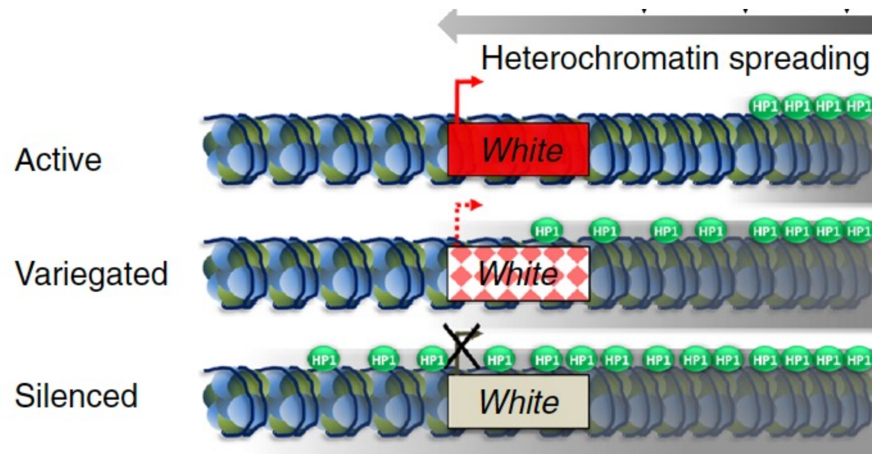
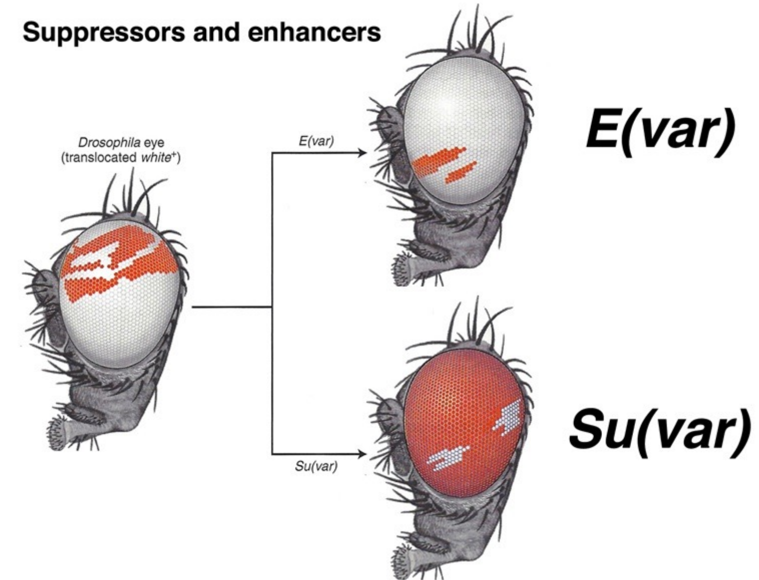
TOPIC 2: HOW TO STUDY EPIGENETIC MODIFICATIONS

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

HOW TO STUDY EPIGENETIC MODIFICATIONS
-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --





D.melanogaster: Su(var)3-9

Epigenetic writer for H3K9me3

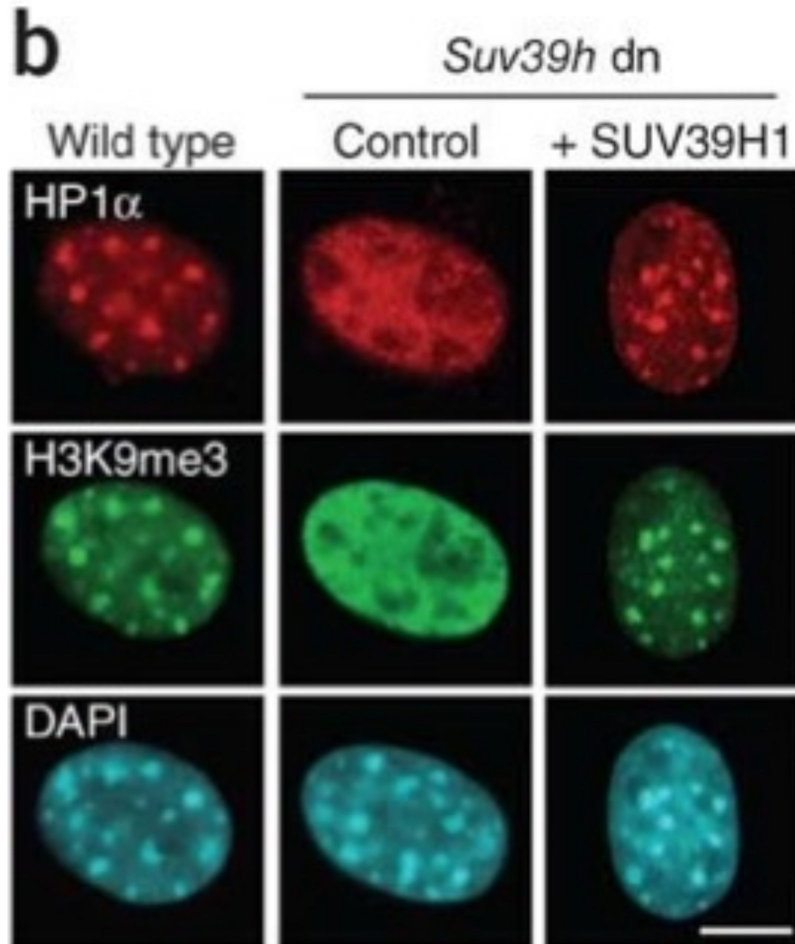
Homologs in:

H.Sapiens: SUV39H1, SUV39H2

M.Musculus: Suv39h1, Suv39h2

S. pombe: Clr4

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/Immunofluorescence)
- Mass spectroscopy

Problem:

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

-DAPI intense regions + DNA-FISH probes

-Chromatin immunoprecipitation

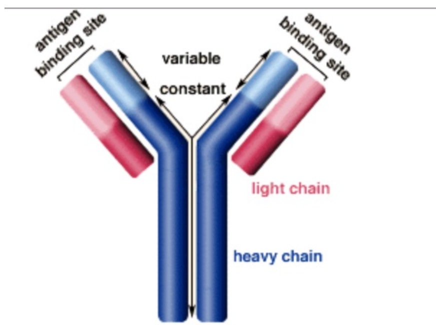
--> Detect histone modifications on specific genomic site of interest (i.e promoter of p53)
--> Detect histone modifications at multiple sites or at the entire genome level

HOW TO STUDY EPIGENETIC MODIFICATIONS

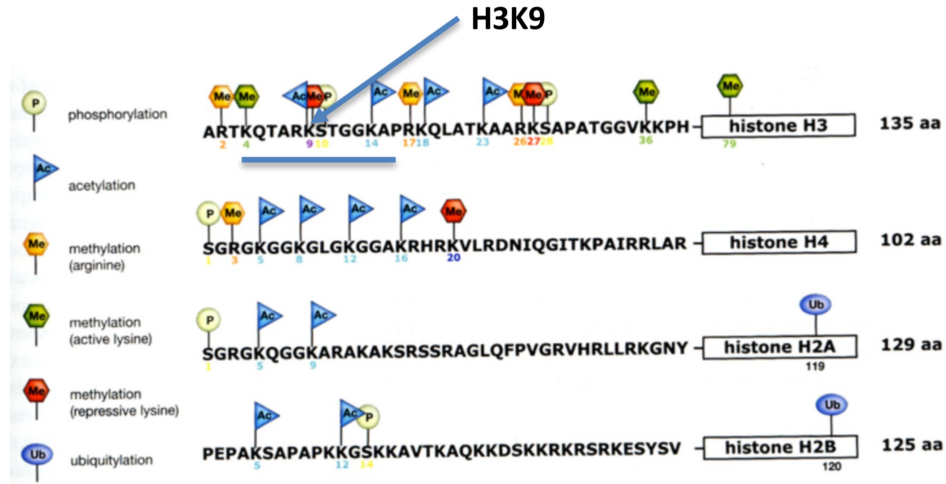
1. Specfic antibodies are central to epigenetic research?
2. Mass spectrometry can read histone codes?
3. How can we identify enzymatic activites of writers?
4. How can we identify epigenetic readers?
5. How can we locate specific histone modifications in the nucleus along the genome?
6. Identification of biological relevance of chromatin modifications

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

1. Generation of antibodies that specifically recognize modified histones



H3K9



ARTKQTARK$_9$STGGKAPRK

CAN WE GENERATE MODIIFCATION SPECIFIC ANTIBODIES??

Histone modification

Lysines: mono-, di- and tri-methylated
Argenines: mono-, di-methylated
Tyrosine, Serine, Threonine: 1 phospho group
Lysines: 1 acetyl group
Lysines: monoubiquitination (not targeted by proteasome)

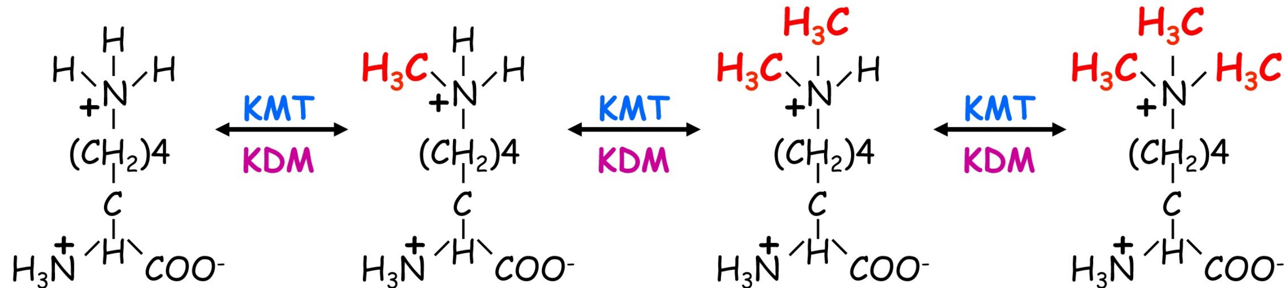
DNA modification

Cytosines: methylation (in CpG contret)

An example: H3K9methylation by Suv39h1 – Peptide length 10-20aa including K9

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

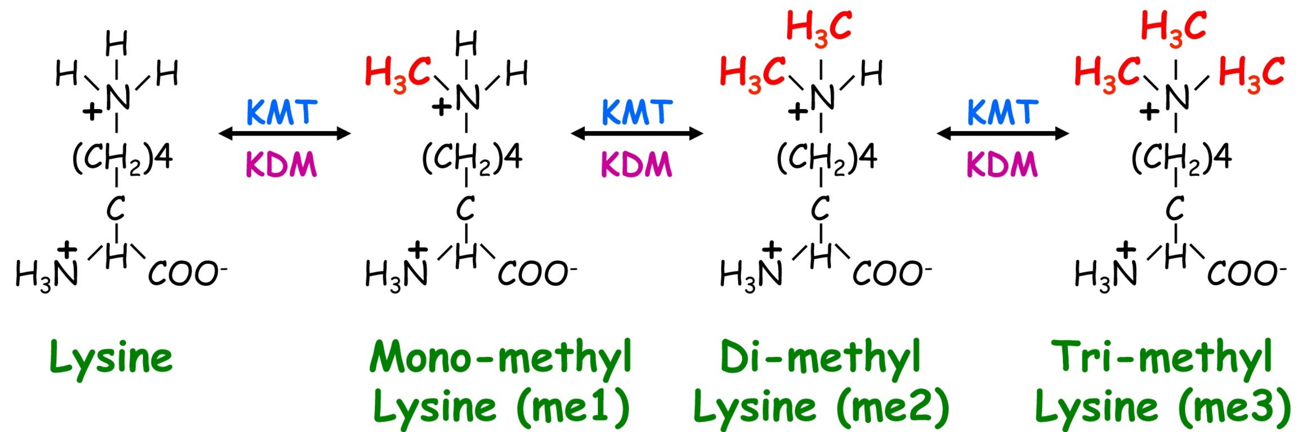
KQTARK$_9$STGGKAP



1. Generation of antibodies that specifically recognize modified histones

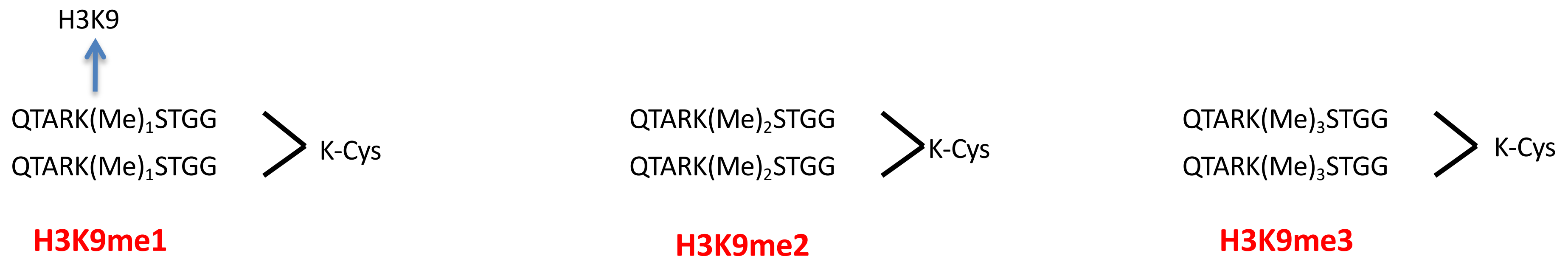
Lysines can be mono-, di- and tri-methylated

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



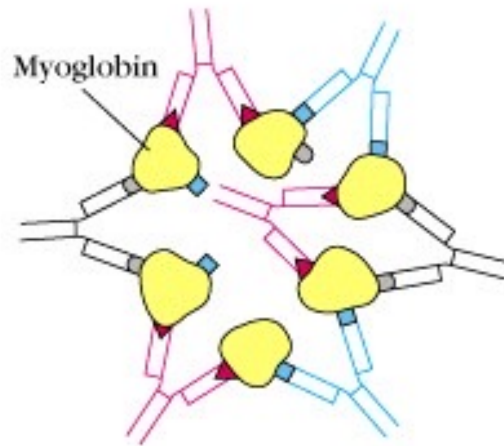
CHEMICAL SYNTHESIS OF BRANCHED PEPTIDES FOR ANTIBODY GENERATION

H3: ARTKQTARK$_9$STGGKAPRK



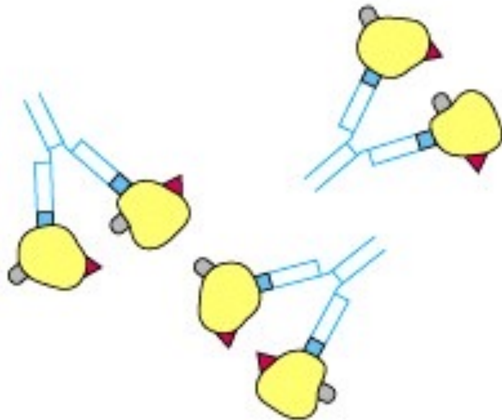
→Branching allows to place histone modifications in close vicinity (2x H3 per octamere)
→→ resembles high concentration of chromatin modification in the nucleus

Generation of polyclonal/monoclonal antibodies



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

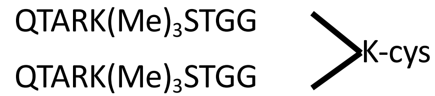
Mix of antibodies with different target sites in peptides used for immunization



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

Single type antibody with single target sites in peptides used for immunization

Generation of polyclonal antibodies



QTARK(Me)$_3$STGG

QTARK(Me)$_3$STGG

K-cys

H3K9me3

1 Antigen synthesis

2 12-week antibody production

3 Polyclonal antibody characterization & purification

Non-immunogenic small molecule

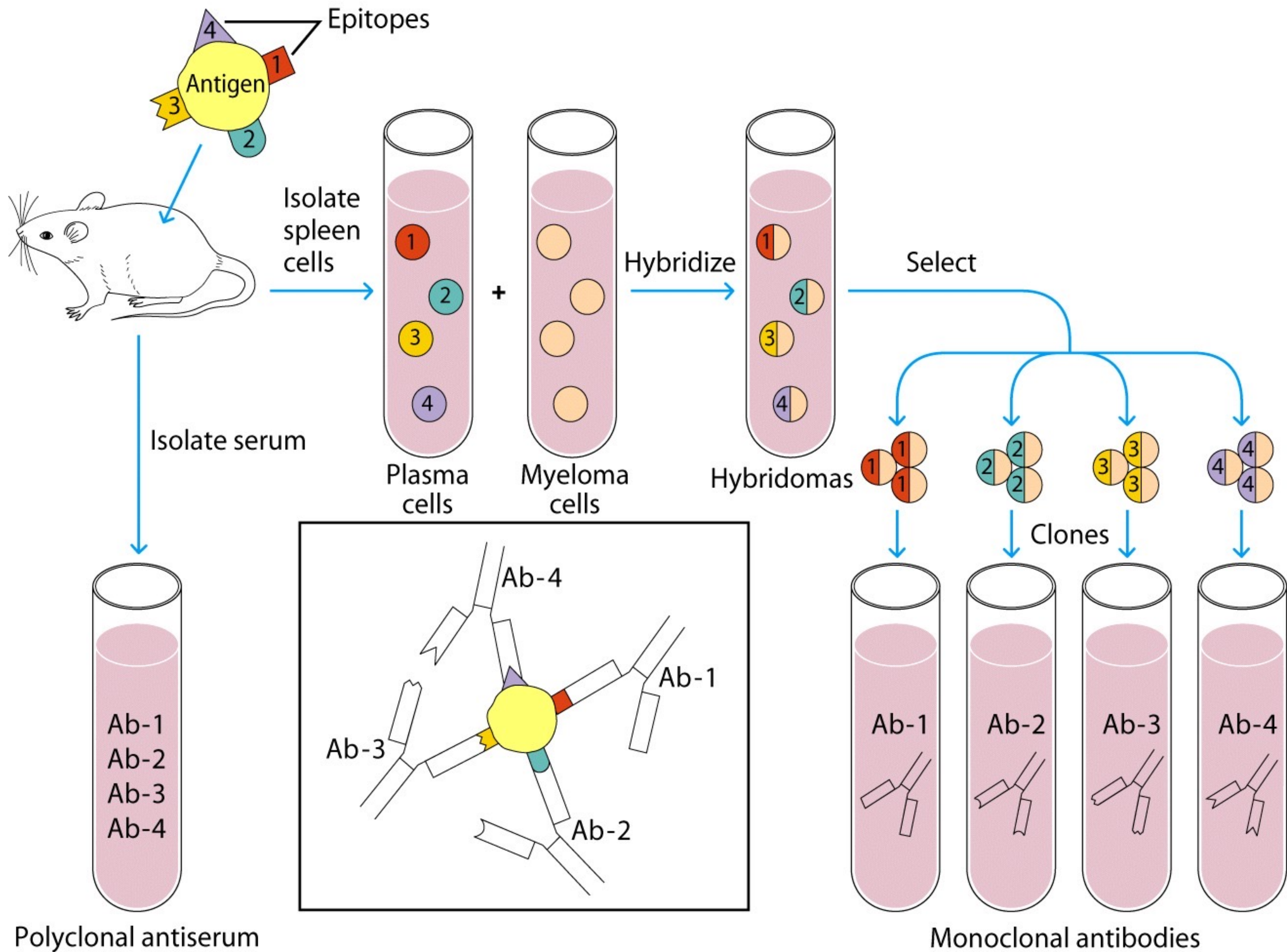
Immunogenic protein carrier

Antigenic conjugate

2 rabbits, 4 immunizations, 2 bleeds per rabbit

ELISA evaluation (titer, affinity, specificity) & amonium sulfate purification

Generation of monoclonal antibodies



Generation of monoclonal antibodies

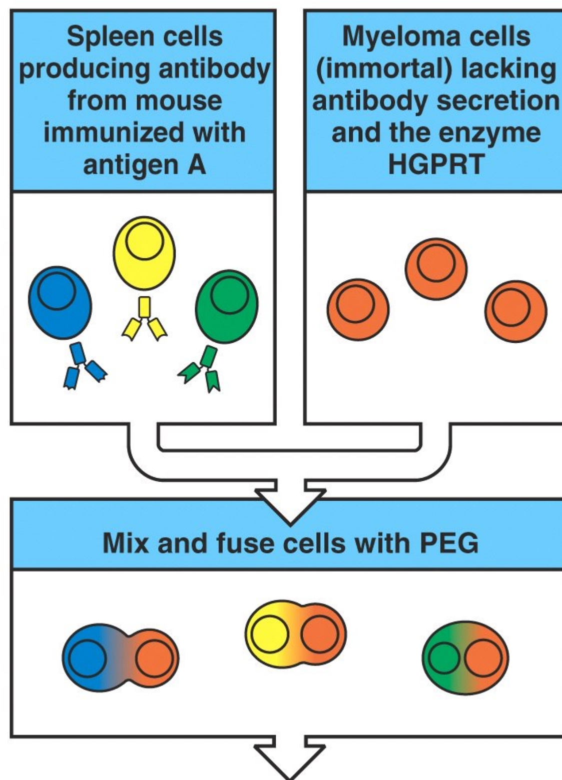


Figure A-14 part 1 of 2 Immunobiology, 6/e. (© Garland Science 2005)

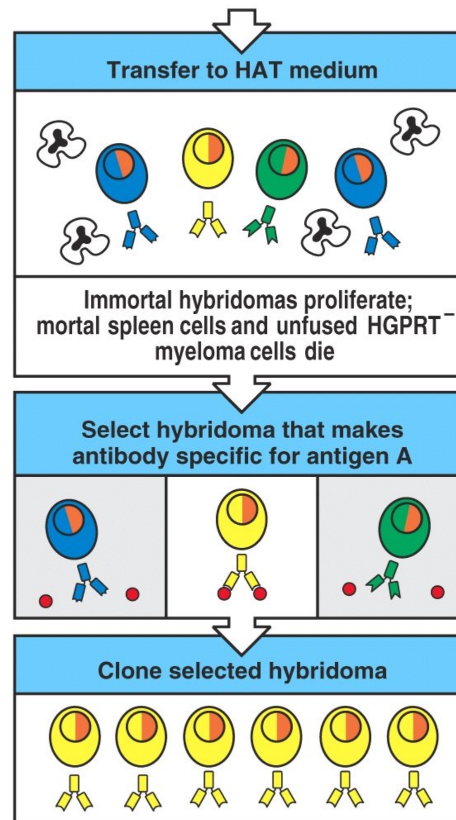
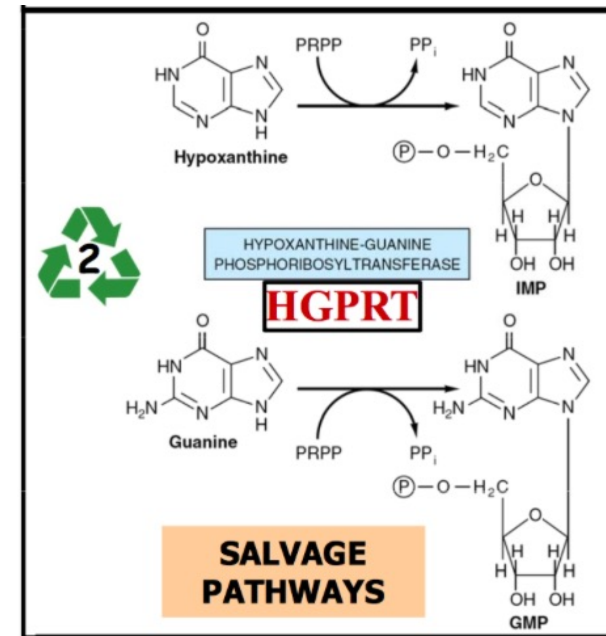


Figure A-14 part 2 of 2 Immunobiology, 6/e. (© Garland Science 2005)



IMP: converted to either adenine (AMP) or guanine (GMP) nucleotides

HAT Medium (hypoxanthine-aminopterin-thymidine medium) is a selection medium for mammalian cell culture, which relies on the combination of aminopterin and hypoxanthine (a purine derivative) and thymidine (a deoxynucleoside). Aminopterin blocks DNA replication by acting as a potent antifolate, specifically inhibiting the enzyme **Dihydrofolate reductase (DHFR)**. This inhibition halts the conversion of dihydrofolate to tetrahydrofolate, a necessary cofactor for the *de novo* synthesis of **purines and thymidine**. The resulting depletion of nucleotide precursors (dNTPs) stalls DNA synthesis, preventing cell division. Hypoxanthine and thymidine provide the essential raw materials for DNA synthesis via the "salvage pathway," allowing cells to bypass the block on de novo nucleotide synthesis imposed by aminopterin, provided that they have the HGPRT gene → 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 event

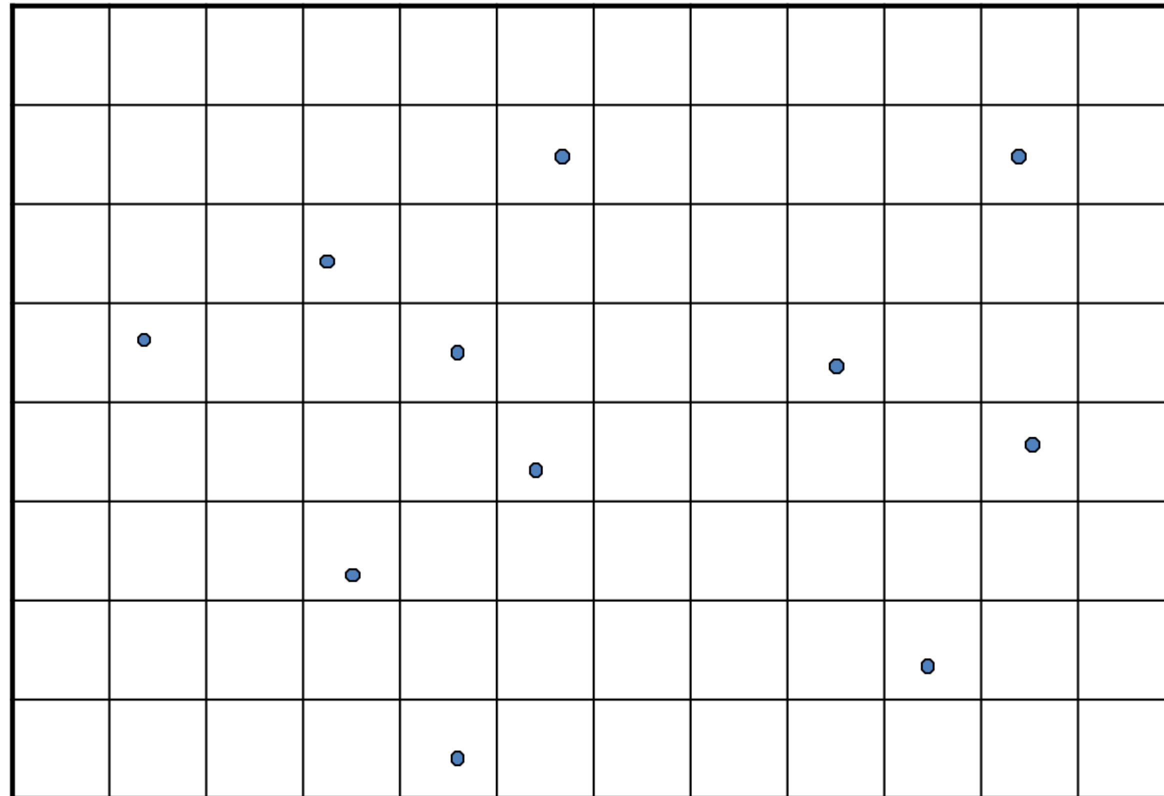
1. Plate at limiting dilution (<1 cell/well) in 96 well plates.

- for example: volume per well: 100 microliter
- use fused cells and prepare cell suspension of ca. 1 cell/ml (0,1 cell per100ul well → 1 cell every 10 wells)

2. Allow clones to expand in 96 well = cell clone or cell line (a population of hybridoma cells derived from a single cell = genetically identical) → Hybridoma line

3. Further expand positive well and test for production of antibody of desired specificity in culture supernatant

Attention:
Each clone produces a monoclonal antibody with specificty for a single epitope presented by the peptide used for immunization



Generation of monoclonal antibodies

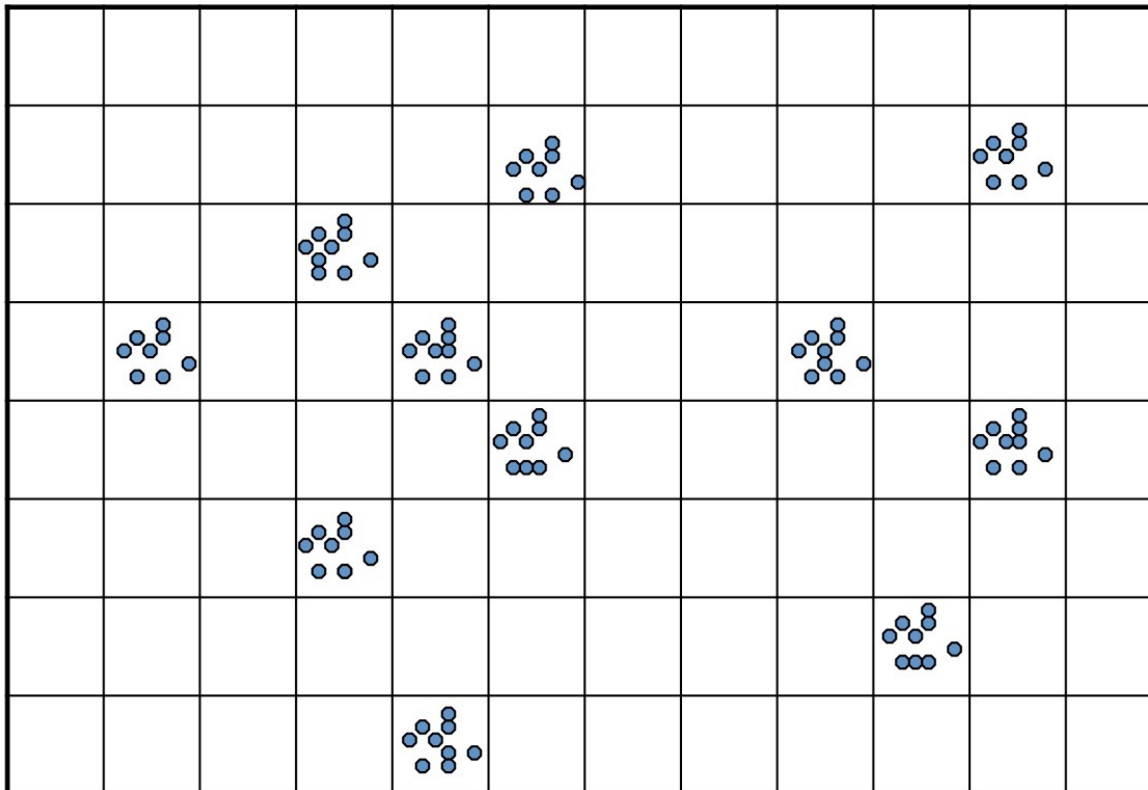
Cloning hybridomas from fusion event

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

- for example: volume per well: 100 microliter
- use fused cells and prepare cell suspension of ca. 1 cell/ml (0,1 cell per well → 1 cell every 10 wells)

2. Allow clones to expand in 96 well = cell clone or cell line (a population of hybridoma cells derived from a single cell = genetically identical) → Hybridoma line

3. Further 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 antibody specificity - biochemically

Peptide Spotted on membrane

QTARK(Me)$_2$STGG
QTARK(Me)$_2$STGG
K-cys

QTARK(Me)$_2$STGG
QTARK(Me)$_2$STGG
K-cys

QTARK(Me)$_3$STGG
QTARK(Me)$_3$STGG
K-cys

QTARKSTGG
QTARKSTGG
K-cys

A

H3-K9 methylation antibodies

H3: K4, K9, K27, K36, K79; H4-K20

unmod, Me, (Me)2, (Me)3

pmol: 50, 10, 2, 0.4, 0

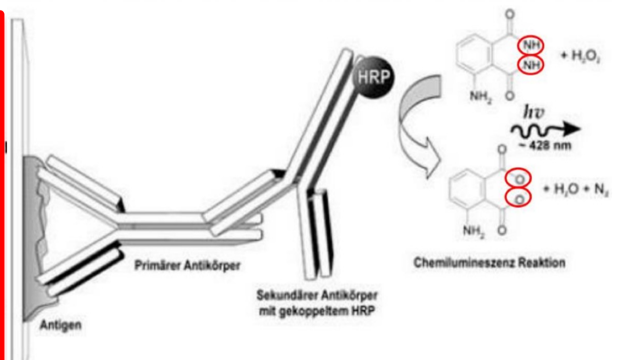
antibody used:
α-2x-mono-meth H3-K9
#4858, IgG, 1:2500
H3(aa 5-15) K9(Me)

antibody used:
α-2x-di-meth H3-K9
#4679, IgG, 1:2500
H3(aa 5-15) K9(Me)2

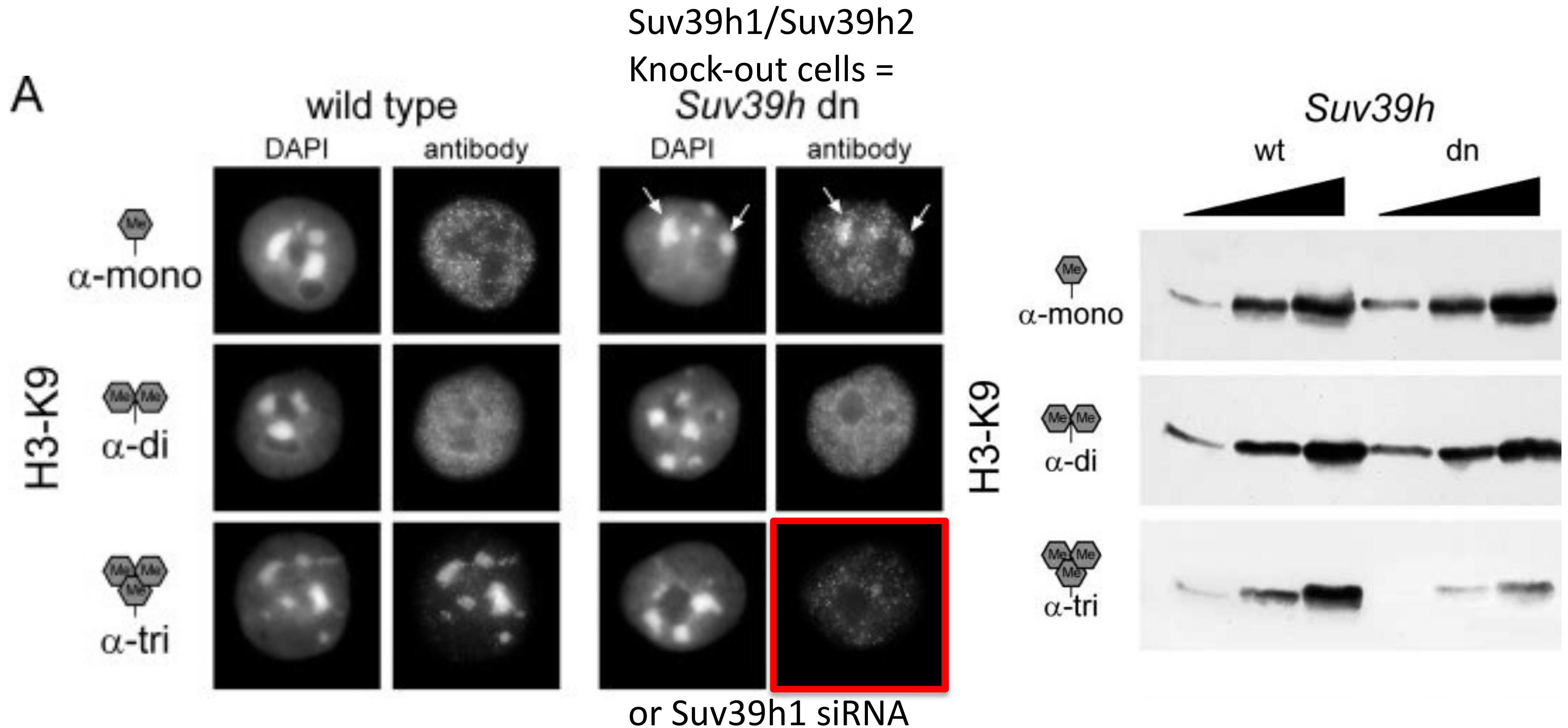
antibody used:
α-2x-tri-meth H3-K9
#4861, IgG, 1:2500
H3(aa 5-15) K9(Me)3

Validation by dot-blot
(peptide used for immunization and control peptides 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



Validation of antibody specificity – in cells (with loss of function experiments)



Semiquantiative information (staining intesity)
Qaulitative infromation on localization of modification (low resolution)

Semiquantiative information
(overall abundance of modification)

In Suv39h DN cells, H3K9me2 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

HOW TO STUDY EPIGENETIC MODIFICATIONS

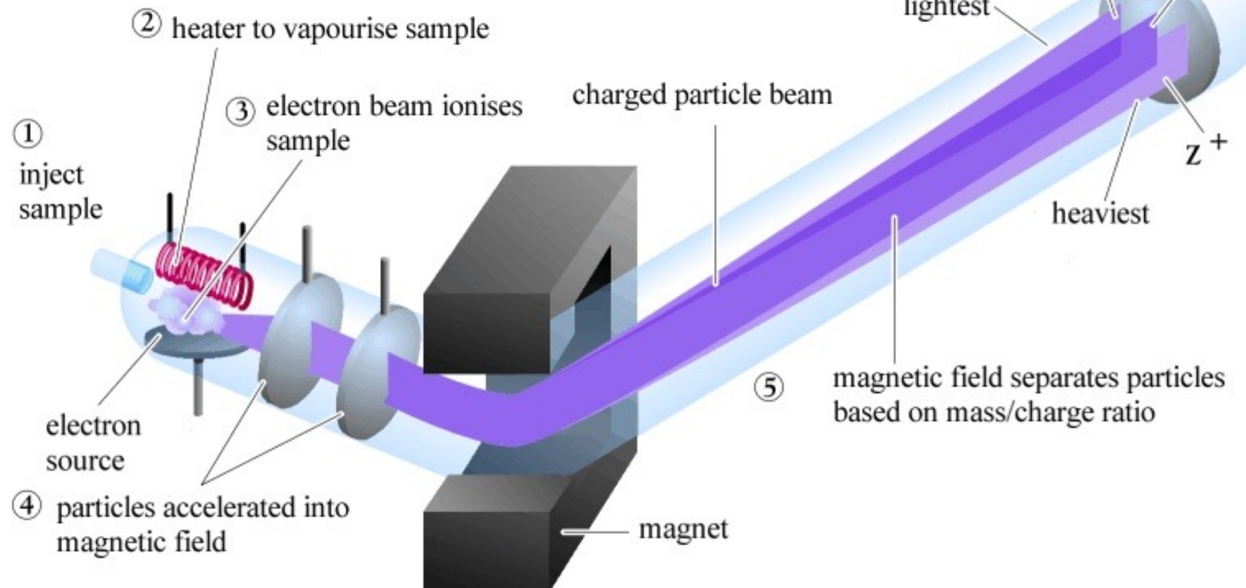
1. **Specfic antibodies are central to epigenetic research?**
2. **Mass spectrometry can read histone codes?**
3. **How can we identify epigenetic readers?**
4. **How can we identify enzymatic activites of writers?**
5. **How can we locate specific histone modifications in the nucleus along the genome?**
6. **Identification of biological relevance of chromatin modifications**

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

2. Studying histone modifications by mass spectrometry

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-change ratio will show different deflection

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

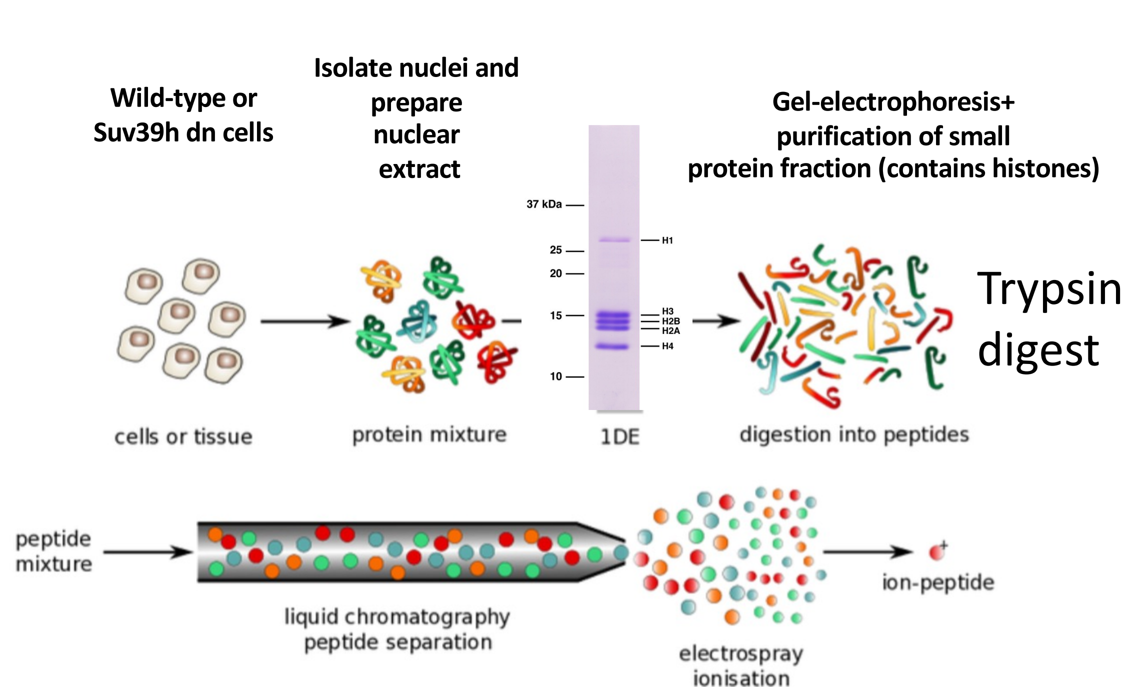
The ions are detected by a mechanism capable of detecting charged particles, such as an electron multiplier. Results are displayed as spectra of the relative abundance of detected ions as a function of the mass-to-charge ratio.

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

MW of all amino acids and all their possible modifications are known = identifiable by mass-to-change ration: also when modifications are present in a series of aminoacids of the particle detected by the detector

MOST IMPORTANT: provide a sample with defined peptide fragments → achieved by digest with proteases that cut peptides at defined positions (comparable with restriction enzymes)

2. Studying histone modifications by mass spectrometry



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

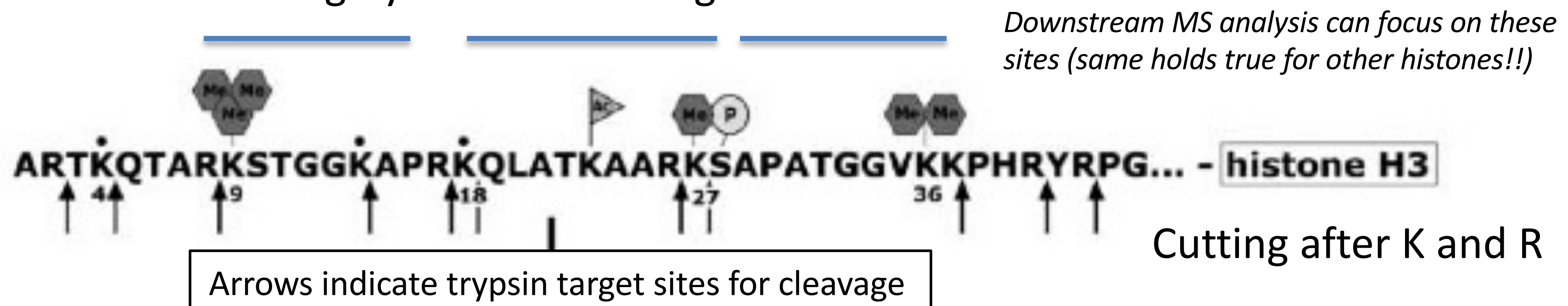
Protease= Trypsin

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

(smaller petides are easier to follow by Mass Spec)

IMPORTANT: we know already all proteins and their amminoacid sequence!!!

That means we can predict all possible small peptide sequences that result from a trypsin cleavage



Downstream MS analysis can focus on these sites (same holds true for other histones!!)

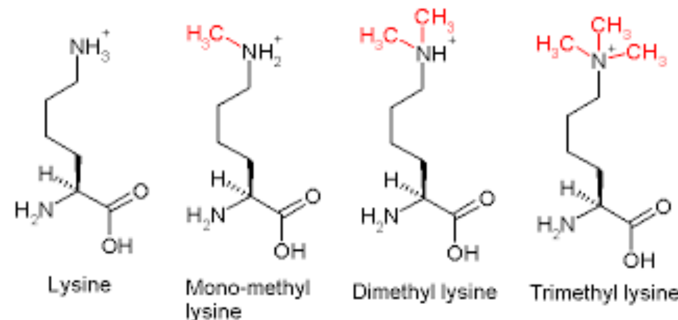
Cutting after K and R

Arrows indicate trypsin target sites for cleavage

Histone modifications change m/z ratio in mass spec experiments

Table of amino acid residues .

Symbol	Structure	Mass (Da
Ala A	-NH.CH.(CH_3).CO-	71.0
Arg R	-NH.CH.[(CH_2)$_3$.NH.C(NH).NH_2].CO-	156.1
Asn N	-NH.CH.(CH_2CONH_2).CO-	114.0
Asp D	-NH.CH.(CH_2COOH).CO-	115.0
Cys C	-NH.CH.(CH_2SH).CO-	103.0
Gln Q	-NH.CH.($CH_2CH_2CONH_2$).CO-	128.1
Glu E	-NH.CH.(CH_2CH_2COOH).CO-	129.0
Gly G	-NH.CH_2.CO-	57.0
His H	-NH.CH.($CH_2C_3H_3N_2$).CO-	137.1
Ile I	-NH.CH.[CH.(CH_3)CH_2.CH_3].CO-	113.1
Leu	-NH.CH.[CH_2CH(CH_3)$_2$].CO-	113.1
Lys K	-NH.CH.[(CH_2)$_4$$NH_2$].CO-	128.1
Met M	-NH.CH.[(CH_2)$_2$.SCH_3].CO-	131.0
Phe F	-NH.CH.(CH_2Ph).CO-	147.1
Pro P	-NH.(CH_2)$_3$.CH.CO-	97.1
Ser S	-NH.CH.(CH_2OH).CO-	87.0
Thr T	-NH.CH.[CH(OH)CH_3].CO-	101.0
Trp W	-NH.CH.[CH_2.C_8H_6N].CO-	186.1
Tyr Y	-NH.CH.[(CH_2).C_6H_4.OH].CO-	163.1
Val V	-NH.CH.[CH(CH_3)$_2$].CO-	99.1



KSTGGK: MW= 577 g/mol

CH_3 : MW= 15 g/mol

Kme_3STGGK: MW=577 – 3 + 45 = 619 g/mol

Kme_2STGGK: MW= 577 – 2 + 30 = 605 g/mol

Kme_1STGGK: MW= 577 – 1 + 15 = 591g/mol

Molecular weight

Protonation changes MW in a defined manner

Ionization

For details: An Introduction to Mass Spectrometry http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm

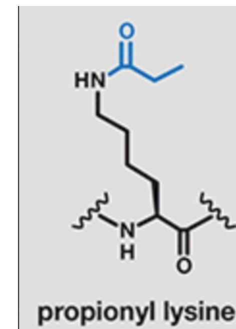
2. Studying histone modifications by mass spectrometry

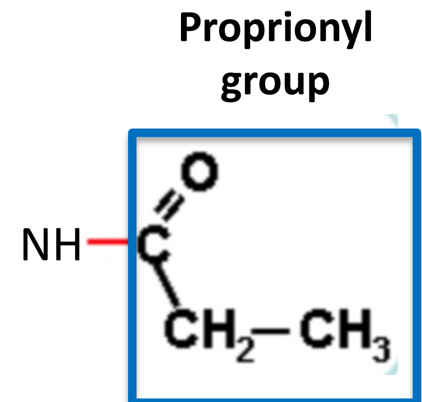
PROBLEM:

1. **Trypsin cleavage after K is not complete because me2 or me3 on lysines block the cutting by Trypsin**

2. **K is frequent in histone tails --> allows cleavages but only at me0, me1 → in a cell population a modification of an aminoacid is not necessarily present at 100% → many small heterogeneous peptides → difficult to analyze**

Trypsin can cleave after Arg (R)
Trypsin can cleave after K-un-methylated; K-me1
Trypsin cannot cleave after K-me2; K-me3





SOLUTION: *In vitro* Propionylation of unmodified or mono-methylated lysine prevents trypsin cleavage

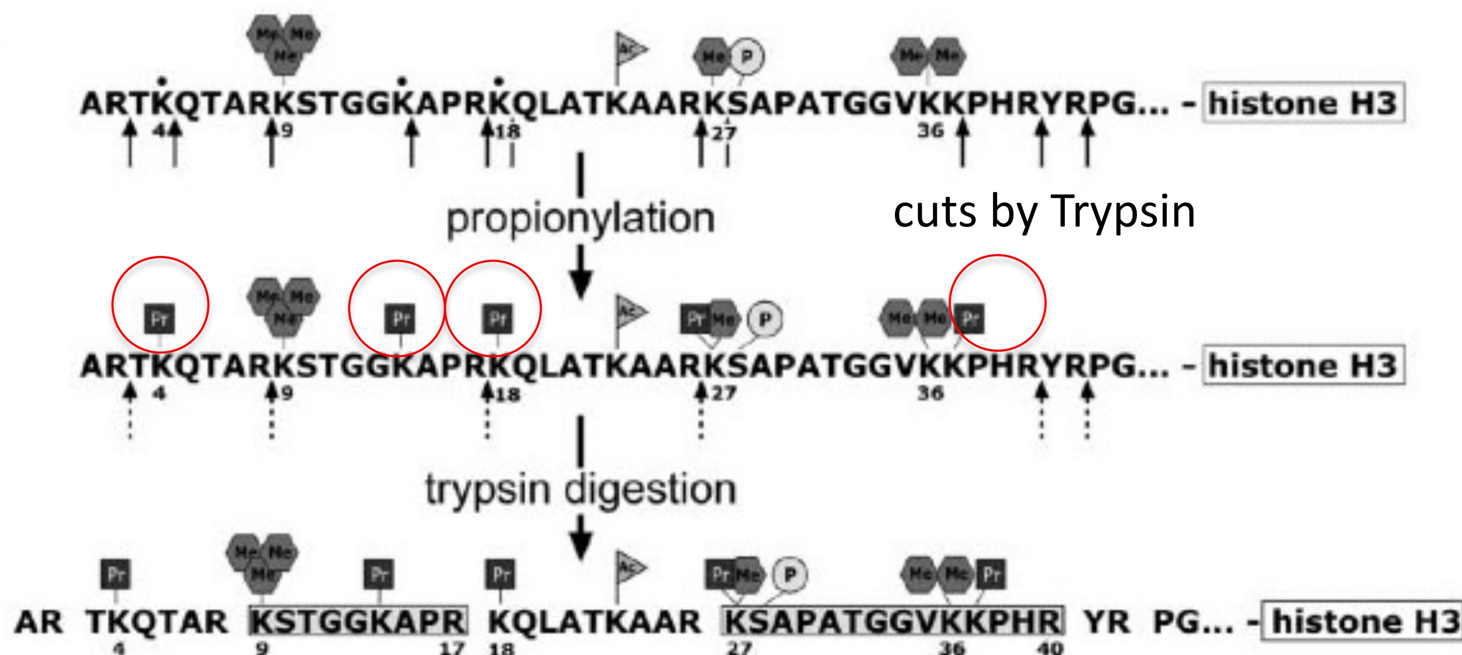
→ Use extract and perform chemical reaction to drive propionylation at all K_{me0}, and K_{me1}

→ now Trypsin can only cut after Arginine.

→ this reduces the complexity of fragments after cutting.

→ creates a mass:charge ratio that allows to differentiate between fragments carrying me0, me1, me2 or me3 marks (me0 + 3x propionyl group; me1 + 2x propionyl group; me2 + 1 propionyl group; me3 + 0x propionyl groups)

2. Studying histone modifications by mass spectrometry



propionylation

bigger and homogeneous fragments

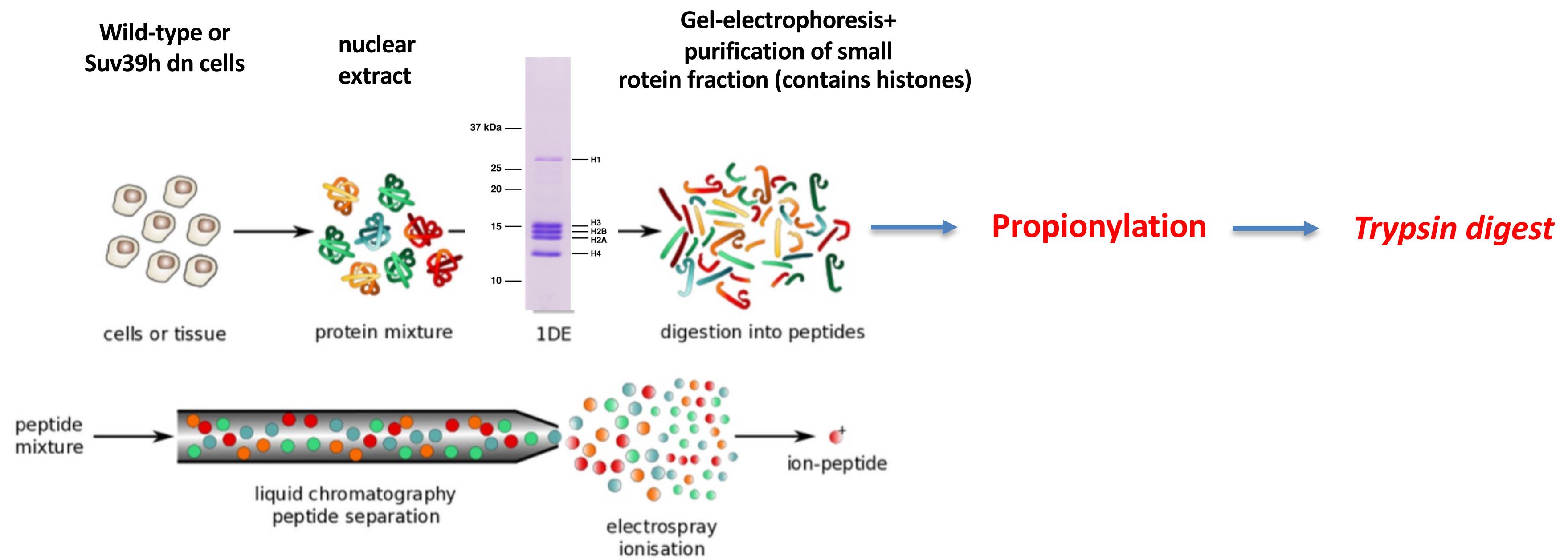
TRYPSIN cuts ONLY at Arg by Trypsin after propionylation

Now Trypsin can no longer cut after Lysines, only after Argenines. This allows a much more uniform cleavage of histone tails

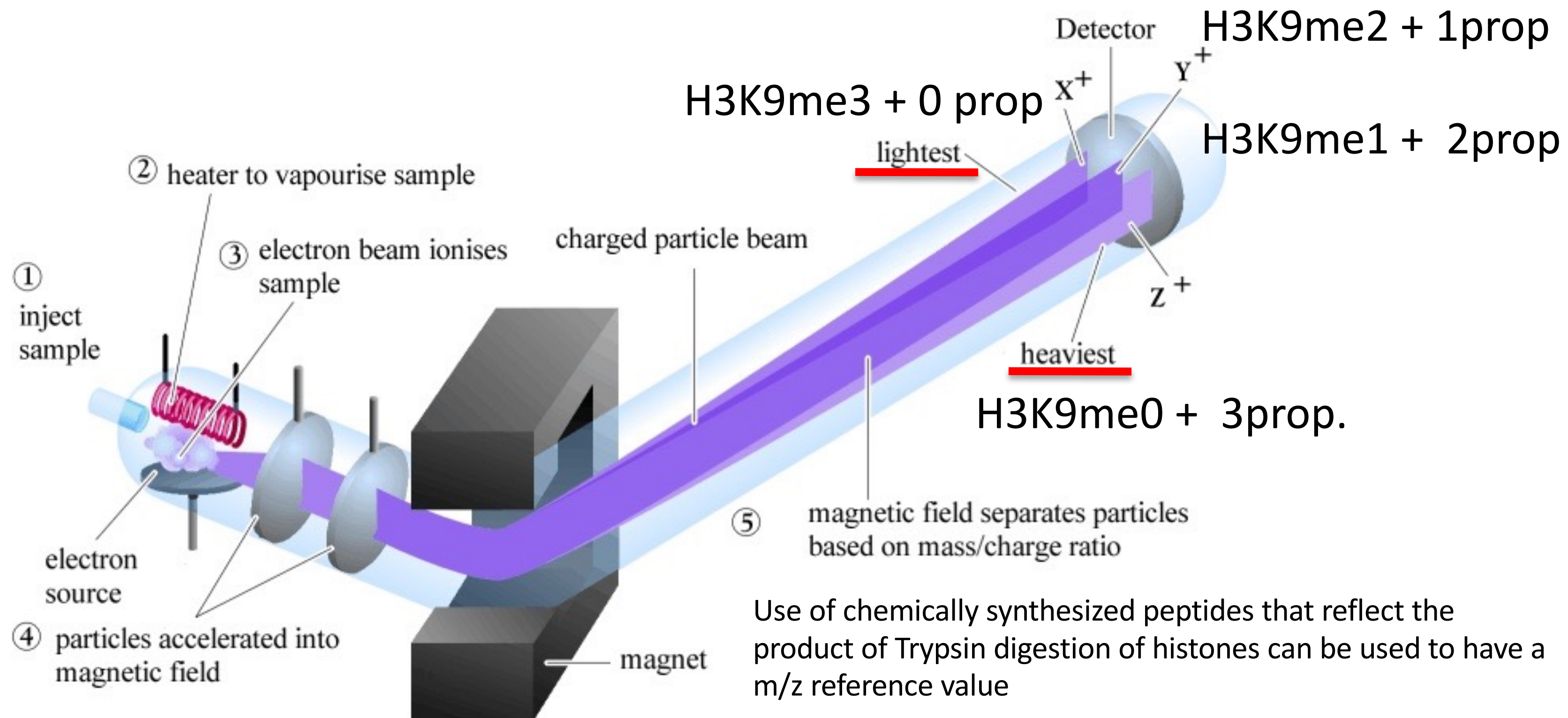
NOTE: The number of methyl-groups/propinyl groups at the H3 peptide decides on the deflection of this specific peptide in mass spectrometry.

Important: All peptides cut at Arg by Trypsin are analyzed contemporarily!!!!

2. Studying histone modifications by mass spectrometry



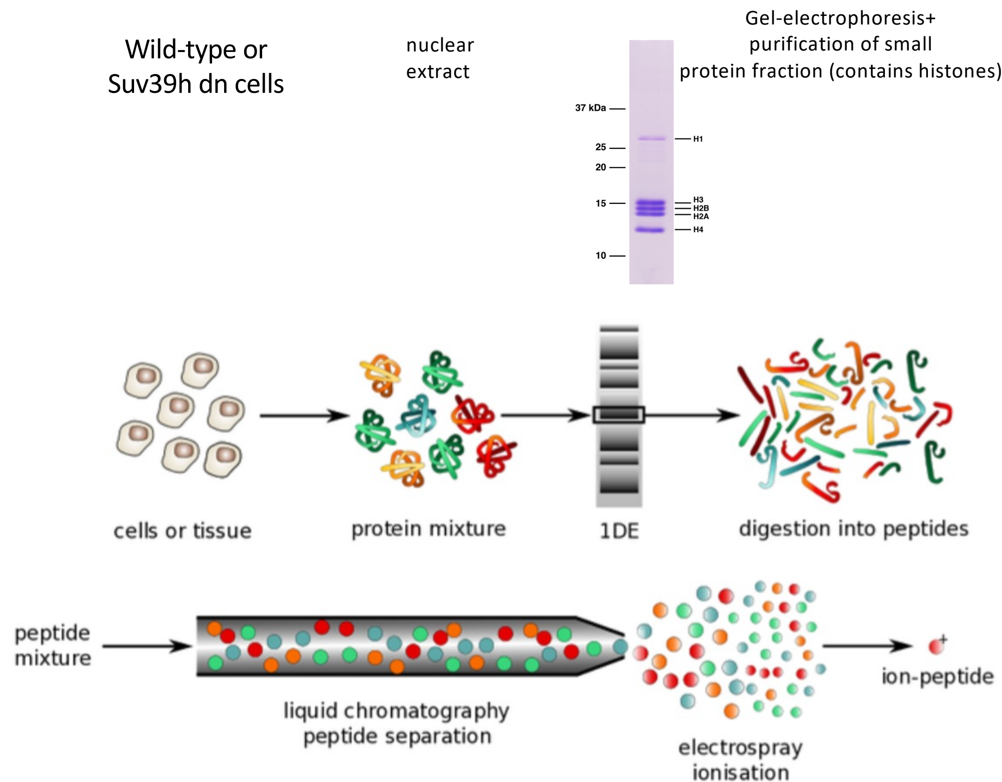
2. Studying histone modifications by mass spectrometry



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

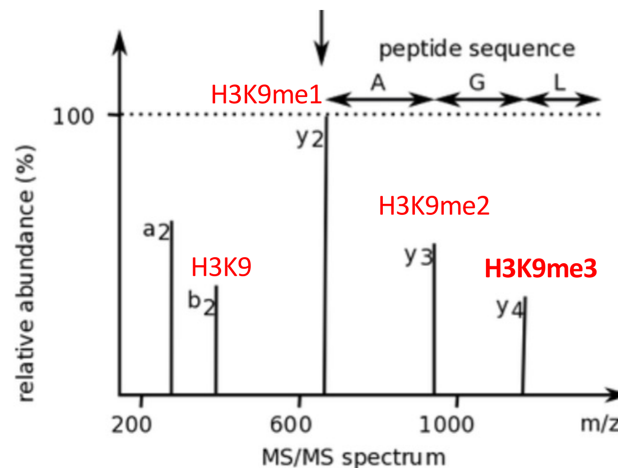
2. Studying histone modifications by mass spectroscopy



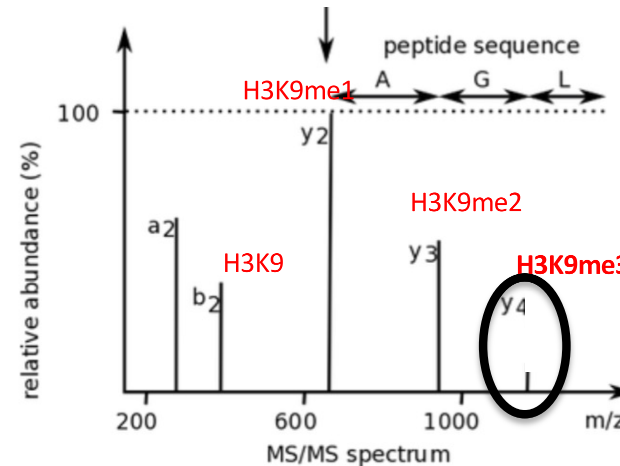
Digestion of fraction containing histones with proteases creates a "library" of small peptides that are derived from histones:
Protease= Trypsin
Trypsin cleaves after every Lysine (K) and Arginine (R)

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

MS

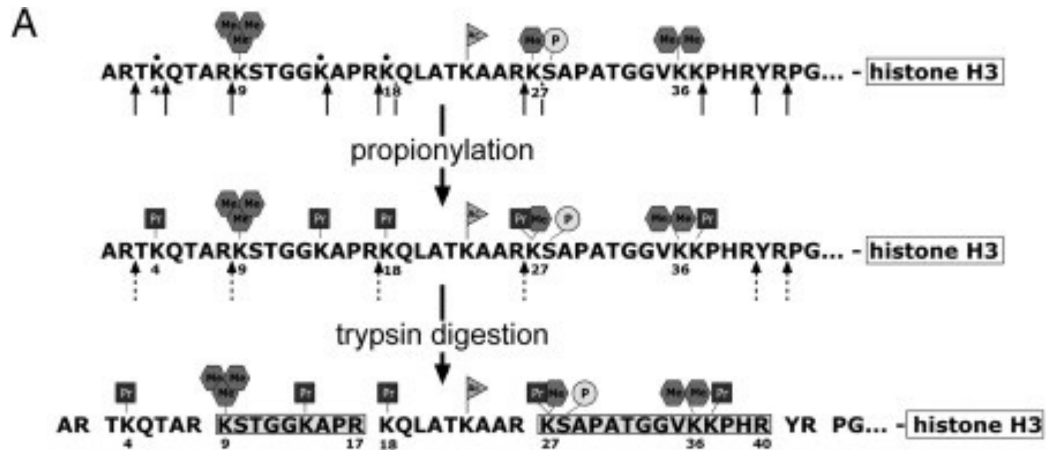


WT profile: Mass/charge ratio

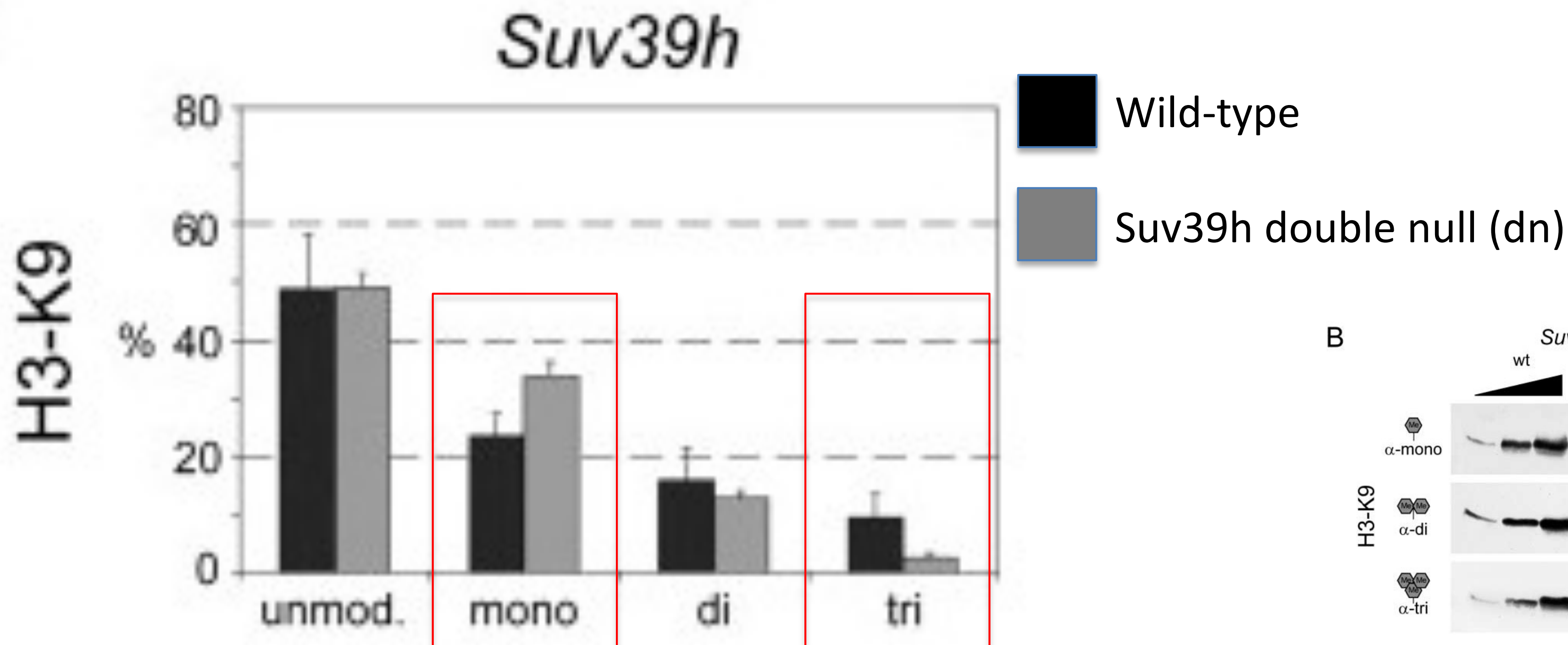


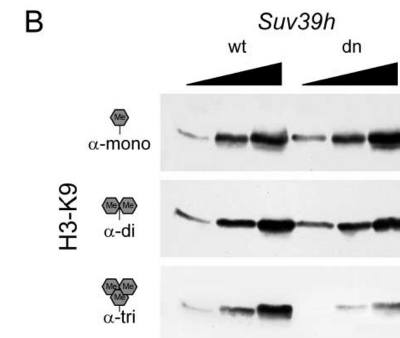
Suv39DN profile: Mass/charge ratio

2. Quantifying histone modifications by mass spectroscopy

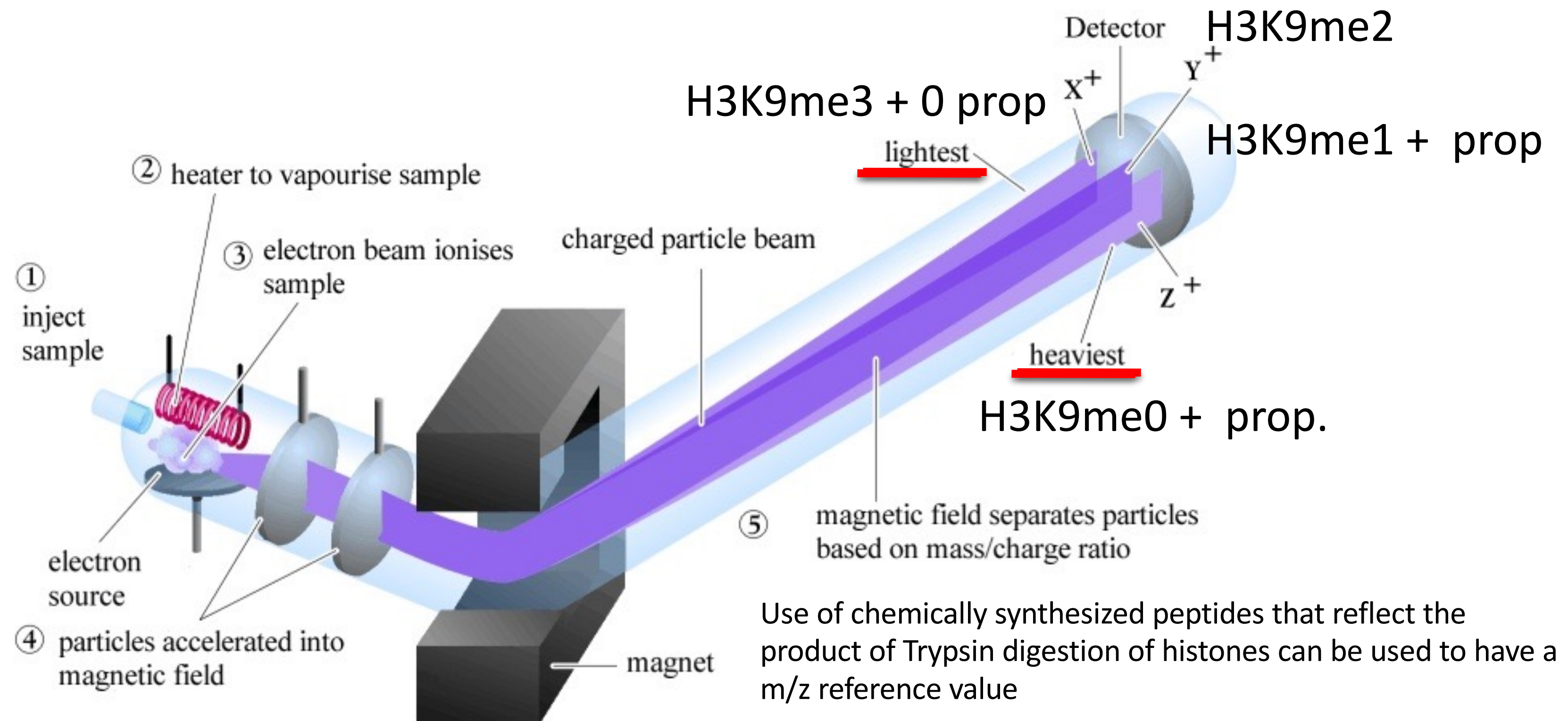


Suv39h1 and Suv39h2 ensure H3K9me3.
H3K9me1 is a preferred substrate for Suv39h1/h2
Suv39h1/h2 methylate H3K9me1
Until reaching the tri-methylated state





2. Studying histone modifications by mass spectroscopy



Use of chemically synthesized peptides that reflect the product of Trypsin digestion of histones can be used to have a m/z reference value

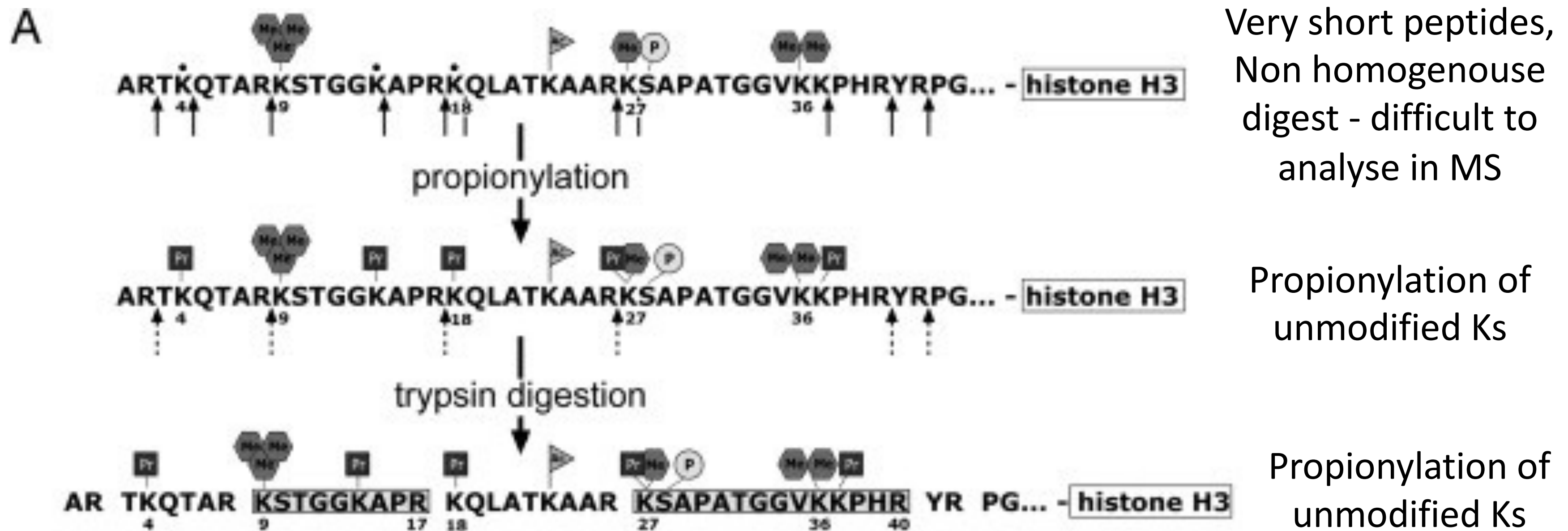
The method allows to quantify the numbers of H3K9me0, H3K9me1 H3K9me2 and H3K9me3 in a sample → We can calculate the % of each histone modification in the given sample.

Important MS is a proteomics approach: The analysis is not limited to H3K9 → all other histone modifications can be quantified in the same analysis. Remember: the histone proteins were isolated prior to MS.

This allows also to discover possible regulatory connections between different histone modifications!!! (for this you always need 2 biologically different samples: genetic model; differentiation, treatment, etc…)

However: Mass spectrometry does not give information on the localization of a particular histone modification along the genome

2. Quantifying histone modifications by mass spectroscopy



Very short peptides, Non homogenouse digest - difficult to analyse in MS

Propionylation of unmodified Ks

Propionylation of unmodified Ks

1. Identifying the enzymatic activity of a histone modifying enzymes (LOF of epigenetic writer)

2. Identification of “modifiable” amminoacids in histone tails

3. Functional link between different types of histone modifications

4. Defining the OVERALL epigenetic status of a cell type (for example: differentiated/stem cell)

5. Quantitative information on the frequency of histone modifications (%)

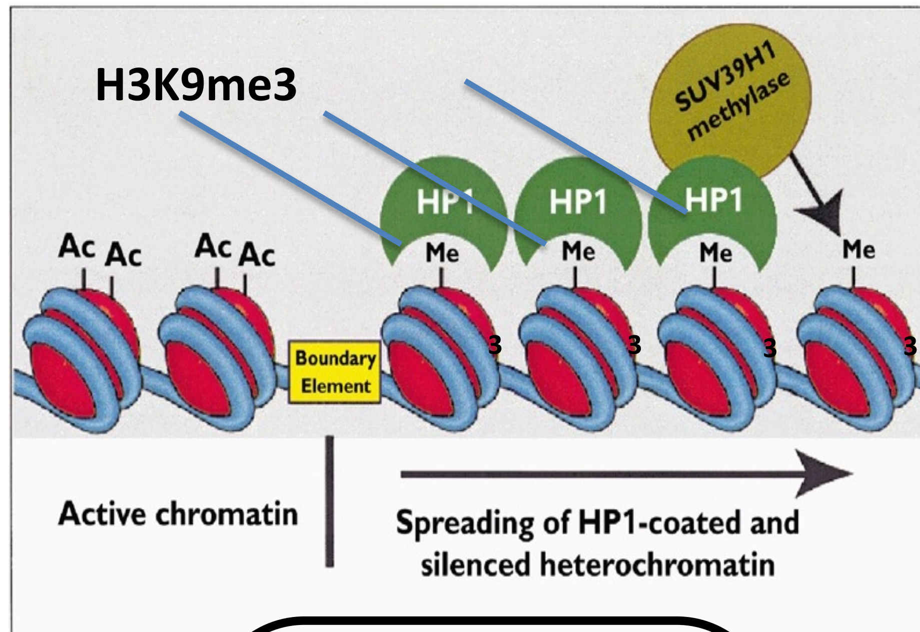
HOW TO STUDY EPIGENETIC MODIFICATIONS

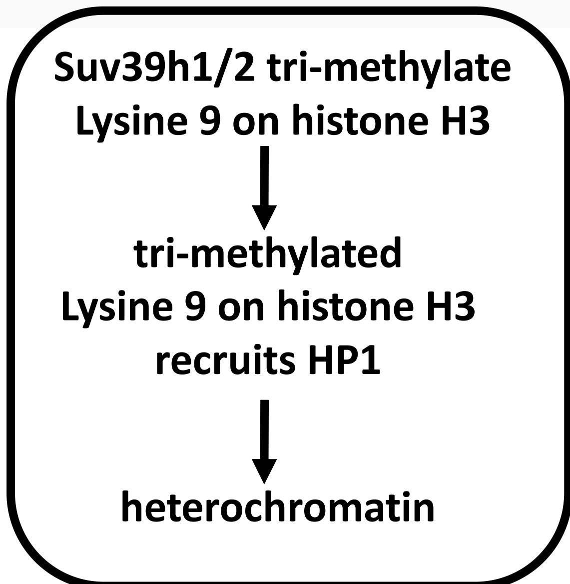
1. **Specfic antibodies are central to epigenetic research?**
2. **Mass spectrometry can read histone codes?**
3. **How can we identify enzymatic activities of writers?**
4. **How can we identify epigenetic readers?**
5. **How can we locate specific histone modifications in the nucleusalong the genome?**
6. **Identification of biological relevance of chromatin modifications**

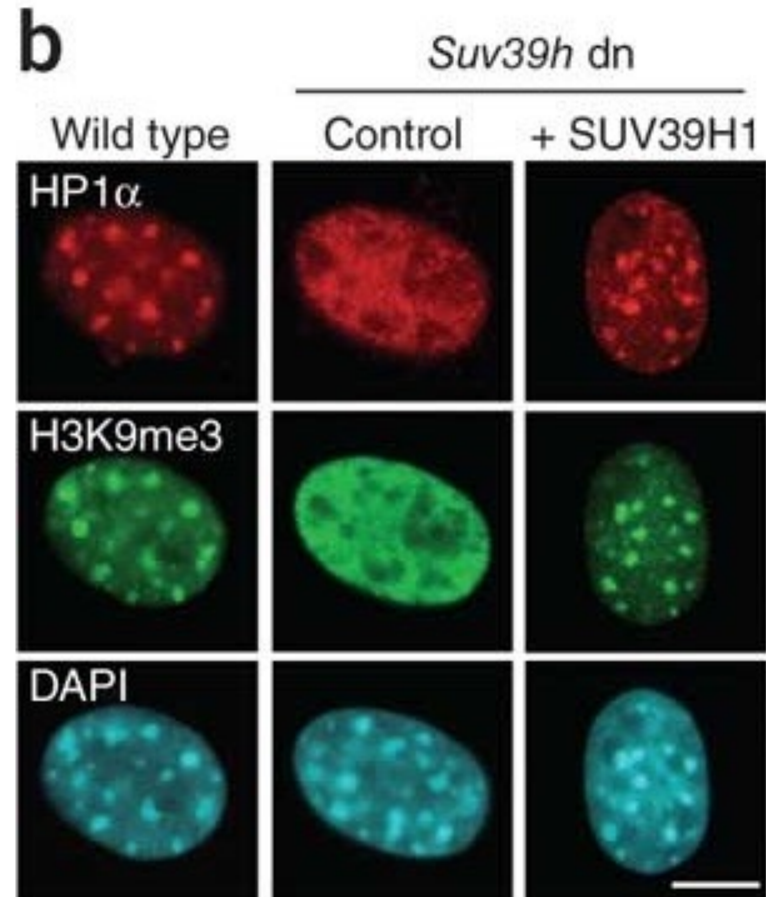
-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

Identification of the enzymatic specificity of SUV39H1

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







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

Identification of the enzymatic specificity of SUV39H1

Candidate gene:
Human
SUV39H1 and SUV39H2

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

Suppressors and enhancers

Wild-type

Drosophila eye (translocated white⁺)

E(var)

Su(var)

Fly with white gene inserted close to centromere (low white expression)

E(var)

Su(var)3-9

Wild-type

A

Su(var)3-9

D

Evidence 2
Suv39h genes
show high sequence
conservation to
a plant gene with
proposed histone-
methyltransferase
activity

chromo

SUV39H1

1

SET

412

249

375

(Hs) SUV39H1
(Mm) Suv39h1
(Mm) Suv39h2
(Dm) Su(var)3-9
(Ce) C15H11.5
(Sp) Clr4
(Hs) EZH2
(Hs) HRX
(Ps) rubisco ls-MT
(At)

Identification of the enzymatic specifcity of SUV39H1

Experiment:

Transfect Hela cells with expression vector for myc-tagged-SUV39H1

Use a commercial anti-myc antibody to immunoprecipitate myc-SUV39H1 → high concentration of SUV39H1

Incubate immunopreciptate with purified histones and S-adenosyl-[methyl-^{14}C]-$_{L}$-methionin as methyl donor

a SUV39H1

SET domain

CHROMO domain

H1, H3, H2A, H2B, H4, 49K, 21K

coomassie

fluorogram

SAM

14

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 (position ??)

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

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

Lachner M[1], O'Carroll D, Rea S, Mechtler K, Jenuwein T.

Identification of H3 Lys9 methyltransferase activity

Experiment: Peptide digestion by Edman sequencing:

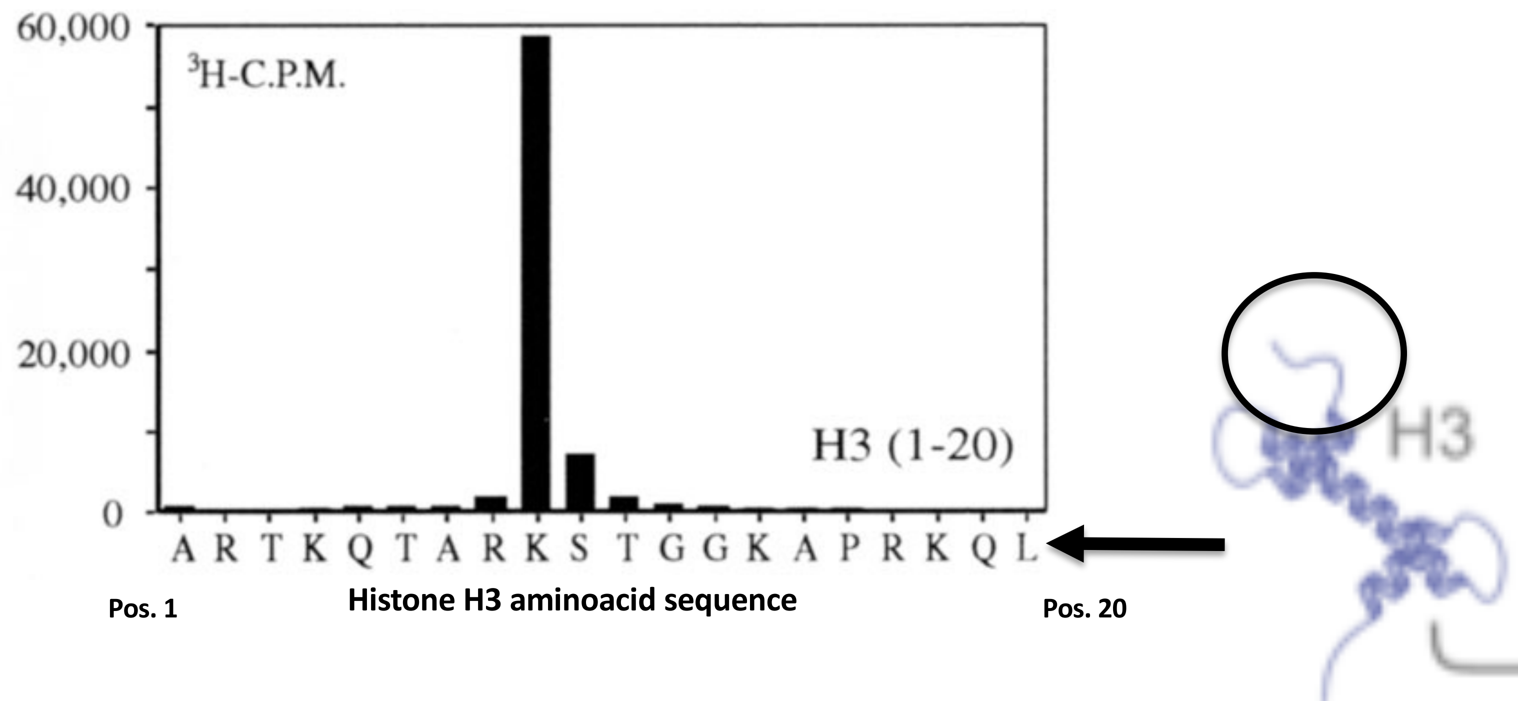
Chemical protein sequencing starting from N-terminal: individual aminoacids are cleaved off step by step

Purify histone H3 by reverse-phase-liquid chromatography after HMTassay

Edman sequencing of histone H3 and quantify the amount of radioactivity per ammino acid

Suv39H1 methylates histone H3 at lysine 9

Alternative: Repeat experiment with non-labelled SAM; cut histone H3 band from gel; purify proteins and perform classic mass spectrometry experiment



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

H3: ARTKQTARKSTGGKAPRK … ARKSA …
4 9 27

Me Me Me

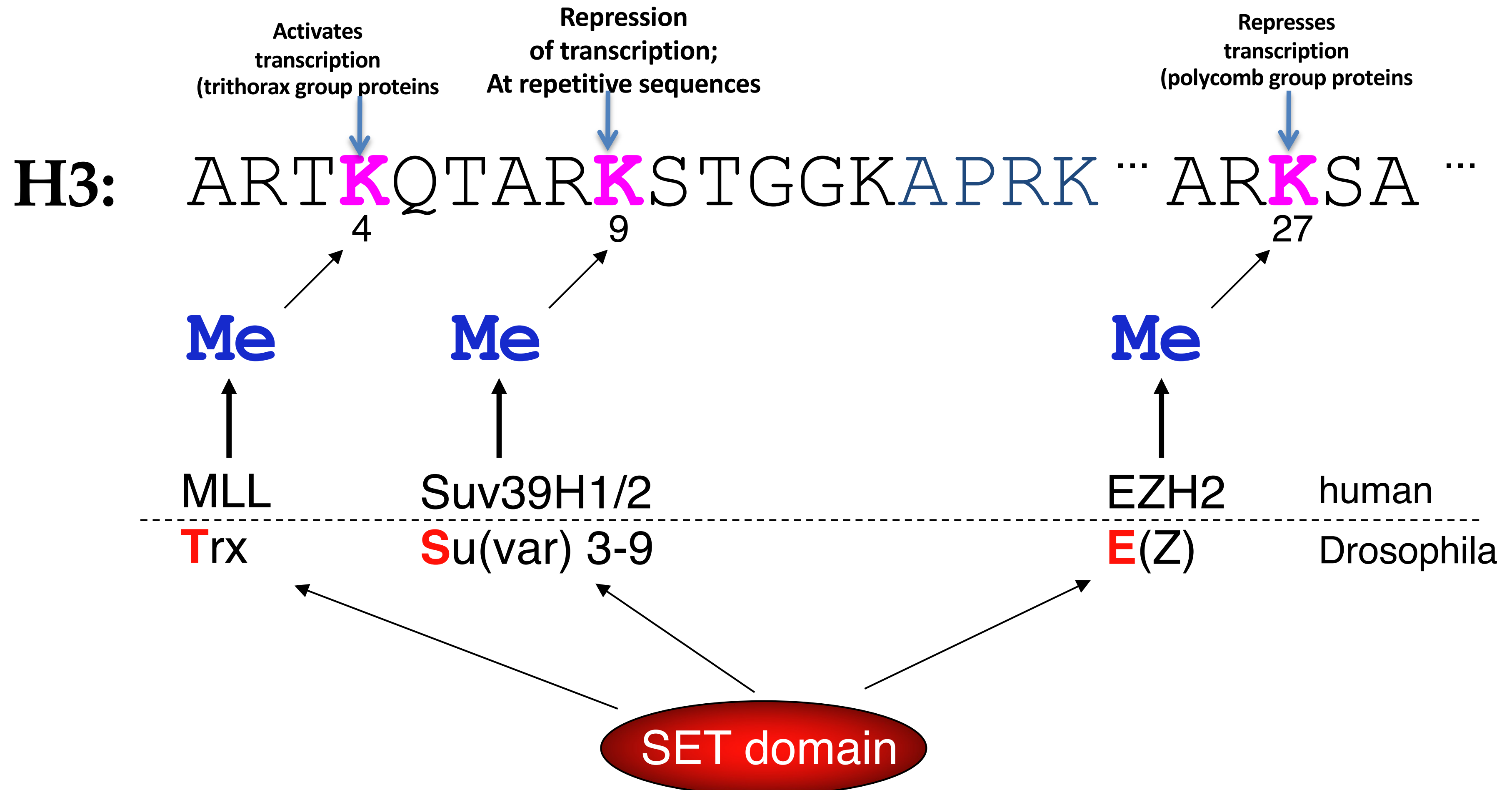
Suv39H1/2 human

Su(var) 3-9 Drosophila

SET domain

Identification of other Histone H3 Lys9 methyltransferases

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



Enzymatic domain is conserved

SET domain protein have sequence specifcity for pepide sequence around target K

- Mutations of some histone methyltransferases → cancer/disease

HOW TO STUDY EPIGENETIC MODIFICATIONS

1. Specfic antibodies are central to epigenetic research?
2. Mass spectrometry can read histone codes?
3. How can we identify enzymatic activites of writers?
4. How can we identify epigenetic readers?
5. How can we locate specific histone modifications in the nucleus along the genome?
6. Identification of biological relevance of chromatin modifications

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

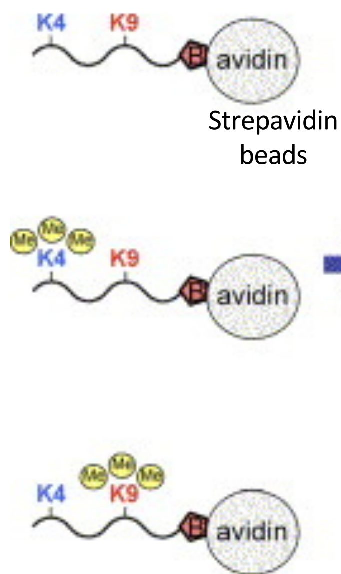
3. How can we identify epigenetic readers?

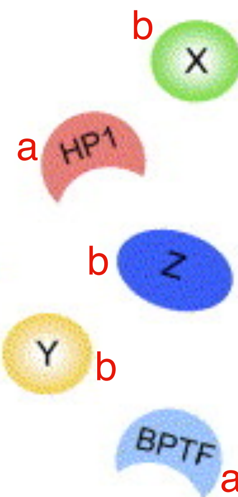
STEP 1: Identifying proteins that bind specifically modified histones

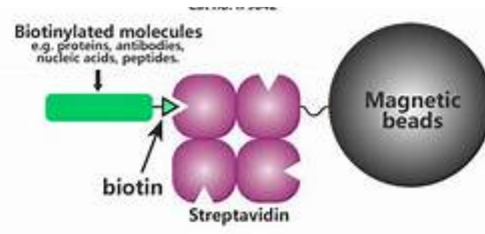
- **histone peptide pulldown assay: protein identification**

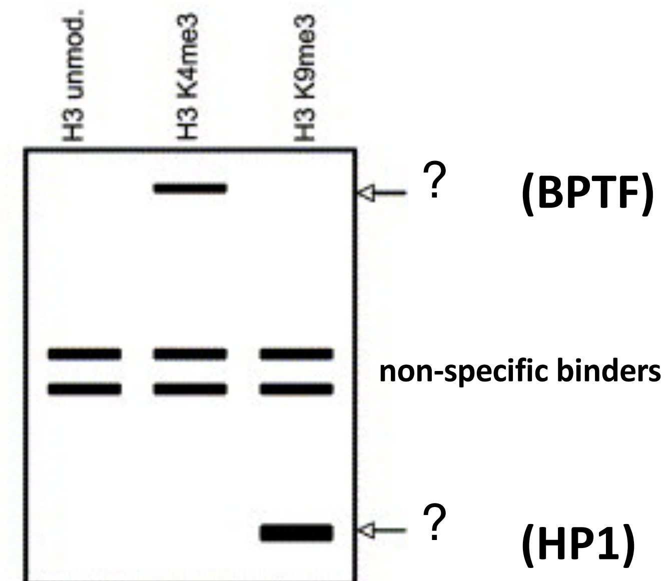
Chemically synthesized peptides, fused to biotin

nuclear extract









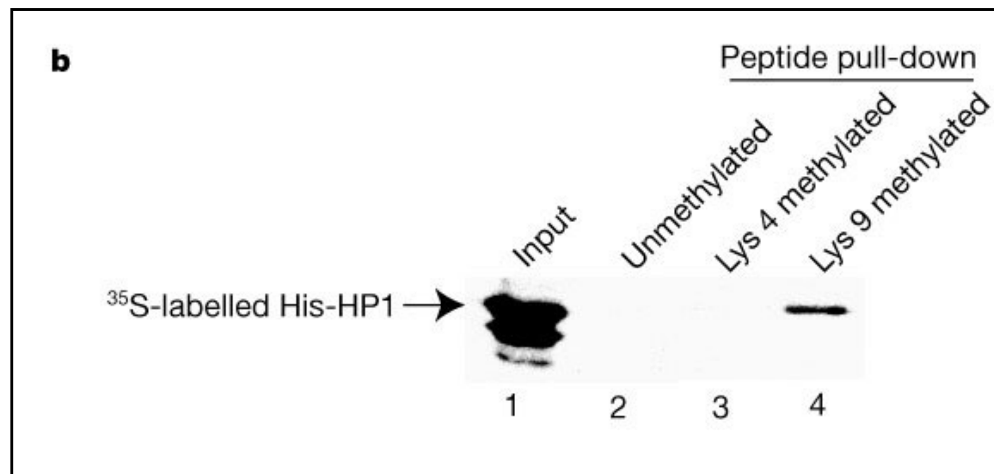
→ stain proteins in gel;
→ cut out specific band,
→ process proteins using protease
→ identify protein by mass spectrometry

3. How can we identify epigenetic readers?

STEP 2: Independent validation of specific interaction of proteins with specifically modified histones

- histone peptide pulldown assay: validation of interaction using recombinant proteins (not extract)

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



Pull-down assay:

-Couple biotinylated histone tail-peptides, carrying specific modifications (methylated, or unmethylated) to stretptavidine coated resin

-Incubate with recombinant ^{35}S-labelled HP1, produced and purified in *E. coli*

-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

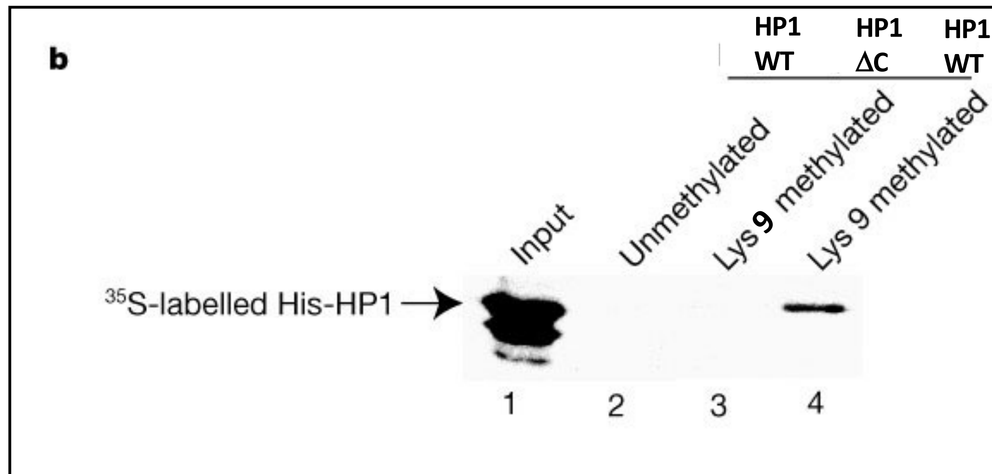
- Lys9-methylated H3 binds to HP1 via the chromodomain motif in HP1

Lys9-methylated H3 binds to HP1

Bannister et al, Nature, 2001

3. How can we identify epigenetic readers?

- histone peptide pulldown assay: validation of interaction and functional analysis

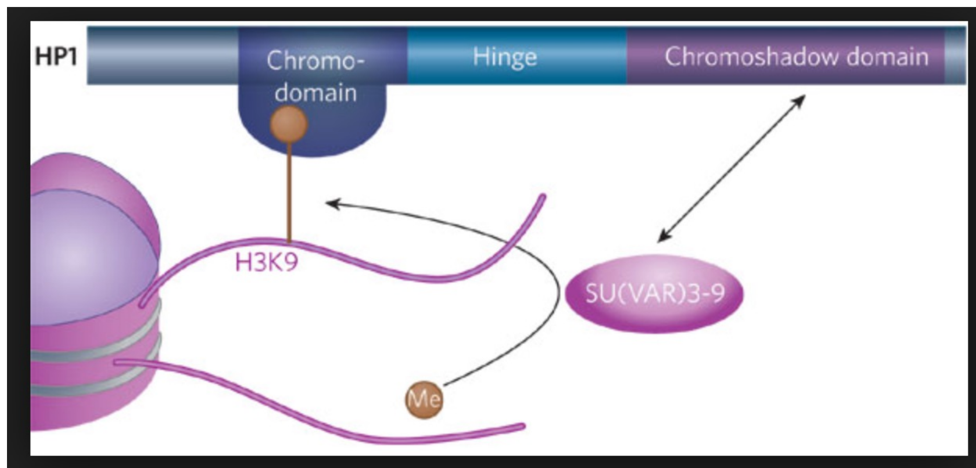


Pull-down assay:
-Couple biotinylated histone tail-peptides, carrying specific modifications (methylated, or unmethylated) to stretptavidine coated resin

-^{35}S-labelled, recombinant, mutant versions of HP1: ΔC (chromo domain deletion);
WT: full length

-Wash resin

-Elute bound proteins, run gel and make radiography



(Delta) ΔC domain (chromodomain) doesn't bind:
= chromo domainbinds modified histone tail

Conclusion: Chromo domain binds methylated histones

General relevance: also other proteins contain chromo domains
→ Candidate epigenetic readers

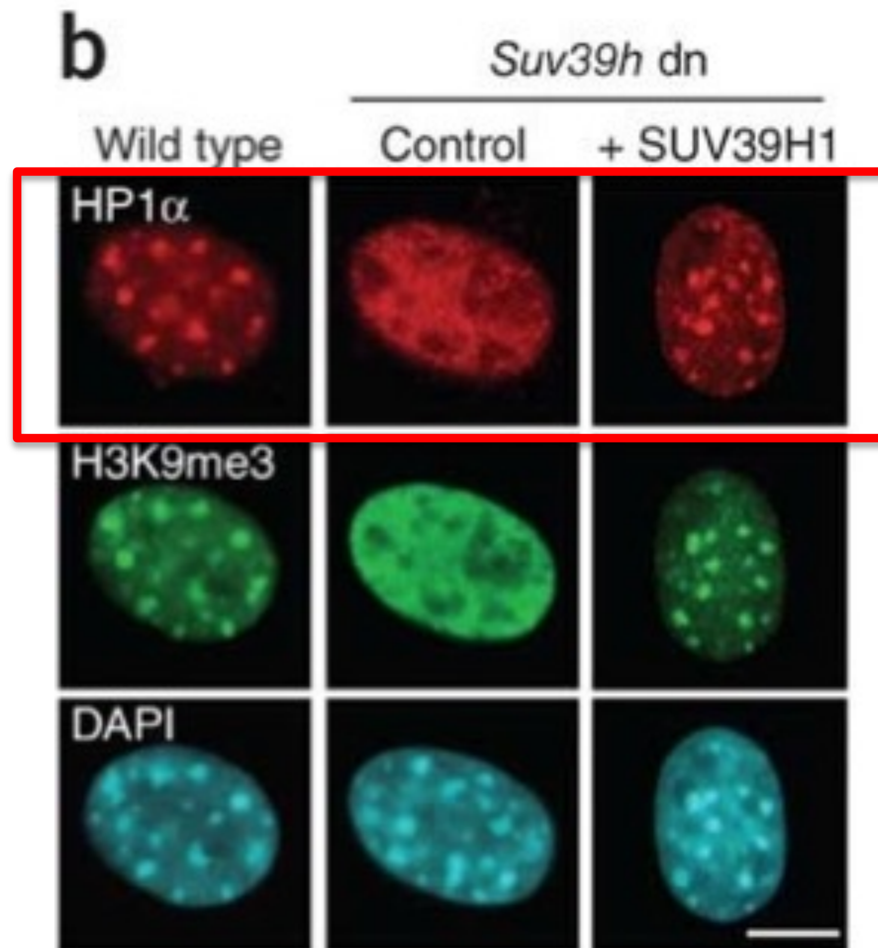
Lys9-methylated H3 binds to the conserved motif called chromodomain

Bannister et al, Nature, 2001

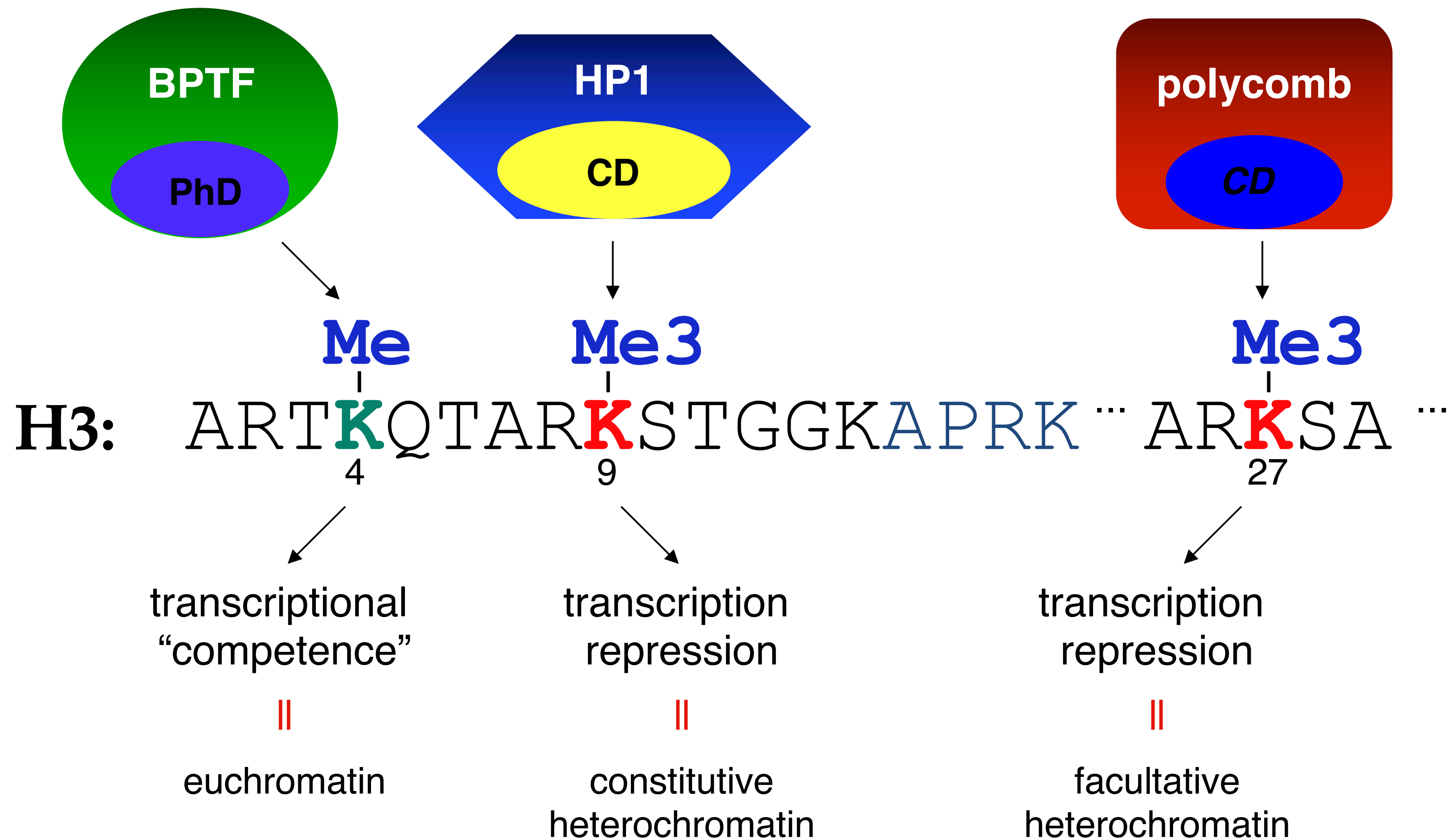
3. How can we identify epigenetic readers?

Generate and antibody against the epigenetic readed

Immunofluorescence to detect colocalization



Site specific methylation of the H3 tail has different functions



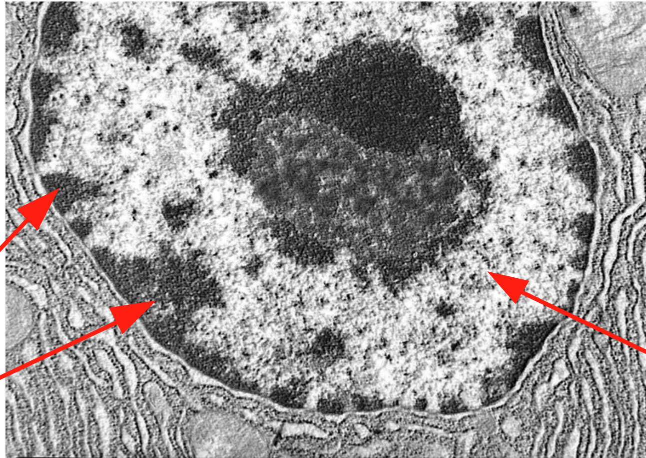
HOW TO STUDY EPIGENETIC MODIFICATIONS

1. Specfic antibodies are central to epigenetic research
2. Mass spectrometry can read histone codes?
3. How can we identify enzymatic activites of writers?
4. How can we identify epigenetic readers?
5. How can we locate specific histone modifications in the nucleus/along the genome?
6. Identification of biological relevance of chromatin modifications

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

Chromatin comes in different flavors

Different types of chromatin



Chromocenter
(aggregates of centromeres
= constitutive heterochromatin)

heterochromatin

euchromatin
(and facultative
heterochromatin)

Constitutive heterochromatin:

- constitute up to 20-30% of nuclear DNA; telomeres, centromeres, and a considerable fraction of repetitive sequences in vertebrates
- highly compacted, replicates late in S phase, (transcriptionally inert; not prone to transcriptio)
- frequently accumulates in nuclear periphery

Euchromatin:

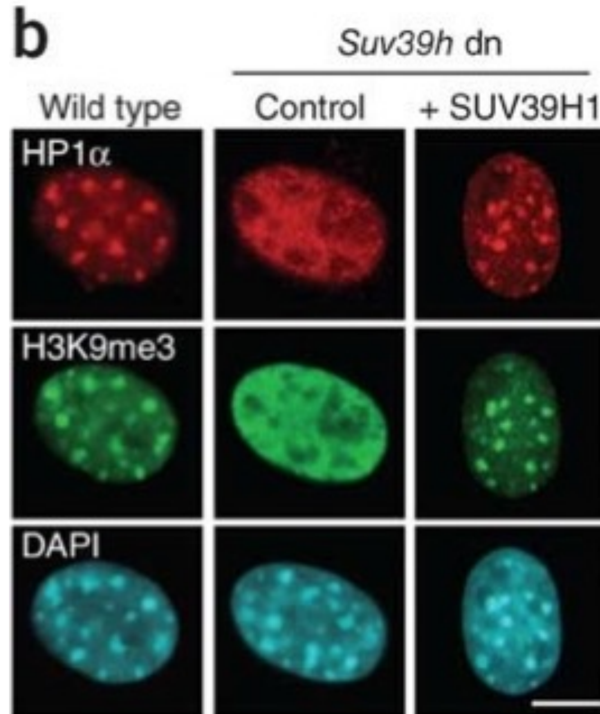
- low condensation state, rich in genes, transcribed replicates early in S phase (variable 30-40%)

Facultative heterochromatin:

- Transcriptionally inactive/silenced but can be activated during development or physiologicallt relevant processes
- Amount is variable
- Most famous example: inactive X chromosome

4. Where do histone modification locate in the cell nucleus?

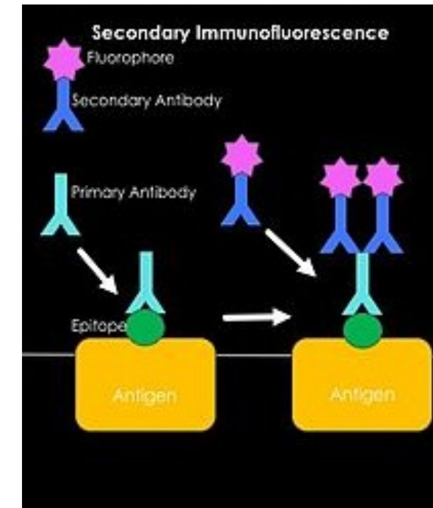
4A. IMMUNOFLUORESCENCE → MACROSCOPIC ANALYSIS
→ Localization of protein across large regions of DNA



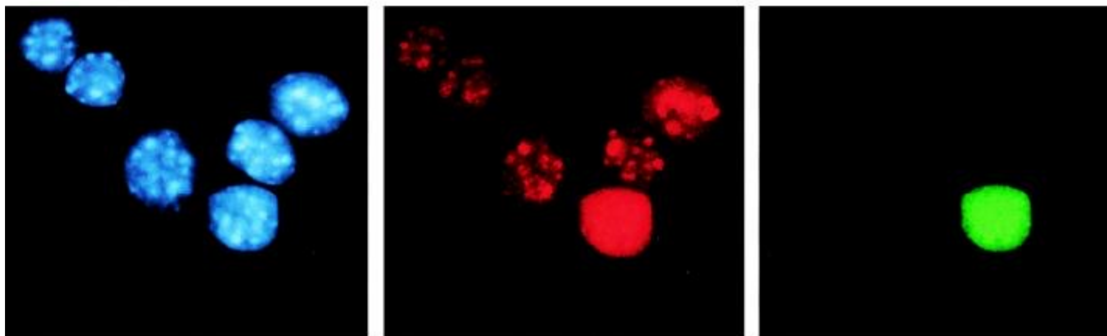
Loss of Suv39h1/h2 causes delocalization fo HP1 from chromocenters (CONSITUTIVE HETEROCHROMATIN)

Chromocenters are subnuclear regions where Multiple centromeres aggregate

Chromocenters are DAPI intense regions and represent tightly packed heterochromatin



Co-immunofluorescene (exicatiation with UV light):
Primary, mouse anti-HP1 + secondary rabbit anti-mouse-Cy3 (emission at 570 nm)
Primary, human anti-H3K9me3 + secondary donkey anti-human Alexa 488 (emission at 488 nm)
DAPI (emission 461 nm)

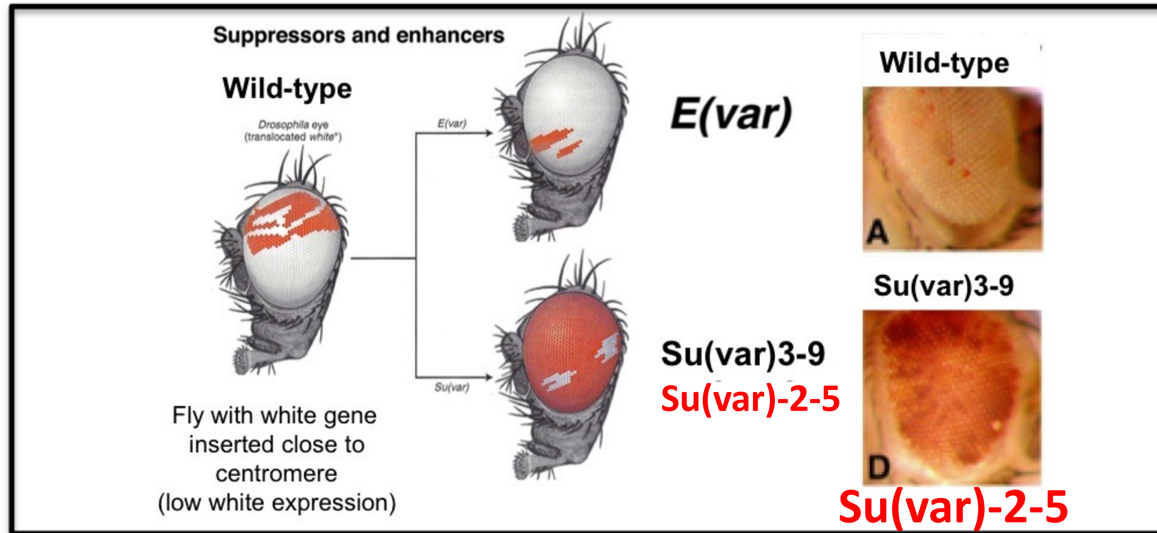


Over-expression of SUV39H1 in Hela cells causes an accumulation of HP1. Stabilization?
Is there a link between Suv39h1 – H3K9me3 and HP1??

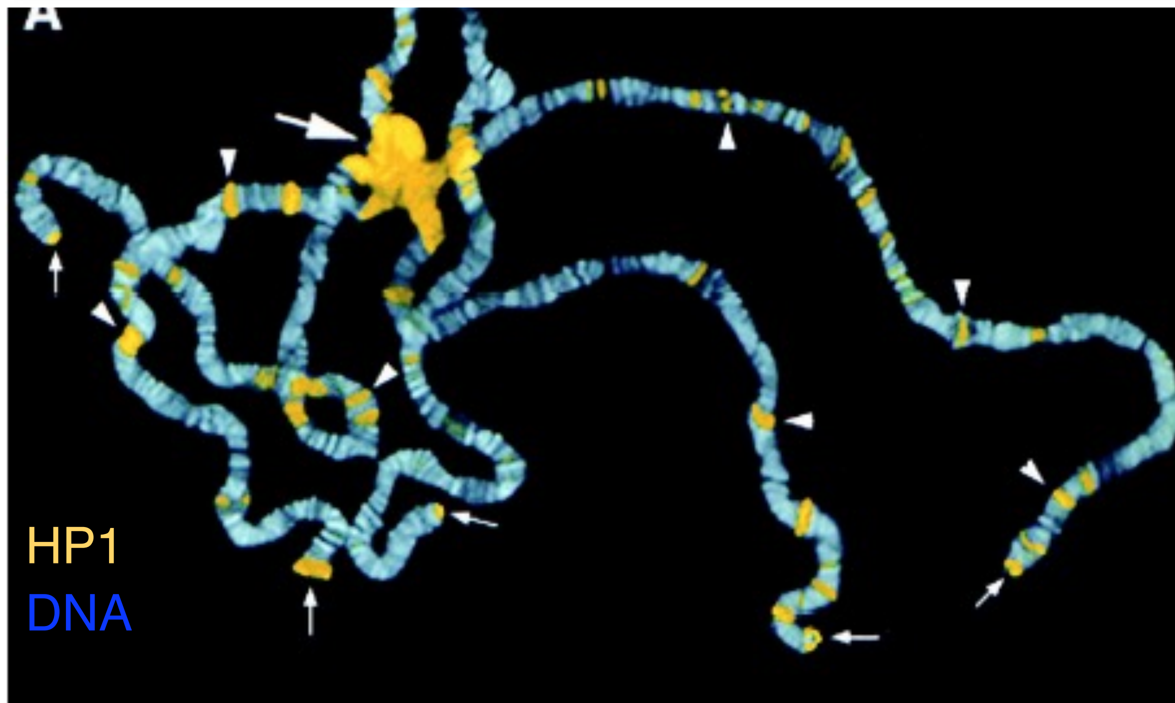
myc-SUV39H1: myc-epitope tagged Suv39H1

Melcher et al, MCB, 2000

4. Where do histone modification locate in the cell nucleus/chromosome?

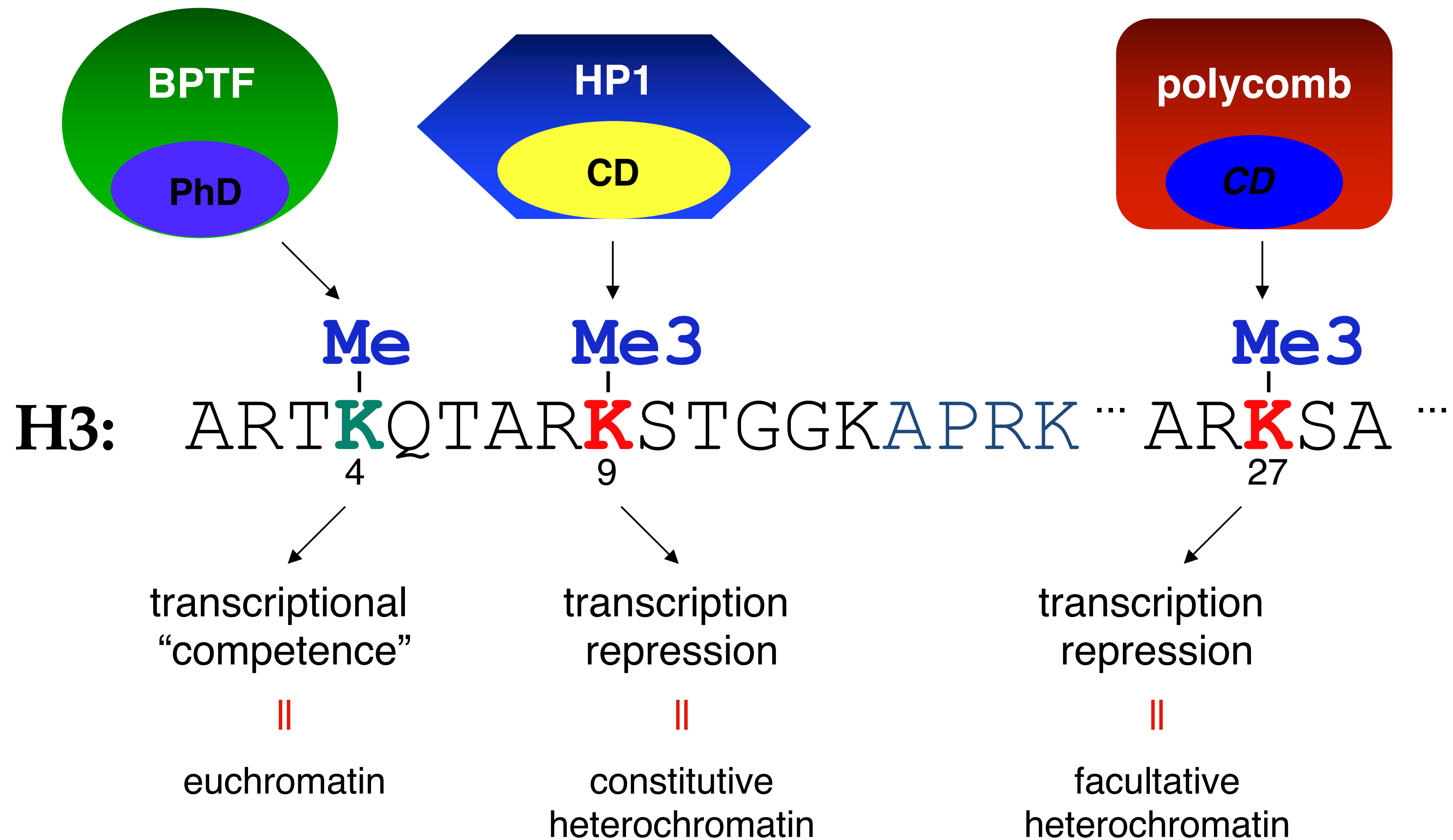


Another evidence for
HP1 and Suv39h interaction
came from Drosophlla



- 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

Site specific methylation of the H3 tail has different functions



4. How to allocate histone modifications and epigentic readers along the DNA?

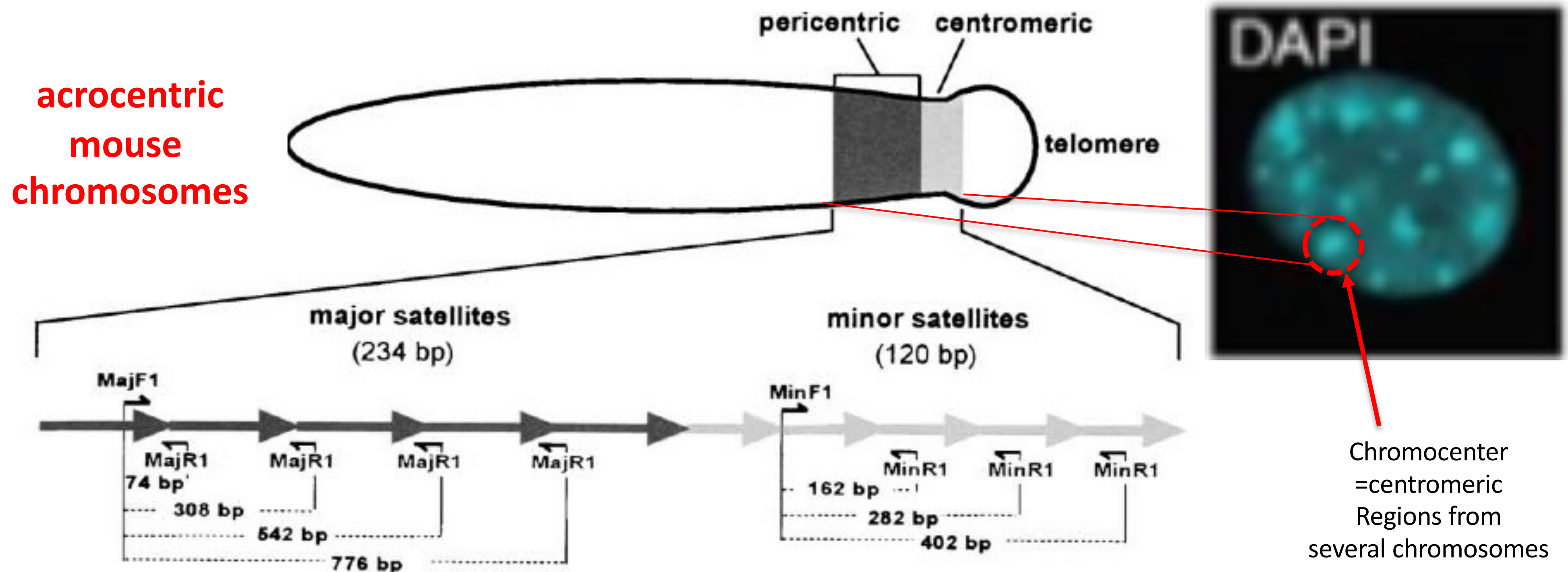
4B. CHROMATIN IMMUNOPRECIPITATION → DETAILLED, SEQUENCE SPECIFIC ANALYSIS

The combination of Immunoprecipitation methods and PCR/Seqeuncing analysis allows to define the histone code at defined sequences

Strategy 1: Target site specific PCR primers

Example: H3K9me3

acrocentric mouse chromosomes



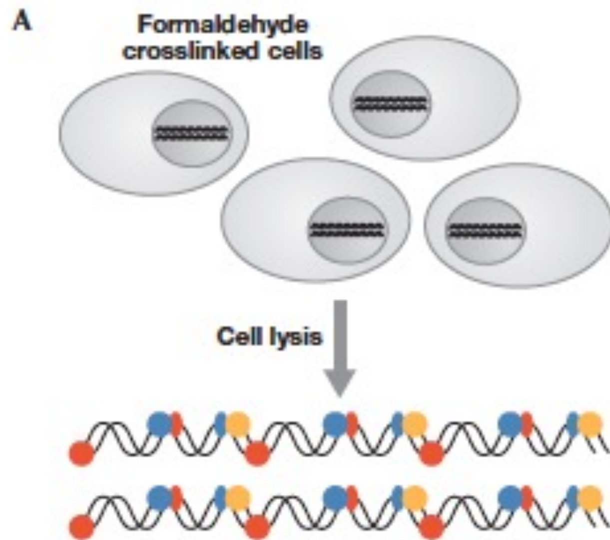
Chromocenter =centromeric Regions from several chromosomes

EXAMPLE: Pericentric heterochromatin in mouse cells

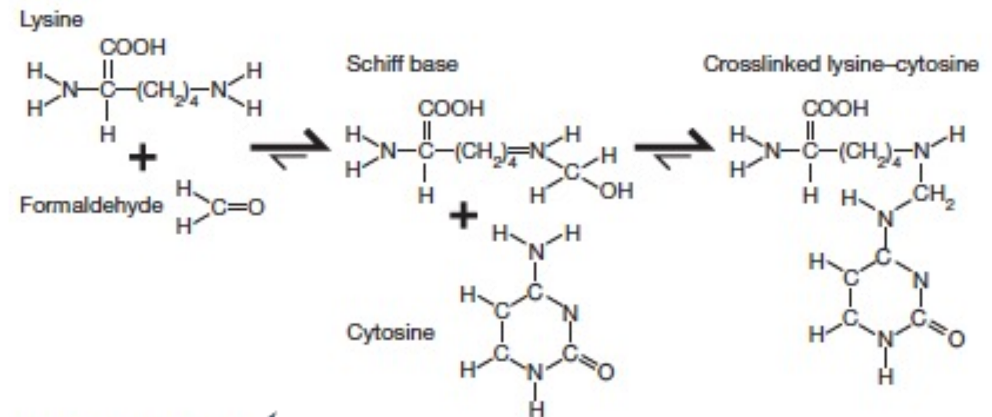
4. How to allocate histone modifications and epigentic readers along the DNA?

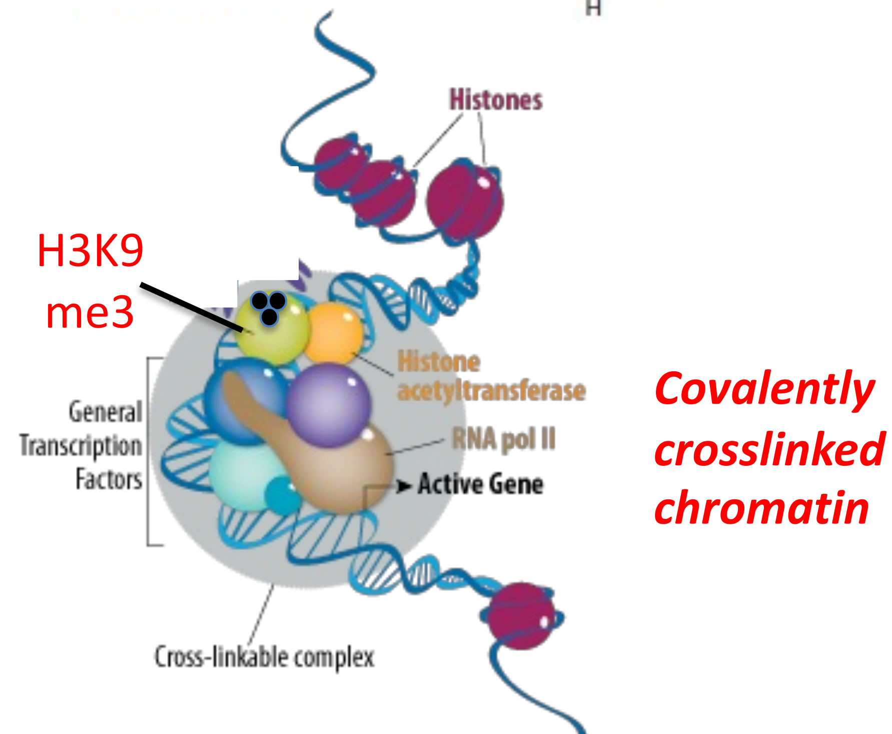
CHROMATIN IMMUNOPRECIPITATION

1. Cross linking with FA



Optimization is crucial.



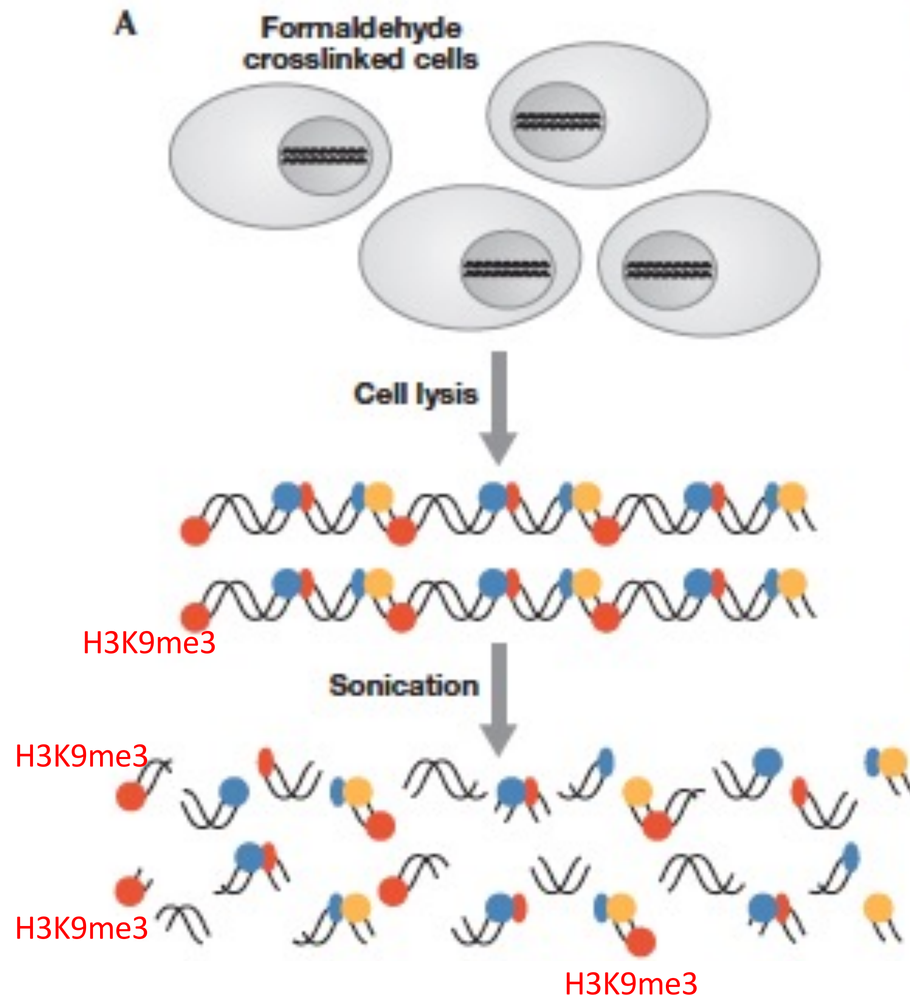


Covalently crosslinked chromatin

4. How to allocate histone modifications and epigentic readers along the DNA?

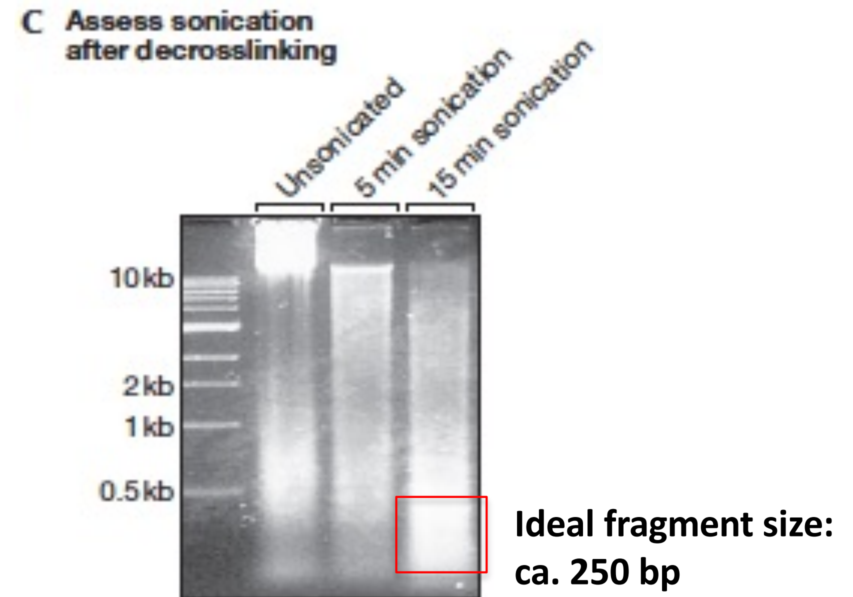
CHROMATIN IMMUNOPRECIPITATION

2. Cross linking followed by sonication (fragmentation of chromatin) – preparation of input material



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

INPUT material



Ideal fragment size: ca. 250 bp

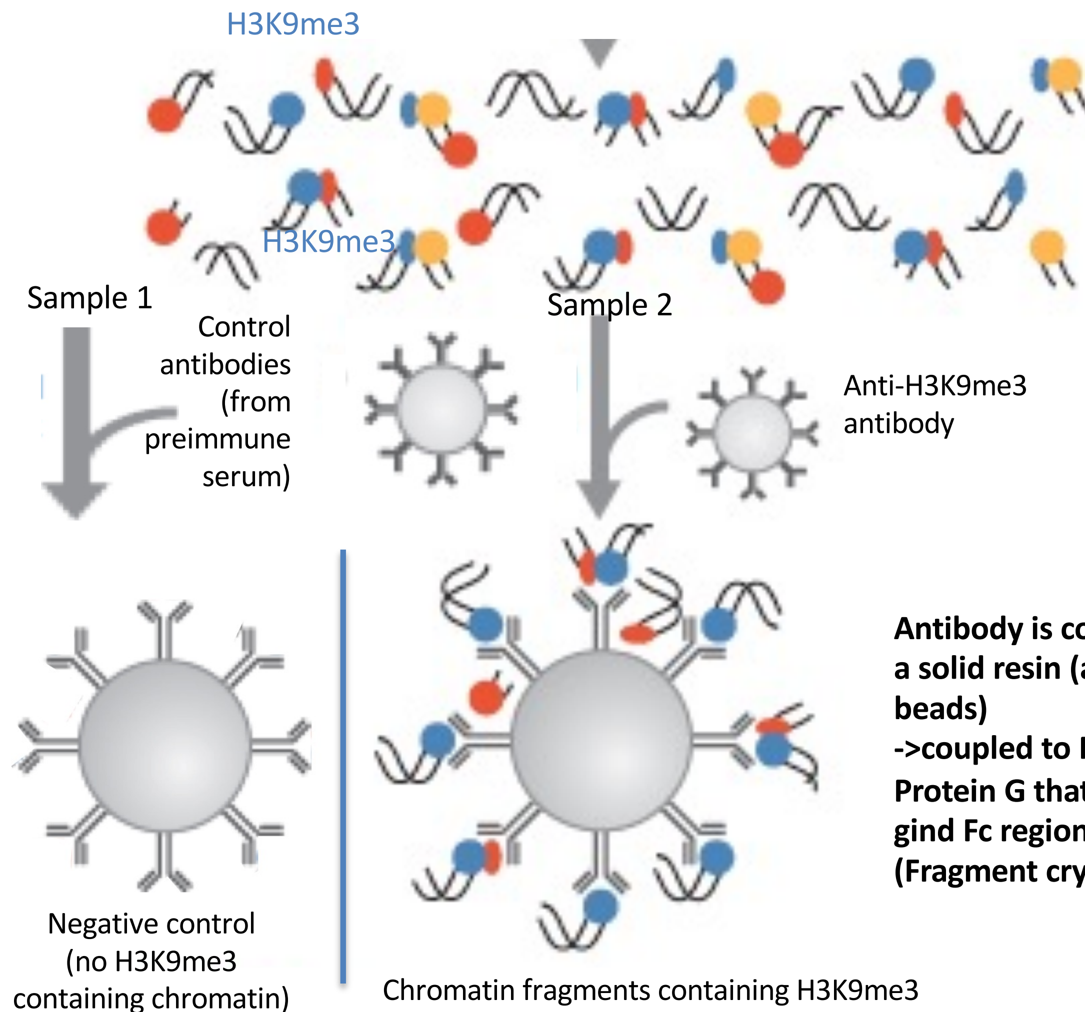
The INPUT material is used to perform immunoprecipitation experiments

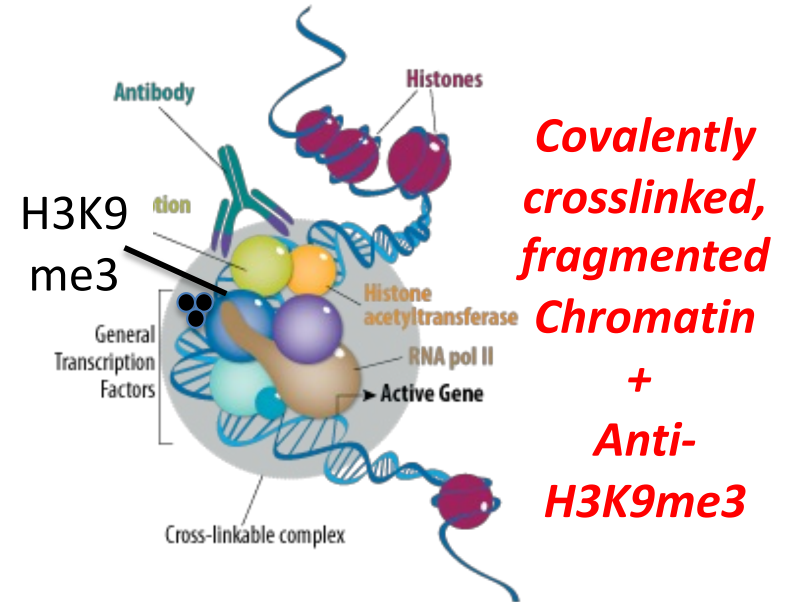
4. How to allocate histone modifications and epigentic readers along the DNA?

CHROMATIN IMMUNOPRECIPITATION

3. Immunoprecipitation (IP)

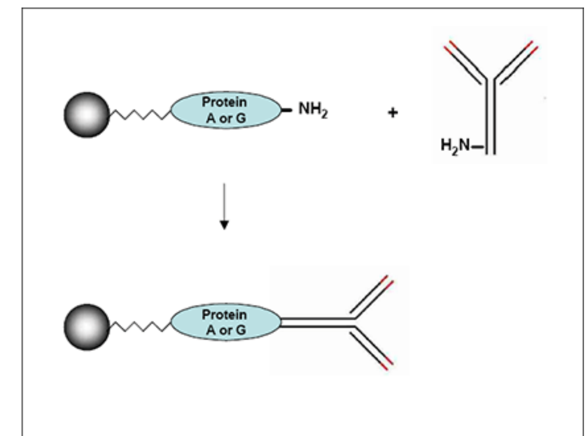
The protein of interest is immunoprecipitated together with the crosslinked DNA: Modified histones; epignetic writers, epigentic readers





Covalently crosslinked, fragmented Chromatin
+
Anti-H3K9me3

Antibody is coupled to a solid resin (agarose, magentic beads)
->coupled to Protein A or Protein G that gind Fc region of antibodies (Fragment crystallizable region)

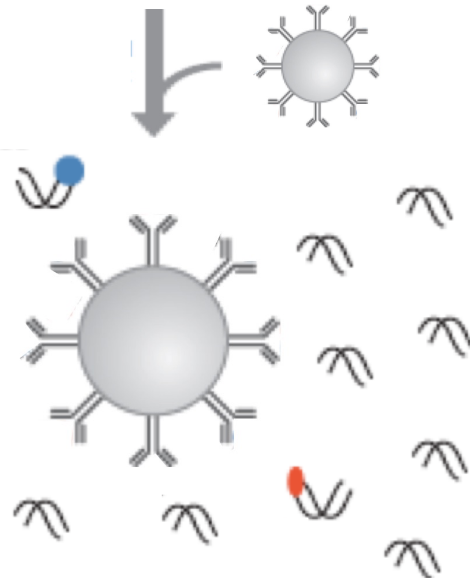


4. How to allocate histone modifications and epigentic readers along the DNA?

CHROMATIN IMMUNOPRECIPITATION

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

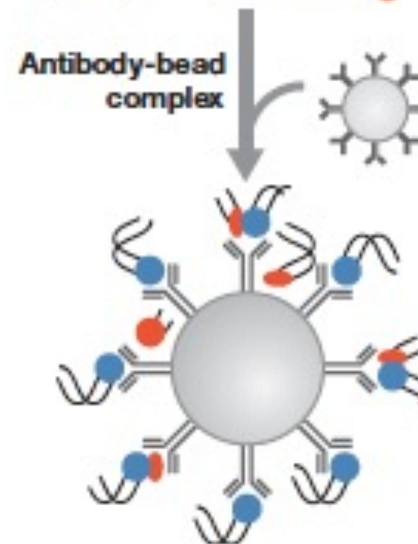
Control antibodies
(from preimmune serum)



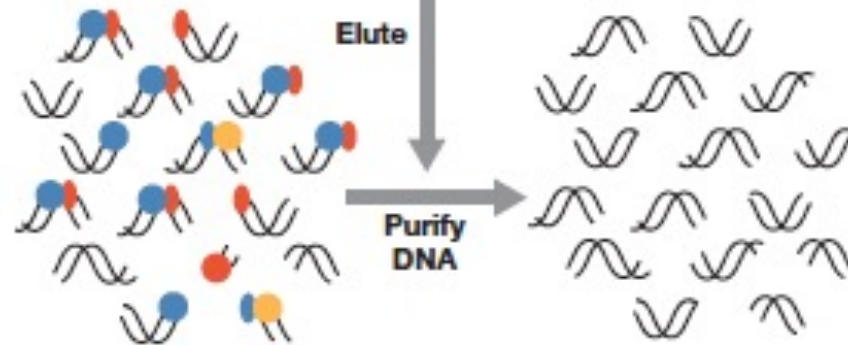
No enrichment
(only non specific binidng)

Unspecific binding (no enrichment of chromatin with histone mdìodifications)

Anti-H3K9me3 antibody



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



Specific binding (enrichment of chromatin with histone mdìodifications

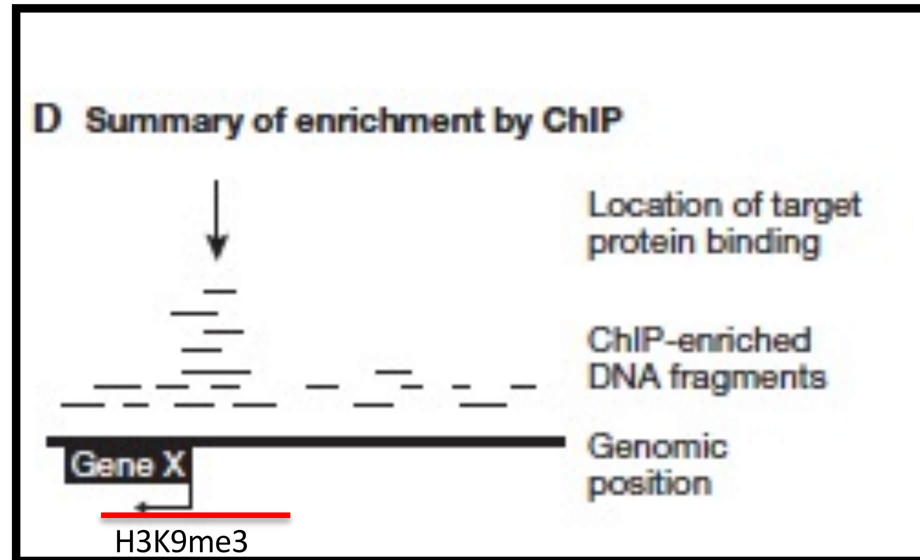
ANALYZE DNA QUANTIFY THE ENRICHMENT OF CHROMATIN CONTAINING H3K9me3
(NGS sequencing, hybridzation, quantitative PCR with target site specific primers)

4. How to allocate histone modifications and epigentic readers along the DNA?

CHROMATIN IMMUNOPRECIPITATION

5. Analysis of ChIP DNA

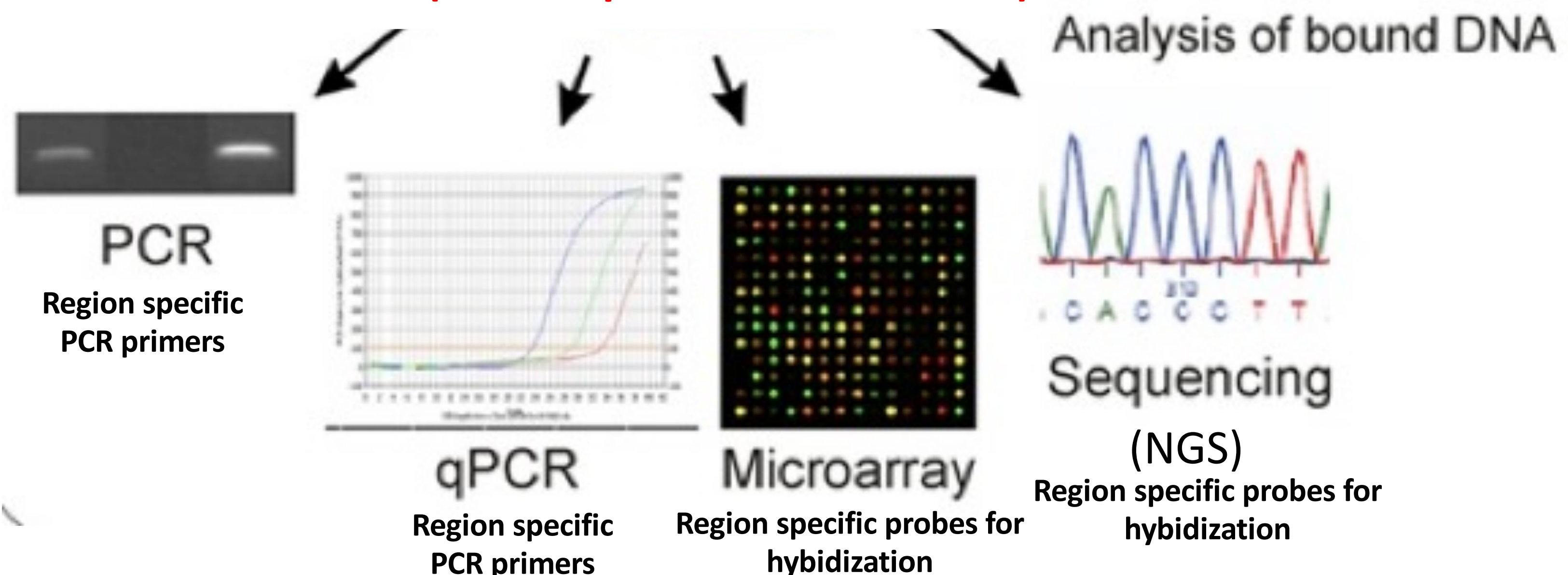
Identification of DNA regions associated with the protein/modification of interest



Purify DNA from chromatin enriched in H3K9me

→ DNA seqeunce give information on localization of histone modification in the genome
→ Direct identification by hybridization or sequecning
→ Direct identification by PCR primer selection

Methods of ChiP analysis (DNA sequence identification)



(NGS)

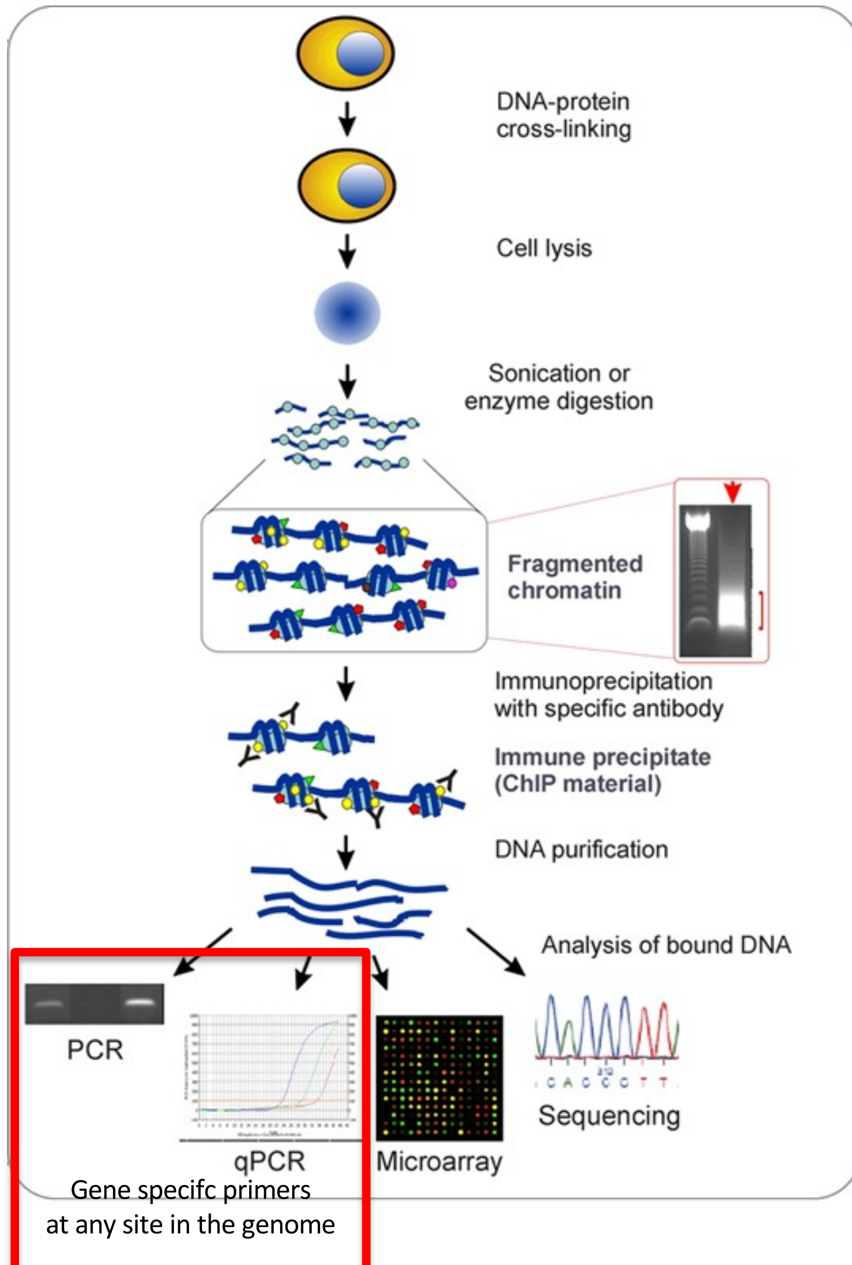
Region specific PCR primers (PCR)

Region specific PCR primers (qPCR)

Region specific probes for hybidization (Microarray)

Region specific probes for hybidization (Sequencing)

CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILLED ANALYSIS
→ Localization of protein at a defined region - seqeunce



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:

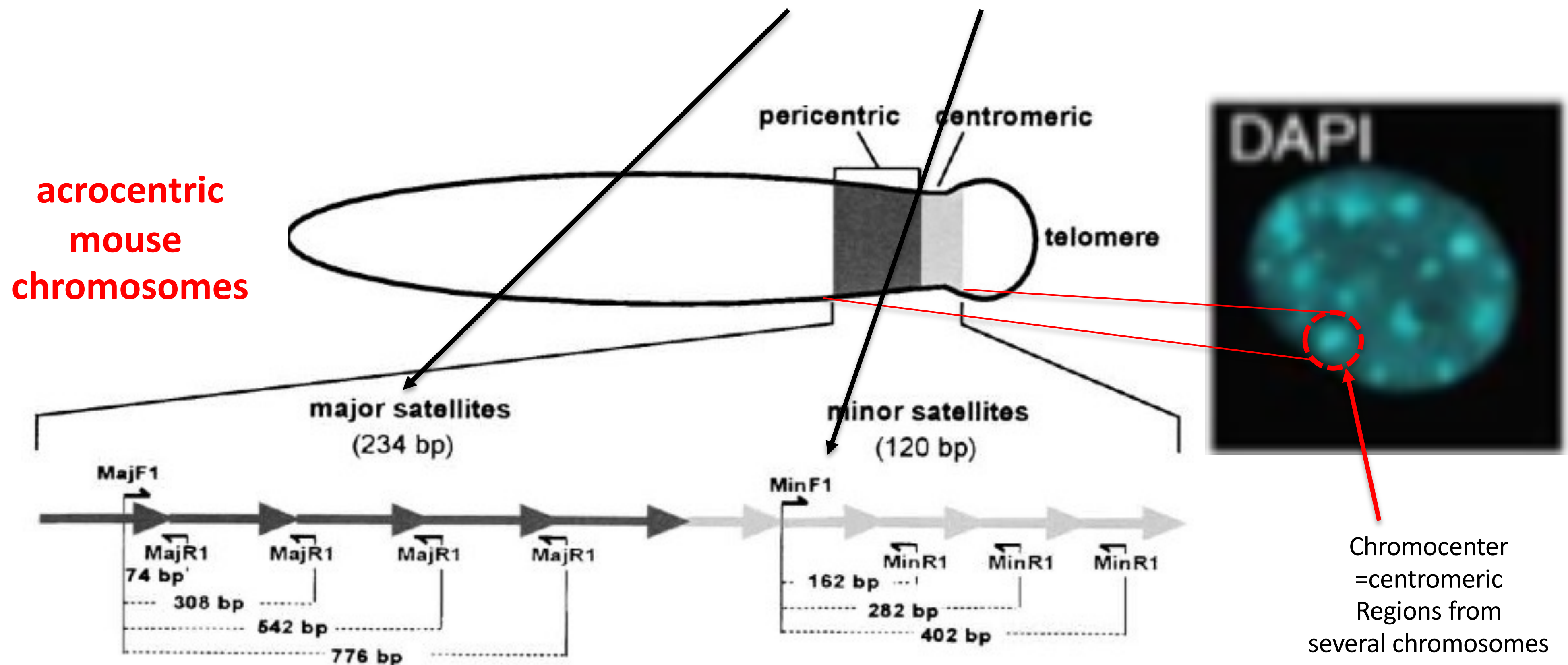
Quantitative PCR: design PCR oligos to amplify defined sequences in immunoprecipitates. Per PCR reaction only one Locus can be examined by real-time PCR

4. How to allocate histone modifications and epigentic readers along the DNA?

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

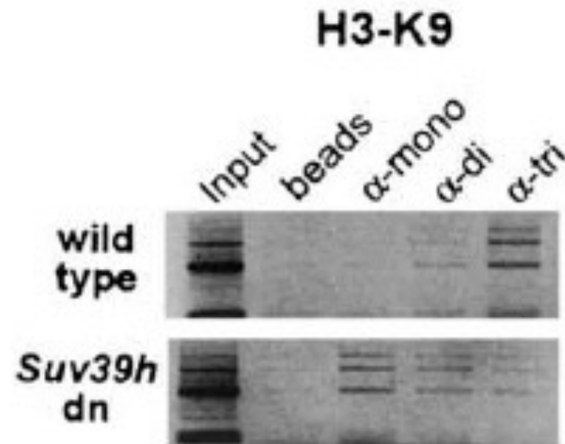
Min/Maj F1…: forward primer in unique region

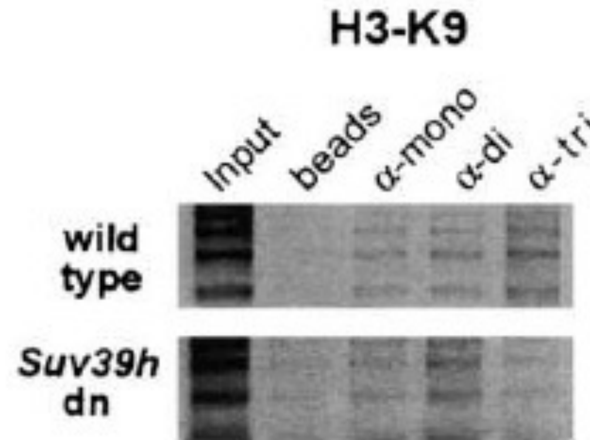
4. How to allocate histone modifications and epigentic readers along the DNA?

CHROMATIN IMMUNOPRECIPITATION (ChIP) COUPLED WITH PCR

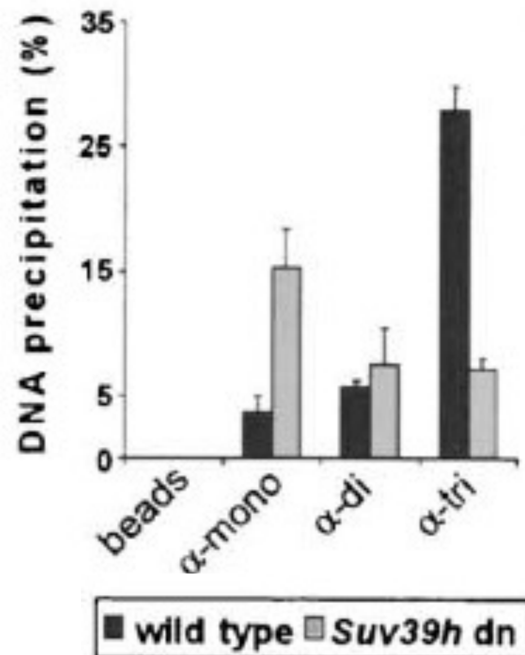
→ H3K9me3 is enriched at pericentric (major+minor) repeats in mouse cells

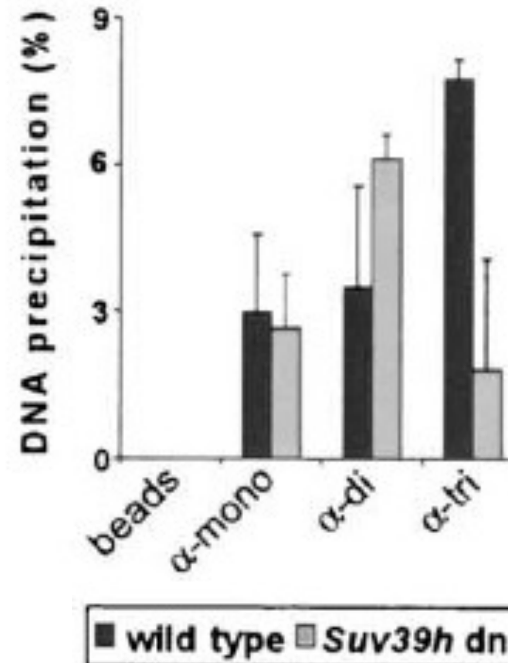
Major satellite repeats **Minor satellite repeats**





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





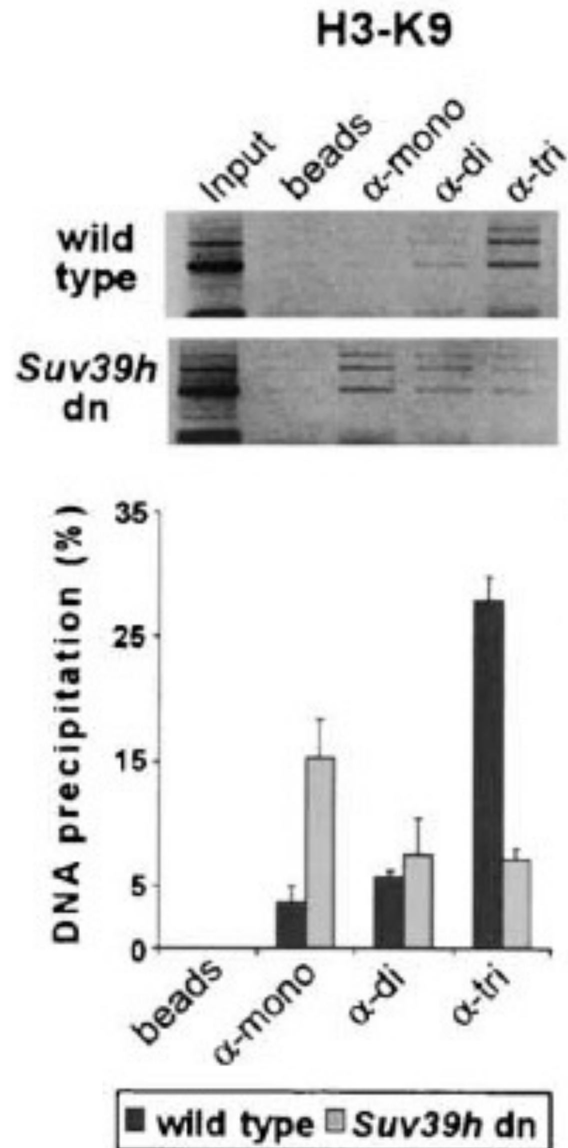
Suv39h1 is required for imposition of H3K9me3 at pericentric repeats

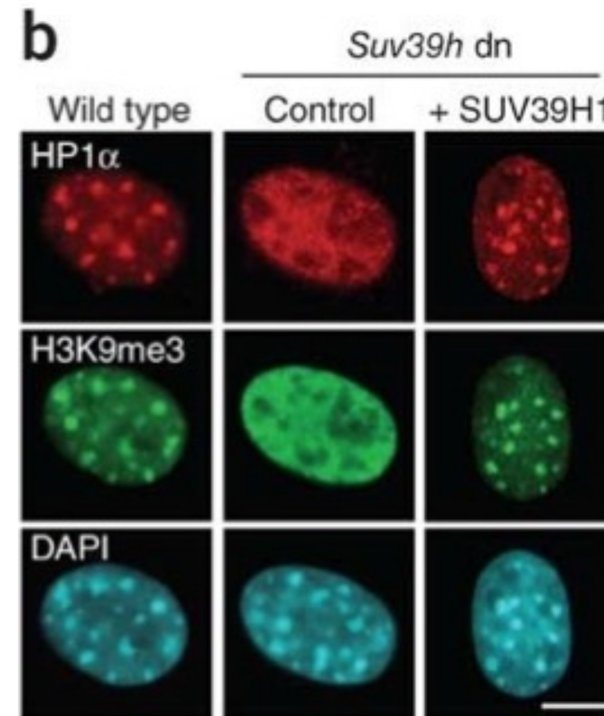
4. How to allocate histone modifications and epigenetic readers along the DNA?

CHROMATIN IMMUNOPRECIPITATION (ChIP) COUPLED WITH PCR

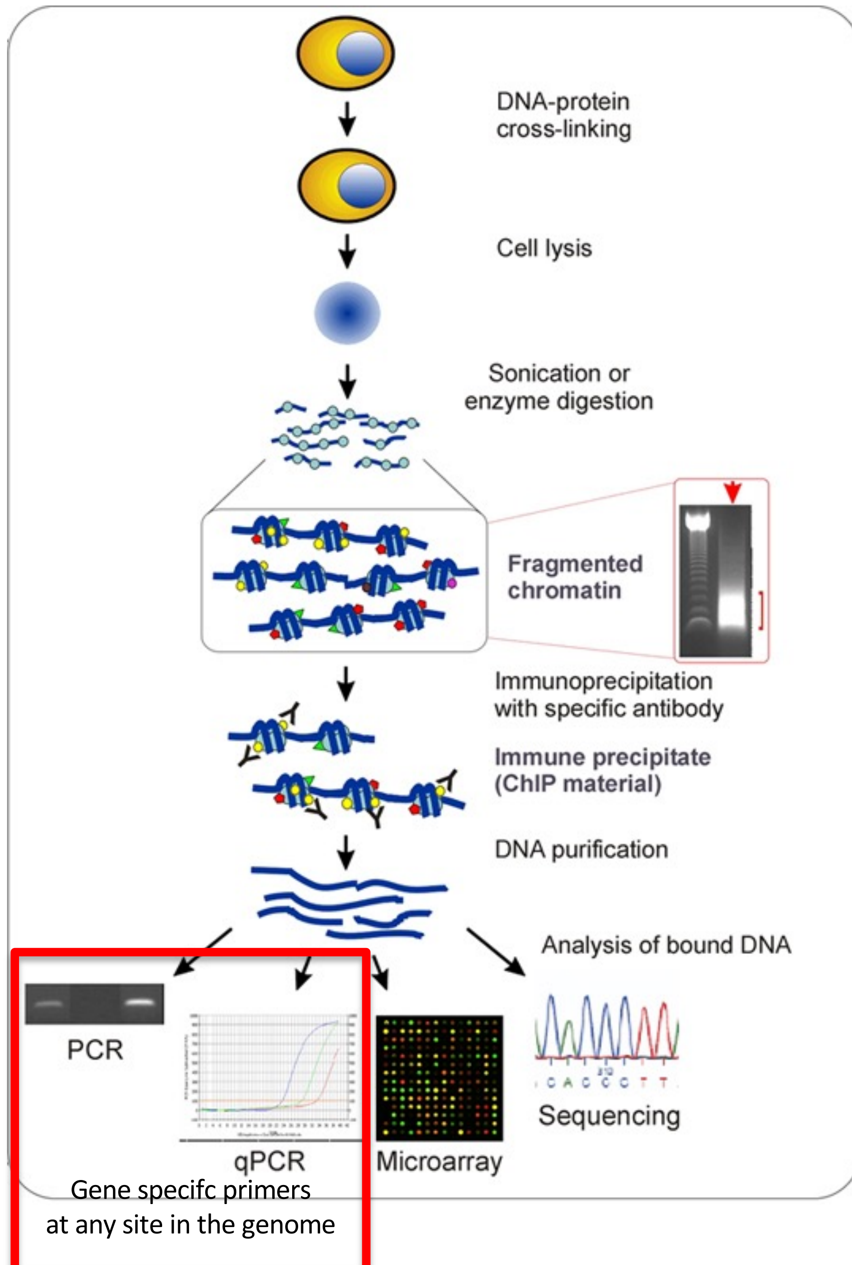
→ H3K9me3 is enriched at pericentric (major+minor) repeats in mouse cells

Major satellite repeats





CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILLED ANALYSIS
→ Localization of protein at a defined region - seqeunce



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

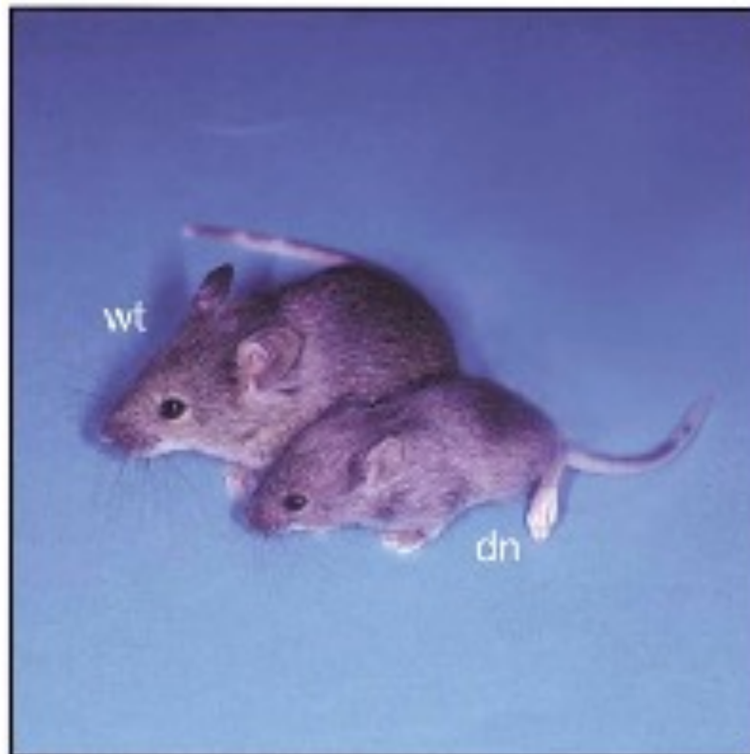
HOW TO STUDY EPIGENETIC MODIFICATIONS

1. **Specfic antibodies are central to epigenetic research**
2. **Mass spectrometry can read histone codes?**
3. **How can we identify enzymatic activites of writers?**
4. **How can we identify epigenetic readers?**
5. **How can we locate specific histone modifications in the nucleus/along the genome?**
6. **Identification of biological relevance of chromatin modifications**

-- STRATEGIES USING SUV39H1 AS A HALLMARK MODEL FOR EPIGENTIC REGUALTION --

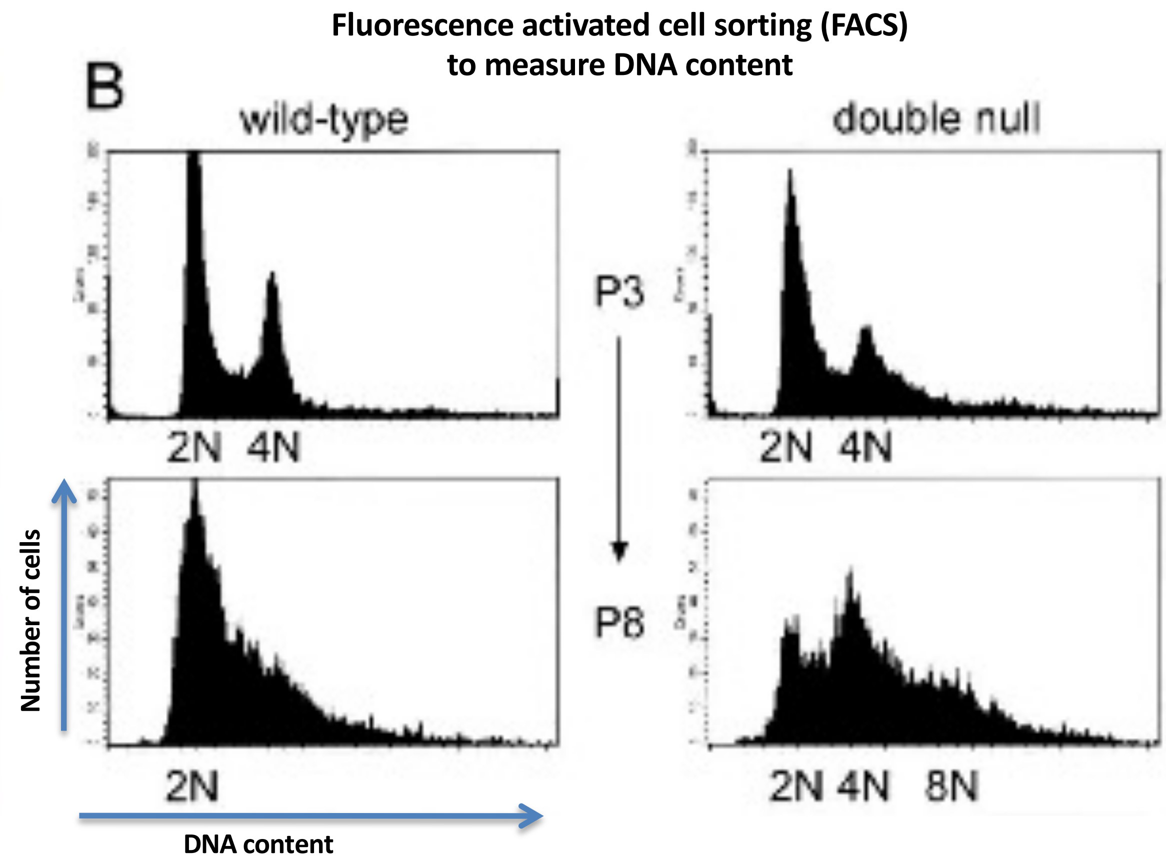
6. IDENTIFICATION OF BIOLOGICAL RELEVANCE OF CHROMATIN MODIFICATIONS

Lack of SUV39h HMTase activity results in genomic instability



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

Fluorescence activated cell sorting (FACS)
to measure DNA content

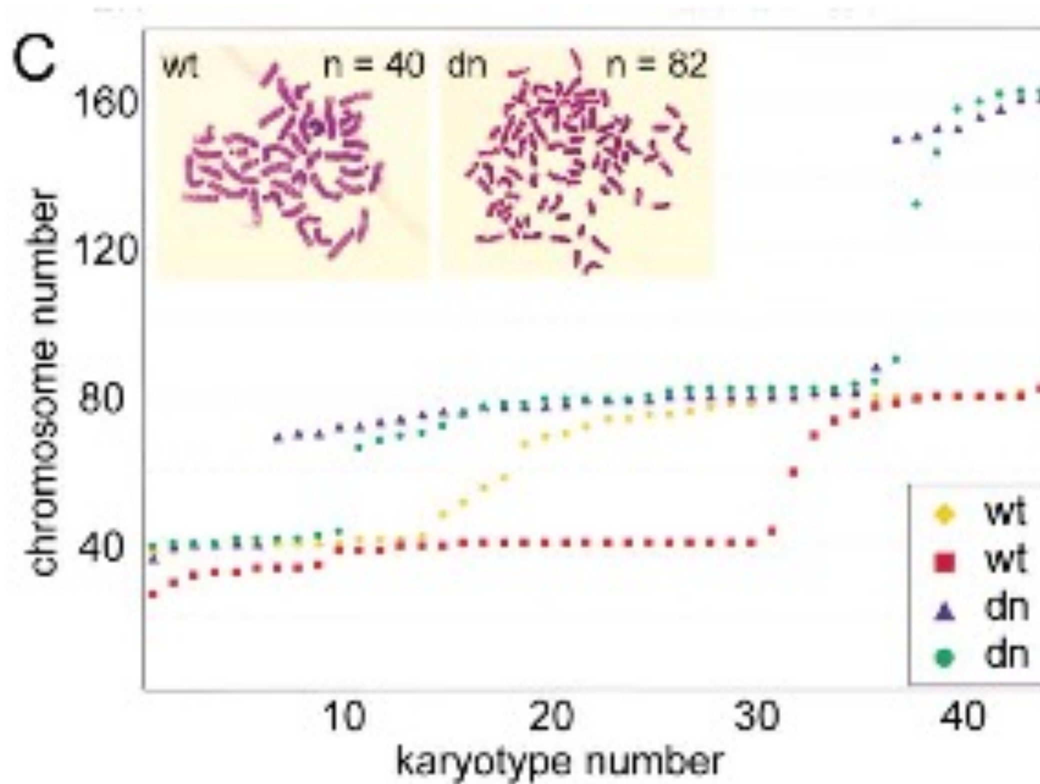


Fibroblasts from Suv39h1/2 null mice
are aneuploidy

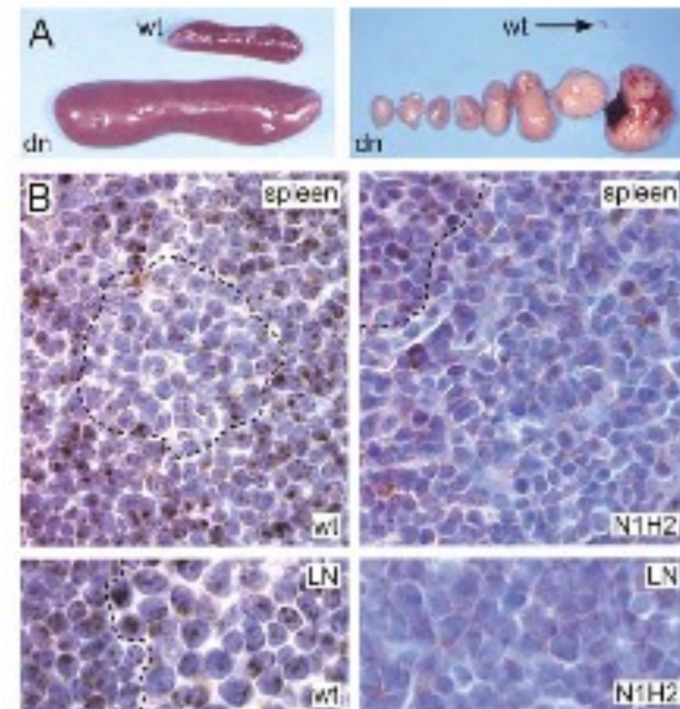
5. IDENTIFICATION OF BIOLOGICAL RELEVANCE OF CHROMATIN MODIFICATIONS

Lack of SUV39h HMTase activity results in genome instability

Loss of Suv39h1/2 results in increased chromosome numbers

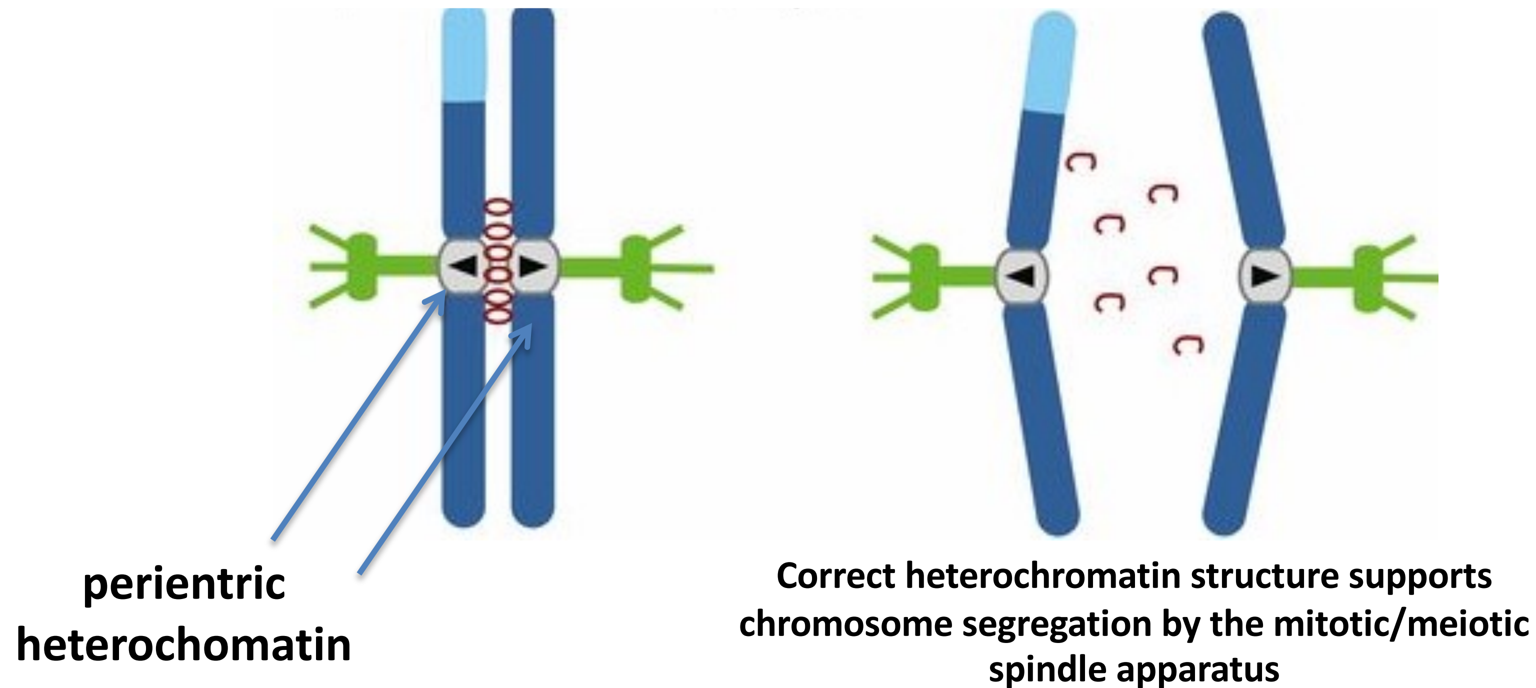


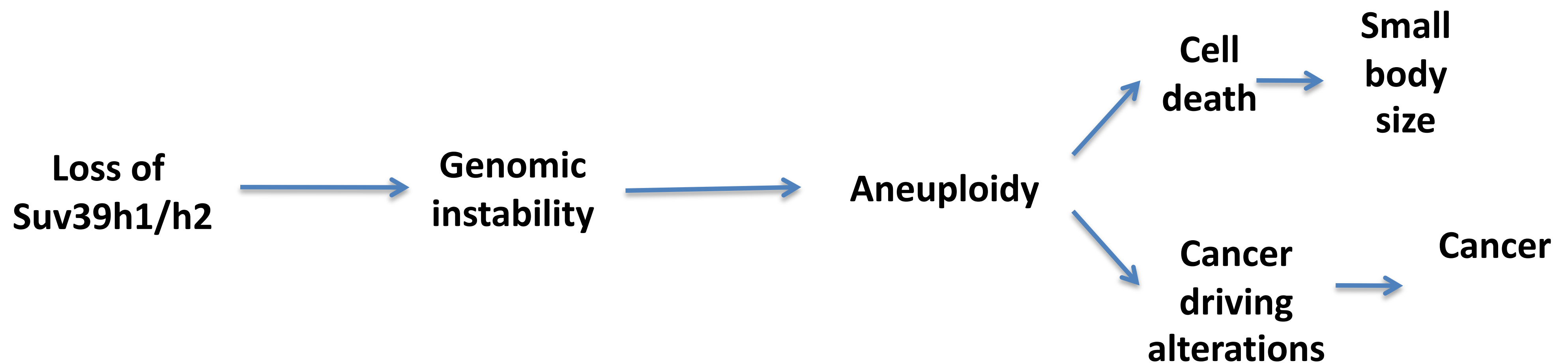
Genomic instability in Suv39h1/2 mice increases lymphomas



5. IDENTIFICATION OF BIOLOGICAL RELEVANCE OF CHROMATIN MODIFICATIONS

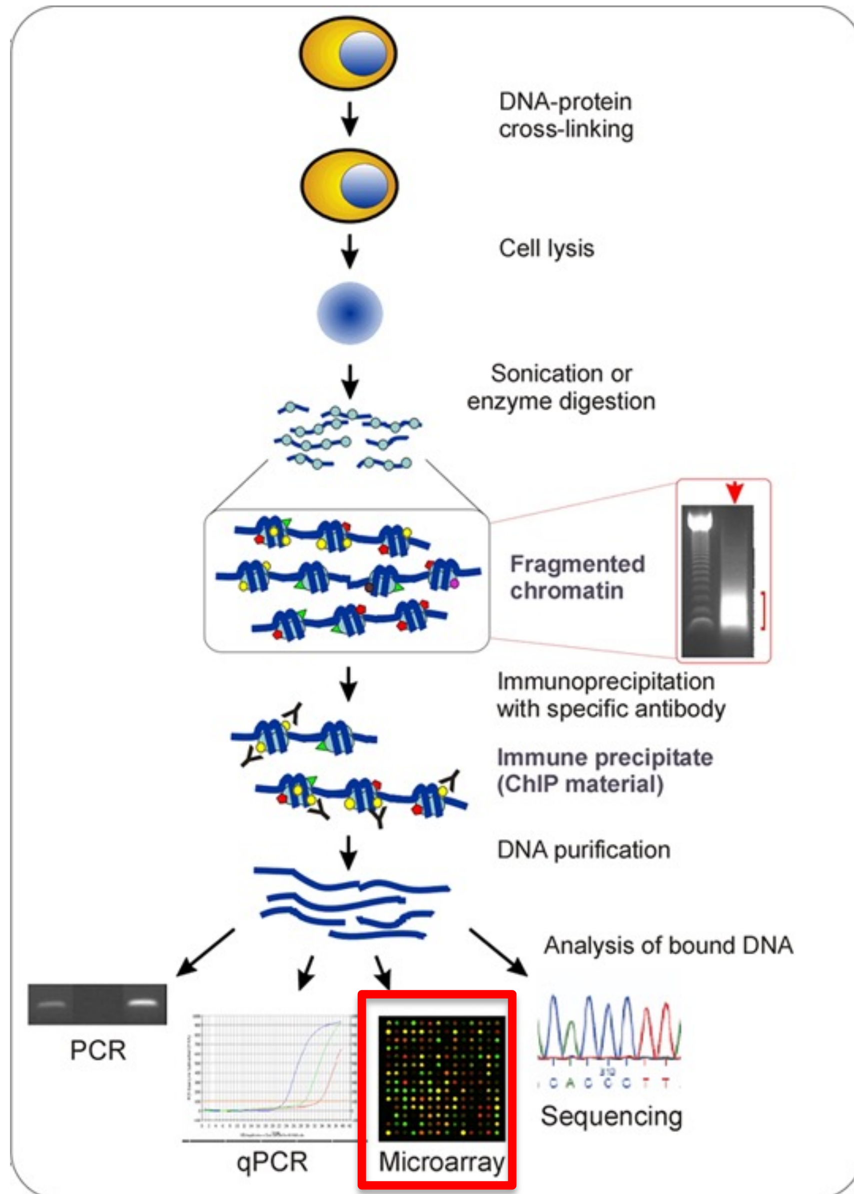
SUV39h HMTase activity is essential for fidelity in mitosis/meiosis





Array hybridization and ChIP-seq to generate maps of epigenomes in different organisms and cell types

CHROMATIN IMMUNOPRECIPITATION (ChIP) → DETAILLED ANALYSIS
→ Localization of protein at a defined region - seqeunce



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

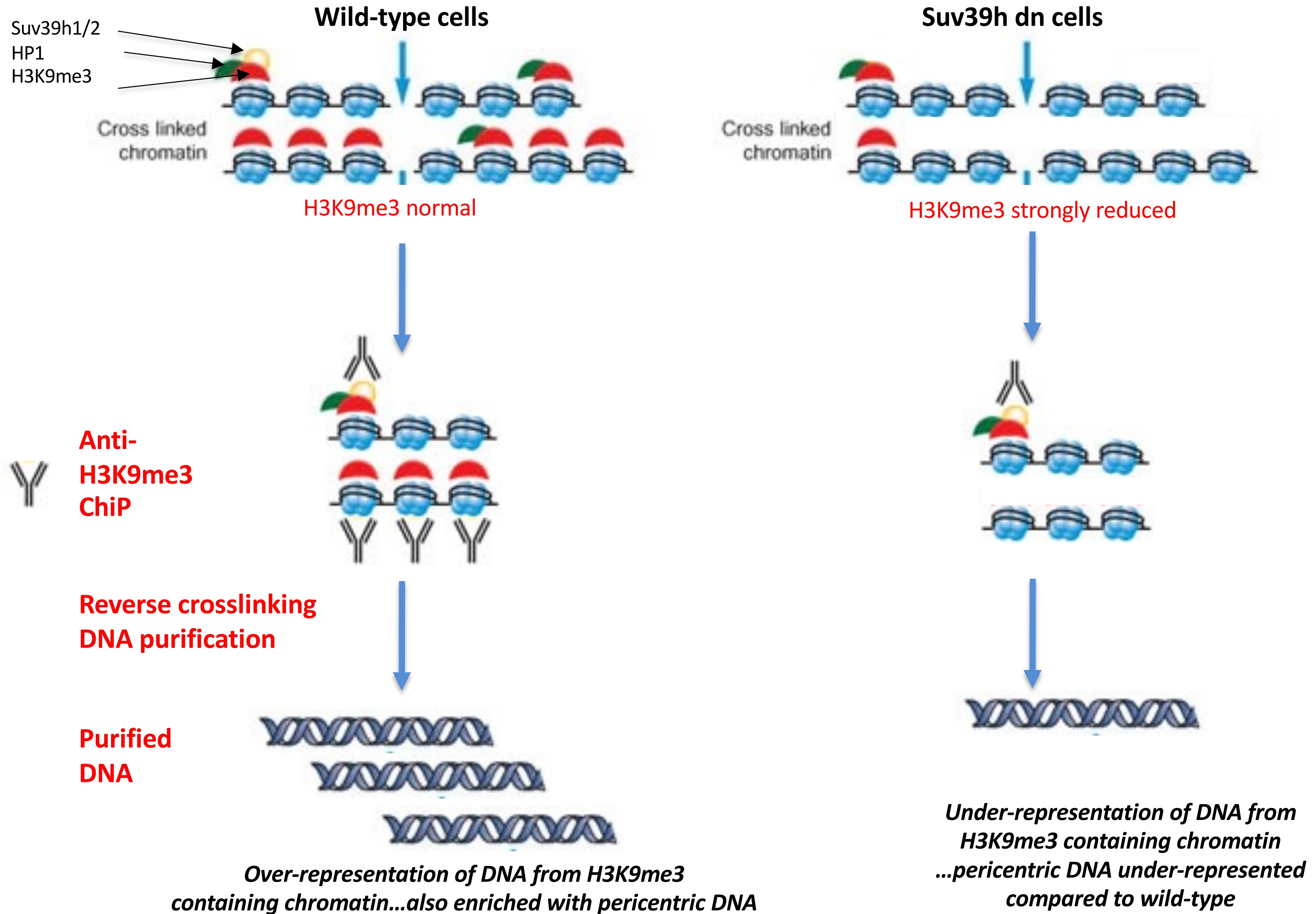
7. Purify DNA and precipitate DNA

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

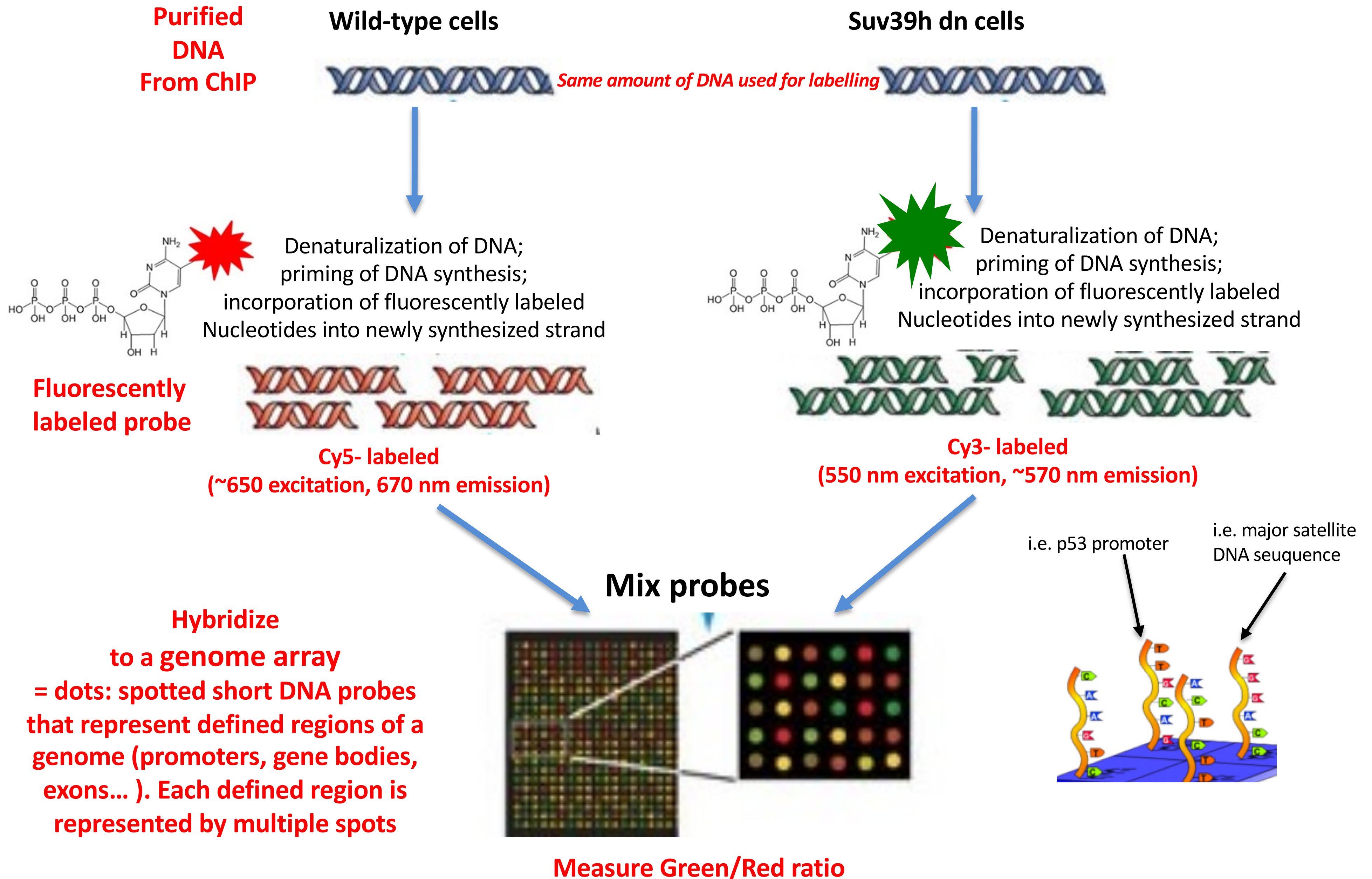
METHODS:

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

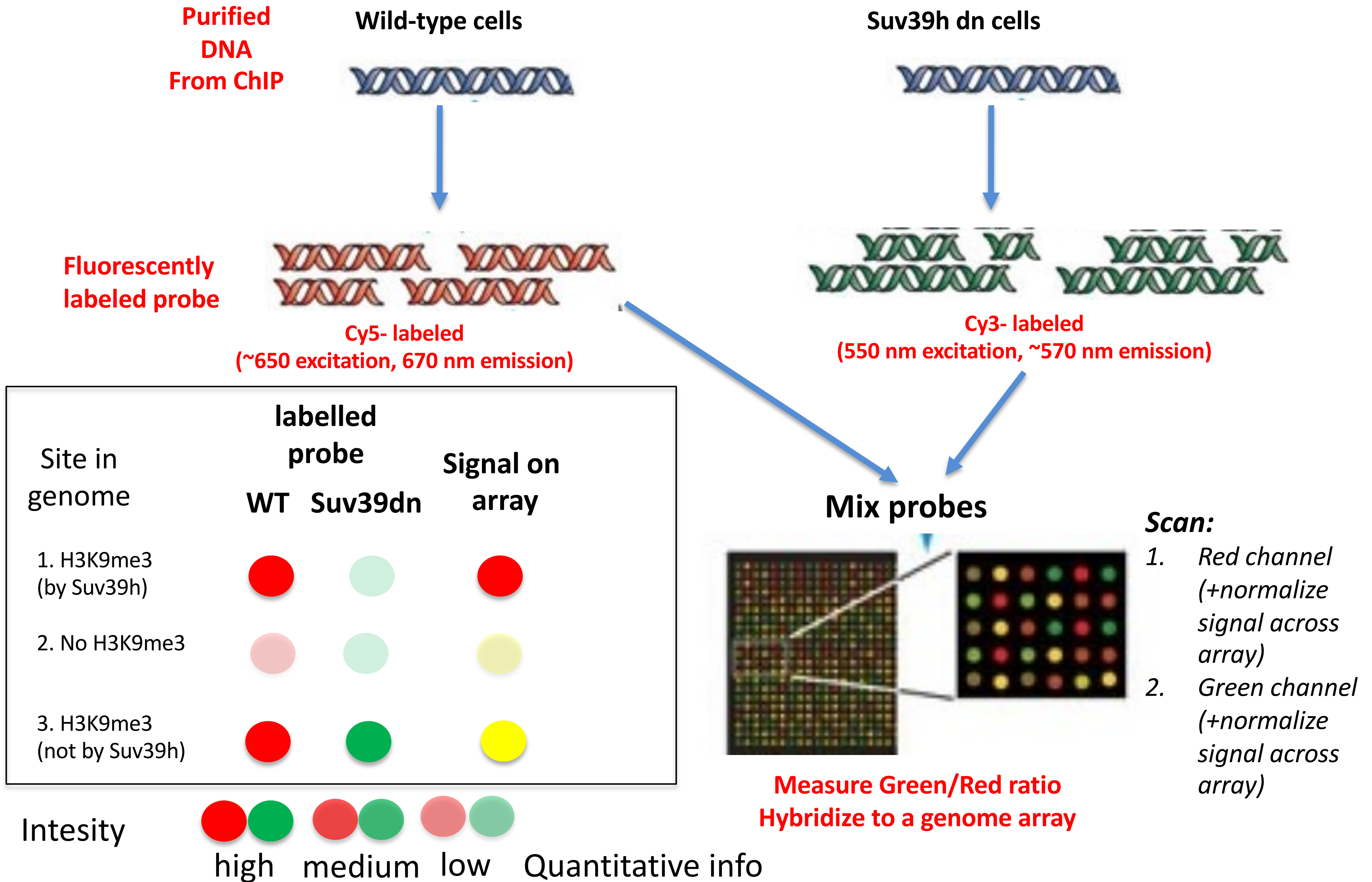
1. ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



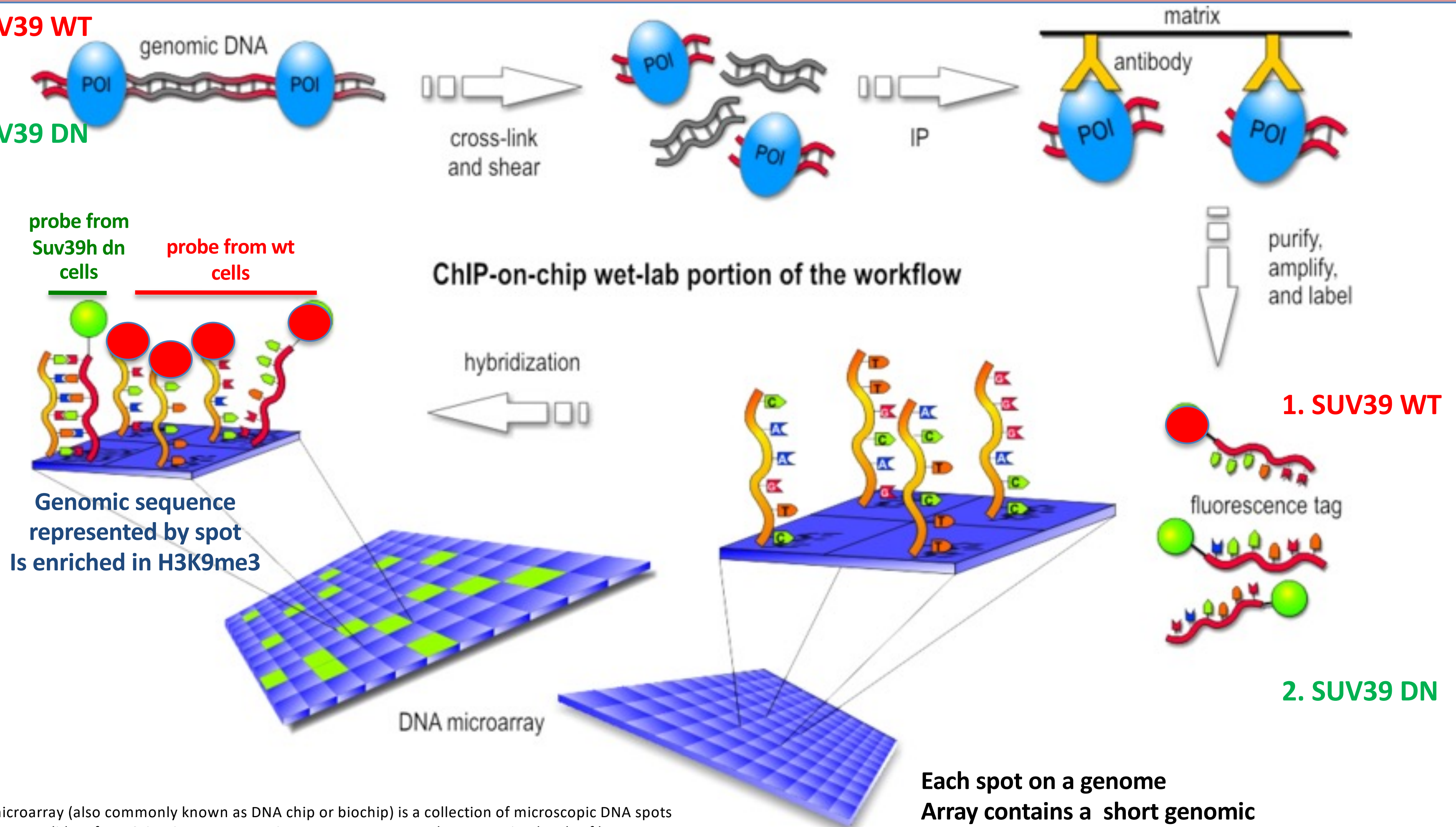
1. ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome



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



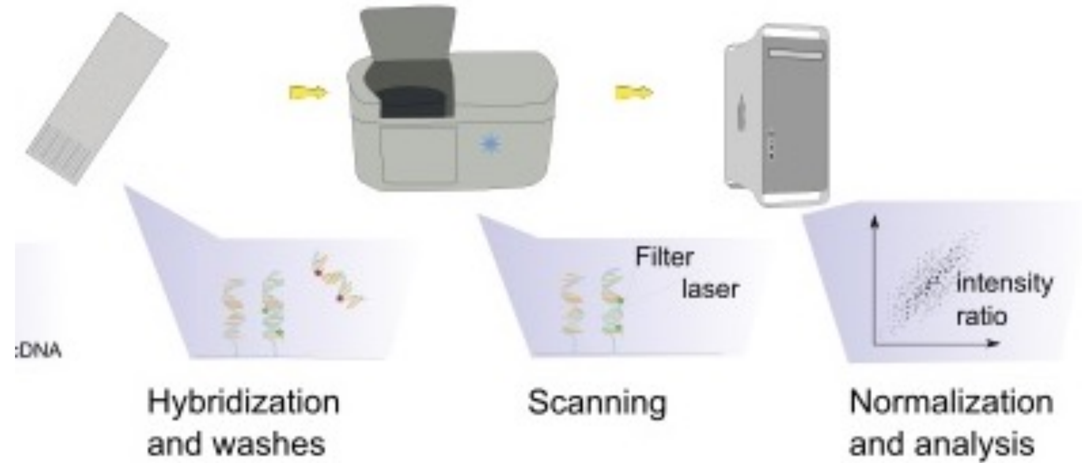
1. ChIP on ChIP: Analysis of epigenetic information across a high number of genomic sites in the genome

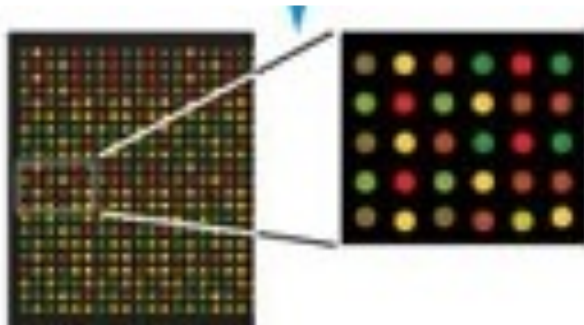


A DNA microarray (also commonly known as DNA chip or biochip) is a collection of microscopic DNA spots attached to a solid surface. Scientists use DNA microarrays to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome. Each DNA spot contains picomoles (10−12 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

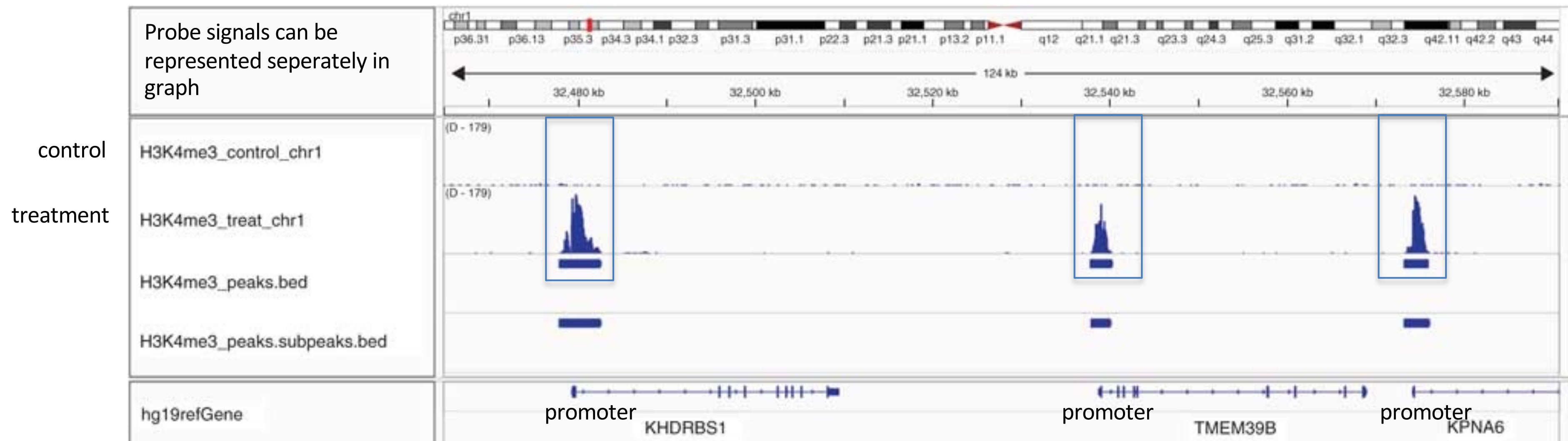
Each spot on a genome
Array contains a short genomic Sequence. Genome arrays cover a large part of the Genome
A ChIP holds ca 20000 spots

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

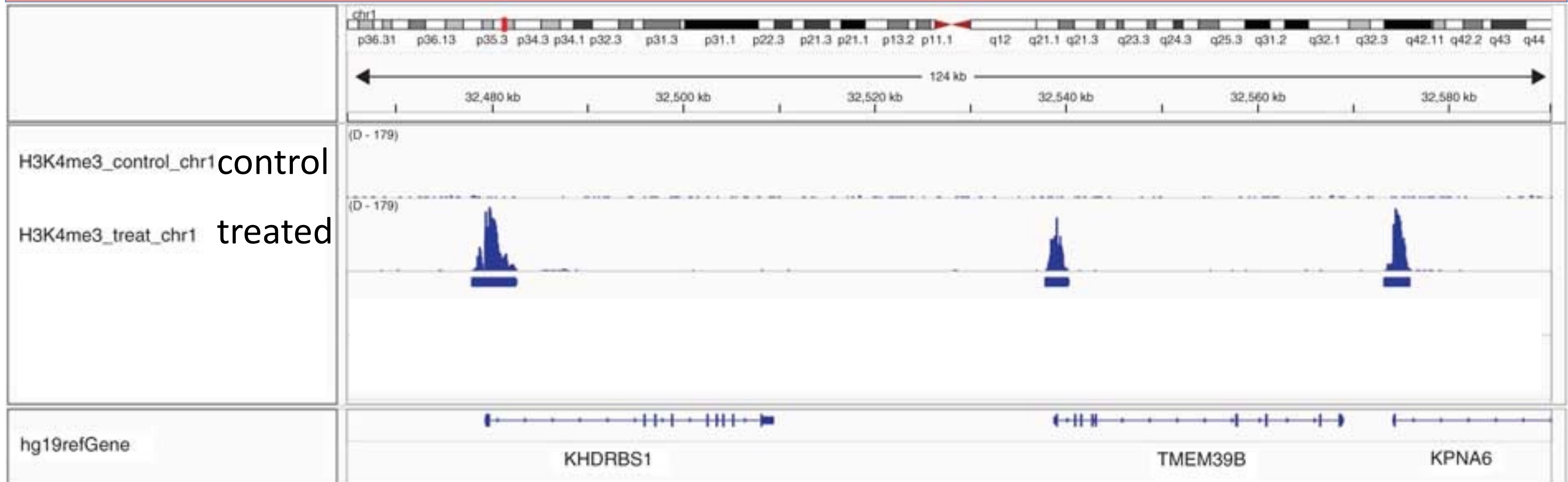




- Arrays do not contain the entire human/mouse genome
- Arrays are enriched for probes (short DNA seqeunces) that map to defined regions.
- Array types are chosen according to a defined experimental need (promoter, enhancer, etc…)
- Interesting elements are represented by multiple probes (along element) → higher resolution



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



Advantage:

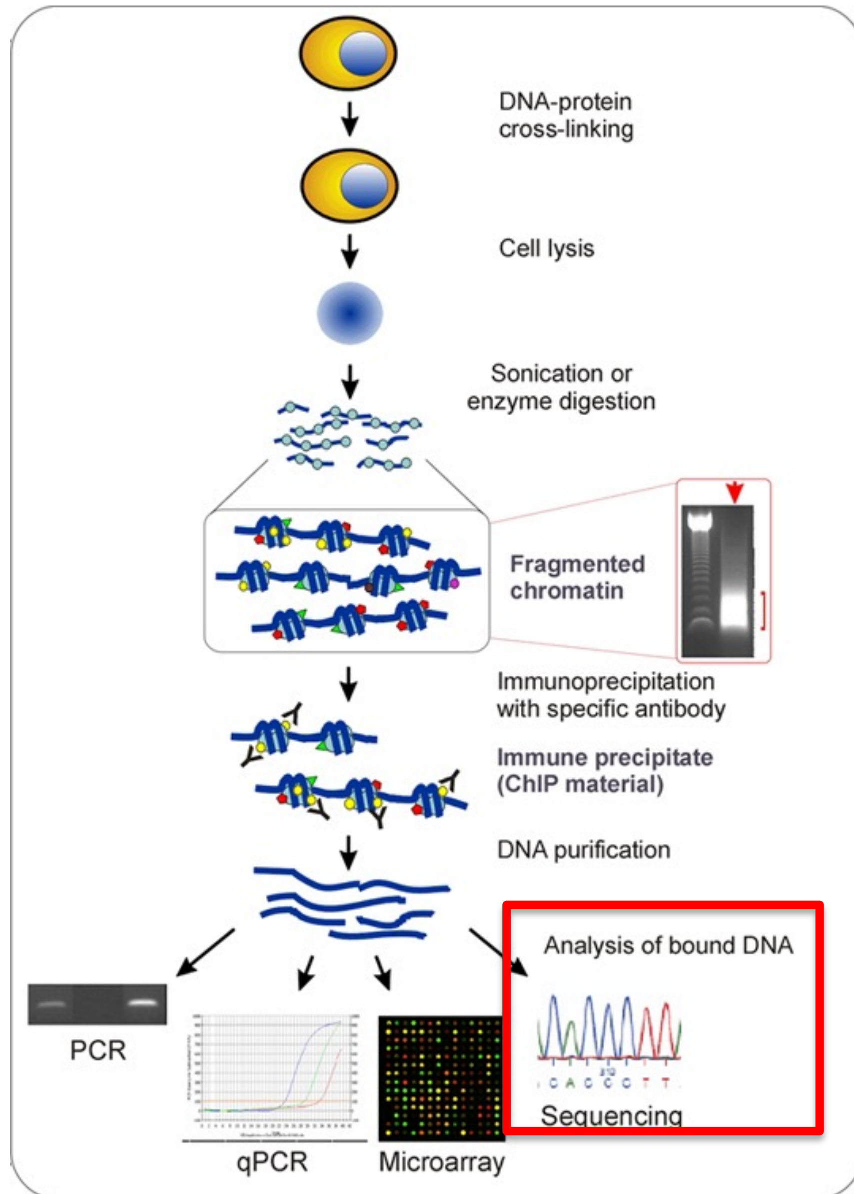
- low tech
- Cheap
- Robust

Disadvantage:

- low resolution (no information on the DNA sequence level)
- Information is restricted to sequences represented by probes – array surface is too small to represent the entire genome
- no data on number of molecules – just proportions;
- laborious to reach a good genome coverage

Already outdated → state of the art: ChIP seq

2. ChIP seq: Analysis of epigenetic information on the single nucleotide level → GENERATION OF GENOME WIDE EPIGENTIC 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

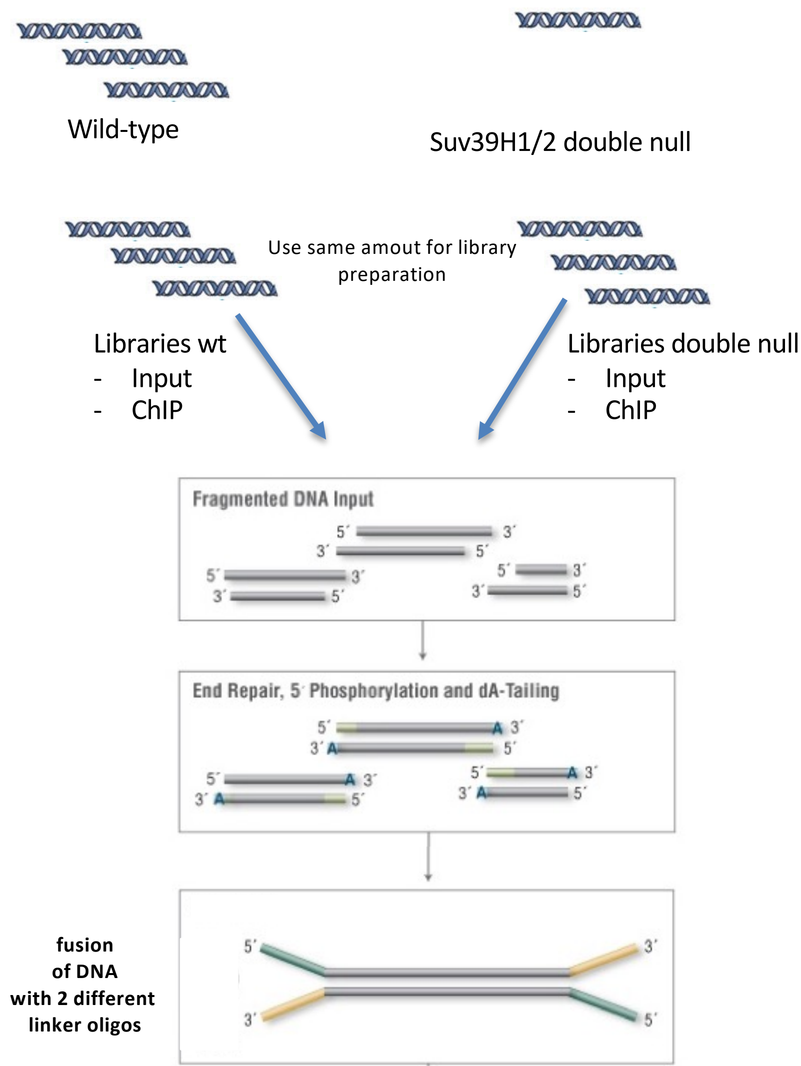
METHODS:

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

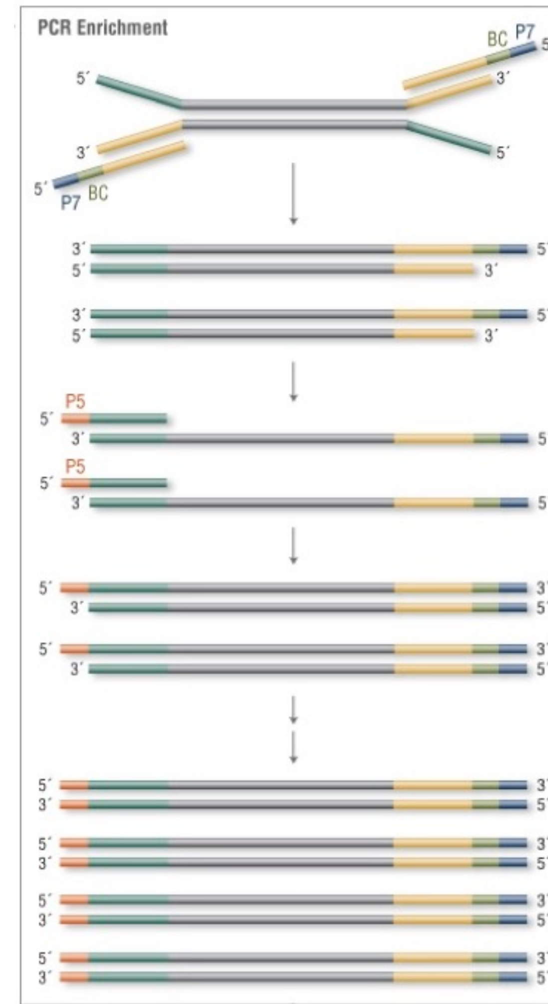
2. ChIP seq: Analysis of epigenetic information on the single nucleotide level
→ GENERATION OF GENOME WIDE EPIGENTIC MAPS

Use immunoprecipitated DNA to generate DNA libraries for massive parallel sequencing:

Purified DNA from anti-H3K9me3 ChIP



fusion of DNA with 2 different linker oligos



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 SEQEUNCING

2. ChIP seq: Analysis of epigenetic information on the single nucleotide level → GENERATION OF GENOME WIDE EPIGENTIC MAPS

Illumina Massively Parallel Sequencing

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

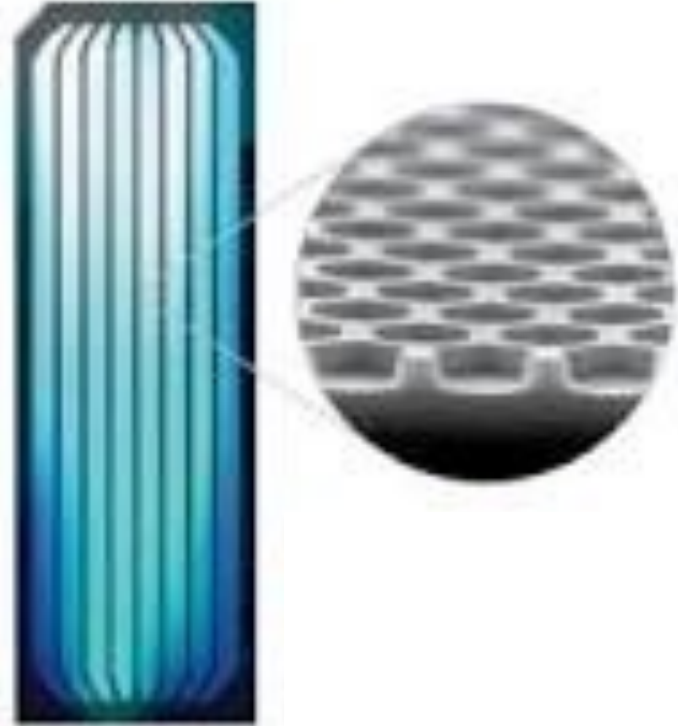
The heart of the Illumina Massive Parallel Sequencer is the "FLOW-CELL". A surface with millions of small wells that allow individual sanger-sequencing reaction in parallel = "massive parallel sequencing". In each well a SINGLE MOLECULE of DNA Is amplified and sequenced

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

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

2. ChIP seq: Analysis of epigenetic information on the single nucleotide level
→ GENERATION OF GENOME WIDE EPIGENTIC MAPS

CLUSTER AMPLIFICATION:



Flow cell contains surface with millions of wells

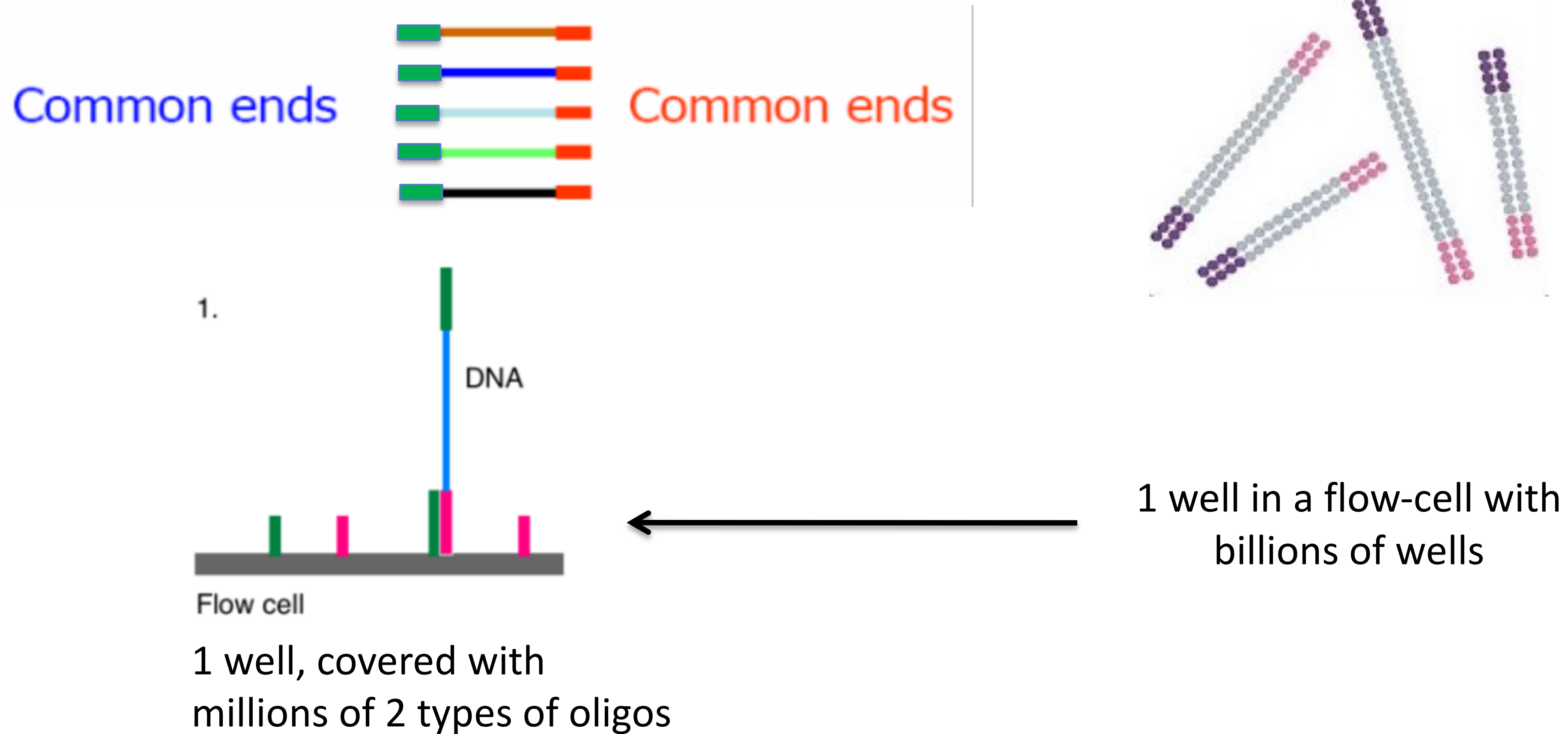
→Each well contains beads mounted with 2 species of oligonucleotides that hybridize with adaptor oligos of DNA library

→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



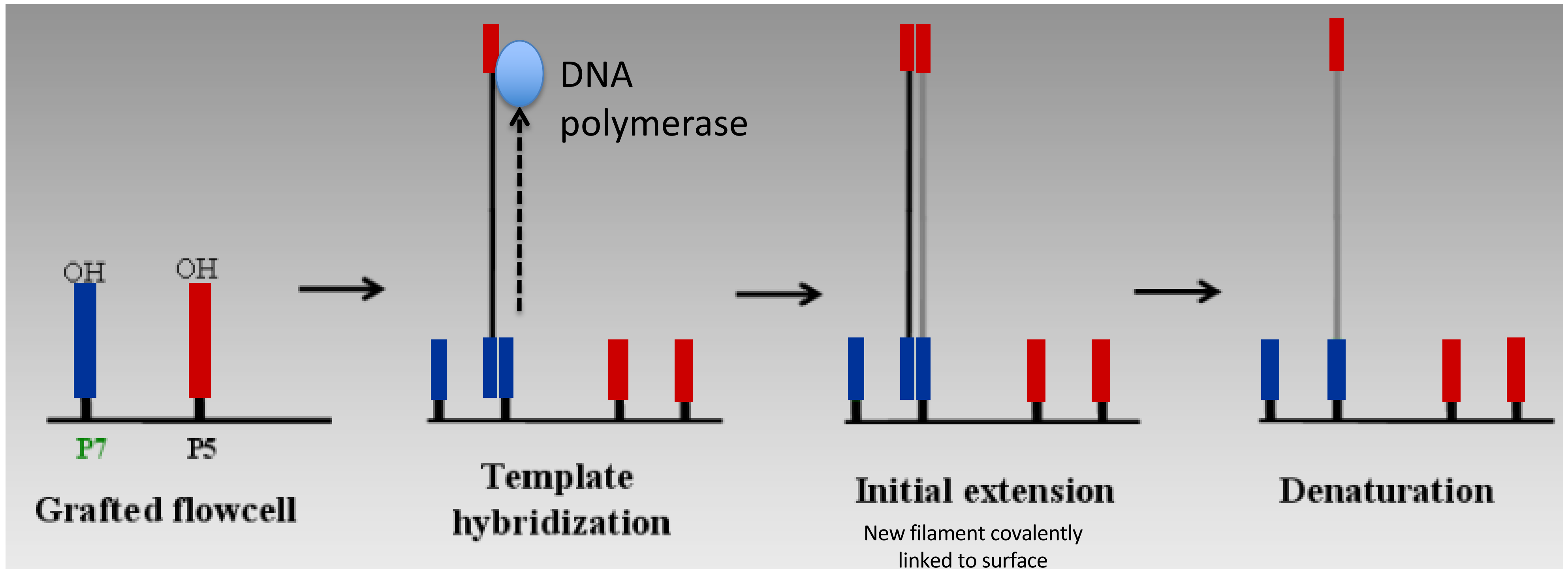
1 well in a flow-cell with billions of wells

1 well, covered with millions of 2 types of oligos

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

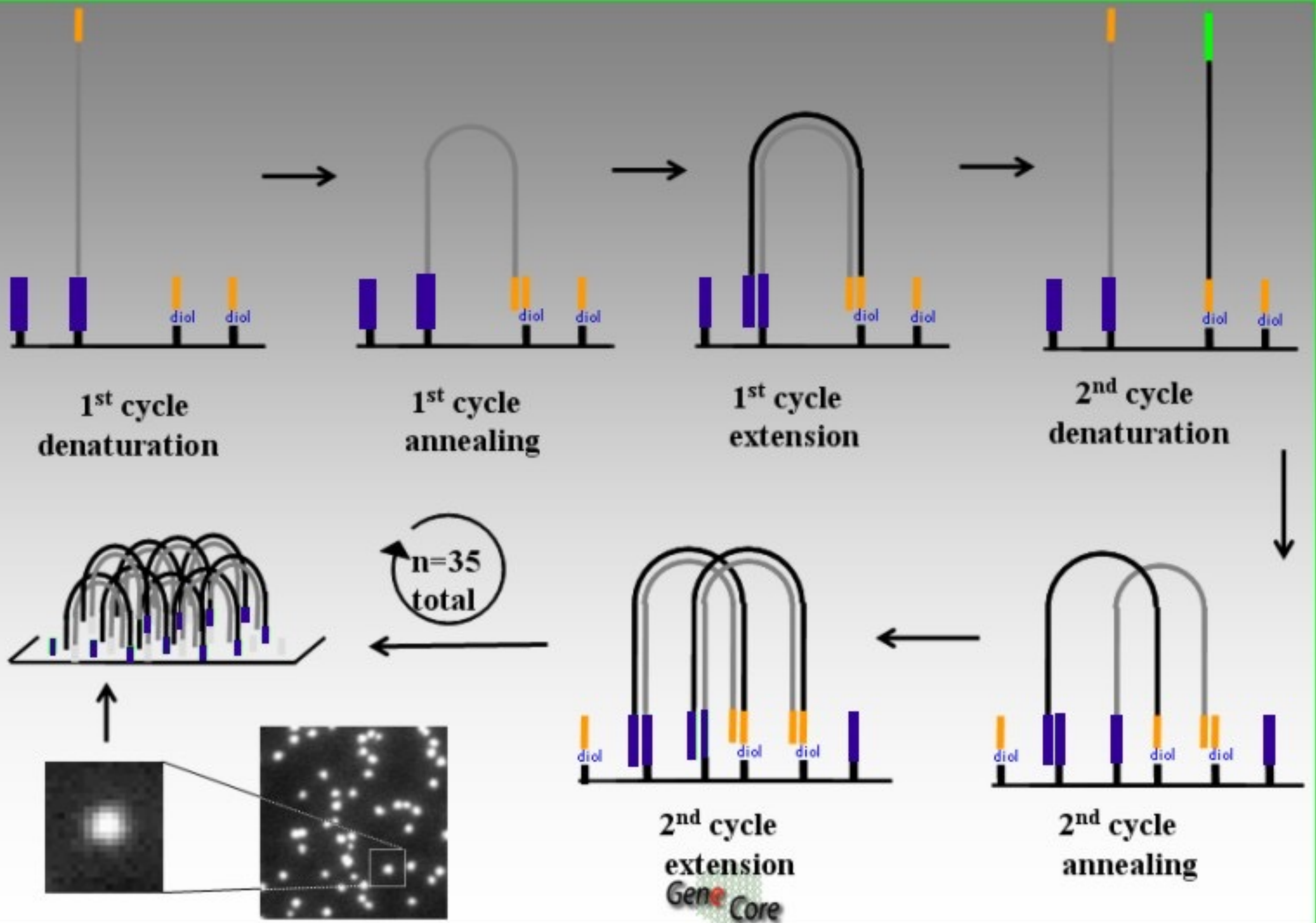
CLUSTER AMPLIFICATION:

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

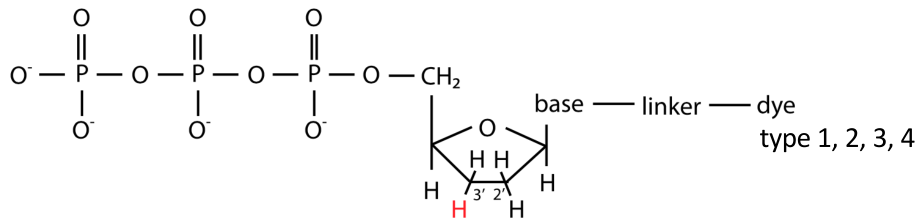


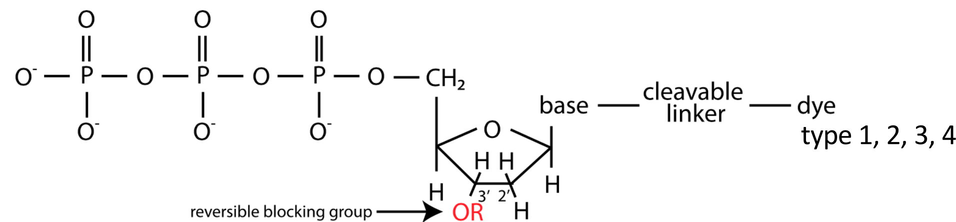
On the surface: complementary oligos

CLUSTER AMPLIFICATION:



DNA SEQUENCING USING REVERSIBLE CHAIN TERMINATORS THAT CARRY A SPECIFIC FLUOROPHOR FOR EACH TYPE OF NUCLEOTIDE





Instead of promoting irreversible primer extension like the Sanger dye terminator method, the reversible chain terminators method uses a cyclic method that consists of nucleotide incorporation, fluorescence imaging and cleavage.
The figure shows a modified nucleotide with a **cleavable dye** and **reversible blocking group. Once the blocking group is removed, a 3'OH is formed and a new nucleotide may come in.**

NOTE: no classic dNTPs are used for sequencing!!!!

Procedure

The steps for such a process can be outlined as follows:

1. Have four dNTP's, each with a different fluorescent marking. These markings should not interfere with base pairing or phosphodiester bond formation.
2. Each dNTP should terminate DNA elongation temporarily with a blocking group on the 3' carbon of the sugar moiety.
3. Upon each cycle, have just one dNTP bind to the elongating strand and emit a fluorescent dye color.
4. Depending on the color emitted, record the particular nucleotide.
5. Cleave the blocking group and fluorescent dye with a palladium-catalyst.
6. Restore a 3' hydroxyl so that the growing strand can now elongate.
7. Repeat from step 1.

Cons

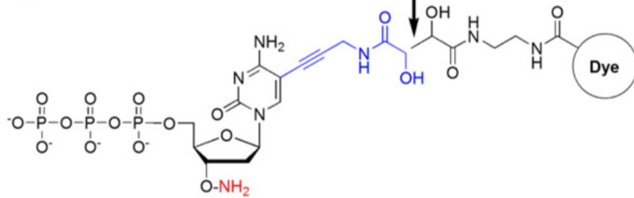
There are some limitations to this method which include:

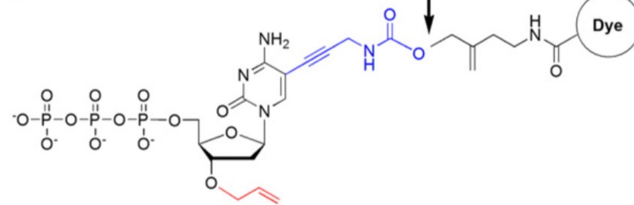
- Incomplete cleavage of blocking groups.
- Difficulties incorporating fluorescent nucleotides.

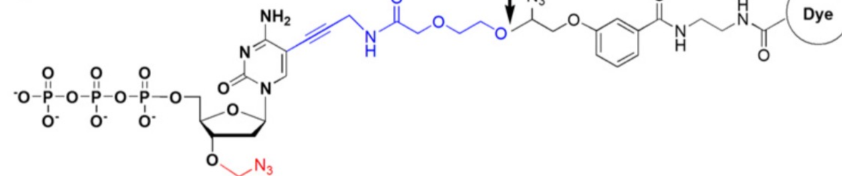
DNA SEQUENCING USING REVERSIBLE CHAIN TERMINATORS:

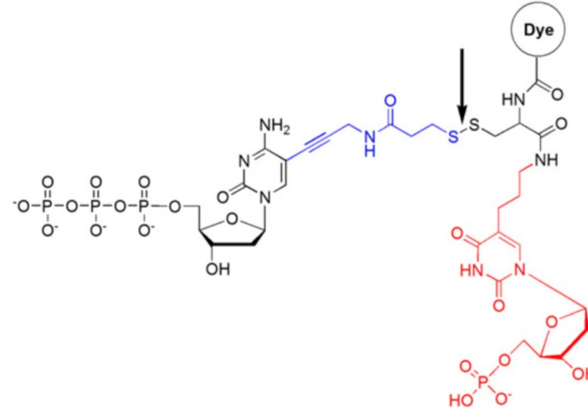
3'-blocked reversible terminators

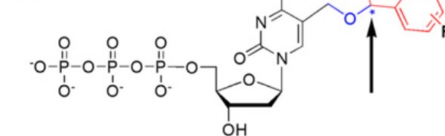
3'-unblocked reversible terminators







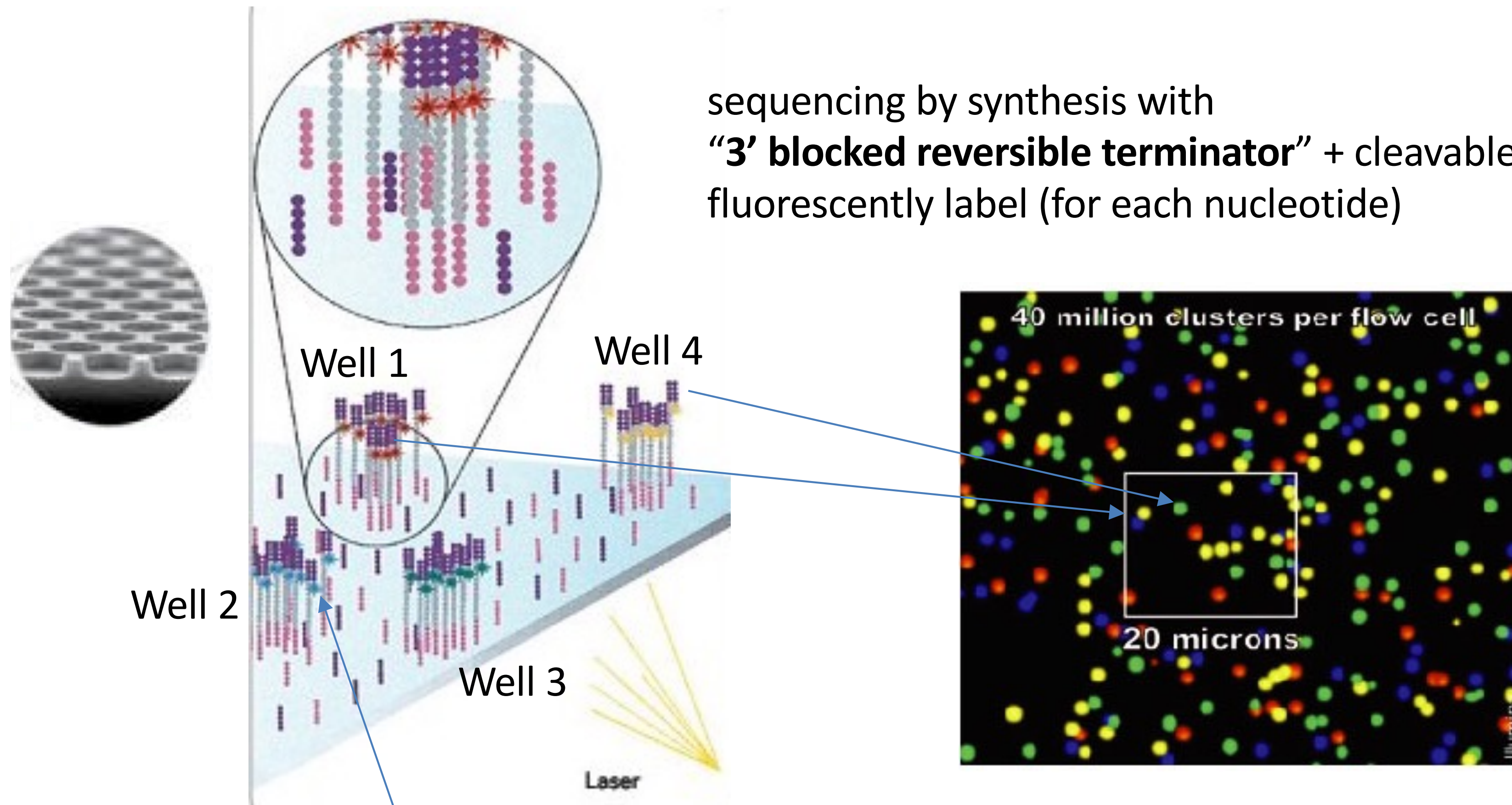




Three different 3’-blocked reversible terminators were shown on the left (A–C) and two 3’-unblocked reversible terminators were shown on the right (D–E).

The chemical structures in red denote the reversible terminating groups. Arrows indicate the site of cleavage separating the fluorescent groups from the nucleotide, and the chemical structures in blue denote the molecular scars that are attached to the base.

Illumina: massive parallel sequencing:



sequencing by synthesis with
"**3' blocked reversible terminator**" + cleavable fluorescently label (for each nucleotide)

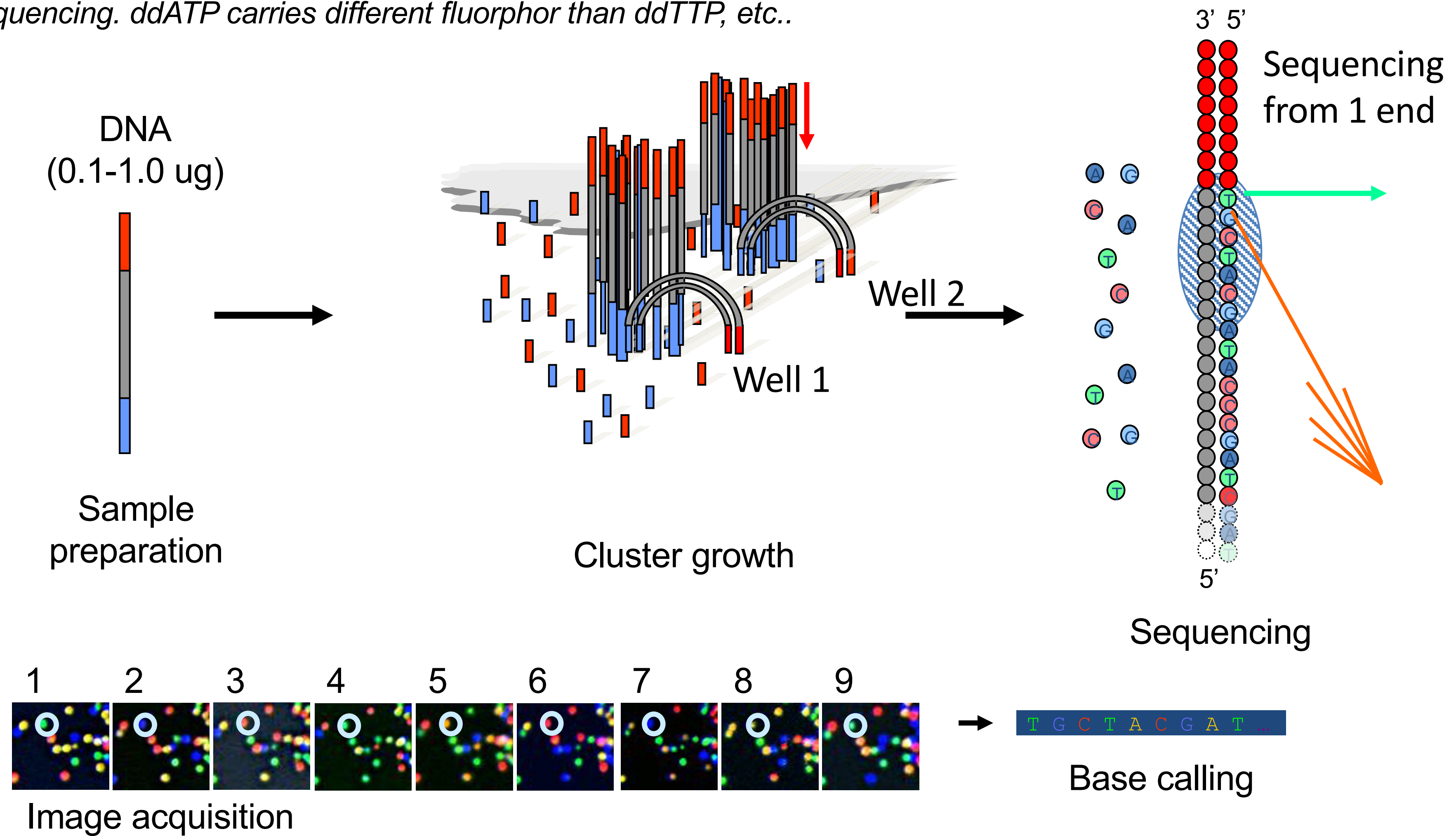
1. **Start of synthesis using primer = incorporation of fluorescent 3'-blocked reversible terminator: synthesis blocked**
2. **Scanning of fluorescent signals of all wells of flow-cell with laser (image)**
3. **Dye cleavage + elimination of reversible blocking group**
4. **wash step**
1. **Repeat steps 1-4 ca. 150x**

illumina.com

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

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 fluorphor than ddTTP, etc..



Illumina: paired end sequencing increases information content

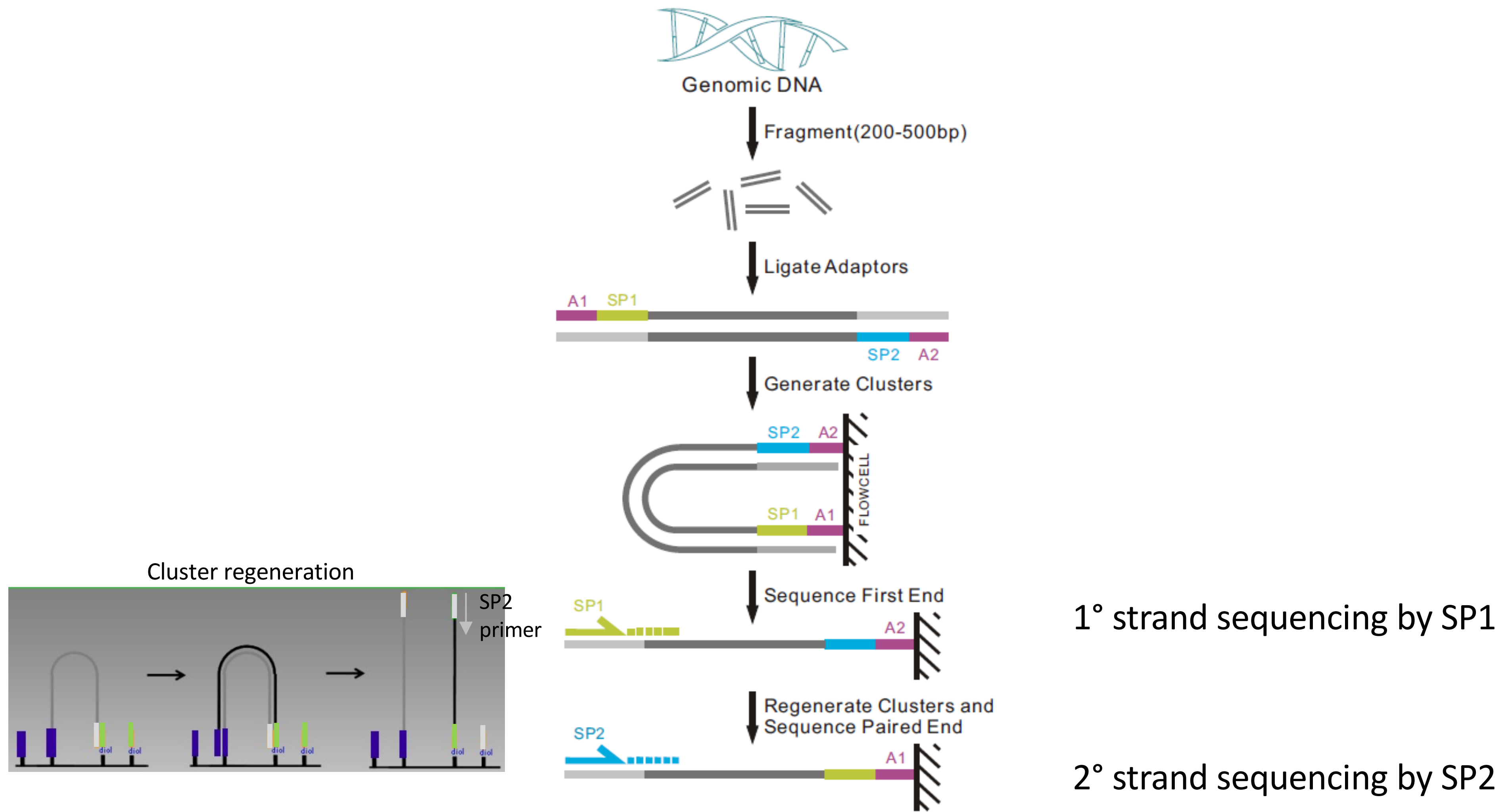


Figure 1-2-1 Pipeline of paired-end sequencing (www.illumina.com)

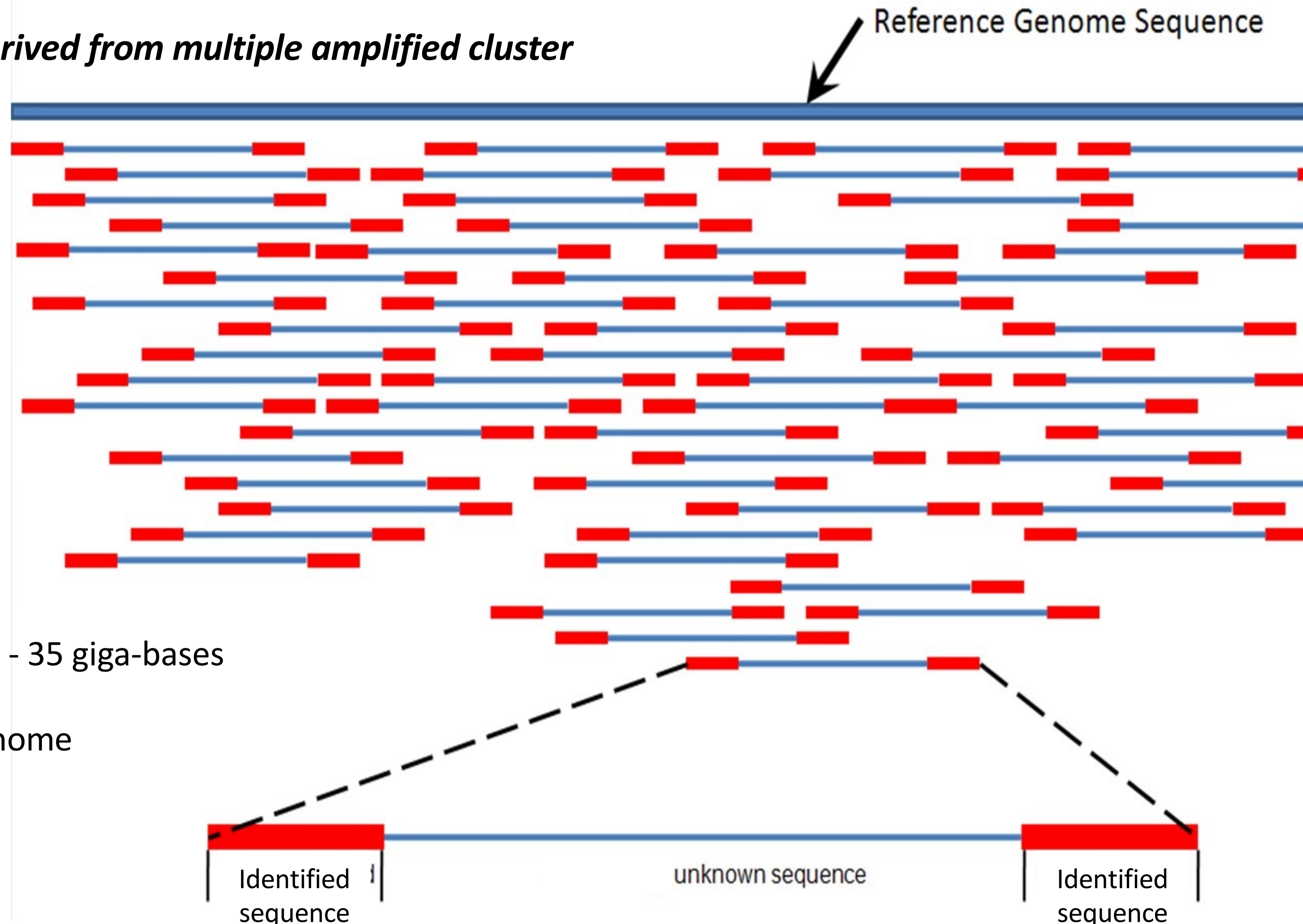
https://www.youtube.com/watch?v=9YxExTSwgPM

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 – enough to cover 3' and 5' ends (representative for a longer genome sequence

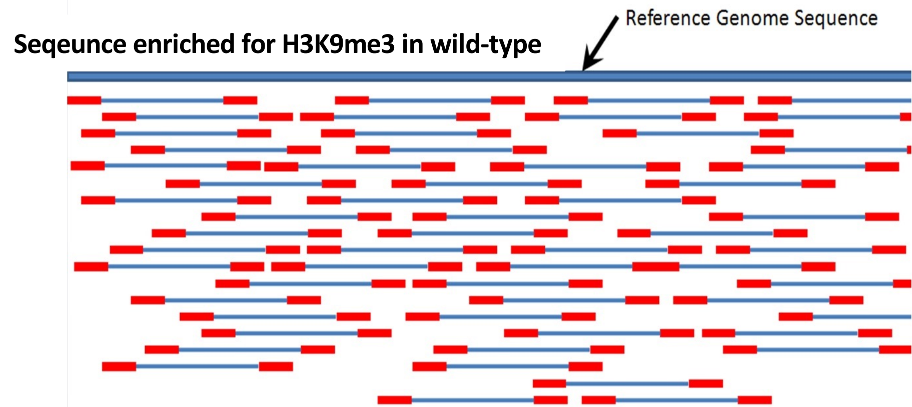
Sequence reads derived from multiple amplified cluster

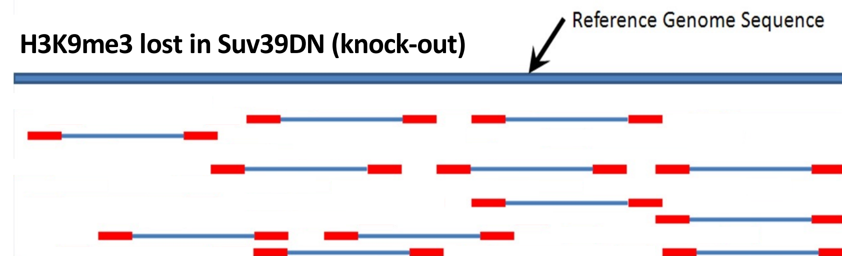


Max. output: 0.5 - 35 giga-bases
$=3.5*10^{10}$
= 10x human genome

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



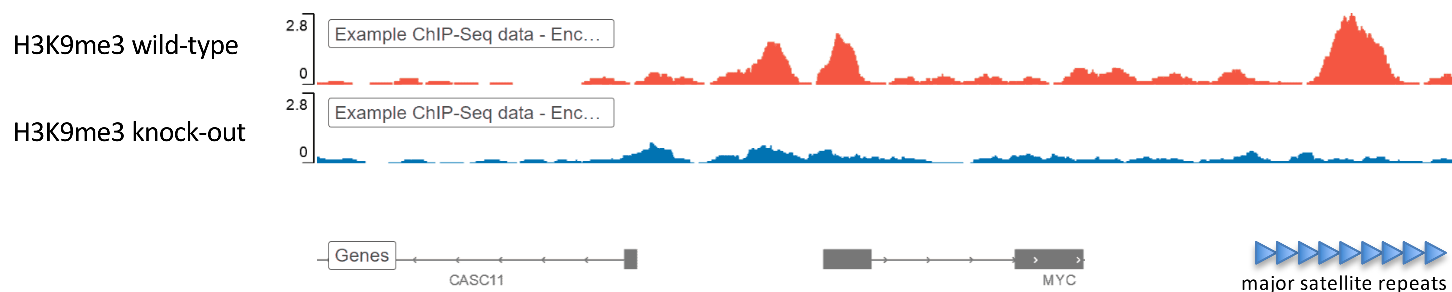


Wildtype

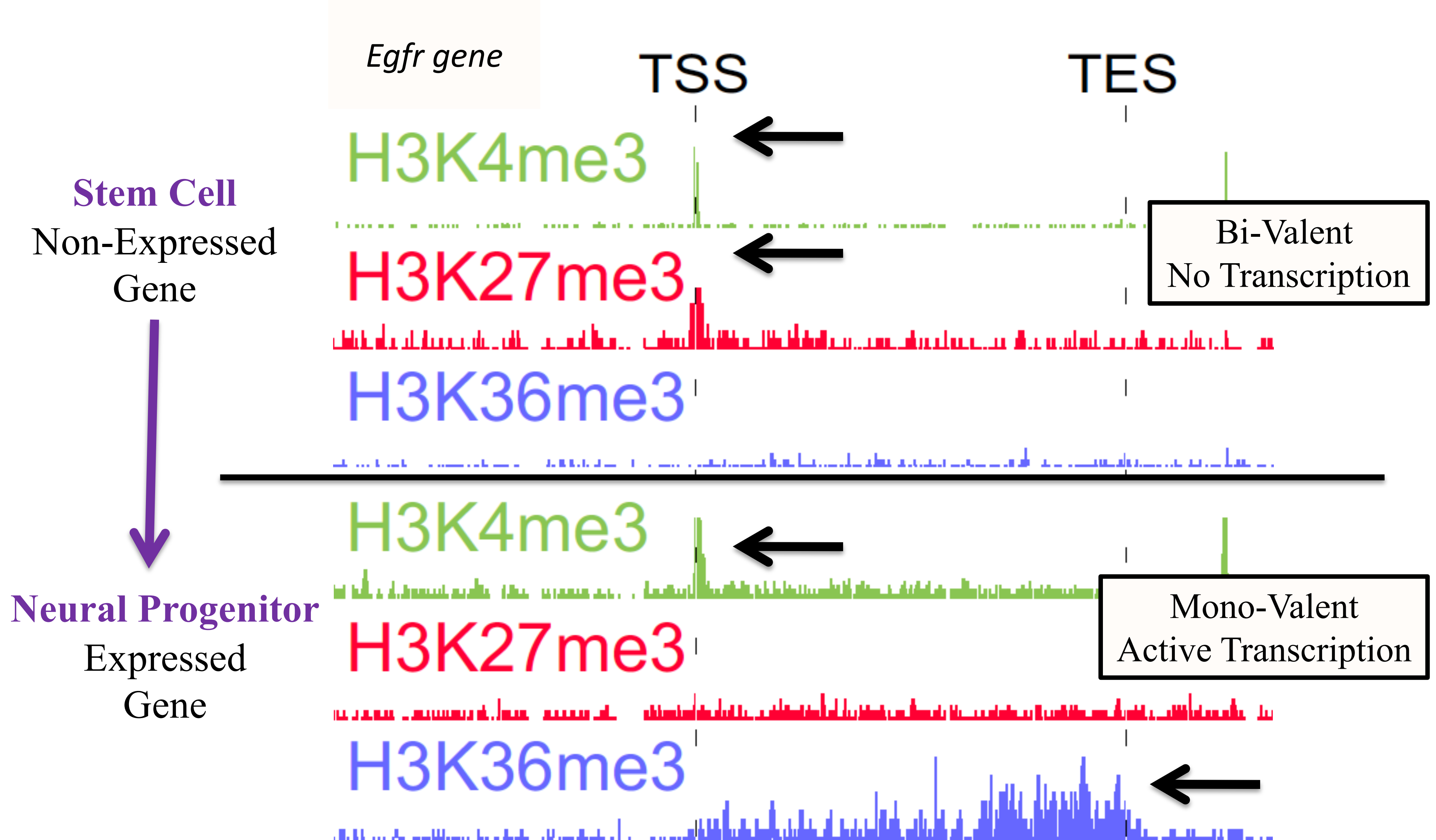
Statistical and quantiative analysis
Reads ChIP vs. Reads Input (genome representation in INPUT)

Suv39DN (knock-out)

Statistical and quantiative analysis
Reads ChIP vs. Reads Input (genome represenetation in INPUT)

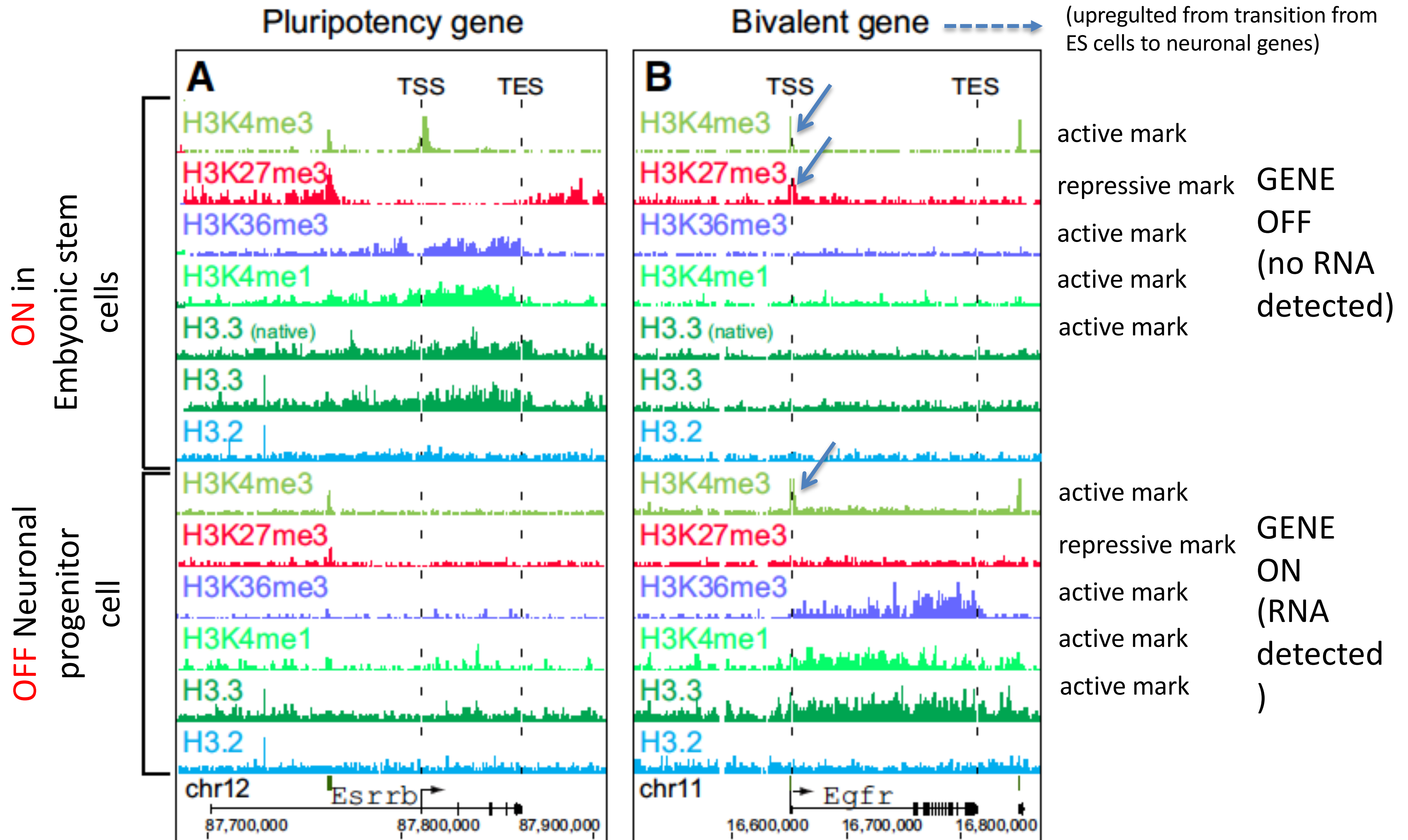


BIOINFORMATICS ANALYSIS:
Mapping ChIP seq reads agins the human genomic sequence



Goldberg et al., Cell, 140: 678-691. 2010

Mapping the epigenetic landscape enables to define "key rules" to define the epigenetic code of active and silent genes



Goldberg et al., Cell, 140: 678-691. 2010

PILE – UP ALIGNEMENT ACROSS THE REFERENCE GENOME