

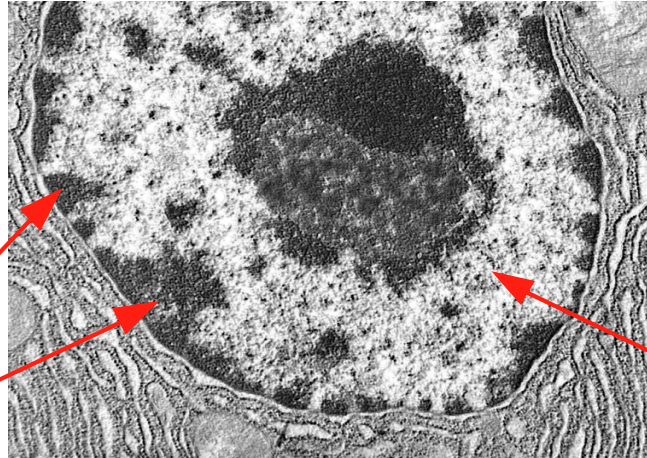
# **HOW TO STUDY EPIGENETIC MODIFICATIONS**

**-- STRATEGIES USING SUV39H1 AS A HALLMARK  
MODEL FOR EPIGENETIC REGULATION --**

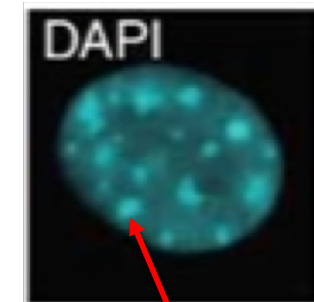
## Chromatin comes in different flavors

### Different types of chromatin

**heterochromatin**



**euchromatin**  
(and facultative  
heterochromatin)



**Chromocenter**  
(aggregates of centromeres  
= constitutive heterochromatin)

### Constitutive heterochromatin:

- constitute ~ 10% of nuclear DNA; telomeres, centromeres, and a considerable fraction of repetitive sequences
- highly compacted, replicates late in S phase, (transcriptionally inert)

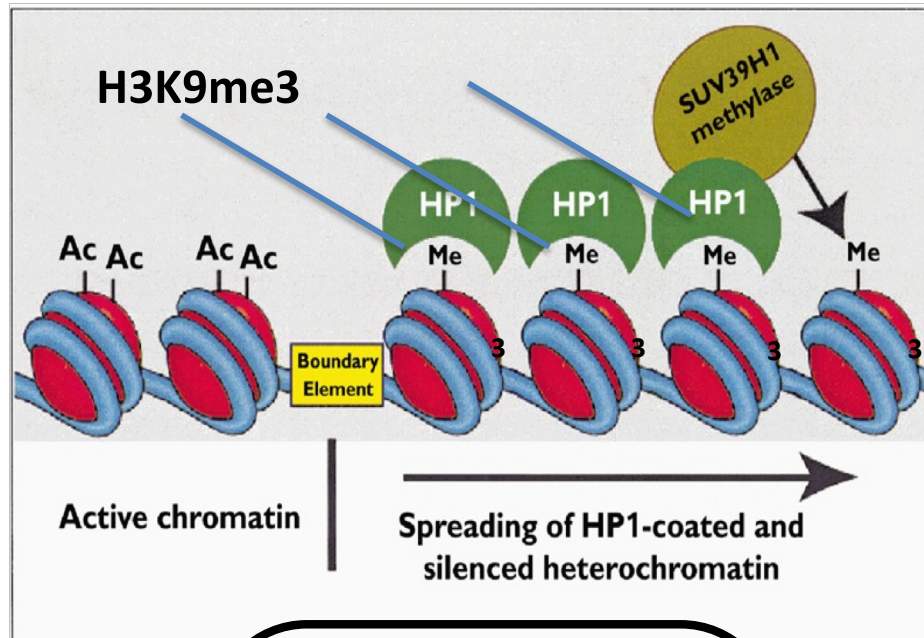
### Euchromatin + facultative heterochromatin:

- constitute ~ 90% of nuclear DNA
- less condensed, rich in genes, replicates early in S phase
- however,**
- only small fraction of euchromatin is transcriptionally active
- the rest is transcriptionally inactive/silenced (but can be activated in certain tissues or developmental stages) → these inactive regions are also known as “facultative heterochromatin”



## Post-translational histone modifications can recruit specialized proteins

Example: SUV39H1 and HP1 form heterochromatin at centromeric and telomeric  
Heterochromatin in flies and vertebrates and SAHFs



Suv39h1/2 tri-methylate

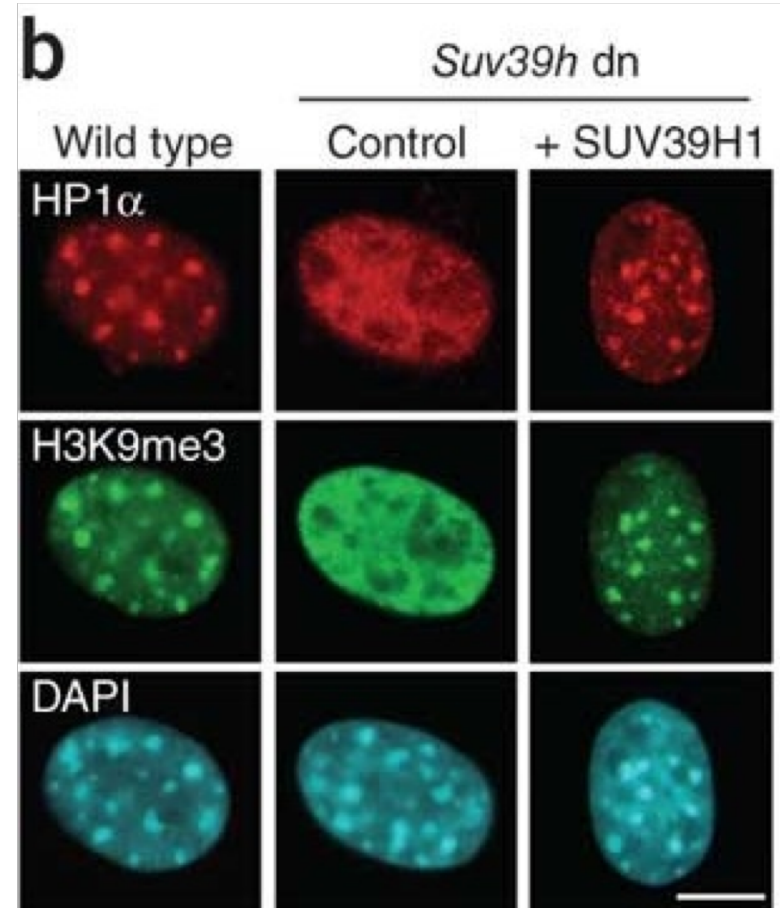
Lysine 9 on histone H3



tri-methylated  
Lysine 9 on histone H3  
recruits HP1



heterochromatin

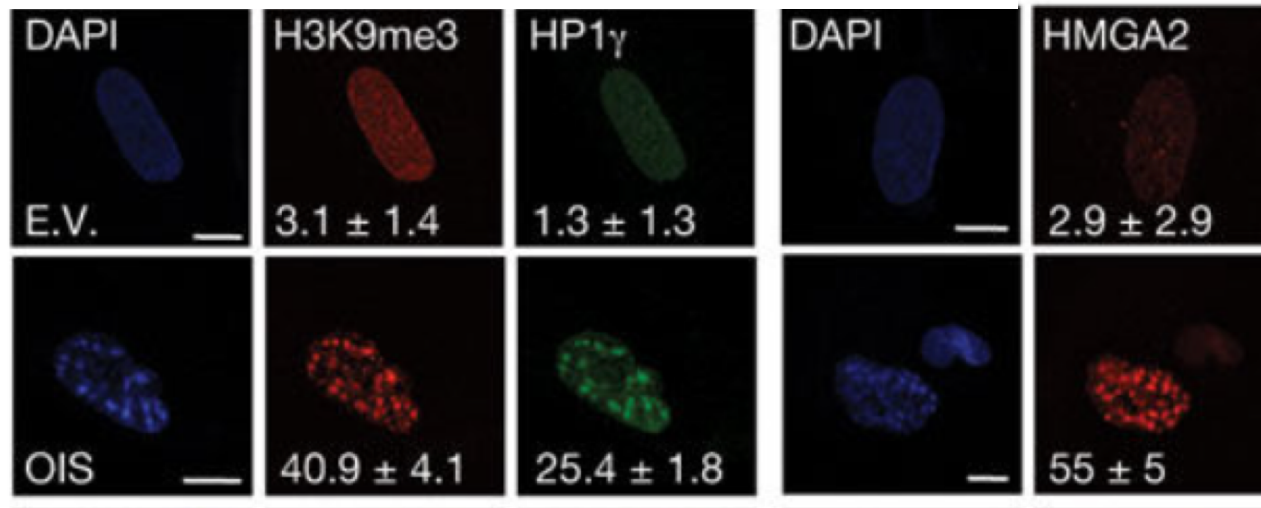


histone modifications can reach high levels in cells and can be visualized by immunofluorescence

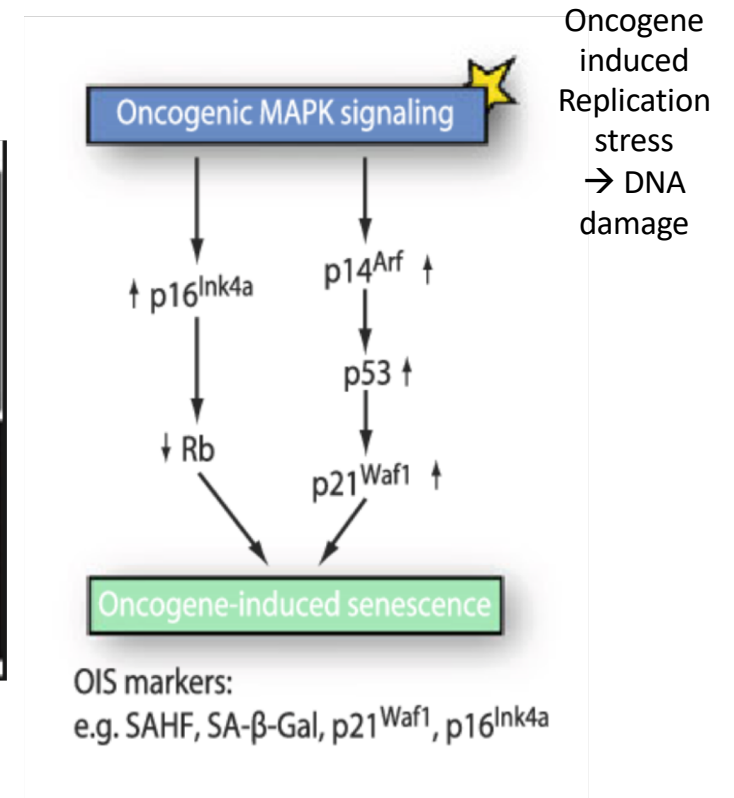
## Post-translational histone modifications can recruit specialized proteins

Example: SUV39H1 and HP1 form heterochromatin at centromeric and telomeric  
Heterochromatin in flies and vertebrates and SAHFs

### CELLULAR STIMULUS IS TRANSLATED IN TO CHROMATIN REGULATION



Senescence activated heterochromatin foci (SAHF)



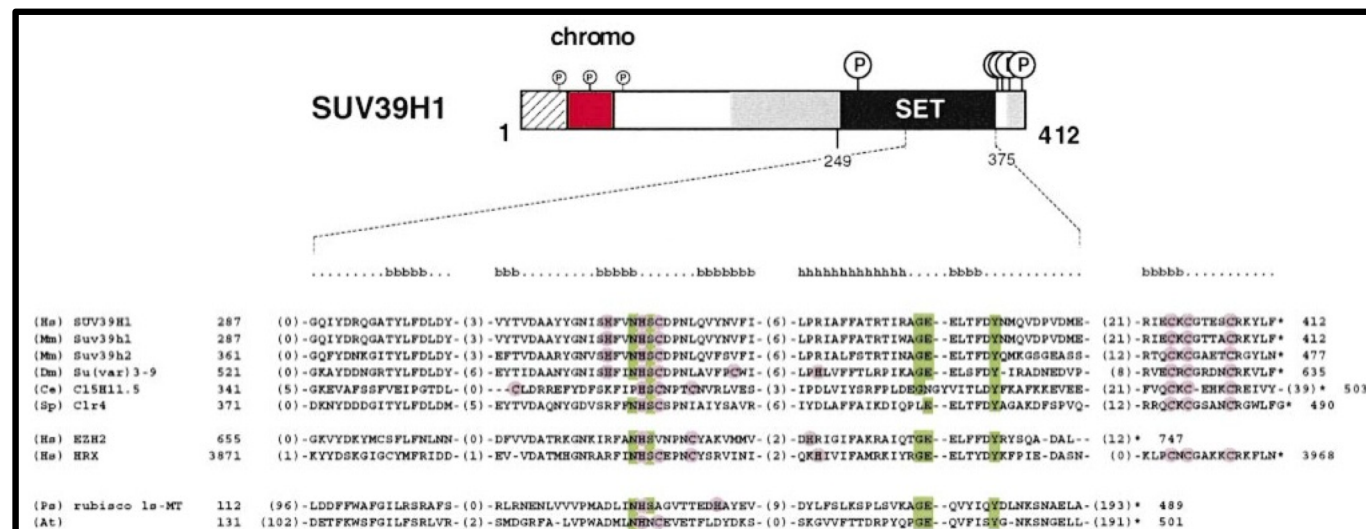
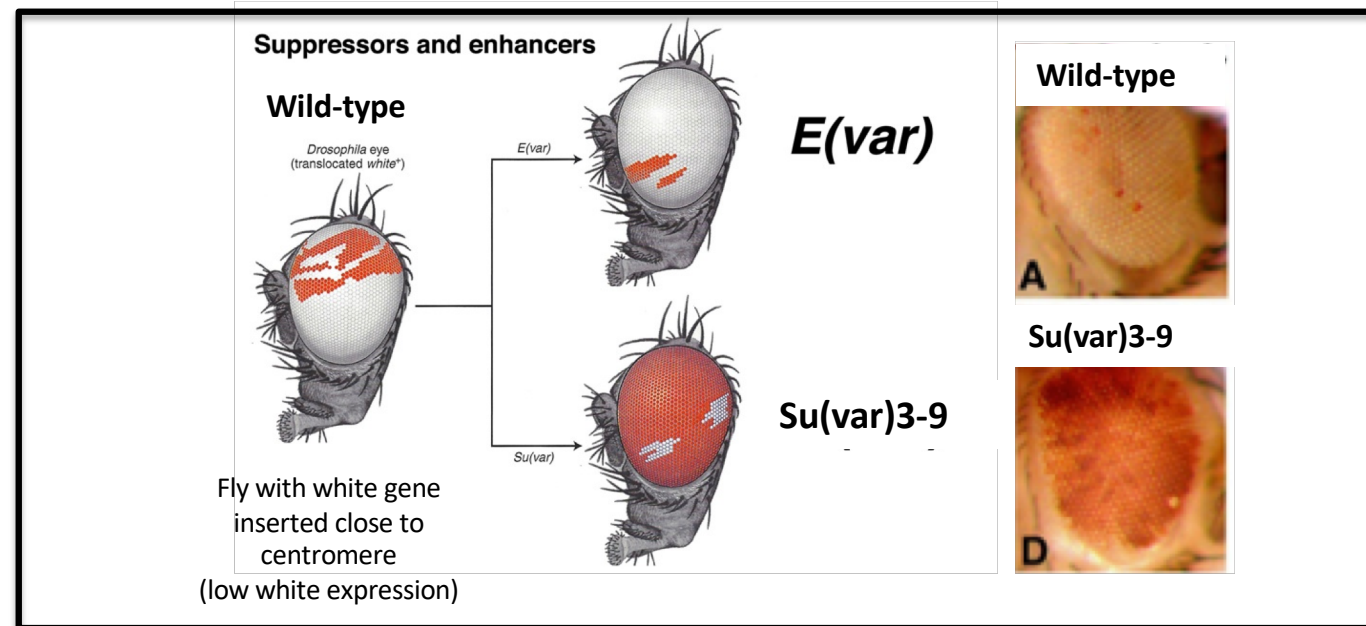
OIS: Oncogene induced senescence

**A historical view: how to understand the enzymatic function of an epigenetic writer**  
**Histones can be acetylated – can they be methylated??**

**Evidence 1**  
**Drosophila**  
**Suv(var)3-9**  
**mutant shows**  
**strong PEV**

## Evidence 2

Suv39h genes show high sequence conservation to a plant gene with proposed histone-methyltransferase activity



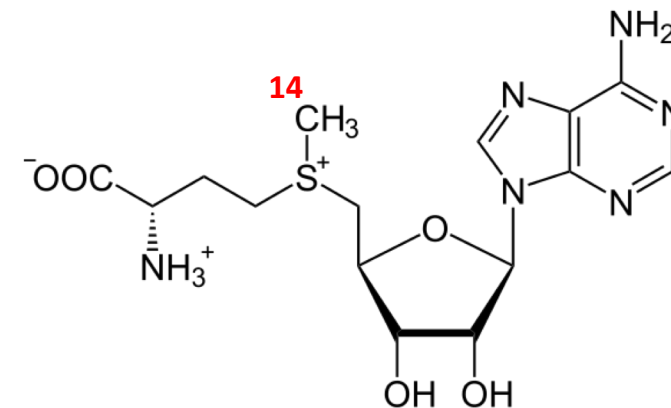
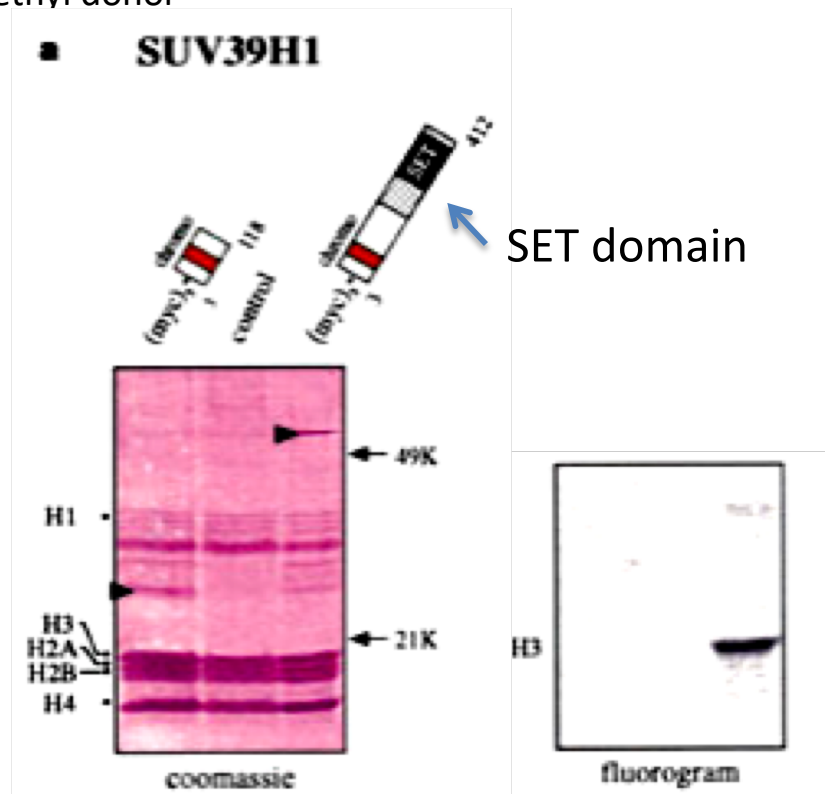
## Identification of H3 Lys9 methyltransferase activity

Experiment:

Overexpression of myc-tagged-SUV39H1 in Hela cells

Use an antibody to immunoprecipitate SUV39H1 → high concentration of SUV39H1

Incubate Immunoprecipitate with purified histones and S-adenosyl-[methyl-<sup>14</sup>C]-L-methionin as methyl donor



SET – domain is required for histone methyl transferases activity

- The SET domain of the SUV39H1 is required for histone methyltransferase activity and this enzyme methylates H3 at Lys9

Nature, 2001 Mar 1;410(6824):116-20.

**Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.**

Lachner M<sup>1</sup>, O'Carroll D, Rea S, Mechtler K, Jenuwein T.

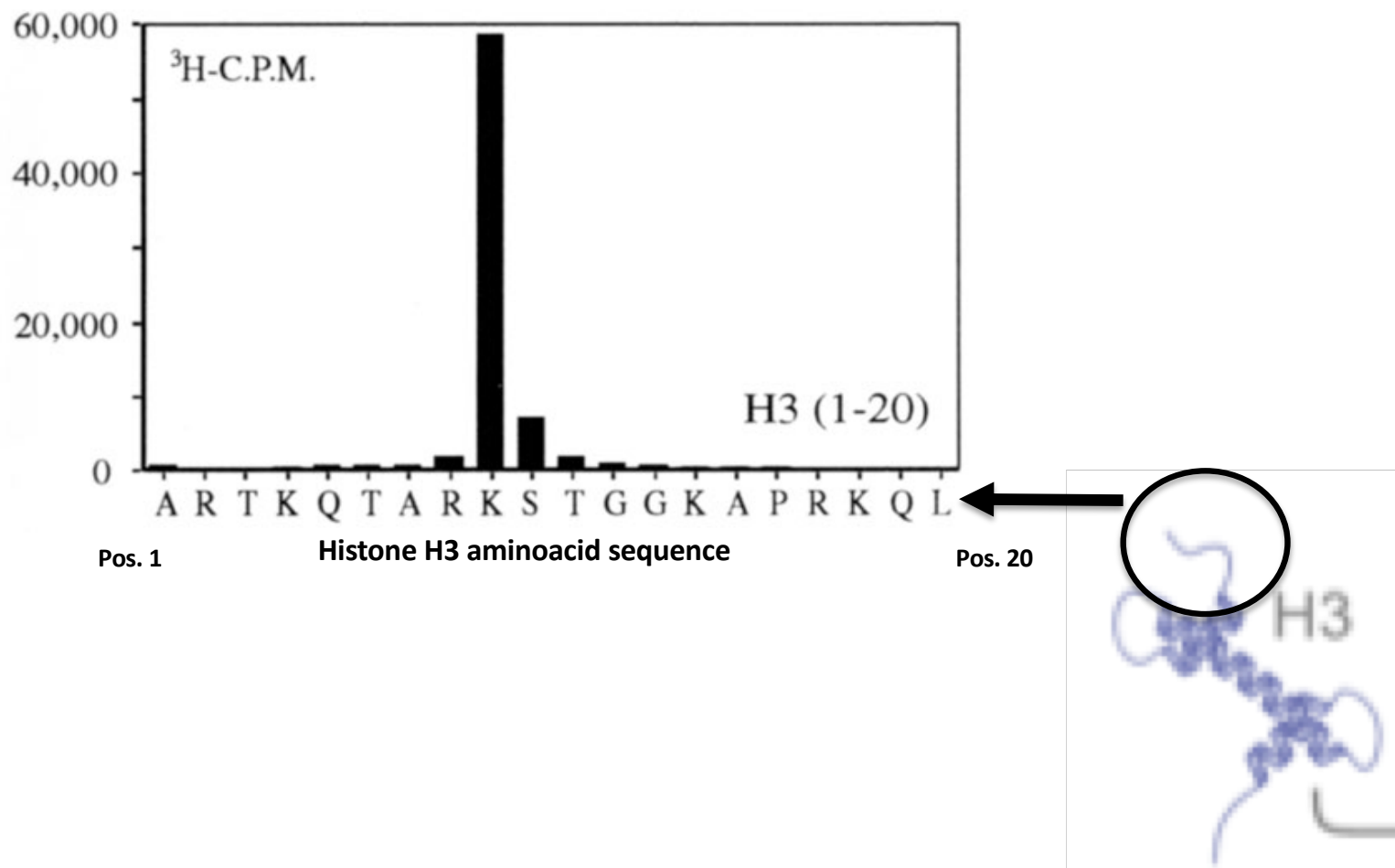
## Identification of H3 Lys9 methyltransferase activity

Experiment:

Purify histone H3 by reverse-phase-liquid chromatography after HMT assay

Sequence histone H3 and quantify the amount of radioactivity per amino acid

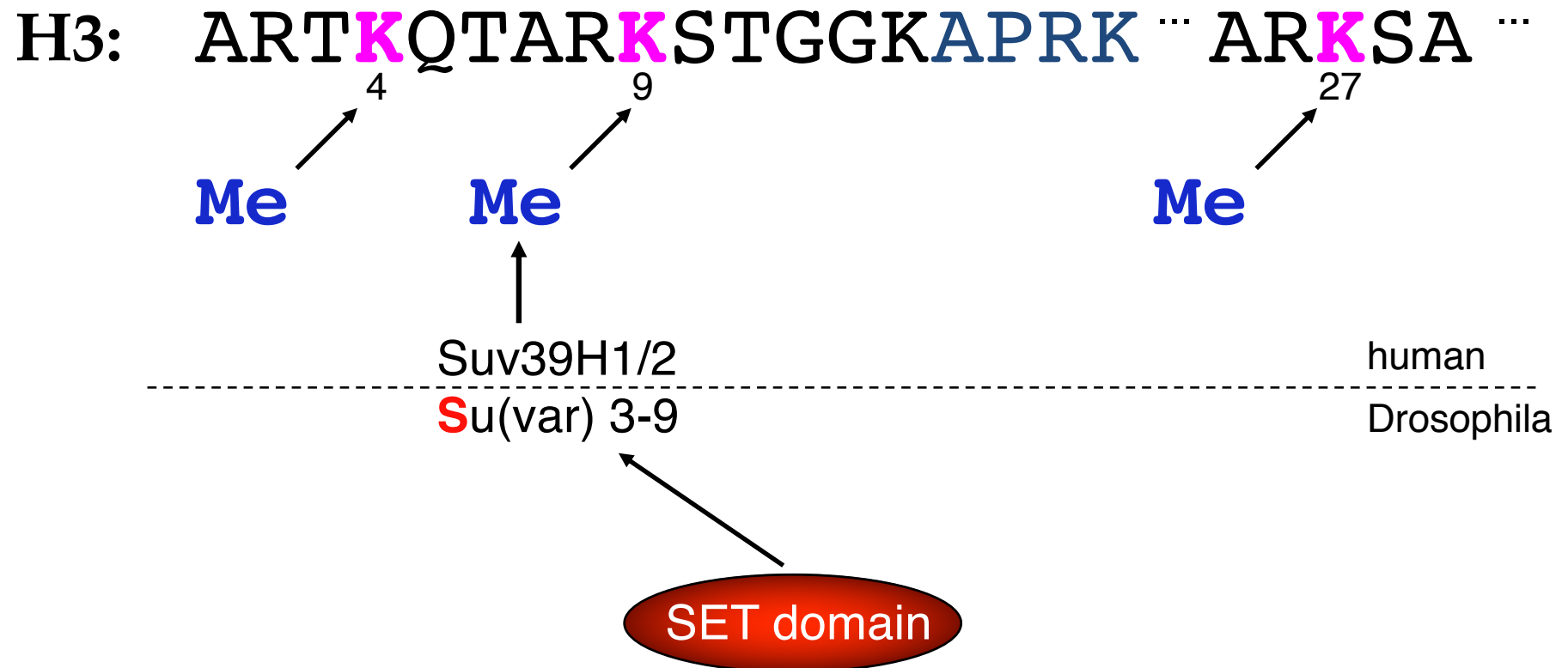
Suv39H1 methylates histone H3 at lysine 9





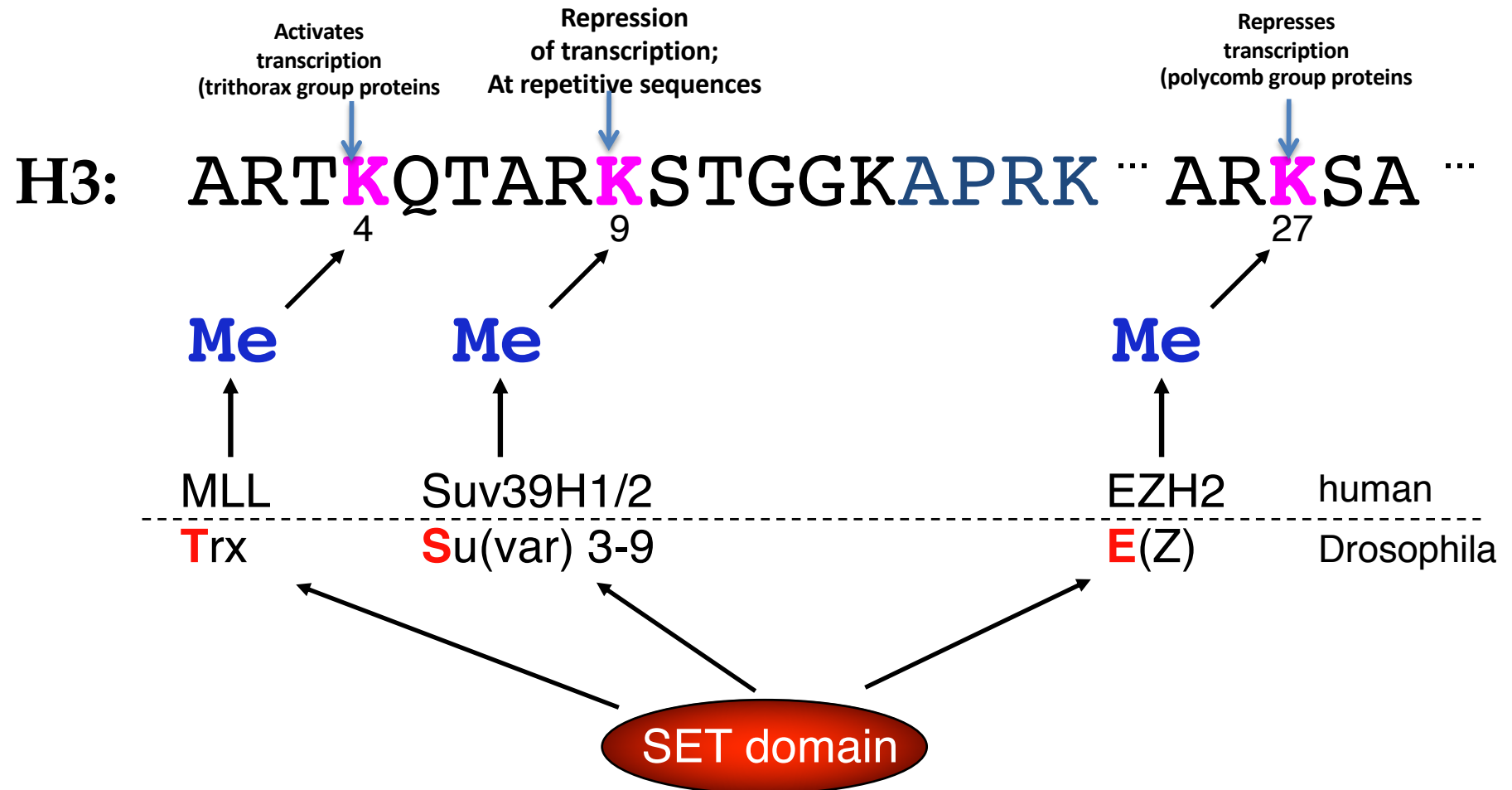
## Identification of other Histone H3 Lys9 methyltransferases

- The SET domain is the conserved catalytic core of histone methyltransferases
- The histone H3 tail has 3 sites for methylation at lysines



## Identification of other Histone H3 Lys9 methyltransferases

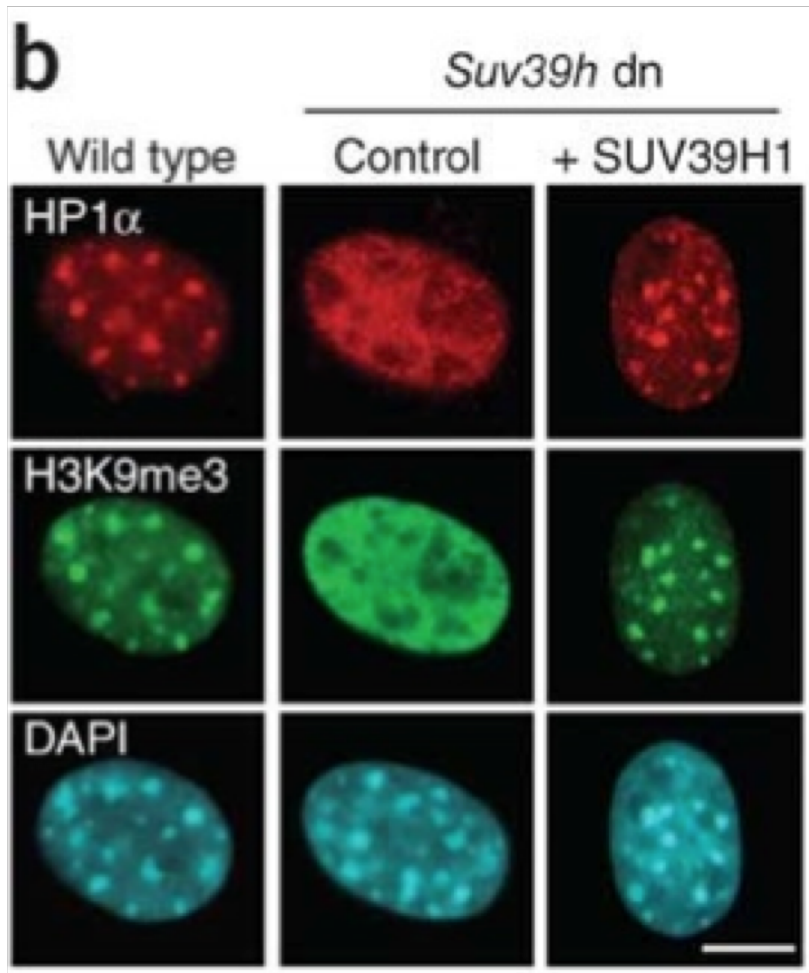
- The SET domain is the conserved catalytic core of histone methyltransferases



- Mutations of some histone methyltransferases → cancer



## Post-translational histone modifications can recruit specialized proteins



### Problem:

1. How can we detect epigenetic modifications?

- Generation of antibodies that specifically recognize modified histones (western blotting/Immunofluorescenza)

- Mass spectroscopy

### Problem:

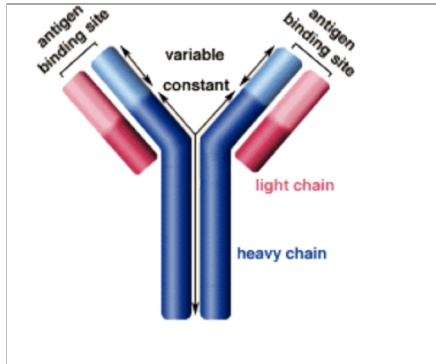
2. How can we locate epigenetic information at defined regions or genes or promoters, etc?

- Chromatin immunoprecipitation

- > Detect histone modifications on single genomic site

- > Detect histone modifications on multiple sites or on the entire genome level

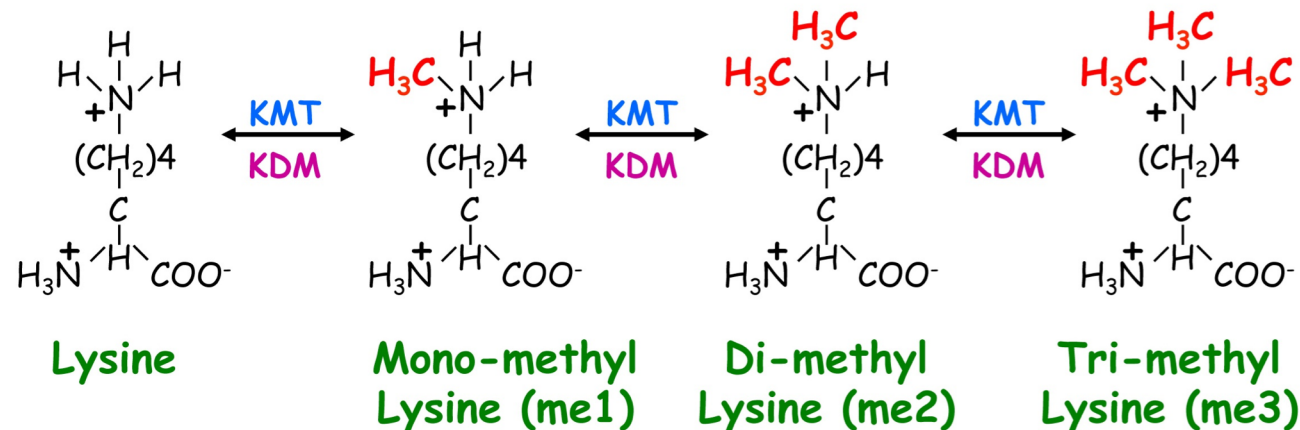
# 1. Generation of antibodies that specifically recognize modified histones



CAN WE GENERATE MODIFICATION SPECIFIC ANTIBODIES??

Lysines can be mono-, di- and tri-methylated  
An example: H3K9methylation by Suv39h1

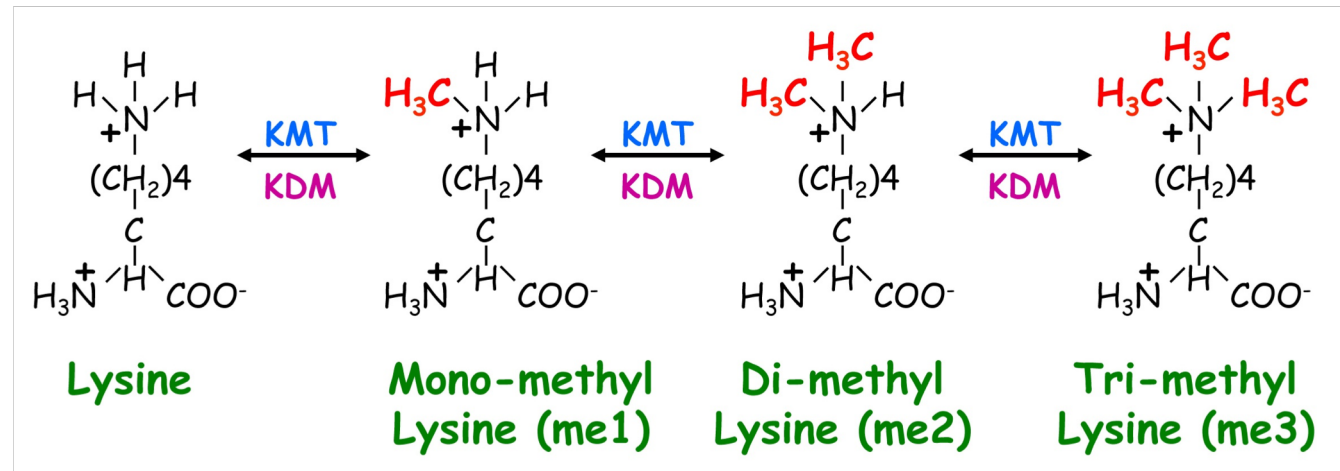
H3K9-specific  
Lysine-HMTases  
(KMTases) can  
mediate mono-,  
di-, tri-methylation



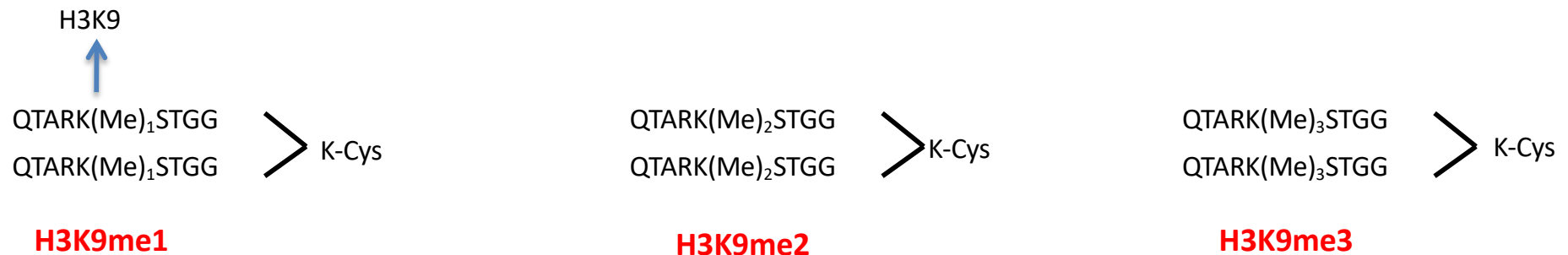
# 1. Generation of antibodies that specifically recognize modified histones

**Lysines can be mono-, di- and tri-methylated**

**H3K9-specific  
Lysine-HMTases  
(KMTases) can  
mediate mono-,  
di-, tri-methylation**



## SYNTHESIS OF BRANCHED PEPTIDES FOR ANTIBODY GENERATION

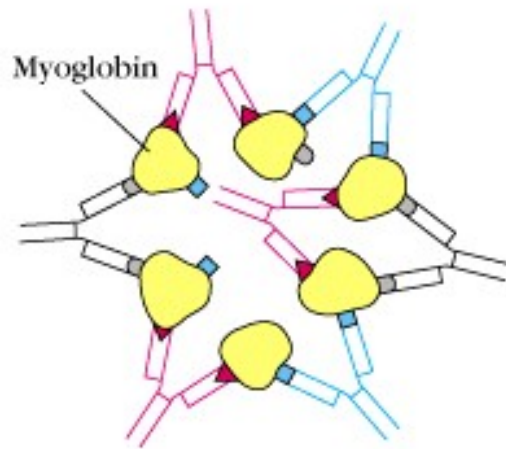


→ Branching allows to place histone modifications in close vicinity  
 → → resembles high concentration of chromatin modification in the nucleus

## Generation of polyclonal/monoclonal antibodies

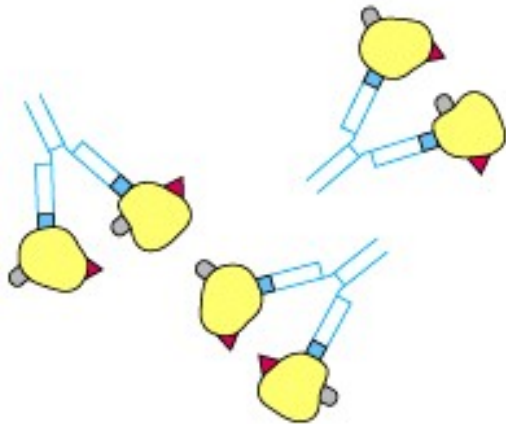
(a)

POLYCLONAL ANTISERUM



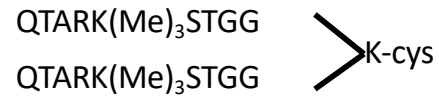
Polyclonal antibodies can form lattices with homogeneous, monomeric protein antigens because each antibody can interact with a different epitope on the antigen.

MONOCLONAL ANTIBODY

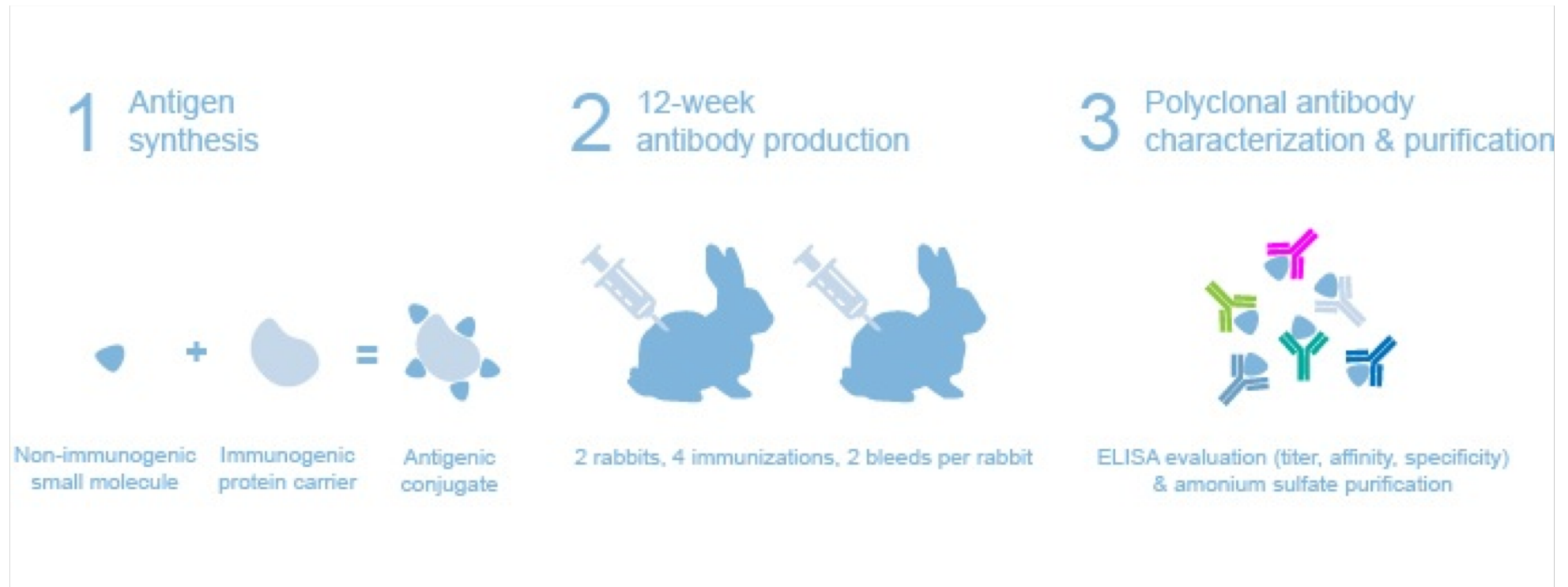


Monoclonal antibodies do not form lattices with homogeneous, monomeric proteins, because only they can bind to only one epitope on the antigen.

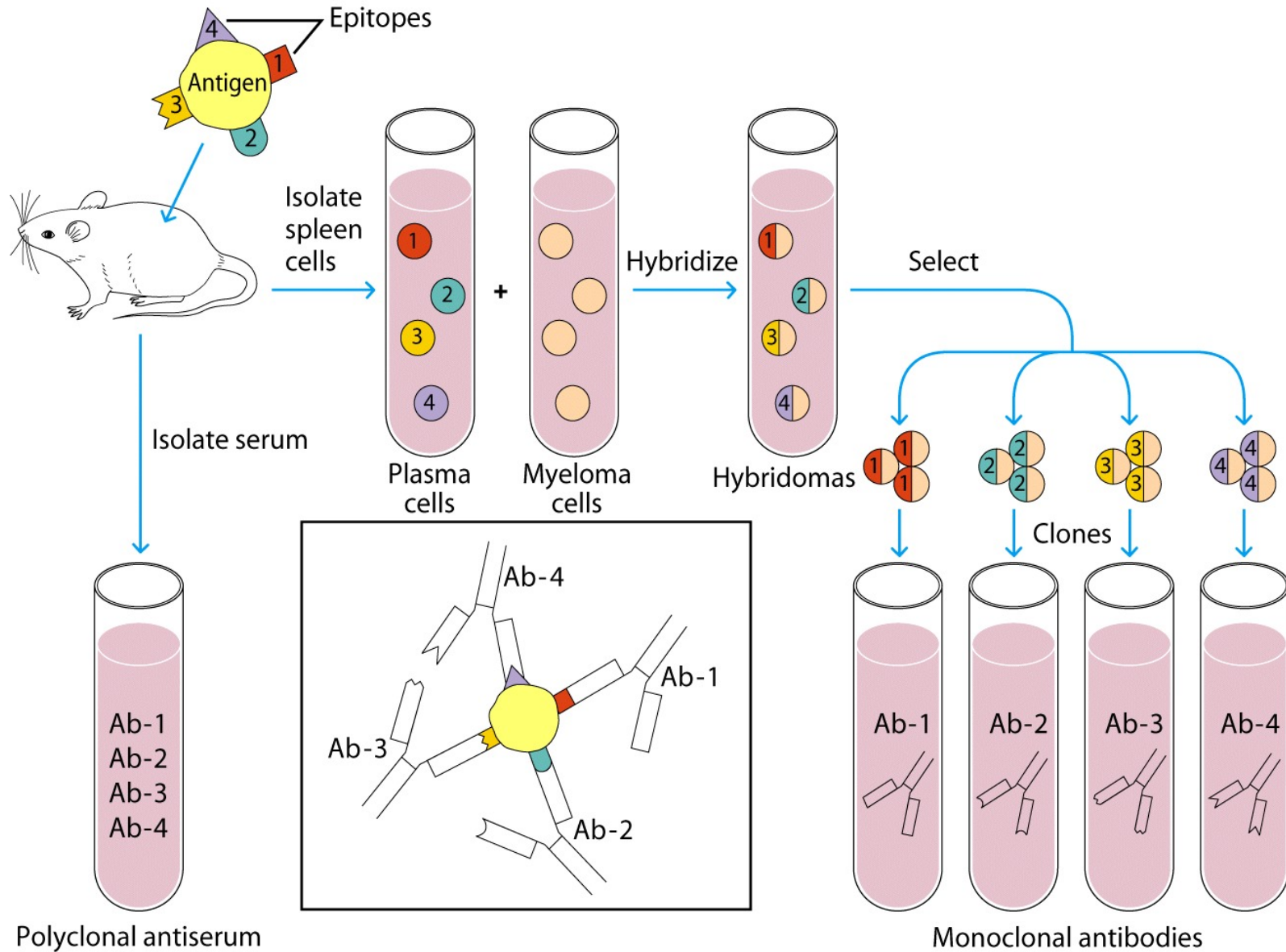
## Generation of polyclonal antibodies



**H3K9me3**



## Generation of monoclonal antibodies





## Generation of monoclonal antibodies

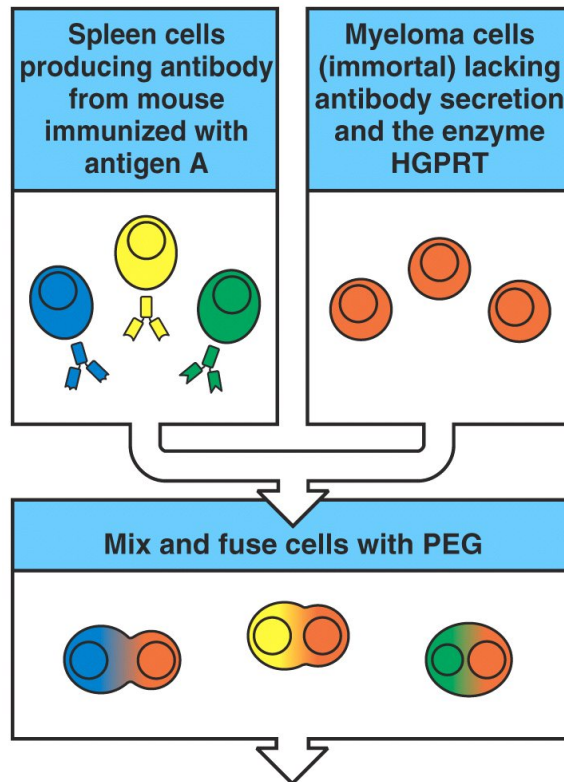


Figure A-14 part 1 of 2 Immunobiology, 6/e. (© Garland Science 2005)

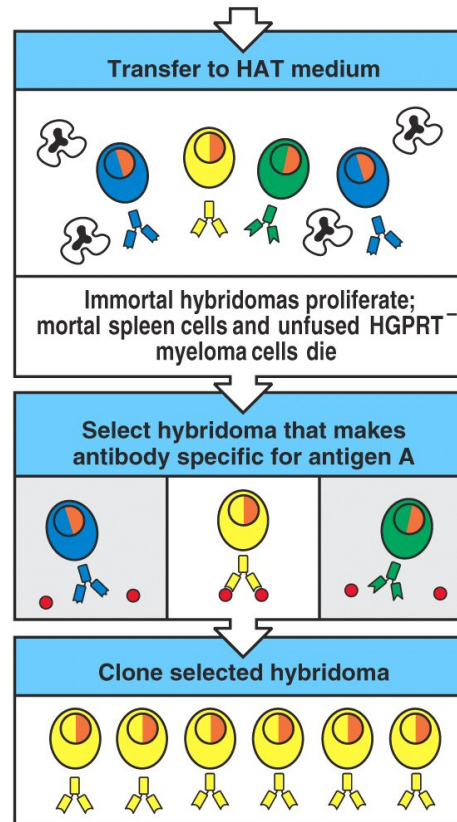
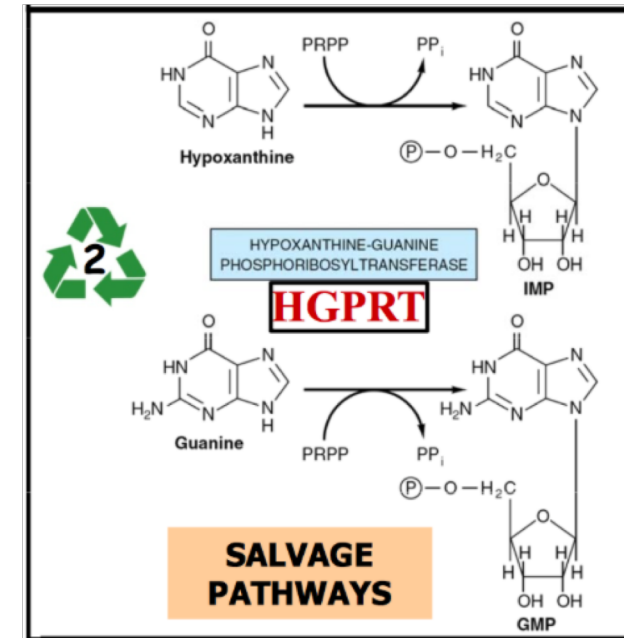


Figure A-14 part 2 of 2 Immunobiology, 6/e. (© Garland Science 2005)



Most purines are recycled rather than degraded. HGPRT can recycle hypoxanthine and guanine to be used in DNA replication

HAT Medium (hypoxanthine-aminopterin-thymidine medium) is a selection medium for mammalian cell culture, which relies on the combination of **aminopterin**, a drug that acts as a powerful folate metabolism inhibitor by inhibiting dihydrofolate reductase, with **hypoxanthine** (a purine derivative) and **thymidine** (a deoxynucleoside) which are intermediates in DNA synthesis. **The trick is that aminopterin blocks DNA de novo synthesis, which is absolutely required for cell division to proceed, but hypoxanthine and thymidine provide cells with the raw material to evade the blockage (the "salvage pathway"), provided that they have the right enzymes, which means having functioning copies of the genes that encode them.** HGPRT: inactivates aminopterin → rescue of DNA replication

HAT medium is often used for preparation of monoclonal antibodies. This process is called Hybridoma technology. Laboratory animals (e.g., mice) are first exposed to an antigen against which we are interested in isolating an antibody. Once splenocytes are isolated from the mammal, the B cells are fused with HGPRT negative, immortalized myeloma cells using polyethylene glycol or the Sendai virus. Fused cells are incubated in the HAT medium. Aminopterin in the medium blocks the de novo pathway. Hence, unfused myeloma cells die, as they cannot produce nucleotides by de novo or salvage pathway. Unfused B cells die as they have a short lifespan. In this way, only the B cell-myeloma hybrids survive. These cells produce antibodies (a property of B cells) and are immortal (a property of myeloma cells). The incubated medium is then diluted into multiwell plates to such an extent that each well contains only 1 cell. Then the supernatant in each well can be checked for desired antibody. Since the antibodies in a well are produced by the same B cell, they will be directed towards the same epitope, and are known as monoclonal antibodies.



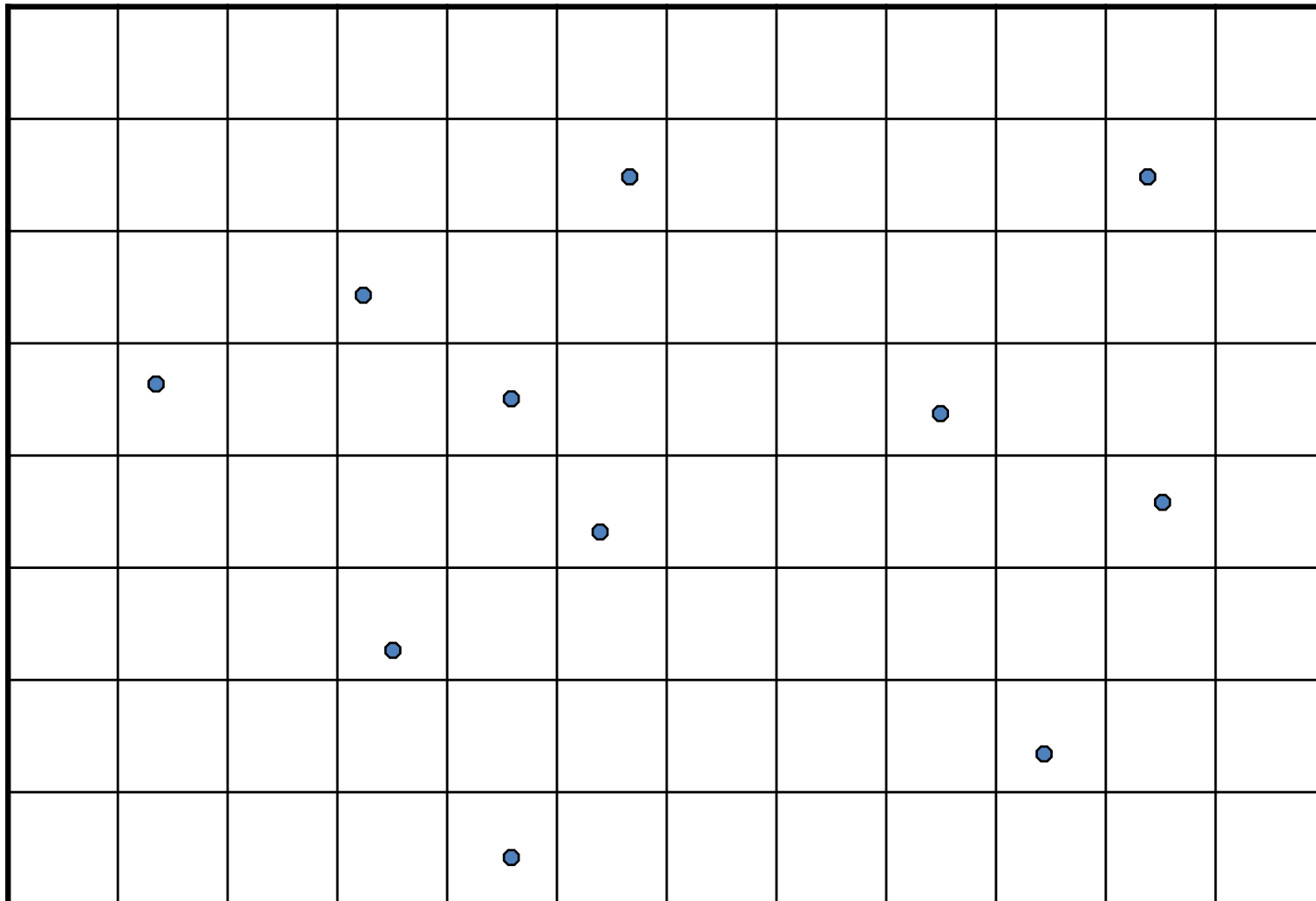
## Generation of monoclonal antibodies

Cloning hybridomas from fusion

Plate at limiting dilution ( $<1$  cell/well) in 96 well plates.

Allow clones to expand.

Expand positive well and test for production of antibody of desired specificity in culture supernatant



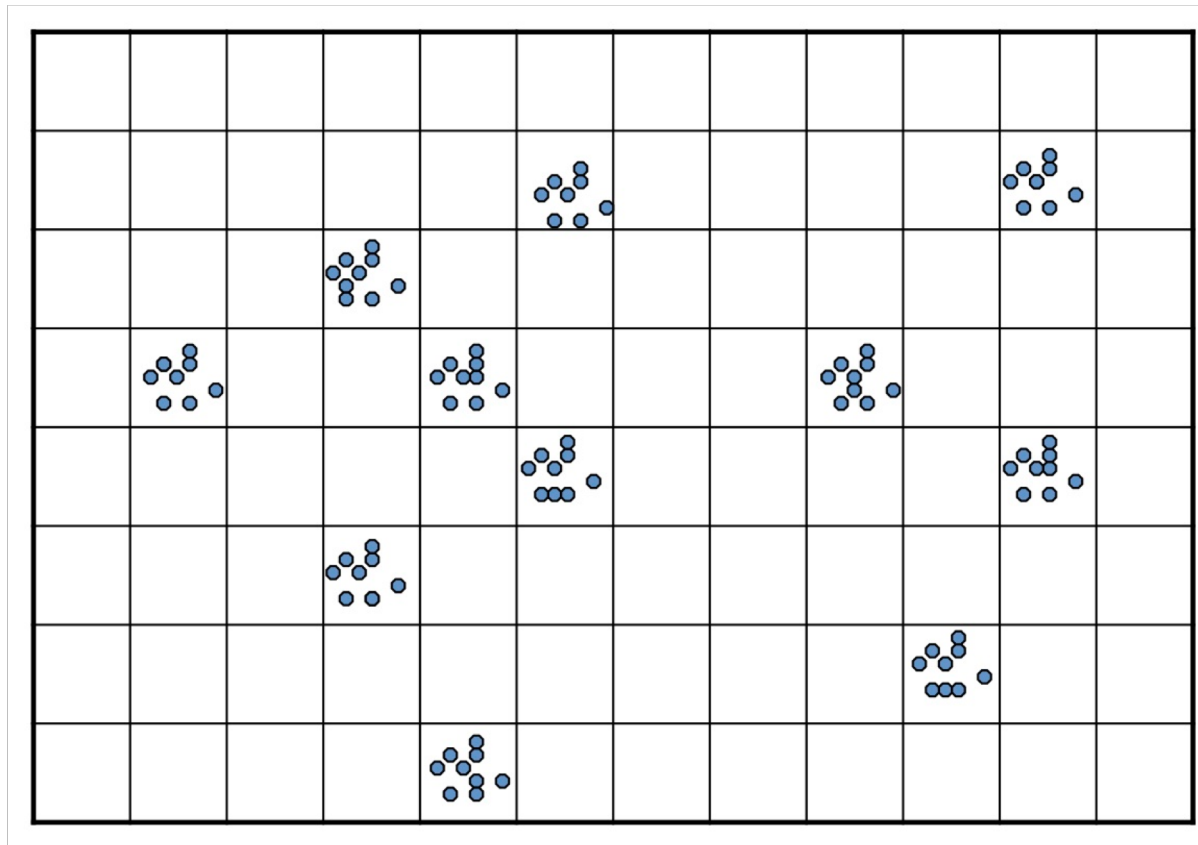
## Generation of monoclonal antibodies

Cloning hybridomas from fusion

Plate at limiting dilution ( $<1$  cell/well) in 96 well plates.

Allow clones to expand.

Expand positive well and test for production of antibody **of desired specificity in culture supernatant**

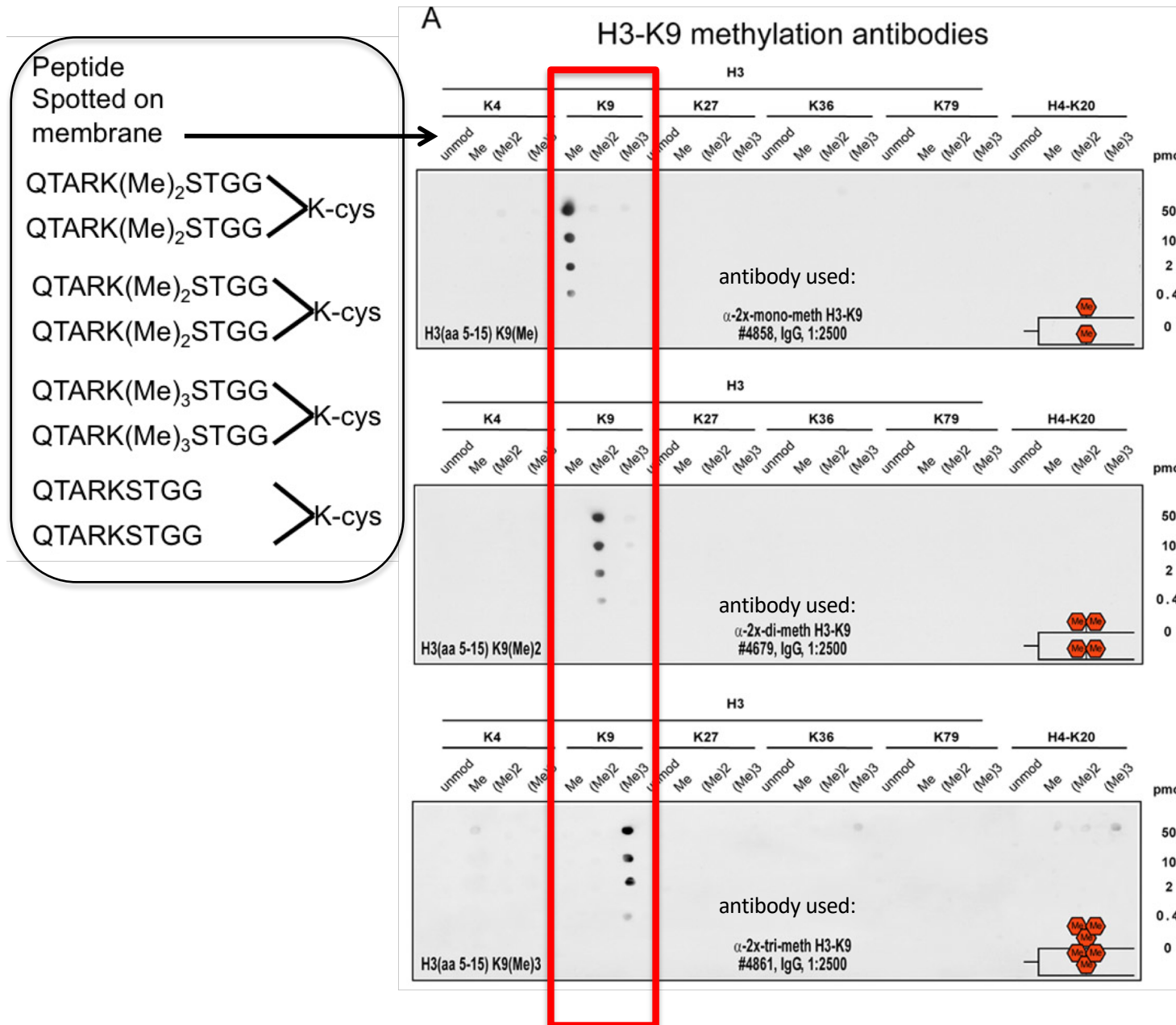


-Hybridoma cells secrete antibody into cell culture medium

-Purification of antibody

-Validation of antibody

## Validation of antibodies

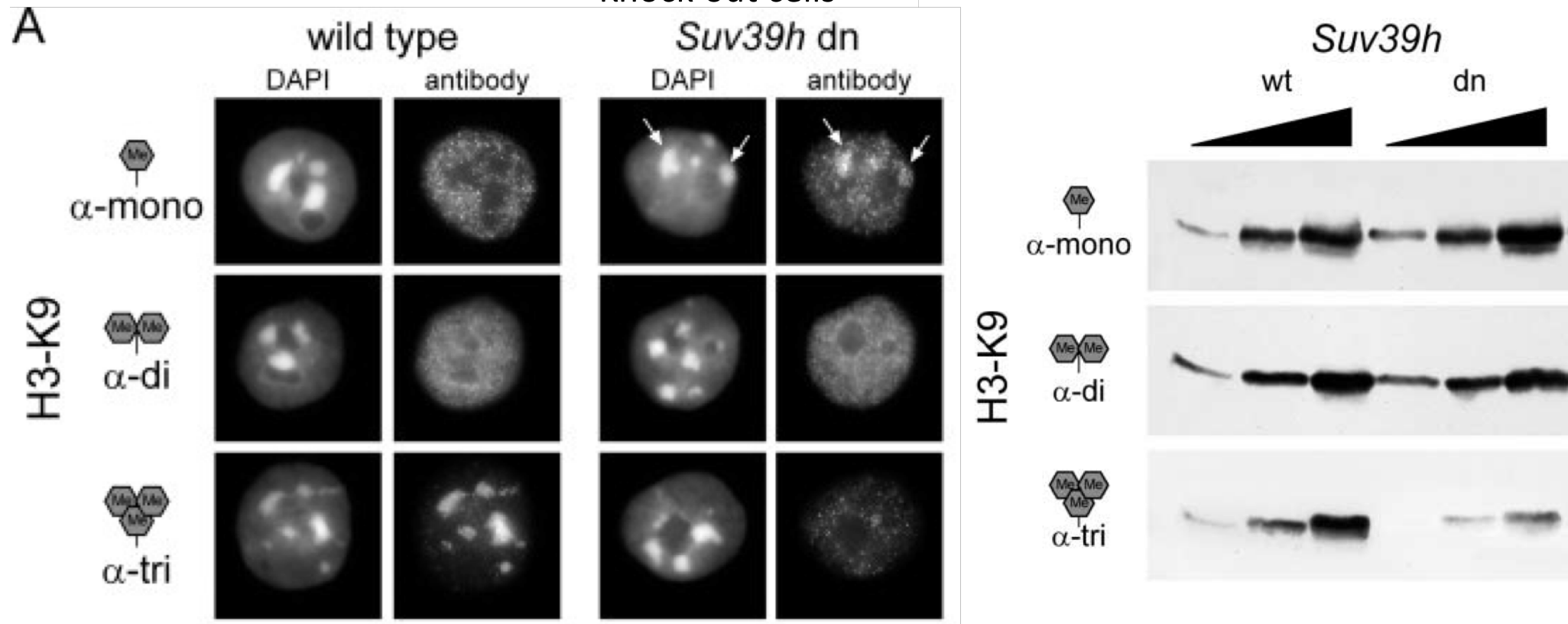


**Validation by dot-blot**  
(pure antigen spotted on membrane)

1. Spot branched peptides at different concentration on membrane
2. Incubate with respective Antibody
3. Incubate with secondary Antibody that is coupled With horseradish peroxidase
4. Add substrate for Horseradish peroxidase
5. Develop

## Obtaining evidence for the specificity of histone modifying enzymes → in cells

Suv39h1/Suv39h2  
Knock-out cells



In *Suv39h* dn cells, H3K9me1 and H3K9me2 are still present.  
*Suv39h1* is a H3K9 specific HMTase that is required to establish the TRI-METHYLATION of H3K9 in mammalian cells

Note, that in the absence of H3K9me3, H3K9me1 is localized at DAPI rich regions

## 2. Studying histone modifications by mass spectroscopy



In a typical MS procedure proteins are ionized, for example by bombarding it with electrons. This may cause some of the sample's molecules to break into charged fragments.

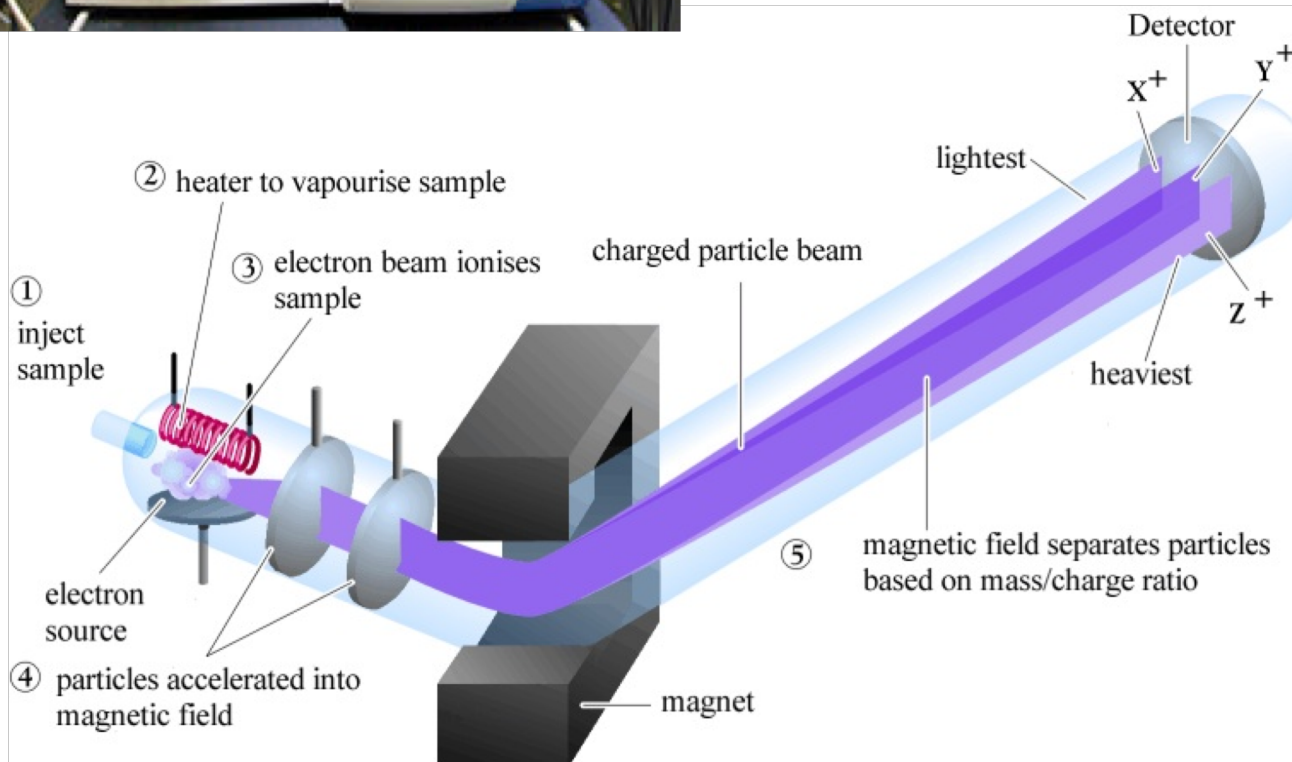
These ions are then separated according to their mass-to-charge ratio, typically by accelerating them and subjecting them to an electric or magnetic field:

Ions of the same mass-to-charge ratio will undergo the same amount of deflection. Ions with different mass-to-charge ratio will show different deflection

→ **mono-methylated H3K9 has different deflection than di- or tri-methylated H3K9**

The ions are detected by a mechanism capable of detecting charged particles, such as an electron multiplier. Results are displayed as spectra of the relative abundance of detected ions as a function of the mass-to-charge ratio.

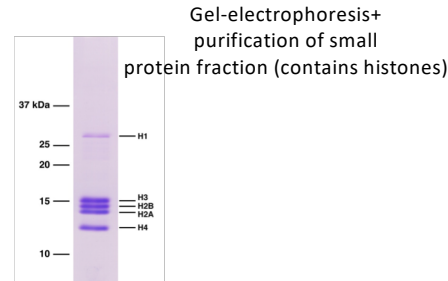
The molecules in the sample can be identified by correlating known masses to the identified masses or through a characteristic fragmentation pattern.



## 2. Studying histone modifications by mass spectrometry

Wild-type or  
Suv39h dn cells

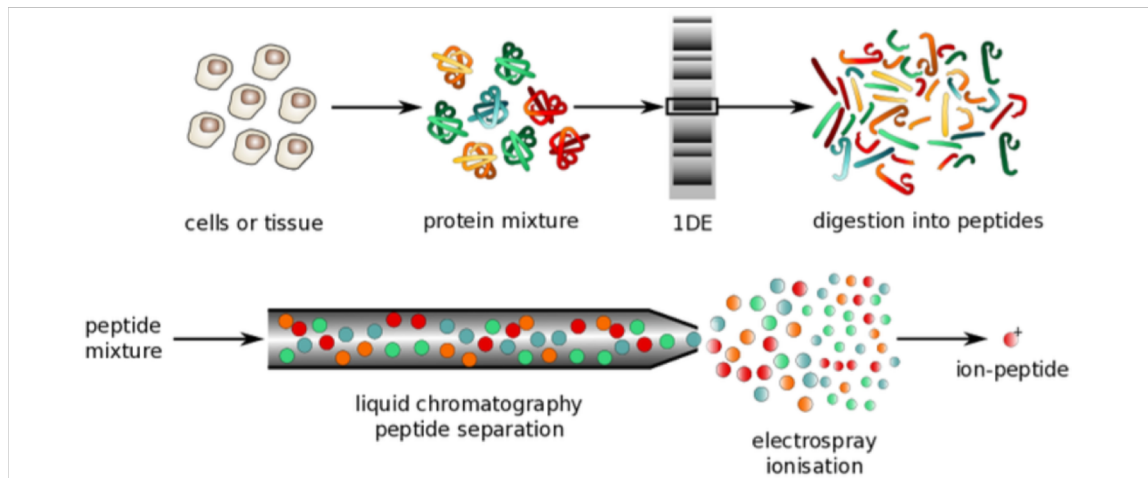
nuclear  
extract



Digestion of fraction  
containing histones with proteases  
creates a “library” of  
small peptides that are  
derived from histones:

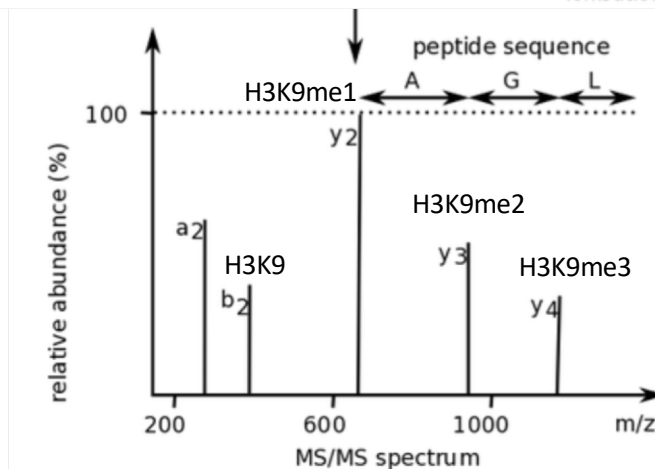
**Protease= Trypsin**

**Trypsin cleaves after every  
Lysine (K) and Arginine (R)**

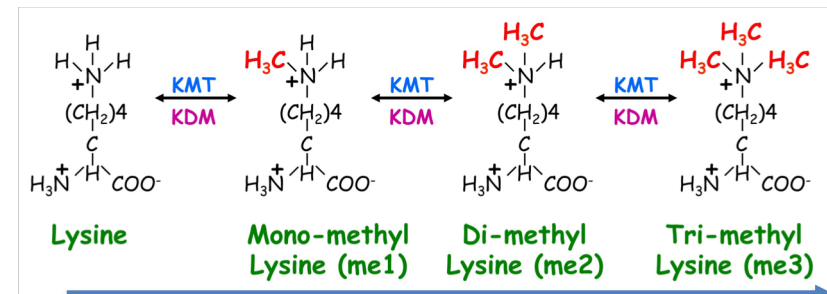


IMPORTANT: we know already  
All proteins and the amino acid  
sequence of all proteins of a cell!!!  
**That means we can predict all  
possible small peptide sequences  
that result from a trypsin cleavage**

**MS**



Mass/charge ratio



Different mass/charge ratio



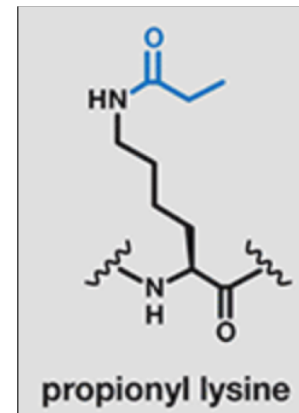
## 2. Studying histone modifications by mass spectroscopy



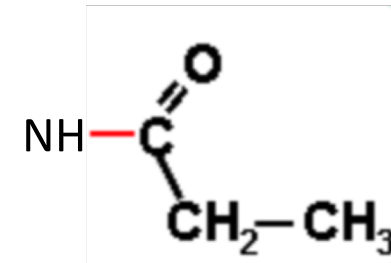
Trypsin can cleave after Arg (R)

Trypsin can cleave after K-unmethylated; K-me1

Trypsin cannot cleave after K-me2; K-me3



Propionyl  
group



### PROBLEM:

1. Trypsin cleavage after K is not complete because me2 or me3 on some Ks block cutting by Trypsin

2. Many cleavages (me0, me1) → small heterogeneous peptides → difficult to analyze

SOLUTION: In vitro Propionylation of unmodified or mono-methylated K prevents trypsin cleavage.

Now Trypsin can only cut after Arginine. This allows a uniform Cleavage of histone tails

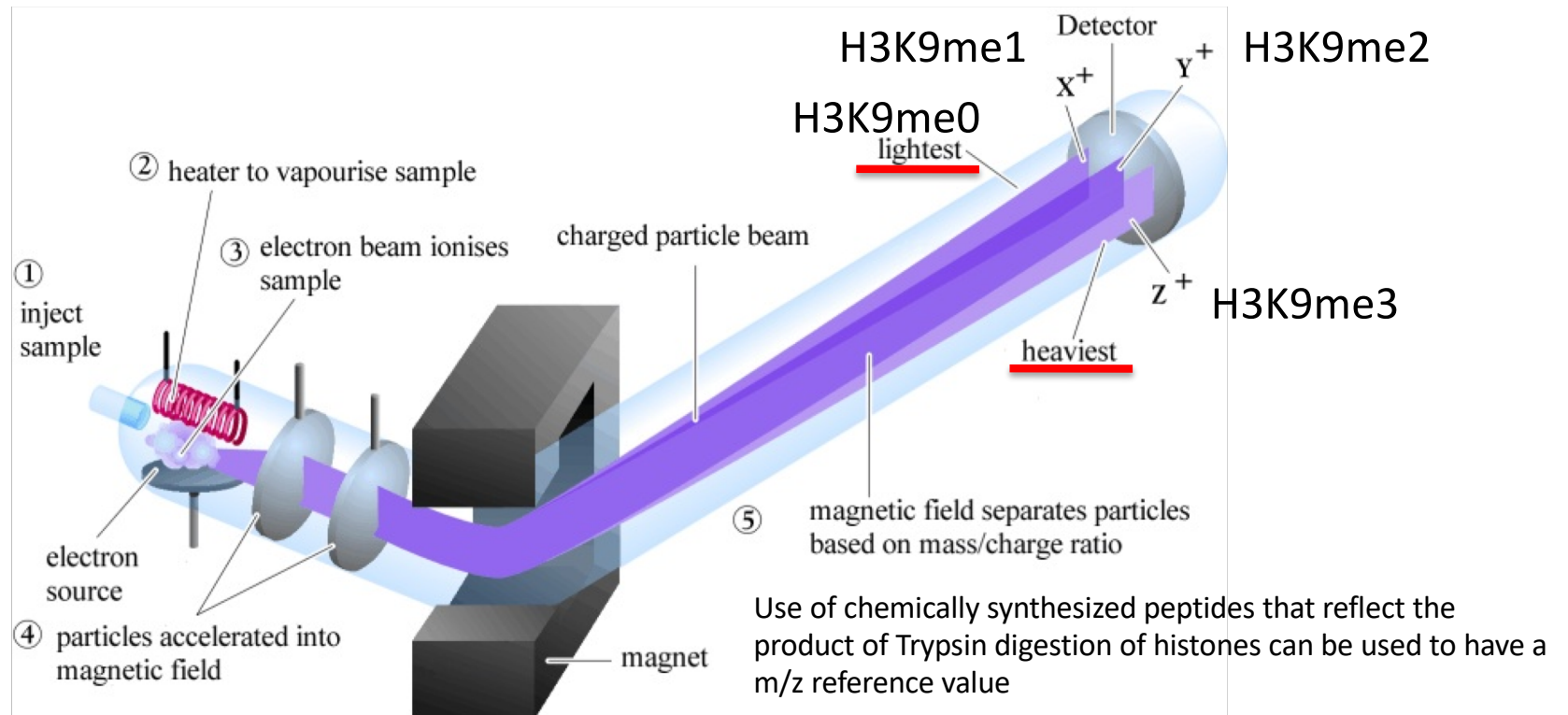




Non-Compete  
C  
T  
p  
ir  
P  
lr  
C

fic

## 2. Studying histone modifications by mass spectroscopy



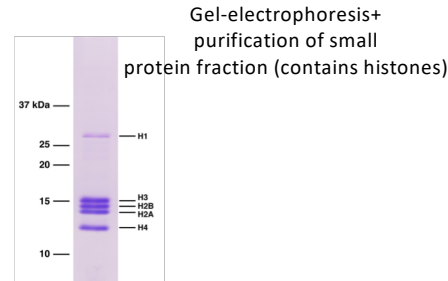
The method allows to quantify the numbers of H3K9me0, H3K9me1 H3K9me2 and H3K9me3 in a sample → We can calculate the % of each histone modification in the given sample.

**Important: The analysis is not limited to H3K9 → other histone modifications can be quantified in the same analysis**

## 2. Studying histone modifications by mass spectroscopy

Wild-type or  
Suv39h dn cells

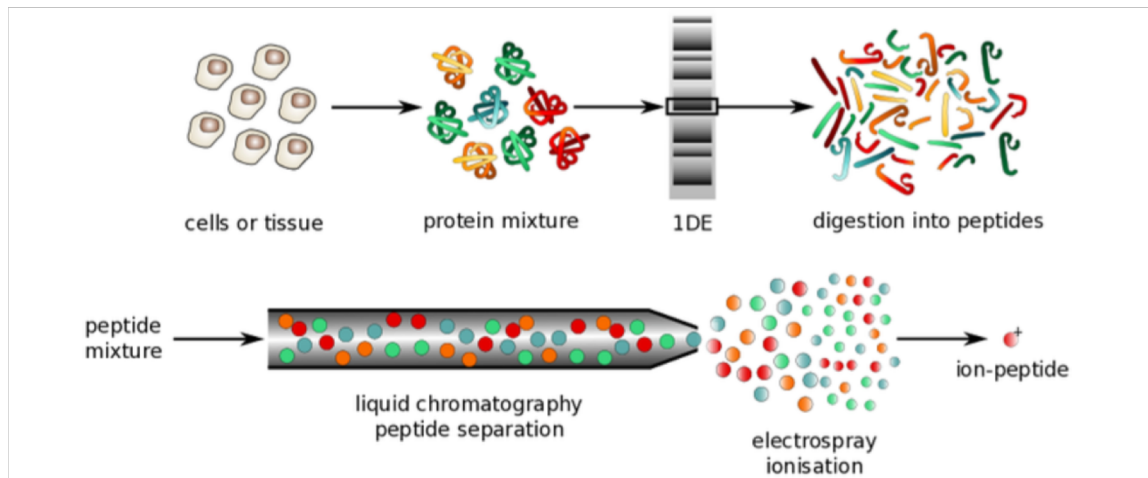
nuclear  
extract



Digestion of fraction  
containing histones with proteases  
creates a “library” of  
small peptides that are  
derived from histones:

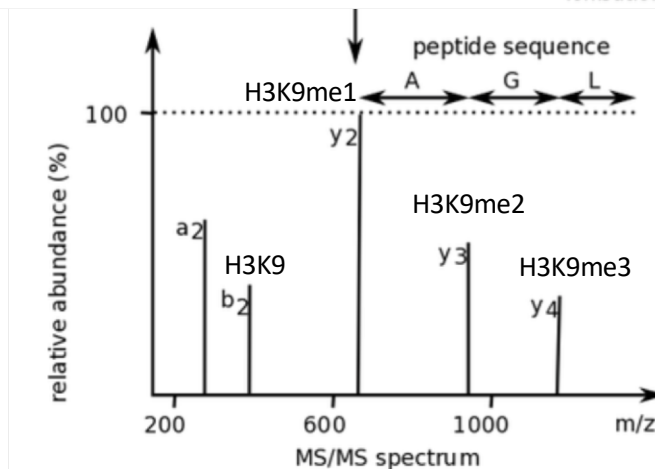
**Protease= Trypsin**

**Trypsin cleaves after every  
Lysine (K) and Arginine (R)**

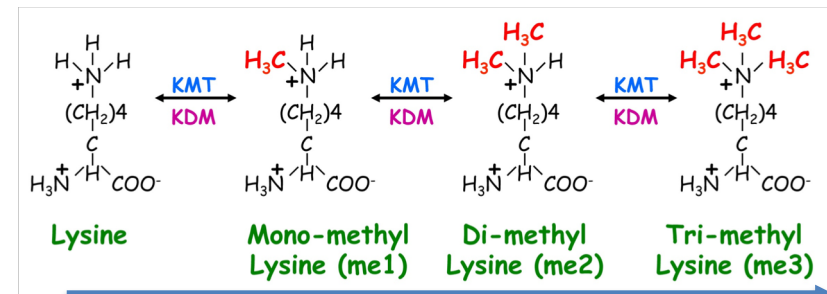


IMPORTANT: we know already  
All proteins and the amino acid  
sequence of all proteins of a cell!!!  
**That means we can predict all  
possible small peptide sequences  
that result from a trypsin cleavage**

**MS**

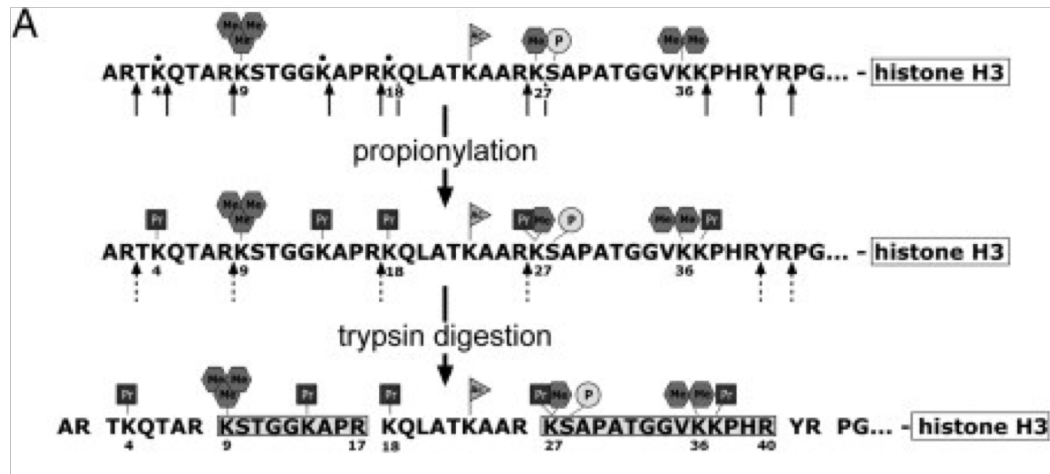


Mass/charge ratio



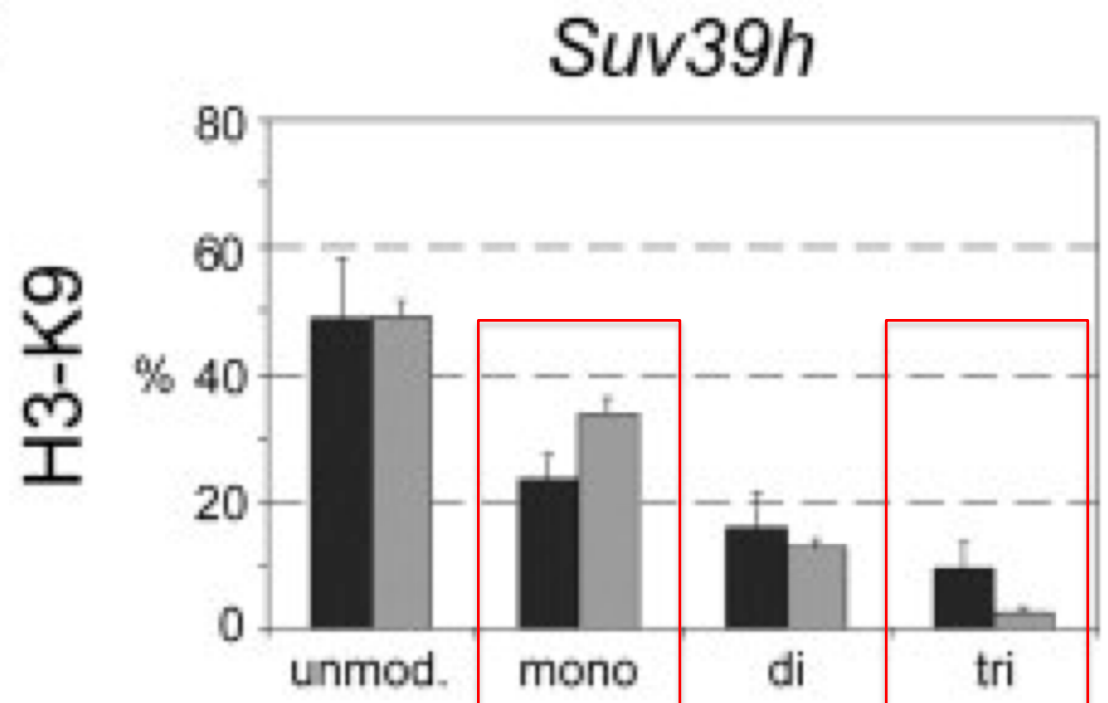
Different mass/charge ratio

## 2. Quantifying histone modifications by mass spectroscopy

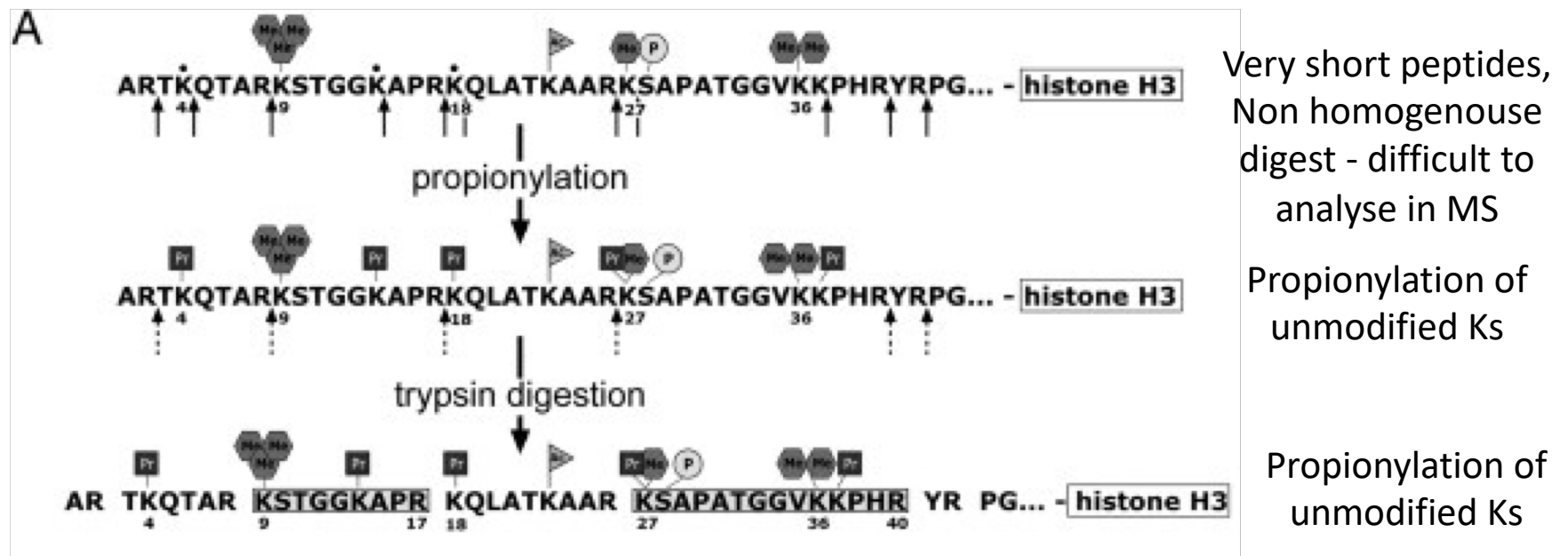


Suv39h1 and Suv39h2 ensure H3K9me3.

H3K9me1 is a preferred substrate for Suv39h1/h2  
Suv39h1/h2 methylate H3K9me1  
Until reaching the tri-methylated state



## 2. Quantifying histone modifications by mass spectroscopy

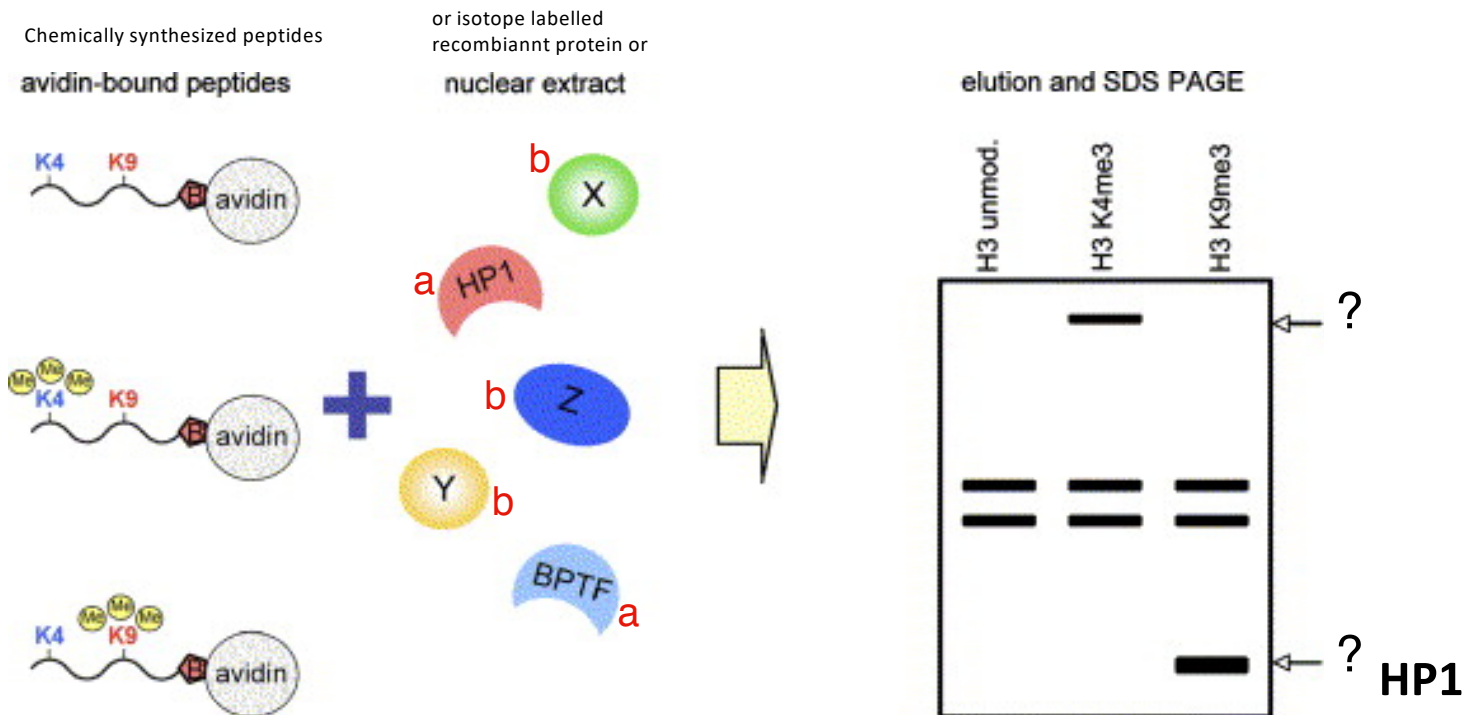


1. Identifying the enzymatic activity of a histone modifying enzymes
2. Identification of “modifiable” aminoacids in histone tails
3. Functional link between histone modifications
4. Defining the ENTIRE epigenetic status of a cell type (for example: differentiated/stem cell)
5. Quantitative information on histone modifications (%)

## HOW CAN EPIGENETIC READERS BE IDENTIFIED?

### Identifying methyl-H3 binding proteins

- histone peptide pulldown assay:



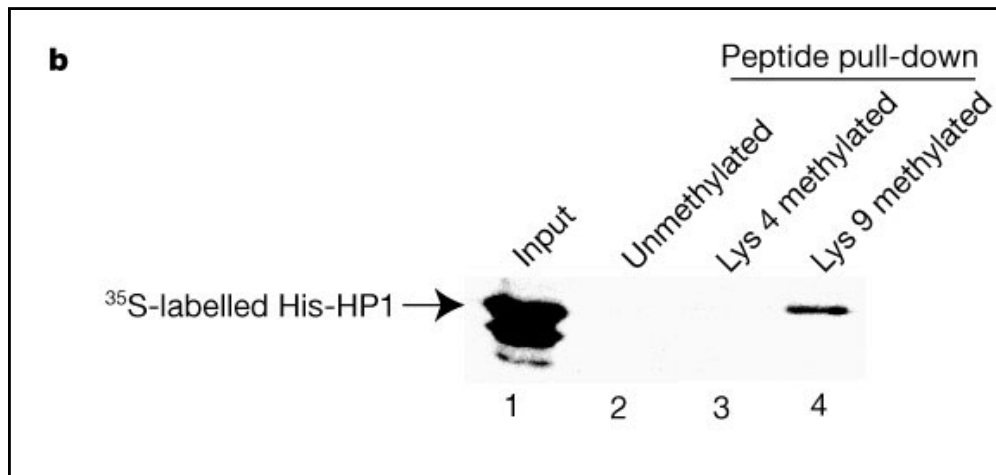
a = candidate approach → identify by Western blotting

b = unbiased approach → identify by Mass Spec



## What are the target sites for Suv39h1 and H3K9me???

- Using the peptide pull-down assay, it was found that Lys9-methylated H3 binds to heterochromatin protein 1 (HP1)



### Pull-down assay:

- Couple modified histone Tail-peptides (methylated, or unmethylated) to resin
- Incubate with recominant  $^{35}\text{S}$ -labelled HP1
- Wash resin
- Elute bound proteins, run gel and make radiography

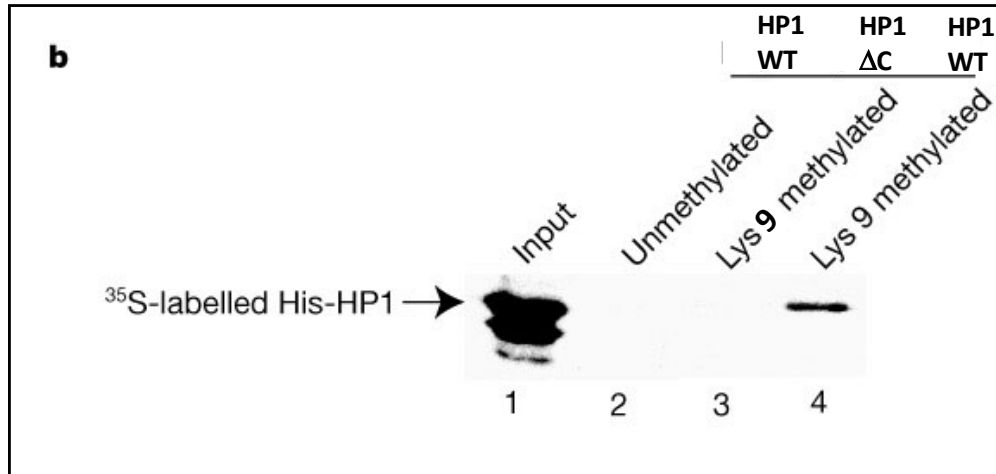
- HP1 is a protein previously identified to be enriched in and important for heterochromatin assembly
- Lys9-methylated H3 binds to HP1 via the chromodomain motif in HP1

**Lys9-methylated H3 binds to the conserved motif called chromodomain**



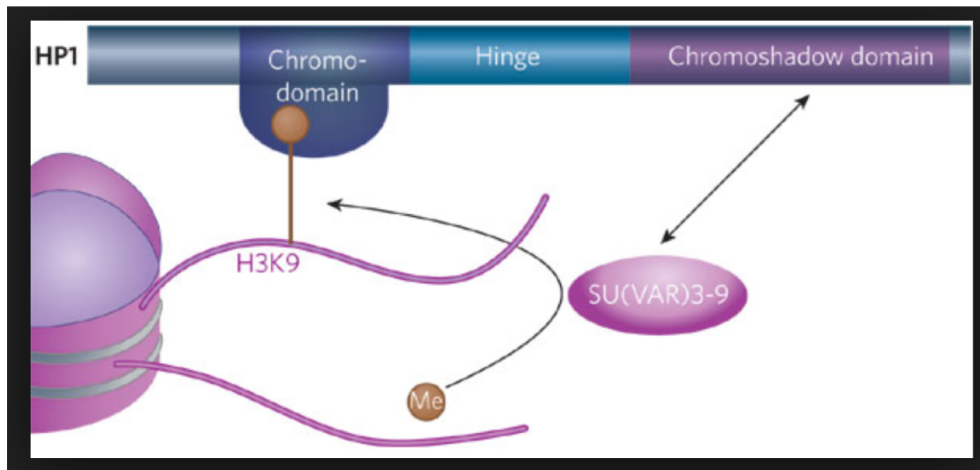
## What are the target sites for Suv39h1 and H3K9me???

- What protein domain is interacting with H3K9me?



### Pull-down assay:

- Couple modified histone Tail-peptides (methylated, or unmethylated) to resin
- Incubate with recominant  **$^{35}\text{S}$ -labelled mutant versions of HP1**
- $-\Delta C$ ;  $\Delta CS$ ;  $\Delta H$**
- Wash resin
- Elute bound proteins, run gel and make radiography

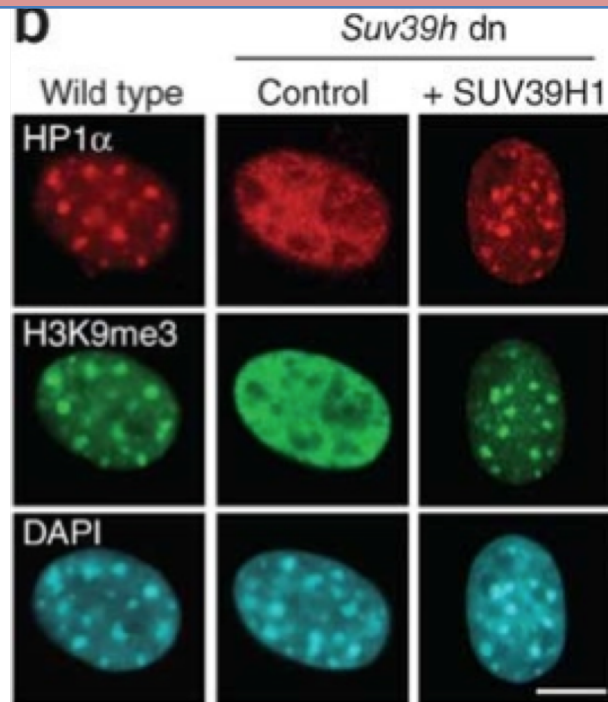


**Lys9-methylated H3 binds to the conserved motif called chromodomain**

## Where does Suv39h1 act in a cell ???

IMMUNOFLUORESCENCE → MACROSCOPIC ANALYSIS

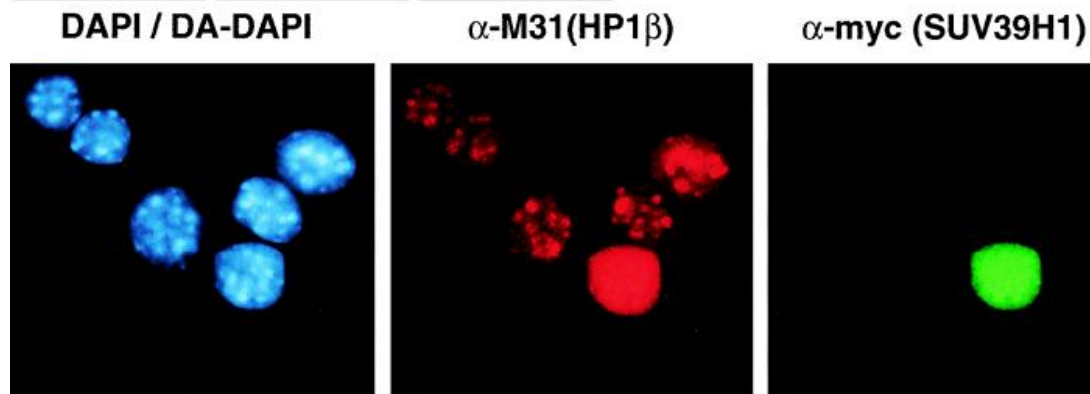
→ Localization of protein across large regions of DNA



Loss of Suv39h1/h2 causes delocalization  
Of HP1 from chromocenters

Chromocenters are subnuclear regions where  
Multiple centromeres aggregate

Chromocenters are DAPI intense regions  
and represent tightly packed heterochromatin

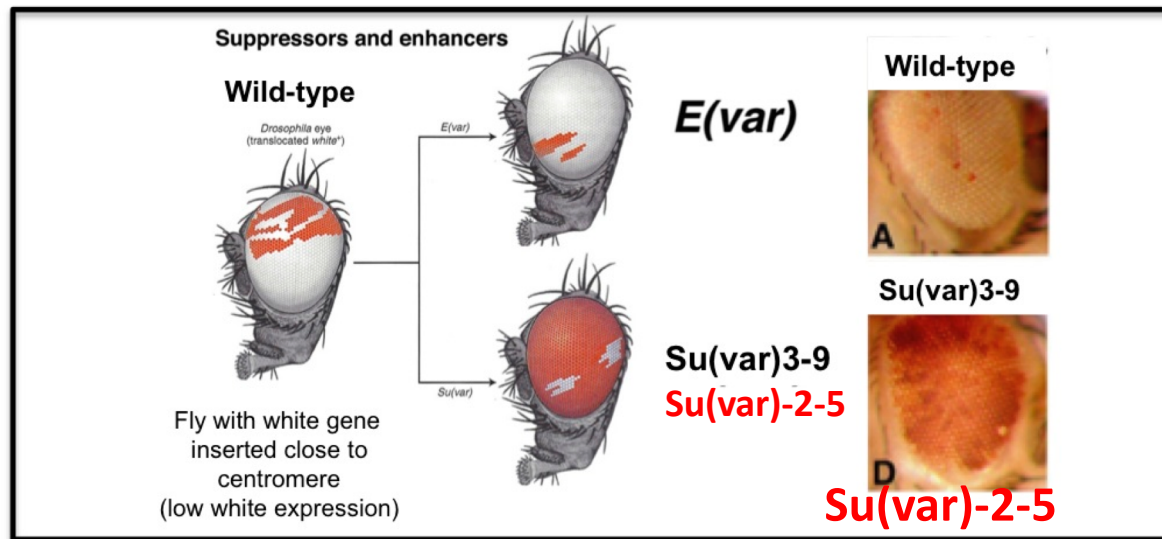


Over-expression of  
SUV39H1 in Hela cells  
causes an accumulation  
of HP1. Is there a link between  
Suv39h1 – H3K9me3 and HP1??

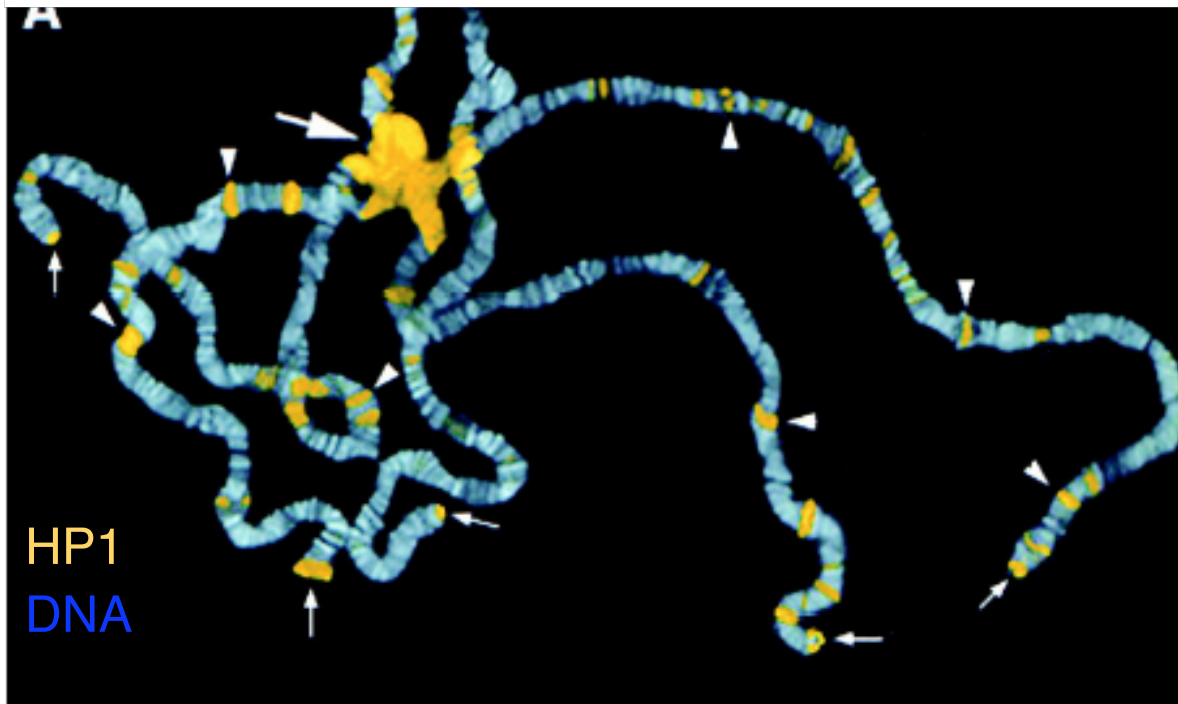
myc-SUV39H1: myc-epitope tagged Suv39H1

Melcher et al, MCB, 2000

## What are the target sites for Suv39h1 and H3K9me???

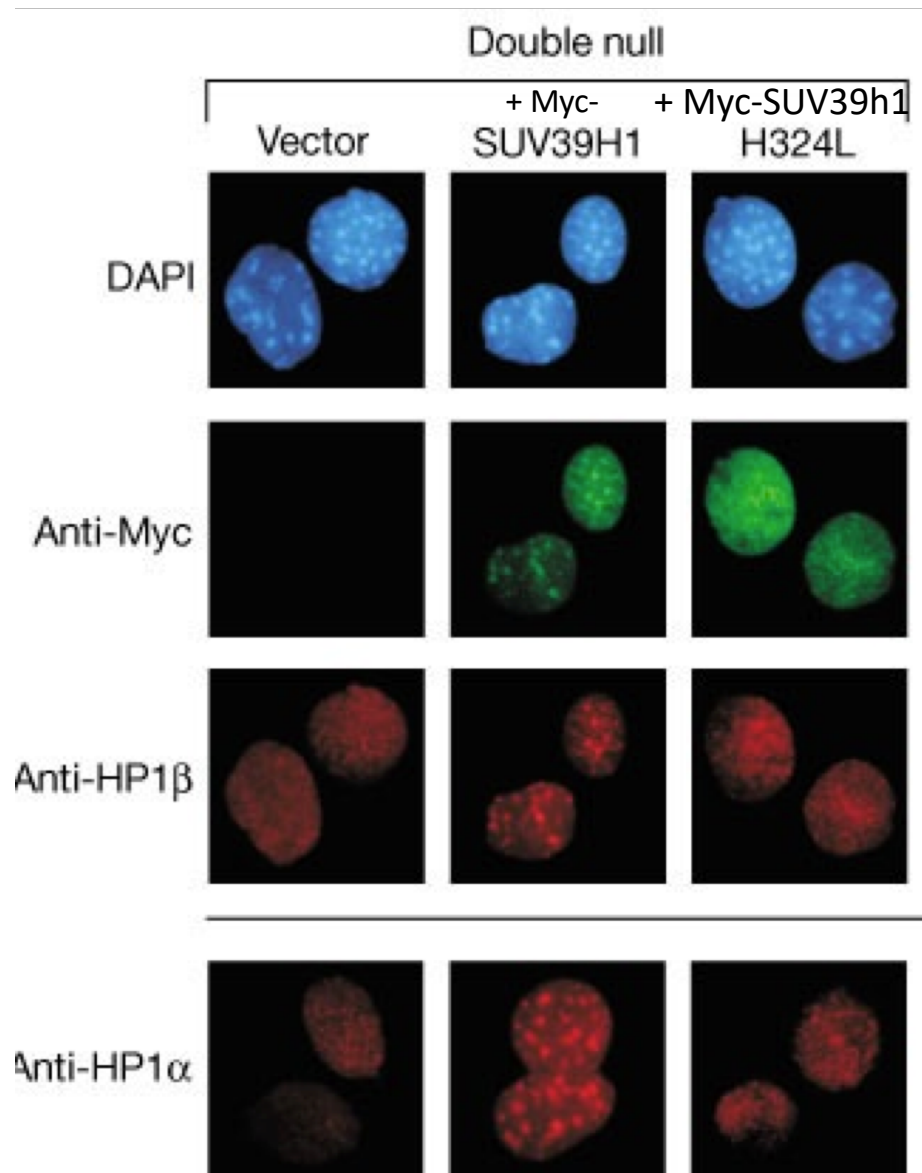


Another evidence for HP1 and Suv39h interaction came from *Drosophila*



- back to early genetics studies in *Drosophila*:
- *Su(var) 2-5* (gene) codes for heterochromatin protein 1
- *Su(var) 2-5* shows similar phenotype like *Su(var)3-9*
- HP1 in *Drosophila* is localized to the chromocenter

## SUV39h HMTase activity is important to build constitutive heterochromatin



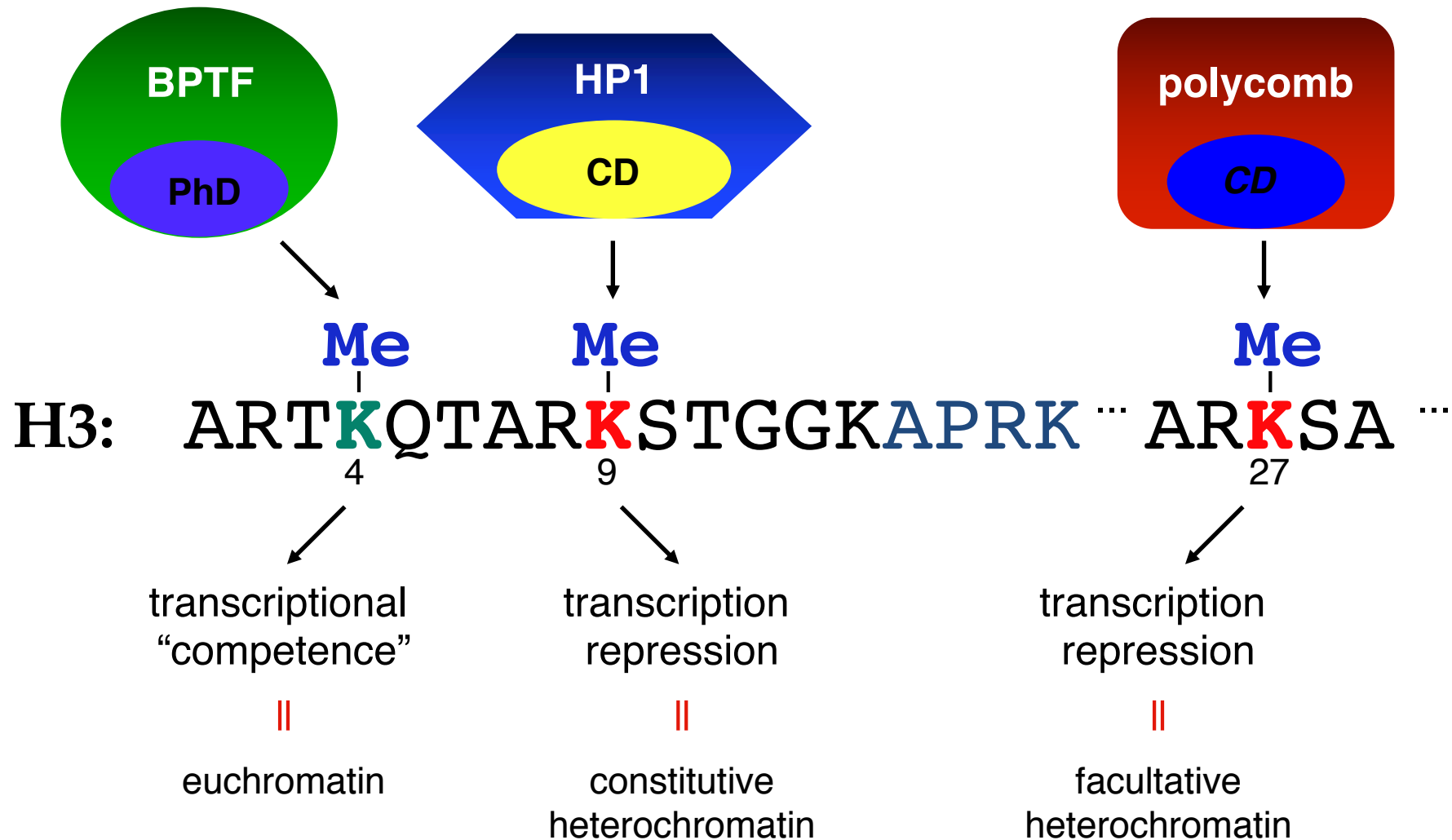
Loss of Suv39h1 expression results in delocalization of HP1 from chromocenters

Re-expression of Suv39h1 rescues Localization of HP1 to chromocenters

Re-expression of Suv39h1 that Contains an enzymatic dead mutation (H324L) in SET domain does not rescue the localization of HP1 to chromocenters

→ The enzymatic activity of Suv39h1 Is required for recruitment of HP1 to chromocenters

## Site specific methylation of the H3 tail has different functions

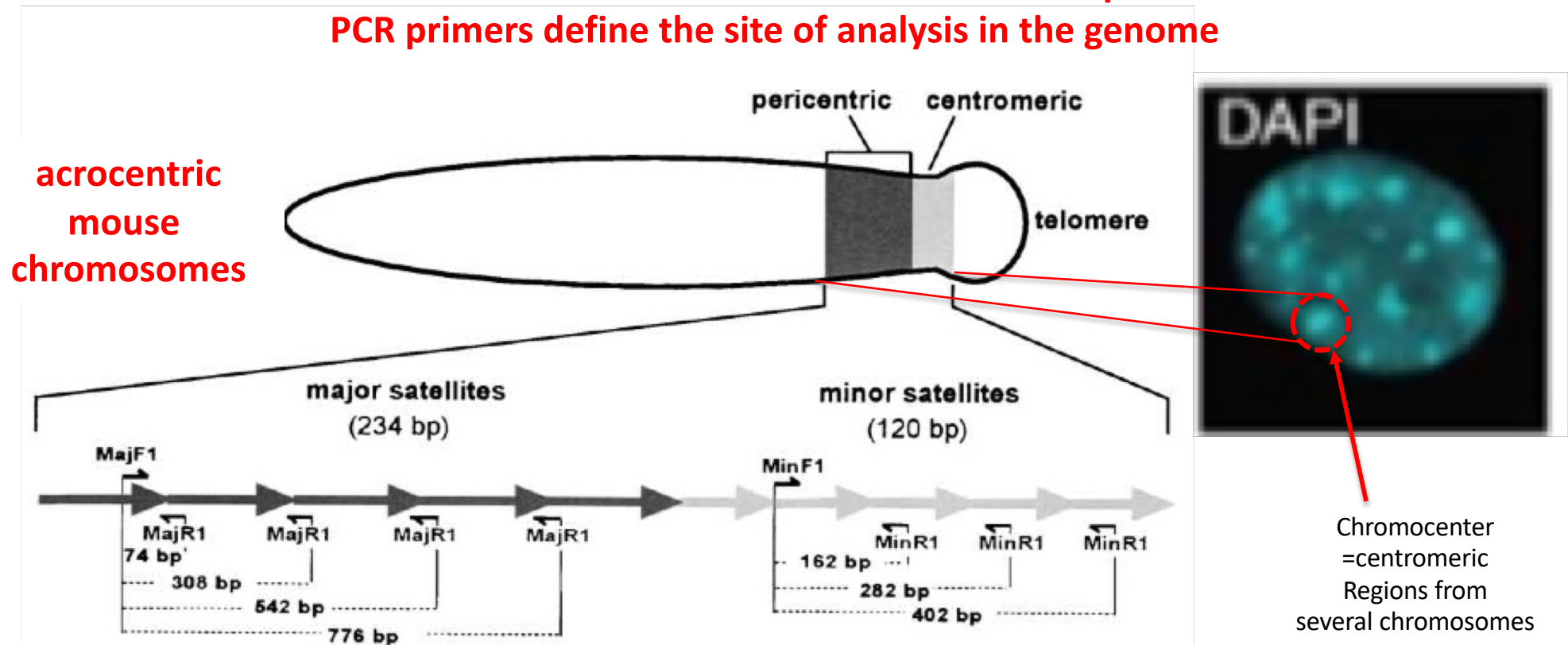


## What are the target sites for Suv39h1 and H3K9me???

CHROMATIN IMMUNOPRECIPITATION → DETAILED ANALYSIS

→ Localization of protein at a defined region - sequence

The combination of Immunoprecipitation methods and PCR analysis allows to define the histone code at defined sequences.  
PCR primers define the site of analysis in the genome

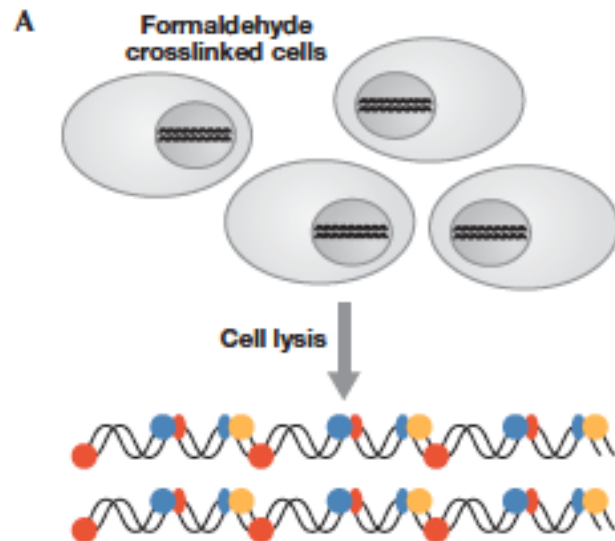


EXAMPLE: Pericentric heterochromatin in mouse cells



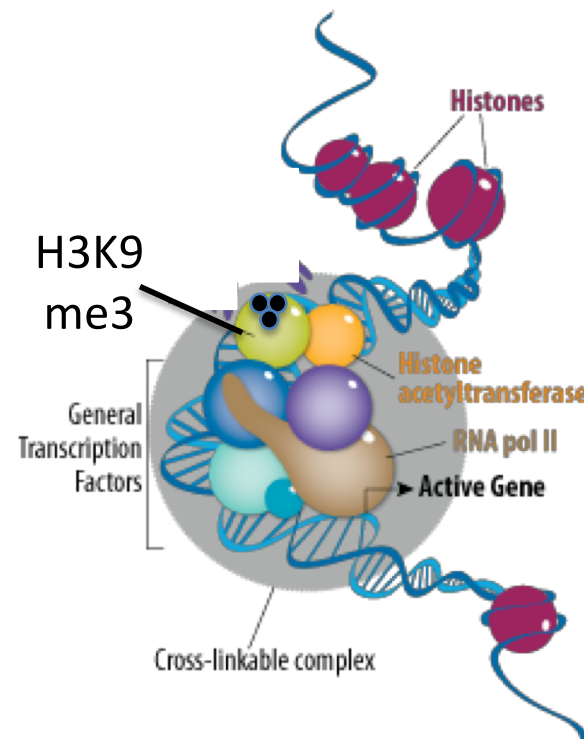
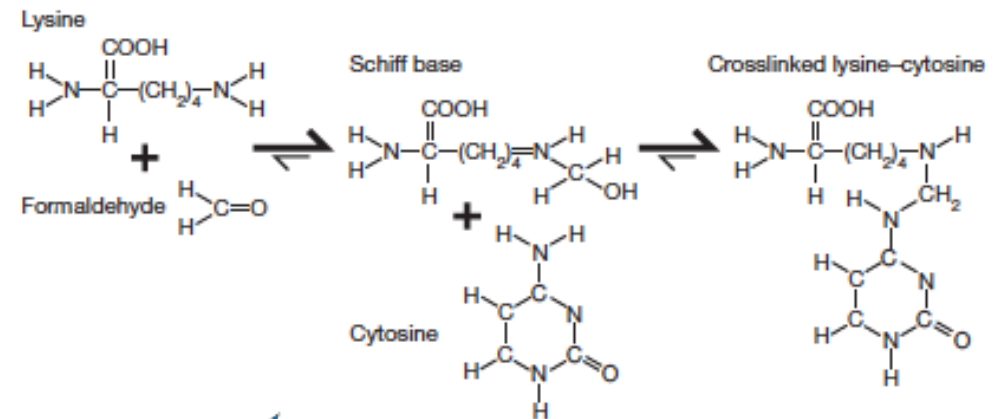
## CHROMATIN IMMUNOPRECIPITATION

### 1. Cross linking with FA



*Optimization is crucial.*

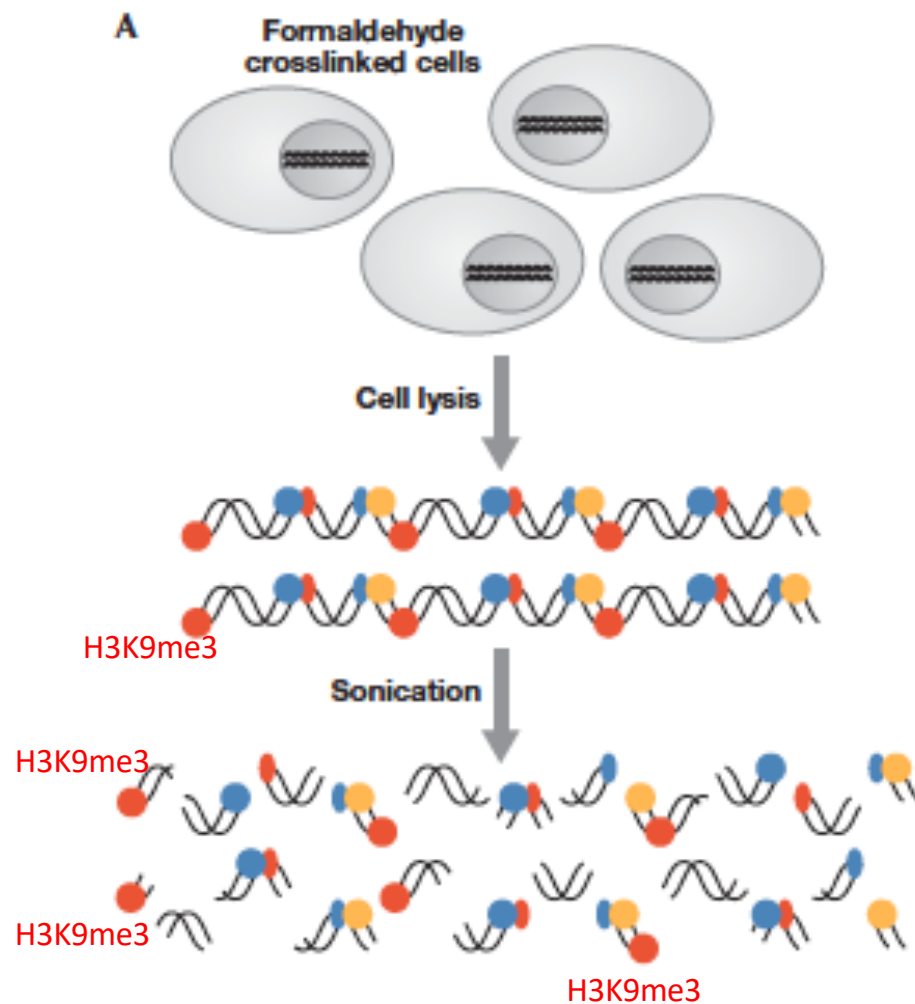
**B** Formaldehyde will crosslink amino or imino groups within 2Å, for example:



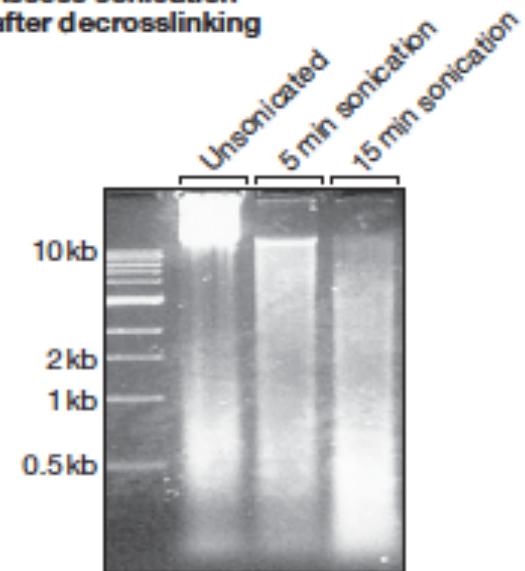
*Covalently crosslinked chromatin*

# CHROMATIN IMMUNOPRECIPITATION

## 2. Cross linking followed by sonication (fragmentation of chromatin)



**C** Assess sonication after decrosslinking

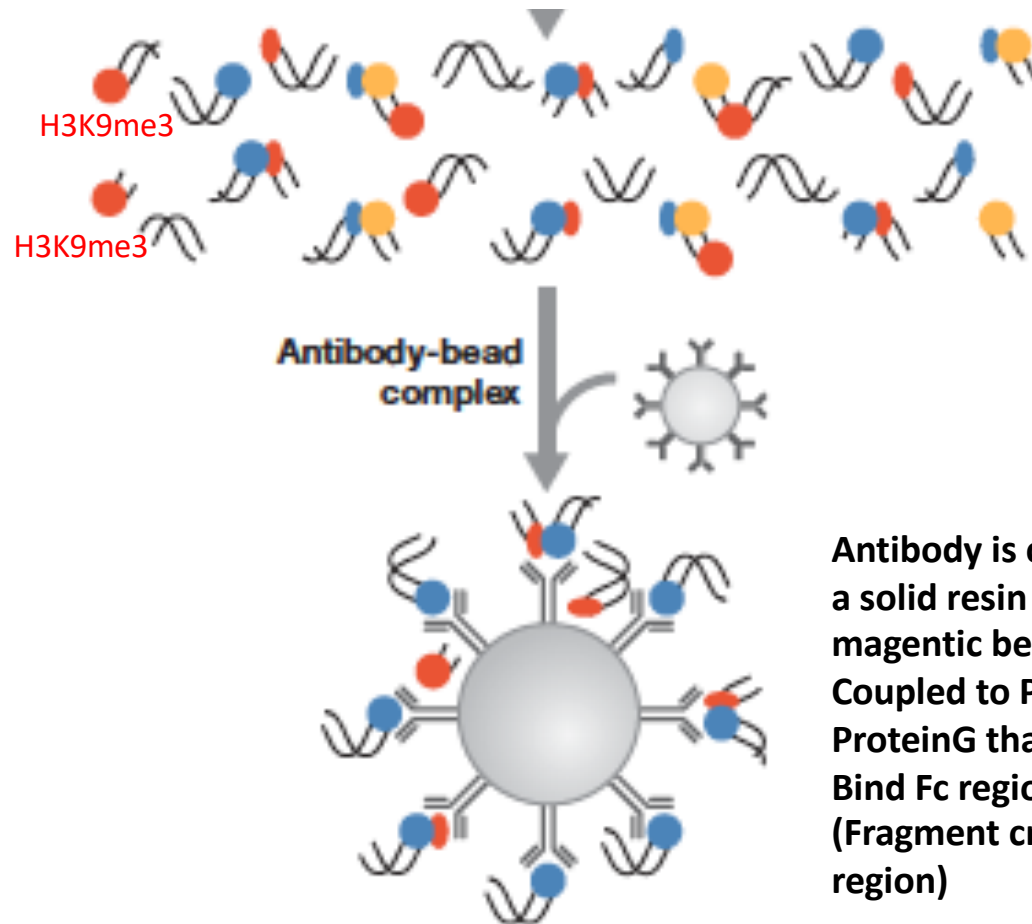


Chromatin is bulky → needs to be cut into small pieces  
to become soluble

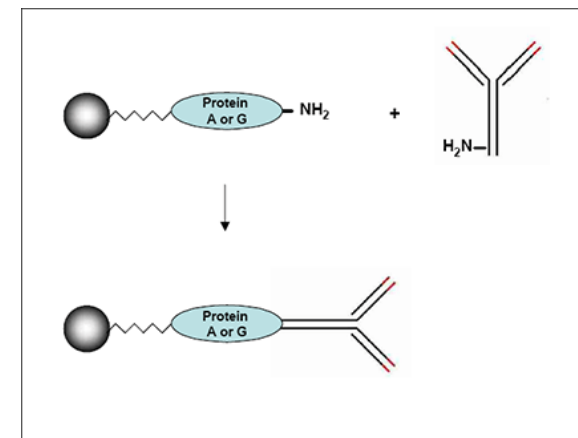
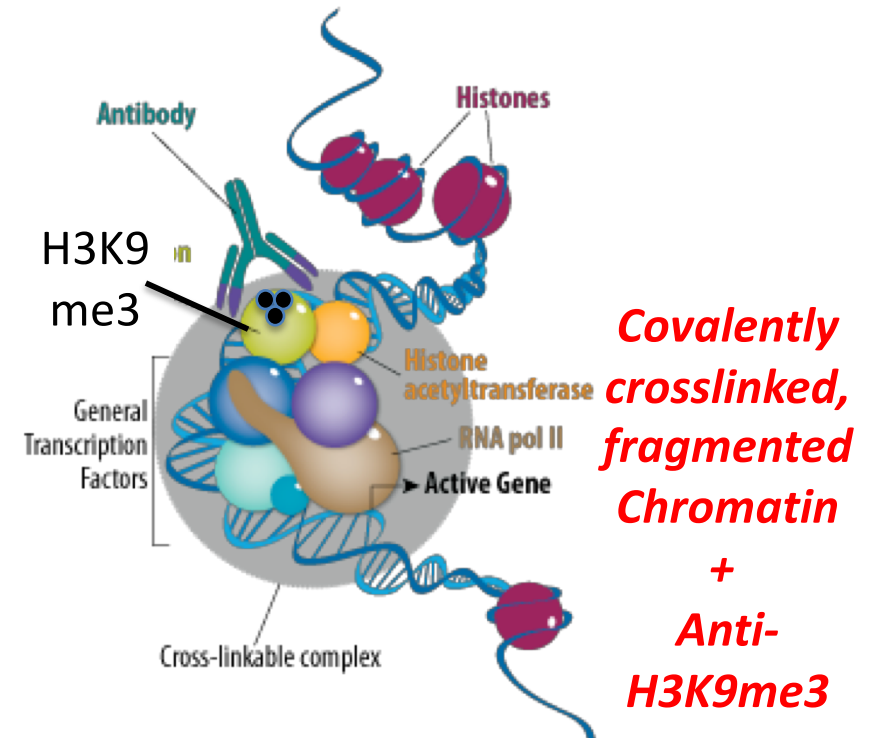
## CHROMATIN IMMUNOPRECIPITATION

### 3. Immunoprecipitation (IP)

The protein of interest is immunoprecipitated together with the crosslinked DNA: Modified histones; epigenetic writers, epigenetic readers



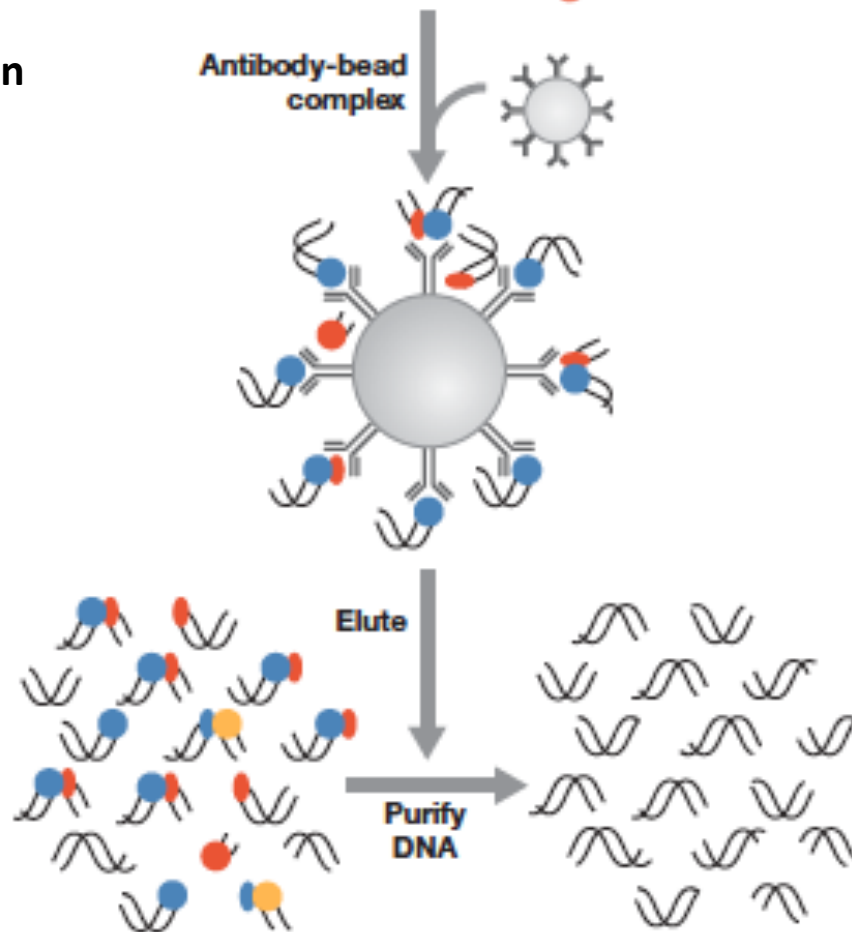
Antibody is coupled to a solid resin (agarose, magnetic beads)  
Coupled to ProteinA or ProteinG that Bind Fc region of antibody (Fragment crystallizable region)



## CHROMATIN IMMUNOPRECIPITATION

### 4. Decrosslinking of PFA crosslinked chromatin and purification of the DNA

- Recover beads by centrifugation
- Wash beads
- Reverse the FA: 65C, high salt concentrations: crosslink break
- RNase and Protease treatment
- Purification of DNA

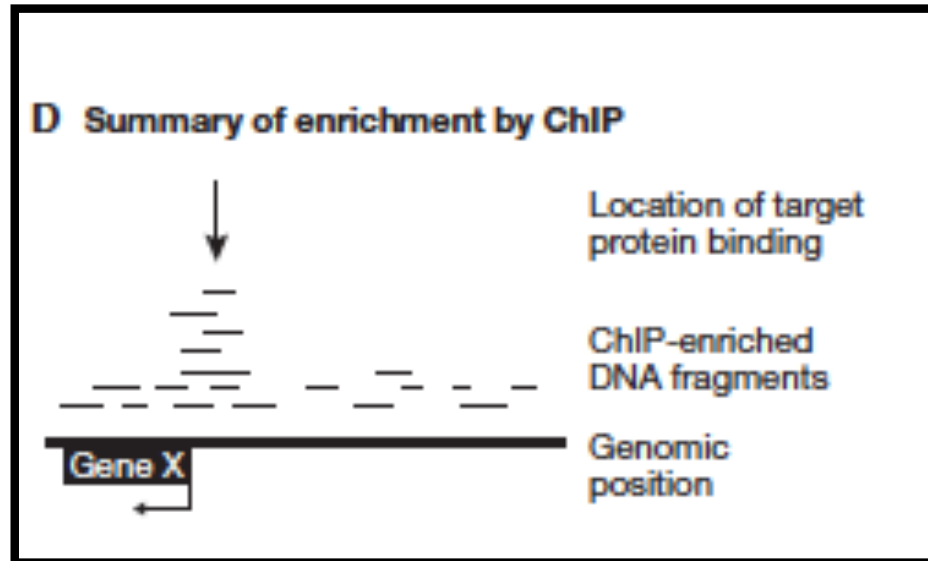


**ANALYZE DNA TO IDENTIFY REGIONS WITH HIGH H3K9me3**  
(sequencing, hybridization, PCR with gene specific primers)

## CHROMATIN IMMUNOPRECIPITATION

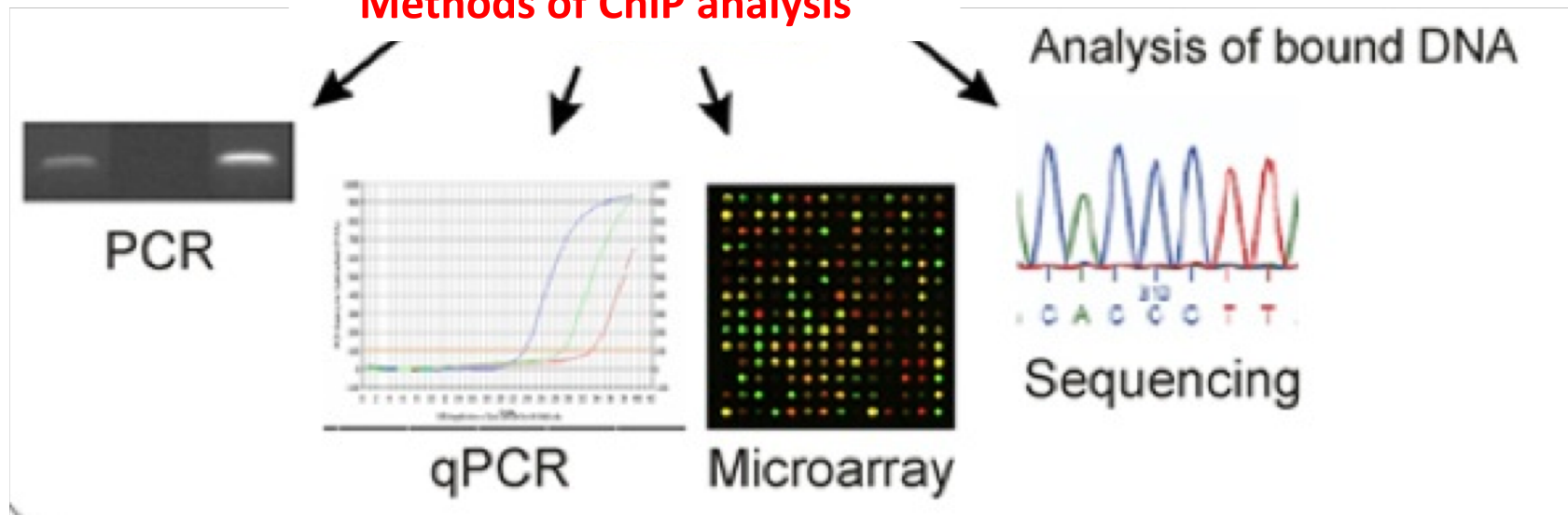
### 5. Analysis of ChIP DNA

Identification of DNA regions associated with the protein/modification of interest



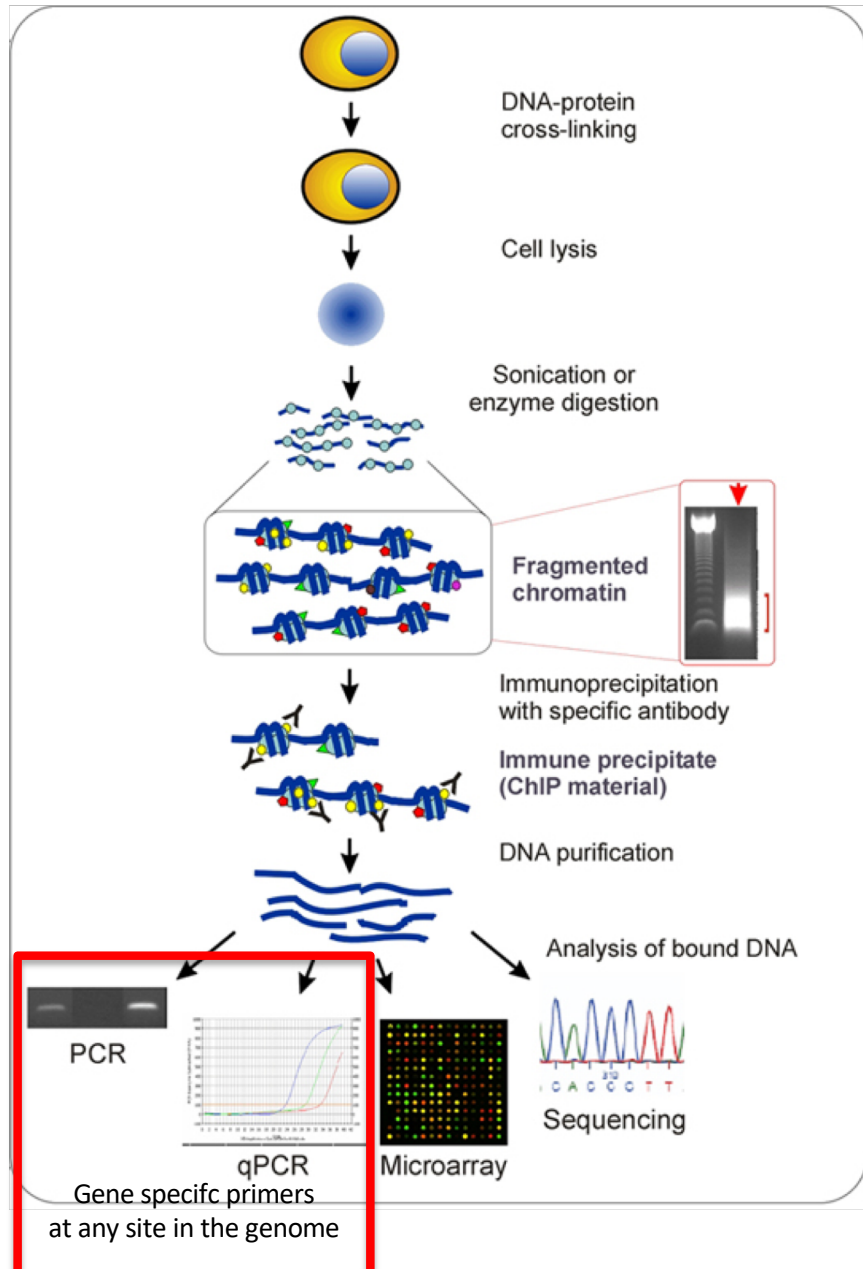
DNA from chromatin enriched in H3K9me  
→ Correspond to precise sites in the genome

### Methods of ChIP analysis



## CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILED ANALYSIS

→ Localization of protein at a defined region - sequence



Cell model system:

i.e. **Wild-type or Suv39 dn cells**

that grow in cell culture dish

1. Crosslink chromatin  
(treatment of cells with Paraformaldehyde)
2. Sonicate crosslinked cells
3. Incubate chromatin fragments with antibodies raised against **H3K9me3**
4. Recover antibody-chromatin complexes using a resin that binds to the constant region of antibodies
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
7. Purify DNA and precipitate DNA
8. Measure the amount of immunoprecipitated DNA  
In control versus Suv39h dn cells

### METHODS:

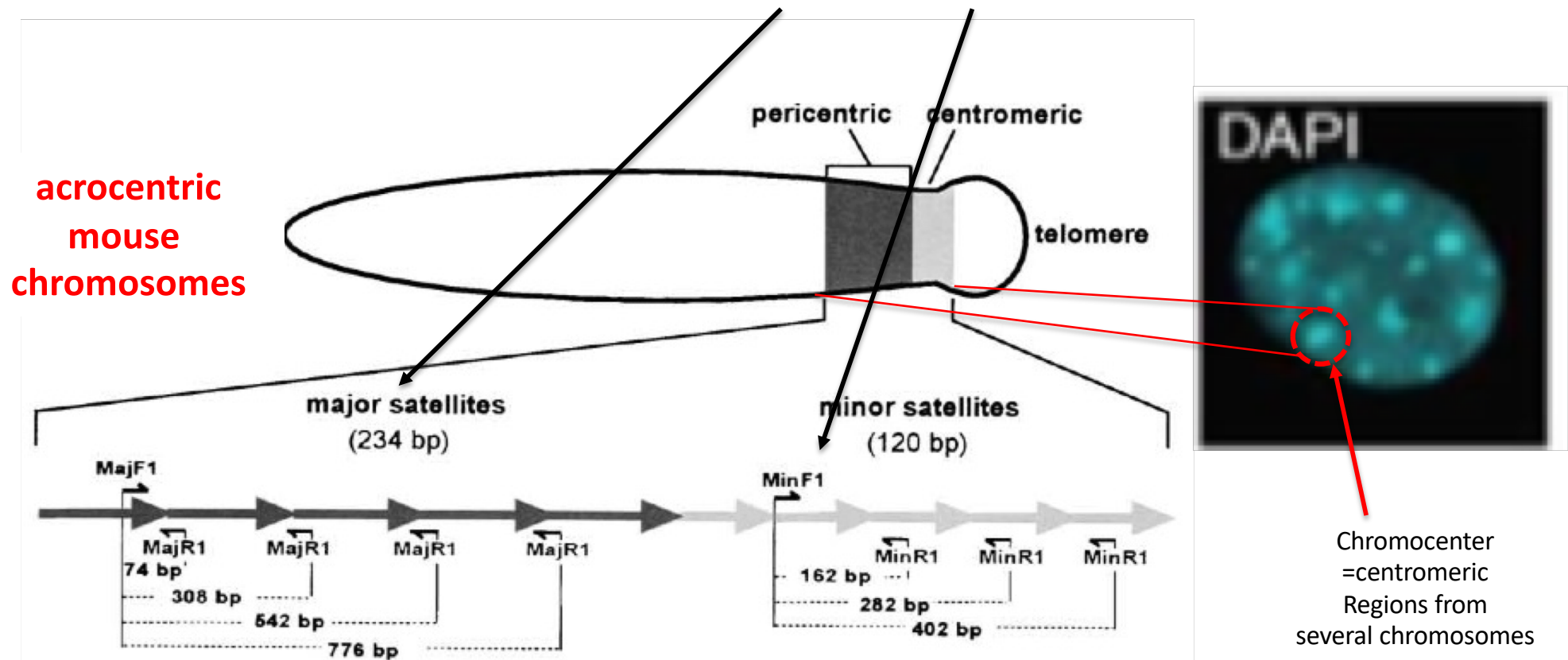
**Quantitative PCR: design PCR oligos to amplify defined sequences in immunoprecipitates. Per PCR reaction only one Locus can be examined by real-time PCR**



## What are the target sites for Suv39h1 and H3K9me???

### CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY QUANTITATIVE PCR

Design PCR oligos that amplify major and minor satellite repeats



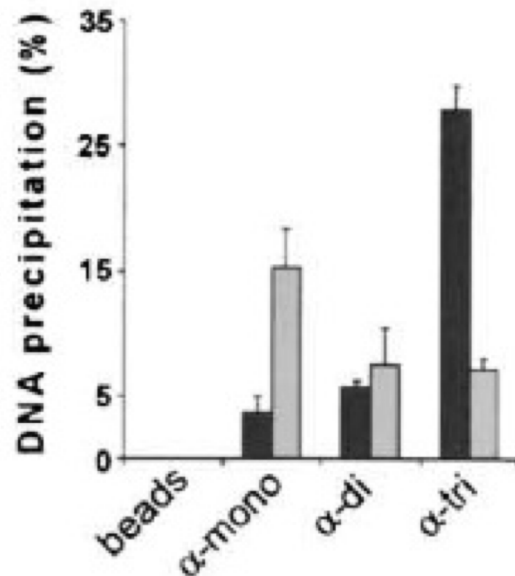
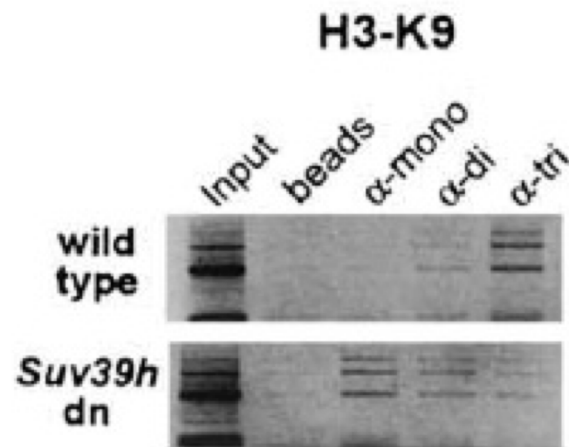
**EXAMPLE: Pericentric heterochromatin in mouse cells**

Min/Maj F1...: forward primer in unique region

## CHROMATIN IMMUNOPRECIPITATION (ChIP) COUPLED WITH PCR

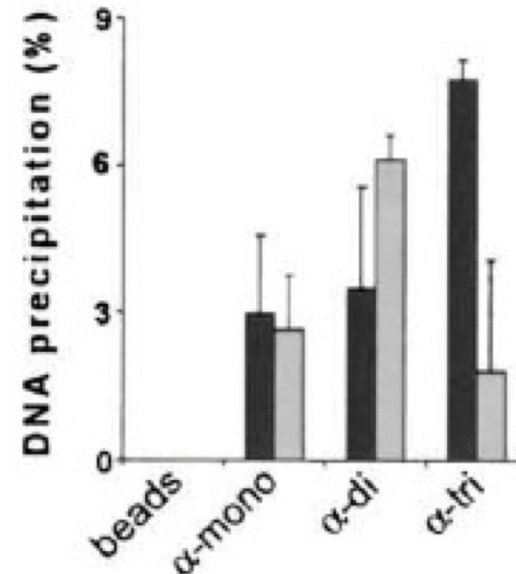
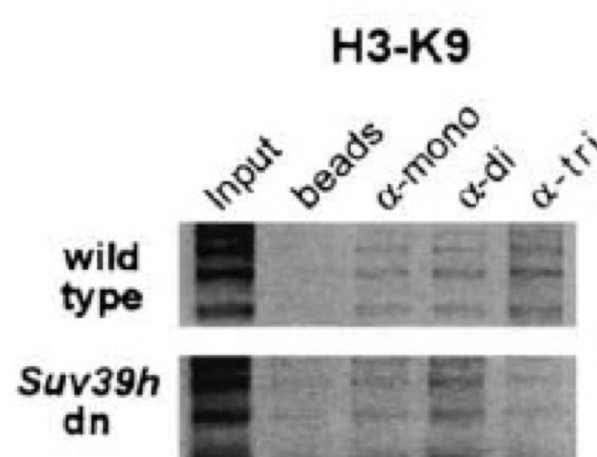
→ H3K9me3 is enriched  
at pericentric (major+minor) repeats in mouse cells

### Major satellite repeats



■ wild type ■ *Suv39h* dn

### Minor satellite repeats

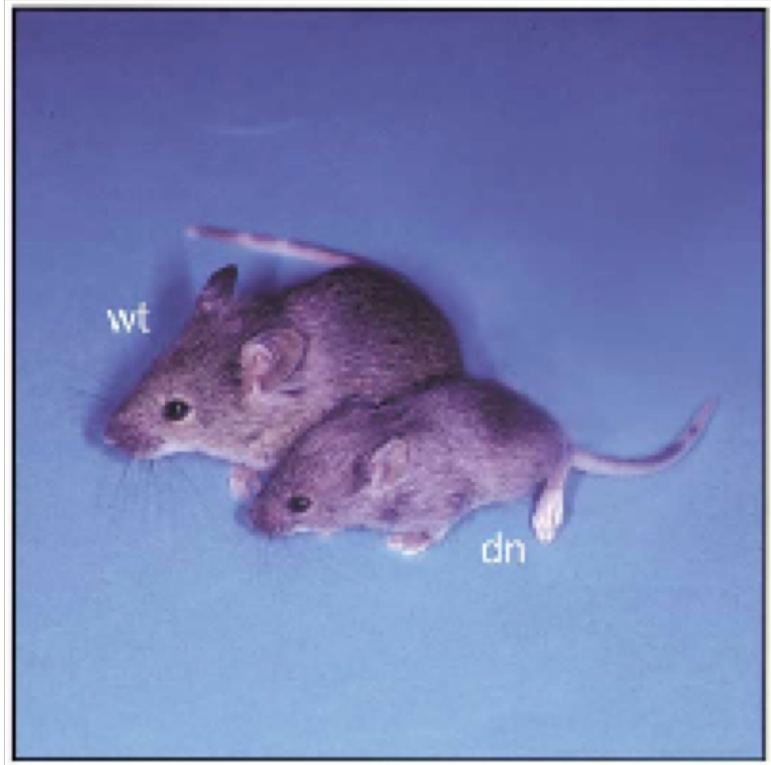


■ wild type ■ *Suv39h* dn

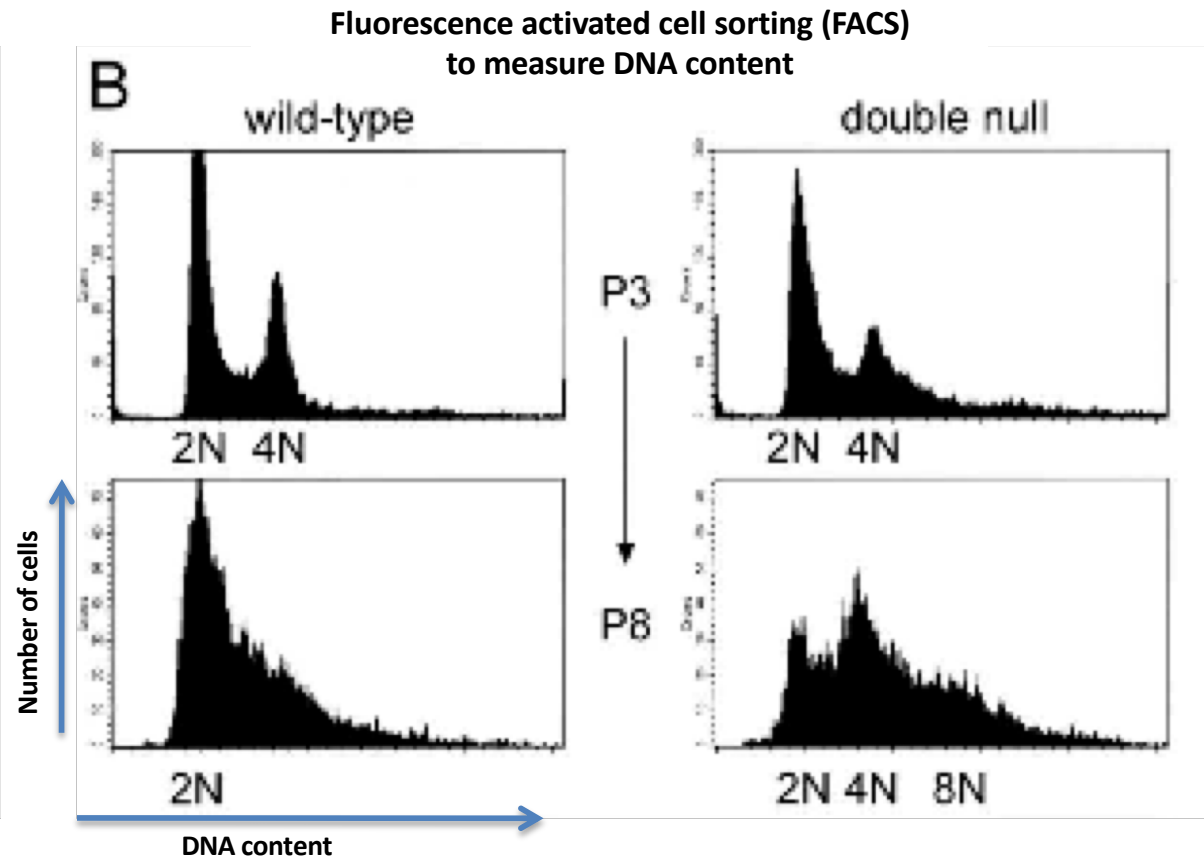
PCR amplification  
of major/minor satellite  
Repeats after ChIP using  
Antibodies that are specific for  
H3K9me1; H3K9me3; H3K9me3

*Suv39h1* is required for imposition  
of H3K9me3 at pericentric repeats

## Lack of SUV39h HMTase activity results in genomic instability



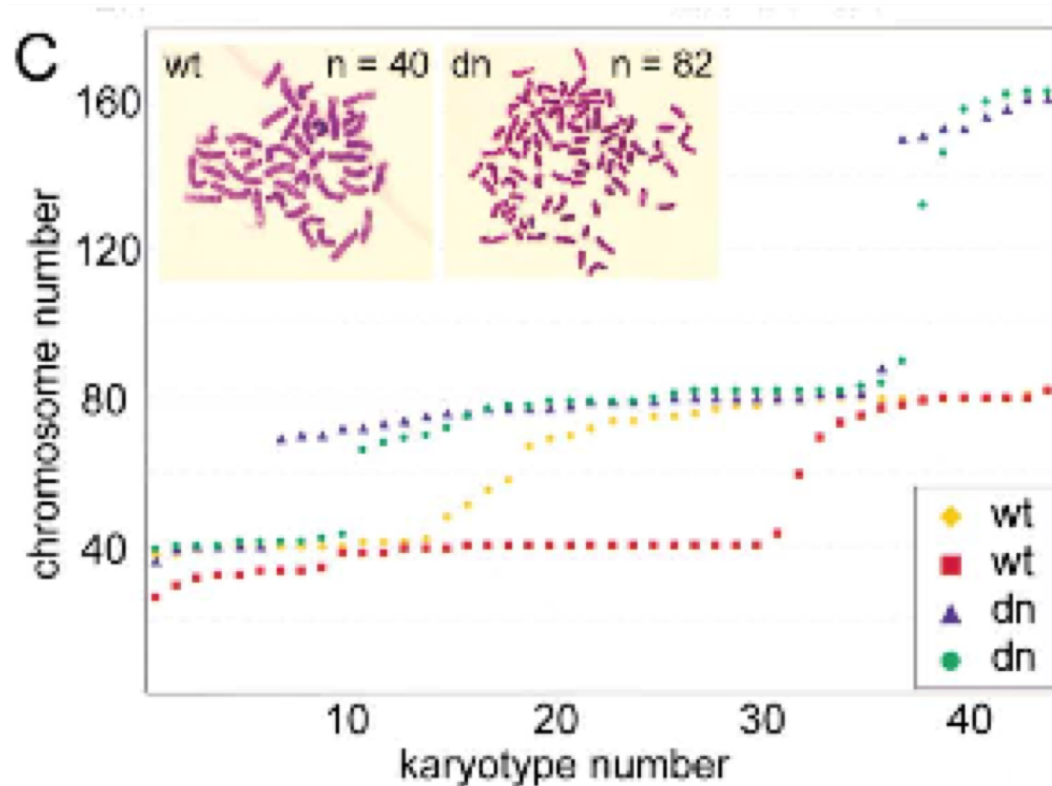
A knock-out model system  
for Suv39h1 and Suv39h2  
- Loss of Suv39h1/2: smaller body size



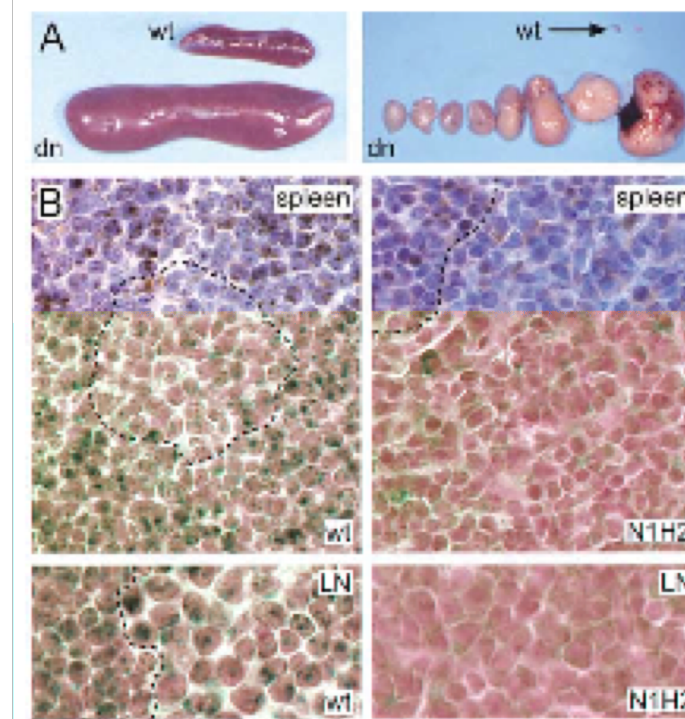
Fibroblasts from Suv39h1/2 null mice  
are aneuploidy

## Lack of SUV39h HMTase activity results in genomic instability

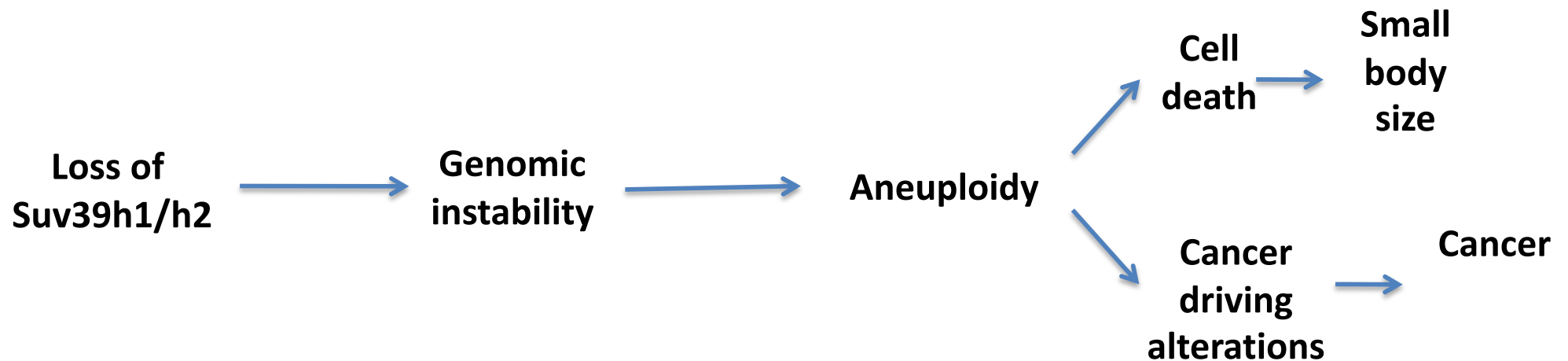
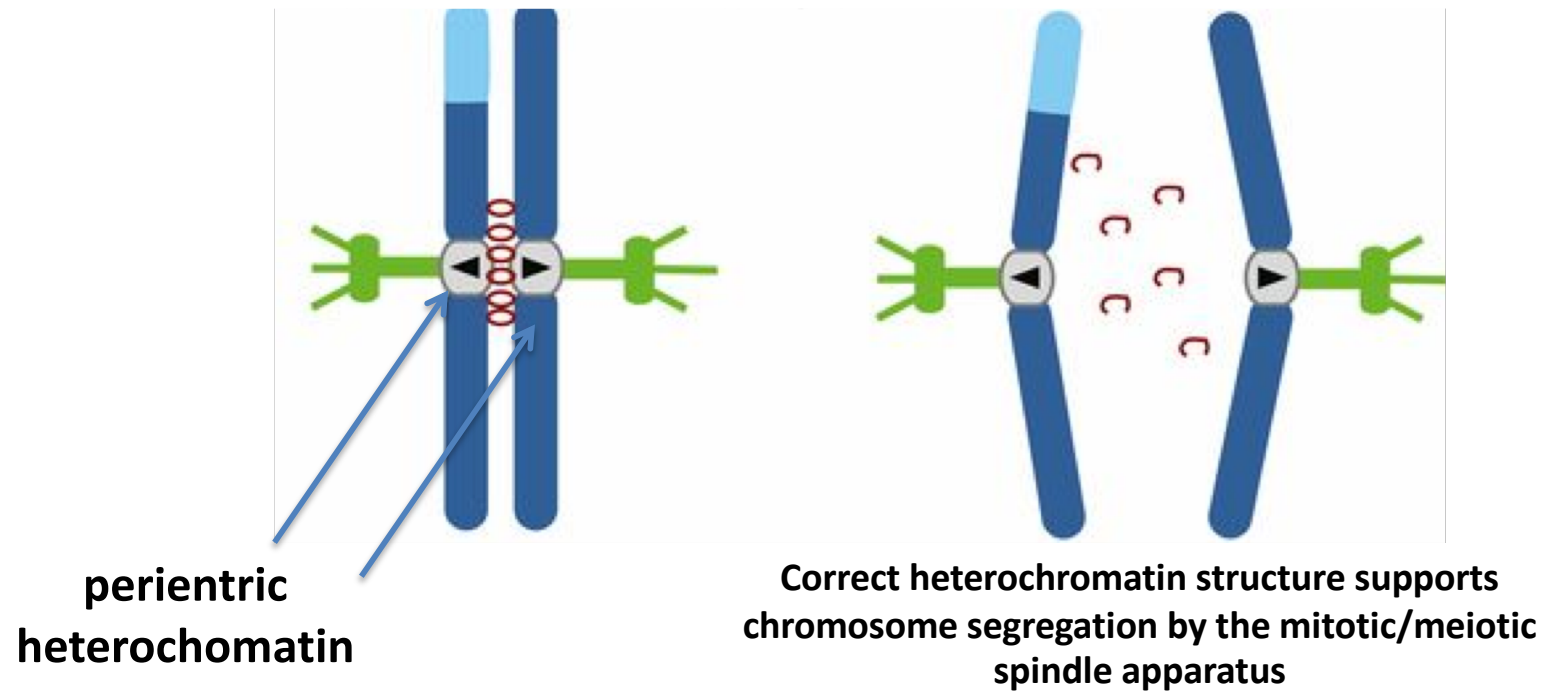
Loss of Suv39h1/2 results in increased chromosome numbers



Genomic instability in Suv39h1/2 mice increases lymphomas



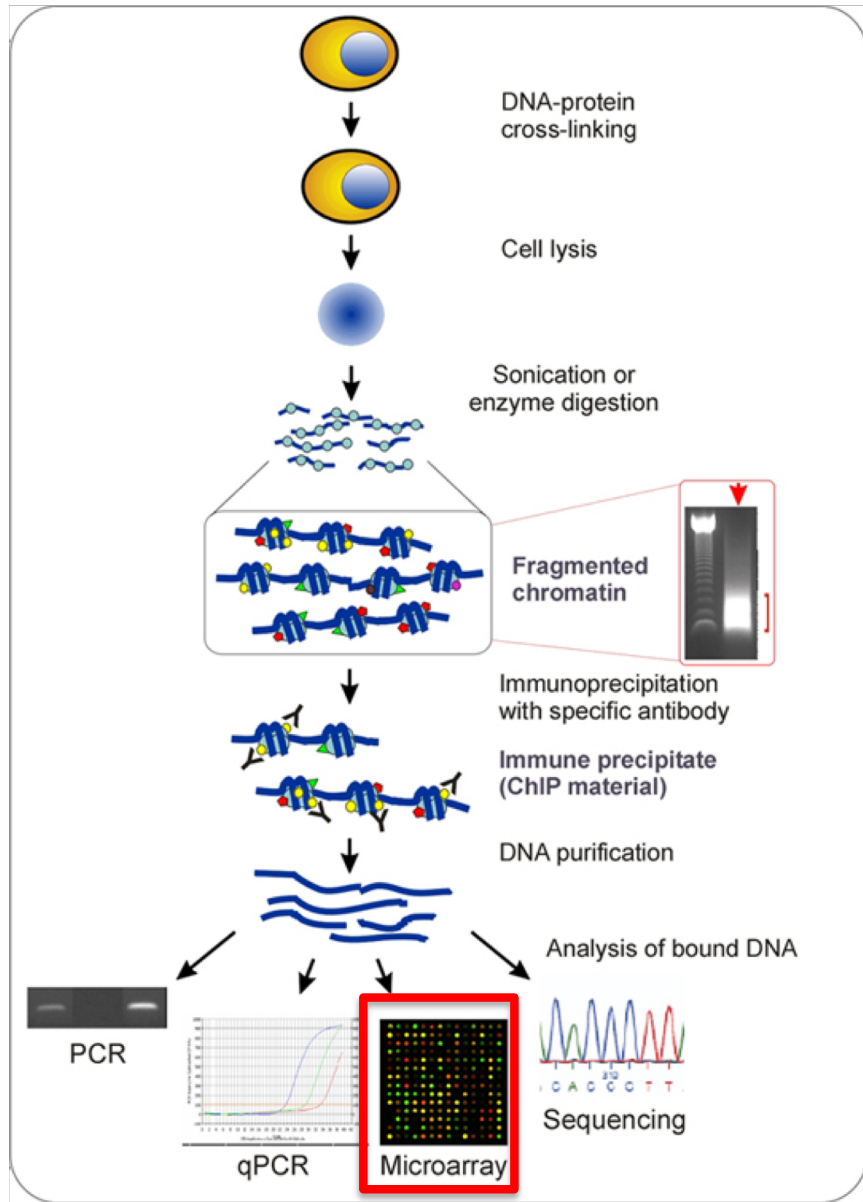
## SUV39h HMTase activity is essential for fidelity in mitosis/meiosis





## CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILED ANALYSIS

→ Localization of protein at a defined region - sequence



Cell model system:

i.e. **Wild-type or Suv39 dn cells**

that grow in cell culture dish

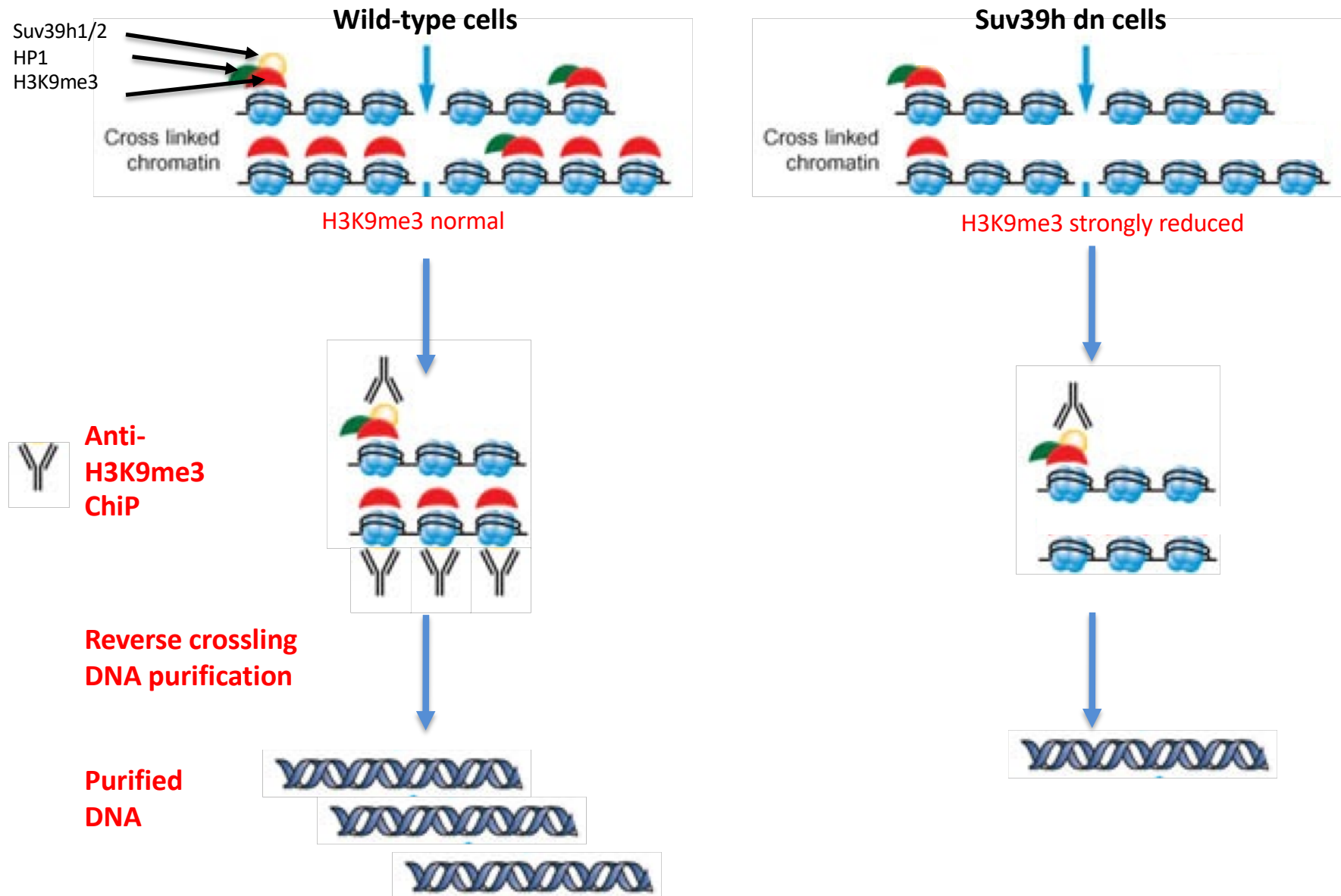
1. Crosslink chromatin  
(treatment of cells with Paraformaldehyde)
2. Sonicate crosslinked cells
3. Incubate chromatin fragments with antibodies raised against **H3K9me3**
4. Recover antibody-chromatin complexes using a resin that binds to the constant region of antibodies
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
7. Purify DNA and precipitate DNA
8. Measure the amount of immunoprecipitated DNA  
In control versus Suv39h dn cells

### METHODS:

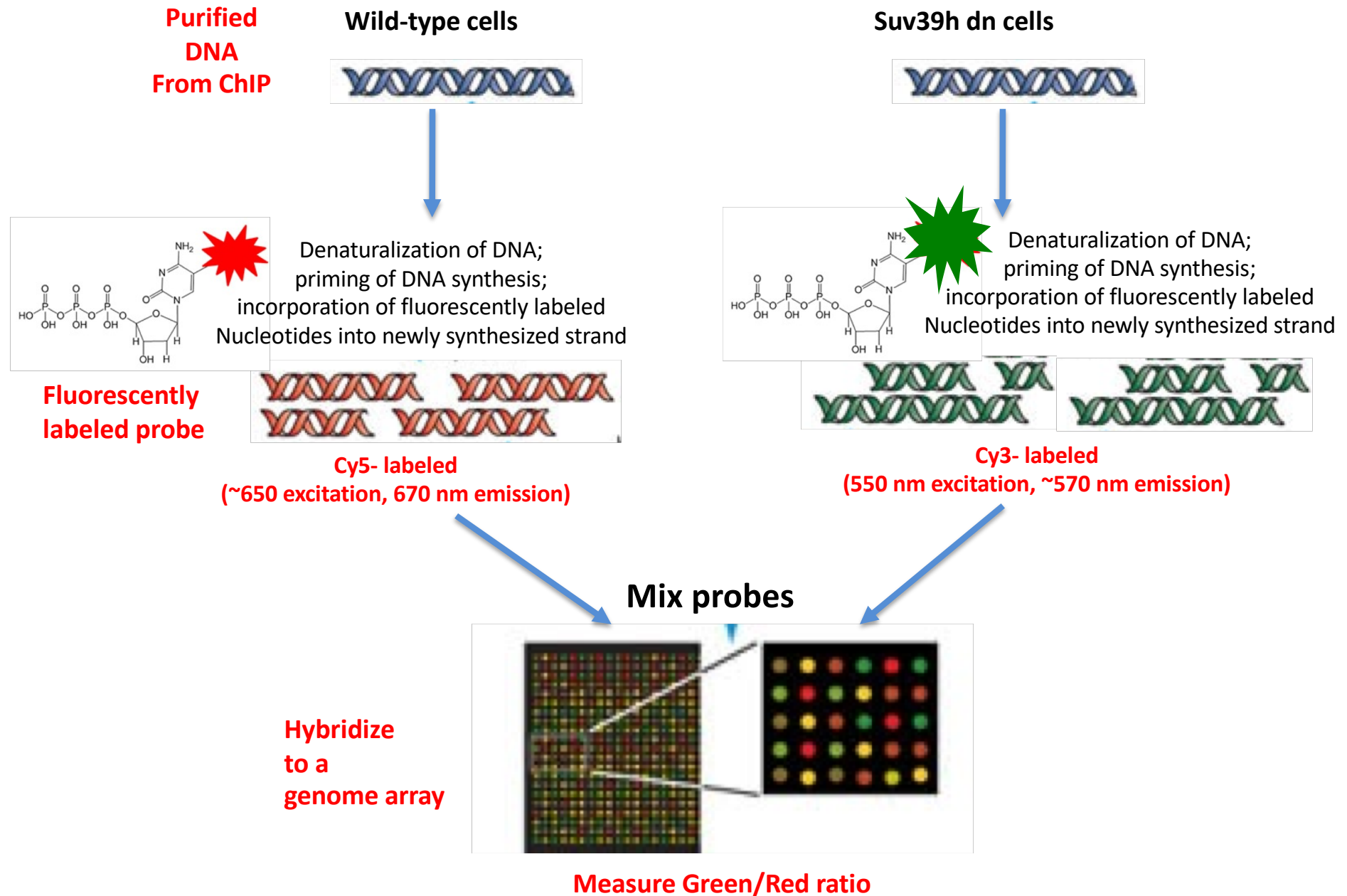
**Quantitative PCR: design PCR oligos to amplify defined sequences in immunoprecipitates. Per PCR reaction only one Locus can be examined by real-time PCR**



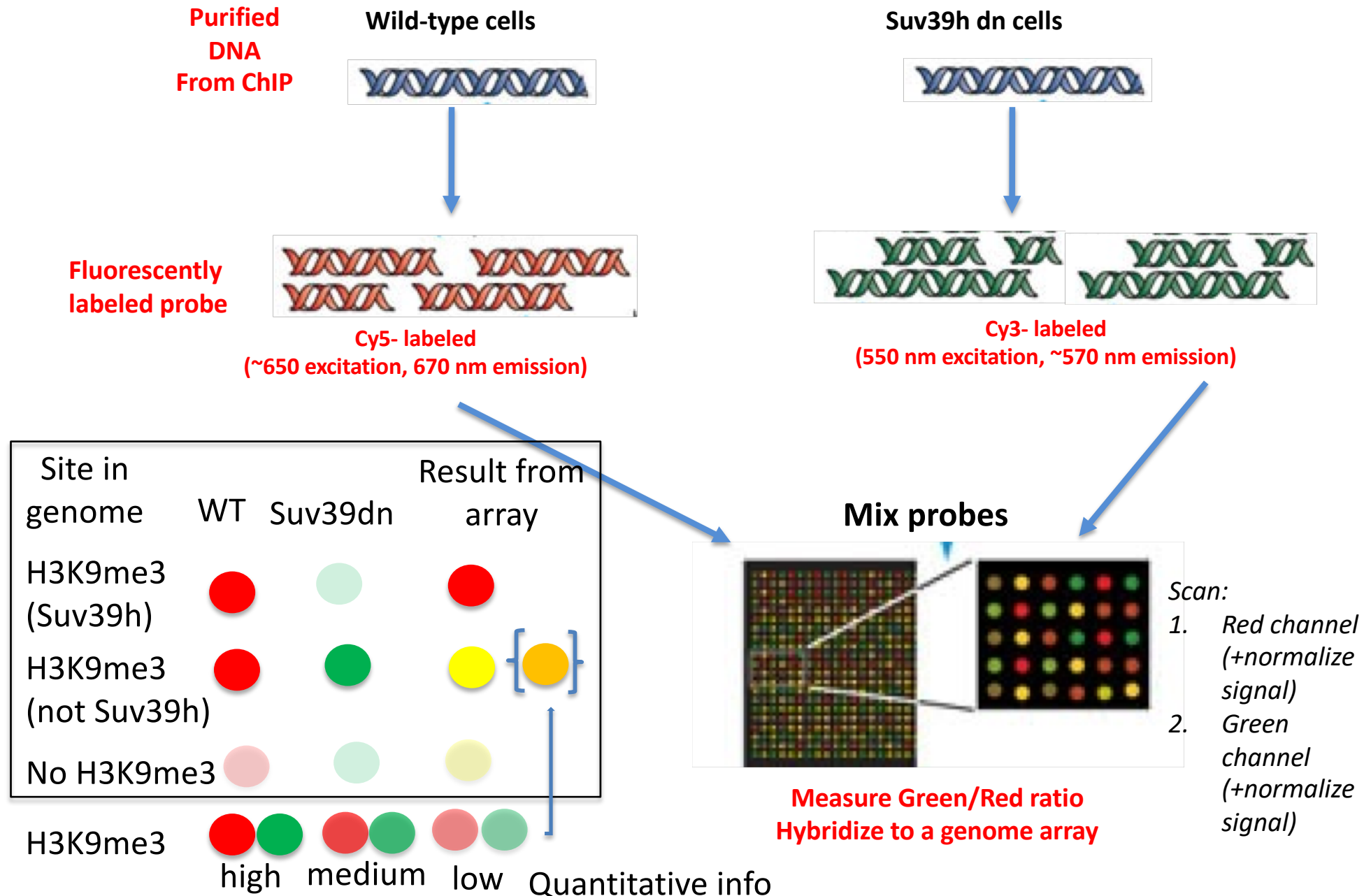
## ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



## ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



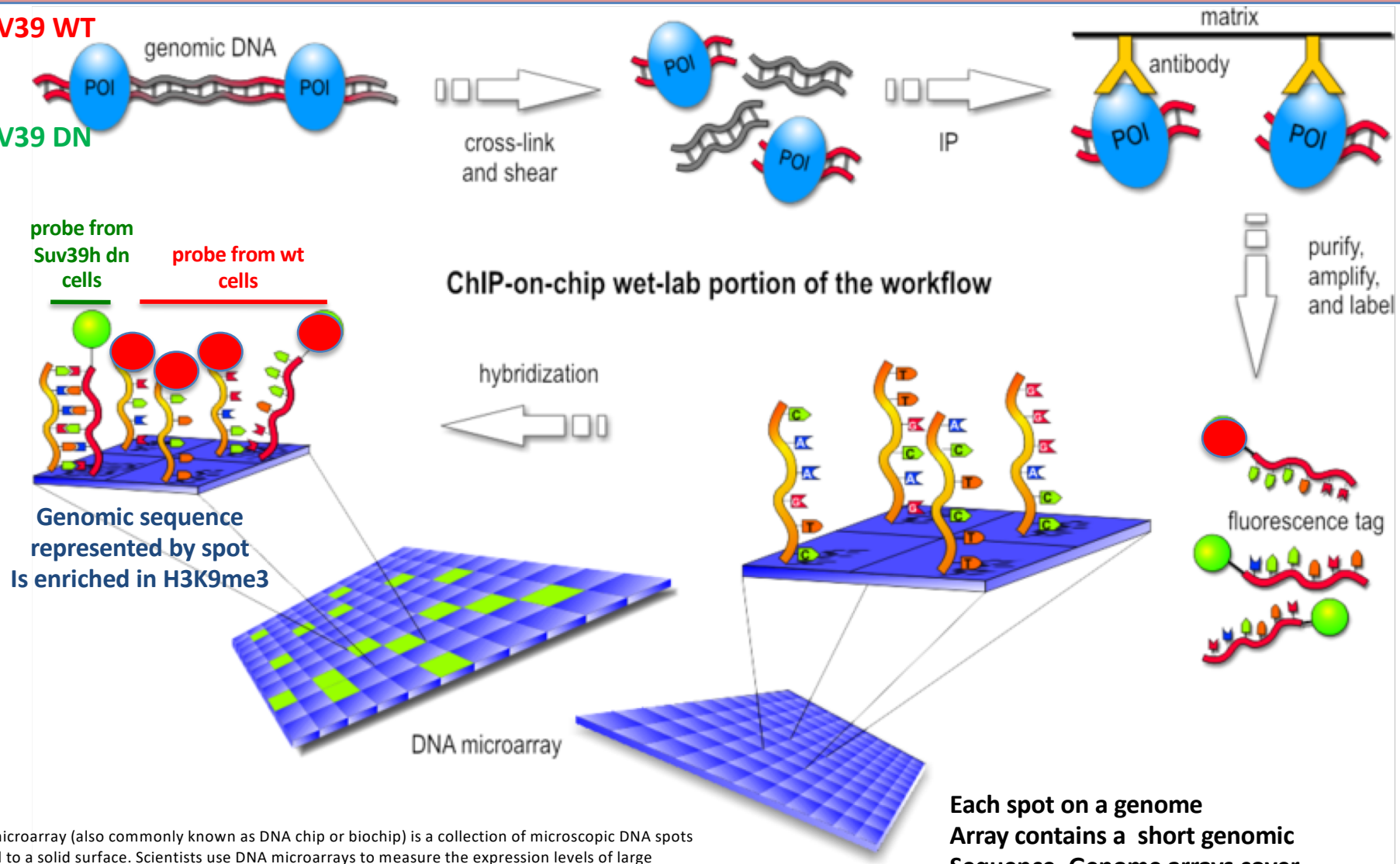
## ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



## ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome

### 1. SUV39 WT

### 2. SUV39 DN

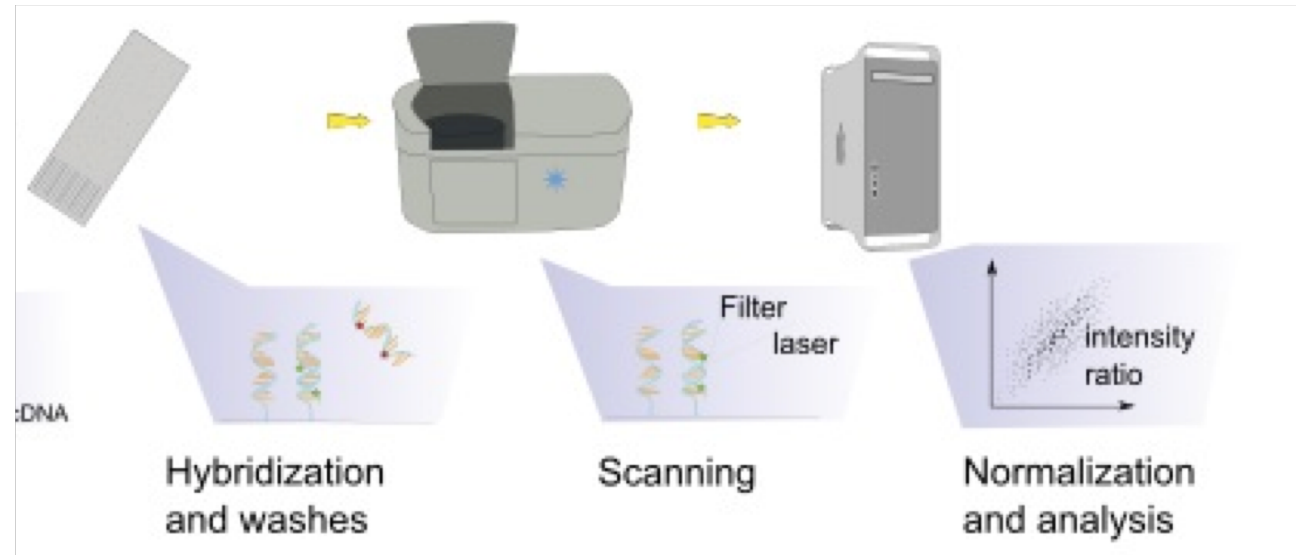
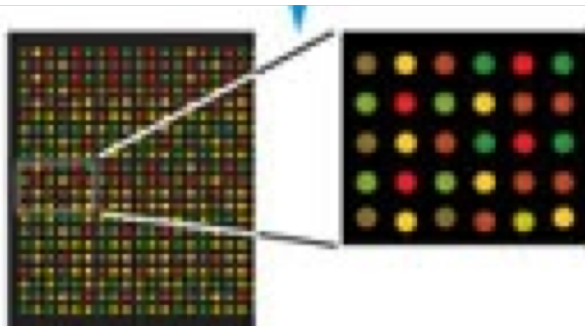


A DNA microarray (also commonly known as DNA chip or biochip) is a collection of microscopic DNA spots attached to a solid surface. Scientists use DNA microarrays to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome. Each DNA spot contains picomoles (10<sup>-12</sup> moles) of a specific DNA sequence, known as probes (or reporters or oligos). These can be a short section of a gene or other DNA element that are used to hybridize a cDNA or cRNA (also called anti-sense RNA) sample (called target) under high-stringency conditions. Probe-target hybridization is usually detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target

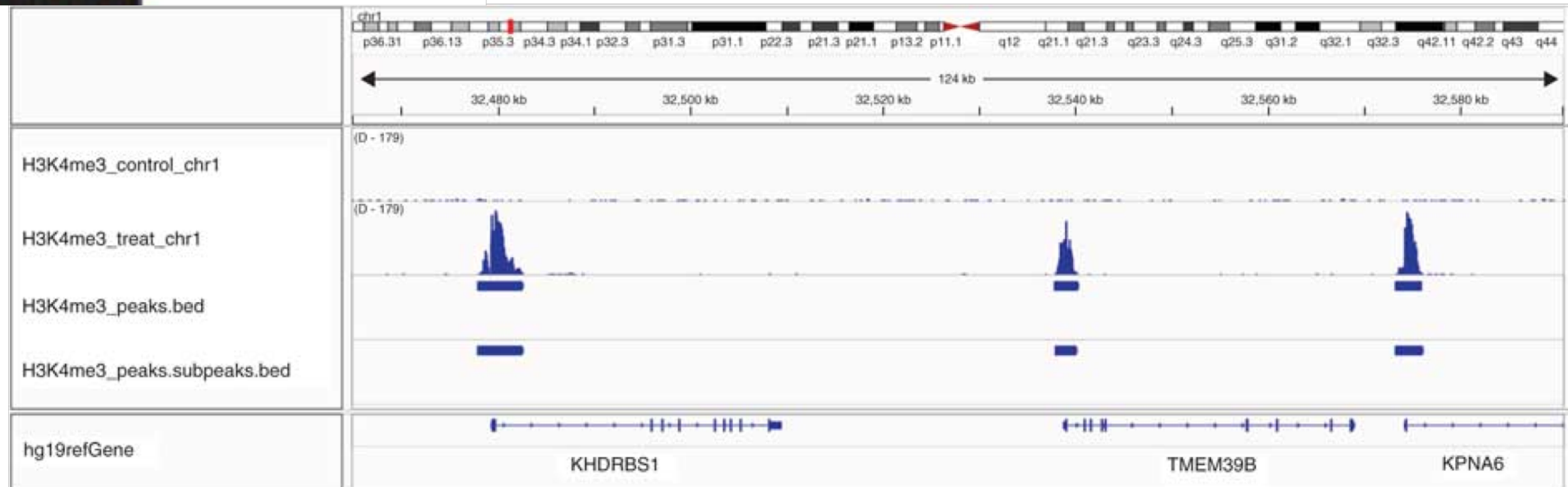
Each spot on a genome  
Array contains a short genomic  
Sequence. Genome arrays cover  
a large part of the  
Genome  
A ChIP holds ca 20000 spots



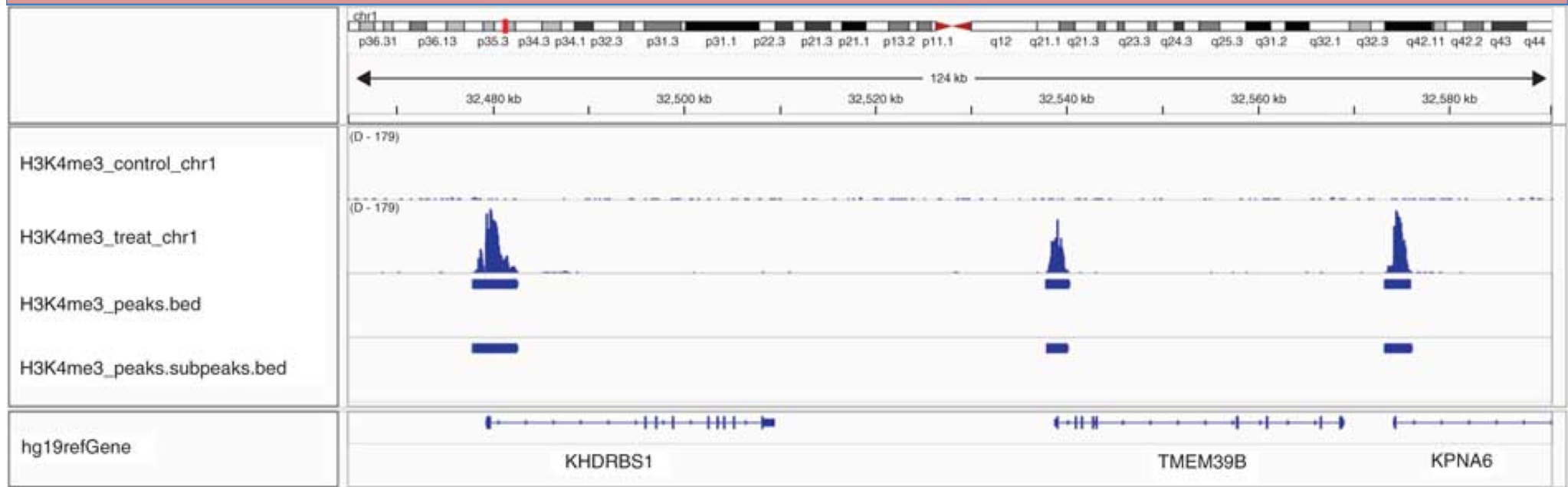
## ChIP on Chip: Analysis of epigenetic information across a high number of genomic sites in the genome



Arrays do not contain the entire human/mouse genome  
Are enriched for particular sequences according to need  
(promoter, enhancer, etc...)



## ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



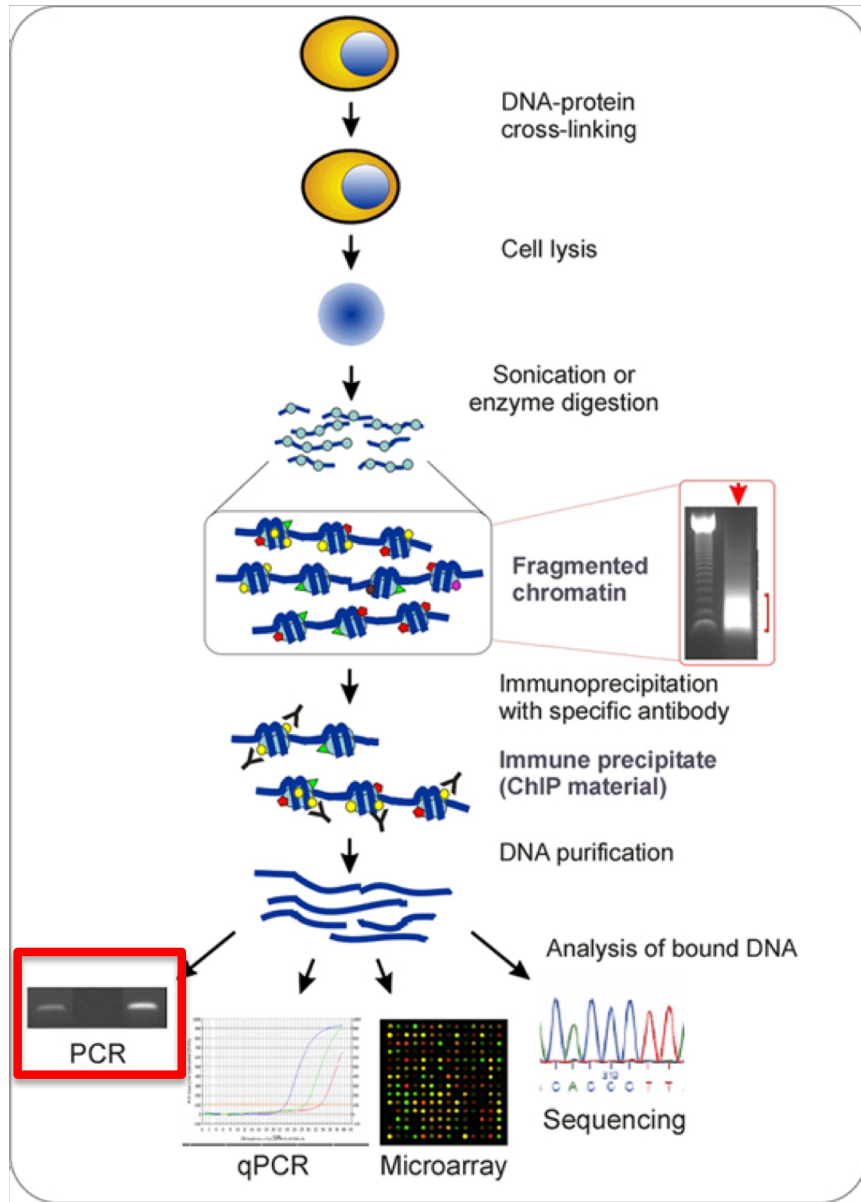
**Advantage:** low tech, cheap

**Disadvantage:** low resolution, no data on number of molecules – just proportions;  
laborious to reach a good genome coverage

Already outdated → state of the art: **ChIP seq**



## ChIP seq: Analysis of epigenetic information on the single nucleotide level → GENERATION OF GENOME WIDE EPIGENETIC MAPS



Cell model system:

i.e. **Wild-type or Suv39 dn cells**

that grow in cell culture dish

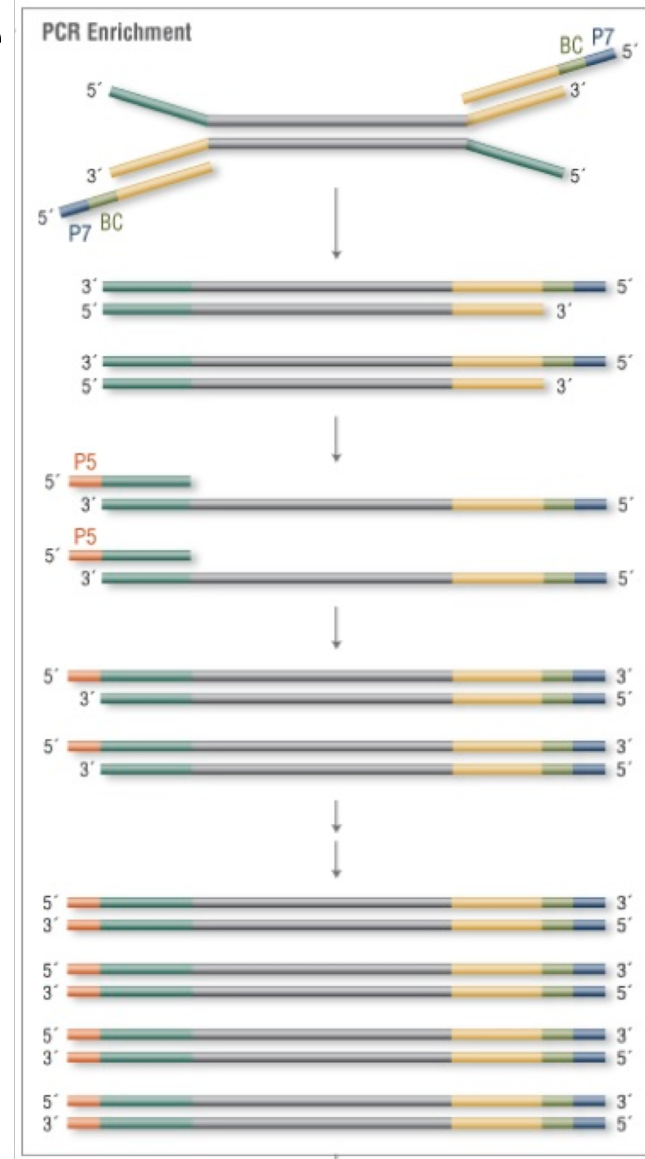
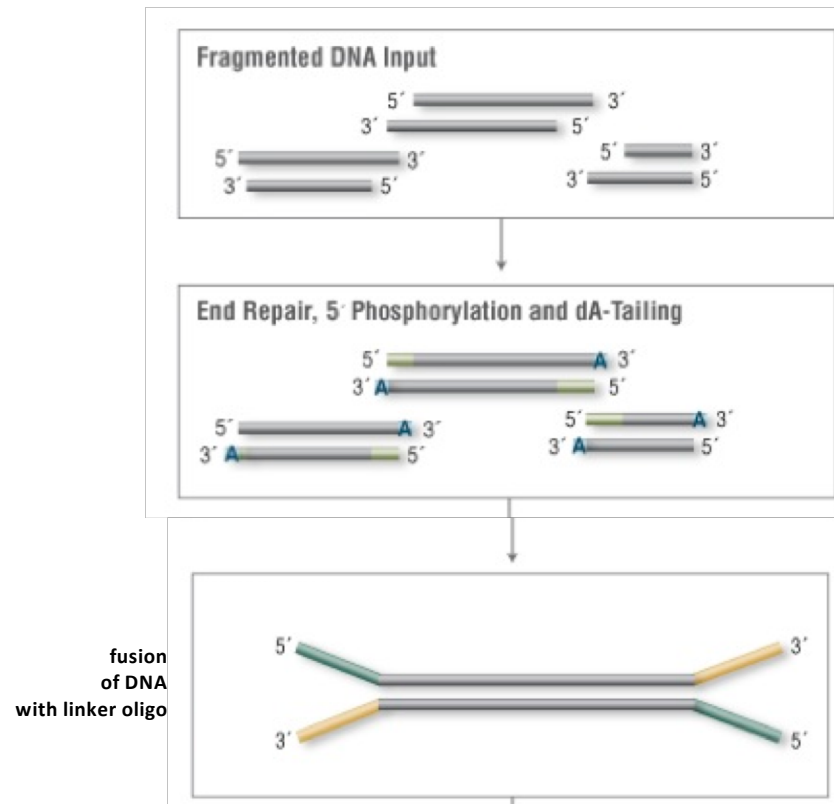
1. Crosslink chromatin  
(treatment of cells with Paraformaldehyde)
2. Sonicate crosslinked cells
3. Incubate chromatin fragments with antibodies raised against **H3K9me3**
4. Recover antibody-chromatin complexes using a resin that binds to the constant region of antibodies
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
5. Elute chromatin at high salt concentration and revert crosslinks at high temperature
6. Digest protein with protease K and RNA with RNase
7. Purify DNA and precipitate DNA
8. Measure the amount of immunoprecipitated DNA  
In control versus Suv39h dn cells

### METHODS:

**Massive parallel sequencing of immunoprecipitated DNA**  
**Permits to obtain epigenetic information on the single nucleotide level**

## ChIP seq: Analysis of epigenetic information on the single nucleotide level → GENERATION OF GENOME WIDE EPIGENETIC MAPS

Use immunoprecipitated DNA to generate DNA libraries:



Linkers serve as uniform primer binding sites. This allows the amplification of the entire DNA library using only 2 types of oligonucleotides

**Amplified library**

**READY FOR MASSIVE PARALLEL SEQUEUNCING**

## ChIP seq: Analysis of epigenetic information on the single nucleotide level → GENERATION OF GENOME WIDE EPIGENETIC MAPS

### Illumina Massively Parallel Sequencing

<https://www.illumina.com/company/video-hub/pfZp5Vgsbw0.html>



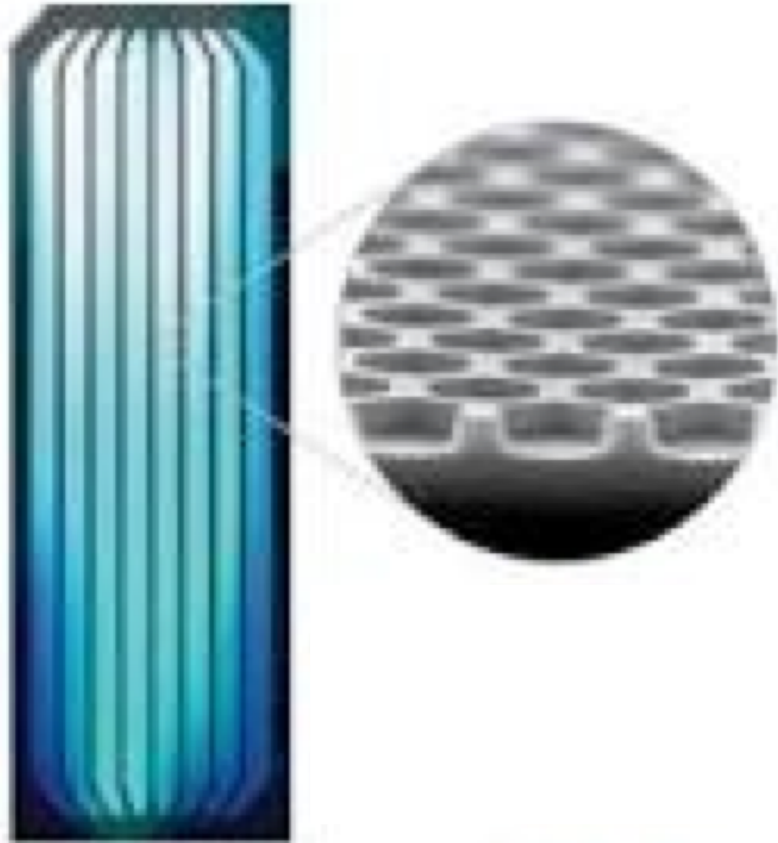
The heart of the Illumina Massive Parallel Sequencer is the “FLOW-CELL”. A surface with millions of small wells that allow thousands of Sanger-sequencing reaction In parallel = “massive parallel sequencing”. In each well a SINGLE MOLECULE of DNA Is amplified and sequenced

Illumina offers the most potent massive sequencing instruments – leader on the market

<https://www.youtube.com/watch?v=pfZp5Vgsbw0>

**ChIP seq: Analysis of epigenetic information on the single nucleotide level**  
**→ GENERATION OF GENOME WIDE EPIGENETIC MAPS**

**CLUSTER AMPLIFICATION:**



**Flow cell contains surface with millions of wells**

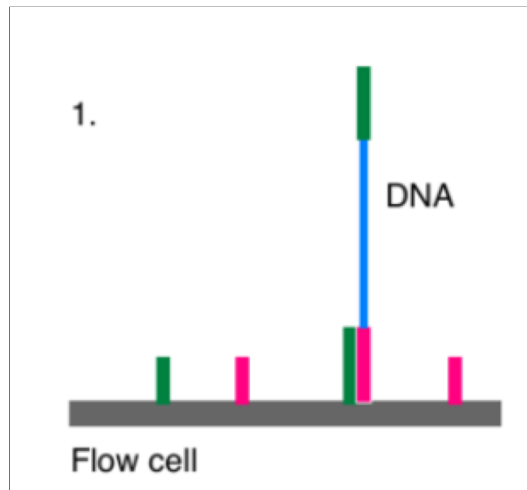
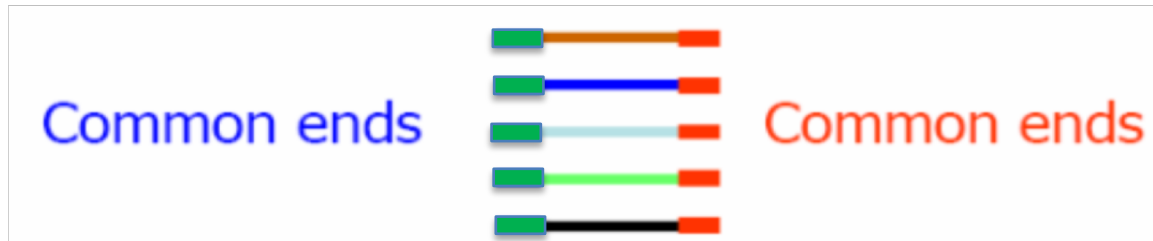
**→ Each well contains beads mounted with 2 species of oligonucleotides that hybridize with adaptor oligos of DNA library**

**→ DNA library will be loaded onto the flow cell in a determined concentration:**

**ONLY ONE MOLECULE OF DNA WILL BE  
PROCESSED FOR SEQUENCING IN A SINGLE  
WELL**

# CLUSTER AMPLIFICATION:

- making DNA library (~300bp fragments)
- ligation of adapters **A** and **B** to the fragments



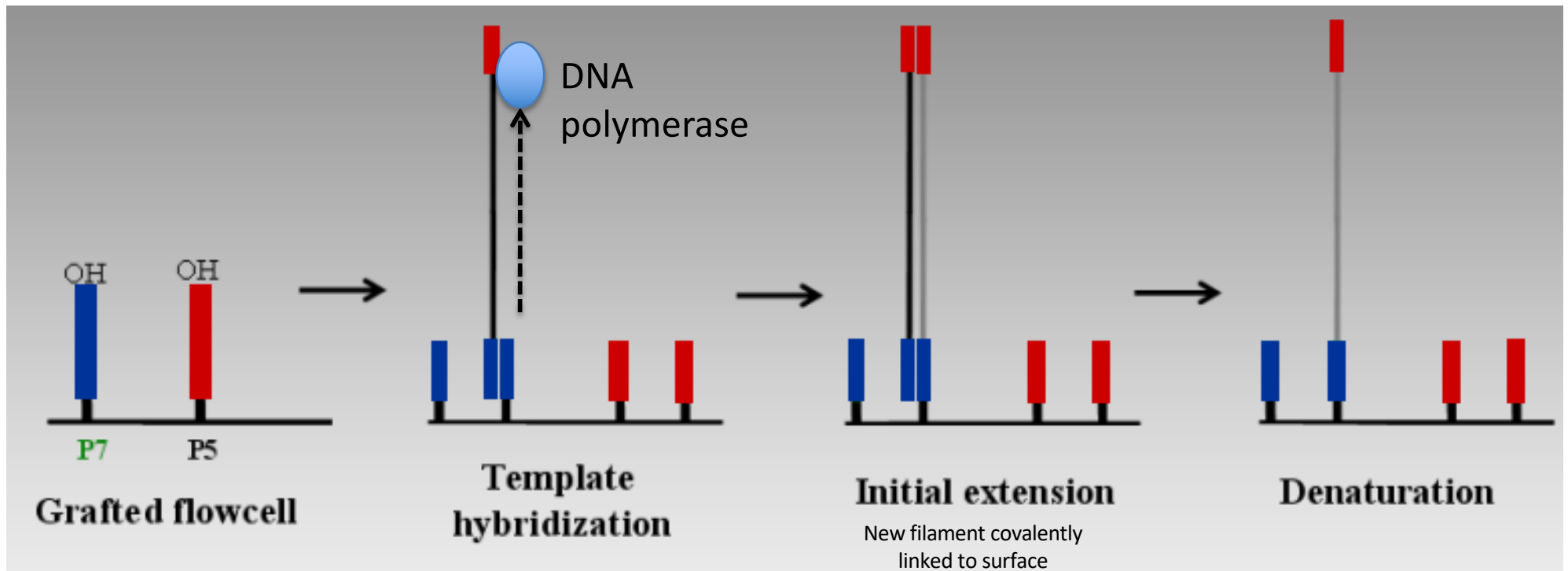
1 well, covered with  
millions of 2 types of oligos

1 well in a flow-cell with  
billions of wells

- **complementary** primers are ligated to the surface
- pairing with ChIP ed ssDNA at random position in the well of the flow cell

# CLUSTER AMPLIFICATION:

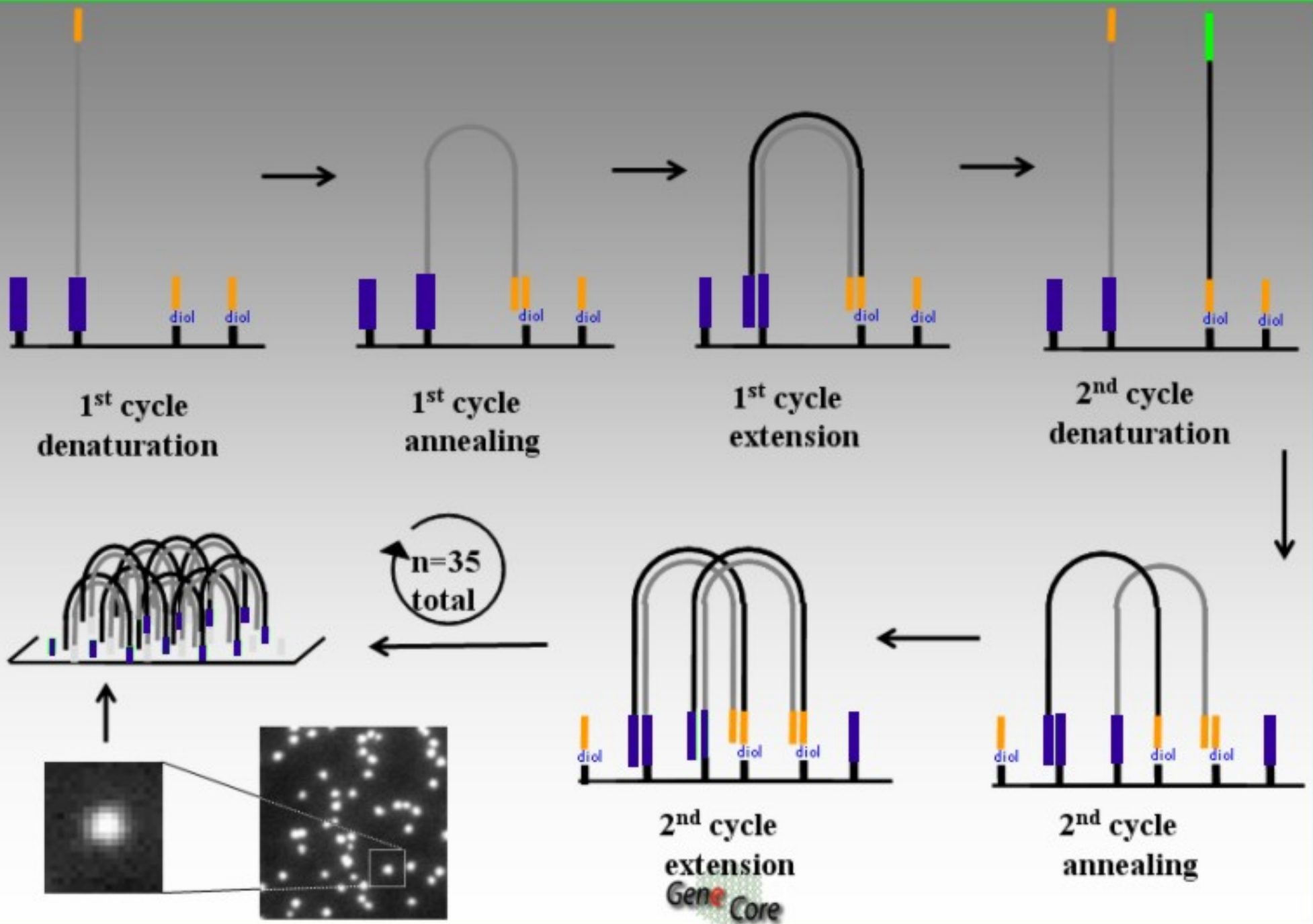
**Bridge amplification:** takes place on surface of beads (each bead is mounted with 2 species of oligos; each oligo can hybridize to a DNA library fragment):  
initiation



On the surface: complementary oligos



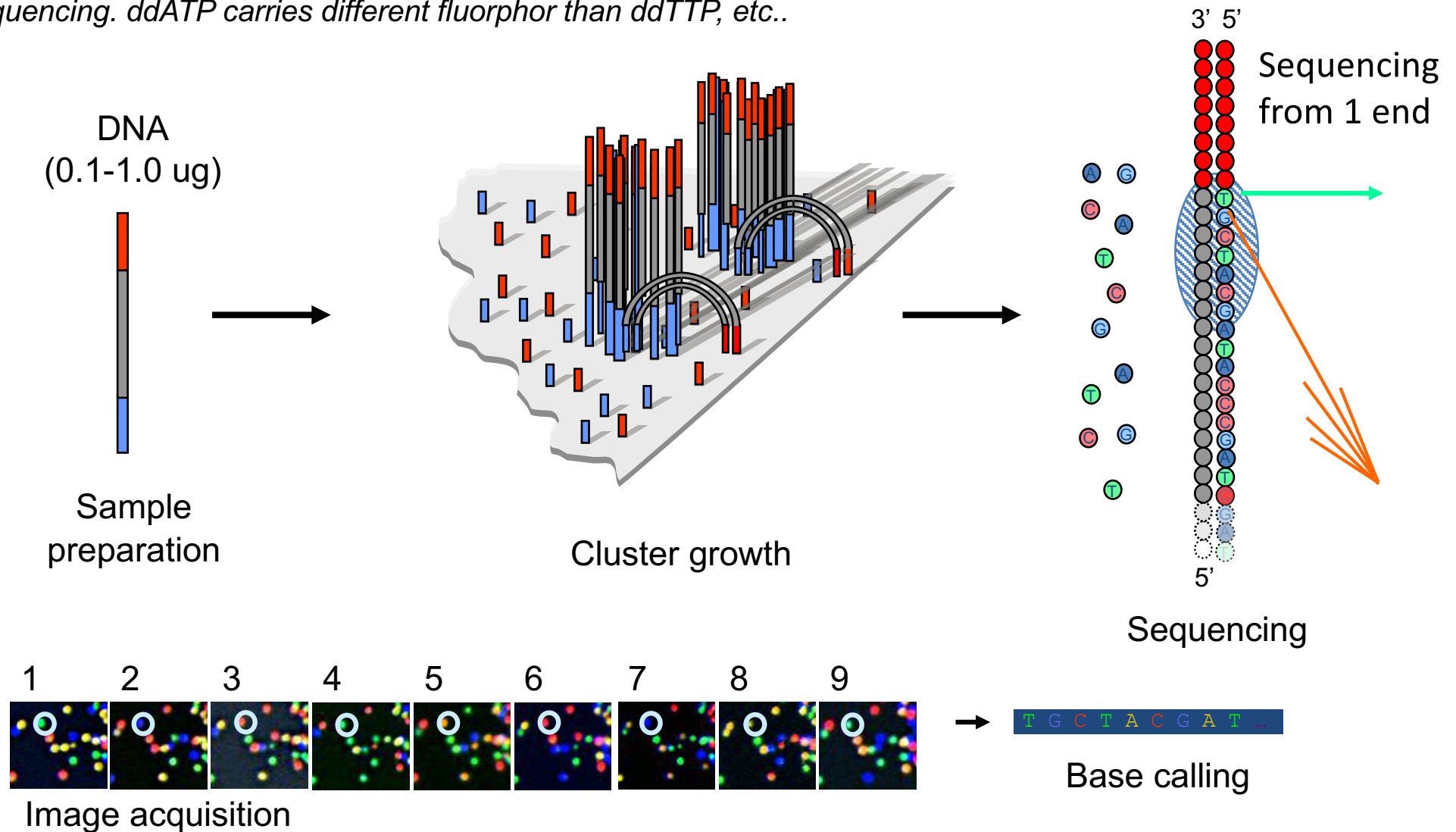
# CLUSTER AMPLIFICATION:



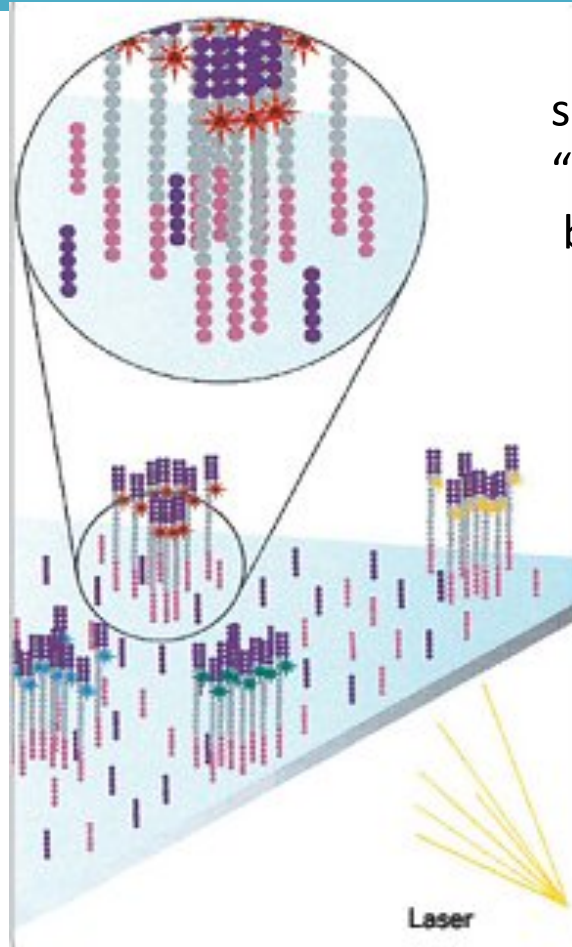
# Illumina Sequencing Technology

## Robust Reversible Terminator Chemistry Foundation

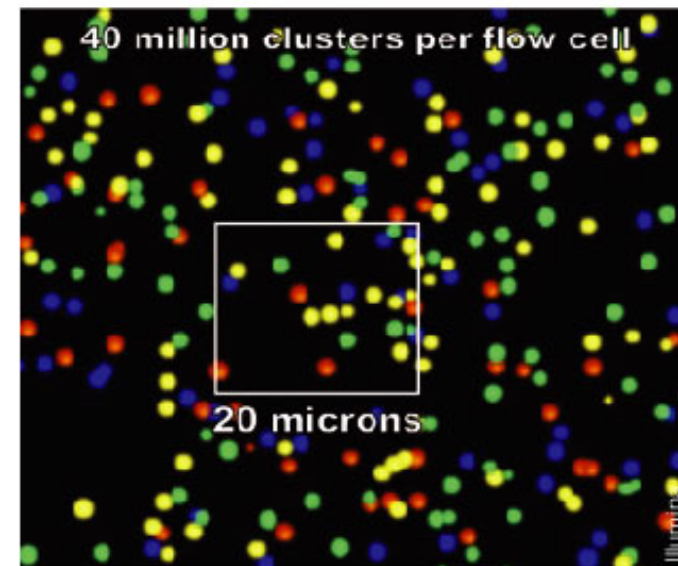
*In each round of sequencing a fluorescently labelled ddNTP will be used for sequencing. ddATP carries different fluorophore than ddTTP, etc..*



# Illumina: massive parallel sequencing:



sequencing by synthesis:  
“**reversible terminator**” nucleotides  
blocked + fluorescently labeled

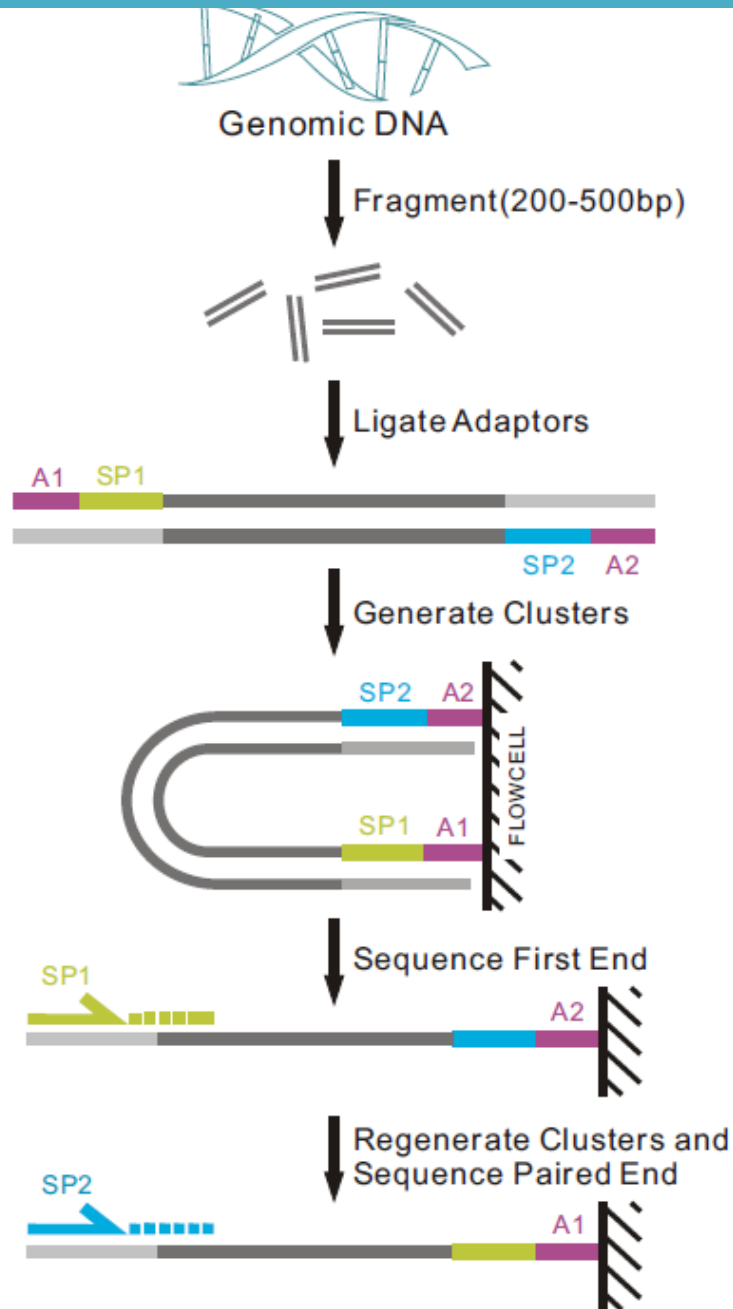


1. Synthesis = incorporation of fluorescent nucleotide: blocking synthesis
2. 4. Scanning of fluorescent signal
3. dye cleavage + elimination
4. wash step

1. Synthesis = incorporation of fluorescent nucleotide: blocking synthesis

**READ LENGTH: ca: 150nt from each primer (2x150nt = 300nt)**

# Illumina: paired end sequencing increases information content



After 1° strand sequencing, A1 anneals to A2 in nanowell-->DNA synthesis --> template strand cleaved off → new strand sequenced

<https://www.youtube.com/watch?v=9YxExTSwgPM>

Figure 1-2-1 Pipeline of paired-end sequencing (www.illumina.com)

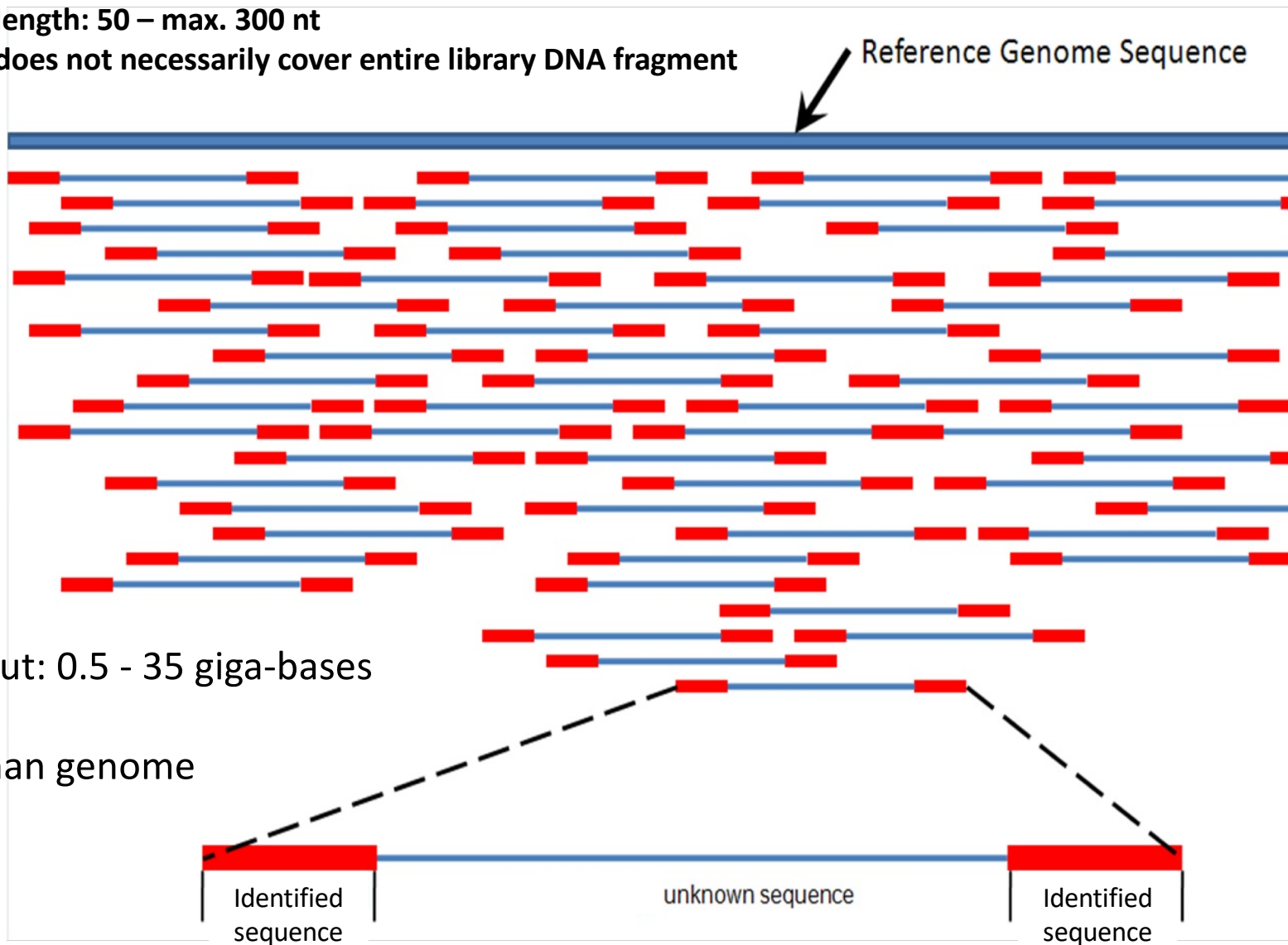
**Data analysis: obtained sequence reads are aligned  
along genomic DNA sequence → high number of reads necessary to obtain  
full sequence coverage**

Read length: 50 – max. 300 nt

Read does not necessarily cover entire library DNA fragment

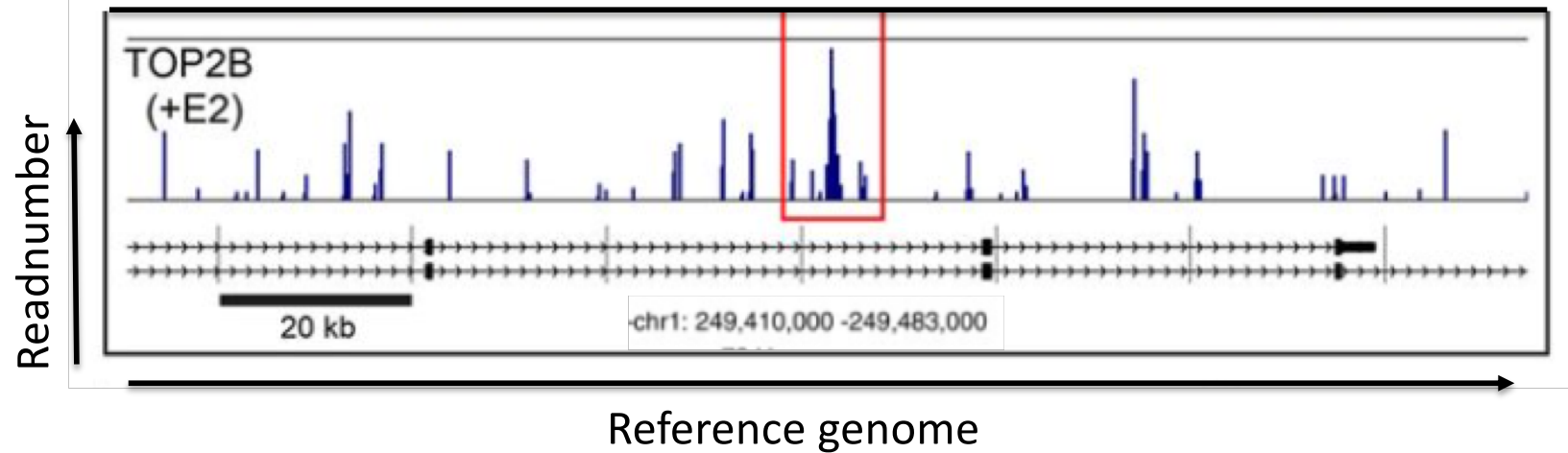
Reference Genome Sequence

Max. output: 0.5 - 35 giga-bases  
 $= 3.5 \times 10^{10}$   
 $= 10 \times$  human genome



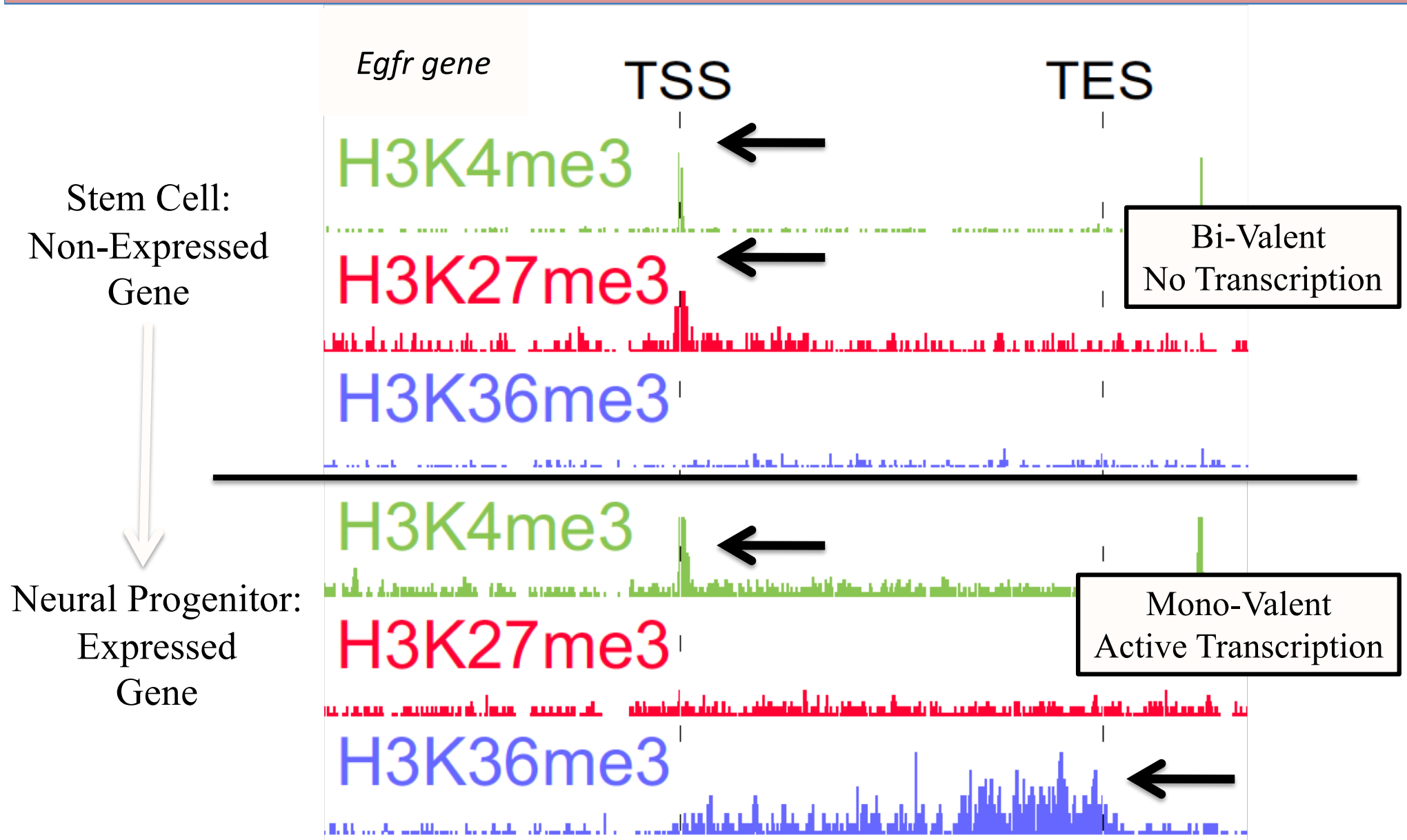
*Sequence derived from one amplified cluster*

**Data analysis: obtained sequence reads are aligned  
along genomic DNA sequence → high number of reads necessary to obtain  
full sequence coverage**

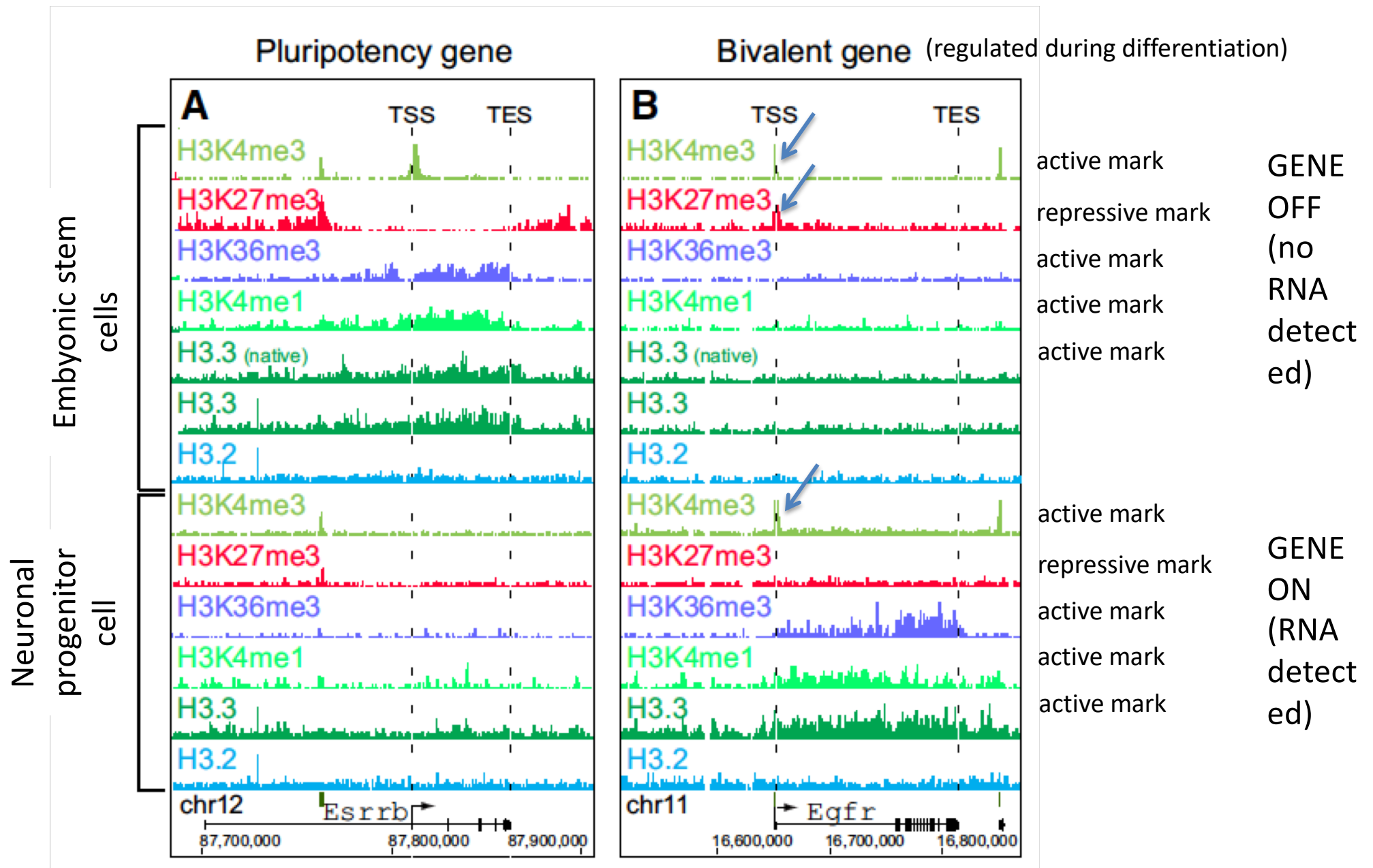




# BIOINFORMATICS ANALYSIS: Mapping ChIP seq reads agins the human genomic sequence

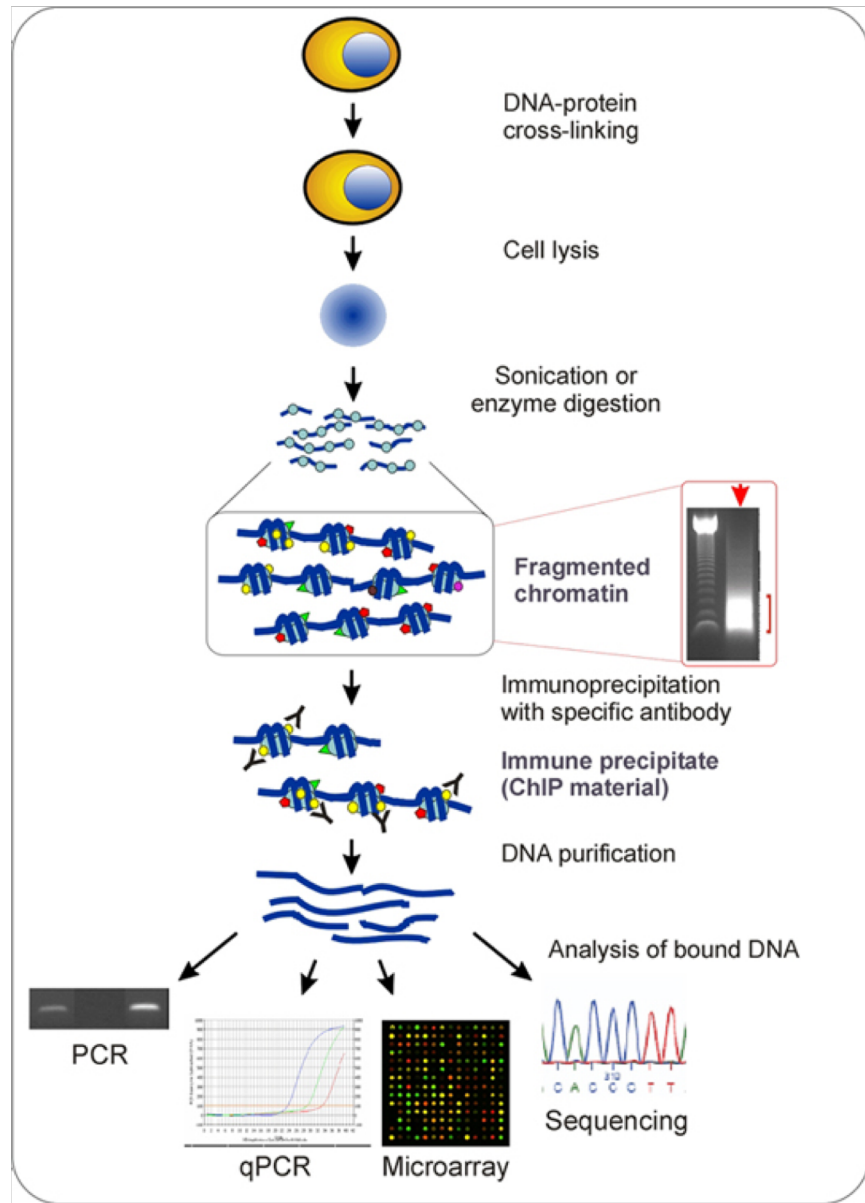


# Mapping the epigenetic landscape enables to define “key rules” to define the epigenetic code of active and silent genes





## CHROMATIN IMMUNOPRECIPITATION → DETAILED ANALYSIS → Localization of protein at a defined region - sequence

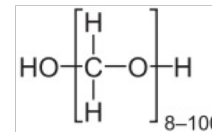


### Cell model system:

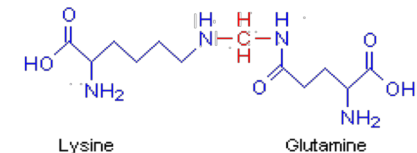
i.e. Wild-type or Suv39 dn cells that grow in cell culture dish

#### 1. Crosslink chromatin

(treatment of cells with Paraformaldehyde)



Paraformaldehyde (PFA)  
is a polymer of formaldehyde,  
And covalently links NH<sub>2</sub> groups  
of biomolecules



Example of crosslinking reaction

#### 2. Sonicate crosslinked cells

→ DNA breaks in a random manner  
until reaching a size of ca. 250 – 500nt

#### 3. Incubate chromatin fragments with antibodies

that are specific for modified histones (i.e. H3K9me3) or  
chromatin modifying enzymes (i.e. Suv39h1); another sample  
with a non-specific, control antibody

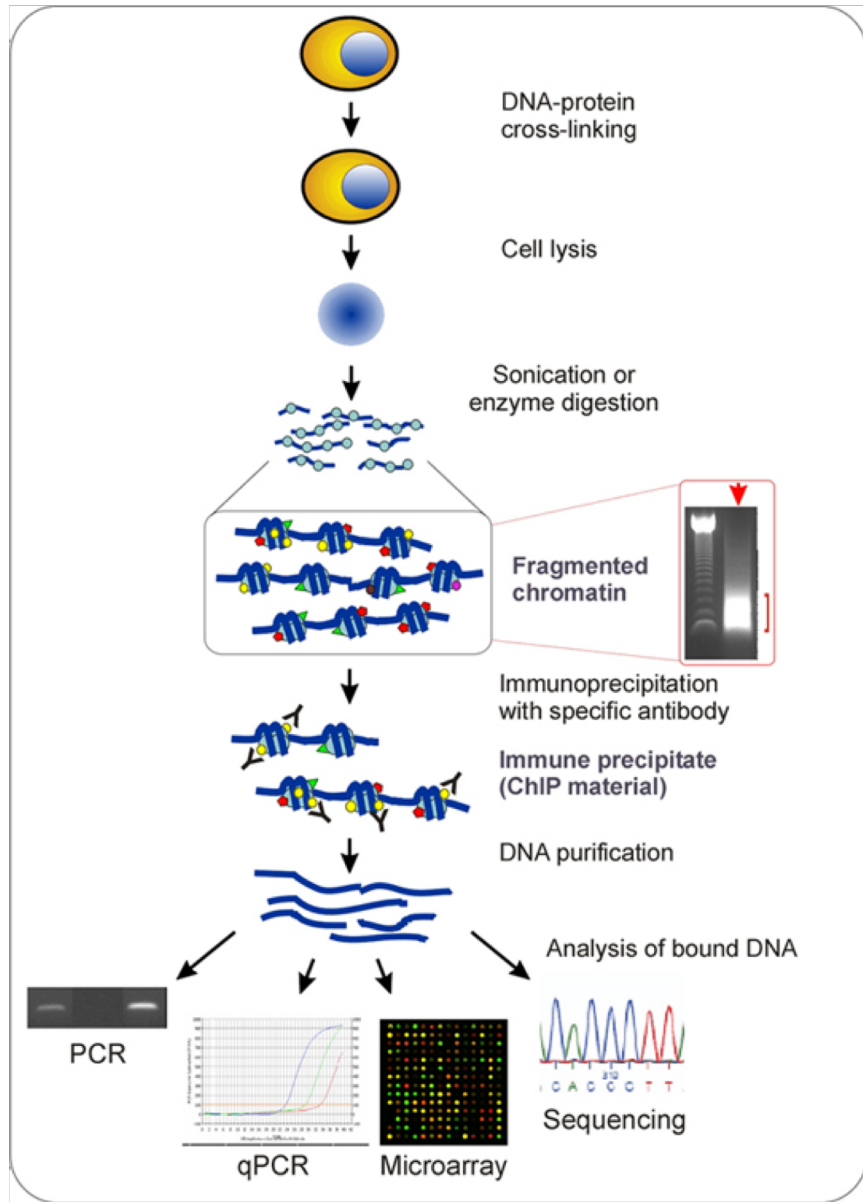
4. Recover antibody-chromatin complexes using  
a resin that binds to the constant region of antibodies

5. Elute chromatin at high salt concentration and  
revert crosslinks at high temperature

6. Digest protein with protease K and RNA with RNase

## CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILED ANALYSIS

→ Localization of protein at a defined region - sequence



5. Elute chromatin at high salt concentration and revert crosslinks at high temperature

6. Digest protein with protease K and RNA with RNase

7. Purify DNA and precipitate DNA

8. Measure the amount of immunoprecipitated DNA  
In control versus Suv39h dn cells

### METHODS:

**A. PCR:** design PCR oligos to amplify defined sequences in immunoprecipitates. Per PCR reaction only one locus can be examined by real-time PCR

### B. ChIP on ChIP:

Immunoprecipitated DNA from control and Suv39h dn cells are labeled with nucleotides that are fluorescently labeled (555nm emission fluorophore or 488 nm emission fluorophore). Samples are mixed and hybridized to an array that contains spotted DNA probes that cover a large fraction of the genome. 488/555 nm ratio is measured → relative enrichment of histone modification at thousands of genomic sites in control/Suv39h dn cells

### B. ChIP Seq:

Immunoprecipitated DNA is sequenced by massive parallel Sequencing. Control vs. Suv39h dn cells reveals enrichment of H3K9me3 at the ENTIRE genome level. single nucleotide levels