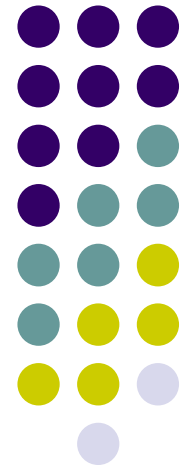
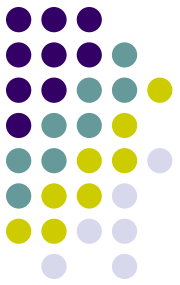


# DNA Methylation and Genome Evolution

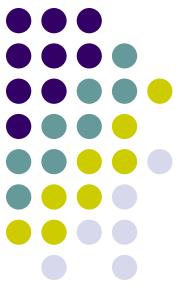
Soojin Yi  
School of Biology  
Georgia Institute of Technology



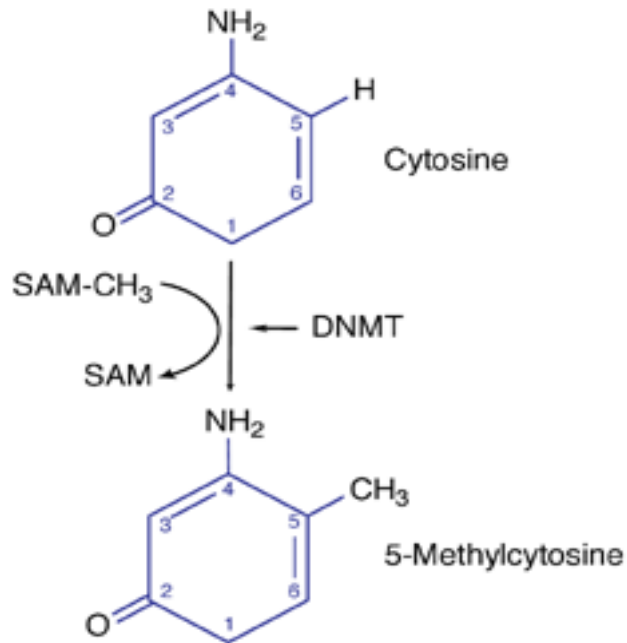


# DNA methylation

- A type of ‘epigenetic’ modification
  - ‘chemical’ marking of a base
- Widespread in the tree of life



# Cytosine methylation



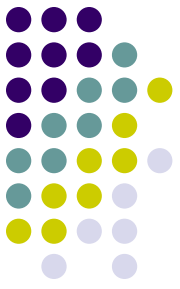
Mechanism of DNA methylation

Expert Reviews in Molecular Medicine

©2002 Cambridge University Press

- Only cytosines followed by guanines in the 3-prime position ('CpG')
- Main mode of DNA methylation in animal genomes

# DNA methylation: 'silencer' or 'dimmer' of gene expression



- Transcriptional gene silencing
- Chromatin compaction
- Suppression of homologous recombination between repeats
- X chromosome inactivation
- Genomic imprinting
- A main source of mutations: affect several aspects of genome evolution

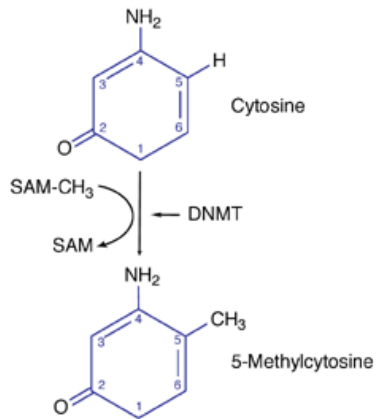
# DNA methylation: highly mutagenic



## 1. methylation

## 2. deamination

## 3. DNA repair

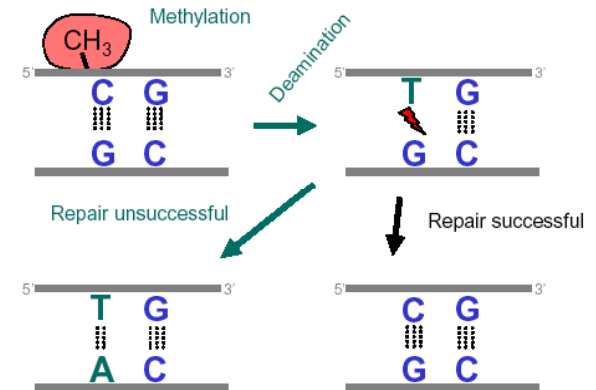


Mechanism of DNA methylation  
Expert Reviews in Molecular Medicine  
©2002 Cambridge University Press

QuickTime™ and a  
TIFF (Uncompressed) decompressor  
are needed to see this picture.

Methyl-cytosine

Thymine



Increase point mutations  
from CpG to TpG (or from CpG to CpA)  
: 'CpG mutation'

# CpG mutation is a major input into mammalian genome evolution



Ratio of CpG to non-CpG substitution in primates: ~15  
[Kim et al. 2006](#); [Elango et al. 2008](#)

The most common mutation in the human genome (25% of all mutations between human and chimpanzee)  
[The chimpanzee genome consortium 2005](#)

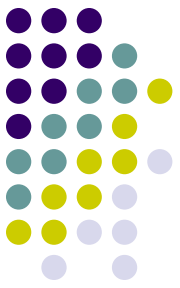
Due to this property, CpG dinucleotides are rare in the human genome (only about 1/5 of what is expected)

# DNA methylation influences several aspects of genome evolution



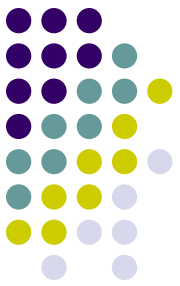
- DNA methylation and molecular clock
  - [Kim et al. 2006 PLoS Genetics](#)
- DNA methylation and compositional evolution
  - [Elango et al. 2008 PLoS Comp. Biol.](#)
- DNA methylation and evolution of vertebrate promoters
  - [Elango & Yi 2008 Mol. Biol. Evol.](#)
- The pattern and functional role of DNA methylation in the honeybee genome
  - [Elango, Hunt et al. \*submitted\*](#)

# DNA methylation influences several aspects of genome evolution



- DNA methylation and molecular clock
  - Kim et al. 2006 PLoS Genetics
- DNA methylation and compositional evolution
  - Elango et al. 2008 PLoS Comp. Biol.
- DNA methylation and evolution of vertebrate promoters
  - Elango & Yi 2008 Mol. Biol. Evol.
- The pattern and functional role of DNA methylation in the honeybee genome
  - Elango, Hunt et al. *submitted*





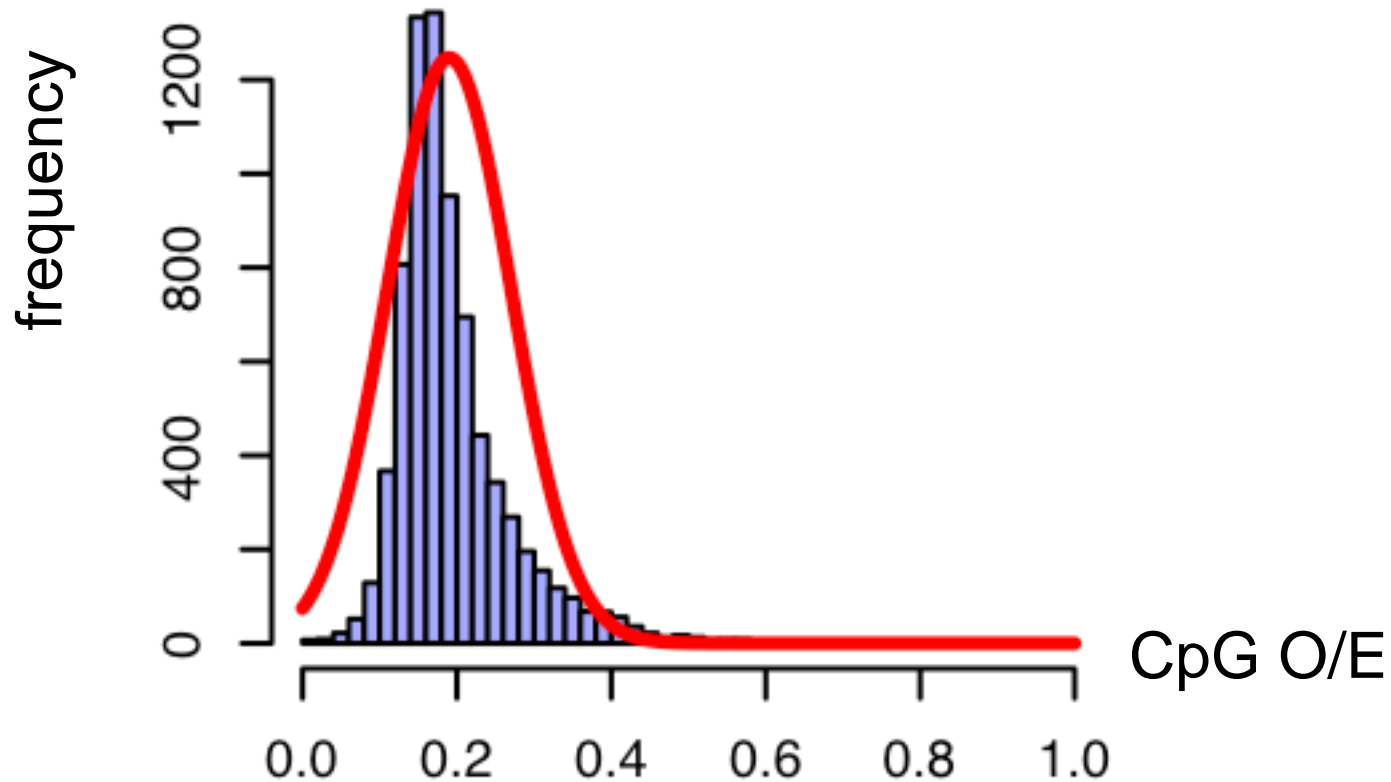
## Normalized CpG Contents (CpG O/E): measure of historical levels of methylation

- Normalized CpG contents

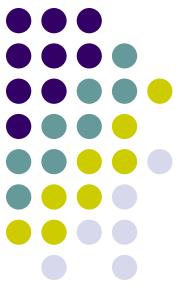
$$CpGO / E = \frac{freq(CpG)_{Obs}}{freq(CpG)_{Exp}}$$

- DNA methylation  $\uparrow$ : CpG O/E  $\downarrow$
- Negatively correlated with the level of genomic methylation (Bird 1980)

# Human genome: most CpG are methylated in the majority of tissues ( 'global DNA methylation' )



*Exceptions to this pattern: promoter regions*

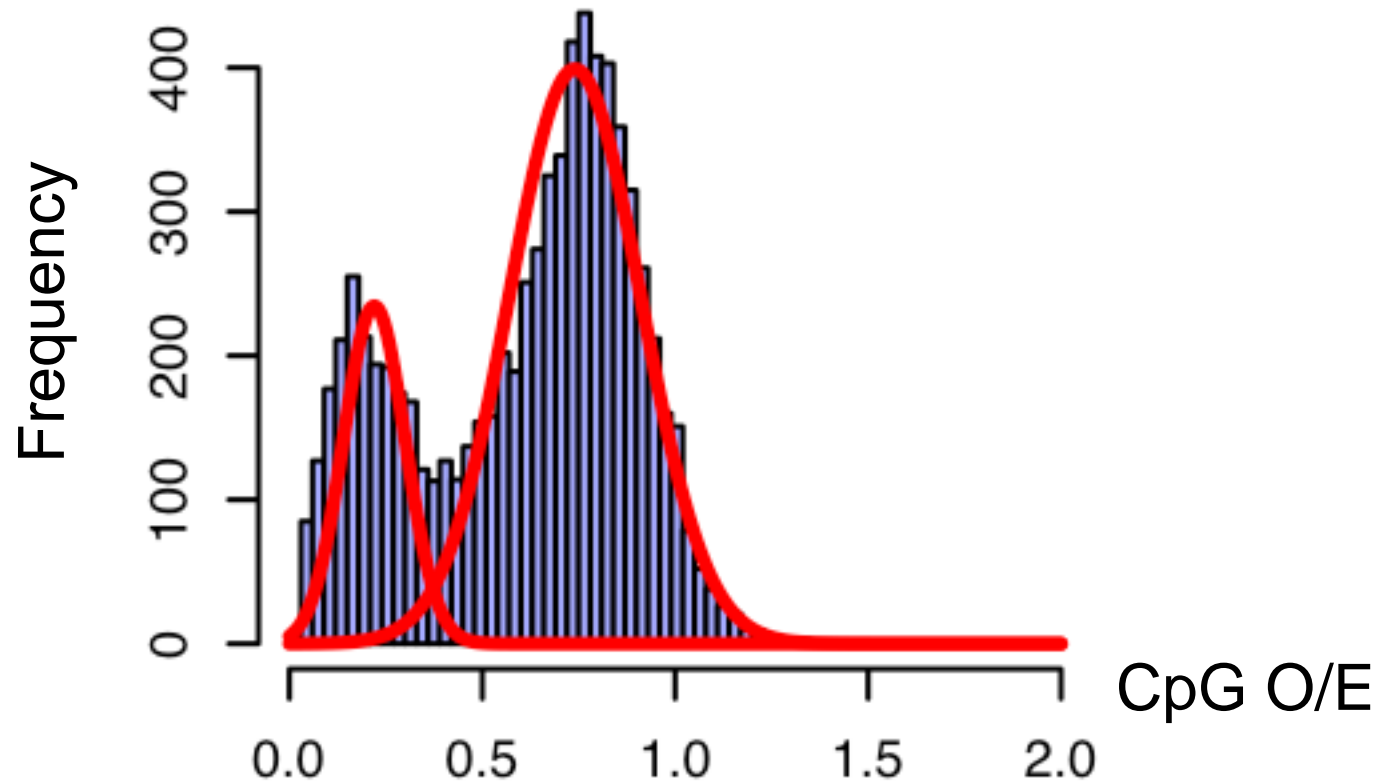


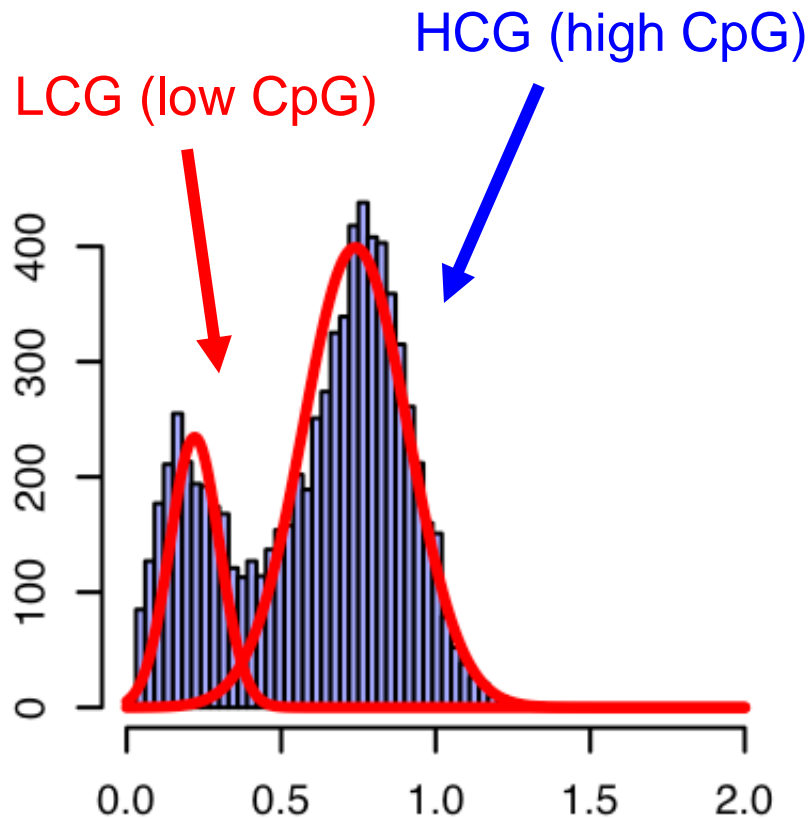
# Promoter DNA methylation and regulation of gene expression

- Promoter methylation: repress transcription
  - Methylated cytosines can recruit specific family of proteins (e.g., MBDs) that can inhibit gene expression
  - For example, MBD2 participates in protein complexes that recruit transcriptional co-repressors, chromatin remodeling proteins, and histone acetylases
- Broadly expressed genes: tend to be hypo-methylated
- tissue-specific genes: tend to be hyper-methylated



# Mammalian promoters: bimodal





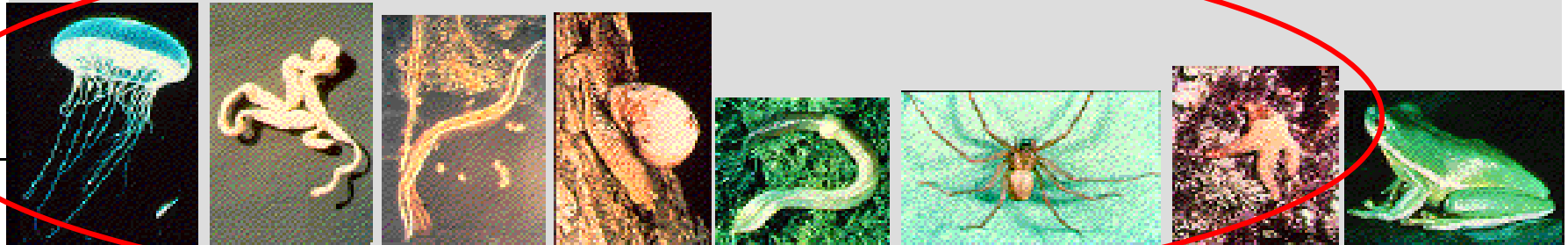
- LCG (low CpG) and HCG (high CpG) promoters
- HCGs are associated with CpG islands
- LCGs and HCGs are functionally distinct
  - HCGs: broadly-expressed genes
  - LCGs: tissue-specific genes

# Evolution of hyper- and hypomethylated human promoters



- When did the bimodality of promoters evolve?
- How and why did the bimodality of promoters originate?

# Distribution of DNA methylation pattern in animal genomes



Cnidaria

Platyhelminthes

Nematoda

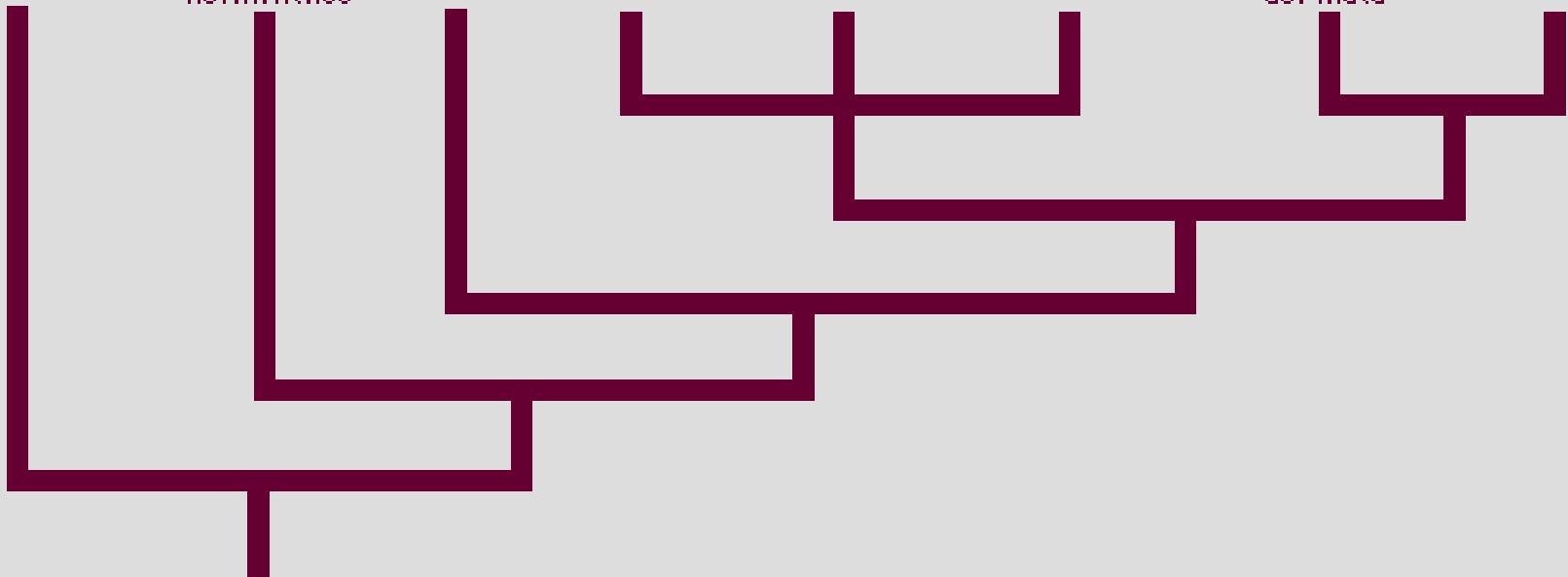
Mollusca

Annelida

Arthropoda

Echinodermata

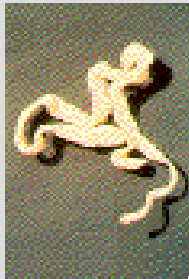
Chordata



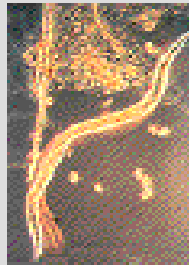
# Distribution of DNA methylation pattern in animal genomes



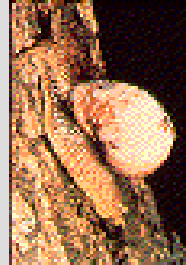
Cnidaria



Platyhelminthes



Nematoda



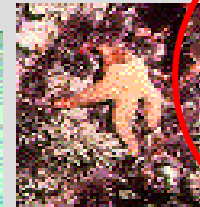
Mollusca



Annelida



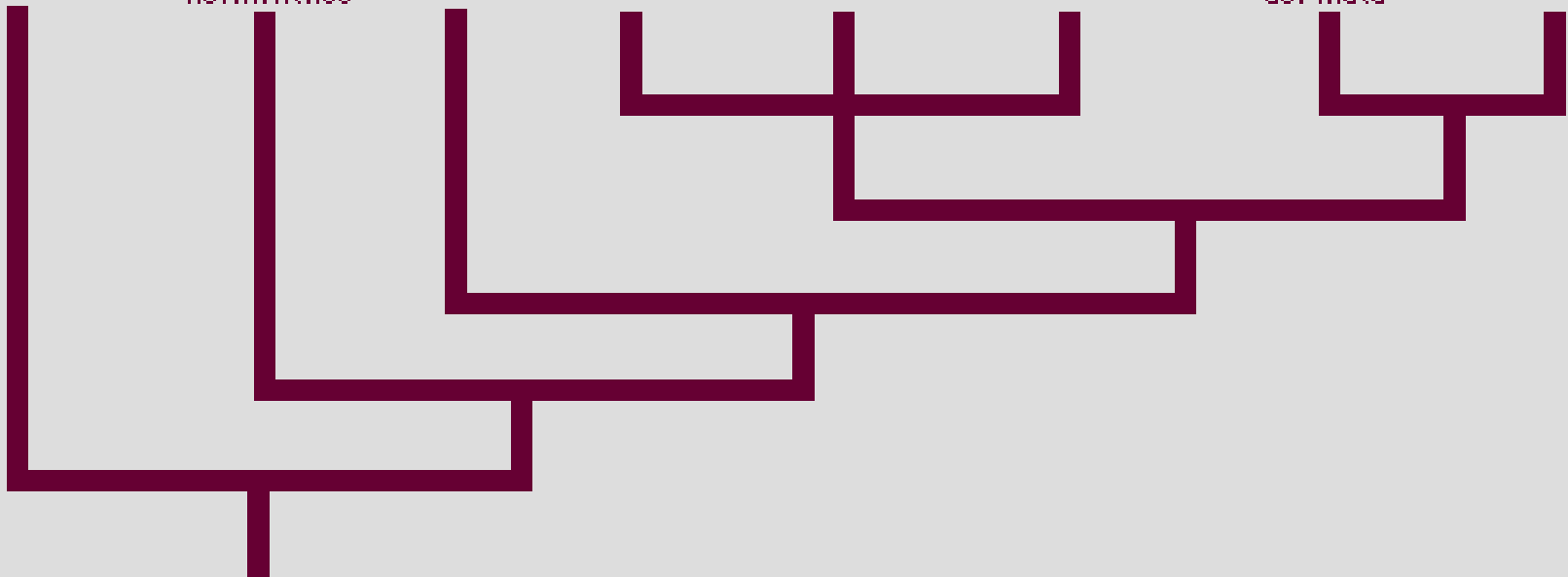
Arthropoda



Echinodermata



Chordata





# Global DNA methylation is vertebrate-specific



Distant Invertebrates: ?

Sea squirt (*Ciona intestinalis*):

Patchy methylation

**Targeted to genes**

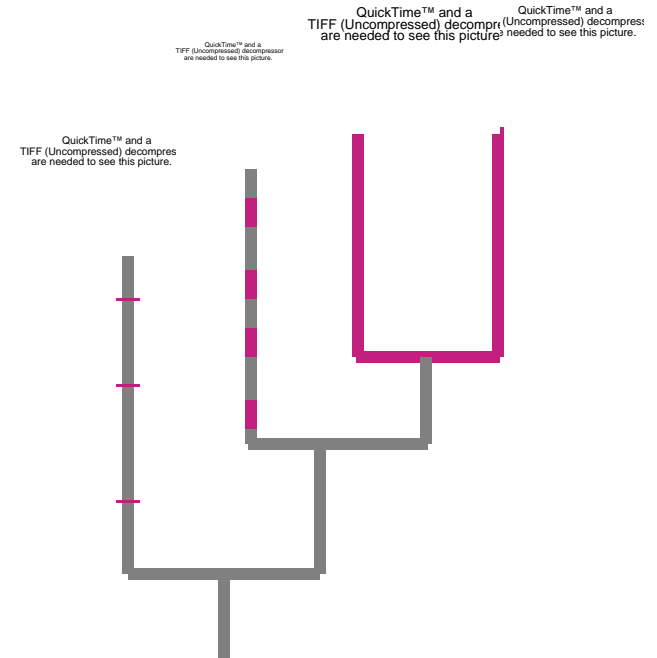
TEs are not methylated

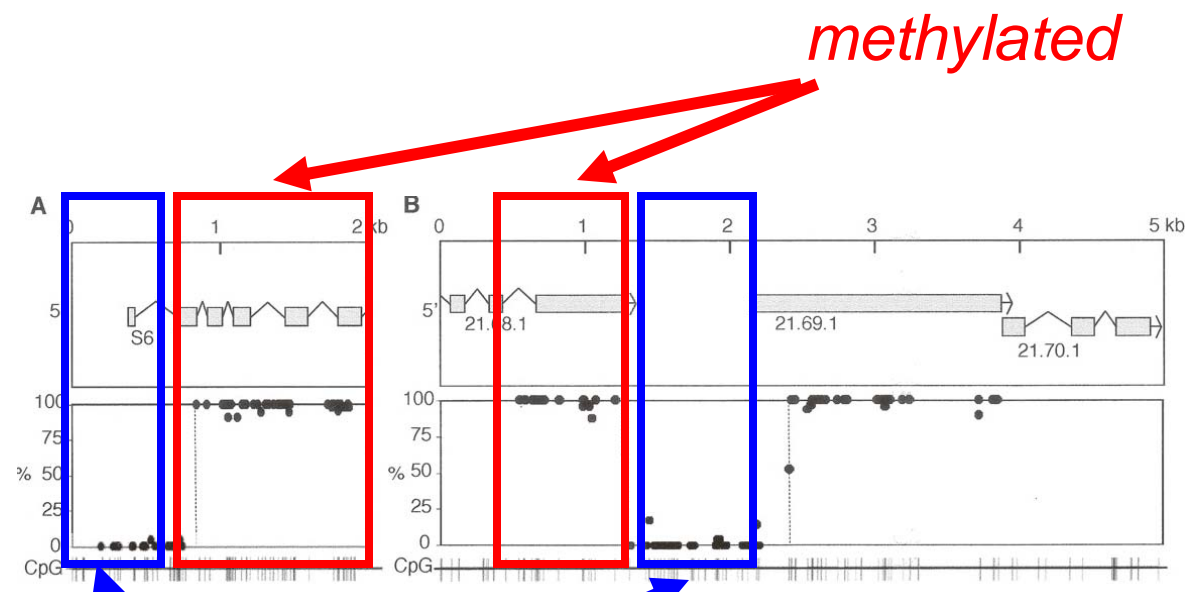
Human:

Global methylation

Except CpG islands

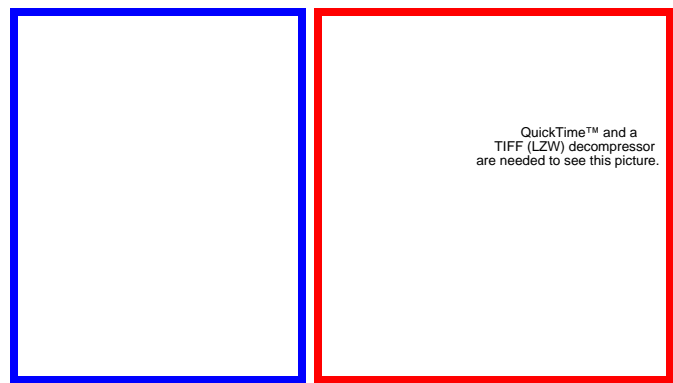
TEs heavily methylated



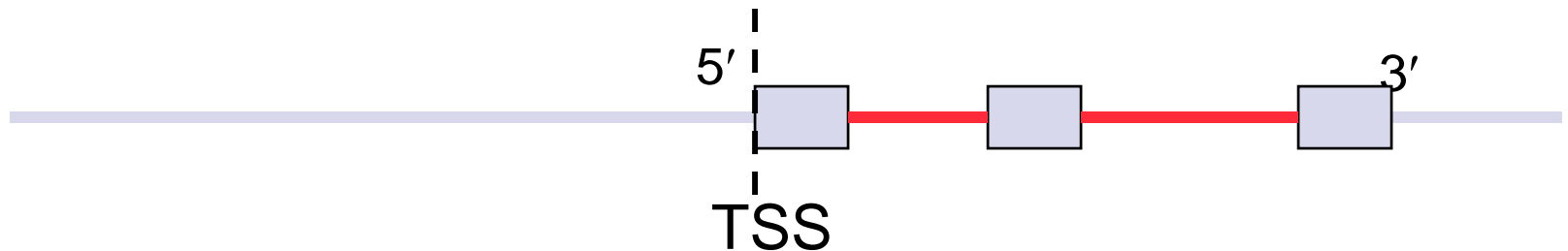
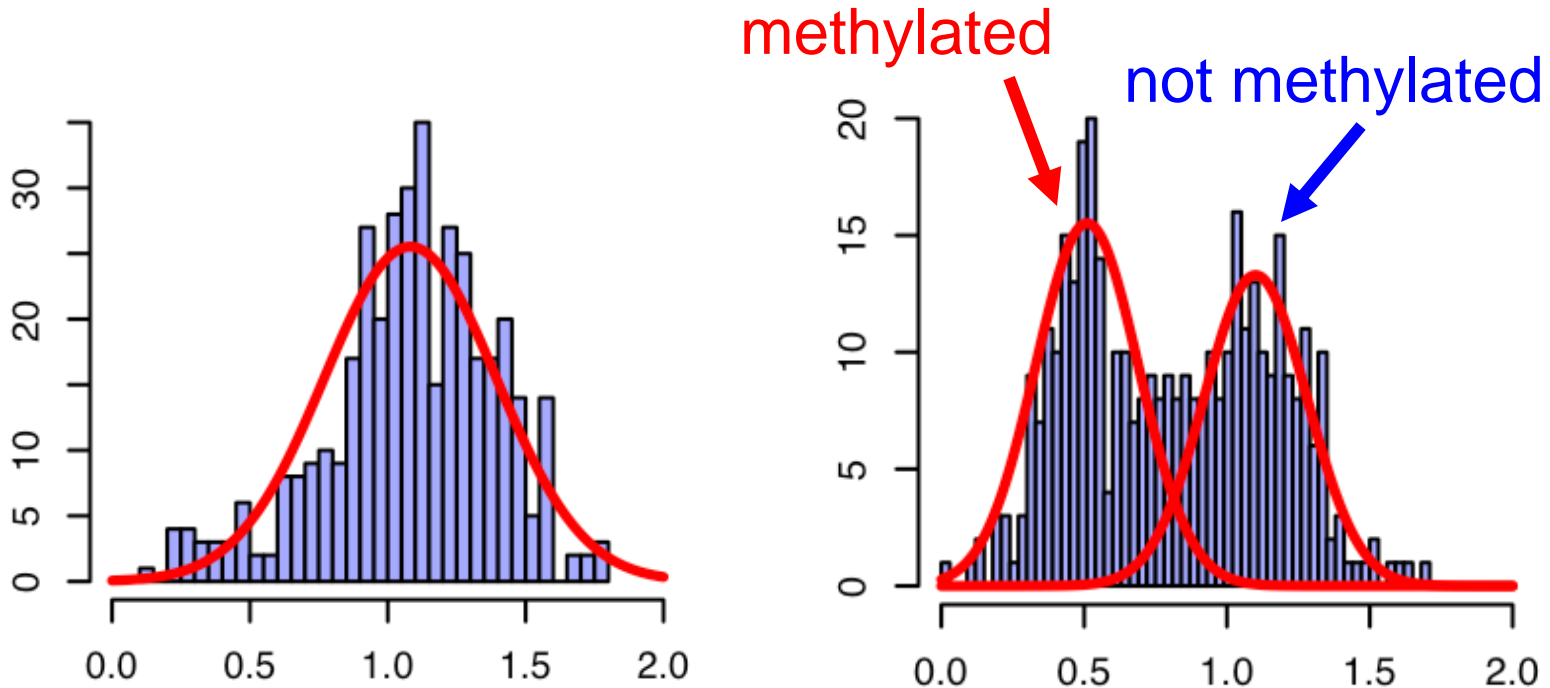
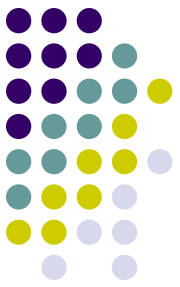


*not methylated*

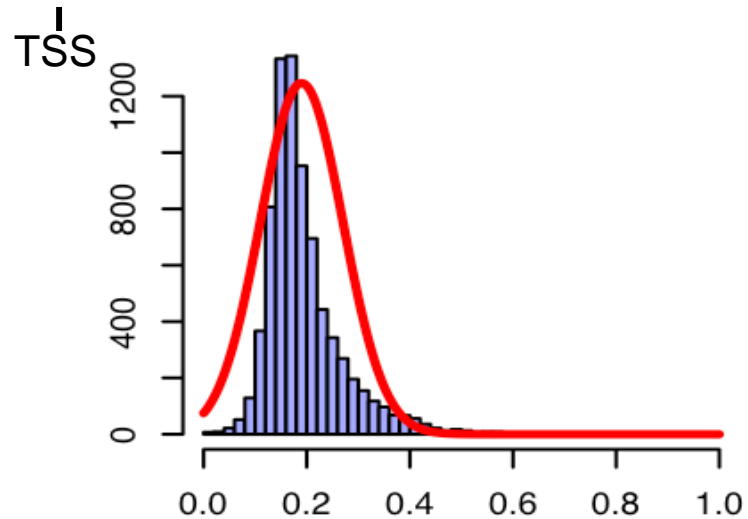
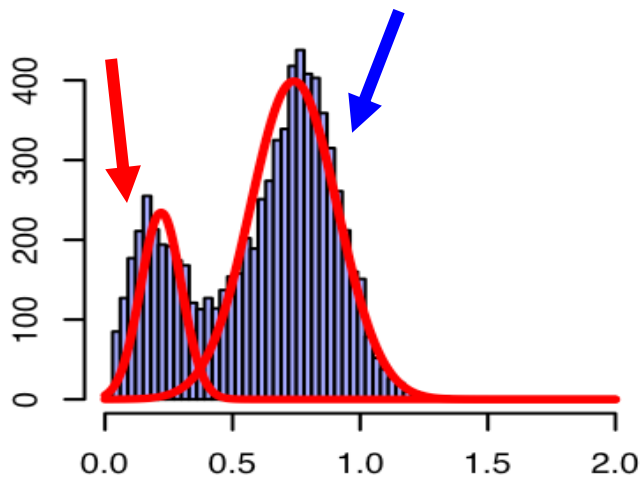
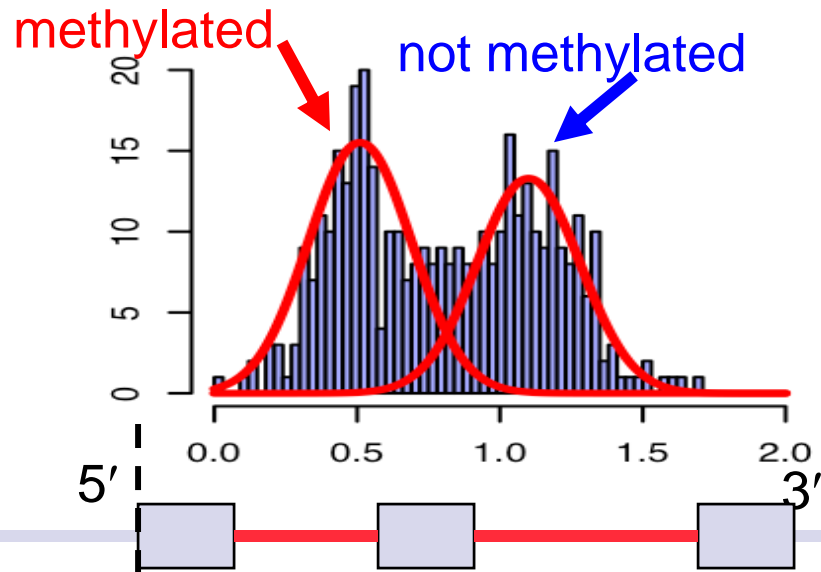
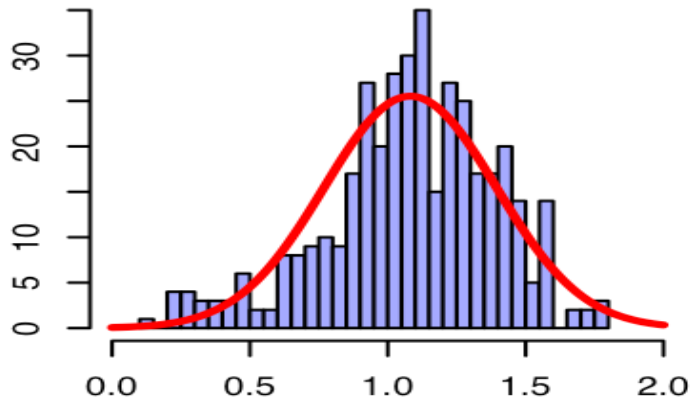
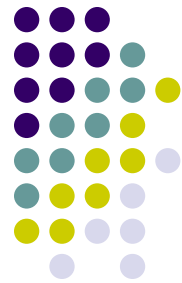
*methylated*

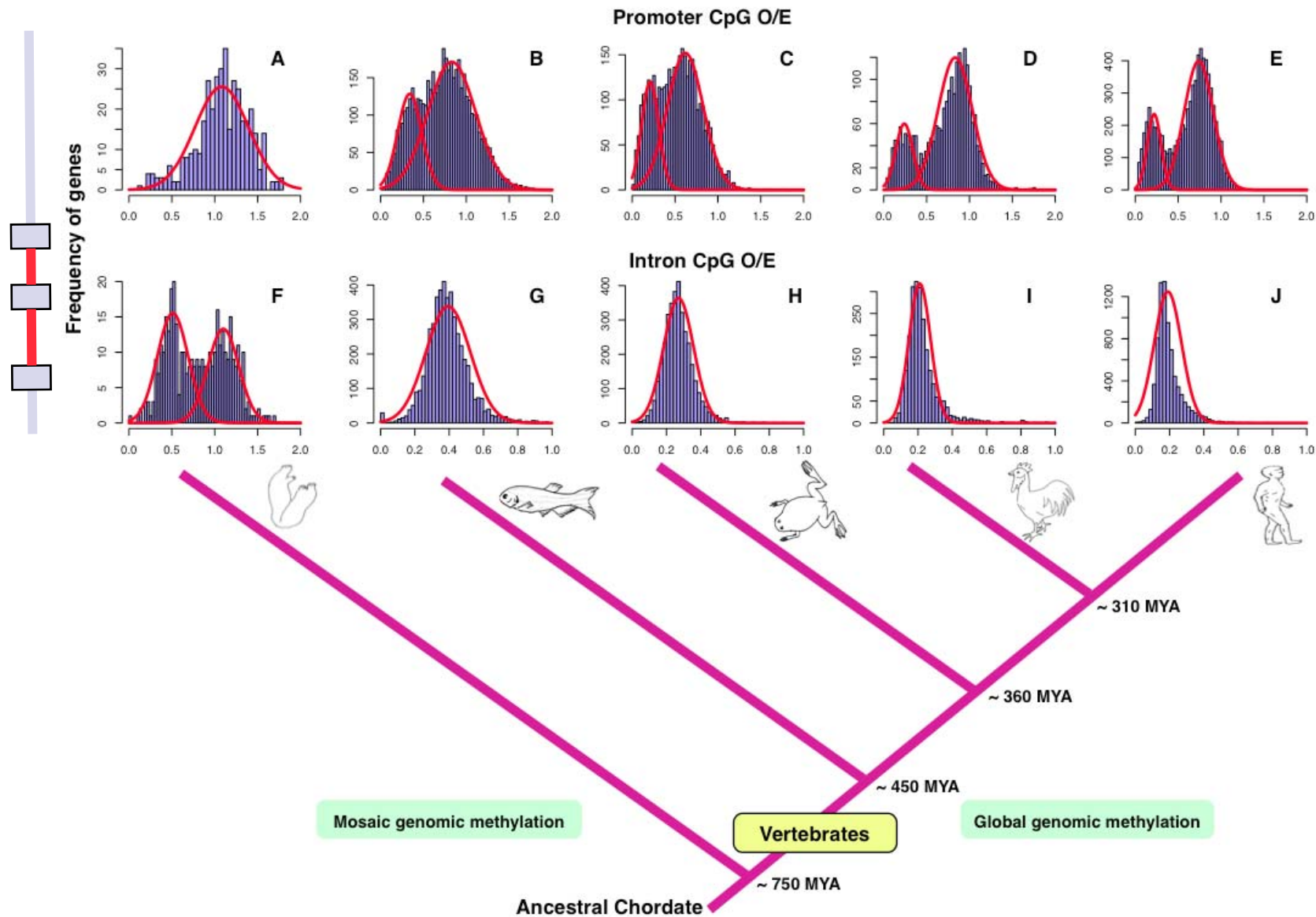


# Distribution of CpG O/E in *C. intestinalis*



# Contrasting patterns of methylation between *C. intestinalis* and *H. sapiens*







QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

Bimodality is a statistically far-better fit to the observed distribution of CpG O/E of vertebrate promoters



QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

Bimodality is a statistically far-better fit to the observed  
distribution of CpG O/E of vertebrate promoters  
LCGs CpG O/E are similar to introns

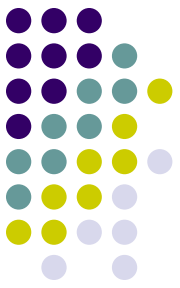


QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

Bimodality is a statistically far-better fit to the observed  
distribution of CpG O/E of vertebrate promoters  
LCGs CpG O/E are similar to introns  
**Are they functionally distinct?**



# Functional consequence of bimodality of vertebrate promoters



- Analyzed EST data

	# tissues	Bottom100	Top100	LCG	HCG
zebrafish	14	21.4 %	42.8 %	28.5 %	35.7 %
frog	21	19.0 %	38.0 %	19.0 %	28.5 %
chicken	18	27.7 %	50.0 %	27.7 %	44.4 %
human	49	14.2 %	67.3 %	20.4 %	61.2 %

Tissue-specific genes are hyper-methylated in diverse vertebrates



QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

Bimodality is a statistically far-better fit to the observed distribution of CpG O/E of vertebrate promoters

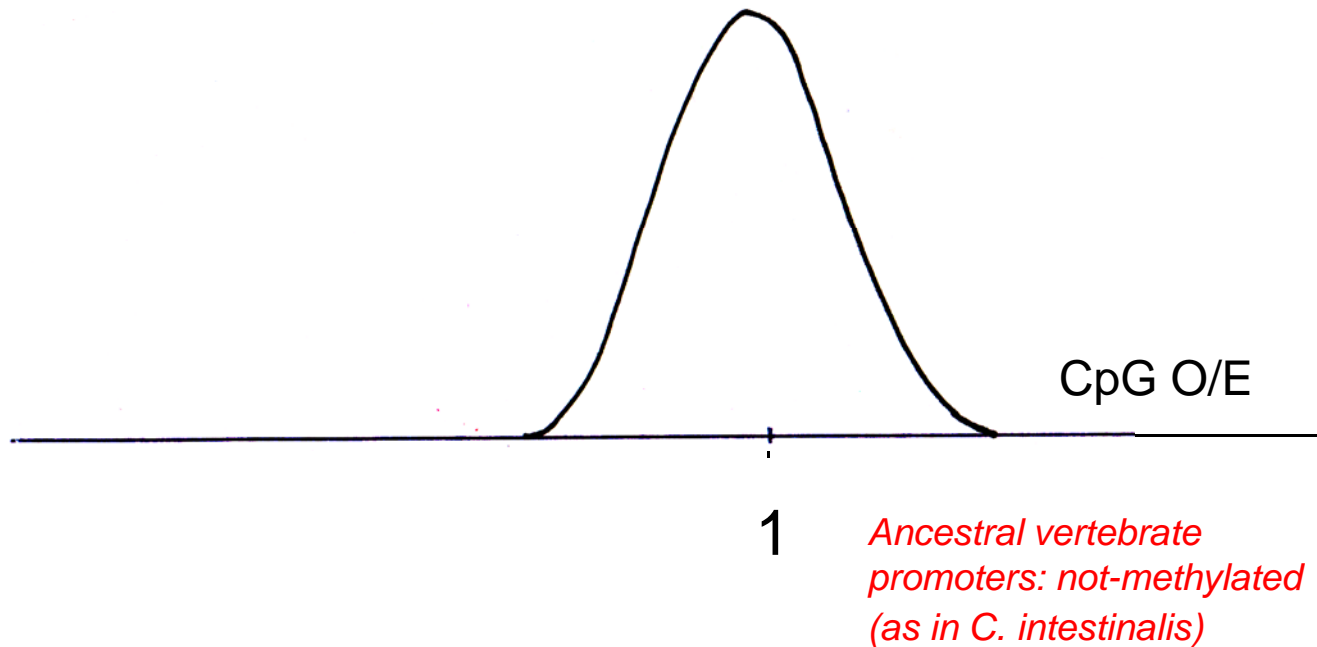
LCGs CpG O/E are similar to introns

**LCG and HCG are functionally distinct:**

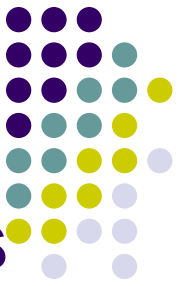
**LCG: tissue-specific**

**HCG: broadly expressed**

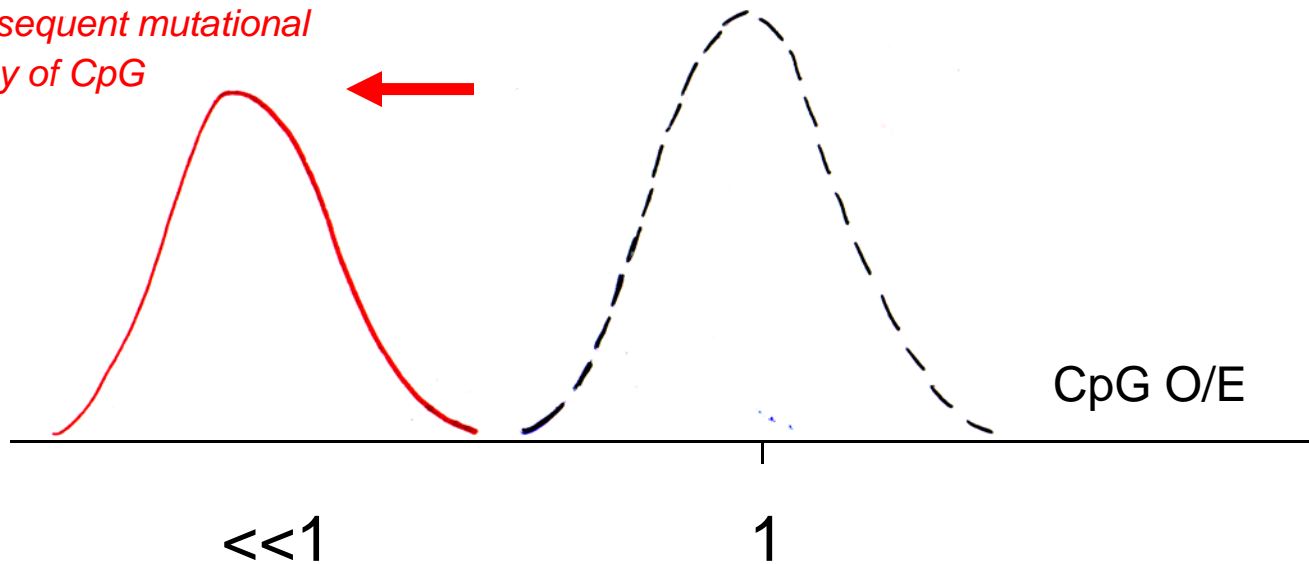
# A model for evolution of promoter bimodality of vertebrates



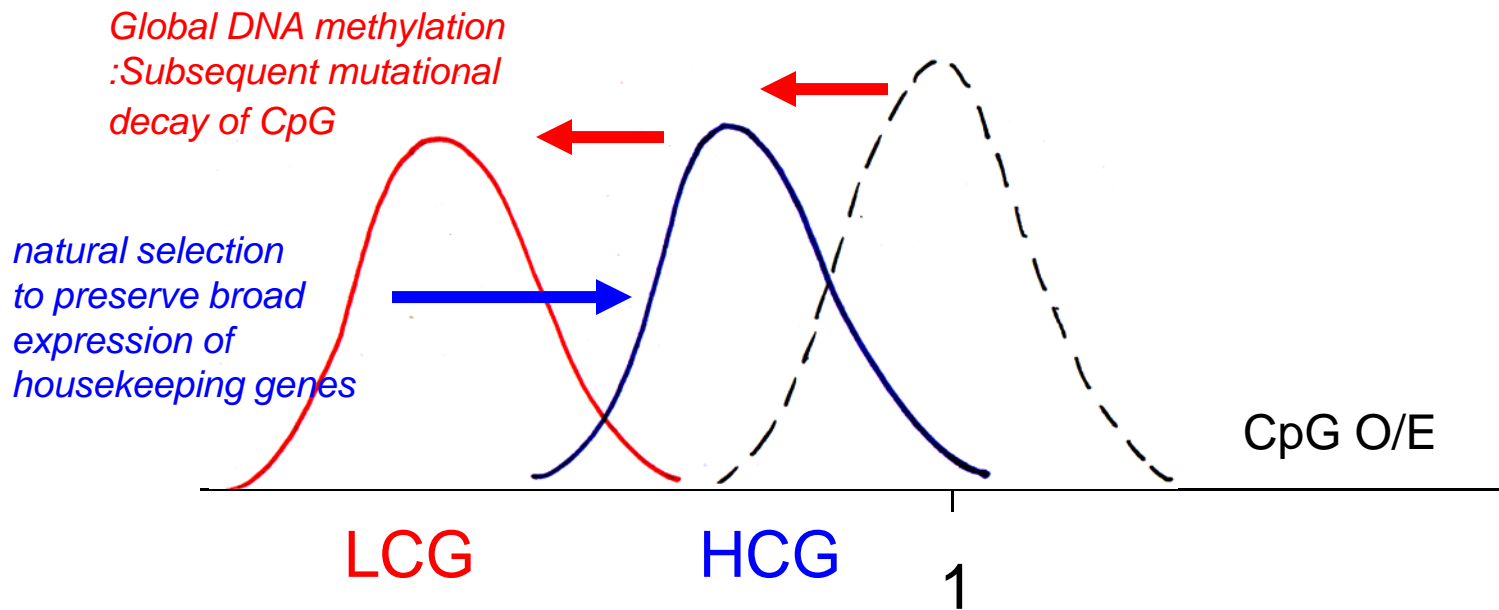
# A model for evolution of promoter bimodality of vertebrates



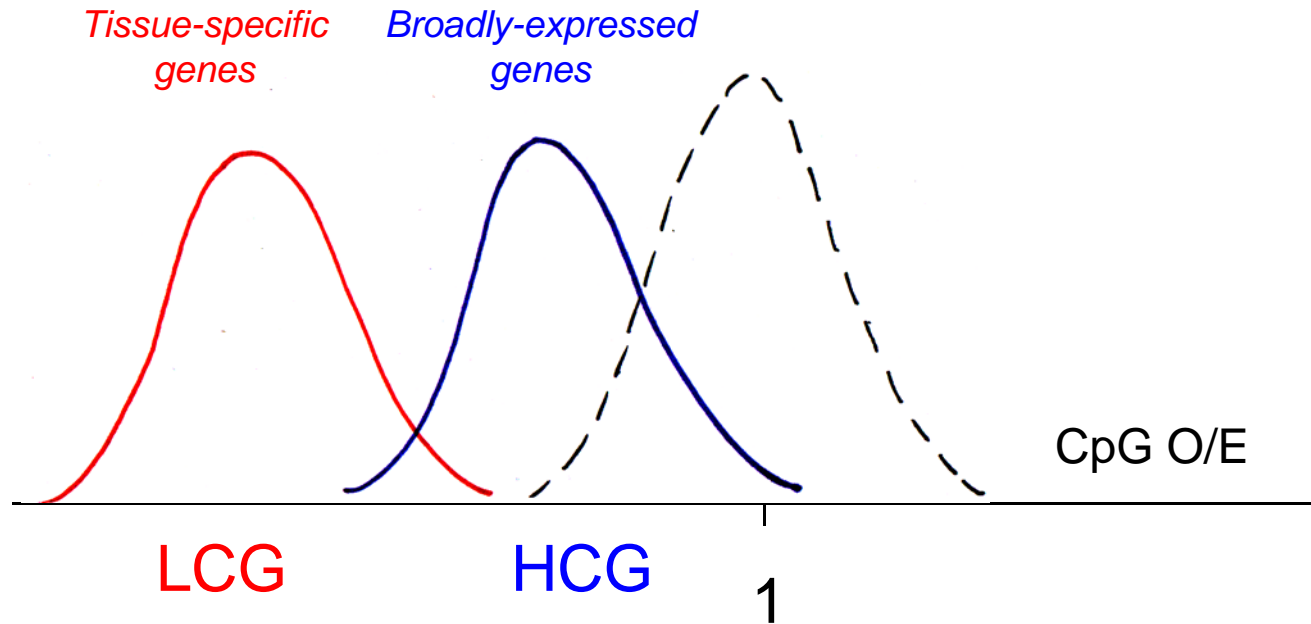
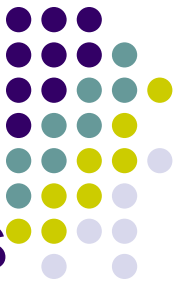
*Global DNA methylation  
:Subsequent mutational  
decay of CpG*



# A model for evolution of promoter bimodality of vertebrates



# A model for evolution of promoter bimodality of vertebrates

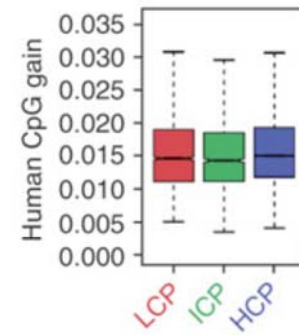
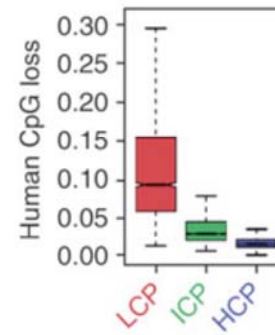
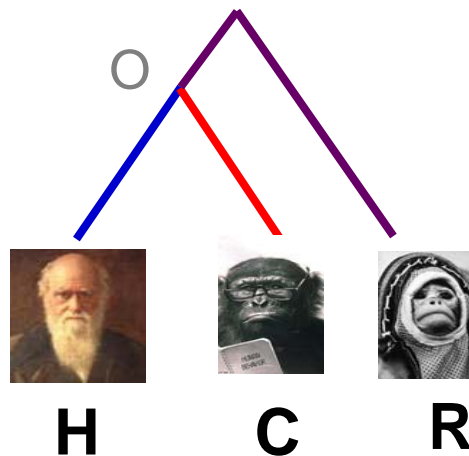


Balance between **mutational decay of LCGs** and **selective preservation of HCGs**

# Evidence for our methylation model for evolution of promoter bimodality I.



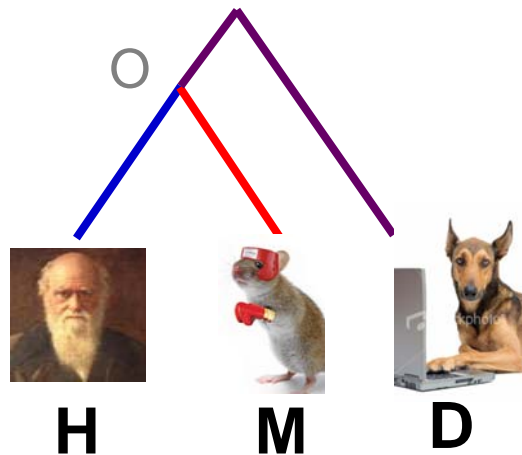
- CpG loss is greater in LCGs than in HCGs
- Comparison of human and chimpanzee genome (using rhesus monkey as an outgroup) Weber et al. Nat. Genet. (2007) 39:457



# Evidence for our methylation model for evolution of promoter bimodality II.



- CpG islands are preferentially lost from the promoters of tissue-specific genes
- Comparison of human and mouse genome (using dog as an outgroup) Jiang et al. Mol. Biol. Evol. (2007) 24:1991







# Conclusions

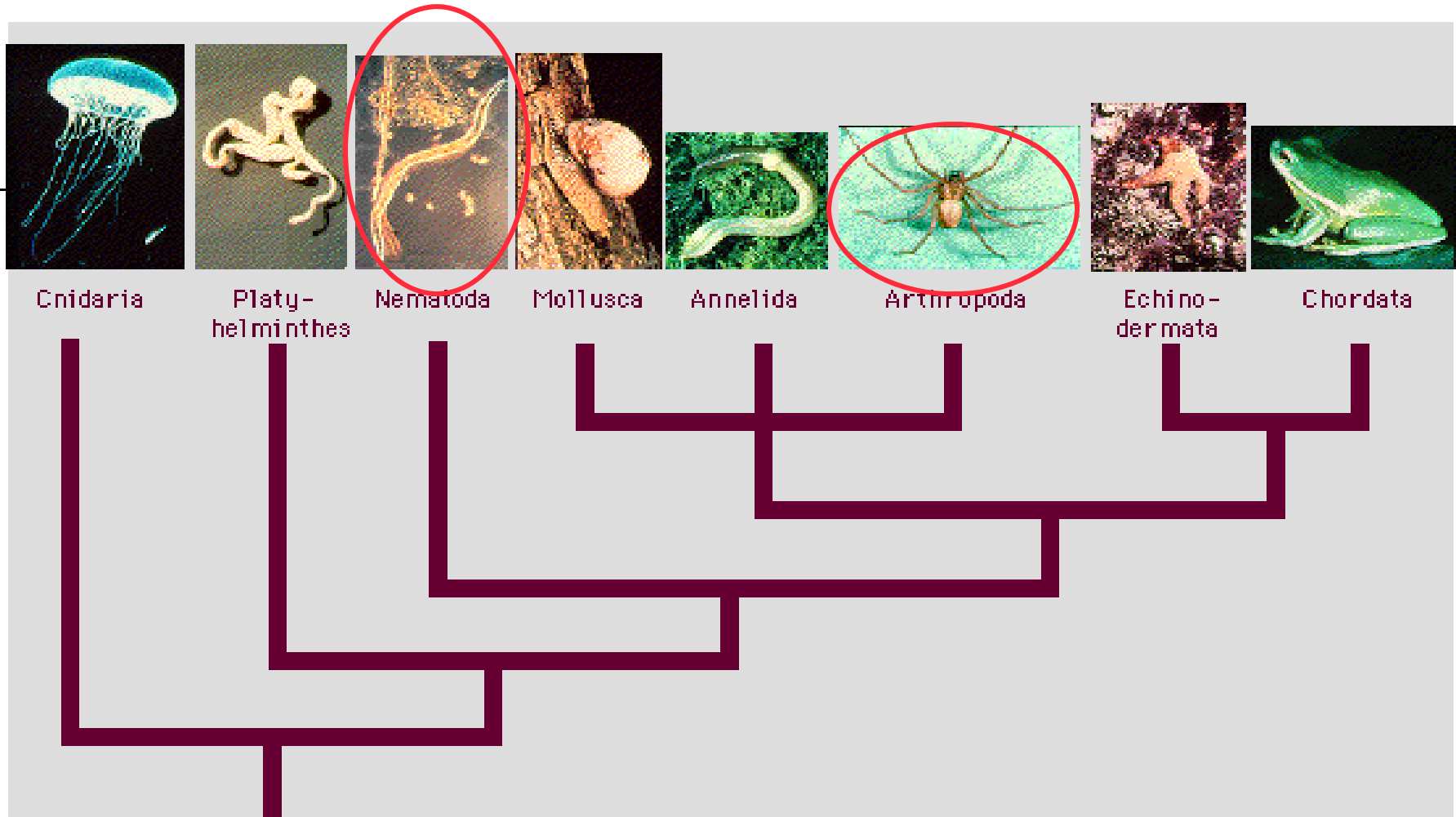
- Vertebrate promoters are bimodal with regard to their CpG contents
- LCG promoters formed due to mutational decay of CpG dinucleotides, caused by the advent of global genomic methylation.
- HCG promoters avoided mutational decay of CpG, and formed CpG islands.

# DNA methylation influences several aspects of genome evolution

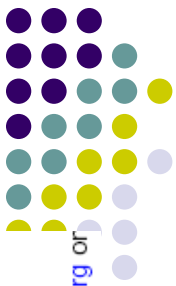


- DNA methylation and molecular clock
  - Kim et al. 2006 PLoS Genetics
- DNA methylation and compositional evolution
  - Elango et al. 2008 PLoS Comp. Biol.
- DNA methylation and evolution of vertebrate promoters
  - Elango & Yi 2008 Mol. Biol. Evol.
- **The pattern and functional role of DNA methylation in the honeybee genome**
  - Elango, Hunt et al. *submitted*

# Distribution of DNA methylation pattern in animal genomes



# Discovery of complete methylation system in honeybee



udy] may occur  
, a small subset of  
fera from *A. m.*  
high differentia-  
Africanized New  
triangles). These  
dM groups from  
)  
potential sites for  
in or near genes  
urred in coding  
lected to cause a  
ange (in a gene  
Although some  
inction, it is also  
hitchhiked with  
under selection.  
val view of the  
*A. mellifera*. Fu-  
large number of  
her investigate  
thin and among  
lection.

eping and Honey  
)9).  
r Bees (CABI, New  
ium, *Nature*, in press.  
ylogen. *Evol.* **37**,  
omy of Honeybees  
dologie (*Celle*) **34**,  
*Celle*) **22**, 627 (1991).  
idologie (*Celle*) **20**,  
(Belknap, Harvard  
pp. 94–95.  
17 (1989).  
767).

morphological U group.  
26. R. E. Page Jr., J. Gadau, M. Beye, *Genetics* **160**, 375  
(2002).

20 July 2006; accepted 5 October 2006  
10.1126/science.1132772

## Functional CpG Methylation System in a Social Insect

Ying Wang,<sup>1</sup> Mireia Jorda,<sup>2</sup> Peter L. Jones,<sup>1</sup> Ryszard Maleszka,<sup>3</sup> Xu Ling,<sup>4</sup>  
Hugh M. Robertson,<sup>5,6,7</sup> Craig A. Mizzen,<sup>1,6</sup> Miguel A. Peinado,<sup>2</sup> Gene E. Robinson<sup>5,6,7\*</sup>

DNA methylation systems are well characterized in vertebrates, but methylation in *Drosophila melanogaster* and other invertebrates remains controversial. Using the recently sequenced honey bee genome, we present a bioinformatic, molecular, and biochemical characterization of a functional DNA methylation system in an insect. We report on catalytically active orthologs of the vertebrate DNA methyltransferases Dnmt1 and Dnmt3a and b, two isoforms that contain a methyl-DNA binding domain, genomic 5-methyl-deoxycytosine, and CpG-methylated genes. The honey bee provides an opportunity to study the roles of methylation in social contexts.

Among the many important functions of CpG DNA methylation, sex-specific regulation of gene expression (imprinting) in vertebrates stands out because it provides insight into intragenomic conflict (1, 2). Provided social insects have CpG methylation, they would be ideal models to further explore the kin-conflict theory of imprinting, because insect societies are composed of many different types of relatives and they interact with each other in many evolutionarily important contexts (2, 3).

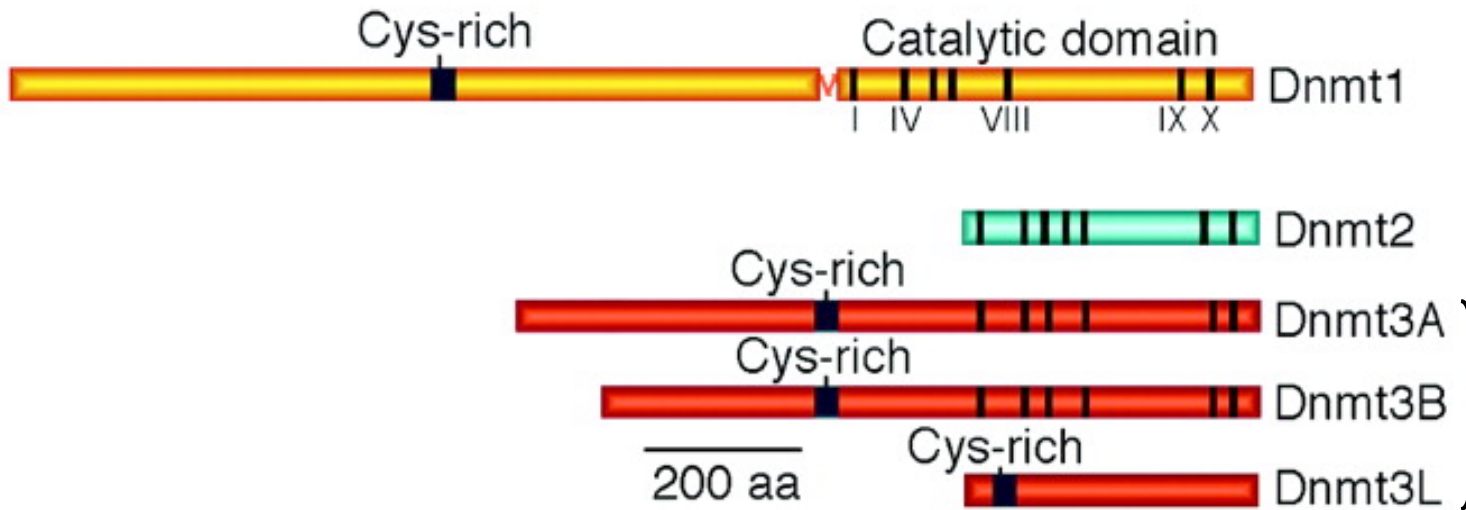
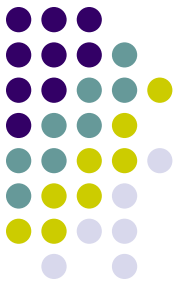
However, although widely conserved from yeast and fungi to plants to vertebrates, DNA methylation in insects is enigmatic. Evidence of CpG-methylated sequences exists for several insect species (4–6), but no bona fide invertebrate deoxycytosine methyltransferases (DNMTs) have been described. Conversely, the model insect *Drosophila melanogaster* shows limited DNA

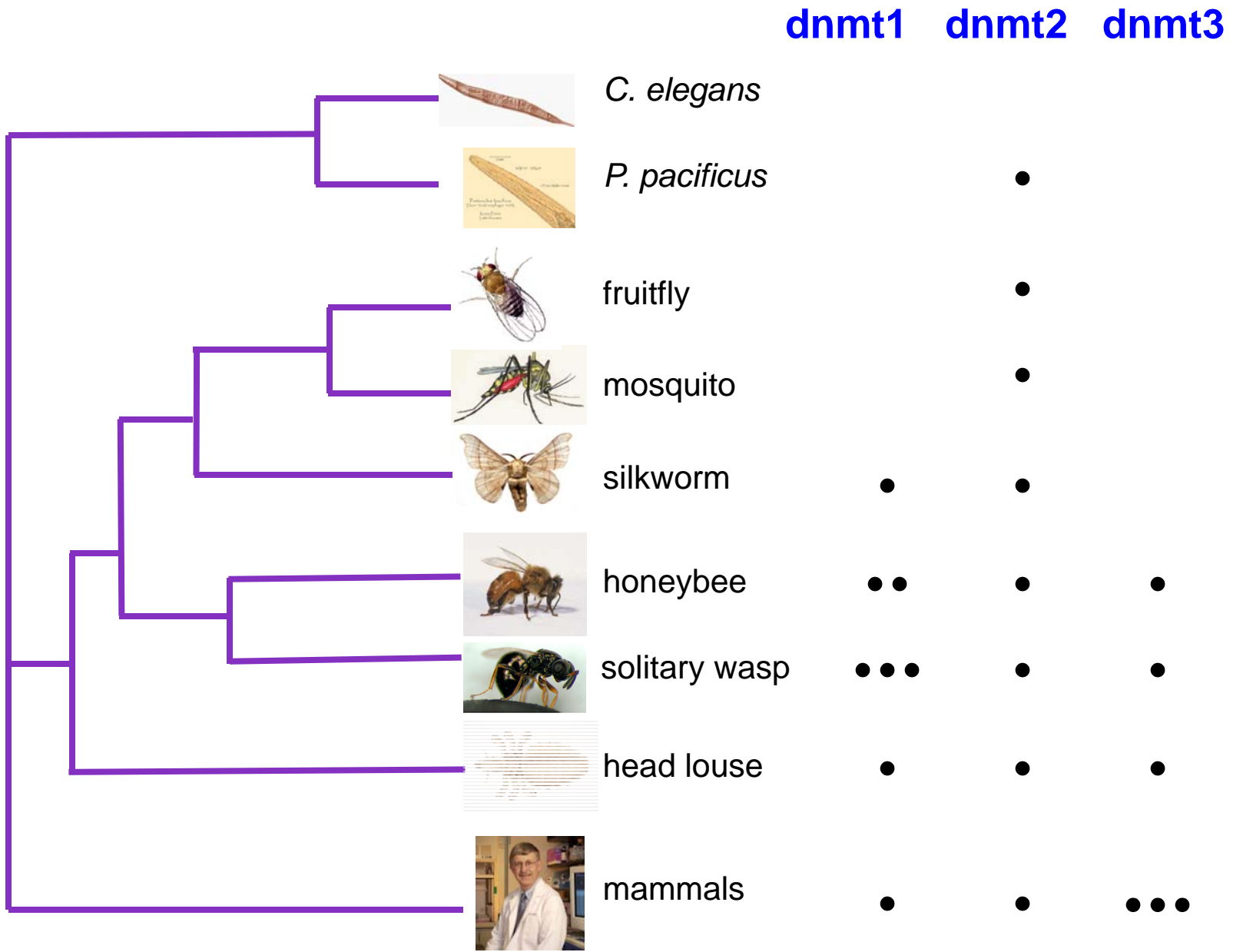
methylation, predominantly in asymmetric CpT and CpA dinucleotides (7), and this is attributed to the only DNMT family member encoded in its genome, dDNMT2, a tRNA<sup>Asp</sup> methyltransferase

<sup>1</sup>Department of Cell and Developmental Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. <sup>2</sup>DIBELL—Cancer Research Institute, L'Hospitalet 08907, Catalonia, Spain. <sup>3</sup>Visual Sciences and Australian Research Council Centre for Molecular Genetics of Development, Research School of Biological Sciences, The Australian National University, Canberra, Australia. <sup>4</sup>Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. <sup>5</sup>Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. <sup>6</sup>Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. <sup>7</sup>Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

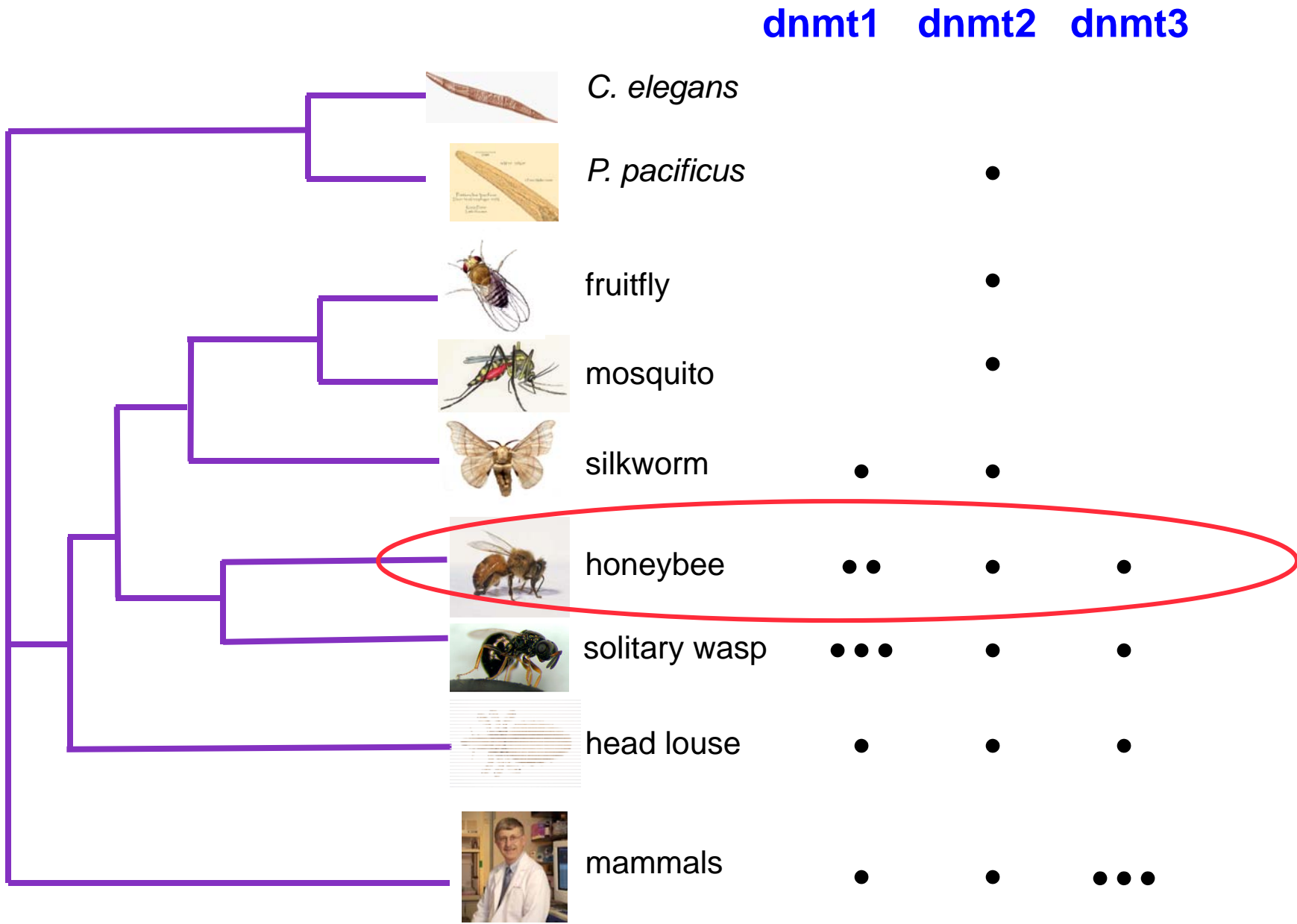
\*To whom correspondence should be addressed. E-mail: generobi@life.uiuc.edu

# Dnmt (DNA methyltransferases) in the human and mouse genome:





Sources: Maleszka 2008; Suzuki & Bird 2008



Sources: Maleszka 2008; Suzuki & Bird 2008

# The pattern and role of DNA methylation in honeybee

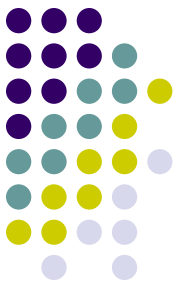


- What is the pattern of genomic methylation in honeybee?
  - Patchy or global?
- What is the role of DNA methylation in honeybee?
  - To silence transposable elements
  - To check transcription of housekeeping genes
  - Related to caste differentiation

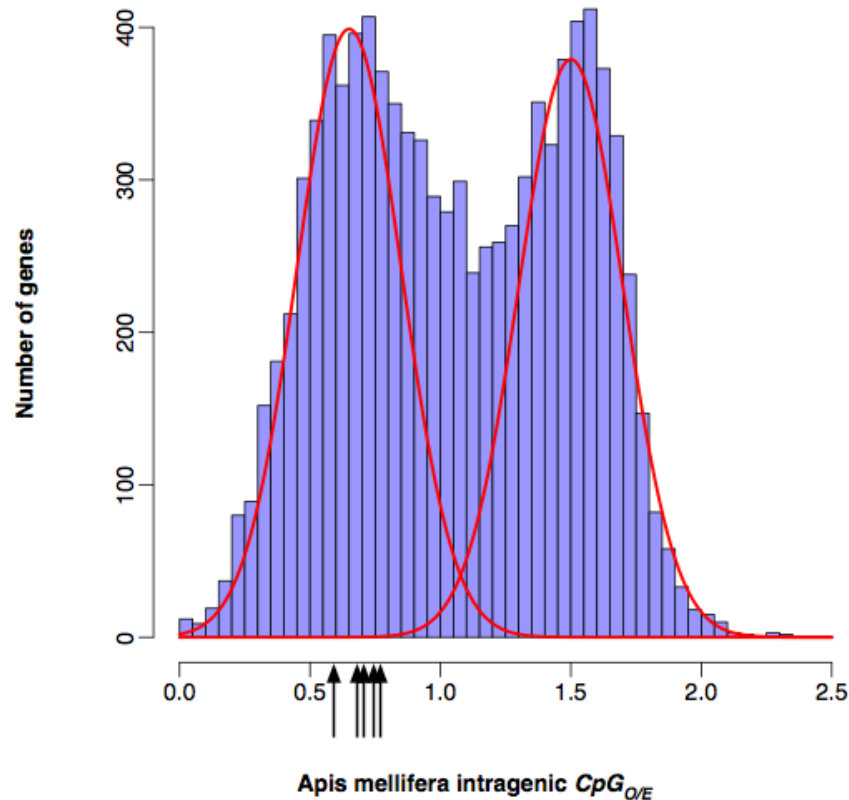


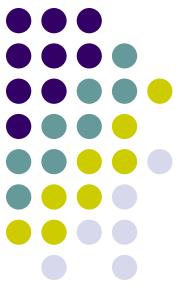
- What is the pattern of genomic methylation in honeybee?
  - Analyzed the distribution of CpG O/E in honeybee genes



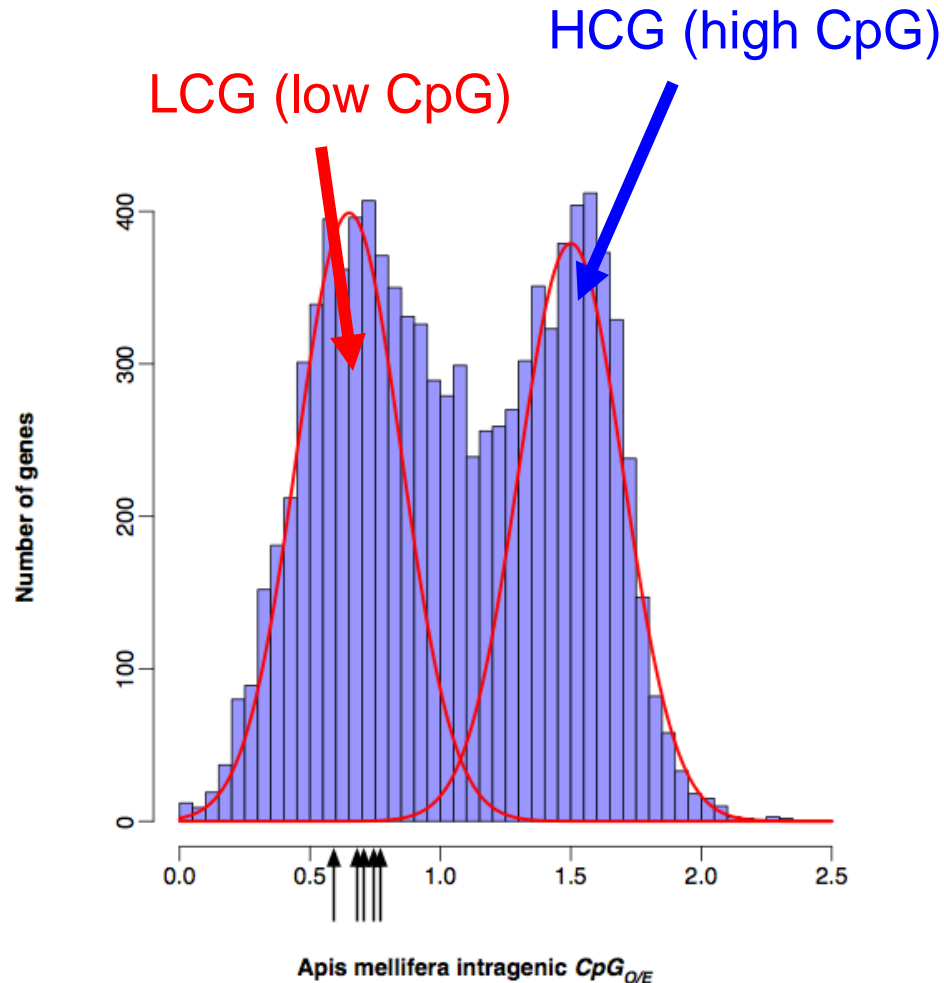


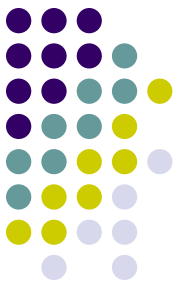
- What is the pattern of genomic methylation in honeybee?
  - **Bimodal** distribution of CpG O/E in intragenic regions





- What is the pattern of genomic methylation in honeybee?
  - **Bimodal** distribution of CpG O/E in intragenic regions

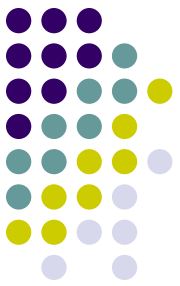




- What is the pattern of genomic methylation in honeybee?
  - Bimodal distribution of CpG O/E in intragenic regions
    - LCG: methylated genes
    - HCG: non-methylated genes
  - Transposable element:
    - Analyzed the mariner element
    - CpG O/E of the mariner element: ~1
    - TE not likely to be a main target of methylation
  - Likely to be 'patchy' methylation (similar to *Ciona*)

- What is role of DNA methylation in honeybee?
  - Analyzed functions of LCG and HCG genes





- What is role of DNA methylation in honeybee?
  - Analyzed functions of LCG and HCG genes
  - LCG and HCG genes enriched in functionally distinct GO terms

LCG	HCG
Translation	Cellular communication
Biosynthesis	Organ and system development
Response to DNA damage	Signal transduction
	Cell adhesion



- What is role of DNA methylation in honeybee?
  - Analyzed functions of LCG and HCG genes
  - LCG and HCG genes enriched in functionally distinct GO terms

<i><b>'housekeeping'</b></i>	<i><b>'development'</b></i>
LCG	HCG
Translation	Cellular communication
Biosynthesis	Organ and system development
Response to DNA damage	Signal transduction
	Cell adhesion

Elango, Hunt et al. *submitted*

- DNA methylation and phenotypic plasticity
- Social insects: model of phenotypic plasticity
- **Highly social (eusocial) behavior**

- Overlap of generations
- Cooperative brood care
- Reproductive division of labor  
*often marked by caste polyphenism*



Worker Adult

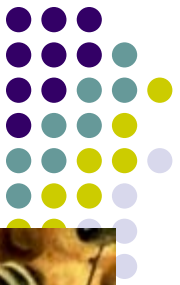


Queen Adult



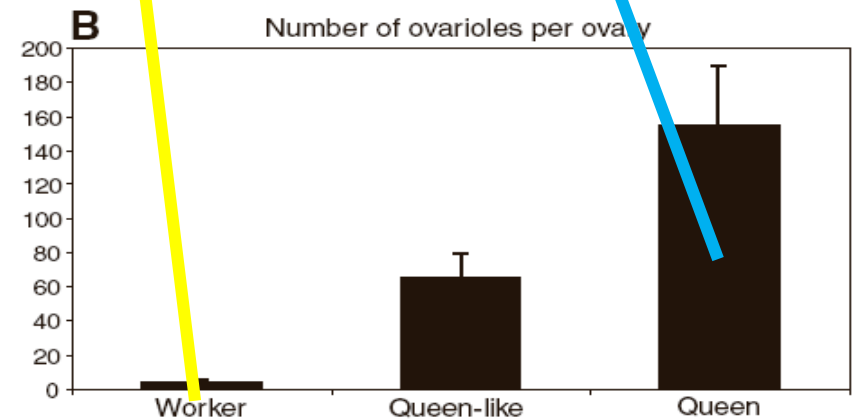
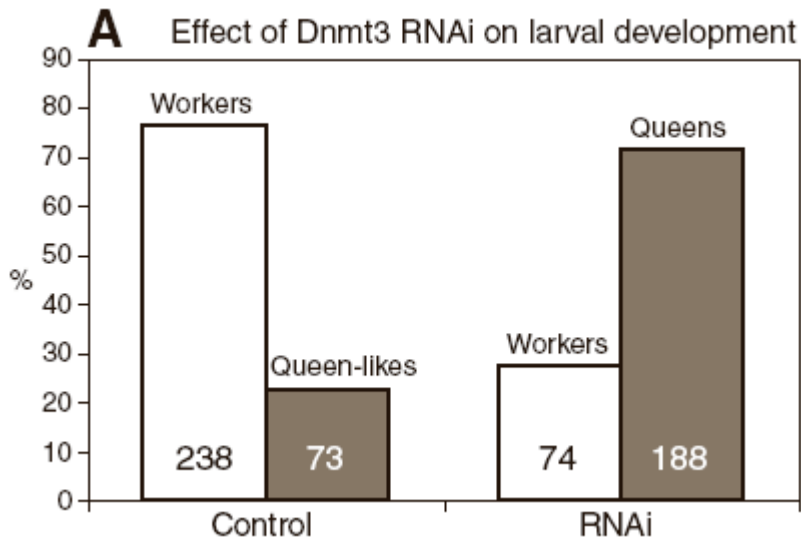


# Dnmt3 silencing in honeybees

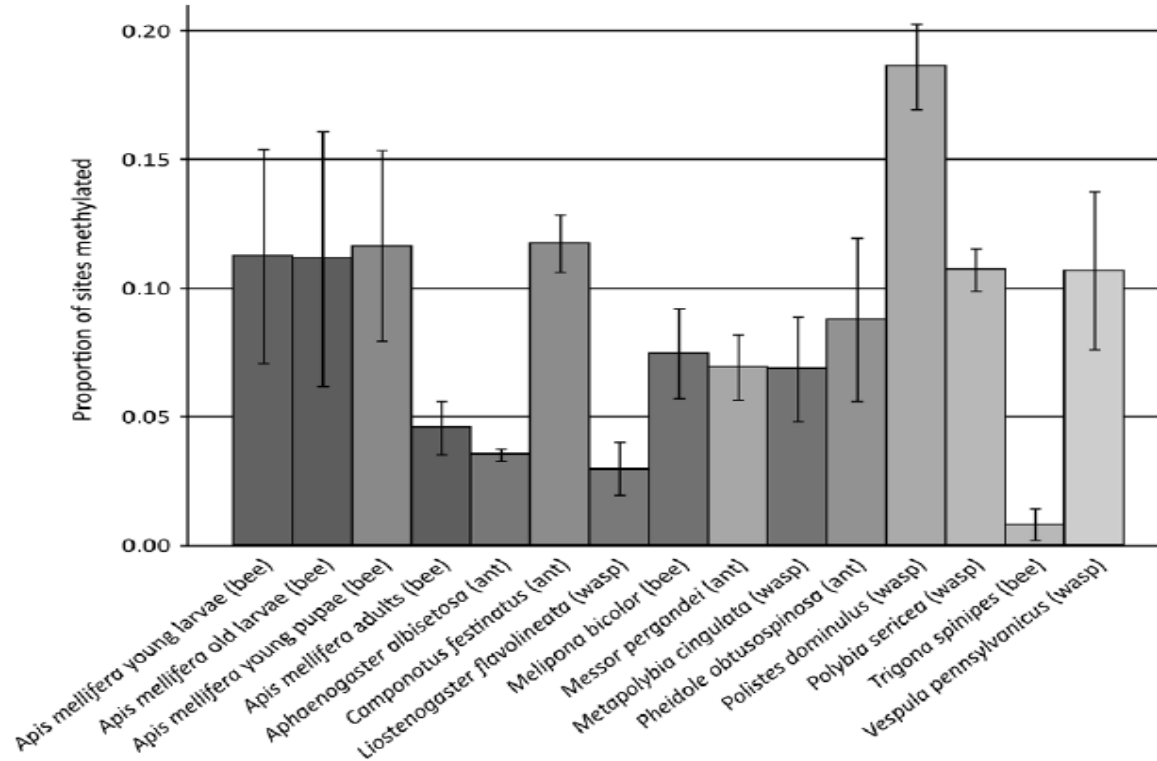
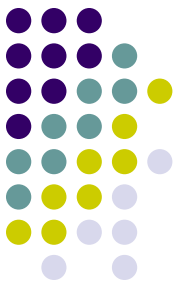


## Nutritional Control of Reproductive Status in Honeybees via DNA Methylation

R. Kucharski,\* J. Maleszka,\* S. Foret, R. Maleszka†  
Science 2008



# Epigenetic regulation: widespread methylation in social insects



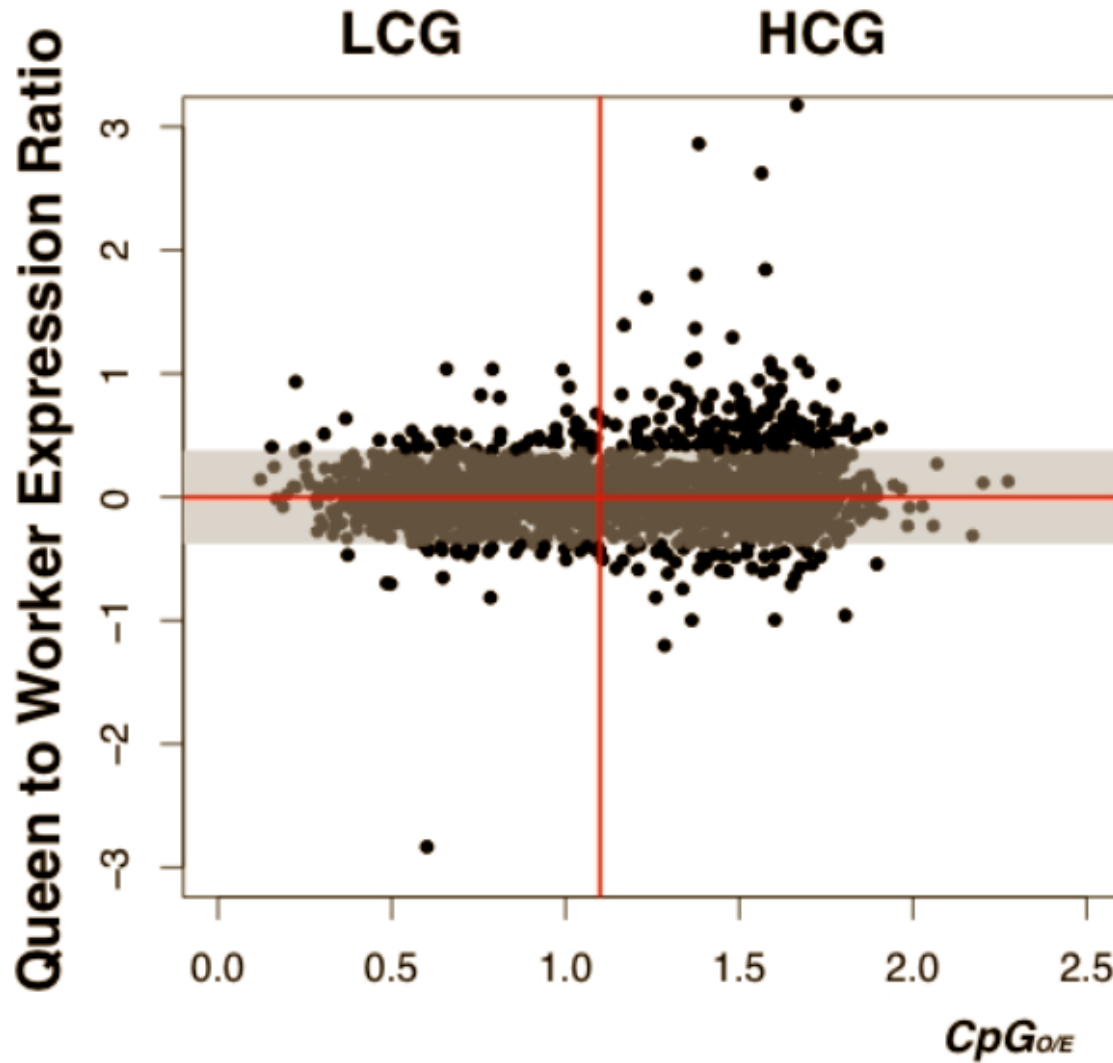
Proportion of methylated restriction sites as the number of AFLP fragments exhibiting evidence of methylation.

Kronforst et al. 2008, Current Biology



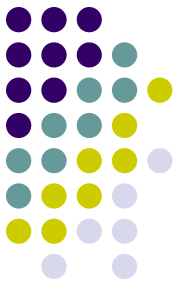
- Methylation status of honeybee genes may be associated with caste-specific genes: e.g. workers develop when queen-specific genes are silenced by methylation
  - Prediction I: Queen-specific genes should be enriched for the hypomethylated class (HCG)
  - Prediction II: If genes hypomethylated in the germline are more accessible to regulation via methylation, caste-specific genes should be enriched for HCG

# Prediction I: Queen-specific genes should be enriched for HCG



$P \sim 0.06$

## Prediction II: caste-specific genes should be enriched for HCG



	LCG	HCG
Caste-specific	75 (127.74)	185 (132.25)
Caste-generic	973 (920.25)	900 (952.75)

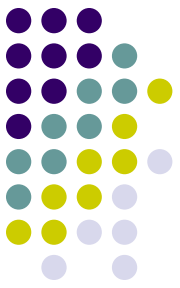
$P < 10^{-10}$   
(Chi-square test)

- Genes with caste-biased expression from small-scale studies overwhelmingly belong to the HCG class (22/28)

# The pattern and role of DNA methylation in honeybee



- What is the pattern of genomic methylation in honeybee? **patchy**
- What is the role of DNA methylation in honeybee?
  - To silence transposable elements ✘
  - To check transcription of housekeeping genes 📄
  - Related to caste differentiation 📄



# Future Directions

- Evolutionary point of bimodality by analyzing early vertebrates (jawless fish) and ancestral chordates such as amphioxus
- Analyses of other invertebrate genomes for the signature of DNA methylation using CpG O/E and direct measurement of methyl-cytosines
- Experimental analyses of DNA methylation in honeybee and yellowjackets
  - Are the levels of DNA methylation and the level of gene expression correlated?
  - Are different castes show different levels of DNA methylation?



# Acknowledgement

- Co-authors, collaborators: Navin Elango, Brendan Hunt, Michael Goodisman
- Lab members/Alumni: Zuogang Peng, Tim Nowack, Varun Kilau, Nitya Sharma, Jeeyoung Lee, Eddie Loh, Ke Xu, Seongho Kim, Charles Warden
- NSF, Alfred P. Sloan Foundation, Blanchard-Milliken Faculty Fellowship

*Thank you for your attention!*





# CpG substitution and CpG islands

Over evolutionary timescale, CpG will gradually be depleted by DNA methylation

CpG islands: clusters of CpG dinucleotides in higher frequency than the rest of the genome

Some ad hoc criteria:

> 200 (or 500)

GC > 55%

CpG Observed more than 50% of the expected

Mostly found in 5' of genes

Some intragenic CpG islands

