GenoToul Bioinfo

The GenoToul Bioinfo facilitie

https://bioinfo.genotoul.fr/

Département MathNum

Claire Hoede

MIA TOULOUSE

Plateforme Bioinfo Occitanie Toulouse

Genotoul Bioinfo

Members

- ~9 FTE
  - including 2.3 system administator
  - 2 CDDs
  - We are looking for 1 CDD IR admin sys / devops :
    - https://jobs.inrae.fr/ot-16752

Plateforme Bioinfo Occitanie Toulouse

Context

- Member of IFB
- Member of GIS Genotoul
- Associated platform France Génomique
  - special partnership with Get-plage
- Member of IR bioinfomics
- CATI Bios4biol

Plateforme Bioinfo Occitanie Toulouse

Our missions

- Providing an infrastructure and environment (support, softwares, etc.) for bioinformatics : https://bioinfo.genotoul.fr/index.php/resources-2/resources/

Projet cofinancé par le Fonds Européen de Développement Régional
Financement dans le cadre de la réponse de l'Union à la pandémie de COVID-19

- Supporting biologists' bioinformatics projects (ANR, PEPR or shorter projects) ~ 25 projects / year

- Train biologists (108 day.trainee in 2022)

Plateforme Bioinfo Occitanie Toulouse

Focus on a datapaper

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A *Bos taurus* sequencing methods benchmark for assembly, haplotyping, and variant calling

Camille Eché[1], Carole Iampietro[1], Clément Birbes[2], Andreea Dréau[2], Claire Kuchly[1], Arnaud Di Franco[2], Christophe Klopp[2], Thomas Faraut[3], Sarah Djebali[3,4], Adrien Castinel[1], Matthias Zytnicki[5], Erwan Denis[1], Mekki Boussaha[6], Cécile Grohs[6], Didier Boichard[6], Christine Gaspin[2,5], Denis Milan[1,3] & Cécile Donnadieu[1]✉

https://www.nature.com/articles/s41597-023-02249-1

L'EUROPE S'ENGAGE L'OCCITANIE AGIT

UNION EUROPÉENNE

Projet cofinancé par le Fonds Européen de Développement Régional

Different sequencing technologies

1 Sample	2 Sequencing technology	3 Library type	4 Dataset
Father	Oxford Nanopore Illumina	Ligation sequencing gDNA 10X Genomics Chromium Hi-C	a,b,c
Mother	Oxford Nanopore Illumina PacBio	Ligation sequencing gDNA 10X Genomics Chromium Hi-C Circular Long Read	a,b,c
Heifer	Oxford Nanopore Illumina PacBio	Ligation sequencing gDNA 10X Genomics Chromium PCR Free 2x250pb Hi-C Circular Long Read Consensus Long Read	a, b, c, d

Fig. 1 Technologies used for the study. The parents and the heifer were sequenced with Oxford Nanopore Technologies on GridION and PromethION, Chromium 10X and the Hi-C method on MiSeq, HiSeq or NovaSeq 6000. The heifer was additionally sequenced with Illumina 2 × 250 bp on NovaSeq 6000 and PacBio Sequel II (CLR and CCS (i.e HiFi reads) mode). For the Trio approach, parent reads (2 × 150bp) from 10X Genomics Chromium datas were used.

Projet cofinancé par le Fonds Européen de Développement Régional

HiFi data yields a genome 500Mpb larger than the reference genome

	ARS-UCD1.2	ONT wtdbg2	10X supernova	HiFi hifiasm	CLR wtdbg2
Pipeline used		a	b	a	a
Data type	CLR	ONT	10X Chromium	CCS	CLR
Quantity	80X	58X	95X	40X	43X
Assembler	Falcon	Wtdbg2	Supernova	Hifiasm	Wtdbg2
Number of contigs	3 077	7 226	26 306	1 444	2 857
Total size	2 700 000 000	2 701 288 401	2 627 892 463	3 244 632 679	2 631 921 359
N50 contigs length	12 000 000	23 641 545	488 571	84 059 894	16 542 341
BUSCO	95.7%*	C:70.2%	C:94.7%	C:95.9%	C:90.0%
Inspector QV **		22.29	27.09	47.25	25.79

Table 2. Summary of heifer produced contigs assemblies. For details about pipeline used in this study, refer to Fig. 2. *BUSCO analysis was performed on polished contigs, **Inspector Quality Value is calculated on reference alignment andreads alignment.

Projet cofinancé par le Fonds Européen de Développement Régional

Workflows used

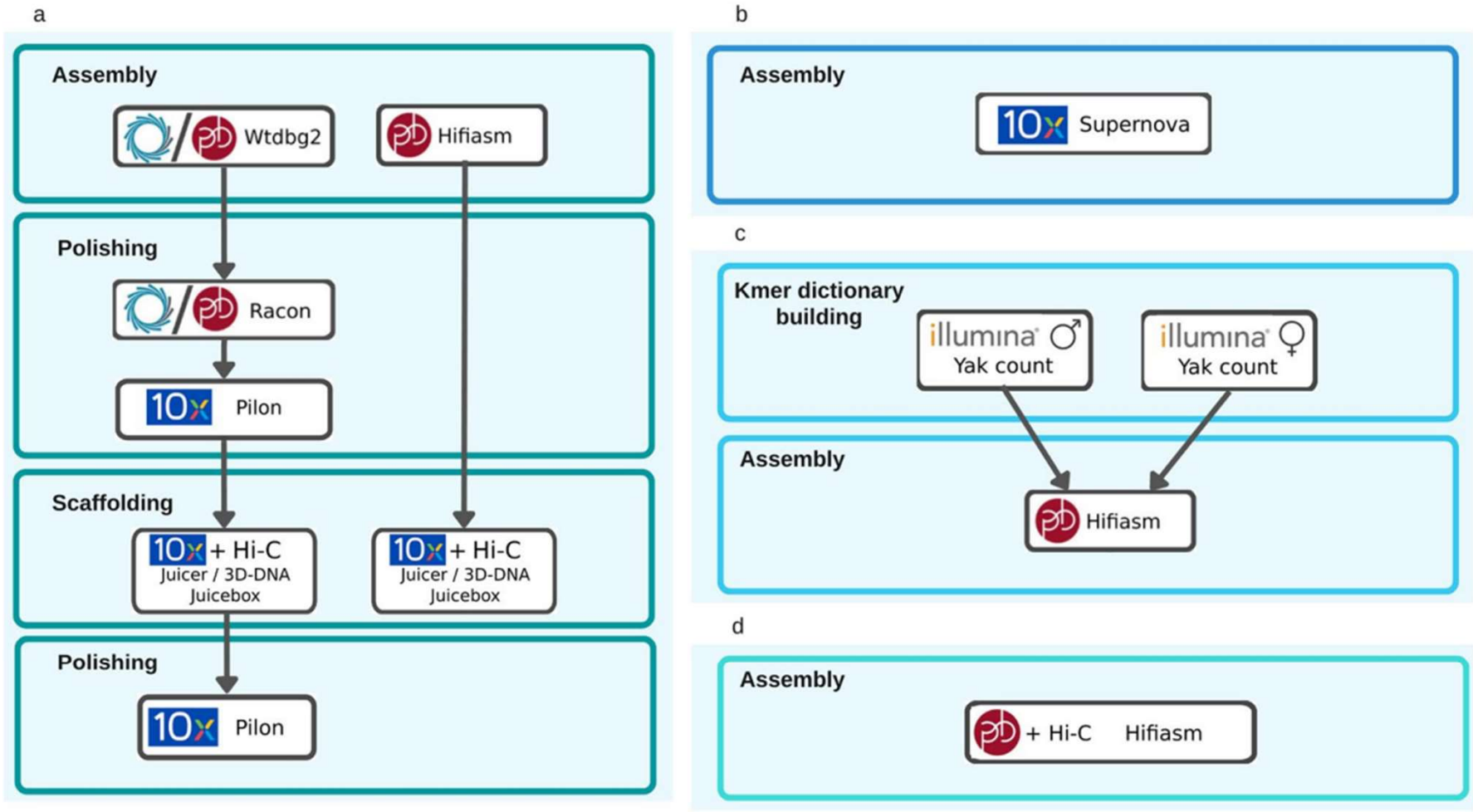

Fig. 2 Details of the 5 pipelines used to produce our assemblies. a-Long reads assemblies from Oxford Nanopore Technologies and Pacific Biosciences followed by polishing step for erroneous assemblies and scaffolding step. b-10X Chromium assembly and scaffolding with Supernova. c-Phased assembly with HiFi and parental illumina reads. d-Phased assembly with HiFi and and Hi-C data.

L'EUROPE S'ENGAGE L'OCCITANIE AGIT
UNION EUROPÉENNE
La Région Occitanie Pyrénées · Méditerranée
Projet cofinancé par le Fonds Européen de Développement Régional

Thanks for your attention

Do you have any questions ?