



GeT experience on long fragments technologies

Denis Milan, INRA & Genotoul, Toulouse, France

<http://get.genotoul.fr>

@get_genotoul

@NanoporeConf | #NanoporeConf

ORGANISED BY
 NANOPORE

GeT Platform at Toulouse (France)



Genomics and Transcriptomics (GeT) Platform of Genotoul hosted by



A strong partnership with Bioinformatics platform



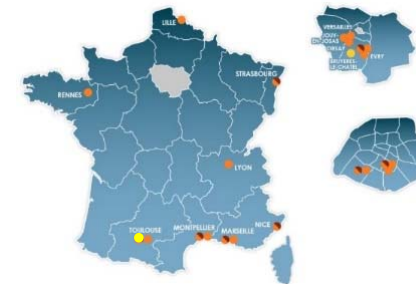
A node of the National Distributed Infrastructure « France Génomique » (60 M€ / 8 y)



Quality certification ISO9001 & NFX 50 900, Propel



A complete portfolio of sequencers at GeT



1

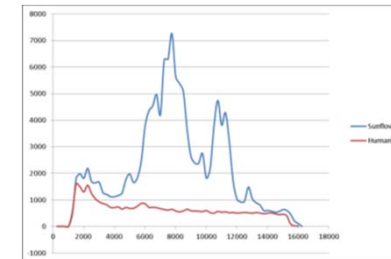
**First experience on
PacBio RSII**

Evolution of Long Fragments sequencers



A need for long fragments sequencing to sequence plant genomes

Sunflower : **30 %** of repeated sequences (LTR)
Human : **8.8 %** of repeated sequences



January 2015 :

- Oxford Nanopore was ever in development
- P6/C4 chemistry of PacBio permits sequencing of longer fragments → Investment in a PacBio RSII

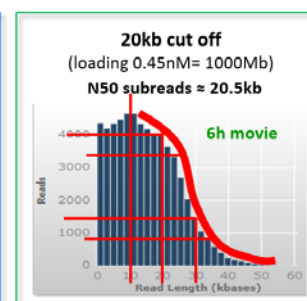
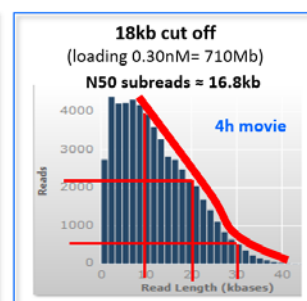
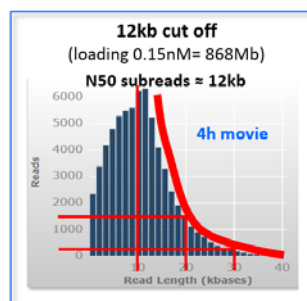
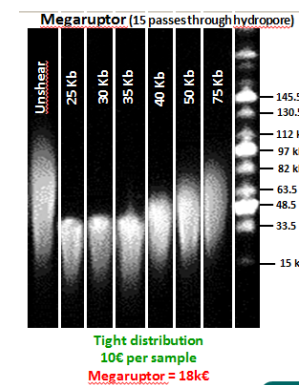


Optimization on PacBio RSII



Improvement of read length :

- Shearing with Megaruptor
- Sizing with Blue Pippin
- Increase of run length



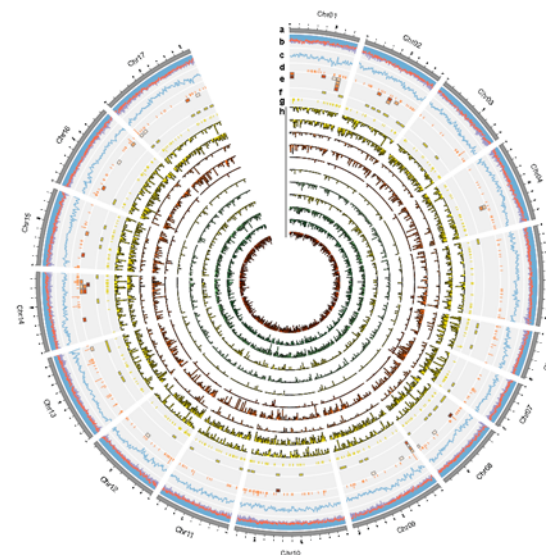
Assemblies of Sunflower genome



- INRA (Sunflower Team) :
Hiseq 127 X
→ 43 % coverage
- International consortium :
454, Hiseq, Genetic and physical map
→ 63 % coverage
- INRA (Sunflower Team) :
PacBio 107 X (407 SMRT)
→ 84 % coverage



13124 contigs N50 = 498 kb
+90 % anchor



Longest subreads

80974 bp
79860 bp
79834 bp
78105 bp
77481 bp



Nicolas Langlade, Stephane Munoz, Jérôme Gouzy, Baptiste Mayjonade,

@NanoporeConf | #NanoporeConf



© 2017 Oxford Nanopore Technologies. All rights reserved. The MinION is for research use only.



2

Test of Minion potential

Back to the future ...

Minion pilot projects

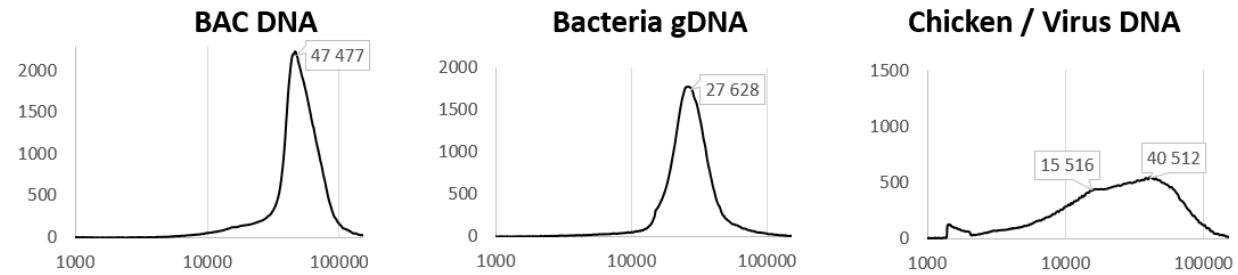


Setting up pilot projects :

- 3 kind of materials : **BAC Clones, bacterial DNA, Virus** infecting animals
- 3 kind of kits : **Rapid, 1D, 2D**
- 3 kind of basecallers : **Minknow** (local), **Metrichor** (cloud), **Albacore** (local)



First results on Rapid Run

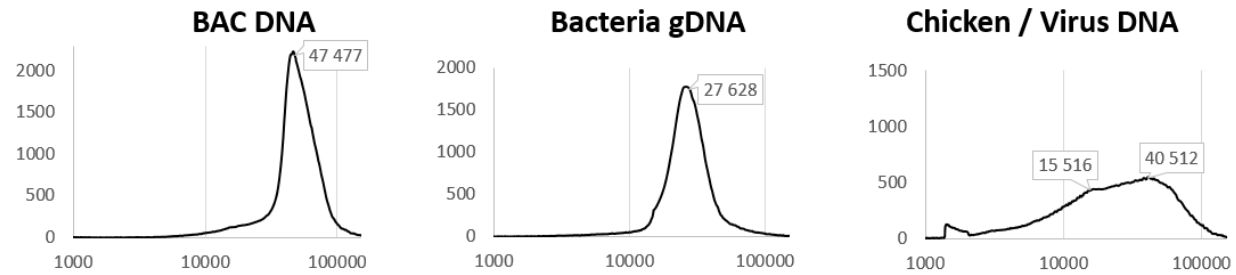


Preparation kit used and Amount of DNA required	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)
Rapid 200 ng	0,1	37,9	0,5	10,1	0,3	33,5

First results on 1D run



Poster 22



Preparation kit used and Amount of DNA required	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)
Rapid 200 ng	0,1	37,9	0,5	10,1	0,3	33,5
1D 1,5µg	2,6	9,6	4,0	8,7	3,9	9,4

First results on 1D run



Poster 22

Guillaume Croville, Guillaume Le Loc'h, Catherine Zanchetta, Maxime Manno, Christophe Klopp, Cécile Donnadieu, Mattias Delpont, Christelle Camus-Bouclainville and Jean-Luc Guérin

Direct whole genome sequencing of poxvirus using Oxford Nanopore MinION

Guillaume Croville¹, Guillaume Le Loc'h¹, Catherine Zanchetta², Maxime Manno³, Christophe Klopp⁴, Cécile Donnadieu⁵, Mattias Delpont⁶, Christelle Camus-Bouclainville⁷ and Jean-Luc Guérin¹

¹ Université de Toulouse, INRAE, UR141, INRAE 1228, 31078 Toulouse, France
² Oxyphac, INRAE 12520, Carrière/Bouclainville, France
³ Patience Biomathématique Genocod, UR275, Mondria et Intelligence Artificielle, INRA Carrière/Bouclainville, France
⁴ INRAE Genotoul

Material & methods

From tracheal lesions to embryonated egg culture on the chorioallantois

Tissue lysis @ 55°C using QIAzol Lysis Reagent

DNA extraction using Phenol:Chloroform:IAA

1D Library prep

Sequencing on a R.2.4 Flow cell

Basecalling

Sequence alignment with Burrows-Wheeler Aligner

Alignments display with IGV

Introduction and objectives

During Fall 2015, two independent cases of fowlpox were diagnosed in commercial layer farms in western France. This disease is caused by an avian poxvirus whose genome is linear and reaches 200 kb in length. They belong to the Poxviridae family (Figure 1) like the causative agent of variola.

Diseased birds showed severe dyspnea and suffocation before death. The main macroscopic lesion was a subcutaneous bleb with severe thickening of the mucosa. Abscesses containing blebs could be observed in any birds. All tracheal mucus and tissues sampled in both farms tested positive for PCR targeting pathogen shared by all avian poxviruses.

The aim of this proof-of-concept study was to generate a full genome sequence, directly from the lesion. The genomes of poxviruses show repeated regions, which may be challenging for genome assembly from short reads. MinION sequencing was therefore assessed to obtain an easy-to-use genome assembly.

Results

A total of 755 838 reads were obtained, of which 610 797 were filtered for quality and analyzed using Metrichor® (Oxford Nanopore Company). The average sequence length was 3.63 kb (Figure 2). A total of 39 625 viral reads were aligned on a reference fowlpox genome sequence using the Burrows-Wheeler Aligner software package² and a consensus sequence of 288 539 bp was obtained with a mean depth of 638X (Figure 3).

Discussion

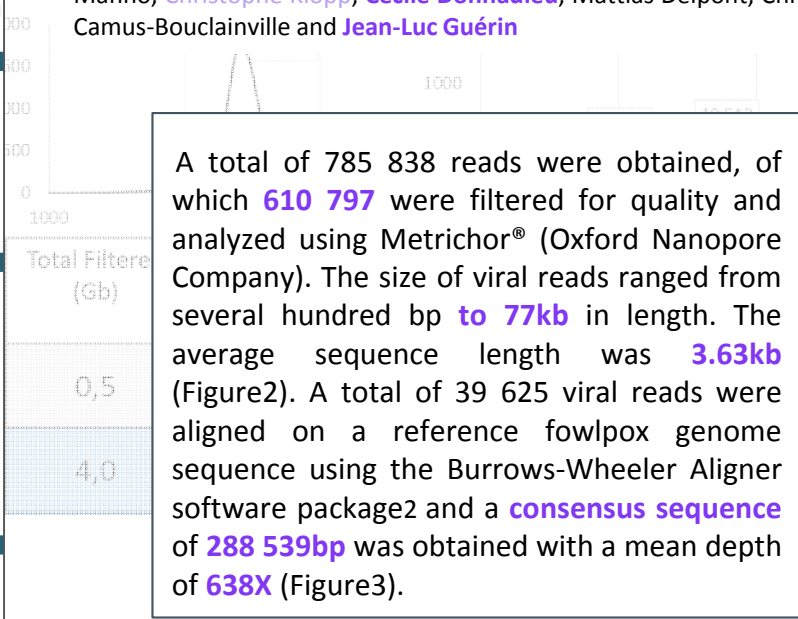
This proof-of-concept study was performed to assess the suitability of MinION sequencing for surveillance of genomes of both medical and veterinary importance.

Using the 1D sequencing ML, we were able to readily generate a full genome sequence of a fowlpox virus. The size of viral reads ranged from several hundred bp to 77 kb in length, allowing an easy assembly.

The complete genome assembly confirmed that this fowlpox virus is clustered with cattle AL and hosts a full length melanconinB virus (EBV) insert. The complete genomic features of this virus remain to be thoroughly explored.

Further sequencing assays are being performed directly from tissues without propagation, and the preliminary results are consistent with the data presented here.

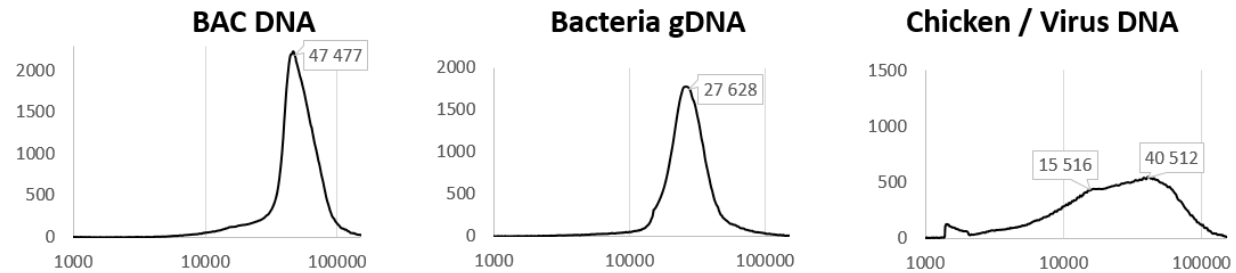
In conclusion, MinION sequencing is suitable for rapid diagnostics of clinical samples and de novo assembly of poxvirus genomes.



A total of 785 838 reads were obtained, of which **610 797** were filtered for quality and analyzed using Metrichor® (Oxford Nanopore Company). The size of viral reads ranged from several hundred bp to **77kb** in length. The average sequence length was **3.63kb** (Figure2). A total of 39 625 viral reads were aligned on a reference fowlpox genome sequence using the Burrows-Wheeler Aligner software package² and a **consensus sequence of 288 539bp** was obtained with a mean depth of **638X** (Figure3).



First results on 1D run



Preparation kit used and Amount of DNA required	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)
Rapid 200 ng	0,1	37,9	0,5	10,1	0,3	33,5
1D 1,5µg	2,6	9,6	4,0	8,7	3,9	9,4

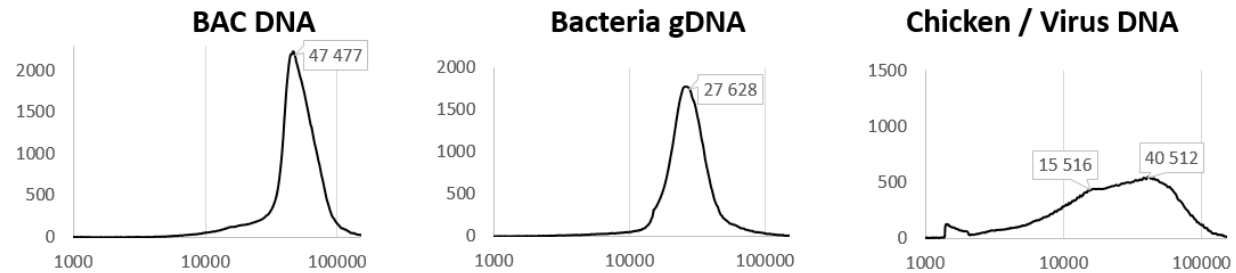
Shearing with Megaruptor (20 kb)



Sizing with Blue Pippin (11-50 kb)



Improvement of 1D to “1D+”



Preparation kit used and Amount of DNA required	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)
Rapid 200 ng	0,1	37,9	0,5	10,1	0,3	33,5
1D 1,5µg	2,6	9,6	4,0	8,7	3,9	9,4
1D+ $\frac{1,5 * \text{mean size}}{8}$		9,3	19,0	3,3	15,1	

Shearing with Megaruptor (20 kb)

@NanoporeConf | #NanoporeConf

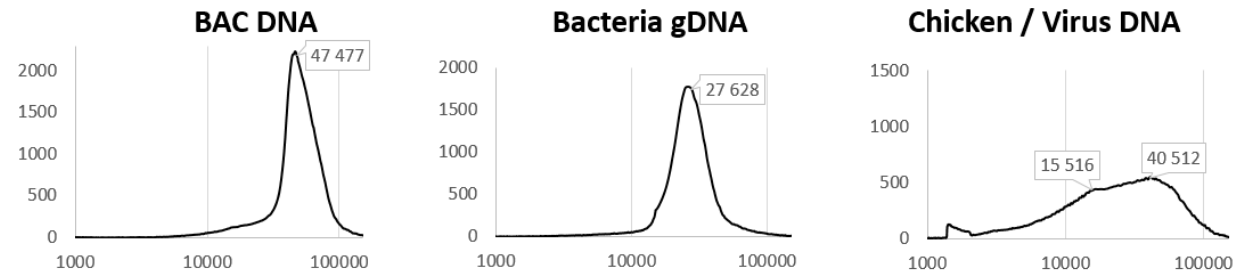


Sizing with Blue Pippin (11-50 kb)

Technologies. All rights reserved. The MinION is for research use only.



Improvement of 1D to “1D+” to “1D++”



Preparation kit used and Amount of DNA required	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)	Total Filtered (Gb)	N50 (Kb)
Rapid 200 ng	0,1	37,9	0,5	10,1	0,3	33,5
1D 1,5µg	2,6	9,6	4,0	8,7	3,9	9,4
1D+ / ++ $\frac{1,5 * \text{mean size}}{8}$			9,3	19,0	3,3	15,1

Shearing with Megaruptor (40 kb)

@NanoporeConf | #NanoporeConf

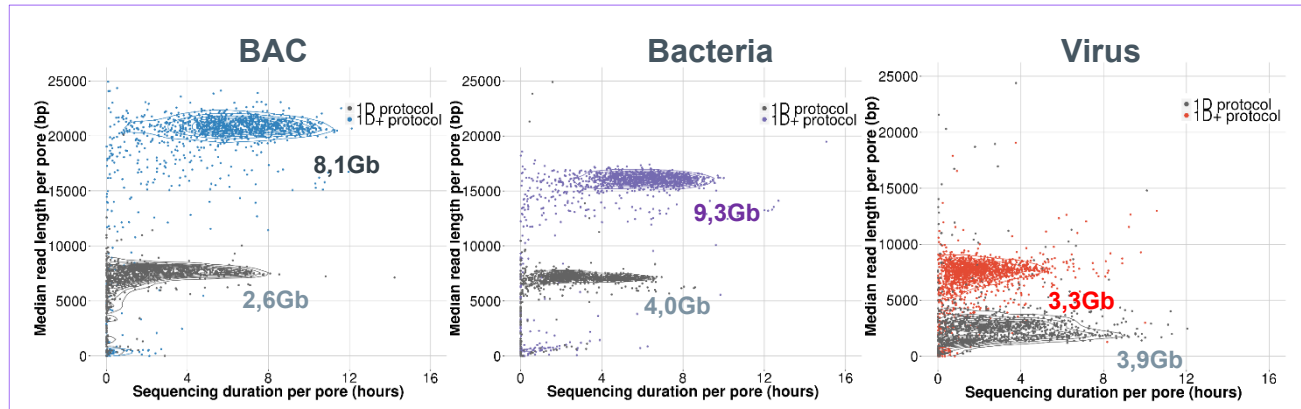


Sizing with Blue Pippin (16-50 kb)

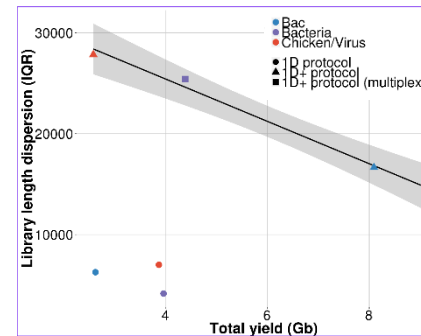
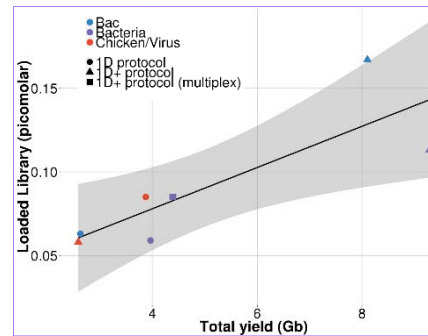
Technologies. All rights reserved. The MinION is for research use only.



Some preliminary feelings to discuss

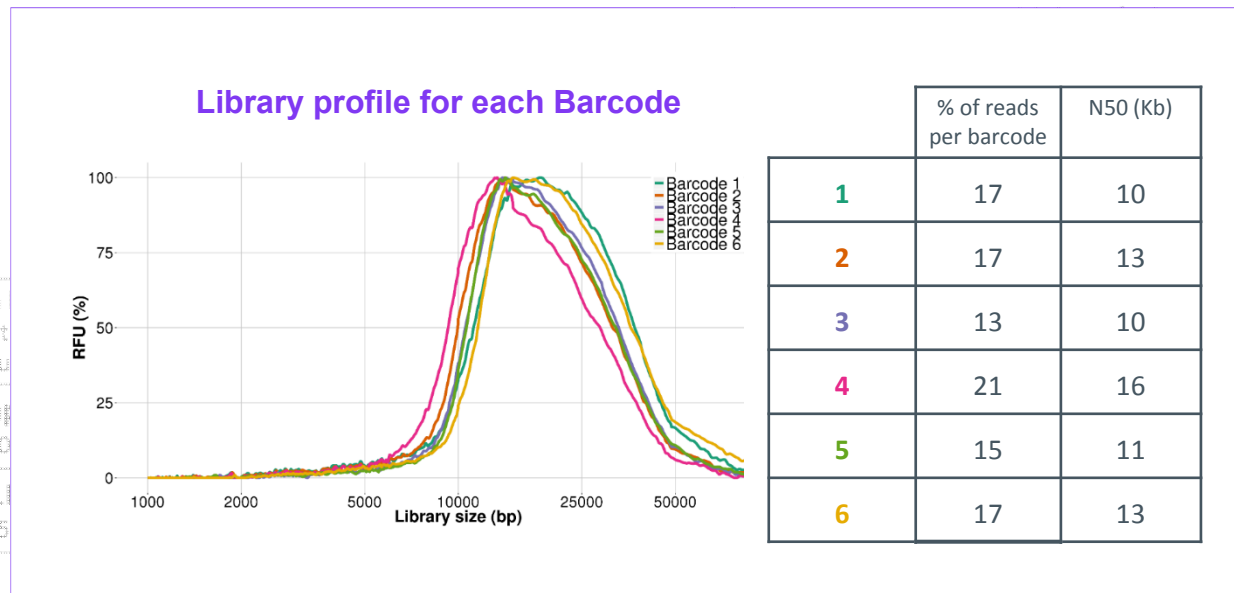


No overloading



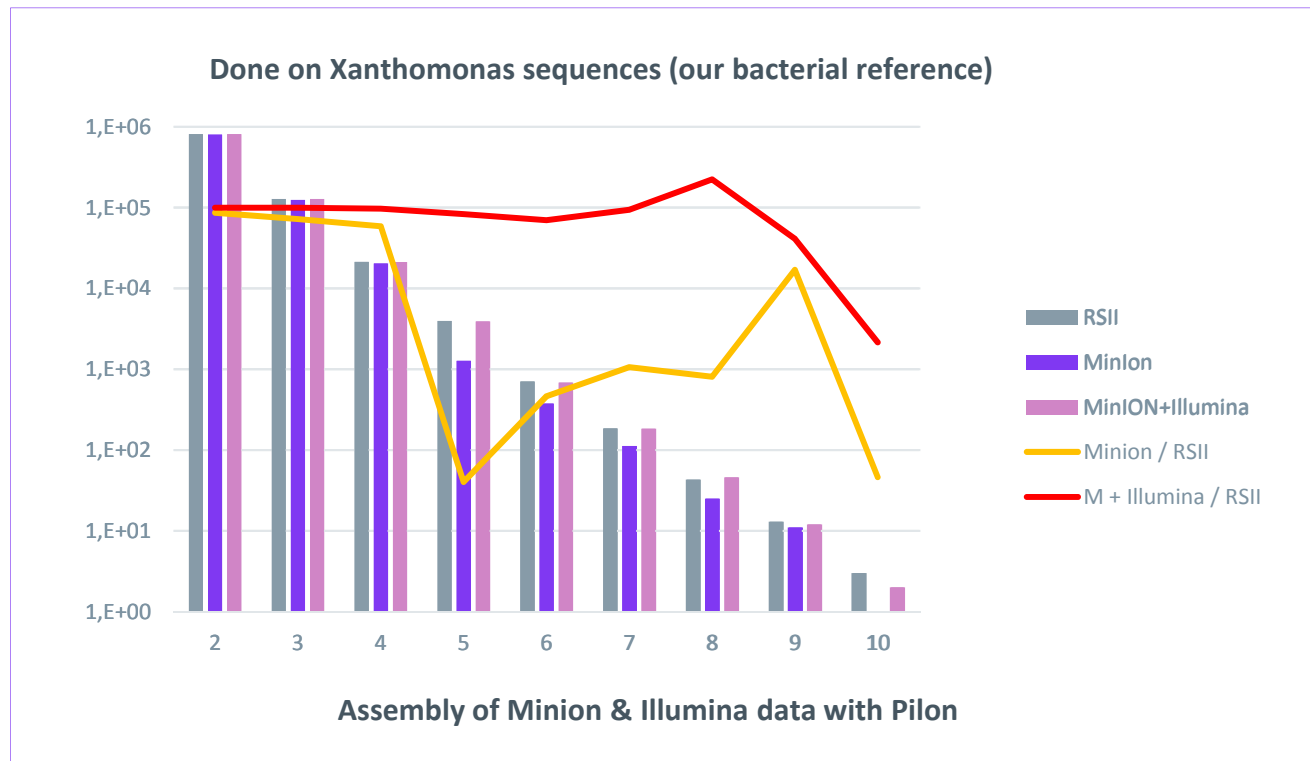
Yield might be correlated to the dispersion of Fragment size

Efficiency of Barcoding (January version of the kit)

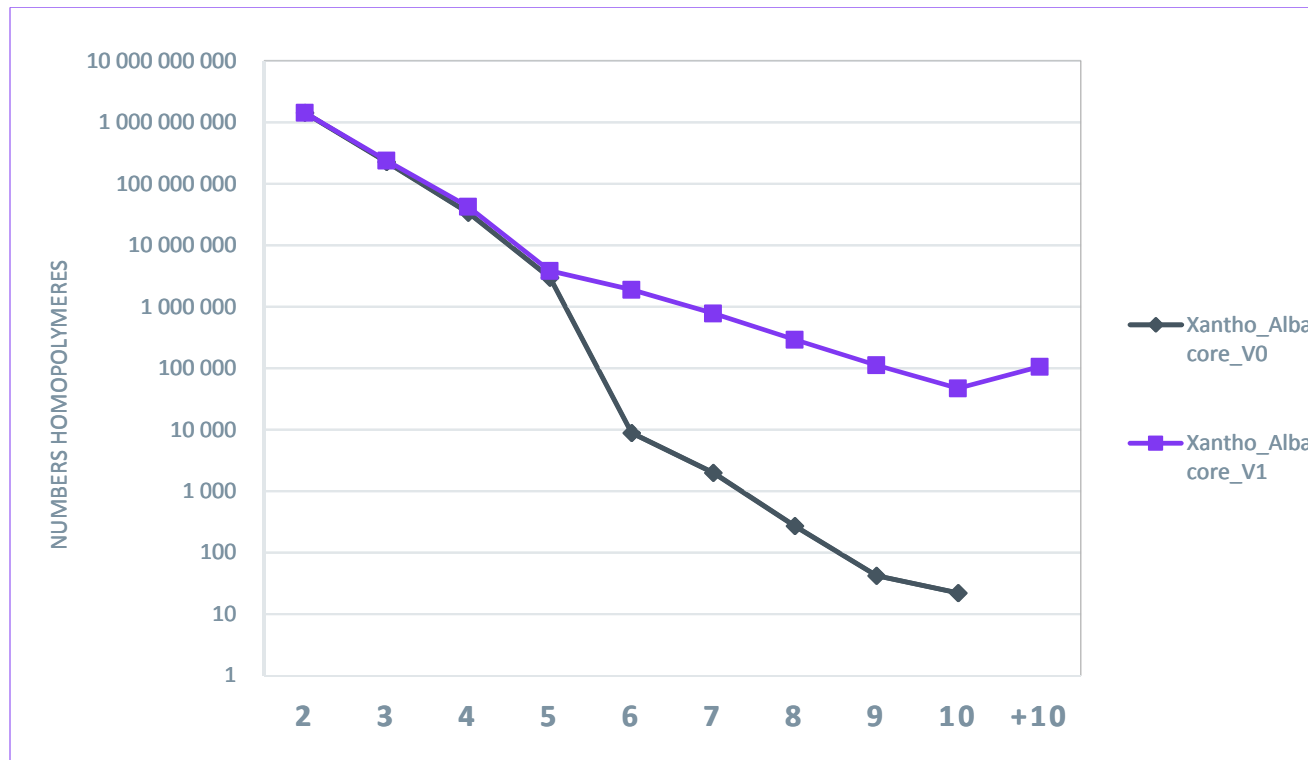


Barcoding is efficient (at least at that level) with homogenous results

Detection of homopolymers



Improvement of Homopolymere sequencing with Albacore v1



Analysis of sequence accuracy on Xanthomonas data set



BUSCO2 : Identification of 148 conserved genes in assembly genome of E Coli

	% Complete genes	% Fragmented genes	% Missing genes
X_pacbio_hgap3	95 %	0 %	5 %
X-Av1-F10_minion_canu	17 %	36 %	47 %
X-Av1-F10_minon_canu_illumina_pilon	95 %	0 %	5 %
X-1D_Metrichor_minion_canu_nanopolish	72 %	15 %	13 %



Poster 6



Sequencing of a local strain of Arabidopsis Thaliana



Sequencing on MinION at 80 x
90 % of the sequence in 20 contigs



Baptiste Mayjonade, Fabrice Roux, Jérôme Gouzy

NUM	187
MIN	5257
MAX	15 602 179
N50 BP	7 697 404
N50 NUM	6
N90 BP	1 036 635
N90 NUM	20
MEAN	644 576
MEDIAN	35 272
BP	120 535 815



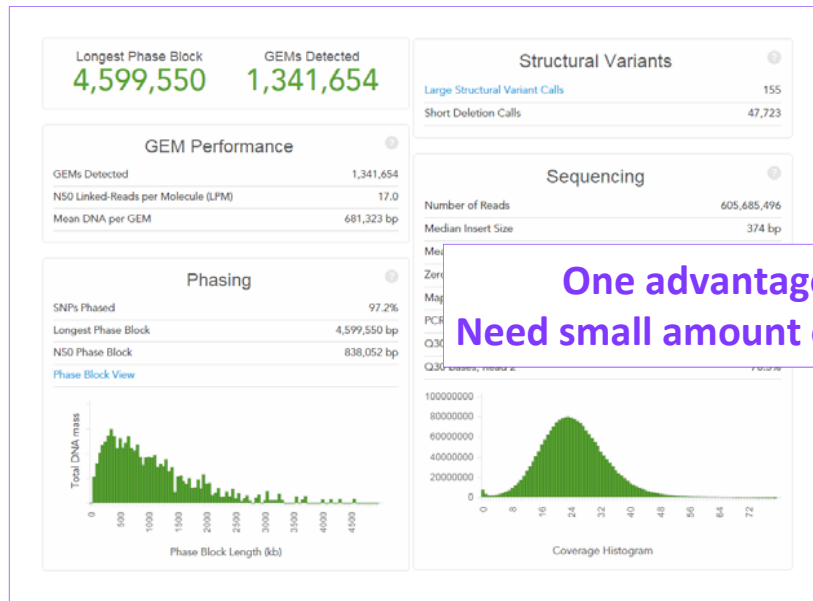
3

Other technologies

Contribution of Chromium from 10Xgenomics to genome sequencing

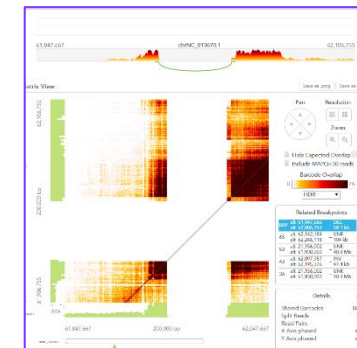


Pilot projects on Rabbit for phasing



20 % of representation

One advantage
Need small amount of DNA



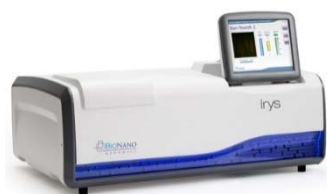
Structural Variants detection

Longest phase block 4.6 Mb

Assembly of Tomato genome using combination of technologies



Chromium (10 X Genomics)



Irys (Bionano)

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

Pilot projects on Tomato for de novo assembly

Technology	N50
Chromium + Illumina (100x)	1.8 Mb
+ Bionano (2 enzymes)	17 Mb



GBF Lab : **Mohammed Zouine**, Pierre Frasse, Mondher Bouzayen
 CNRGV lab (Bionano) : **Sandrine Arribat**, William Marrande, **Hélène Bergès**



4

Conclusions

Conclusions



Oxford Nanopore has a great potential

- Very long reads
- A simple possibility to provide the technology at various scales (from SmidgION to PromethION)
- A potential to distinguish modified bases
- A potential to challenge Illumina for the Cost / Gb

Different technologies available for long reads

- PacBio : A golden standard up to now for long reads (no bias in sequence errors)
- 10x : Low amount of DNA, Illumina accuracy, low cost in complement of Illumina
- ONT : A greater potential of evolution, but up to now a less mature technology

Key points

- Ultra long DNA is required, when possible in large amount
- The infrastructure for the analysis of data should not be underestimated
- An interest to combine different technologies depending on applications, requested specifications & money available

Thanks to ...



INRA GeT Platform :



Cécile Donnadieu, Catherine Zanchetta, Pauline Heuillard, Olivier Bouchez *et al*
Maxime Manno, Claire Kuchly, Céline Vandecasteele

INRA Bioinfo platform :



Christophe Klopp, Christine Gaspin

INRA CNRGV :



Sandrine Arribat, William Marrande, Caroline Callot,
Stéphane Cauet, Hélène Bergès

INRA/CNRS LIPM :



Baptiste Mayjonade, Stéphane Munoz, Nicolas Anglade,
Fabrice Roux, Jérôme Guozy

INRA/ENSAT : GBF :



Mohammed Zouine, Pierre Frasse, Mondher Bouzayen

INRA GenPhySE :



Julie Demars, Isabelle Hochu

ENV/INRA :



Guillaume Croville, Guillaume Le Loc'h, Mattias Delpont, and Jean-Luc Guérin

Present at London Calling



- Cécile Donnadieu (Head of GeT)

Poster 6



- Catherine Zanchetta (Minion Developments)

Poster 6



- Jean-Luc Guérin (PI Virology project)

Poster 22

@NanoporeConf | #NanoporeConf

© 2017 Oxford Nanopore Technologies. All rights reserved. The MinION is for research use only.



LC
2017
LONDON
CALLING

@NanoporeConf | #NanoporeConf

© 2017 Oxford Nanopore Technologies. All rights reserved. The MinION is for research use only.

