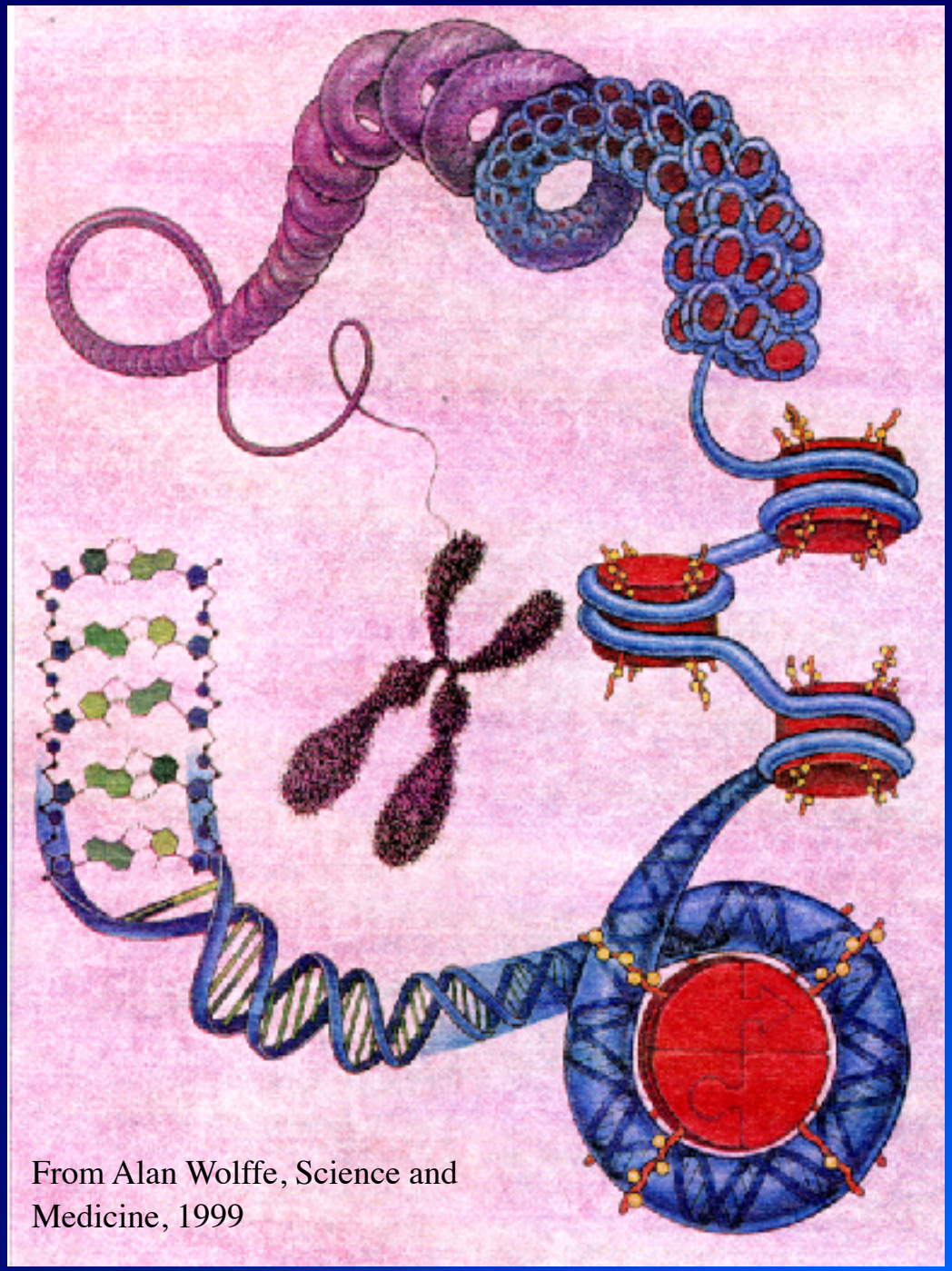
DNA Methylation and Cancer

October 25, 2016

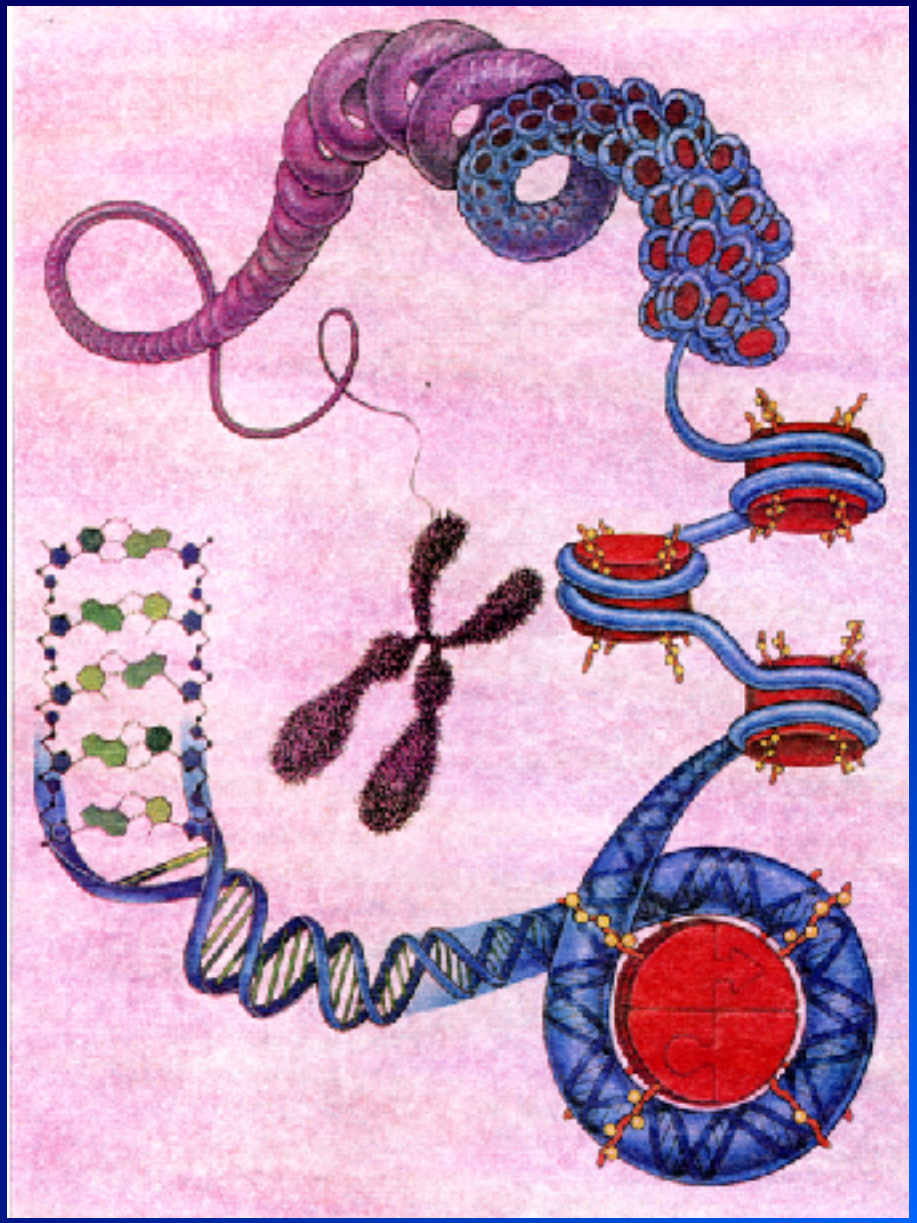
Dominic Smiraglia, Ph.D.
Department of Cancer Genetics



From Alan Wolffe, Science and Medicine, 1999

Vital Statistics

- Human genome contains 3 billion bp
- ~ 50,000 different genes
- 1 bp of DNA is .34 nM long
- That's 2 meters of DNA in every diploid cell
- Each base pair typed side by side would fill 30 encyclopedia volumes
- 2 haploid genomes join at fertilization
- ~ 10 trillion nucleated cells in the human body



How do we exist?

- Complex systems to faithfully replicate the genome
- DNA repair system that proof-reads and fixes errors and breaks
- Complex system of chromatin formation that protects, packages, and compartmentalizes the genome
    - Contributes to patterns of gene expression
- Individual cells “commit suicide” if things get too messed up

In cancer, problems occur with each of these systems

"The best way not to get mutations is to not divide. Unfortunately, one of the consequences of being a multicellular organism is that some of our tissues have to divide a lot."

-Jerry W. Shay, 2001 *Ageing vulnerability: causes and interventions. Wiley, Chichester (Novartis Foundation Symposium 235) p 98.*

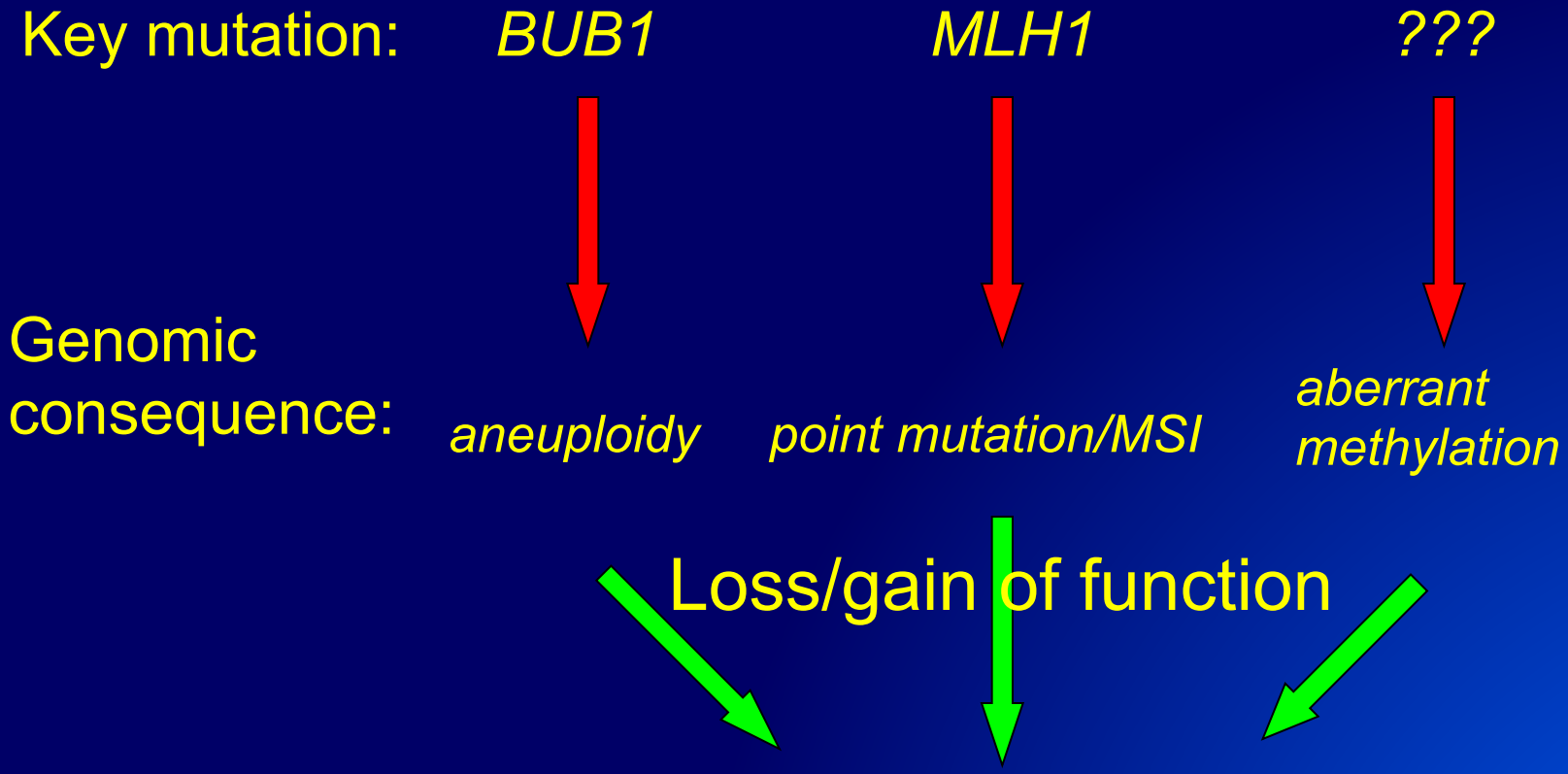
<u>Genetic mutation</u>

- Aneuploidy
- Loss of heterozygosity
- Point mutations
- Translocations

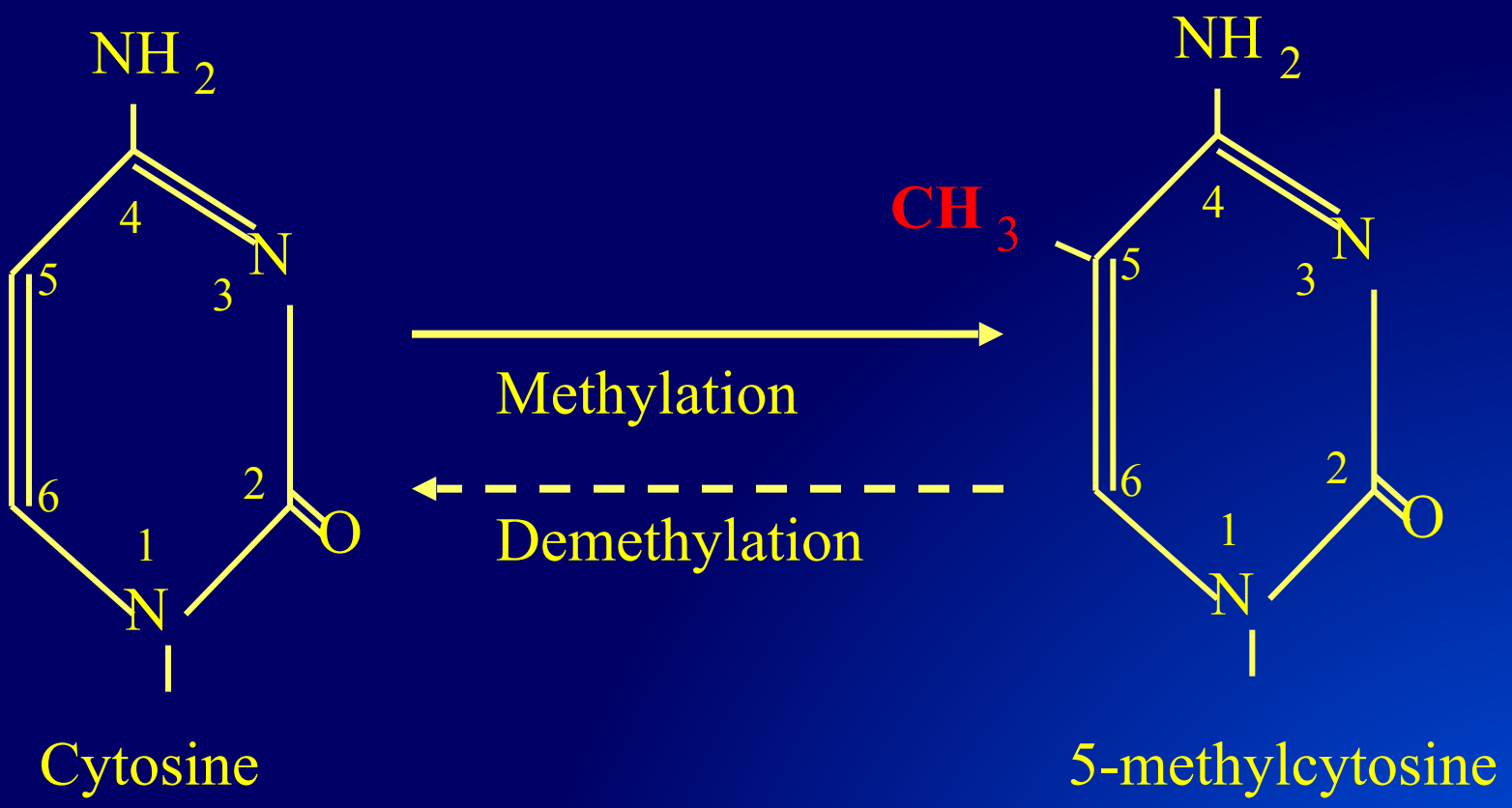
<u>Epigenetic mutation</u>

- Hypomethylation of repeats
- Hypermethylation of CpG islands
- Histone methylation
- Histone acetylation/deacetylation

Genomic instability



DNA methylation is an epigenetic modification of the DNA



Does not change the coding of the DNA, hence epigenetic

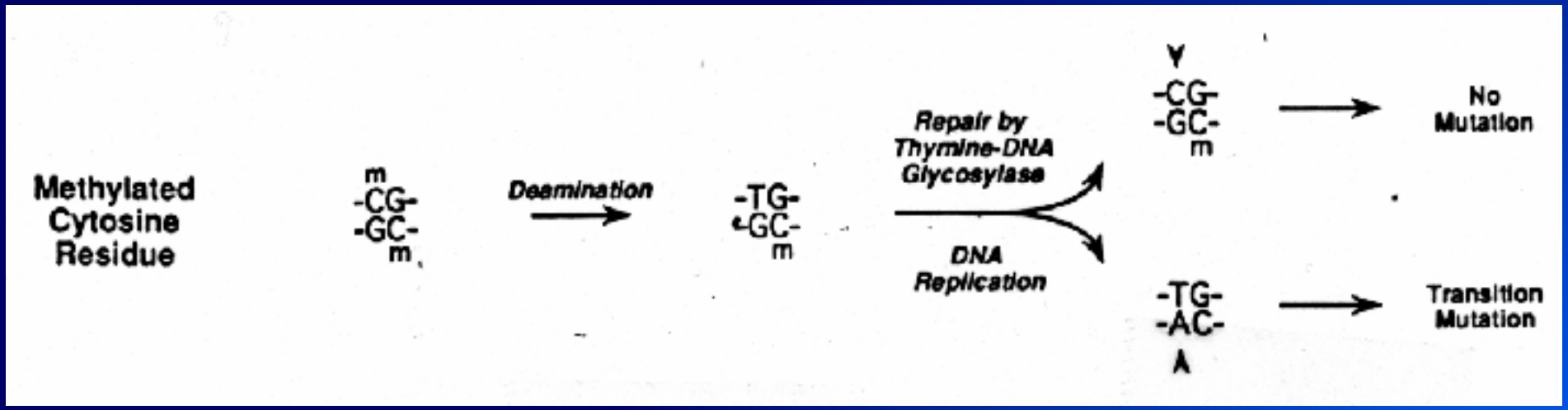
DNA Methylation – basic points

- Critical to normal development; tightly regulated
- Occurs almost exclusively at CpG dinucleotides
- Uses s-adenosyl methionine (SAM) as methyl donor
- CpGs in repetitive elements highly methylated
- CpG islands generally unmethylated
    - Exceptions are the inactive X chromosome and imprinted genes

Number of methylation targets in the genome

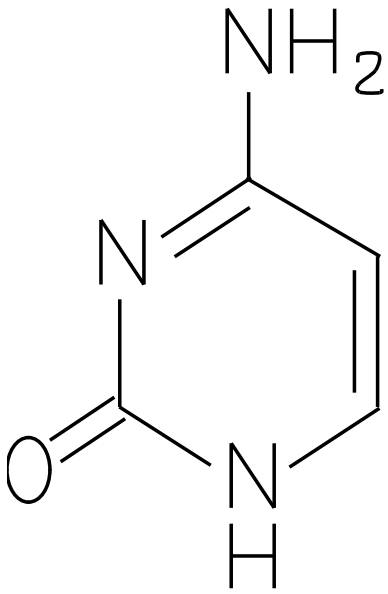
- There are 16 possible dinucleotides
- CpG should make up ~1/16th of the genome
  - However, only about 1/80th is CpG

DNA Methylation and Mutation

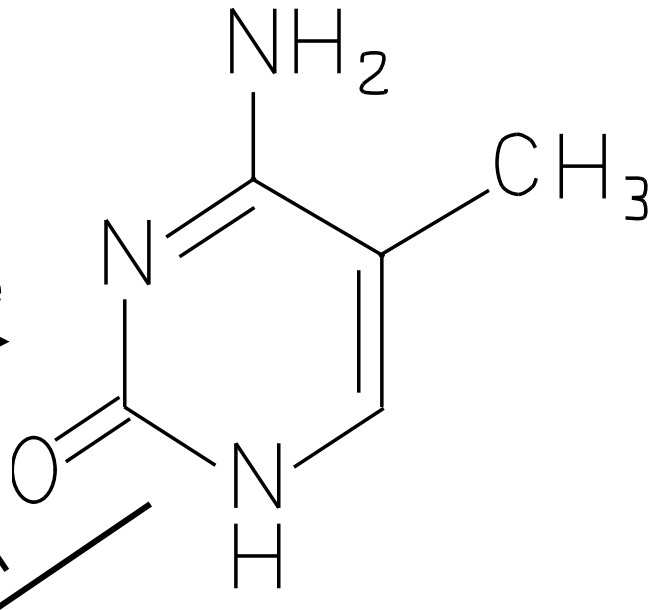


50% of p53 mutations involve C:T transitions at CpG dinucleotides

Cytosine



5-Methylcytosine

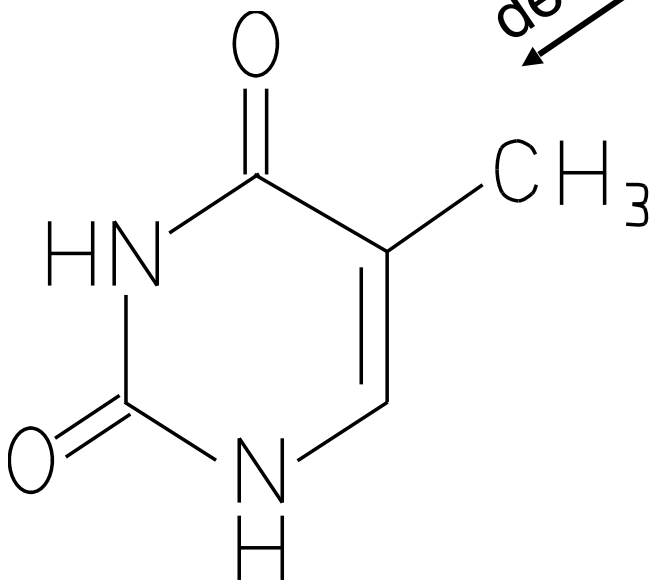


DNA methyltransferase

S-adenosyl methionine

deamination

Thymine



- Because of this deamination reaction, C:T transitions occur at a 10 – 40 fold higher rate than others

- Reduced the amount of CpG dinucleotides to ~ 5x less than expected

Distribution of methylation targets in the genome

- ~90% are found scattered throughout repetitive elements
  - mostly methylated
  - Reduces risk of deletion by recombination
  - Reduces insertional mutation by limiting transposon activity
  - Important for centromere stability

Repetitive element silencing

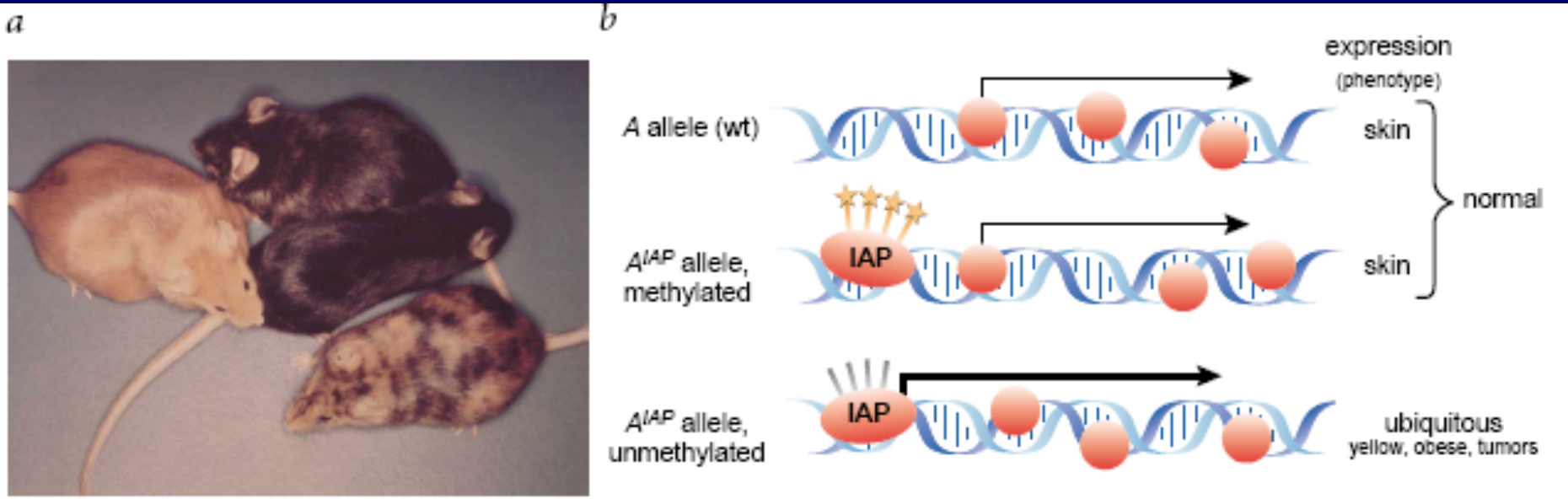


Fig. 2 Epigenetic effects on mouse coat color. *a*, Mice segregating the A^{hvy} and *a* alleles show variegated coat color. The A^{hvy} allele[186], as well as other dominant *A* alleles, is formed by the insertion of IAP into the *agouti* locus[166–168]. These alleles are also designated as A^{IAP}. Reproduced with permission from ref. 168. *b*, The methylation status of the IAP element determines expression of the *agouti* gene. When the element is methylated, the gene is expressed only in the skin, similar to expression of the wildtype allele. Hypomethylation of the element generates an ubiquitously expressed transcript that causes the yellow coat color, obesity and tumors.

- Methylation silences the strong IAP promoter
- Lack of methylation allows use of IAP promoter
  - High level ubiquitous expression of Agouti gene

Jaenisch and Bird Nat Genet Sup 33:245

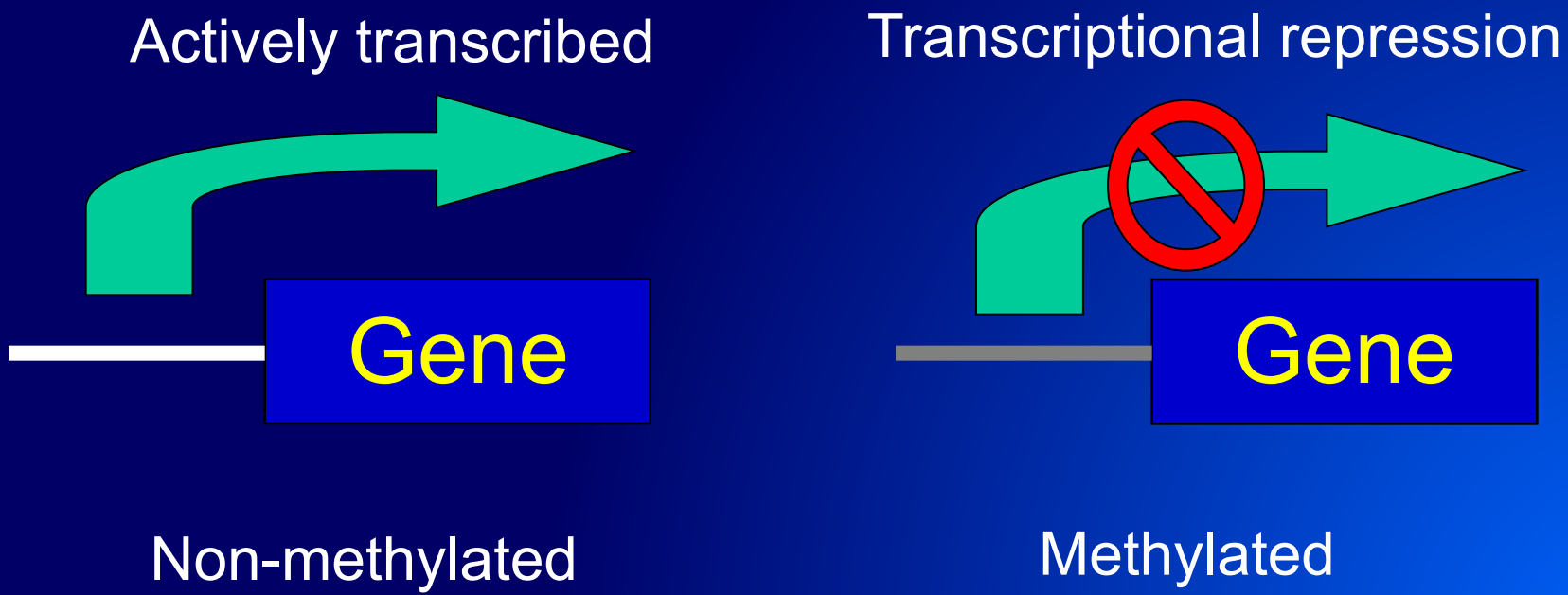
Distribution of methylation targets in the genome

- 10% are found in CpG islands
- CpG islands are short stretches with almost the expected frequency of CpG
  - Mostly unmethylated
  - Mainly found in the promoter regions of genes

- Why have CpG islands not been depleted of CpGs by deamination over evolutionary time?
  - Because they were protected from methylation
  - Probably because they are important
  - Suggests we probably should study them

Consequence

-Methylated CpG islands associated with a closed chromatin structure



DNA Methylation – cancer

- Genome wide hypomethylation of normally methylated CpG dinucleotides throughout the genome
- Tumor suppressor genes can be inactivated by CpG island hypermethylation in their promoters
  - Analogous to loss of function mutations (genetic means)

Transcription, Acetylation, and Methylation

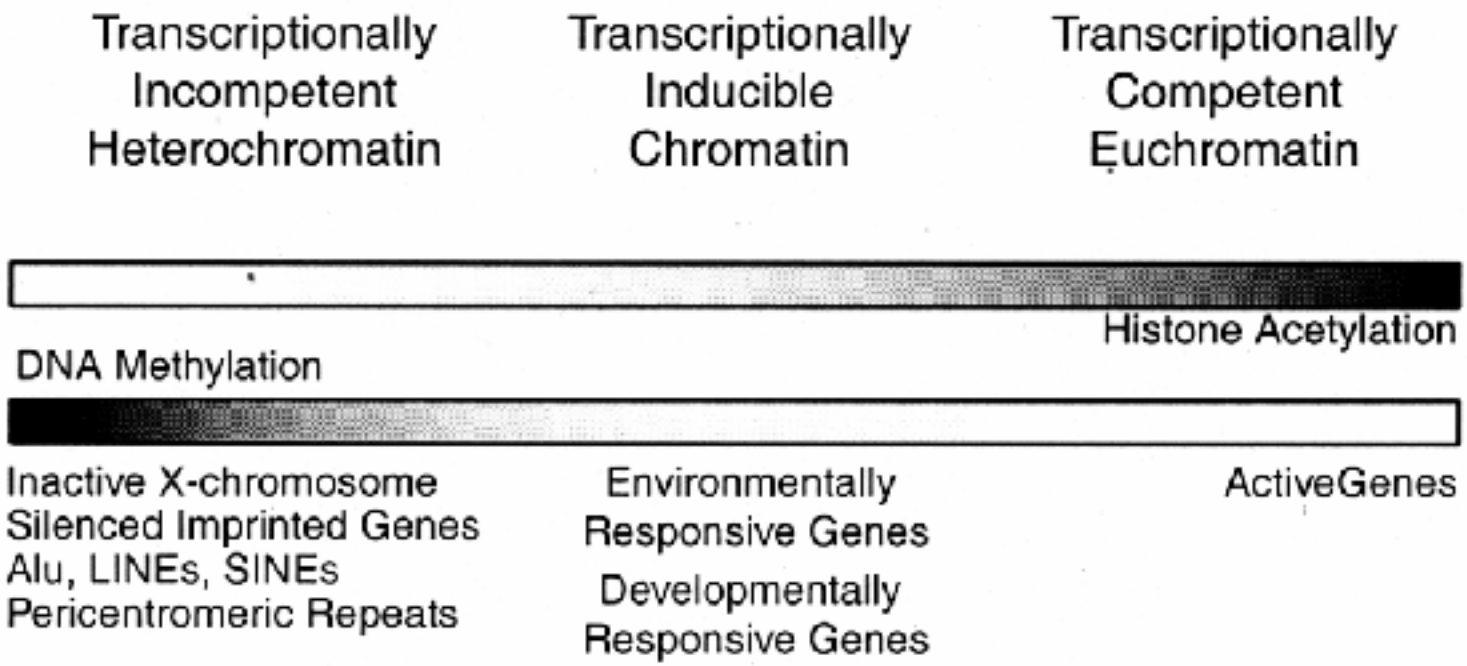


Figure 1 The transcriptional rheostat. DNA methylation and histone acetylation help to establish chromatin states that either foster or inhibit transcription. The shaded bars represent the inverse correlation between these two epigenetic modifications. While a cell will use histone acetylation status to modulate gene expression, DNA methylation primarily serves as a transcriptionally repressive 'lock'

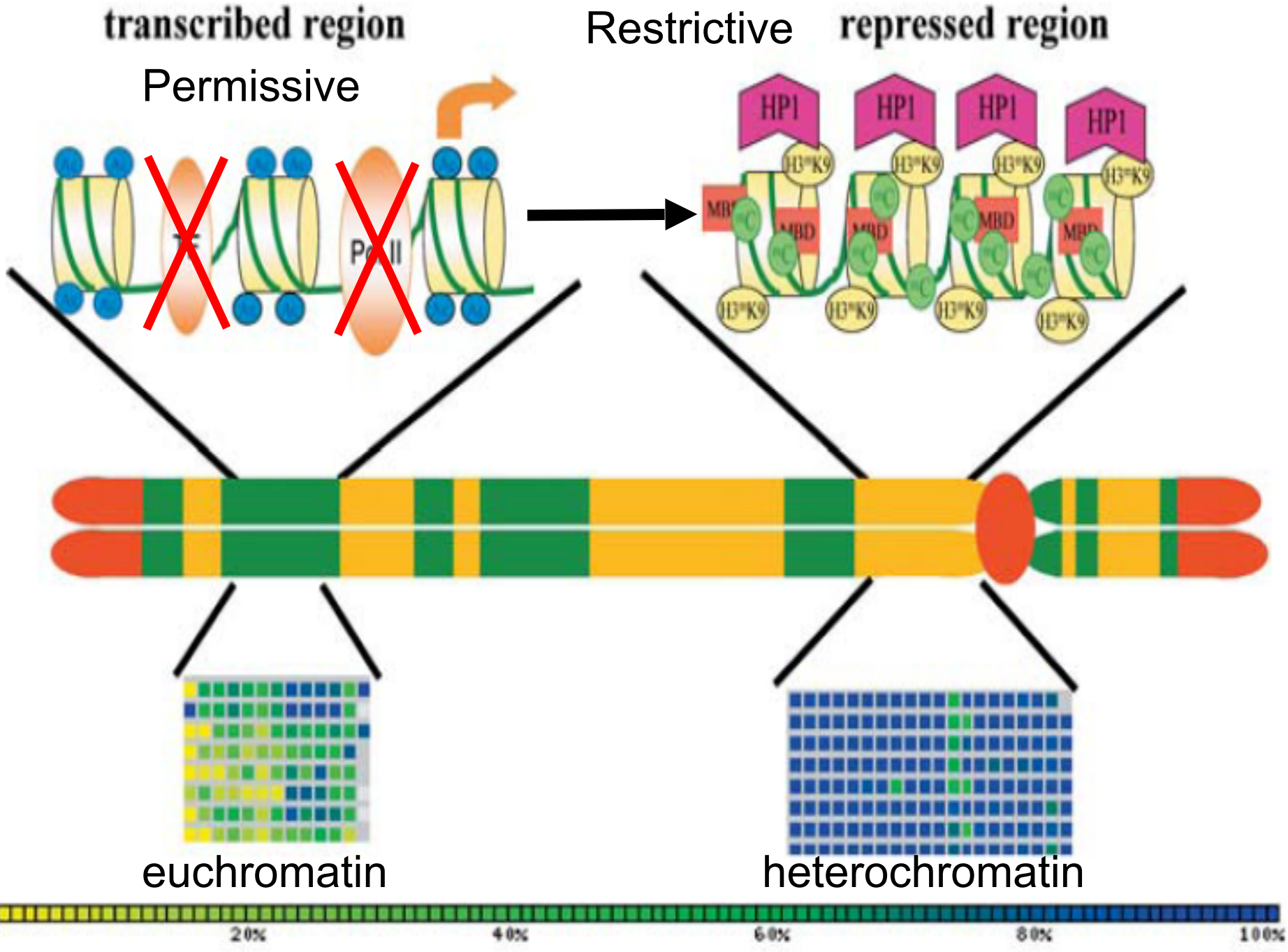
Key components

- Methyl binding domain proteins (MBDs 1-4 and MeCP2)
  - Bind only methylated CpG
  - Recruit protein complexes with HDAC and transcription repression activity

- Histone methylases (Suvar39)
  - H3K9me (histone H3 Lysine 9 methylation)

- HP1
  - Heterochromatin binding protein 1
  - Provides a crucial link between histone methylation and DNA methylation

Herman et al Cell Mol Life Sci 61:2571



Summary 1

- DNA methylation is one type of epigenetic modification of DNA
- Tightly regulated, critical for normal development
  - Imprinting
  - X-inactivation
  - Centromeric stability
  - Silencing of retroviral/IAP promoters
  - Reduces potential for somatic recombination and insertional mutation by transposon activity
- CpG islands are generally unmethylated
  - Often promoter regions
  - Allow for expression
- Regulation breaks down in cancer
  - Hypomethylation of repetitive elements
  - Hypermethylaion of CpG islands

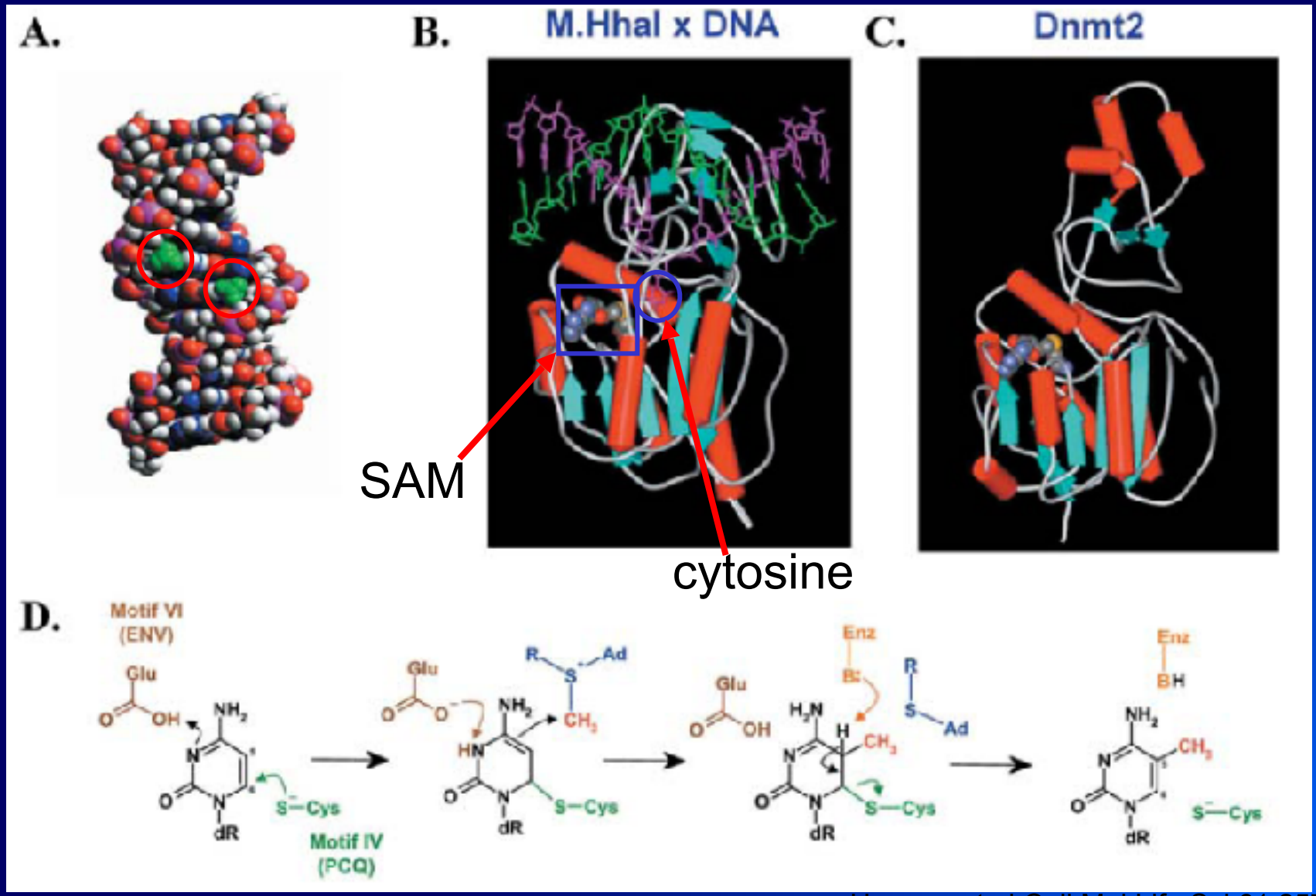
Summary 1

- DNA methylation tightly linked with histone modifications
  - Induces and is induced by H3K9me
  - Induces histone deacetylation

- Contributes to setting up euchromatic/heterochromatic regions
  - Euchromatin
    - Lack of DNA methylaiton
    - Acetylated histones
    - Permissive to gene expression
  - Heterochromatin
    - High degree of DNA methylation
    - Deactylated and methylated histones
    - Restrictive to gene expression

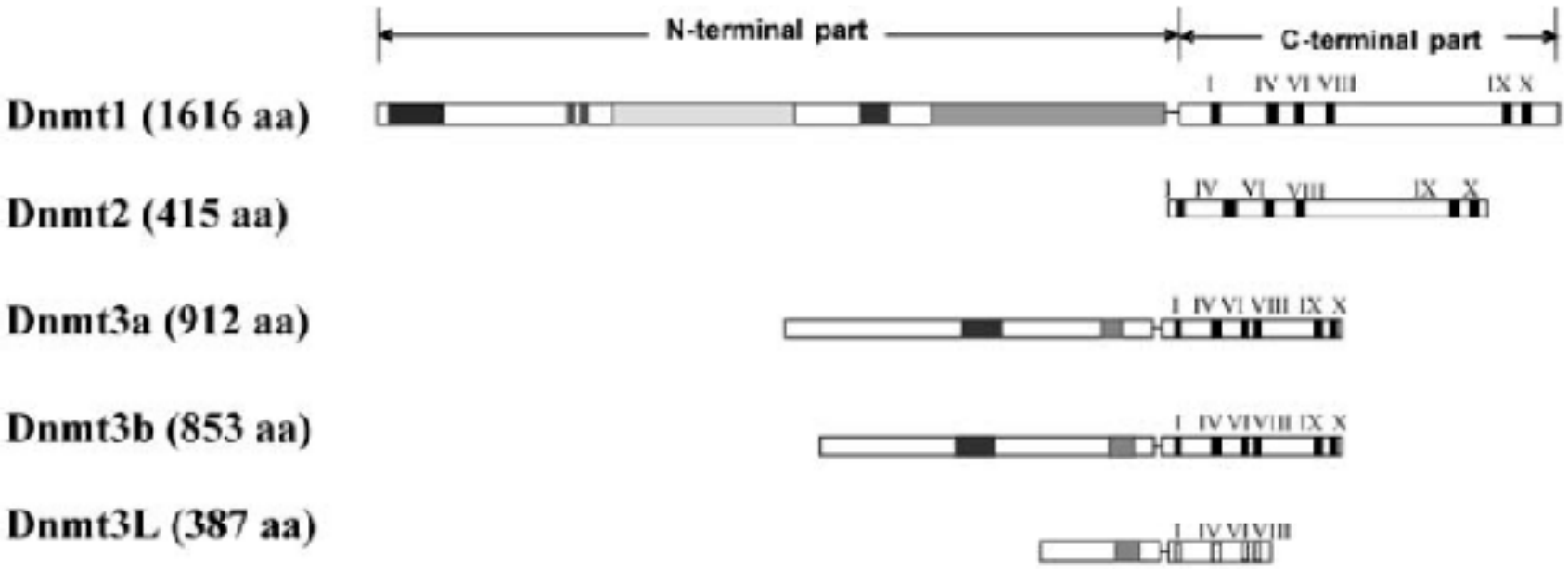
Nitty-gritty

Mammalian DNA methyltransferases



Herman et al Cell Mol Life Sci 61:2571

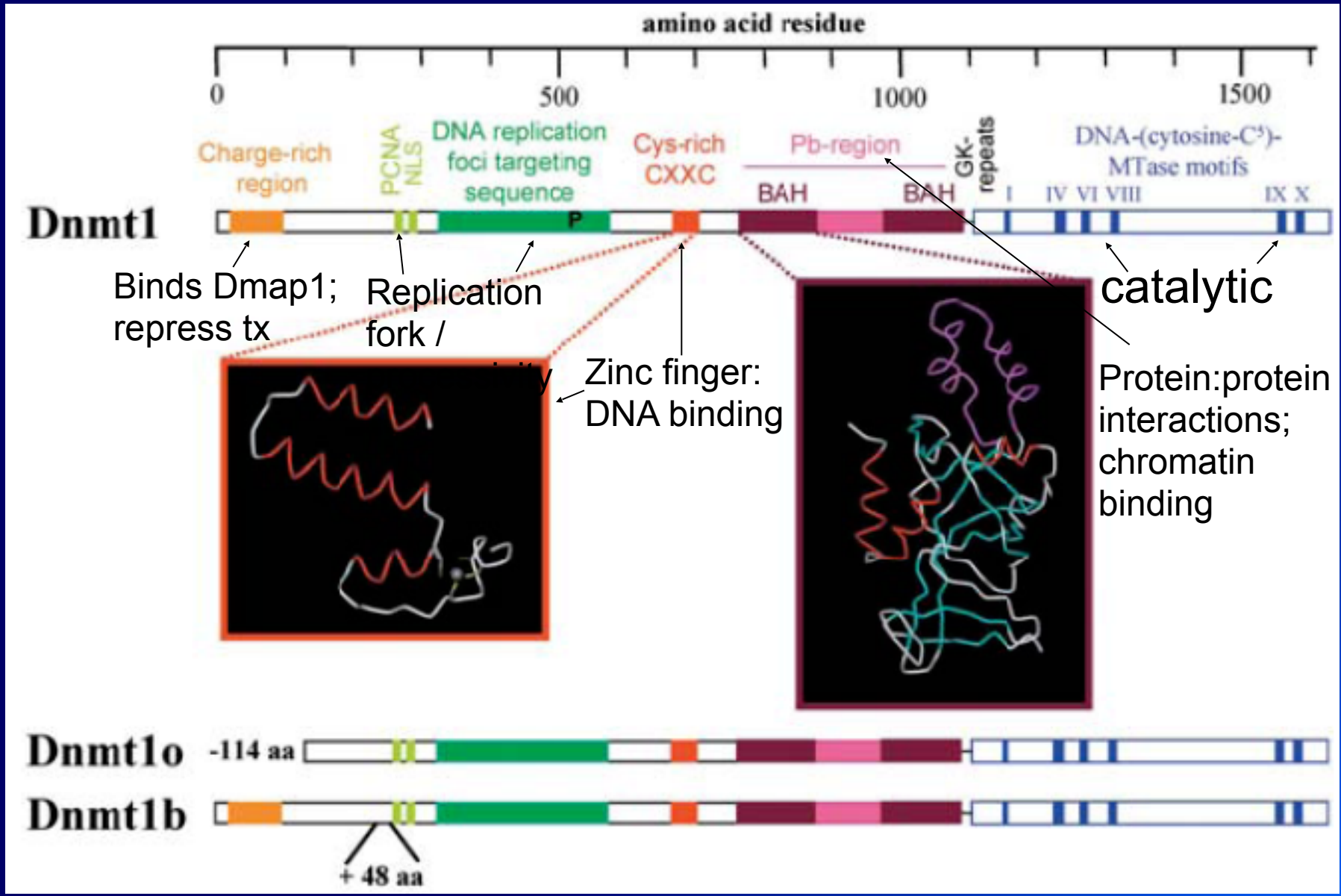
Mammalian DNA methyltransferases



- C-term portions have strong homology to M.HhaI methylase
  - Business end of the enzyme - catalyic activity
- N-term portions are regulatory
  - Protein-protein interactions
  - DNA binding

Herman et al Cell Mol Life Sci 61:2571

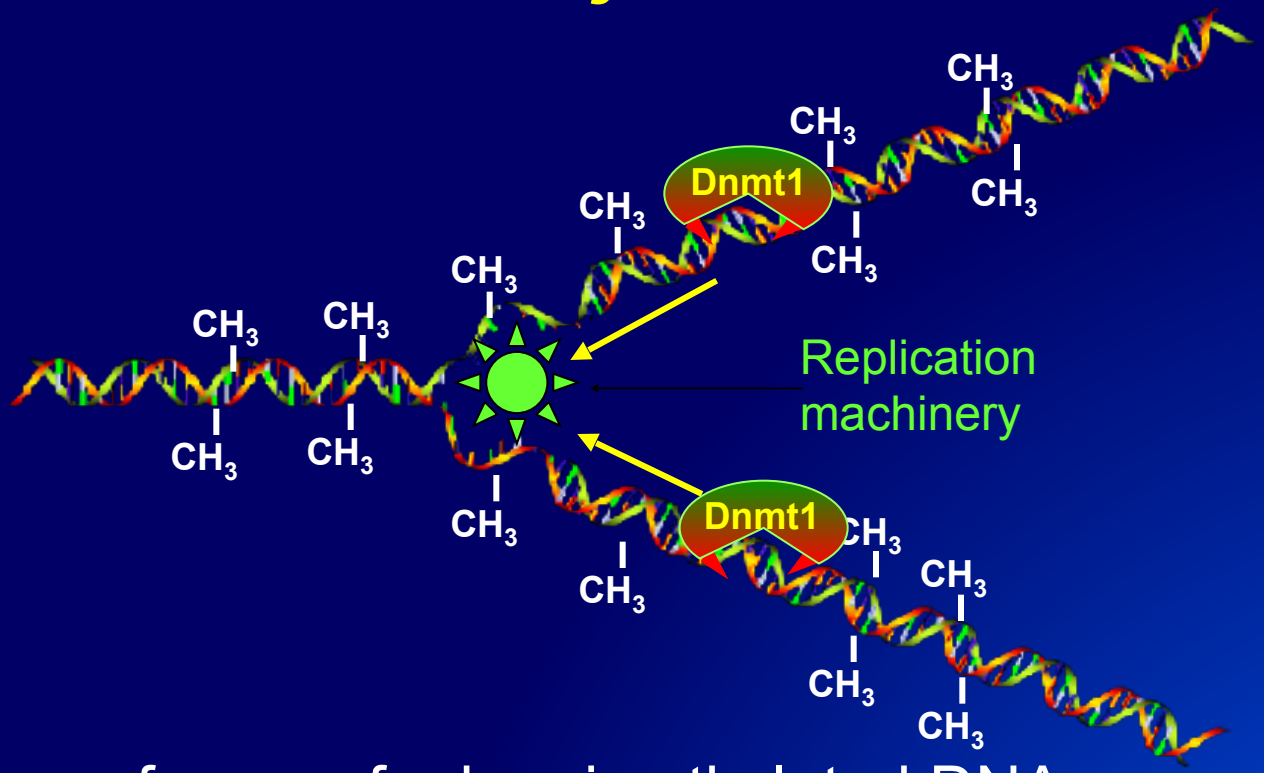
Dnmt1 - *keeps on keepin' on*



The maintenance methyltransferase

Herman et al Cell Mol Life Sci 61:2571

Dnmt1 major function



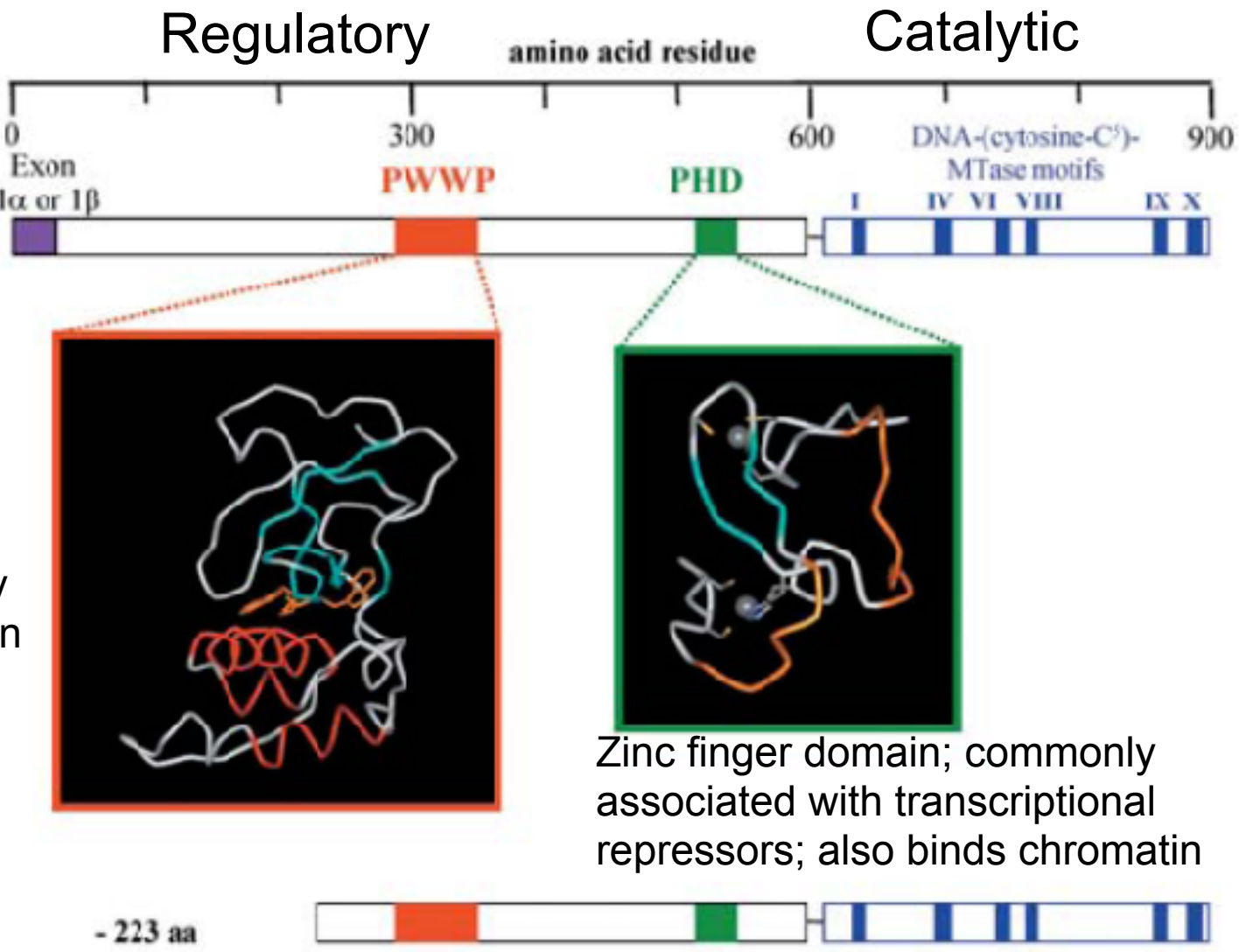
- Strong preference for hemimethylated DNA
  - Methylates nascent strand at positions where template strand is methylated
  - Highly processive
  - High affinity for cruciform structures *in vitro*
    - Will *de novo* methylate

Dnmt1 has lots of binding partners

- PCNA – replication fork/ processivity factor
- p21WAF- CDK inhibitor (cell cycle regulator)
- RB – cell cycle regulator
- E2F1 – transcription factor
- SUV39H1 – histone methyltransferase
- HDAC1+2 – histone deacetylases
- MBD2+3, MeCP2 – methyl binding domain proteins
- HP1 – heterochromatin binding protein (H3mK9)
- Dnmt3a+3b – *de novo* DNA methyltransferases
- Dmap1 – transcriptional repressor

- Perhaps different partners in different settings
  - S-phase vs. G_0
  - Maintenance vs. *de novo* vs. transcriptional repression

Dnmt3a, 3b – *the 'do over' enzymes*



Strongly basic region; help with DNA binding; may target to chromatin

Zinc finger domain; commonly associated with transcriptional repressors; also binds chromatin

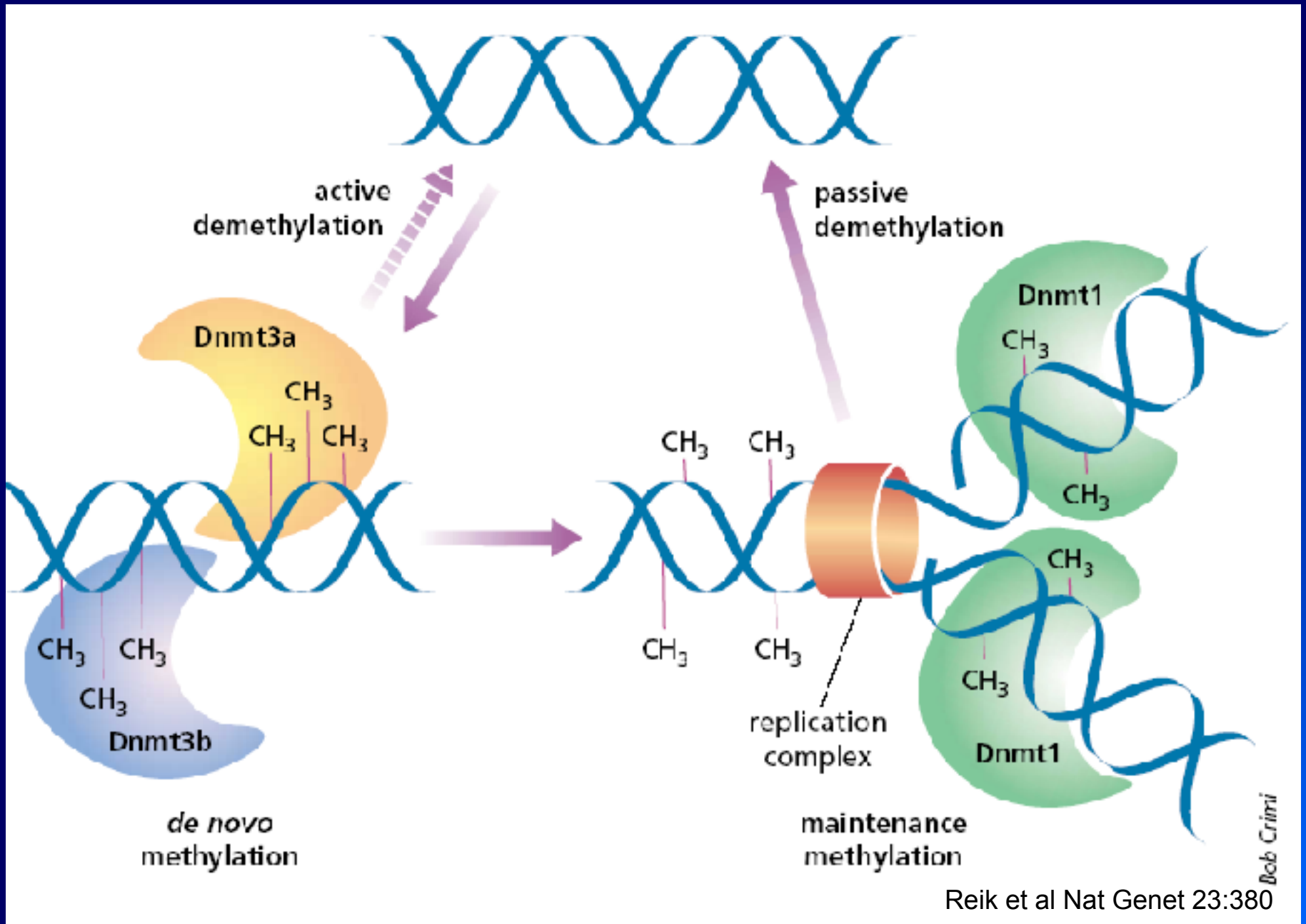
The *de novo* methyltransferases

Herman et al Cell Mol Life Sci 61:2571

Dnmt3a, 3b differences from Dnmt1

- No preference for hemimethylated DNA
  - *de novo* methylation ability
  - Expression of 3a in Drosophila caused methylation
  - Dnmt1 can be stimulated into *de novo* activity by methyl CpGs in *cis and interaction with 3a or 3b*
    - May account for methylation 'spreading'
- Expression
  - Dnmt1 expressed in nearly all cells, all time points
  - 3a, 3b expressed high levels in early embryogenesis and gametogenesis
    - Expressed at lower levels in adult cells

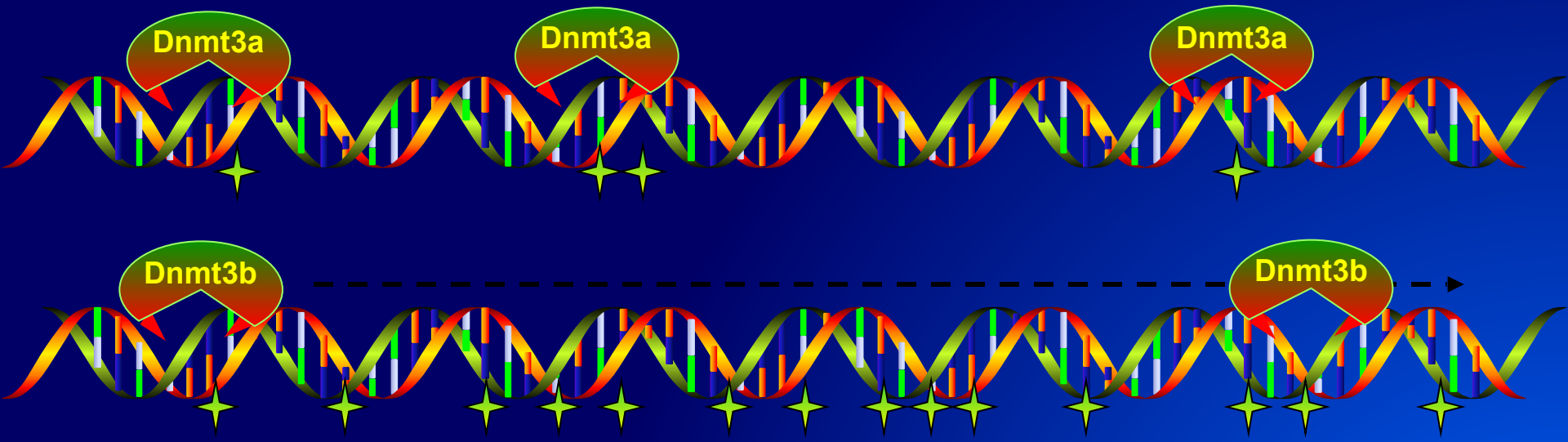
The methylation cycle



Reik et al Nat Genet 23:380

Form dictates function

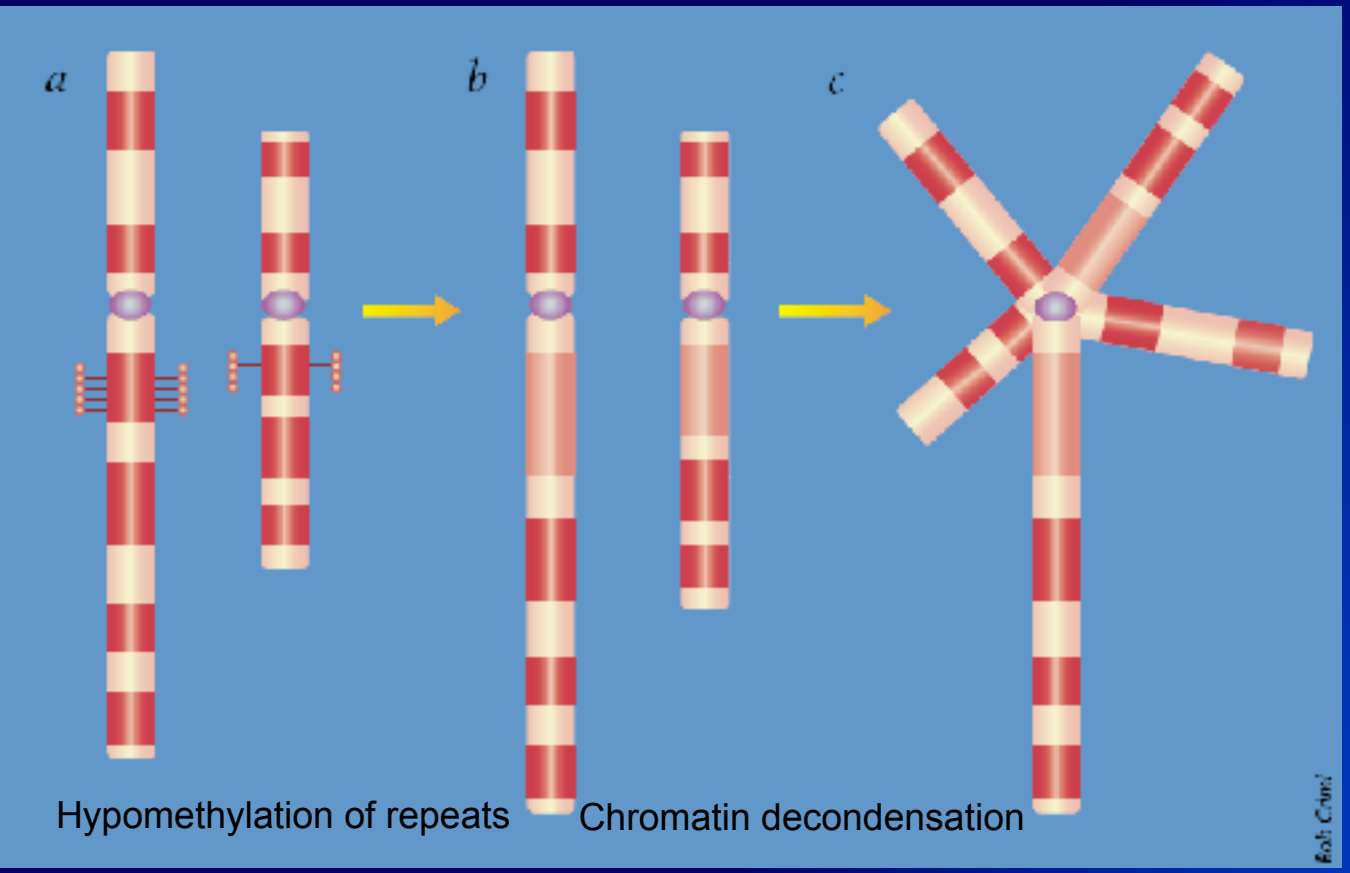
- Small differences in the catalytic sites
  - Dnmt3a is a distributive enzyme
  - Dnmt3b is a processive enzyme
    - Higher intrinsic activity



- Dnmt3a methylates single copy genes
  - Targeted small regions
- Dnmt3b methylates pericentromeric regions
  - Long stretches with high CpG content

Dnmt3b mutations

- ICF syndrome (immunodeficiency, centromere instability, and facial anomalies)
  - Hypomethylated pericentromeric regions.



Involves chr 1, 9 and 16

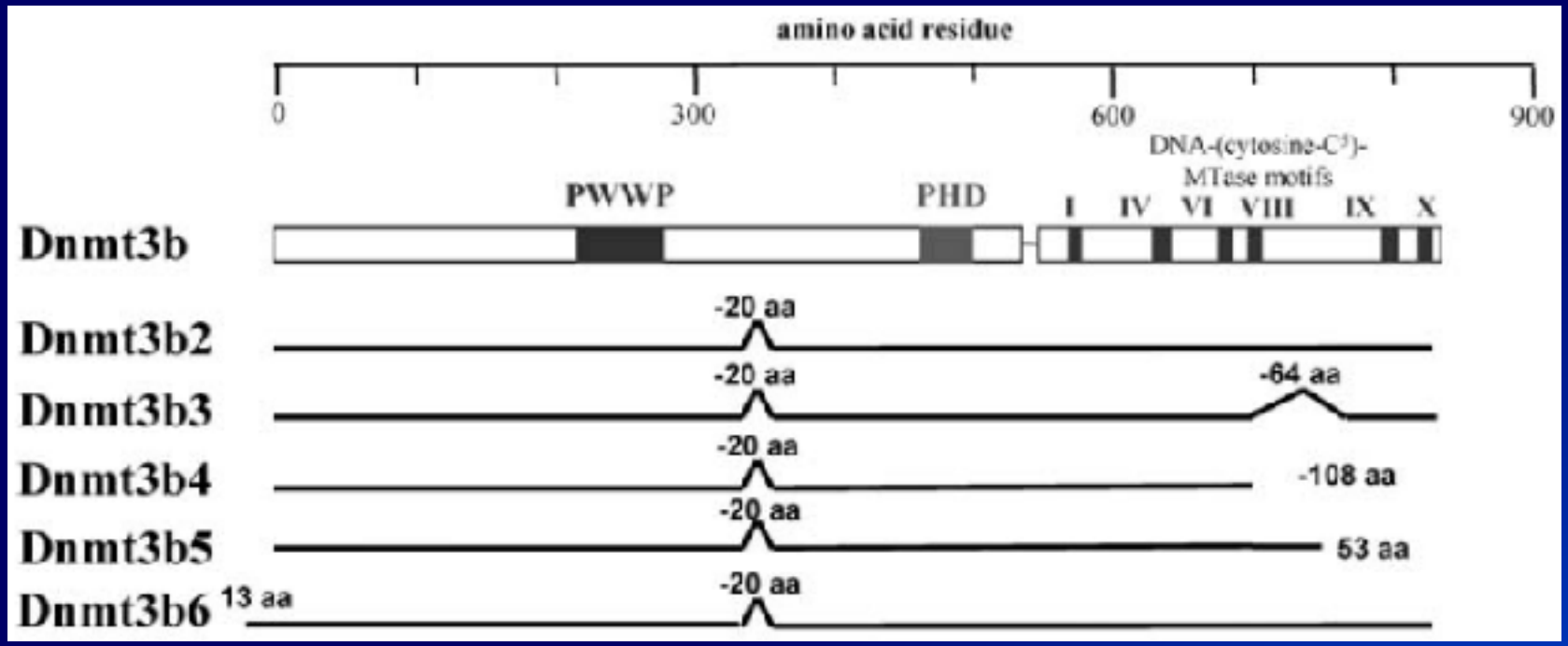
Reik et al Nat Genet 23:380

Dnmt3a isoforms

- Dnmt3a – long form
  - Major isoform in adult tissues
  - Targeted to heterochromatin

- Dnmt3a2 – short form
  - Missing the N-term 223 amino acids of the long form
  - Major form during embryogenesis
  - Targeted to euchromatin
    - Fits with known role in focused methylation of single copy genes

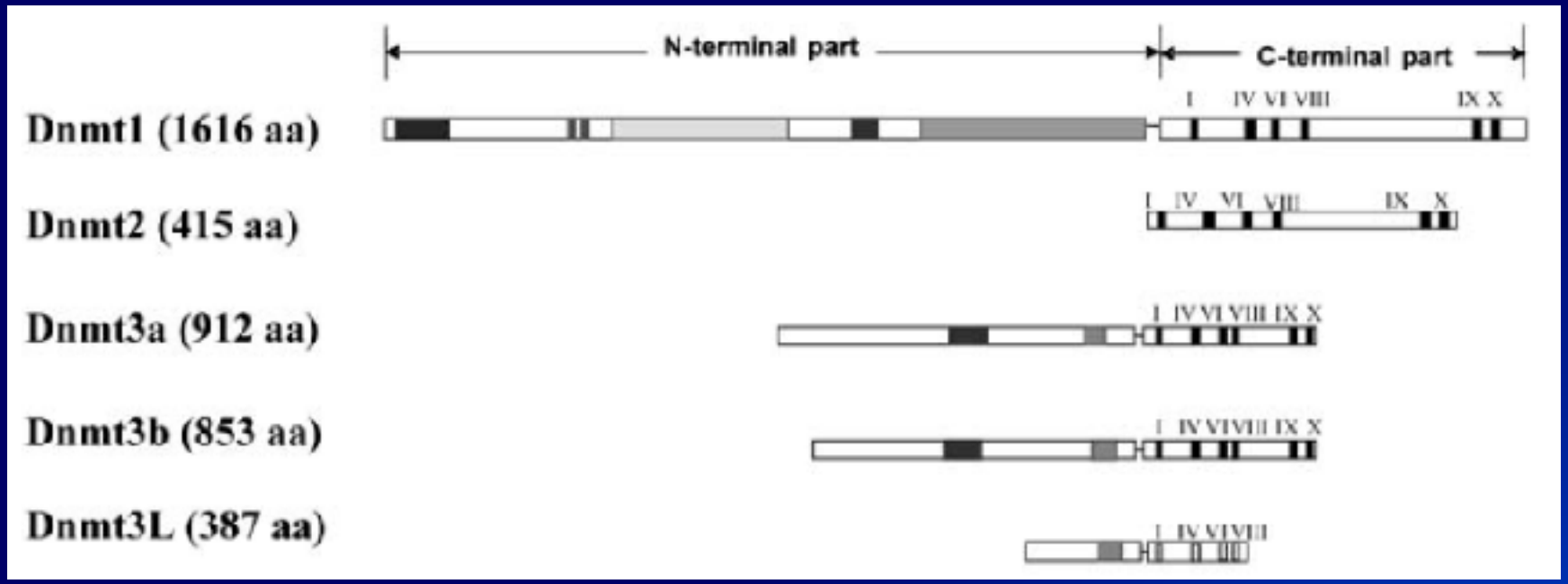
Dnmt3b isoforms



- 3b and 3b2 are enzymatically active
- 3b3, 3b4, 3b5 are enzymatically inactive
  - Still targeted to pericentromeric regions
  - May have regulatory role
- Little known about 3b6

Herman et al Cell Mol Life Sci 61:2571

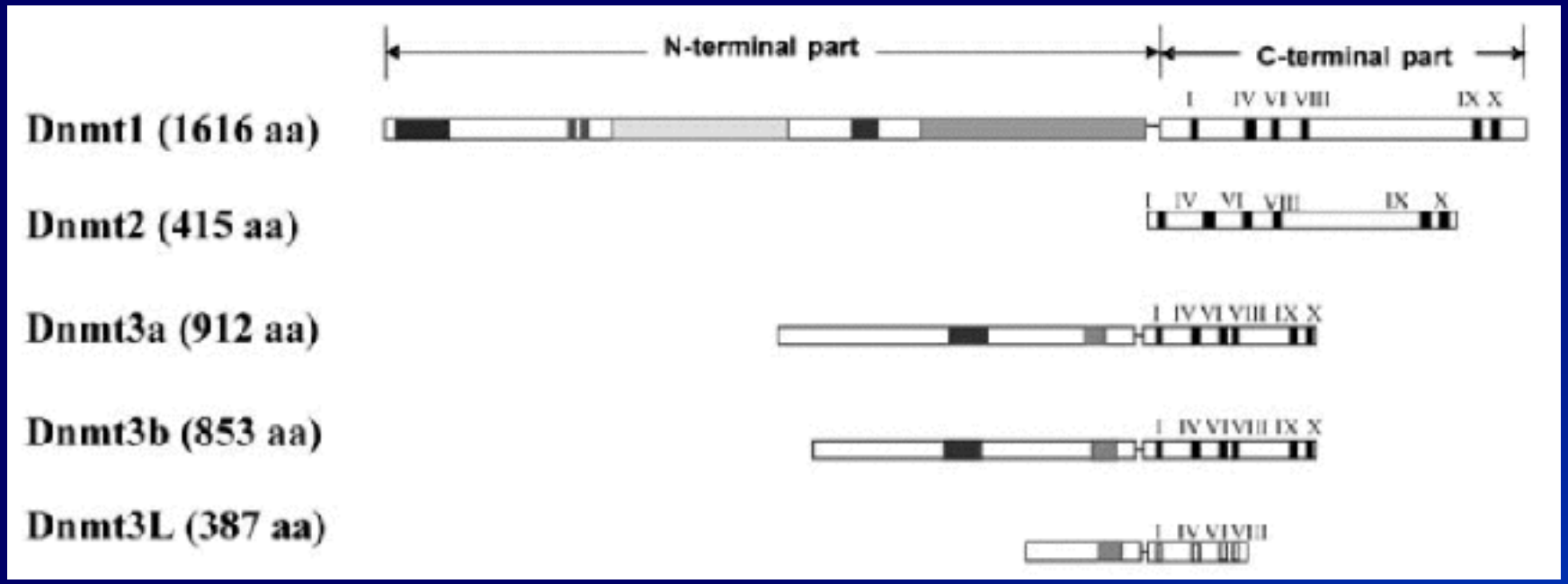
Dnmt3L – *Mother's little helper*



- Lacks PWWP domain targeting to chromatin and DNA binding
- Retains PHD domain associated with transcriptional repressors
- Lacks enzymatic activity
  - Conserved motifs have inactivating point mutations
- Interacts with Dnmt3a
  - Stimulates the rather low enzymatic activity
  - Required for maternal imprinting

Herman et al Cell Mol Life Sci 61:2571

Dnmt2 – *the ugly stepchild*



Herman et al Cell Mol Life Sci 61:2571

- Catalytic domain; no regulatory domain
- Ubiquitously expressed
- No *in vivo* activity
- Recently very low level *in vitro* activity found

Recent evidence suggests it methylates RNA

Possibly, interaction with an unknown protein will stimulate activity

Summary 2

- Dnmt1 is major DNA methyltransferase in adult somatic cells
  - High affinity for hemi-methylated DNA; maintenance
  - Found at the replication fork
  - Highly processive; binds PCNA
  - Can *de novo* methylate
  - Binds many proteins involved in gene regulation and chromatin formation
- Dnmt3a, 3b are the major *de novo* methyltransferases
  - No special affinity for hemi-methylated DNA
  - High expression in early embryo and germ cells
  - Dnmt3a is distributive and used for focal methylation in single copy genes
  - Dnmt3b is processive with higher activity
    - Methylation of repeats in pericentromeric regions
    - Mutations cause ICF syndrome; expanded chromatin structure near centromeres

Summary 2

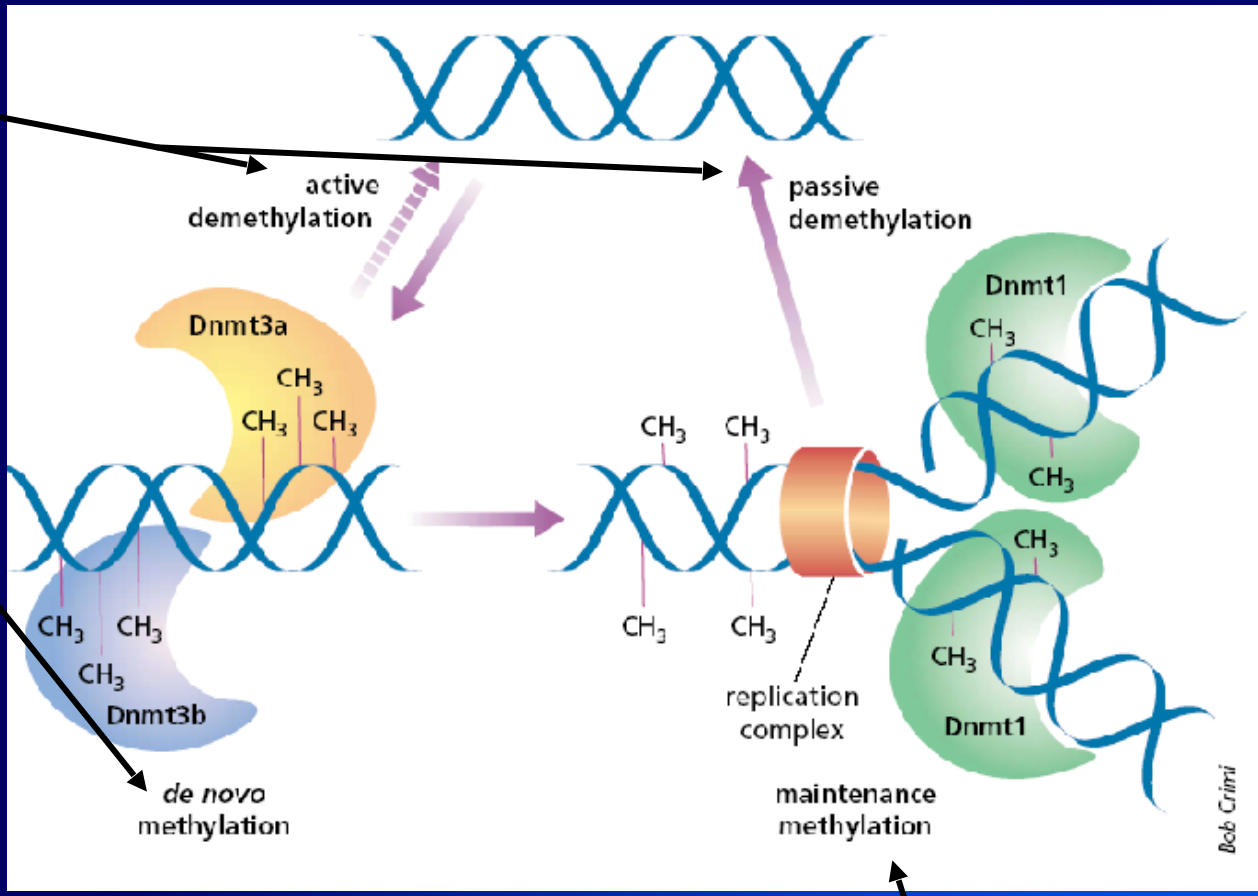
- Dnmt3L has no enzymatic activity
  - Non-consensus catalytic regions
  - Binds to Dnmt3a and Dnmt3b
  - Increases the enzymatic activity of Dnmt3a
  - Interaction with Dnmt3a required for maternal imprinting

- Dnmt2 has no enzymatic activity demonstrated *in vivo*
  - Appears to have functional catalytic region
  - No regulatory regions
  - Might require protein interaction to have activity

Begin the Begin

How does this process get started?

Early embryogenesis/ gametogenesis. Establish the pattern.



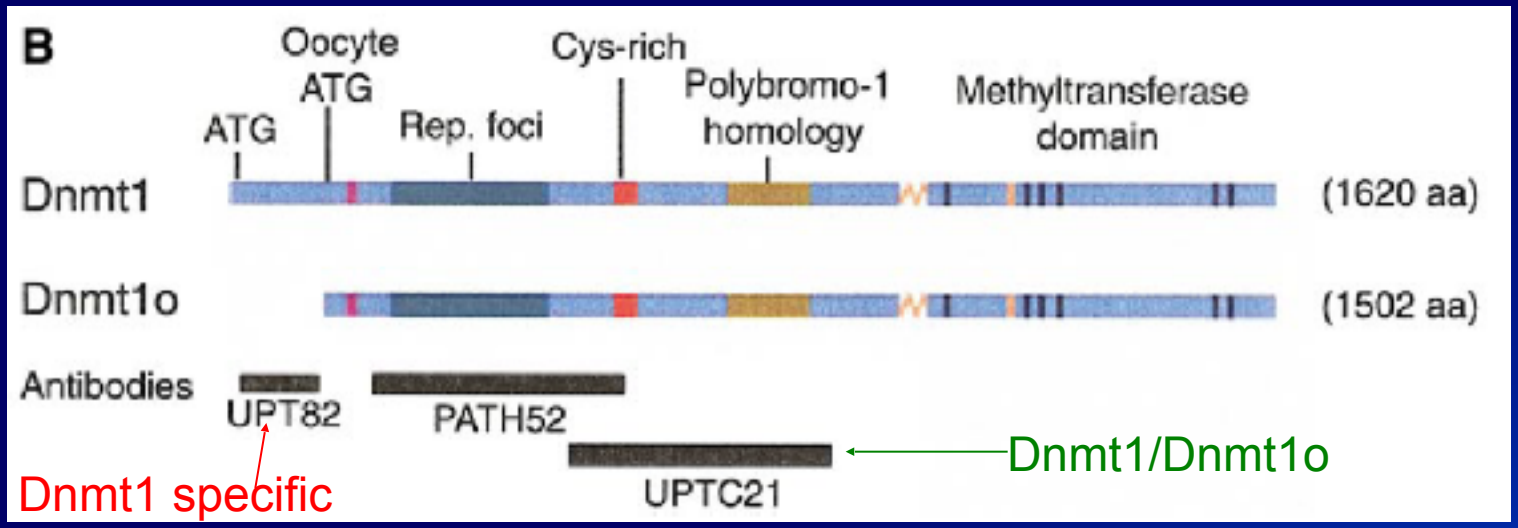
Maintain the pattern

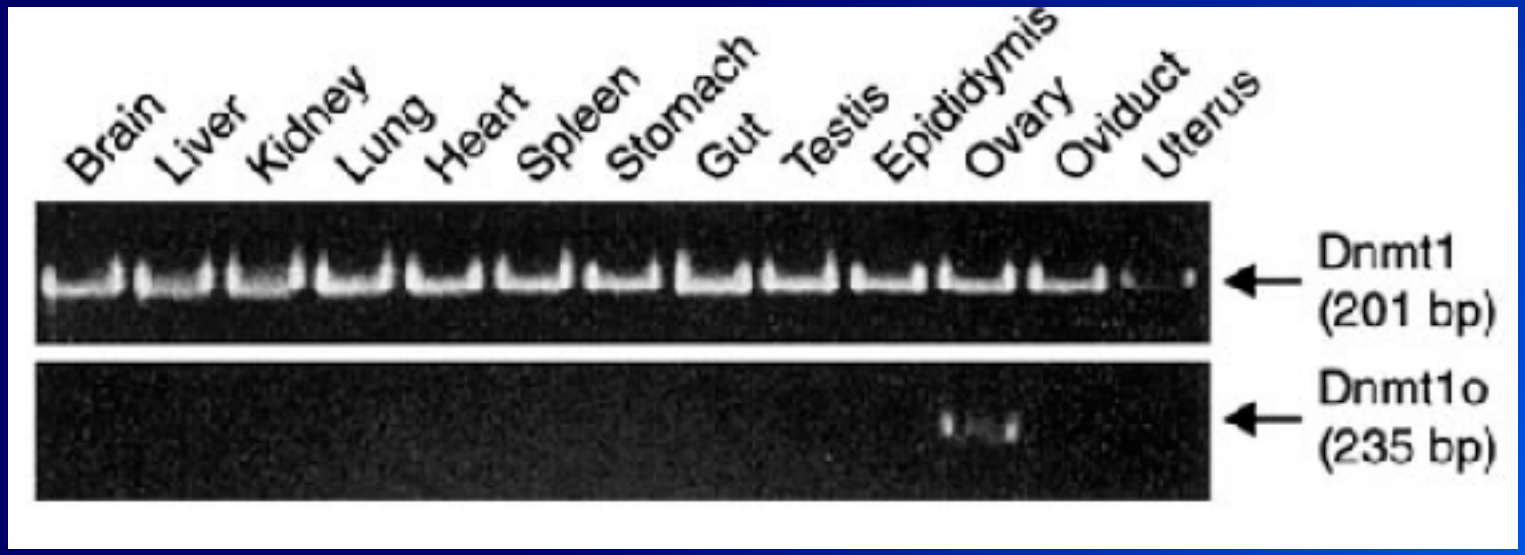
Demethylation

- Passive – replicating the genome without maintaining methylation on the nascent strand
  - Occurs in the female pronucleus and the early embryo up to implantation
    - Dnmt1 is locked away in the cytoplasm, later released into the nucleus

- Active – quick enzymatic removal of methyl groups without DNA replication
  - Exact mechanism is unclear
  - Occurs in male pronucleus
  - In somatic cells at other time points
    - Differentiation from stem cells

Dnmt1 distribution in oocytes and early embryo



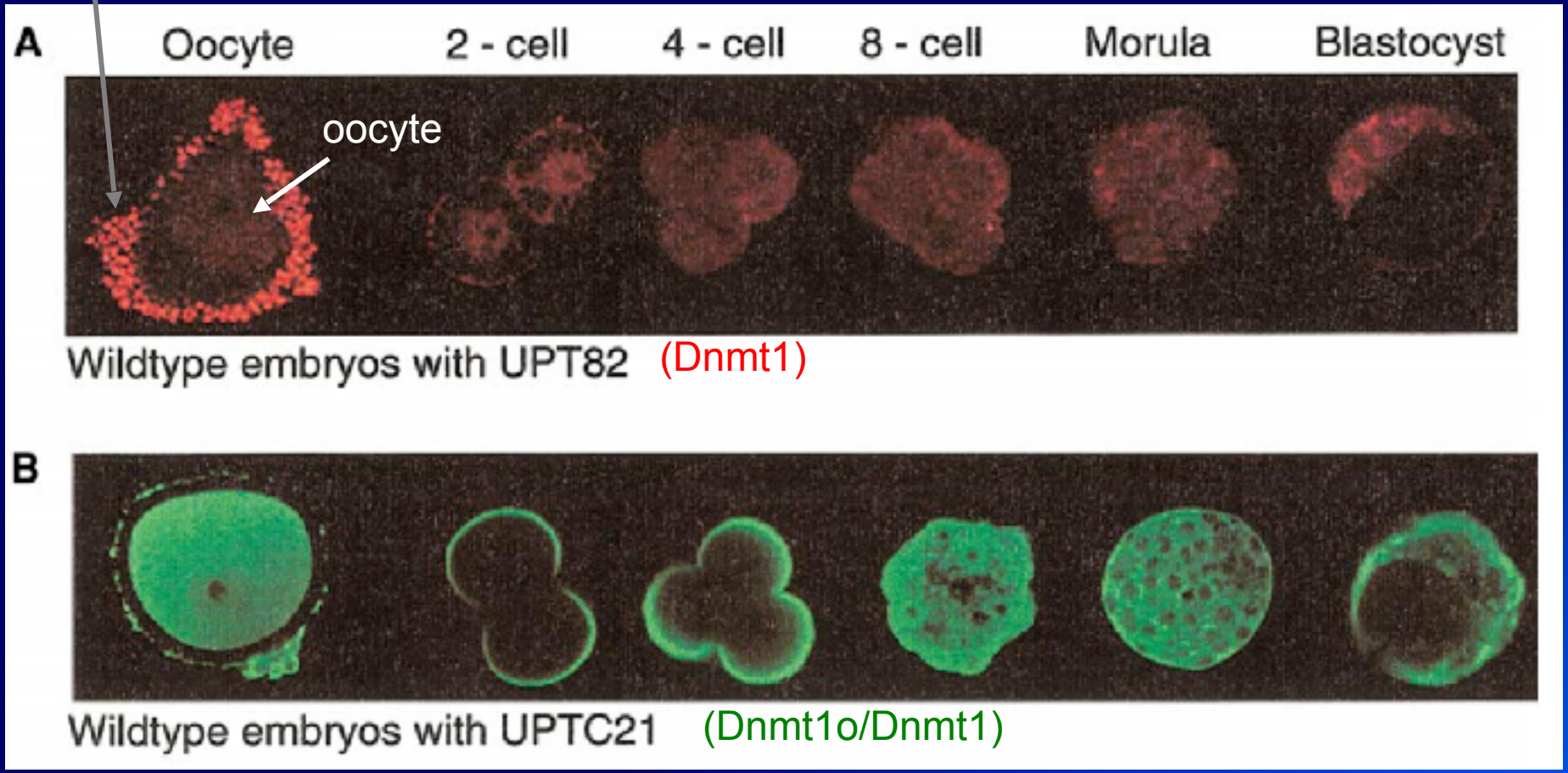


Ratnam et al. Developmental Biology 245:304-314

Dnmt1 distribution in oocytes and early embryo

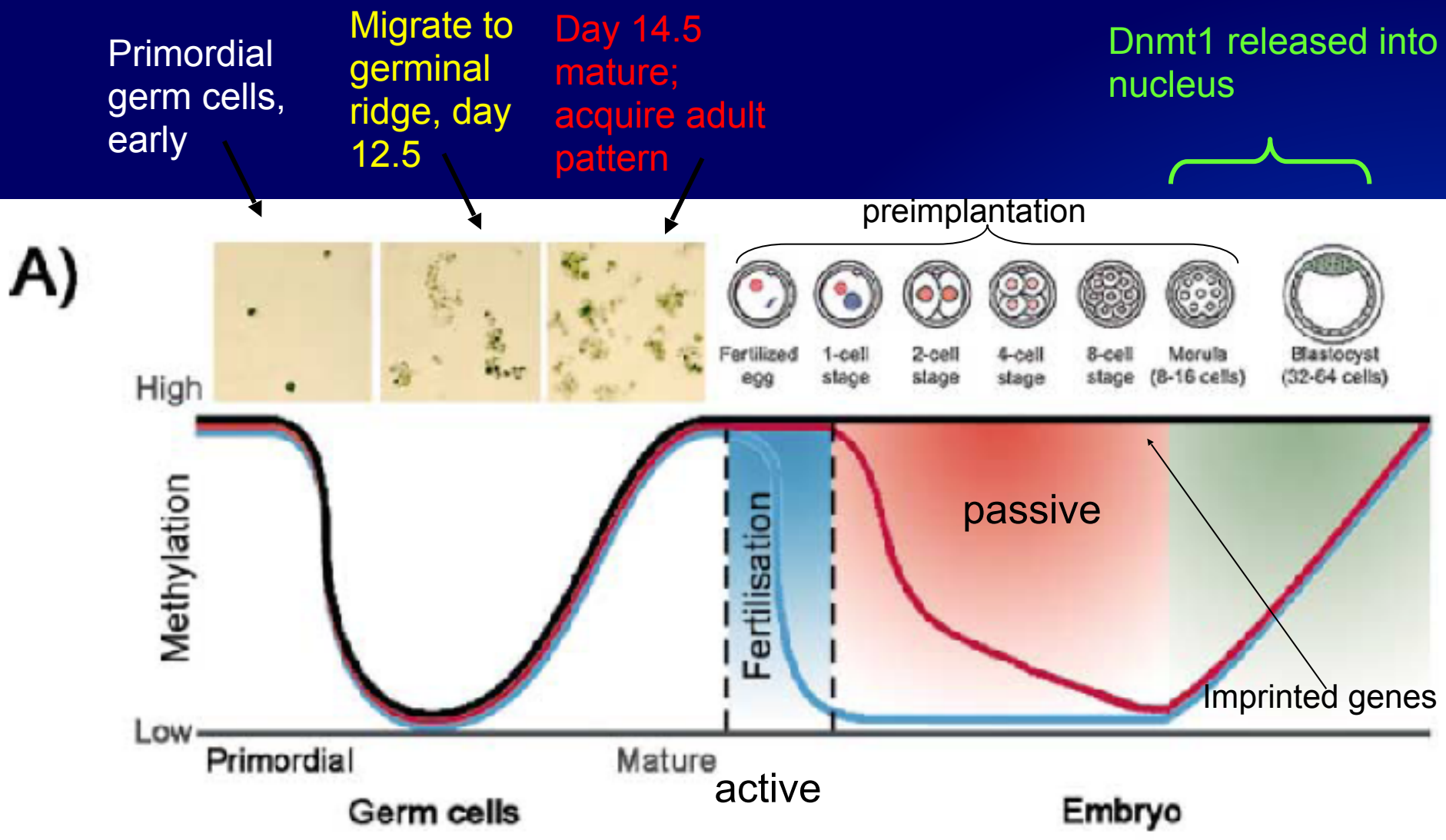
Somatic granulosa cells

Ratnam et al. Developmental Biology 245:304-314



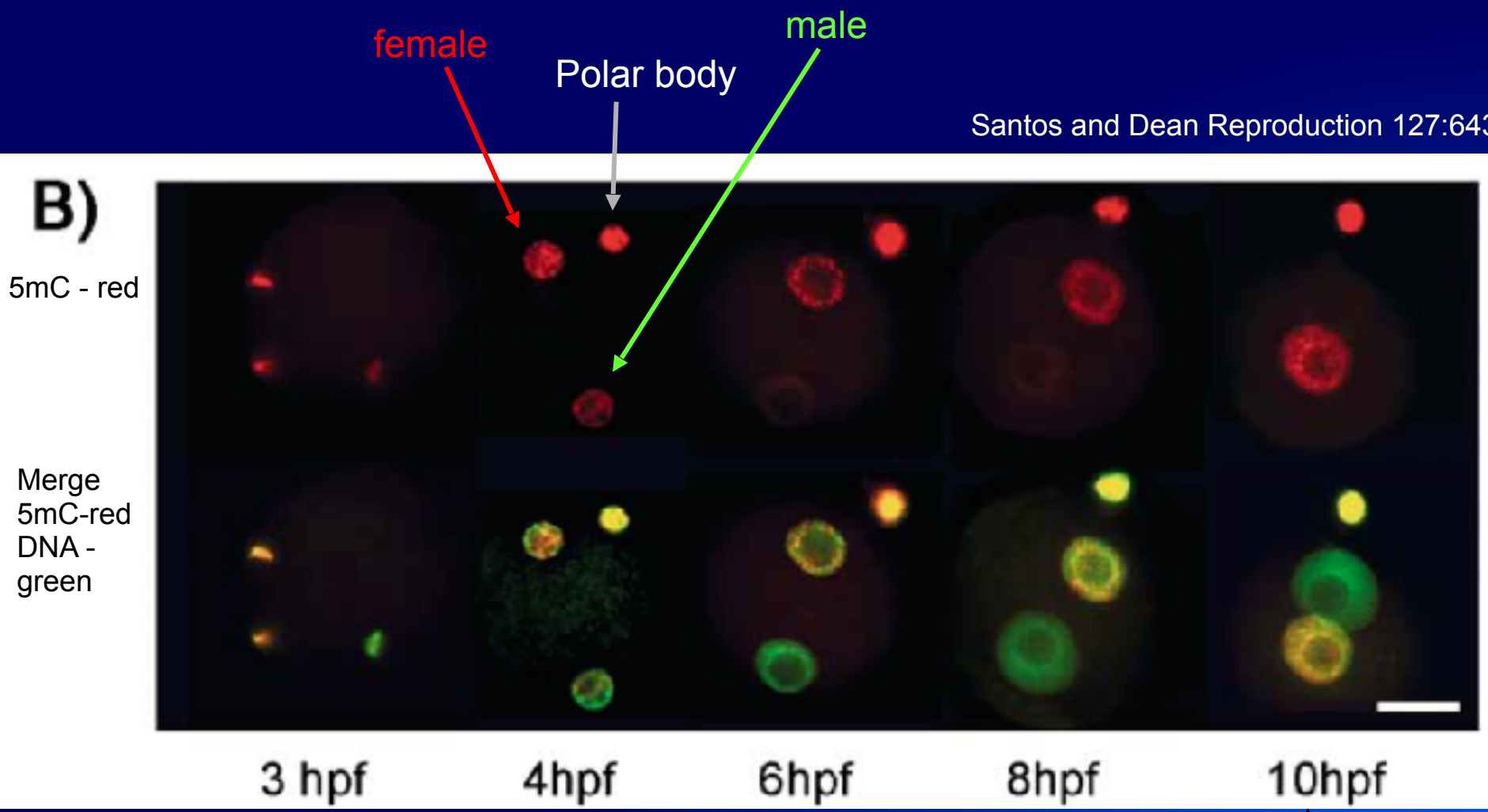
- Very little Dnmt1 up to morula stage
- Plenty of Dnmt1o
- Dnmt1o excluded from the nucleus up to morula stage
  - Dnmt1o binds cytoskeletal annexin V

DNA methylation in germ cells and early embryo



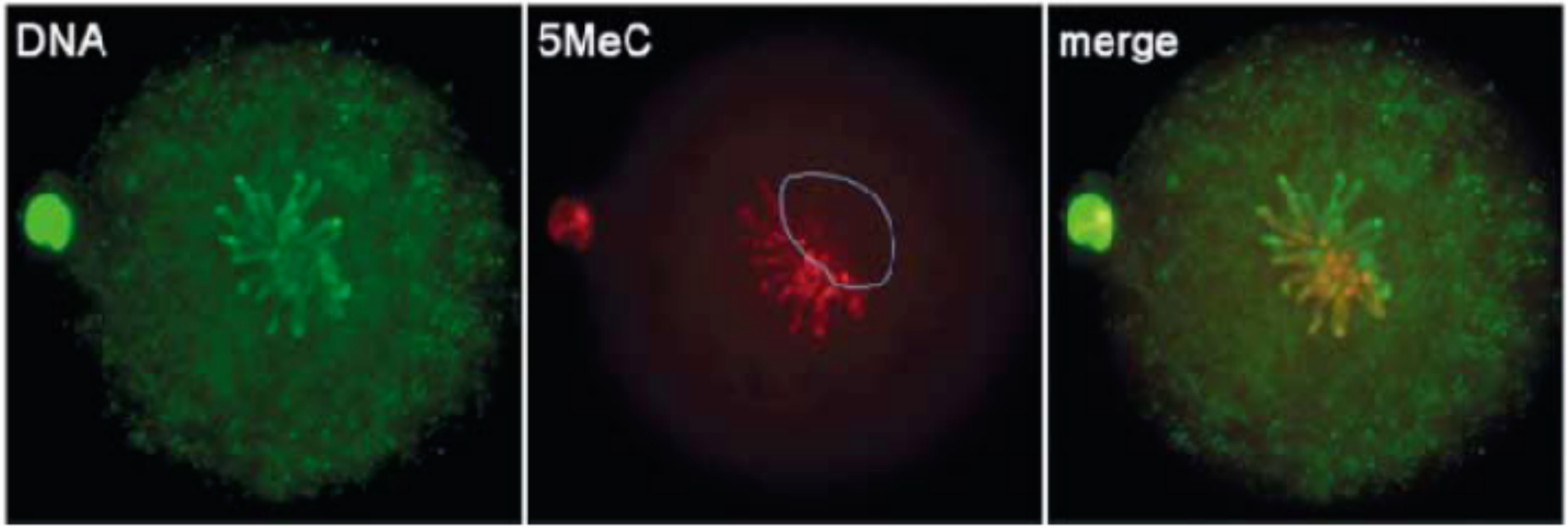
Santos and Dean Reproduction 127:643

DNA methylation in germ cells and early embryo



Santos and Dean Reproduction 127:643

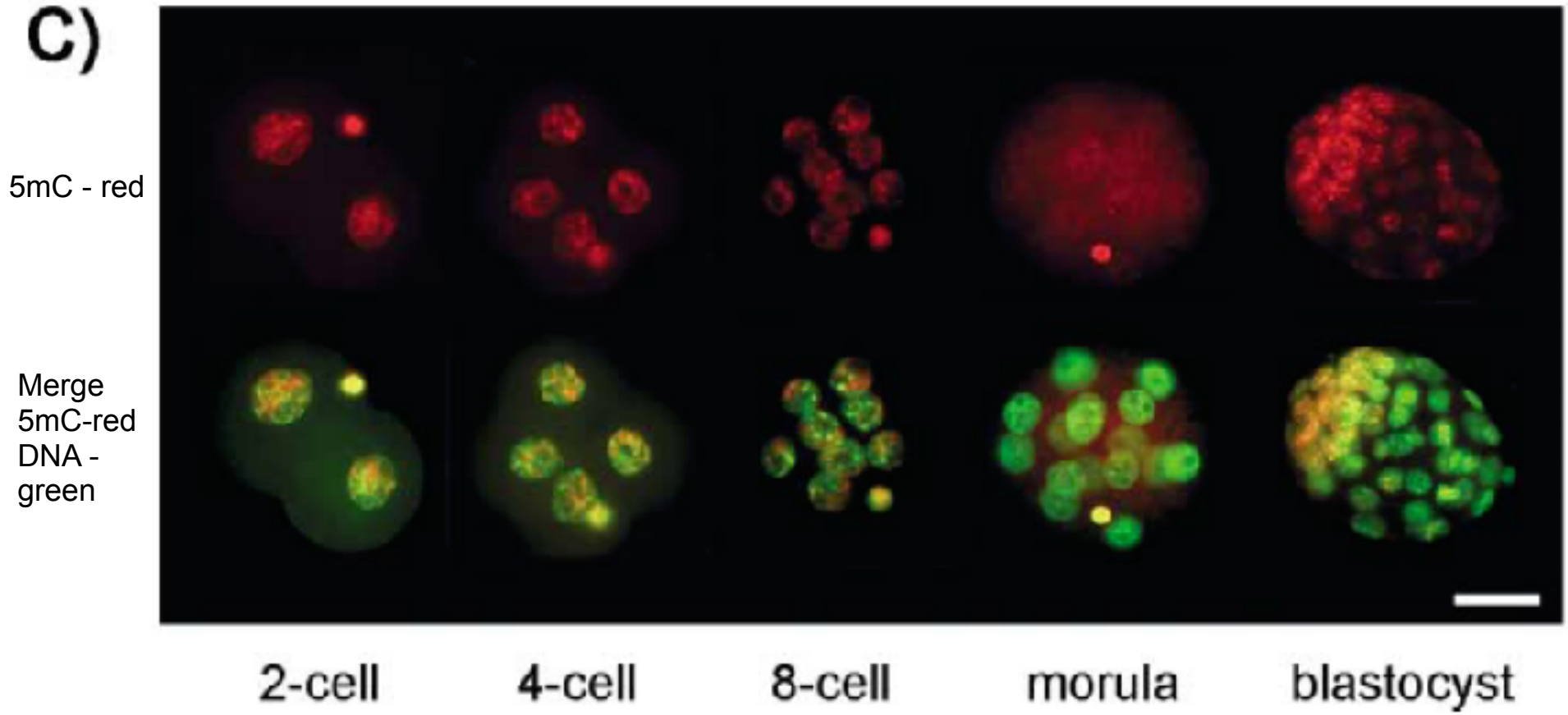
DNA methylation in germ cells and early embryo



Male and female genomes align at syngamy

Santos and Dean Reproduction 127:643

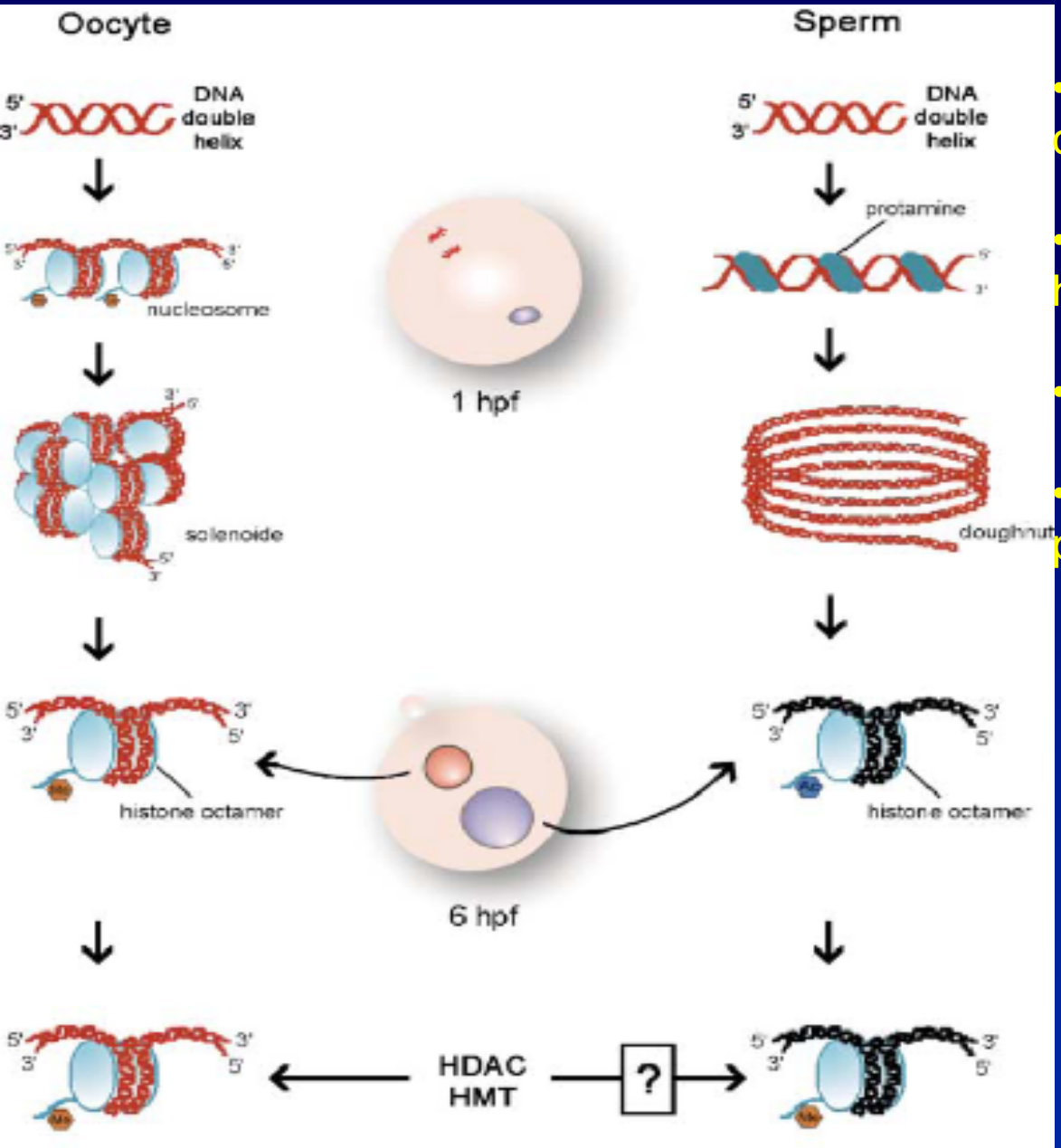
DNA methylation in germ cells and early embryo



Progressive loss of methylation in the female genome

Santos and Dean Reproduction 127:643

DNA methylation in germ cells and early embryo

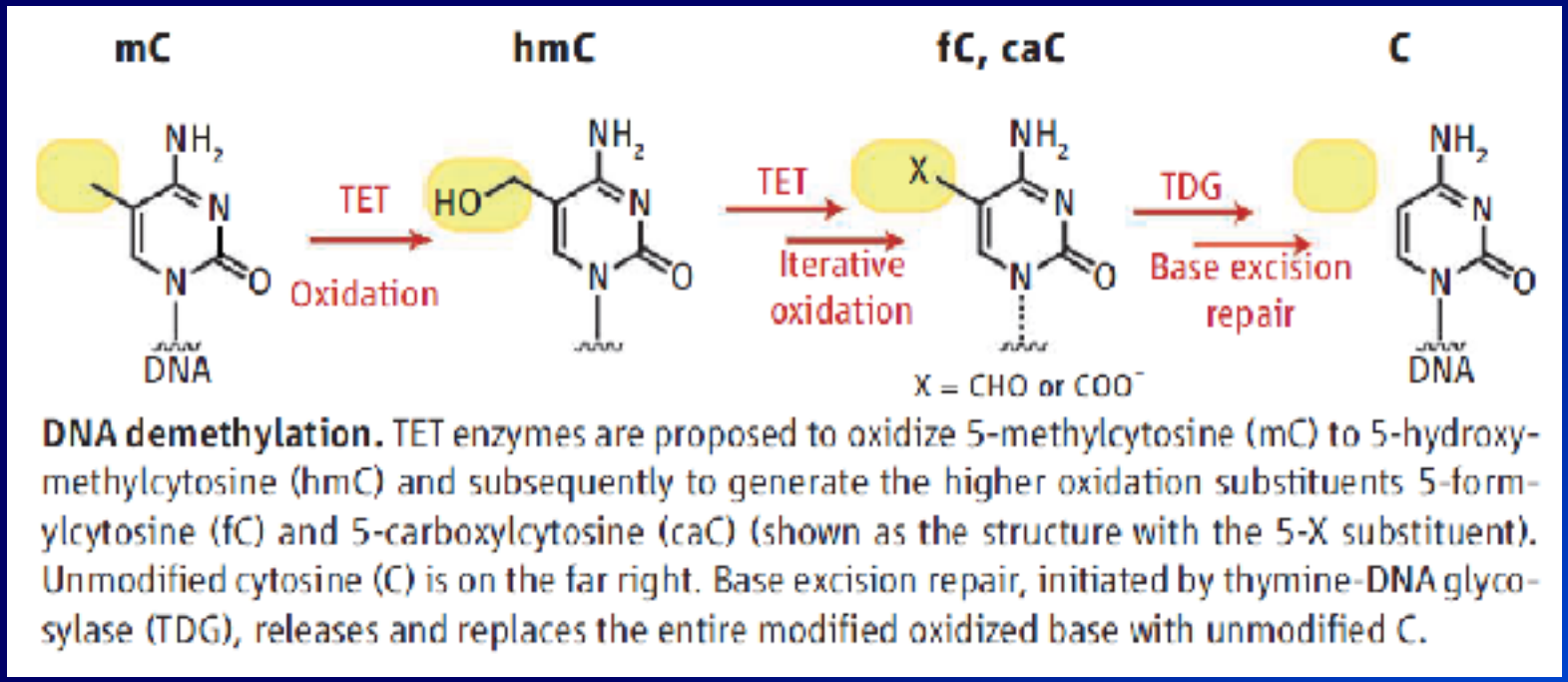


- Sperm genome has unique chromatin structure
- Protamine packaging instead of histones
- Both genomes must unwind
- Sperm genome must replace protamines with histones
  - Acquires acetylated histones
  - Male genome demethylated

Santos and Dean Reproduction 127:643

Ten eleven translocation (TET) proteins

Family of proteins that are 2-oxoglutarate and Fe(II)-dependent dioxygenases which all have the capacity to convert 5mC to 5hmC and further to 5fC to 5caC.



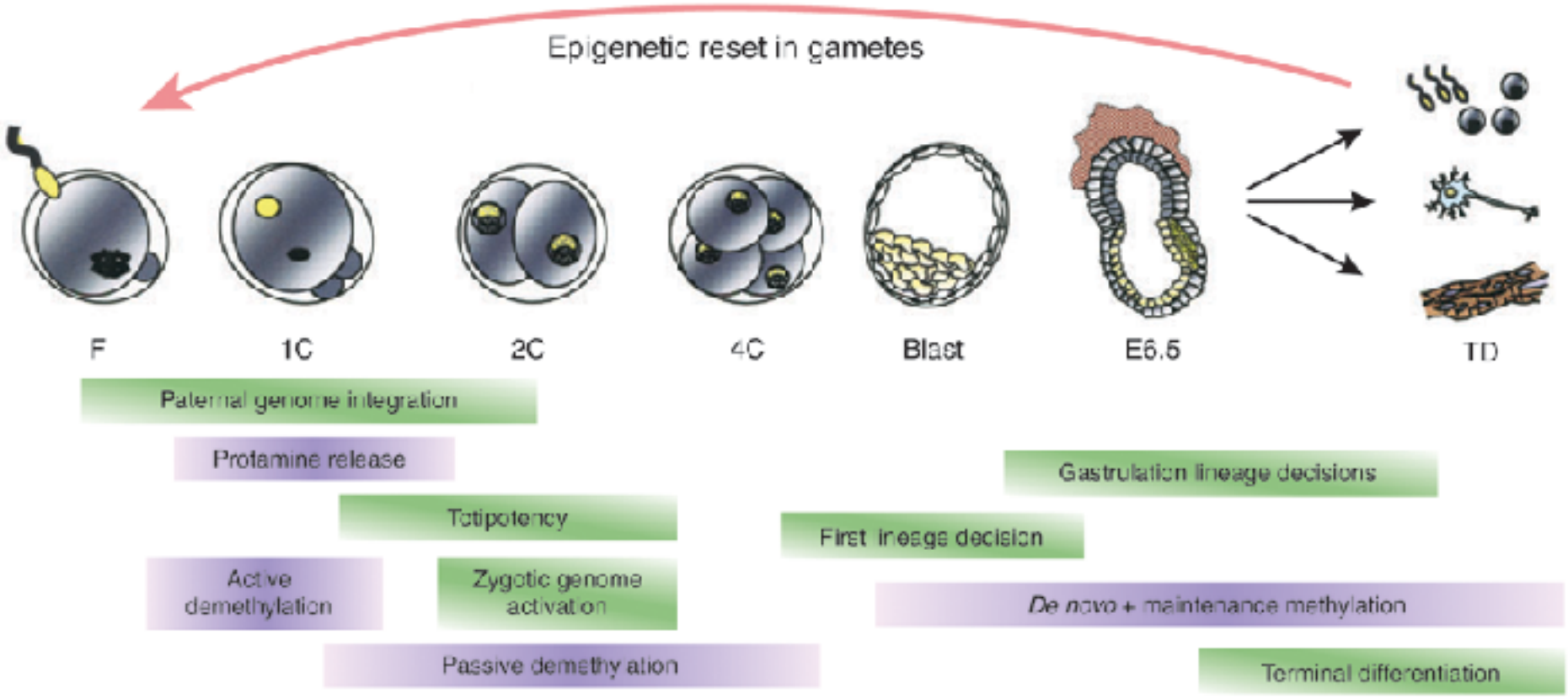
DNA demethylation. TET enzymes are proposed to oxidize 5-methylcytosine (mC) to 5-hydroxymethylcytosine (hmC) and subsequently to generate the higher oxidation substituents 5-formylcytosine (fC) and 5-carboxylcytosine (caC) (shown as the structure with the 5-X substituent). Unmodified cytosine (C) is on the far right. Base excision repair, initiated by thymine-DNA glycosylase (TDG), releases and replaces the entire modified oxidized base with unmodified C.

TET proteins cannot convert 5caC to C. A glycosylase must be involved.

Depletion of TDG leads to accumulation of 5caC in mouse ESCs. (He *et al.*, 2011 *Science*)

Image adapted from Nabel and Kohli, 2011, *Science*

Summary 3



Mager and Bartolomei Nature Genetics 37(11):1194-1200

Summary 3

- Both primordial germ cells and early embryo erase most methylation
  - "reset" the DNA methylation pattern
  - Centromeric and imprinted sequences excluded in early embryo

- Passive and active demethylation of the female and male genomes after fertilization
  - Active demethylation of male genome
    - Likely associated with replacing protamine based DNA packaging with histone based
    - Involves oxidation of 5mC to 5hmC via activity of TET proteins
  - Passive demethylation of female genome
    - Exclusion of Dnmt1 and Dnmt1o from the nucleus
    - Replication occurs without maintenance of methylation

Summary 3

- Methylation resets at implantation
  - *de novo* methylation by Dnmt3a and Dnmt3b
  - Maintained by the release of stores of Dnmt1o into the nucleus
    - Subsequent new production of Dnmt1

The way we do the things we do

DNA methylation in cancer - Primary observations

- Global hypomethylation
  - Demethylation of repetitive elements (passive/active?)
    - Increase chromosomal instability
    - Increase somatic recombination (LOH)
    - Ectopic expression of genes
      - Driven by parasitic DNAs' promoters
      - Demethylation of CpG island promoter

- Localized hypermethylation
  - Normally unmethylated CpG islands become methylated
    - ~29,000 CpG islands in human
    - Generally in promoter regions
    - Associated with gene silencing
      - Silencing of tumor suppressor genes like *p16*, *GSTP1*

How to detect methylation of a gene
Three basic strategies:

1) Methylation sensitive restriction enzymes

- Some restriction enzymes will not cut if the DNA is methylated
  - *Hpa*II will cut CCGG, but not C^{me}CGG
    - Isoschizomer *Msp*I cuts same site regardless of methylation

2) Bisulphite treatment

- Chemically converts cytosine to uracil
  - 5-methylcytosine is protected from conversion
- Creates a sequence difference between methylated and unmethylated DNA

3) Immunoprecipitation/Affinity

- Antibody to 5-methyl-C/bind to MBD proteins; purify and look at the enriched sequences

Scanning vs. candidate gene approaches

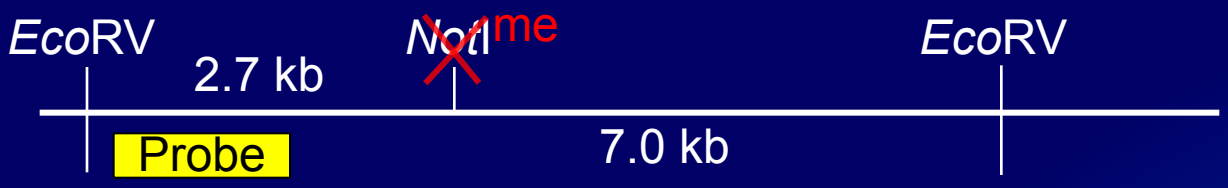
- Candidate gene approach
  - Ask if specific gene(s) of interest is methylated
  - Example: p16 in pancreatic cancer
    - Often one allele deleted, one allele present
    - No expression
    - Methylation cause of lack of expression in normal allele

- Scanning approach
  - What loci are affected by CpG island methylation?
  - How severe is the aberrant methylation phenotype?
  - Study hundreds or thousands of loci in unbiased way

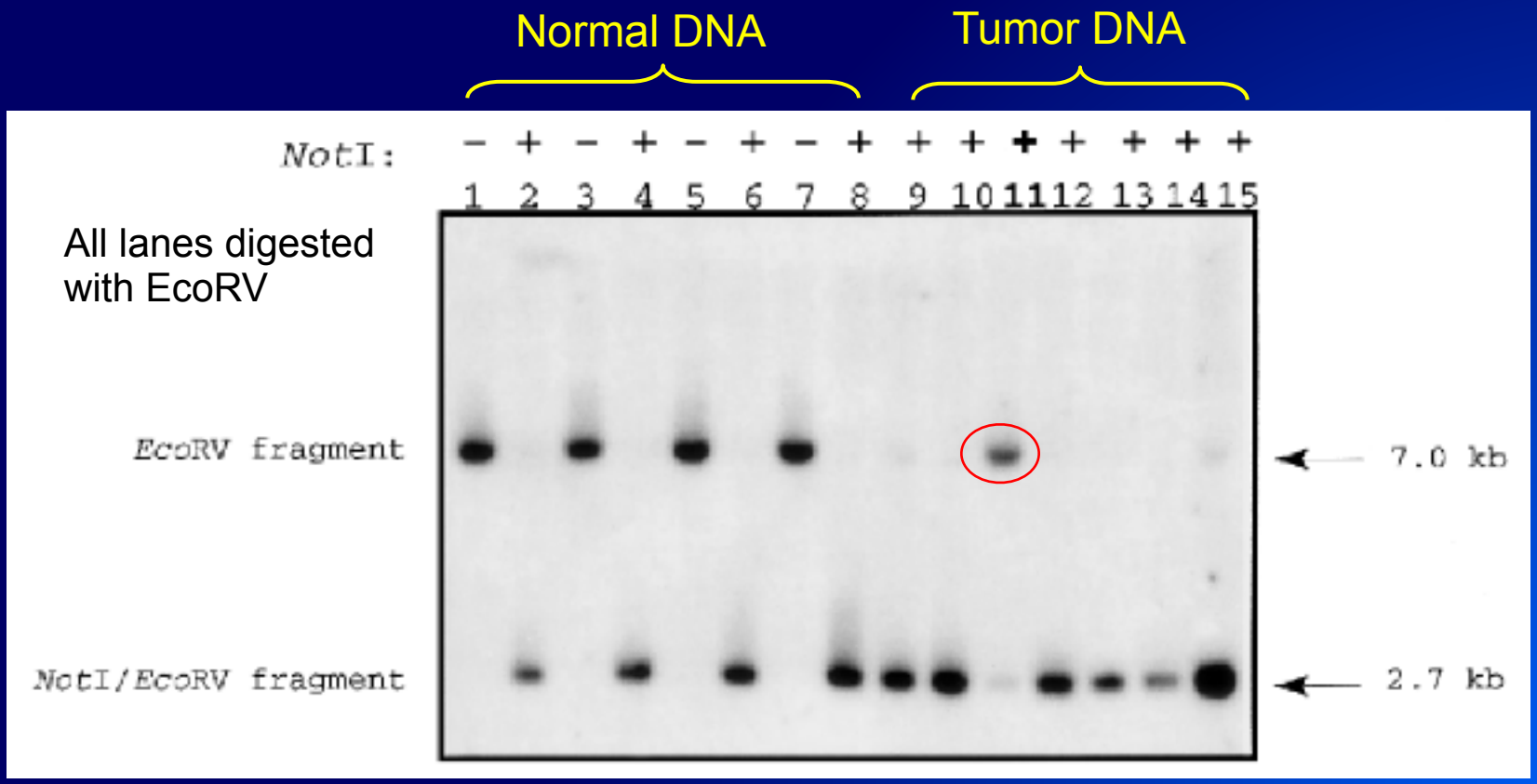
- Candidate approach often provides detailed info for a single locus
  - Can be applied to many samples; highly sensitive
- Scanning approach provides less detailed info for many loci
  - Cost and complexity limit number of samples; less sensitive

Methylation sensitive restriction enzymes: Candidate

Candidate Single sequence Southern blot. Does the methylation sensitive enzyme cut?

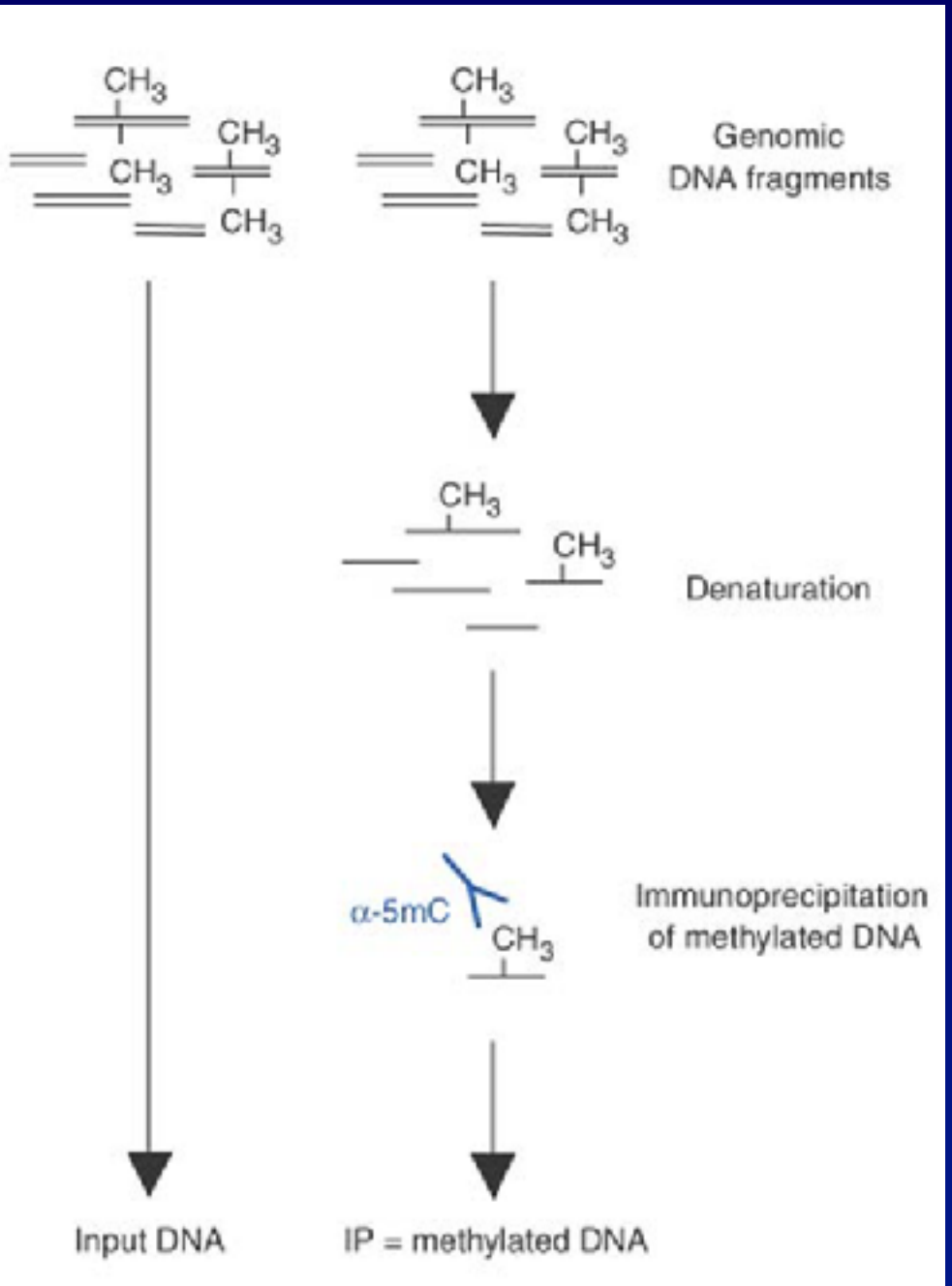


Estrogen receptor promoter CpG island

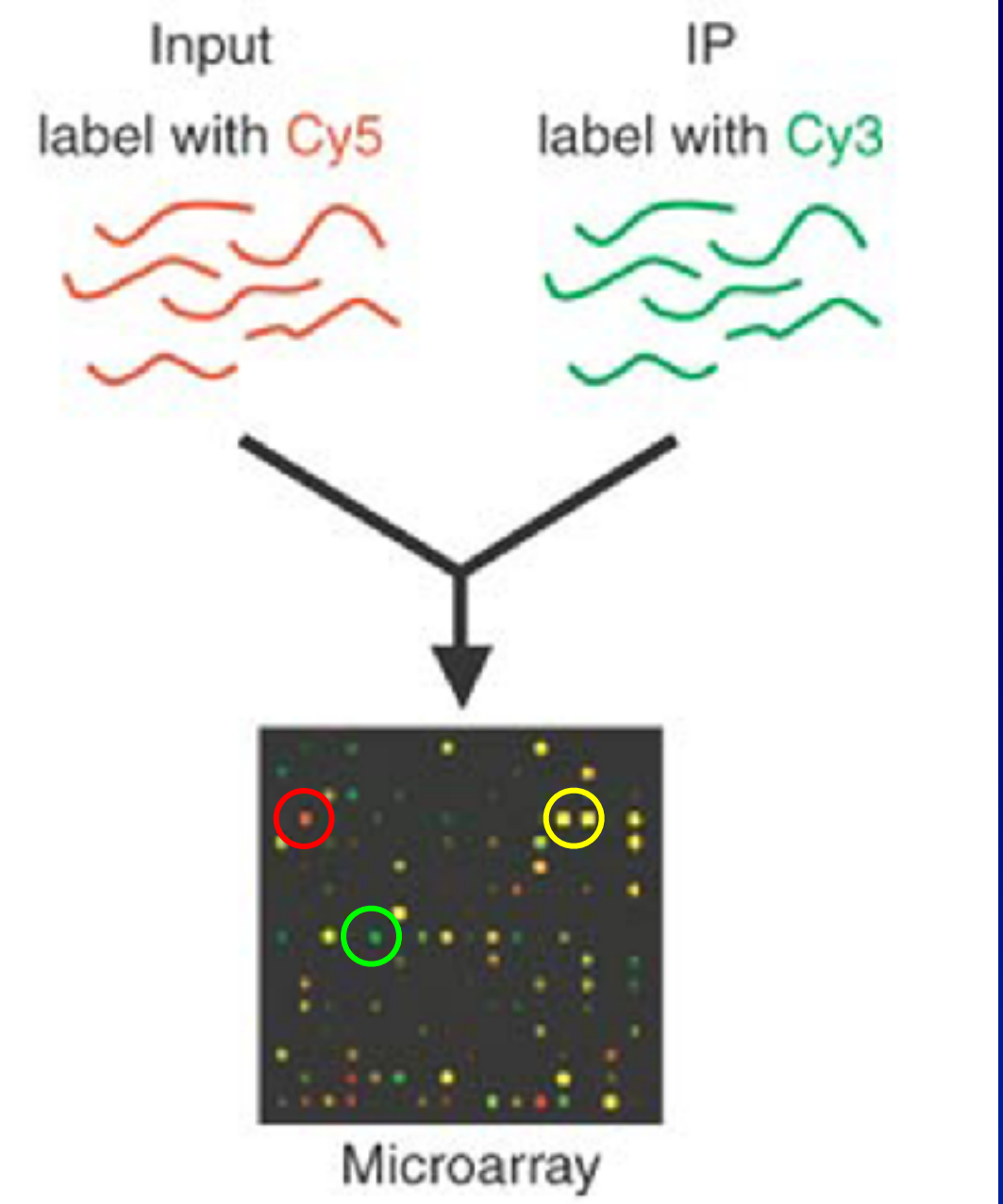


Smiraglia et al Genomics 58:254

Methylated DNA immunoprecipitation (MeDIP): Scanning



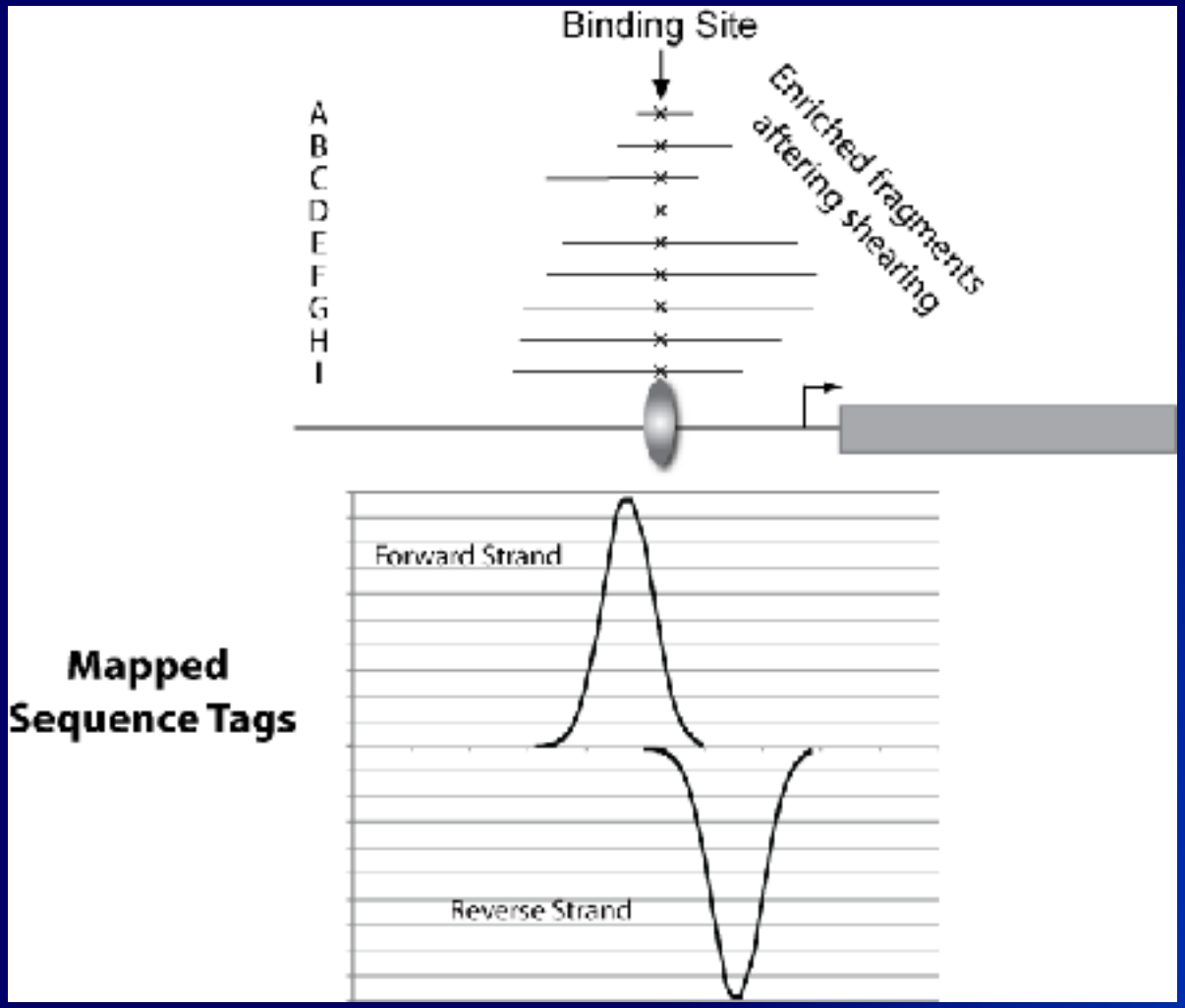
- Sonicate genomic DNA to get into ~500bp fragments
- Denature DNA to make epitope accessible
- IP with antibody to 5mC
- Collect starting DNA (Input DNA), and the DNA bound to the antibody (IP DNA)
- IP DNA should be enriched for methylated molecules compared to the input DNA
- How do you measure this??



- Label the IP DNA green and the Input DNA red
- Hybridize to a microarray with all CpG islands
- Red spots indicate a CpG island NOT enriched by the IP, therefore NOT methylated
- Yellow spots indicate a CpG island partially enriched by the IP, therefore partially methylated
- Green spots indicate a CpG island strongly enriched by the IP, therefore highly methylated

Next Generation **Sequencing – even better approach**

- Massively parallel sequencing of short tags of the enriched DNA compared to background (non-enriched)



- Methylation is indicated by many short seq tags mapping to a region in the enriched DNA

Sodium Bisulfite Treatment

- Converts unmethylated C to T with PCR amplification
- Methylated cytosines remain cytosines

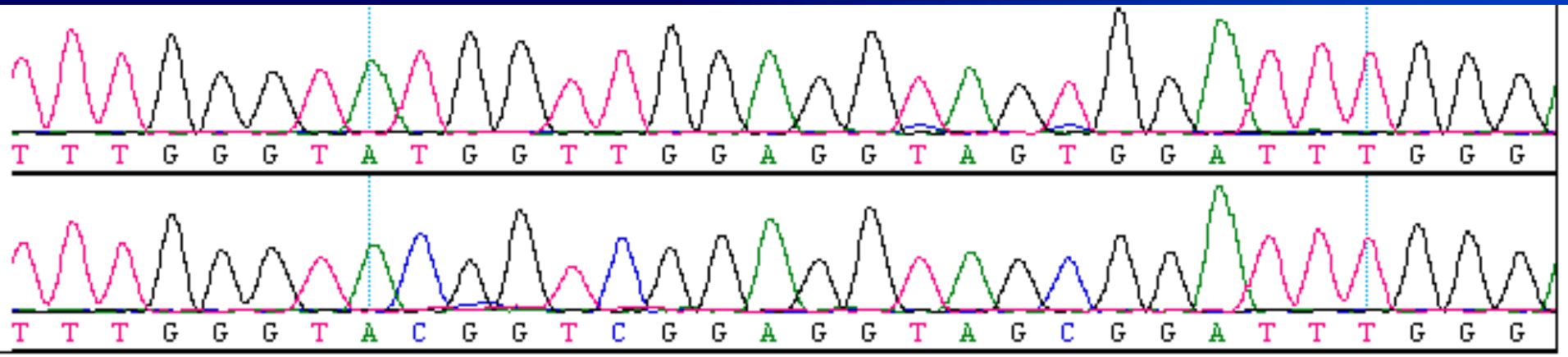
GATC….A^{m}CGTACGT….GATC

GATT….A^{m}CGTATGT….GATT

Bisulphite sequencing: Candidate

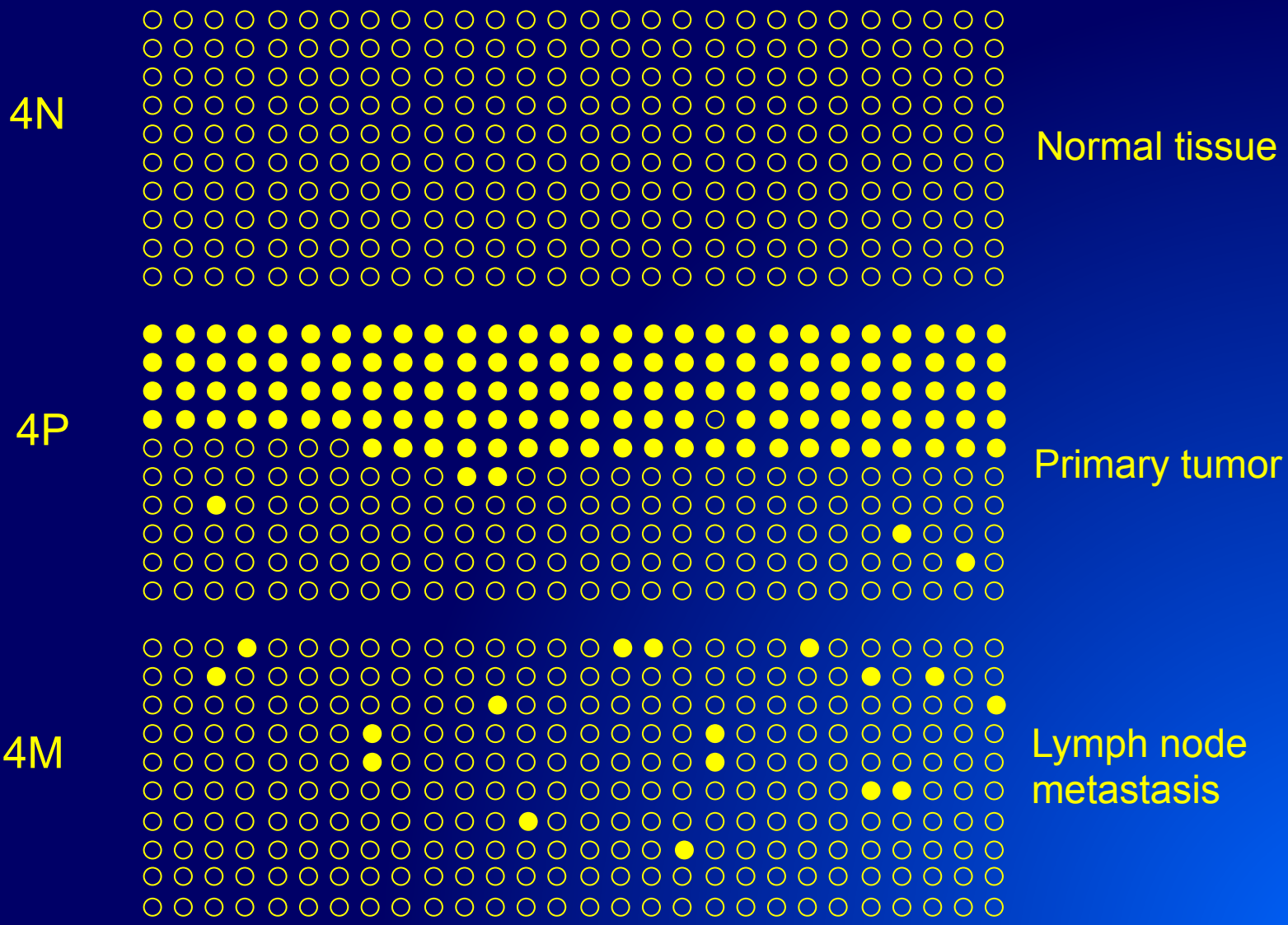
GATT….ACGTACGTACGTACGT ….GATT

GATT….ATGTATGTATGTATGT ….GATT



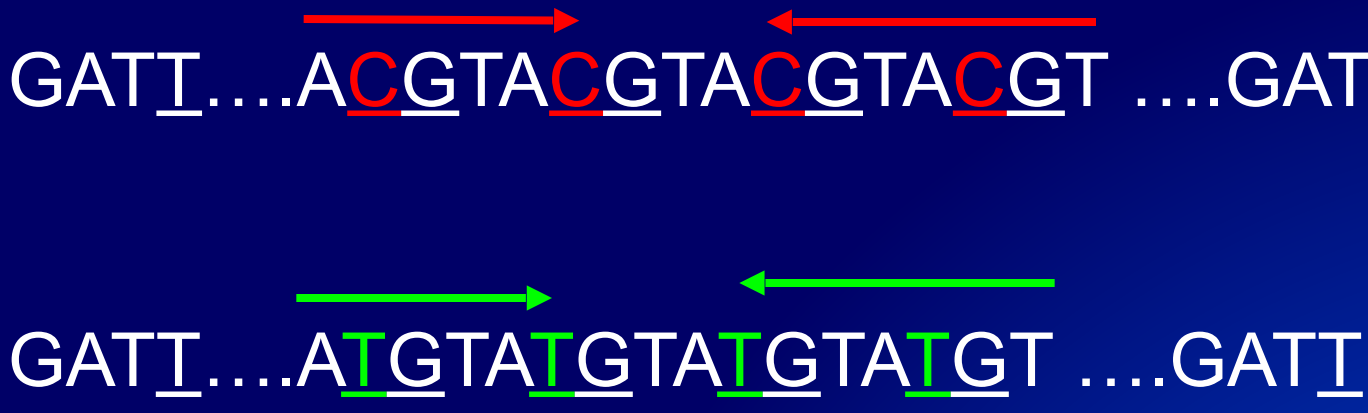
Example Bisulphite sequencing data

p16 methylation in head and neck cancer



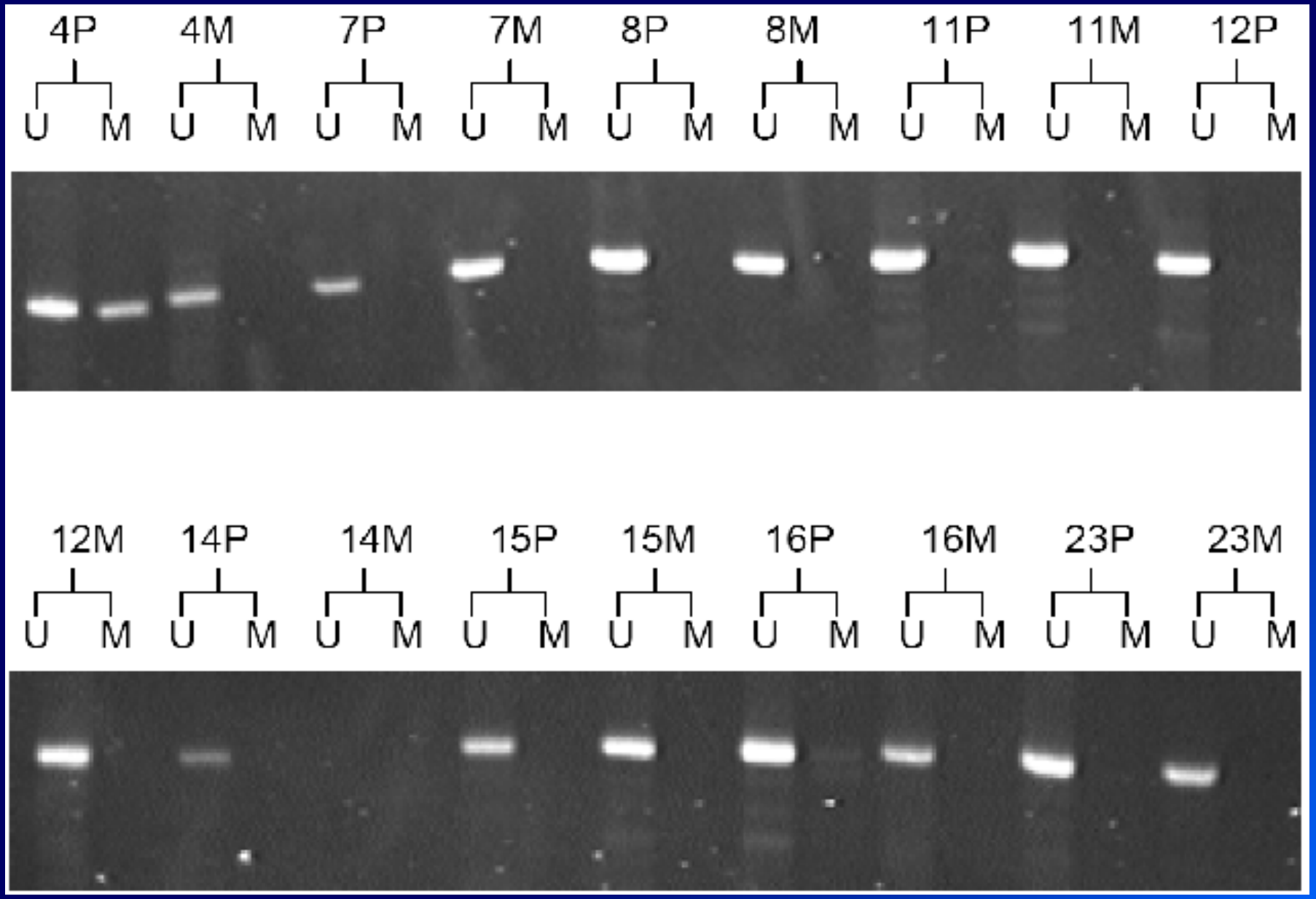
Smiraglia et al J Med Genet 40:25

Methylation Specific PCR (MSP) : Candidate



- Design two primer sets:
  - one specific for methylated sequence: will ONLY amplify methylated molecules
  - one specific for unmethylated sequence: will ONLY amplify unmethylated molecules
- Non-quantitative
  - Simply look for presence or absence of band using methylated primer set
    - Unmethylated primer set acts as control for locus integrety

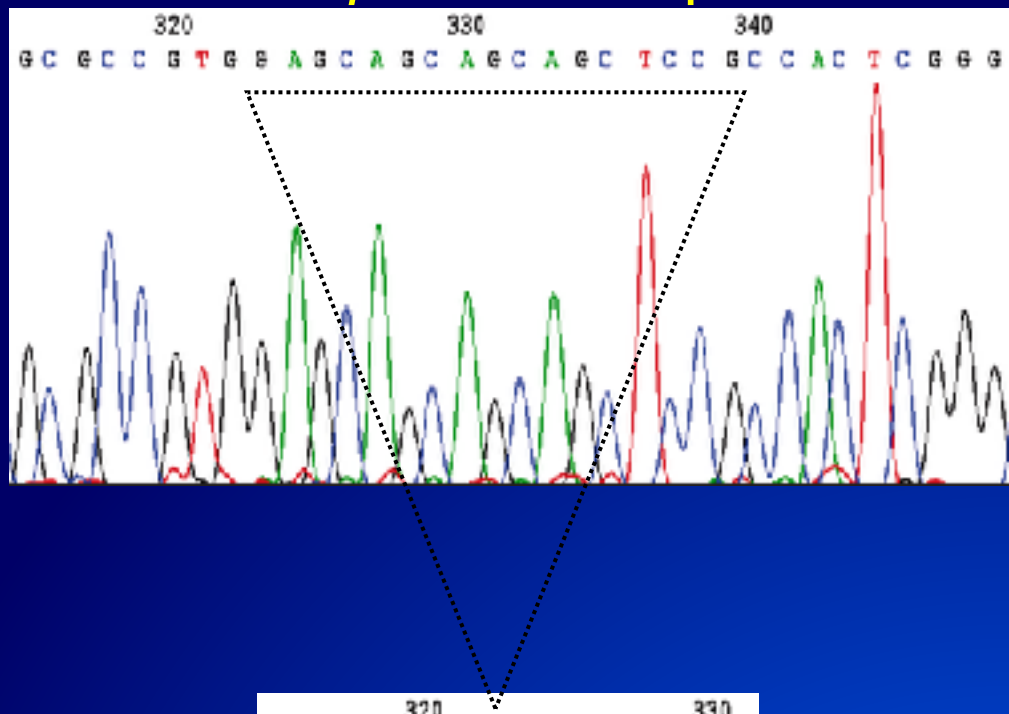
MSP analysis of p16 in head and neck cancer

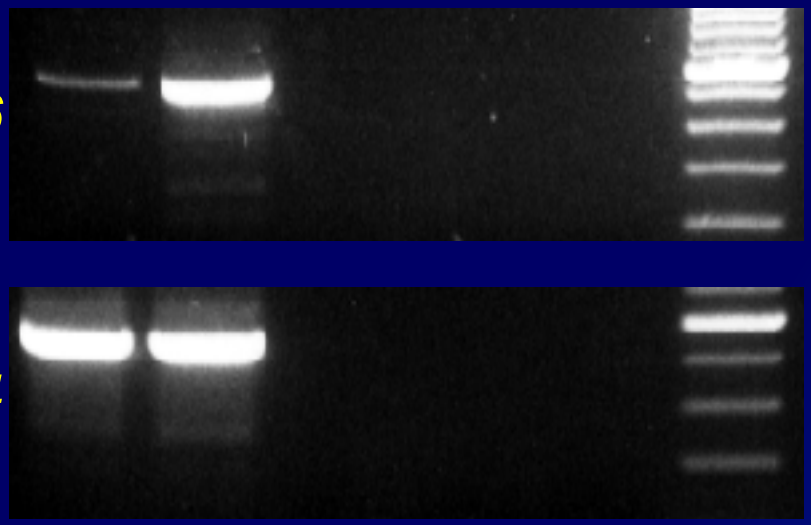


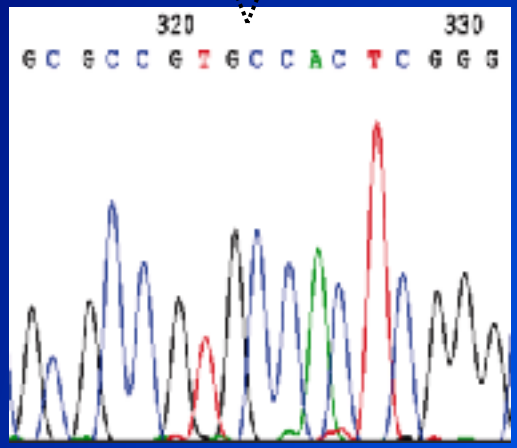
Smiraglia et al J Med Genet 40:25

Multiple routes of *p16* loss of function

4P *p16* transcript







4M *p16* transcript

17 bp deletion; nonsense mutation after amino acid 60, early in exon 2

Smiraglia et al J Med Genet 40:25

Multiple routes of *p16* loss of function

- Patient 4 head and neck primary tumor has CpG island methylation
  - MSP and bisulfite sequencing demonstrate methylation
  - Lack of transcription – loss of p16 function

- Lymph node metastasis from same patient does not have methylation
  - MSP and bisulfite sequencing both demonstrate lack of methylation
  - Plenty of transcription
  - Deletion in exon 2 causing non-functional truncation – loss of p16 function

How can the primary and metastatic tumors from the same patient have different mechanisms for loss of p16 function?

Dynamic CpG Island hypermethylation

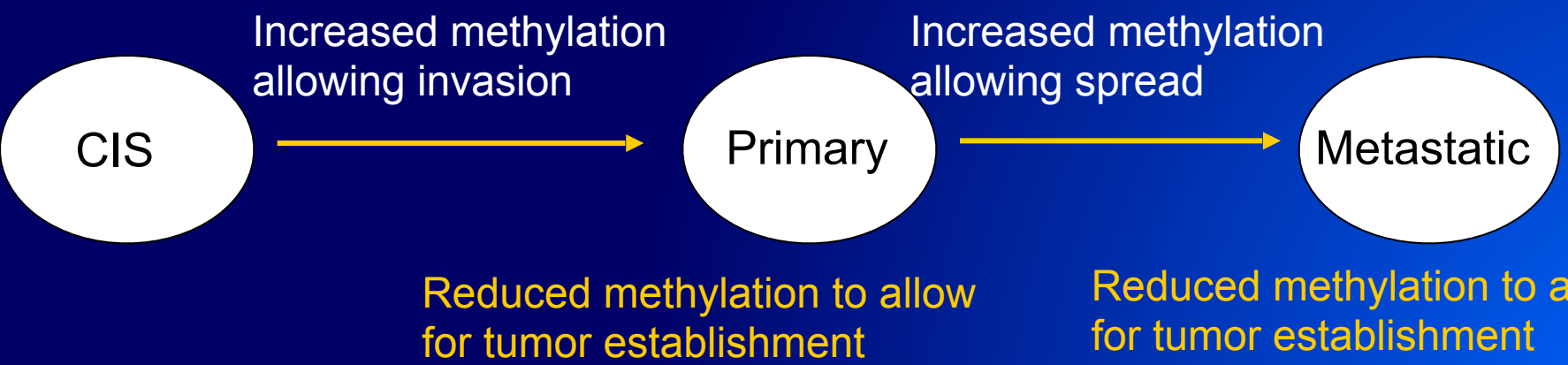
Could methylation of CpG islands be erased to give the cells an advantage when metastasizing?

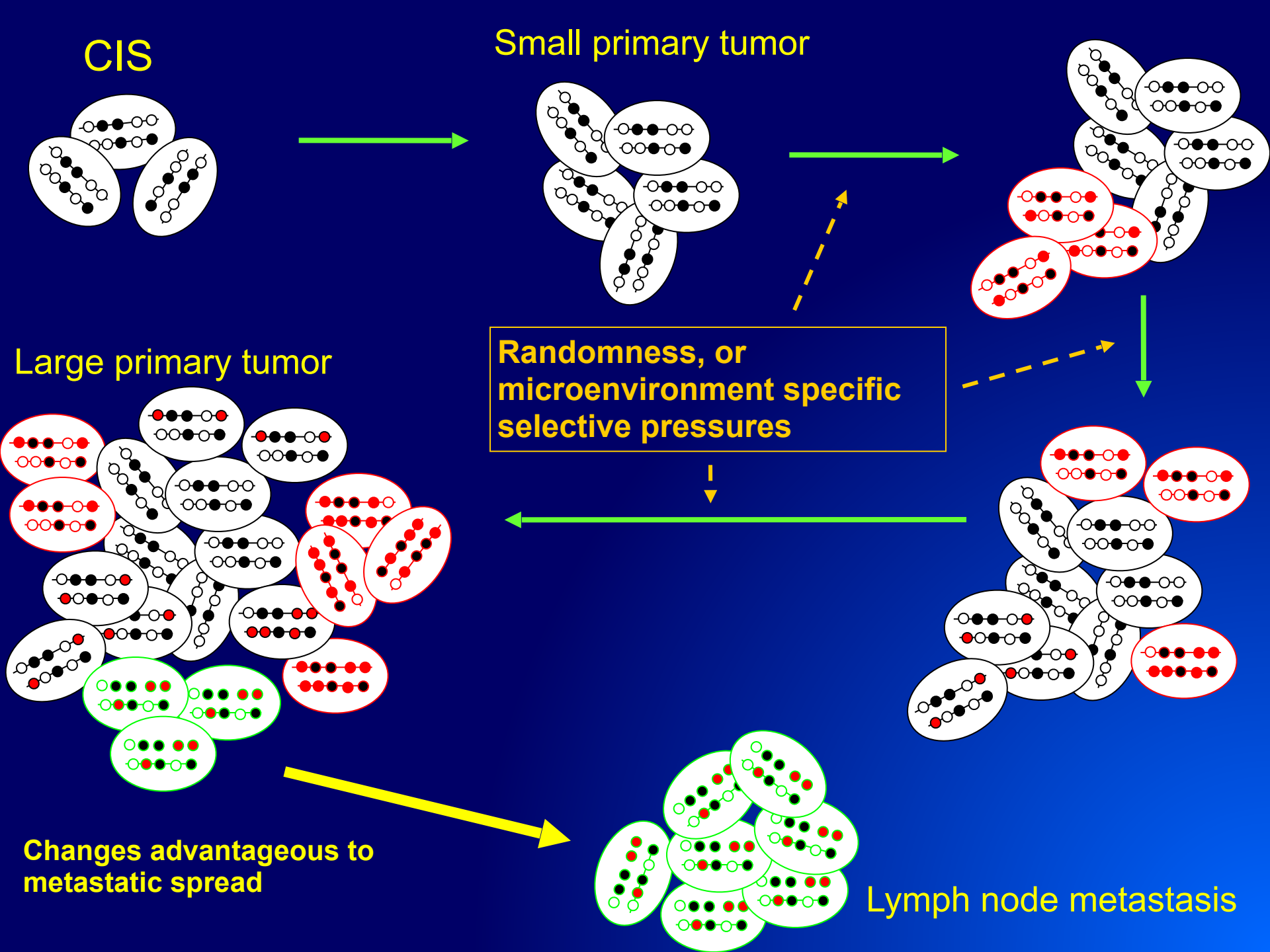
- Idea would be unique to epigenetic changes
  - Active demethyltation of a specific promoter?

- E-Cadherin (E-cad): involved in cell-cell adhesion
  - useful to cells growing in tumor mass
  - detrimental to invasiveness
  - detrimental to cells breaking away for metastasis
  - useful to establishment of tumor at metastatic site

Dynamic CpG Island hypermethylation

- Prostate cancer cell line monolayer
  - heterogeneous hypermethylation of E-cad
- In model system of basement membrane invasion
  - increased hypermethylation of E-cad
- In model system of three-dimensional tumor growth
  - reduced methylation of E-Cad




CIS
Small primary tumor
Large primary tumor
Randomness, or microenvironment specific selective pressures
Changes advantageous to metastatic spread
Lymph node metastasis

Summary 4

- Global hypomethylation of bulk chromatin
- Localized hypermethylation of CpG islands
- Silences tumor suppressor genes if CpG island is in the promoter
- Detect methylation – methylation sensitive enzymes
  - Southern blot – candidate gene
  - RLGS – scanning approach
- Detect methylation – bisulphite treatment
  - Bisulphite sequencing
  - Methylation sensitive PCR
- Scanning approaches
  - Limited information about many loci
- Candidate approaches
  - Detailed information about few loci

Summary 4

- Cancer cell doesn't care why there is loss of function of TSG
  - Genetic or epigenetic reasons give the same result
- Epigenetic changes can be reversed
  - Cancer cells may take advantage of this for certain genes
    - More hypothetical than proven fact, at this point…
- Tumors are heterogeneous
  - Different subpopulations may use different routes of loss of function