNOW 1.1 / Albacore 1.0.1

Complex genome

Ganoderma boninense – Oil palm pathogen



2 nucleus
Heterozygous

Illumina results :

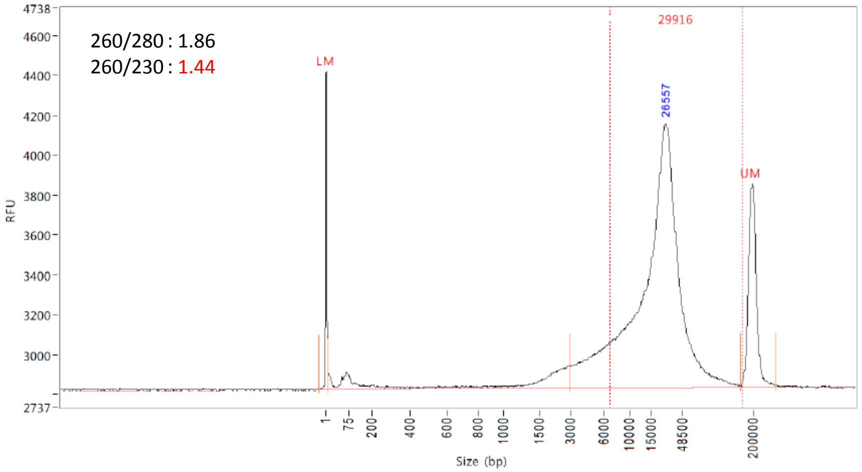
Number of contigs	**10164**
Longest contig	170 kb
N50 contig lenght	**12 Kb**
L50 contig count	1367

Repeated regions > 10 kb

50 Mb

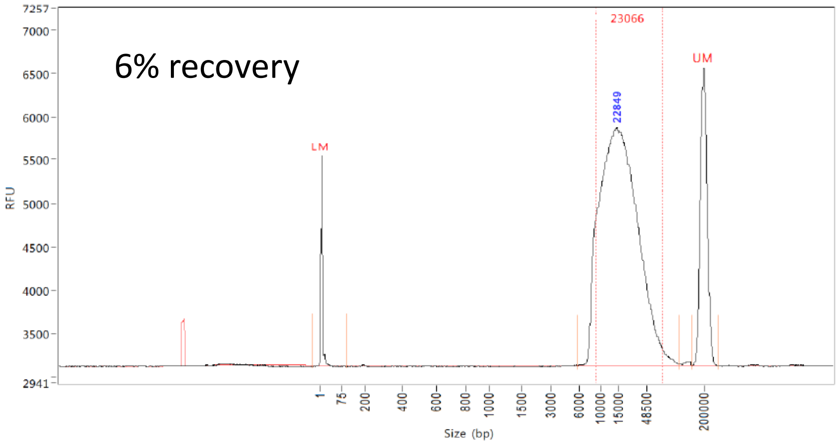
Pacbio vs Contaminated DNA

The polymerase stops



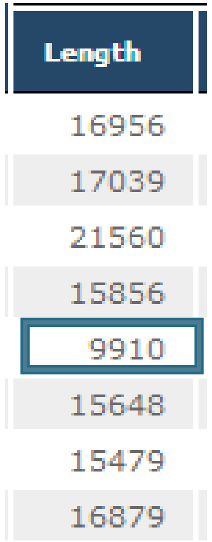
260/280 : 1.86
260/230 : 1.44

Pacbio

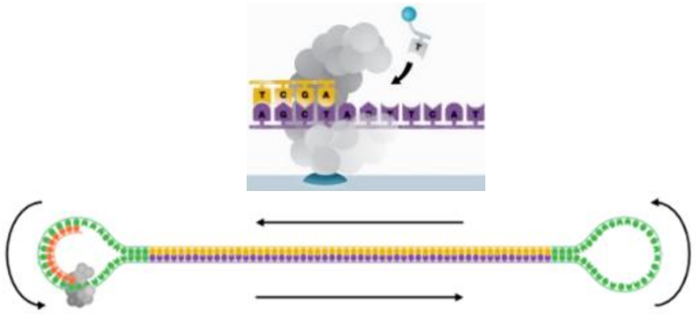


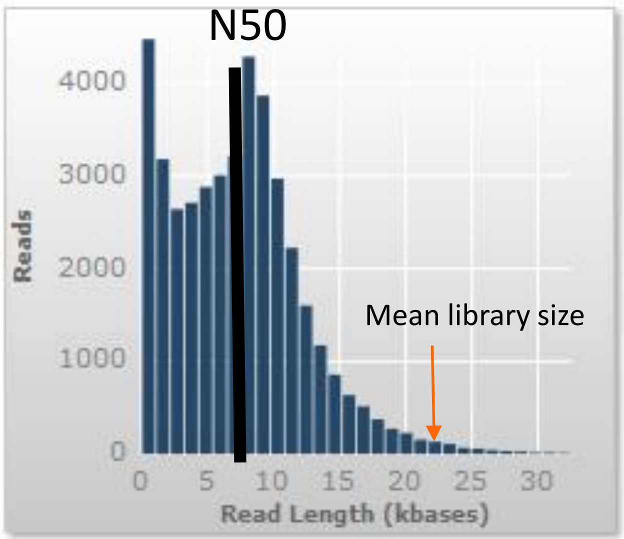
6% recovery

Polymerase Reads



Length
16956
17039
21560
15856
9910
15648
15479
16879





N50

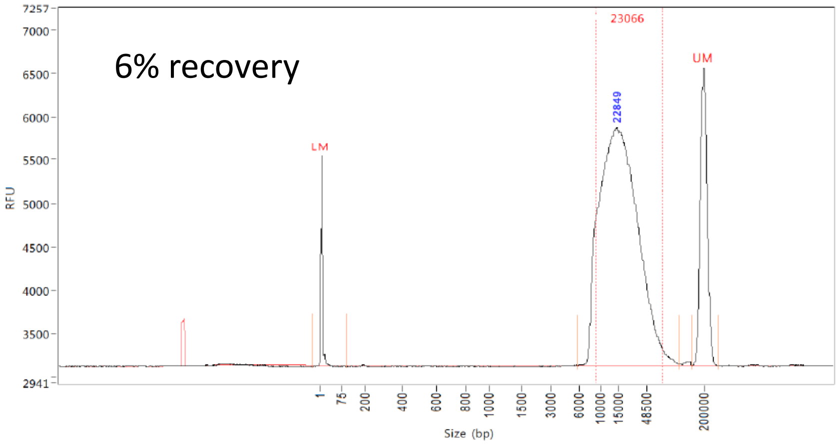
Mean library size

Pacbio/MinION vs Contaminated DNA

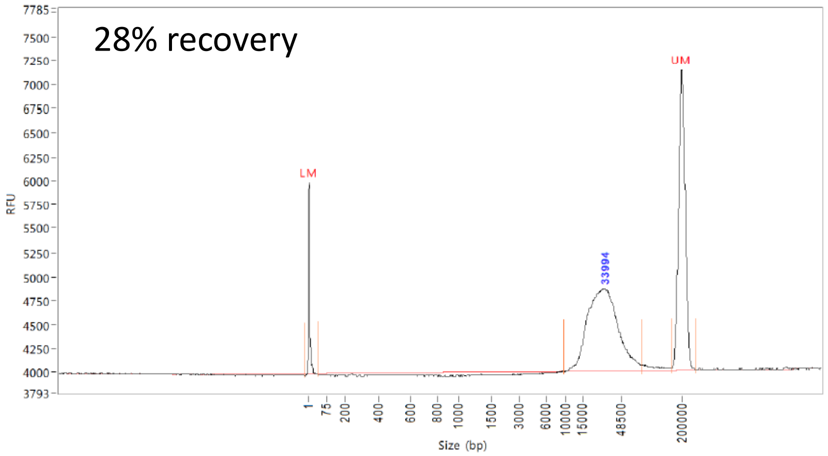
The MinION keep going when the Pacbio polymerase stops

	Pacbio	Nanopore
Library prep	5 days (€€€€)	**1 day (€)**

Pacbio


6% recovery

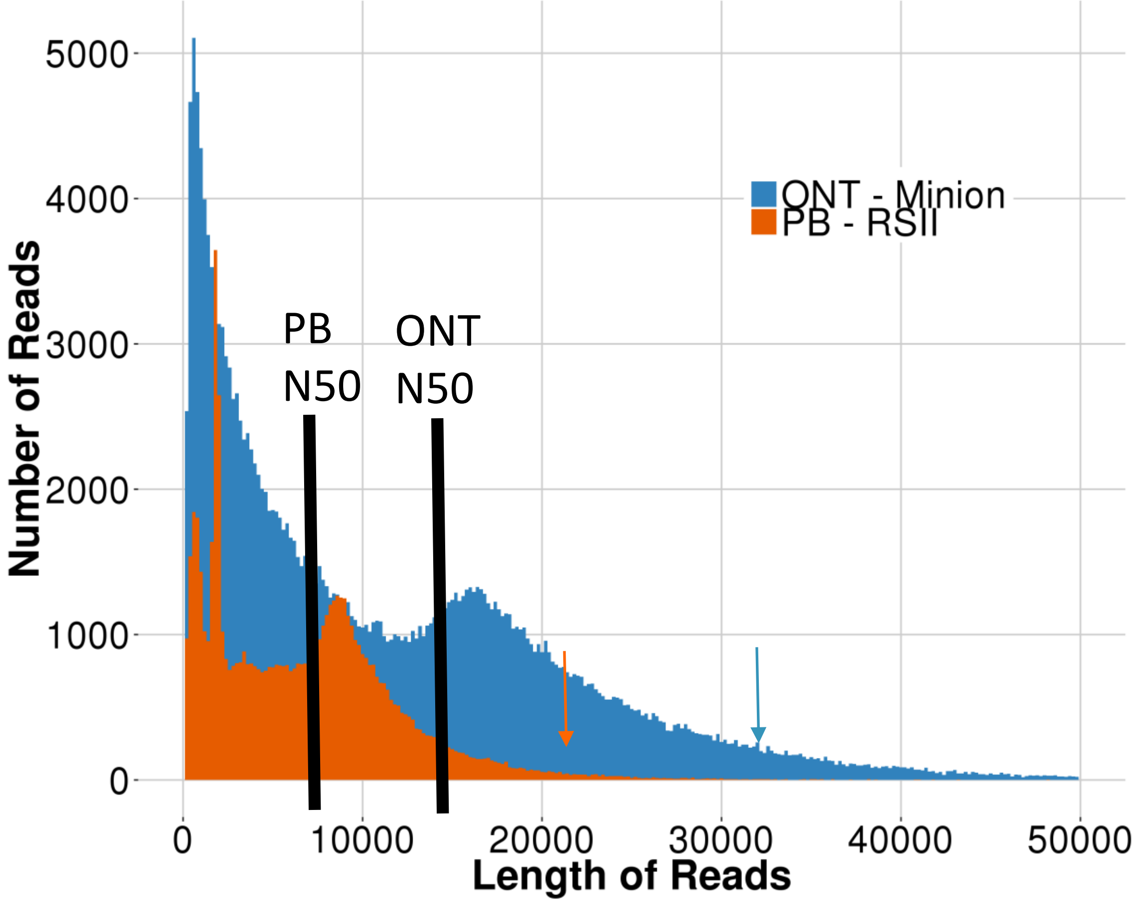
MinION


28% recovery

Contaminated DNA + Complex genome

Technologies comparison

	Pacbio	Nanopore
Library prep	5 days (€€€€)	**1 day (€)**
Reagents	6 SMRT	**1 FC**
Quantity of data	4.4 Gb	4.3 Gb
WindowMaker Haplomerger 1x Pilon	*Ref + Alt*	*Ref + Alt*
Genes complete	92 %	89 %
Proteins found	89 %	86 %



Close results

- 2 or 3 x Pilon ?

- Systematic errors ?
- Contaminant/Signal ?

- Tools ?

Conclusion

- **MinION technology : Long DNA fragments = long reads**

Sample → Extraction → Lib Prep → Sequencing → Bio informatics → Bio statistics

- **Depending on the biological question, the MinION can give a partial or a complete answer**

- **Most part of the time, carry out study using multiple technologies will be the best option**

- **To go further and improve genomes assembly :**

  scaffolding technologies (10X Genomics, BioNano)…

Tomato genome

Technology	N50
PacBio (RSII 70 x)	3.2 Mb
+ Bionano (2 enzymes)	32 Mb
+ Chromium + Illumina (100x)	45 Mb

Remerciements

Maxime MANNO, Céline Vandecasteele, Claire Kuchly, Cécile Donnadieu, Olivier Bouchez, Gérald Salin, Alain Roulet, Céline Roques

Christophe Klopp

Frédéric Breton, Alexandra Vaillant, Leatizia Camus-Kulandaivelu

Tan joon sheong, Sharifah Shahrul Rabiah Syed Alwee, Kwan Yen Yen

Baptiste Mayjonade, Jérôme Gouzy, Fabrice Roux

Guillaume Croville, Jean-luc Guerin

Caroline Callot, Stéphane Cauet, Hélène Berges

Corinne Cruaud

Mohammed Zouine, Pierre Frasse, Mondher Bouzayen