"New techniques" of modification of the genomes and epigenomes (NTMGE)

Yves Bertheau

“New techniques” of modification of the genomes and epigenomes (NTMGE)

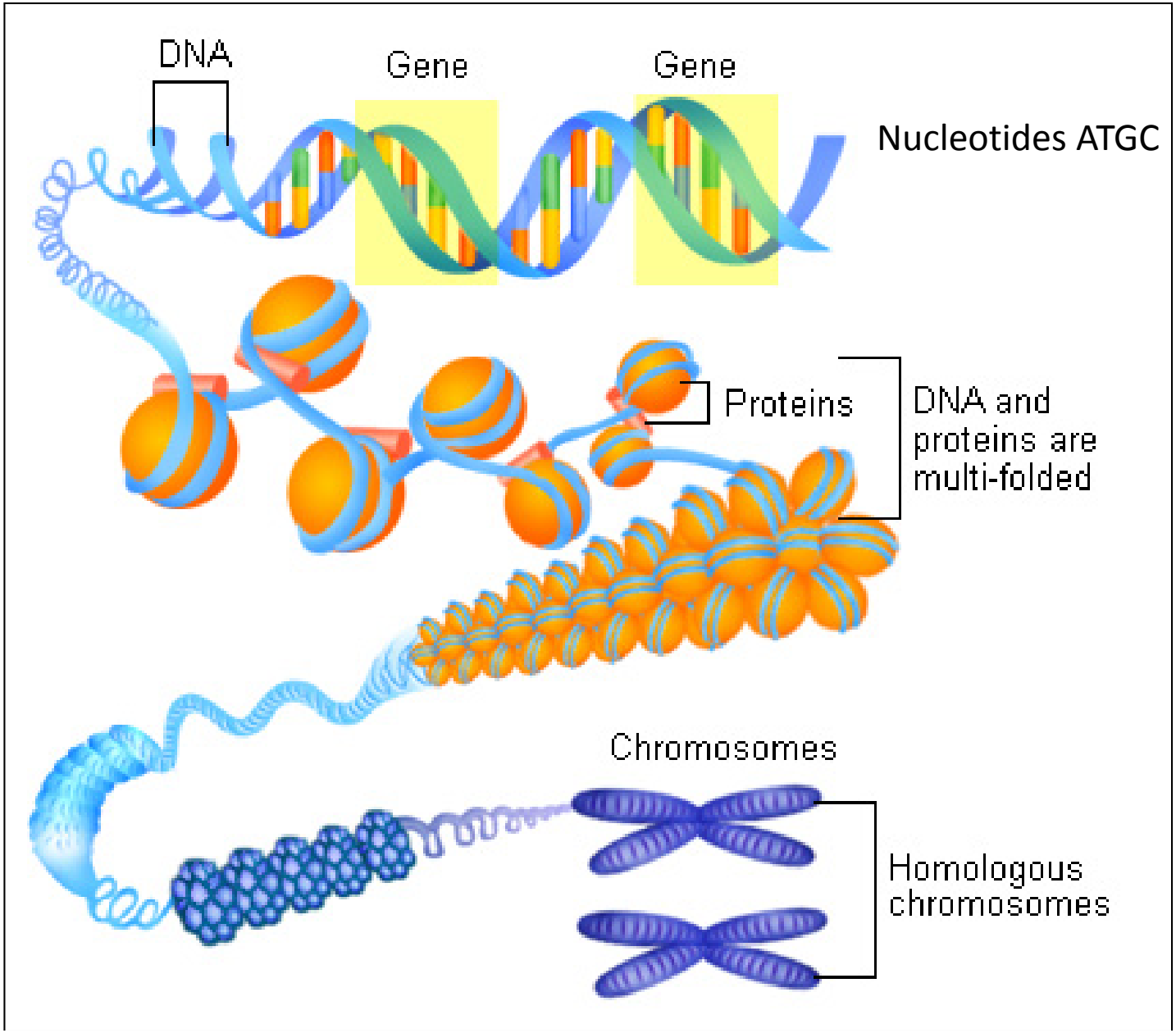
Yves Bertheau

IFOAM EU Workshop
2017/07/06
Echzell (Germany)

MUSÉUM NATIONAL D'HISTOIRE NATURELLE
INRA SCIENCE & IMPACT
CESCO Centre des Sciences de la Conservation Équipe "Socio-Écosystèmes"
cnrs
UPMC SORBONNE UNIVERSITÉS

REMINDER ON GENOMES, EPIGENOMES AND EPITRANSCRIPTOME

Chromosomes' structures and changes



DNA: Point mutation, indels, substitution
by e.g. chemicals, irradiation, DNA repair systems such as NHEJ, MMEJ…

Transversion, conversion

Translocation, inversion, duplication/amplification, insertion by e.g. transposition, polyploidization…

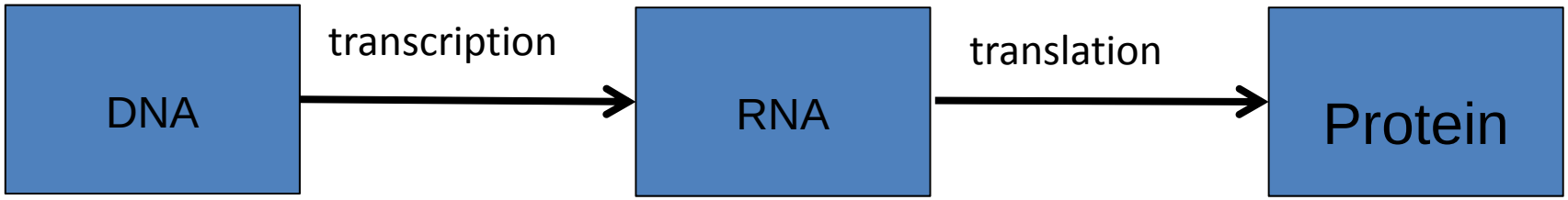
DNA and chromatin (histone and non histone proteins): epimutation

RNA: epimutation (epitranscriptome)

High compaction of DNA, due to RNA, proteins…
Opening needed to produce proteins…

Human genome ± 1.9m

“Central dogma of biology”



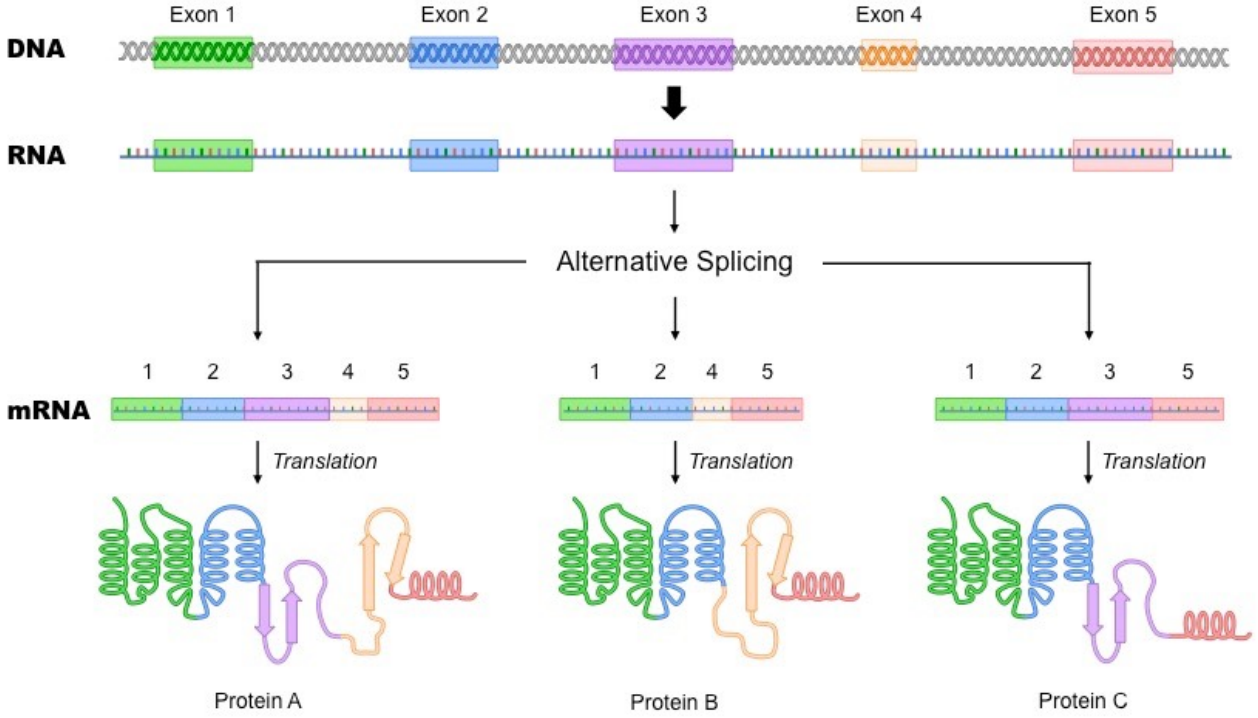
RNA transcripts as well as their abundance are representative of a cell at the moment…

But in reality:

- alternative splicing generates many isoforms (eukaryotes: splicing from exons/introns of a “encoding sequence”),
- indeed different abundances for various transcribed are present,
- the transcriptome is a mixture of transcripts of all genes,
- Exon skipping may produce aberrant proteins and disease…

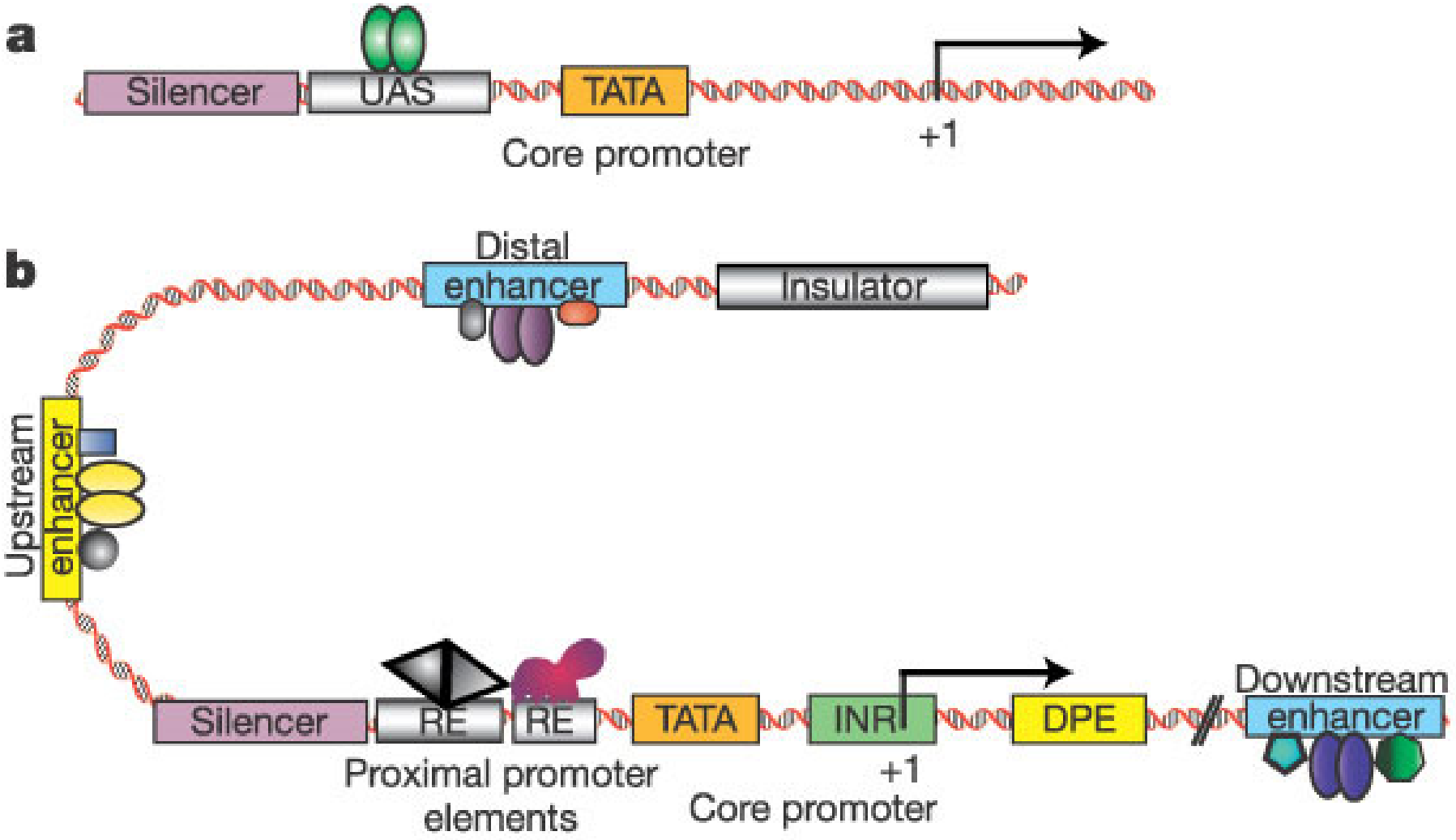
For instance: the human transcriptome is made up of tens of thousands of transcripts of the 20,000 genes (plants: 25-40,000 genes)

Alternative splicing and non-coding DNA



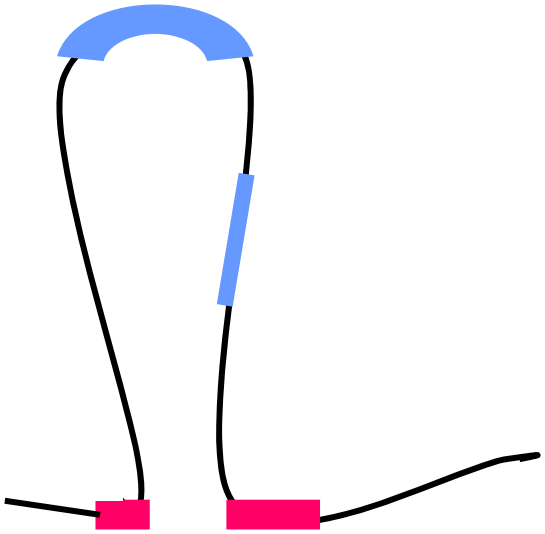
- Production of small RNA involved in gene regulation,
- Several proteins can originate from a gene (see e.g. ENCODE Human genome project)
- Large non-coding regions (ca 95% of human genome…) with rather unknown functions (“dark matter”), with or without selection pressure…
- Non-coding regions involved in gene regulation in complex networks of genetic interactions.

Eukaryotic transcriptional unit: very complex with distal effects…



Levine et Tjian Nature, 2003

Scaffold/matrix attached regions S/MARs



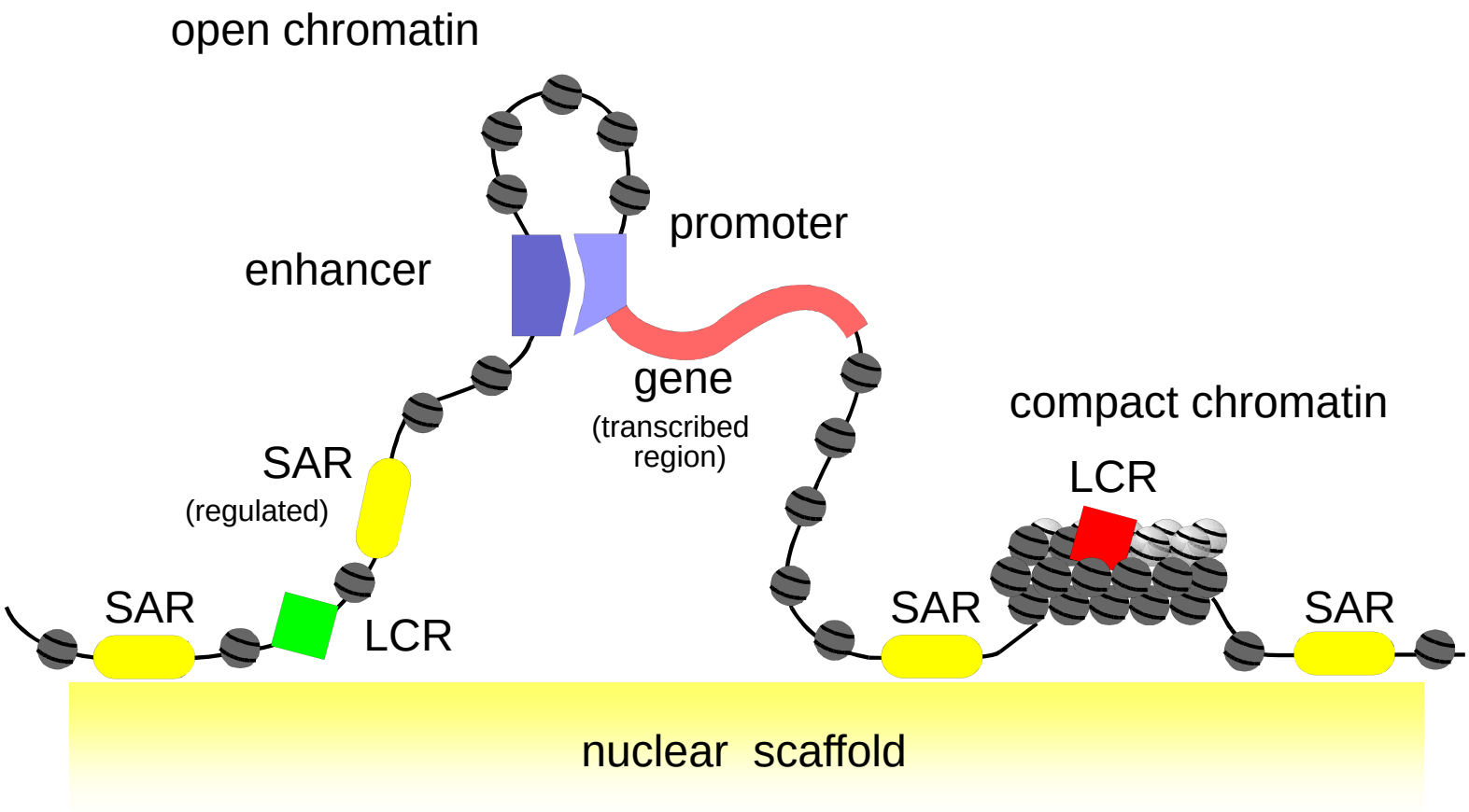
S/MARs
genes
residual DNA

Scaffold/matrix attached regions (S/MARs) are regions of the DNA strand that are found the basis of chromatin loops. They anchor the DNA to the proteinaceous nuclear matrix.

Each loop is considered to be a functional domain.

S/MARs may act as border elements and thus, protect gene expression from position effects.

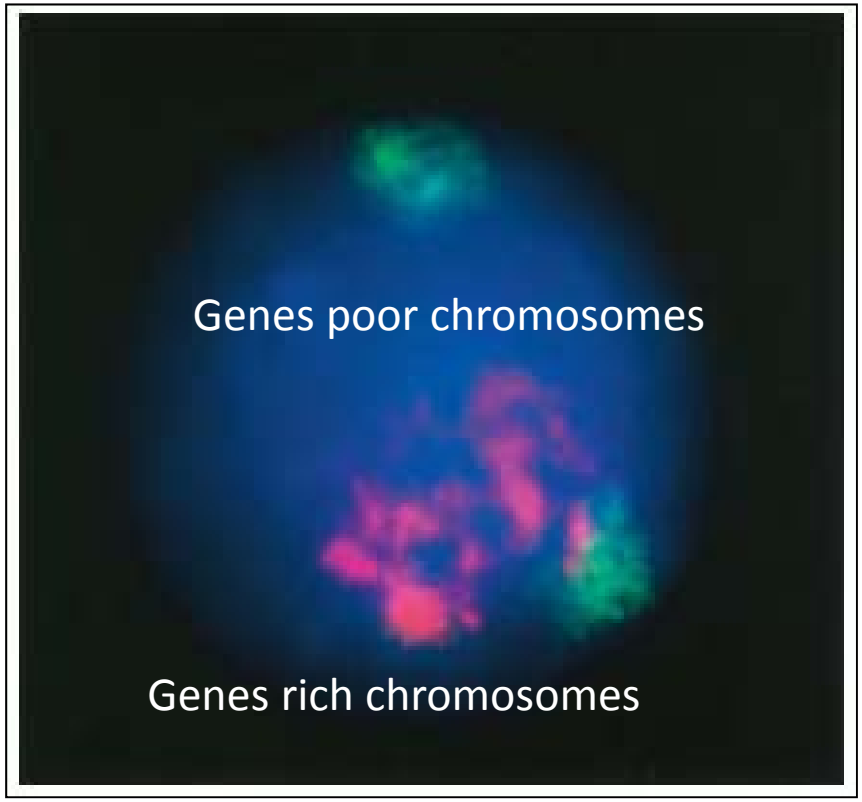
Scaffold/matrix attached regions S/MARs



J. Bode / E. Wingender 1993

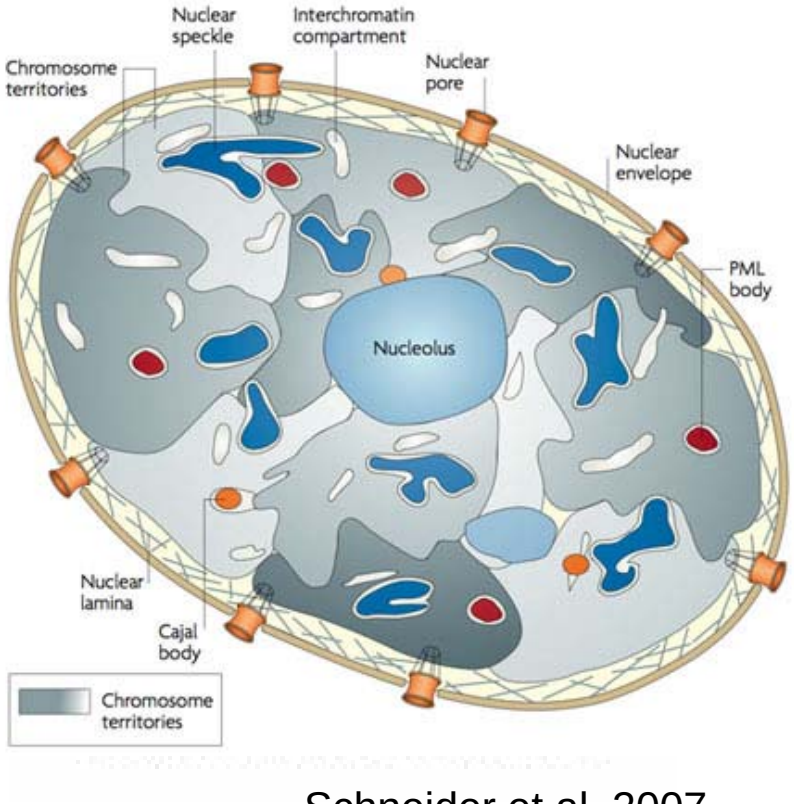
Chromosomal structure capture: interactions between different loci on several chromosomes

Schneider et al. 2007

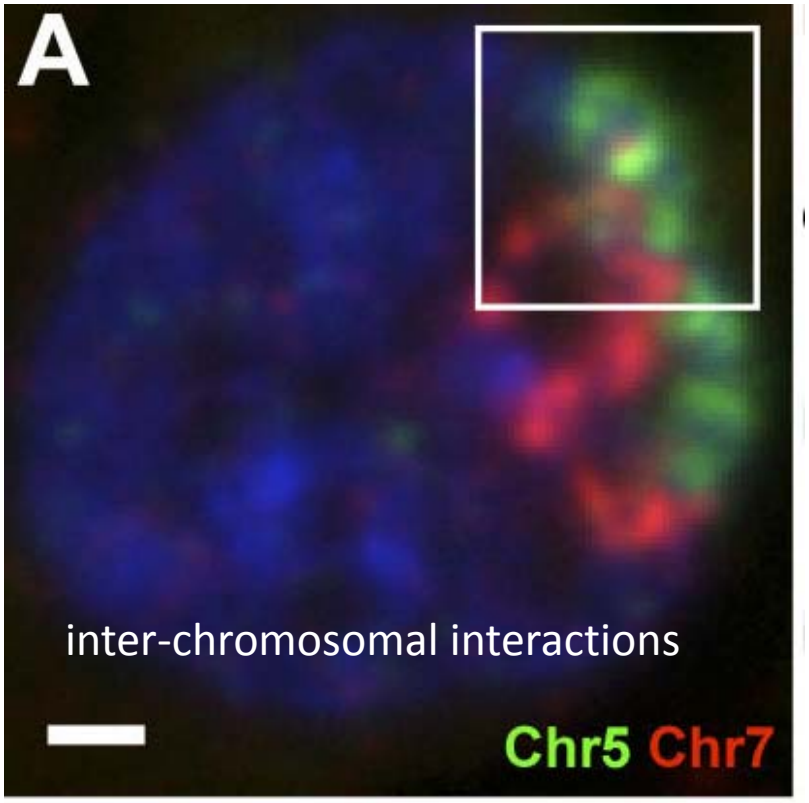


Spector et al. 2003

Chromosomal structure capture: interactions between different loci on several chromosomes



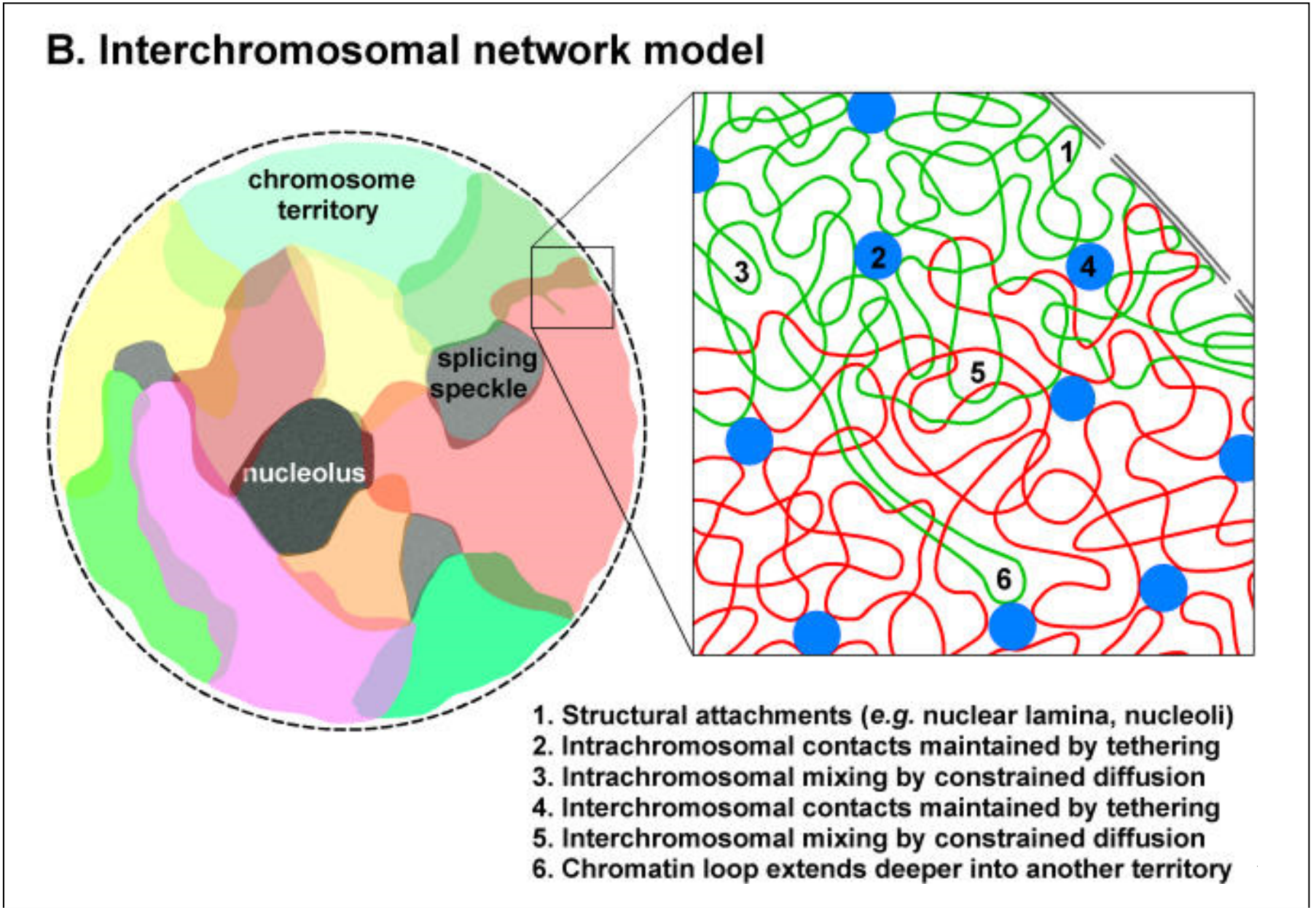
Schneider et al, 2007



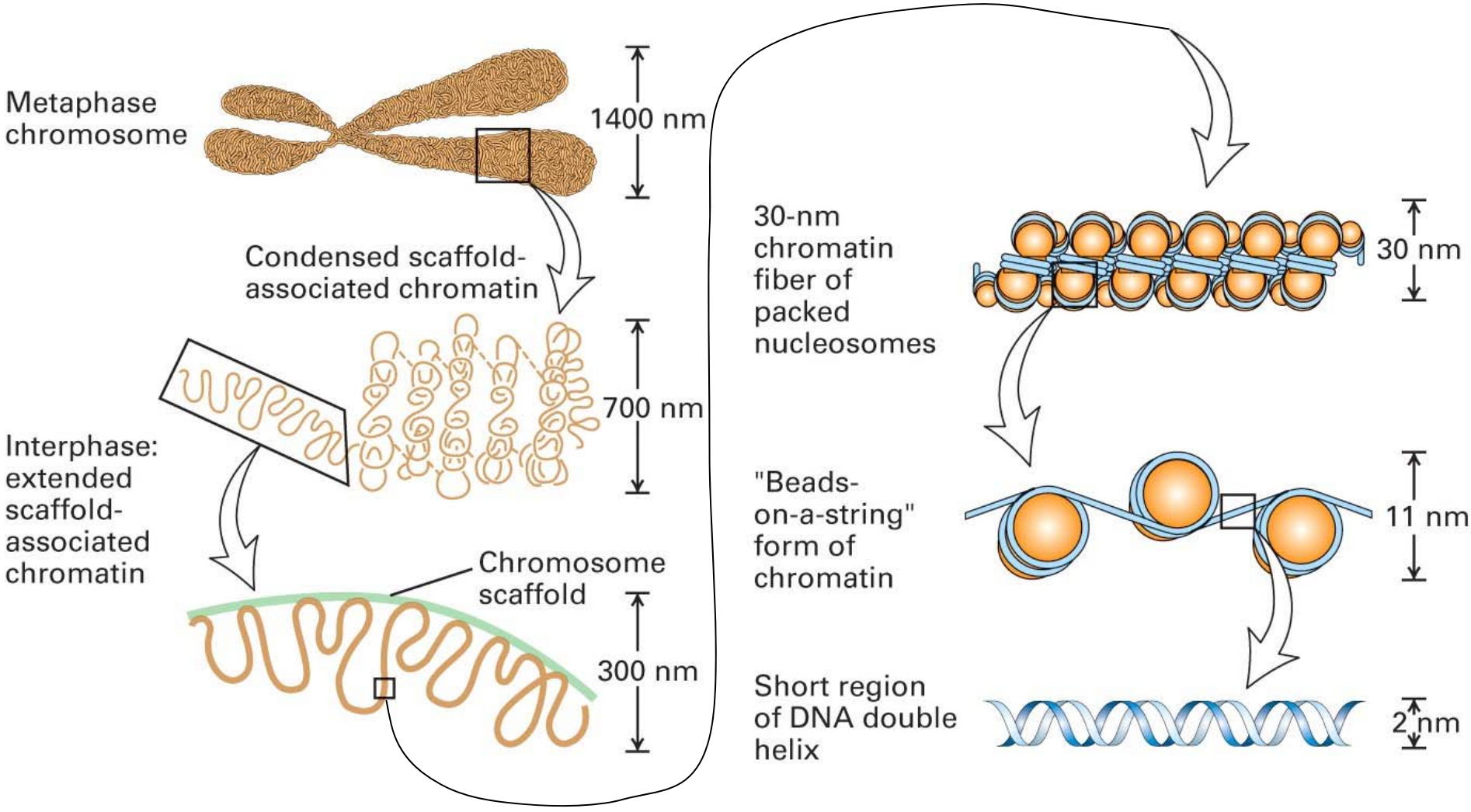
Branco et Pombo, 2006

Very large holes necessary for the delivery of all NTMGE reagents

Complex interaction networks



Chromatine structure, epimutations and gene regulation…



Reminder about genomes and epigenomes

- Very large non-coding DNA parts,

e.g.: human 95% (from the “junk DNA” to a new *Terra incognita),*

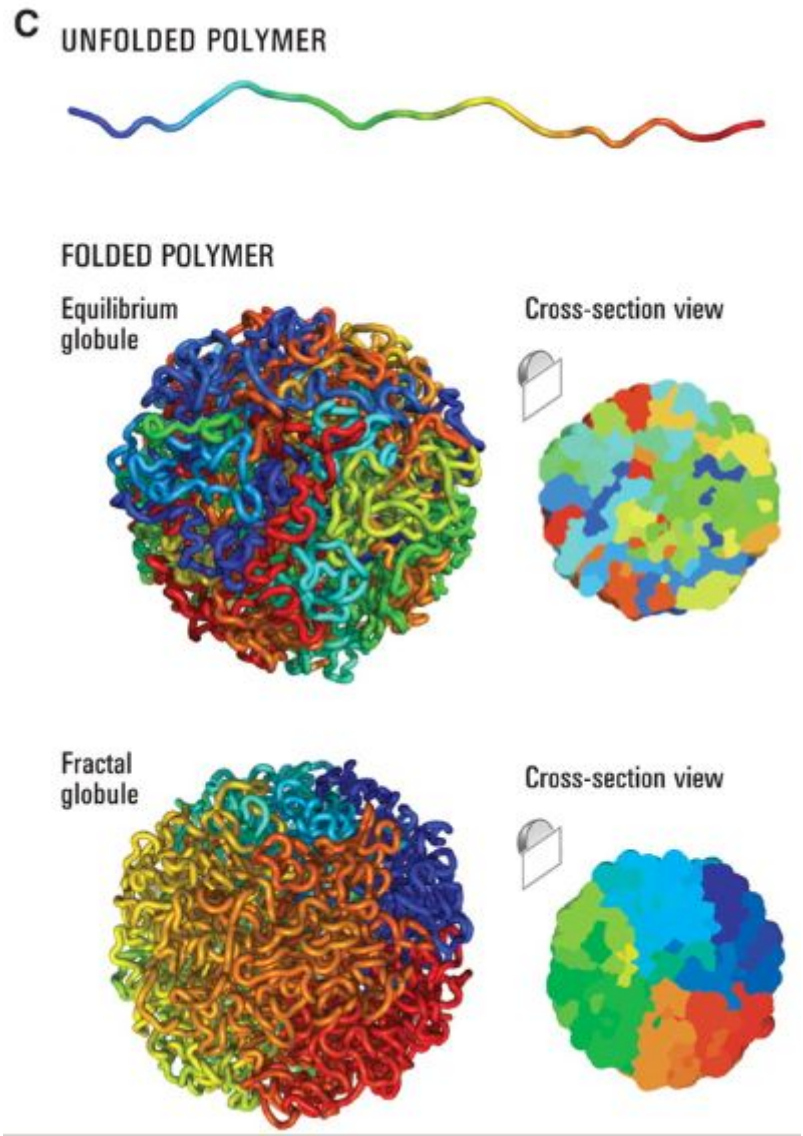
- A single nucleotide change may affect function and/or regulation, gene splicing… both at short and long distances (e.g. DNA modification induces epigenetic changes)

- Very slight correlation between linear and spatial organizations,

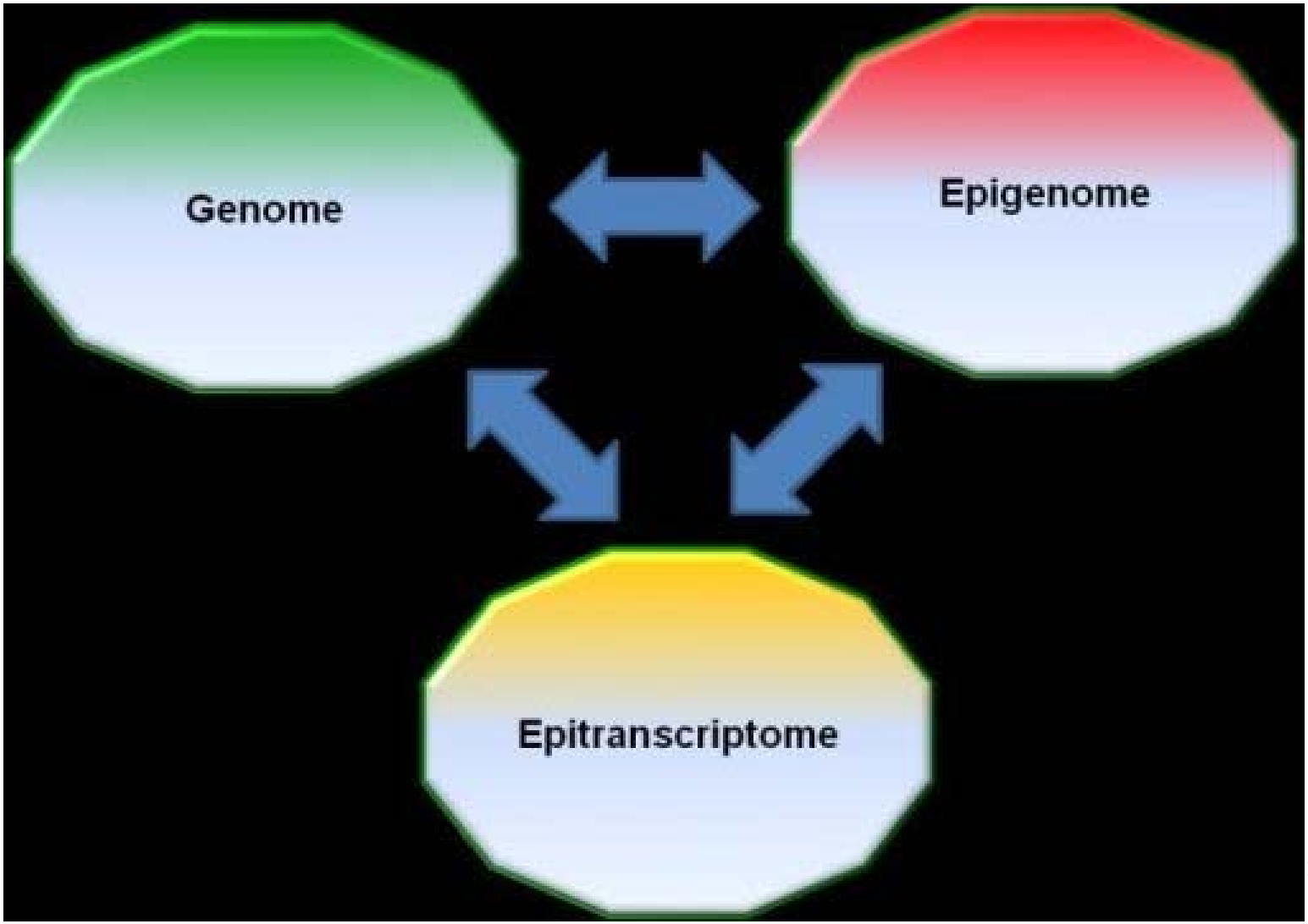
- Epigenomes (proteins, DNA and RNA [= epitranscriptome]):

“*now we know from where we have to start*” (conclusion of the EFSA meeting, June 2016)

- We are thus very far from mechanical and linear molecular biology of the 70s’ used for the communication about NTMGE



Interactions between domains very badly known



A few plants' specificities

- Not mobile,
- organization of the genomes (nuclear, chloroplast and mitochondria) and ploidies,
- large differences genotypic and phenotypic between mono- and dicotyledonous and between species' varieties, particularly Elite lines (e.g. corn ca. 20% of differences with strong linkage disequilibrium…),
- cell (to remove for transformations) walls of recalcitrant plants for regenerations (or use of immature embryos and mosaicism...),
- A few cells (germ lines) of apical and axillary meristems are preserved from mutations and epimutations (analogy with animal germlines isolation at the beginning of development), some derived cells used to maintain organisms' homogeneity and plant longevity (see e.g. the Lausanne Napoléon's tree)…

Conclusion

- All DNA, proteins and RNA changes (including point mutations) may have effects either locally or distally,
- Changes can be easily or not detectable by e.g. phenotypic or genotypic changes, but face numerous issues of sequencing platforms, assembly and comparison software, ata bases errors, ack of sufficient quality control and quality assurance procedures, gold references standards,
- Most – if not all – of the artificial induced mutations left genomic / epigenomic scars usable in techniques' identifications and differentiation of *in vitro* and *in vivo* changes,
- DNA changes, such as insertion, are inducing, generally distal, epimutation(s),
- Focuses on nuclear genomes do no provide insights on plastids' genomes,
- More and more epimutations (DNA, RNA and proteins) are thought to be inheritable by generally unknown mechanisms (= "adaptation", part of heterosis?...),
- Changes of fitness may be invisible in the greenhouses and further local fields testing, but visible in other circumstances…

The molecular biology as used in the communication about NTMGE takes into consideration only the tip of the iceberg

DNA changes (introduced or not) are only a part of the uncertainties and risk assessment: the GMO paradigm only based on DNA nucleotides since decades has to integrate epigenetics and complex interactions networks, with their largely unknown sectors

NTMGE

(“NEW TECHNIQUES” OF MODIFICATION OF THE GENOMES AND EPIGENOMES)

Initial work in the EU on NTMGE

- Zinc finger nuclease technology (ZFN1-ZFN3) + TALEN+ meganucleases (then Crispr-endonuclease)
- Oligonucleotide directed mutagenesis (ODM)
- Cisgenesis/ Intragenesis vs. Transgenesis
- RNA-dependent DNA Methylation (RdDM)
- Grafting (GM rootstock / scion)
- Agro-infiltration (Agro-infiltration "*sensu stricto*", Agro-infection, Floral-dip i.e. plant transformation)
- Reverse breeding
- Synthetic biology

Working group: 2007-2012

Lack of consensus: report not made publicly available

But several experts used some of its "positive" claims

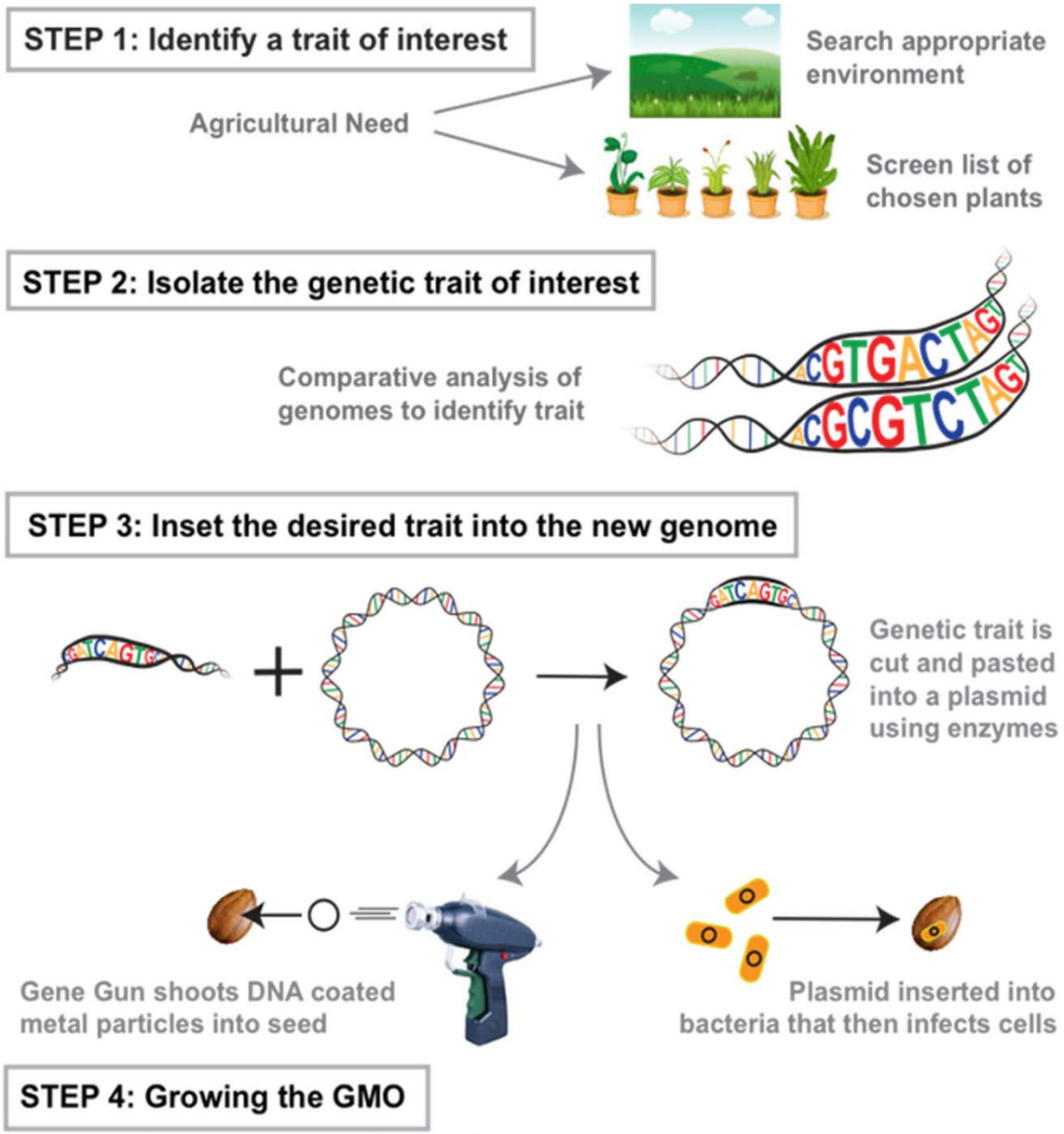
NTMGE: contextualisation

- A suite of agricultural technical developments, often referred to as revolutions (Middle Ages, Green...)
- Technical development of molecular biology: GMO, then the NTMGE, synthetic biology... without disruptive change,
- Public Affairs that led to the refusal by the citizens of certain techniques and leading to regulatory questions, (GMO or not, exempted or not?),
- Communities asking for more natural food commodities, a wording embedded by companies,
- A society where the progress and innovation are felt as a source of happiness with both private and public stakeholders (ministries of environment and agriculture, GSTC vs. ... Assessment Agency), scientists defending institutional or not interests, politicians with pre-established choices due to socio-economic *a priori*...

Globally: a new step in science's politicisation and scientification of politics

RELATED TECHNIQUES
(USED TO PRODUCE BOTH GMO AND NTMGE PRODUCTS)

GMO production



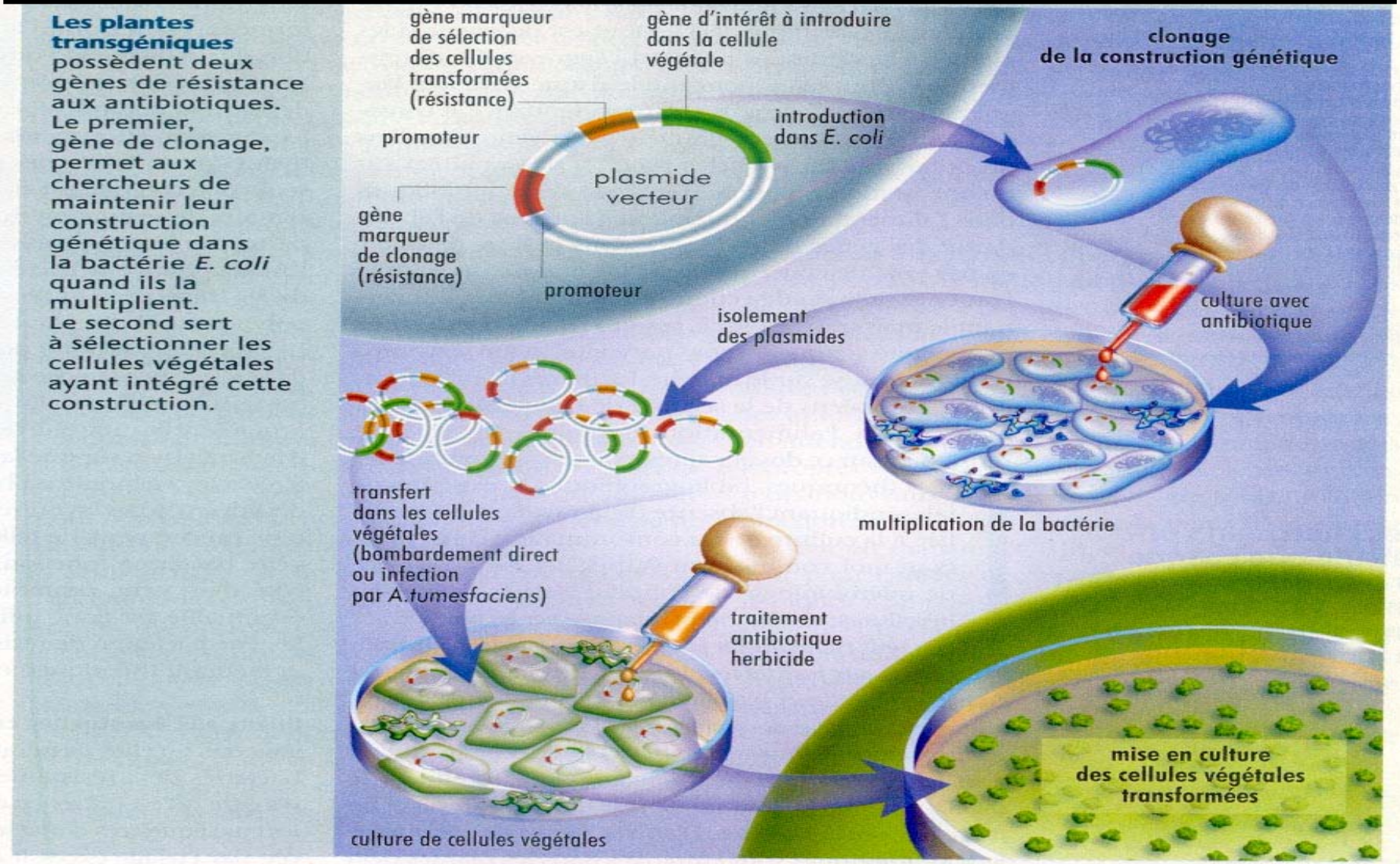
Alternatives:
gene regulation
or sequence modification

Alternatives: identification
Of sequences' changes or epimutations (DNA or chromatine)

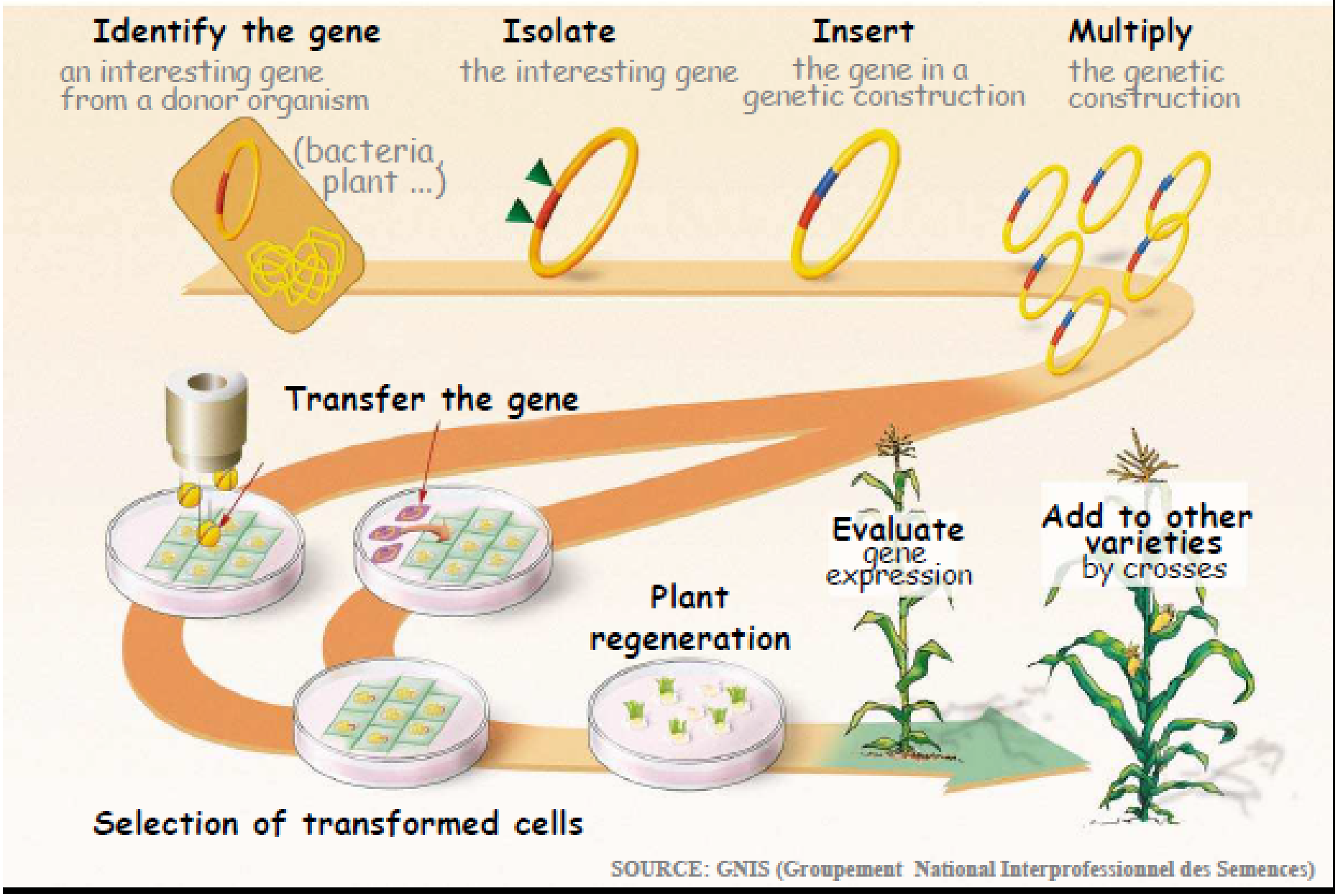
Sequences' changes
or epigenetic
insertion or deletion / indel
sequences in one or more locations…

http://sitn.hms.harvard.edu/flash/2015/how-to-make-a-gmo/

PRINCIPALES ETAPES DE LA TRANSGENESE VEGETALE



The steps involved in genetic modification



Common techniques to current GMOs and NTMGE

The NTMGE require the use of "old technologies" used for transgenic GMOs, already sold:

- protoplastisation, vectorization (very large proteins inducing huge holes in membranes, genome and plasmid of Agrobacterium leftovers…), cell cultures with their somaclonal variants, selection systems of modified cells with or without recombination systems, regeneration of the few not “recalcitrant” plants (thus with the same spectrum of species for GMOs and NTGME),
- all stressful techniques induced mutations and epimutations (up to 35% for cell cultures)
  - Hard to detect (missing reliable software and reference genomes, huge differences between Elite varieties of a species, systematic biases with each NGS platform) point or indels mutations, especially in non-coding or repeated regions, issues with translocations and inversions...
  - Difficult to discard (insufficient number of backcrosses by companies, co-segregations according to the characters, non Mendelian inheritable regions) leaving millions of bp not “cleaned" while the result is hard to control (issues of missing reliable software and reference genomes)

International Conference held in London in October 2016: labs are desperately looking for “good chefs”, well trained and regret the lack of training schools for future chefs…

Conclusion

- Numerous mutations and epimutations (of DNA, proteins and RNA) are induced by the technics used for both producing current GMOs and NTMGE, all scars usable in technics identification
- Several steps (e.g. modified cells selection tools with or without recombination tools, reagents delivery: Agrobacterium still the most efficient…) can leave foreign DNA traces,
- Purified proteins or RNA with contaminating DNA (see e.g. the issue of commercial PCR enzymes),

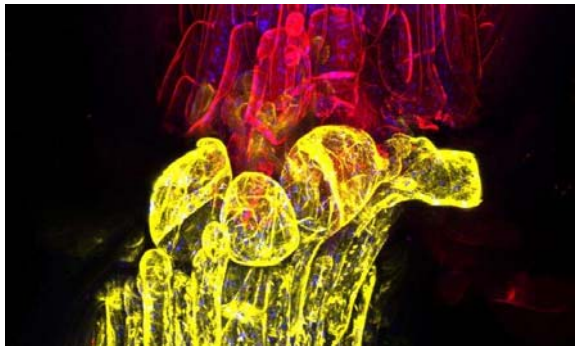
SOME ISSUES WITH NTMGE...

Graft: interactions between scion and (e.g. GM) rootstock

- Circulation of pathogens, proteins (e.g. Cry1Ac), DNA, ARN (siRNA and mRNA), hormones…
- Gene regulation (silencing…), proteins synthesis of rootstock in scions,
- Communication of genomes with establishment of epialleles
  http://phys.org/news/2016-01-grafted-genomes.html#jCp

The products of the non-GM scion (e.g. fruits) cannot be considered as not being influenced by the GM rootstock

Credit: Charles Melnyk/University of Cambridge



Sam van Aken, Syracuse University, New York.

RNAi

- Main issue: efficient delivery (agro-infiltration, biolistic, viruses) with different scars,
- Interaction with non-target organisms,
- Numerous off-targets

CRISPR-endonuclease

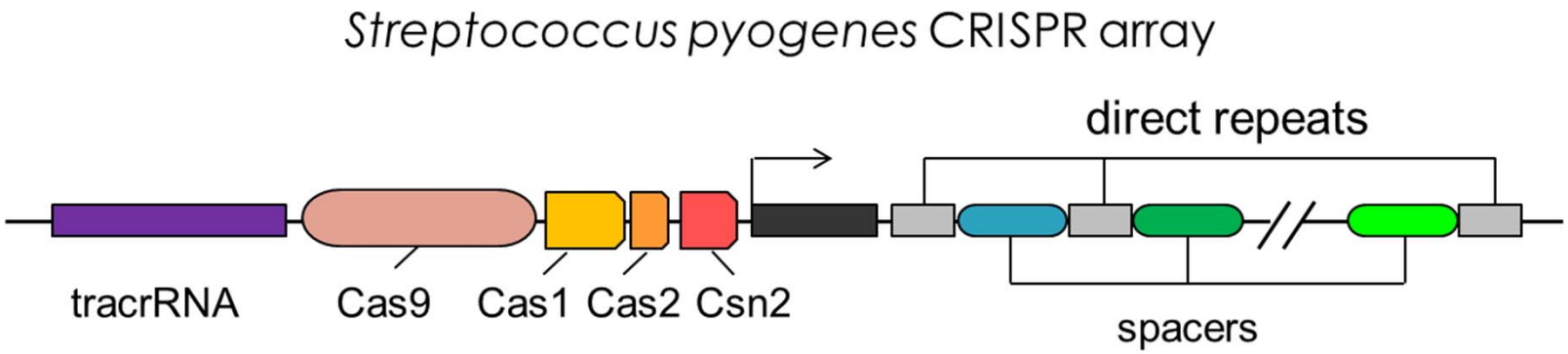
(6 classes, 19 sub-classes with functions slightly or fully unknown)

A bacterial “acquired adaptive immunity” against phages described since1987,
adapted for modifying the prokaryotic genomes in 2012, then to eukaryotic ones in 2013

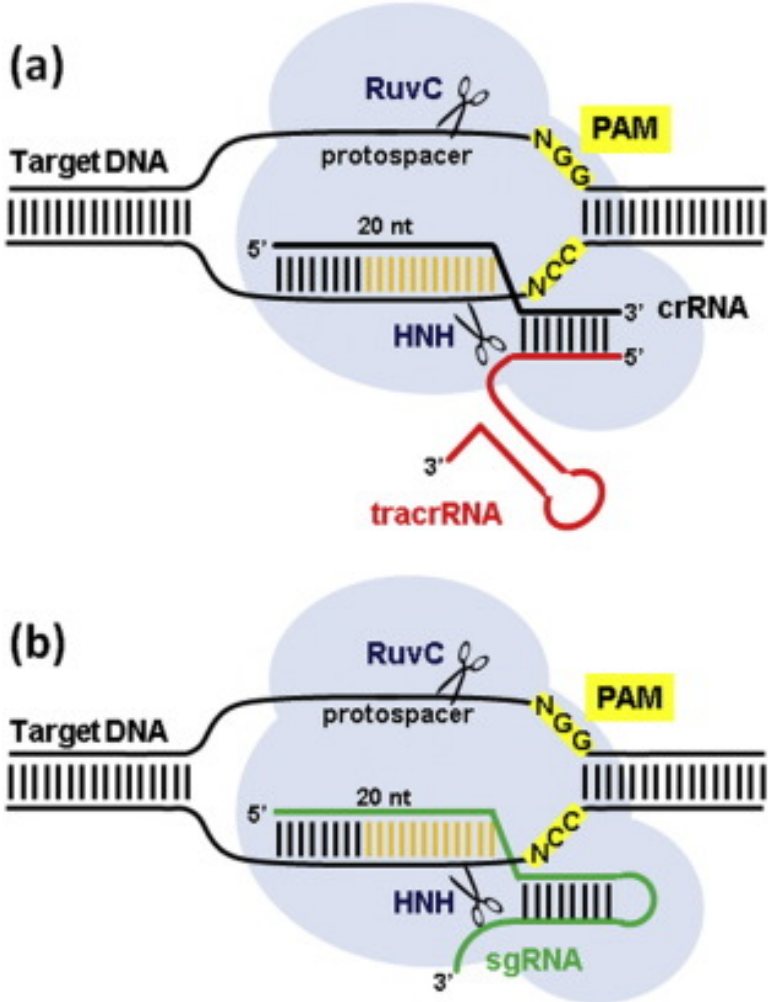
Evolutive convergence with animal and plants’ systems of genome stabilization
such as piRNA and transposable elements, small RNA and DICER…

Numerous limitations of use (species, genome regions…)

Main specific issues: unintended changes known as off-targets and exon skipping (providing aberrant proteins)



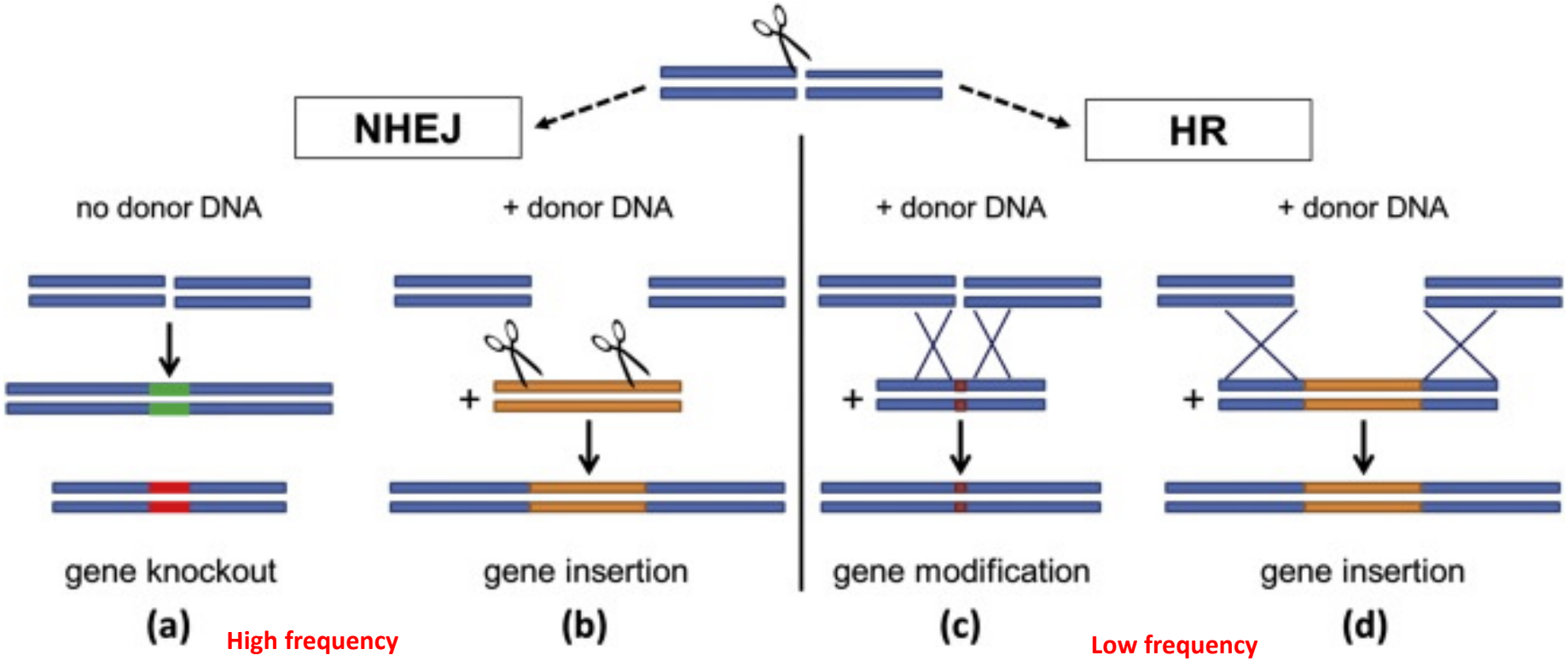
Crispr-Cas9



- Strong limitations of action in the genomes by anchoring PAM sequences what explains the races for new other nucleases using different PAM sequences e.g.: Cas9, C2c1, Cpf1 (RNA cutting), or modification to bring a PAM...
- Size limitations of inserts see e.g.: for human genes
- Numerous Off targets because SDN = sequence directed mutagenesis (sequences homologies)
- Reduction of off-target by changes - > nickase activity , reducing the quantities of reagents
- C2C2 (Cas13a) and RCas9 changes of RNA...

RNA-guided DNA cleavage by Cas9. (a) In the native system, the Cas9 protein (light blue) is guided by a structure formed by a CRISPR RNA (crRNA, in black), which contains a 20-nt segment determining target specificity, and a trans-activating CRISPR RNA...
Luisa Bortesi, Rainer Fischer Biotechnology Advances, Volume 33, Issue 1, 2015, 41–52

NTMGE: based on initially DSB DNA repair systems



Numerous unintended random mutations and epimutations, exon skipping, difficult to accurately
predict, hardly detectable even with whole genome sequencing and to discard by e.g. backcrosses

Main sources of uncertainties to be considered in health and environmental risks' assessment

- The trait(s) and its (their) regulation(s),
- The NTMGE, its variant and/or combinations, which cause different unintended effects,
- The type, quality and reliability of unintented changes monitoring,
- The mutations, epimutations and gene regulation changes due to the related techniques,
- The off-targets of NTMGE,
- The exon-skipping of NTMGE,
- The verified absence of nucleic acids in purified nucleases used in ribonucleoprotein complexes,
- The type, quality and reliability of procedures to discard (e.g. by back-crossing) and then monitor resting unintended changes.

RISK EVALUATION

Some rapid considerations about risk assessment

- As observed for current transgenesis: many unexpected effects for the NTMGE (in addition to the related techniques seen above), of the same nature and new: off-targets and exon-skipping
  - All procedures for GMOs should therefore apply to products NTMGE (= minimal),
  - new procedures have to developed about epigenetics issues (DNA, proteins and RNA = epitranscriptomics)
- Sequencing techniques, assembly and comparison software are still unreliable, data bases of sequences and annotations are still hugely biased,
- Number of backcrosses used by companies to discard mutations and epimutations are in practice drastically below the requested minimum number, tacking not into account issues of co-segregation, of non-Mendelian regions (particularly present in Elite cultivars), and the huge differences of sequences and regulations between Elite cultivars, all together with a lack of Gold reference genomes,
- nucleic acids can pass the intestinal barrier and regulate the expression of genes of the feedig host,
- a genetic change (in coding or non-coding regions) leads almost systematically to epimutations,
- a single modified nucleotide can have repercussions on several mega base pairs of distance,
- you can't compare a local modification to a natural mutation that if we study the entire genome for changes in
- absence of guidelines of assessment for epigenetics. EFSA meeting of june 2016 : .../... *The main take-home message from the colloquium was to ask and seek answers to those questions that will increase our understanding of epigenetics. What do epigenetic modifications mean? How do we study them? What is the size of such modifications that we need worry about? Dr Robert Feil of the National Centre for Scientific Research (CNRS), France, said: "We have had some very good discussions and I think this will certainly help us to formulate more precisely what these questions are and to formulate how we want to go ahead."*

Whether this is for health or the environment, complete dossiers as for GMO therefore appears as the minimum to be required to evaluate the potential risks of these techniques in development for over 10 years and to prepare effective follow-up post-marketing (specific and general surveillance)

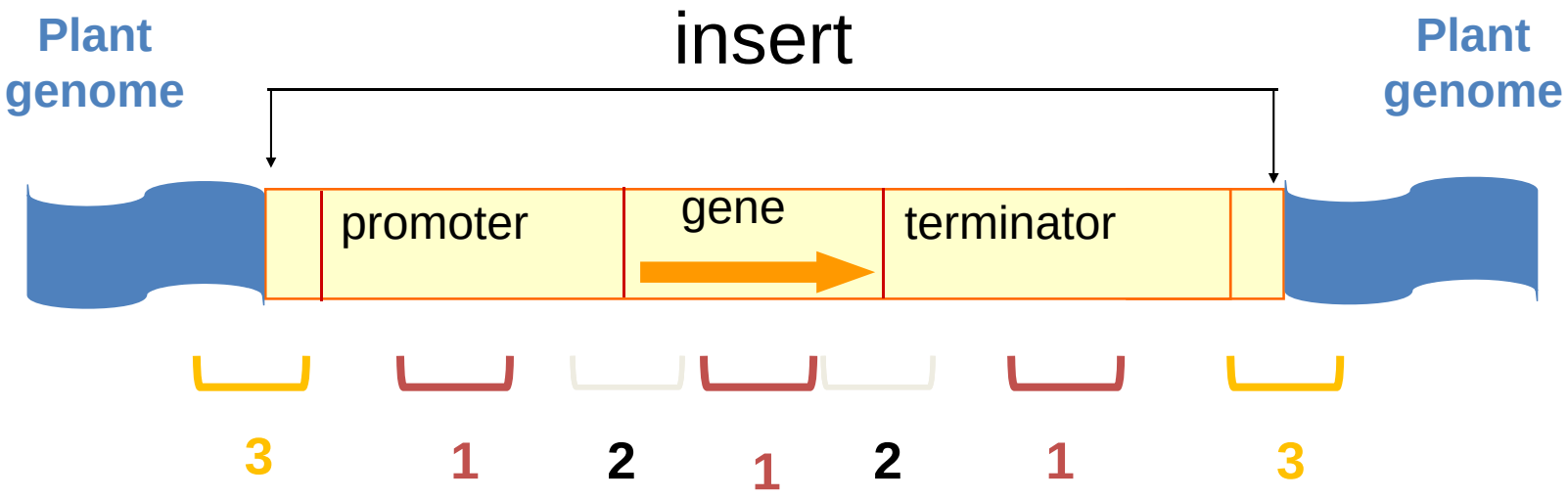
Dossier not as costly as claimed by companies (see the CIMMYT's publication)

Conclusion

- Related techniques induces numerous inheritable scars usable in differentiating *in vitro* and *in vivo* mutagenesis,
- NTGME specific needs (e.g. PAMs, safe harbor regions…) provide signatures which can be combined with mutagenized parts for identifying the NTMGE used,
- Off-targets and exon skipping both provide NTMGE signatures

NTMGE TRACEABILITY, DETECTION, AND IDENTIFICATION OF PRODUCTS AND TECHNIQUES USED

The routine GMOs detection method: PCR – Potential targets



1: Screening (P35S / Tnos / nptII…)
2: Construct-specific test
3: Event-specific test
4: Plant reference genes (from 1 per plant species to one for all)
5. Donor organism when needed (INRA's CaMV test saved tons of seeds)

Current unit: haploid genome
Request of companies and EC: to change into mass / kernels based unit for facilitating coexistence (important change for decreasing GMO contents of stacked genes)

GMO screening (taxa & controls)

Negative

GMO detection

Positive

Authorized?

GMO identification

Illegal

No

Yes

Assay individual ingredients

GMO quantification

No need for labeling

Less than 0.9%

Labeling required

More than 0.9%

4 kinds of targets: ubiquitous elements (screening), taxa reference genes, construct specific, identification by edge fragments (construct / plan genome)

NTMGE product traceability: a cleverly maintained confusion

- Traceability (ISO standard): easy and cheap, only depending on the goodwill of retailers: *Traceability* is the ability to identify and trace the history, distribution, location, and application of products, parts, materials, and services. A traceability system records and follows the trail as products, parts, materials, and services come from suppliers and are processed and ultimately distributed as final products and services.
- Detection: easy and relatively cheap depending on the techniques used:
  - Action or process of discovery, put highlight or note something.
  - Very easy to achieve when the target is known (Eve, patent, databases...), unknown GMOs: different methods the results sometimes lightweight convergent and facilitated by DSS
  - methods use: phenotype (ex: seedling tolerance to a herbicide, immunological methods, PCR, LCR, Qβ replicase, SNPLex, LAMP, (methods for omics) spectroscopy to the field and laboratories according to ese methods (ex) (: PCR and LAMP to the field)
  - No sequence limit size (from a nucleotide to large sequences)
- Quantification:
  - quantitative methods available,
  - For qualitative methods: the content can be reliably determined towards a threshold with methods of sub-sampling as for example those used in seeds' certification.

Identification of the used NTMGE / the owner of the product

Generally speaking, stresses leave traces in genomes and epigenomes

- Profiles of mutations and epimutations (all techniques, scars and off-targets, exon-skipping…) vectorization traces, DNA contaminants of RNP, mutations hot-spots, hypermutations…
- Specific NTMGE signatures,
- systems of molecular structuring, cell lines profiling and “life history” (see for instance cancer cells typing)...
- From the “*matrix approach*” (used for unknown GMOs) to
  - unambiguous / univocal signatures...
  - *in vitro* and *in vivo* mutagenesis origins.

A scientific peer-reviewed paper in preparation…

NTMGE product traceability

- European Research projects on NTGME have to be launched, as in the 90s' for current GMOs, for a proof of concept,
- Research labs will then transfer techniques, particularly those using univocal signatures, to routine enforcement labs,
- Patents and scientific papers watch to feed databases as usual,
- Cost are not expected to increase, not impacting consumers,
  - Even those using NGS (see Oxford Nanopore claims),
  - EC funded research projects provide rapidly the proof of concept, after a necessary moratorium period due to the current delay in starting the research projects,
  - Companies supply, as for the current GMOs, the NTMGE products' signatures, defined by the research labs,
  - Standardization and inter-laboratory trials are performed by e.g. ENGL,
  - By the application by the supply chains' actors of their current traceability procedures and controls.

Conclusion

- Detection methods are available: from simplex to multiplex,
- Inherited targets are available: from techniques scars and NTMGE specific ones to general cell lines profiling,
- Costs are not expected to increase provided:

NTMGE AND FEEDING THE WORLD

NTMGE and rapid plant breeding?

- The NTMGE: really a revolution?
- A plus in varietal selection?
  - Reduction of the time of screening (greenhouse vs. fields)? selection and integration of new traits?
  - New characters? Homologies and variations between species and cultivars? Mono features / oligo / multigenic? QTL? A variation in a variety is not necessarily in the same way expressed in another (effect of the genetic background of the environment...), epigenetics of culture conditions of parents influence the expression of characters in progenies (cf. drought resistance...),
  - an ability to improve the resilience of crops in the face of climate change? Emerging pests and diseases?
  - Oriented products consumers rather than farmers? The effect of other supply chains actors…

An economy of the promise as for GM of the 1990s and cloning 20 years ago

METAPHORS, RHETORIC AND COMMUNICATION

Some misleading metaphors…

SDN = **S**equence **D**irected **N**uclease: do not expect an unique and precise cut by “molecular scissors”

But a series of cuts (with numerous small pieces to join by an uncontrolled error prone repair mechanism)

“Genome editing”… Do you expect known languages in electronic writing?

You actually have to edit handwritten texts between different unknown languages

The specific promised modification?

Numerous 'off target’ due to ‘ricochets’ among homologous sequences

In fact it is as Stalin’s organs

Providing such cells and genomes to reconstruct

Targeted mutagenesis: you imagine one clean ‘one shot’?

i.e. numerous side-effects due to both related techniques and NTGME

Communication

- Over-simplified (erasing uncertainties, without statistics...) as recommended in the guidelines on risks' communication,
- Marketing strategies: decoy and lure, loss leader,
- Economy of promise (as previously done for cloning, Gos derived pharmaceuticals from GMOs… several decades ago)
- Reproducibility issues, absence of standardization, quality control, claims not supported by the results (see e.g. US NIST's *Genome Editing Consortium*) and investment bubbles,

For an ethical questioning of the debate around the new genome modification techniques

- Confusion maintained between science and innovation, scientist and expert, applications such as “gene drive” and NTGME techniques,
- a semantic and rhetorical struggle as to the question of nature, outside the context of evolution,
- an attempt to privatize of the notion of general interest...
- A capture of rent: see the “mosquito gene drive” strategy vs the previous public policies of suppressing mosquitos’ multiplication sites, as cited by the WHO director in 2016,
- Virtually unheard among the experts: the link of interest (including institutional), and the need for neutrality and transparency,
- a condescension policies of several experts with regard to laymen who don't stick to their view,
- an approach by the technological HOW, not by the WHY and for WHO,
- technical limitations defavouring again the agricultural and ordinary biodiversities,
- unknown and unenforced ethics guidelines,
- A politicization of science and a scientification of politics…

As recently recalled by Emmanuelle Carpenter "for some applications, there is still the problem of the so-called mutations "off-target"- i.e. unintended-, which may never be completely excluded.

Jennifer Doudna : "*This really underscores the fact that we don't know enough about nature to anticipate all the ways that nature has come up with to manipulate DNA. Nature's had a lot longer than we have to be tinkering. We really need fundamental research to uncover those basic mechanisms that drive the development of technologies in the future.*"

After years in expertise committees, particularly in the HCB, I like to remember this sentence of *Pierre Gilles de Gennes* (Nobel Prize):

« Vous savez, les experts sont souvent comme les militaires. Ils sont experts de la dernière guerre mais pas de la prochaine… »

"You know, the experts are often like the military. They are experts of the last war but not the next…"