Epigenetics and palms
What’s beyond the double helix...

Alain RIVAL, Estelle JALIGOT, Thierry BEULE, Frédérique RICHAUD, Pascal ILBERT

UMR DIADE
Palm Development Group
Cirad BioS
Montpellier
France.

XII EUNOPS Meeting, Sanremo, Italy, May 12-13th, 2012
What is Epigenetics?

« The interactions of genes with their environment, which bring the phenotype into being ».  


“Changes in gene expression not attributable to nucleotide sequence variation”.

“Things written in pen you can’t change. That’s DNA,” says geneticist Danielle Reed. “Things written in pencil you can. That’s epigenetics.”
Several new sequencing technologies are emerging that have the potential to provide **increases in throughput and reductions in cost**.

Companies such as 454 Life Sciences, Solexa, and Helicos Biosciences all have **competing technologies**, vying to be widely adopted for the next generation of sequencing machines.

The success of genome sequencing project has created wide-spread interest in exploring **epigenomes** in order to elucidate how the genome executes the information it holds.
• A mammal diploid cell contains in average 6000 Mb of DNA, corresponding to 1.8 meters, which is packed within a nucleus of 6 microns diameter, following a strongly organized packaging.

• Each chromosome yields one single molecule of linear DNA which is folded several times and distributed over the centromere.

• Within the nucleus, DNA is never encountered under a free form but associated to other molecules, principally histones. Histones are small basic proteins which are present with the same proportion as DNA.

• Other proteins associated to DNA are non histones acidic proteins, which account for 10 to 30% of the whole structure.

• This DNA-Protein complex is called chromatin (as it is easily stained) and it constitutes the chromosomes.
DNA increasing orders of packing
Heterochromatin refers to regions of the genome that have low gene density, contain satellite repeat elements and are late replicating.

Euchromatin is a lightly packed form of chromatin that is rich in gene concentration, and is often under active transcription. Euchromatin comprises the most active portion of the genome within the cell nucleus.

Heterochromatin and euchromatin are associated with distinct DNA methylation and histone modification patterns that correlate with particular states of gene activity.
Acetylation of histones loosens the chromatin, facilitating replication and transcription.

Methylated histones hold the DNA more tightly, thus restricting access and impeding transcription.

Methylation of Lysine 4 and Lysine-27 on Histone3 may be involved in development.

Specifically, there is fewer methylated Lysine 27 in the chromatin of differentiated cells.

Lysine-4 methylation acts to promote transcription by recruiting nucleosome remodeling enzymes and histone acetylase.
Transcriptionally active chromatin regions tend to be hyperacetylated and hypomethylated.

If a region of DNA or a gene is destined for silencing, chromatin remodeling enzymes such as histone deacetylases and ATP-dependent chromatin remodelers likely begin the gene silencing process.

One or more of these activities may recruit DNA methyltransferase resulting in DNA methylation, followed finally by recruitment of the methyl-CpG binding proteins.

The region of DNA will then be heritably maintained in an inactive state.
Epigenomics is the large scale study of epigenetic marks on the genome including:

- Covalent modifications of histone tails (acetylation, methylation, phosphorylation, ubiquitination)
- DNA methylation.
- Small RNAs

Epigenetic components are all amenable to genome-wide studies

Integrated approaches that correlate gene expression with DNA methylation and chromatin profiles are being designed.
From genome to epigenome

Murell et al, Human Molecular Genetics 2005 14 (Review Issue 1):
Distribution patterns and transcription activity

Detailed distribution patterns and transcription activity (vertical blue bars) in a gene-rich region (A) and a repeat-rich region (B).

Red boxes: genes; Arrows indicate the direction of transcription.

Distribution of genes, repetitive sequences, DNA methylation, siRNAs, H3K27me3, and low nucleosome density (LND) regions in Arabidopsis.

The chromosomal distributions use chromosome 1 as an example. The x axis shows chromosomal position.

The small RNAs world

- These tiny RNAs (~21-26 nt) induce silencing through homologous sequence interactions
- They can control mRNA stability or translation, or target epigenetic modifications to specific regions of the genome.
- Small RNAs and evolutionarily conserved RNA-mediated silencing pathways have established a new paradigm for understanding eukaryotic gene regulation and revealed novel host defenses to viruses and transposons.
MicroRNAs (miRNAs) and transacting siRNAs (tasiRNAs) are primarily involved in regulating gene expression and plant development,

siRNAs play a major role in defending the genome against the proliferation of invading viruses and endogenous transposable elements.

The function of the fourth type of sRNAs, natural-antisense siRNAs (nat-siRNAs), is not entirely clear but is likely related to plant stress responses.

Zhang et al., PNAS 104, 4536 (2007)
Deciphering the small RNAs machinery

RNA-directed DNA methylation (RdDM), first discovered in plants requires a dsRNA that is processed to 21-24 nt small RNAs.

In Arabidopsis, links between locus specific small RNAs, DNA methyltransferases, and histone modifications, including deacetylation and histone H3 lysine 9 (H3K9) methylation have been identified (see red pathway).

Some small RNAs might target native promoters of endogenous genes (see dashed blue pathway).

RNAi and epigenetic alterations of the genome, such as DNA methylation and histone modifications.
Putative pathway for RNA directed DNA methylation in \textit{A. thaliana}. Target loci (in this case tandemly repeated sequences; coloured arrows) recruit an RNA polymerase IV complex consisting of NRPD1A and NRPD2 through an unknown mechanism, and this results in the generation of a single-stranded RNA (ssRNA) species. This ssRNA is converted to double-stranded RNA (dsRNA) by the RNA-dependent RNA polymerase RDR2.

The dsRNA is then processed into 24-nucleotide siRNAs by DCL3. The siRNAs are subsequently loaded into the protein AGO4, which associates with another form of the RNA polymerase IV complex, NRPD1B-NRPD2. AGO4 that is ‘programmed’ with siRNAs can then locate homologous genomic sequences and guide the protein DRM2, which has \textit{de novo} cytosine methyltransferase activity. Targeting of DRM2 to DNA sequences also involves the chromatin remodelling protein DRD1.

The 'epigenetic code' considerably extends the information potential of the genetic code.

Thus, one genome can generate many 'epigenomes' as the fertilised egg progresses through development and translates its information into a multitude of cell fates.

The transcriptomes of an organism are continually changing in response to developmental and environmental cues.

The epigenome is not static and can be molded by developmental signals, environmental perturbations, and disease states.

Therefore, many epigenomes will need to be sequenced for a single organism, making epigenome sequencing perhaps even more challenging than genome sequencing.
The search for epigenetic factors of flower development in oil palm

Sylvie Doulbeau (IRD), Frédérique Aberlenc (IRD), James Tregear (IRD), Estelle Jaligot (Cirad), Pascal Ilbert (Cirad), Thierry Beulé (Cirad)
Rationale

- The commercial development of large scale propagation of oil palm is hampered by the generation of somaclonal variants affecting the floral architecture: the « mantled » genotype.

- Field observation of clonal plantings on the long term and results from DNA markers analysis (Rival et al, 1998) are consistent with an epigenetic origin for somaclonal variation.

- The “mantled” phenotype is correlated with a global hypomethylation of genomic DNA, which was revealed through a whole genome approach (Jaligot et al, 2000) and the use of Methylation-Sensitive DNA markers, such as MS-RFLPS and MSAPs (Jaligot et al, 2002; 2004).
RESEARCH PAPER

Isolation and expression analysis of genes encoding MET, CMT, and DRM methyltransferases in oil palm (*Elaeis guineensis* Jacq.) in relation to the ‘mantled’ somaclonal variation

Alain Rival¹••*, Estelle Jaligot¹••, Thierry Beulé¹ and E. Jean Finnegan²

¹ CIRAD, IRD, UMR DIAPC, BP 64501, F-34394 Montpellier, Cedex 5, France
² CSIRO Plant Industry, GPO Box 1600, ACT 2602, Canberra, Australia

Received 23 April 2008; Revised 9 June 2008; Accepted 10 June 2008
Establishment and Maintenance of DNA methylation

siRNA-generating pathway

Establishment

DRM1/DRM2

Maintenance

CG

CNG

CHH (asymmetric)

DDM1 Chromatin remodelling

HDMS Histone deacetylation

MET1

CMT3

KYP Histone H3K9 methylation

Unknown protein Histone H3K9 methylation

siRNAs

RDR2 DCL3 RPD1 AGO4

RDR6 SDE2 SDE3 AGO1

OTHER?
Several oil palm MADS box genes have shown differential expression patterns according to the presence of *mantled* abnormality.

Alterations in expression affect not only B-type, but also C, D and E-type genes.

Whorls 2, 3 and 5 are affected by homeotic changes.
Methylation around MADS Box candidate genes

Reduced expression of genes Eg DEF1 and EgGLO2 (B type), EgAG2 (C and D type) and EgS1 (type E) in abnormal oil palm flowers

Thank you for your kind attention ...

alain.rival@cirad.fr