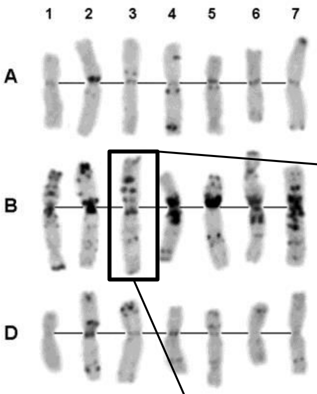
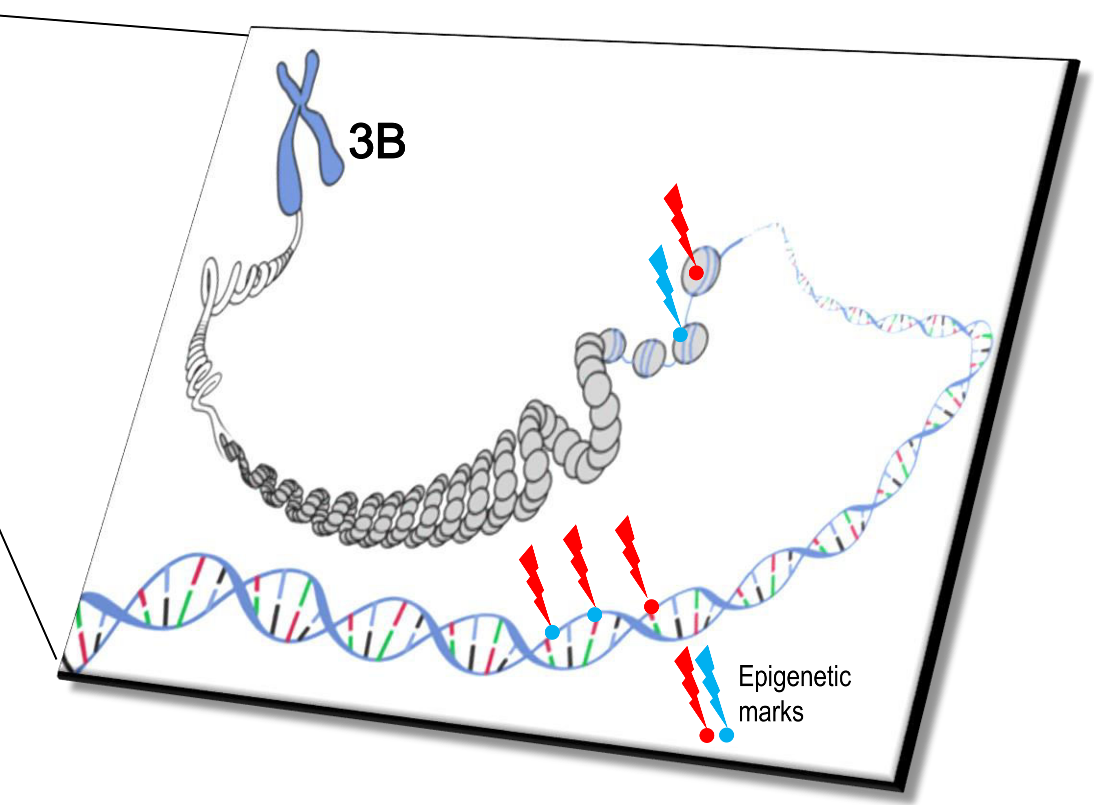
Epi3B

Epigenomic maps of the chromosome 3B





Team leader: Etienne Paux

Aurélie Evrard

Joint Research Unit on the Genetics, Diversity and Ecophysiology of Cereals (GDEC), Clermont-Ferrand, France

Outline

- Introduction on epigenetics and chromatin states
- Aim and strategy of the epi3B project
- First results from mapping and peak calling

Epigenetic examples

The term « epigenetics» describes (heritable?) patterns of phenotypic variation that are not attributable to differences in DNA sequence.

(Eichten *et al*., 2014)

Vernalization in temperate cereals

A winter wheat crop in February 2013 at INRA-GDEC site. (photo by S. Toillon)

Seedlings required an exposure to low temperature in order to ensure flowering and grain development in spring and summer.

- ❖ Activation of the key locus gene *VRN1*
- ❖ Alteration of 2 epigenetic marks in barley

Epigenetic examples

The term « epigenetics» describes heritable patterns of phenotypic variation that are not solely attributable to differences in DNA sequence.

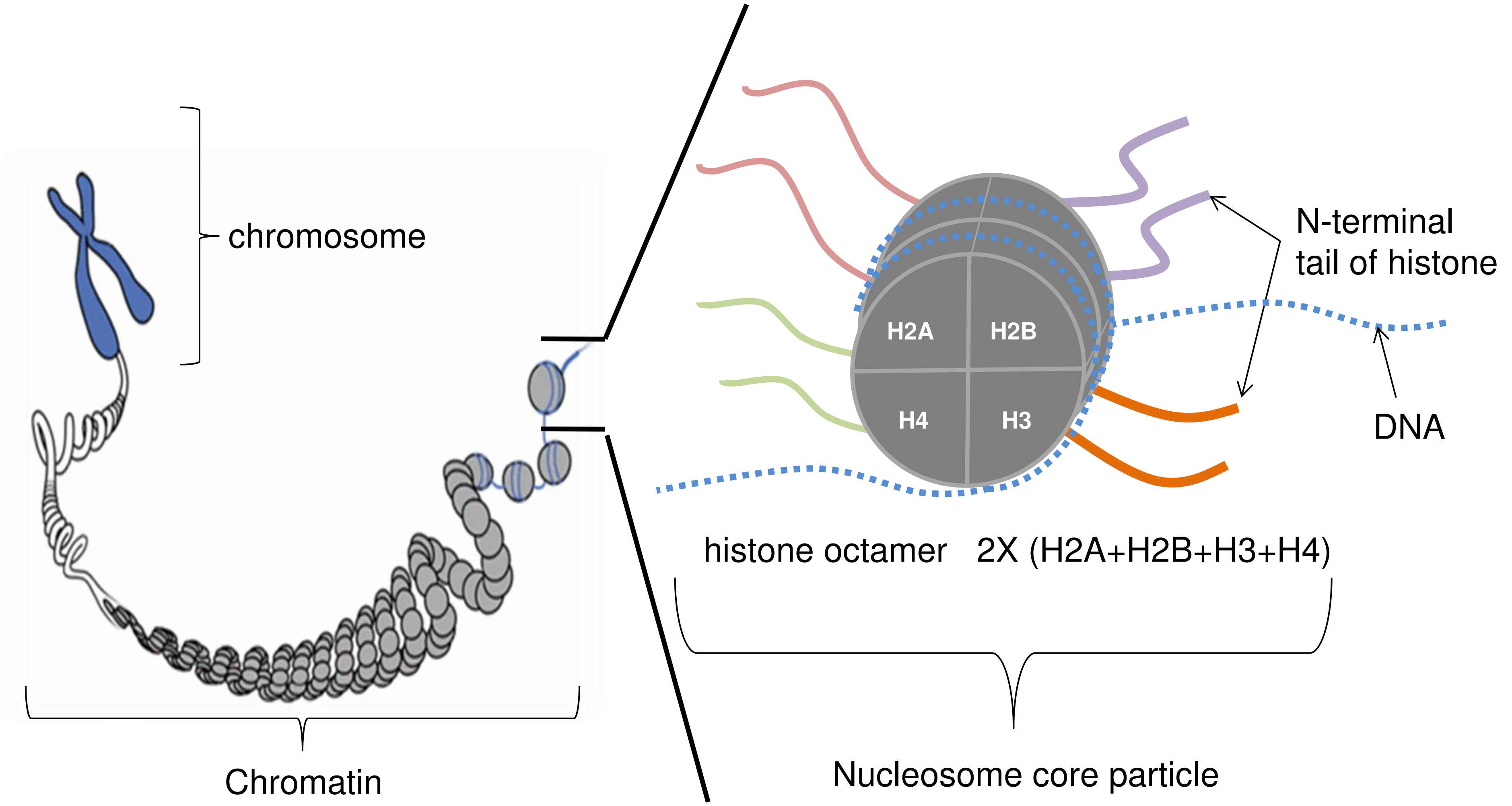
(Eichten *et al.*, 2014)

Hybrid vigour in plants

Superior performance in yield or biomass of an offspring compared to its 2 distinct parental lines.

Parent 1 ear (left), hybrid ear (middle), parent 2 ear (right).
University of Nebraska-Lincoln, 2004.

Chromatin & histones



chromosome

Chromatin

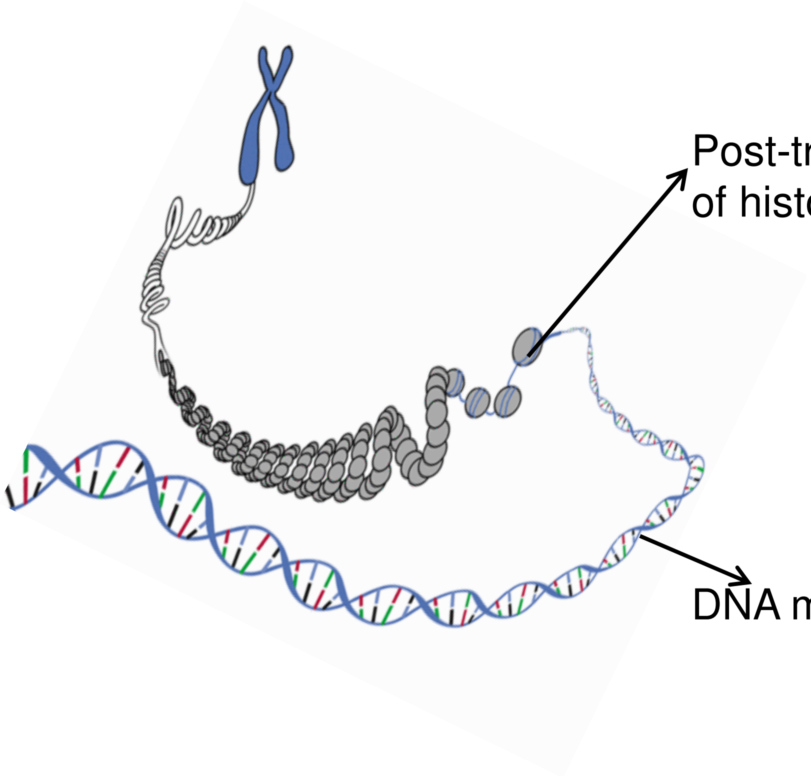
N-terminal
tail of histone

DNA

histone octamer 2X (H2A+H2B+H3+H4)

Nucleosome core particle

Chromatin & epigenetic marks



Post-traductional modifications of histones

Acetylation
Ubiquitination
Méthylation
Phosphorylation
Sumoylation ...

DNA modifications

Cytosine methylation

Involved in different processes such as

- regulation of transcription
- replication
- recombination
- maintenance of genome integrity

H3 tail

ARTKQTARKSTGGKAPRKQLATKAAR**K**SAPATGGVKK......... amino acids

23456789 **27** 36 ..... position

H3K4me2
H3K4me3

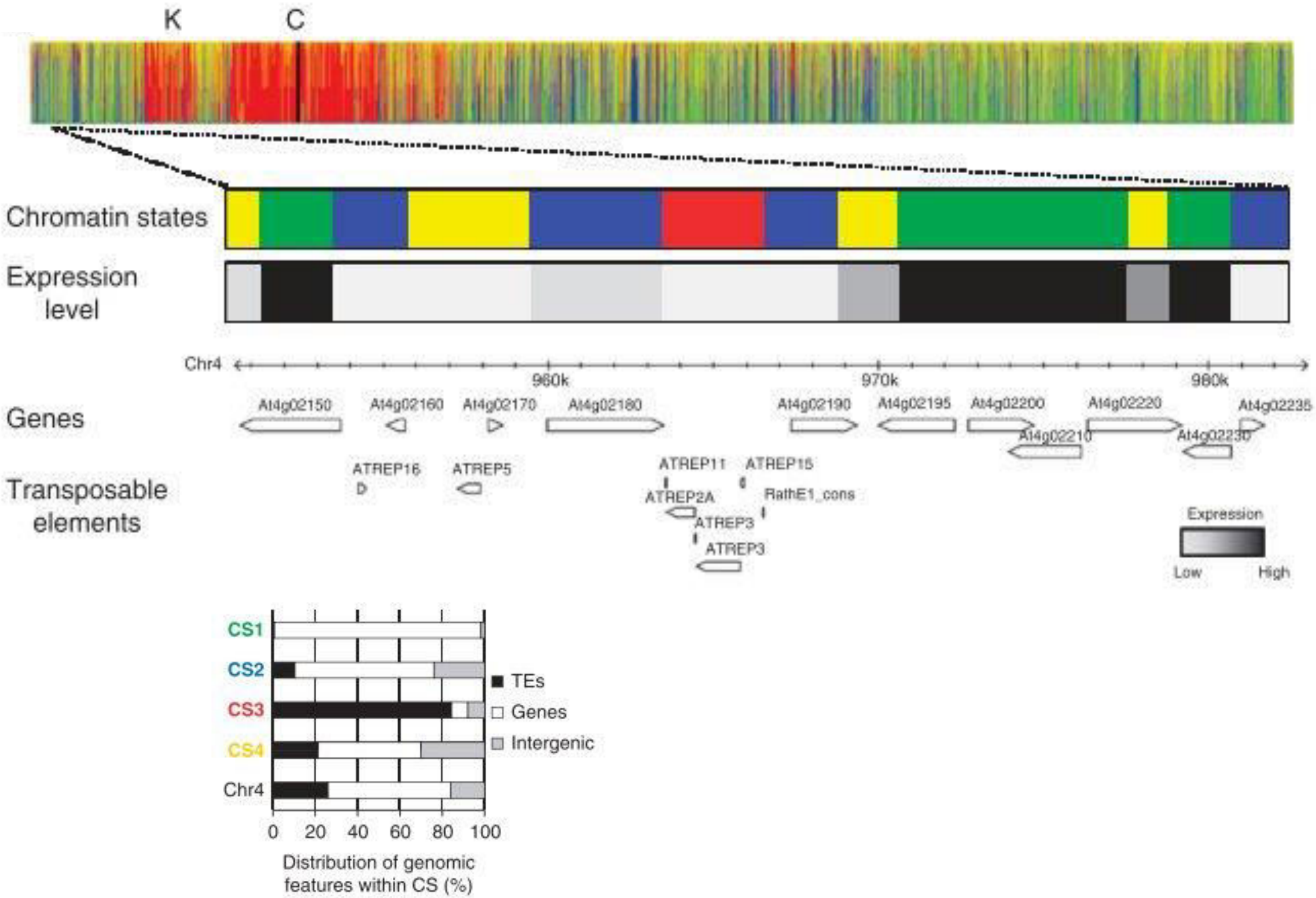
H3K27me3 H3K36me3

Picture from textbook Anatomy & Physiology, OpenStax College

Chromatin states

Eukaryotic chromatin : euchromatin and heterochromatin domains

Chromatin states define by the position of marks along features of a chromosome.



Roudier *et al.*, 2011

Challenges

In bread wheat genome-wide studies have been limited due to :

- the lack of reference sequences
- large genome size & allohexaploidy
- high repeat content (>80%)

~ 17,000 Mb

120-fold *A. thaliana* genome
45-fold rice genome

Aim of the Epi3B project

In bread wheat genome-wide studies have been limited due to :

- the lack of reference sequences
- large genome size & allohexaploidy
- high repeat content (>80%)

~ 17,000 Mb

120-fold *A. thaliana* genome
45-fold rice genome

BUT

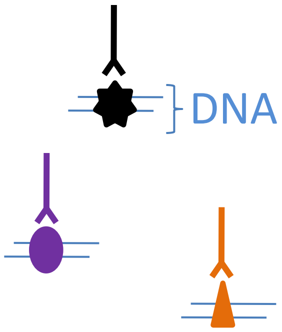
The first reference sequence of the chromosome 3B is now available.

The aim of the Epi3B project is **to explore the epigenomic landscape** along the chromosome 3B

- Explore the localization/combination of epigenetic marks around genes
- Define chromatin states along the chromosome
- Investigate the distribution patterns of the epigenomic marks according the gene density, gene expression, recombination hotspots or other features of the gene space.

Methodology

Chromatin immunoprecipitation (ChIP) methylated DNA immunoprecipitation (MeDIP) from wheat leaves

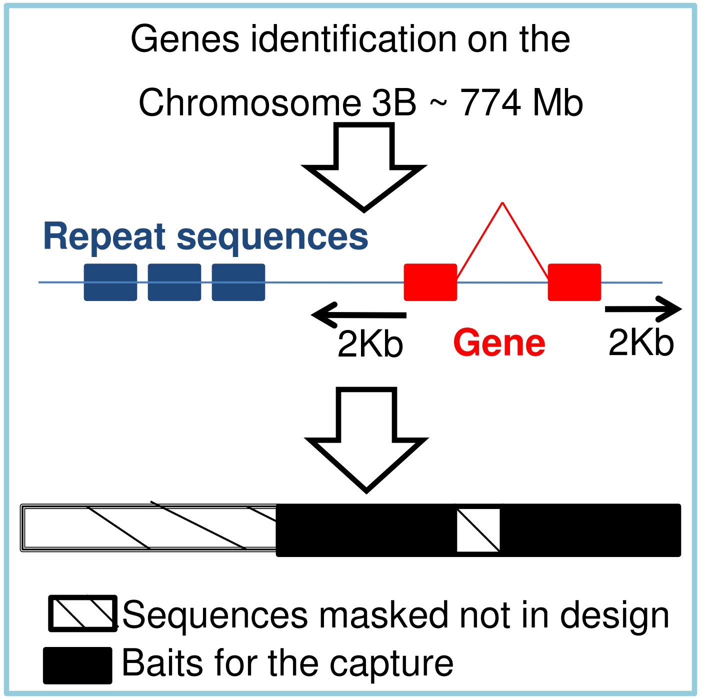


Histone 3
H3K4me2
H3K4me3
H3K36me3
H3K27me3
5mCytosine

- Collection of whole genome fraction according to each antibody
- Negative control for each ChIP (specificity of beads to antibodies only)

Methodology

- We aim to target (only) genes on the 3B chromosome (liquid hybridization to capture sequences)



Genes identification on the

Chromosome 3B ~ 774 Mb

Repeat sequences

2Kb **Gene** 2Kb

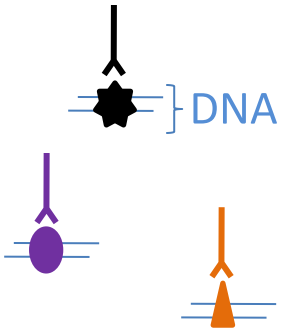
▨ Sequences masked not in design

■ Baits for the capture

Size capture sequences : 36Mb

Methodology

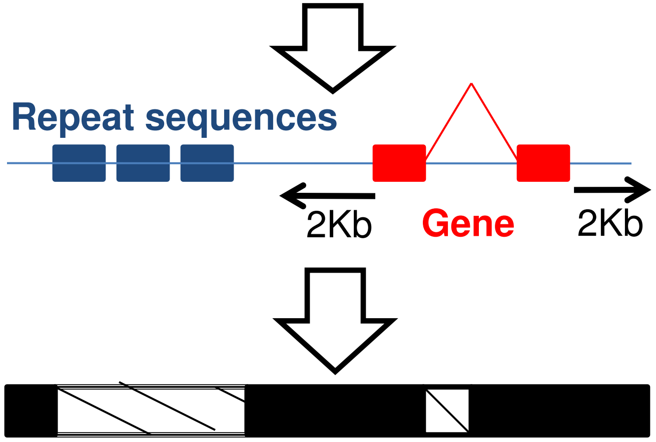
Chromatin immunoprecipitation (ChIP) methylated DNA immunoprecipitation (MeDIP)



Histone 3
H3K4me2
H3K4me3
H3K36me3
H3K27me3
5mCytosine

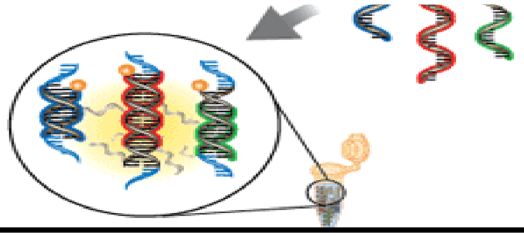
Genes identification on the

Chromosome 3B ~ 774 Mb



▨ Sequences masked not in design
■ Baits for the capture

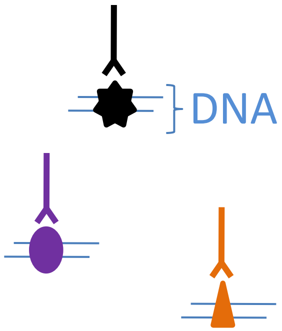
3-Sequences capture: liquid hybridization



Enrichment of epigenetic marks of the 3B.

Methodology

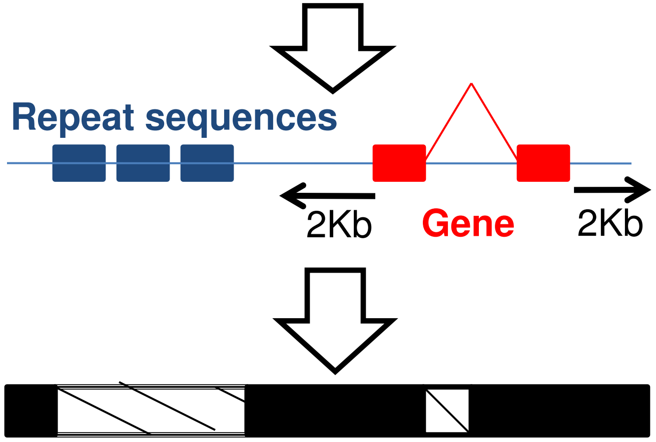
Chromatin immunoprecipitation (ChIP) methylated DNA immunoprecipitation (MeDIP)



Histone 3
H3K4me2
H3K4me3
H3K36me3
H3K27me3
5mCytosine

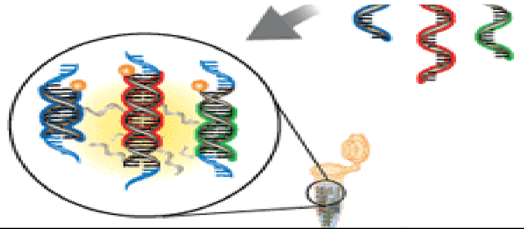
Genes identification on the

Chromosome 3B ~ 774 Mb



▨ Sequences masked not in design
■ Baits for the capture

Sequences capture: liquid hybridization



Sequencing

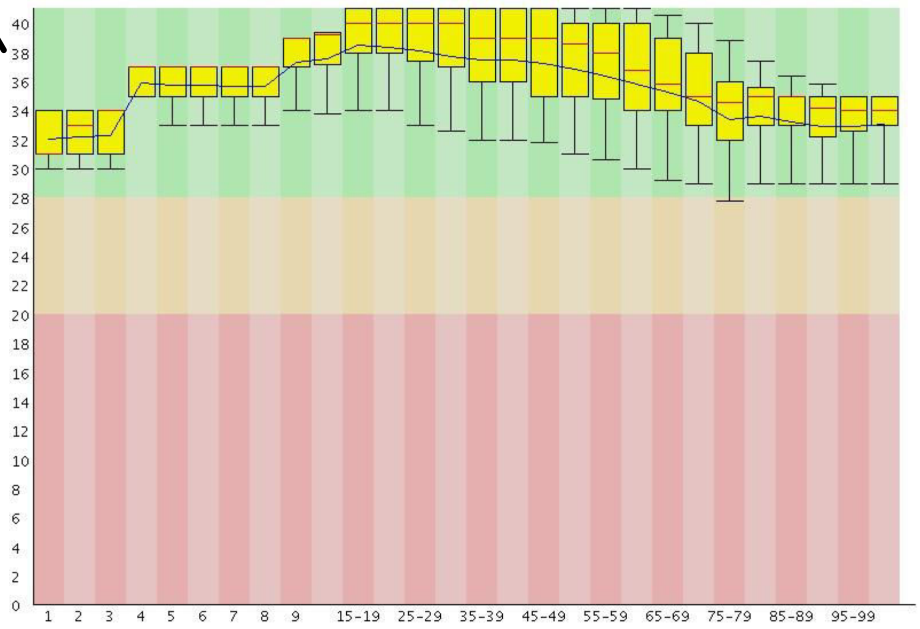
Bioinformatic analysis

Quality sequencing

Each sample has an average **of 103 millions quality trimmed read pairs** (QC >30).

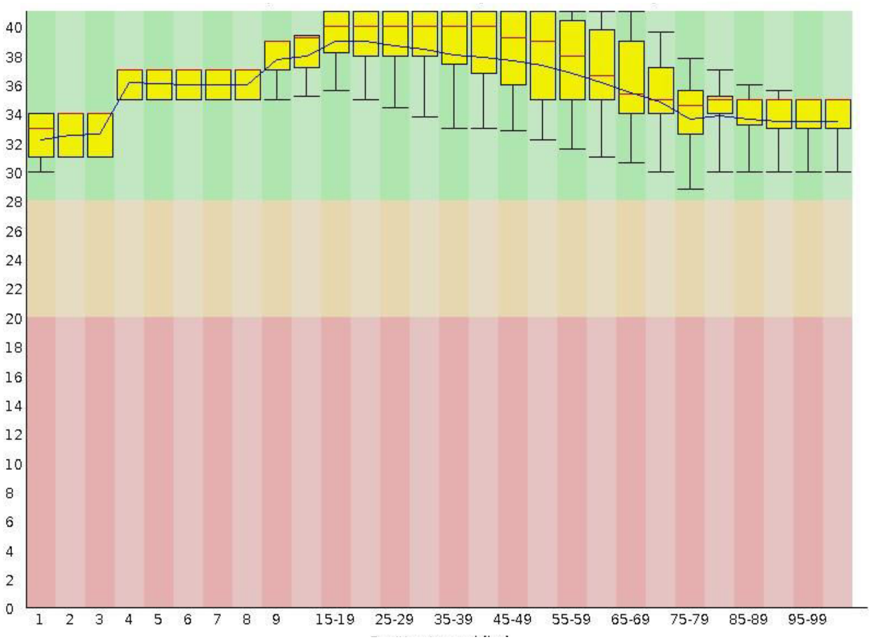
6 samples in duplicate : Histone 3, H3K4me2,H3K4me3, H3K36me3, H3K27me3, 5mCytosine

Read quality



Position in read (bp)

Read Forward sample H3 (rep1)



Position in read (bp)

Read Reverse sample H3 (rep1)

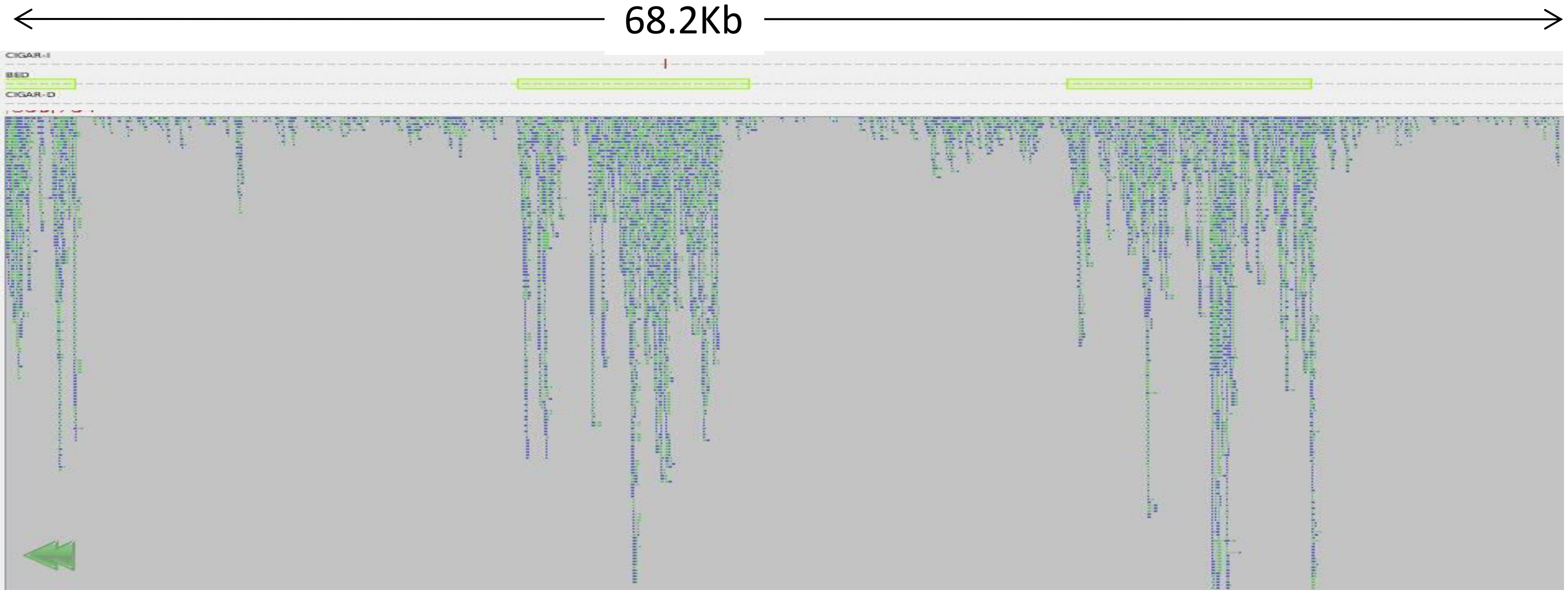
Capture efficiency

36MB of capture (=gene fraction) targeted on 3B → 0.21%

Capture efficiency on-target → 21%

Capture enrichment (from 0.21% to 21%) → 100X

% bases covered on the capture → 99%



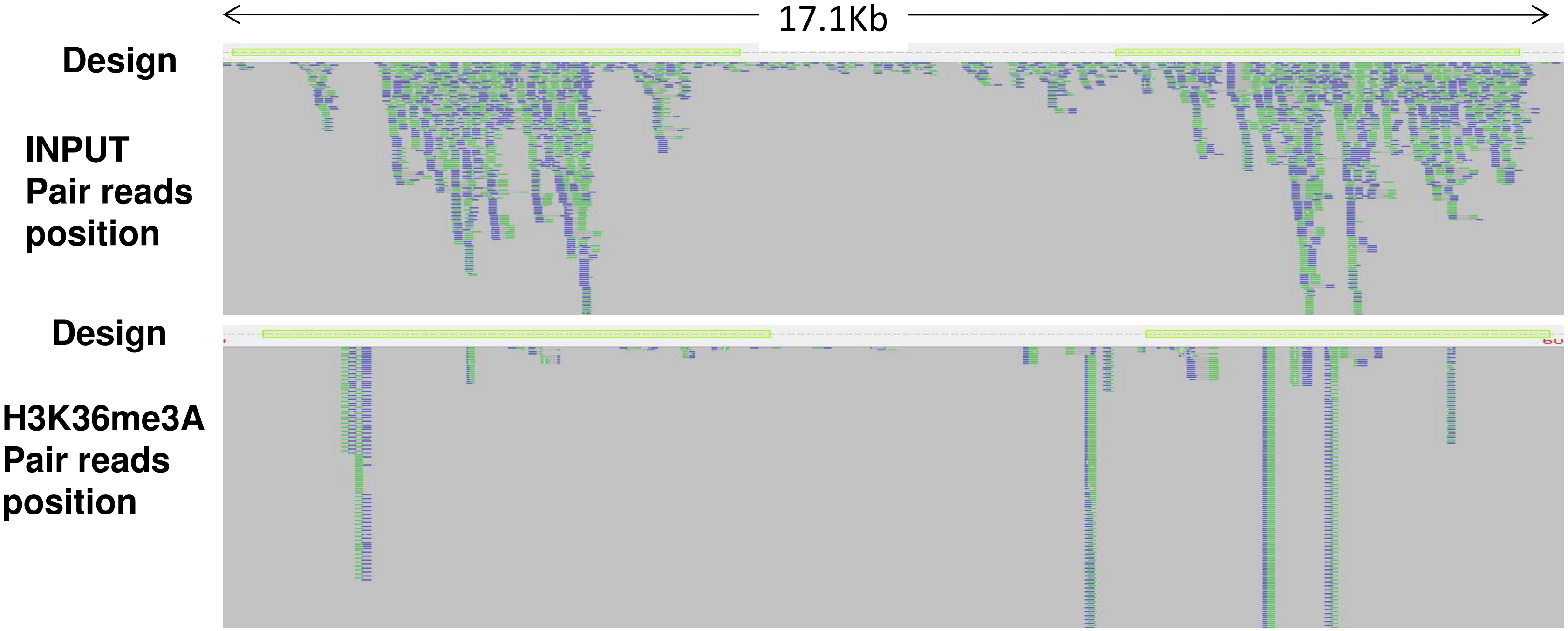
Capture design

INPUT
Pair reads
position

Results Mapping

- bwa with no mismatch in seeds and 2 mismatches allowed in read

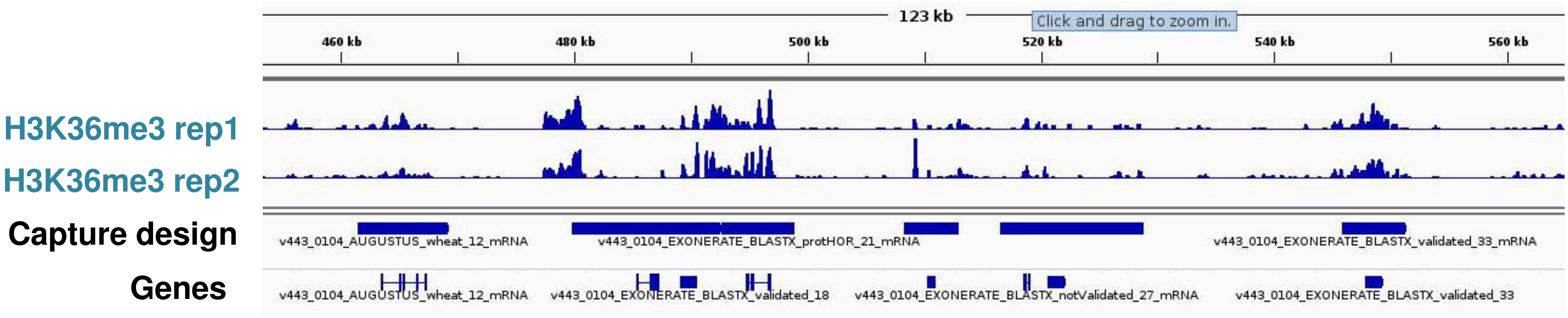
	H3K36me3 (rep 1)	H3K36me3 (rep 2)	H3K4me2 (rep1)	H3K4me2 (rep 2)	met Cyt (rep1)	met Cyt (rep2)
% coverage along capture design	40	38	13	13	4	5



Tablet visualization, Milne *et al.*, 2013

Results peak calling using macs2

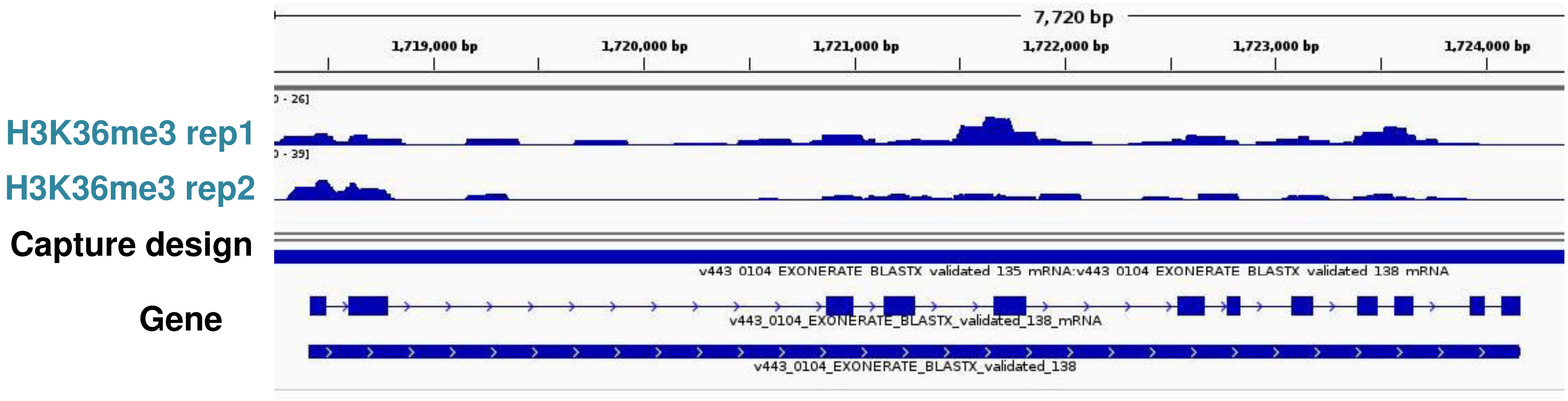
	H3K36me3 (rep 1)	H3K36me3 (rep 2)	H3K4me2 (rep1)	H3K4me2 (rep 2)	met Cyt (rep1)	met Cyt (rep2)
Spearman coefficient IDR (Irreproducibility discovery rate)	0.78		0.87		0.97	
Nb peaks common between biological replicates on capture design	2915		2247		395	



IGV visualization, Thorvaldsdóttir *et al.*, 2012

Results peak calling using macs2

	H3K36me3 (rep 1)	H3K36me3 (rep 2)	H3K4me2 (rep1)	H3K4me2 (rep 2)	met Cyt (rep1)	met Cyt (rep2)
Spearman coefficient IDR (Irreproducibility discovery rate)	0.78		0.87		0.97	
Nb peaks common between biological replicates on capture design	2915		2247		395	



IGV visualization, Thorvaldsdóttir *et al.*, 2012

Conclusions & future work

- All chIP and MeDIP experiments and data sequencing were successfully achieved

- Reads mapping and peak detection for each sample are currently running (test other softwares for peak calling –CCAT, FindPeak, SICER)

The Epi3B project aims

- to define chromatin states in a polyploid species

- to better characterize the genic space

- to determine the impact of the epigenetic marks on the transcriptional level of genes

Ultimately, these epigenomic maps will be useful resource to identify the structural and functional organization of the chromosome 3B in a polyploid context.

Acknowledgements

Group SEVEN

GDEC
Génétique Diversité Ecophysiologie des Céréales

Helene Rimbert, Biogemma

GENTYANE
PLATE-FORME de
GENOTYPAGE et SÉQUENÇAGE en AUVERGNE

GReD
GENETIQUE REPRODUCTION & DEVELOPPEMENT

Pr. Christophe Tatout
Dr. Aline Probst
Matthias Benoit