



Capra hircus / Ovis aries Axis1 - Assembly

Clément Birbes Christophe Klopp Claire Kuchly









Assembly production

Capra hircus:

- 7 CLR:

- 7 Sequenced
- 7 Assembled and scaffolded
- 7 Available assemblies

- 1 Phased:

- 1 Sequenced
- 1 Assembled
- Phasing validation

Ovis aries:

- 9 CLR:

- 9 Sequenced
- 7 Assembled and scaffolded
- 7 Available assemblies

- 1 Phased:

- 1 Sequenced
- 1 Assembled
- Phasing validation







Seqoccin Q20+ Assembly (Test on Bos taurus)



	Flye Q20+	Hifiasm Q20+	Flye CLR	Wtdbg2 CLR	Flye ONT	Wtdbg2 ONT	Flye Hifi	Hifiasm HiFi
Coverage	30X	30X	33X	33X	58X	58X	41X	41X
Number of scaffolds	2 398	17150	3 466	2 857	2 744	7 226	18159	905
Total size (Gb)	2.70	2.22	2.66	2.63	2.68	2.70	3.44	3.21
Longest contig (Mb)	76.9	4.8	70.5	56.6	95.23	70.7	11.5	167.7
N50 scaffold length (Mb)	28.5	0.17	19.6	16.5	31.4	23.7	0.75	86.9
L50 scaffold count	30	3570	38	46	26	35	948	15
BUSCO	95.8%	56.9%	95.0%	90.0%	85.7%	70.2%	93.5%	96.0%

- Q20+ hifiasm : No good results so far

- Q20+ flye: Results as good as CLR flye + correction CLR

ONT flye + correction short reads

- Q20+: Good consensus sequences quality with flye (HiFi equivalent) Missing repeated parts (ONT / CLR equivalent)

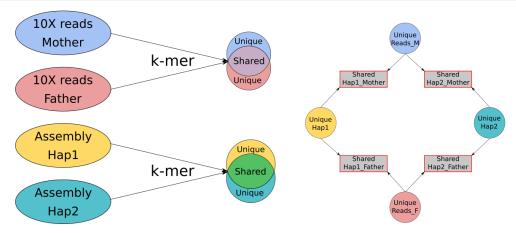






Seqoccin Phased assembly (Test on Bos taurus)

	ARS- UCD1.2	Hifiasm consensus	Parent-HiFi Hap1	Parent-HiFi Hap2	HiC-HiFi Hap1	HiC-HiFi Hap2
Number of contigs	3 077	1 444	2 871	2 300	2 658	2 136
Total size (Gb)	2.7	3.2	3.2	3.1	3.1	3.2
N50 contigs length (Mb)	12	84.1	71.6	69.2	80.1	71.6
BUSCO	95.7 %	95.9 %	95.8 %	95.3 %	95.8 %	95.7 %
Phasing ratio			97.3 %	96.7 %	62.6 %	60.5 %
Contigs phasing ratio			97.5 %	96.9 %	84.6 %	85.6 %













SeqOccIn

Phasing assemblies from SNPs (Test on Bos taurus)

- Use hapcut2 (&LongShot + WhatsHap) to phase assemblies
- No additional sequencing data
- Work with CLR, CCS, Nanopore

	ADC LICDA 2	l lificom component	lla-a1	llo2
	ARS-UCD1.2	Hifiasm consensus	Hap1	Hap2
Number of contigs	3 077	1 444	3 773	3 009
Total size (Gb)	2.7	3.2	3.5	3.6
N50 contigs length (Mb)	12	84.1	63.2	68.5
BUSCO	95.7 %	95.9 %	95.8 %	95.7 %
Phasing ratio			51.3 %	51.7 %
Number of contigs Total size (Gb) N50 contigs length (Mb) BUSCO Phasing ratio Contigs phasing ratio			60.6 %	59.4 %

- First results (from Arnaud) give good metrics but bad phasing
- Test on new Hapcut2 version











Partenaires du projet

Coordination

Cécile Donnadieu **Christine Gaspin** Carole lampietro Denis Milan





































syngenta MAÏSADOUR





Axe1 Génomique

Clément Birbes Arnaud Di-Franco Andreea Dréau Camille Eché **Thomas Faraut** Carole lampietro Christophe Klopp Claire Kuchly Camille Marcuzzo **Amandine Suin** Matthias Zytnicki

Axe2 Epigénétique

Remy Félix Serres Paul Terzian Celine Vandecasteele Christophe Klopp

Axe3 Métagénomique

Adrien Castinel Jean Mainguy Olivier Bouchez Géraldine Pascal Claire Hoede

Axe 4 Transversal

Amandine Broha Abdias-Archimede Towe-Patipe Erwan Denis Romain Therville Didier Laborie Céline Noirot Gérald Salin Marie-Stéphane Trotard

GenPhySE

Julie Demars Cédric Cabau **Sylvie Combes Patrice Dehais Thomas Faraut** Katia Feve Nathalie Iannucelli Sophie Leroux Géraldine Pascal Frédérique Pitel

MIAT

Matthias Zytnicki

GABI

Didier Boichard Mekki Boussaha Sébastien Fritz Cécile Grohs

Le Moulon

Alain Charcosset Johann Joets Delphine Madur Stéphane Nicolas Rémi Séraphin Clémentine Vitte











































