



HAL
open science

TAPS for Multimodal Epigenomic Profiling in Livestock A Comparison with ONT, WGBS, and EM-Seq

Céline Vandecasteele, Julie Demars, Sophie Leroux, Nathalie Iannuccelli, Carole Iampietro, Craig Marshall, Travis Sanders, Aaron Garnett, Martin Ranik, Kristina Giorda, et al.

► To cite this version:

Céline Vandecasteele, Julie Demars, Sophie Leroux, Nathalie Iannuccelli, Carole Iampietro, et al.. TAPS for Multimodal Epigenomic Profiling in Livestock A Comparison with ONT, WGBS, and EM-Seq. Plant and animal genome conference 32, Jan 2025, San Diego, United States. 2025. <hal-04911060>

HAL Id: hal-04911060

<https://hal.inrae.fr/hal-04911060v1>

Submitted on 24 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

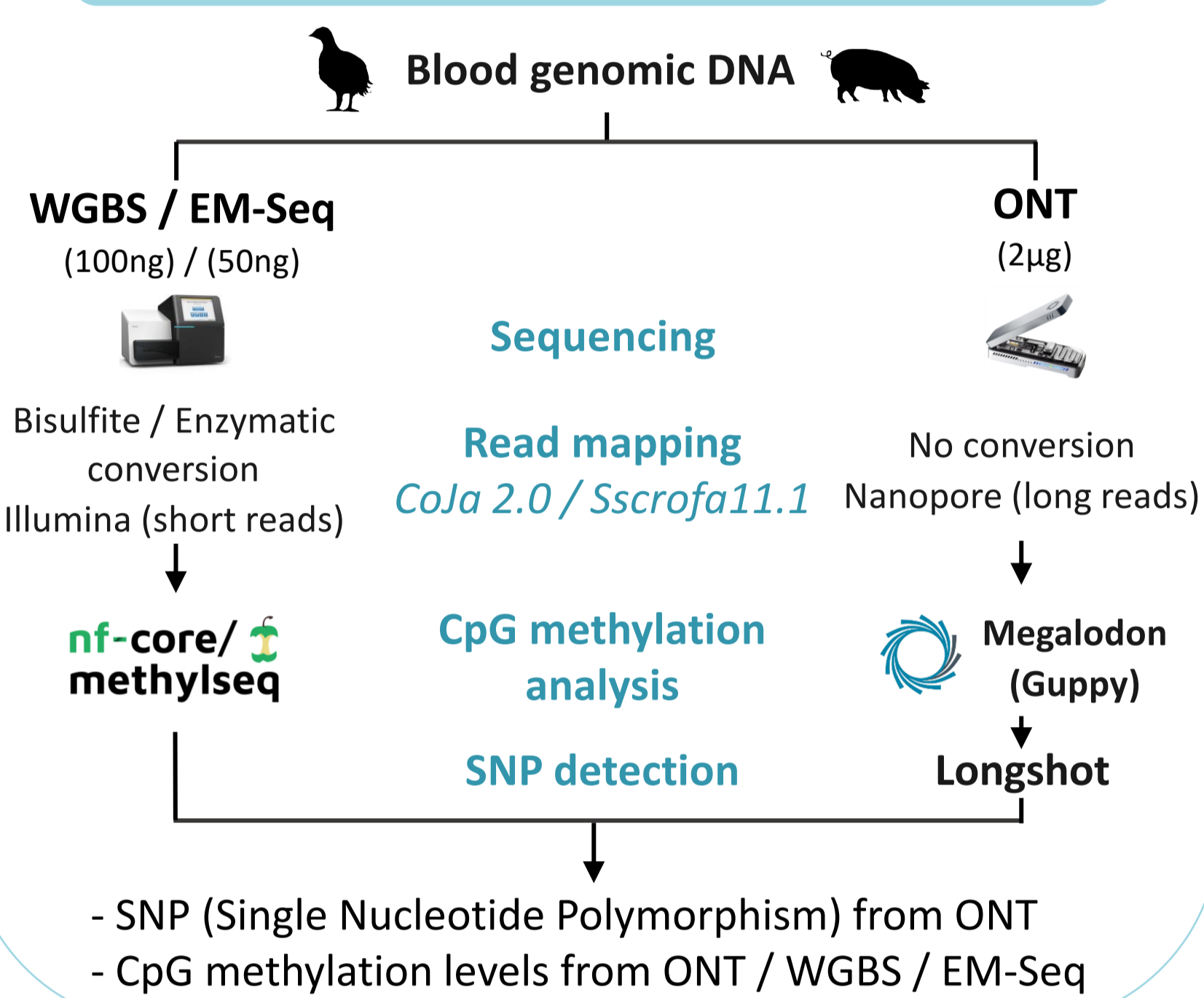
Céline Vandecasteele¹, Julie Demars², Sophie Leroux², Nathalie Iannuccelli², Carole Iampietro¹, Craig Marshall³, Travis Sanders³, Aaron Garnett³, Martin Ranik³, Kristina Giorda³, Eduard Casas³, Thomas Harrison³, Doug Wendel³, Brian Kudlow³, Frédérique Pitel²

¹ INRAE, US 1426, GeT-PlaGe, Genotoul, France Genomique, Université Fédérale de Toulouse, Castanet-Tolosan, France
² GenPhySE, Université de Toulouse, INRAE, ENVT, 31326, Castanet-Tolosan, France
³ Watchmaker Genomics, Boulder, Colorado, USA

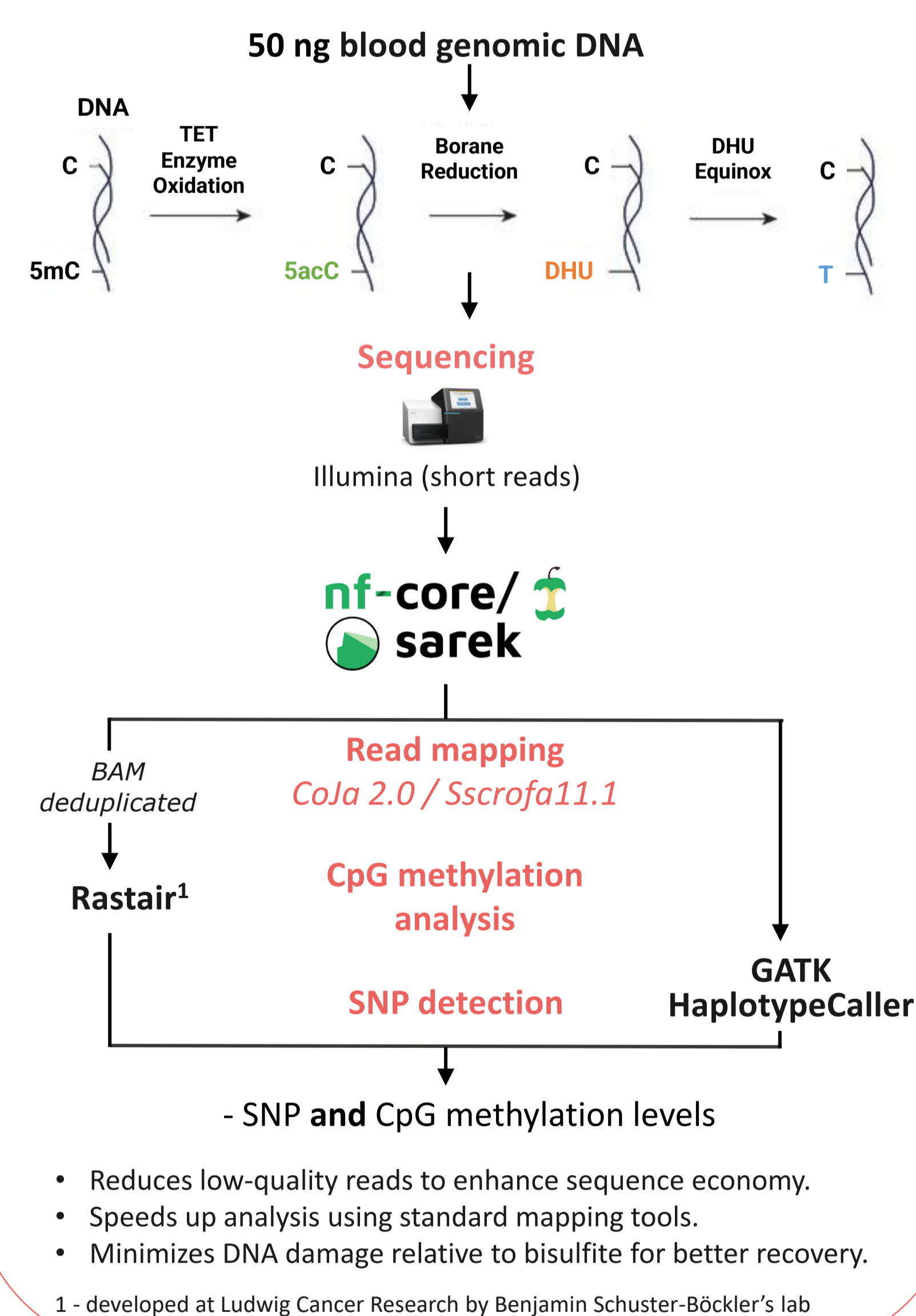
Introduction

Genetic and genomic selection advances have significantly optimized livestock performance over the past decades. However, genetics accounts for only a portion of the observed phenotypic variability, while much of this variability—often attributed to environmental factors—remains inaccessible to genetic approaches. One of the questions in breeding today is thus whether integrating epigenetic markers into prediction models could improve their accuracy or even whether epigenetics should be included in genetic evaluation models for breeding candidates. The availability of a large-scale epigenotyping tool is crucial for answering this question and eventually implementing epigenotyping in animal breeding. Here we evaluated CpG methylation detection using an advanced prototype of TAPS (TET-assisted pyridine borane sequencing) against existing data from Whole Genome Bisulfite Sequencing (WGBS), EM-seq (Enzymatic Methyl-seq), and ONT (Oxford Nanopore Technology) in two farm species: pigs (*Sus scrofa*) and quails (*Coturnix japonica*).

Standard methods



TAPS



Conclusion

This pilot study provides very encouraging results for the use of TAPS in epi/genotyping analyses in farm animals, with robust sequencing and mapping metrics and comparable CpG methylation profiles and SNP genotypes between all technologies.

Available data sets / TAPS results

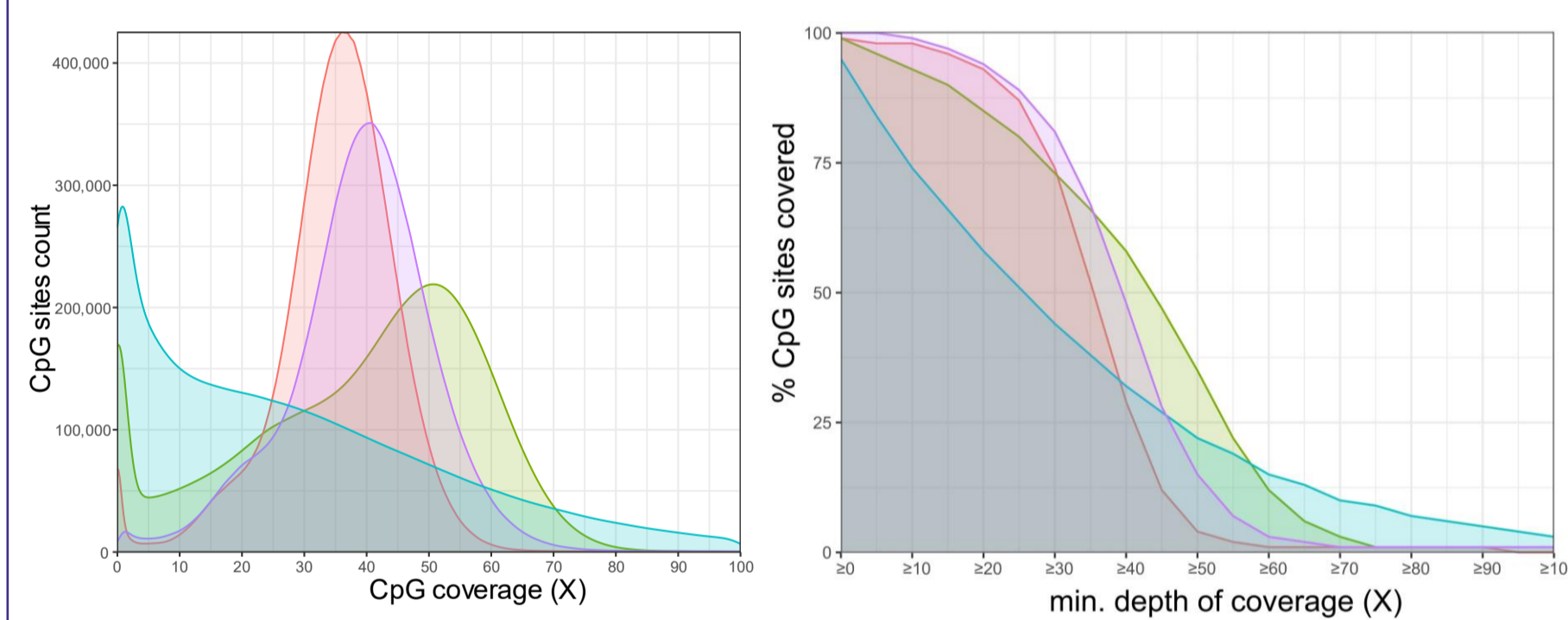
		WGBS	EM-Seq	ONT	MicroArray	TAPS
<i>Sus scrofa</i>	Sample 1	ND	148 Gb	192 Gb	650 k Array	287 Gb*
	Sample 2	238 Gb	ND	152 Gb	650 k Array	269 Gb*
<i>Coturnix japonica</i>	Sample 1	119 Gb	ND	55 Gb	ND	94 Gb*
	Sample 2	283 Gb	90 Gb	65 Gb	ND	94 Gb*

*3 replicates

Results

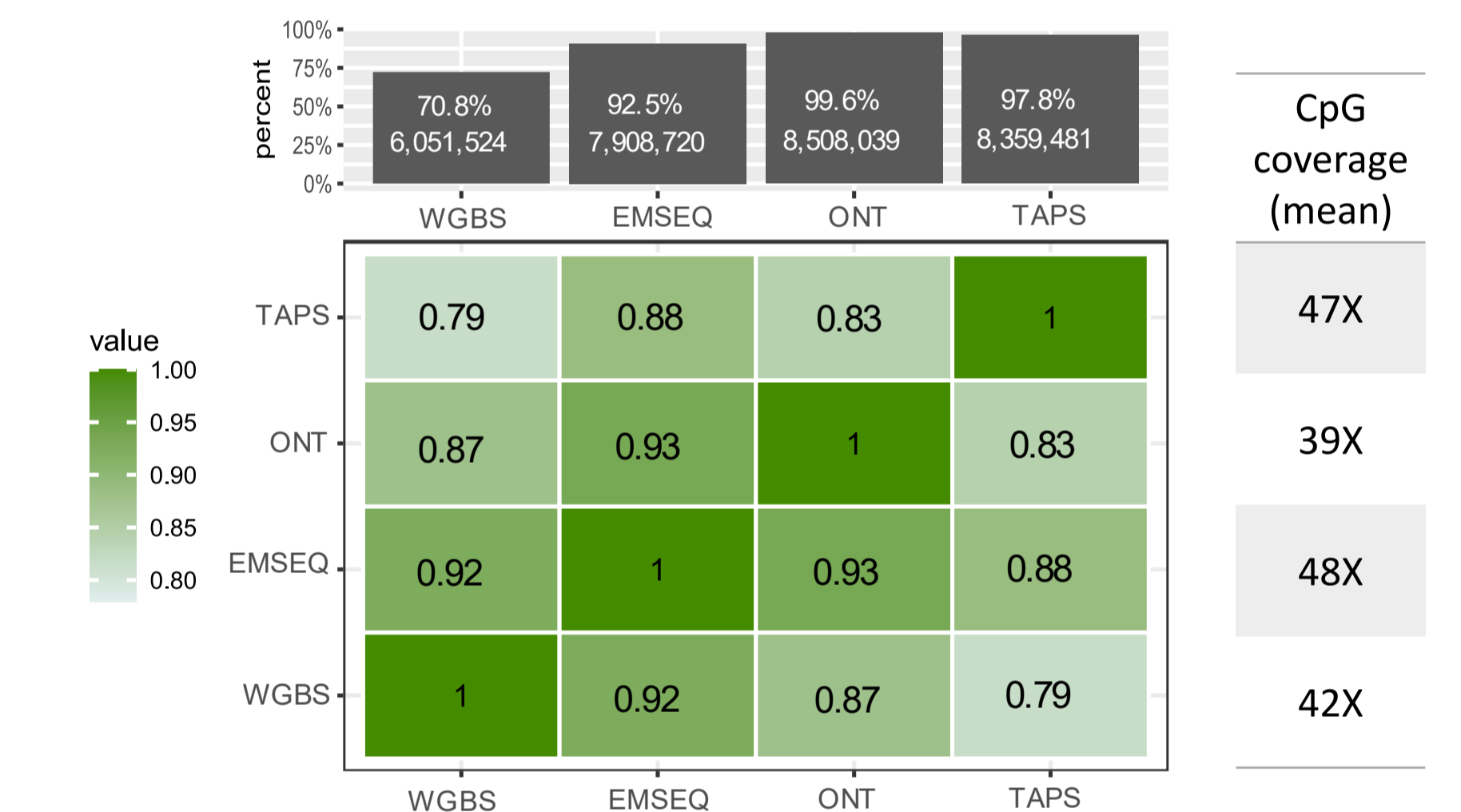
Methylation analyses (Quail sample 2)

At the 10X threshold, the WGBS approach covers less than 75% of CpG sites, while EM-Seq covers around 90%, and both ONT and TAPS achieve nearly 100% coverage.



Coverage of CpG sites across technologies

The distribution of CpG site coverage and the percentage of CpG sites covered at a minimum depth of coverage



Pearson correlation matrix of CpG methylation level per-site across technologies (Filtering data: 10x < depth < 100x). Barplots above the matrix show the number of CpG sites covered by 10 or more reads and their percentage compared to the reference genome. Average CpG coverage by technology before depth filtering is shown on the right.

SNP analyses (pigs)

Filtering variants post genotyping with stringent criteria

- 1 - Keep only variants present on 650K for SSC1 to SSC18, n=524,117 variants
- 2 - Remove duplicates corresponding to CpG identified with rastair, n=453,192 variants
- 3 - Keep only bi allelic SNPs, n=350,011 variants
- 4 - Keep only SNPs QUAL>300 & <1300, n=155,285 variants
- 5 - Keep only SNPs with meanDP >15 across samples, n=129,211 variants

AB72734224	ONT	Homoz Ref	Hetero	Homoz Alt
TAPS	67871	1009	232	
Homoz Ref	52	55790	2810	
Hetero	1	20	1140	
Homoz Alt				

AB72734224	650K	Homoz Ref	Hetero	Homoz Alt
TAPS	68358	714	128	
Homoz Ref	192	55570	2992	
Hetero	1	20	1143	
Homoz Alt				

AB72734233	ONT	Homoz Ref	Hetero	Homoz Alt
TAPS	27802	1767	98	
Homoz Ref	81	96344	2217	
Hetero	1	78	537	
Homoz Alt				

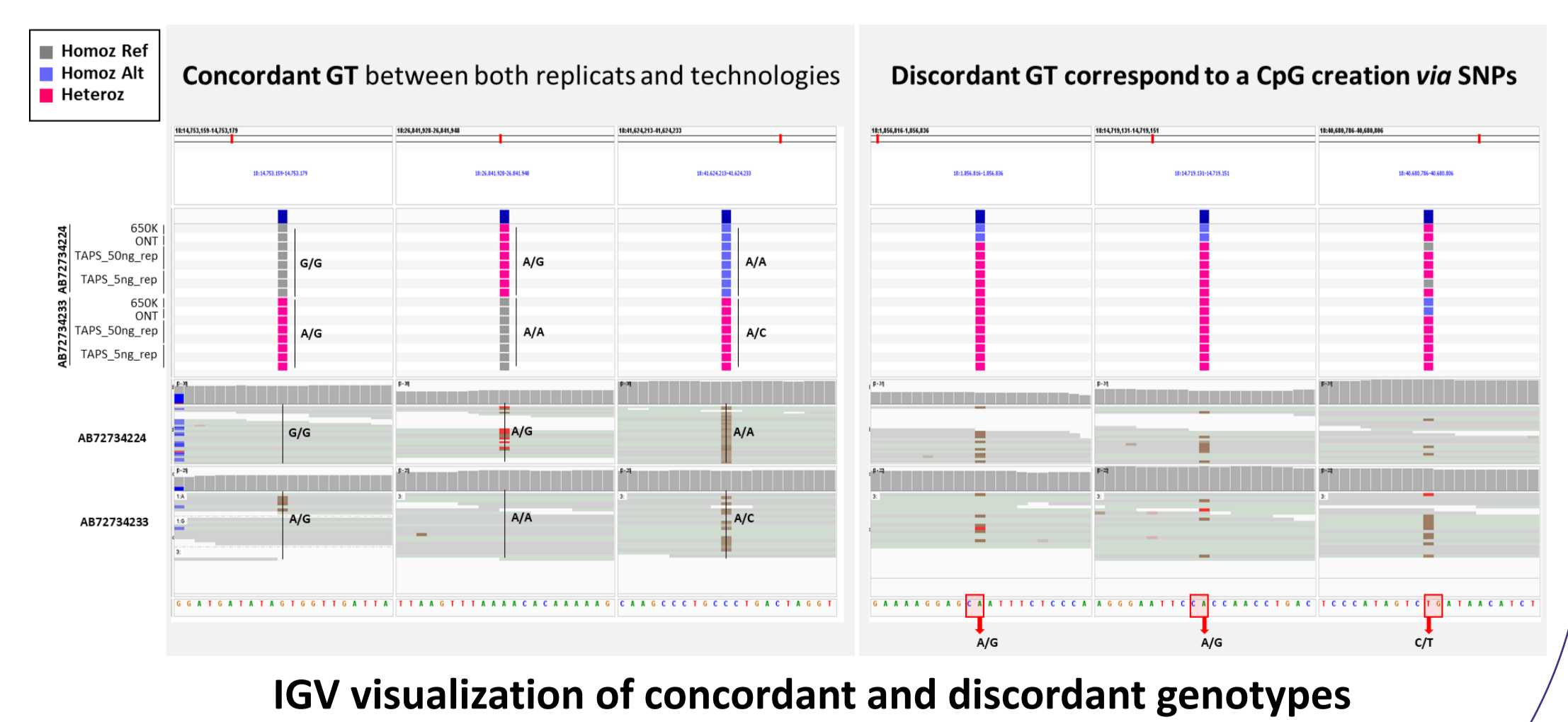
AB72734233	650K	Homoz Ref	Hetero	Homoz Alt
TAPS	28173	1480	43	
Homoz Ref	1186	96645	2669	
Hetero	3	77	537	
Homoz Alt				

Genotypes comparisons

The excess of heterozygous in TAPS libraries can be explained by *de novo* SNV-driven CpGs

	Replicates	ONT vs. TAPS	650K vs. TAPS
AB72734224_A	0.9889	0.9680	0.9686
AB72734224_B	0.9867	0.9678	0.9684
AB72734224_C	0.9870	0.9663	0.9670
AB72734233_A	0.9773	0.9671	0.9576
AB72734233_B	0.9884	0.9681	0.9585
AB72734233_C	0.9807	0.9709	0.9614

Genotypes correlations between TAPS libraries replicates and across technologies



→ A custom script was developed to resolve variant calling complexities associated with *de novo* SNV-driven CpGs. Contact Support@watchmakergenomics.com for details.