Epigenetic databases in cattle and the prediction of phenotypes from genotypes

Jean-Paul Renard — Académie d'Agriculture de France - Paris

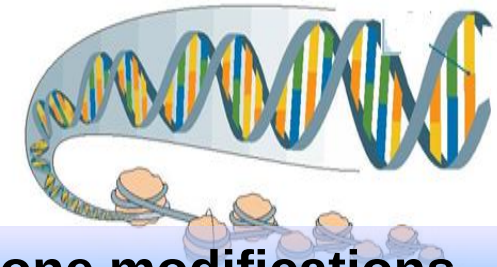
Hélène Jammes - Hélène Kiefer — INRA UMR 1198- Jouy en Josas

Eve Devinoy — INRA GABI- 1113- Jouy en Josas

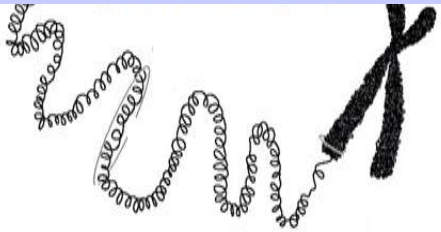
66th EAAP Annual Meeting
30 August – 4th september 2015
Warsaw, Poland

Between genotype and phenotype…

…3 levels of genes regulation…



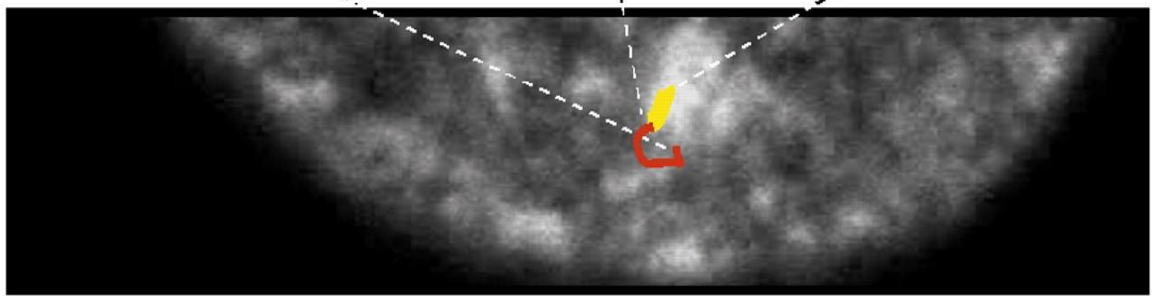
Histone modifications
DNA methylation
Non coding RNAs
Nuclear compartments



Sequence

Chromatin

Nuclear organization



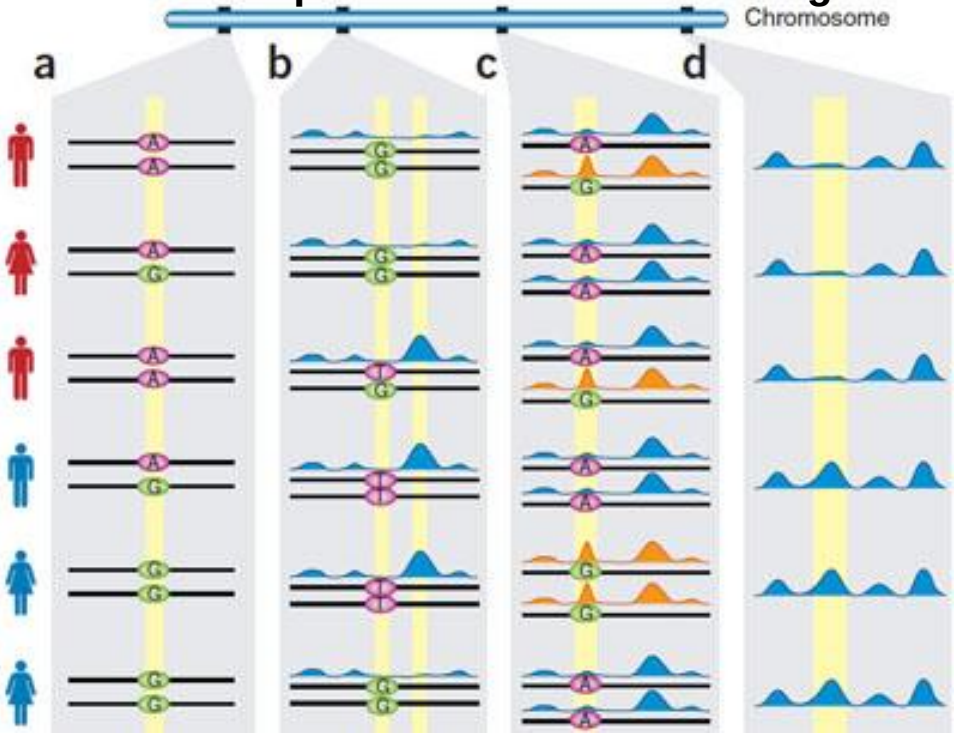
van Driel, R. *et al.*, 2003 Journal of Cell Science

…Controlled by epigenetic information

Epigenetic information in the prediction of traits from GWAS..

in human

ex: interpreting non coding genetic variation in complex human traits at a single locus



Involves the use of reference epigenome maps of primary and cultured cells

Ward and Kellis, 2012, Nat.Botech.

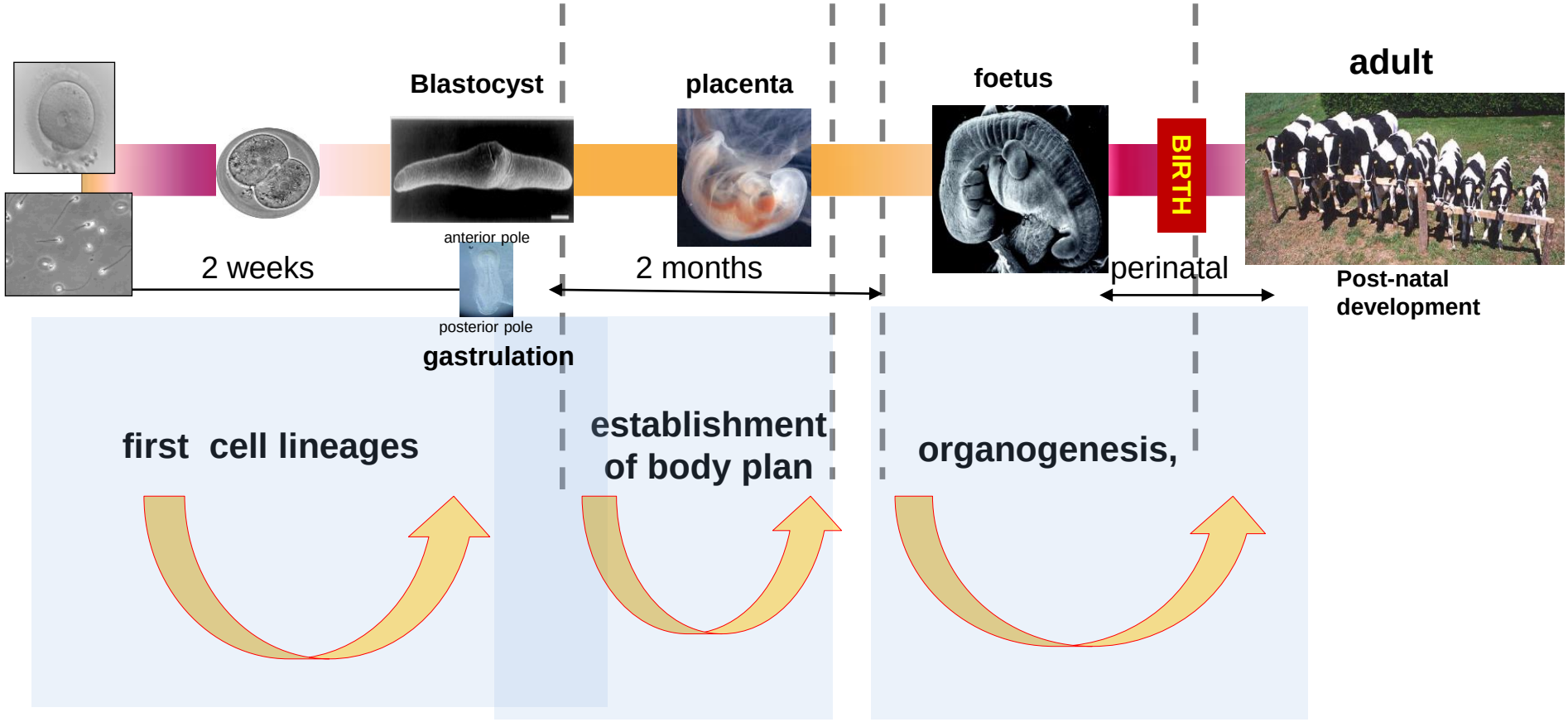
a- Genetic association with an organismal trait (**GWAs)**

b- Genetic association with a molecular trait (**eQTL**)

c- Genetic association allelic activity (**ASE)**

d-molecular biomarker for an physiological trait (**EpiWAs**)

waves of epigenetic reprogramming shape transcriptomic patterns during both embryonic, foetal and post natal periods



Measuring the contribution of epigenetics to the variability of phenotype in the bovine species

First step: simplify the approach using

- **genetically identical animals: through cloning**
- **DNA methylation**

Second step: develop a wide genome approach using key tissues: blood, liver, muscle placenta, mammary gland

Third step: combine phenotypic and epigenetic data

Outcome:
- **extension to individuals from non related genotypes**
- **integration into a bovine data platform**

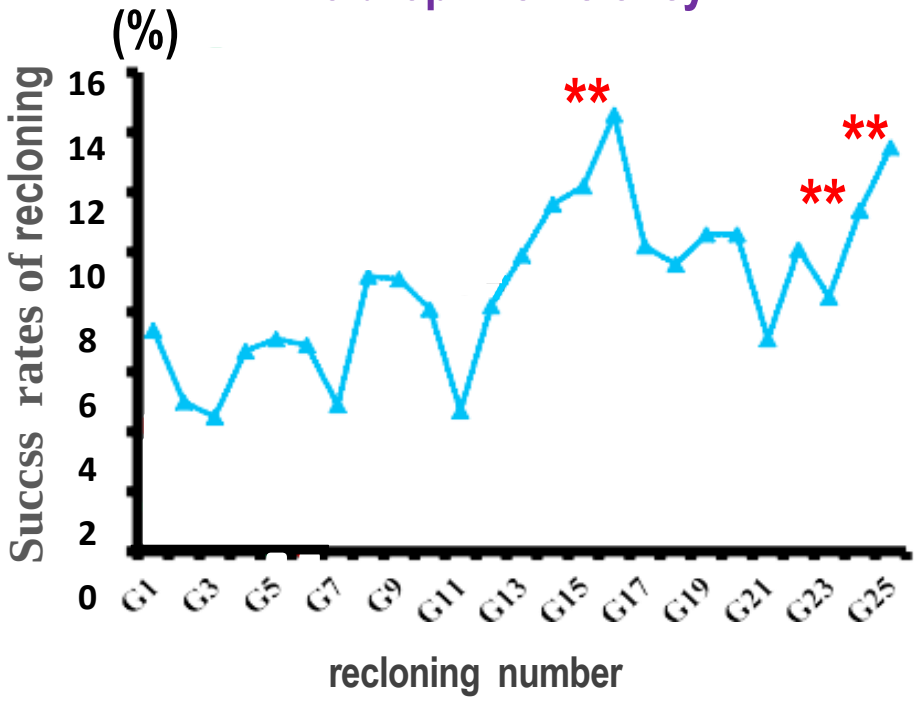
Cloning techniques do not select for somatic cells more amenable to full reprogramming

mouse cloning

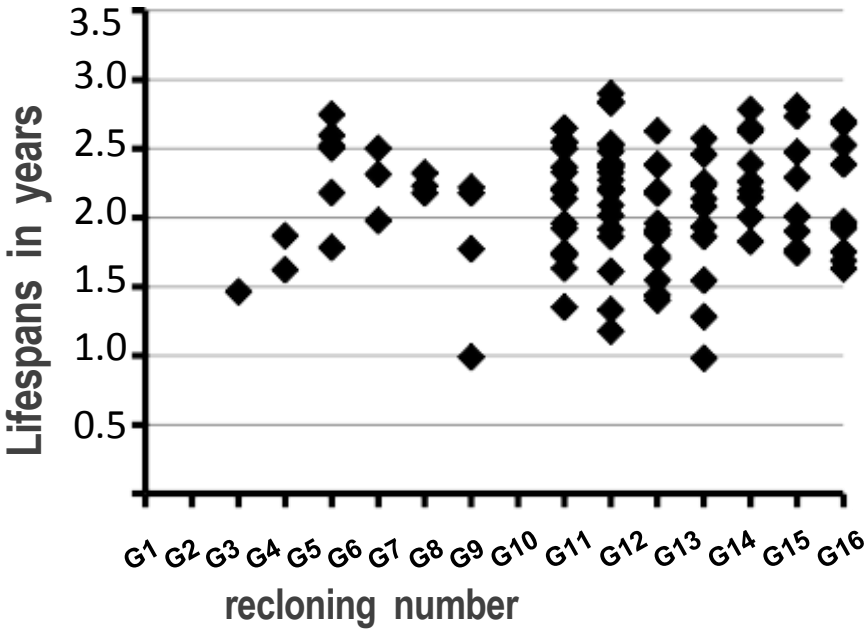
-optimisation of in vitro conditions (empirical)

-recloning from adult cloned mouse

no drop in efficiency



no impact on viability

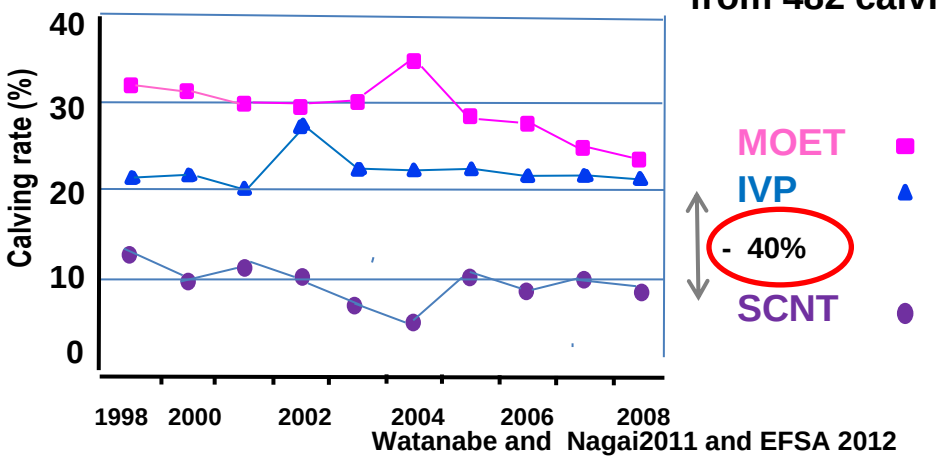


Wakayama et al. 2013 Cell Stem cell

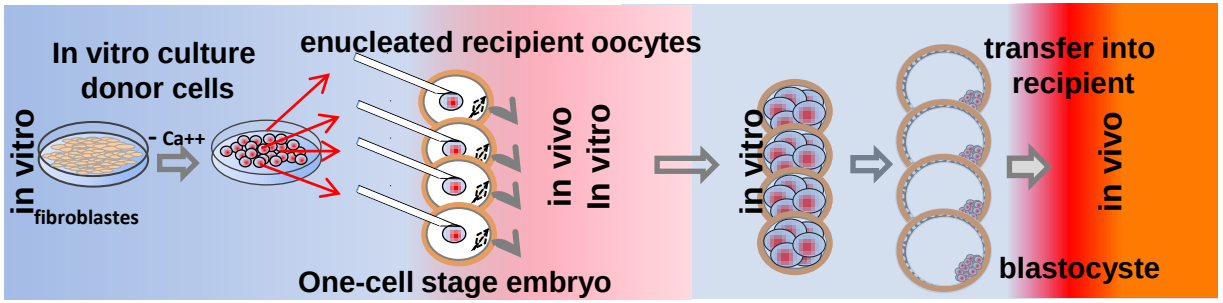
Bovine cloning: a way to profile epivariants in a context of high environmental constraints applied on the genome.

relatively efficiency at birth 40%

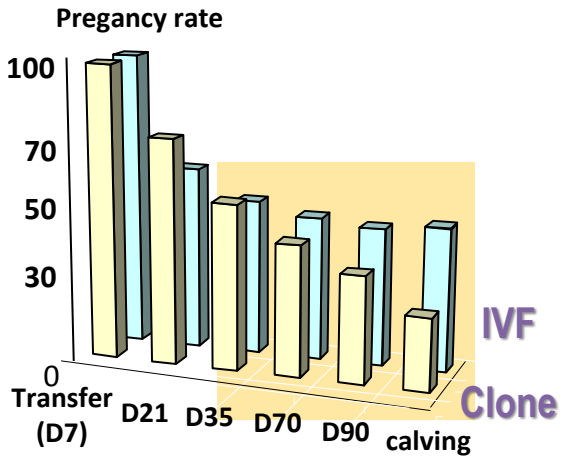
same nuclear genome



different nuclear microenvironment from the one-cell stage



— in vitro (culture) — in vivo (oocyte) - in vitro (culture) ——→ in vivo →



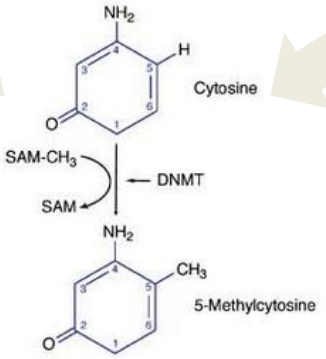
9 months- long gestation

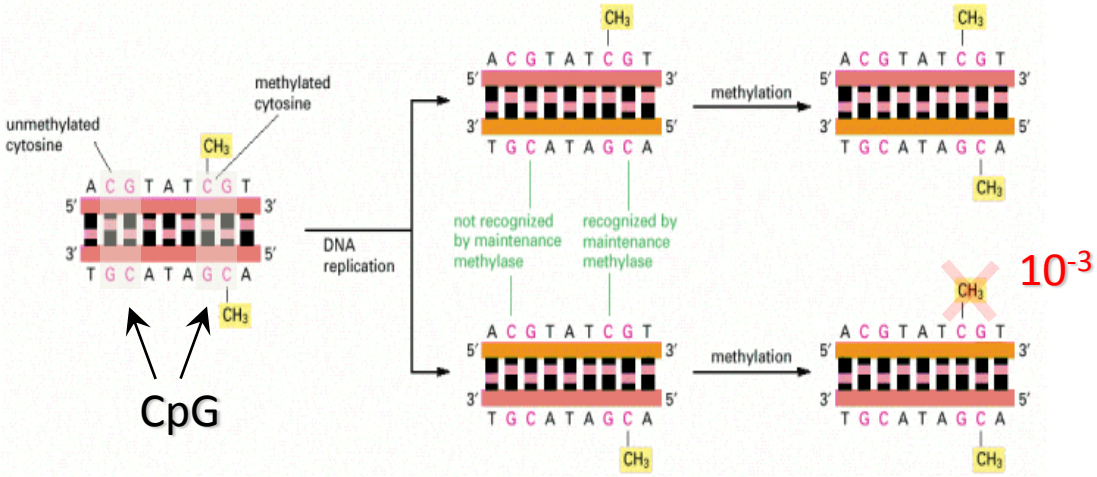
DNA methylation is highly dependant on genomic sequence

Environment

Sequence

Stochastic





Active DNA methylation

- Dnmt1: maintenance

Inheritable (mitosis)

- Dnmt3: de novo

in response to stimuli

DNA methylation loss

- Passive
- Active

Tools used to characterize, identify and measure DNA methylation variations

1. Global analysis

Cytosine methylation electrophoretic assay

LUMA (Luminometric Methylation Assay)

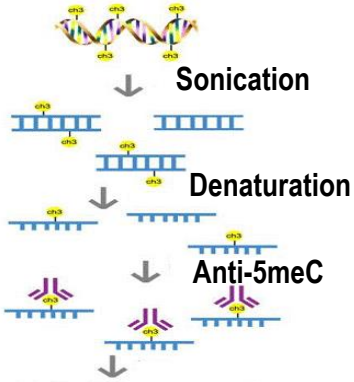
Immunocytochemistry

2. Whole genome analysis

MeDIP/Chip

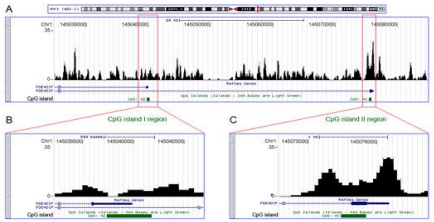
MeDIP/sequencing

RRBS (Reduced RepresentativeBisulfite Sequencing)



Methylated DNA analysis

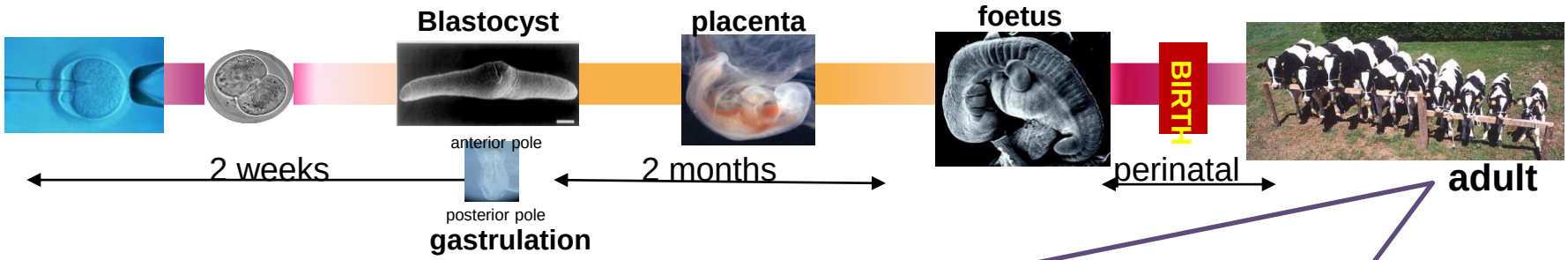
Analyse by Chip hybridization or sequencing



3. Sequence specific approach

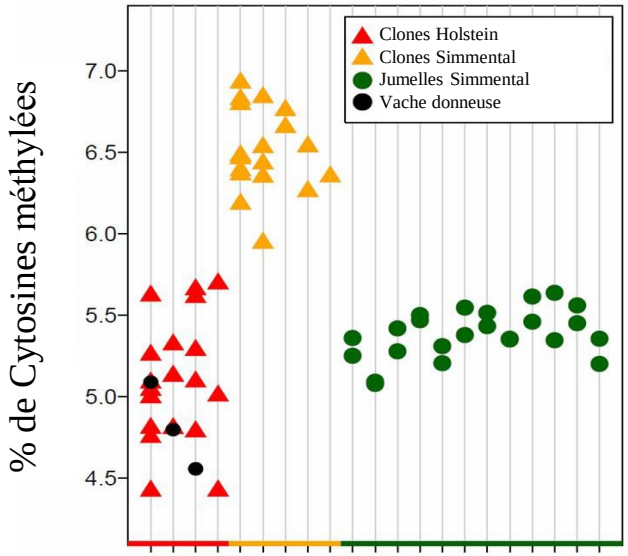
Bisulfite conversion /pyrosequencing

Bovine clones are genetically identical but epigenomically variants

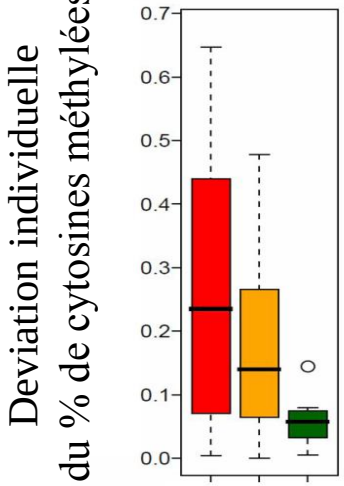


Variability of global methylation rate in lymphocytes of adult bovine clones

A



B



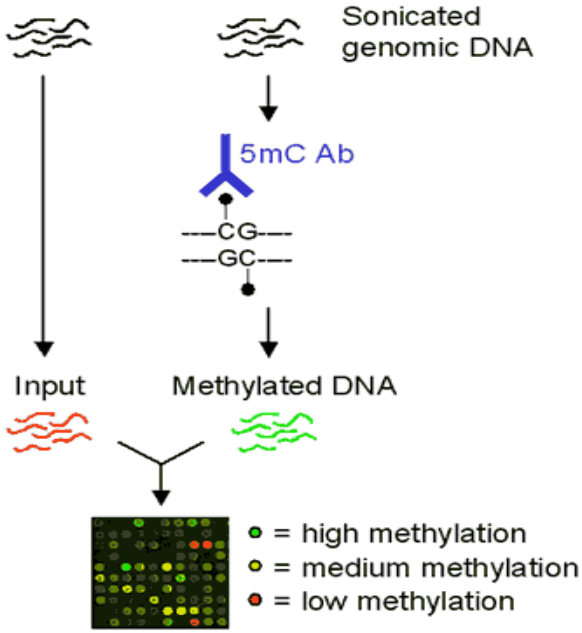
Holstein, 4 genotypes

Simmental, 5 genotypes

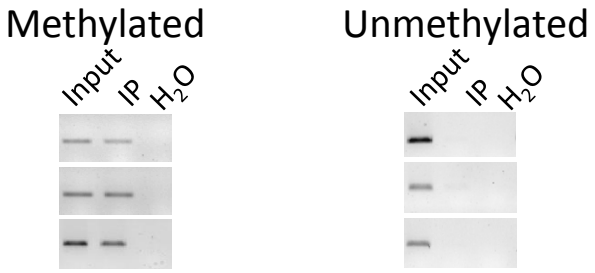
Twins pairs, n=13 Simmental

Cytosine methylation electrophoretic assay ***De Montera et al., 2010***

Epigenetic signatures at perinatal stage in liver

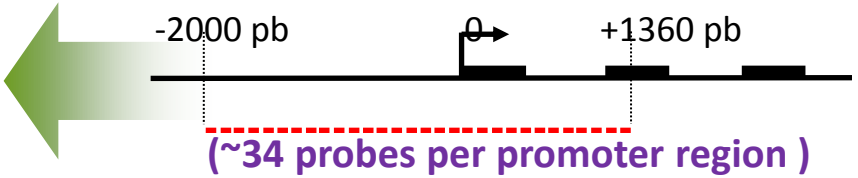


From : http://www.epigenome-noe.net/researchtools/protocol.php?protid=33



bovine microarray targeting

21 416 genes



(~34 probes per promoter region)

+ 12 larger regions

MeDIP/Chip- Kiefer et al. submitted

Data analysis framework

1

Identification of methylated probes

ChIPmix R package 720 000

2

Identification of highly methylated regions

containing clusters of enriched probes in at least one condition

3

Identification of 83 Differentially Methylated Regions (DMRs) between clones and AI

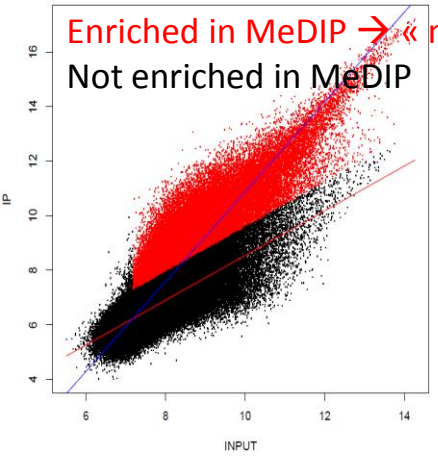
SpatStat R package

4

Validation by pyrosequencin
80% regiond vamidat

Enriched in MeDIP → « methylated »

Not enriched in MeDIP



Data analysis framework

1

Identification of methylated probes among 720 000 probes

ChIPmix R package

2

Identification of 15885 highly methylated regions

containing clusters of enriched probes in at least one condition

3

Identification of 83 Differentially Methylated Regions (DMRs) between clones and AI

SpatStat R package

4

Validation by pyrosequencing

Data analysis framework

1

Identification of methylated probes

ChIPmix R package

2

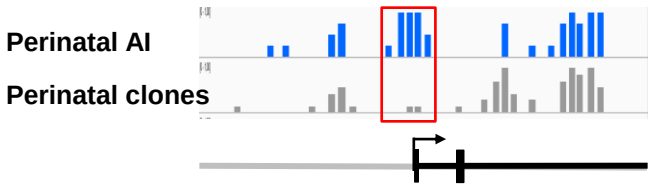
Identification of highly methylated regions

containing clusters of enriched probes in at least one condition

3

Identification of 83 Differentially Methylated Regions (DMRs) between clones and AI

SpatStat R package



4

Validation by pyrosequencing

Data analysis framework

1

Identification of methylated probes

ChIPmix R package

2

Identification of highly methylated regions

containing clusters of enriched probes in at least one condition

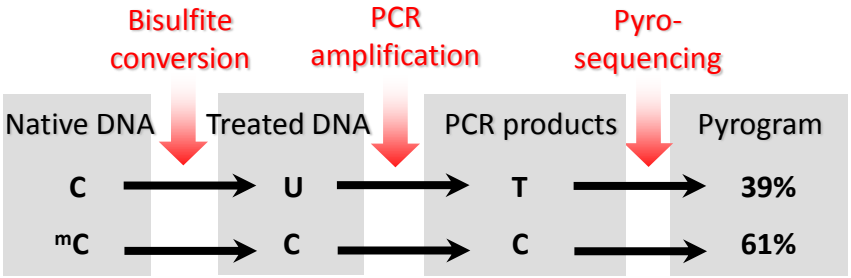
3

Identification of 83 Differentially Methylated Regions (DMRs) between clones and AI

SpatStat R package

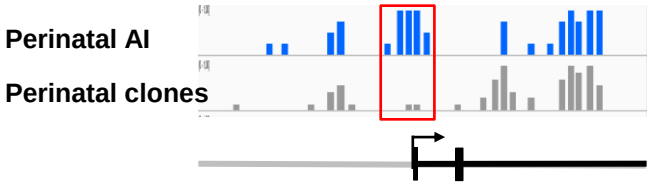
4

Validation by pyrosequencing

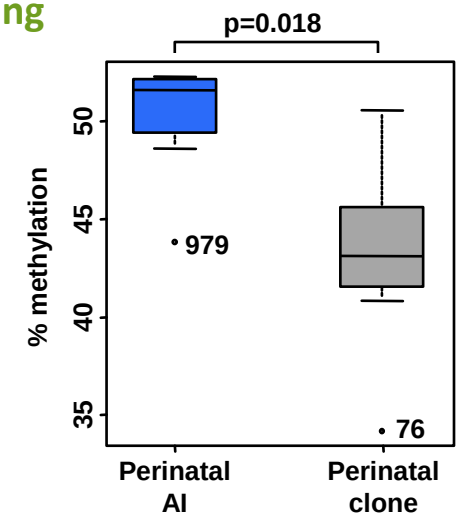


One example (among hundred others) of DMR methylation rate correlated to phenotypic markers of liver physiology

Microarray data

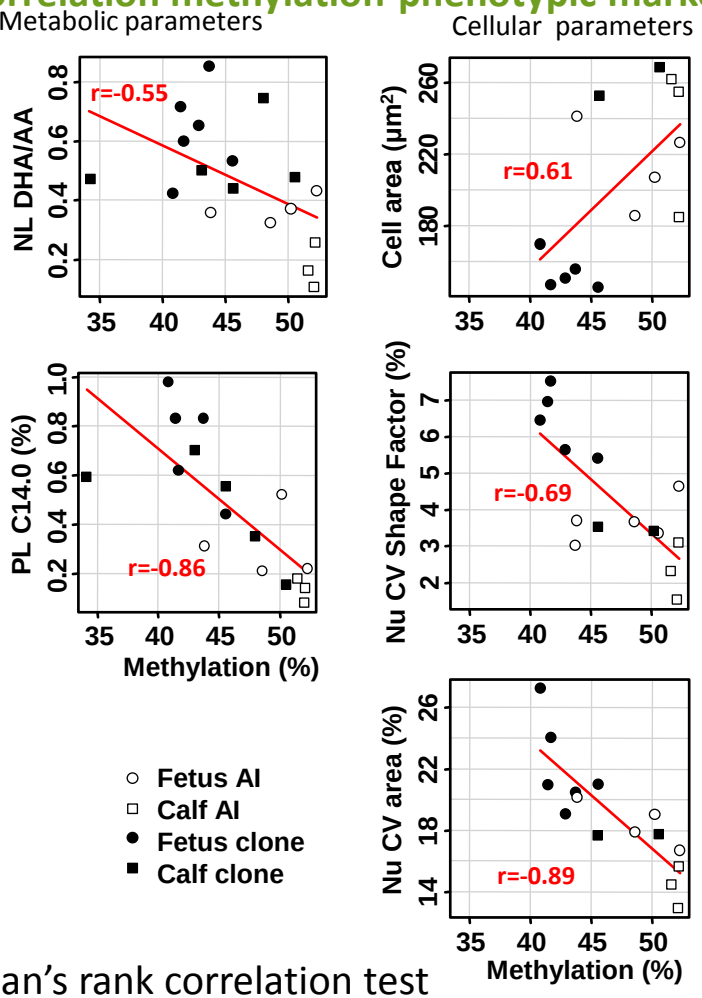


Pyrosequencing validation



Permutation test

Correlation methylation-phenotypic markers



Spearman's rank correlation test

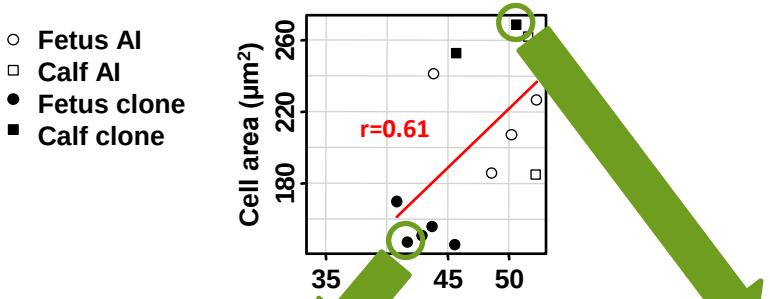
An example of DMR methylation rate correlated to phenotypic markers of the liver's physiology

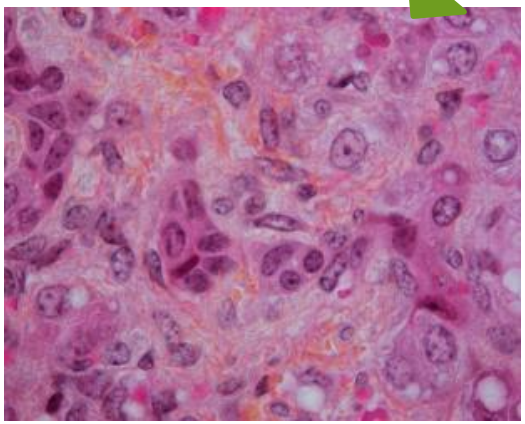
to predict early lactation-associated hepatosteatosis in dairy cows

Ongoing validation from the blood of dairy cows at early lactation.

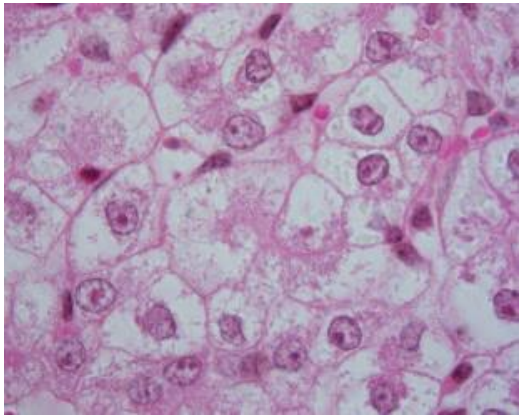
- In human, the associated gene shows differential methylation in the blood of newly diagnosed, drug naive patients with type 2 diabetes Canivell et al. 2014, PLoS one)

Correlation methylation-phenotypic markers



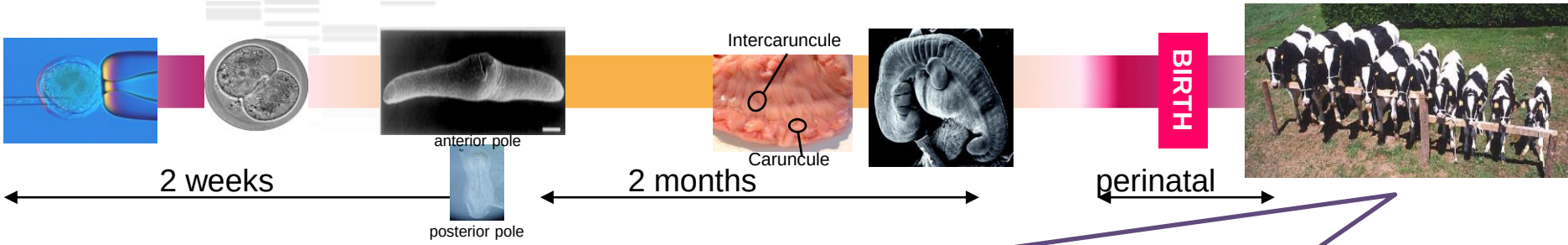


Clone with no glycogen storage (average size of hepatocytes is smaller)

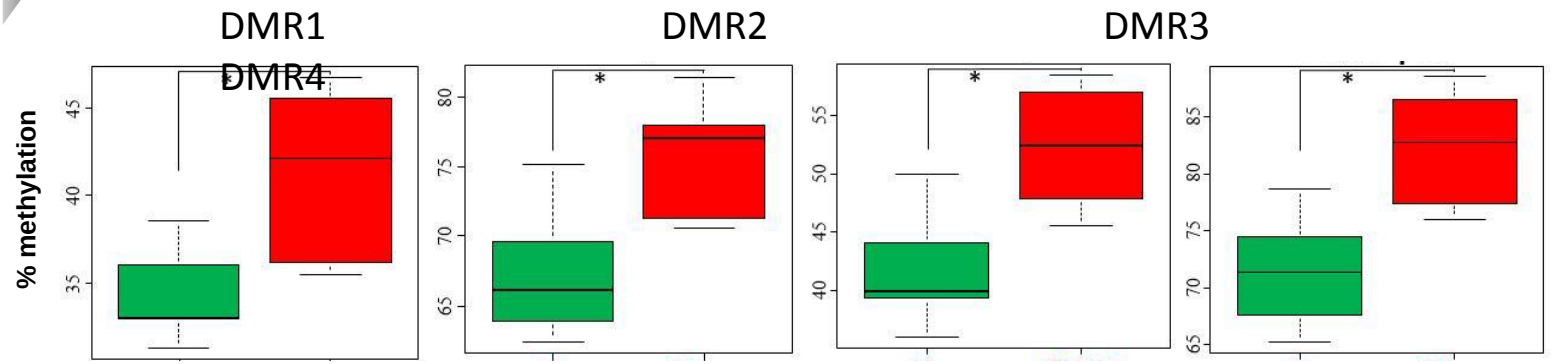


Clone with normal glycogen storage (average size of hepatocytes is similar to AI)

Extension to muscle epigenomic analysis: screening AI and cloned adults



Muscle biopsy at 18th months ~ 500 µg genomic DNA



*p<0.05
Permutation test

AI controls, 18 months
n=9

Healthy clones, 18 months
n=8

Kiefer et al. submitted

A first step towards the tracability of cloned products ?

and from now …

capture more Data

Genomic data
phenotypical data
environmental data

on an individual basis

Genome and functions

- Genomics (QTL, SNPs…
- **Epigenomics**
- Metabolomics
- Ruminomics

Traits

size, growth,

Productive traits

Reproductive traits

puberty
ovulation rates
fertilisation
embryo death
fetal losses…..

Phenomes

- more data
- higher resolution

environment

Animal management

Animal health/treatment

What happened during the reproductive cycle?

•What was the outcome?

- more data
- higher resolution

through FANG

through ATOL

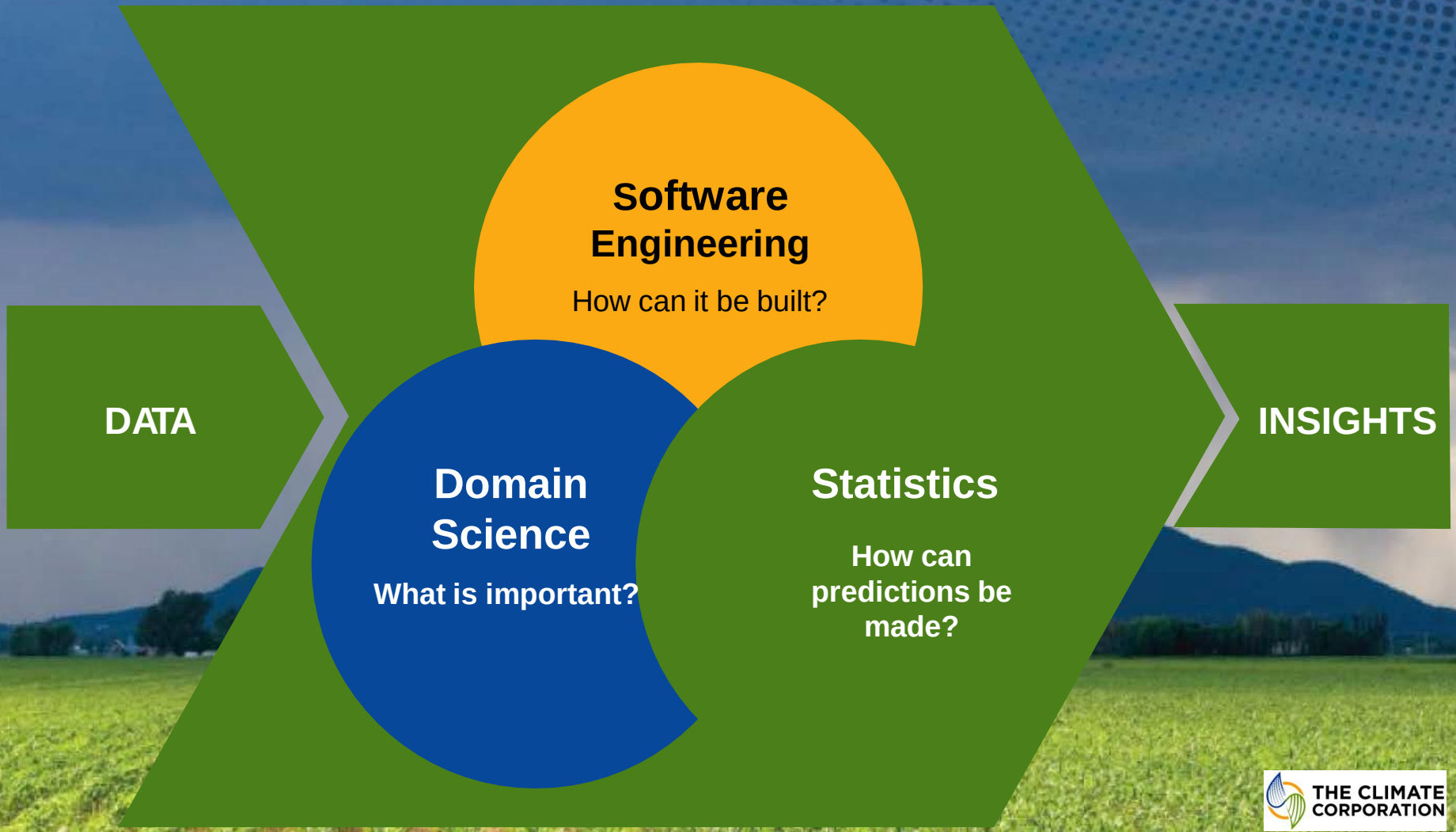
through EOL

The individual animal within its environement as the reference

In agriculture

Leading Data Science Platform in the Industry

Models:
Predict the Future

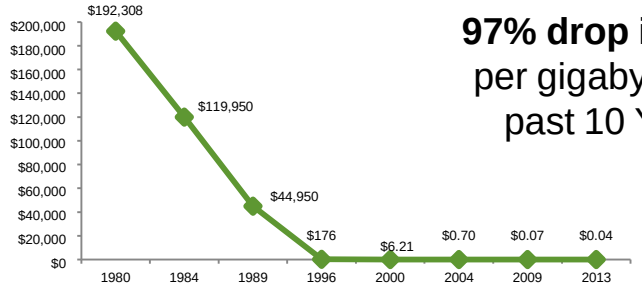


THE CLIMATE CORPORATION

Computing Applications Make Data Relevant to Animal production. They are already economically operational in Agriculture

Storing Data is Getting Cheaper

Average Price Per Gigabyte, 1980-2013



97% drop in price per gigabyte over past 10 Years

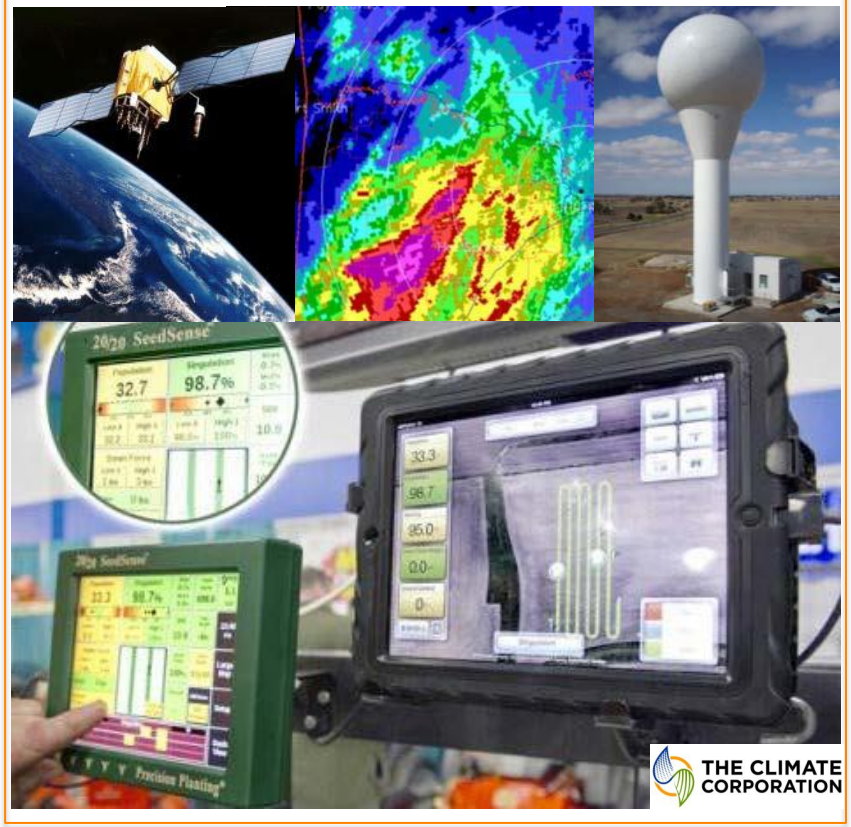
Source: Wayback Machine (Statistics Brain)

Low Cost Computing is Proliferating

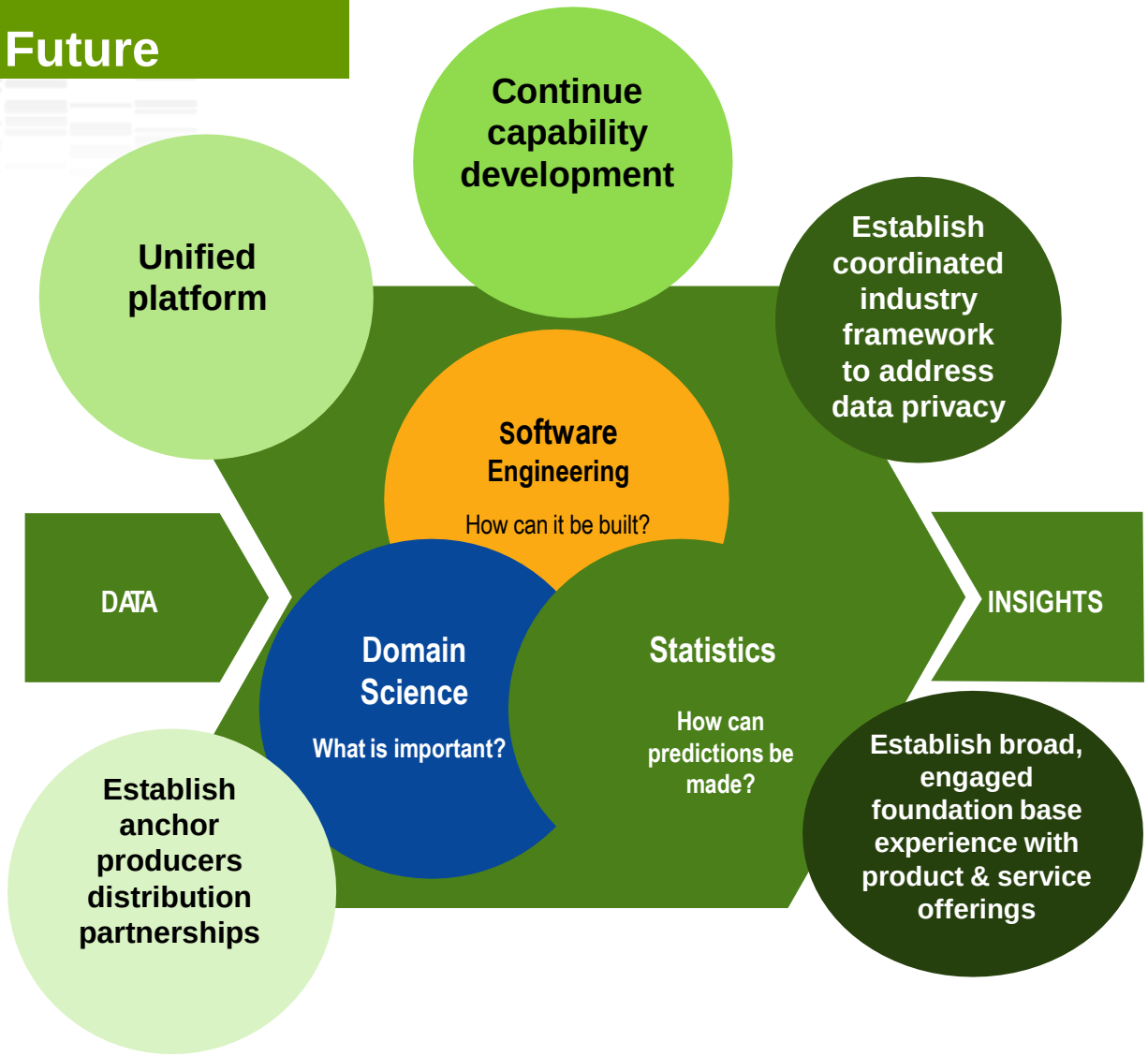
Distributed Computing Allows Multiple Computers to Use More Data and Solve a Problem Together

ex: Hadoop platform for Big data management

In Agriculture, Environmental and Operating Data Are Taking Shape



Models: Predict the Future



UMR 1198 Biologie du Developpement et Reproduction

SABRE

Agence Nationale de la Recherche ANR

APIS-GENE

Allice ÉLEVAGE | INNOVATION | SERVICE

Poster n°15-21249

Poster n°14-21257

Hélène Jammes

Hélène Kiefer

Evelyne Camplion

Audrey Prézelin

Luc Jouneau

Maxime Gasselin

Hala Al Adhami

François Piumi

Jean Philippe Perrier

- Oral session n°37wed.17h

Coll. i Biostatistical Analysis
Marie-Laure Magniette
AgroParisTech

Coll. Methylation Analysis (RRBS) CNRS, Strasbourg

EpigRAni partners

- Eve Devinoy, GABI, INRA, Jouy en Josas
- Dominique Rocha, GABI, INRA, Jouy en Josas
- Marion Boutinaud, Pegase, INRA, Rennes
- Jorg Tost, CEA/CNG, Evry