Biología Molecular aplicada al diagnóstico clínico
Módulo I: Clase 1

Introducción a la biología molecular y su aplicación a la Medicina

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

15.04.2019

La **Genética** es la ciencia que estudia 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

La **Medicina Genómica** tiene como objetivo mejorar la calidad de la práctica médica orientando el cuidado pacientes hacia una medicina individualizada, predictiva y preventiva basándose en la informacion genómica de cada individuo.

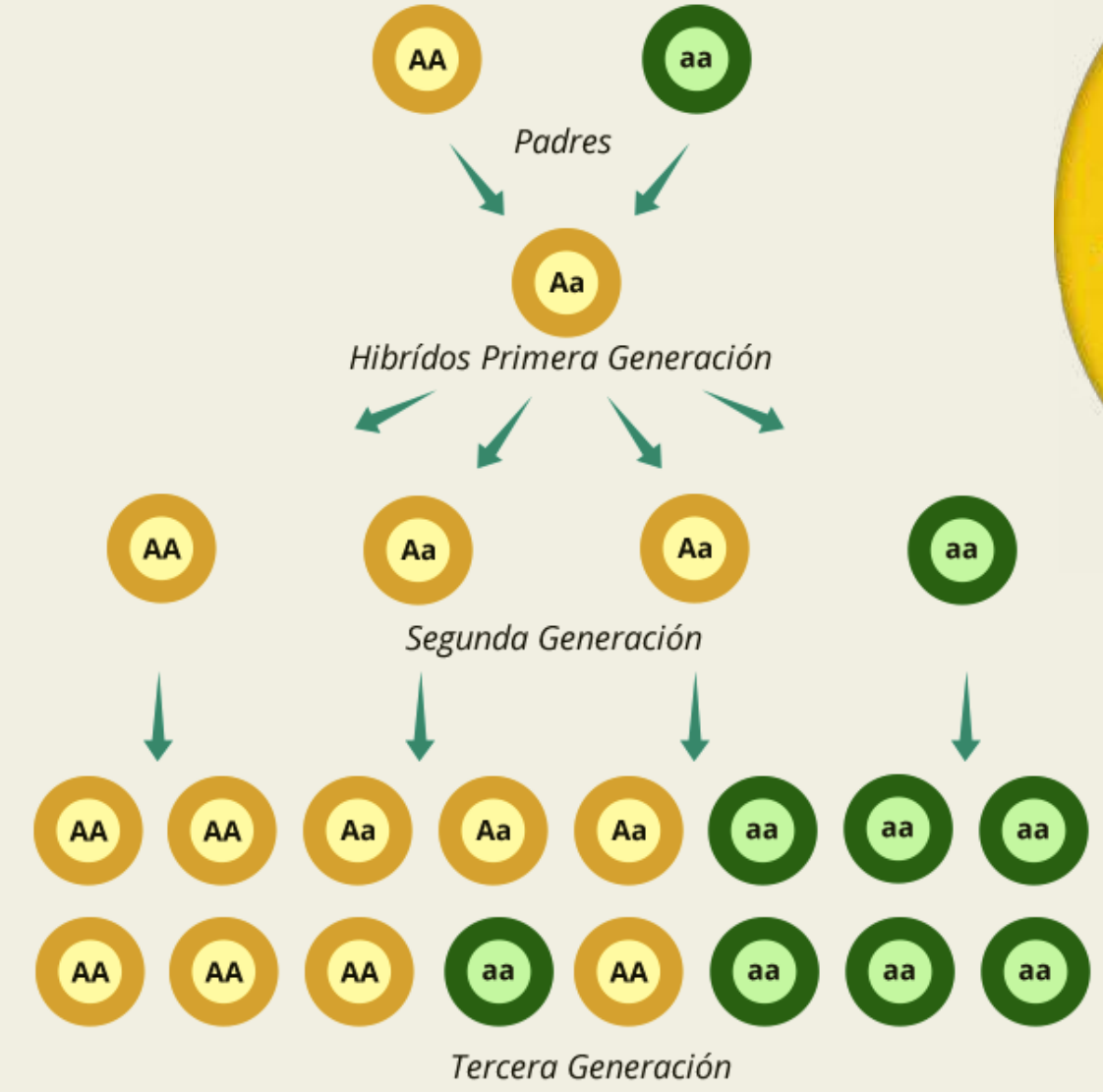
Héritas
Medicina de precisión
CIBIC + INDEAR

Las tres leyes de la genética de Mendel

Primera ley
Cuando se cruzan dos variedades de individuos de raza pura para un determinado caracter, todos los híbridos de la primer generación son iguales.

Segunda ley
Cuando se cruzan variedades de la primera generación entre sí, se obtienen semillas amarillas y verdes en la proporción 3:1
(75% amarillas y 25% verdes).

Tercera ley
Cuando se cruzan plantas que difieren en dos caracteres (dihíbridos), cuyo genotipo, por ejemplo es **AaAa**, se originan cuatro tipos distintos, que se combinaron de todas las formas posibles. En total se obtienen 16 genotipos posibles.



Gregor Mendel
(1822-1884)
Principios de la genetica
1866

Héritas
Medicina de precisión
CIBIC + INDEAR

1865
Mendel
documents patterns of heredity in pea plants

1902
Sutton and Boveri
propose chromosome theory of heredity

1927
Muller
shows that X-rays induce mutations

1930s
Hämmerling
shows that hereditary information is contained in the nuclei of eukaryotic cells

1931
McClintock
demonstrates genetic recombination in corn

1944
Avery, McLeod, and McCarty
show that DNA is the "transforming principle" responsible for heredity

1952
Hershey and Chase
use radioactive labeling to prove that DNA is responsible for heredity

1990s
Genome
sequencing projects begin

1869
Miescher
first identifies DNA ("nuclein")

1915
Morgan
and his "Fly Room" colleagues confirm the chromosome theory of heredity

1928
Griffith's
"transformation experiments" transform non-pathogenic bacteria strains to pathogenic

1941
Beadle and Tatum
describe the "one gene–one enzyme" hypothesis

1950
Chargaff
discovers that A = T and C = G (Chargaff's rules)

1953
Watson and Crick
propose the double helix

1961
Jacob and Monod
propose the

Francis Crick James Watson Maurice Wilkins Rosalind Franklin

Dogma central de la Biología
(Pre-Genomic Era)

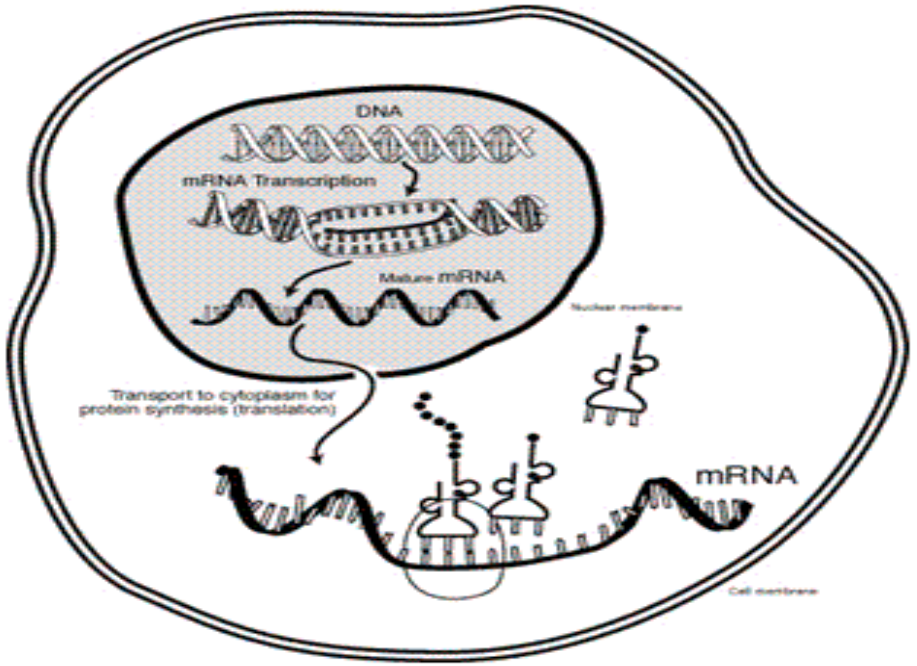
Francis Crick

James Watson

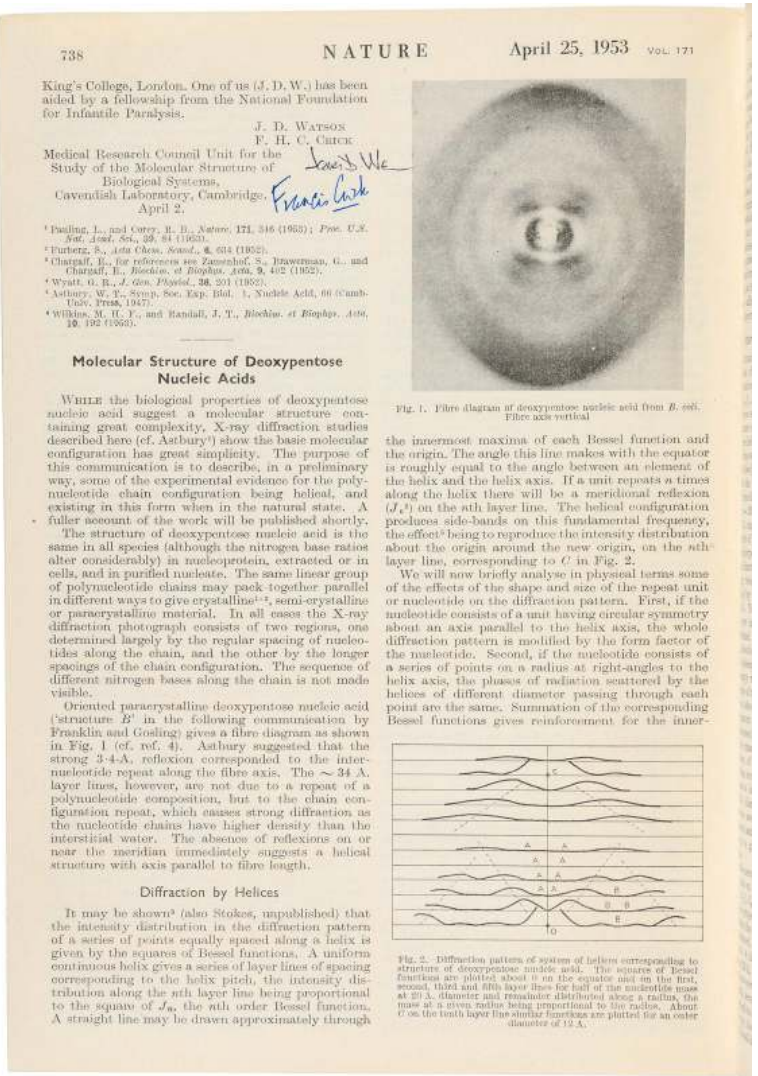
Maurice Wilkins

Rosalind Franklin

Fisiologia – Medicina
Nobel Prize 1962



DNA
↓
transcription
↓
RNA
↓
translation
↓
Protein



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.

Molecular Structure of Deoxypentose Nucleic Acids

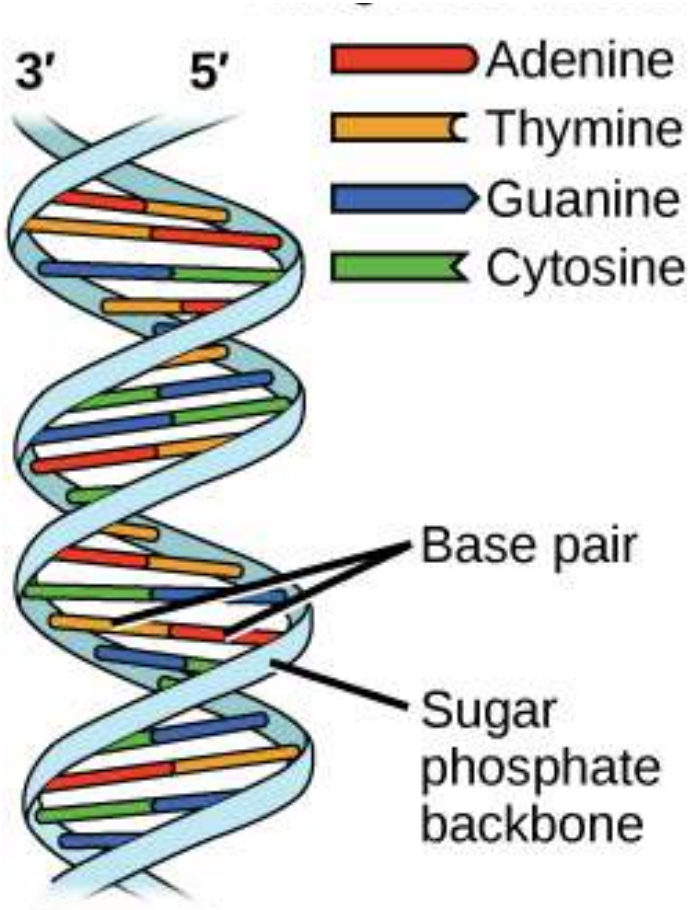
Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

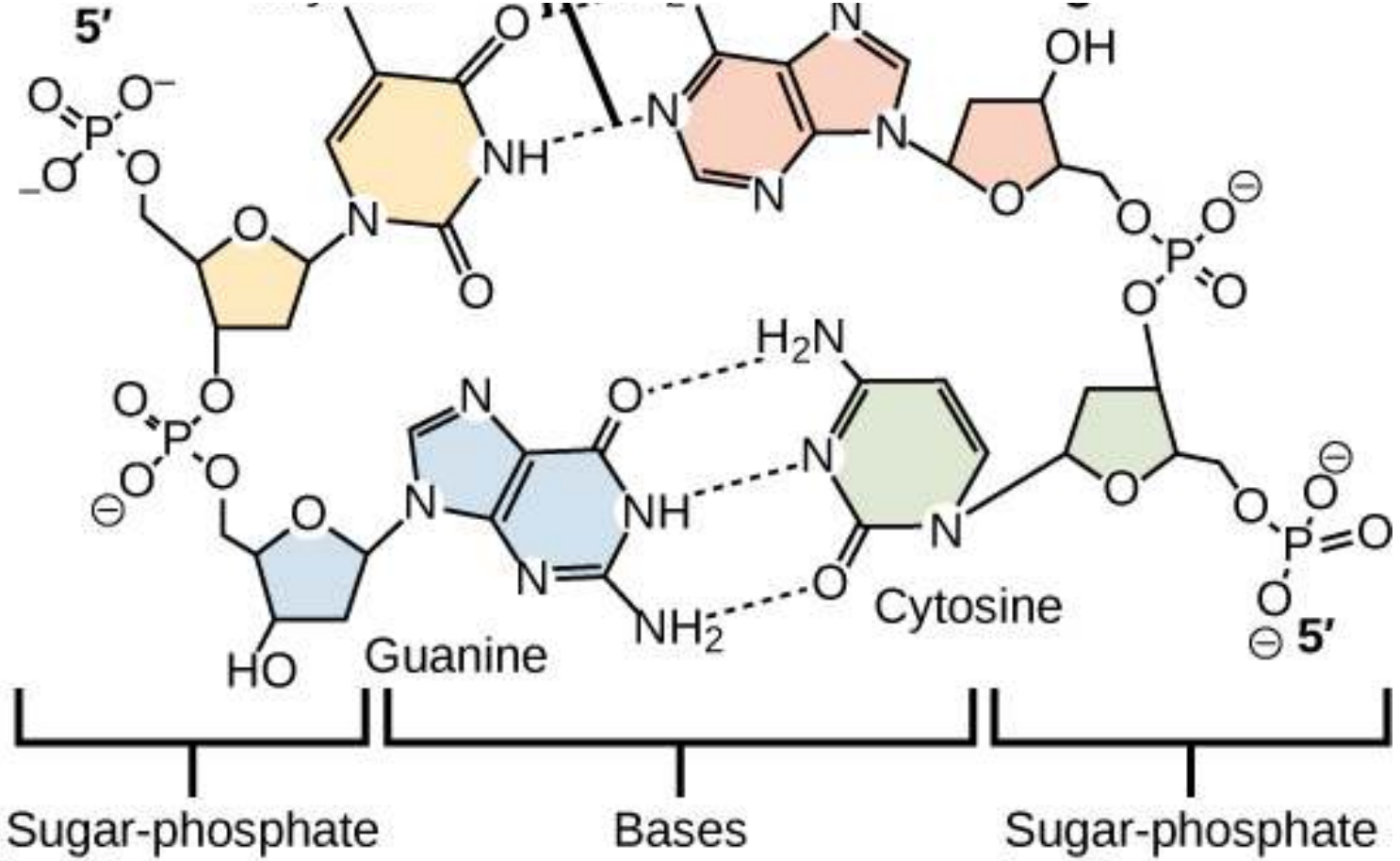
Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid.

Edwin Schrödinger
(1887-1961)

1933 Noble Prize (Fisica) : aportes a la mecanica cuántica moderna
1943: Lectures series – ***What is life***?

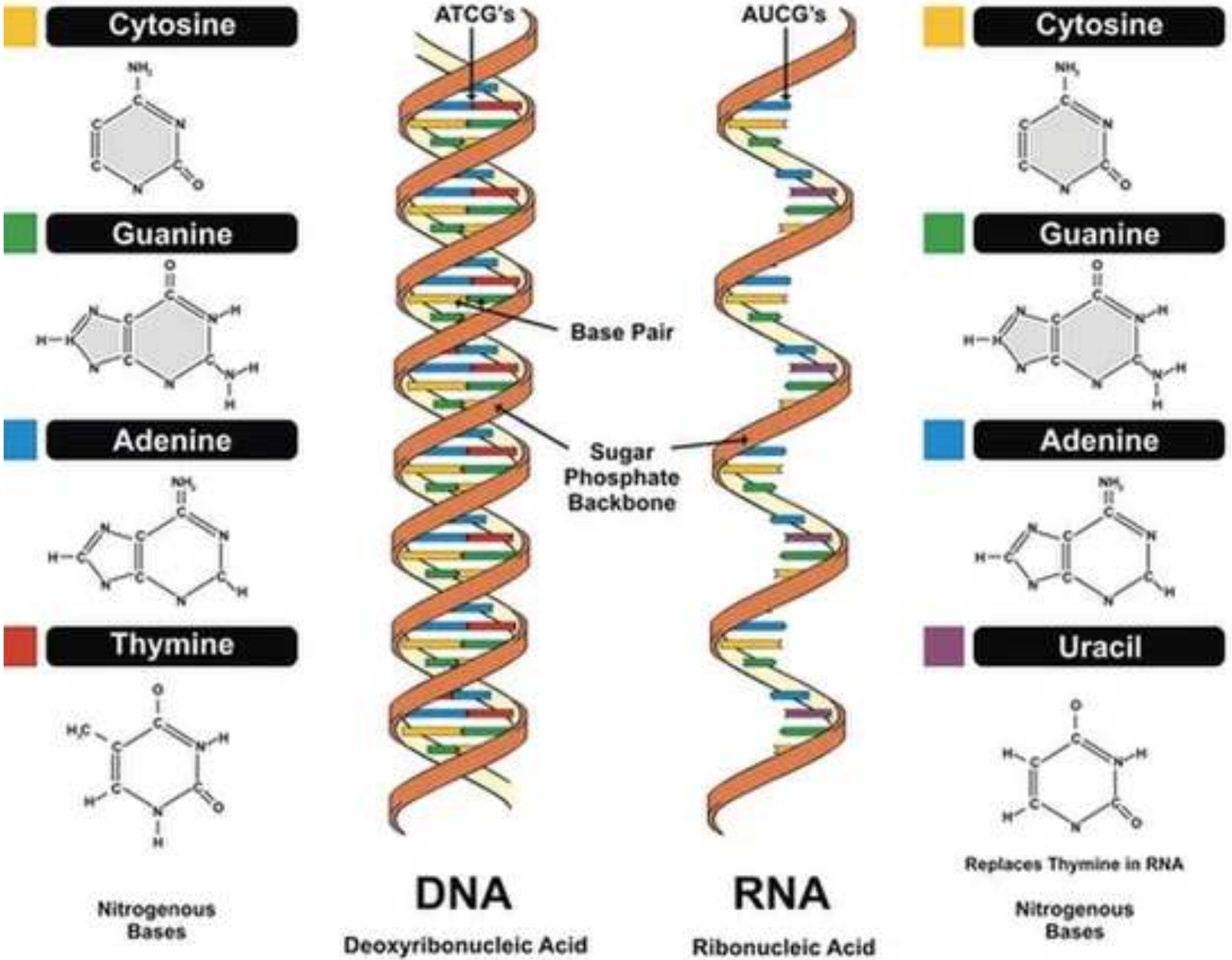
" Because so much information had to be packed into every cell, it must be packed in a ***hereditary code-script*** *embedded in the molecular fabric of chromosomes.*
To understand life, then, we would have to identify these molecules, and crack their code"





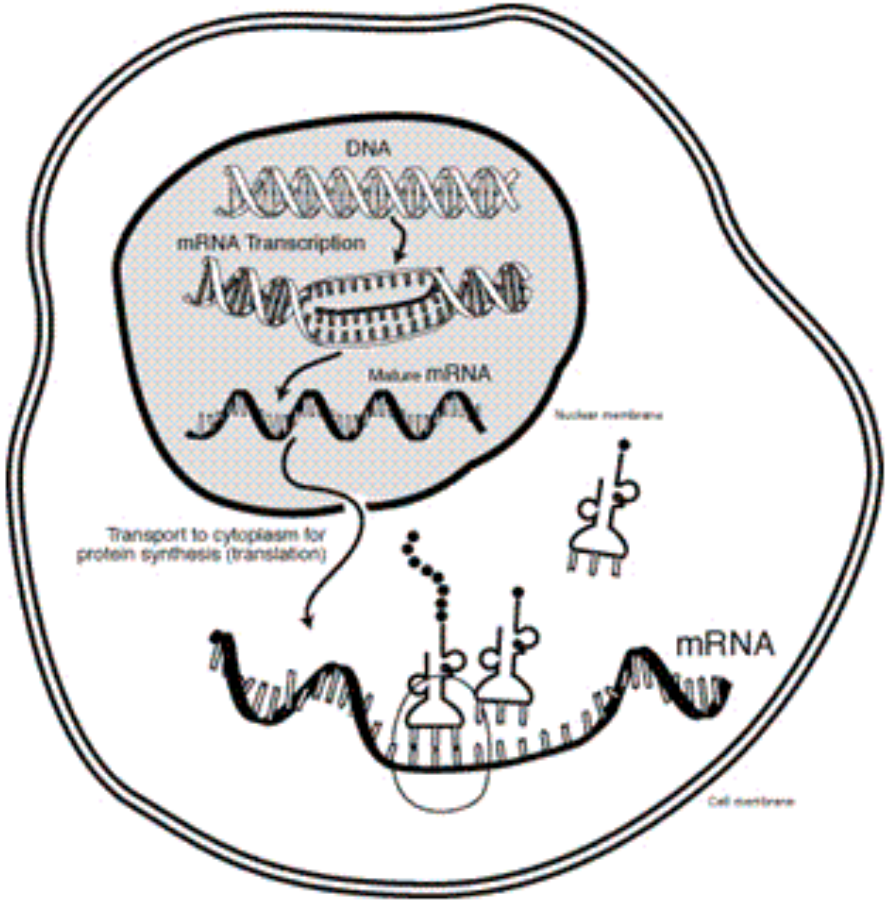
Héritas
Medicina de precisión
CIBIC + INDEAR

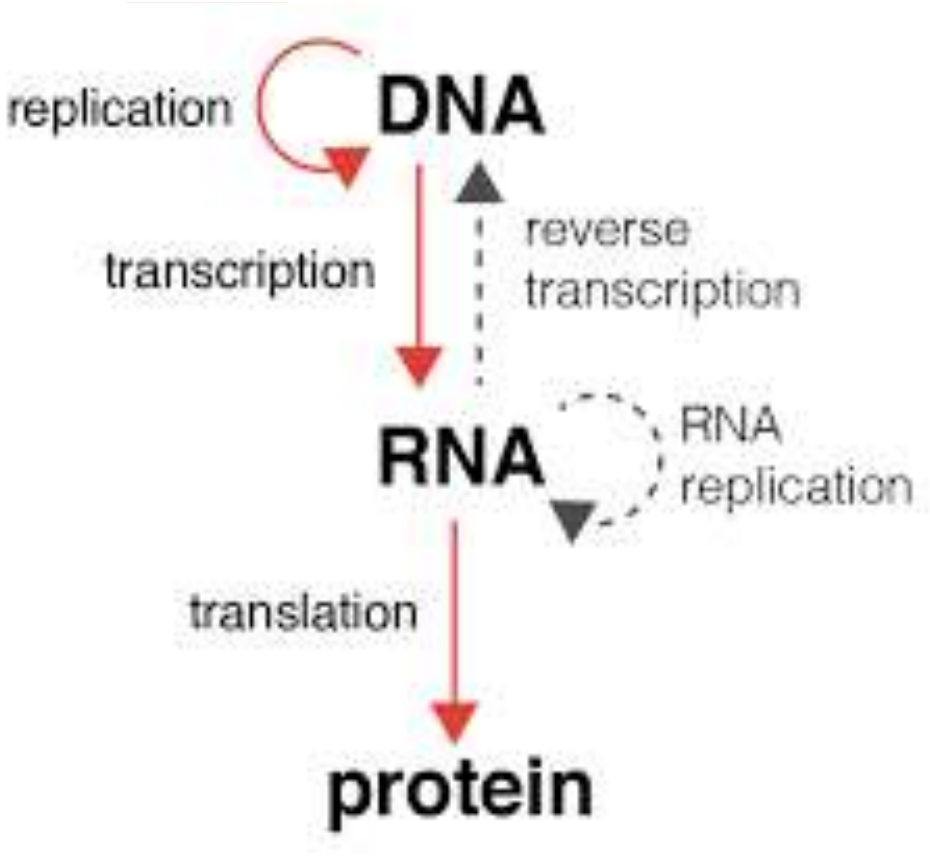
Estructura del DNA /RNA



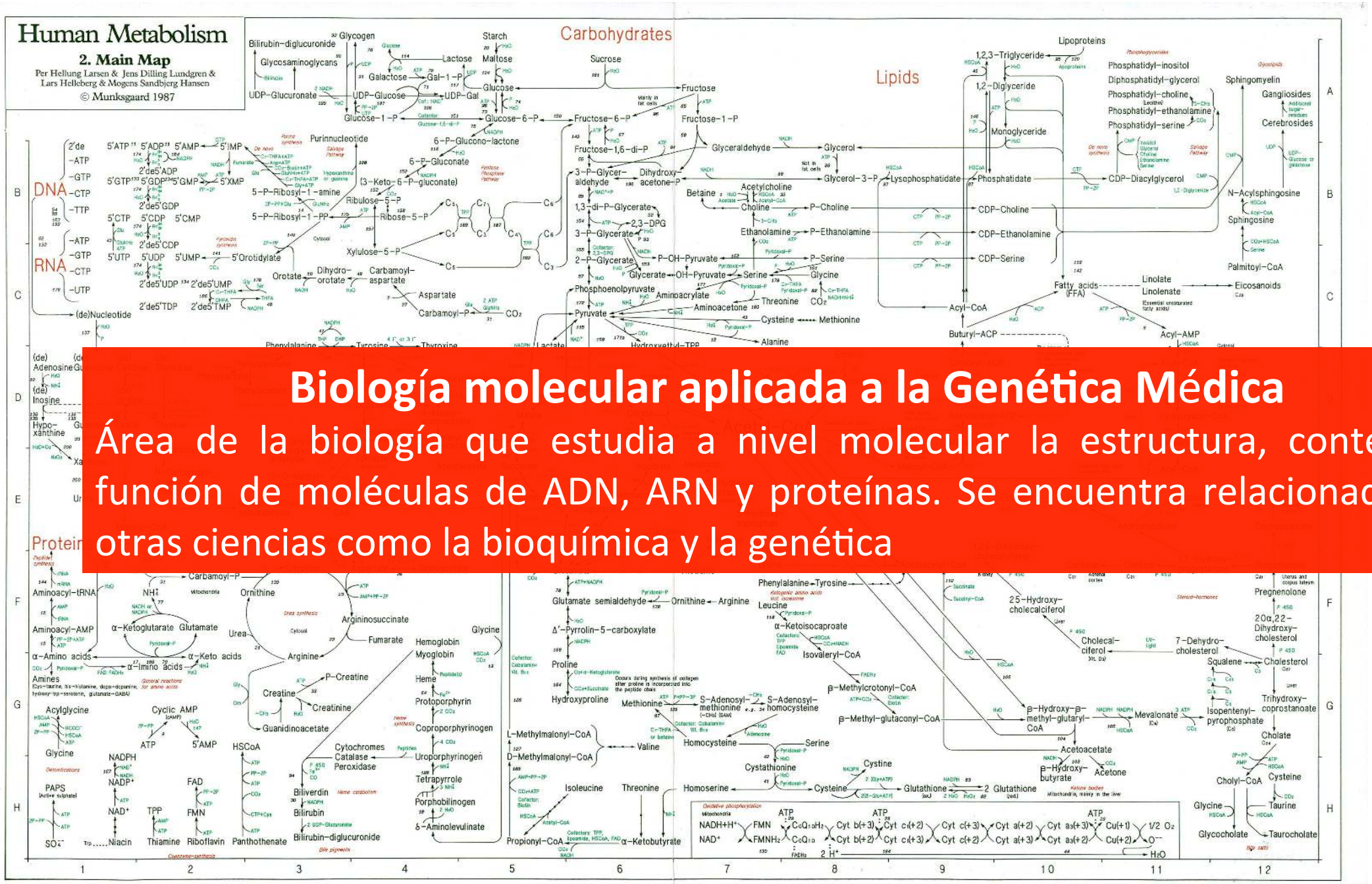
Héritas
Medicina de precisión
CIBIC + INDEAR

Dogma central de la biología *molecular*



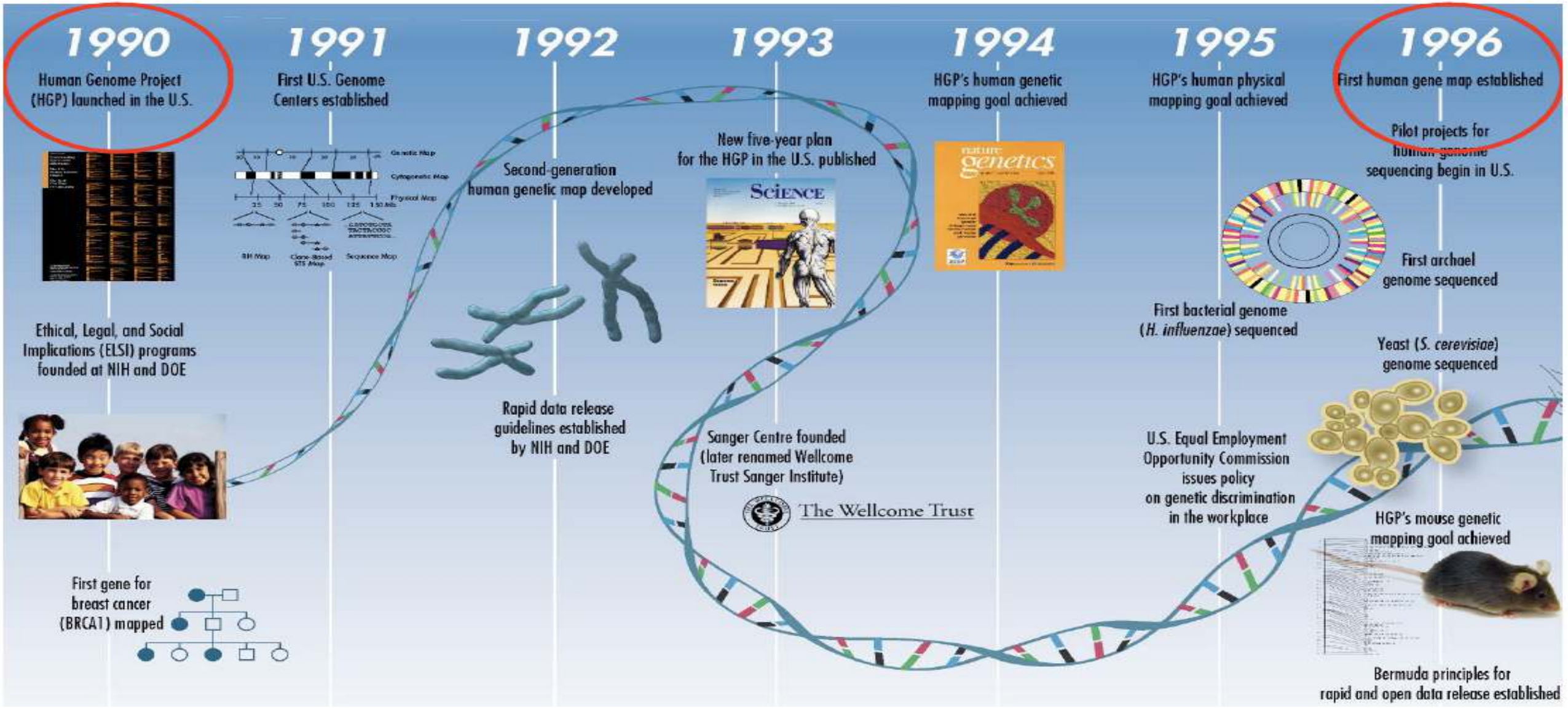


Héritas
Medicina de precisión
CIBIC + INDEAR

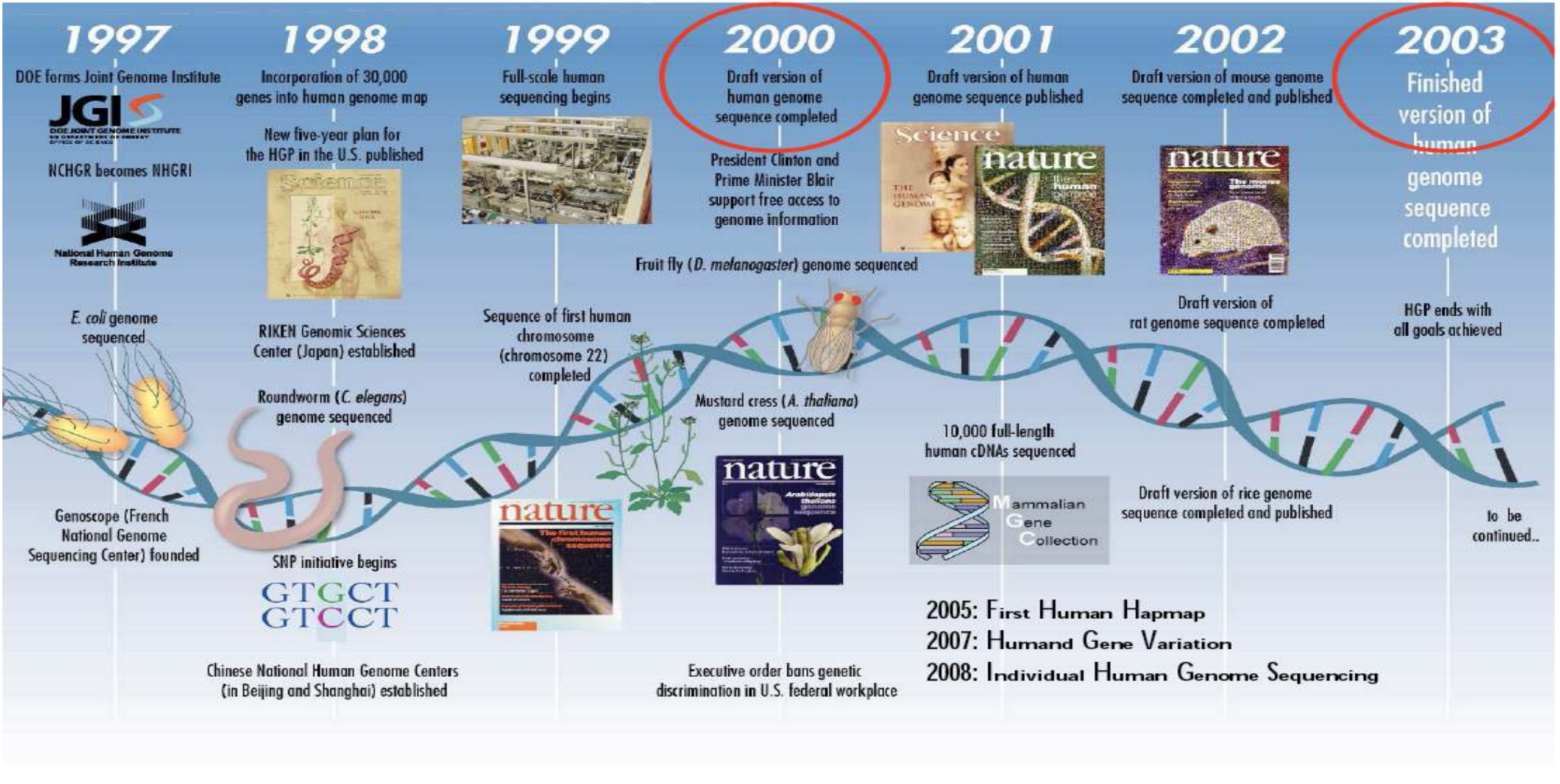


Héritas
Medicina de precisión
CIBIC + INDEAR

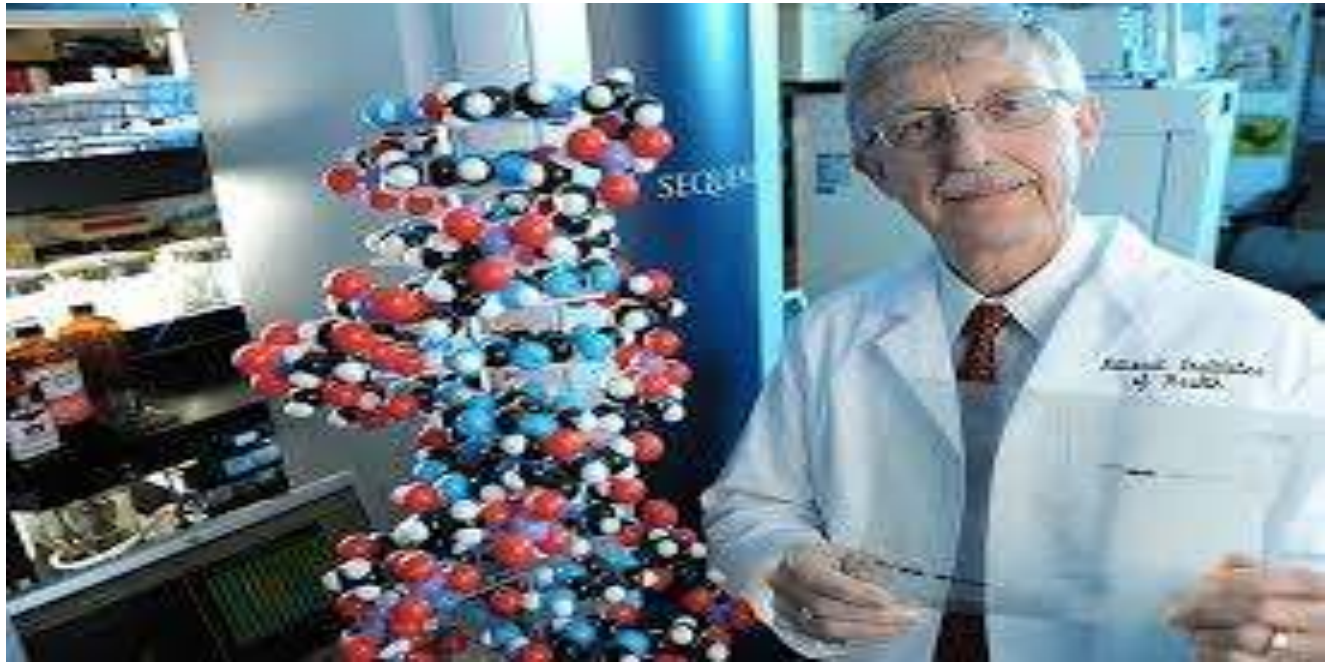
HISTORIA DE LA GENOMICA: EL PROYECTO GENOMA HUMANO



HISTORIA DE LA GENOMICA: EL PROYECTO GENOMA HUMANO



"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
(1990 – 2000/2003)

Héritas
Medicina de precisión
CIBIC + INDEAR

HISTORIA DE LA GENOMICA: EL 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
- Transferencia de tecnologías asociadas al sector privado
- Definir principios éticos, legales y sociales

Héritas
Medicina de precisión
CIBIC + INDEAR

Humans have fewer genes

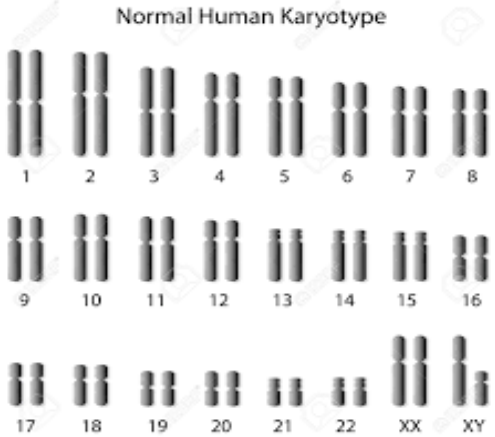
In Thursday's issue of the journal Nature, researchers who decoded the human genome concluded that people have only 20,000 to 25,000 genes, a drop from the 30,000 to 40,000 estimated in 2001.

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

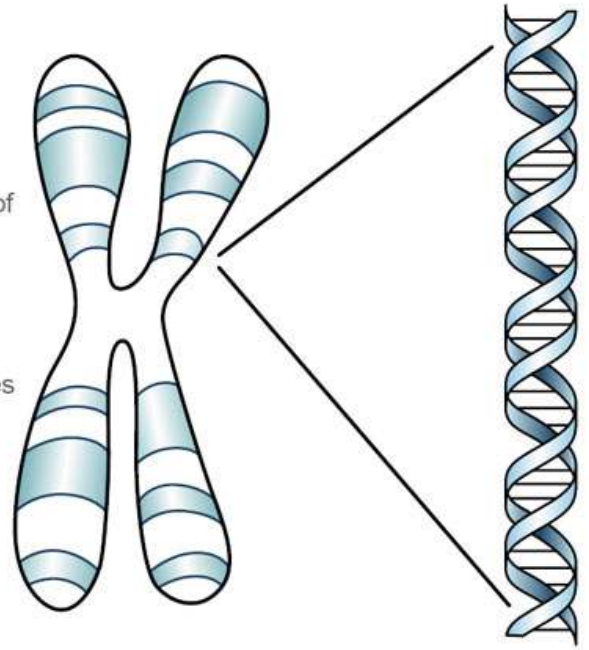
SOURCE: Nature

AP

HISTORIA DE LA GENOMICA: EL PROYECTO GENOMA HUMANO



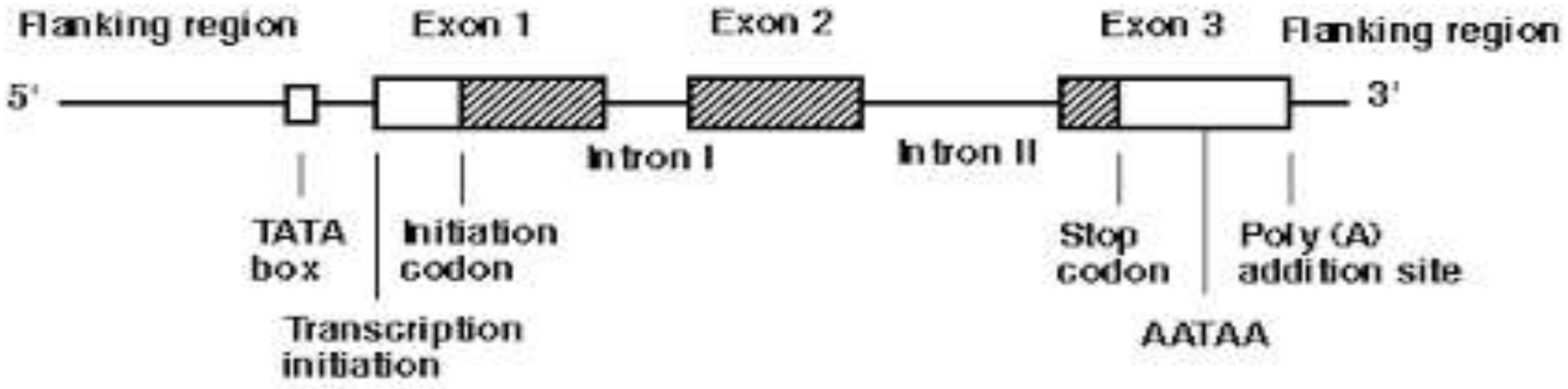
We all have 23 pairs of chromosomes. One pair of chromosomes determines our sex. The other 22 pairs of chromosomes are non-sex chromosomes and determine things like hair color and our eye color.



Gene Each chromosome is made up of many genes. Genes are made of a section of a long molecule called DNA. Genes carry the genetic information.

Chromosome

DNA DNA codes the genetic information on a gene.



Héritas
Medicina de precisión
CIBIC + INDEAR

Whole Genome (3 billones pb) 25,000 genes
1-2% exones 23% intrones 75% intergenómico

Héritas
Medicina de precisión
CIBIC + INDEAR

25,000 genes
(60 millones pb)
1-2% *exones*

Se estiman alrededor de $10x10^6$ SNPs
Single **N**ucleotide **P**olimorphism

Mutaciones vs. polimorfismos

Definiciones

nature
articles

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

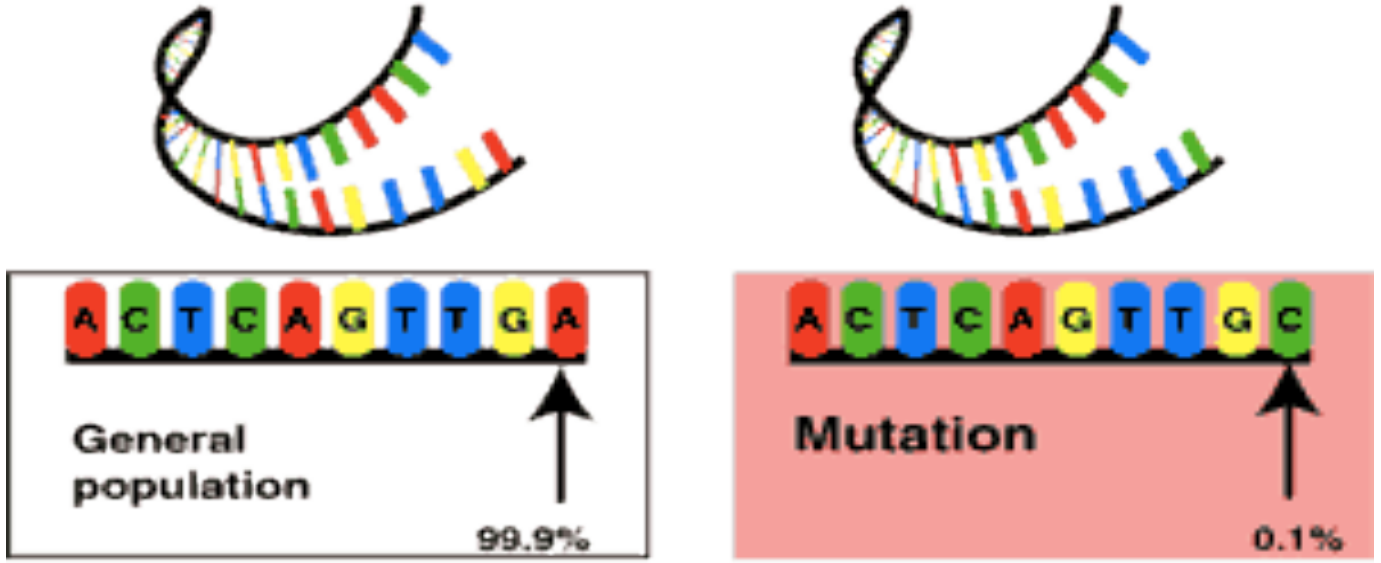
The International SNP Map Working Group*

* 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.

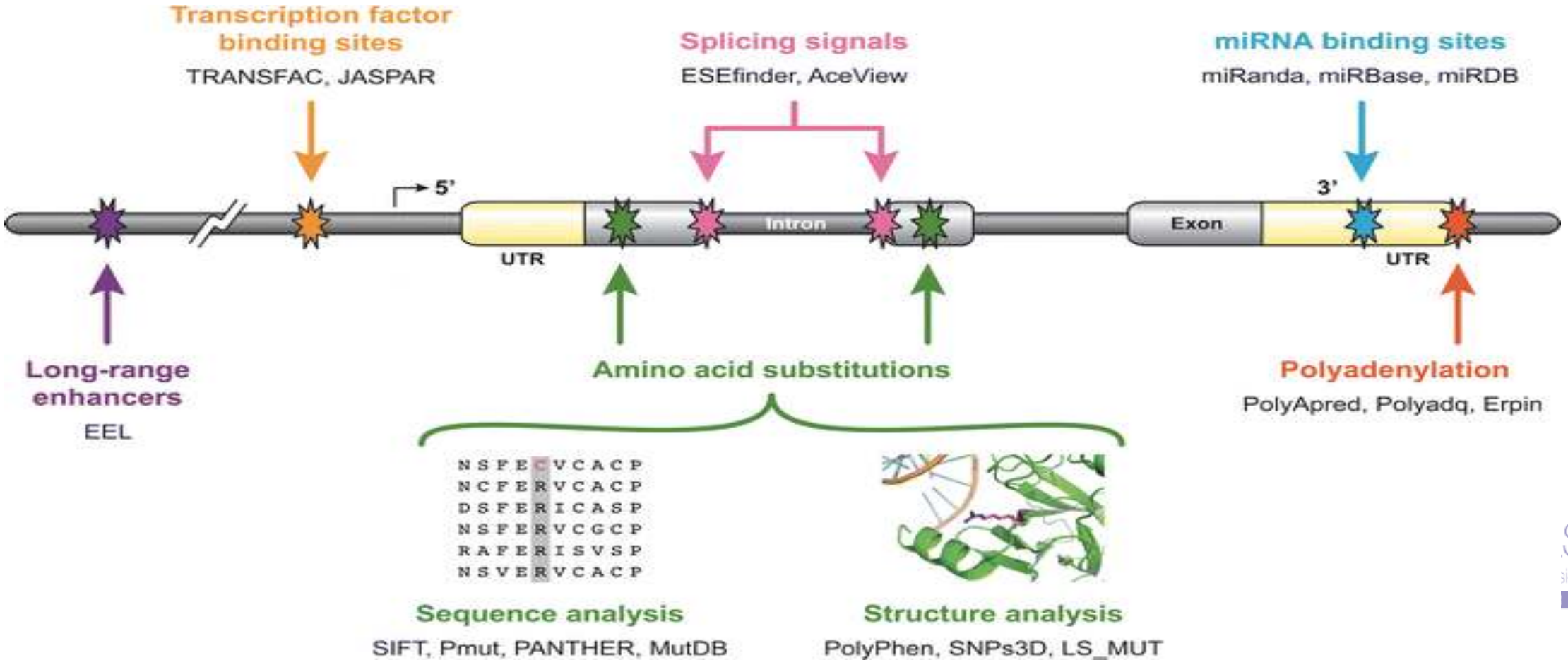
Asociados a la respuesta de los individuos a enfermedades herditarias, enfermedades infecciosas y respuesta a fármacos



Héritas
Medicina de precisión
CIBIC + INDEAR

SNP codificantes: se localizan en la secuencia codificante pueden modificar o no la cadena de aminoácidos de proteinas

SNP no-codificantes: en regiones no codificantes. Consecuencias en el proceso de traducción (*splicing*, factores de transcripción) > Reguladores



Hallazgos

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

  ➔ Genes reguladores
  ➔ Genes con mas de una función
  ➔ Moléculas reguladoras (ADN,ARN,mARN,siARN)

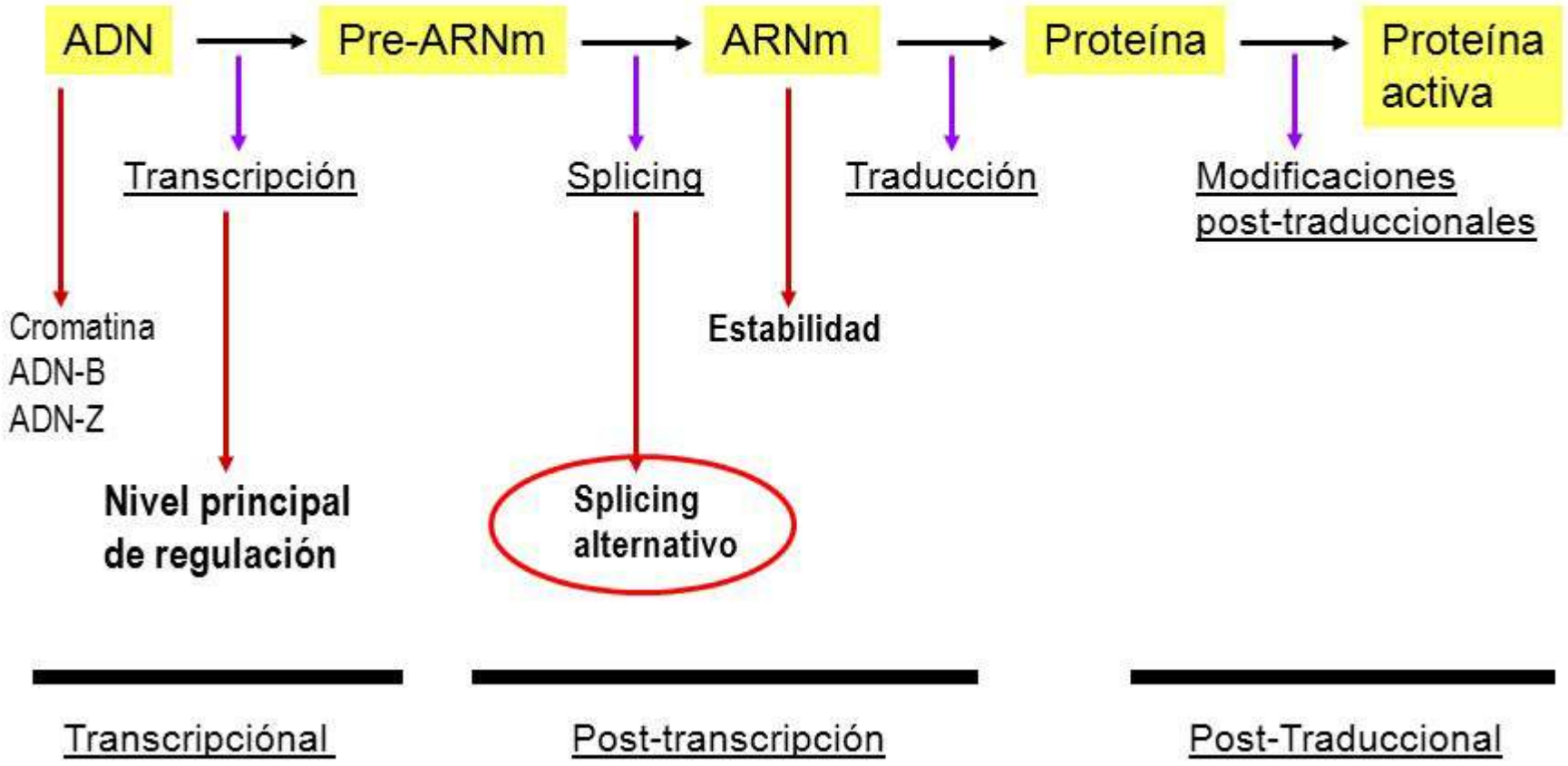
  La mayoría de los genes son **polimórficos**
  estrategia biológica para la evolución y para la supervivencia

- El estudio del genoma está relacionado con las enfermedades multifactoriales y comprender esta relación requiere conocimiento y herramientas de análisis

Regulación Genética

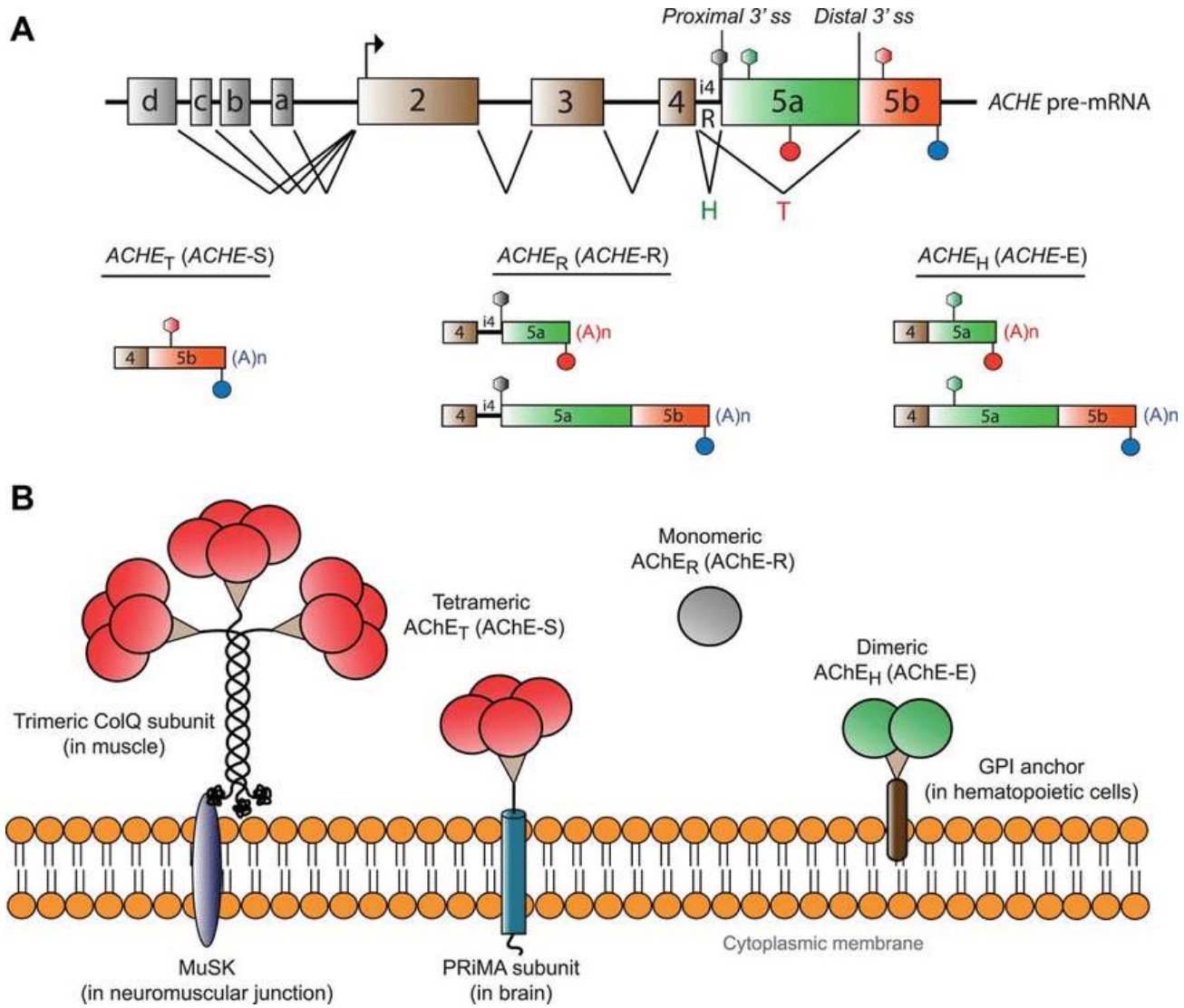
Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Genética
Clasificación



Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Genética – *Splicing alternativo*



Héritas
Medicina de precisión
CIBIC + INDEAR

REGULACION EPIGENETICA

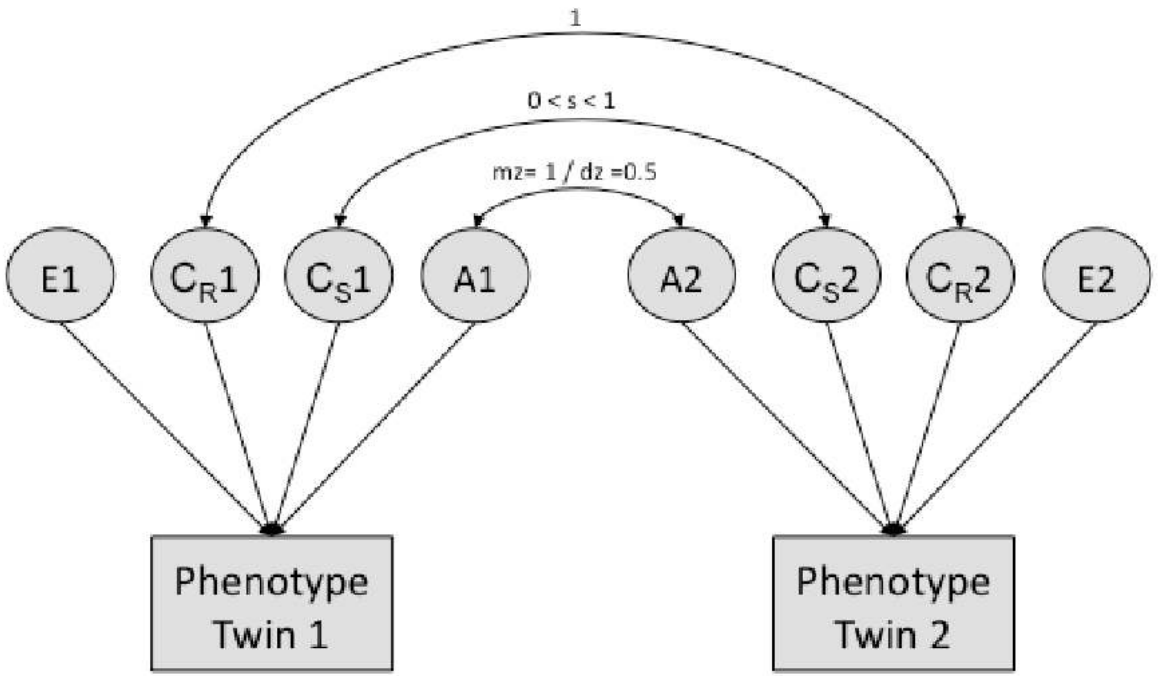
Héritas
Medicina de precisión
CIBIC + INDEAR

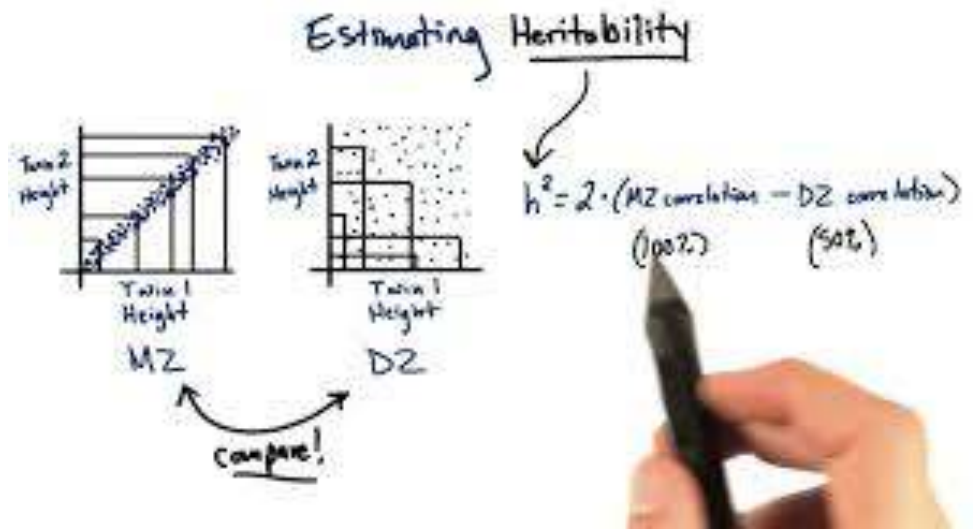
Cambios heredables de la expresión génica que ocurren *sin* que se presenten modificaciones en la secuencia de ADN

Principales mecanismos epigenéticos

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

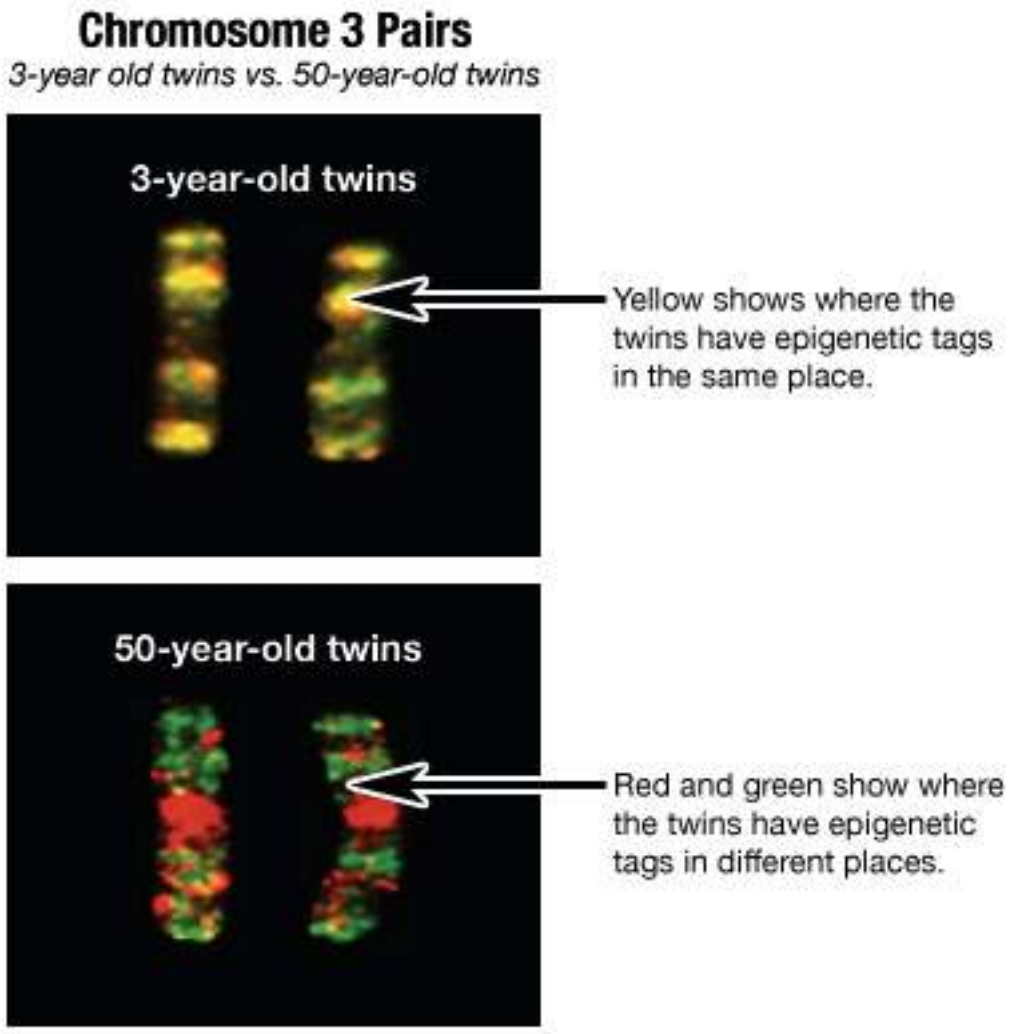
Héritas
Medicina de precisión
CIBIC + INDEAR


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!

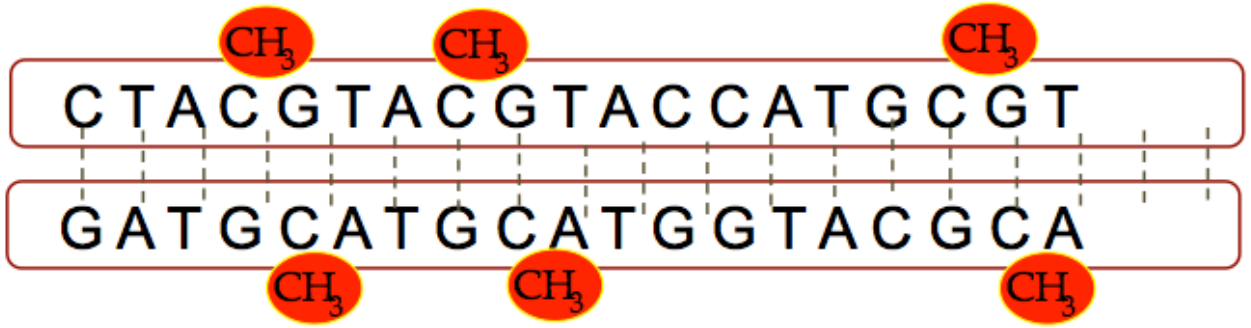
www.ncbi.nlm.nih.gov

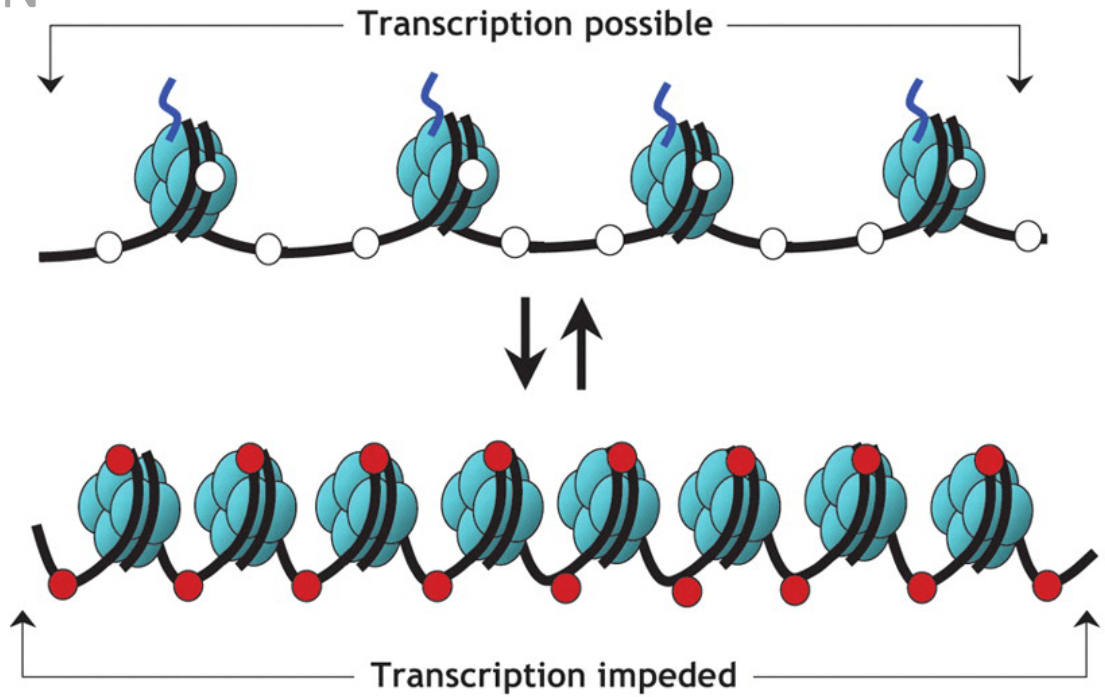
Epigenetics of discordant monozygotic twins: implications for disease



Héritas
Medicina de precisión
CIBIC + INDEAR

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





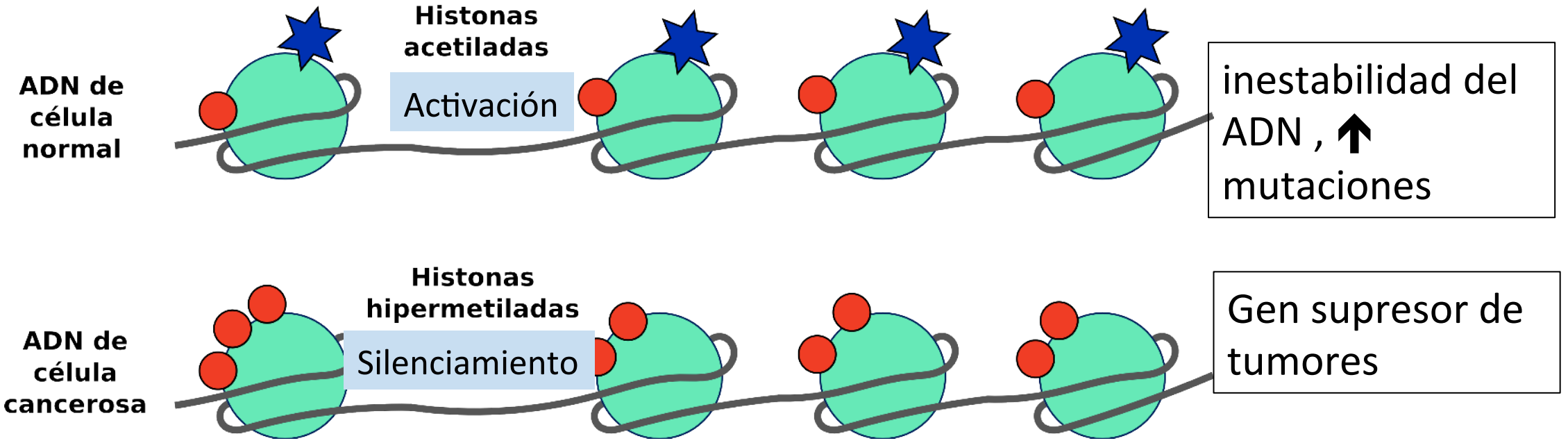
Ocurre generalmente en Citosinas, especialmente en nucleótidos emparejados CpG (Dímeros metilados CpG)

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

Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Epigenética

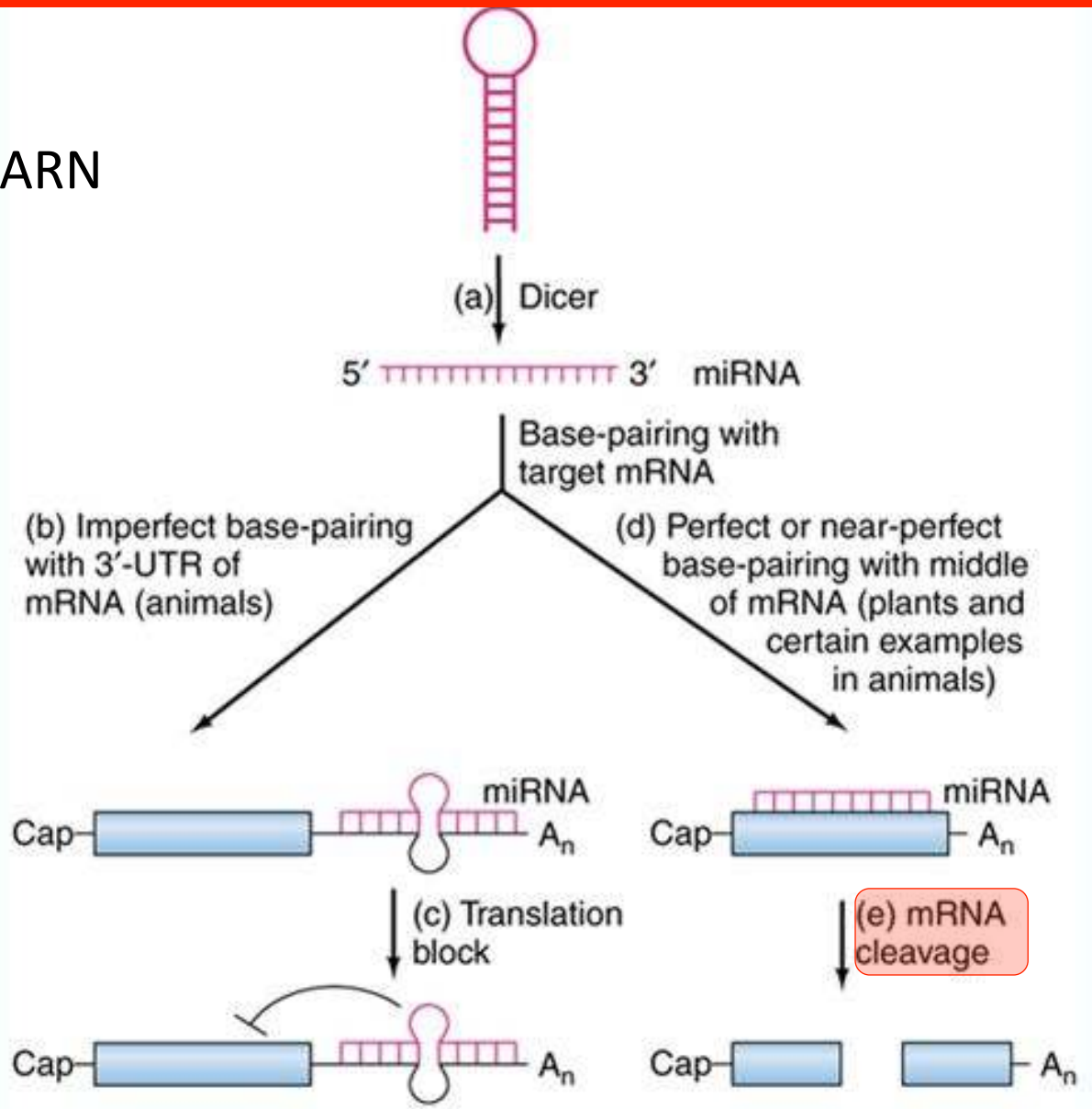
- Metilación del ADN
- Modificación post-traduccional de Histonas
- Silenciamiento de genes mediados por ARN



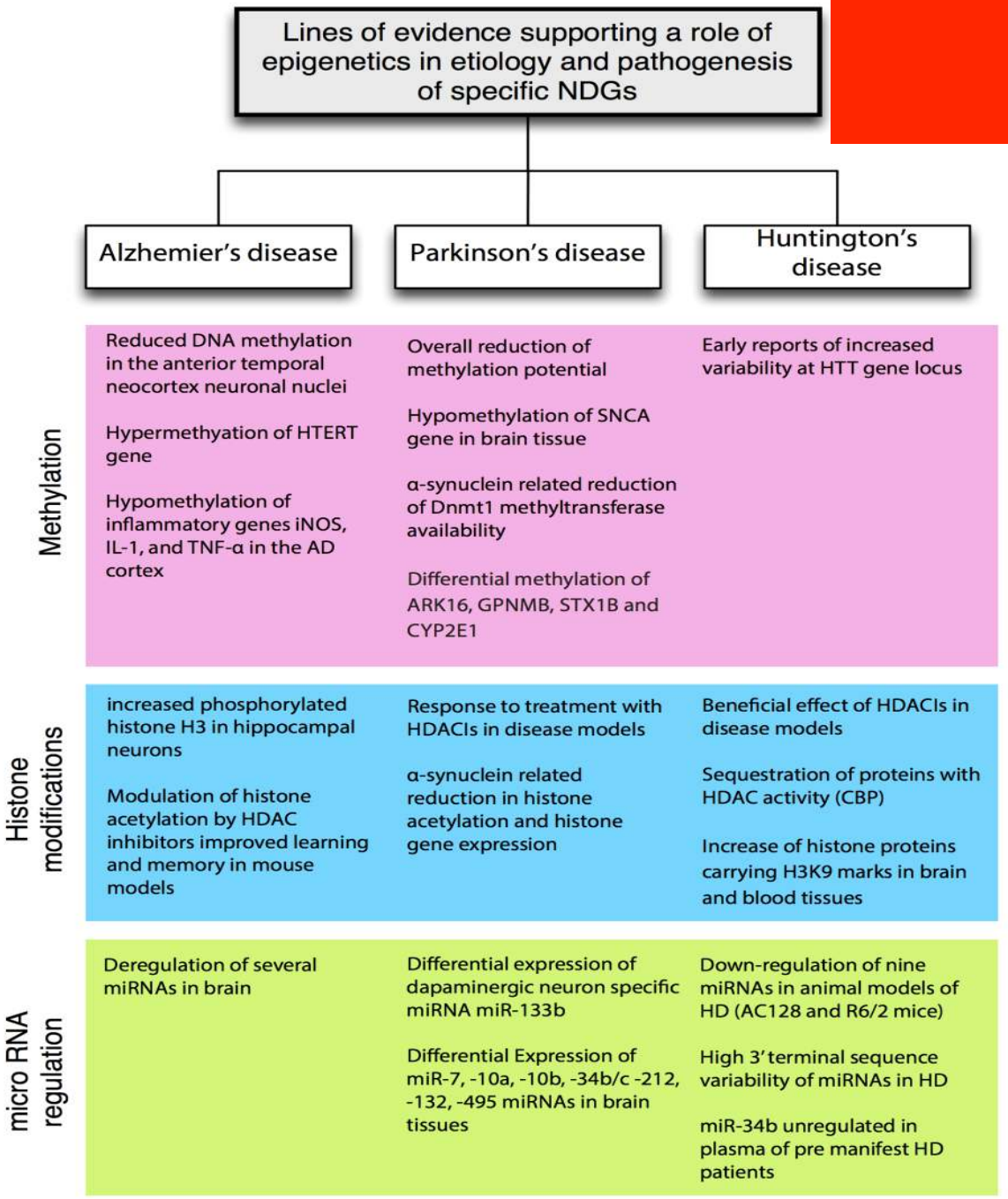
Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Epigenética

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

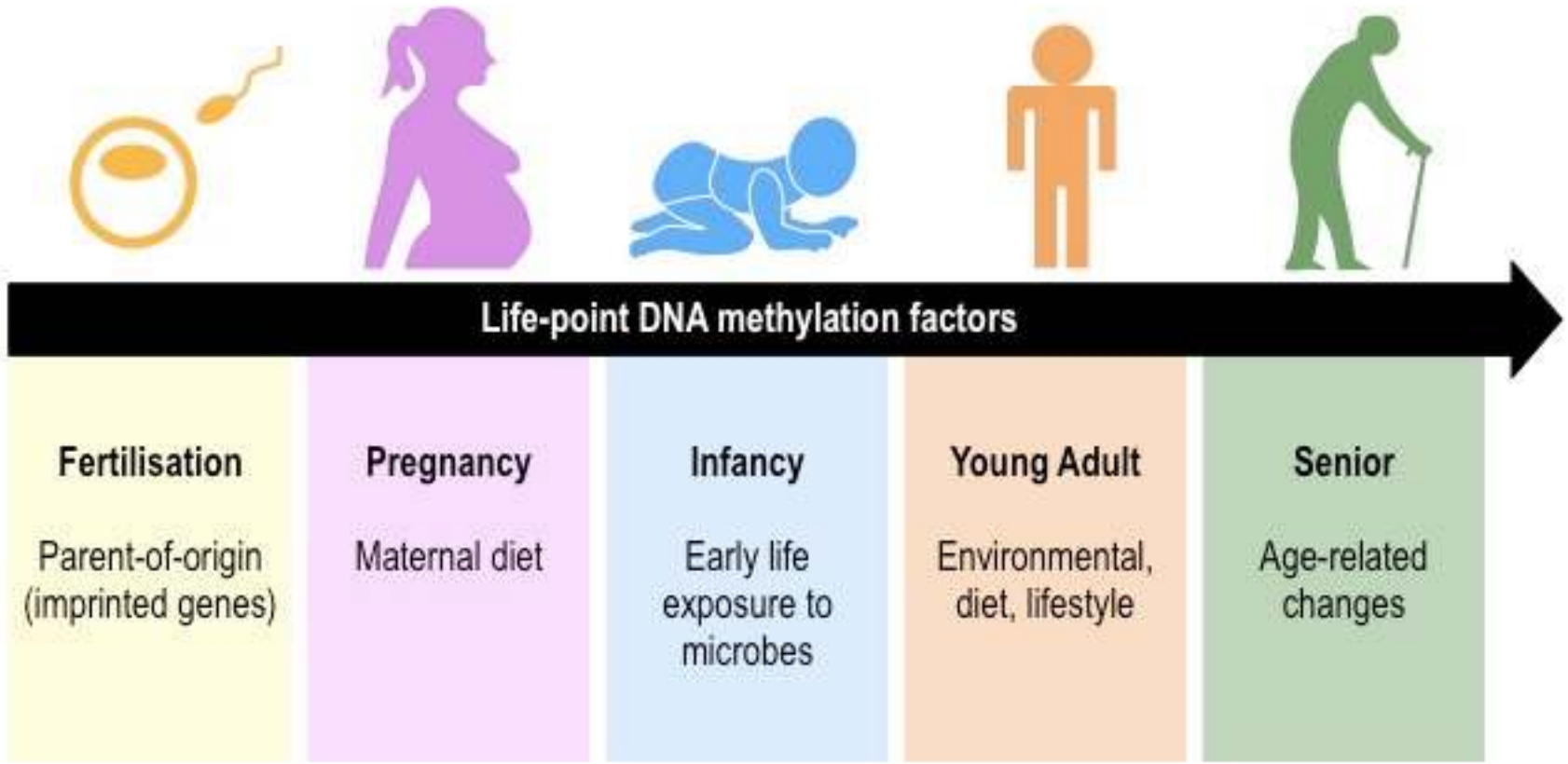


Regulación Epigenética



Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Epigenética



Héritas
Medicina de precisión
CIBIC + INDEAR

Regulación Epigenética

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

Héritas
Medicina de precisión
CIBIC + INDEAR

Aplicaciones de la *Biología Molecular* en Genética Médica

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

Héritas
Medicina de precisión
CIBIC + INDEAR

Enfermedades hereditarias

Conjunto de enfermedades genéticas cuya característica principal es su supervivencia de generación en generación, transmitiéndose de padres a hijos.

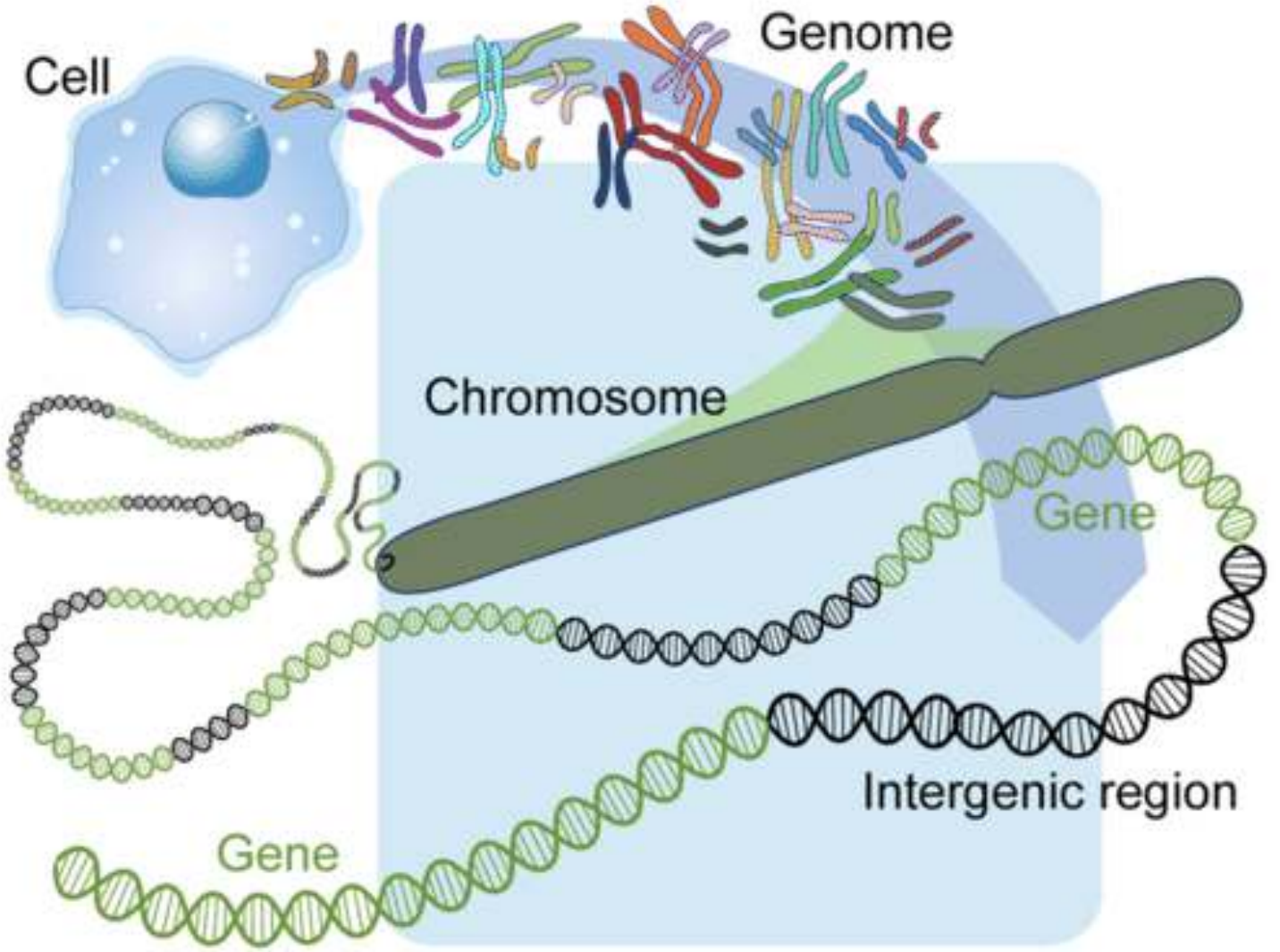
Héritas
Medicina de precisión
CIBIC + INDEAR

- **Cromosómicas**
  Anormalidades numéricas
  Anormalidades estructurales

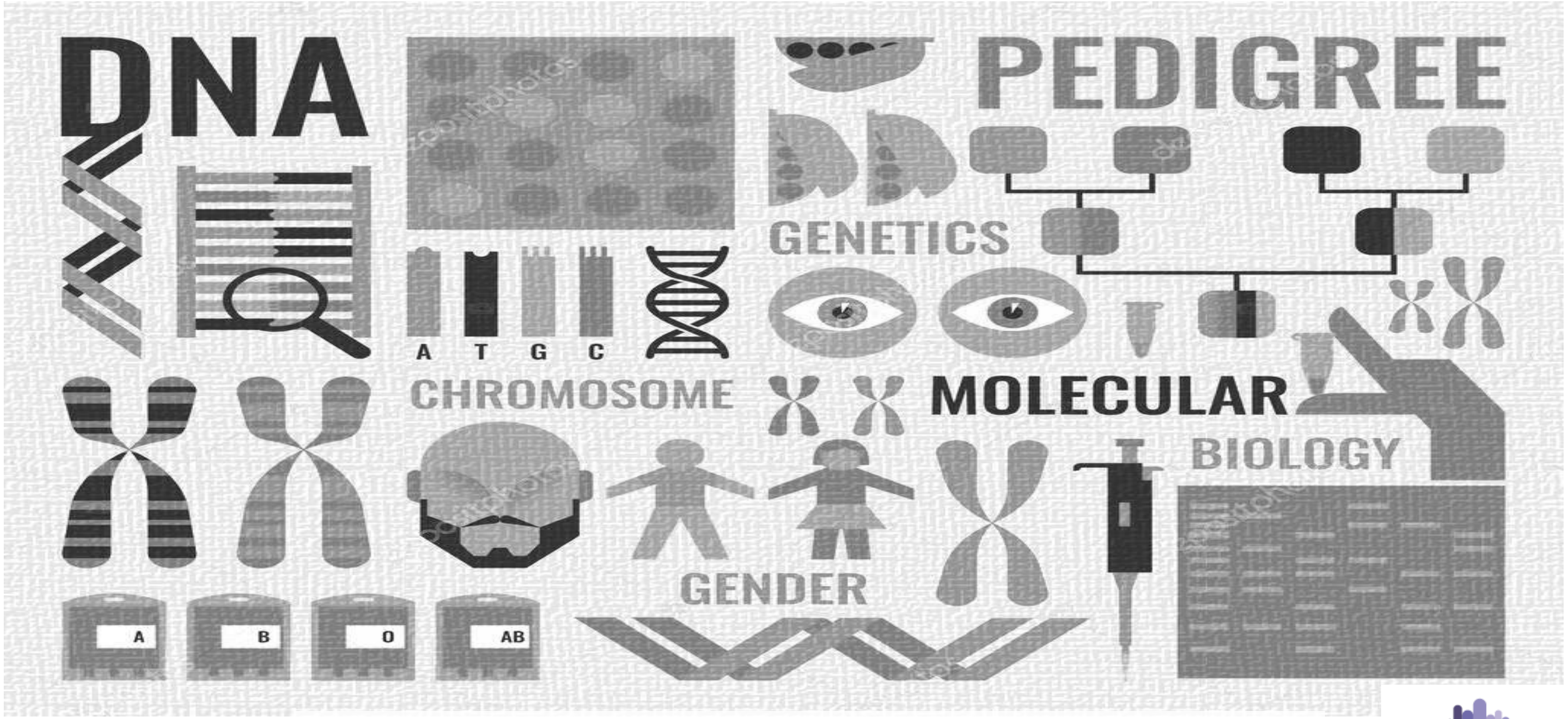
- **Monogénicas**
  Herencia Mendeliana

- **Multifactoriales**
  Poligénicas
  GxE

- **Mitocondriales** (herencia materna)



Héritas
Medicina de precisión
CIBIC + INDEAR



Enfermedades hereditarias Mendelianas

Tipos de Herencia

Héritas
Medicina de precisión
CIBIC + INDEAR

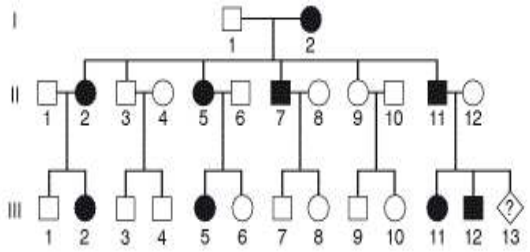
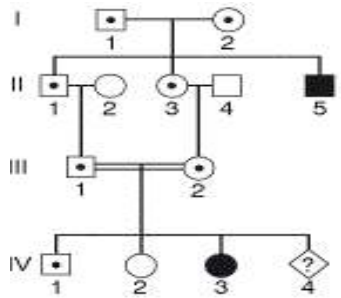
Enfermedades hereditarias Mendelianas
Árbol genealógico

El *pedigree* es una forma de análisis genético en donde el médico genetista hace un diagrama que muestra a un individuo con una característica estudiada y todos sus familiares conocidos.

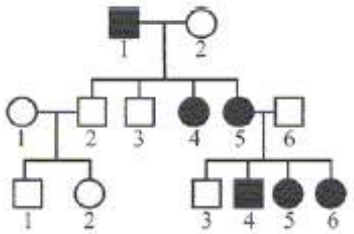
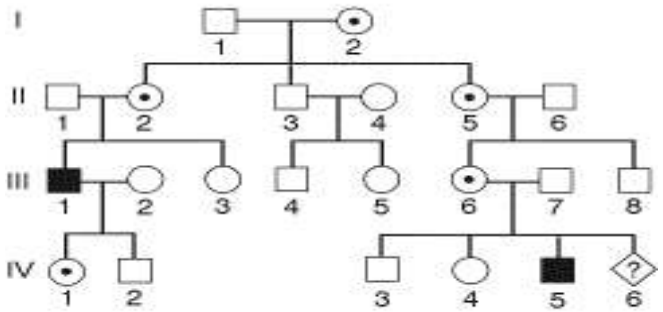
El *pedigree* indica la presencia o ausencia de esta característica y si es aplicable la variación de expresión de la misma

Propósito facilitar el análisis genético de una característica examinando su posible **patrón de herencia** en una familia en particular

Héritas
Medicina de precisión
CIBIC + INDEAR

	PEDIGREE	PROGENITOR AFECTADO	CON-SANGUINEIDAD	PROBABILIDAD DE TRANSMISION (EJEMPLOS)
AUTOSOMICA DOMINANTE		SI	NO	50% - PROGENITOR AFECTADO (NF1, MARFAN, NOONAN)
AUTOSOMICA RECESIVA		NO (PORTADORES ASINTOMATICOS OBLIGADOS)	SI	25% - AFECTADO (FIBROSIS QUISTICA, AME)

Héritas
Medicina de precisión
CIBIC + INDEAR

	PEDIGREE	PROGENITOR AFECTADO	MUJERES	EJEMPLOS
LIGADO A X DOMINANTE		AMBOS SEXOS	AFECTADAS	SME RETT, HIPOFOSFATEMIA
LIGADO A X RECESIVA		SOLO VARONES	PORTADORAS	HEMOFILIA, E. FABRY

Héritas
Medicina de precisión
CIBIC + INDEAR

EXISTEN DESVIACIONES DE LOS PATRONES MENDELIANOS CLÁSICOS…

- Penetrancia incompleta
- Expresividad variable
- Mutaciones de-novo
- Mosaicismo germinal
- Impronta genética
- Heterogeneidad de locus
- Mutaciones dinámicas

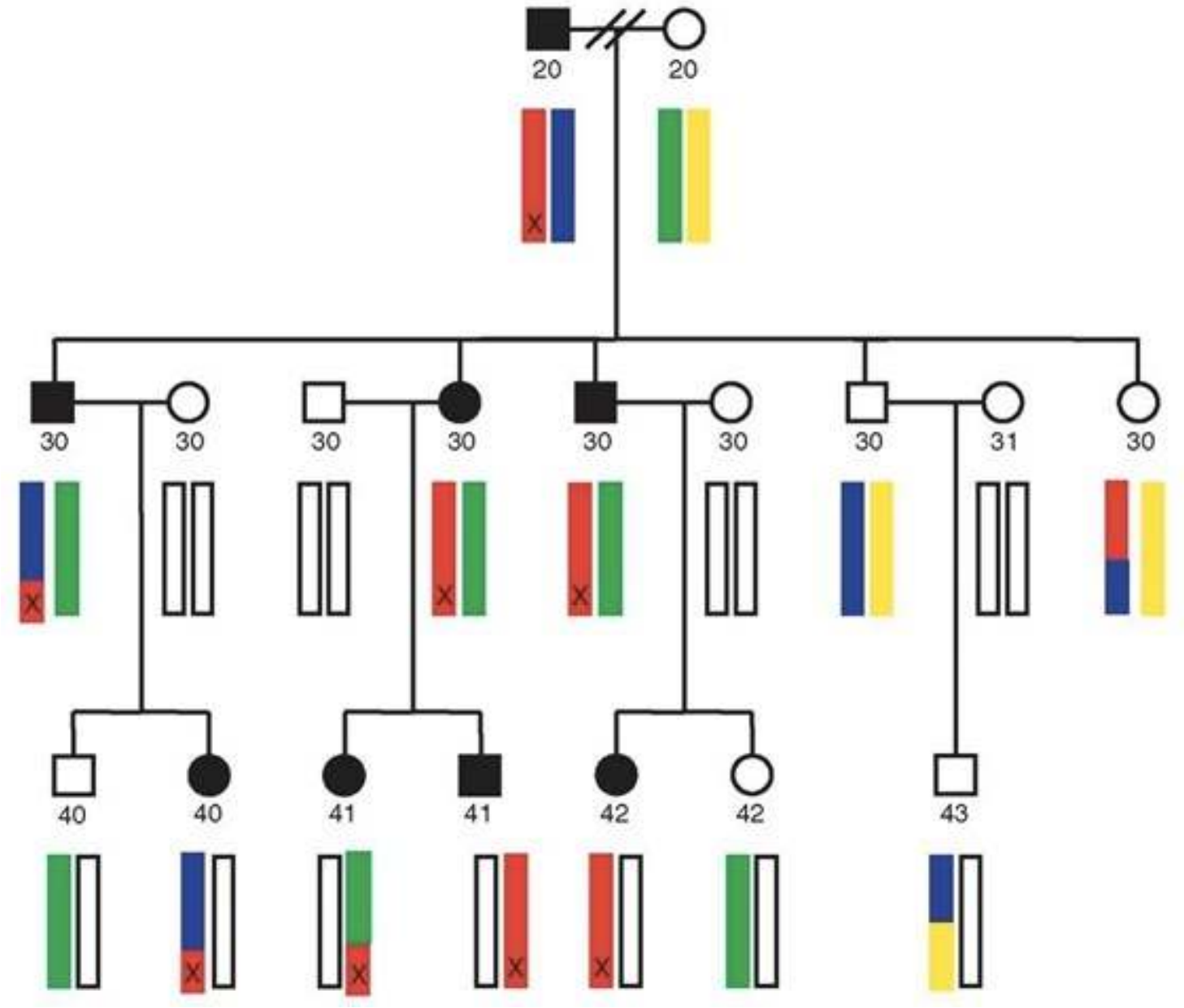
Generalmente los patrones de herencia no pueden ser definidos

Héritas
Medicina de precisión
CIBIC + INDEAR

COMO SE ESTUDIABAN LAS ENFERMEDADES MENDELIANAS PREVIO A LA FINALIZACION DEL PROYECTO GENOMA HUMANO?

Héritas
Medicina de precisión
CIBIC + INDEAR

Enfermedades hereditarias Mendelianas - Estudios de Ligamiento

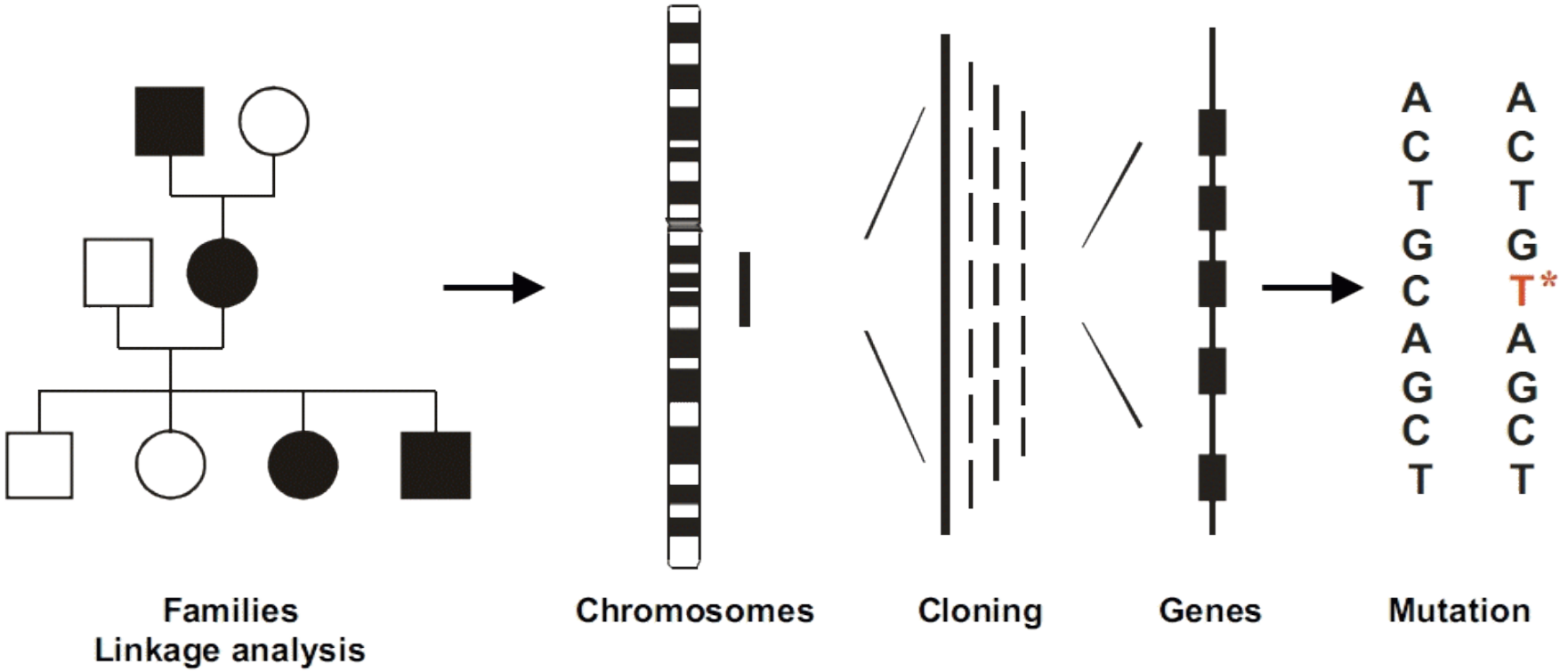


Héritas
Medicina de precisión
CIBIC + INDEAR

Gene mapping

Gene identification

Functional studies



Estudios funcionales para **confirmar la causalidad** de la mutacion identificada

Héritas
Medicina de precisión
CIBIC + INDEAR

Estudios de Ligamiento

Limitaciones

- Tiempo de realizacion: 3-6 meses
- Altamente laborioso
- Familias numerosas
- Polimorfismos informativos (microsatélites)
- Suficiente cantidad de meiosis informativas (por lo menos 3 generaciones)
- Requiere posterior *finemapping* para posteior secuenciacion de genes candidatos

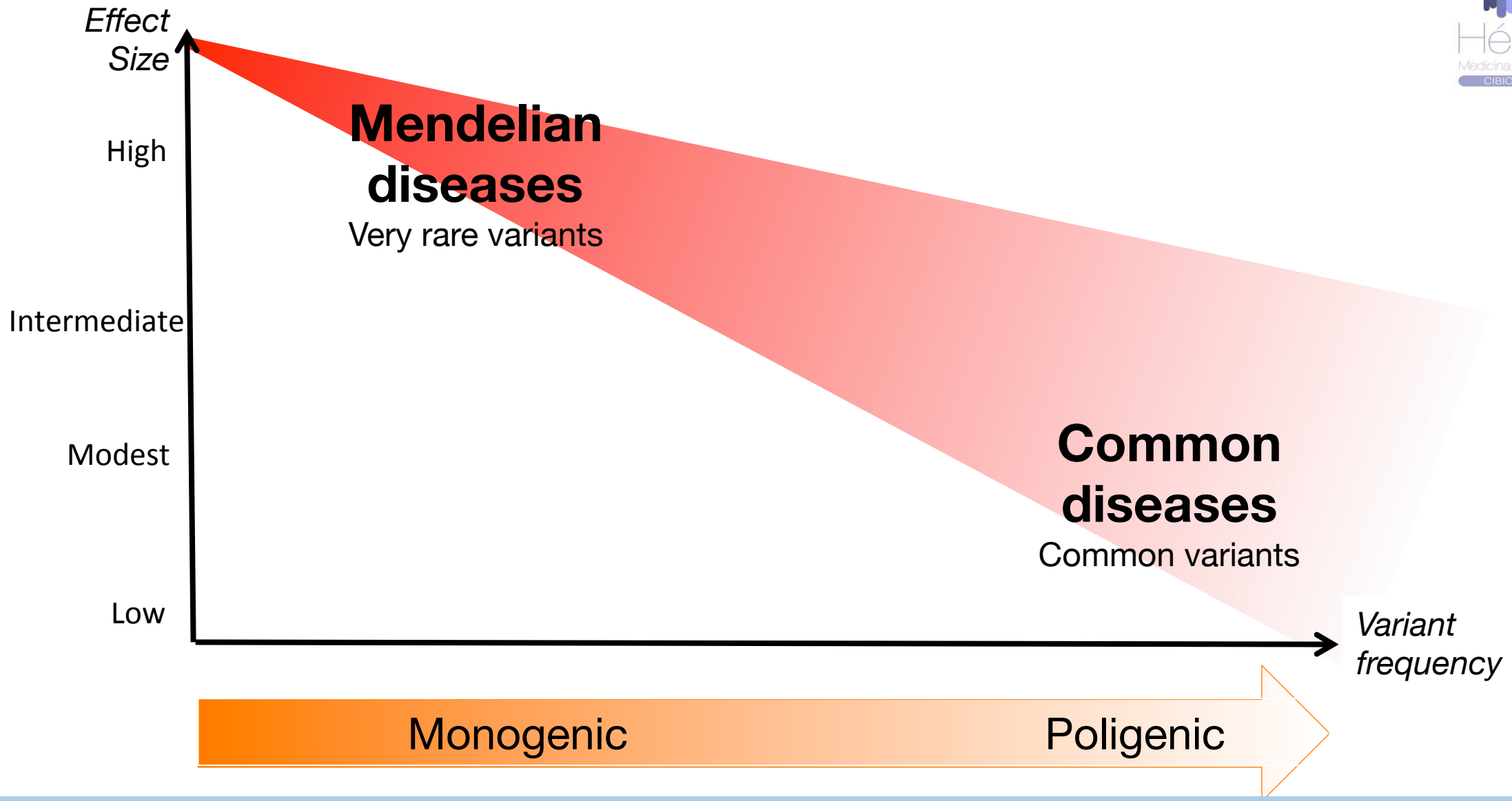
Héritas
Medicina de precisión
CIBIC + INDEAR

COMO ESTUDIAMOS LAS ENFERMEDADES MENDELIANAS EN LA ERA POSTGENOMICA?

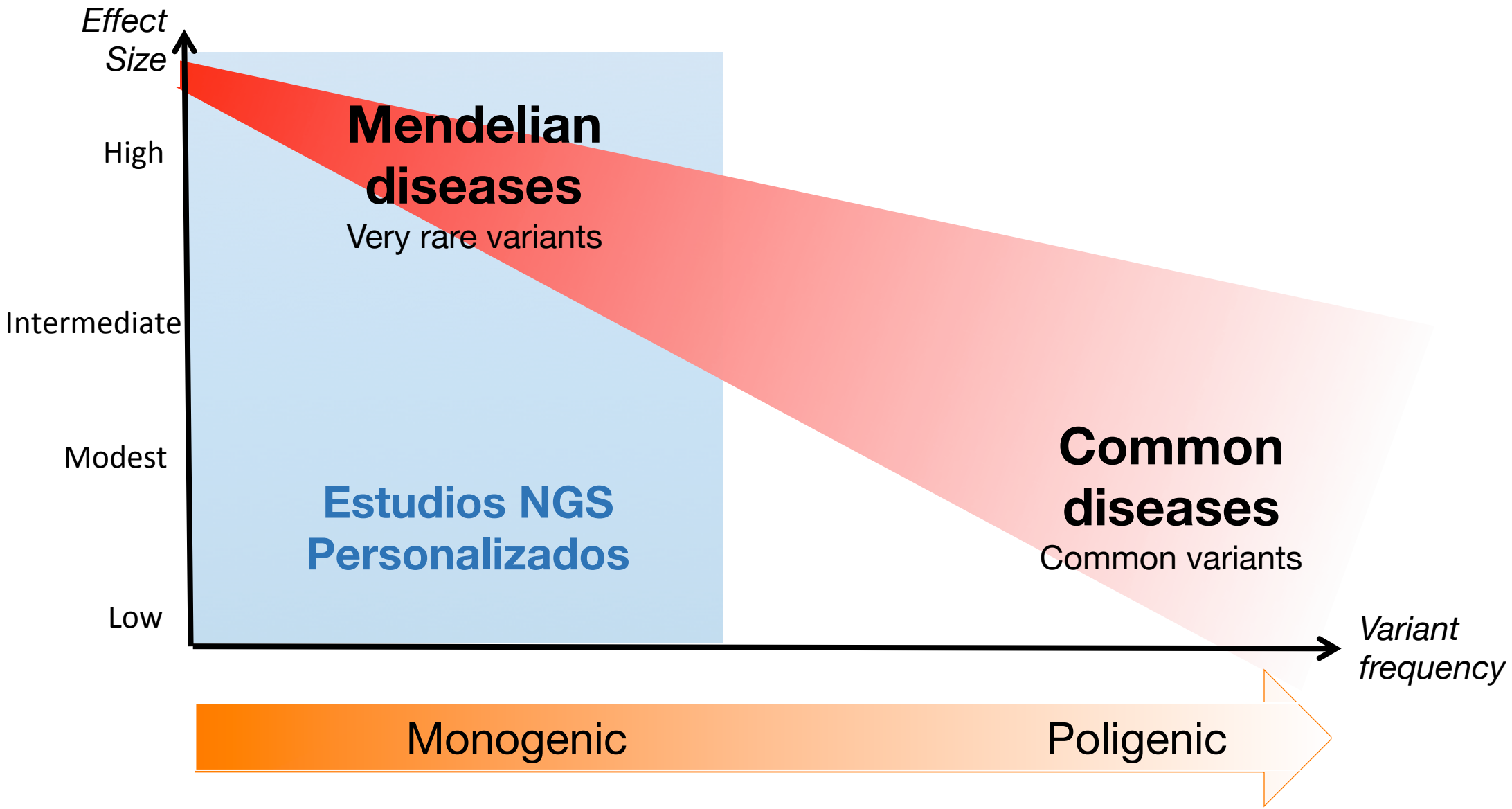
Héritas
Medicina de precisión
CIBIC + INDEAR

EL PROYECTO GENOMA HUMANO
1990-2003

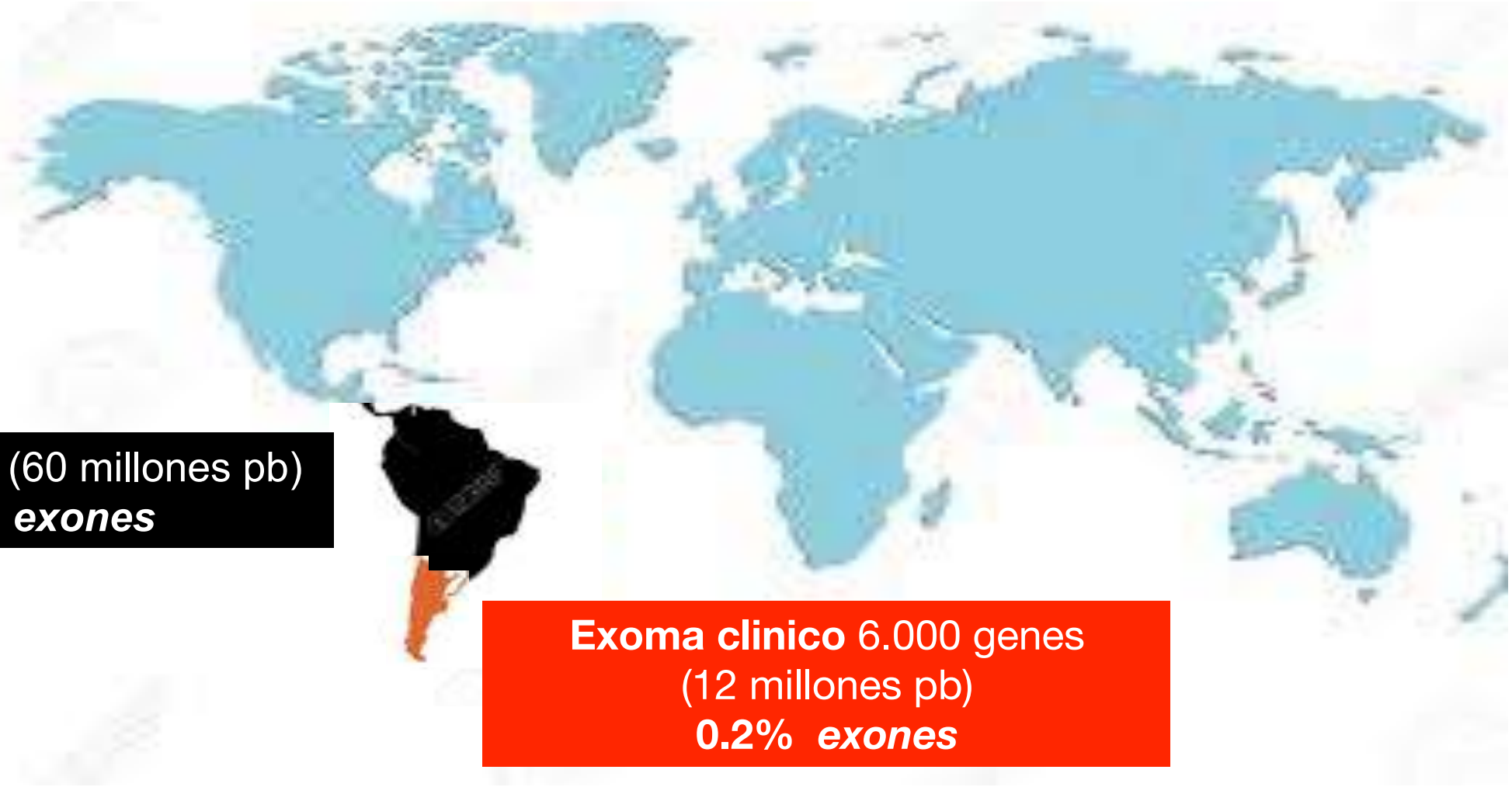
Héritas
Medicina de precisión
CIBIC + INDEAR



En función al tipo de patología sospechada, se utilizarán diferentes técnicas de biología molecular


Effect Size
High
Intermediate
Modest
Low
Mendelian diseases
Very rare variants
Estudios NGS Personalizados
Common diseases
Common variants
Variant frequency
Monogenic
Poligenic

Whole Genome (3 billones pb) 25,000 genes
1-2% exones 23% introns 75% intergenómico



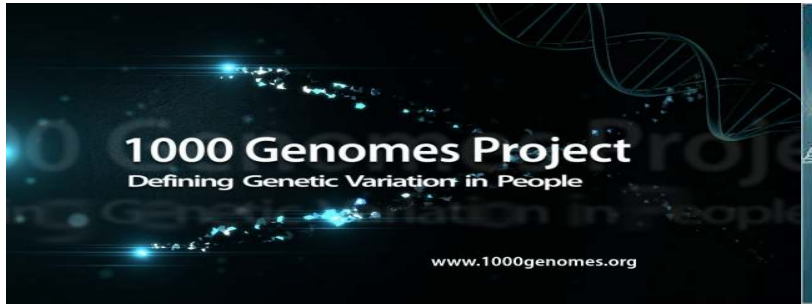
25,000 genes (60 millones pb)
1-2% *exones*

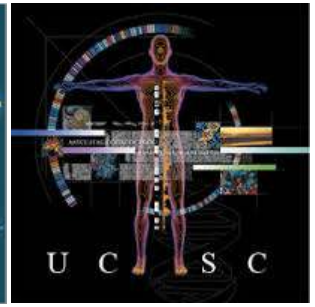
Exoma clinico 6.000 genes
(12 millones pb)
0.2% *exones*

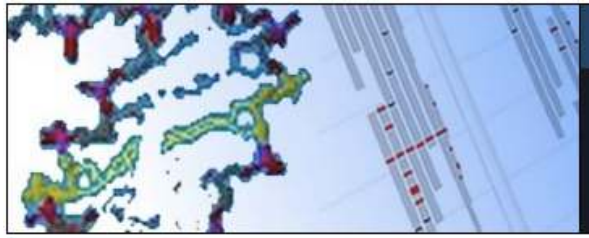
Estudios NGS

Base 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

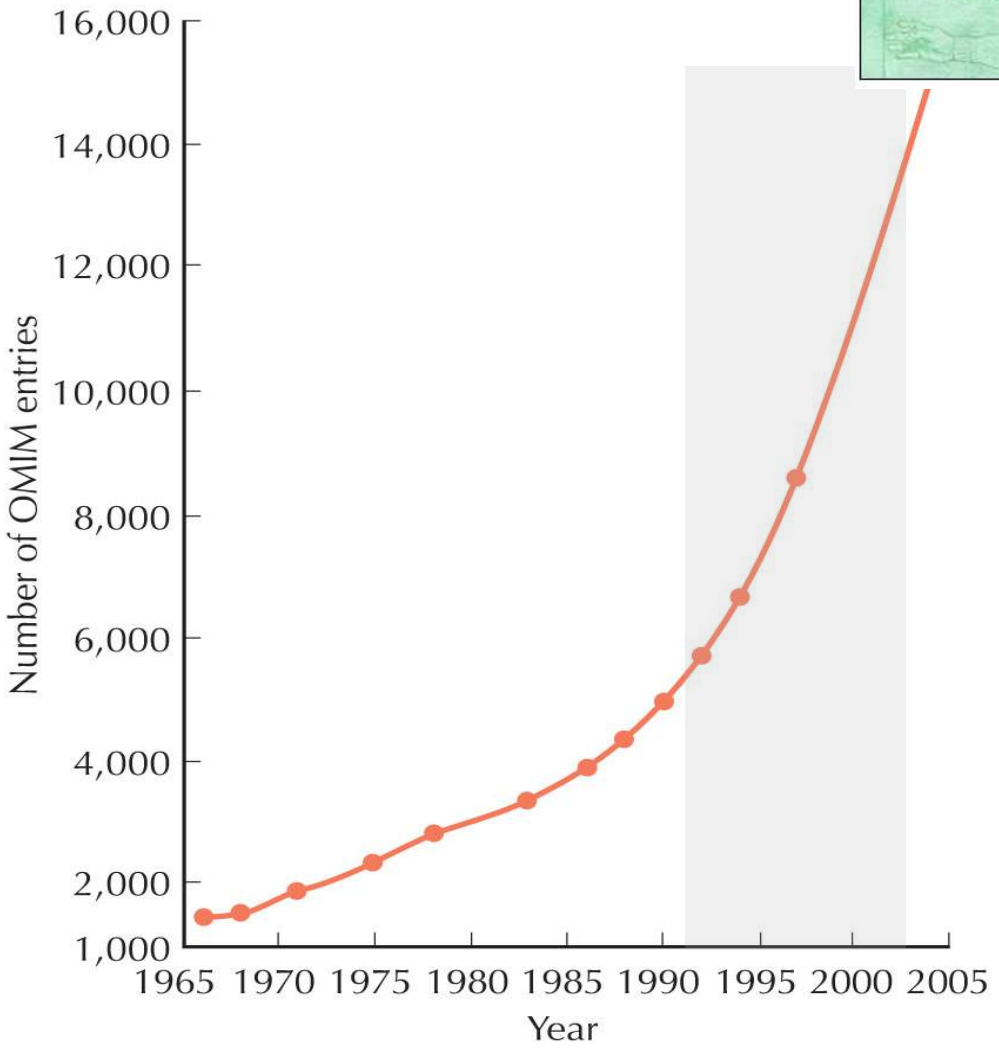
OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh. Its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

En función al tipo de patología sospechada, se definen diferentes técnicas de biología molecular

Next Generation Sequencing
(NGS)

Héritas
Medicina de precisión
CIBIC + INDEAR

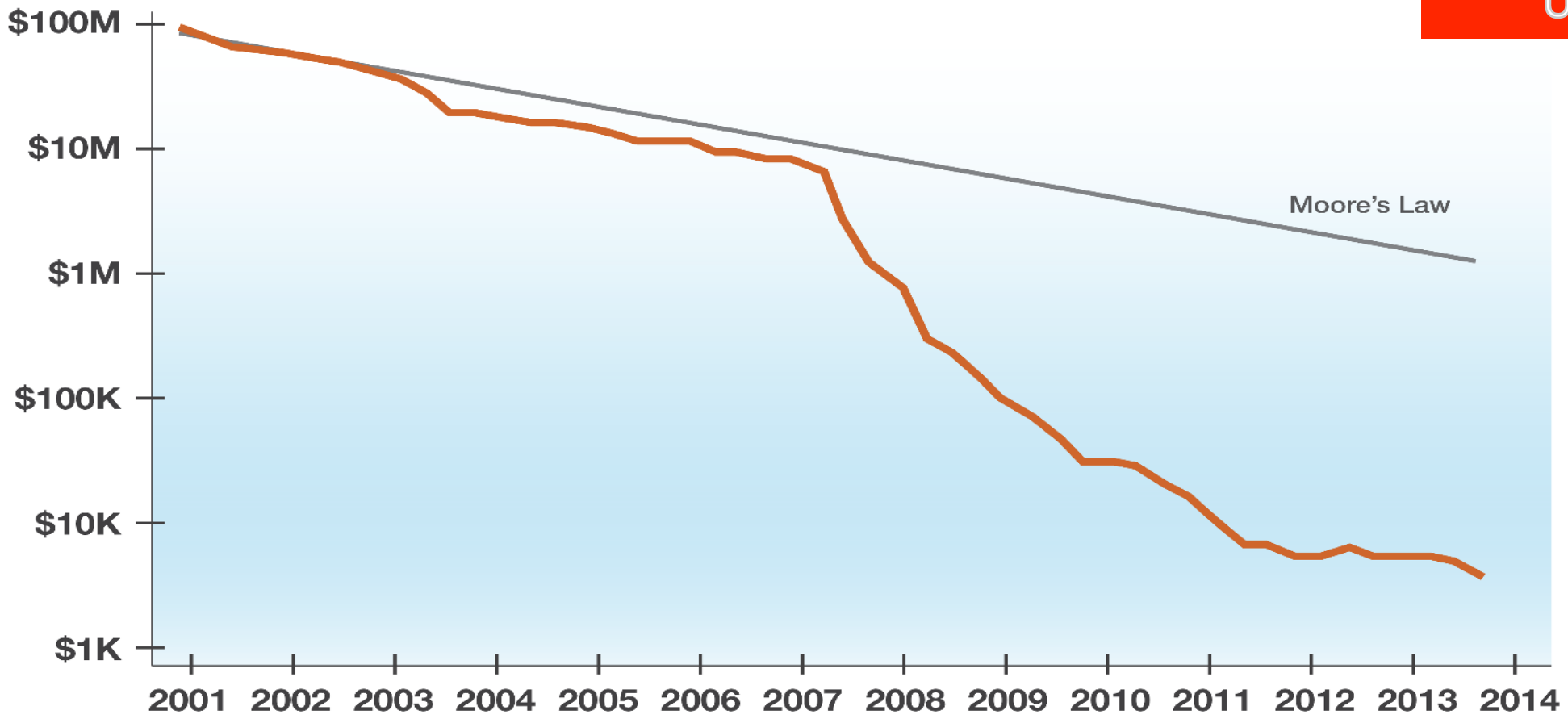
1990
PGH - 13a
U$D 3,000,000,000

Falling fast

In the first few years after the end of the Human Genome Project, the cost of genome sequencing roughly followed Moore's law, which predicts exponential declines in computing costs. After 2007, sequencing costs dropped precipitously.

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

Cost Per Genome

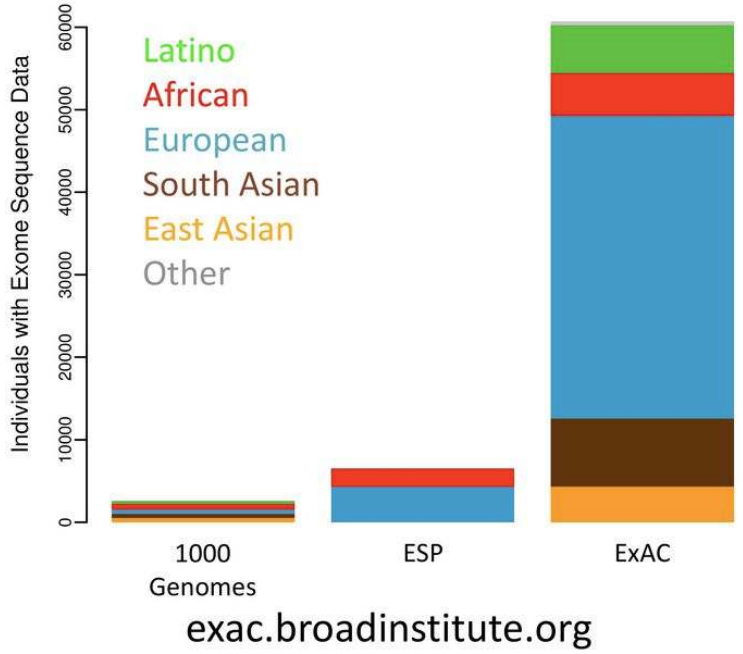


Héritas
Medicina de precisión
CIBIC + INDEAR

exac.broadinstitute.org

ExAC Browser

Population frequency Databases ExAc / GnomAD



Estandarizacion de diversos datasets de proyectos de secuenciacion existentes

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

Héritas
Medicina de precisión
CIBIC + INDEAR

Análisis bioinformático / **Interpretación de variantes**

- ✓ **OMIM (≅4000 genes, >6200 fenotipos GENES)**
- ✓ **Databases ExAc / GnomAD**
- ✓ **Clasificación de variantes (ACMG)**

- ☑ **Bening**
- ☑ **Likely-bening**
- ☑ **Variant of unknown significance**
- ☑ **Likely-patogenic**
- ☑ **Pathogenic**

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very strong
Population data	MAF is too high for disorder BA1/BS1 OR observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affectede statistically increased over controls PS4	
Computational data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other data		Reputable source w/out shared data = benign BP6 Found in case with an alternate cause BP5	Reputable source = pathogenic PP5 Patient's phenotype or FH highly specific for gene PP4			

Programas de Controles externos para validacionde estudios NGS

EMQN The European Molecular Genetics Quality Network

COLLEGE of AMERICAN PATHOLOGISTS

Harvard Business Review
РОССИЯ

ANALYTICS

Data Scientist: The Sexiest Job of the 21st Century

Conclusiones

- Costos de estudios moleculares cada vez más accesibles
- Algoritmos diagnósticos clínicos → **QUE**
- Algoritmos diangósticos moleculares → **COMO**
- Equipos interdiciplinarios para diagnóstico de pacientes

Héritas
Medicina de precisión
CIBIC + INDEAR

Gracias !!!

Héritas
Medicina de precisión
CIBIC + INDEAR