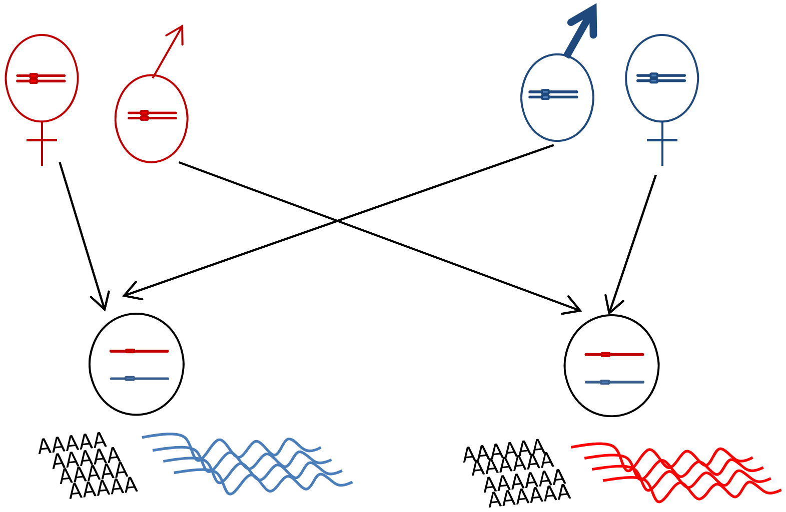
Analyse de méthylation de l'ADN dans l'embryon de poulet par séquençage bisulfite

Frédérique Pitel – Céline Noirot

Experimental design enabling to detect parent-of-origin dependent expression

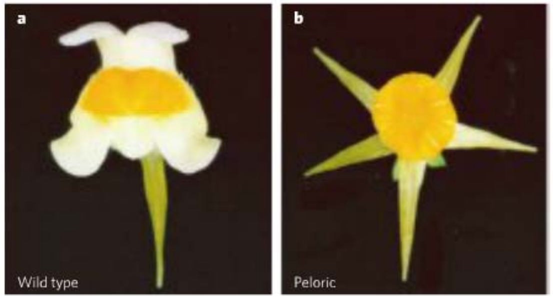
- Parental SNPs to deduce origin of the allele

- Reciprocal cross
  - 12 embryos (E 4.5) from each cross (6 males and 6 females)
  - RNASeq



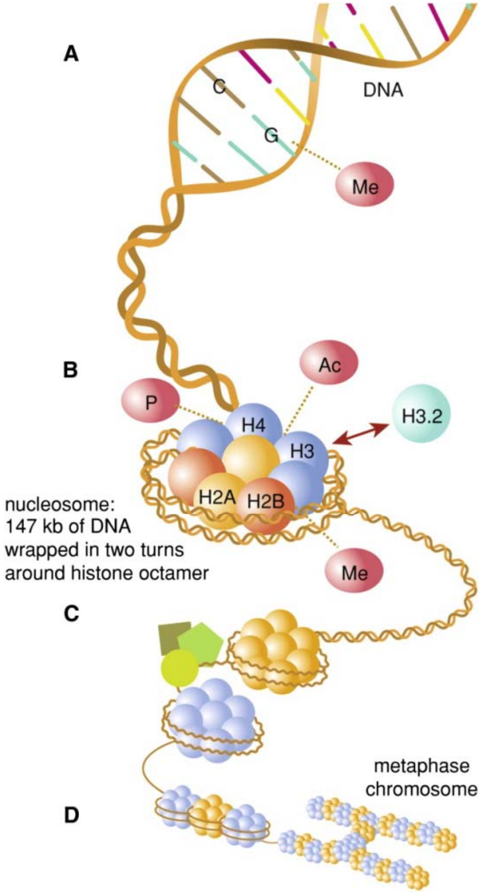
- Search for regions with transcripts presenting a bias in expression
  - Imprinting ?
  - Allele Specific Expression

Methylation analyses



Cubas et al, 1999 (Bird 2007)

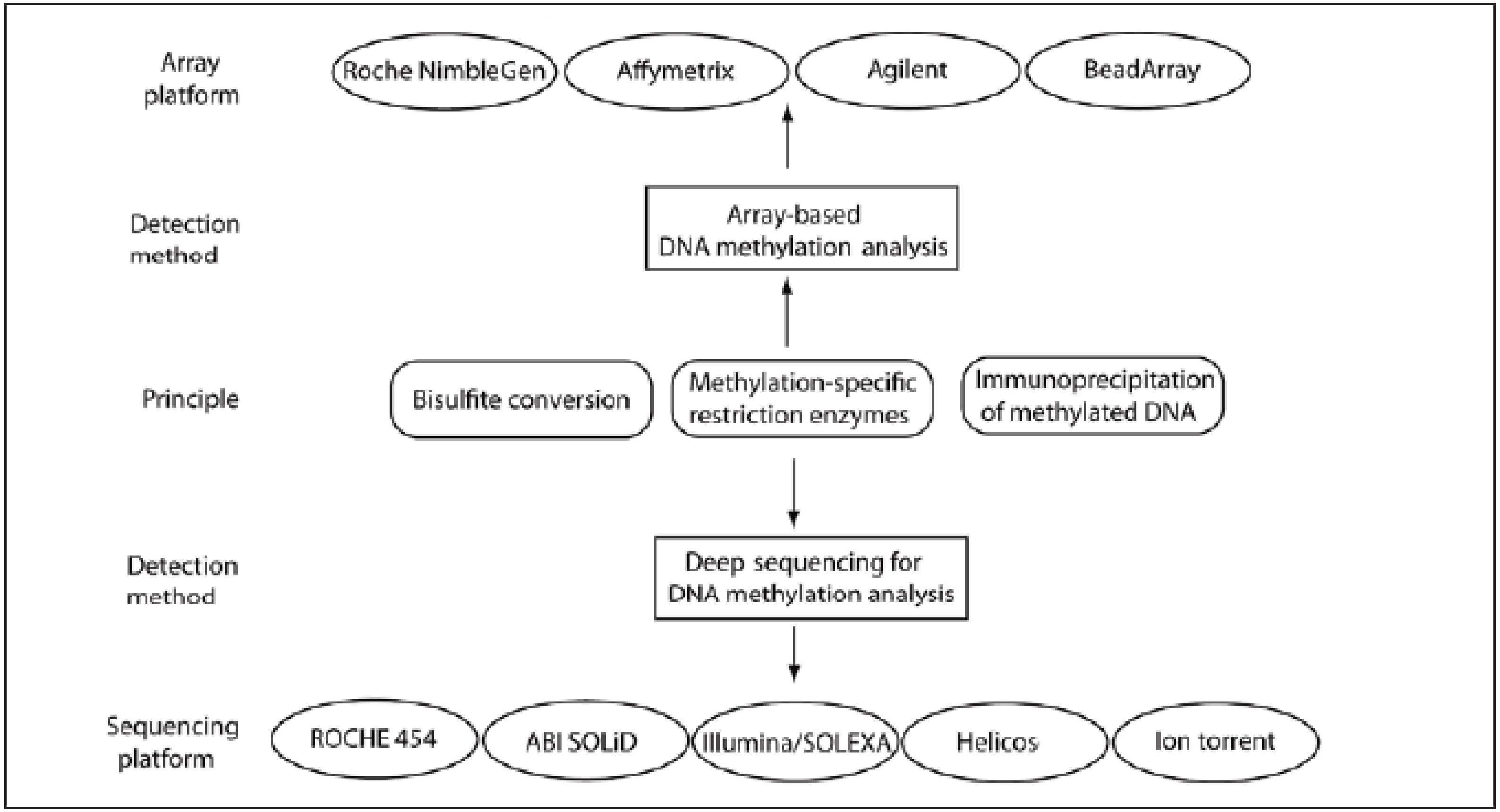
Morgan et al, 1999



Denk & McMahon 2012

- Genome-wide investigation of DNA methylation at single-base resolution in the chicken embryo
    - Methylome description
    - Comparison expression / methylation level

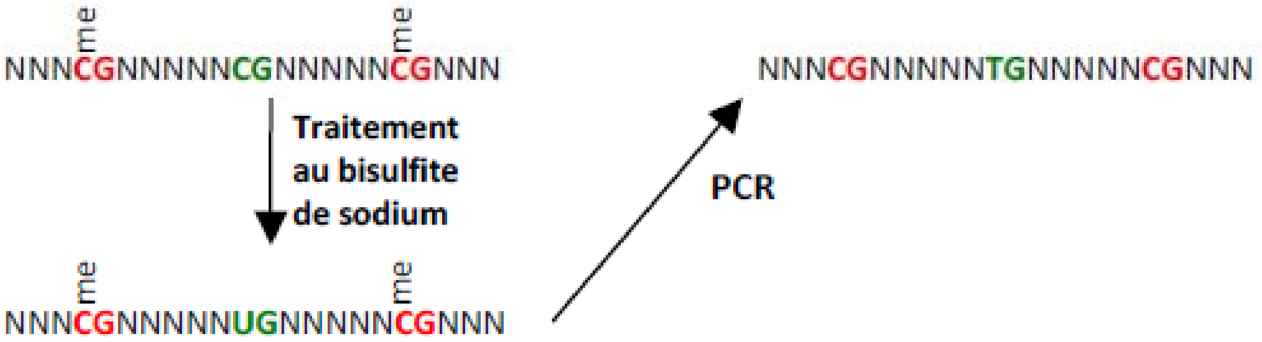
Technologies for methylation studies



Bisulfite Sequencing (BS-seq)

- Two embryos as a test (RNA + DNA Seq)
  - NextFlex bisulfite sequencing kit (BIOO Scientific Corp)
  - Control DNA: pUC19 DNA (Zymo Research), methylated or not

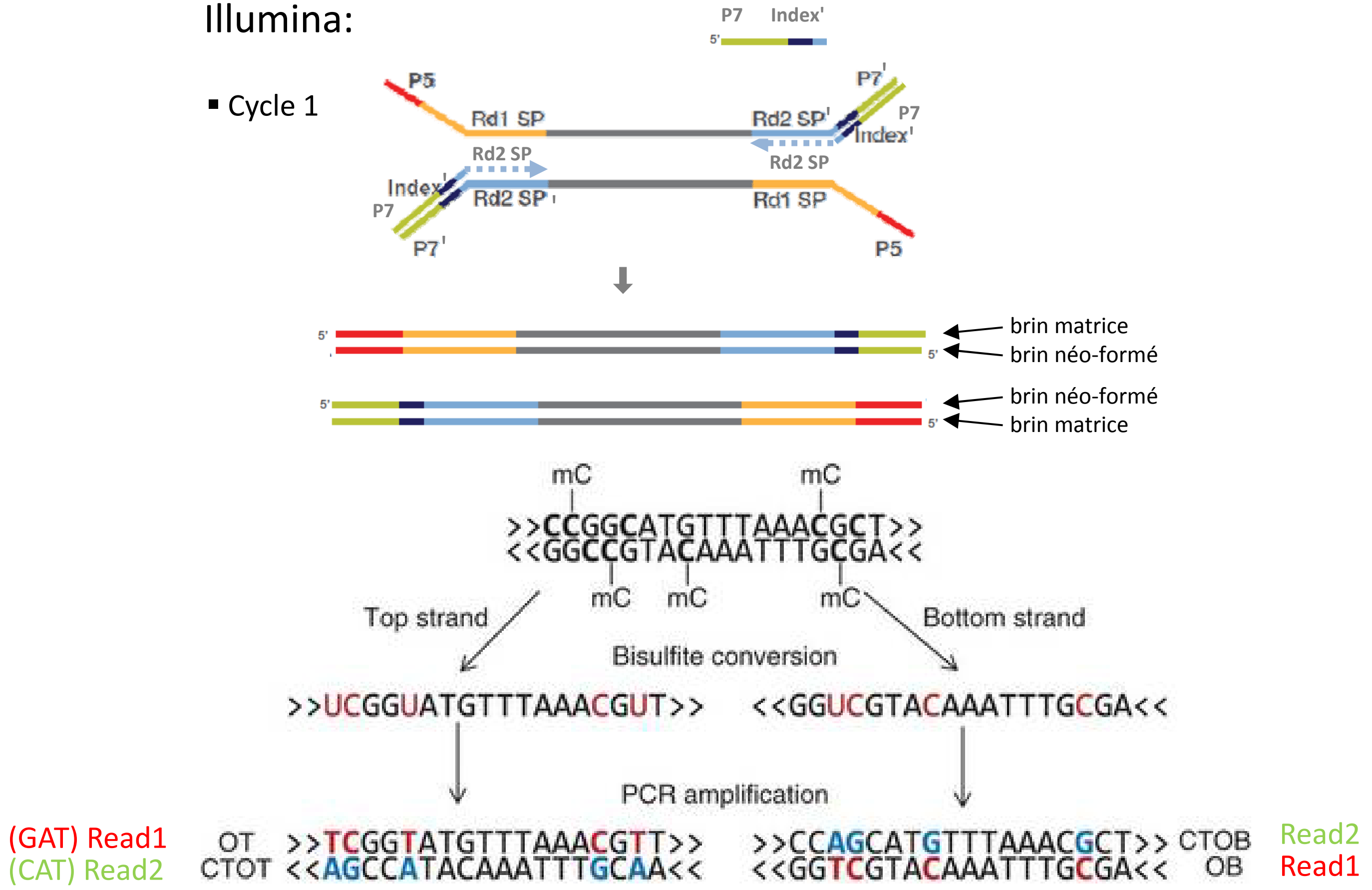
- Methylated C not converted by bisulfite



Methylation analysis (adapters ligation – conversion – PCR)

Illumina:

- Cycle 1



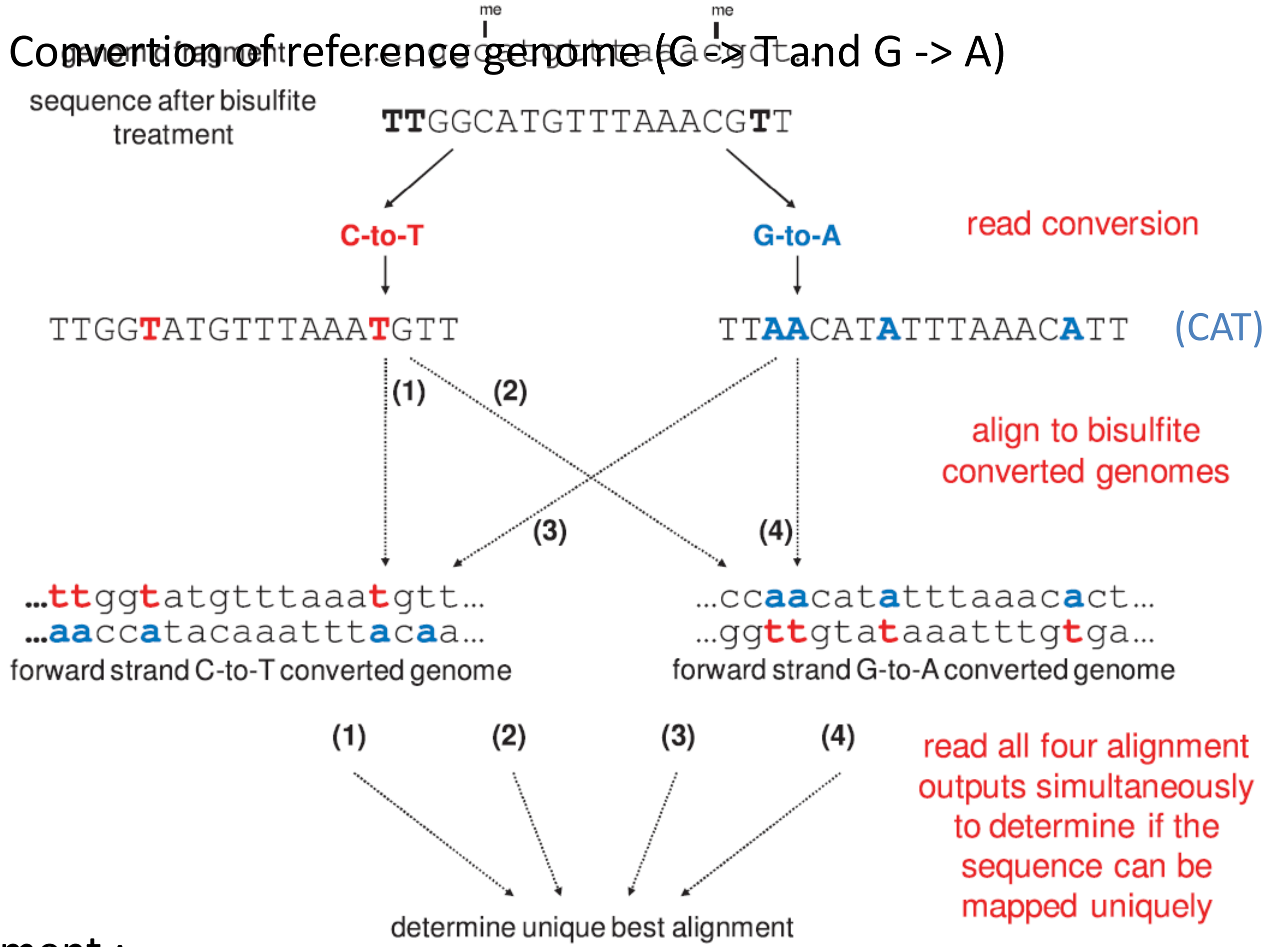
Krueger et al 2012

OT : Original Top
OB : Original Bottom

CTOT : Complementary to Original Top
CTOB : Complementary to Original Bottom

Bismark

Convertion of reference genome (C -> T and G -> A)



(GAT)

(CAT)

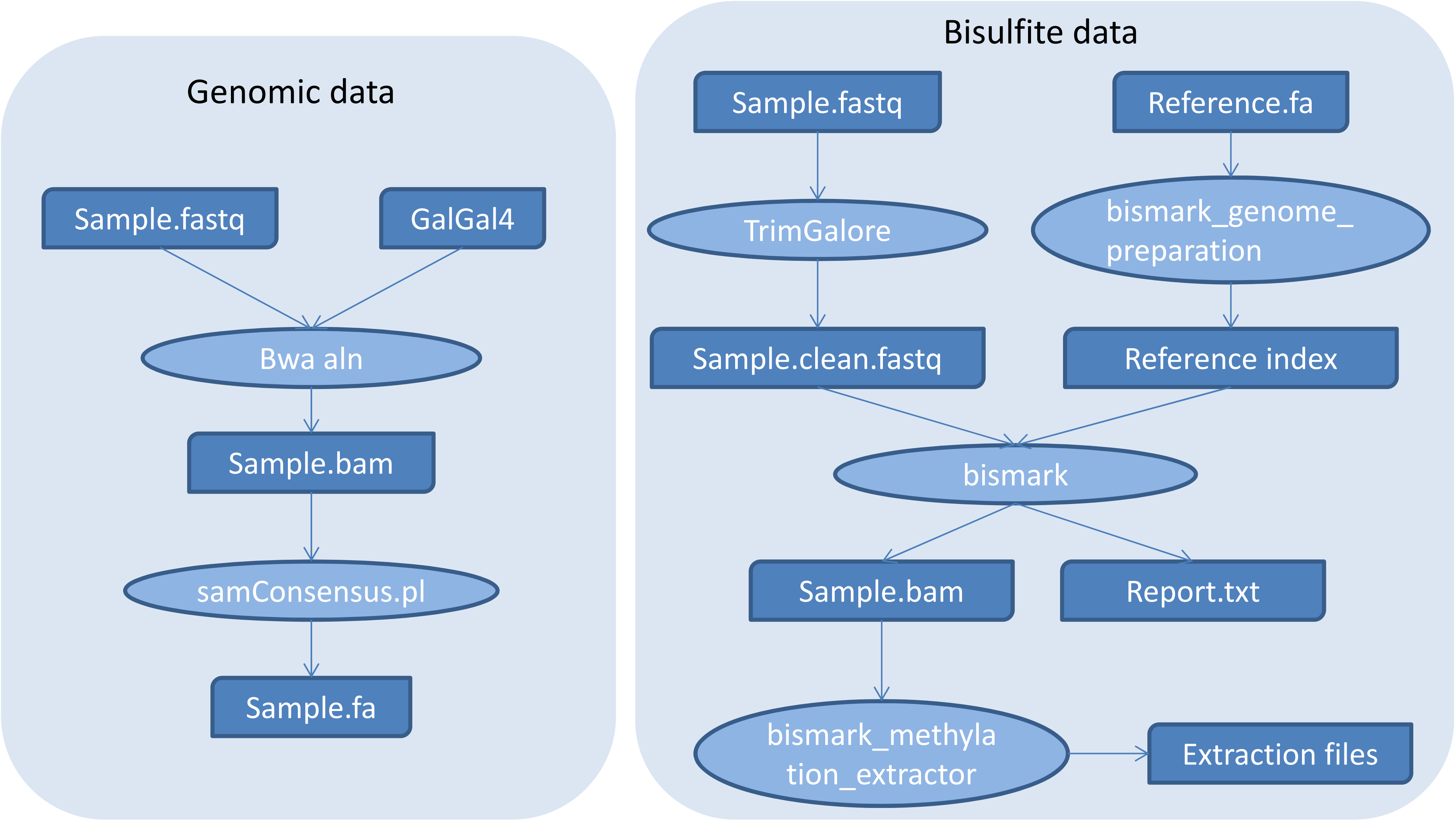
Directionnal paired alignment :

Read1 (OB or OT)

Read2 (CTOB or CTOT)

Krueger et al 2012

Bioinformatics traitments


Genomic data

Sample.fastq

GalGal4

Bwa aln

Sample.bam

samConsensus.pl

Sample.fa

Bisulfite data

Sample.fastq

Reference.fa

TrimGalore

bismark_genome_preparation

Sample.clean.fastq

Reference index

bismark

Sample.bam

Report.txt

bismark_methylation_extractor

Extraction files

Bioinformatics traitments (DNA)

- Reference genome : ICGSC Gallus_gallus-4.0/galGal4

- Genomic sample for both embryos
  - Emb1 : 180M reads paired (93.4% mapped)
  - Emb2 : 209M reads paired (93.1% mapped)

- Generation of new reference sample specific
  - > Reduce putative biais due to SNPs

Bioinformatics traitments (BS-seq)

- Cleanning : TrimGalore
  - Based on Cutadapt
  - Discard reads which length < 20bp

- BS-seq Data (nb reads paired)

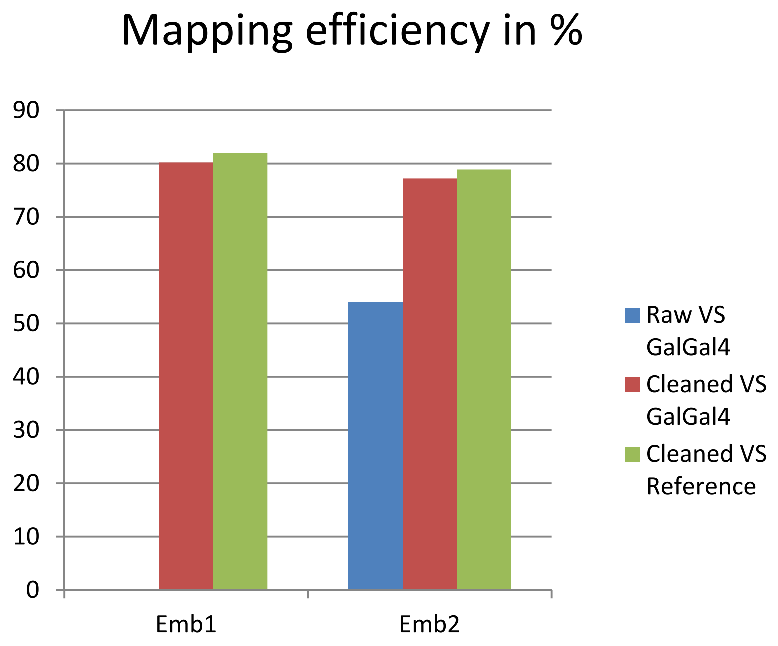
	Raw	Trimmed	Removed
Emb 1	157,230,624	68,430,229	1,099,914
Emb 2	164,728,172	70,697,424	918,796

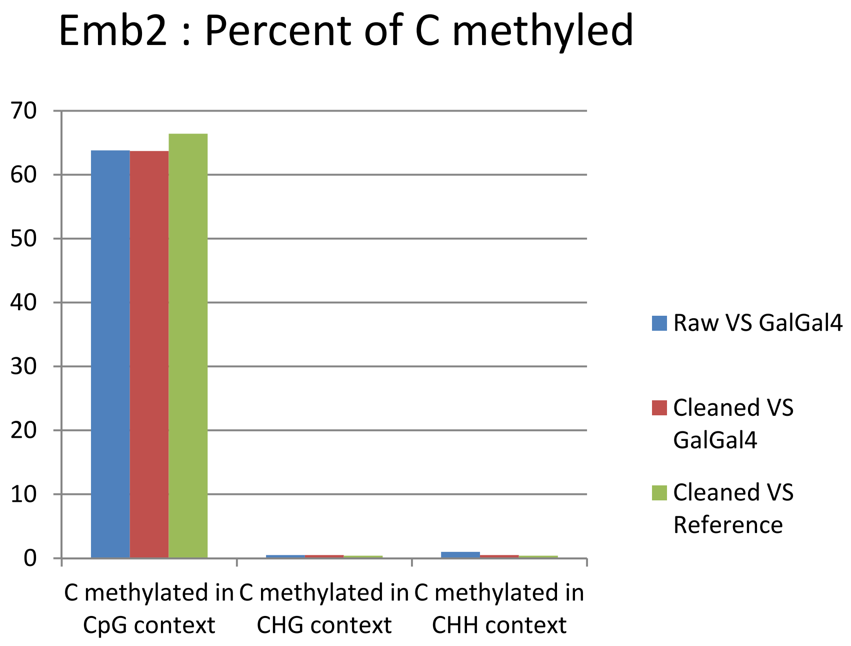
Bioinformatics traitments (BS-seq)

- Tests realized :
  - Alignment of raw data on GalGal4 (Emb2)
  - Alignment of cleaned data on GalGal 4 (Emb 1 & 2)
  - Alignment of cleaned data on sample reference (Emb 1 & 2)

Alignment results

- Outputs : Bam + Report.txt
- Depth = 14.3 x





Methylation calling results

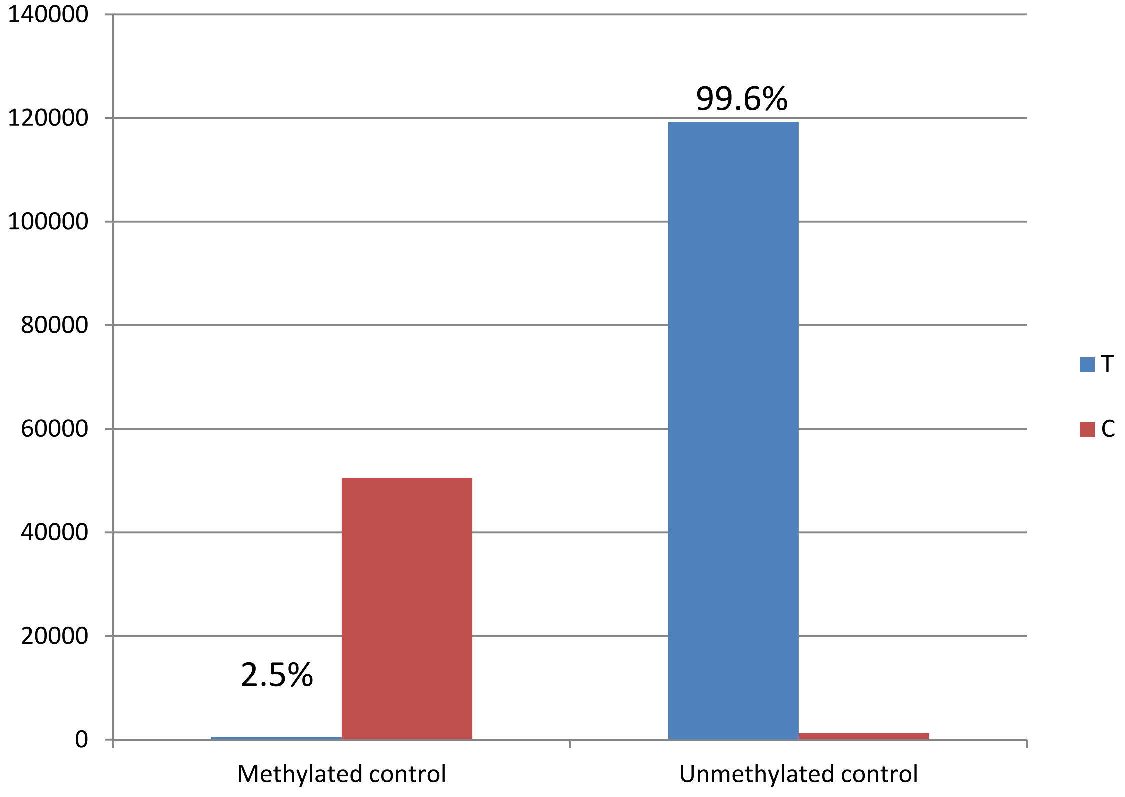
- Extraction of methylation with bismark_methylation_extractor
- Options :

  --report --bedGraph --counts

  --cytosine_report --no_overlap --CX_context
- Outputs:
  - report.txt
  - BEDgaph
  - CHG_OB.txt, CHG_OT.txt, CHH_OB.txt, CHH_OT.txt, CpG_OB.txt, CpG_OT.txt
  - CX_report.txt

Controls

pUC19 DNA (Zymo Research), methylated or not



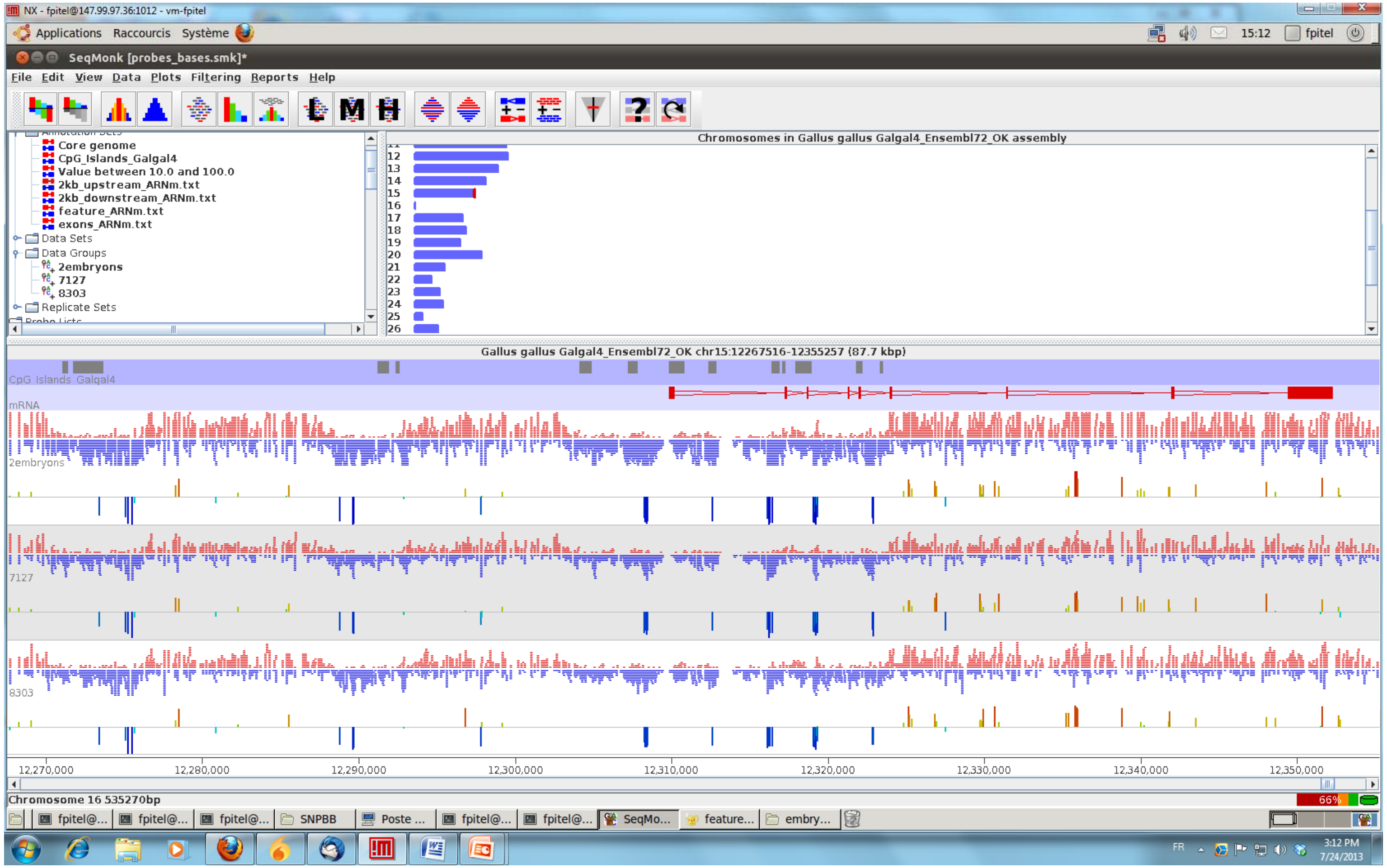
CpG

Overconversion

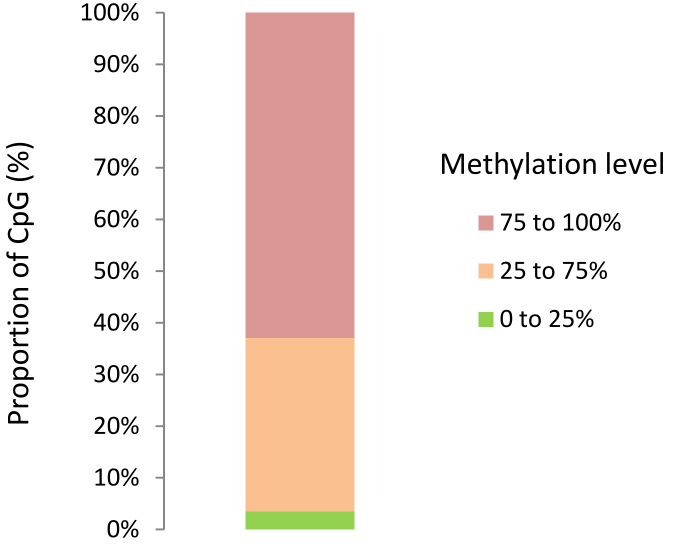
Conversion

Non CpG		
Non converted C	*2593*	*0.40228432*
Total	*644569*	

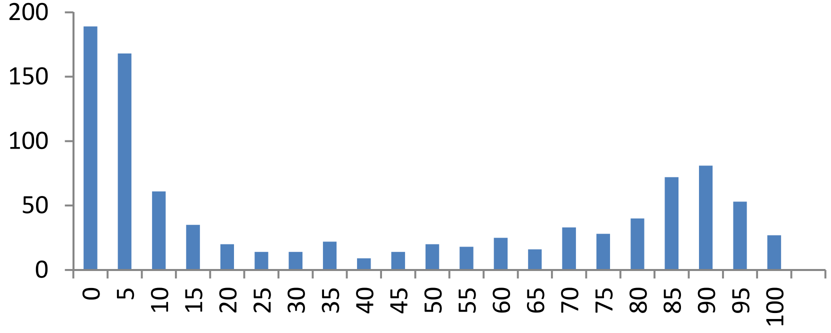
SeqMonk visualization



Methylation level per genomic feature

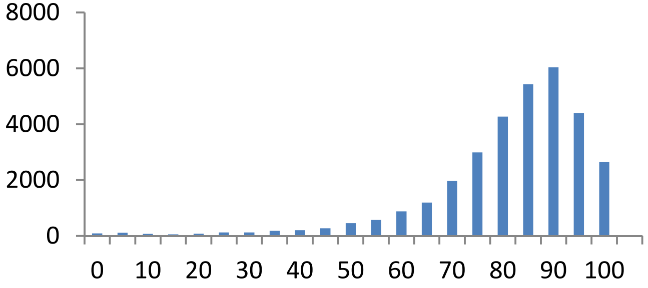


CpG islands

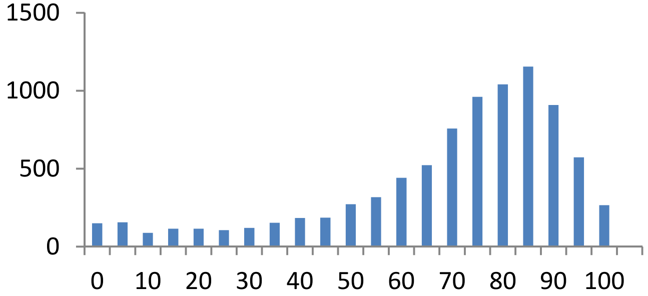


- Similar to other vertebrates

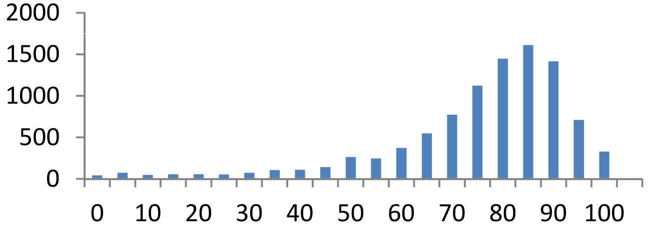
Exons



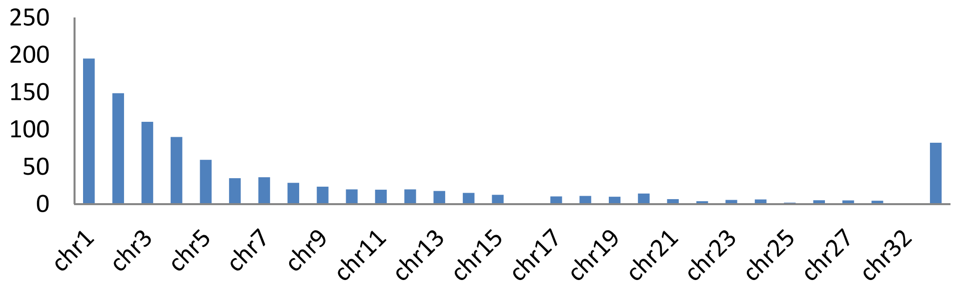
Upstream

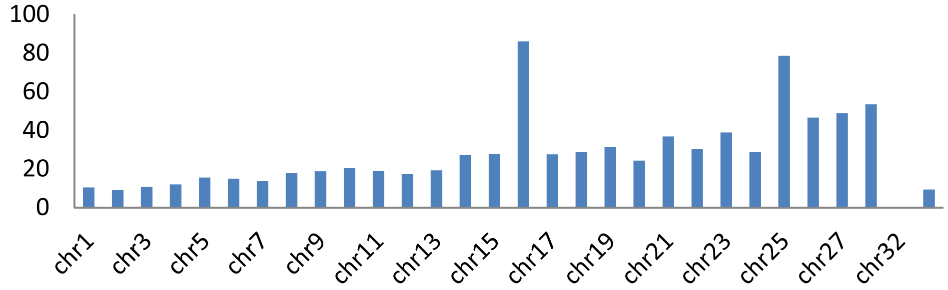


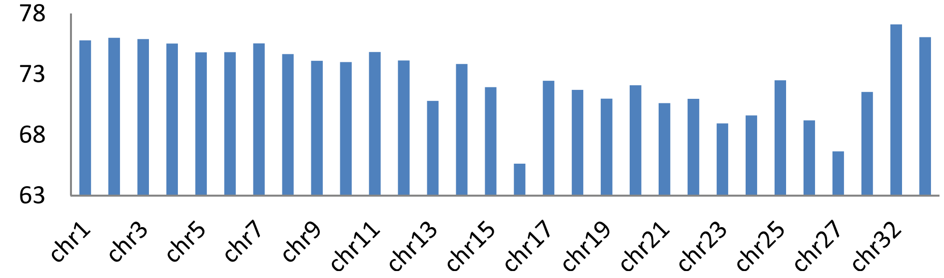
Downstream



Methylation level per chromosome


Chromosome length (Mb)


Genes / Mb


Methylation level

- Expected distribution along the genome

Conclusion

- Method OK
    - Many thanks to Genotoul (PlaGe & Bioinfo)
    - High quality DNA
    - RRBS for sufficient depth

- Pipeline (NG6) in progress

- Confirmation methods available at GeT-PlaGe (Pyromark,…)

- Other embryos to be sequenced

- EpiTherm program underway

Acknowledgments

UE PEAT
Tours
David Gourichon

URA
Tours
Catherine Beaumont

PRC
Tours
Vincent Coustham
Anne Collin

UMR GABI
Jouy en Josas
Tatiana Zerjal
Bertrand Bed'hom

LGC
Toulouse
Laure Frésard
Sophie Leroux
Mireille Morisson
Bertrand Servin
Magali San Cristobal
Alain Vignal

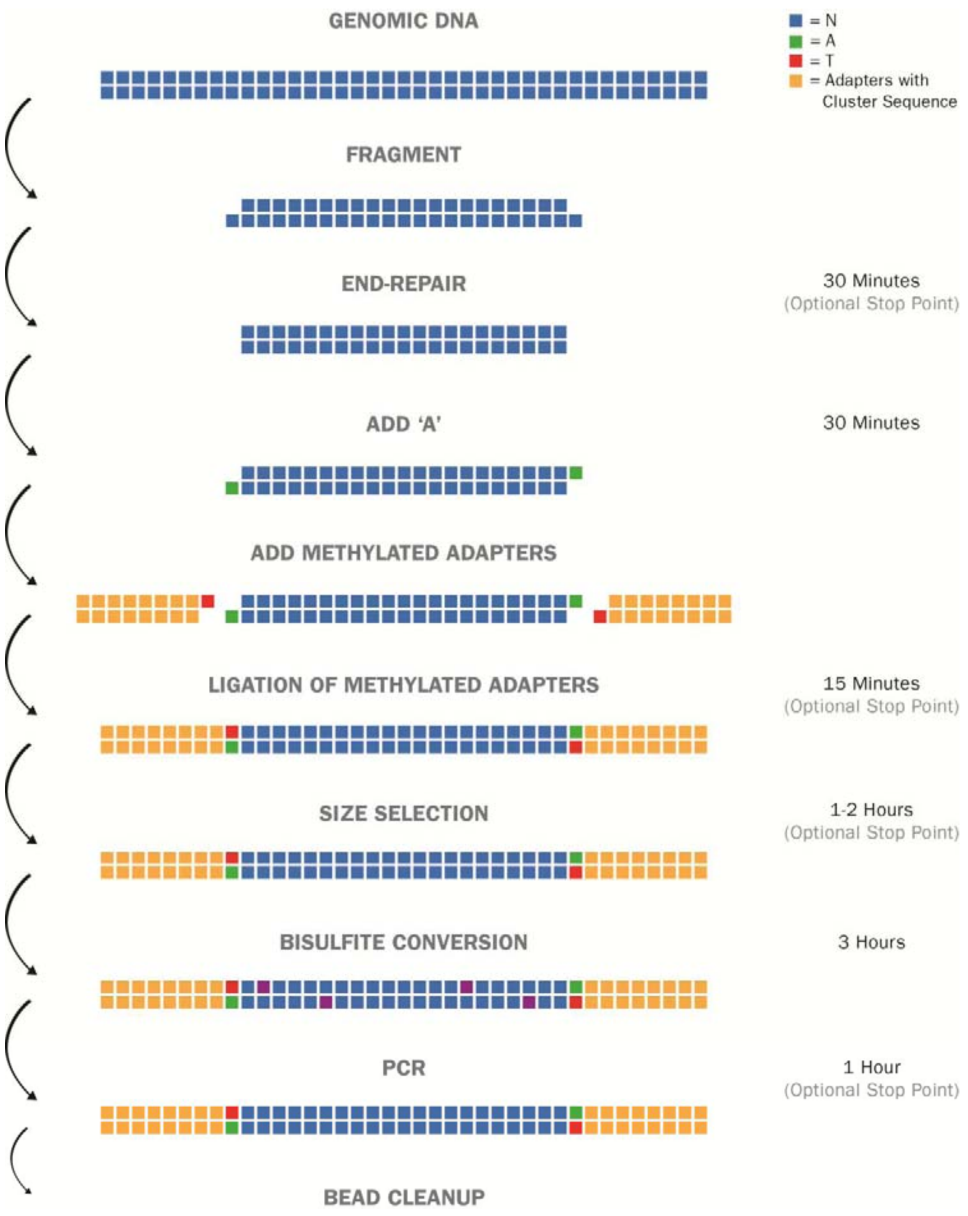
SIGENAE
Patrice Dehais

UMR PEGASE
Rennes
Sandrine Lagarrigue

Roslin Institute
Edinburgh
Pete Kaiser

GenoToul / PlaGe
Toulouse
Olivier Bouchez
Nathalie Marsaud
Diane Esquerré
Gérald Salin
Céline Noirot

Fundings
Animal Genetics Department
Region Midi-Pyrénées
ANR EpiBird



🟦 = N
🟩 = A
🟥 = T
🟧 = Adapters with Cluster Sequence

GENOMIC DNA

FRAGMENT

END-REPAIR
30 Minutes
(Optional Stop Point)

ADD 'A'
30 Minutes

ADD METHYLATED ADAPTERS

LIGATION OF METHYLATED ADAPTERS
15 Minutes
(Optional Stop Point)

SIZE SELECTION
1-2 Hours
(Optional Stop Point)

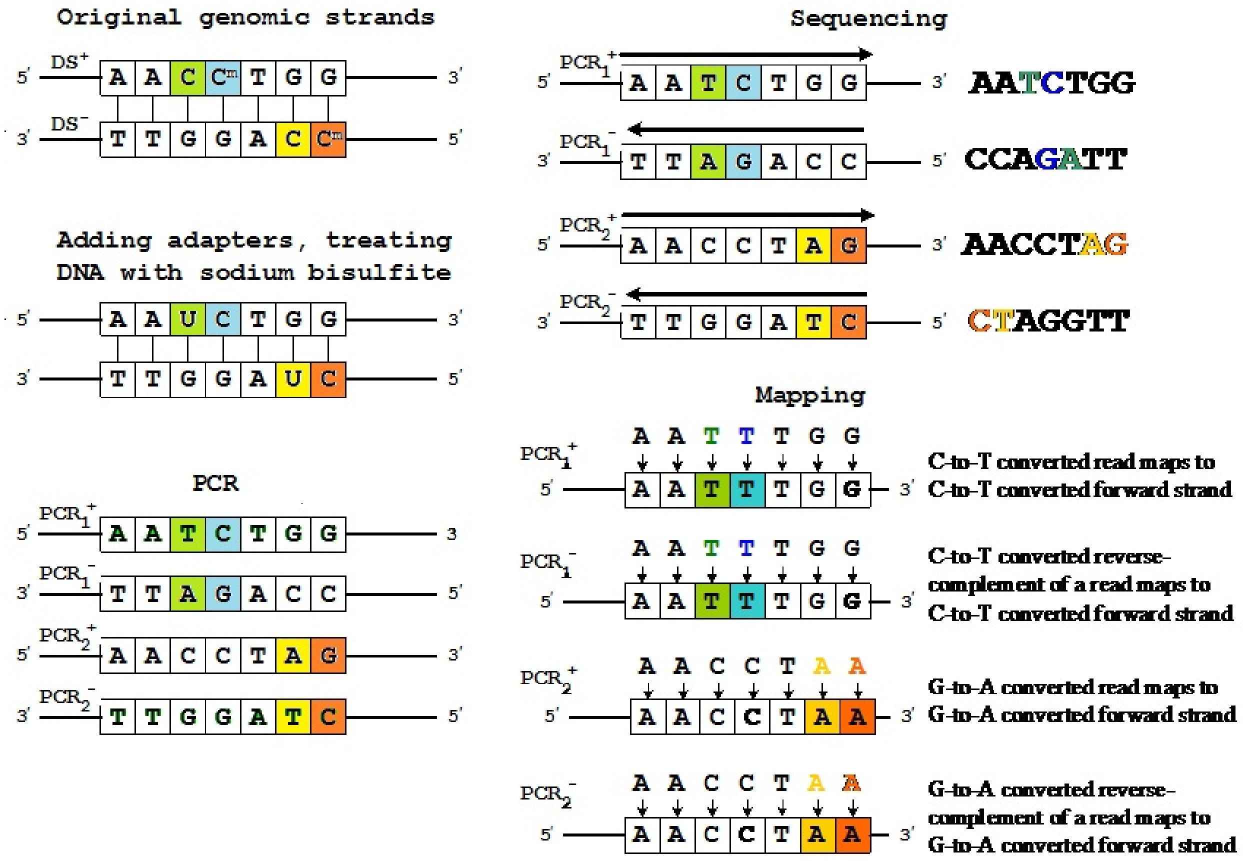
BISULFITE CONVERSION
3 Hours

PCR
1 Hour
(Optional Stop Point)

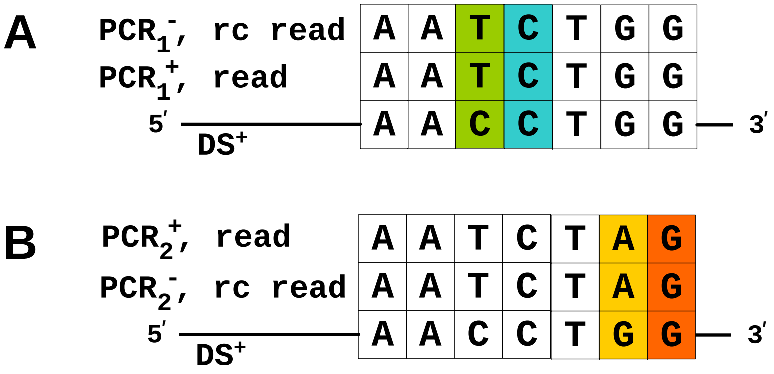
BEAD CLEANUP

BIOO Scientific Corp

BS-seq


Original genomic strands
DS+ 5' A A C Cm T G G 3'
DS- 3' T T G G A C Cm 5'
Adding adapters, treating DNA with sodium bisulfite
5' A A U C T G G 3'
3' T T G G A U C 5'
PCR
PCR1+ 5' A A T C T G G 3
PCR1- 3' T T A G A C C 5'
PCR2+ 5' A A C C T A G 3'
PCR2- 3' T T G G A T C 5'
Sequencing
PCR1+ 5' A A T C T G G 3' AATCTGG
PCR1- 3' T T A G A C C 5' CCAGATT
PCR2+ 5' A A C C T A G 3' AACCTAG
PCR2- 3' T T G G A T C 5' CTAGGTT
Mapping
PCR1+ A A T T T G G 5' A A T T T G G 3'
C-to-T converted read maps to C-to-T converted forward strand
PCR1- A A T T T G G 5' A A T T T G G 3'
C-to-T converted reverse-complement of a read maps to C-to-T converted forward strand
PCR2+ A A C C T A A 5' A A C C T A A 3'
G-to-A converted read maps to G-to-A converted forward strand
PCR2- A A C C T A A 5' A A C C T A A 3'
G-to-A converted reverse-complement of a read maps to G-to-A converted forward strand

BS-seq



A

PCR$_1^-$, rc read	A	A	T	C	T	G	G
PCR$_1^+$, read	A	A	T	C	T	G	G
5′ DS$^+$	A	A	C	C	T	G	G	3′

B

PCR$_2^+$, read	A	A	T	C	T	A	G
PCR$_2^-$, rc read	A	A	T	C	T	A	G
5′ DS$^+$	A	A	C	C	T	G	G	3′