Biología Molecular aplicada al Diagnóstico Médico
Módulo I: Clase 1

Introducción a la Biología Molecular y su aplicación a la Medicina

Bioq. Ma. Florencia Gosso, PhD
Círculo Médico de Rosario
19.03.2018

La **biología molecular** constituye el área de la biología que estudia la estructura, contexto y función de moléculas de ADN, ARN y proteínas.

Se encuentra relacionada con otras ciencias como la *bioquímica* y la *genética*.

En la actualidad, área de soporte fundamental en el contexto de la **Genética Médica**

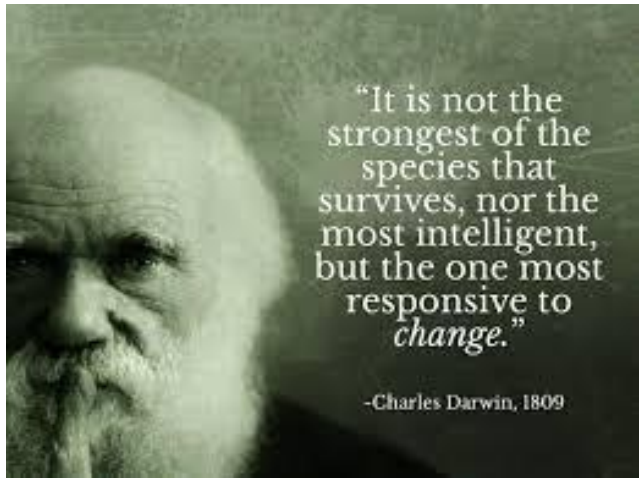
Héritas
Medicina de precisión
CIBIC + INDEAR

La **Genética** es la ciencia que estudia la *variación*, la *diversidad biológica* y la *herencia*

La **Genética Médica** estudia los aspectos genéticos en la especie humana y su relación con la salud y la enfermedad, así como su aplicación al diagnóstico, pronóstico y asesoramiento de enfermos y familiares

Héritas
Medicina de precisión
CIBIC + INDEAR

Hitos de la Biología Molecular
(Pre-Genomic Era)



1859
Charles Darwin
"Origin of Species"

1859
George Mendel
Leyes Mendelianas
de la Herencia

1871
Miescher
Descubrimiento del ADN
("nucleína")

Héritas
Medicina de precisión
CIBIC + INDEAR

1953

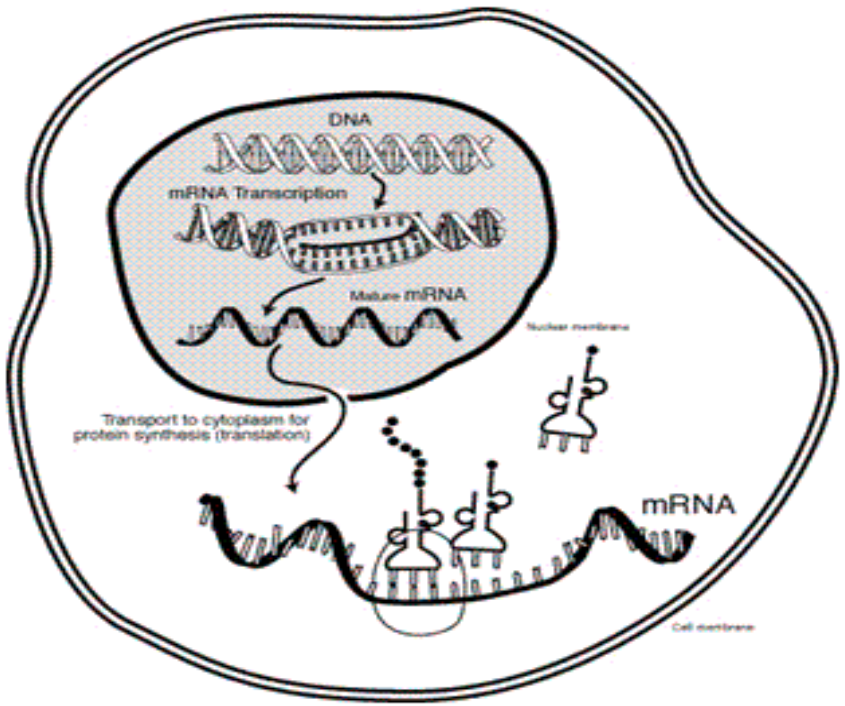
Hitos de la Biología Molecular *(Pre-Genomic Era)*

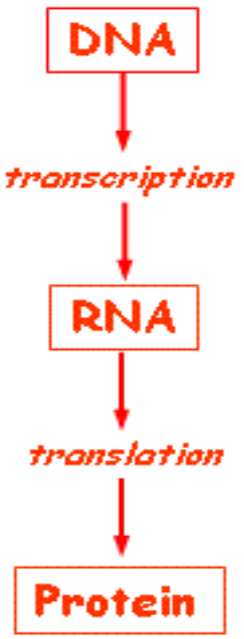
Francis Crick

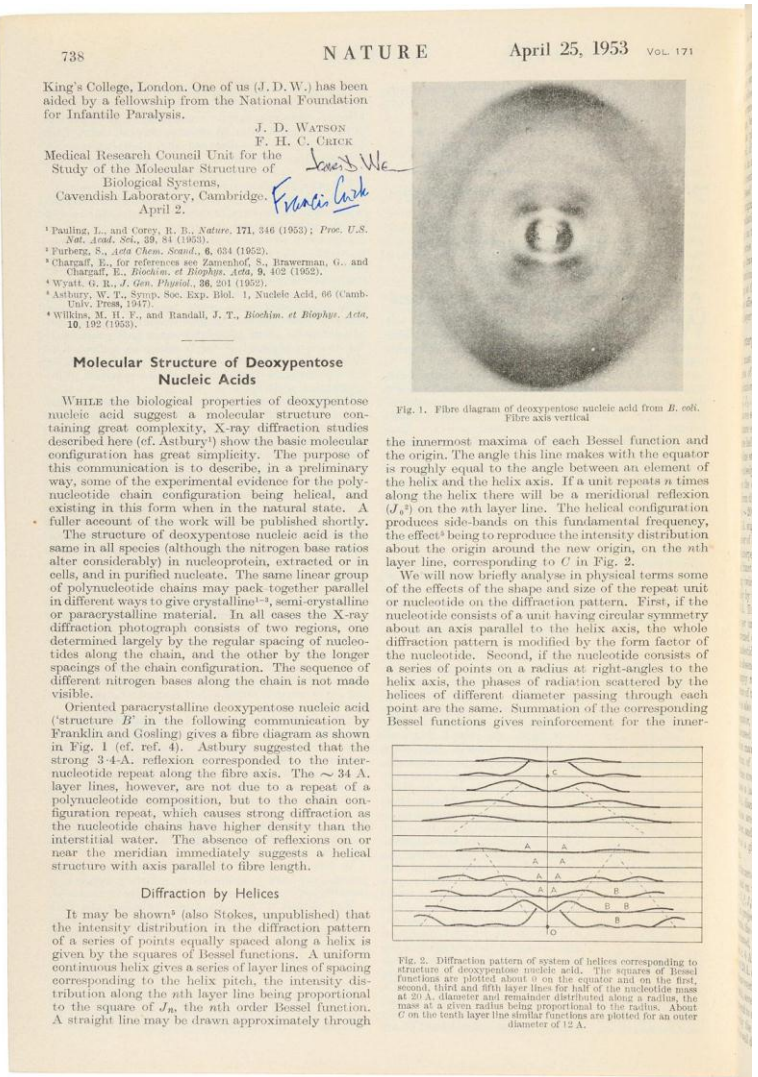
James Watson

Maurice Wilkins

Rosalind Franklin







738 NATURE April 25, 1953 Vol. 171

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. Watson
F. H. C. Crick

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge. April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Cambridge Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

While the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1–3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure *B*' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-Å. reflexion corresponded to the internucleotide repeat along the fibre axis. The ~ 34 Å. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the nth layer line being proportional to the square of J_n, the nth order Bessel function. A straight line may be drawn approximately through

Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats n times along the helix there will be a meridional reflexion (J_n) on the nth layer line. The helical configuration produces side-bands on this fundamental frequency, the effect[6] being to reproduce the intensity distribution about the origin around the new origin, on the nth layer line, corresponding to C in Fig. 2.

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about 0 on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 Å. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About C on the tenth layer line similar functions are plotted for an outer diameter of 12 Å.

Dogma central de la Biología
(Pre-Genomic Era

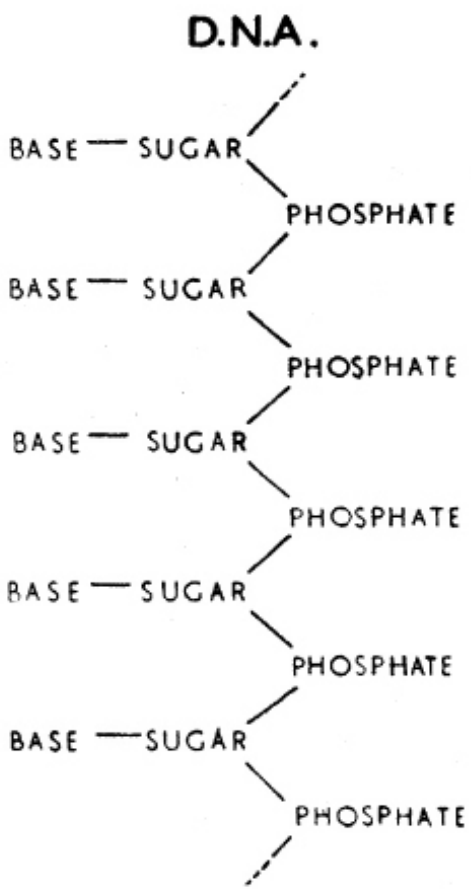


Fig. 1. Chemical formula of a single chain of deoxyribonucleic acid

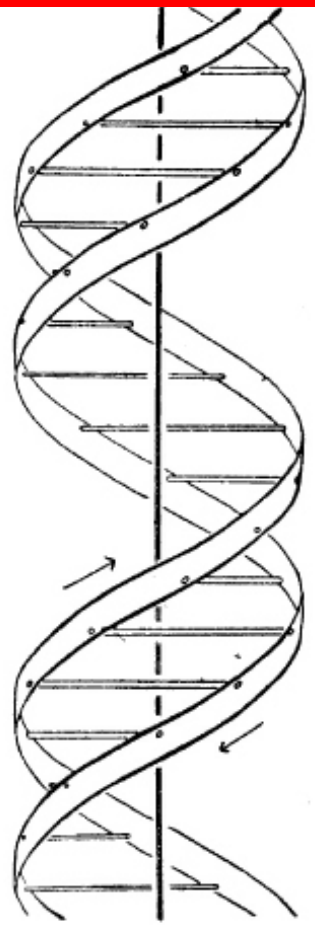
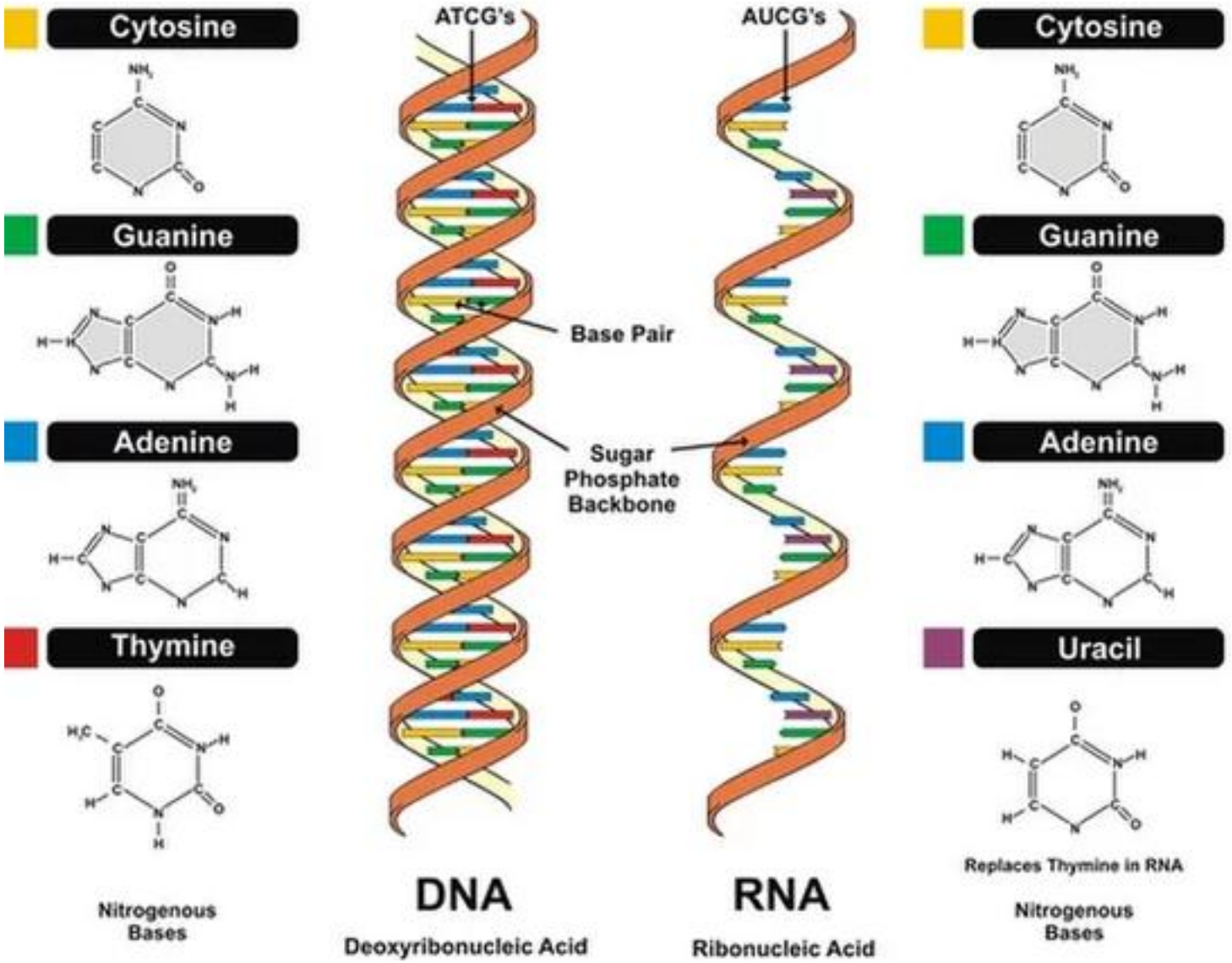


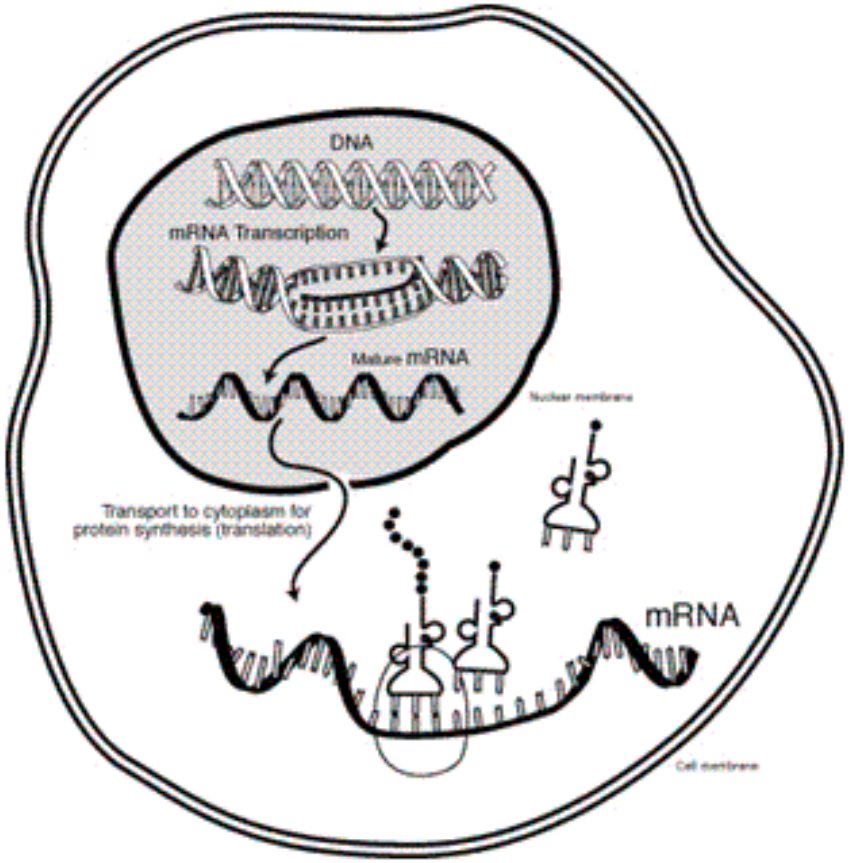
Fig. 2. This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

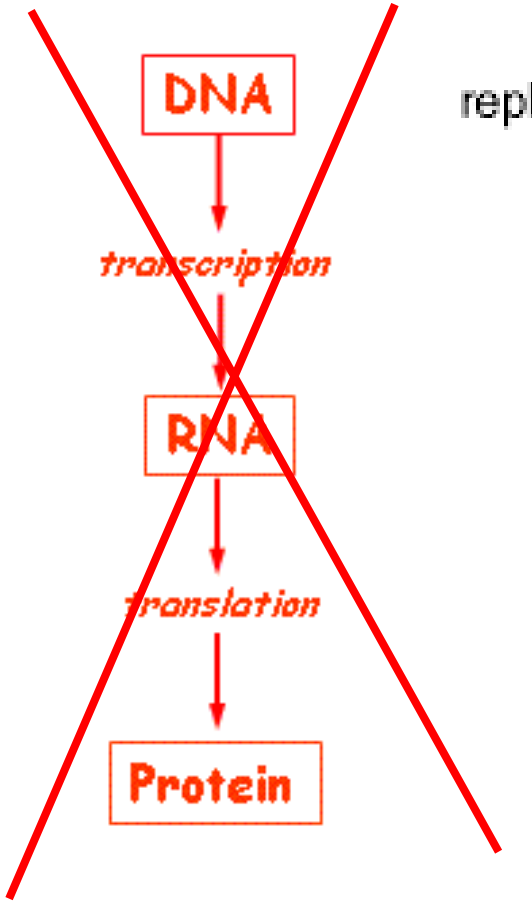
Estructura del DNA /RNA

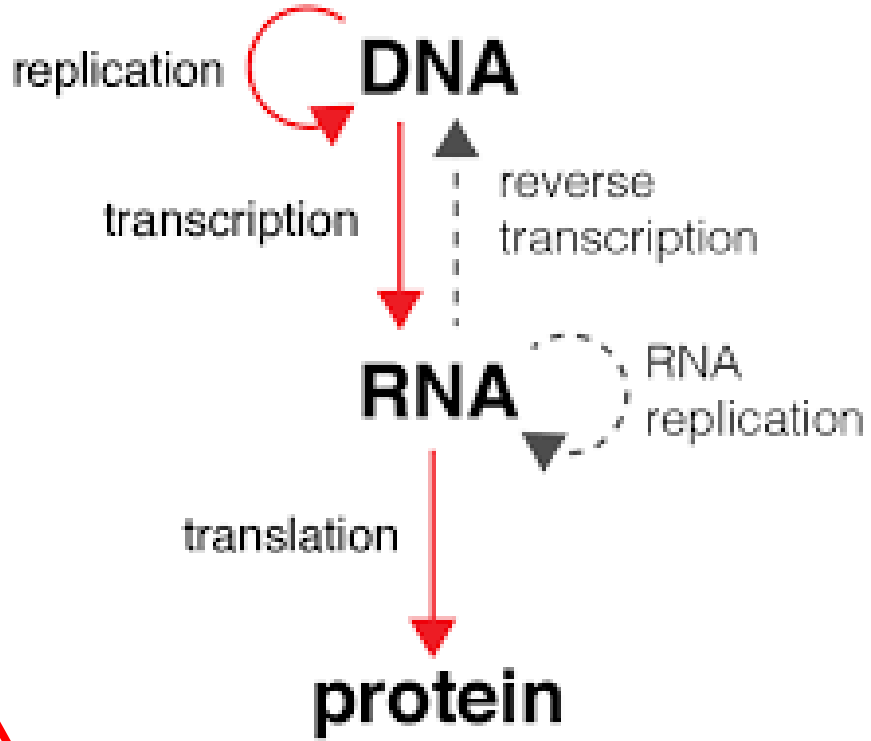


Héritas
Medicina de precisión
CIBIC + INDEAR

Dogma central de la biología molecular

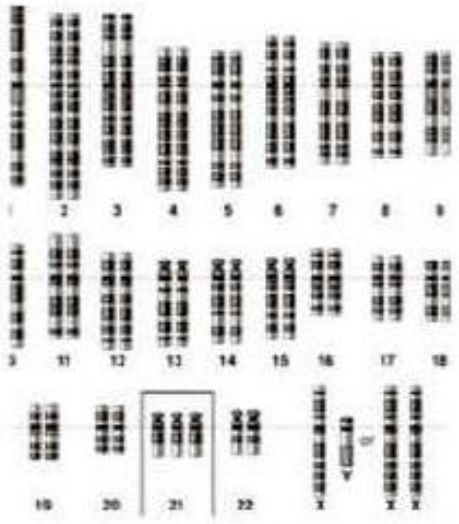






Héritas
Medicina de precisión
CIBIC + INDEAR

Hitos de la Biología Molecular
(Pre-Genomic Era)



1959
Jerome Lejeune
Síndrome de Down

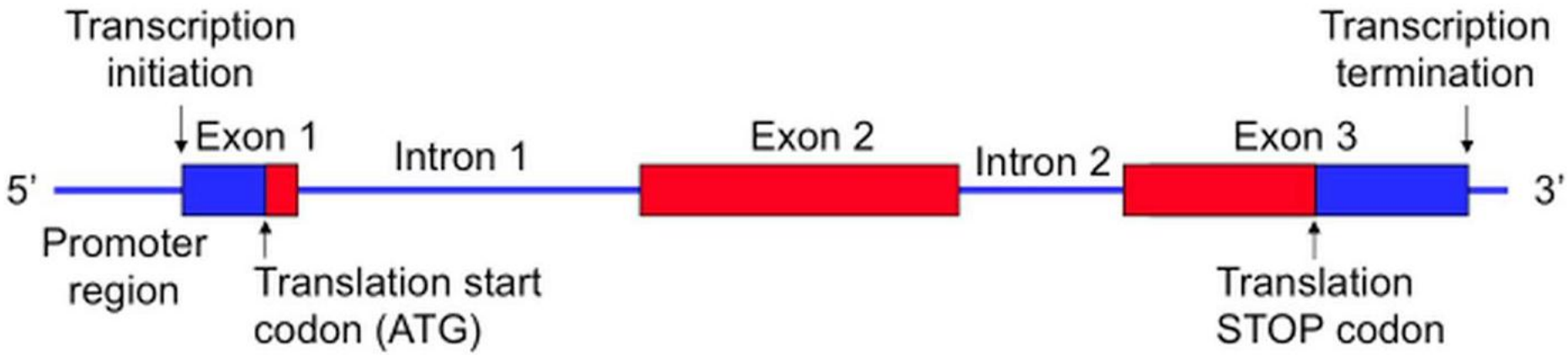
1983-85
Kary Mullis
PCR
Polimerase chain reaction

1990
Proyecto Genoma Humano

1996
Clonación (Dolly)

Héritas
Medicina de precisión
CIBIC + INDEAR

Estructura del gen en Eucariotas
(Pre-Genomic Era)



Héritas
Medicina de precisión
CIBIC + INDEAR

PROYECTO GENOMA HUMANO

1990 - 2003

"Adquirir la **información** fundamental sobre nuestro material genético para profundizar en el conocimiento de la genética humana y el *papel de los distintos genes en la salud y la enfermedad*"

Francis Collins
Director Human Genome Project

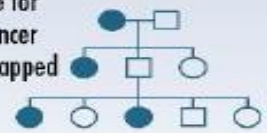
Héritas
Medicina de precisión
CIBIC + INDEAR

Hitos de la Biología Molecular *(Genomic Era)*

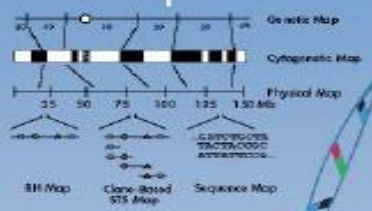
1990
Human Genome Project (HGP) launched in the U.S.

Ethical, Legal, and Social Implications (ELSI) programs founded at NIH and DOE

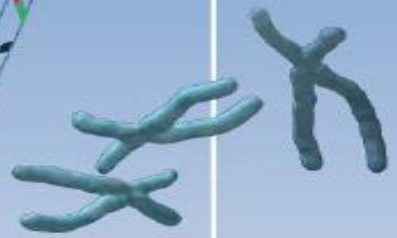
First gene for breast cancer (BRCA1) mapped



1991
First U.S. Genome Centers established



1992
Second-generation human genetic map developed



Rapid data release guidelines established by NIH and DOE

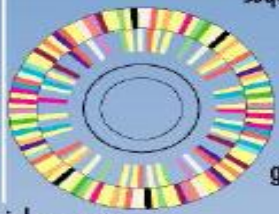
1993
New five-year plan for the HGP in the U.S. published

Sanger Centre founded (later renamed Wellcome Trust Sanger Institute)

The Wellcome Trust

1994
HGP's human genetic mapping goal achieved

1995
HGP's human physical mapping goal achieved



First bacterial genome (*H. influenzae*) sequenced

U.S. Equal Employment Opportunity Commission issues policy on genetic discrimination in the workplace

1996
First human gene map established

Pilot projects for human genome sequencing begin in U.S.

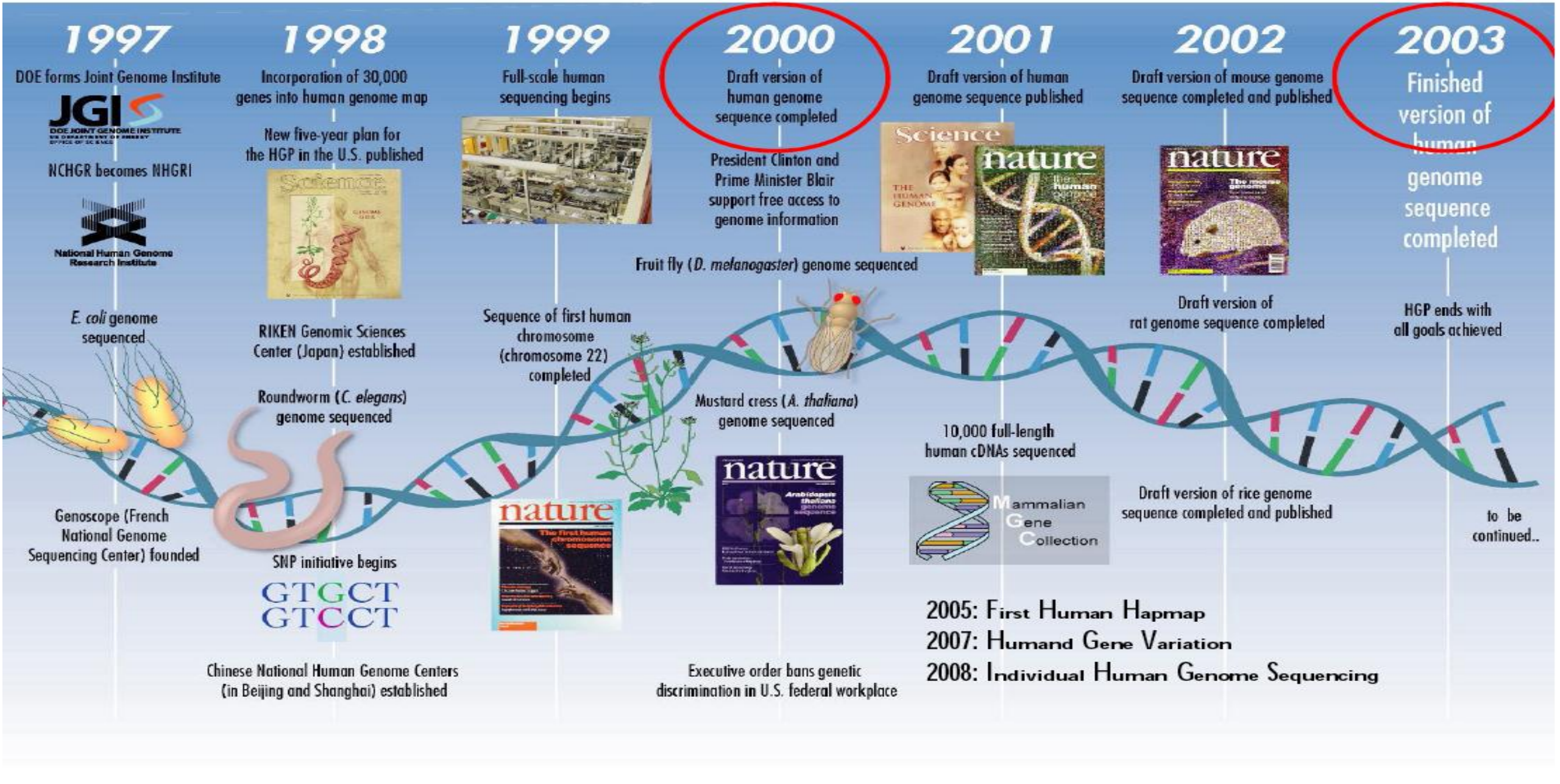
First archael genome sequenced

Yeast (*S. cerevisiae*) genome sequenced

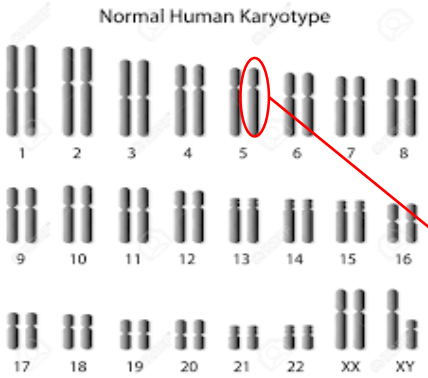
HGP's mouse genetic mapping goal achieved

Bermuda principles for rapid and open data release established

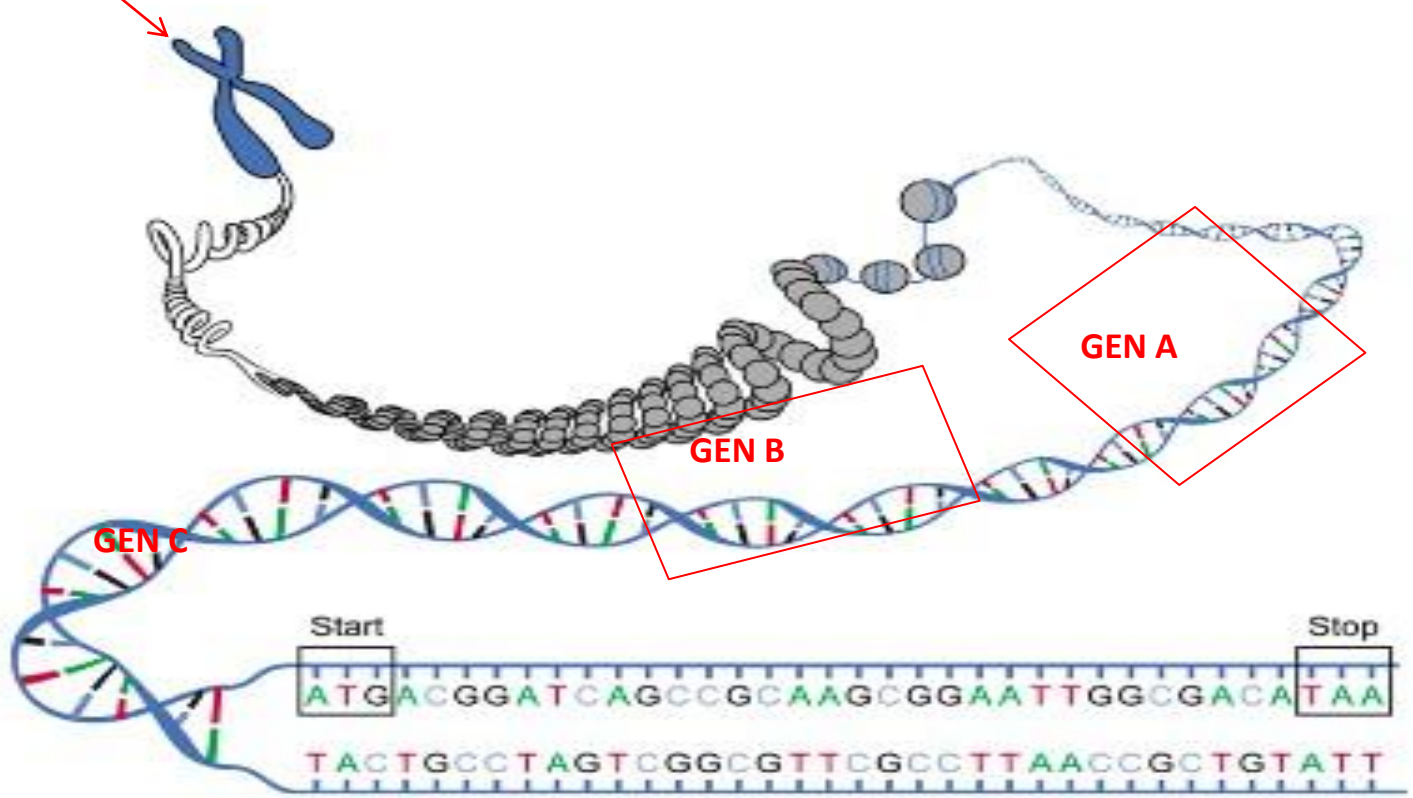
HISTORIA DE LA GENOMICA: EL PROYECTO GENOMA HUMANO



Proyecto genoma humano
OBJETIVOS



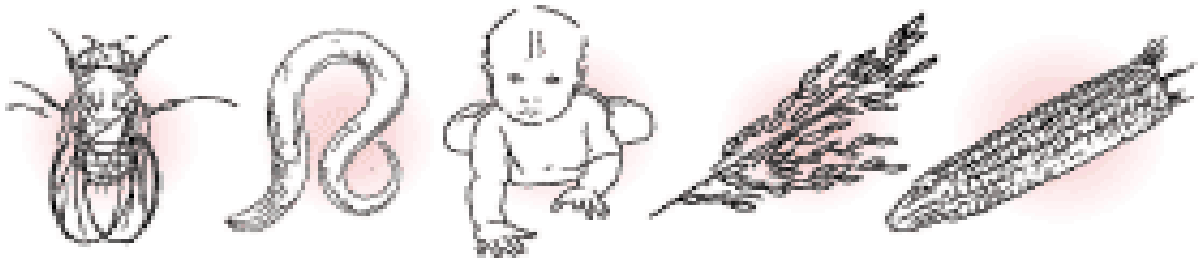
- Identificación de ≈ 30.000 genes
- Determinar la secuencia de ADN (3 billones pb)
- Generación de Bases de Datos públicas
- Mejorar herramientas para análisis de datos
- Redefinir principios éticos, legales y sociales



Héritas
Medicina de precisión
CIBIC + INDEAR

Humans have fewer genes

In Thursday's issue of the journal Nature, researchers who decoded the human genome concluded that people have only 20,000 to 25,000 genes, a drop from the 30,000 to 40,000 estimated in 2001.

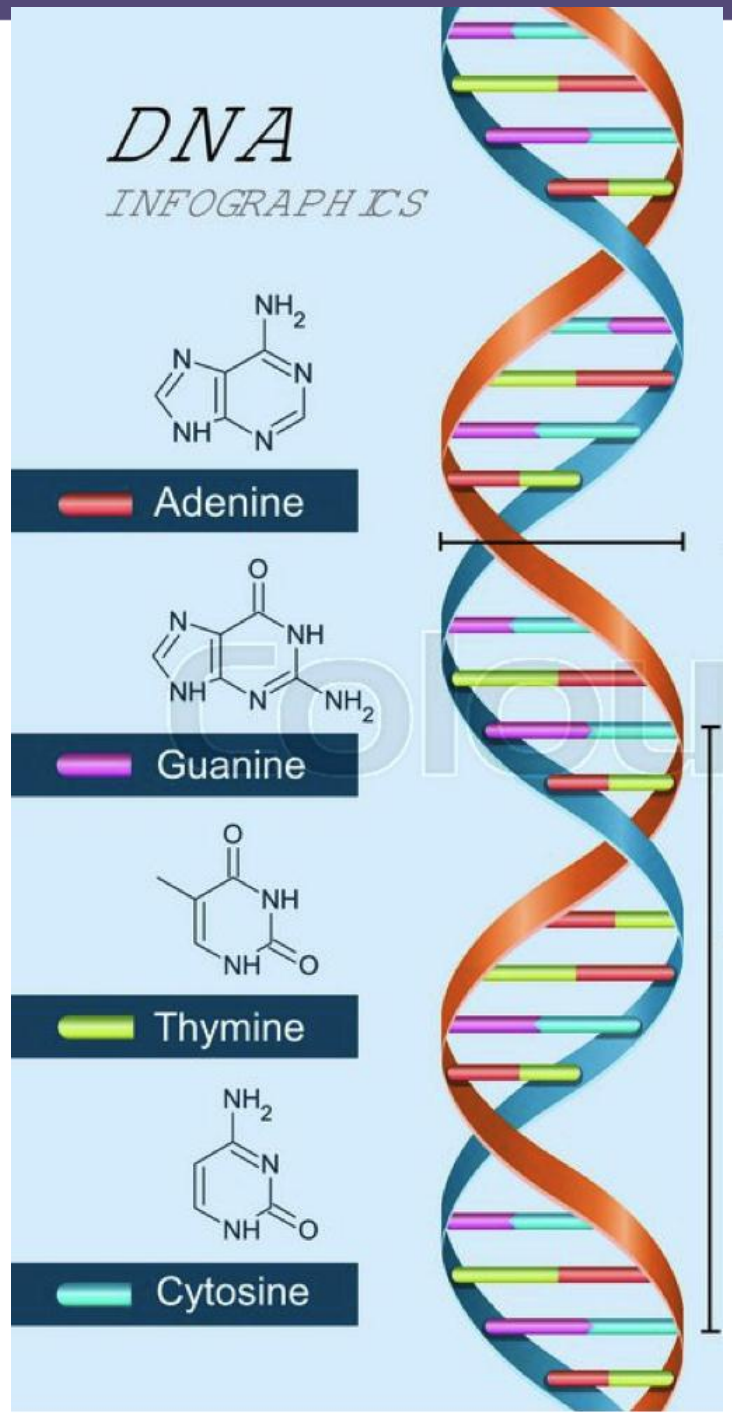


	Fruit fly	C. Elegans	Humans	Rice	Maize
Genes	13,600	19,500	20,000 - 25,000	45,000	50,000

SOURCE: Nature

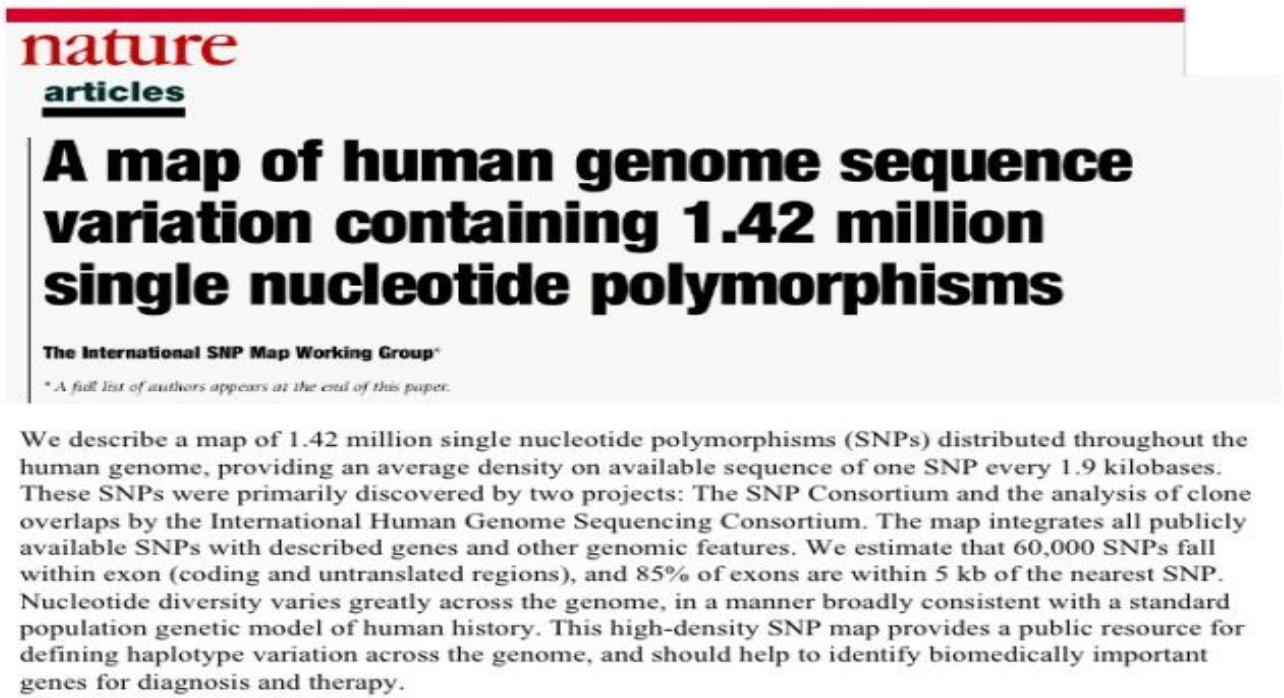
AP

Héritas
Medicina de precisión
CIBIC + INDEAR



- DNA: 3 billones de pb
- Entre 20,000-25,000 genes (50% con función desconocida)
- 98-99% del genoma es no-codificante
- 99,9% idéntico entre individuos
- Compartimos el 99,9% de nuestra secuencia génica con nuestros congéneres, y más del 99% con otras especies
- Se estiman alrededor de $10x10^6$ SNPs
  Single **N**ucleotide **P**olimorphism (1 SNP c/1000 pb)

Héritas
Medicina de precisión
CIBIC + INDEAR



nature
articles

A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms

The International SNP Map Working Group*

* *A full list of authors appears at the end of this paper.*

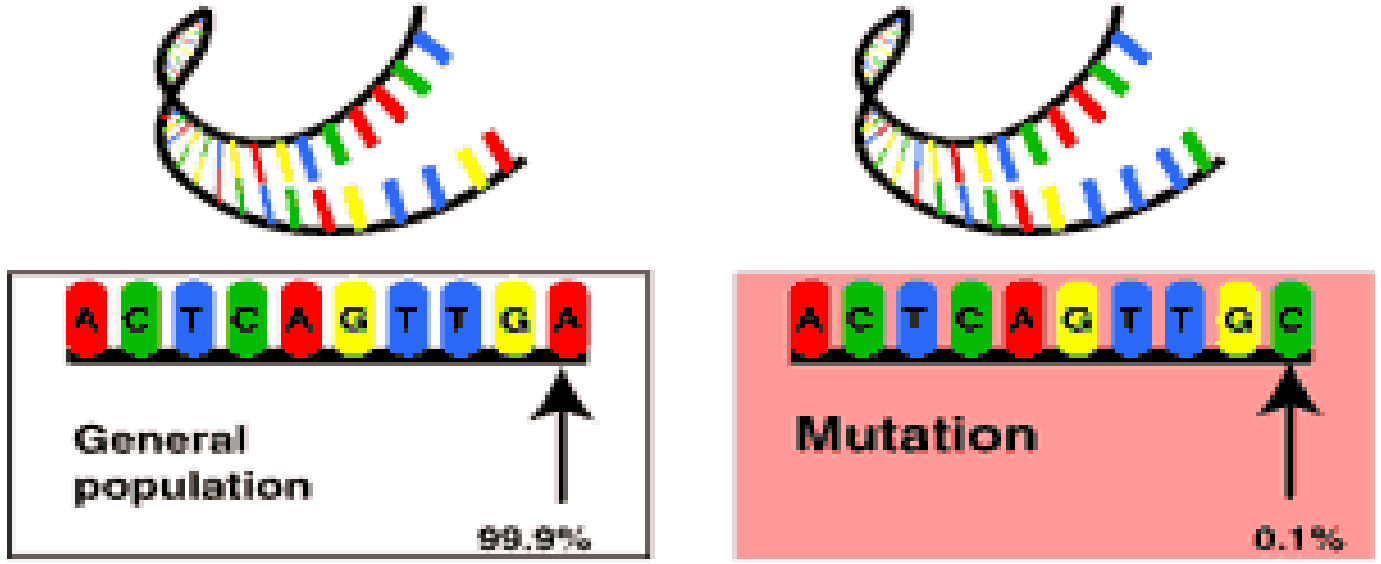
We describe a map of 1.42 million single nucleotide polymorphisms (SNPs) distributed throughout the human genome, providing an average density on available sequence of one SNP every 1.9 kilobases. These SNPs were primarily discovered by two projects: The SNP Consortium and the analysis of clone overlaps by the International Human Genome Sequencing Consortium. The map integrates all publicly available SNPs with described genes and other genomic features. We estimate that 60,000 SNPs fall within exon (coding and untranslated regions), and 85% of exons are within 5 kb of the nearest SNP. Nucleotide diversity varies greatly across the genome, in a manner broadly consistent with a standard population genetic model of human history. This high-density SNP map provides a public resource for defining haplotype variation across the genome, and should help to identify biomedically important genes for diagnosis and therapy.

Los SNP constituyen hasta el 90% de todas las variaciones genómicas humanas, y aparecen en promedio, cada 1,000 pb

Estas variaciones en la secuencia del ADN pueden afectar la respuesta de los individuos a enfermedades herditarias, enfermedades infecciosas y respuesta a fármacos

Héritas
Medicina de precisión
CIBIC + INDEAR

POLIMORFISMO: cambio en la secuencia del ADN en donde al menos dos secuencias diferentes pueden estar presentes, cada secuencia presente en al menos 1% de la población sin asociarse en forma directa con el desarrollo de enfermedad



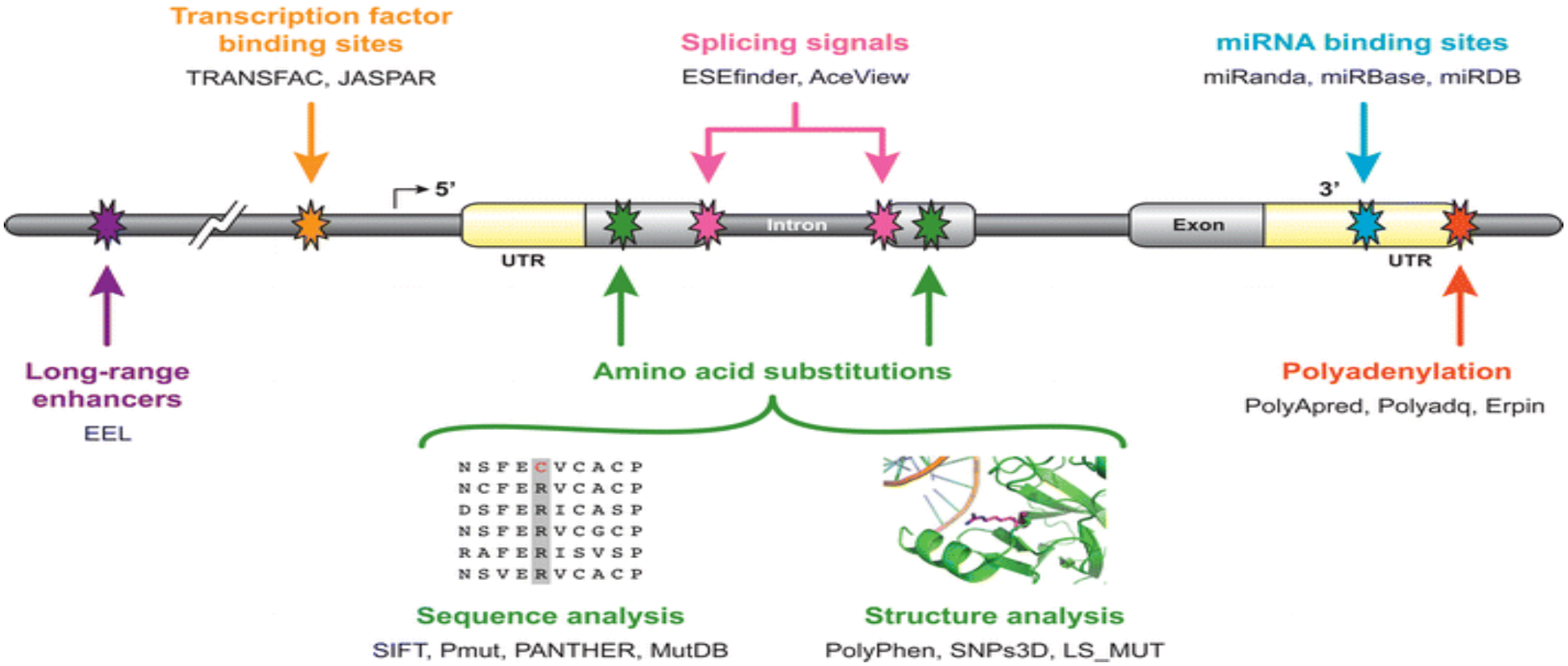
MUTACIÓN: MAF (Minor allele frequency) < 1%

POLIMORFISMO: MAF >1%, existen por lo menos 2 alelos

Héritas
Medicina de precisión
CIBIC + INDEAR

SNP codificantes: se localizan en la secuencia codificante pueden modificar o no la cadena de aminoácidos de proteínas> **Variabilidad genética poblacional**

SNP no-codificantes: en regiones intronicas, 5'UTR, 3'UTR, promotora. Consecuencias en el proceso de traducción (*splicing*, factores de transcripción) > **Reguladores**



Héritas
Medicina de precisión
CIBIC + INDEAR

Hallazgos

- La complejidad del genoma no radica en el nro. de genes, sino en la *interacción* entre ellos

  ➔ Genes reguladores
  ➔ Genes con más de una función
  ➔ Moléculas reguladoras (ADN,ARN,mARN,siARN)
  ➔ La mayoría de los genes son polimórficos (variabilidad genética)

Estrategia biológica para la evolución y para la supervivencia
(HGP 2001)

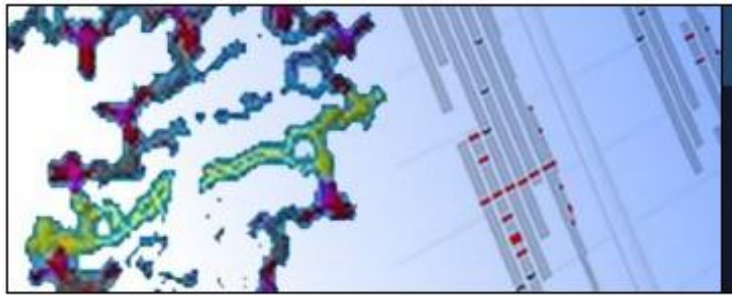
"It is not the strongest of the species that survives, nor the most intelligent, but the one most responsive to change"
(Charles Darwin, 1809)

Héritas
Medicina de precisión
CIBIC + INDEAR

Proyecto genoma humano
BASES DE DATOS

Héritas
Medicina de precisión
CIBIC + INDEAR

Proyecto genoma humano
BASES DE DATOS



dbSNP

Database of single nucleotide polymorphisms (SNPs) and multiple small-scale variations that include insertions/deletions, microsatellites, and non-polymorphic variants.

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

OMIM

OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh. Its official home is omim.org.

Héritas
Medicina de precisión
CIBIC + INDEAR

EL PROYECTO GENOMA HUMANO

Héritas
Medicina de precisión
CIBIC + INDEAR

2003

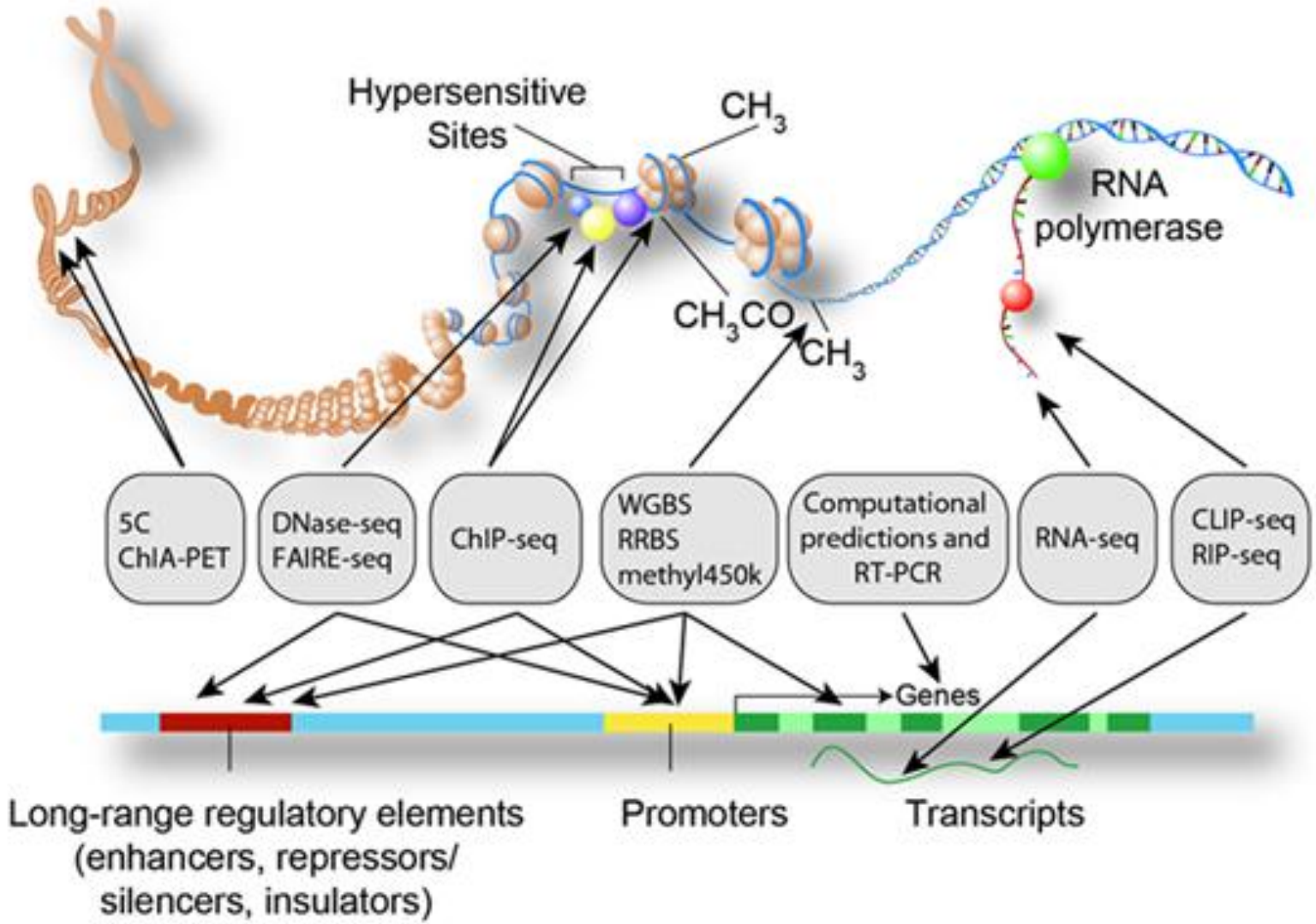
Hitos de la Biología Molecular

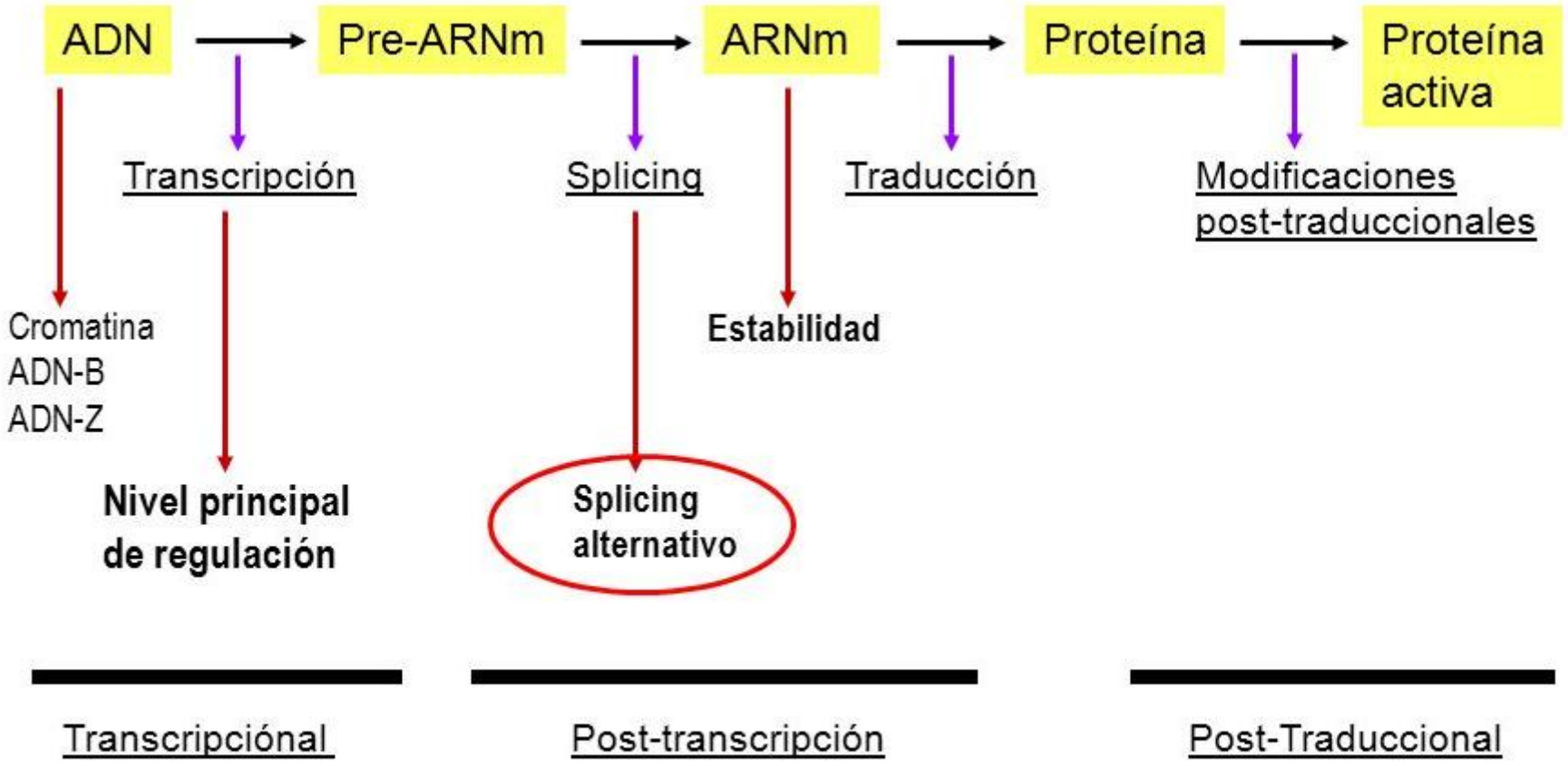
ENCODE (ENCyclopedia Of DNA Elemens)

National Human Genome Research Institute

OBJETIVOS

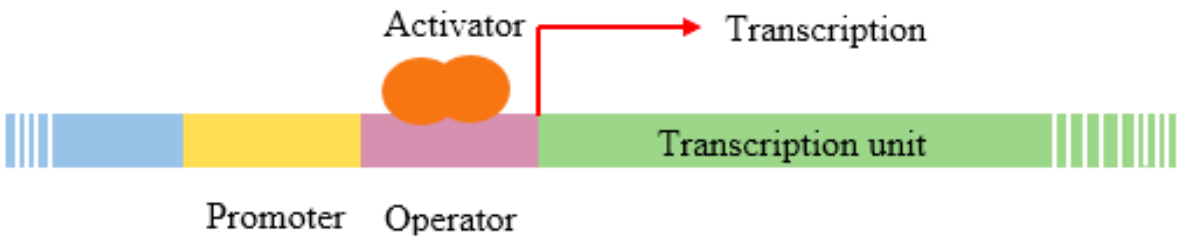
Identificación de regiones de transcripción, factores asociados, estructura de la cromatina y modificación de histonas

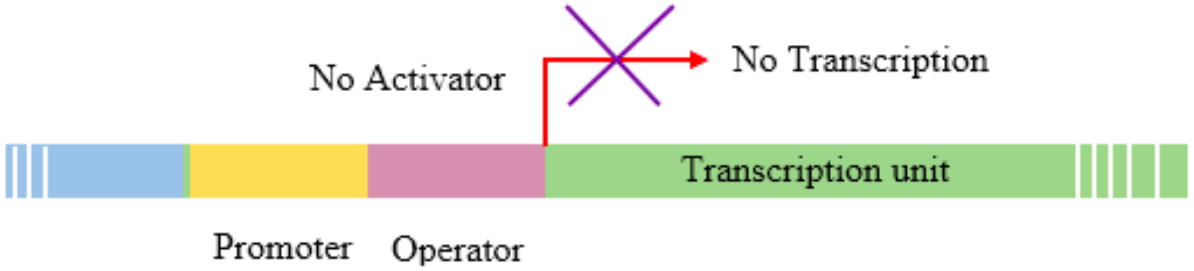




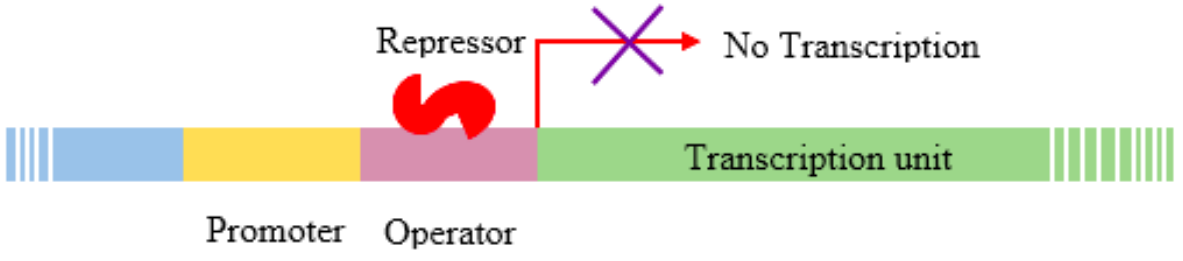
REGULACION GENÉTICA

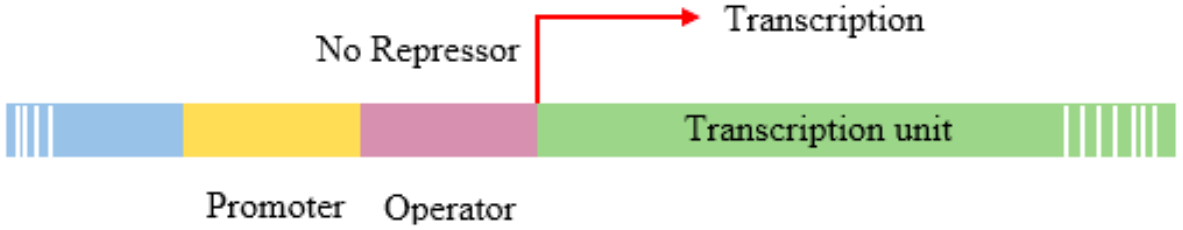
Héritas
Medicina de precisión
CIBIC + INDEAR





REGULACIÓN POSITIVA



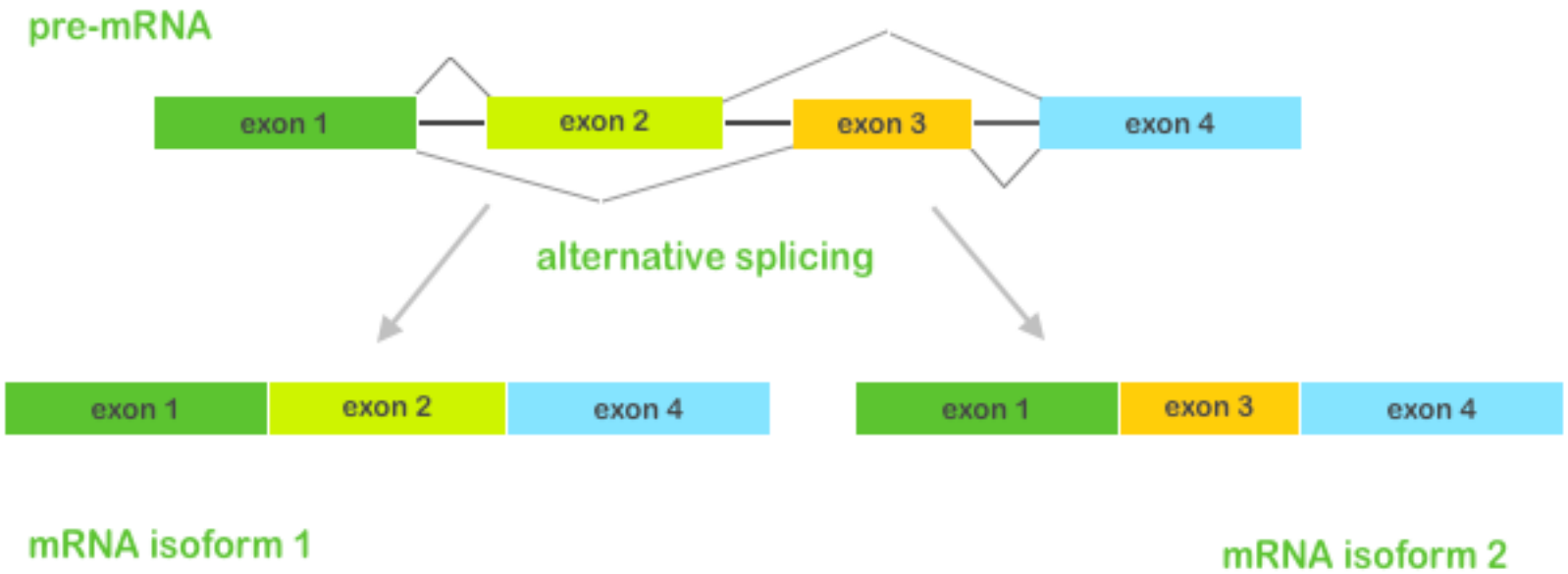


REGULACIÓN NEGATIVA

Héritas
Medicina de precisión
CIBIC + INDEAR

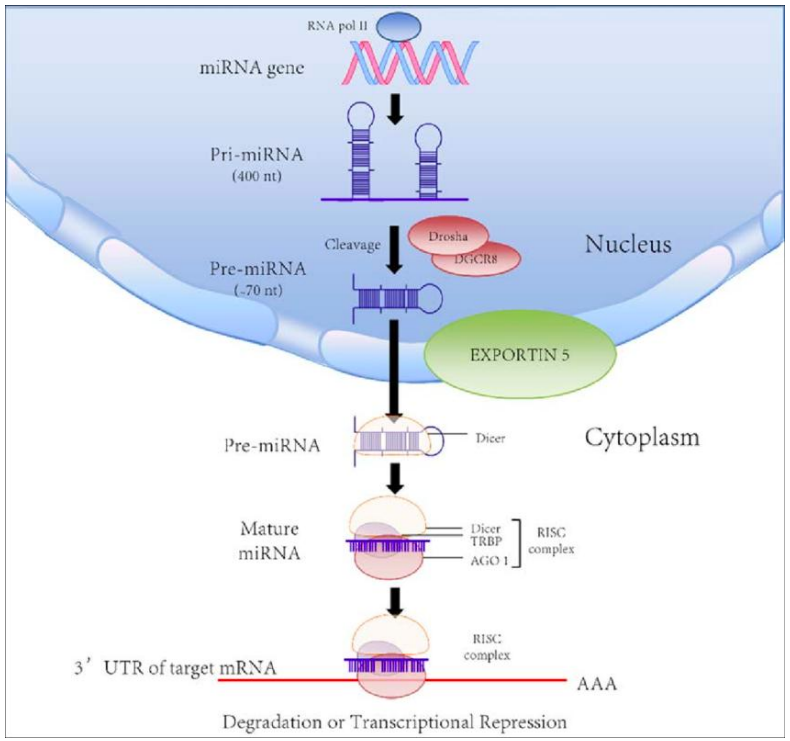
Splicing altenativo del RNA mensajero

- "Sintonía fina" de la expresión genética a nivel post-trascripcional
- Variabilidad de proteínas a partir de un número finito de genes
- Expressión diferencial de proteínas



Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación de la expresión genética mediada por microRNAs (miARNs)



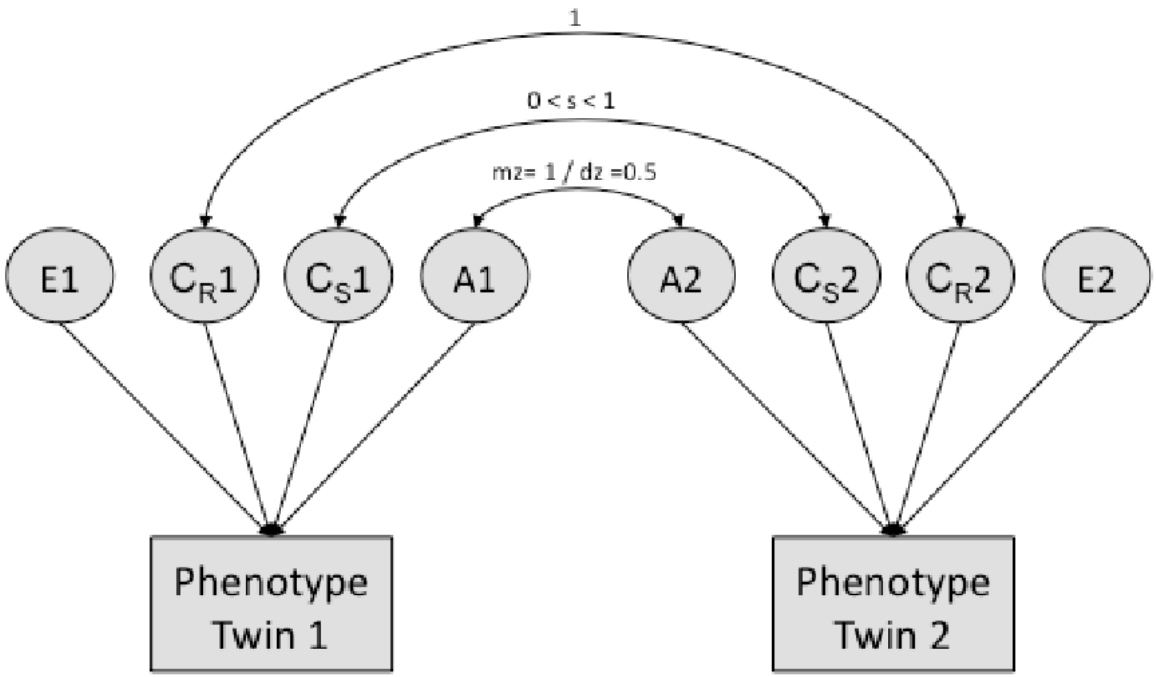
- Secuencias cortas de mRNA (18-26 nt)
- Interacción en region 3'UTR
- Esenciales durante embriogénesis

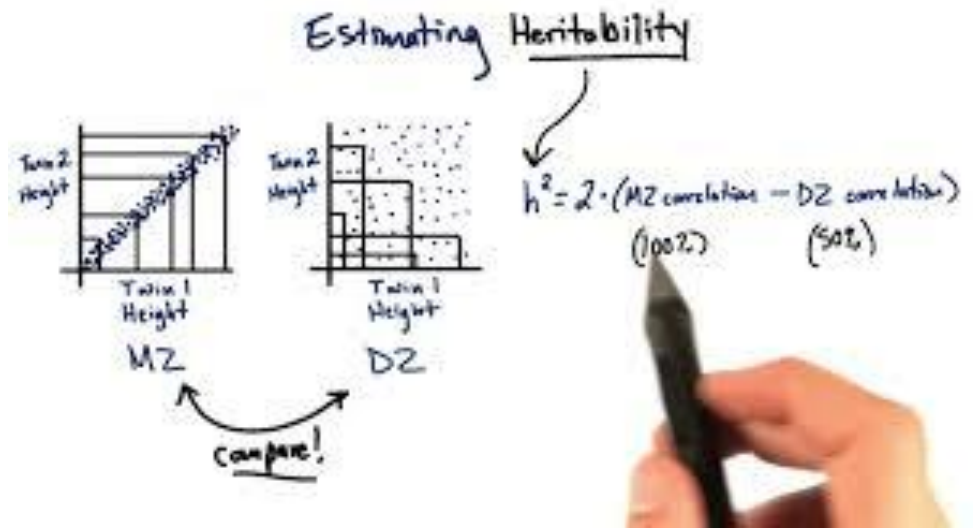
MicroARNs en la enfermedad

Metastatic miRNAs	The epithelial to mesenchymal transition (EMT) cause metastasis after initation and progression in tumor cells and this miRNAs play roll in this process	Some of them: miR-7,50-52 miR-17/20,53,54 miR-22,55-57 miR-30,58,59
Oncogenic miRNAs	Called “oncomirs” promote tumor development by negatively inhibiting tumor suppressor genes and/or genes that control cell differentiation or apoptosis	Some of them:miR10-b, miR-17, miR-21,miR-155
Tumor suppressor miRNAs	Target of them is oncogenes in cell differentiation, invasion, apoptosis, proliferation, metastasis	Some of them: Let 7 family, miR 34b, miR145, miR205, miR200

REGULACION EPIGENETICA

Héritas
Medicina de precisión
CIBIC + INDEAR


1
0 < s < 1
mz= 1 / dz =0.5
E1
C_R1
C_S1
A1
A2
C_S2
C_R2
E2
Phenotype Twin 1
Phenotype Twin 2


Estimating Heritability
Twin 2 Height
Twin 1 Height
MZ
DZ
compare!
$h^2 = 2 \cdot$ (MZ correlation − DZ correlation)
(100%)
(50%)

Cambios heredables de la expresión génica que ocurren *sin* que se presenten modificaciones en la secuencia de ADN

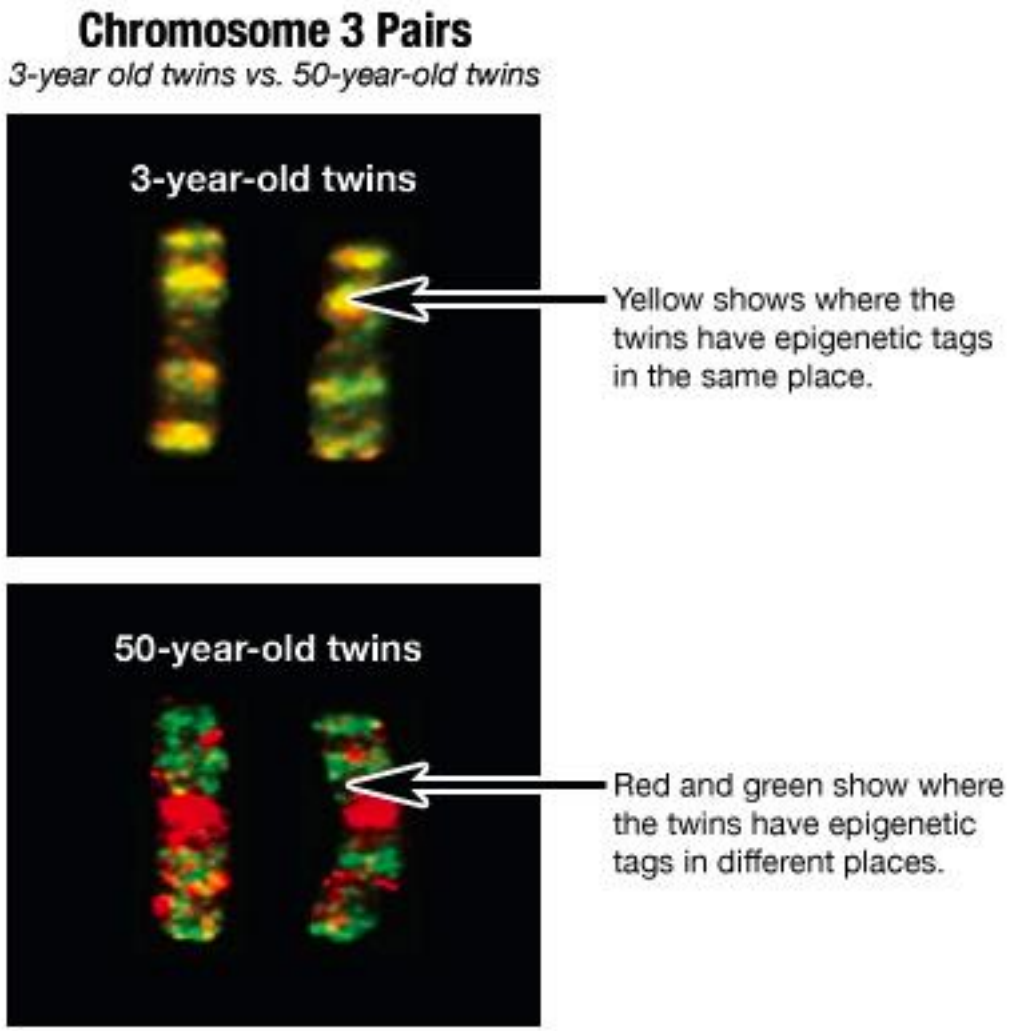
Principales mecanismos epigenéticos

- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por microARNs

Héritas
Medicina de precisión
CIBIC + INDEAR

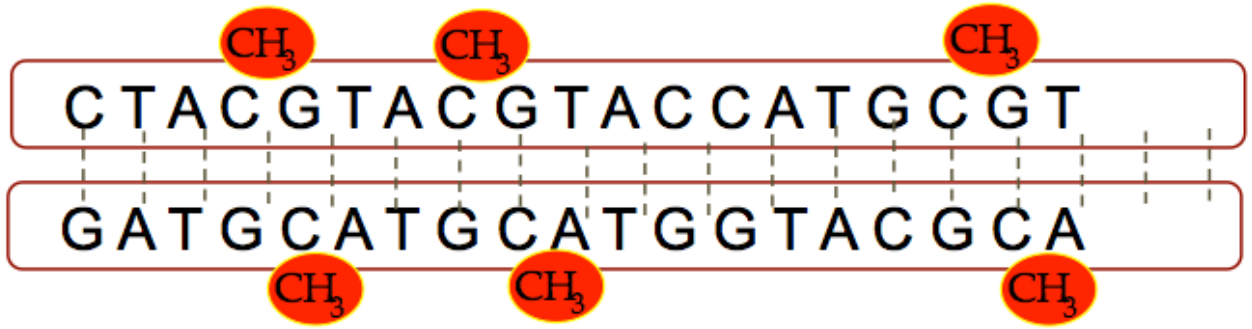
www.ncbi.nlm.nih.gov

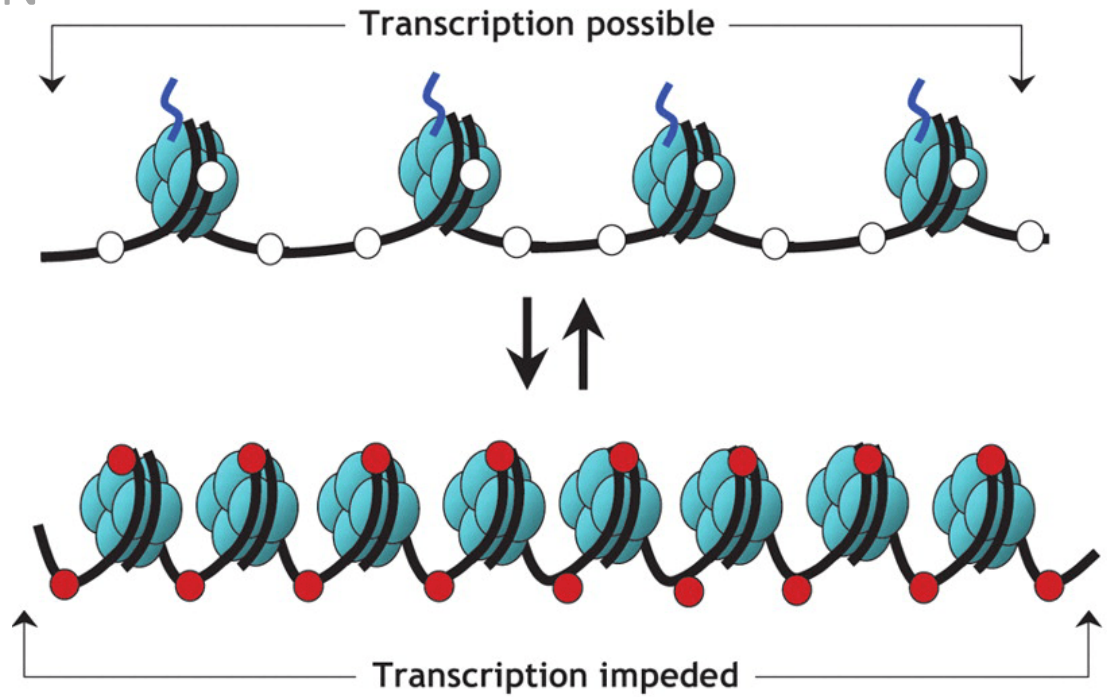
Epigenetics of discordant monozygotic twins: implications for disease



Héritas
Medicina de precisión
CIBIC + INDEAR

- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por microARN

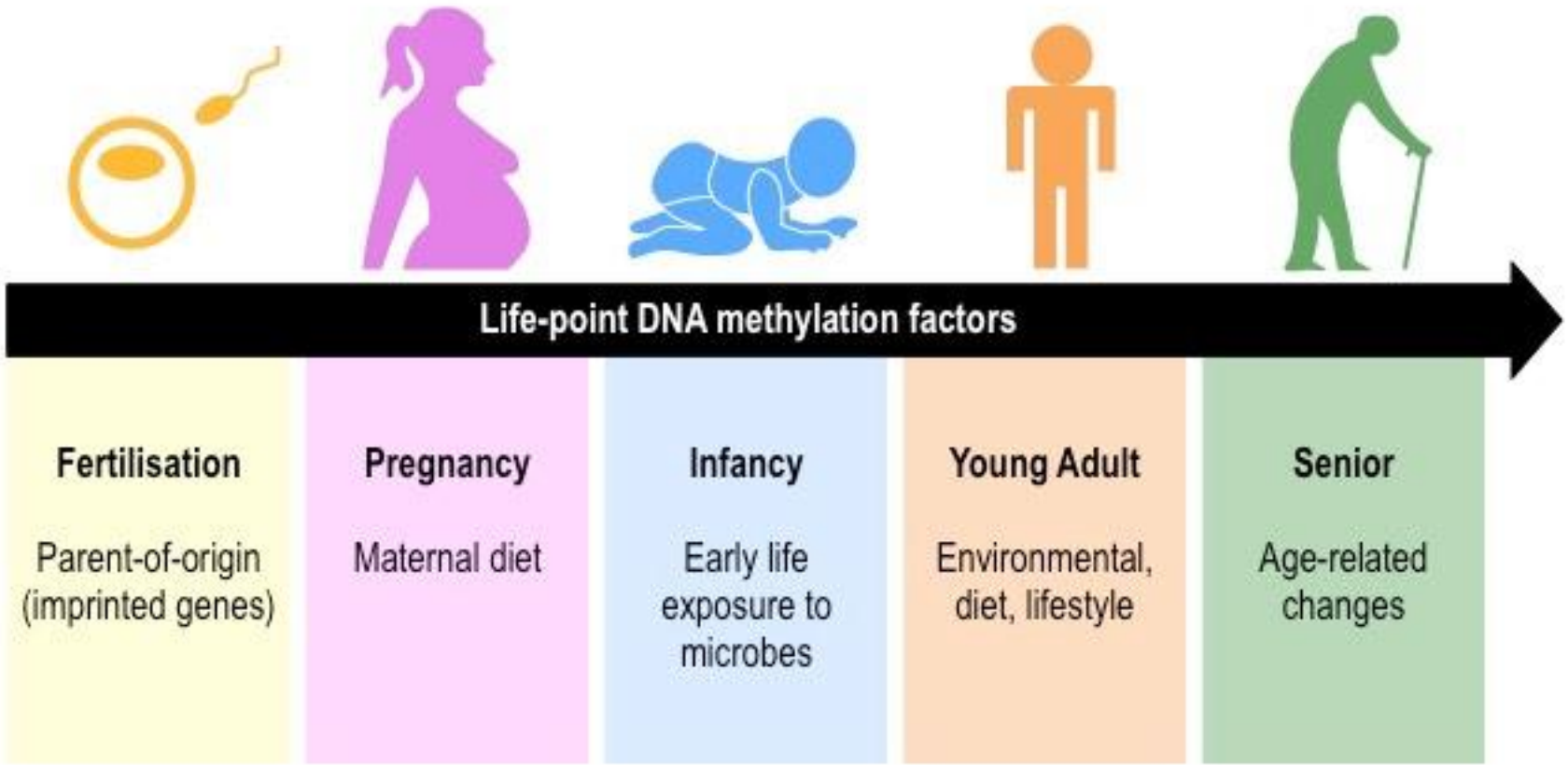




Ocurre generalmente en Citosinas, especialmente en nucleótidos emparejados CpG (Dímeros metilados CpG)

Químicamente muy estable
Mecanismo de silenciamiento de genes, impronta genómica, inactivación cromosoma X

Héritas
Medicina de precisión
CIBIC + INDEAR



Héritas
Medicina de precisión
CIBIC + INDEAR

Metilación aberrante

Cáncer: hipometilación (inestabilidad del ADN, activación de oncogenes) e hipermetilación (mutación de genes y silenciamiento de genes supresores de tumores)

Mutaciones

Síndromes de X frágil, Bockwith Wiedemann, Prader Willi/Abgelman, ICF, ATRX, Rull

Metilación fisiológica

Inactivación del cromosoma X

Silenciamiento de transposones

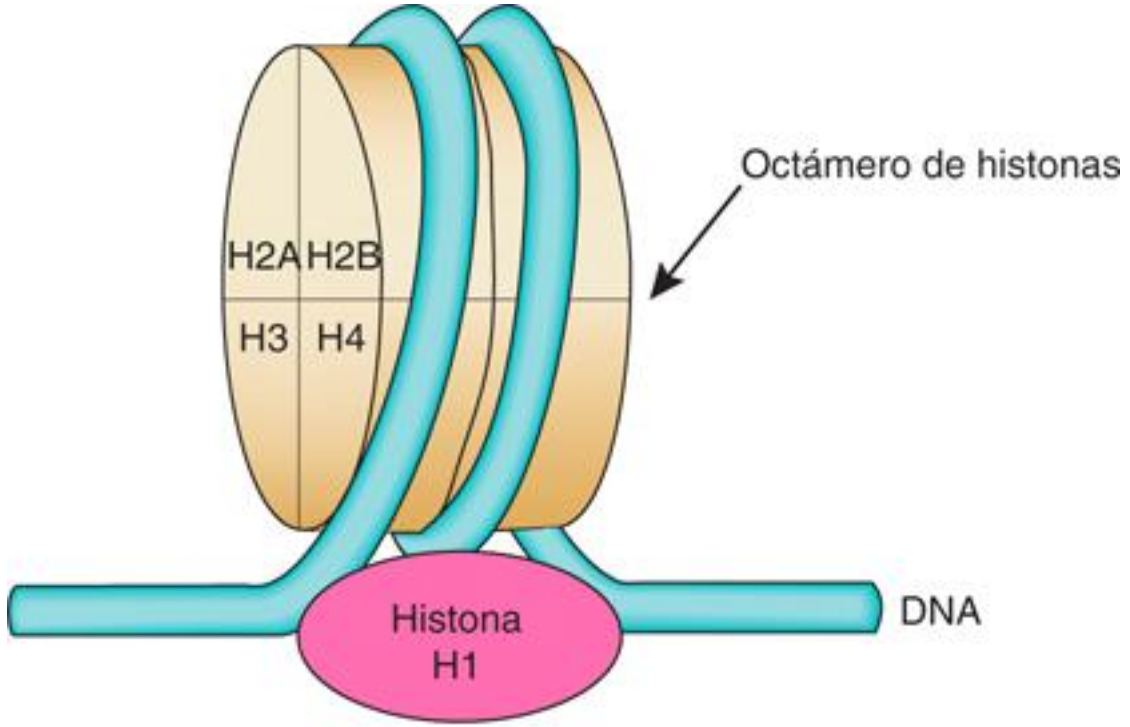
Mantenimiento de la estabilidad cromosómica

Modulación de la estructura de la cromatina

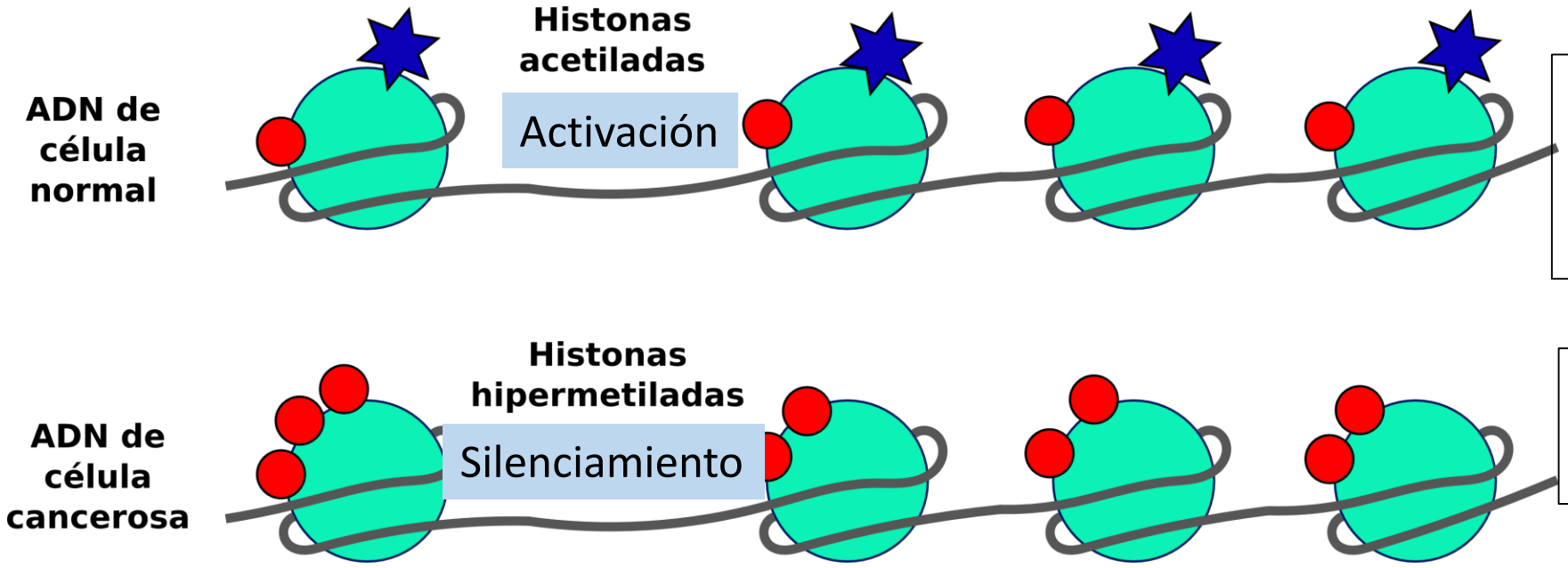
Regulación transcripcional

- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por microARN

Histonas: unidades fundamentales de la cromatina
CORE: histonas del cuerpo central del *Nucleosoma* (H2A, H2b, H3, H4)
LINKER: Histonas de enlace (H1)

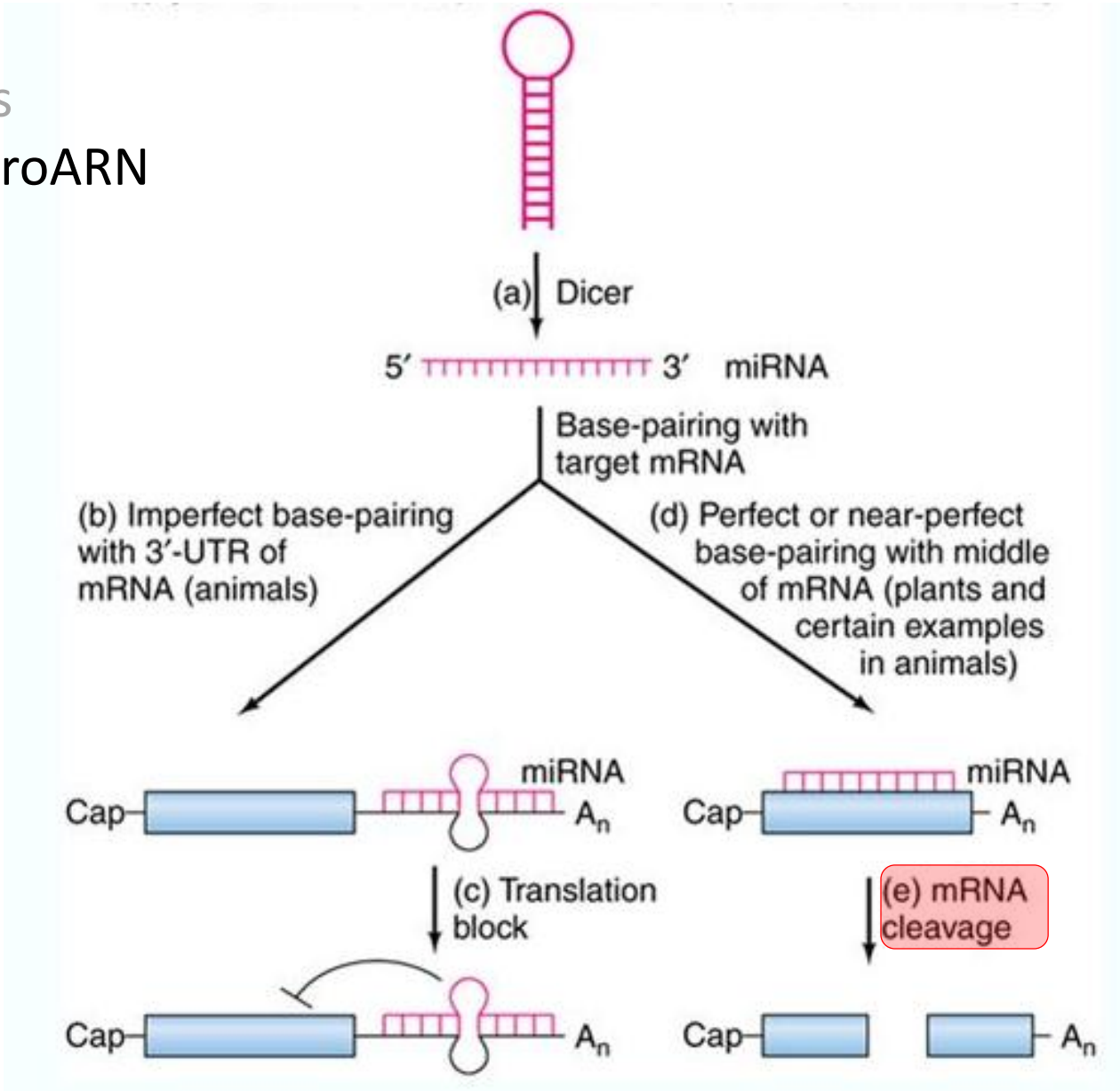


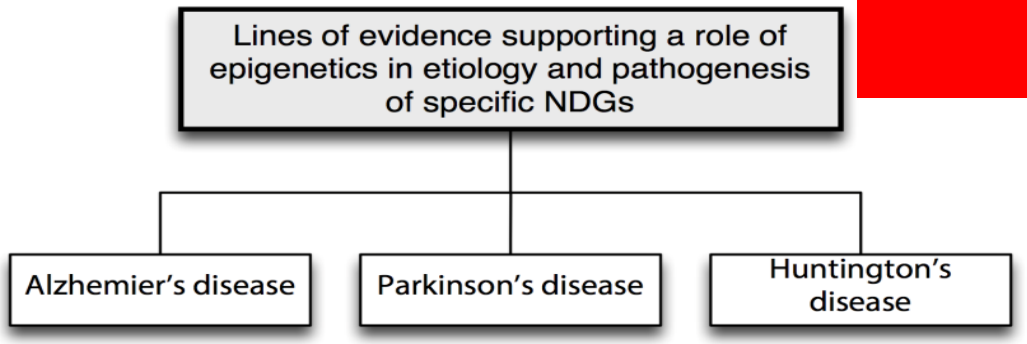
- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por microARN



Héritas
Medicina de precisión
CIBIC + INDEAR

- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por microARN





	Alzhemier's disease	Parkinson's disease	Huntington's disease
Methylation	Reduced DNA methylation in the anterior temporal neocortex neuronal nuclei Hypermethyation of HTERT gene Hypomethylation of inflammatory genes iNOS, IL-1, and TNF-α in the AD cortex	Overall reduction of methylation potential Hypomethylation of SNCA gene in brain tissue α-synuclein related reduction of Dnmt1 methyltransferase availability Differential methylation of ARK16, GPNMB, STX1B and CYP2E1	Early reports of increased variability at HTT gene locus
Histone modifications	increased phosphorylated histone H3 in hippocampal neurons Modulation of histone acetylation by HDAC inhibitors improved learning and memory in mouse models	Response to treatment with HDACIs in disease models α-synuclein related reduction in histone acetylation and histone gene expression	Beneficial effect of HDACIs in disease models Sequestration of proteins with HDAC activity (CBP) Increase of histone proteins carrying H3K9 marks in brain and blood tissues
micro RNA regulation	Deregulation of several miRNAs in brain	Differential expression of dapaminergic neuron specific miRNA miR-133b Differential Expression of miR-7, -10a, -10b, -34b/c -212, -132, -495 miRNAs in brain tissues	Down-regulation of nine miRNAs in animal models of HD (AC128 and R6/2 mice) High 3'terminal sequence variability of miRNAs in HD miR-34b unregulated in plasma of pre manifest HD patients

Héritas
Medicina de precisión
CIBIC + INDEAR

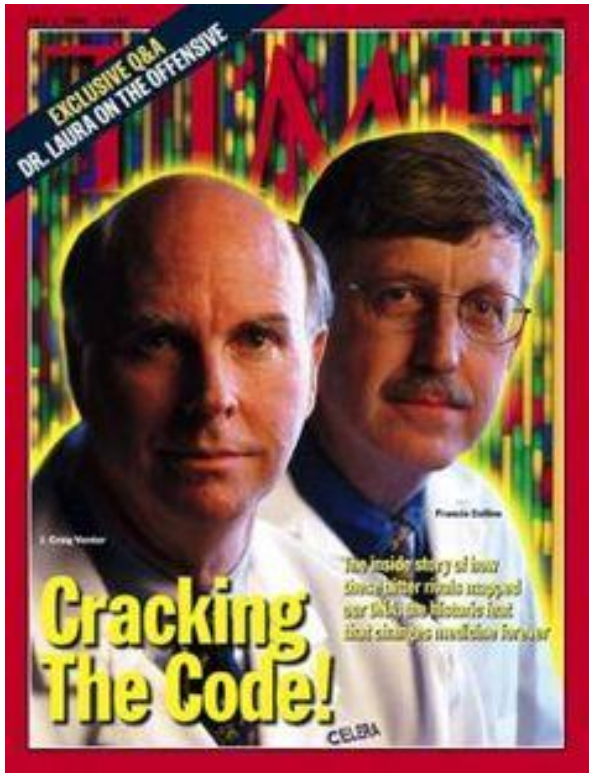
Aplicaciones de la *Biología Molecular* en Genética Médica

- Genética de Poblaciones
- Medicina Forense
- Enfermedades Hereditarias Mendelianas
- Infectología
- Farmacogenética
- Cáncer hereditario y somático
- Fertilidad y Reproducción

Héritas
Medicina de precisión
CIBIC + INDEAR

ENFERMEDADES HEREDITARIAS
COMO ESTUDIARLAS?

Héritas
Medicina de precisión
CIBIC + INDEAR

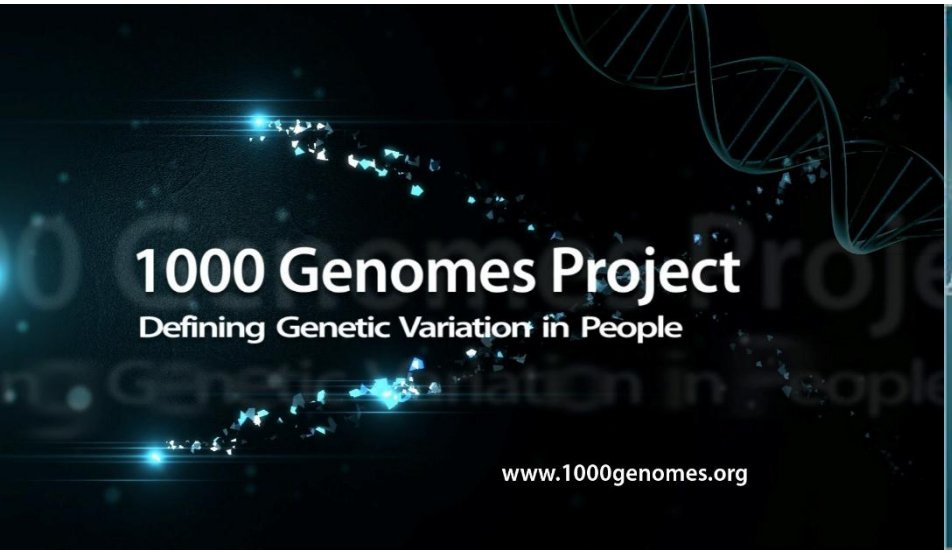


1990 – 2003

Proyecto Genoma Humano

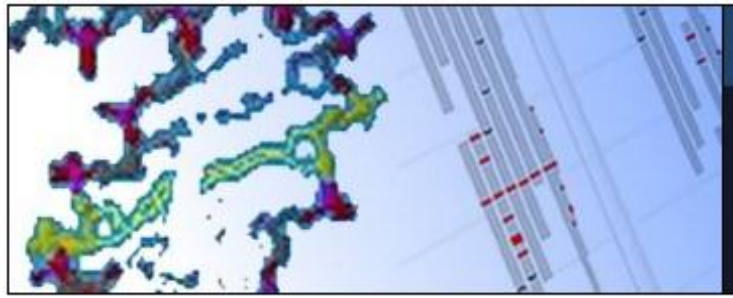
U$D 3,000,000,000

Héritas
Medicina de precisión
CIBIC + INDEAR



BASES DE DATOS

Héritas
Medicina de precisión
CIBIC + INDEAR



dbSNP

Database of single nucleotide polymorphisms (SNPs) and multiple small-scale variations that include insertions/deletions, microsatellites, and non-polymorphic variants.

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

OMIM

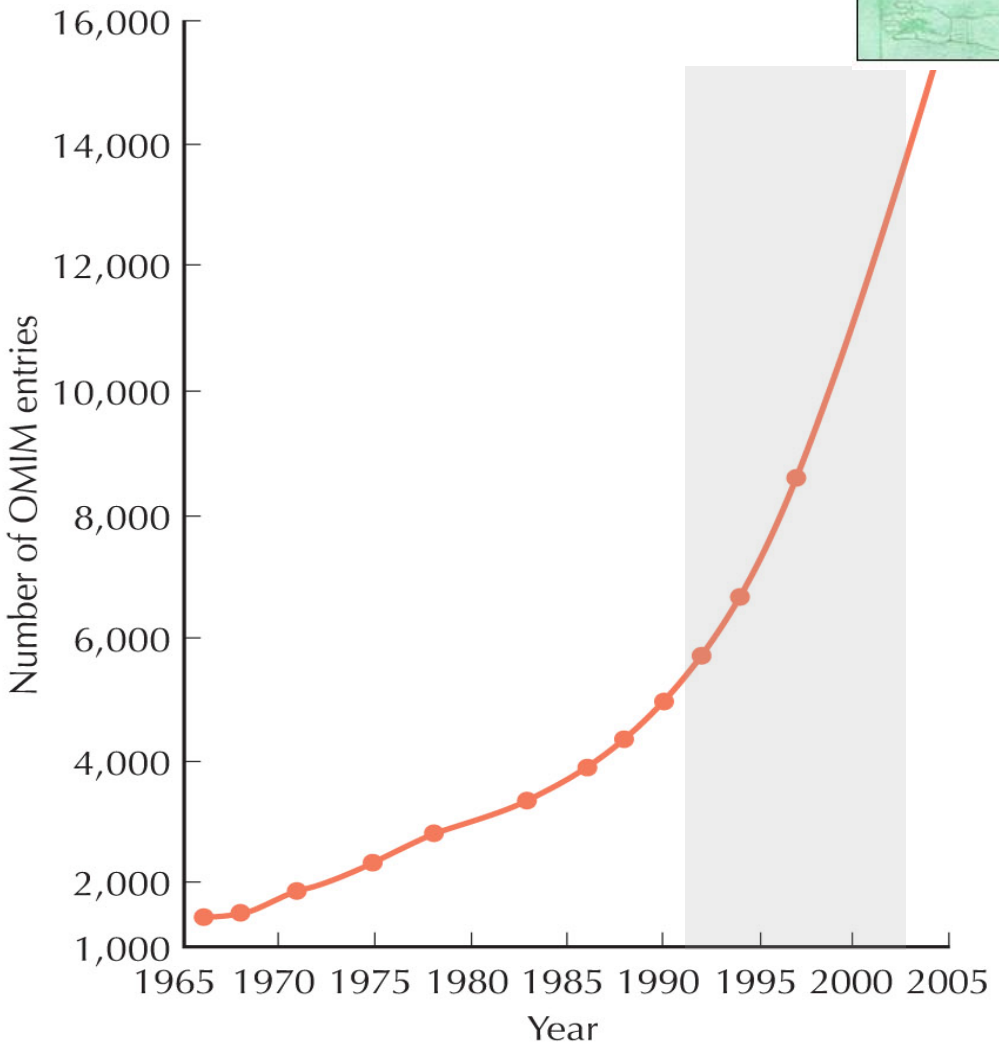
OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh. Its official home is omim.org.

Héritas
Medicina de precisión
CIBIC + INDEAR

Enfermedades hereditarias Mendelianas
Bases de Datos

OMIM

OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh. Its official home is omim.org.



OMIM Entry Statistics

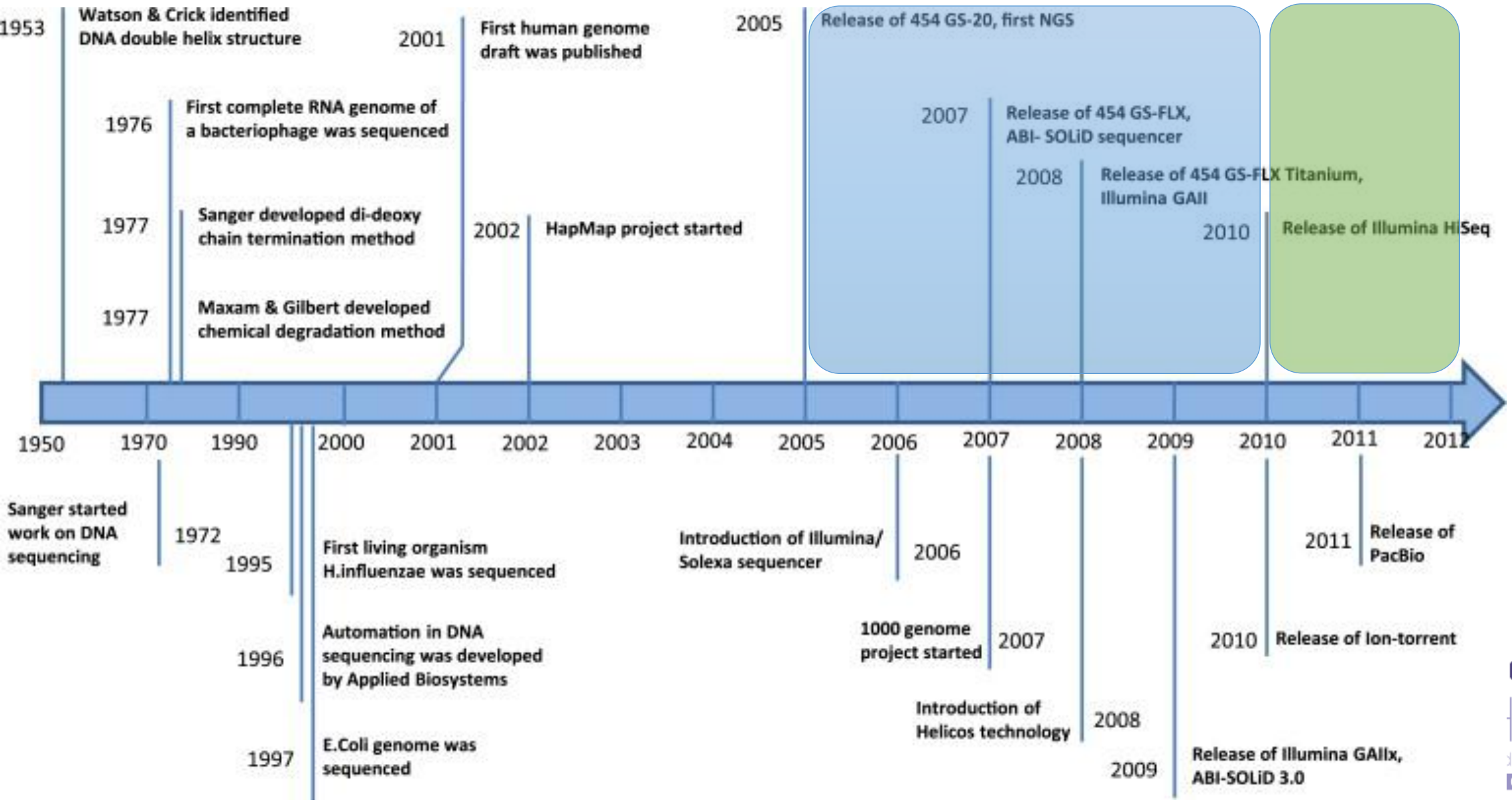
Number of Entries in OMIM (Updated April 15th, 2016) :

Prefix	Autosomal	X Linked	Y Linked	Mitochondrial	Totals
* Gene description	14,450	705	48	35	15,238
+ Gene and phenotype, combined	81	2	0	2	85
# Phenotype description, molecular basis known	4,369	303	4	29	4,705
% Phenotype description or locus, molecular basis unknown	1,495	126	5	0	1,626
Other, mainly phenotypes with suspected mendelian basis	1,693	112	2	0	1,807
Totals	22,088	1,248	59	66	23,461

Héritas
Medicina de precisión
CIBIC + INDEAR

Next Generation Sequencing (NGS)

Héritas
Medicina de precisión
CIBIC + INDEAR

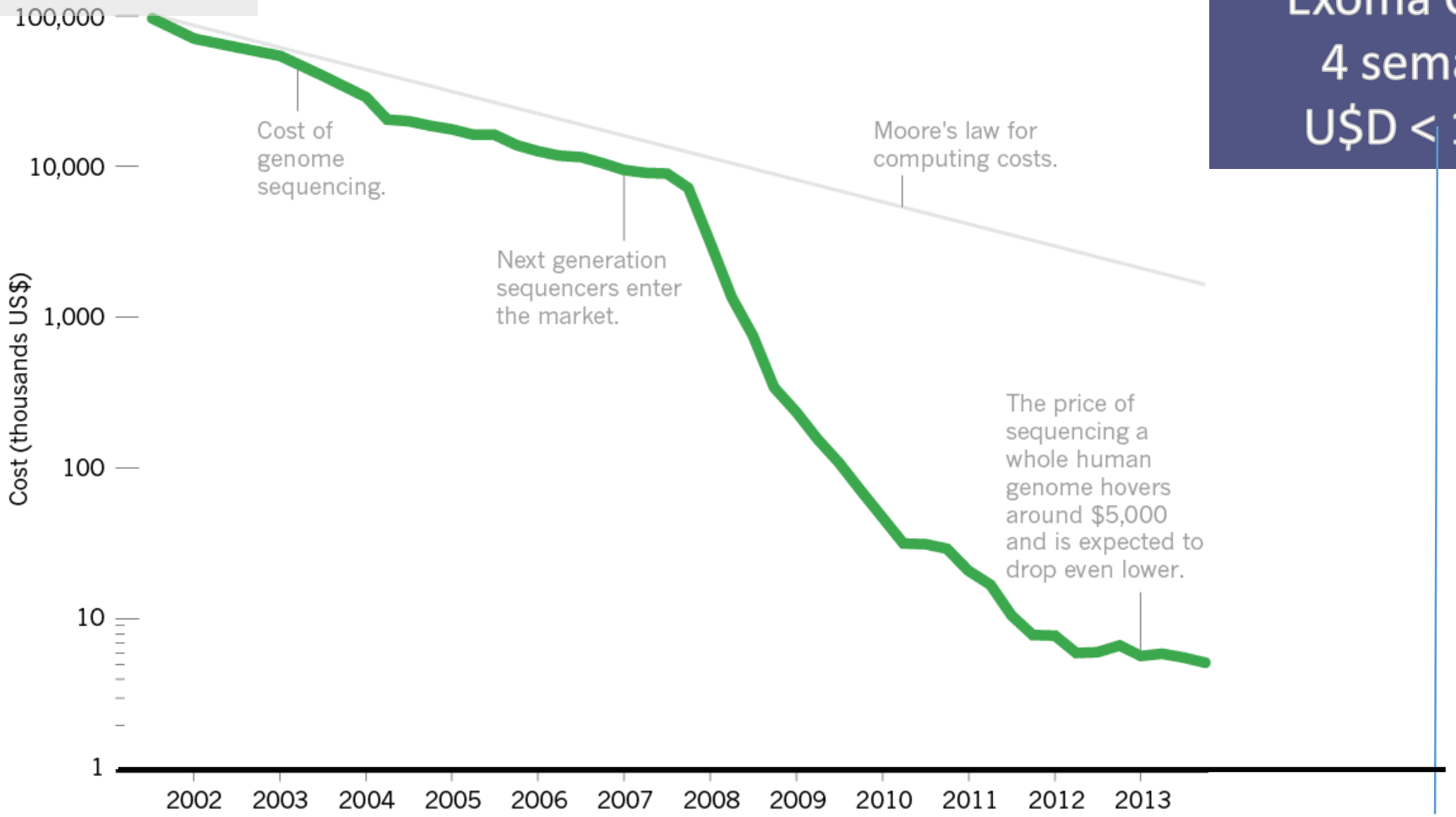


léritas
dicina de precisión
CIBIC + INDEAR

1990
Proyecto Genoma Humano
13 años
U$D 3,000,000,000

Enfermedades hereditarias Mendelianas
NGS

2018
Exoma Clínico
4 semanas
U$D < 1,000



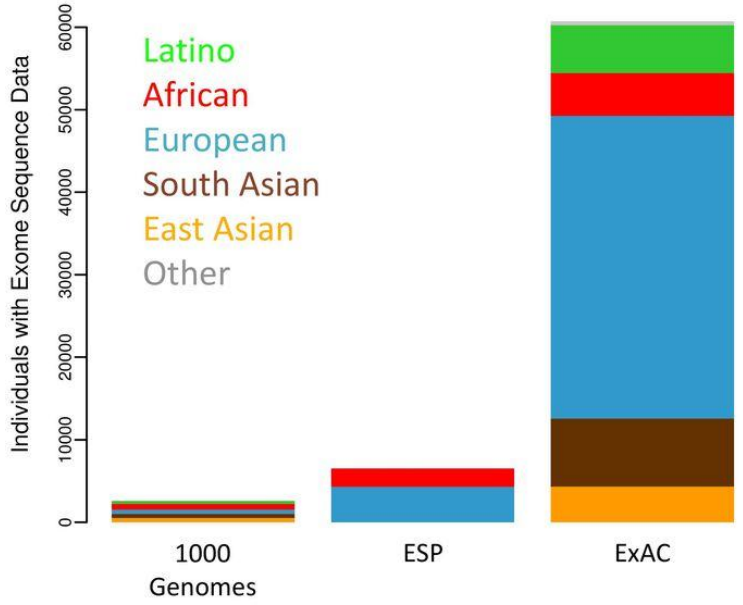
Héritas
Medicina de precisión
CIBIC + INDEAR

exac.broadinstitute.org

ExAC Browser

Daniel MacArthur

Population frequency Databases ExAc / GnomAD



exac.broadinstitute.org

Estandarización de diversos datasets de proyectos de secuenciación existentes

- ✓ Generación de grandes grupos control (individuos sanos)
- ✓ Herramienta indispensable para interpretación de variantes en el contexto de estudios NGS

Héritas
Medicina de precisión
CIBIC + INDEAR

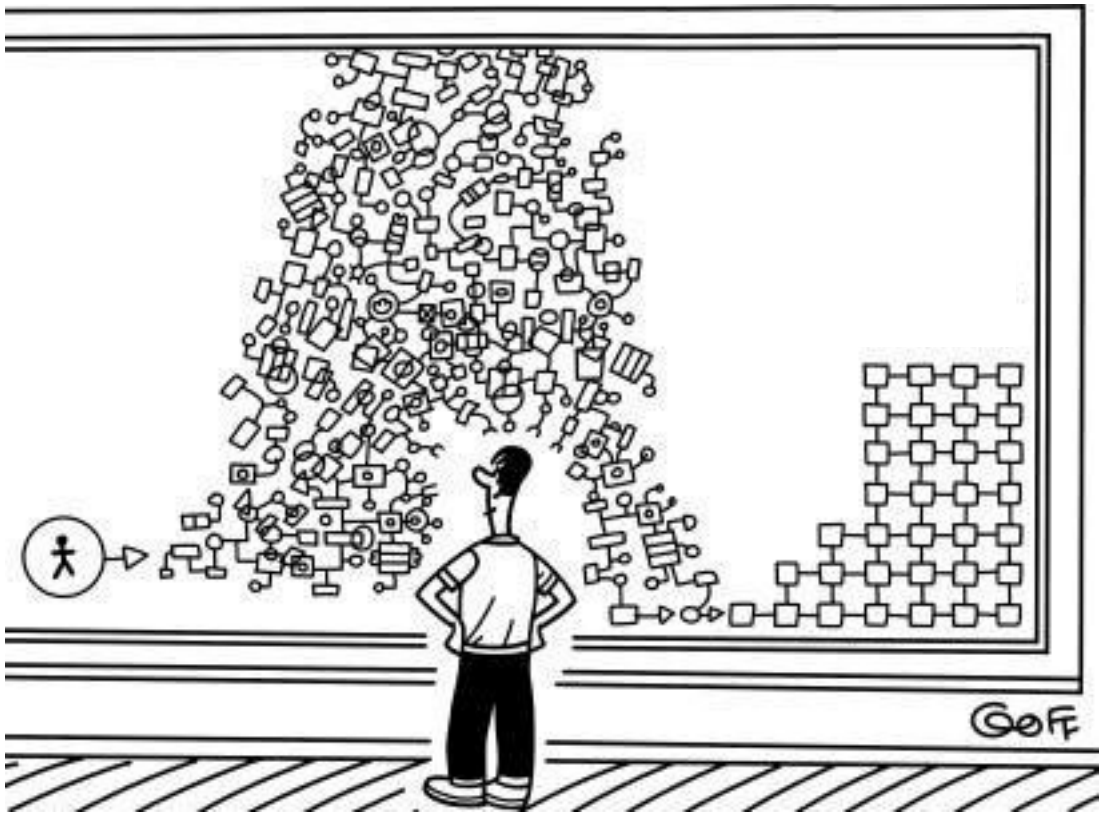
¿COMO UTILIZAR EFICIENTEMENTE LA INFORMACION GENERADA VIA NGS ?

Héritas
Medicina de precisión
CIBIC + INDEAR

Harvard Business Review

ANALYTICS

Data Scientist: The Sexiest Job of the 21st Century



Héritas
Medicina de precisión
CIBIC + INDEAR

Gracias !!!

Héritas
Medicina de precisión
CIBIC + INDEAR