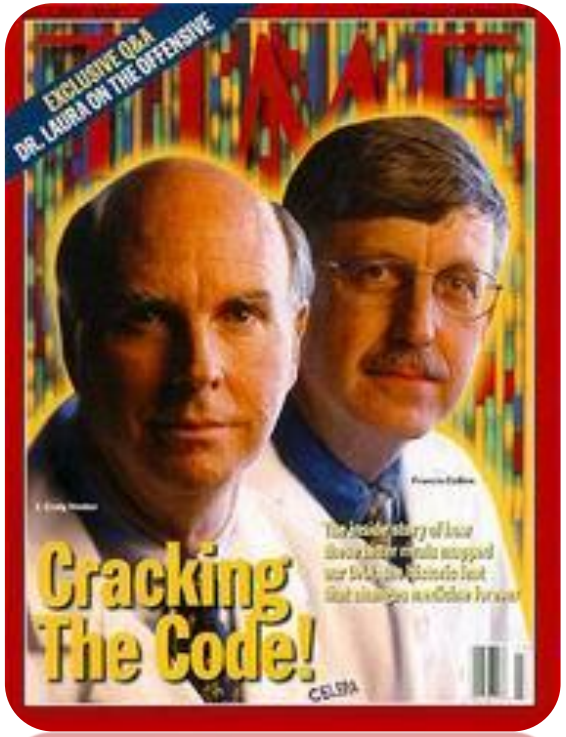
Librerías y Tecnologías de Secuenciación del ADN



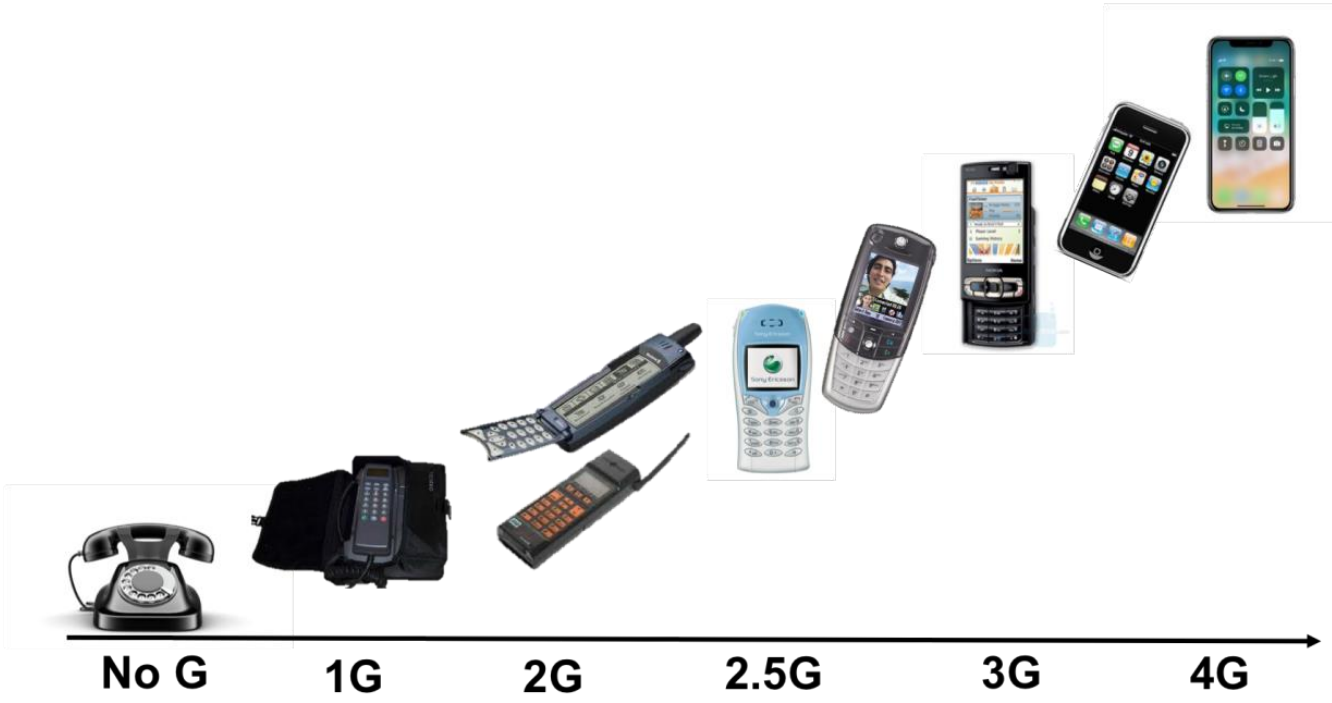
Rafael Fort

Secuenciación del ADN

- **Secuenciación de 1^{era} generación**
  - Método de Fred Sanger
    - Método enzimático de secuenciación

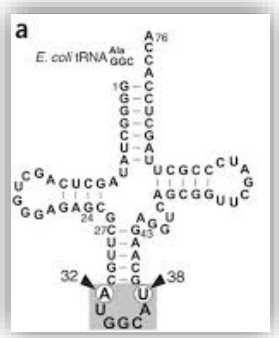
Proyecto Genoma Humano

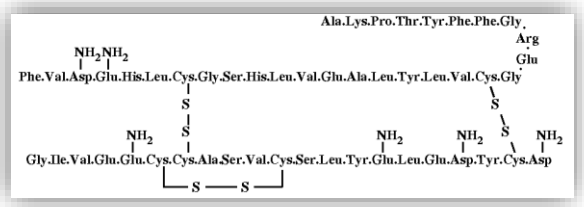
- Secuenciación de 1^{era} generación



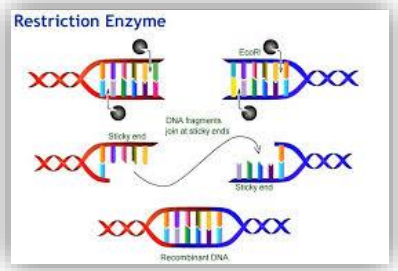
- ***Secuenciación de Nueva Generación***
  - Secuenciación de **2da y 3^{era} Generación** …

Breve reseña histórica de la Secuenciación

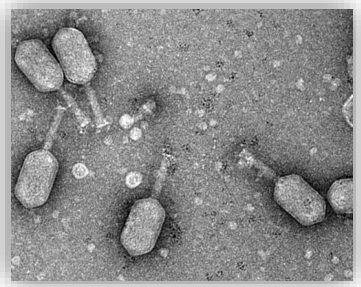




- **1949-1955 - Sanger** secuencia la **primer proteína**
  - Insulina
- **1965 -** Primer molécula de **ARNt** secuenciada (140 kg de levadura <- 1 gramo de ARNt puro)**.**
  - Se necesitaron de 5 personas trabajando durante 3 años para determinar los 76 nucleótidos.



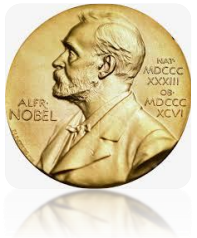
- **1970 -** Se descubren la **enzimas de restricción tipo II**
  - Cortan el ADN de manera secuencia específica de 4-6 pb.



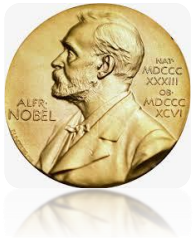
- **1972 - Walter Fiers** secuenció el **primer gen** de ARN el operon lac de *E.coli*.
- **1977 - Maxam AM, Gilbert W.** A new method for sequencing DNA. *Proc Natl Acad Sci*
- **1977 - Sanger F, Nicklen S, Coulson AR.** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci*

¿Quiénes fueron los pioneros en esta materia?

- **Frederick Sanger 1918 – 2013**

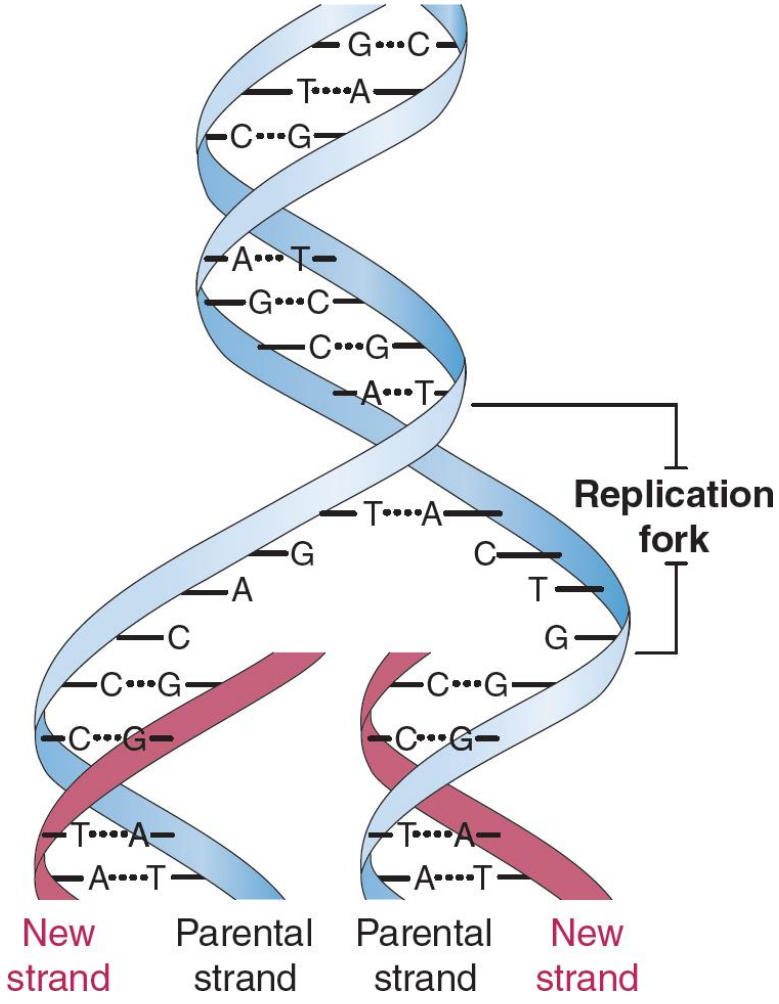


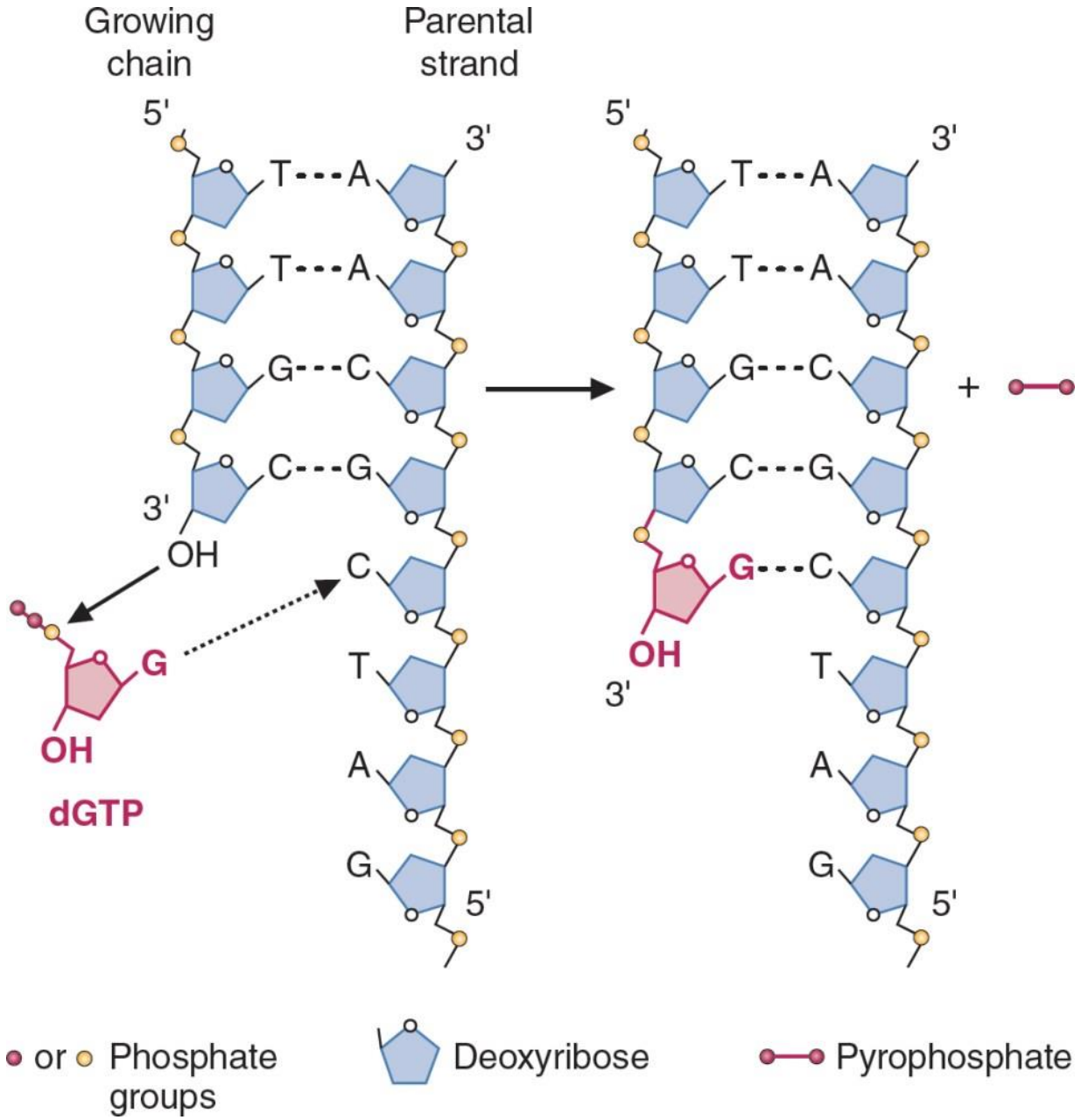
- **Walter Gilbert 1932**



  - Ganaron el premio Nobel en 1980

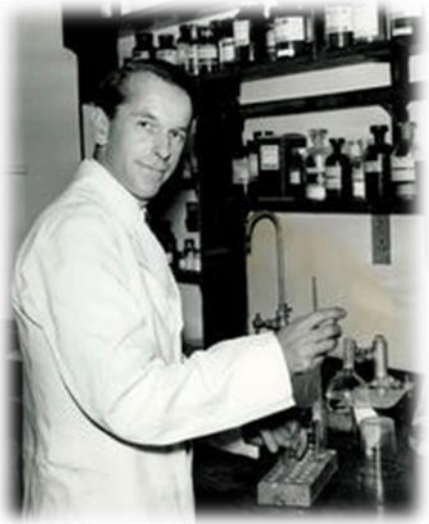
¿Cómo secuenciamos el ADN?

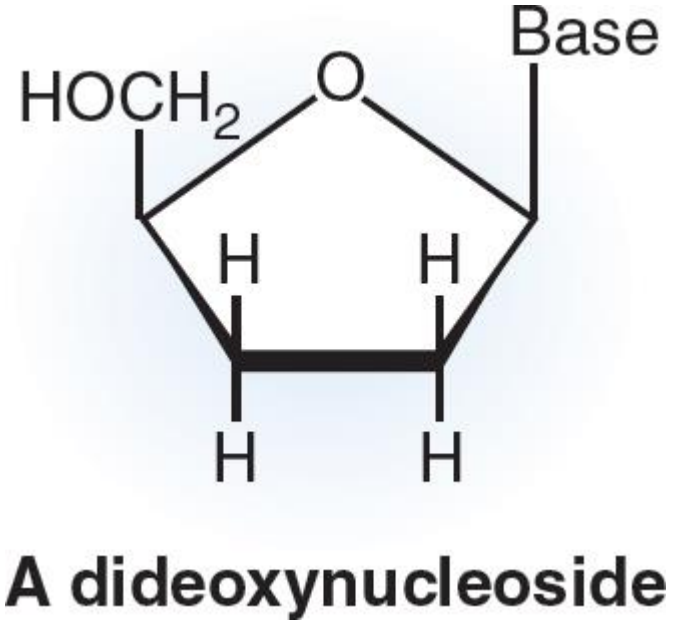




https://basicmedicalkey.com/synthesis-of-dna/

Secuenciación del ADN





A dideoxynucleoside

- **Método Enzimático de secuenciación (síntesis) - “Terminación de Cadena”**
  - **Fred Sanger (1977)**
  - **Utiliza trazas de dideoxynucleótidos para realizar la secuenciación, estos producen la interrupción enzimática de la prolongación de la cadena de ADN.**

Método de Secuenciación de Fred Sanger

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 12, pp. 5463–5467, December 1977
Biochemistry

DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage φX174)

F. SANGER, S. NICKLEN, AND A. R. COULSON

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Contributed by F. Sanger, October 3, 1977

ABSTRACT **A new method for determining nucleotide sequences in DNA is described. It is similar to the "plus and minus" method [Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* 94, 441–448] but makes use of the 2′,3′-dideoxy and arabinonucleoside analogues of the normal deoxynucleoside triphosphates, which act as specific chain-terminating inhibitors of DNA polymerase. The technique has been applied to the DNA of bacteriophage φX174 and is more rapid and more accurate than either the plus or the minus method.**

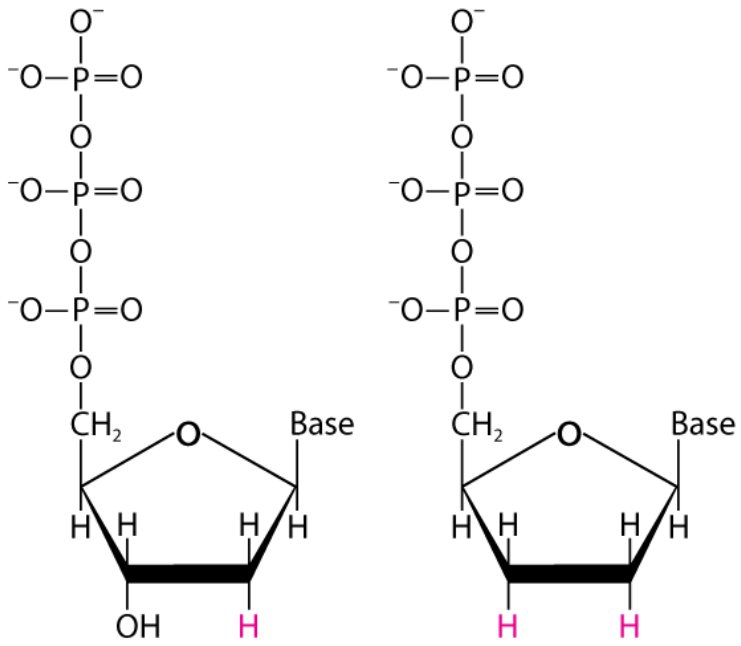
The "plus and minus" method (1) is a relatively rapid and simple technique that has made possible the determination of the sequence of the genome of bacteriophage φX174 (2). It depends on the use of DNA polymerase to transcribe specific regions of the DNA under controlled conditions. Although the method is considerably more rapid and simple than other available techniques, neither the "plus" nor the "minus" method is completely accurate, and in order to establish a sequence both must be used together, and sometimes confirmatory data are necessary. W. M. Barnes (*J. Mol. Biol.*, in press) has recently developed a third method, involving ribo-substitution, which has certain advantages over the plus and minus method, but this has not yet been extensively exploited.

Another rapid and simple method that depends on specific chemical degradation of the DNA has recently been described by Maxam and Gilbert (3), and this has also been used extensively for DNA sequencing. It has the advantage over the plus and minus method that it can be applied to double-stranded DNA, but it requires a strand separation or equivalent fractionation of each restriction enzyme fragment studied, which

a stereoisomer of ribose in which the 3′-hydroxyl group is oriented in *trans* position with respect to the 2′-hydroxyl group. The arabinosyl (ara) nucleotides act as chain terminating inhibitors of *Escherichia coli* DNA polymerase I in a manner comparable to ddT (4), although synthesized chains ending in 3′ araC can be further extended by some mammalian DNA polymerases (5). In order to obtain a suitable pattern of bands from which an extensive sequence can be read it is necessary to have a ratio of terminating triphosphate to normal triphosphate such that only partial incorporation of the terminator occurs. For the dideoxy derivatives this ratio is about 100, and for the arabinosyl derivatives about 5000.

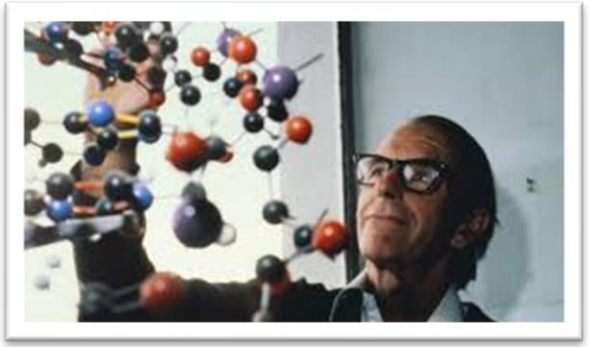
METHODS

Preparation of the Triphosphate Analogues. The preparation of ddTTP has been described (6, 7), and the material is now commercially available. ddA has been prepared by McCarthy *et al.* (8). We essentially followed their procedure and used the methods of Tener (9) and of Hoard and Ott (10) to convert it to the triphosphate, which was then purified on DEAE-Sephadex, using a 0.1–1.0 M gradient of triethylamine carbonate at pH 8.4. The preparation of ddGTP and ddCTP has not been described previously; however we applied the same method as that used for ddATP and obtained solutions having the requisite terminating activities. The yields were very low and this can hardly be regarded as adequate chemical characterization. However, there can be little doubt that the activity was due to the dideoxy derivatives.

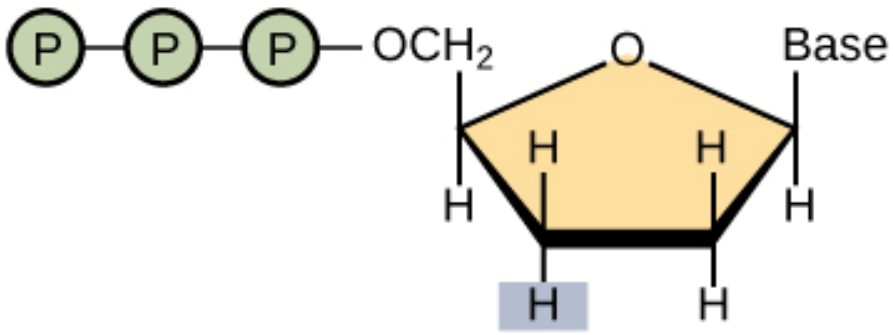


Deoxyribonucleoside triphosphate (dNTP)

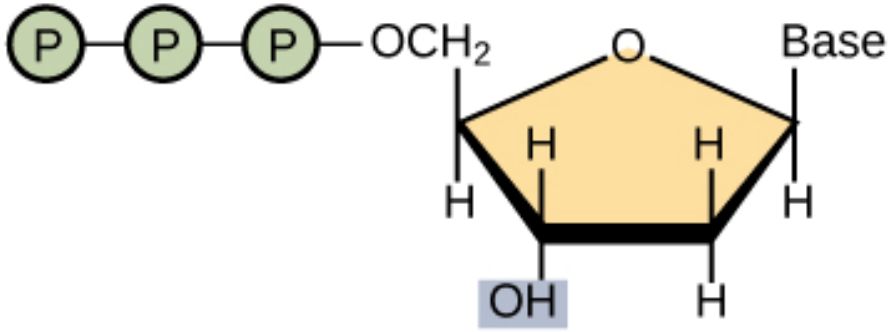
Dideoxyribonucleoside triphosphate (ddNTP)



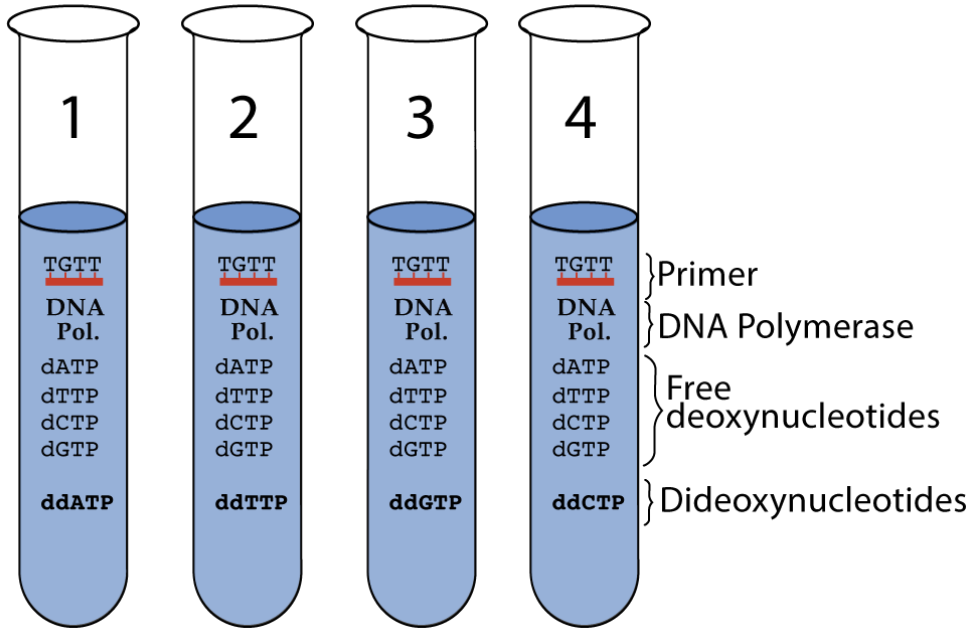
Método de Secuenciación de Fred Sanger



Dideoxynucleotide (ddNTP)

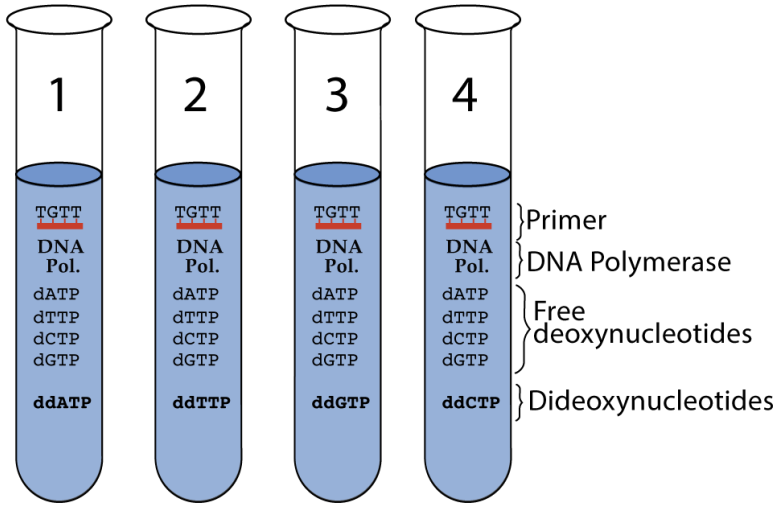


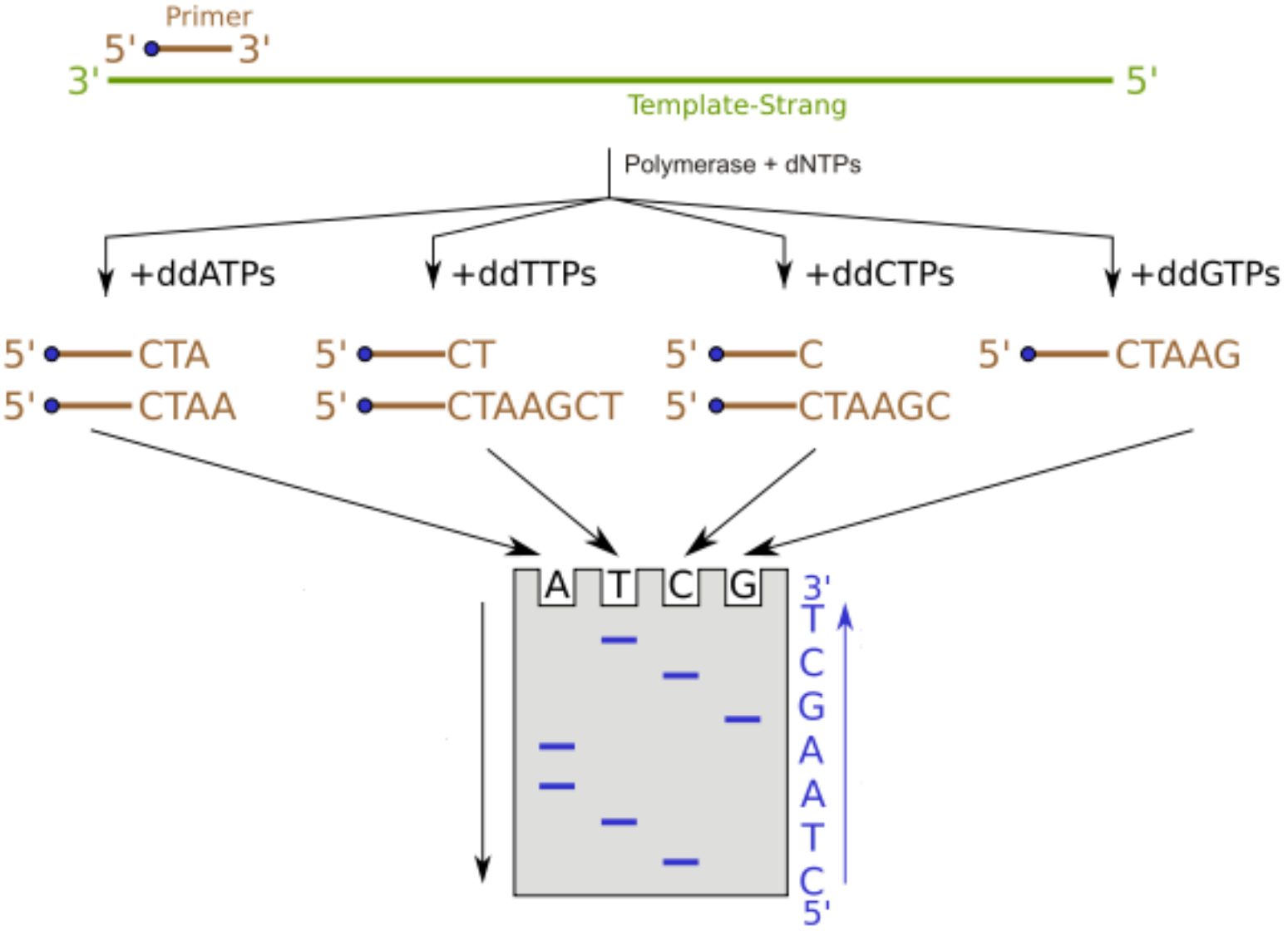
Deoxynucleotide (dNTP)



https://sciencemusicvideos.com/ap-biology/genetic-engineering-and-biotechnology/sequencing/

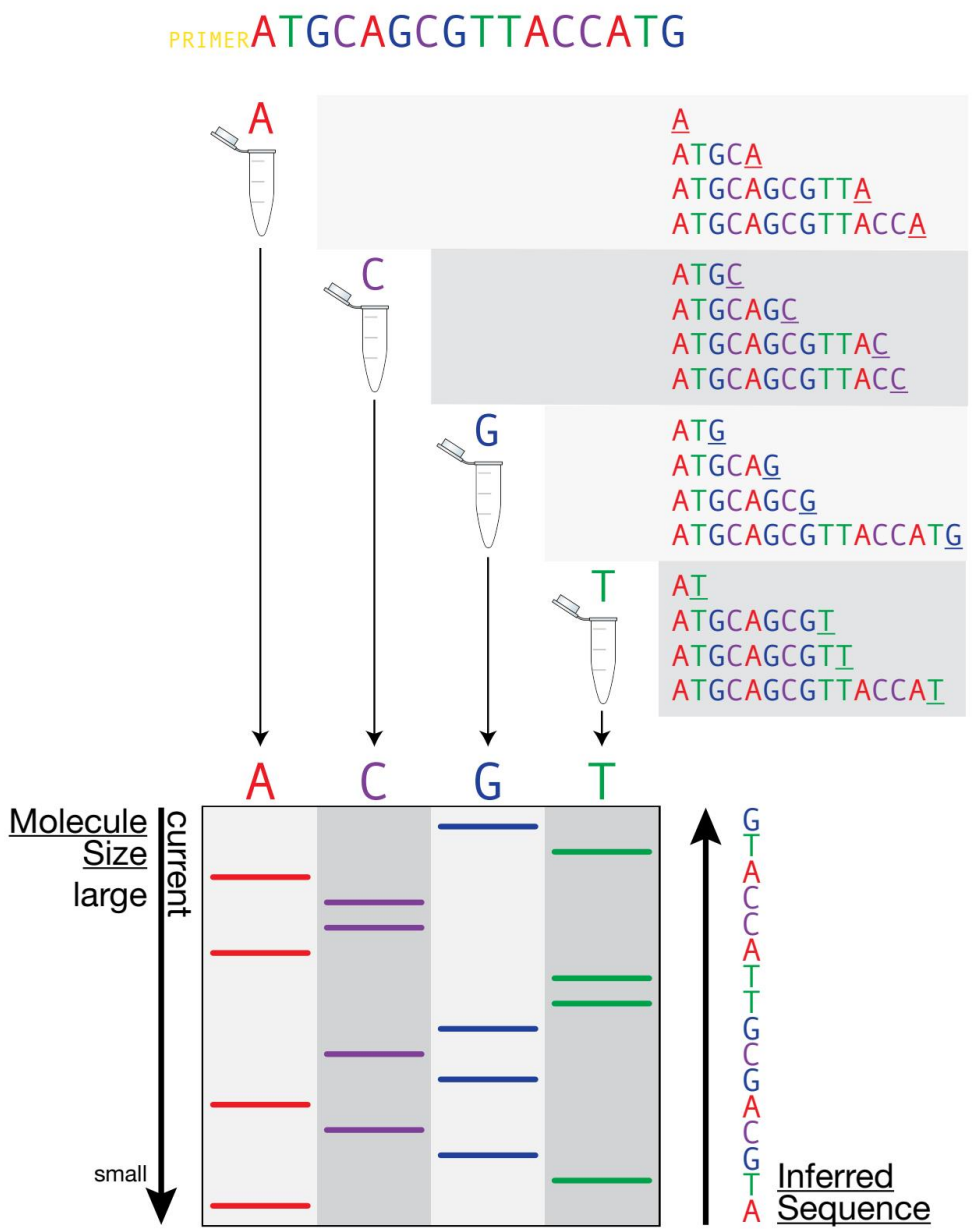
Método de Secuenciación de Fred Sanger





https://sciencemusicvideos.com/ap-biology/genetic-engineering-and-biotechnology/sequencing/

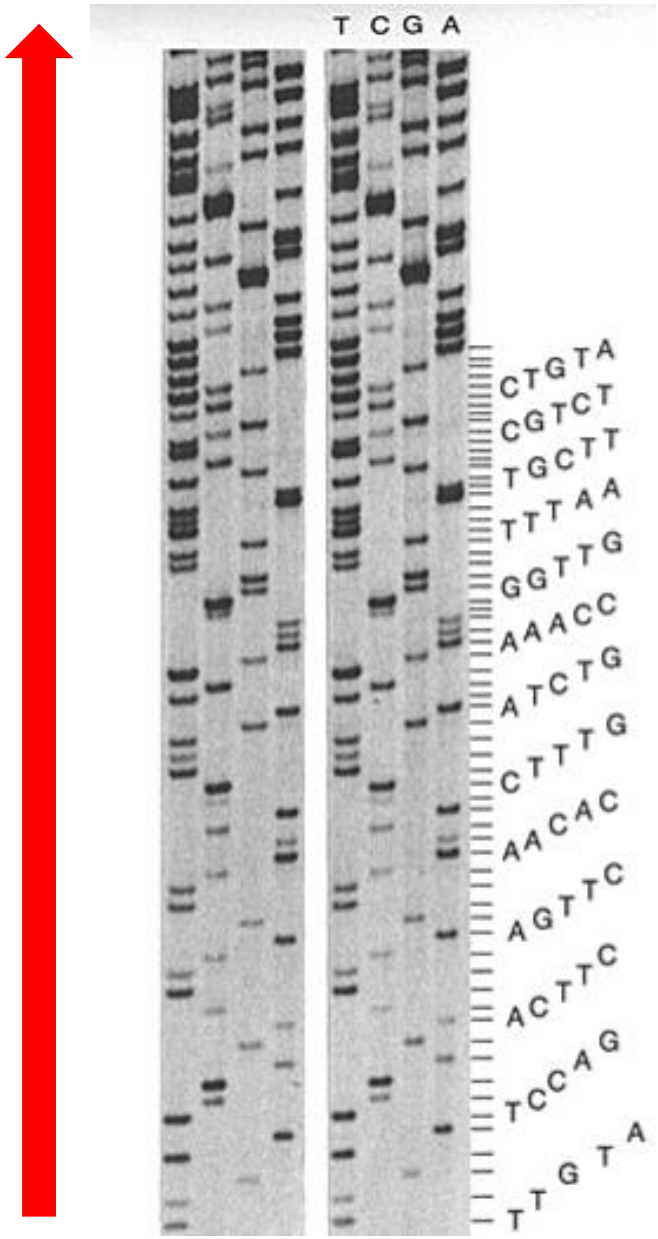
Método de Secuenciación de Fred Sanger

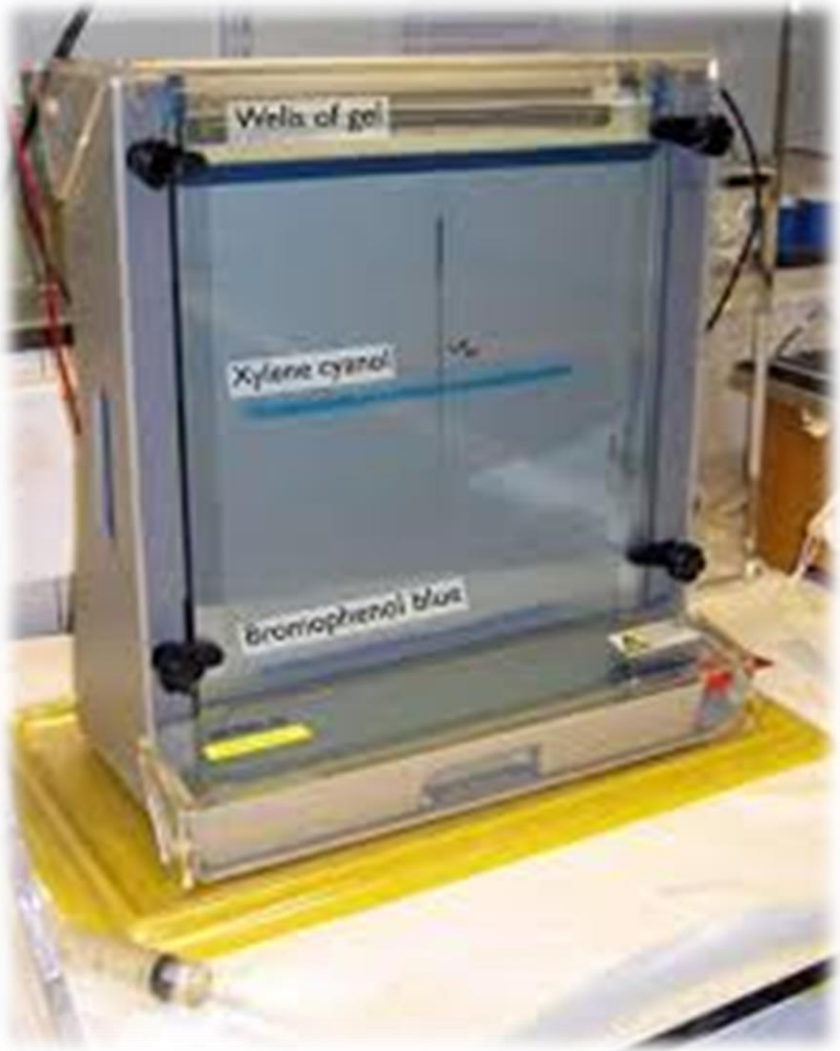


- **Cada reacción tendrá:**
  - **ADN Polimerasa**
  - **dNTPs marcados radiactivamente con fósforo 32**
  - **Trazas de un ddNTP (ddATP, ddCTP, ddGTP y ddTTP)**
  - **Oligonucleótido (primer)**

- **Cada una de las condiciones (tubos) es sometida a separación en gel de electrophoresis (poliacrilamida).**
- **Finalmente se revela el marcado en placa fotográfica.**

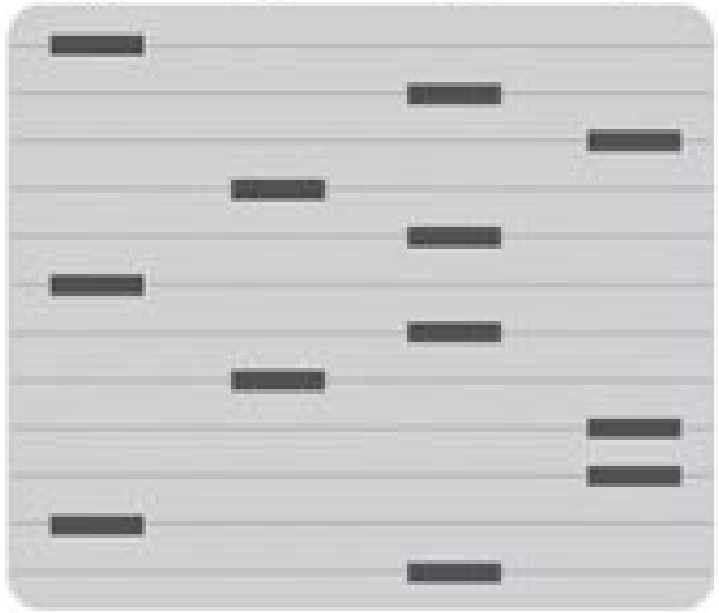
- 4 carriles uno por cada nucleótido (A, T, C y G).
- Revelado en placa fotográfica.
- Escalera de fragmentos que difieren en un solo nucleótido.
- Leemos desde abajo hacia arriba la secuencia, desde el fragmento más corto al más largo.



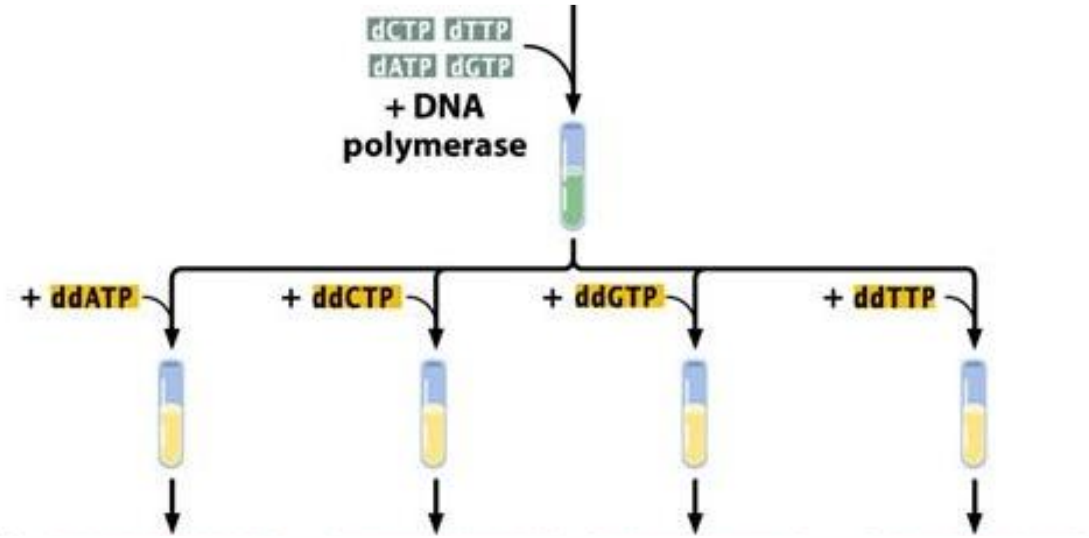


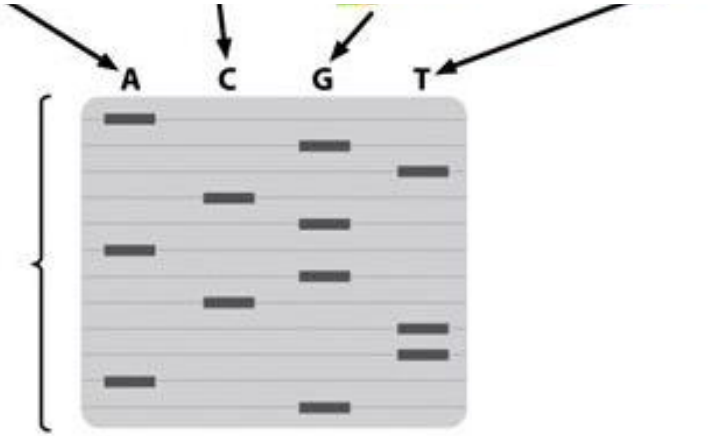
Ejercicio: secuenciación Sanger

ddNTP: **A** **C** **G** **T**



Genetics: A Conceptual Approach, Third Edition
© 2009 W. H. Freeman and Company





1. ¿Cuál es la secuencia obtenida?

A: 5’ AGTCGAGCTTAG 3’

B: 5’ CTAAGCTCGACT 3’

C: 5’ GATTCGAGCTGA 3´

D: 5’ TCAGCTCGAATC 3’

2. ¿Cuál es la secuencia molde secuenciada?

A: 5’ AGTCGAGCTTAG 3’

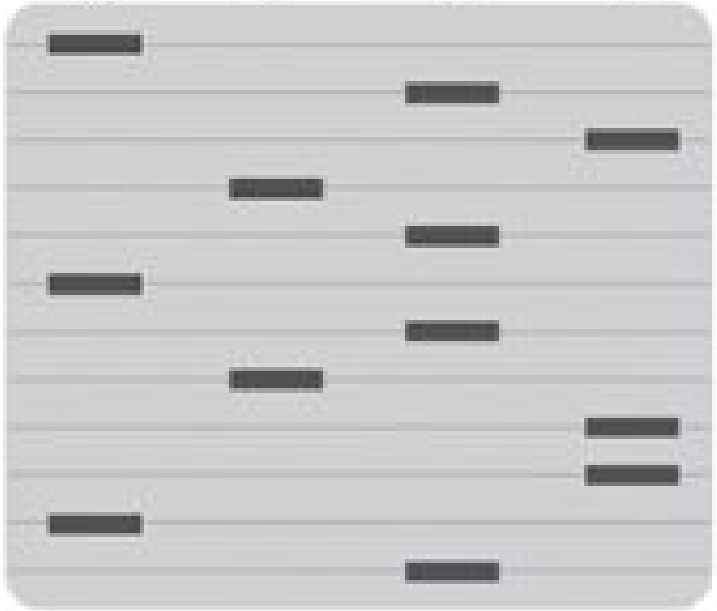
B: 5’ CTAAGCTCGACT 3’

C: 5’ GATTCGAGCTGA 3´

D: 5’ TCAGCTCGAATC 3’

Hebra molde

ddNTP: **A** **C** **G** **T**



1. ¿Cuál es la secuencia obtenida?

A: 5’ AGTCGAGCTTAG 3’

B: 5’ CTAAGCTCGACT 3’

C: 5’ GATTCGAGCTGA 3´

D: 5’ TCAGCTCGAATC 3’

2. ¿Cuál es la secuencia molde secuenciada?

A: 5’ AGTCGAGCTTAG 3’

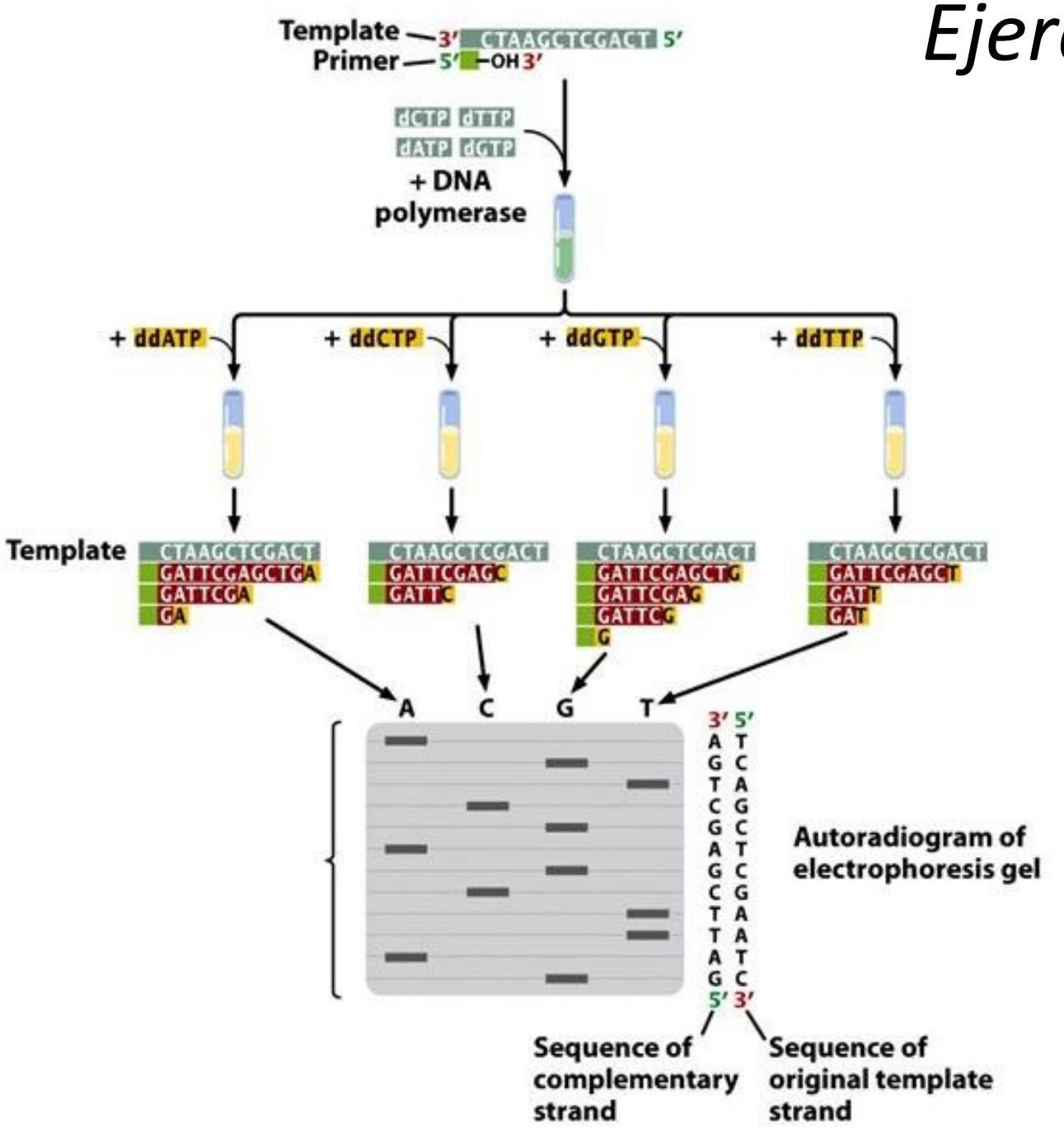
B: 5’ CTAAGCTCGACT 3’

C: 5’ GATTCGAGCTGA 3´

D: 5’ TCAGCTCGAATC 3’

Genetics: A Conceptual Approach, Third Edition
© 2009 W. H. Freeman and Company

Ejercicio: Resolución



1. ¿Cuál es la secuencia obtenida?

A: 5’ AGTCGAGCTTAG 3’

B: 5’ CTAAGCTCGACT 3’

C: 5’ GATTCGAGCTGA 3´

D: 5’ TCAGCTCGAATC 3’

2. ¿Cuál es la secuencia molde secuenciada?

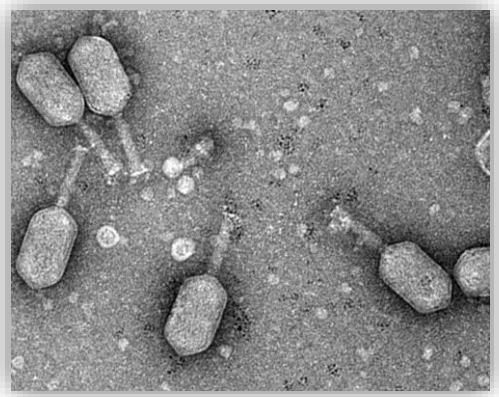
A: 5’ AGTCGAGCTTAG 3’

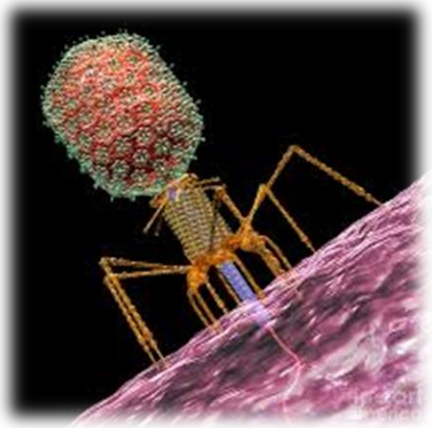
B: 5’ CTAAGCTCGACT 3’

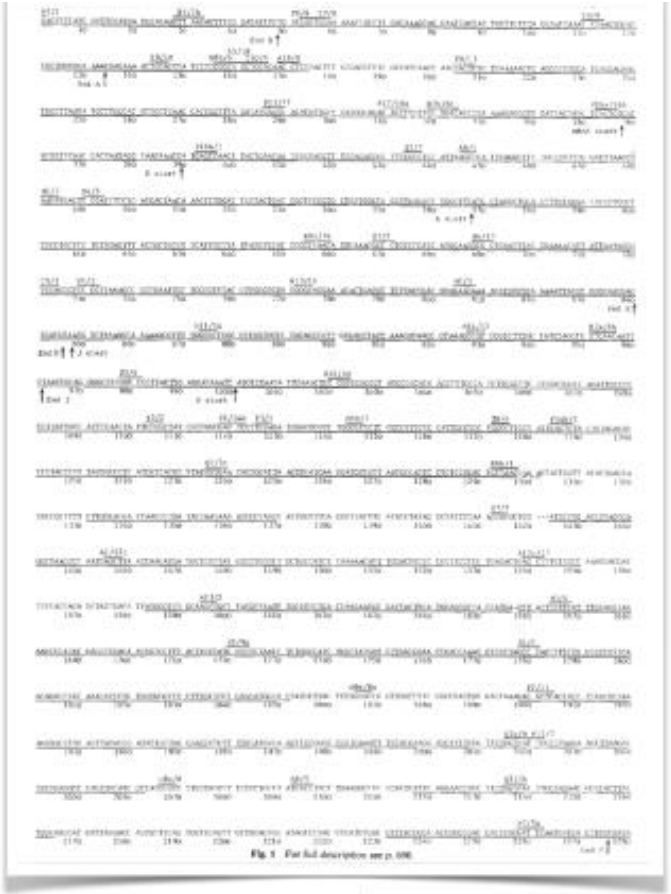
C: 5’ GATTCGAGCTGA 3´

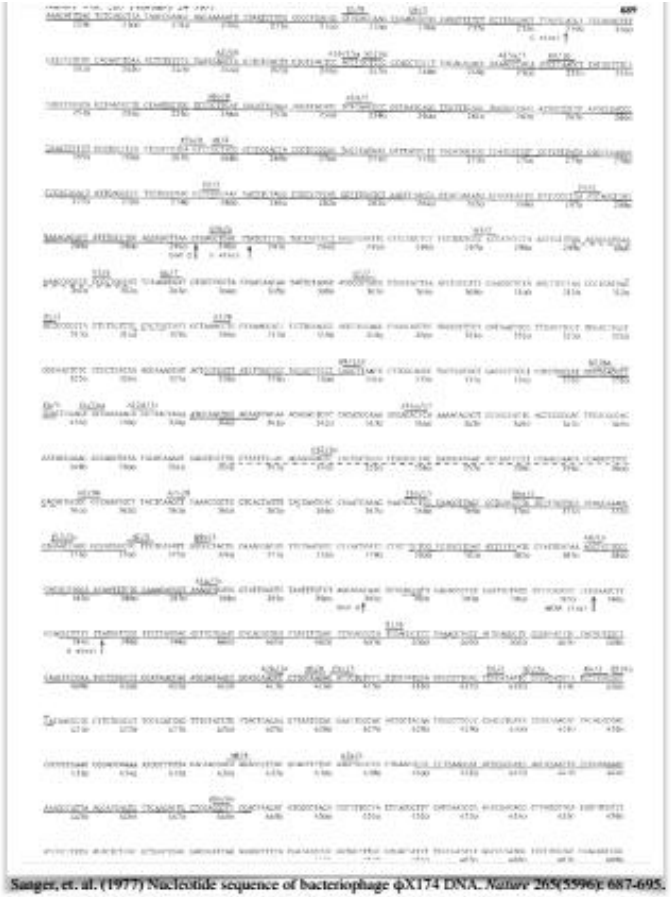
D: 5’ TCAGCTCGAATC 3’

Primer Genoma Secuenciado









Sanger, et. al. (1977) Nucleotide sequence of bacteriophage ϕX174 DNA. *Nature* 265(5596): 687-695.

ΦX Genome
1977

5.368 pb

Secuenciación del ADN

- **Secuenciación de 1^{era} generación**
  - Método de Fred Sanger
    - Método enzimático de secuenciación

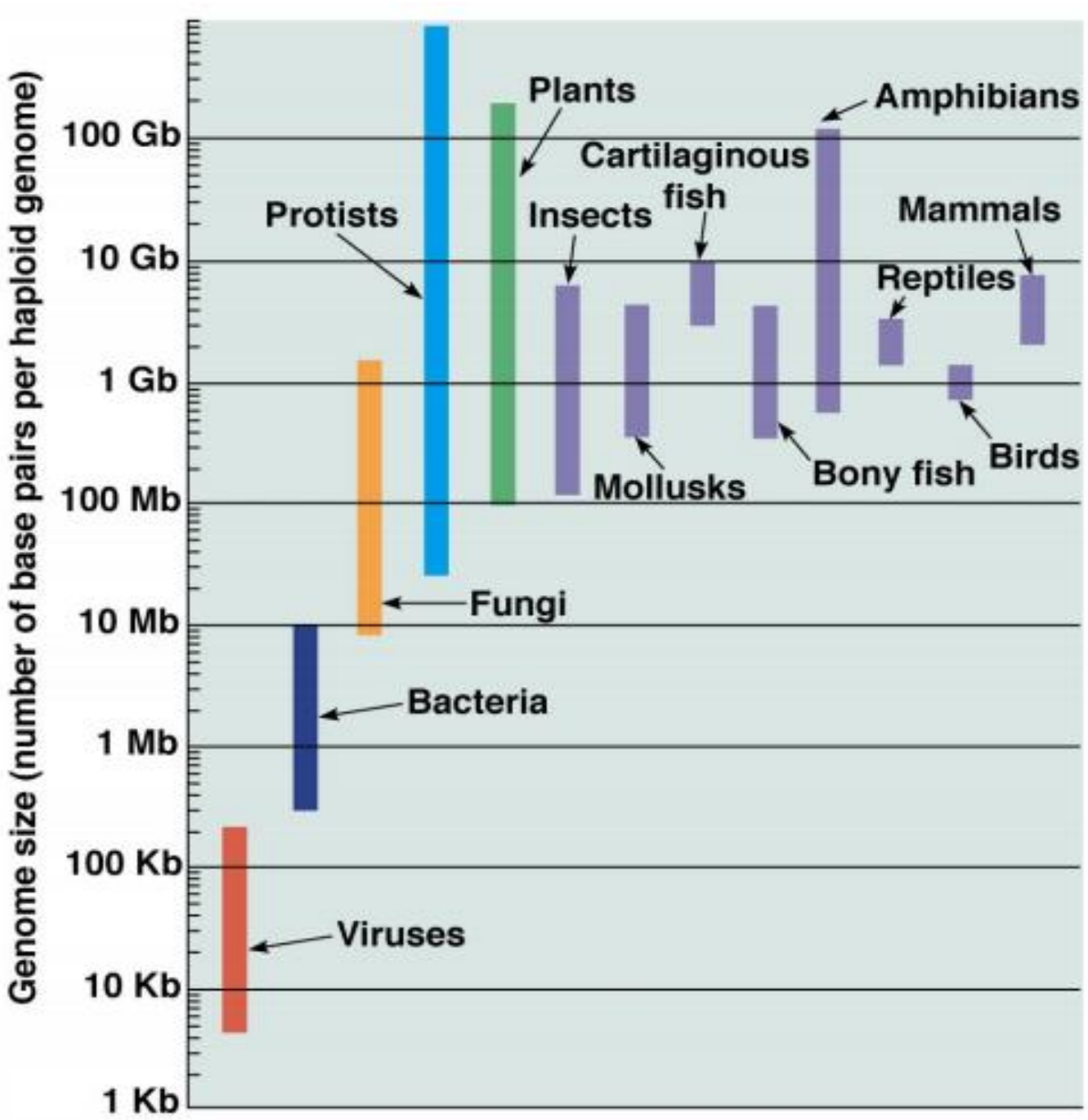
Proyecto Genoma Humano

- **Secuenciación de 1^{era} generación**

- ***Secuenciación de Nueva Generación***
  - Secuenciación de **2da y 3^{era} Generación** …

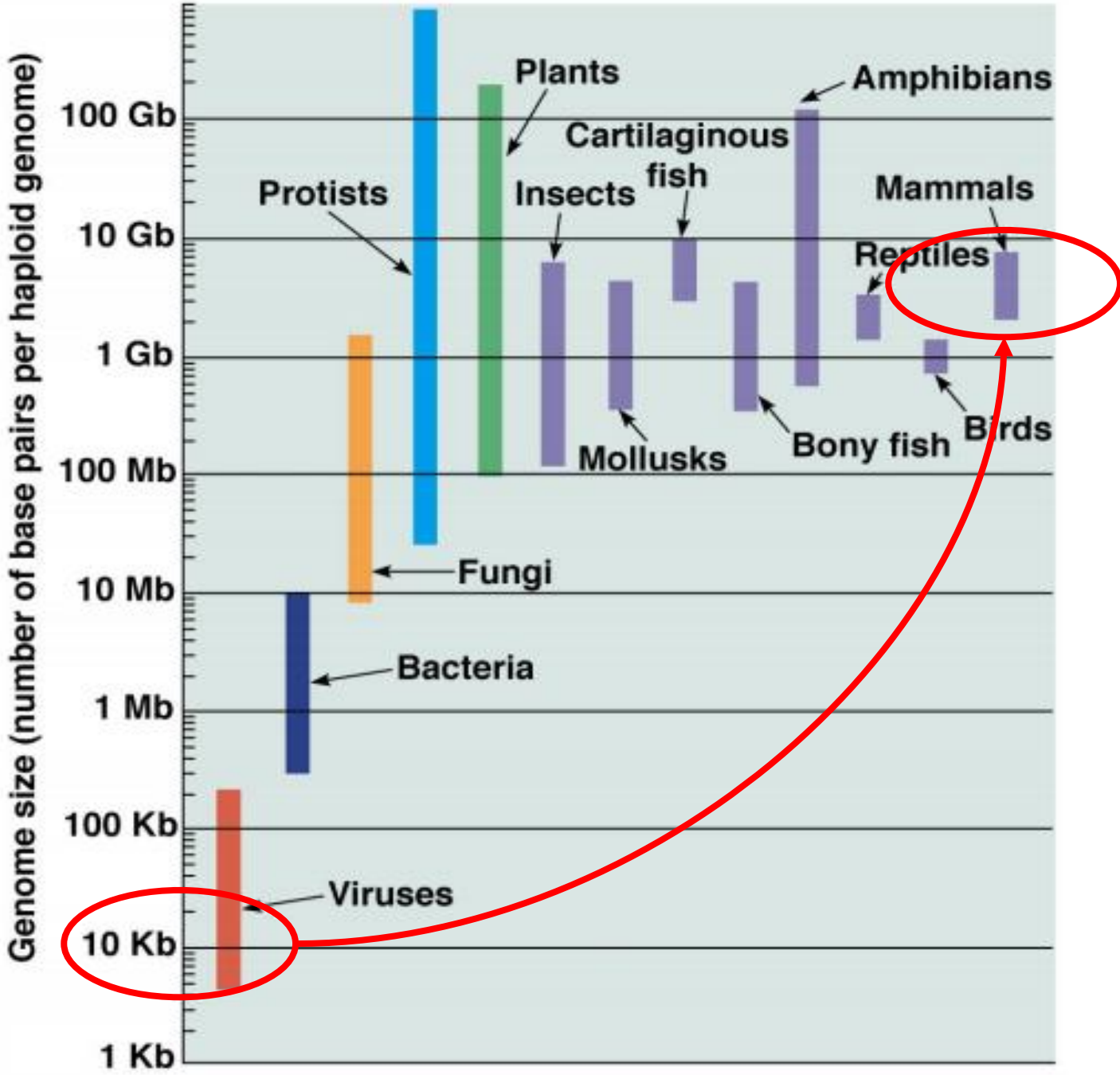
Genomas

Prefix	Symbol	Meaning	Multiple of base unit
kilo-	k	thousand	1,000
mega-	M	million	1,000,000
giga-	G	billion	1,000,000,000



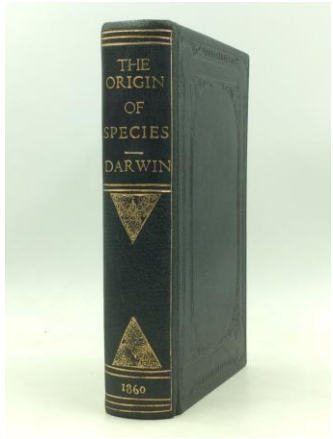
Genomas

Prefix	Symbol	Meaning	Multiple of base unit
kilo-	k	thousand	1,000
mega-	M	million	1,000,000
giga-	G	billion	1,000,000,000



Secuenciar el Genoma Humano

- **Son 3.000.000.000 pares de bases.**
- **Fuente Calibri 12 …. Palabra de 9 km de largo …**
- **200 libros aprox.**



Genoma Humano es aprox. 526.315 veces mas grande que el genoma del bacteriófago !!!

Automatización de la Secuenciación de Sanger

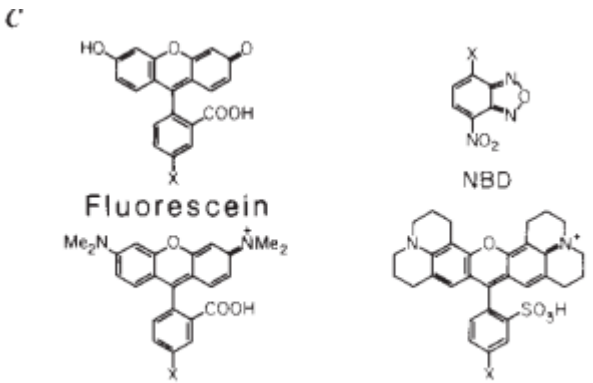
Fluorescence detection in automated DNA sequence analysis

Lloyd M. Smith, Jane Z. Sanders, Robert J. Kaiser, Peter Hughes, Chris Dodd, Charles R. Connell*, Cheryl Heiner*, Stephen B. H. Kent & Leroy E. Hood

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA
* Applied Biosystems, Inc., Foster City, California 94404, USA

directly from this information. Both these techniques are highly effective but are also very labour-intensive, fairly expensive and involve the use of radioisotopes. For these reasons, and because many genes remain to be sequenced (there are 3×10^9 bases in the human genome alone), we have undertaken the development of an automated and non-isotopic method of DNA sequence analysis.

c



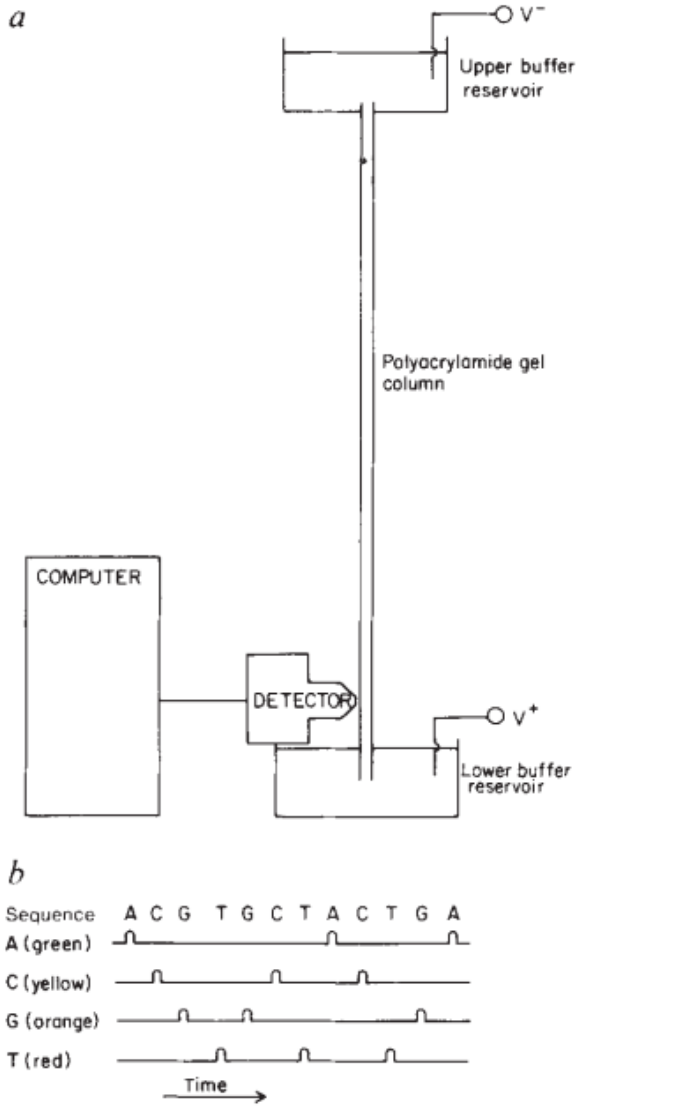
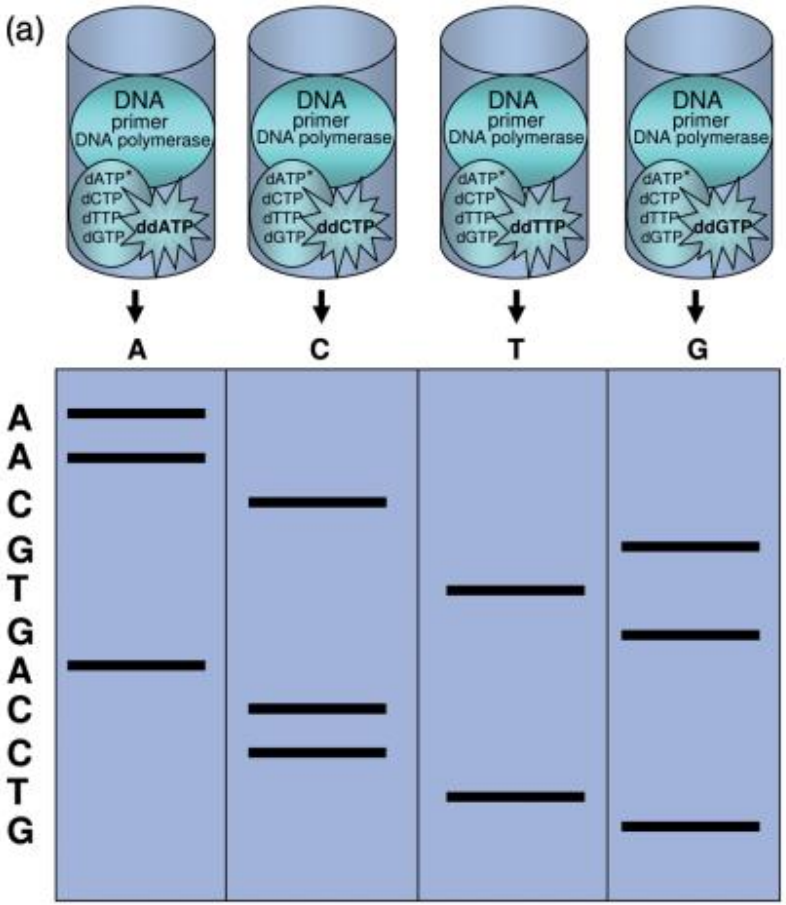


Fig. 1 *a*, Simplified diagram of the automated DNA sequenator. *b*, Idealized output from the automated DNA sequenator.

NATURE VOL. 321 12 JUNE 1986

Automatización de la Secuenciación de Sanger

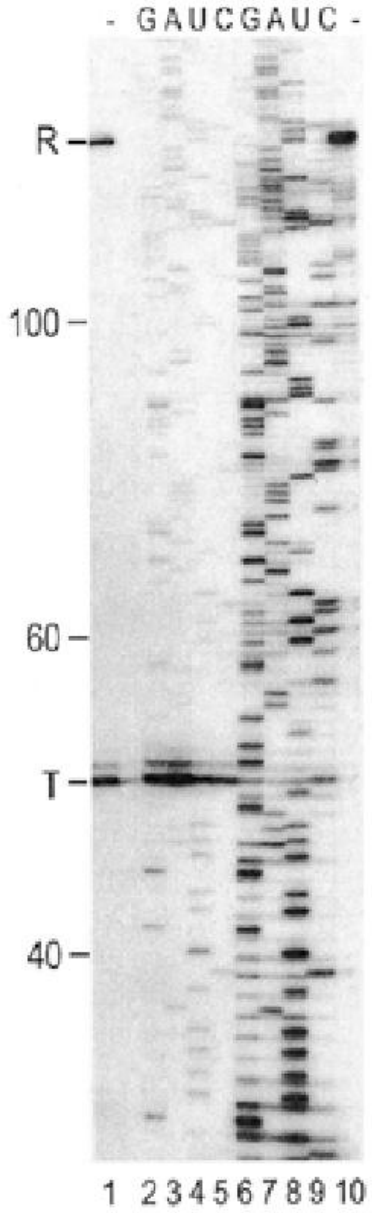


Automatización de la Secuenciación de Sanger

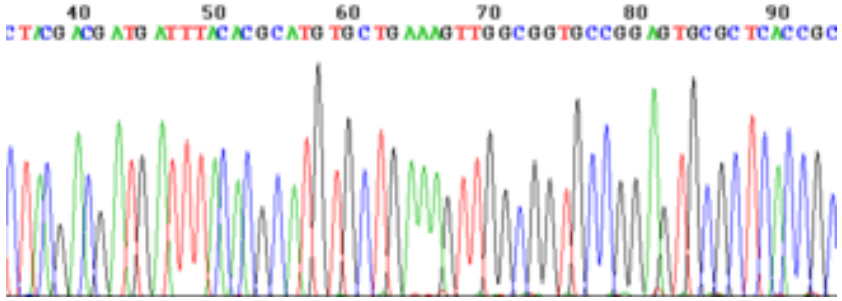
- **L. Hood and M. Hunkapiller se asociaron con Applied Biosystems, Inc. (ABI) 1987.**
- **ABI 370 (16 geles electroforesis en paralelo)**
- **200-400 pares de bases**

- **Realizaron dos grandes mejoras:**
  - **Marcación fluorescente de los ddNTPs.**
  - **Análisis informático de los datos.**

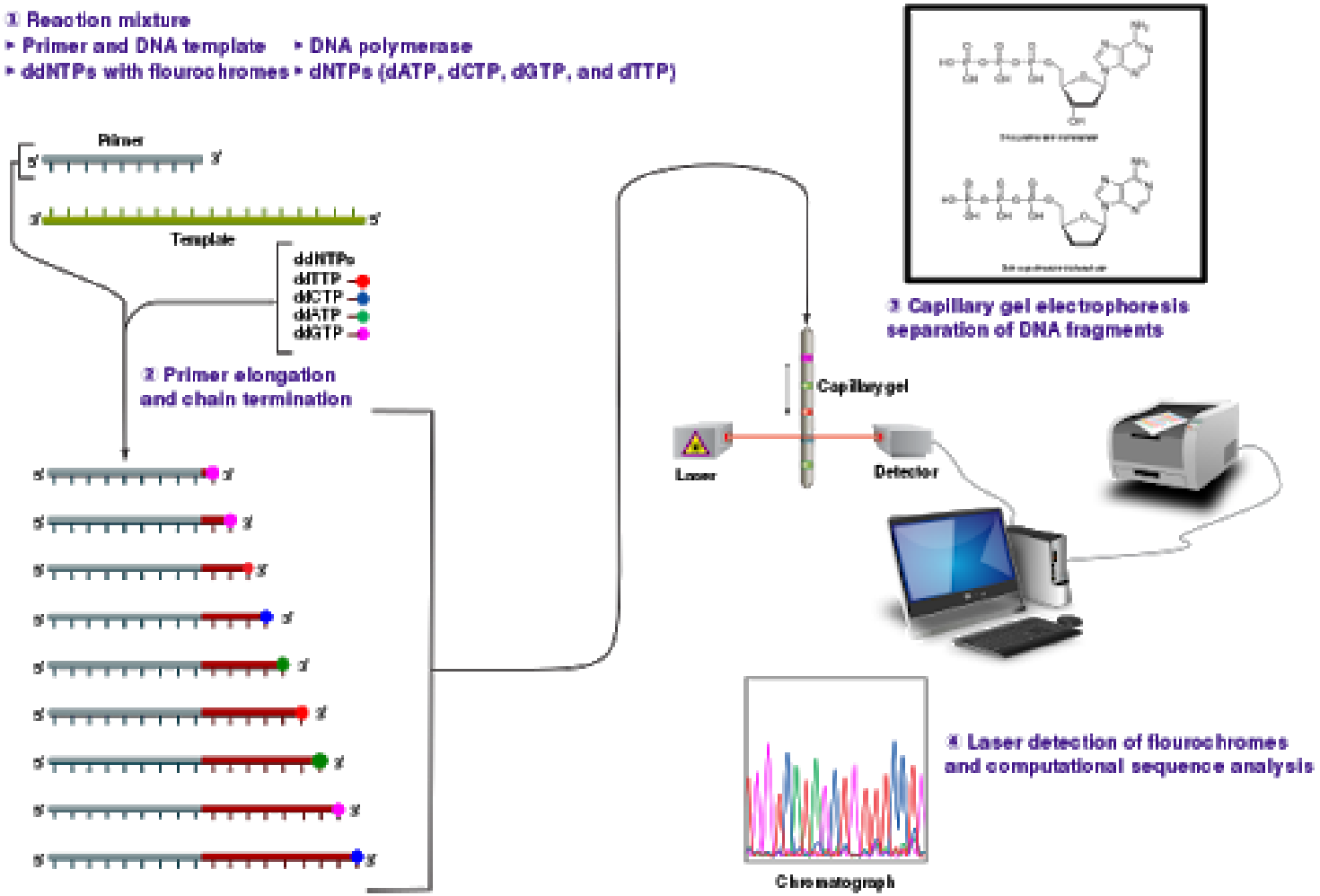
Gel de secuenciación de electroforesis



Applied Biosystems 3500 Genetic Analyzer



Applied Biosystems 3500 Genetic Analyzer



- **Largo de lecturas de 500 – 1000 pb.**
- **Paralelización de reacciones, hasta 96 capilares.**
- **Procesa desde 0.5 a 2 Mbases por día.**
  - **Bajó de 18hrs a 3hrs de corrida**
- **Costo por Mbases aprox. USD 800.**

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC30412/

Volume 4 Number 11 November 1977 Nucleic Acids Research

Sequence data handling by computer

4992–4999 Nucleic Acids Research, 1995, Vol. 23, No. 24 © 1995 Oxford University Press

A new DNA sequence assembly program

James K. Bonfield, Kathryn F. Smith and Rodger Staden*

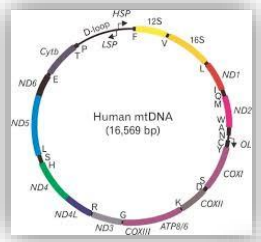
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received September 29, 1995; Revised and Accepted November 15, 1995

ties for storage, editing and analysis of both DNA and amino acid sequences.
A magnetic tape containing these programs is available on request.

- Ensamblado de secuencias al definir **solapamientos entre fragmentos de ADN**.
- Traducción de secuencias de ADN en los **6 marcos abiertos de lectura**.
- Uso de **codones**, **%GC**, definición de **sitios para enzimas de restricción**.
- El ensamblador de Staden aumentó 10 veces la capacidad de secuenciación dando lugar a la **era Genómica**.

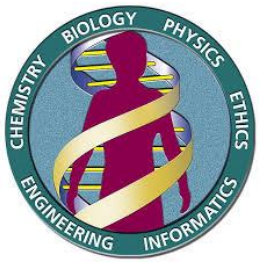
http://staden.sourceforge.net/



Automatización de la Secuenciación de Sanger

- **La secuenciación de Sanger se utilizó para la secuenciación en diferentes proyectos.**
  - **1er Genoma mitocondrial humano (1981)**
    - **Anderson S, et al. Sequence and organization of the human mitochondrial genome. Nature.**
  - **1er Genoma bacteriano (Haemophilus influenza, 1995)**
    - **Fleischmann RD, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science.**
  - **1er Genom Eucariota (Saccharomyces cerevisiae, 1996)**
    - **Goffeau A, et al. Life with 6000 genes. Science.**
  - **1er Genoma Multicelular (Caenorhabditis elegans, 1998)**
    - **TCESC. Genome sequence of the nematode C. elegans: a platform for investigating biology. Science.**
  - **1er Genoma de Planta (Arabidopsis thaliana, 2000)**
    - **TAGI. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature.**
  - **Genoma Insecto (Drosophila melanogaster, 2000)**
    - **Adams MD, et al. The genome sequence of Drosophila melanogaster. Science.**

Proyecto Genoma Humano (PGH)



- **El Proyecto Genoma Humano fue un proyecto de investigación científica con el objetivo de determinar la secuencia del ADN, identificar y cartografiar los aproximadamente 25.000 genes del genoma humano desde un punto de vista físico y funcional.**

- **El proyecto, dotado con 3000 millones de dólares, fue fundado en 1990 bajo la dirección del doctor Francis Collins. Involucró centros de investigación de los Estados Unidos, Canadá, Nueva Zelanda, Gran Bretaña y España.**
- **Se utilizó un método de secuenciación desarrollado por Frederick Sanger.**
- **Borrador inicial del genoma fue terminado en el año 2000.**
- **En febrero de 2001, las dos prestigiosas publicaciones científicas estadounidenses, Nature y Science, publicaron la secuenciación definitiva del Genoma Humano, con un 99.9% de fiabilidad.**
- **El genoma completo fue presentado en abril del 2003.**

Secuenciación Jerárquica

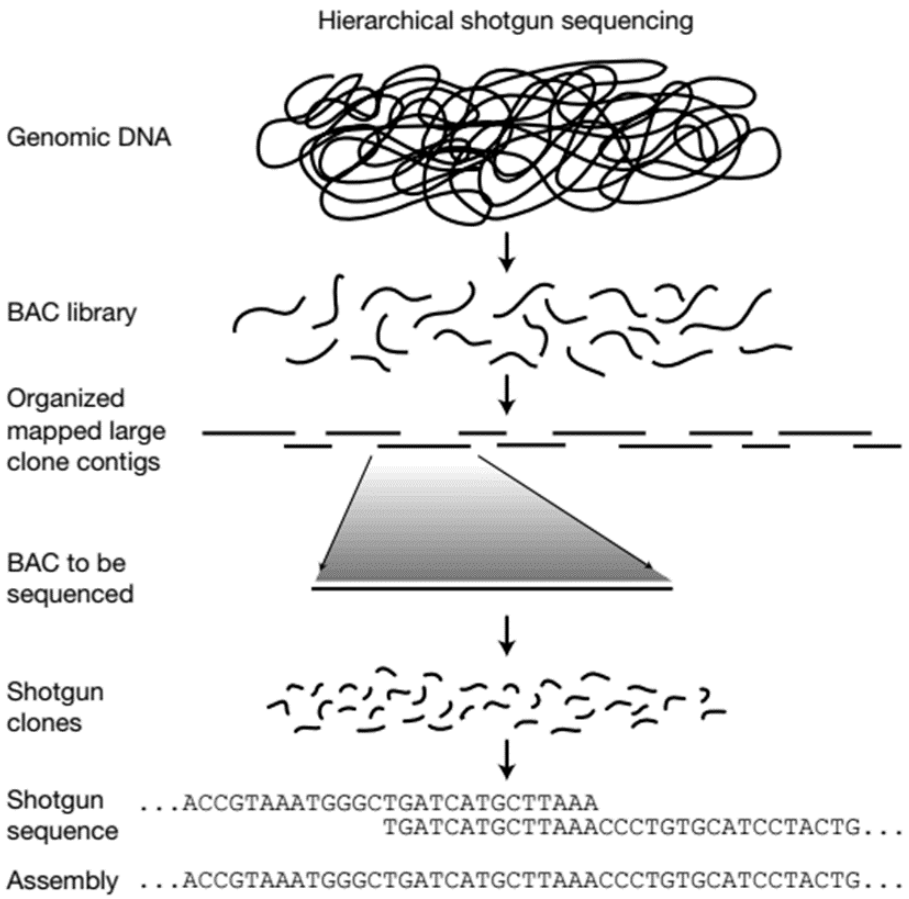


Figure 2 Idealized representation of the hierarchical shotgun sequencing strategy. A library is constructed by fragmenting the target genome and cloning it into a large-fragment cloning vector; here, BAC vectors are shown. The genomic DNA fragments represented in the library are then organized into a physical map and individual BAC clones are selected and sequenced by the random shotgun strategy. Finally, the clone sequences are assembled to reconstruct the sequence of the genome.

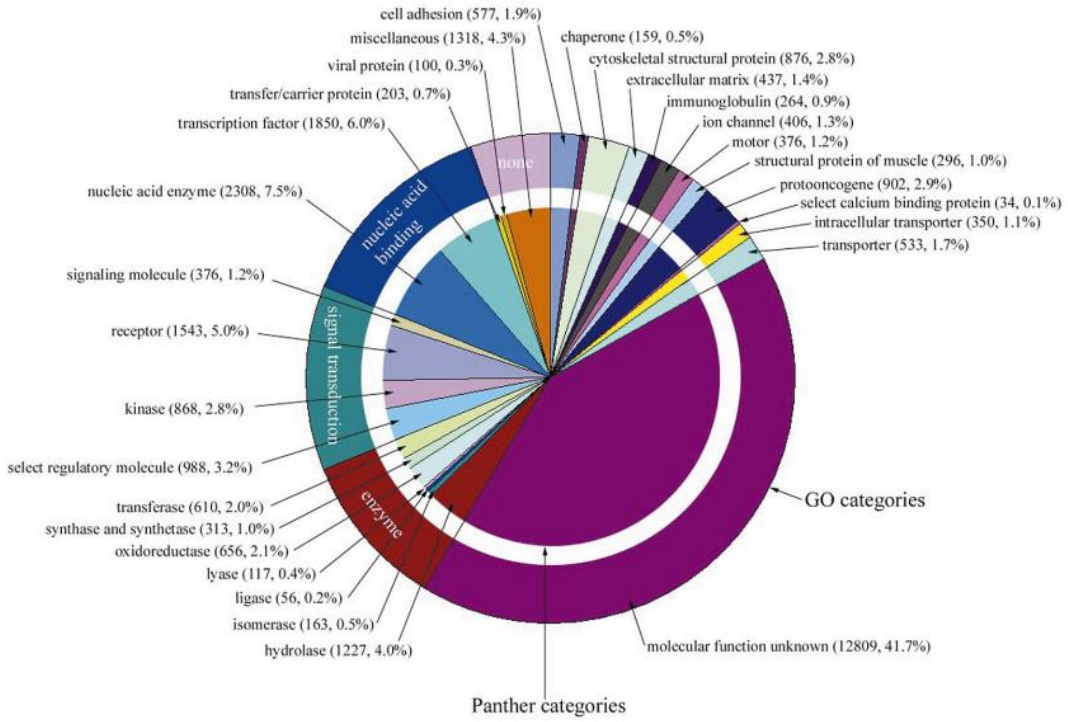
Figure 3 The automated production line for sample preparation at the Whitehead Institute, Center for Genome Research. The system consists of custom-designed factory-style conveyor belt robots that perform all functions from purifying DNA from bacterial cultures through setting up and purifying sequencing reactions.

Proyecto Genoma Humano (PGH)

articles

Initial sequencing and analysis of the human genome

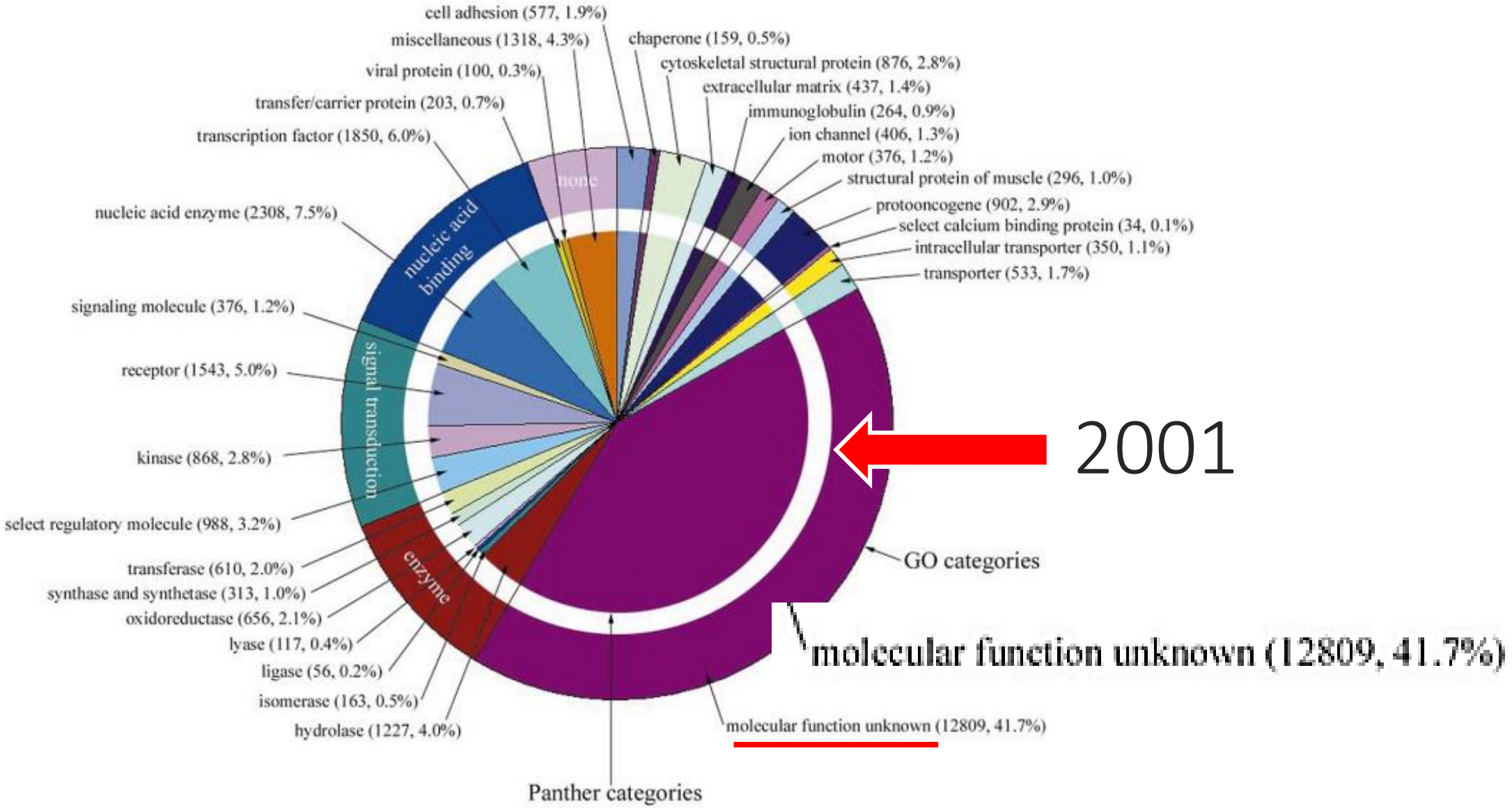
International Human Genome Sequencing Consortium*



The Sequence of the Human Genome

J. Craig Venter,[1]* Mark D. Adams,[1] Eugene W. Myers,[1] Peter W. Li,[1] Richard J. Mural,[1]
Granger G. Sutton,[1] Hamilton O. Smith,[1] Mark Yandell,[1] Cheryl A. Evans,[1] Robert A. Holt,[1]
Jeannine D. Gocayne,[1] Peter Amanatides,[1] Richard M. Ballew,[1] Daniel H. Huson,[1]
Jennifer Russo Wortman,[1] Qing Zhang,[1] Chinnappa D. Kodira,[1] Xiangqun H. Zheng,[1] Lin Chen,[1]
Marian Skupski,[1] Gangadharan Subramanian,[1] Paul D. Thomas,[1] Jinghui Zhang,[1]
George L. Gabor Miklos,[2] Catherine Nelson,[3] Samuel Broder,[1] Andrew G. Clark,[4] Joe Nadeau,[5]
Victor A. McKusick,[6] Norton Zinder,[7] Arnold J. Levine,[7] Richard J. Roberts,[8] Mel Simon,[9]
Carolyn Slayman,[10] Michael Hunkapiller,[11] Randall Bolanos,[1] Arthur Delcher,[1] Ian Dew,[1] Daniel Fasulo,[1]
Michael Flanigan,[1] Liliana Florea,[1] Aaron Halpern,[1] Sridhar Hannenhalli,[1] Saul Kravitz,[1] Samuel Levy,[1]
Clark Mobarry,[1] Knut Reinert,[1] Karin Remington,[1] Jane Abu-Threideh,[1] Ellen Beasley,[1] Kendra Biddick,[1]
Vivien Bonazzi,[1] Rhonda Brandon,[1] Michele Cargill,[1] Ishwar Chandramouliswaran,[1] Rosane Charlab,[1]
Kabir Chaturvedi,[1] Zuoming Deng,[1] Valentina Di Francesco,[1] Patrick Dunn,[1] Karen Eilbeck,[1]
Carlos Evangelista,[1] Andrei E. Gabrielian,[1] Weiniu Gan,[1] Wangmao Ge,[1] Fangcheng Gong,[1] Zhiping Gu,[1]
Ping Guan,[1] Thomas J. Heiman,[1] Maureen E. Higgins,[1] Rui-Ru Ji,[1] Zhaoxi Ke,[1] Karen A. Ketchum,[1]
Zhongwu Lai,[1] Yiding Lei,[1] Zhenya Li,[1] Jiayin Li,[1] Yong Liang,[1] Xiaoying Lin,[1] Fu Lu,[1]
Gennady V. Merkulov,[1] Natalia Milshina,[1] Helen M. Moore,[1] Ashwinikumar K Naik,[1]
Vaibhav A. Narayan,[1] Beena Neelam,[1] Deborah Nusskern,[1] Douglas B. Rusch,[1] Steven Salzberg,[12]
Wei Shao,[1] Bixiong Shue,[1] Jingtao Sun,[1] Zhen Yuan Wang,[1] Aihui Wang,[1] Xin Wang,[1] Jian Wang,[1]
Ming-Hui Wei,[1] Ron Wides,[13] Chunlin Xiao,[1] Chunhua Yan,[1] Alison Yao,[1] Jane Ye,[1] Ming Zhan,[1]
Weiqing Zhang,[1] Hongyu Zhang,[1] Qi Zhao,[1] Liansheng Zheng,[1] Fei Zhong,[1] Wenyan Zhong,[1]
Shiaoping C. Zhu,[1] Shaying Zhao,[12] Dennis Gilbert,[1] Suzanna Baumhueter,[1] Gene Spier,[1]
Christine Carter,[1] Anibal Cravchik,[1] Trevor Woodage,[1] Feroze Ali,[1] Huijin An,[1] Aderonke Awe,[1]
Danita Baldwin,[1] Holly Baden,[1] Mary Barnstead,[1] Ian Barrow,[1] Karen Beeson,[1] Dana Busam,[1]
Amy Carver,[1] Angela Center,[1] Ming Lai Cheng,[1] Liz Curry,[1] Steve Danaher,[1] Lionel Davenport,[1]
Raymond Desilets,[1] Susanne Dietz,[1] Kristina Dodson,[1] Lisa Doup,[1] Steven Ferriera,[1] Neha Garg,[1]
Andres Gluecksmann,[1] Brit Hart,[1] Jason Haynes,[1] Charles Haynes,[1] Cheryl Heiner,[1] Suzanne Hladun,[1]
Damon Hostin,[1] Jarrett Houck,[1] Timothy Howland,[1] Chinyere Ibegwam,[1] Jeffery Johnson,[1]
Francis Kalush,[1] Lesley Kline,[1] Shashi Koduru,[1] Amy Love,[1] Felecia Mann,[1] David May,[1]
Steven McCawley,[1] Tina McIntosh,[1] Ivy McMullen,[1] Mee Moy,[1] Linda Moy,[1] Brian Murphy,[1]
Keith Nelson,[1] Cynthia Pfannkoch,[1] Eric Pratts,[1] Vinita Puri,[1] Hina Qureshi,[1] Matthew Reardon,[1]
Robert Rodriguez,[1] Yu-Hui Rogers,[1] Deanna Romblad,[1] Bob Ruhfel,[1] Richard Scott,[1] Cynthia Sitter,[1]
Michelle Smallwood,[1] Erin Stewart,[1] Renee Strong,[1] Ellen Suh,[1] Reginald Thomas,[1] Ni Ni Tint,[1]
Sukyee Tse,[1] Claire Vech,[1] Gary Wang,[1] Jeremy Wetter,[1] Sherita Williams,[1] Monica Williams,[1]
Sandra Windsor,[1] Emily Winn-Deen,[1] Keriellen Wolfe,[1] Jayshree Zaveri,[1] Karena Zaveri,[1]
Josep F. Abril,[14] Roderic Guigó,[14] Michael J. Campbell,[1] Kimmen V. Sjolander,[1] Brian Karlak,[1]
Anish Kejariwal,[1] Huaiyu Mi,[1] Betty Lazareva,[1] Thomas Hatton,[1] Apurva Narechania,[1] Karen Diemer,[1]
Anushya Muruganujan,[1] Nan Guo,[1] Shinji Sato,[1] Vineet Bafna,[1] Sorin Istrail,[1] Ross Lippert,[1]
Russell Schwartz,[1] Brian Walenz,[1] Shibu Yooseph,[1] David Allen,[1] Anand Basu,[1] James Baxendale,[1]
Louis Blick,[1] Marcelo Caminha,[1] John Carnes-Stine,[1] Parris Caulk,[1] Yen-Hui Chiang,[1] My Coyne,[1]
Carl Dahlke,[1] Anne Deslattes Mays,[1] Maria Dombroski,[1] Michael Donnelly,[1] Dale Ely,[1] Shiva Esparham,[1]
Carl Fosler,[1] Harold Gire,[1] Stephen Glanowski,[1] Kenneth Glasser,[1] Anna Glodek,[1] Mark Gorokhov,[1]
Ken Graham,[1] Barry Gropman,[1] Michael Harris,[1] Jeremy Heil,[1] Scott Henderson,[1] Jeffrey Hoover,[1]
Donald Jennings,[1] Catherine Jordan,[1] James Jordan,[1] John Kasha,[1] Leonid Kagan,[1] Cheryl Kraft,[1]
Alexander Levitsky,[1] Mark Lewis,[1] Xiangjun Liu,[1] John Lopez,[1] Daniel Ma,[1] William Majoros,[1]
Joe McDaniel,[1] Sean Murphy,[1] Matthew Newman,[1] Trung Nguyen,[1] Ngoc Nguyen,[1] Marc Nodell,[1]
Sue Pan,[1] Jim Peck,[1] Marshall Peterson,[1] William Rowe,[1] Robert Sanders,[1] John Scott,[1]
Michael Simpson,[1] Thomas Smith,[1] Arlan Sprague,[1] Timothy Stockwell,[1] Russell Turner,[1] Eli Venter,[1]
Mei Wang,[1] Meiyuan Wen,[1] David Wu,[1] Mitchell Wu,[1] Ashley Xia,[1] Ali Zandieh,[1] Xiaohong Zhu[1]

Proyecto Genoma Humano (PGH)



2001

Proyecto Genoma Humano (PGH)

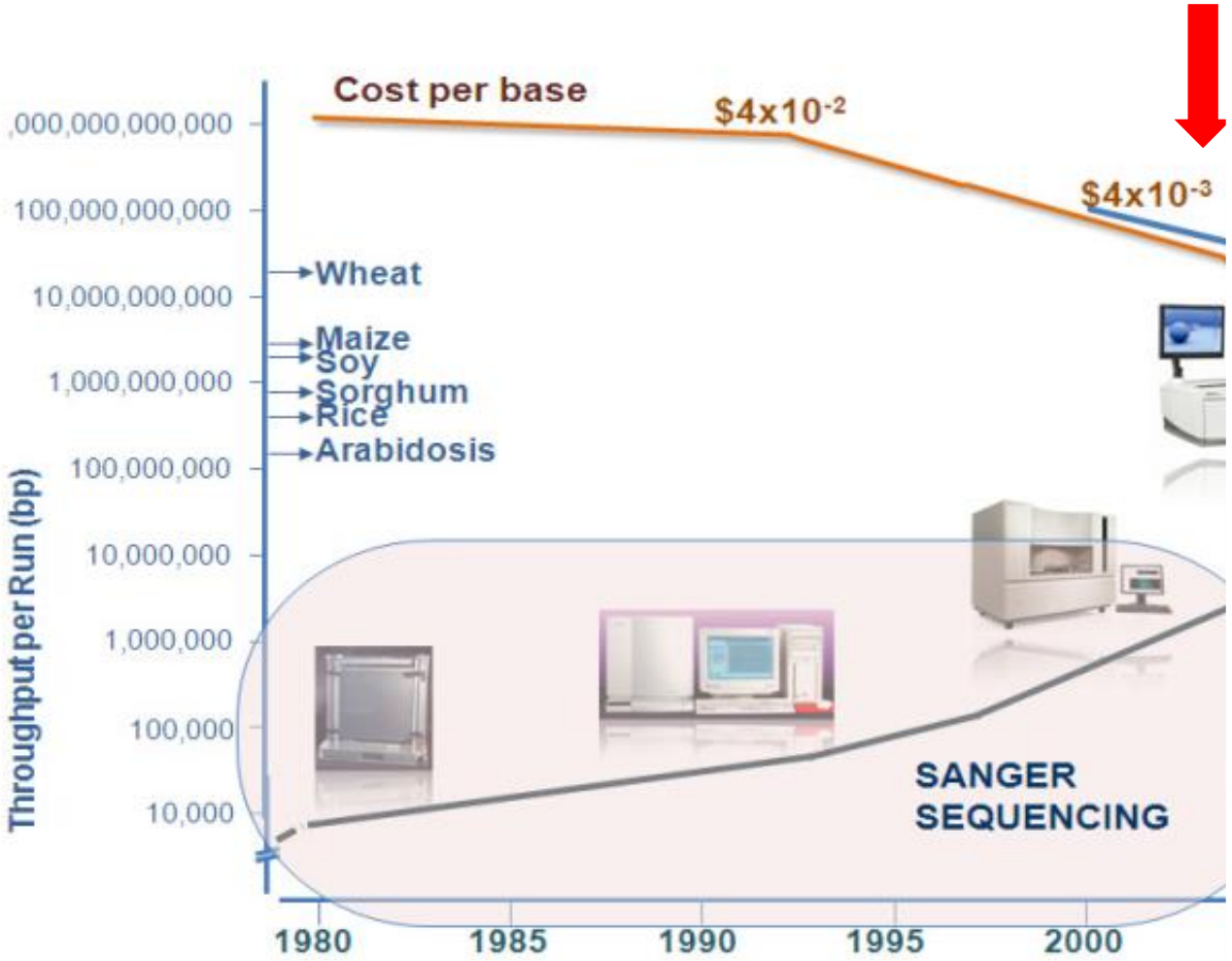
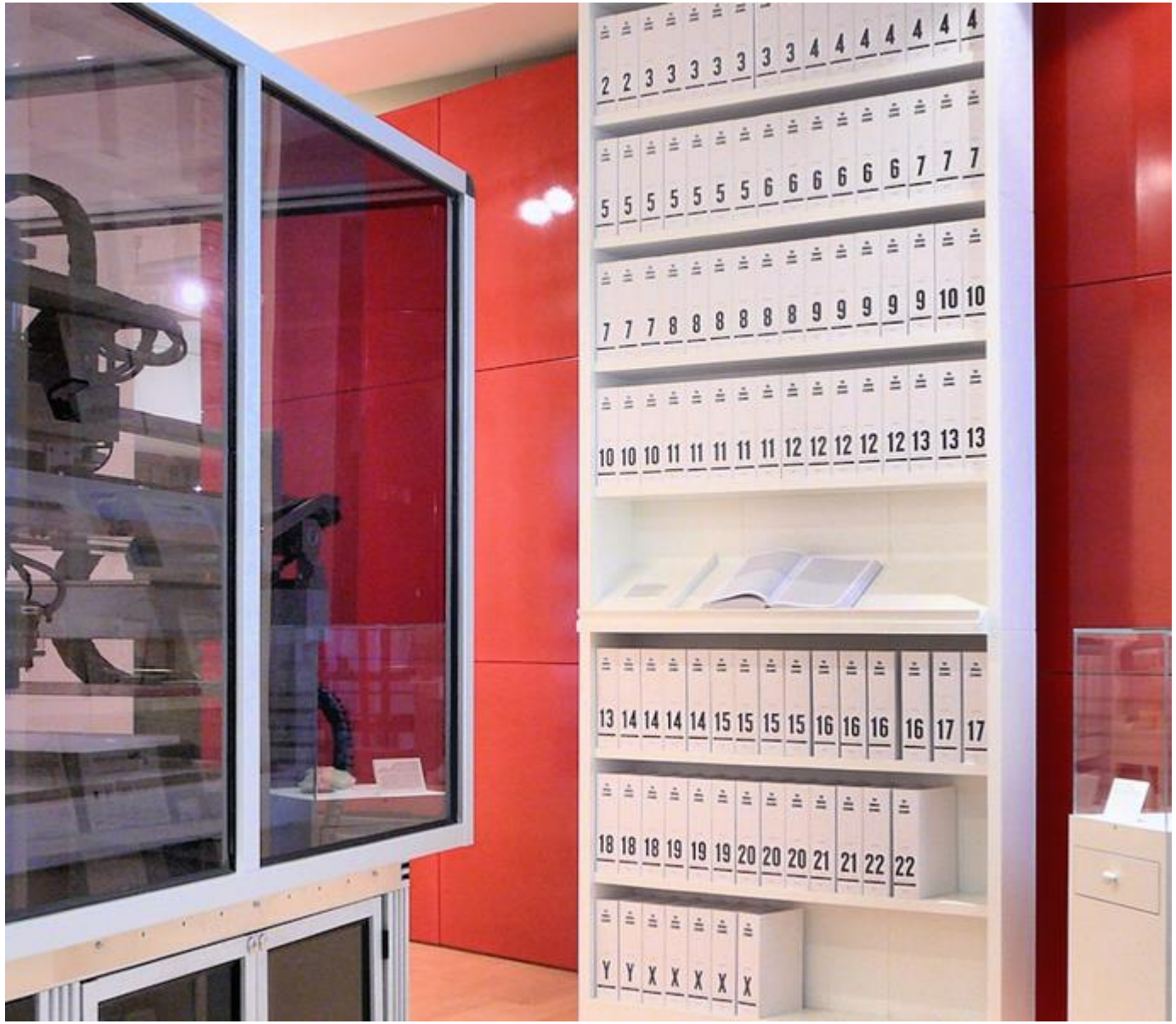
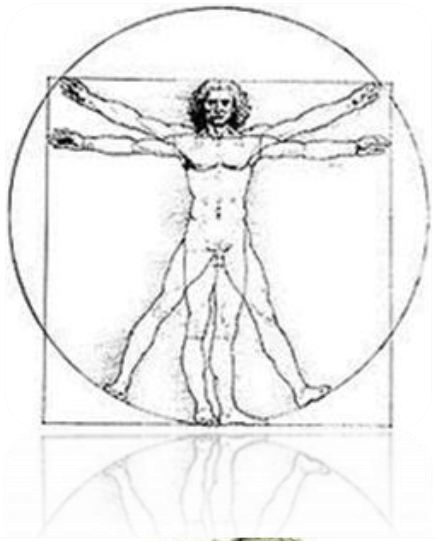


Fig. 1. Increase in maximum throughput per run in sequencing platforms from 1980 to 2011. Throughput per run based on Stratton *et al* (2009) and Glenn (2011).

Victor Llaca (April 20th 2012). Methods and Applications, IntechOpen, DOI: 10.5772/37918

Proyecto Genoma Humano (PGH)



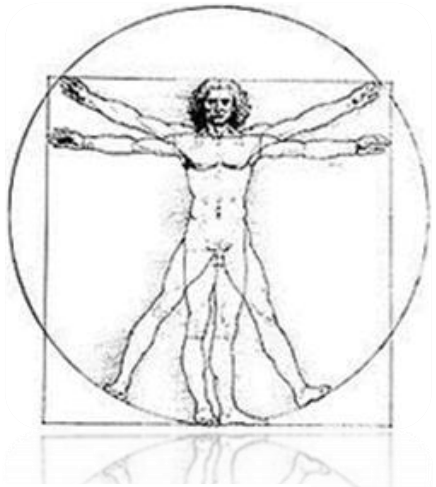


3000 millones de dólares

Proyecto Genoma Humano (PGH)

How important is R&D for economic growth?

By Heather Hall | November 15, 2019



El Proyecto del Genoma Humano, consorcio que trazó un mapa del genoma humano en 1990-2001 (inversión: 3000 millones de dólares). **Por cada dólar invertido**, se estima que **hubo un retorno de 141 dólares** en nuevos medicamentos, productos, servicios y empleo.

https://www.rdworldonline.com/how-important-is-rd-for-economic-growth/

Proyecto Genoma Humano (PGH)

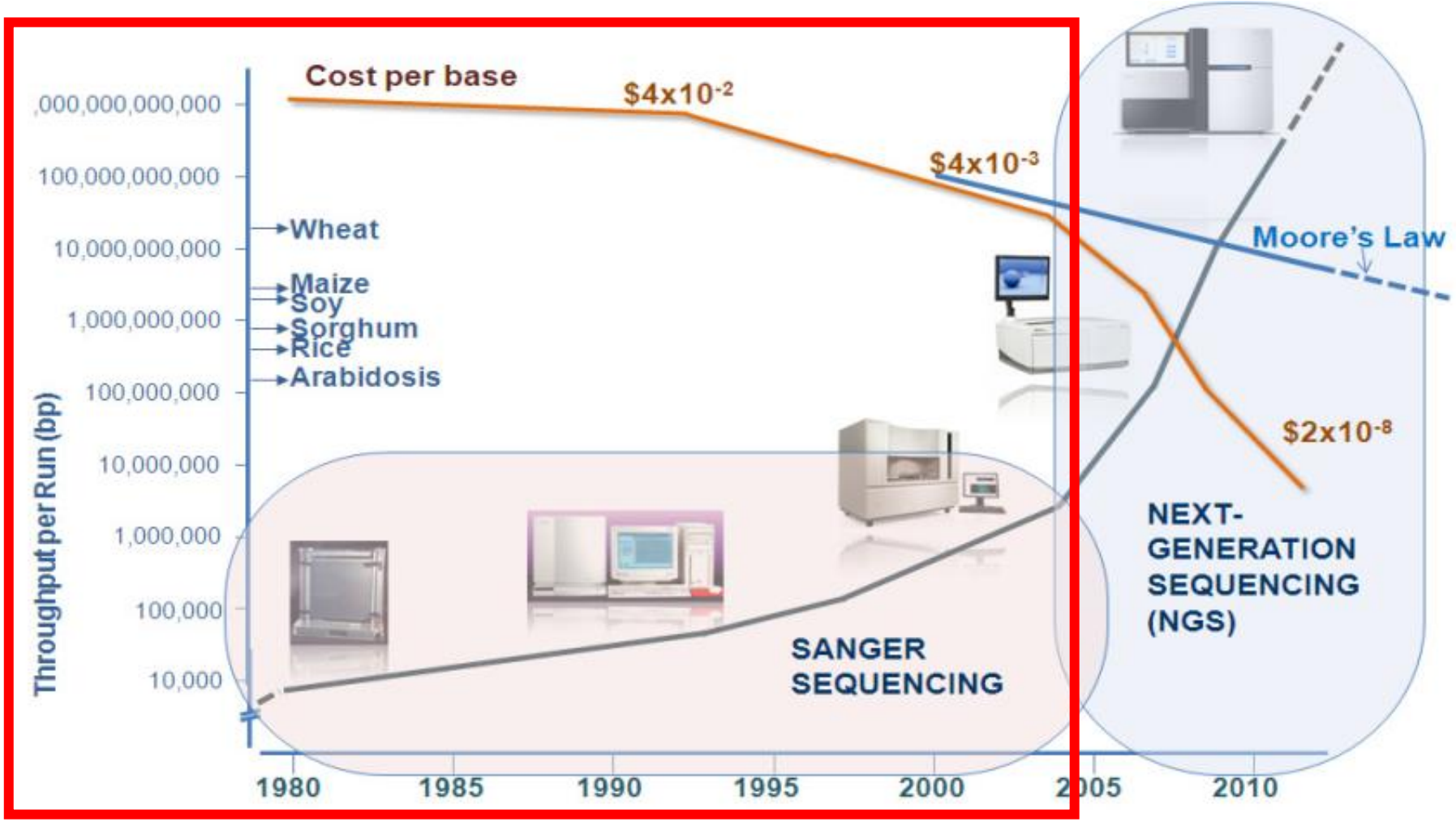


Fig. 1. Increase in maximum throughput per run in sequencing platforms from 1980 to 2011. Throughput per run based on Stratton *et al* (2009) and Glenn (2011).

Victor Llaca (April 20th 2012). Methods and Applications, IntechOpen, DOI: 10.5772/37918

GENE SEQUENCING

The Race for the $1000 Genome

Fast, cheap genetic analyses will soon become a reality, and the consequences—good and bad—will affect everybody

17 MARCH 2006 VOL 311 **SCIENCE** www.sciencemag.org

National Human Genome Research Institute (NHGRI)

How to get genomes at one ten-thousandth the cost

Jeffery A Schloss

The NHGRI's Advanced DNA Sequencing Technology program is spearheading the development of platforms that will bring routine whole-genome sequencing closer to reality.

NATURE BIOTECHNOLOGY VOLUME 26 NUMBER 10 OCTOBER 2008

The research community has responded enthusiastically to the initial and subsequent calls for proposals; more than 160 grant applications have been submitted to date. In a total of five rounds of research grant awards, the NHGRI has committed more than $100 million to 50 research teams. The funded projects have

La Nueva Generación . . .

Secuenciación del ADN

- **Secuenciación de 1^{era} generación**
  - Método de Fred Sanger
    - Método enzimático de secuenciación

Proyecto Genoma Humano

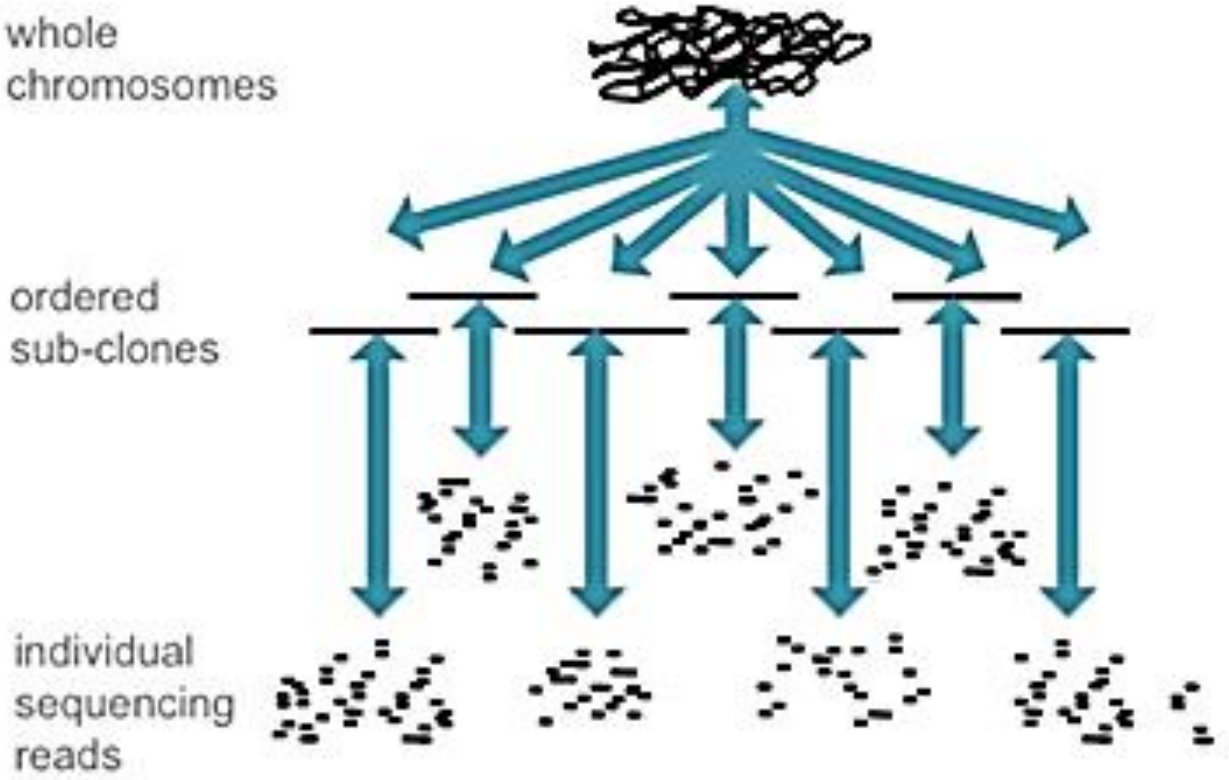
- Secuenciación de 1^{era} generación

- ***Secuenciación de Nueva Generación***
  - **Secuenciación de 2da** y 3^{era} Generación …

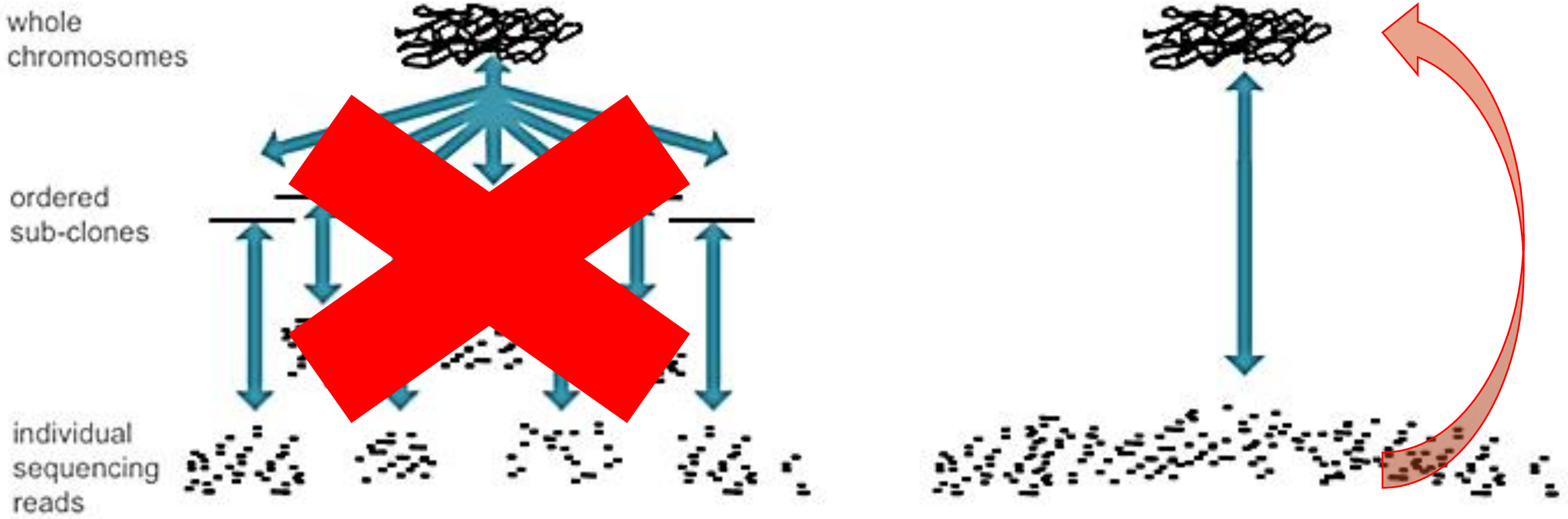
¿Qué trae de “nuevo” la nueva generación?

- El término **secuenciación de nueva generación** (**NGS**) hace referencia a las tecnologías de **secuenciación de ADN de alto rendimiento no basadas en Sanger**. **Millones o billones** de **hebras de ADN** pueden ser **secuenciadas en paralelo**.

NGS vs Sanger Sec

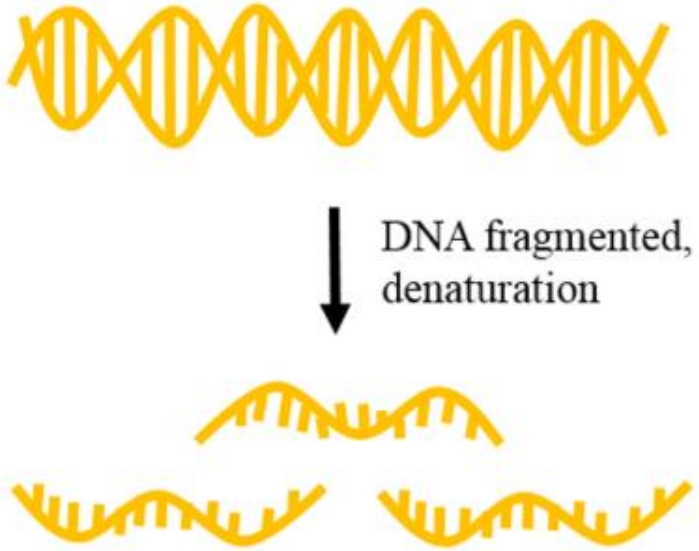


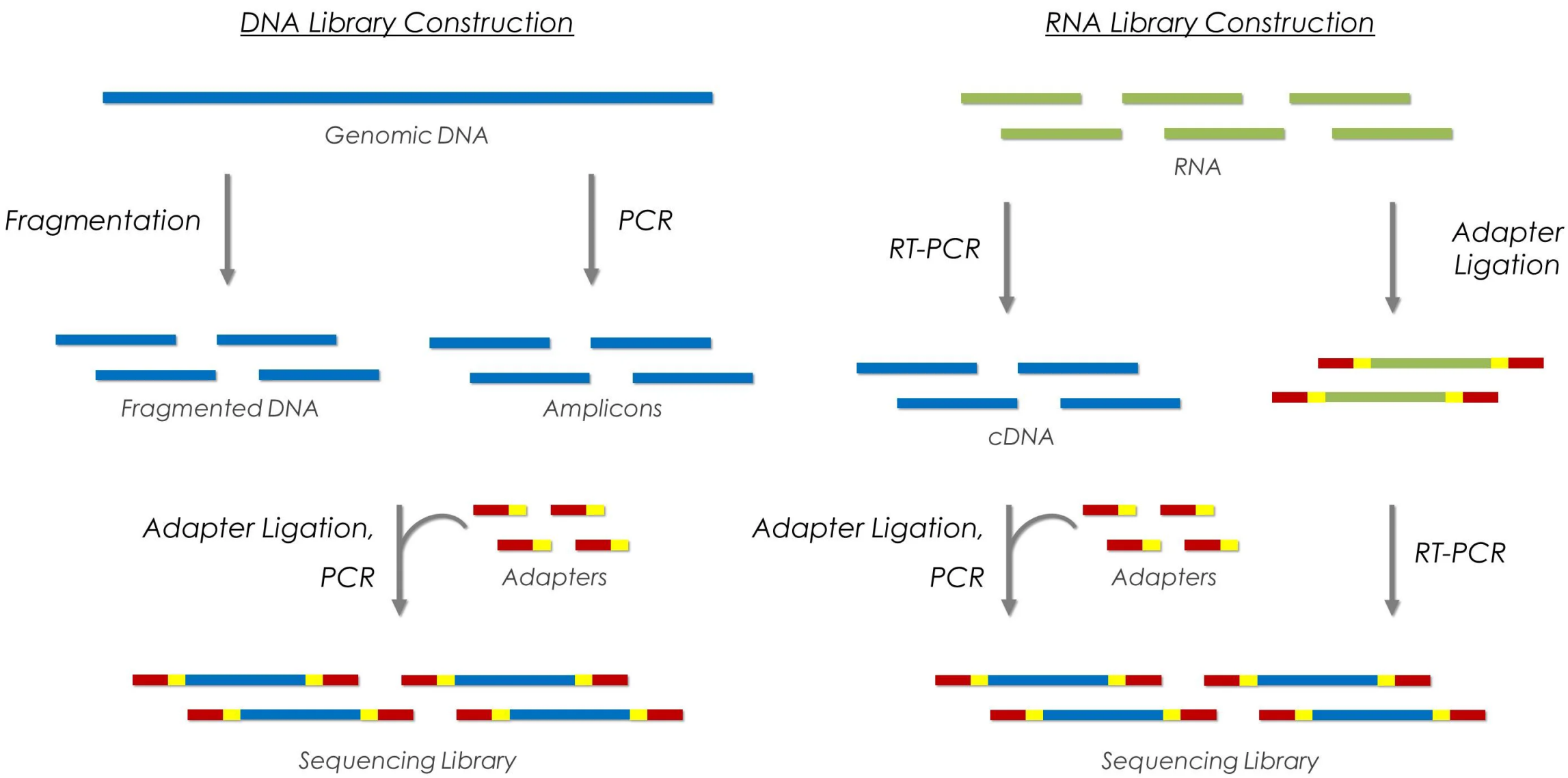
NGS vs Sanger Sec



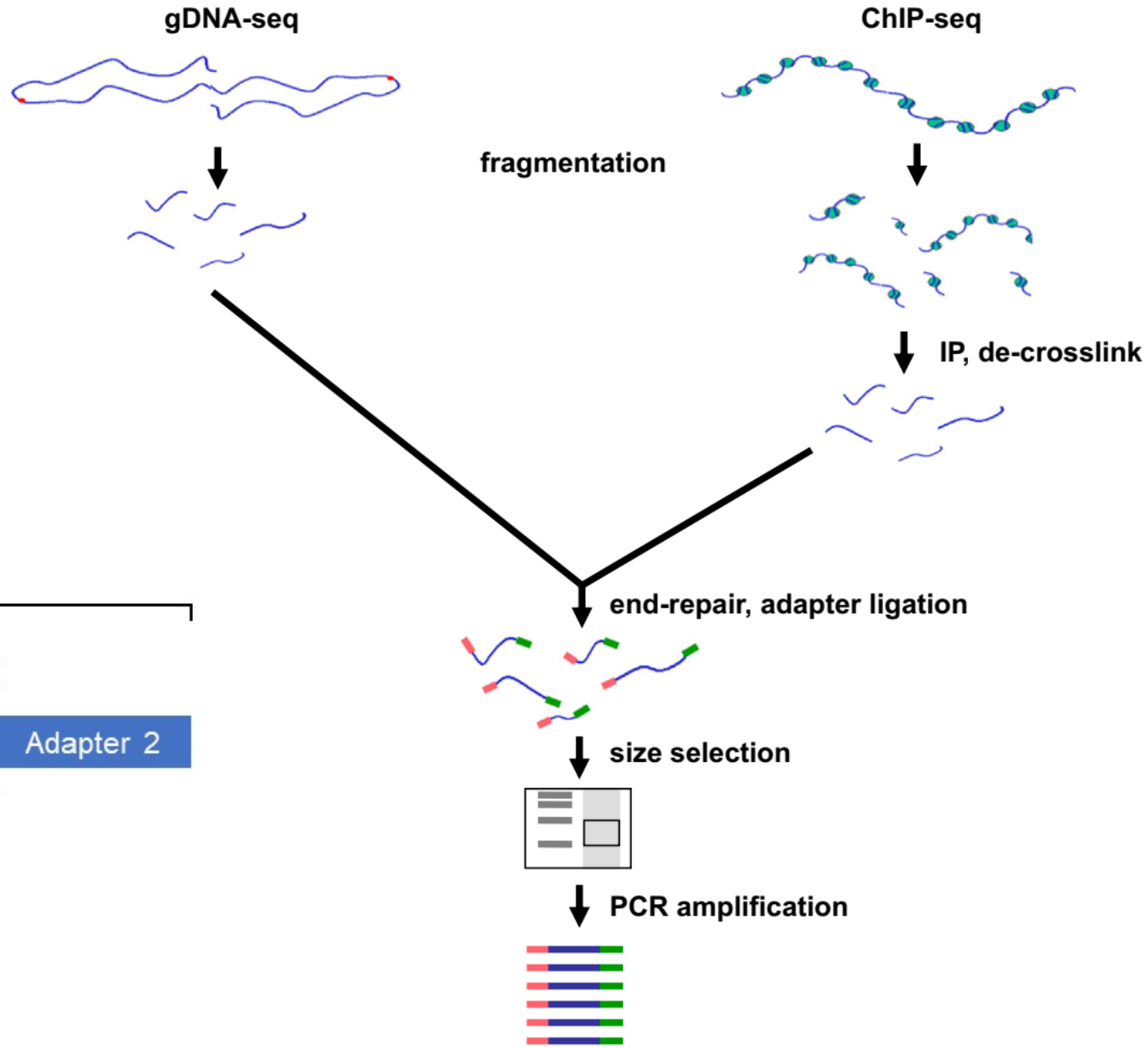
Sin embargo, una desventaja de las **tecnologías NGS** fue sus **lecturas relativamente cortas (35-100 pb)**, lo que requirió el desarrollo de nuevos algoritmos de alineación

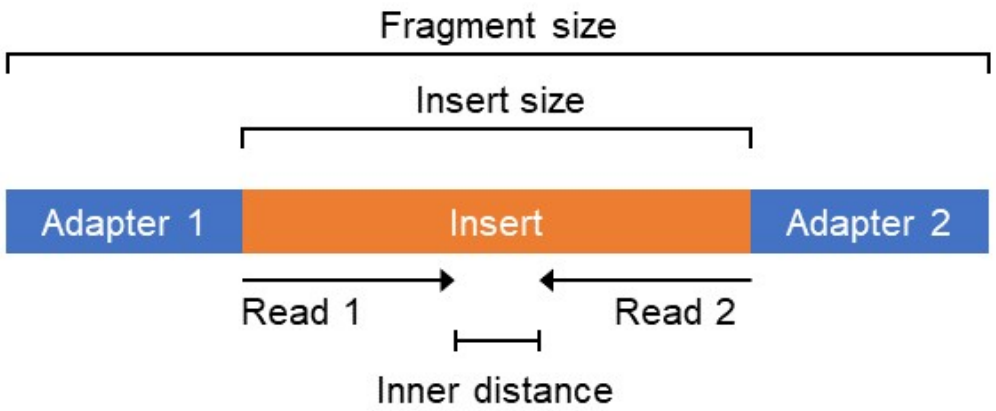
¿Cómo secuenciamos un fragmento de ADN no conocido?



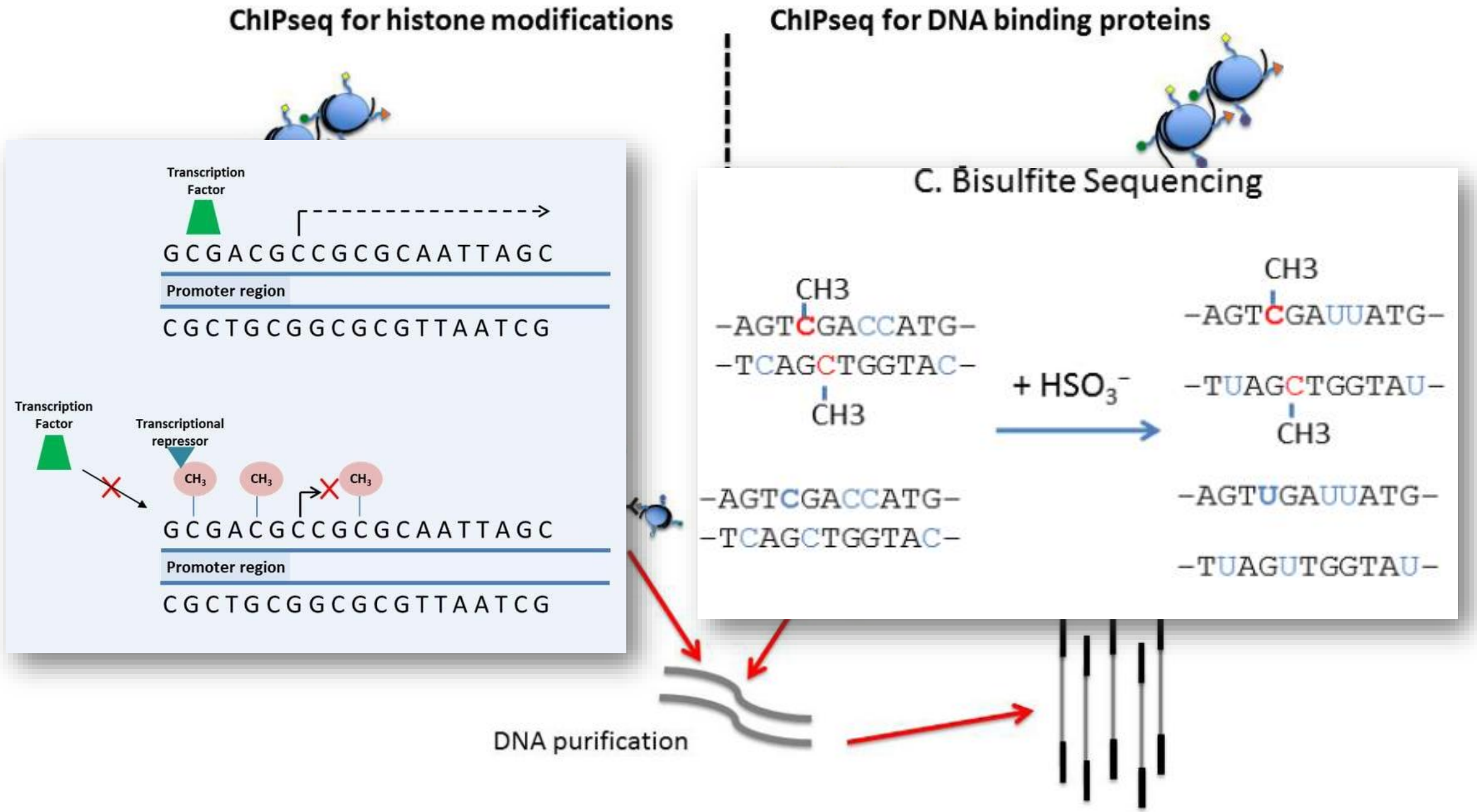

DNA Library Construction
Genomic DNA
Fragmentation
PCR
Fragmented DNA
Amplicons
Adapter Ligation, PCR
Adapters
Sequencing Library
RNA Library Construction
RNA
RT-PCR
Adapter Ligation
cDNA
Adapter Ligation, PCR
Adapters
RT-PCR
Sequencing Library

Librerías de ADN



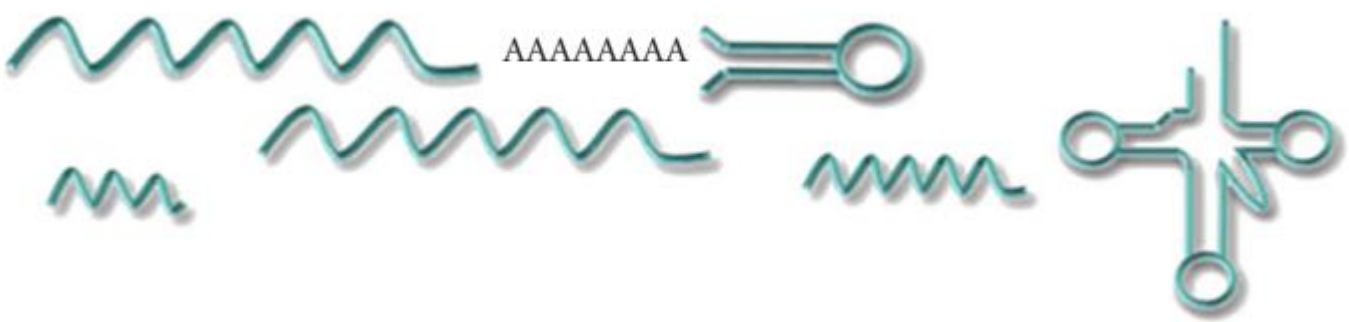


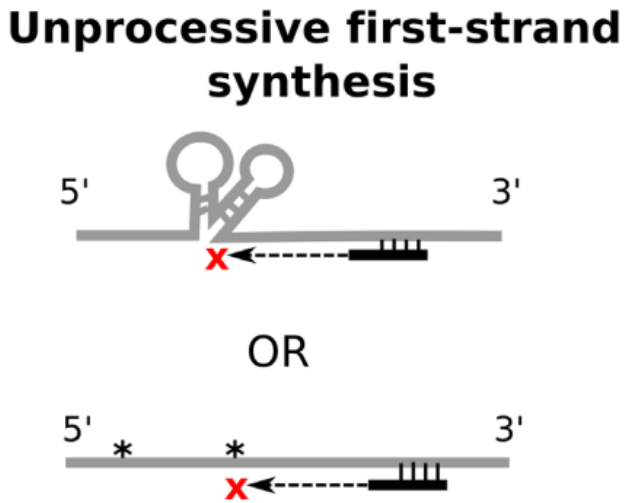
van Dijk EL, et al. Exp Cell Res. 2014

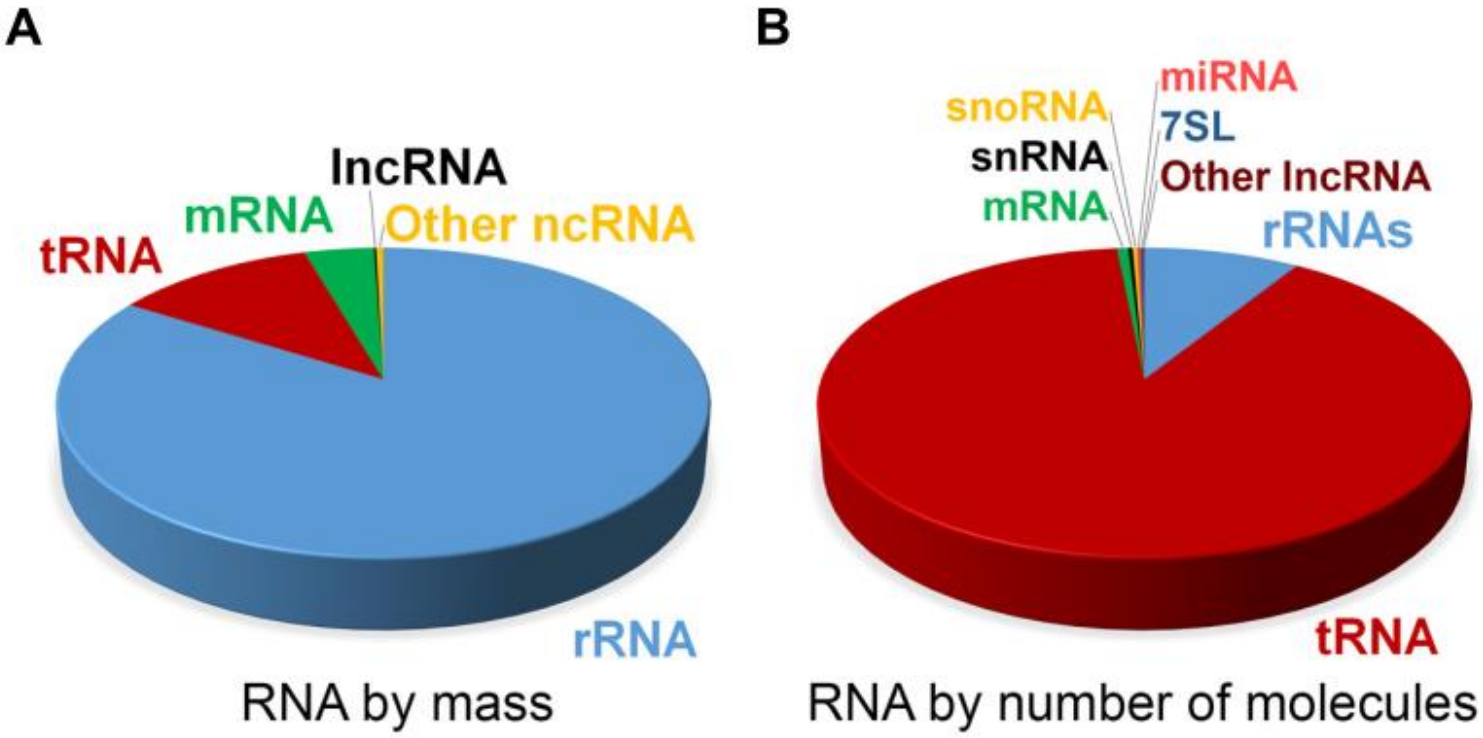


Head SR, et al. Biotechniques. 2014.

Diversidad de ARNs en la células

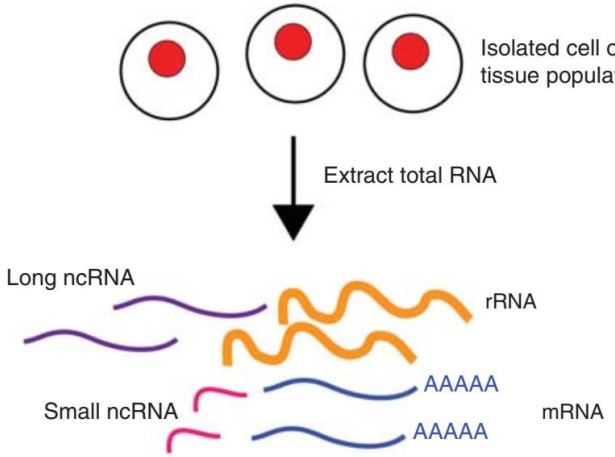






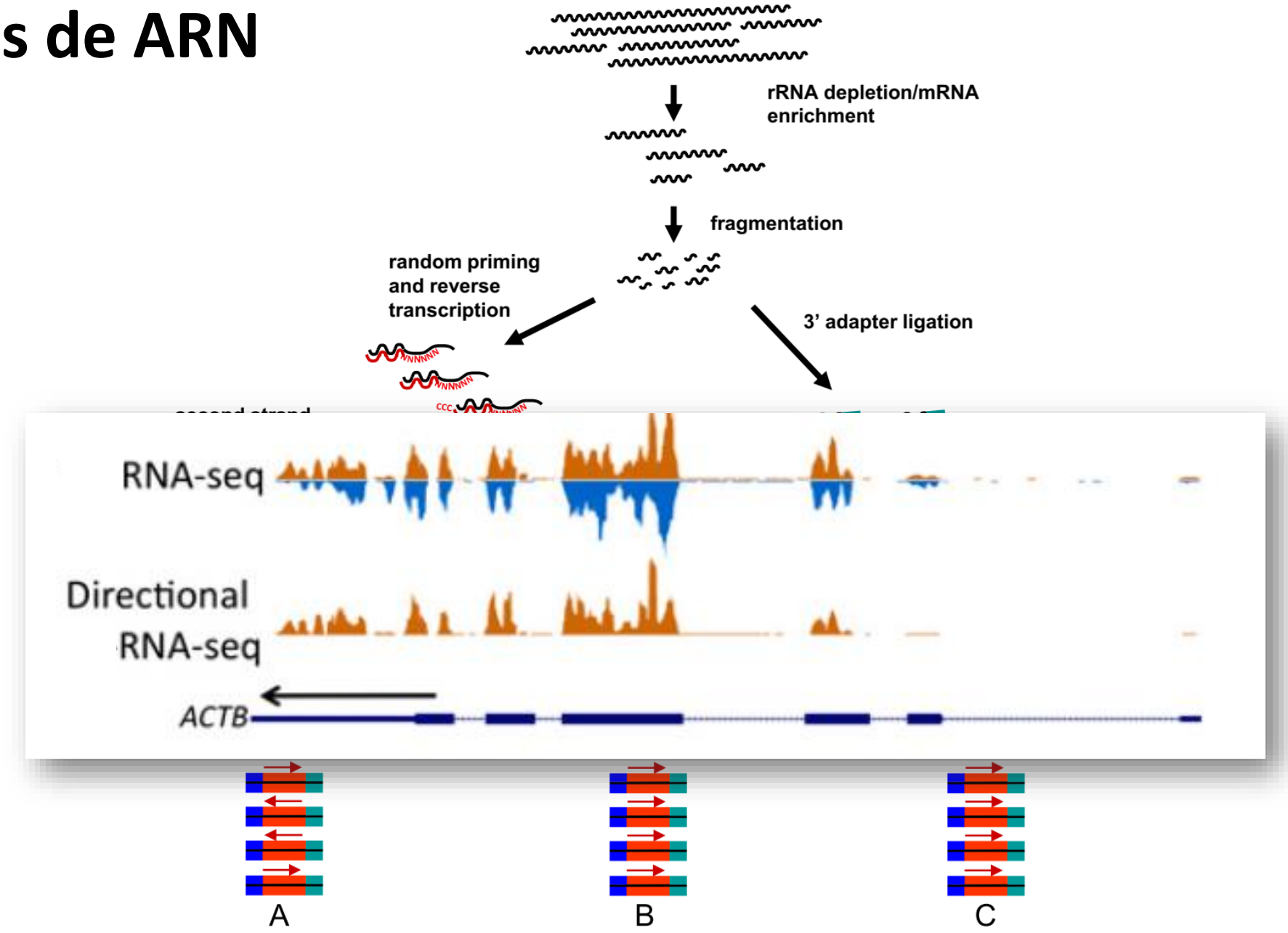
Palazzo AF., & Lee ES., et al. 2015

¿Cómo seleccionamos los ARNs de interés?



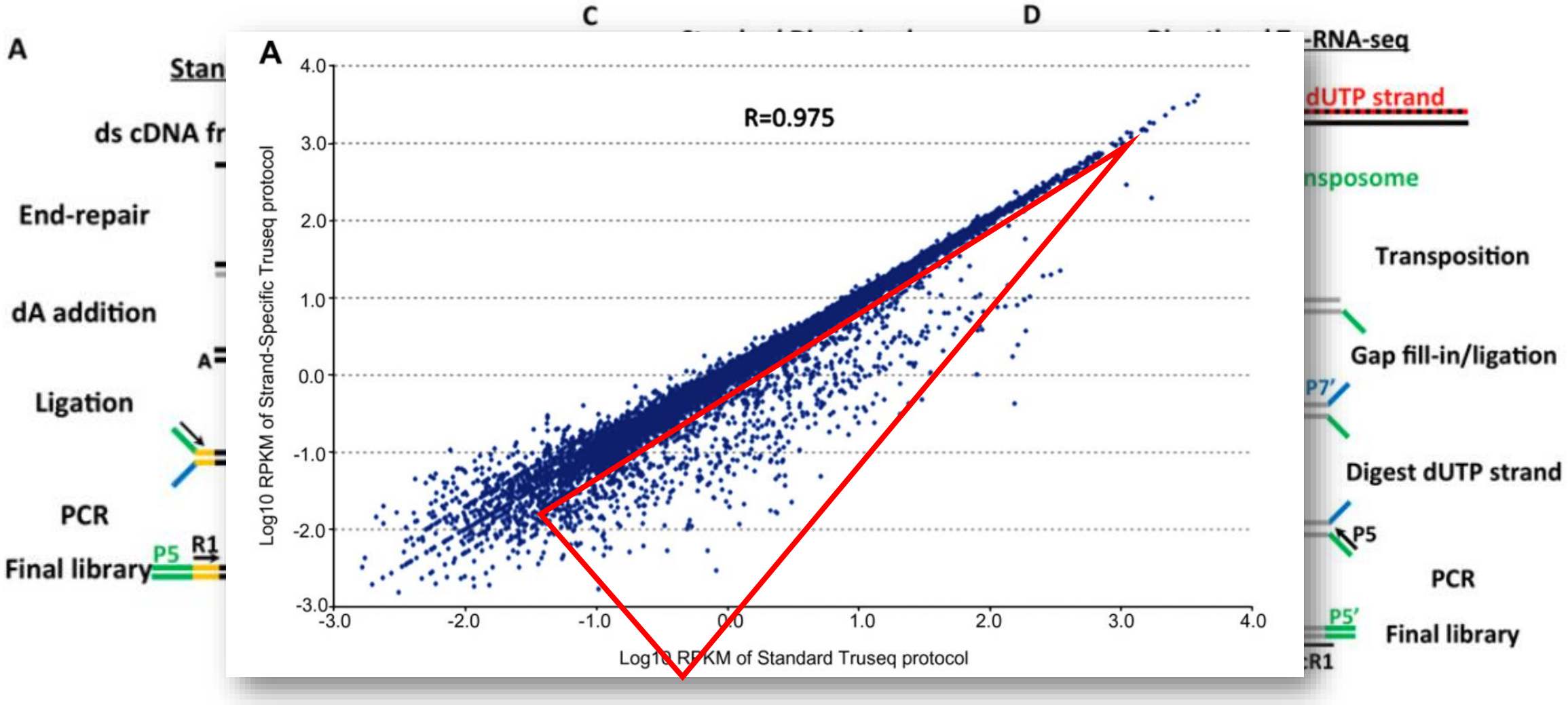
Kukurba KR, et al. Cold Spring Harb Protoc. 2015

Librerías de ARN



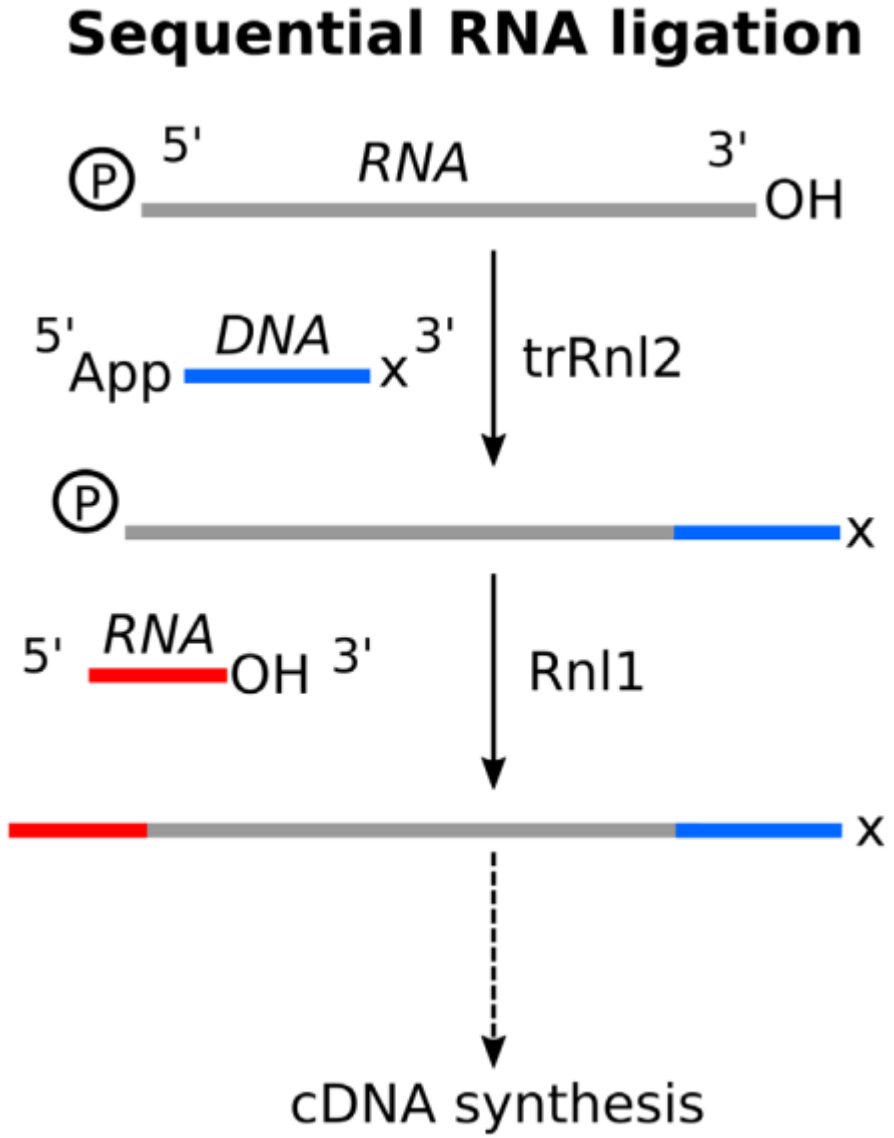
van Dijk EL, et al. Exp Cell Res. 2014

Librerías de ARN



Gertz J, et al. Genome Res. 2012; Sultan M, et al. Biochem Biophys Res Commun. 2012

Librerías de ARN



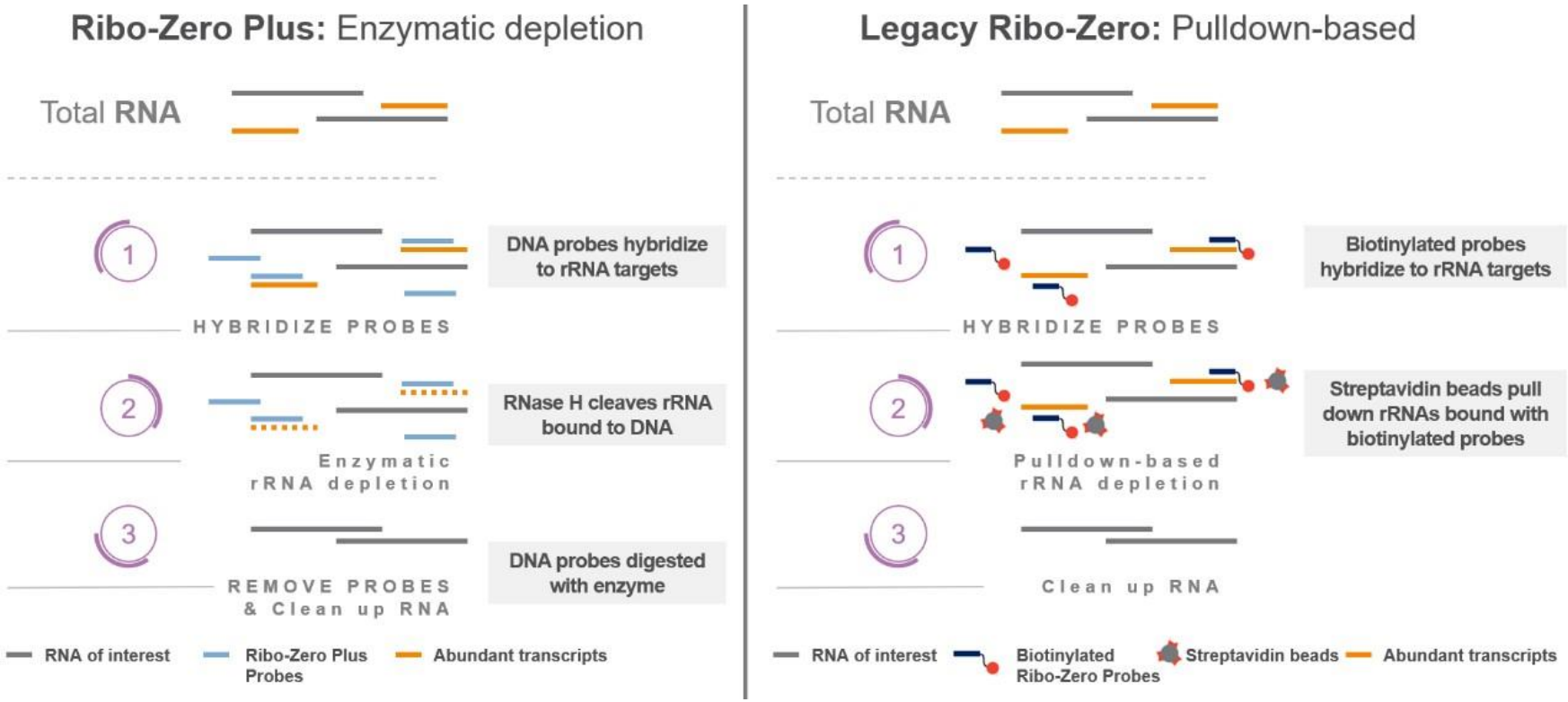
Gertz J, et al. Genome Res. 2012; Sultan M, et al. Biochem Biophys Res Commun. 2012

Estrategias para seleccionar ARNs de interés

Poly-A

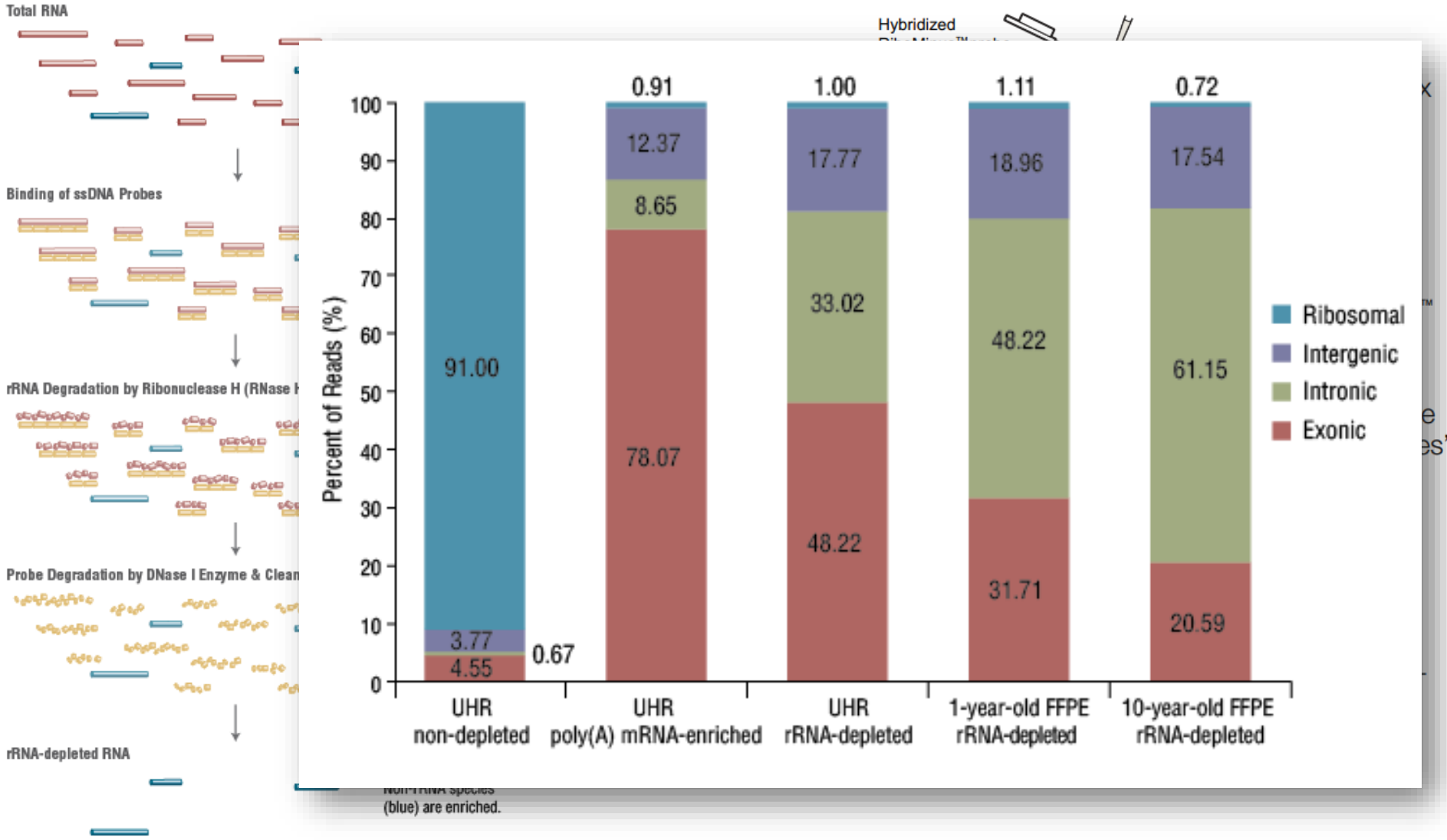
Tratamiento con DNAsas

Poly-A y sin Poly-A

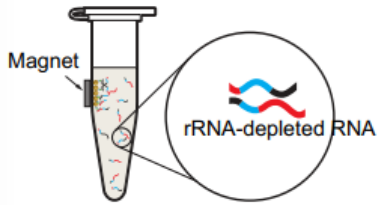


Ura, H., et al. Sci Rep (2022)

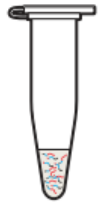
Estrategias para seleccionar ARNs de interés



10 minutes denature
20 minutes hybridization



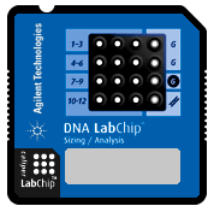
1 minute remove

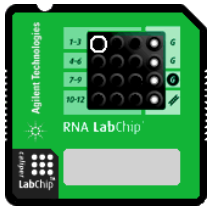


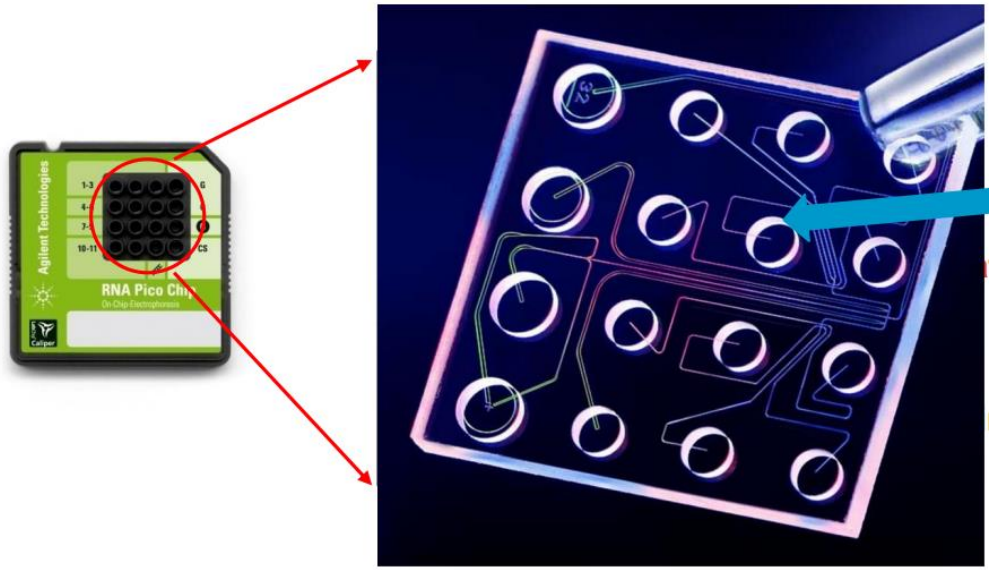
25 minute bead clean-up

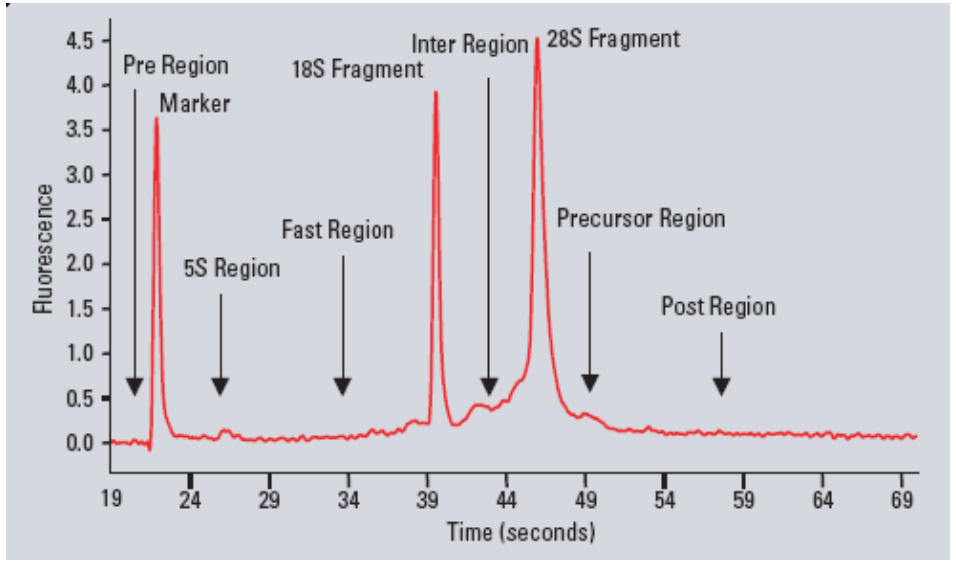
Ura, H., et al. Sci Rep (2022)

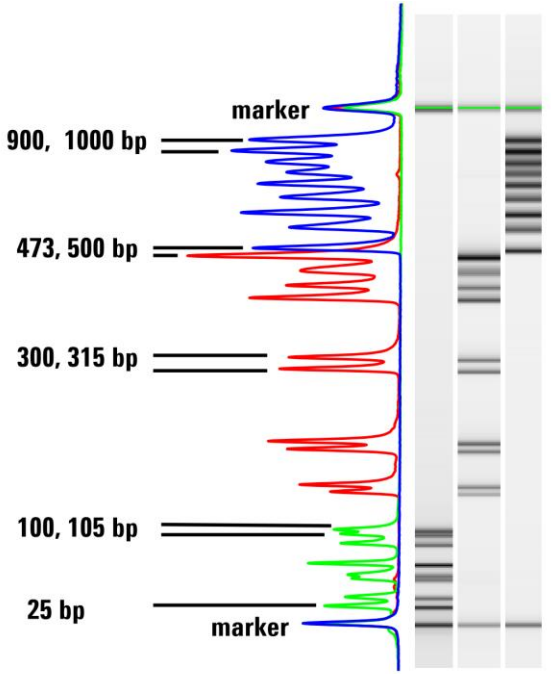
Análisis de Integridad y cuantificación de ARN y ADN - Bioanalyzer 2100



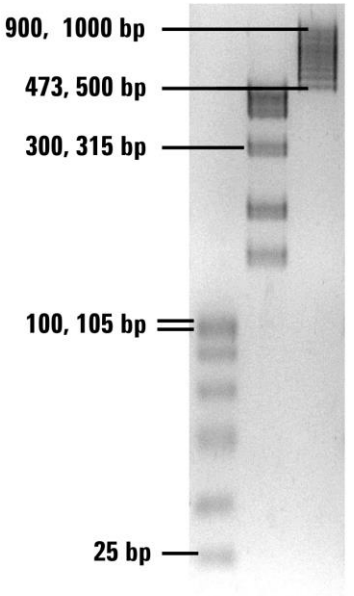








2100 bioanalyzer data
Gel-like image

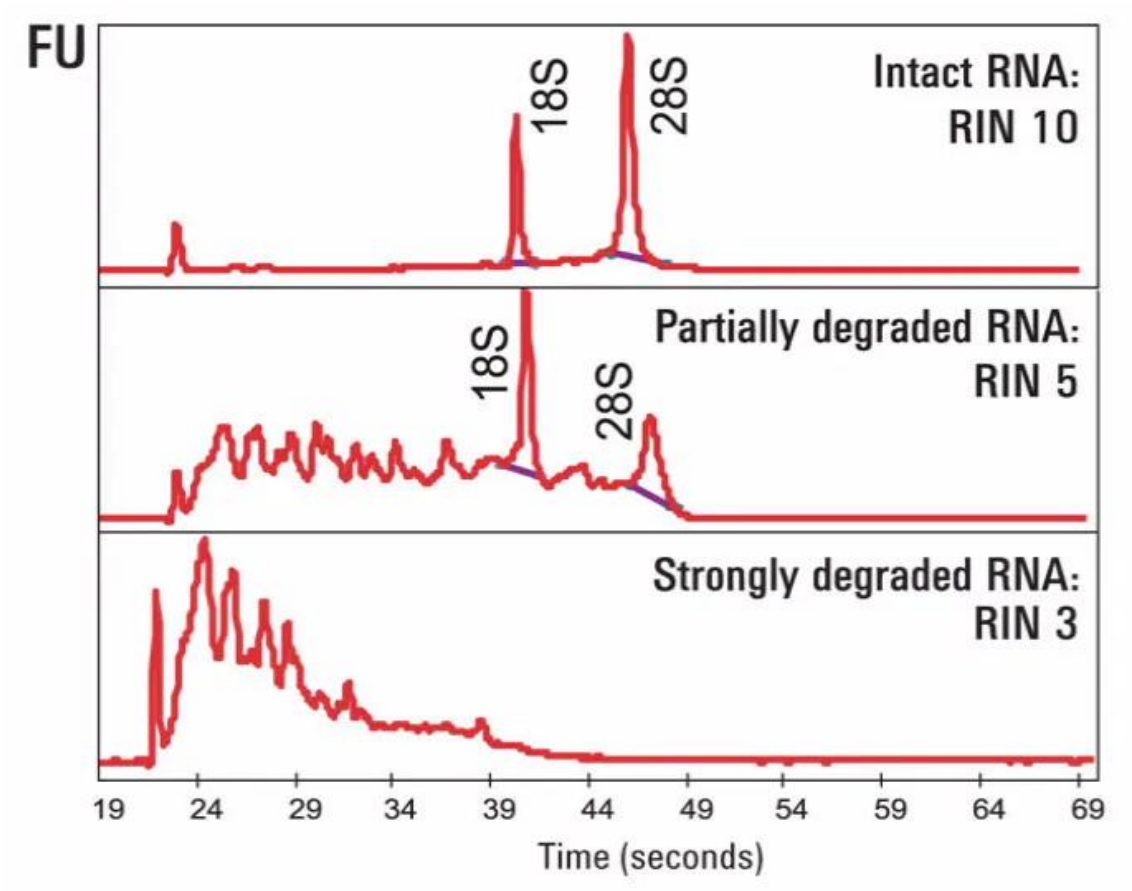


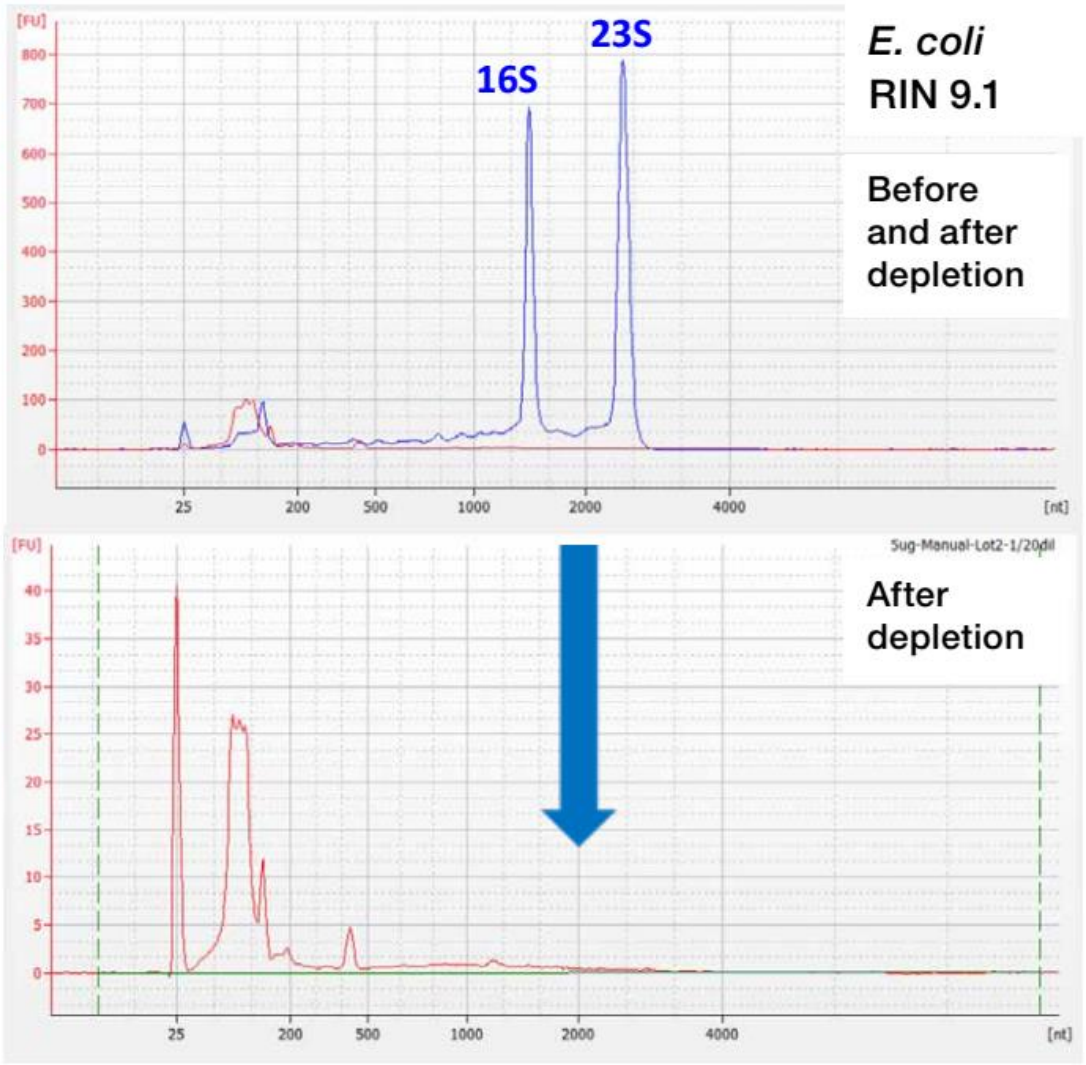
2 % agarose gel stained
with Ethidiumbromide

Análisis de ARN y ADN con Bioanalyzer 2100 - RIN

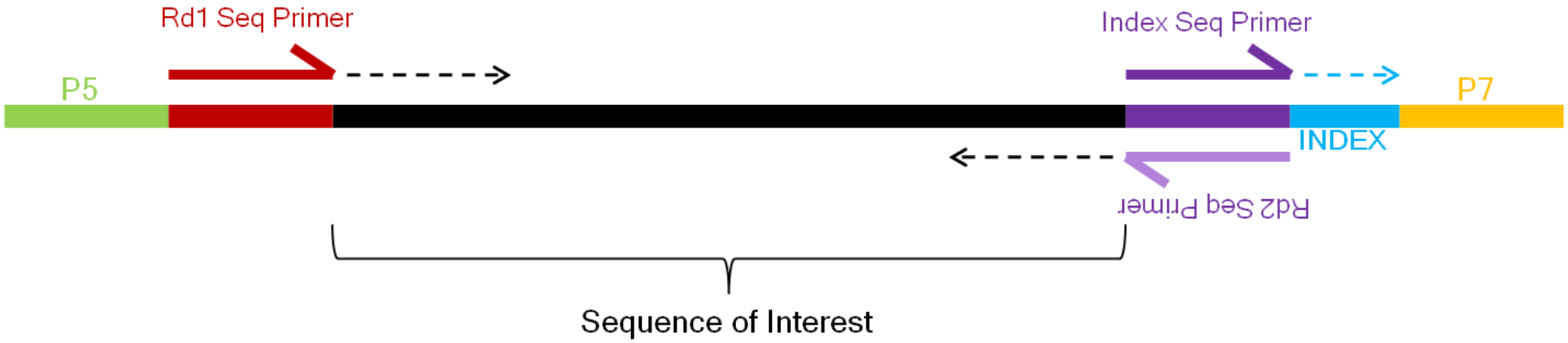
El **número de integridad del ARN (RIN)**

Medida de la integridad del ARN, un indicador de la integridad del ARNm

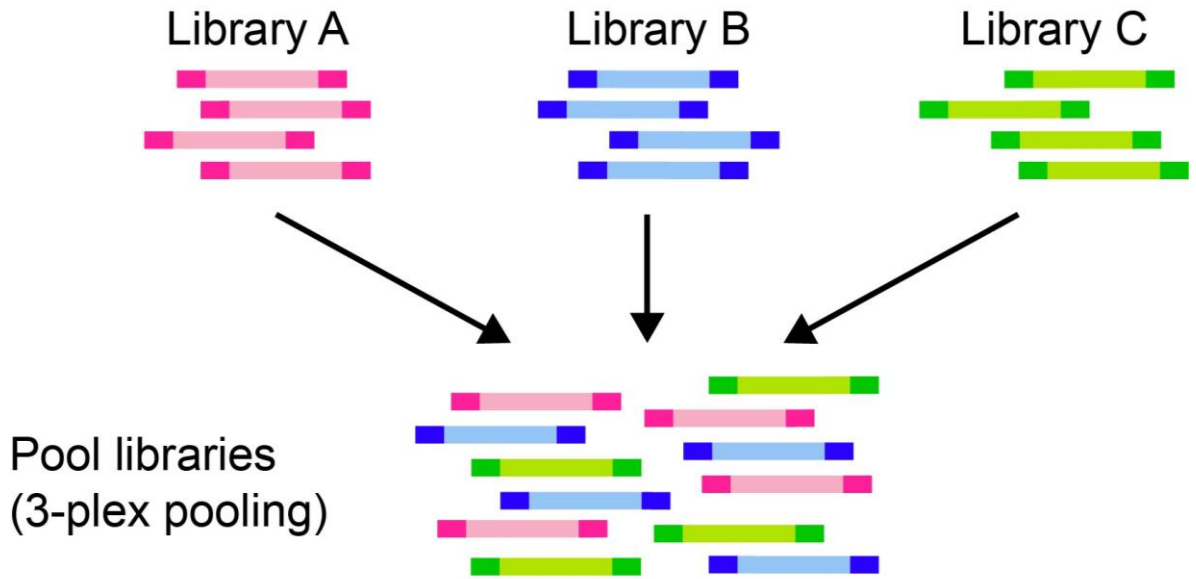


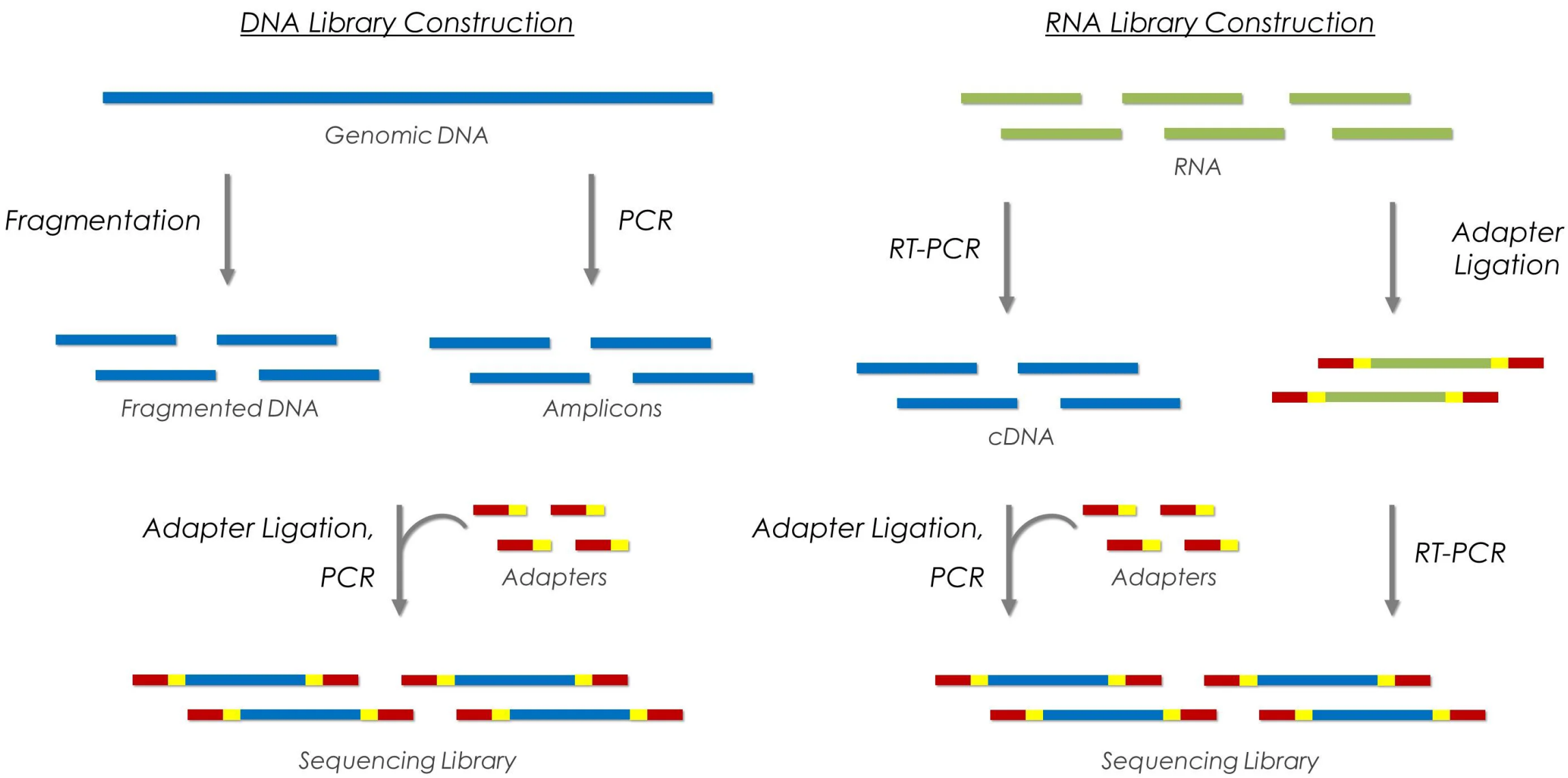


STRUCTURE DETAILS

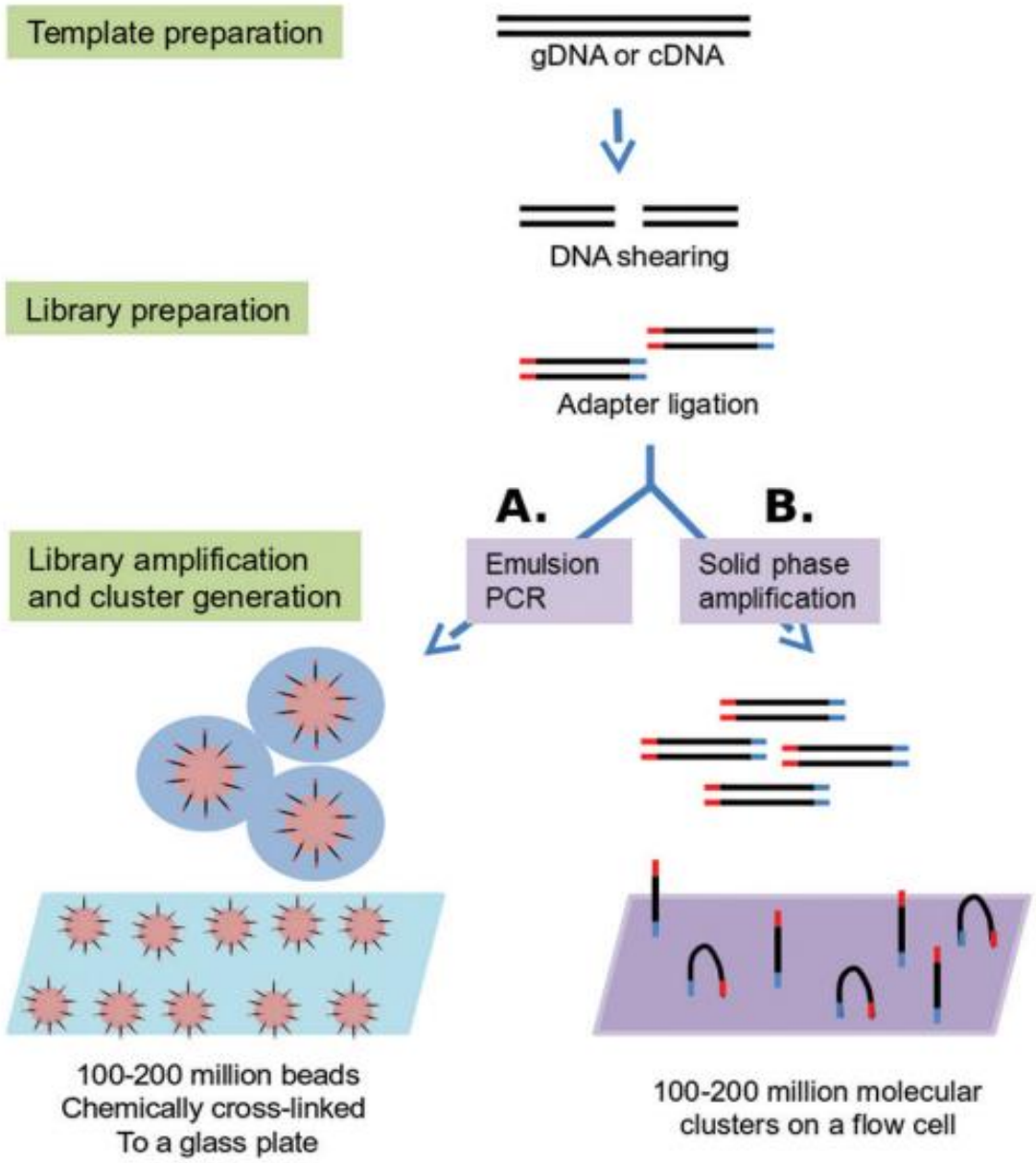


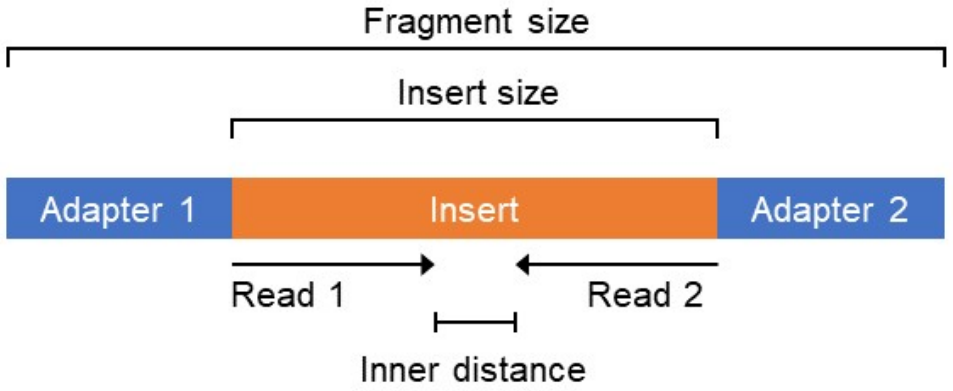
Barcode: Durante la preparación de la biblioteca antes de la secuenciación, se asigna un **código de barras específico** (**6-8 nucleótidos [nt] de longitud**) a cada muestra.

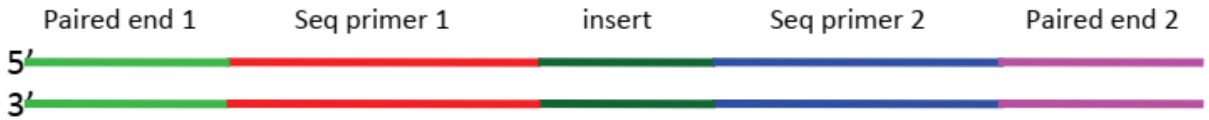



DNA Library Construction
Genomic DNA
Fragmentation
PCR
Fragmented DNA
Amplicons
Adapter Ligation, PCR
Adapters
Sequencing Library
RNA Library Construction
RNA
RT-PCR
Adapter Ligation
cDNA
Adapter Ligation, PCR
Adapters
RT-PCR
Sequencing Library

Preparación de Librerías





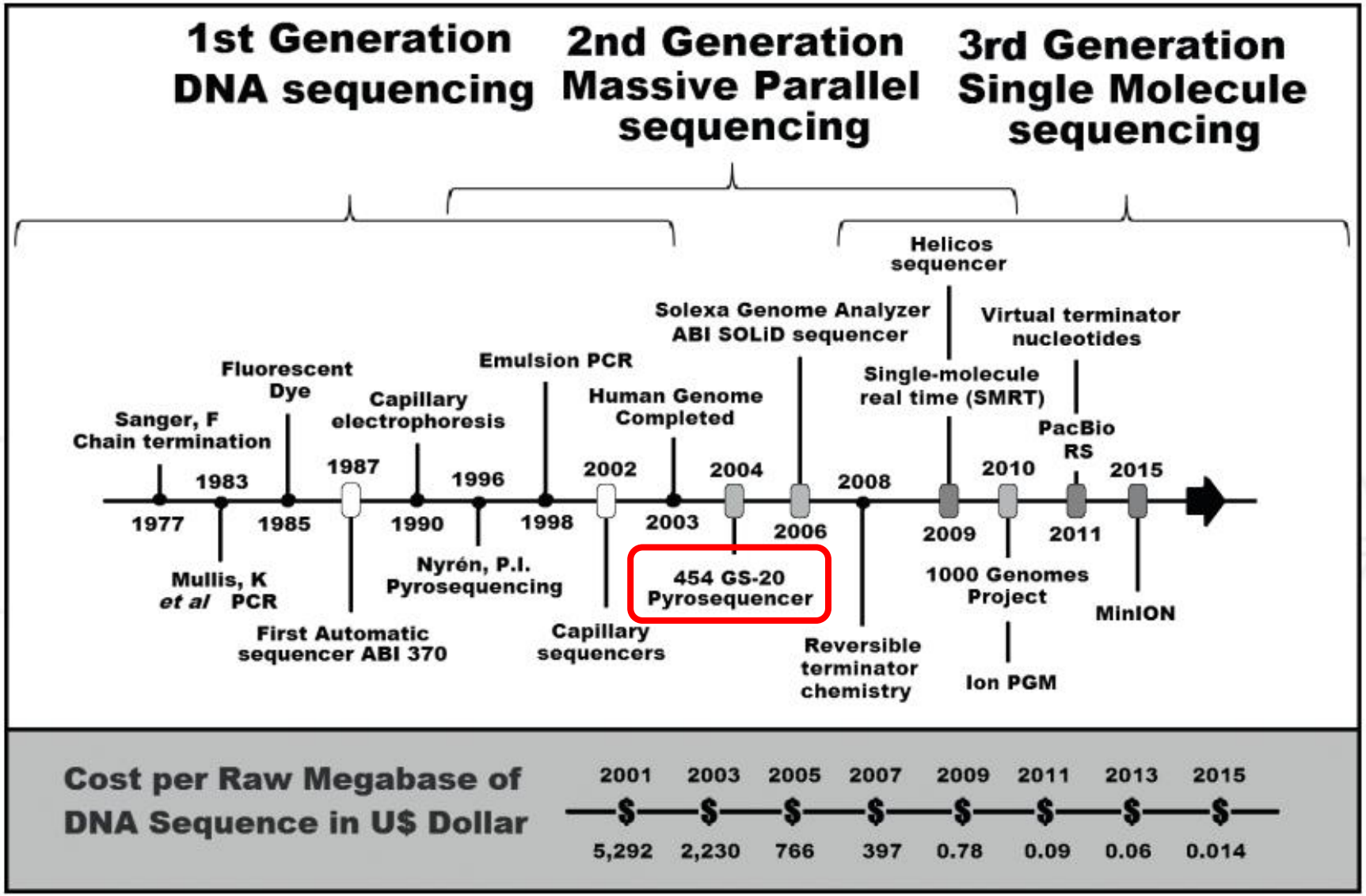


Mutagenesis vol, 29 no, 5 pp, 303–310 doi:10,1093/mutage/geu031

REVIEW

Next-generation sequencing technologies: breaking the sound barrier of human genetics

Plataformas de Secuenciación



http://dx.doi.org/10.5772/intechopen.69337

Nature. 2005 September 15; 437(7057): 376–380.

Genome Sequencing in Open Microfabricated High Density Picoliter Reactors

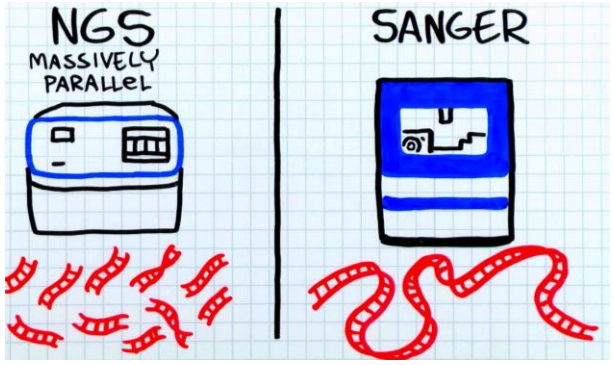
Marcel Margulies[1,*], Michael Egholm[1,*], William E. Altman[1], Said Attiya[1], Joel S. Bader[1], Lisa A. Bemben[1], Jan Berka[1], Michael S. Braverman[1], Yi-Ju Chen[1], Zhoutao Chen[1], Scott B. Dewell[1], Lei Du[1], Joseph M. Fierro[1], Xavier V. Gomes[1], Brian C. Goodwin[1], Wen He[1], Scott Helgesen[1], Chun He Ho[1], Gerard P. Irzyk[1], Szilveszter C. Jando[1], Maria L.I. Alenquer[1], Thomas P. Jarvie[1], Kshama B. Jirage[1], Jong-Bum Kim[1], James R. Knight[1], Janna R. Lanza[1], John H. Leamon[1], Steven M. Lefkowitz[1], Ming Lei[1], Jing Li[1], Kenton L. Lohman[1], Hong Lu[1], Vinod B. Makhijani[1], Keith E. McDade[1], Michael P. McKenna[1], Eugene W. Myers[3], Elizabeth Nickerson[1], John R. Nobile[1], Ramona Plant[1], Bernard P. Puc[1], Michael T. Ronan[1], George T. Roth[1], Gary J. Sarkis[1], Jan Fredrik Simons[1], John W. Simpson[1], Maithreyan Srinivasan[1], Karrie R. Tartaro[1], Alexander Tomasz[4], Kari A. Vogt[1], Greg A. Volkmer[1], Shally H. Wang[1], Yong Wang[1], Michael P. Weiner[2], Pengguang Yu[1], Richard F. Begley[1], and Jonathan M. Rothberg[1]

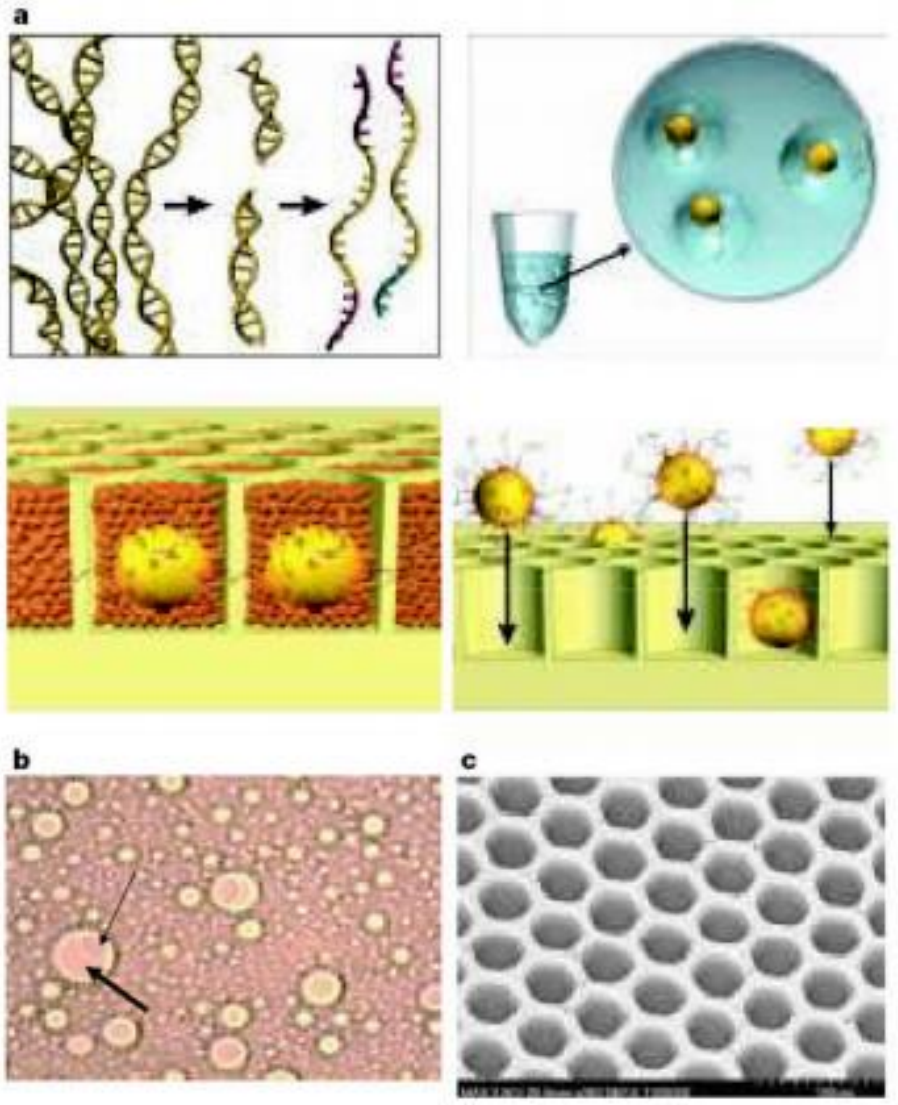
1 *454 Life Sciences Corp., 20 Commercial St., Branford, CT 06405, USA.*

2 *The Rothberg Institute For Childhood Diseases, 530 Whitfield St., Guilford, CT 06437, USA.*

3 *University of California, Berkeley, CA 94720, USA.*

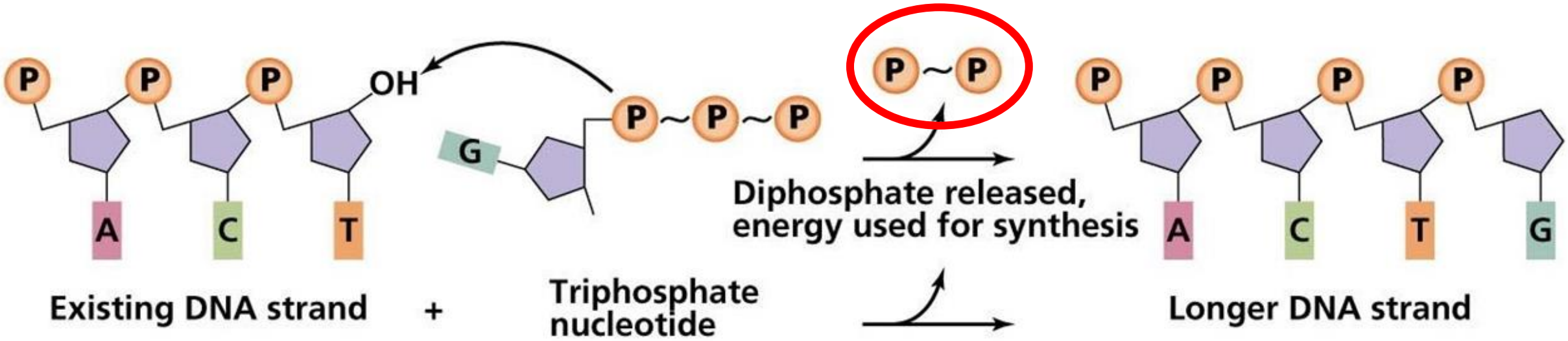
4 *Laboratory of Microbiology, The Rockefeller University, New York, NY 10021, USA.*





Escalado de la reacción de PCR y Secuenciación a picolitros (10^{-12}) gran reducción de costo y tiempo.

454: secuenciación (Pirosecuenciación)



Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings.

454: secuenciación (Pirosecuenciación)

- **Medición de síntesis de pirofosfato**:
  - proceso de dos enzimas en el que se usa **ATP sulfurilasa** para convertir el **pirofosfato** en **ATP**, que luego se usa como sustrato para **luciferasa**, produciendo **luz proporcional a la cantidad de pirofosfato**.

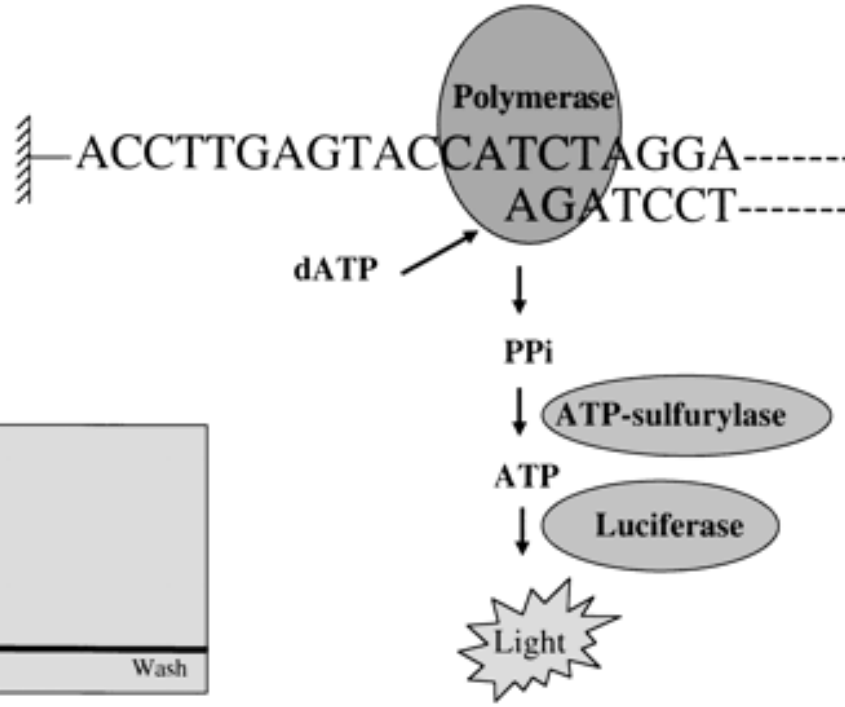
Anal Biochem. 1985 Dec;151(2):504-9.

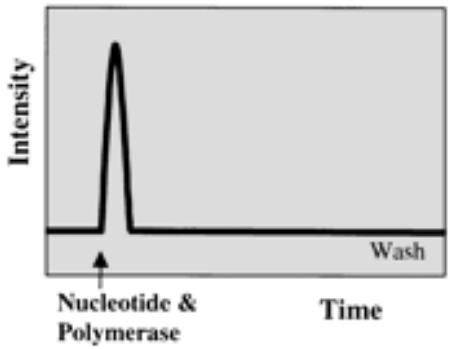
Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis.

Nyrén P, Lundin A.

Abstract

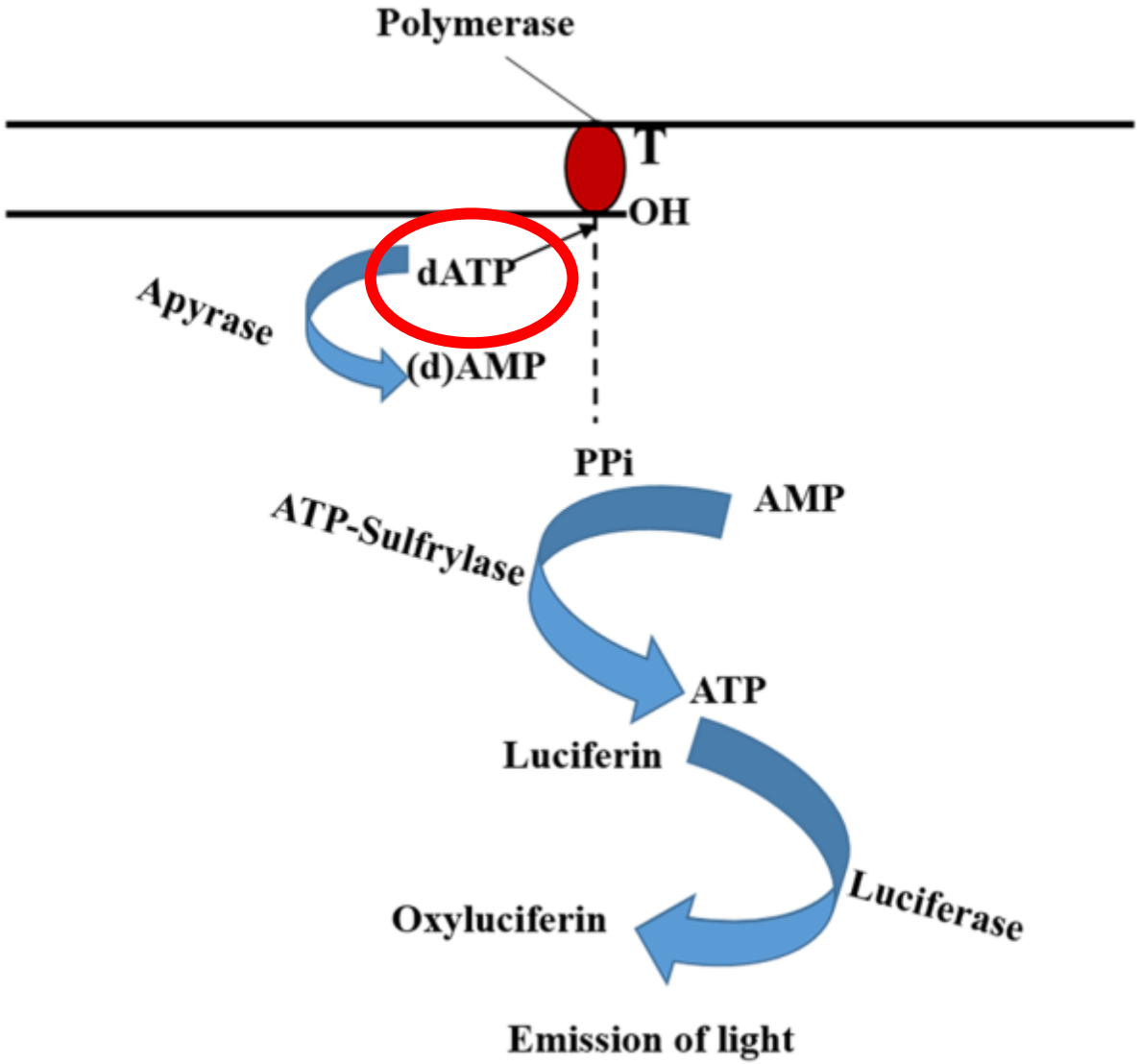
A sensitive method for the analysis of inorganic pyrophosphate (PPi) which utilizes the enzymes ATP sulfurylase and firefly is described. The assay is based on continuous monitoring of the ATP formed in the ATP sulfurylase reaction using purified luciferase. The assay can be completed in less than 2 s and is not affected by inorganic phosphate. The method has been continuous monitoring of formation of PPi in Rhodospirillum rubrum chromatophores. The assay is extremely sensitive, the range of the assay being 1 X 10(-9) - 5 X 10(-7) M PPi. It is suitable for routine applications. It is also possible to use the n determination of low amounts of adenosine 5'-phosphosulfate.

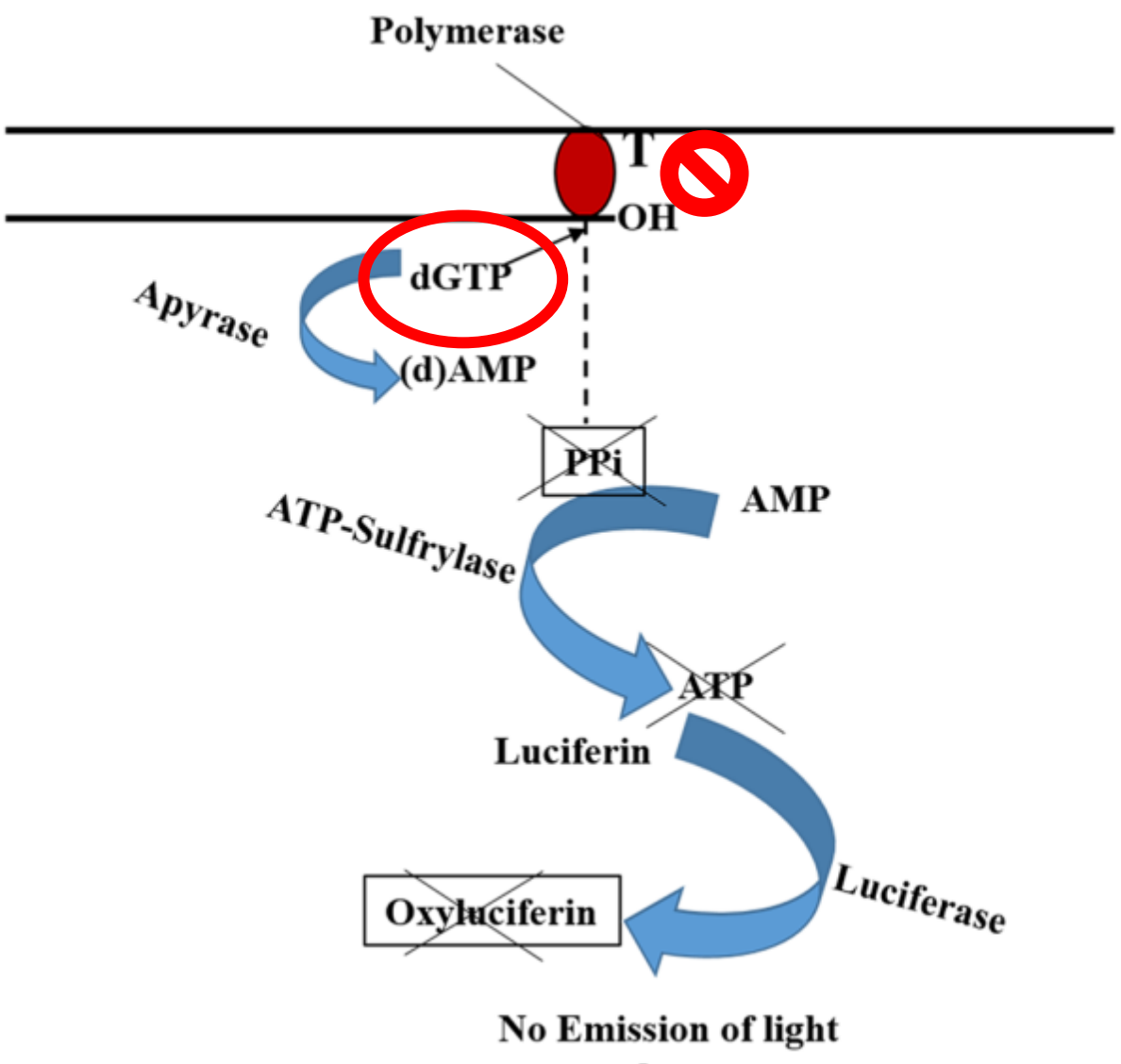




- Principales beneficios:
  - **Nucleótidos no modificados.**
  - **Medición en tiempo real.**

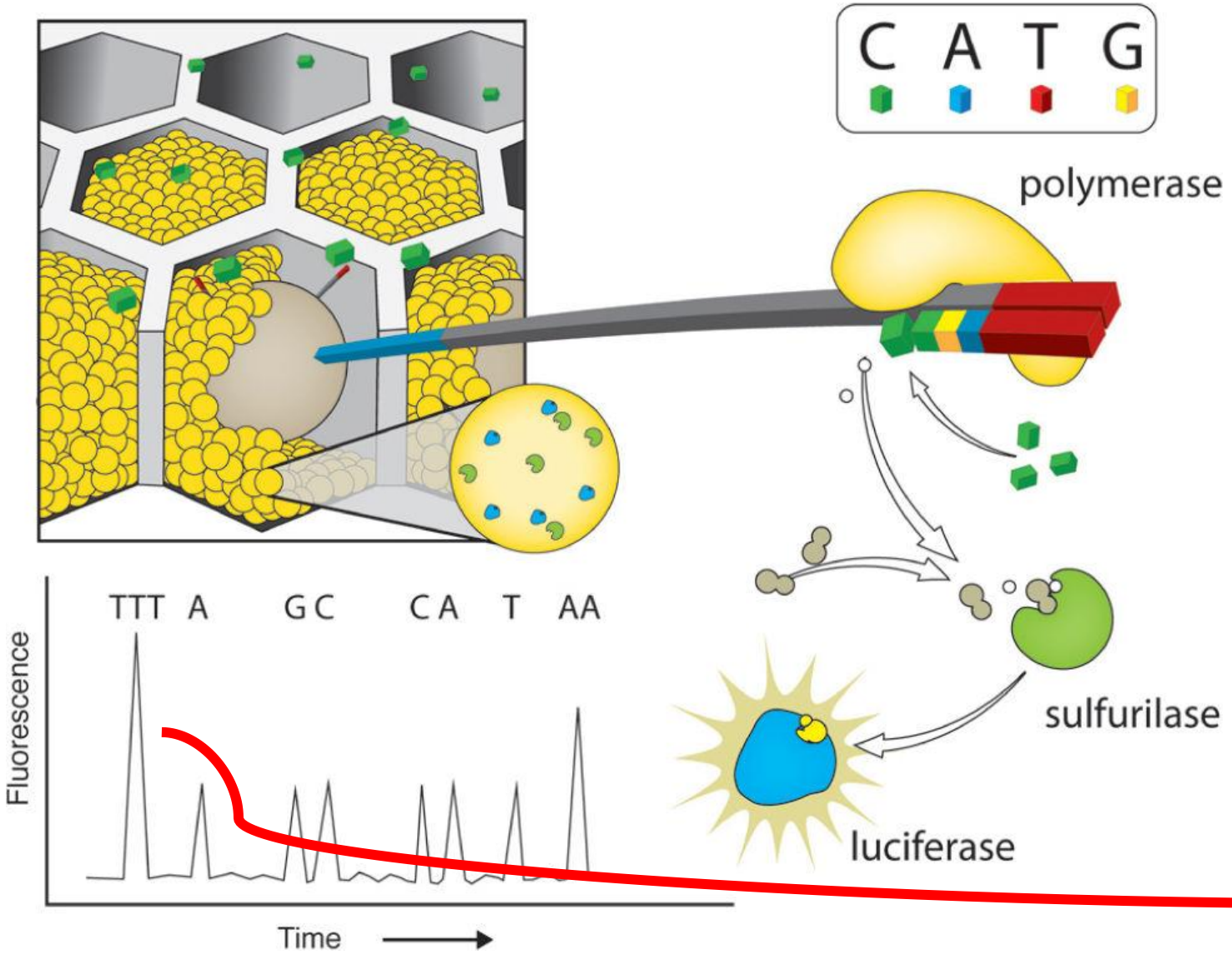
454: secuenciación (Pirosecuenciación)

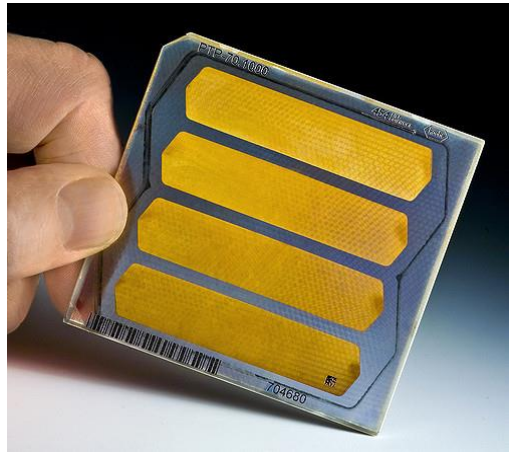


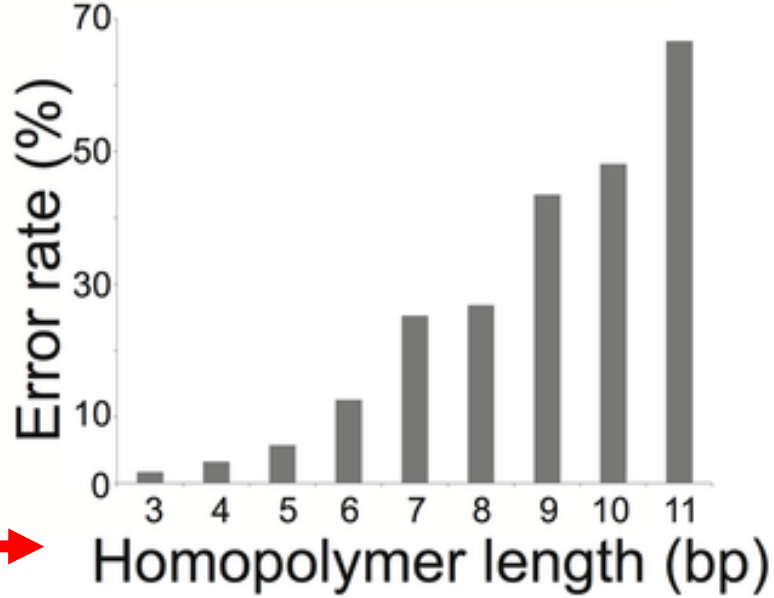


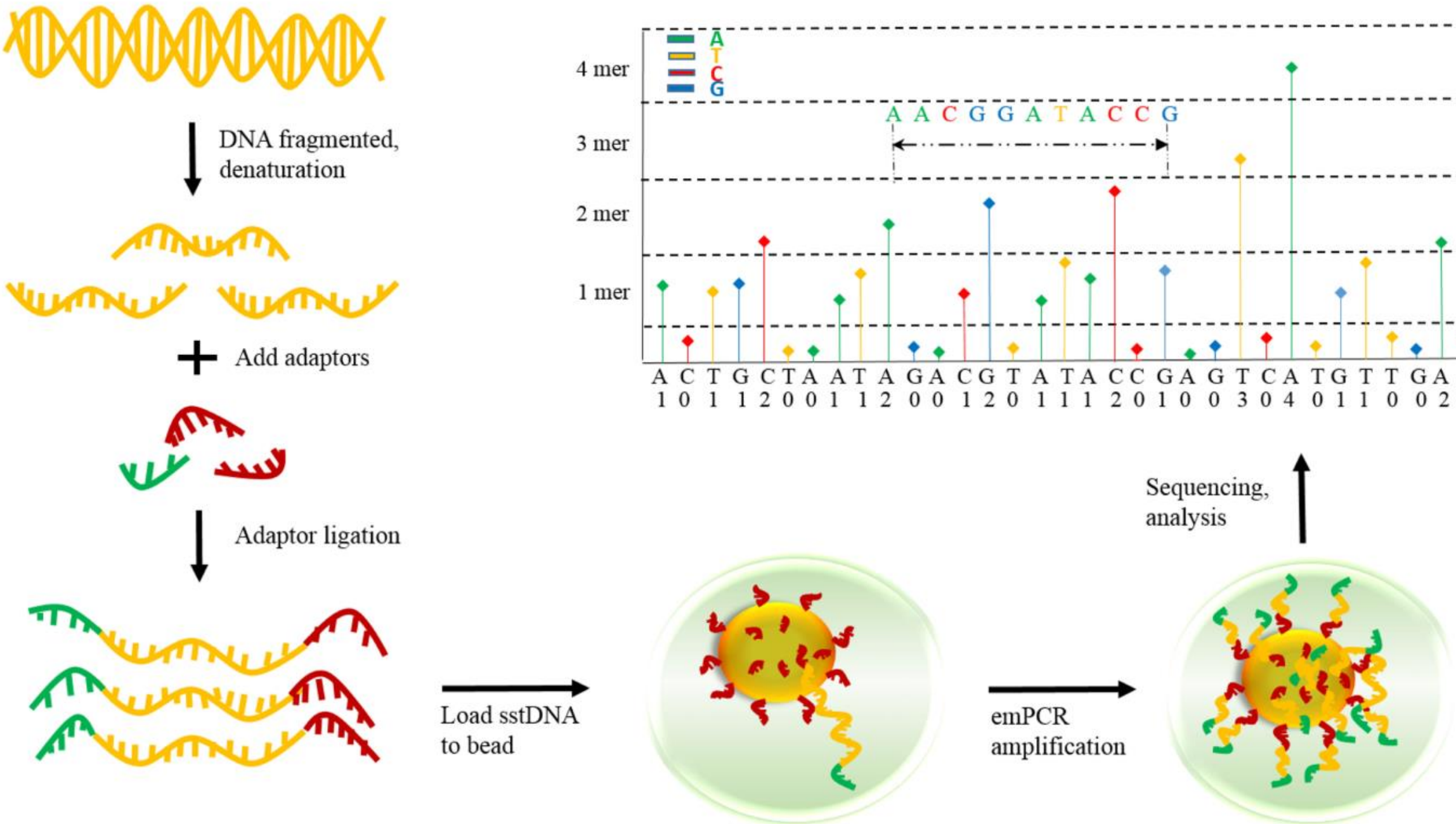
Ambardar S,et al. High Throughput Sequencing: An Overview of Sequencing Chemistry. Indian J Microbiol. 2016

454: secuenciación (Pirosecuenciación)





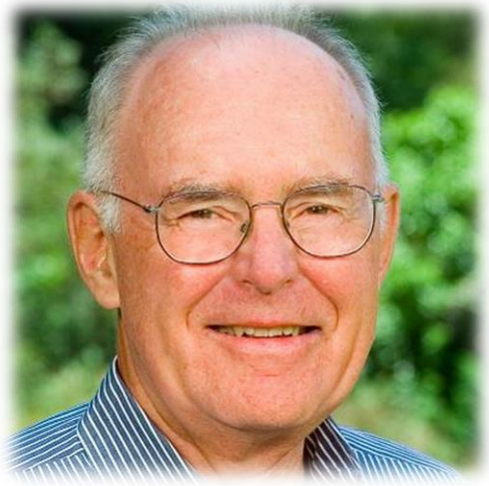


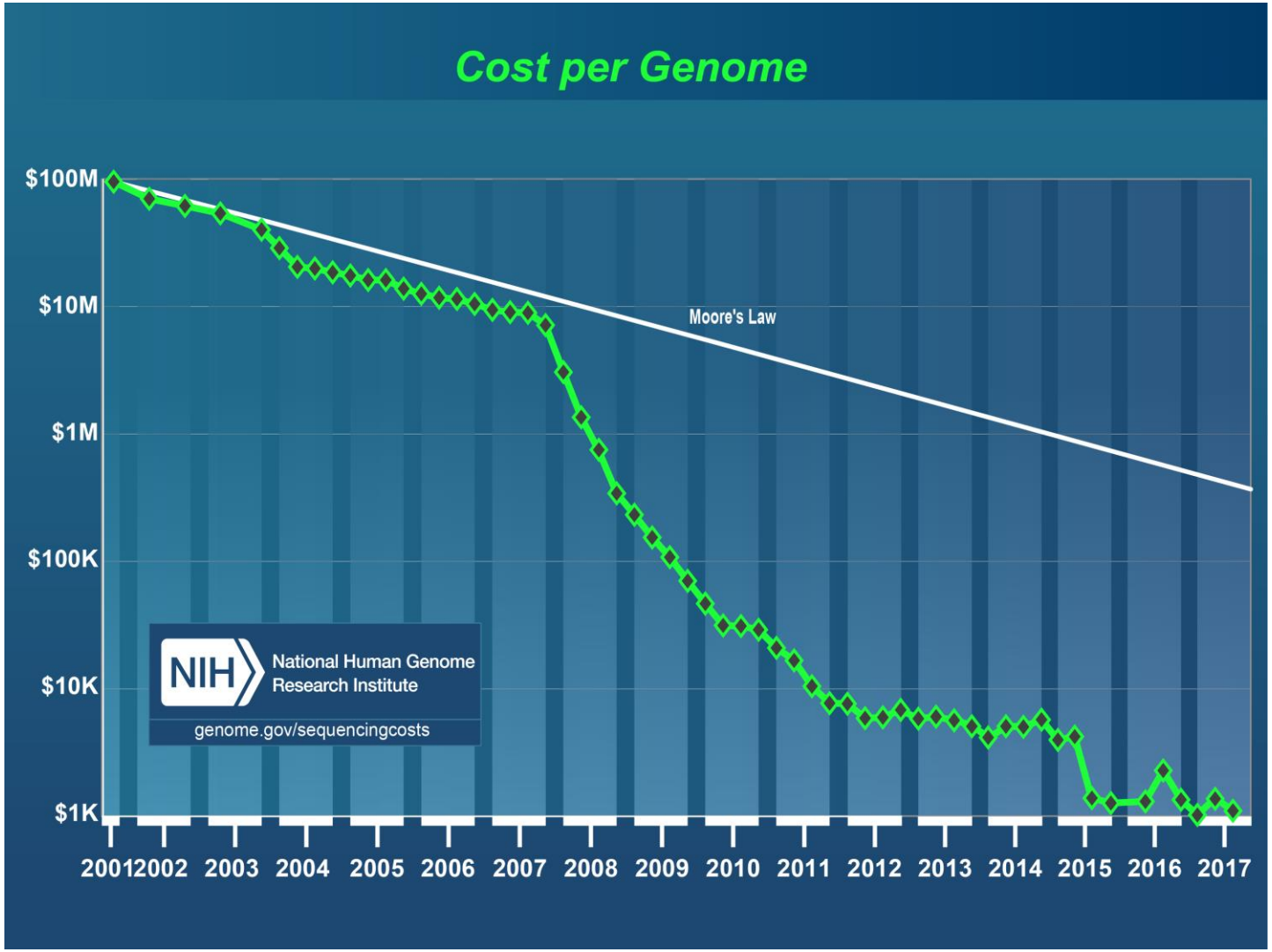


Ka-Chun Wong, et al. 2019. DNA Sequencing Technologies: Sequencing Data Protocols and Bioinformatics Tools.

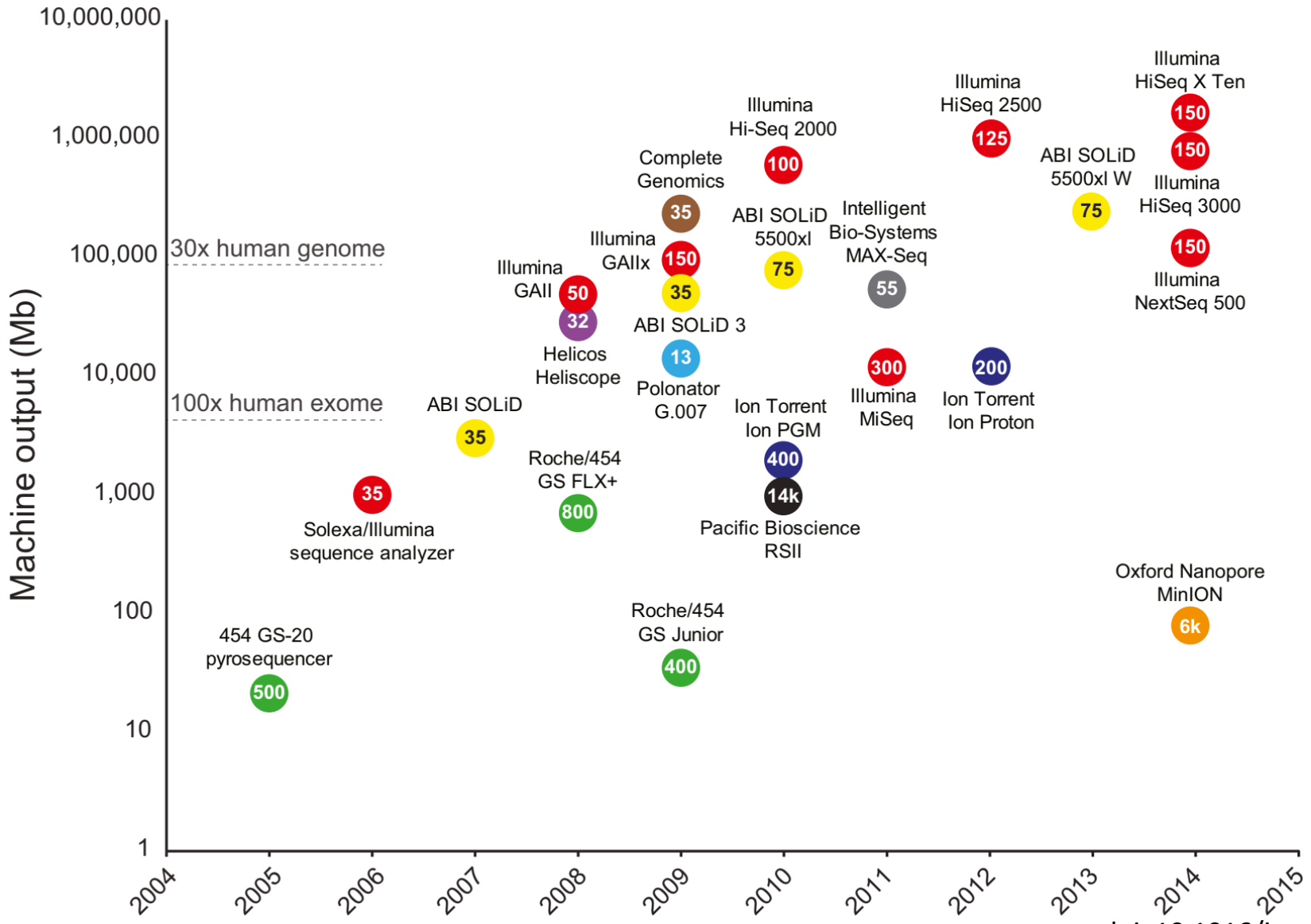
Secuenciadores vs Ley de Moore

Expresa que aproximadamente cada dos años se duplica el número de transistores en un microprocesador.



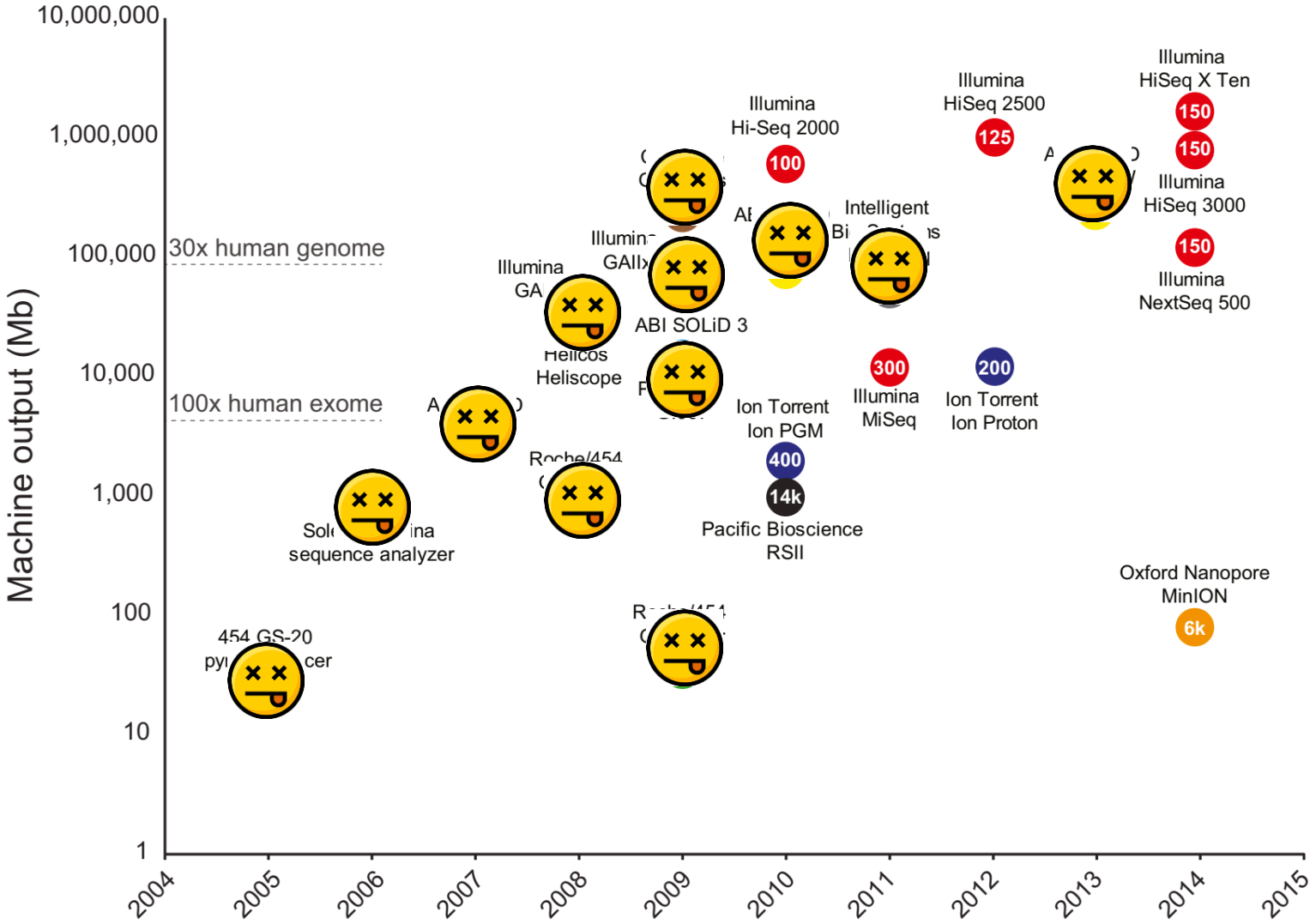


Plataformas de Secuenciación



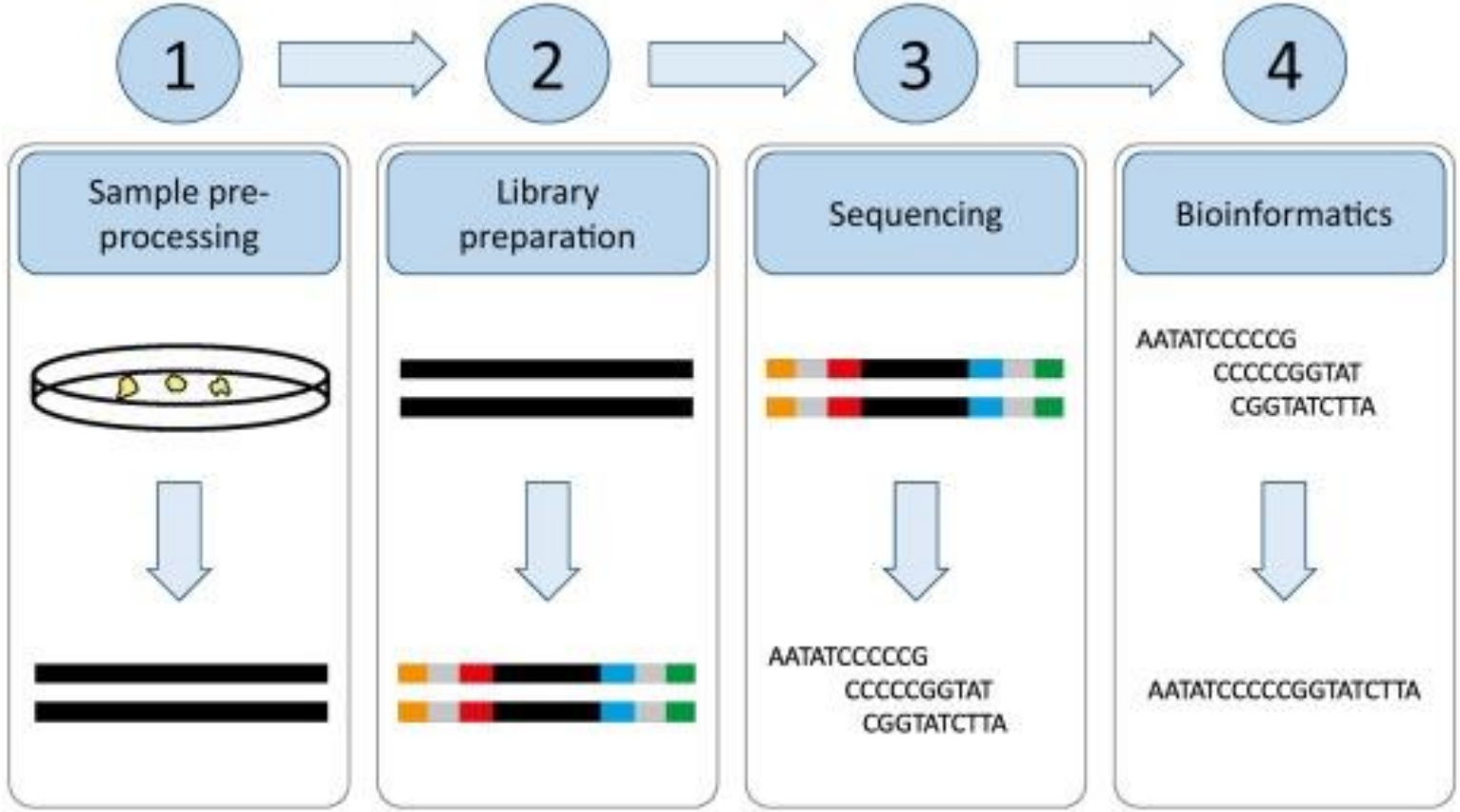
doi: 10.1016/j.molcel.2015.05.004

Plataformas de Secuenciación



doi: 10.1016/j.molcel.2015.05.004

2da parte del Teórico


1
2
3
4
Sample pre-processing
Library preparation
Sequencing
Bioinformatics
AATATCCCCCG
CCCCCGGTAT
CGGTATCTTA
AATATCCCCCG
CCCCCGGTAT
CGGTATCTTA
AATATCCCCCGGTATCTTA


Illumina P5 Adapter
Illumina P7 Adapter
i5 index
i7 index
Index 2 (i5) Sequencing Primer *
Read 1 Sequencing Primer
Read 1
Index 1 (i7) Sequencing Primer
Index 2 (i5) Sequencing Primer *
Read 2
Read 2 Sequencing Primer

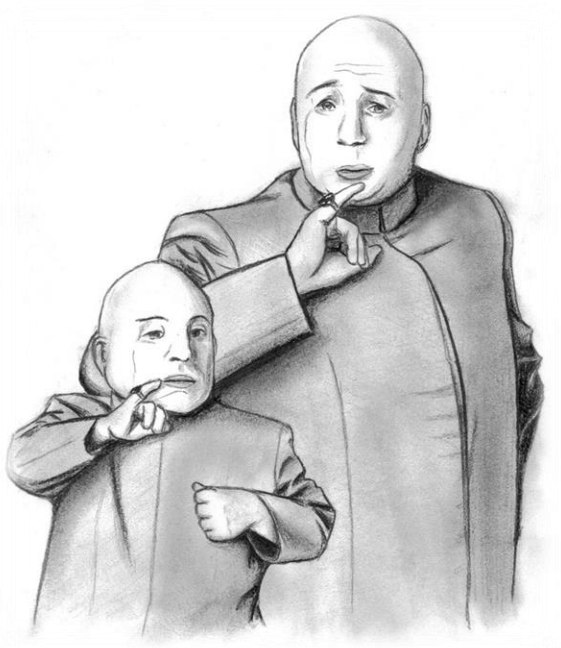
¿Cómo se clasifican las NGS?

- **Por generaciones:**
  - **1^{era}: Secuenciación de Sanger.**
  - **2da: Amplificación por PCR (emulsion o puente).**
  - **3^{era}: Secuenciación de moléculas únicas.**

¿Cómo se clasifican las NGS?

- **Por generaciones:**
  - **1^{era}: Secuenciación de Sanger.**
  - **2da: Amplificación por PCR (emulsion o puente).**
  - **3^{era}: Secuenciación de moléculas únicas.**

- **Por el tamaño de la lectura (read):**
  - **Cortos (35-600 pb)**
    - **Secuenciamiento por síntesis (SBS).**
    - **Secuenciamiento por ligación (SBL).**
  - **Largos (1-100 kb)**
    - **Single-molecule.**



Características de los Secuenciadores de 2da Generación

- **Necesitan de un paso de amplificación clonal.**
  - **Emulsion PCR. (Emulsión)**
  - **Bridge PCR. (Puente)**
- **Producen lecturas cortas.**
  - **Single-end. (un Extremo)**
  - **Paired-end. (ambos Extremos)**
- **Se clasifican en:**
  - **Secuenciación por síntesis.**
  - **Secuenciación por ligación.**

illumina®

ion torrent

by life technologies™

QIAGEN

Roche 454 SEQUENCING

AB applied biosystems™

Complete genomics

Características de los Secuenciadores de 2da Generación

- **Necesitan de un paso de amplificación clonal.**
  - **Emulsion PCR. (Emulsión)**
  - **Bridge PCR. (Puente)**
- **Producen lecturas cortas.**
  - **Single-end. (un Extremo)**
  - **Paired-end. (ambos Extremos)**
- **Se clasifican en:**
  - **Secuenciación por síntesis.**
  - **Secuenciación por ligación.**

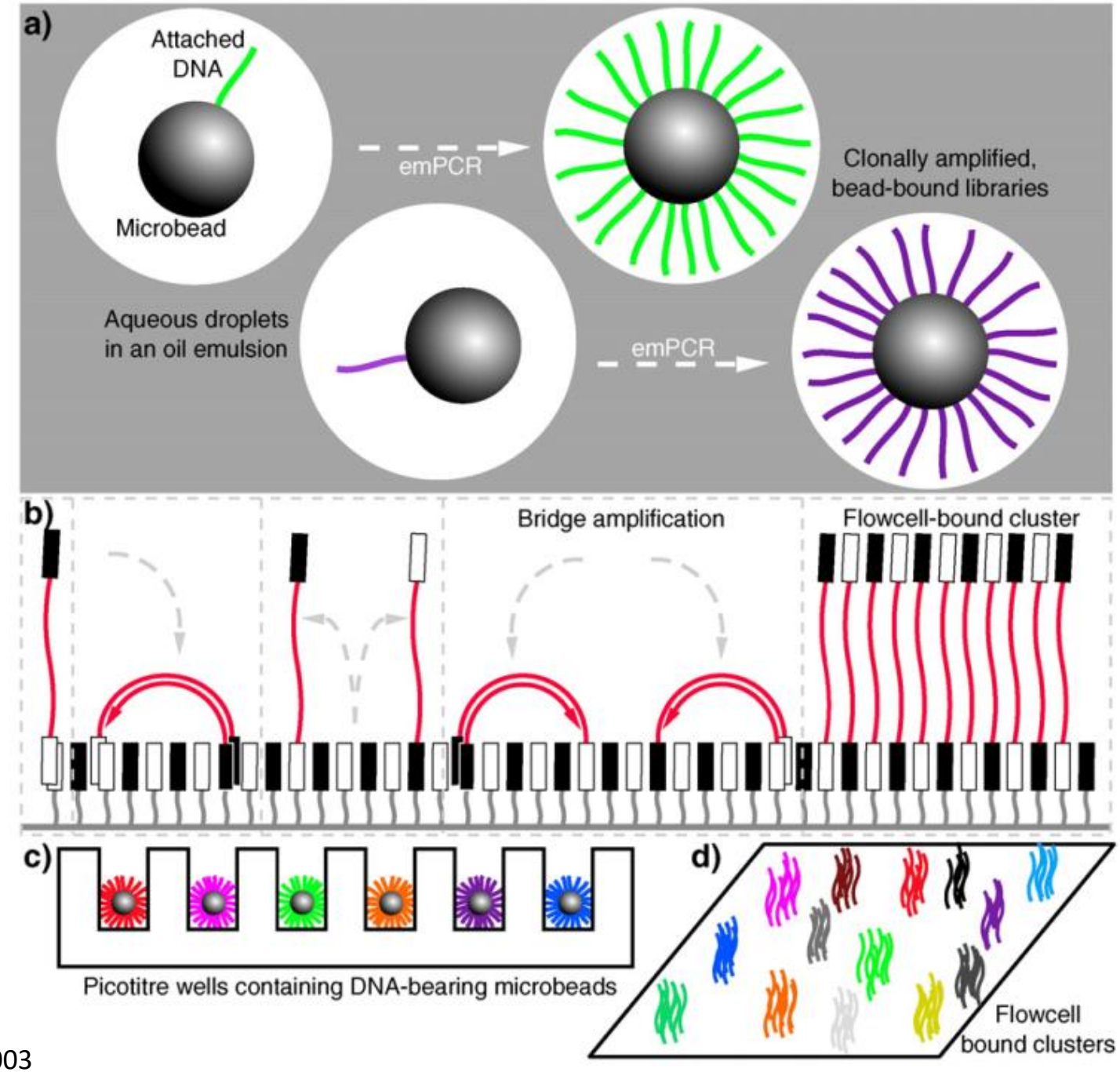
illumina®

ion torrent

by life technologies™

Emulsion PCR

Bridge PCR



https://doi.org/10.1016/j.ygeno.2015.11.003

Secuenciación por síntesis (SBS).

- **En la secuenciación por SBS se utiliza una polimerasa y se sensa una señal (fluoróforo o un cambio en la concentración iónica) que identifica la incorporación de un nucleótido en una cadena que se está sintetizando**

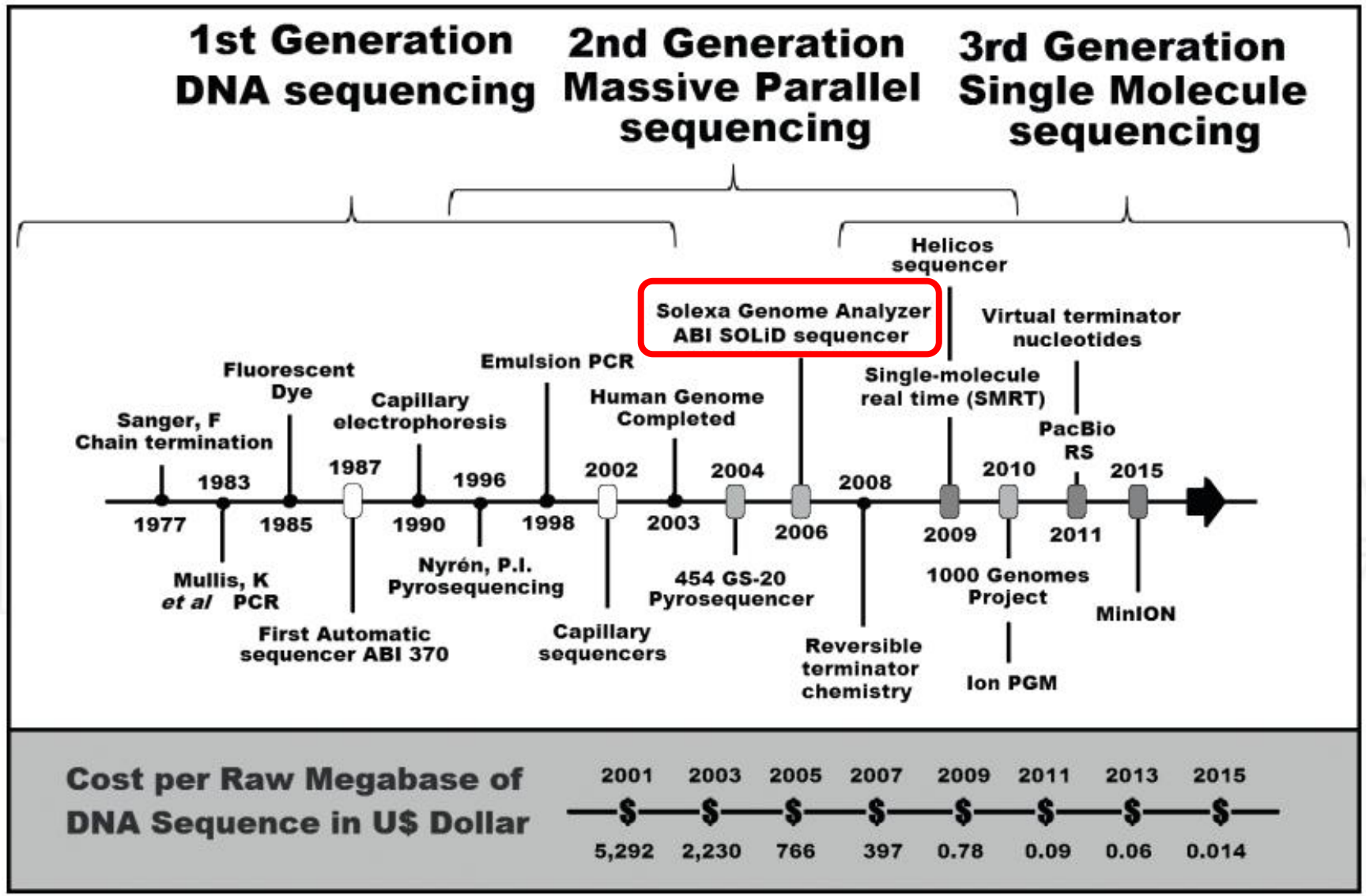
QIAGEN

Roche 454 SEQUENCING

ion torrent by life technologies

illumina®

Plataformas de Secuenciación

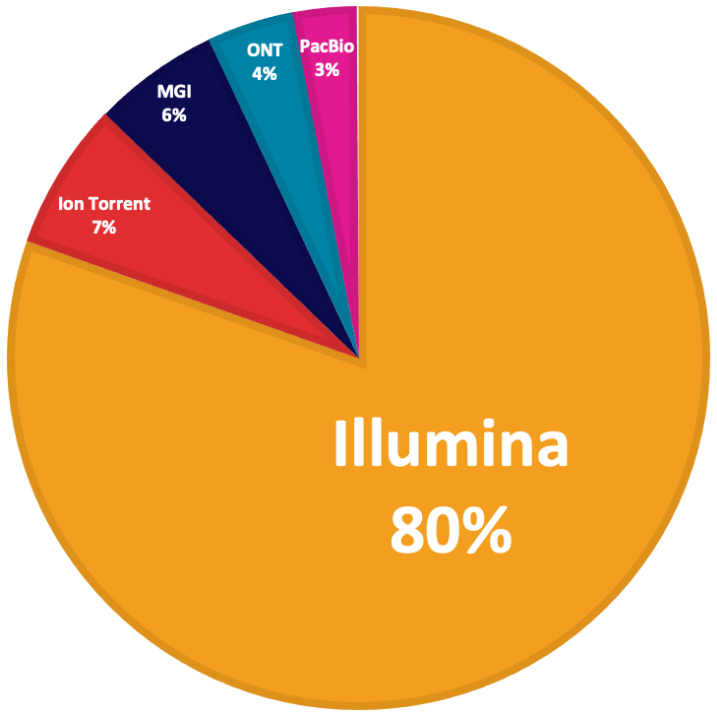


http://dx.doi.org/10.5772/intechopen.69337

Illumina

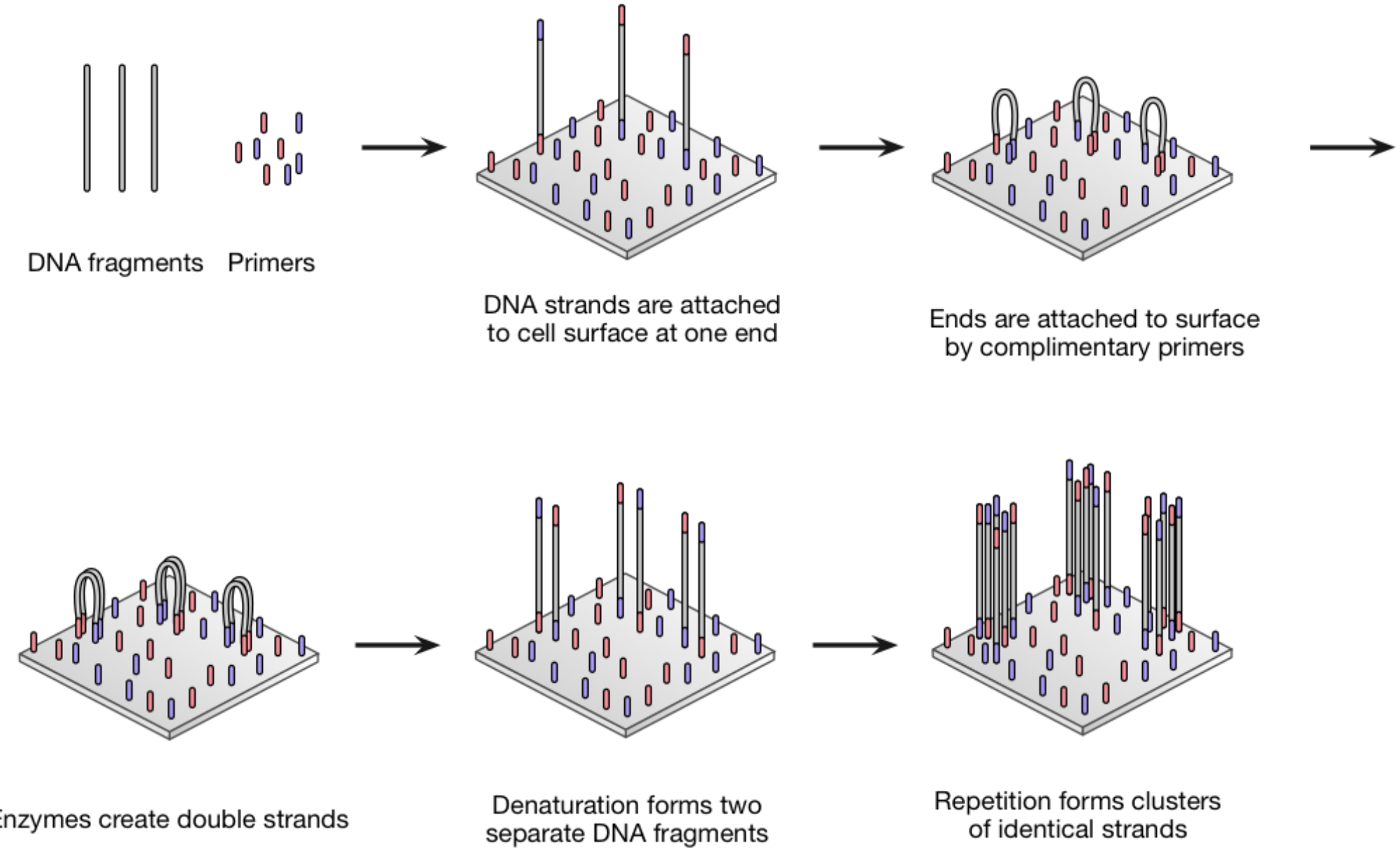
Illumina/Solexa lanzaron el Genome Analyzer II en 2006, y los avances en la tecnología de Illumina han marcado el ritmo de las reducciones en el costo.

La plataforma Illumina actualmente domina el mercado HTS.

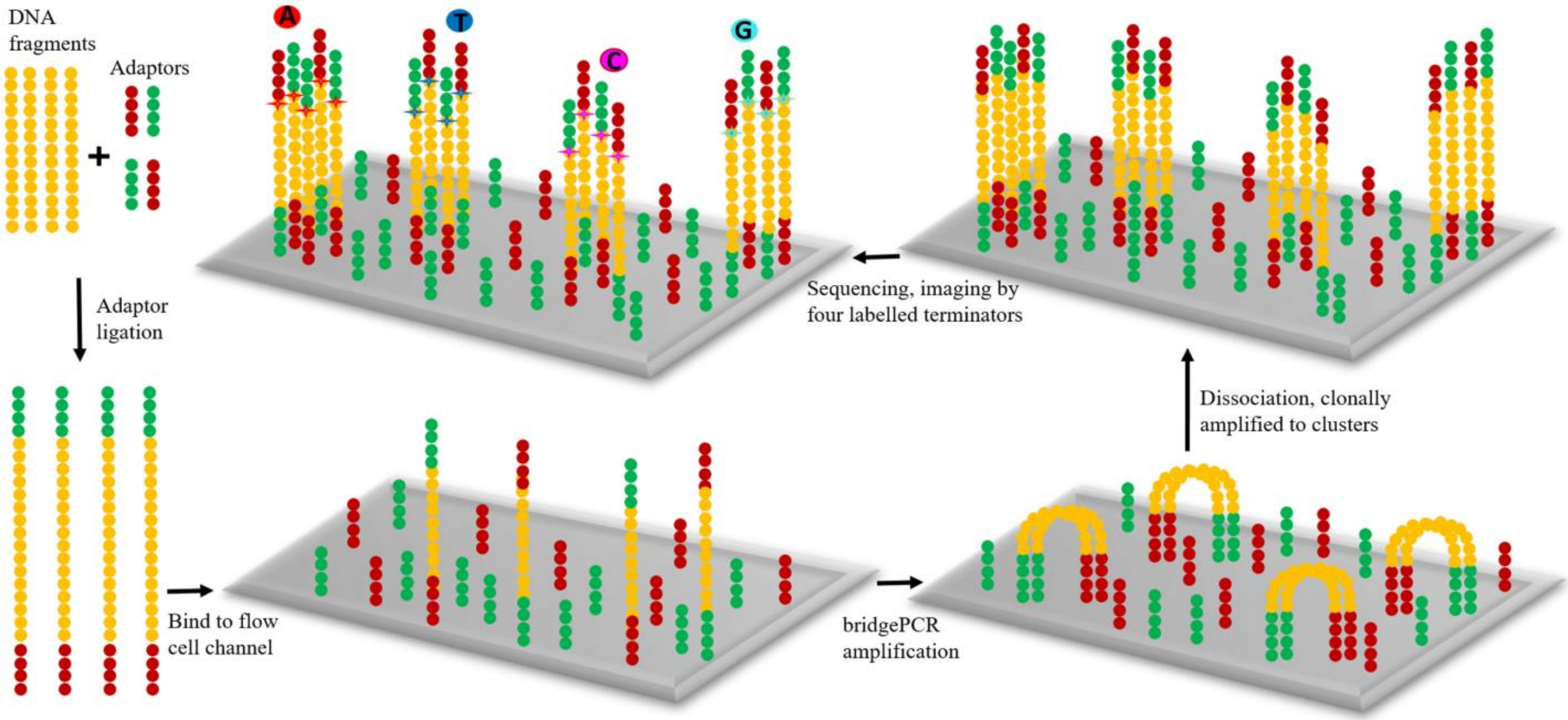


- **Amplificación clonal** de fragmentos de **ADN ligados al adaptador en la superficie** de un portaobjetos de vidrio.
- **Terminación cíclica reversible**.
- **Rondas progresivas de incorporación de las bases**.
- **Imágenes fluorescentes** para determinar el **nucleótido** agregado.
- La plataforma **MiSeq y HiSeqs** son las más establecidas

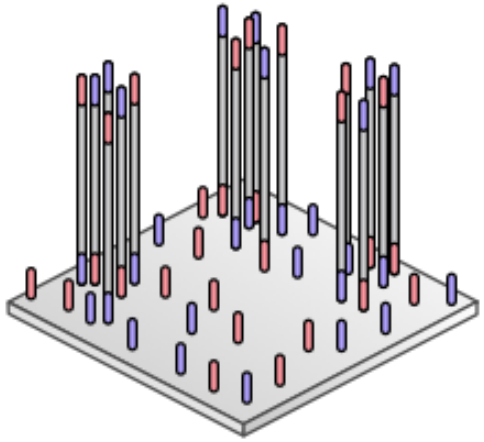
Illumina: amplificación en cluster ó en puente (Bridge)



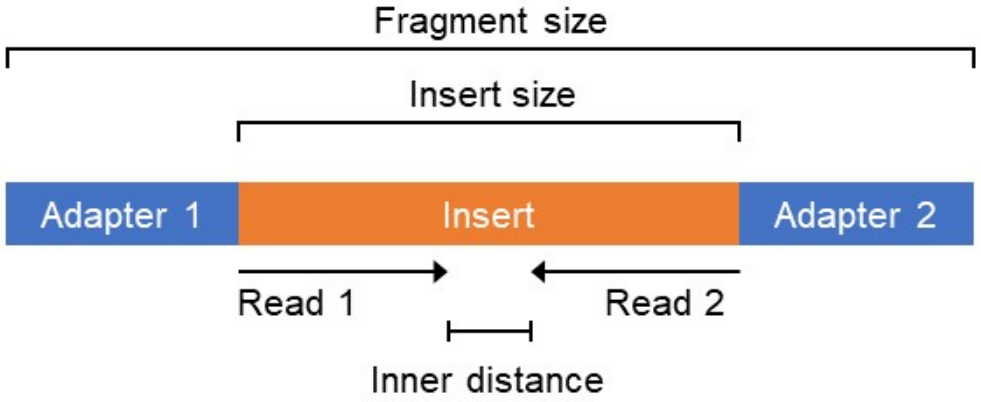
Illumina: amplificación en cluster ó en puente (Bridge)



Illumina: amplificación en cluster ó en puente (Bridge)

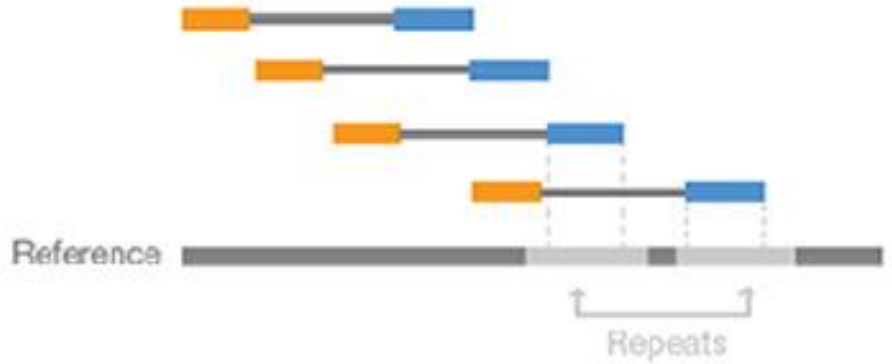


Repetition forms clusters of identical strands

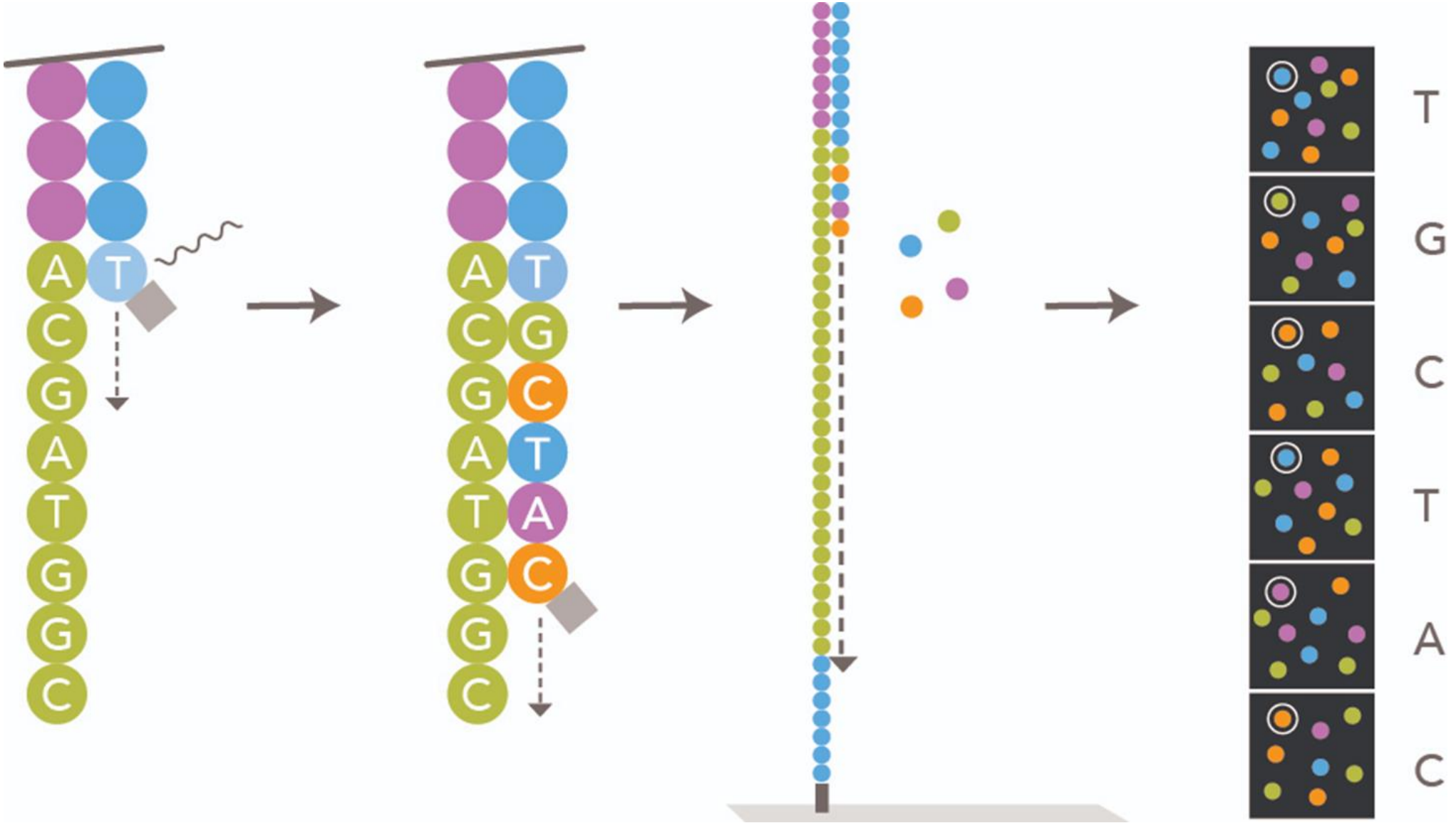


Paired-End Reads

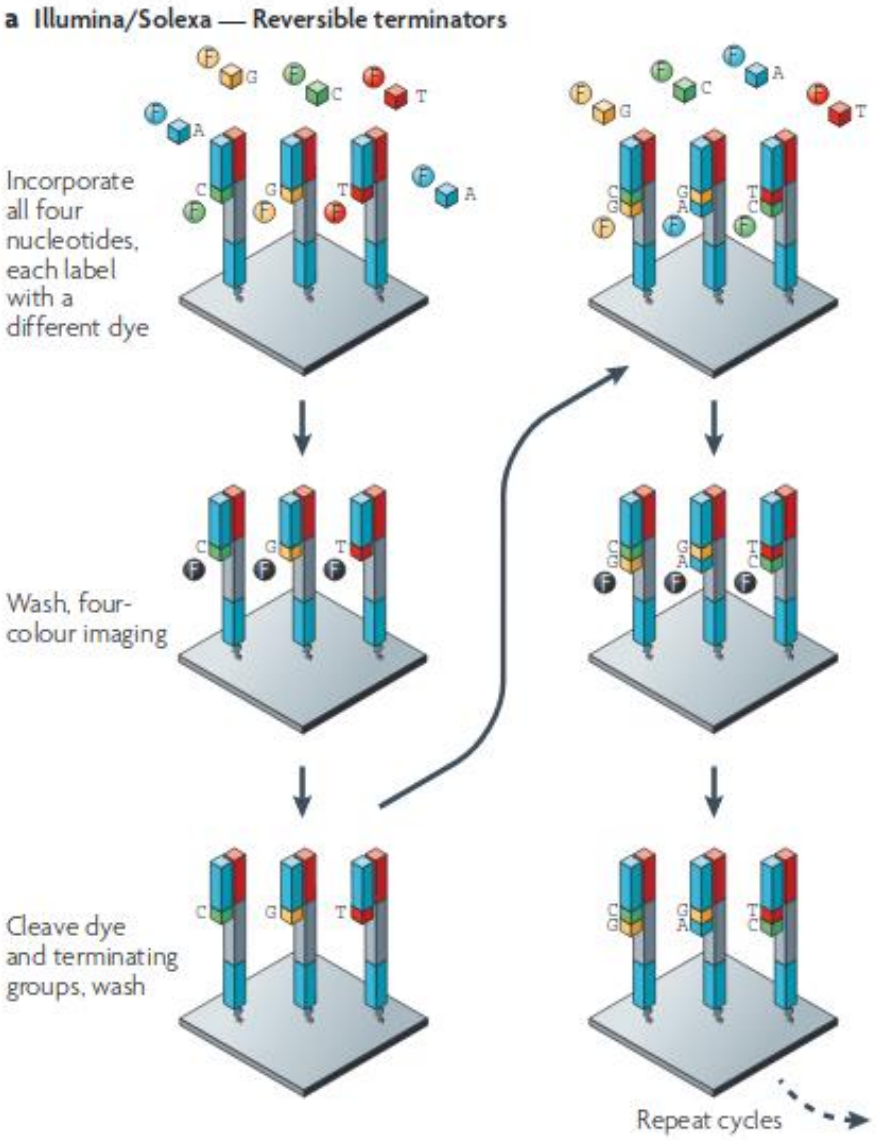
Alignment to the Reference Sequence

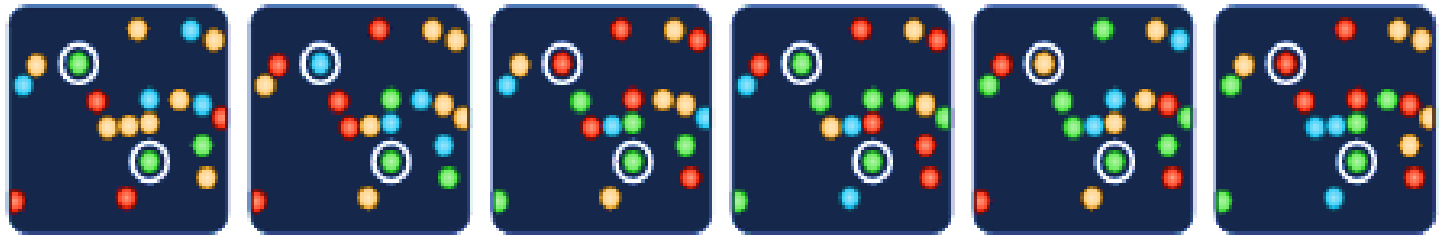


Illumina: secuenciación y análisis



Illumina: secuenciación y análisis





Top: CATCGT
Bottom: CCCCCC

Illumina: secuenciación y análisis

Nucleotide Flows

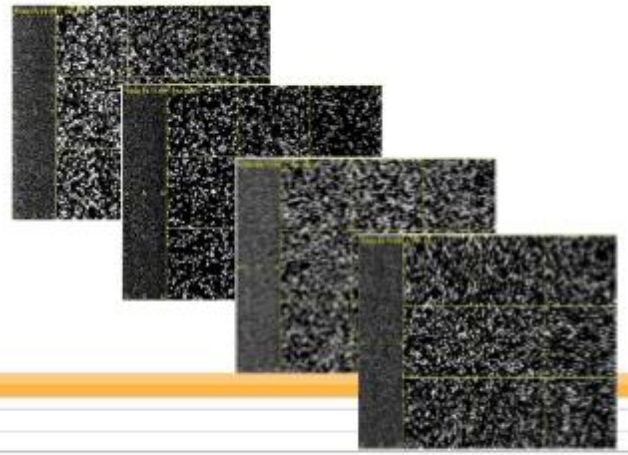
Raw Images

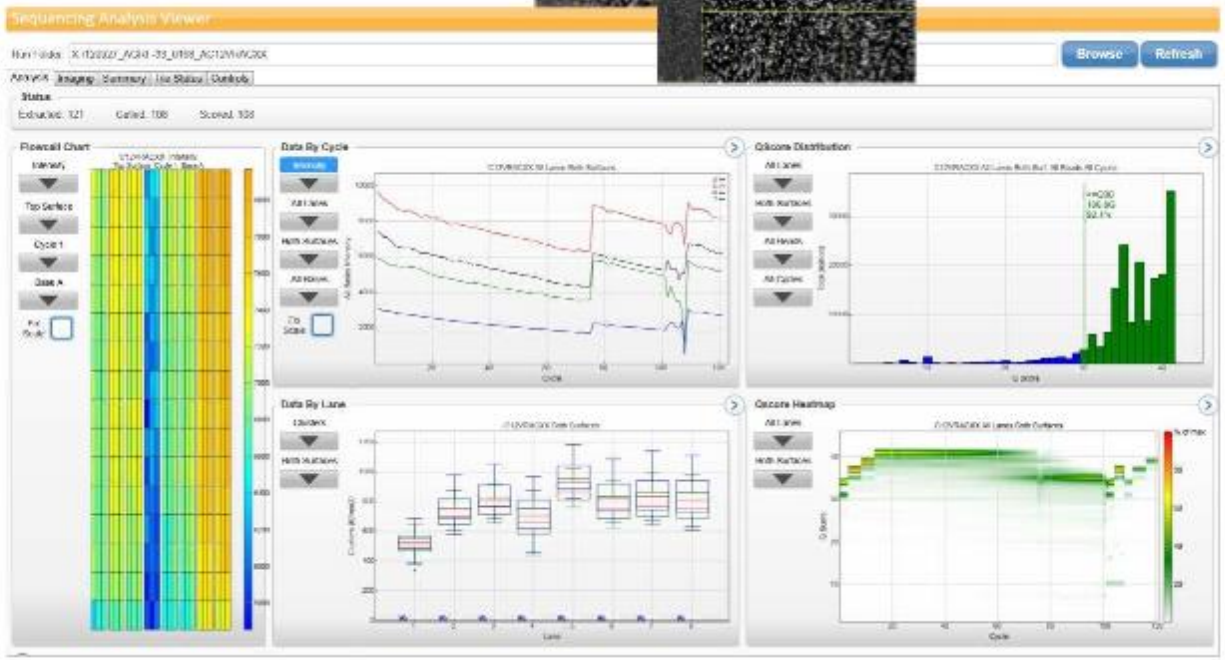
Image Processing

Base-calling

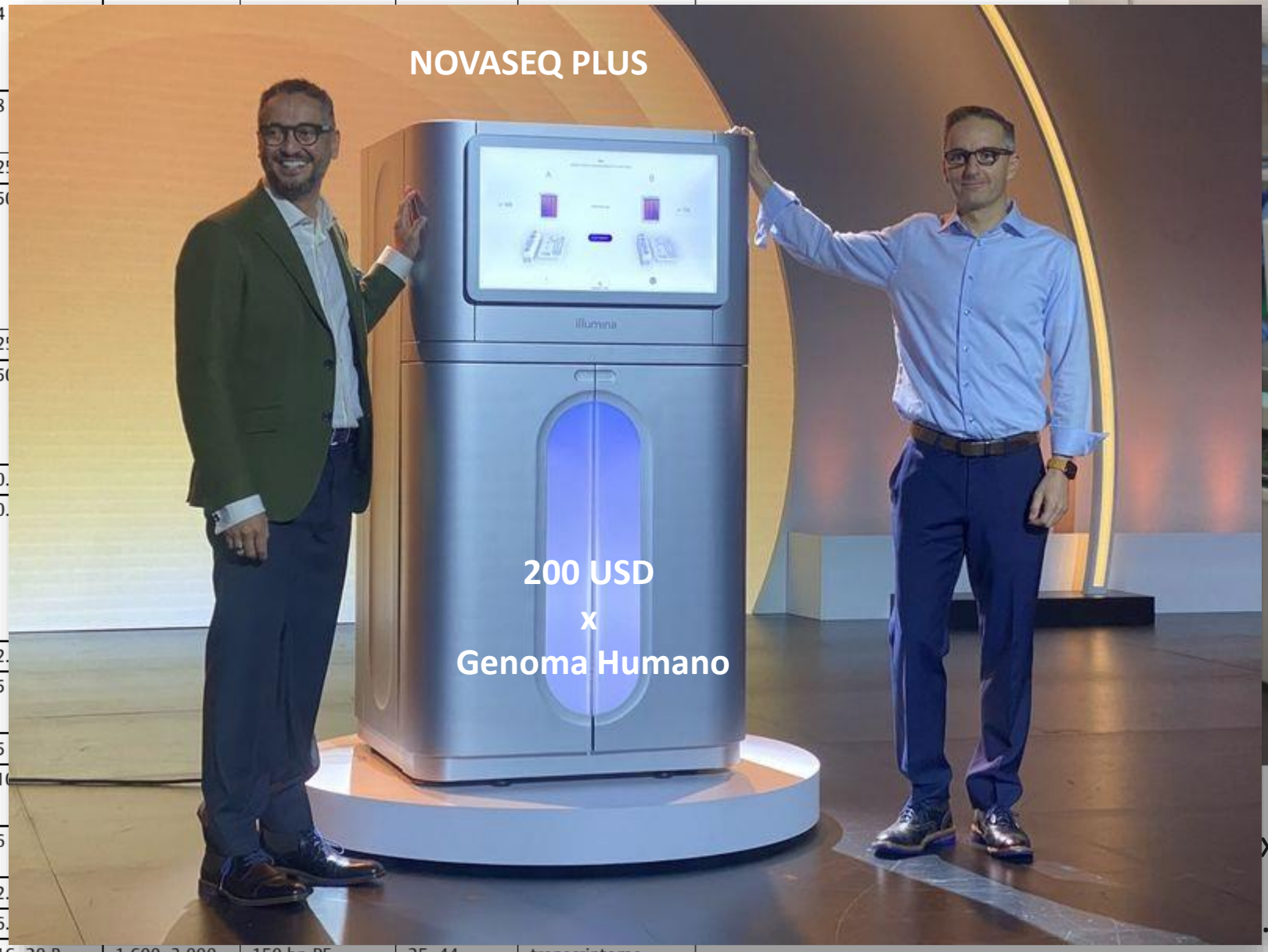
Quality Filtering

.bcl





	Instrument	Maximum reads per run	Maximum output range (gigabases)	Read length	Sequencing run time (hours)	Main applications
Illumina	Iseq 100	4				
		8				
	MiniSeq[a]	2				
		5				
	MiSeq[b,c]	2				
		5				
	NextSeq 550[a,b]	0.				
		0.				
	HiSeq 3000	2.				
		5				
	HiSeq 4000	5				
		1				
	HiSeqX[d]	6				
	NovaSeq 6000	2.				
		6.				
		16–20 B	1,600–3,000	150-bp PE	25–44	transcriptome sequencing, methylation sequencing

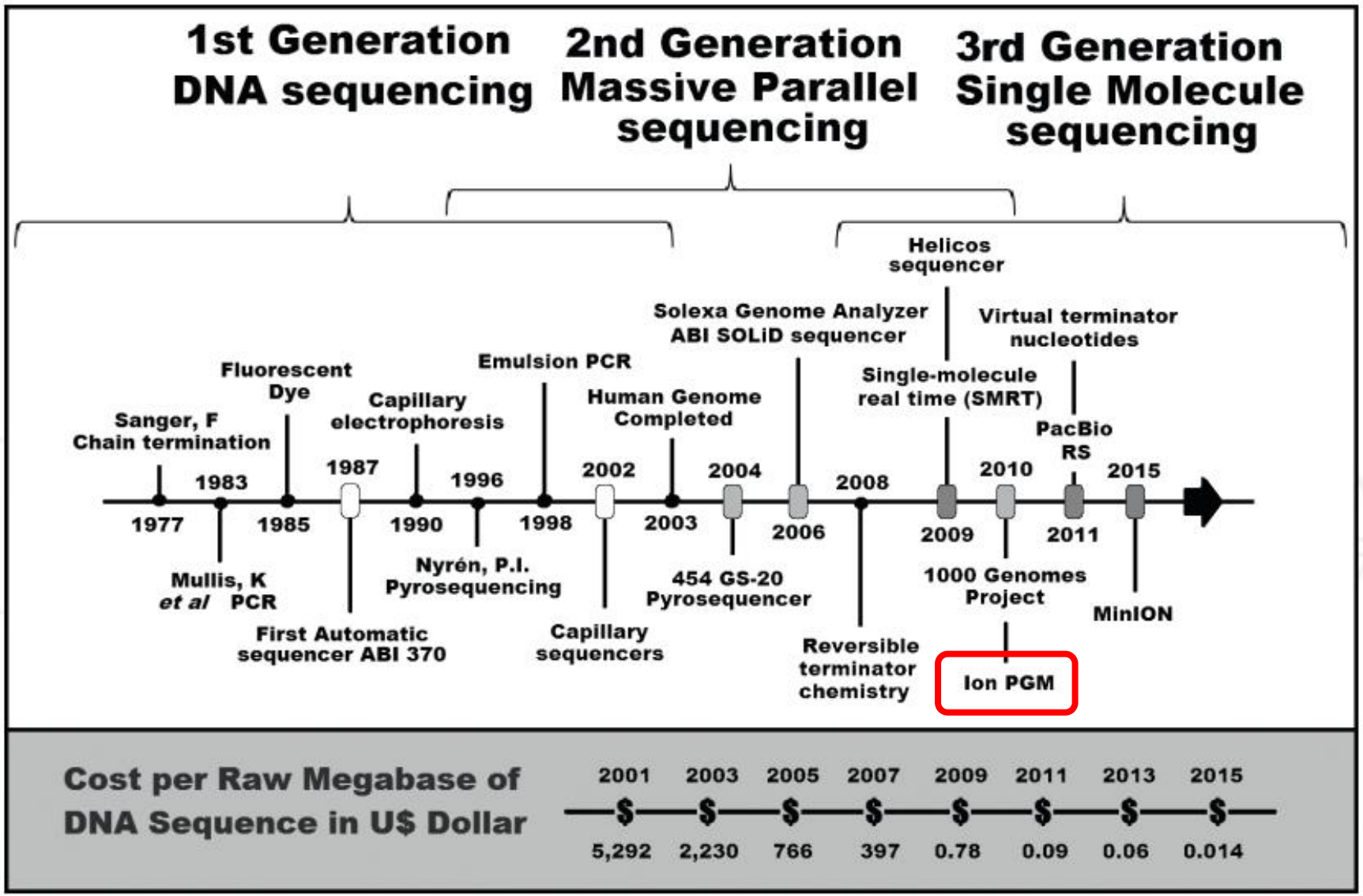


x) en 48hrs

. 25.000 USD

https://doi.org/10.1055/s-0039-1688446

Plataformas de Secuenciación



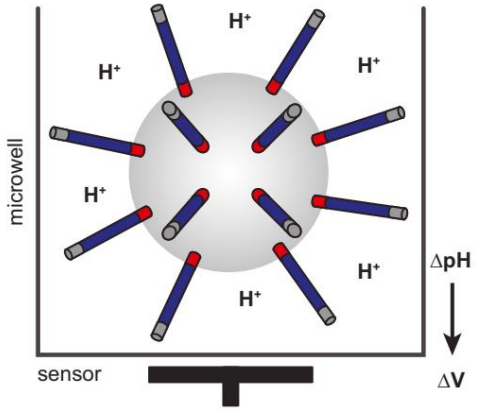
http://dx.doi.org/10.5772/intechopen.69337

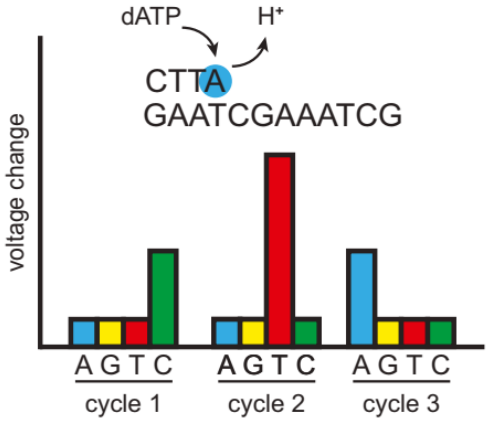
Ion Torrent

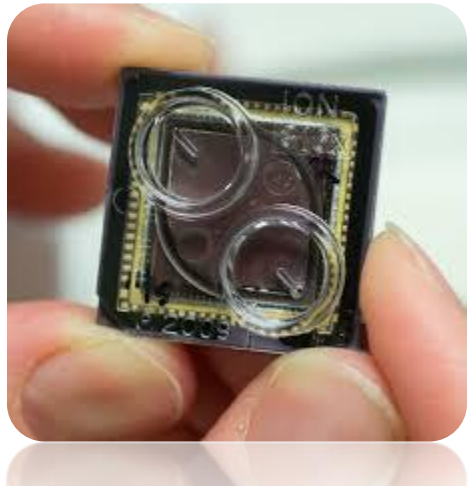
Life Technologies/ThermoFisher/Ion Torrent comercializó la tecnología de secuenciación de semiconductores de **Ion Torrent en 2010**, secuenciador Ion PGM de mesada.

La **emulsión-PCR** se usa para amplificar clonalmente fragmentos de ADN ligados al adaptador en la superficie de las perlas.

- El Ion Torrent fue la primera plataforma NGS **sin detección óptica**, acelerando drásticamente las ejecuciones de secuenciación y reduciendo los costos.
- Ion Torrent **detecta los iones H +** que se liberan a medida que se incorpora cada dNTP.

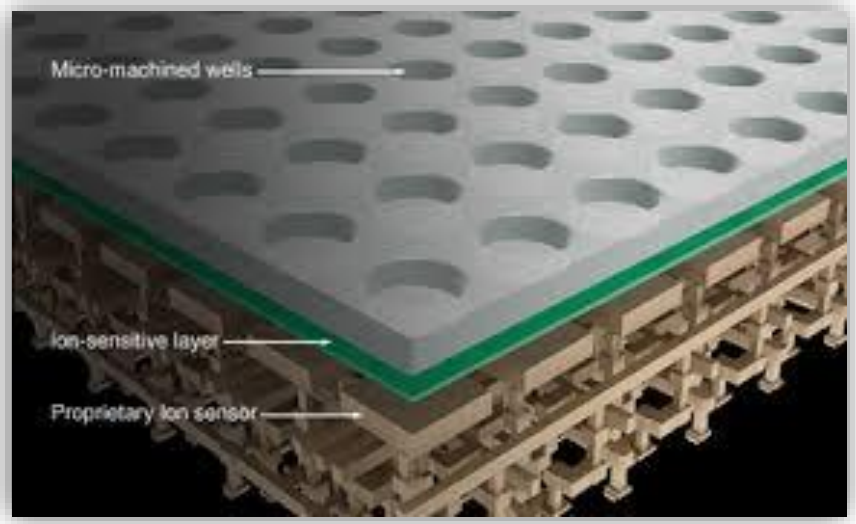


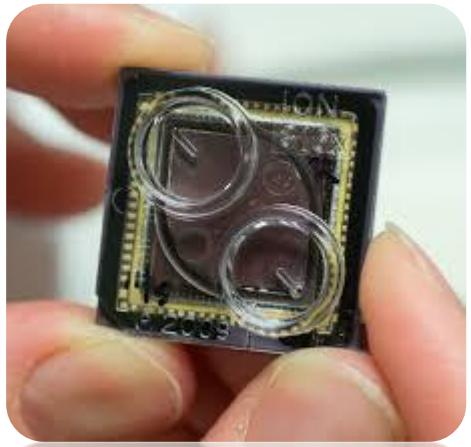




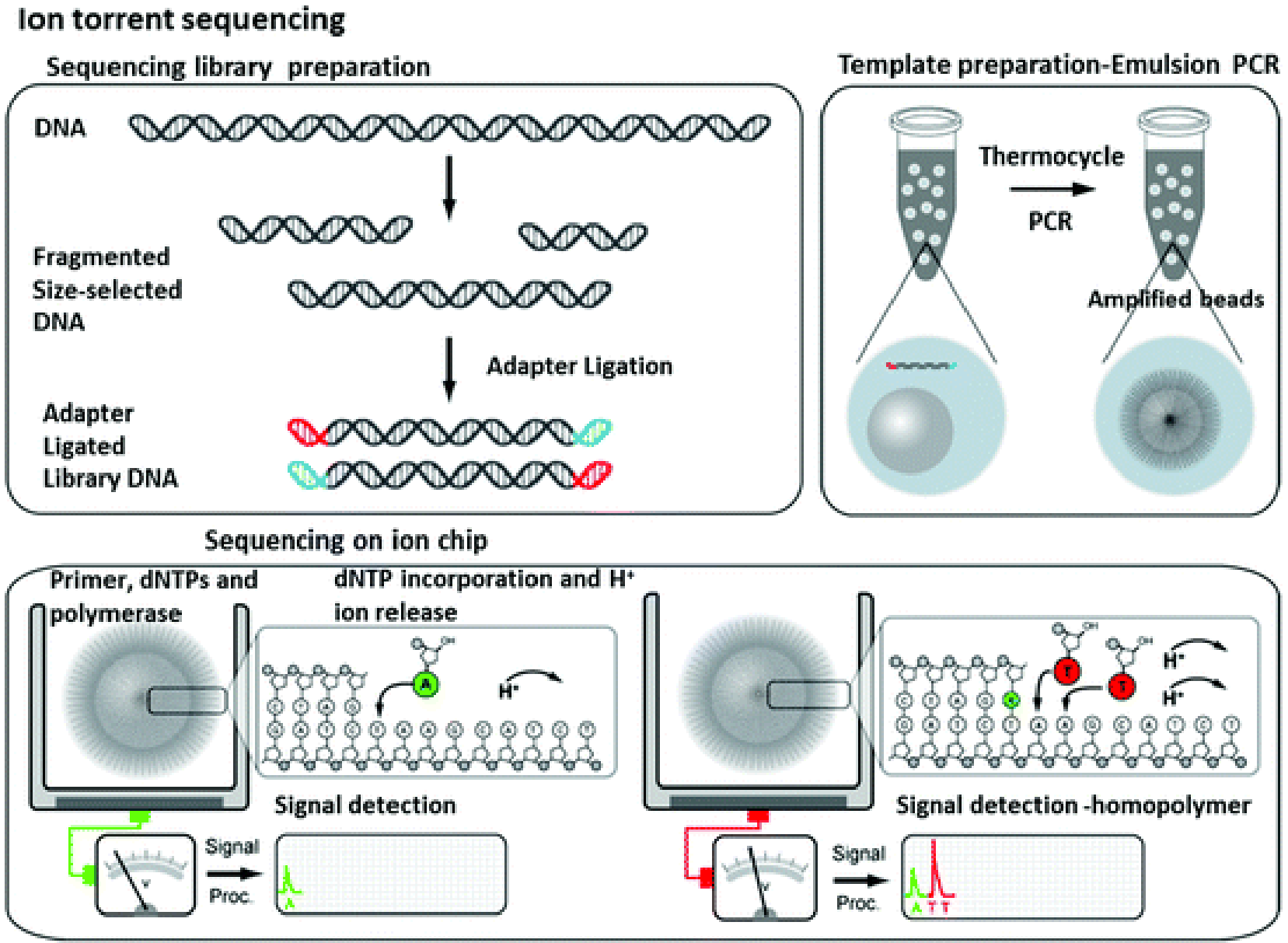
Ion Torrent

- Las **perlas (Beads) se distribuyen posteriormente en micropocillos** donde se produce una reacción de secuenciación por síntesis.
- Ion Torrent **detecta los iones H +** que se liberan a medida que se incorpora cada dNTP.
- Estos **cambios de pH** son detectados por un sensor ubicado en la parte inferior del micropocillo y **convertido en una señal de voltaje**.
- La **señal de voltaje es proporcional al número de bases incorporadas**.
- La **velocidad de secuenciación, 2-8 horas** dependiendo de la máquina y el chip utilizado, hace que estos secuenciadores sean particularmente útiles para aplicaciones clínicas.

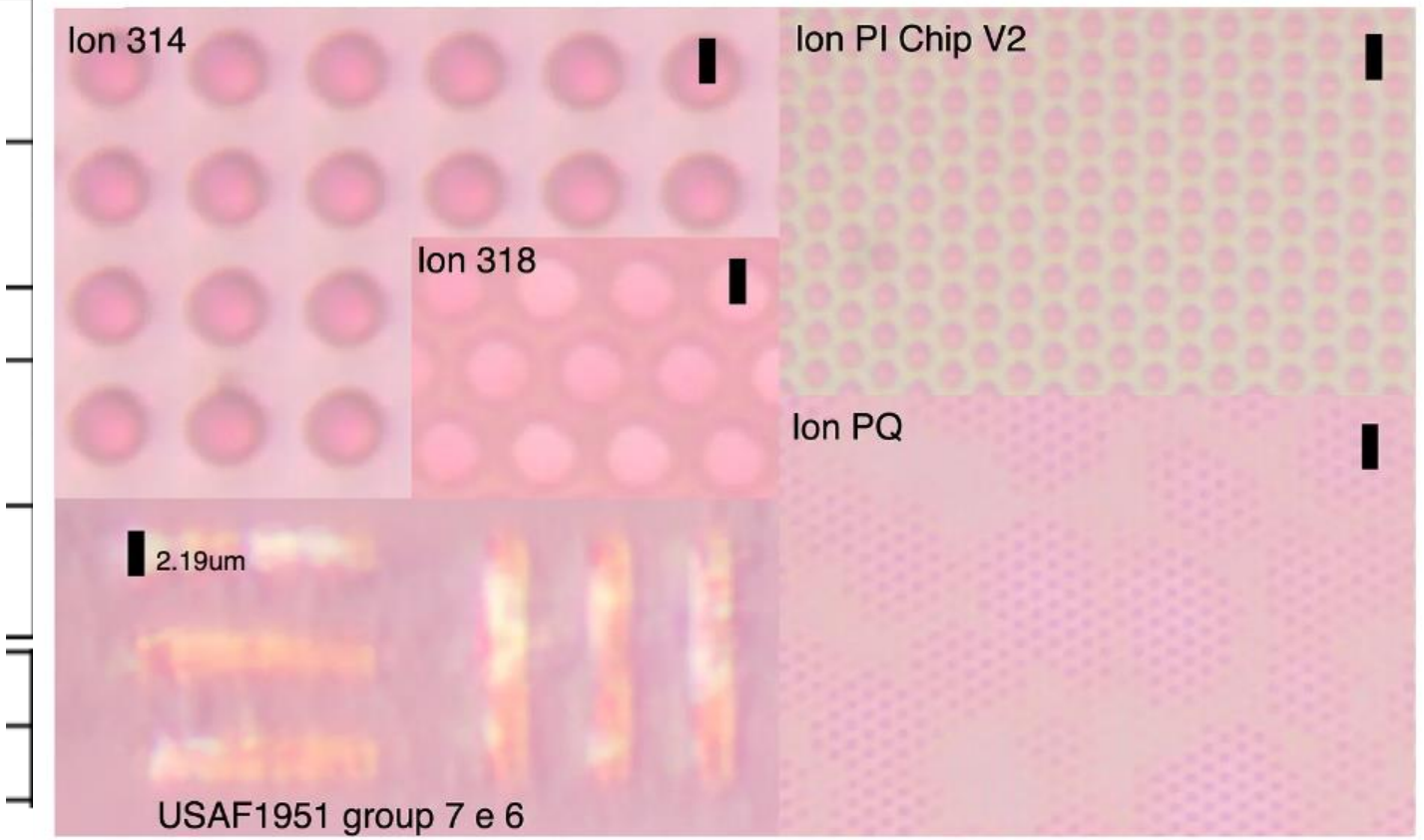




Ion Torrent


Ion torrent sequencing
Sequencing library preparation
DNA
Fragmented
Size-selected
DNA
Adapter Ligation
Adapter
Ligated
Library DNA
Template preparation-Emulsion PCR
Thermocycle
PCR
Amplified beads
Sequencing on ion chip
Primer, dNTPs and
polymerase
dNTP incorporation and H+
ion release
Signal detection
Signal
Proc.
Signal detection -homopolymer
Signal
Proc.

	Instrument	Maximum reads per run	Maximum output range (gigabases)	Read length	Sequencing run time (hours)	Main applications

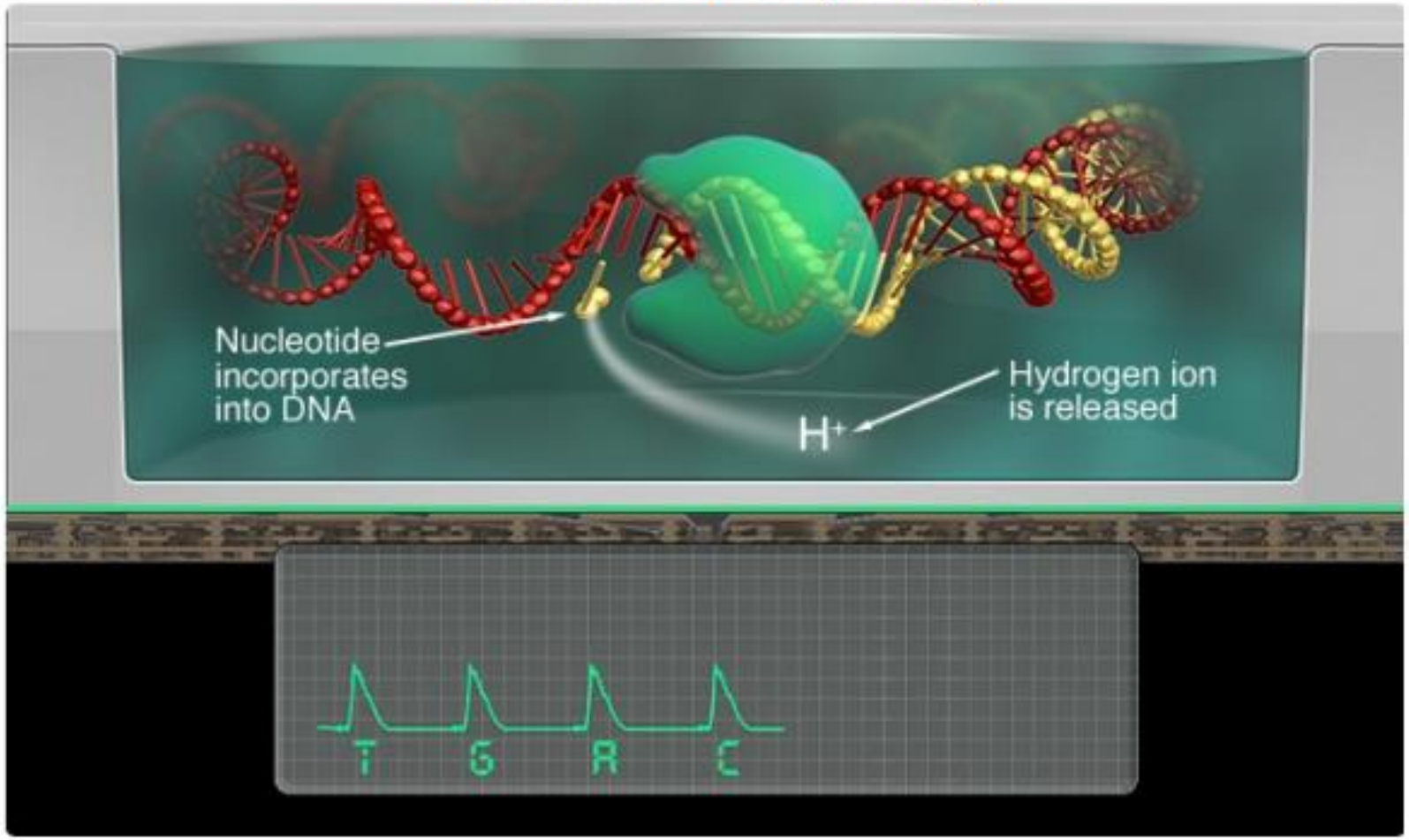


Ion 314 (~3um wells), Ion 318 (~3um wells), Ion PI (~1.5um wells), Ion PQ (<0.7um wells)

https://doi.org/10.1055/s-0039-1688446

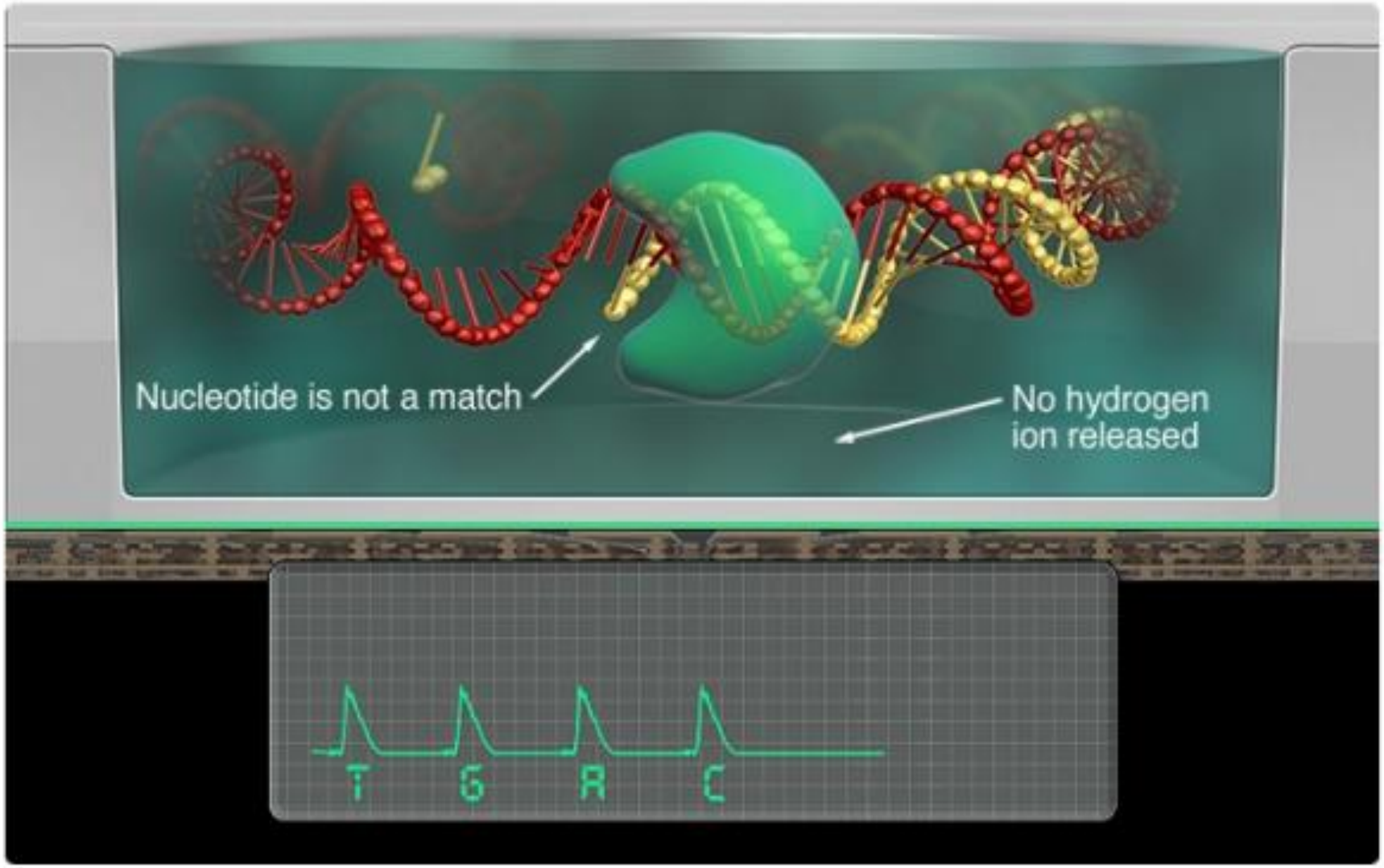
Ion Torrent

4 nucleotides flow sequentially

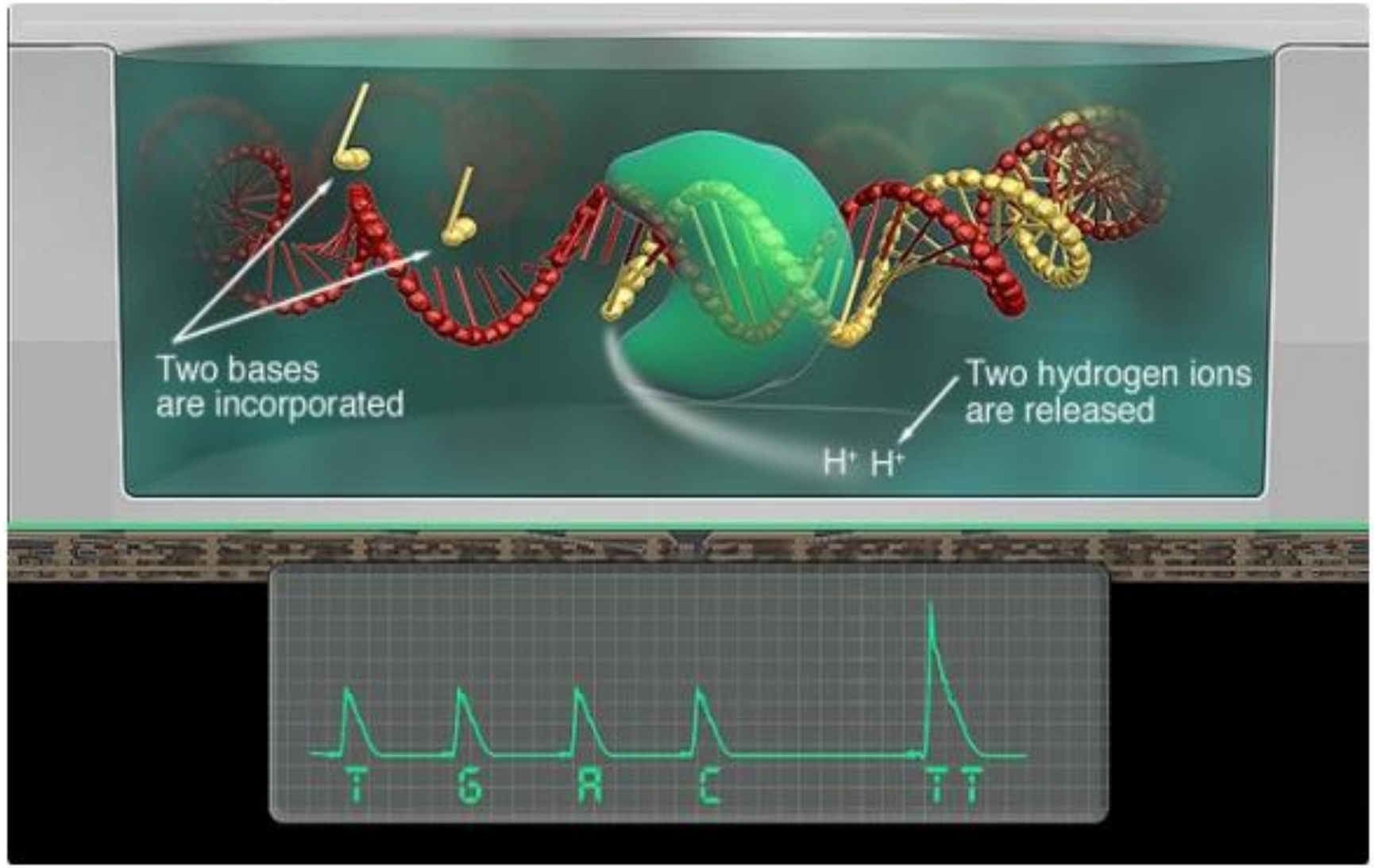


No camera, just a pH sensor

Ion Torrent

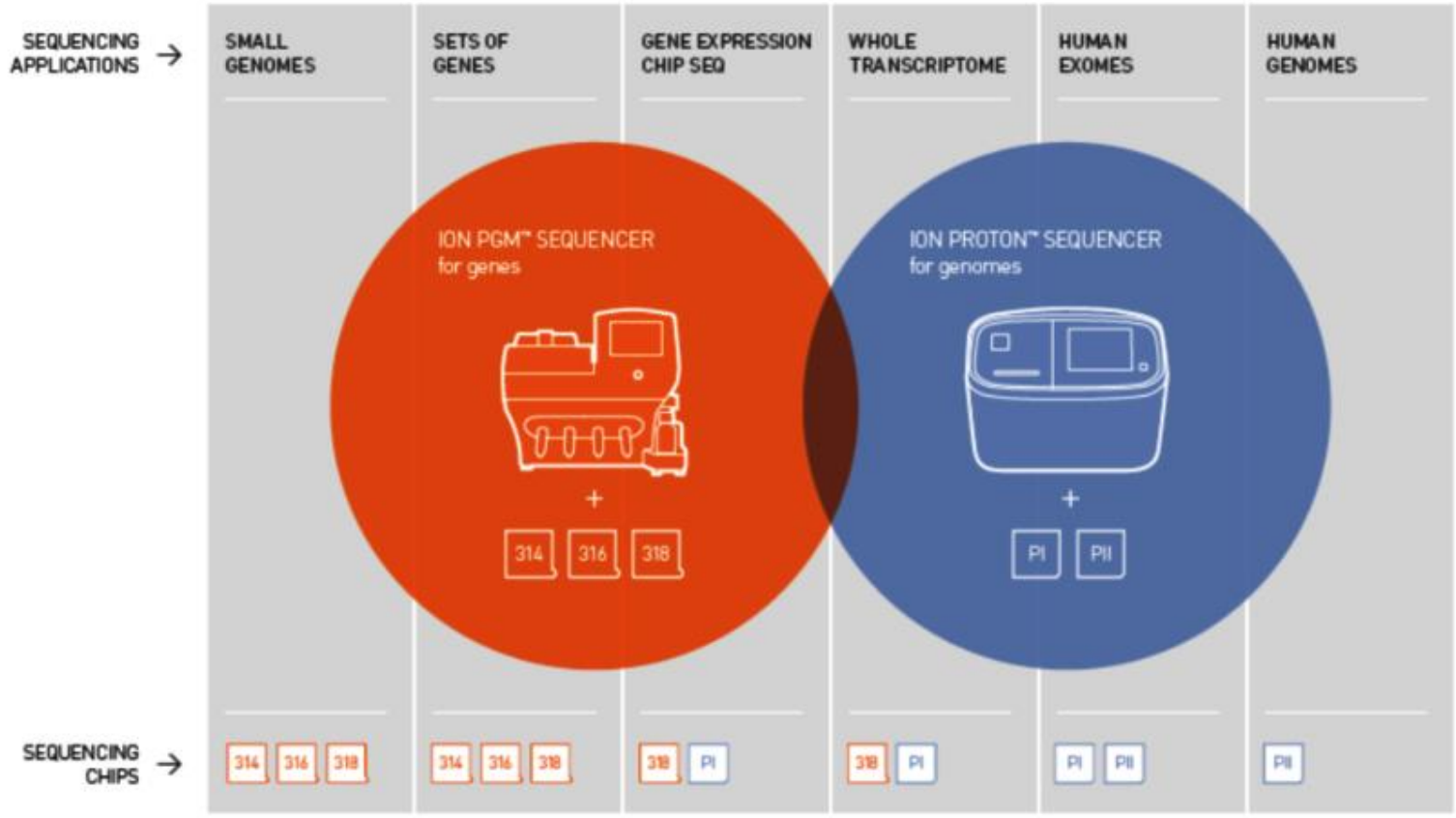


Ion Torrent


Two bases
are incorporated
Two hydrogen ions
are released
H+ H+
T G A C T T

ION S5

10-20 x PGM



Ion 510™ Chip
2–3 M reads

Ion 520™ Chip
3–6 M reads

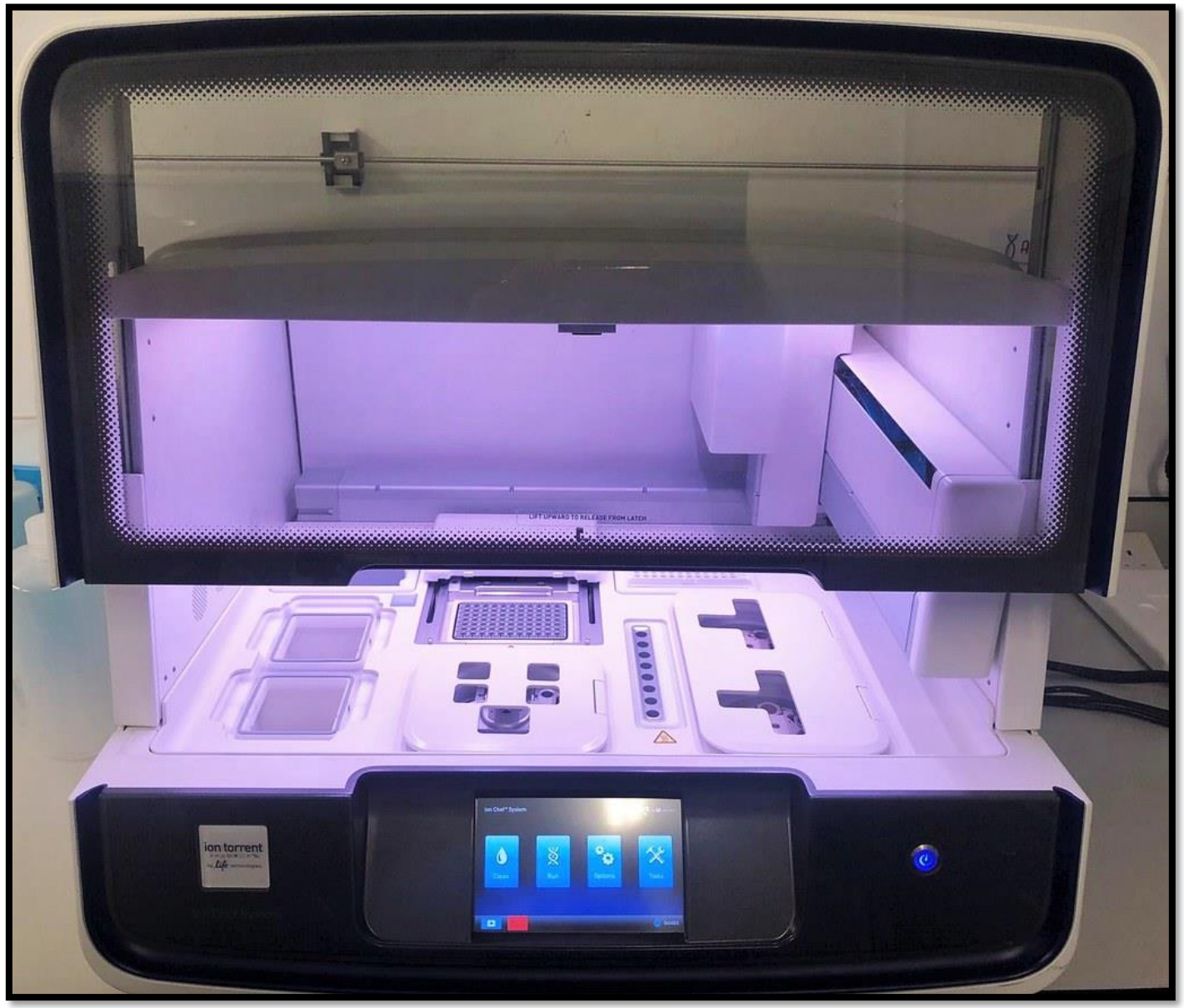
Ion 530™ Chip
15–20 M reads

Ion 540™ Chip
60–80 M reads

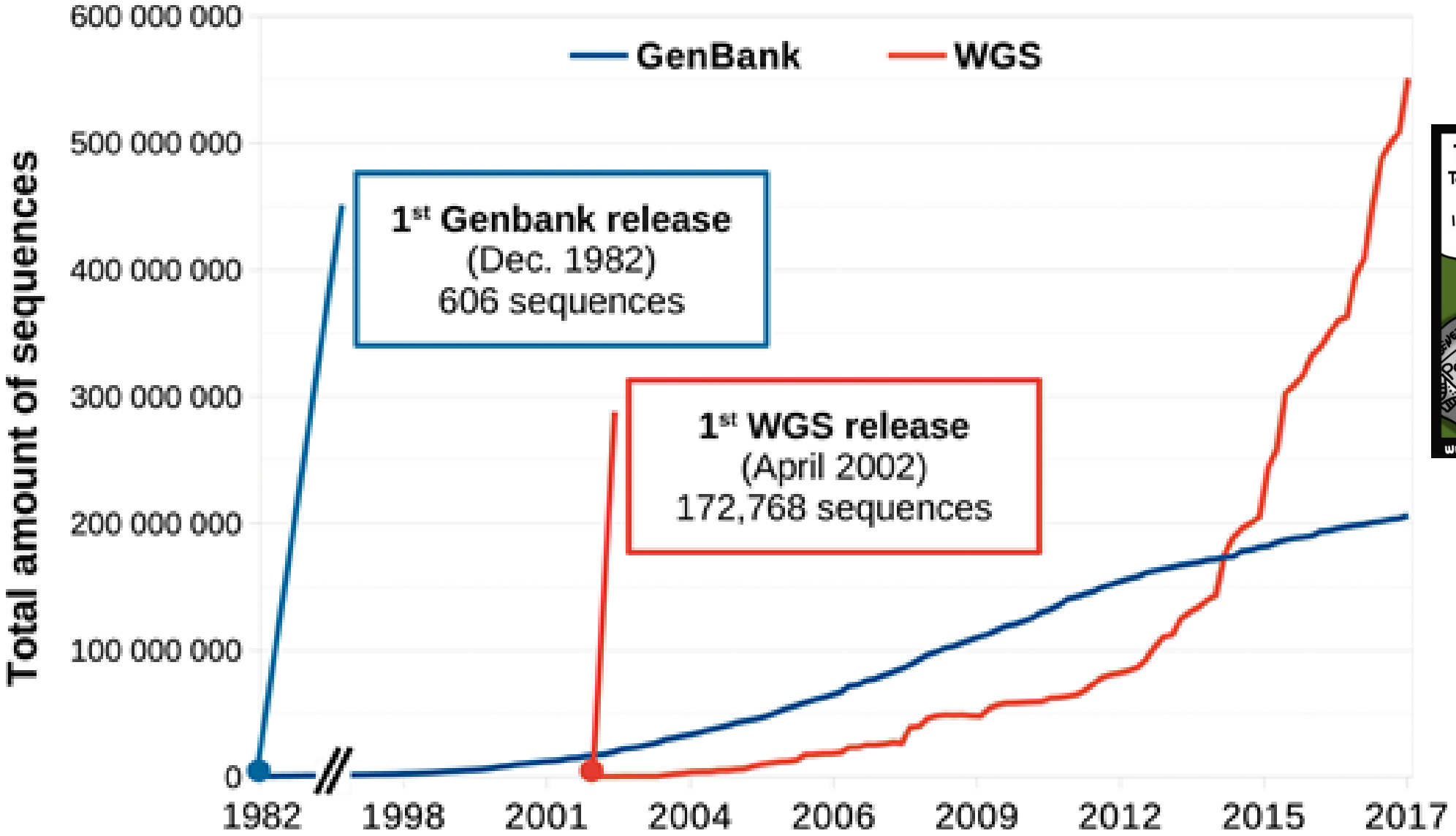
New

Ion 550™ Chip
100–130 M reads

Automización en la construcción de Librerías

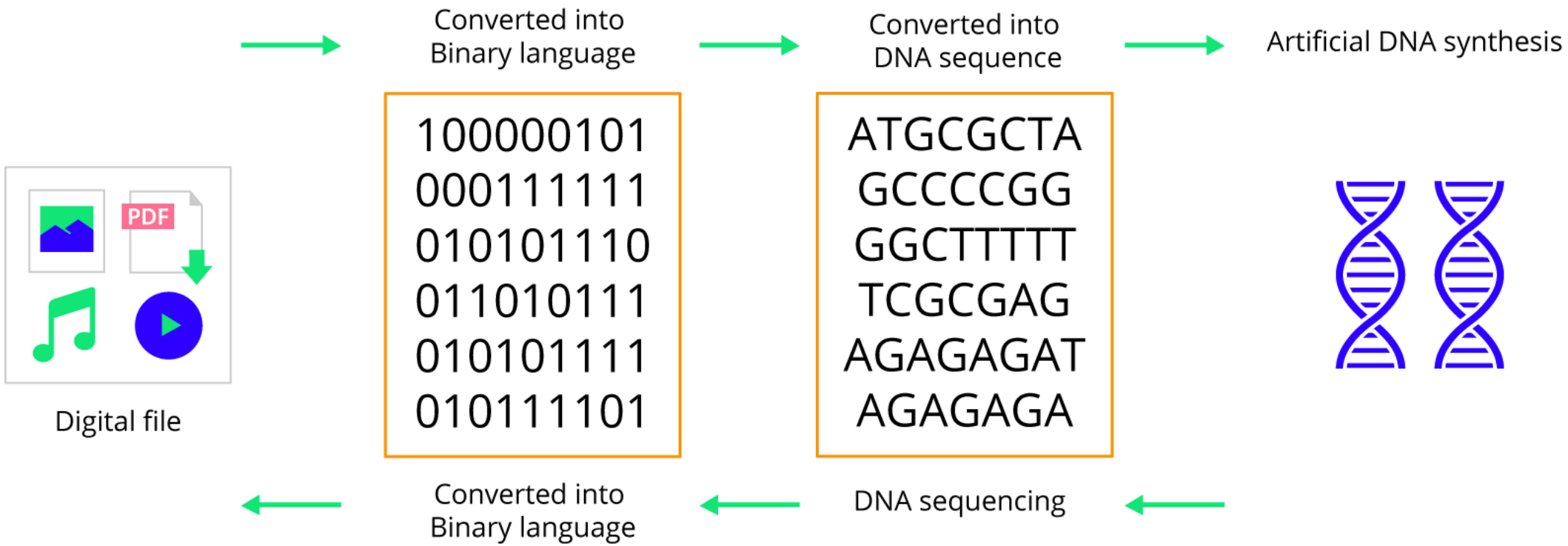


ION CHEF


Total amount of sequences
600 000 000
500 000 000
400 000 000
300 000 000
200 000 000
100 000 000
0
1982
1998
2001
2004
2006
2009
2012
2015
2017
GenBank
WGS
1st Genbank release
(Dec. 1982)
606 sequences
1st WGS release
(April 2002)
172,768 sequences


Too much pressure...
Too much expectations...
I can't stand it anymore!
NEWS
GENOMIC NEWS
POWER OF DNA
www.biocomicals.com

Figure 1 - *Process of DNA digital data storage*



Adapted from Source: Genetic Education, DNA digital data storage by Dr Tushar Chauhan

Original Image

Image Reconstructed From Bacteria

LETTER

doi:10.1038/nature23017

CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria

Seth L. Shipman[1,2,3], Jeff Nivala[1,3], Jeffrey D. Macklis[2] & George M. Church[1,3]

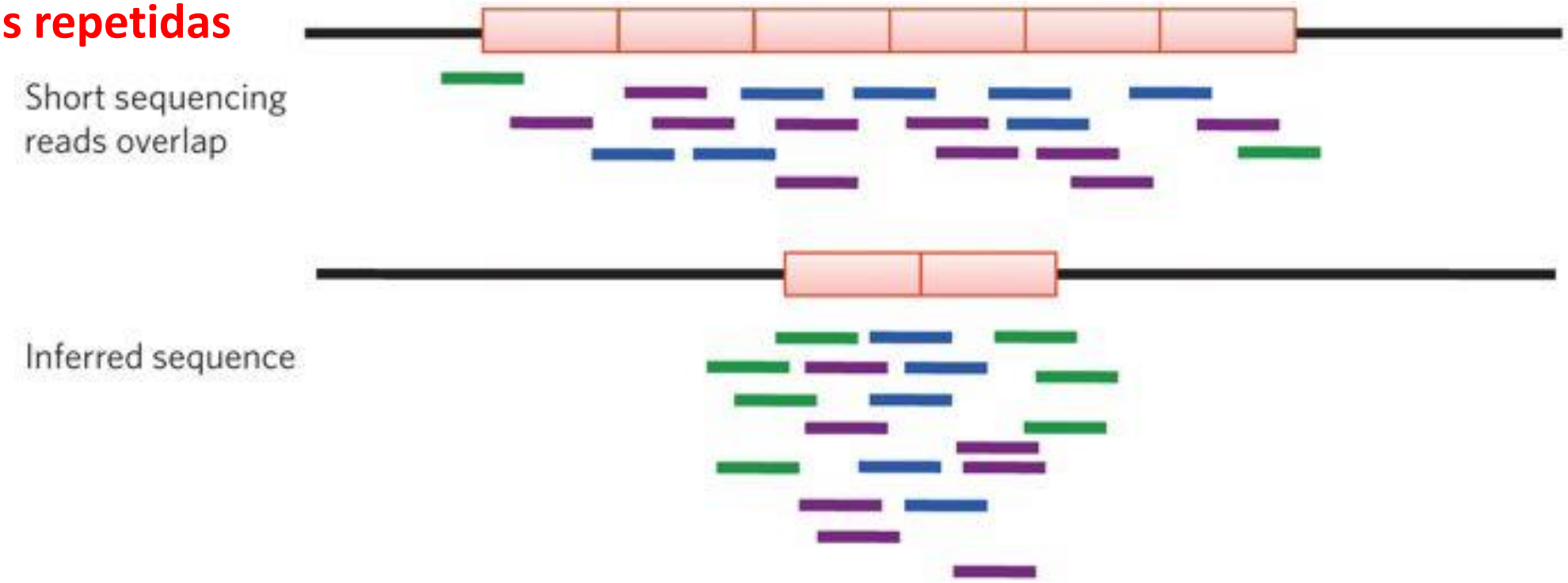
https://www.technologyreview.com/2017/07/12/150497/scientists-used-crispr-to-put-a-gif-inside-living-dna/ Shipman, S., *et al. Nature*, (2017)

¿Está todo resuelto?

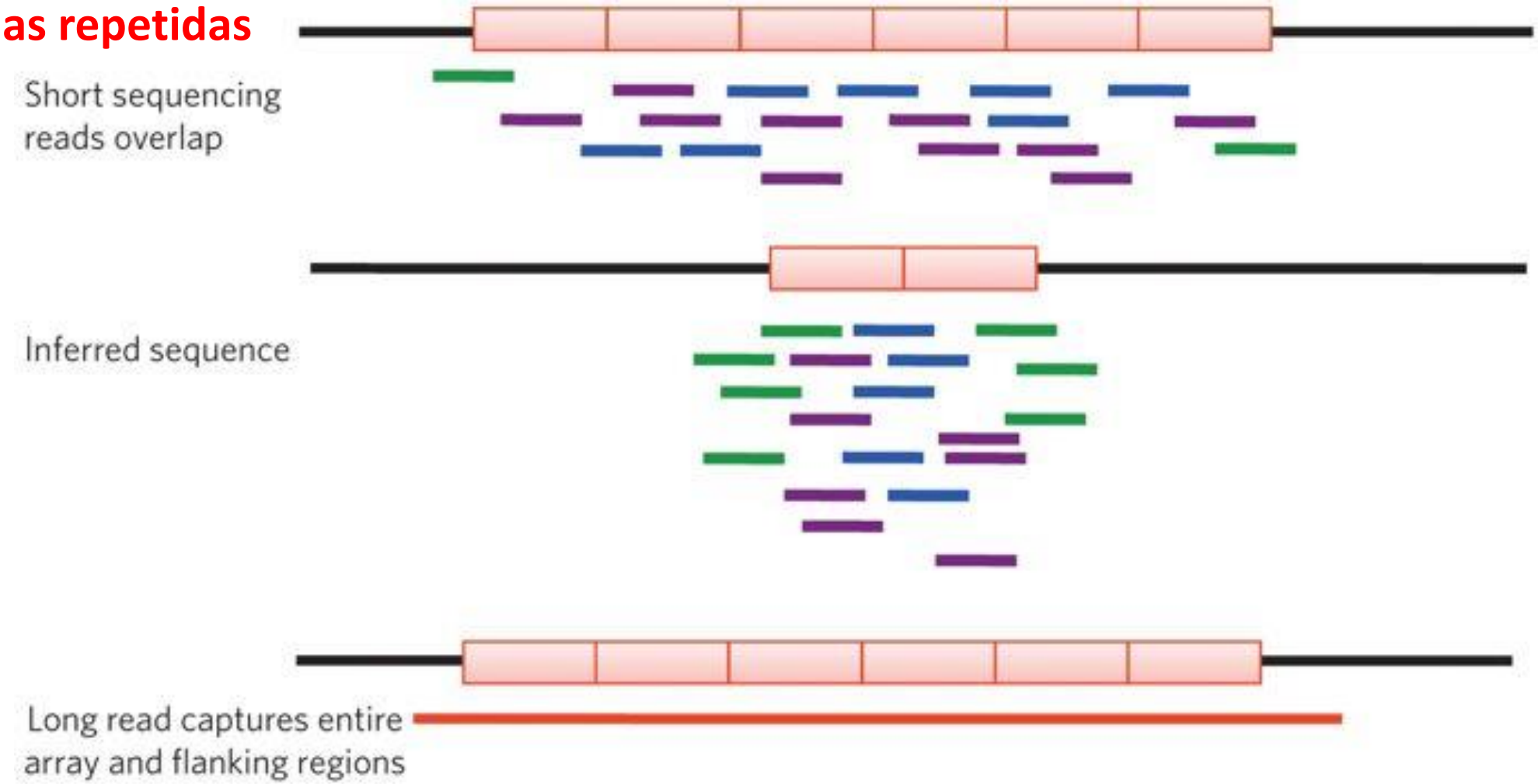
Igual plataforma diferente librería

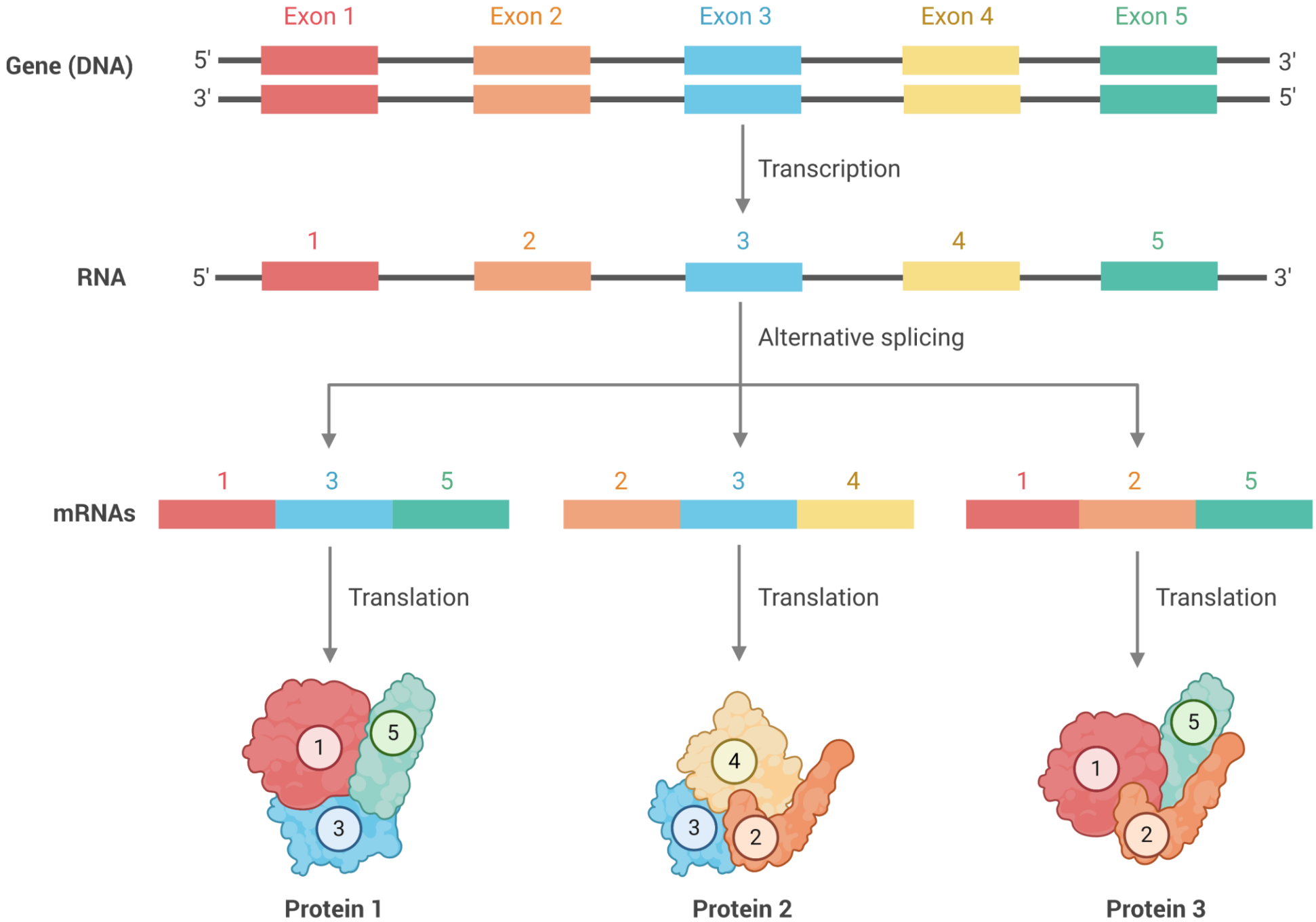
¿Está todo resuelto?

Secuencias repetidas

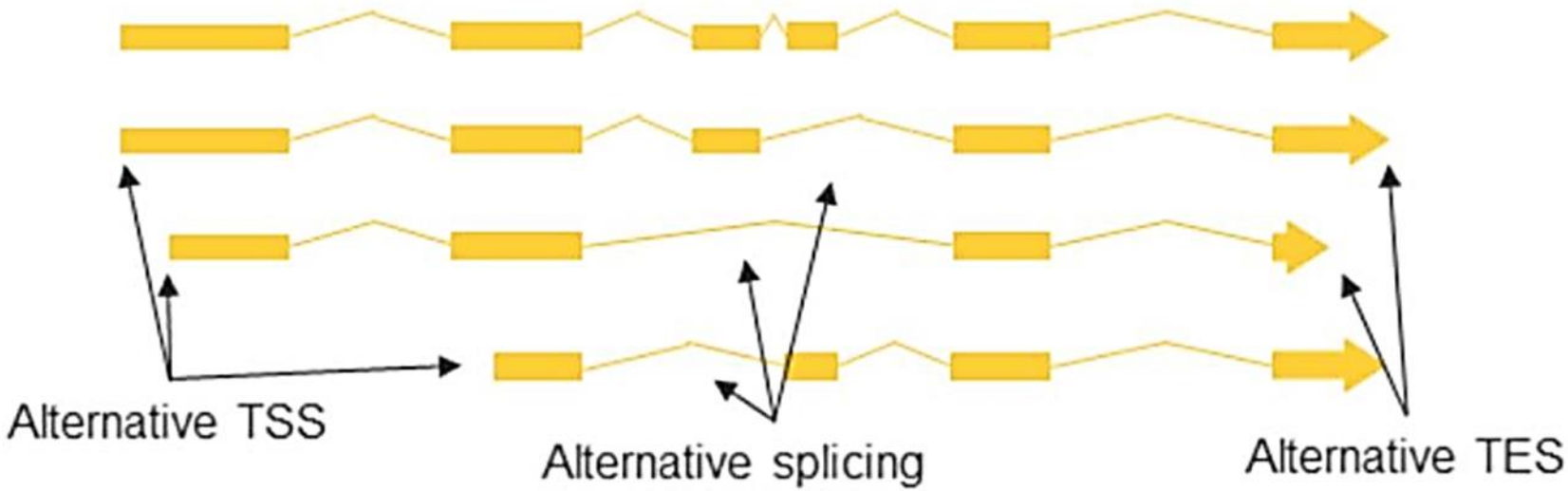


Secuencias repetidas




Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
Gene (DNA)
5'
3'
3'
5'
Transcription
RNA
1
2
3
4
5
5'
3'
Alternative splicing
mRNAs
1 3 5
2 3 4
1 2 5
Translation
Translation
Translation
Protein 1
Protein 2
Protein 3

Isoformas de transcritos



Short-read sequencing

Long-read sequencing

Secuenciación del ADN

- **Secuenciación de 1^{era} generación**
  - Método de Fred Sanger
    - Método enzimático de secuenciación

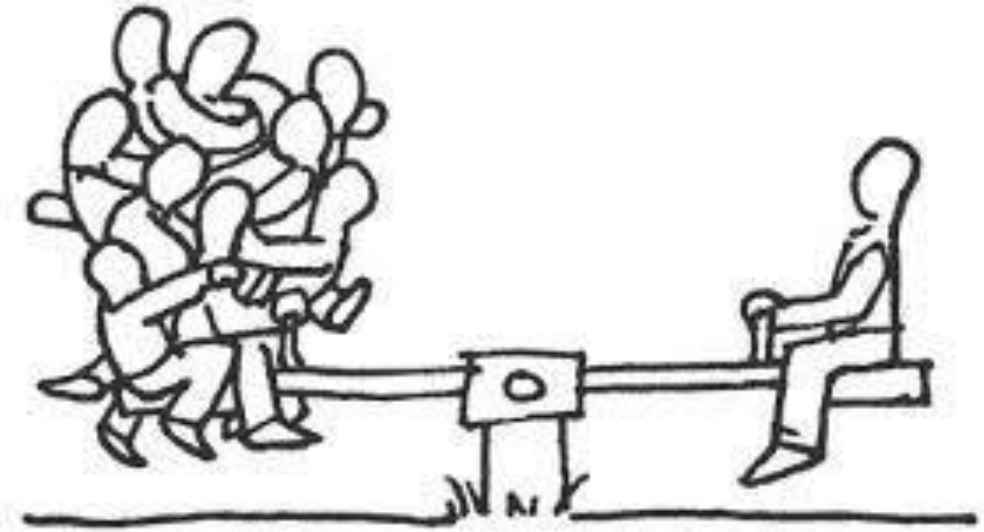
Proyecto Genoma Humano

- Secuenciación de 1^{era} generación

- ***Secuenciación de Nueva Generación***
  - Secuenciación de 2da y **3^{era} Generación** …

Secuenciadores de 3^{era} Generación

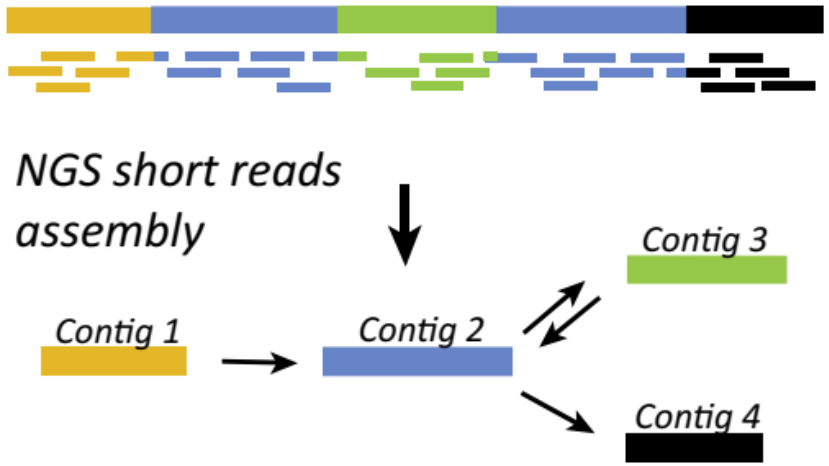
Amplificación por PCR vs molecula única



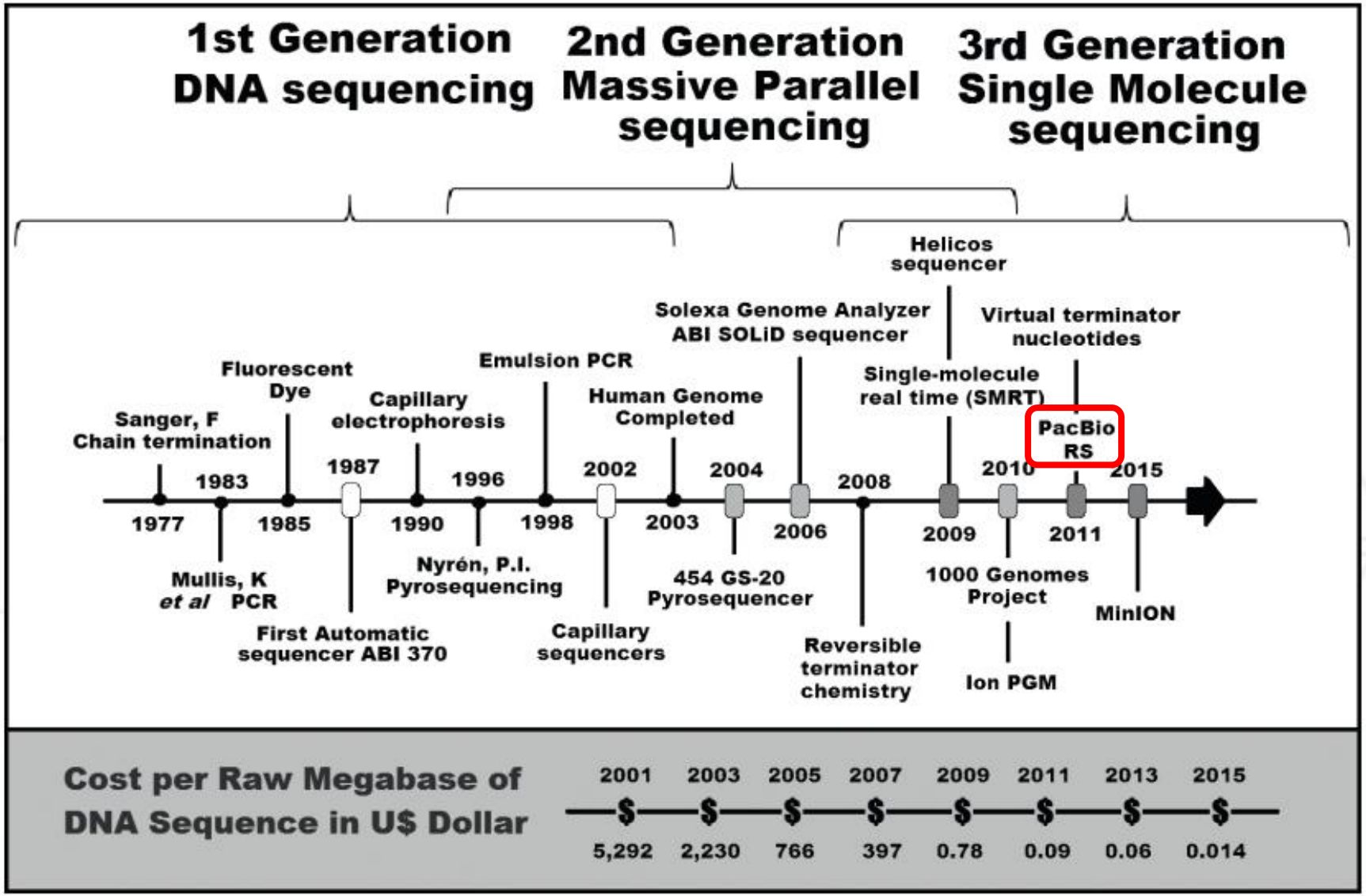
Oxford NANOPORE Technologies®

PACBIO®

Largos vs cortos



Plataformas de Secuenciación



http://dx.doi.org/10.5772/intechopen.69337

Pacific Biosciences (PacBio): SMRT

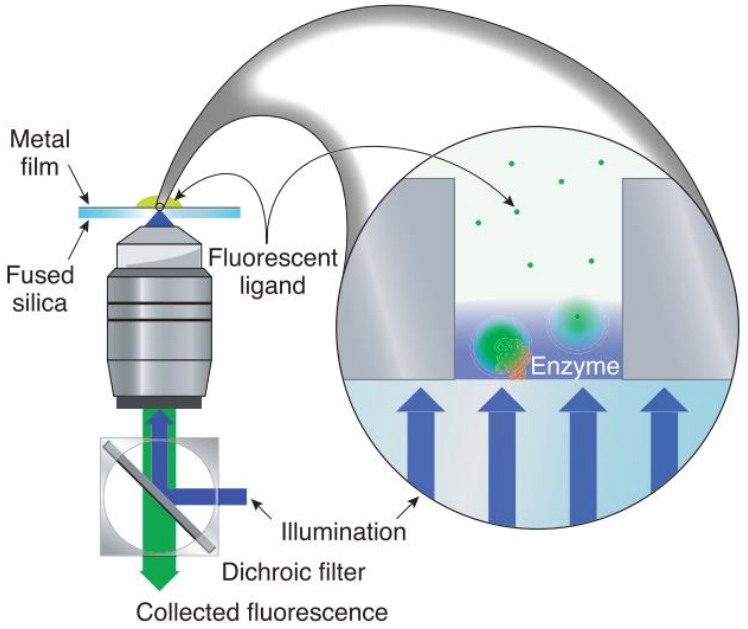
Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations

M. J. Levene,[1] J. Korlach,[1,2] S. W. Turner,[1]* M. Foquet,[1] H. G. Craighead,[1] W. W. Webb[1]†

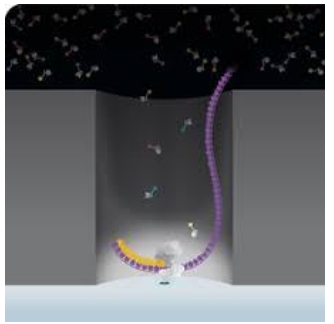
Optical approaches for observing the dynamics of single molecules have required pico- to nanomolar concentrations of fluorophore in order to isolate individual molecules. However, many biologically relevant processes occur at micromolar ligand concentrations, necessitating a reduction in the conventional observation volume by three orders of magnitude. We show that arrays of zero-mode waveguides consisting of subwavelength holes in a metal film provide a simple and highly parallel means for studying single-molecule dynamics at micromolar concentrations with microsecond temporal resolution. We present observations of DNA polymerase activity as an example of the effectiveness of zero-mode waveguides for performing single-molecule experiments at high concentrations.

31 JANUARY 2003 VOL 299 SCIENCE

Fig. 1. An apparatus for single-molecule enzymology using zero-mode waveguides.

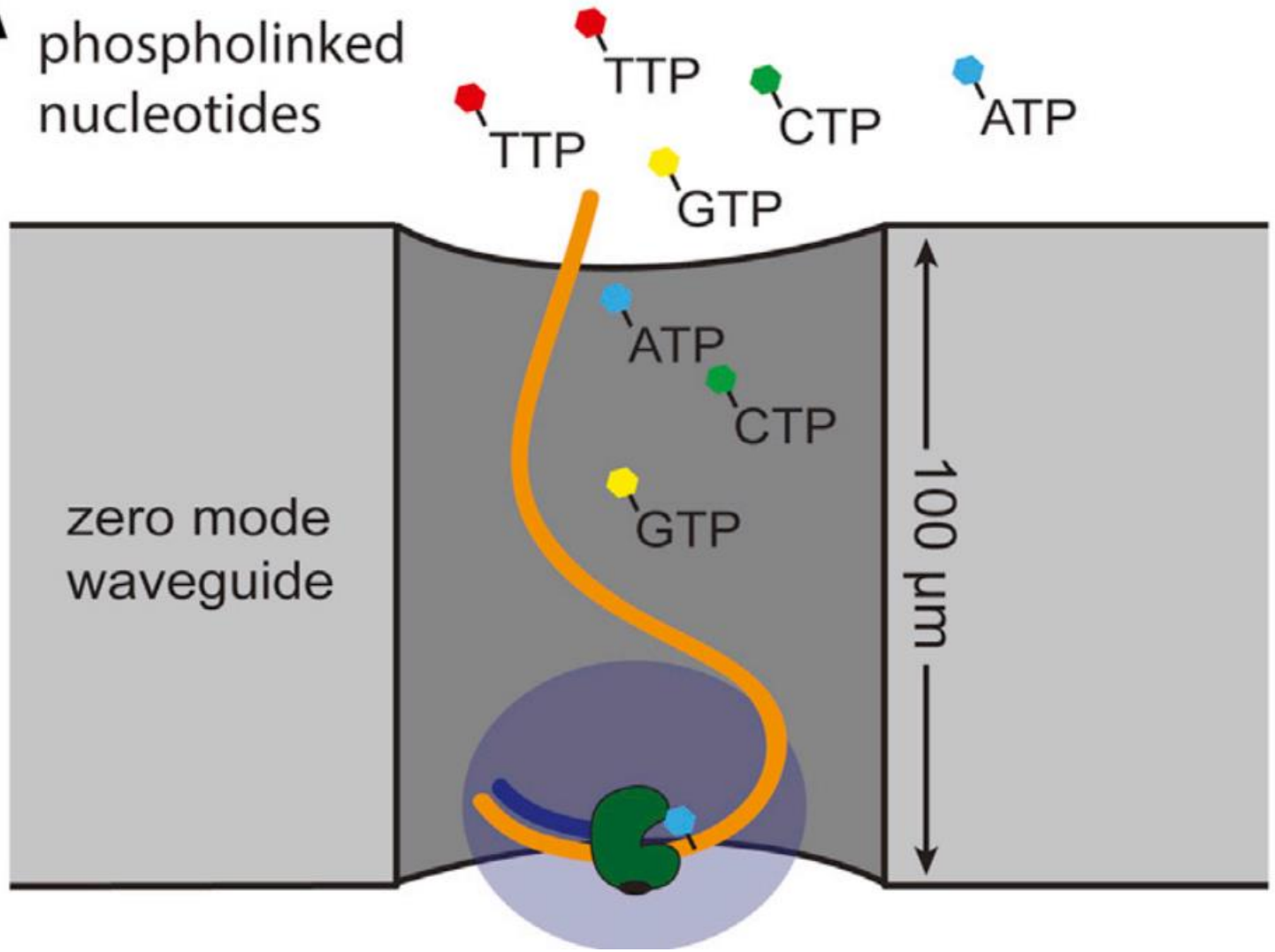


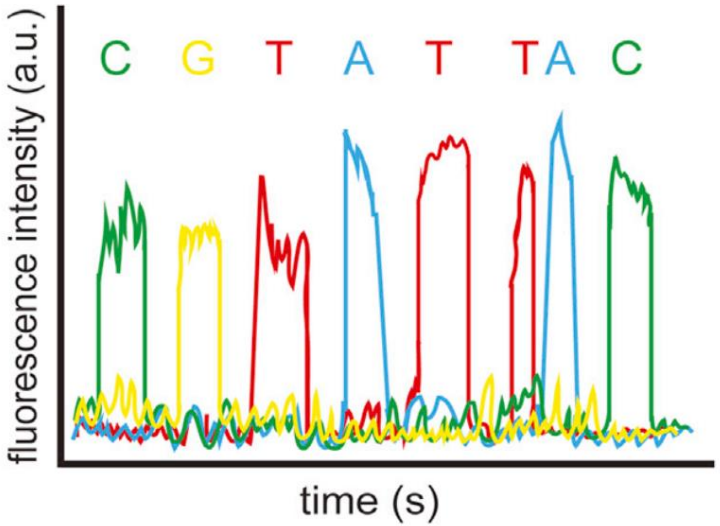
Pacific Biosciences (PacBio) : SMRT



- **La secuenciación de una sola molécula en tiempo real (SMRT) fue iniciada por Nanofluidics, Inc. y comercializada por Pacific Biosciences en 2010.**
  - La preparación del molde implica la **unión de adaptadores de horquilla a los extremos de las moléculas** de ADN o ADNc, generando un molde capeado (SMRT-bell).
  - Se **evita la amplificación clonal**, permitiendo la secuencia directa del ADN.
  - La **síntesis de ADN ocurre en cámaras del tamaño de un zeptolitro** (ZMW, zeptolitro = 10^{-21}), en las que **una sola polimerasa** se inmoviliza en el fondo de la cámara.
  - La polimerización ocurre continuamente con **nucleótidos fluorescentes** y la secuencia de ADN se puede leer en tiempo real a partir del **registro de las señales fluorescentes**.
  - Las **lecturas (reads) promedio son> 14 kb**, pero las lecturas individuales pueden ser de **hasta 60-100 kb**

Pacific Biosciences: SMRT



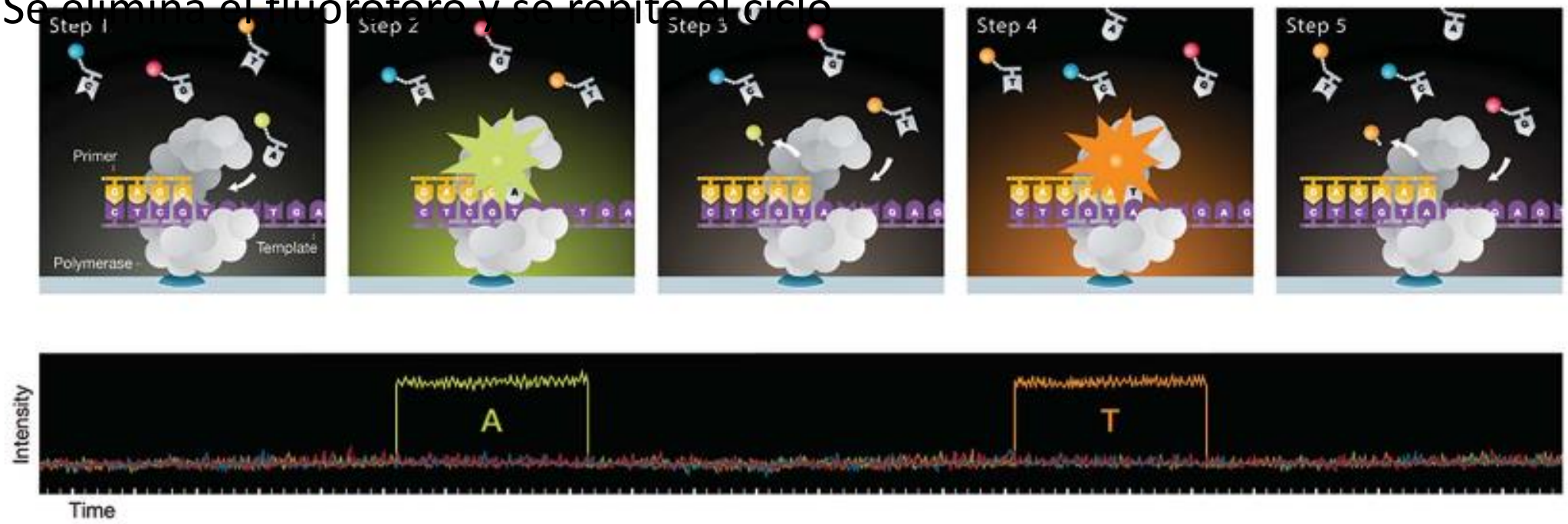


Lecturas de 10-15kb hasta 80kb

Error *base calling* aprox. 10%

Pacific Biosciences: SMRT

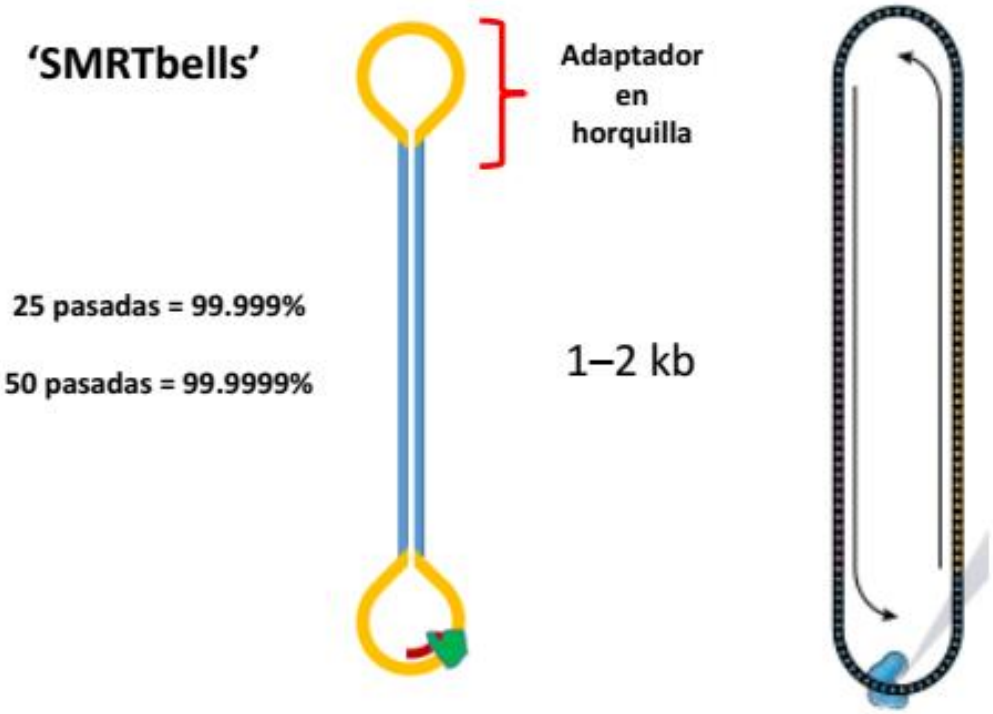
- Los nucleótidos incorporados se detectan mediante la detección de imágenes en tiempo real durante la síntesis de la cadena.
  - El ADN molde formándose en el **sitio activo de la polimerasa**, se puede registrar un pulso de intensidad de fluorescencia para el fluoróforo correspondiente.
  - Tal emisión de luz dura algunos milisegundos y será detectada por el sensor ZMW.
  - Se elimina el fluoróforo y se repite el ciclo



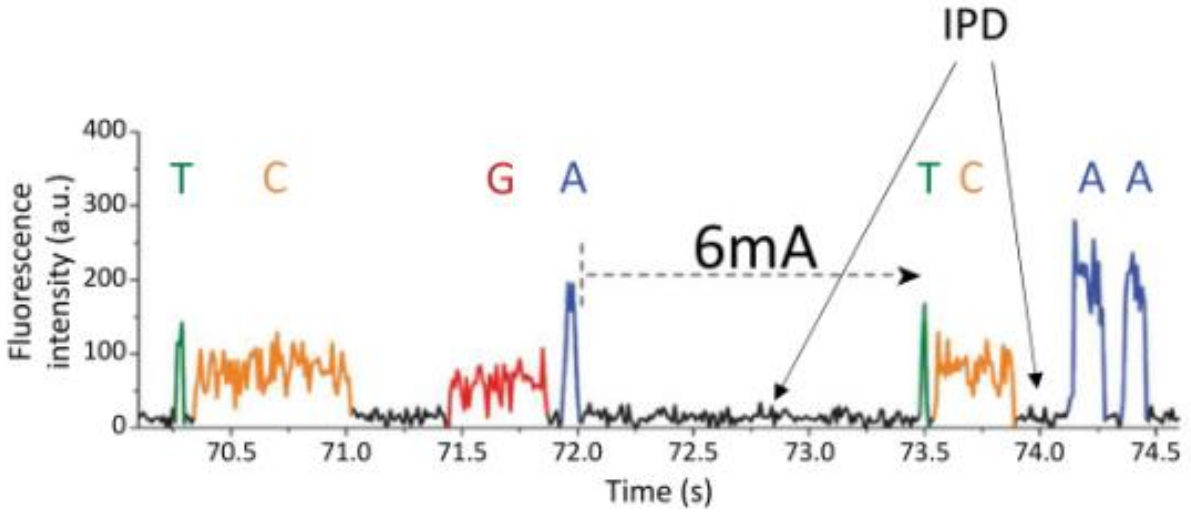
10.1146/annurev-anchem-062012-092628

Pacific Biosciences: SMRT

- Los nucleótidos incorporados se detectan mediante la detección de imágenes en tiempo real durante la síntesis de la cadena.
  - El ADN molde formándose en el **sitio activo de la polimerasa**, se puede registrar un pulso de intensidad de fluorescencia para el fluoróforo correspondiente.
  - Tal emisión de luz dura algunos milisegundos y será detectada por el sensor ZMW.

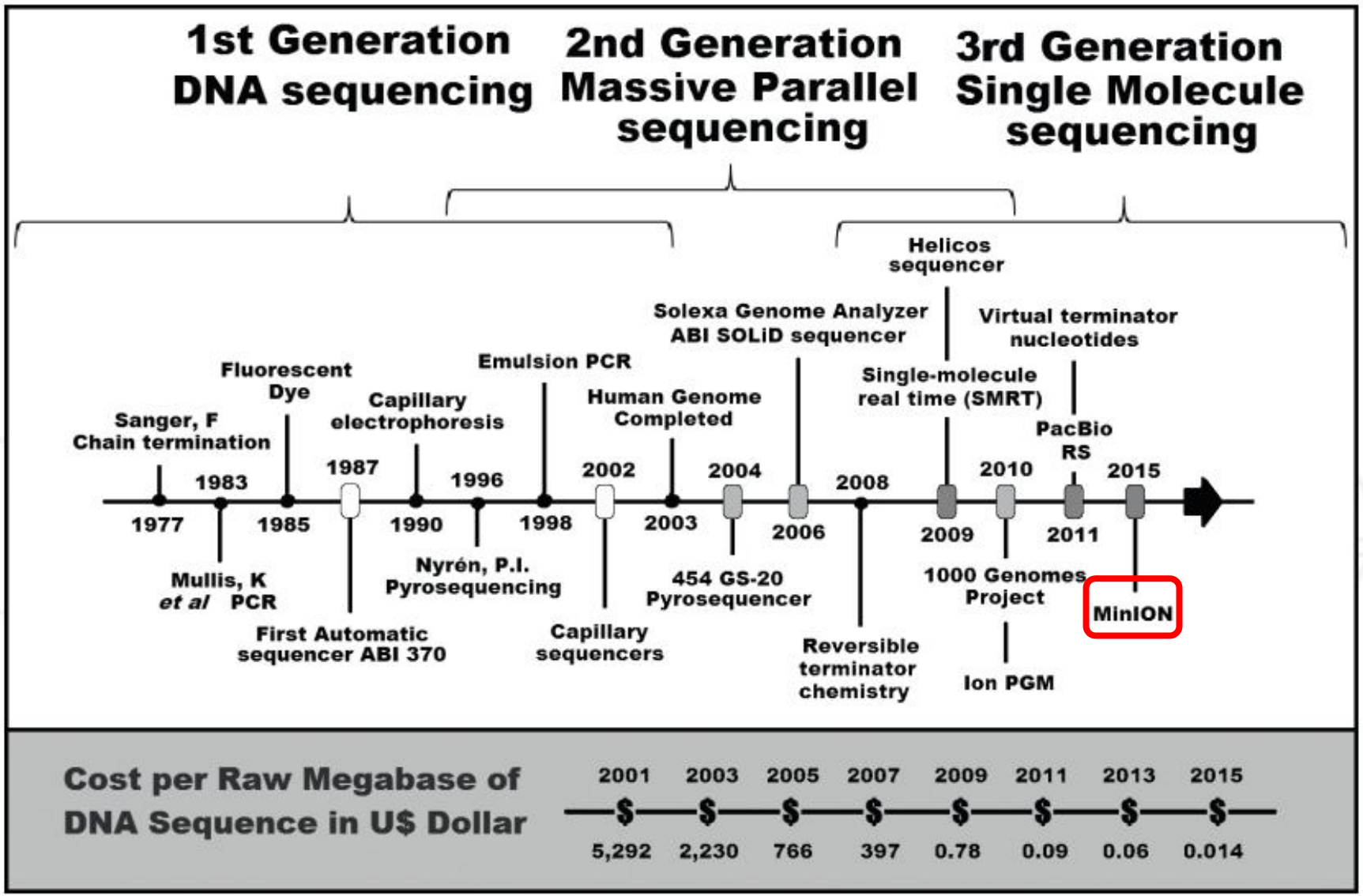


https://doi.org/10.1016/j.tig.2018.05.008



10.1146/annurev-anchem-062012-092628

Plataformas de Secuenciación

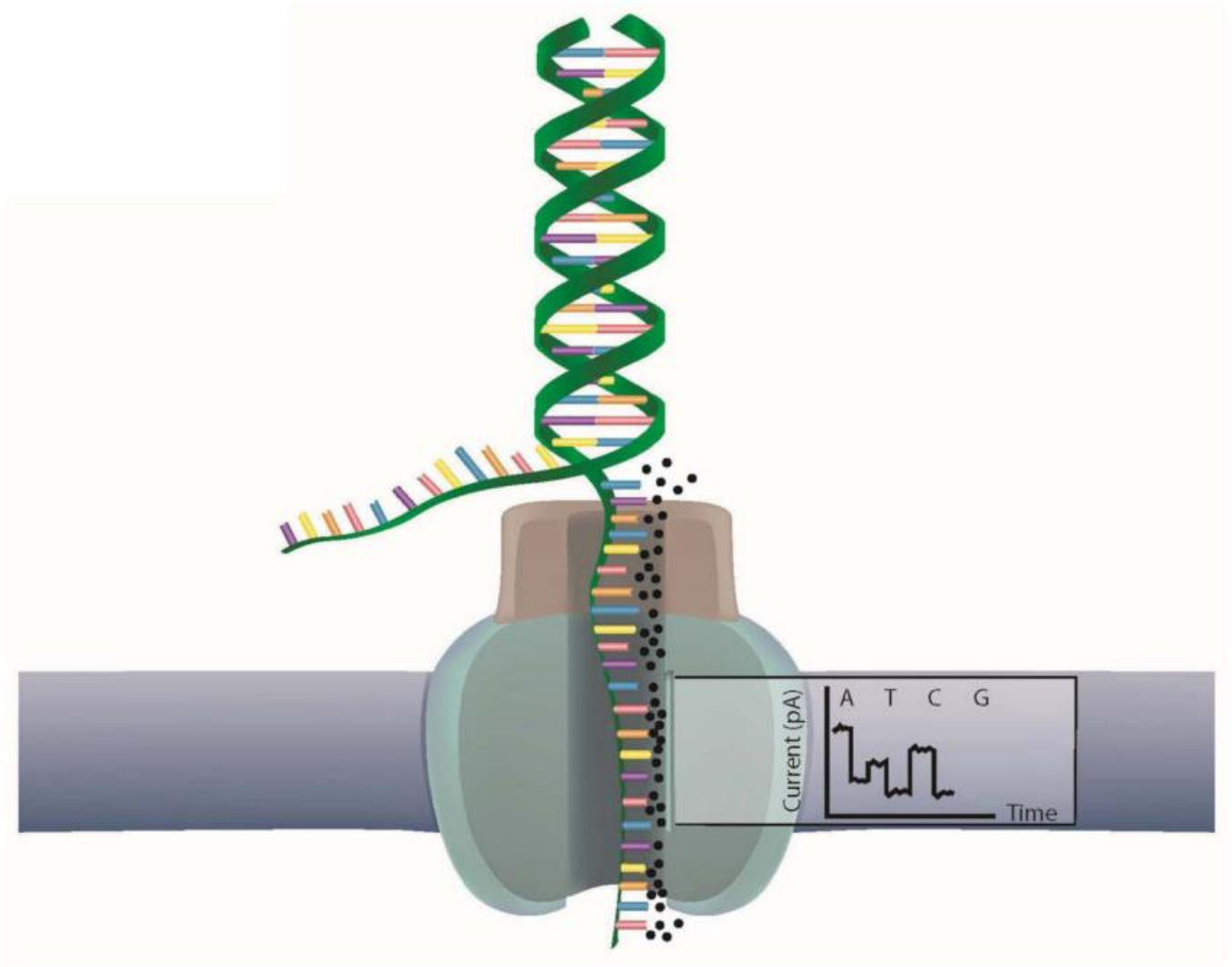


http://dx.doi.org/10.5772/intechopen.69337

Oxford Nanopore Technologies

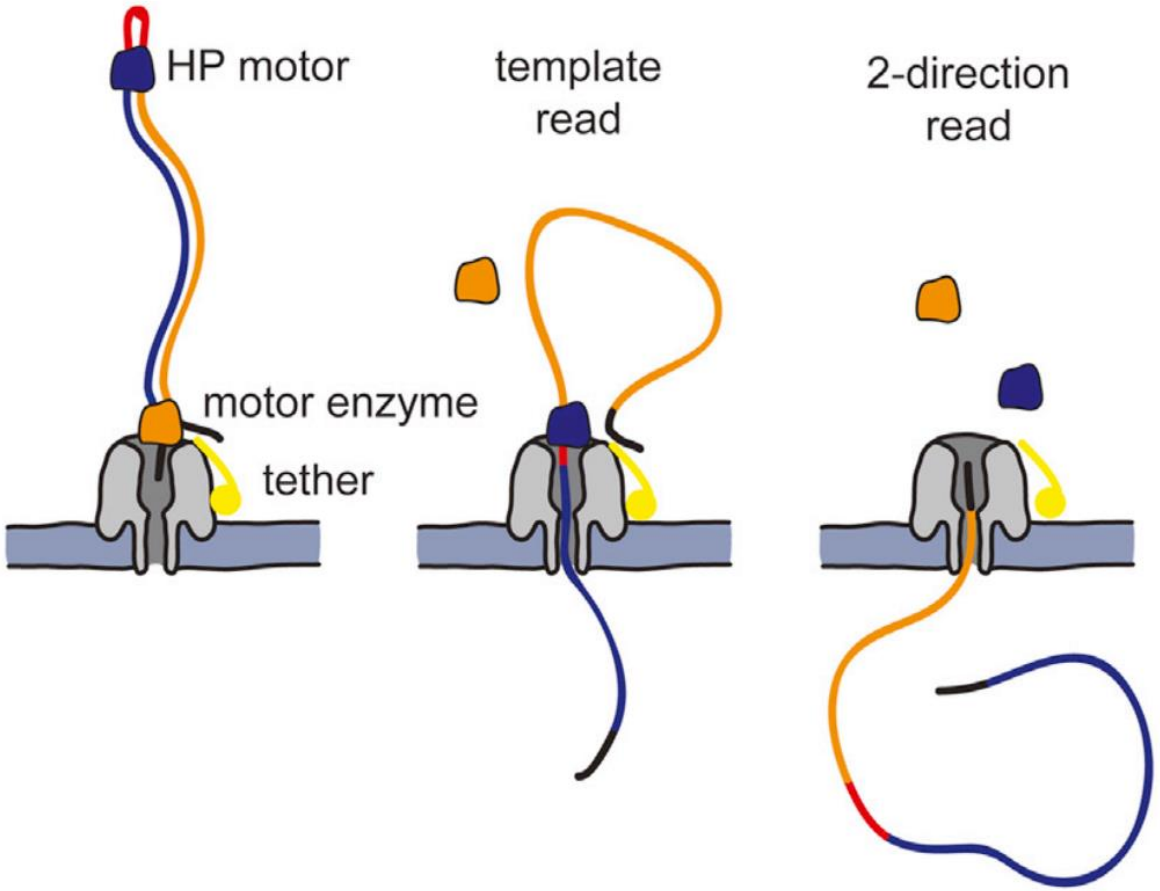
- La tecnología de **Oxford Nanopore (2014)**, es una **célula de flujo de secuenciación con cientos de microporos independientes**, cada uno con una **bicapa sintética perforada por nanoporos biológicos**.
- La secuenciación se logra **midiendo cambios característicos en la corriente inducidos por como las bases pasan a través del poro** mediante una proteína motora molecular.
  - La **preparación de la biblioteca es mínima**, la ligación de adaptadores al igual que la secuenciación SMRT, esta metodología de preparación de la biblioteca se puede hacer con o sin amplificación por PCR.
  - El **primer adaptador está unido a una enzima motora**.
  - El primer dispositivo comercialmente disponible para la secuenciación de nanoporos es el **MinION, un secuenciador portátil alimentado por USB.**
  - **Corre en 48 horas** y puede producir **> 90 Mb** de datos con lecturas medias y máximas de **6 kb y >60 kb**, respectivamente.

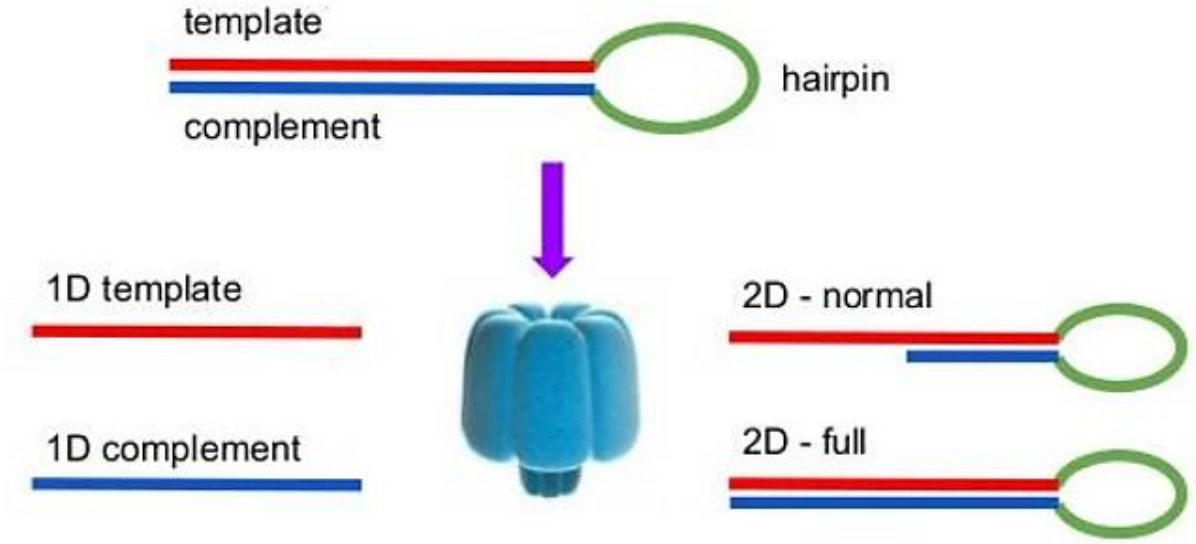
Oxford Nanopore Technologies

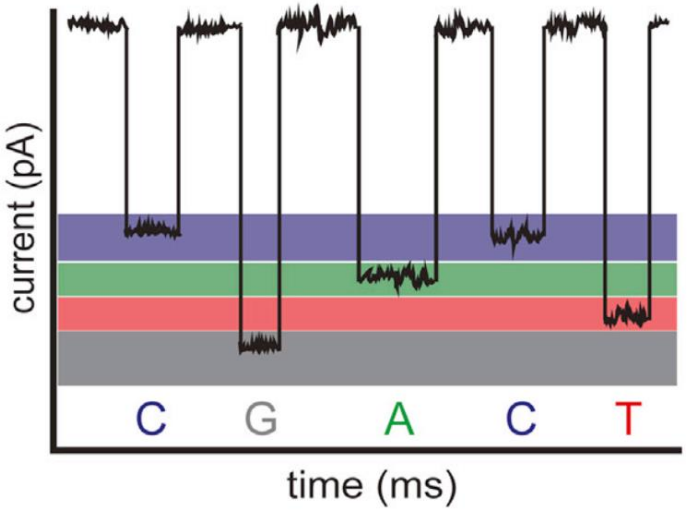


10.1161/CIRCRESAHA.113.300939

Oxford Nanopore Technologies







Lecturas de hasta 1Mb

Error *base calling* aprox. 5-15% 1D y aprox. 3% 2D

The potential impact of nanopore sequencing on human genetics

Matthew W. Loose*

School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK

*To whom correspondence should be addressed. Tel: +44 1158230358; Fax: +44 1158230313; Email: matt.loose@nottingham.ac.uk

	MinION	GridION	PromethION
			
Flow cell number	1	5	48
Integrated compute	No	Yes	Yes
Channels per flow cell	512	512	3000
Nanopores per flow cell[a]	2048	2048	6000
Yield per flow cell (theoretical/reported)	20 Gbp/2.3 Gbp	20 Gbp/2.3 Gbp	50 Gbp/6.7 Gbp[b]
Approx time to 30× human genome[c]	9 days/78 days	2 days/16 days	<2 h/<7 h

Ebola virus 2015

Scientists in the UK have just produced a DNA read that is over 2.3 million bases in length! 😲

bbc.com

World's longest DNA sequence decoded

A team of UK scientists have claimed the record for decoding the world's longest DNA sequence.

SCIENTI

OPEN

Nanopo
and Gen
Internat

Received: 1 August 2017
Accepted: 11 December 2017
Published online: 21 December 2017

Sarah L. Castro-Walla
H. Rubins[6], Alexa B. R
Douglas J. Botkin[11], Timothy A. Stephenson[12], Sissel Juul[13], Daniel J. Turner[13], Fernando Izquierdo[13], Scot Federman[2,3], Doug Stryke[2,3], Sneha Somasekar[2,3], Noah Alexander[7], Guixia Yu[2,3], Christopher E. Mason[7,14,15] & Aaron S. Burton[16]

sequencing for

Oxford NANOPORE technologies

COVID-19 Updates

PRODUCTS SERVICES APPLICATIONS GET STARTED RESOURCES

About VolTRAX

Oxford Nanopore offers a range of options for converting your original biological sample to a form ready for application into a nanopore sensing device.

Oxford Nanopore has developed VolTRAX – a small device designed to perform library preparation automatically, so that a user can get a biological sample ready for analysis, hands-free. VolTRAX is designed as an alternative to a range of lab equipment, to allow consistent and varied, automated library prep options.

Buy VolTRAX >

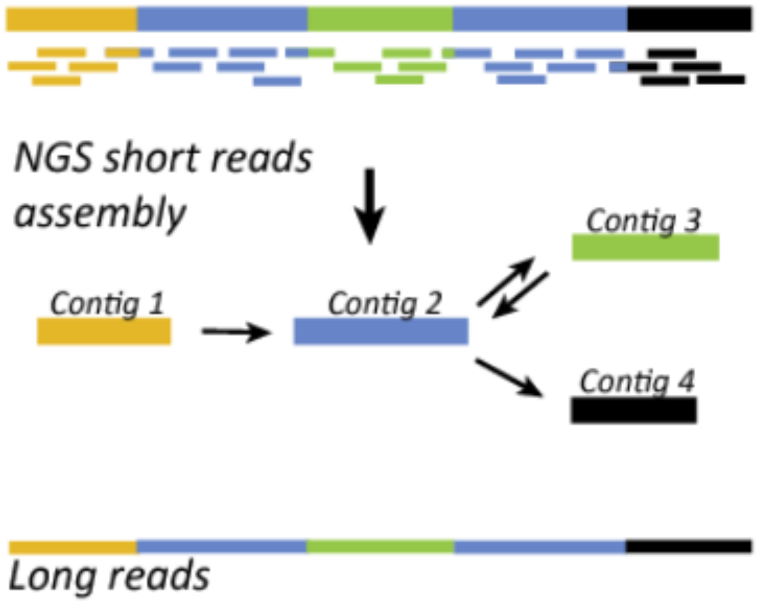
VolTRAX:

automated sample an

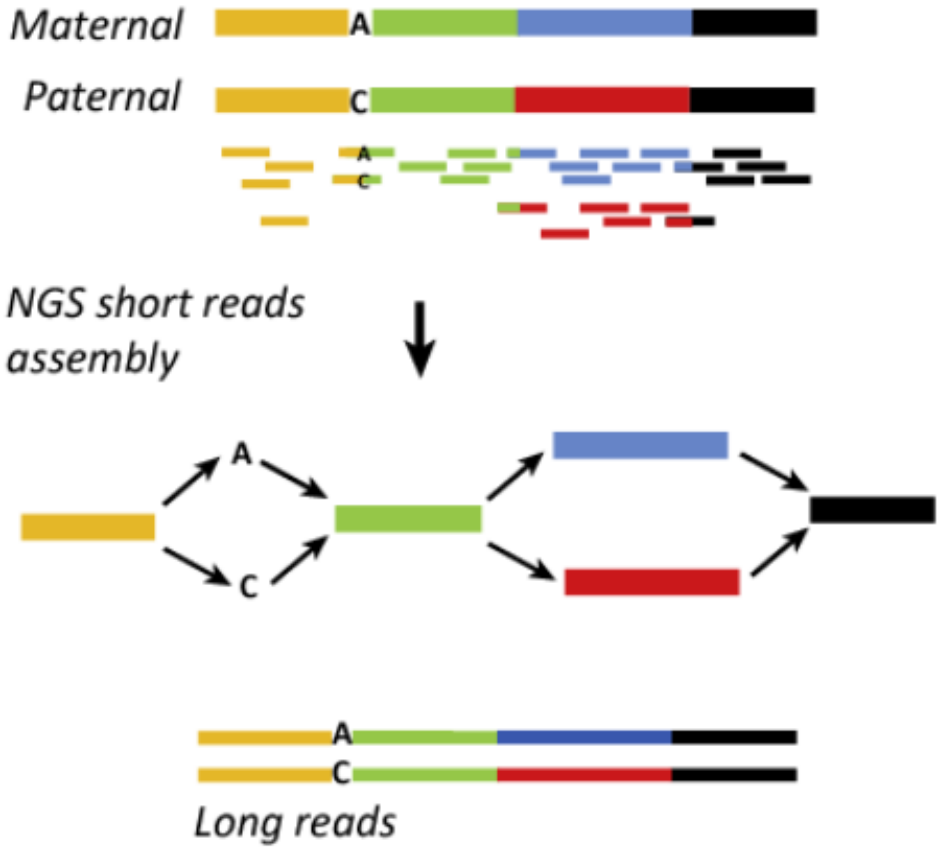
without human i

Application	Advantages/disadvantages	Ideal platform at human scale	References
Whole Genome Sequencing and assembly	Highly contiguous genome assemblies Ultra-long reads should further increase contiguity and could close existing gaps in assemblies Challenges remain for delivery of high molecular weight DNA to the sequencing nanopores Direct detection of DNA modifications possible DNA amounts required are high (single flow cell requires 1-1.5 ug DNA) Assembly tools computationally expensive	GridION PromethION	(35,40,41,52,53)
Structural variants	Long reads will, by definition, improve the detection of structural variants in genomes Ultra-long read preparations will aid structural variant detection but require high molecular weight DNA	GridION PromethION	(34,37,47)
Sequence variation	Variant calling with Nanopore data is possible, but enhanced by the use of post base calling polishing (nanopolish), or short read correction Homopolymers were not called by early Nanopore base callers. Homopolymer base callers are now available (Albacore—Oxford Nanopore Technologies)	GridION PromethION	(20,27)
Phasing	Phasing across ever larger regions of the human genome will be facilitated by longer reads	GridION PromethION	(34,35,51,54)
Targeted Sequencing	Targeted methods for DNA capture prior to sequencing utilising Cas9 suggest that individual regions of the genome can be captured for sequencing.In principle, selective sequencing ('read until') could enable dynamic selection of target during a sequencing run.	All	(48,55)
Direct Sequencing of RNA	A unique property of nanopore sequencers is the ability to directly sequence RNA, including the detection of modifications. Currently, direct RNA requires 500 ng polyA-tailed mRNA as input.	MinION GridION PromethION	(49,50)

Secuencias repetidas

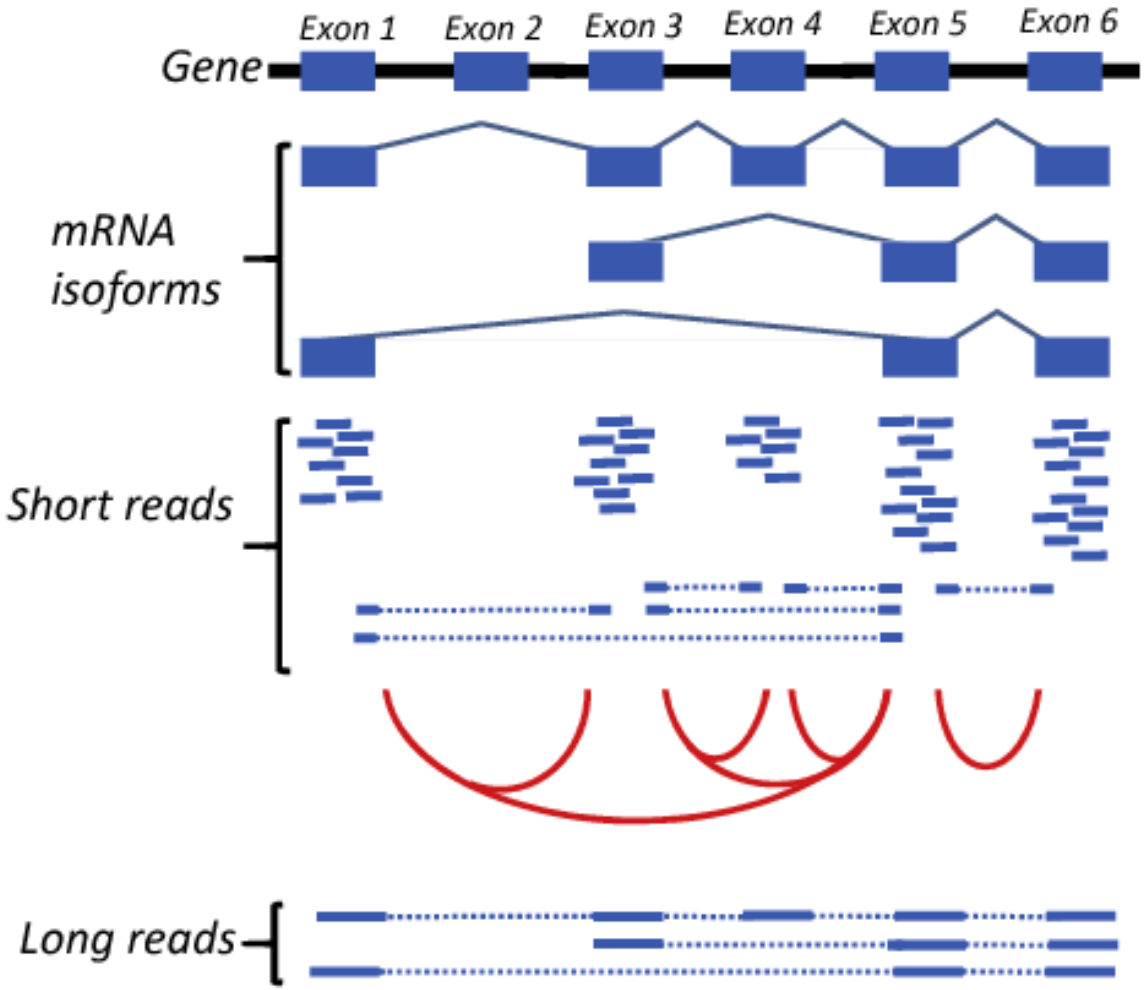


Fase y Haplotipos

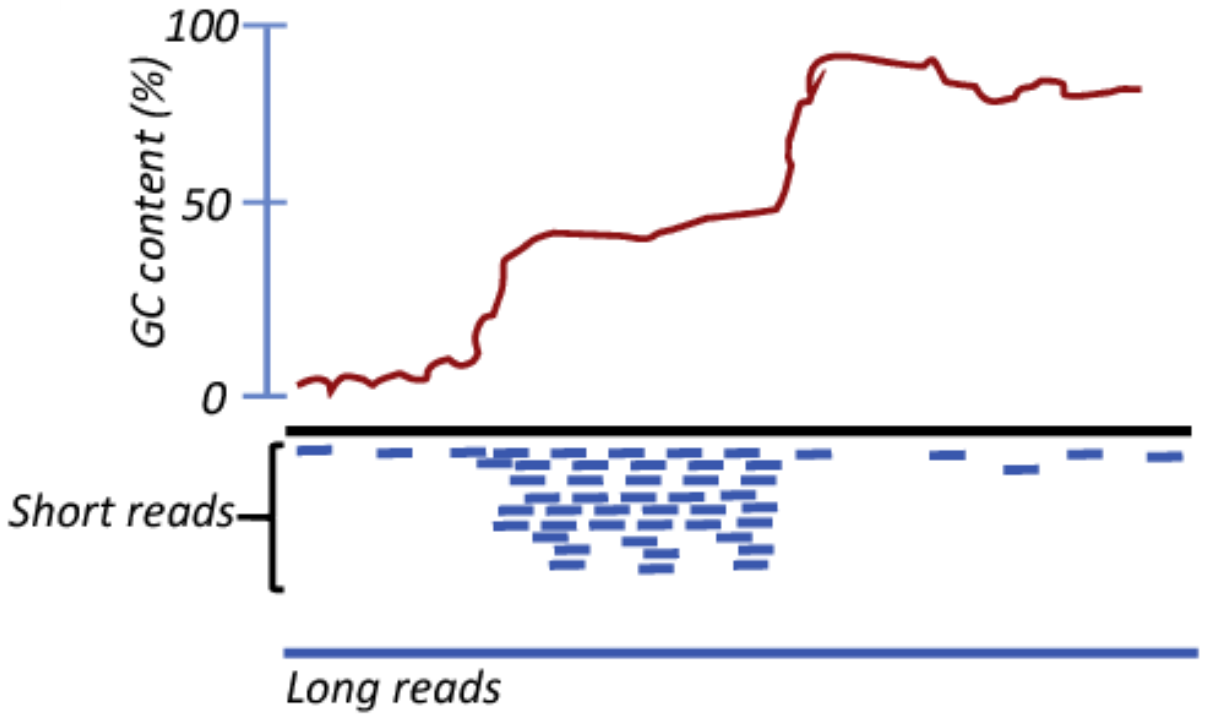


https://doi.org/10.1055/s-0039-1688446

Isoformas de Transcritos



Regiones de %GC extremos



https://doi.org/10.1055/s-0039-1688446

RESEARCH ARTICLE

HUMAN GENOMICS

2022

The complete sequence of a human genome

Sergey Nurk[1]†, Sergey Koren[1]†, Arang Rhie[1]†, Mikko Rautiainen[1]†, Andrey V. Bzikadze[2], Alla Mikheenko[3], Mitchell R. Vollger[4], Nicolas Altemose[5], Lev Uralsky[6,7], Ariel Gershman[8], Sergey Aganezov[9]†, Savannah J. Hoyt[10], Mark Diekhans[11], Glennis A. Logsdon[4], Michael Alonge[9], Stylianos E. Antonarakis[12], Matthew Borchers[13], Gerard G. Bouffard[14], Shelise Y. Brooks[14], Gina V. Caldas[15], Nae-Chyun Chen[9], Haoyu Cheng[16,17], Chen-Shan Chin[18], William Chow[19], Leonardo G. de Lima[13], Philip C. Dishuck[4], Richard Durbin[19,20], Tatiana Dvorkina[3], Ian T. Fiddes[21], Giulio Formenti[22,23], Robert S. Fulton[24], Arkarachai Fungtammasan[18], Erik Garrison[11,25], Patrick G. S. Grady[10], Tina A. Graves-Lindsay[26], Ira M. Hall[27], Nancy F. Hansen[28], Gabrielle A. Hartley[10], Marina Haukness[11], Kerstin Howe[19], Michael W. Hunkapiller[29], Chirag Jain[1,30], Miten Jain[11], Erich D. Jarvis[22,23], Peter Kerpedjiev[31], Melanie Kirsche[9], Mikhail Kolmogorov[32], Jonas Korlach[29], Milinn Kremitzki[26], Heng Li[16,17], Valerie V. Maduro[33], Tobias Marschall[34], Ann M. McCartney[1], Jennifer McDaniel[35], Danny E. Miller[4,36], James C. Mullikin[14,28], Eugene W. Myers[37], Nathan D. Olson[35], Benedict Paten[11], Paul Peluso[29], Pavel A. Pevzner[32], David Porubsky[4], Tamara Potapova[13], Evgeny I. Rogaev[6,7,38,39], Jeffrey A. Rosenfeld[40], Steven L. Salzberg[9,41], Valerie A. Schneider[42], Fritz J. Sedlazeck[43], Kishwar Shafin[11], Colin J. Shew[44], Alaina Shumate[41], Ying Sims[19], Arian F. A. Smit[45], Daniela C. Soto[44], Ivan Sović[29,46], Jessica M. Storer[45], Aaron Streets[5,47], Beth A. Sullivan[48], Françoise Thibaud-Nissen[42], James Torrance[19], Justin Wagner[35], Brian P. Walenz[1], Aaron Wenger[29], Jonathan M. D. Wood[19], Chunlin Xiao[42], Stephanie M. Yan[49], Alice C. Young[14], Samantha Zarate[9], Urvashi Surti[50], Rajiv C. McCoy[49], Megan Y. Dennis[44], Ivan A. Alexandrov[3,7,51], Jennifer L. Gerton[13,52], Rachel J. O'Neill[10], Winston Timp[8,41], Justin M. Zook[35], Michael C. Schatz[9,49], Evan E. Eichler[4,53]*, Karen H. Miga[11,54]*, Adam M. Phillippy[1]*

El consorcio Telómero a telómero (T2T) presenta una secuencia completa de 3055 millones de pares de bases de un genoma humano.
Introduce casi 200 millones de pares de bases (8% del genoma) que contienen 1956 genes, 99 de los cuales se prevé que codifiquen proteínas.

Regiones resueltas:

- satélites centroméricos
- duplicaciones segmentarias
- brazos cortos de los cinco cromosomas acrocéntricos,

https://doi.org/10.1126/science.abj6987

STATISTICS	GRCH38	T2T-CHM13	DIFFERENCE (±%)
Summary			
Assembled bases (Gbp)	2.92	3.05	+4.5
Unplaced bases (Mbp)	11.42	0	−100.0
Gap bases (Mbp)	120.31	0	−100.0
Number of contigs	949	24	−97.5
Contig NG50 (Mbp)	56.41	154.26	+173.5
Number of issues	230	46	−80.0
Issues (Mbp)	230.43	8.18	−96.5
Gene annotation			
Number of genes	60,090	63,494	+5.7
Protein coding	19,890	19,969	+0.4
Number of exclusive genes	263	3,604	
Protein coding	63	140	
Number of transcripts	228,597	233,615	+2.2
Protein coding	84,277	86,245	+2.3
Number of exclusive transcripts	1,708	6,693	
Protein coding	829	2,780	
Segmental duplications			
Percentage of segmental duplications (%)	5.00	6.61	
Segmental duplication bases (Mbp)	151.71	201.93	+33.1
Number of segmental duplications	24097	41528	+72.3
RepeatMasker			
Percentage of repeats (%)	51.89	53.94	
Repeat bases (Mbp)	1,516.37	1,647.81	+8.7
Long interspersed nuclear elements	626.33	631.64	+0.8
Short interspersed nuclear elements	386.48	390.27	+1.0
Long terminal repeats	267.52	269.91	+0.9
Satellite	76.51	150.42	+96.6
DNA	108.53	109.35	+0.8
Simple repeat	36.5	77.69	+112.9
Low complexity	6.16	6.44	+4.6
Retroposon	4.51	4.65	+3.3
rRNA	0.21	1.71	+730.4

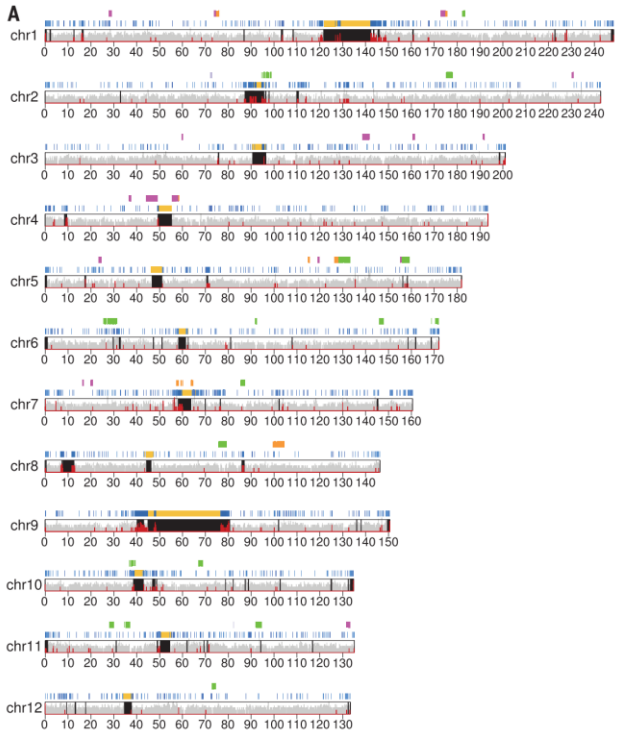
RESEARCH ARTICLE

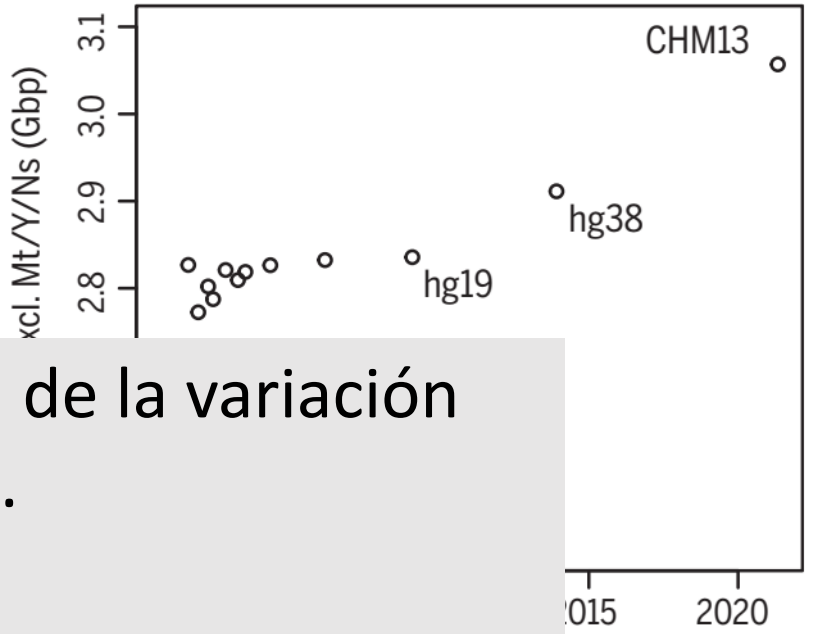
HUMAN GENOMICS

The complete sequence of a human genome

Sergey Nurk[1]†, Sergey Koren[1]†, Arang Rhie[1]†, Mikko Rautiainen[1]†, Andrey V. Bzikadze[2], Alla Mikheenko[3], Mitchell R. Vollger[4], Nicolas Altemose[5], Lev Uralsky[6,7], Ariel Gershman[8], Sergey Aganezov[9]‡, Savannah J. Hoyt[10], Mark Diekhans[11], Glennis A. Logsdon[4], Michael Alonge[9], Stylianos E. Antonarakis[12], Matthew Borchers[13], Gerard G. Bouffard[14], Shelise Y. Brooks[14], Gina V. Caldas[15], Nae-Chyun Chen[9], Haoyu Cheng[16,17], Chen-Shan Chin[18], William Chow[19], Leonardo G. de Lima[13], Philip C. Dishuck[4], Richard Durbin[19,20], Tatiana Dvorkina[3], Ian T. Fiddes[21], Giulio Formenti[22,23], Robert S. Fulton[24], Arkarachai Fungtammasan[18], Erik Garrison[11,25], Patrick G. S. Grady[10], Tina A. Graves-Lindsay[26], Ira M. Hall[27], Nancy F. Hansen[28], Gabrielle A. Hartley[10], Marina Haukness[11], Kerstin Howe[19], Michael W. Hunkapiller[29], Chirag Jain[1,30], Miten Jain[11], Erich D. Jarvis[22,23], Peter Kerpedjiev[31], Melanie Kirsche[9], Mikhail Kolmogorov[32], Jonas Korlach[29], Milinn Kremitzki[26], Heng Li[16,17], Valerie V. Maduro[33], Tobias Marschall[34], Ann M. McCartney[1], Jennifer McDaniel[35], Danny E. Miller[4,36], James C. Mullikin[14,28], Eugene W. Myers[37], Nathan D. Olson[35], Benedict Paten[11], Paul Peluso[29], Pavel A. Pevzner[32], David Porubsky[4], Tamara Potapova[13], Evgeny I. Rogaev[6,7,38,39], Jeffrey A. Rosenfeld[40], Steven L. Salzberg[9,41], Valerie A. Schneider[42], Fritz J. Sedlazeck[43], Kishwar Shafin[11], Colin J. Shew[44], Alaina Shumate[41], Ying Sims[19], Arian F. A. Smit[45], Daniela C. Soto[44], Ivan Sović[29,46], Jessica M. Storer[45], Aaron Streets[5,47], Beth A. Sullivan[48], Françoise Thibaud-Nissen[42], James Torrance[19], Justin Wagner[35], Brian P. Walenz[1], Aaron Wenger[29], Jonathan M. D. Wood[19], Chunlin Xiao[42], Stephanie M. Yan[49], Alice C. Young[14], Samantha Zarate[9], Urvashi Surti[50], Rajiv C. McCoy[49], Megan Y. Dennis[44], Ivan A. Alexandrov[3,7,51], Jennifer L. Gerton[13,52], Rachel J. O'Neill[10], Winston Timp[8,41], Justin M. Zook[35], Michael C. Schatz[9,49], Evan E. Eichler[4,53]*, Karen H. Miga[11,54]*, Adam M. Phillippy[1]*

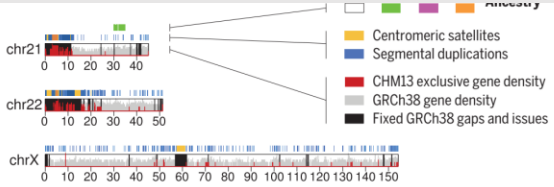
2022





No capta la diversidad completa de la variación genética humana.

The Human Pangenome Reference Consortium

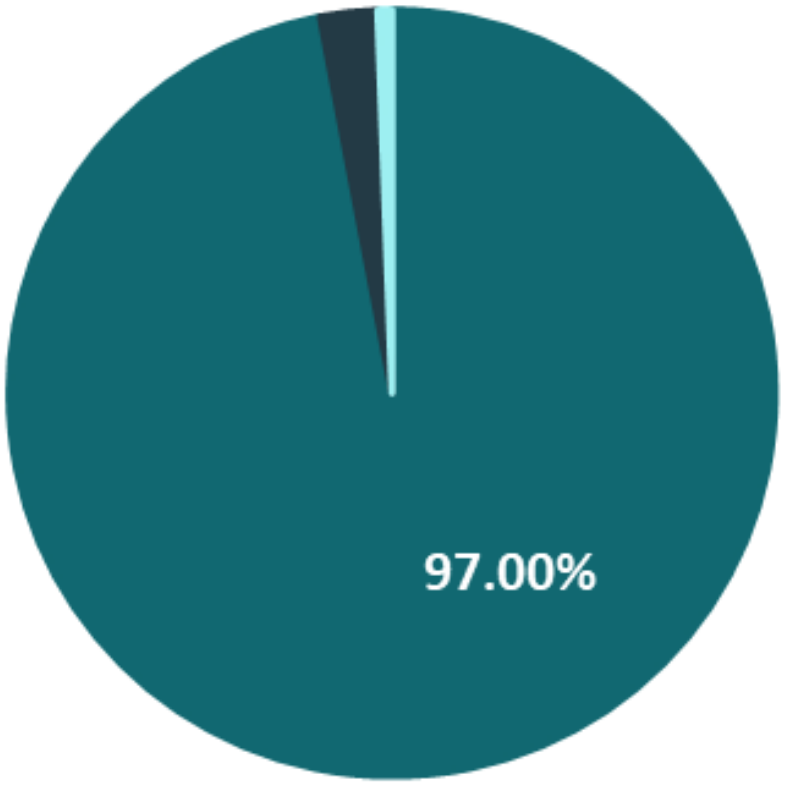


¿Por qué quiero secuenciar?

¿Qué quiero secuenciar?

Metodología y Plataforma

Global Sequencing Market, Segmentation By Type Of Sequencing, 2018, Percentage (%)



Next Generation Sequencing

Third Generation Sequencing

15.5% crecimiento anual 2014 - 2018

Sanger Sequencing

Source: The Business Research Company

https://www.thebusinessresearchcompany.com/press-release/sequencing-market-segments

Metodologías de Secuenciación

- **La estrategia específica empleada por cada plataforma determina la calidad, cantidad y sesgos del resultado datos de secuencia y la utilidad de la plataforma para particulares aplicaciones**

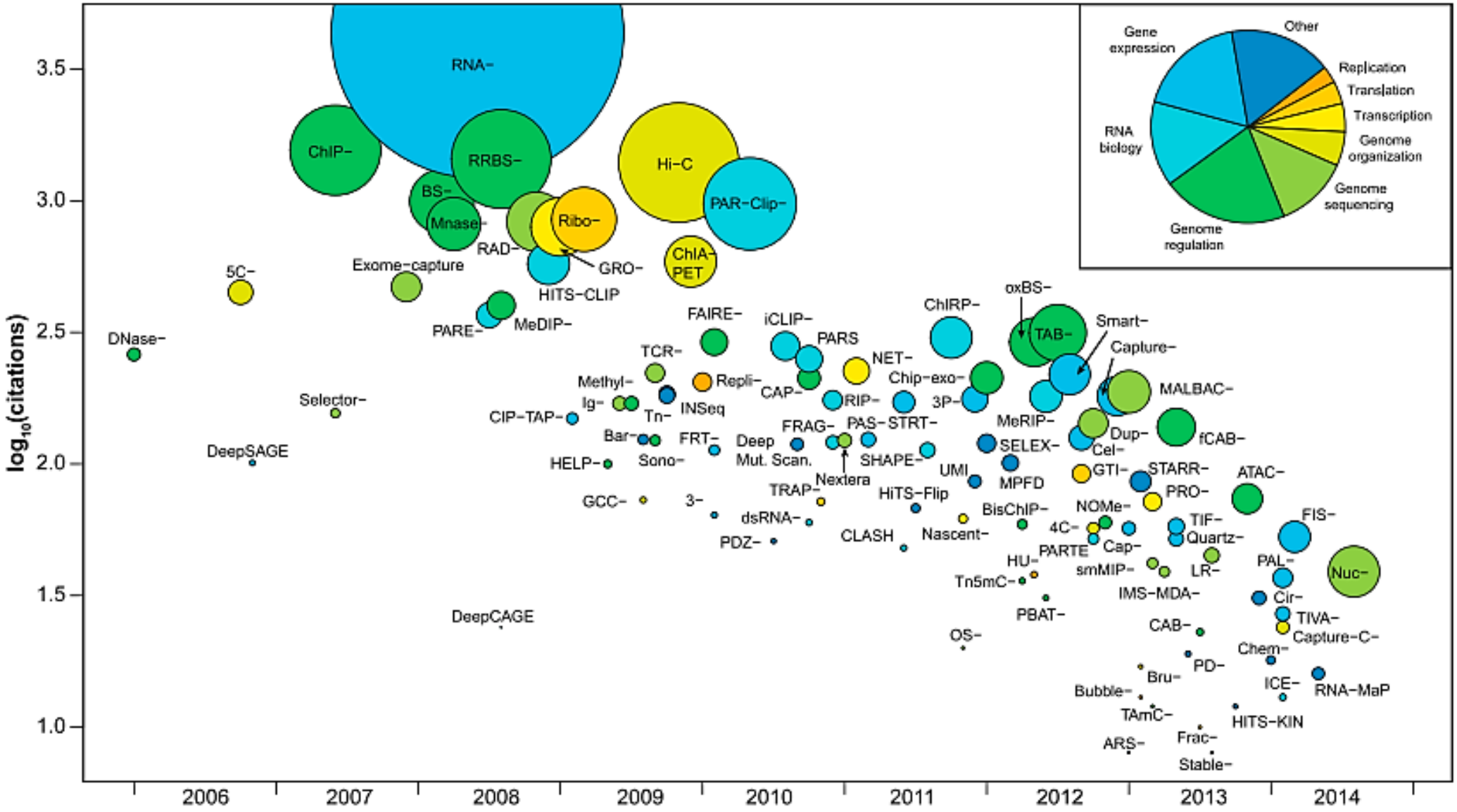


Figure 4. Overview of Selected HTS Applications

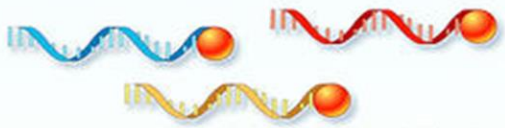
Publication date of a representative article describing a method versus the number of citations that the article received. Methods are colored by category, and the size of the data point is proportional to publication rate (citations/months). The inset indicates the color key as well the proportion of methods in each group. For clarity, seq has been omitted from the labels.

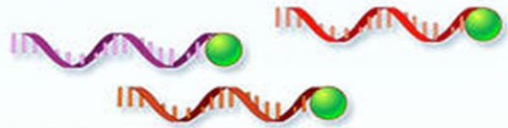
doi: 10.1016/j.molcel.2015.05.004

DNA MICROARRAY

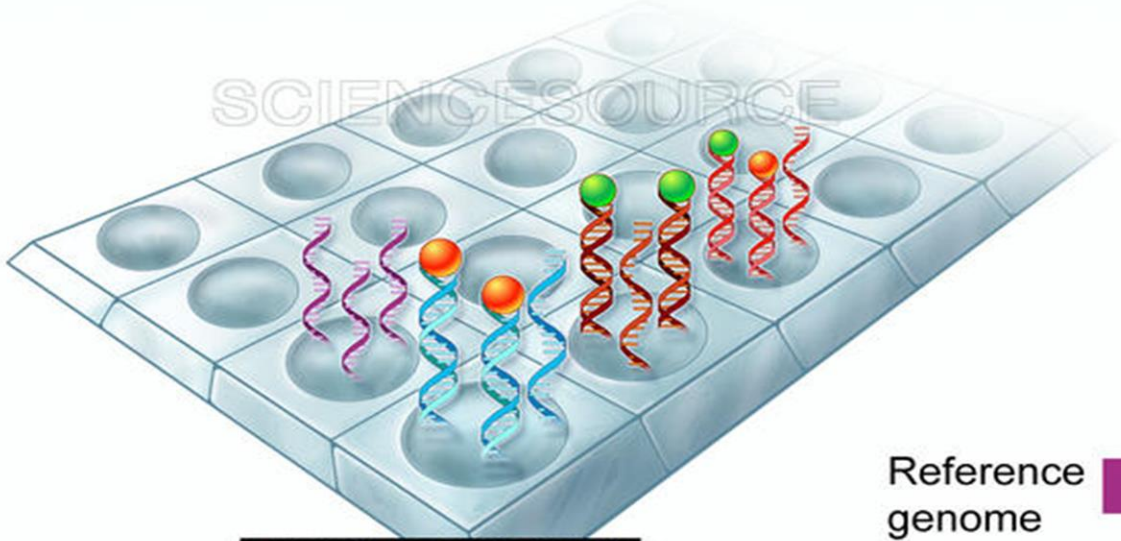
cDNA sample 1

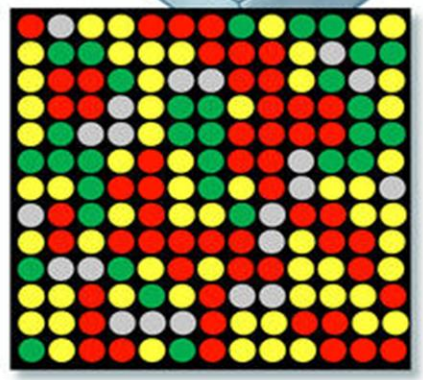
cDNA sample 2





Fluorescent tag



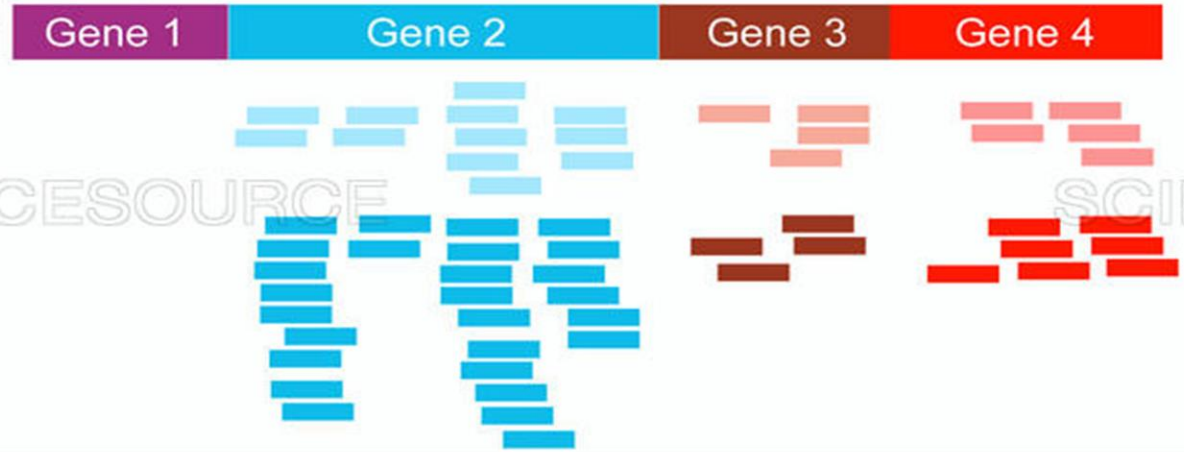


RNA-Seq

cDNA sample 1

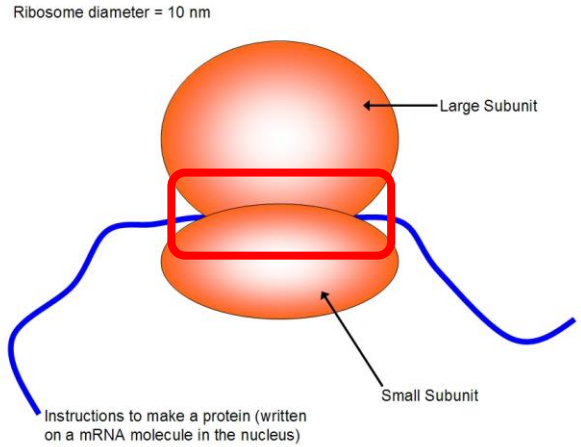
cDNA sample 2

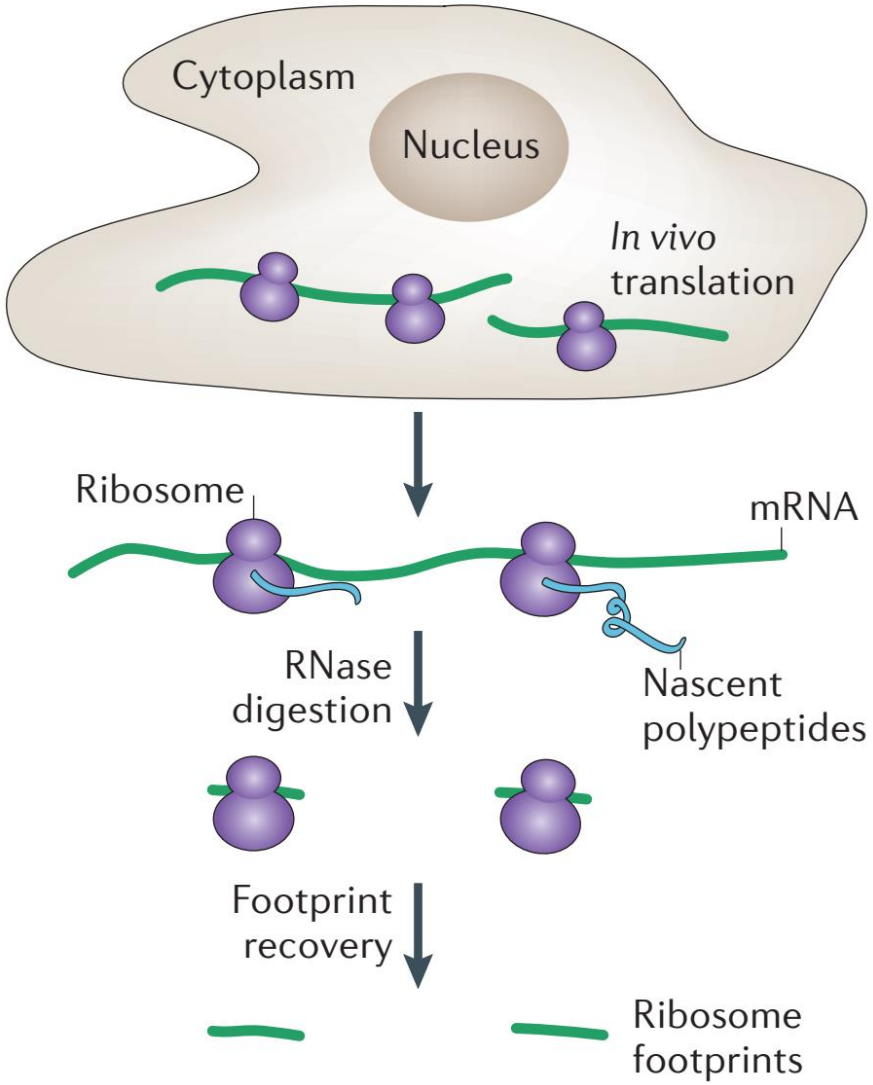
Reference genome

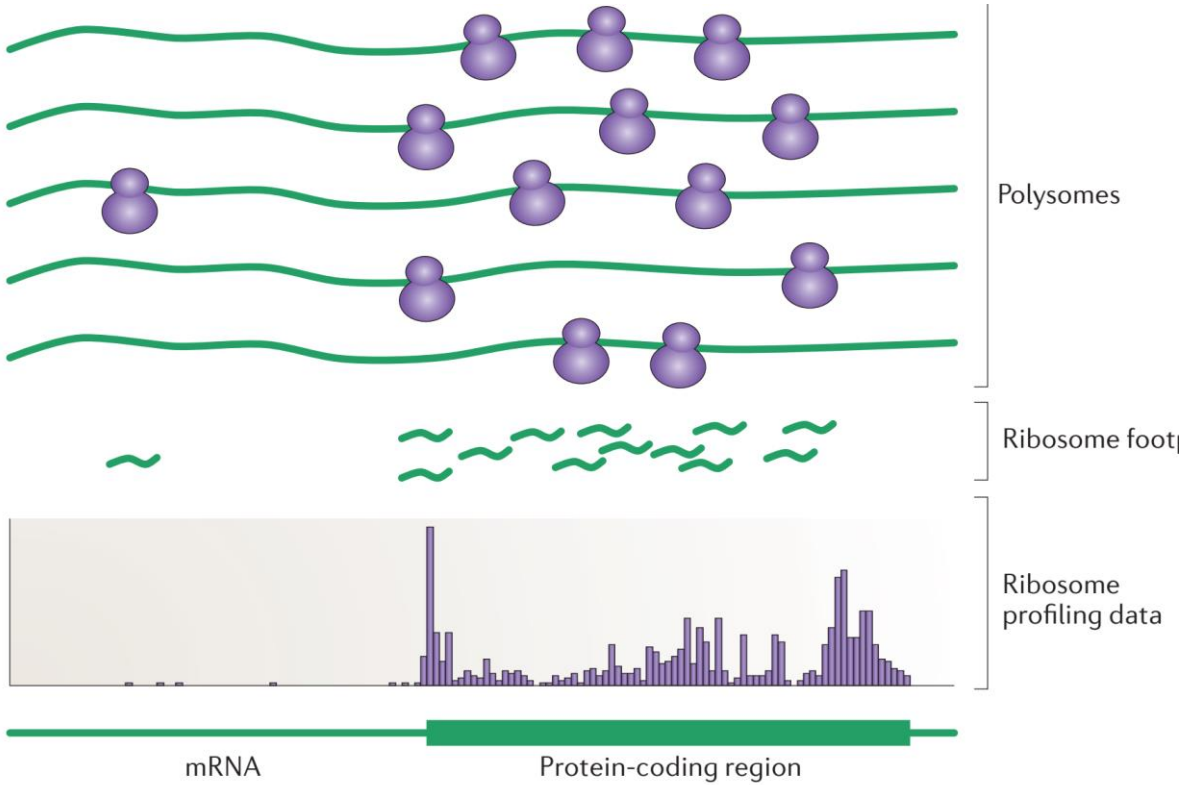


SCIENCESOURCE

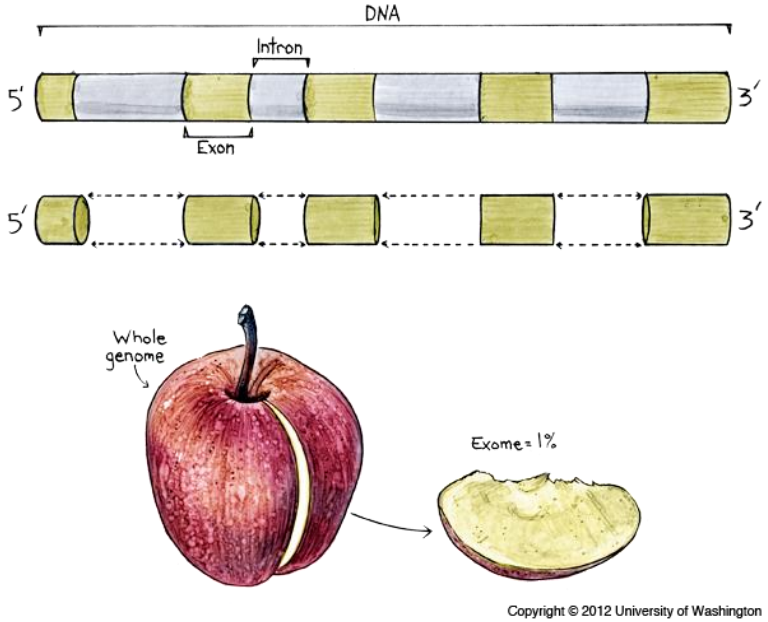
estrategiaS de secuenciaCIÓN: RIBO-SEQ

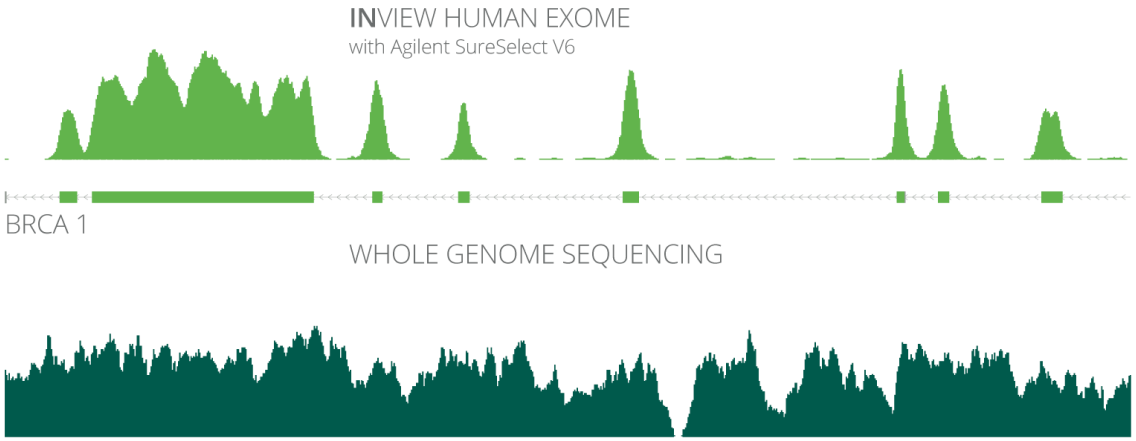






Aplicaciones: estudio de los genes.





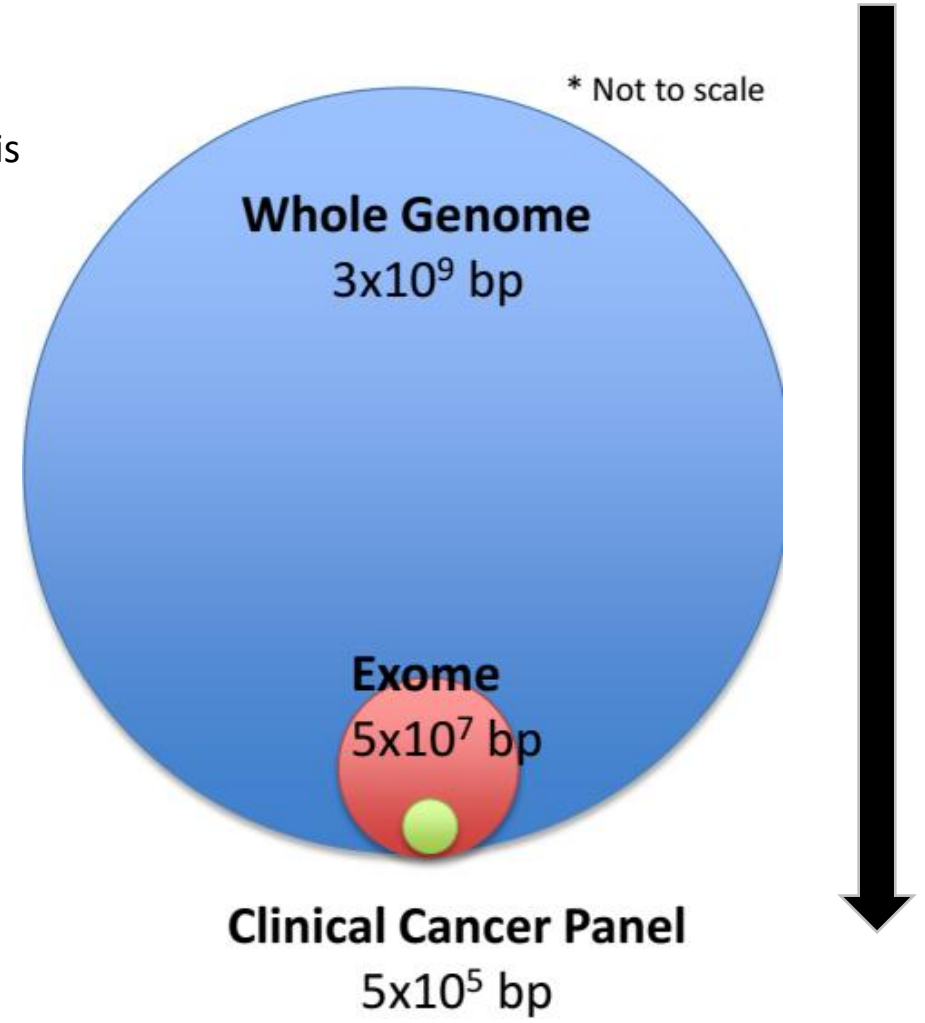
Cuánto quiero secuenciar?

$$$
Costos

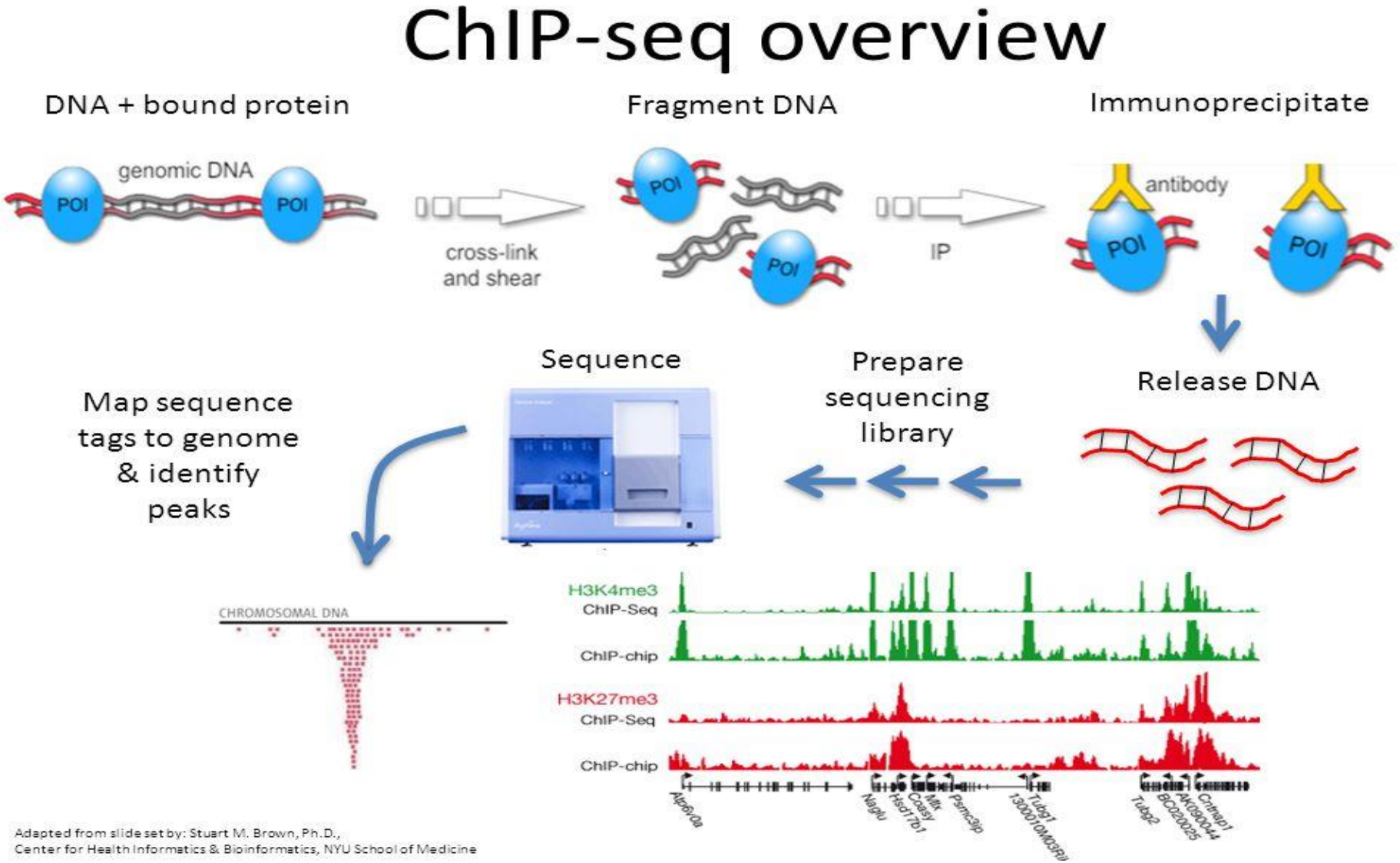
- **Genoma**:
  - Ventajas: CNVs, Fusiones, No codificante, regiones reguladoras, sin hipótesis previa.
  - Desventajas: Muy caro, mas laborioso (Bioinformática) análisis.
- **Exoma**:
  - Ventajas: detección de mutaciones en cualquier codón, CNVs
  - Desventajas: Caro, mas laborioso, mucha información.
- **Paneles de Genes**:

  Entre 50 – 300 genes (Amplicones)

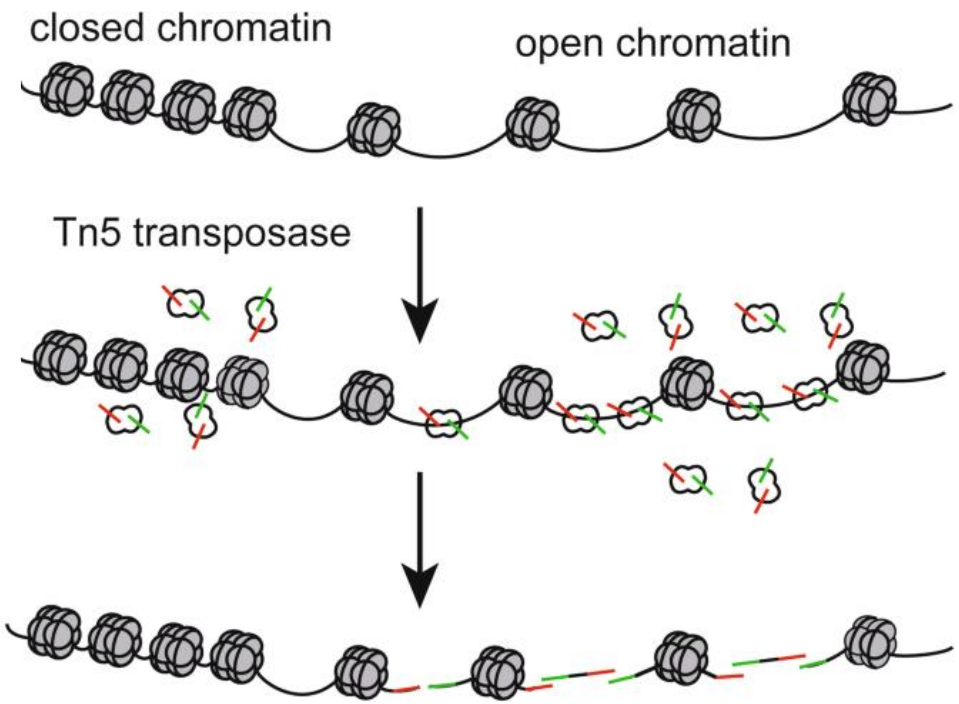
  - Ventajas: barato y rápido análisis.
  - Desventajas: Acotado y limitado en variantes estructurales.

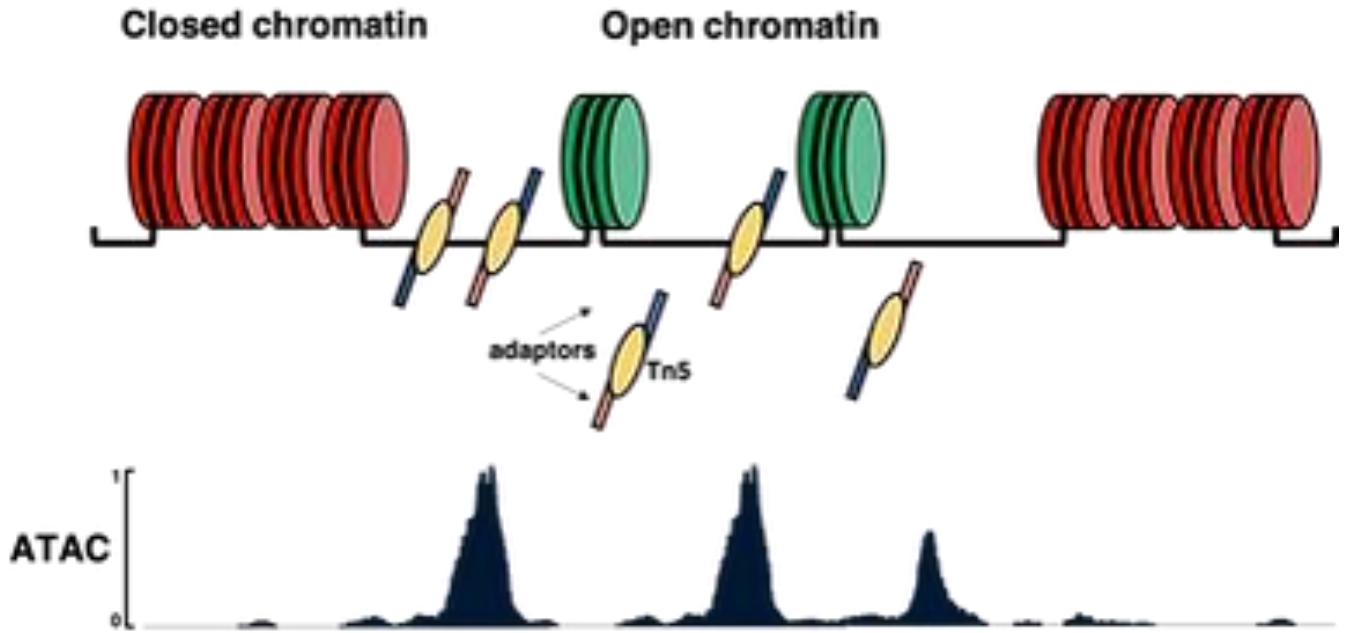


estrategia de secuenciación: CHIP-SEQ



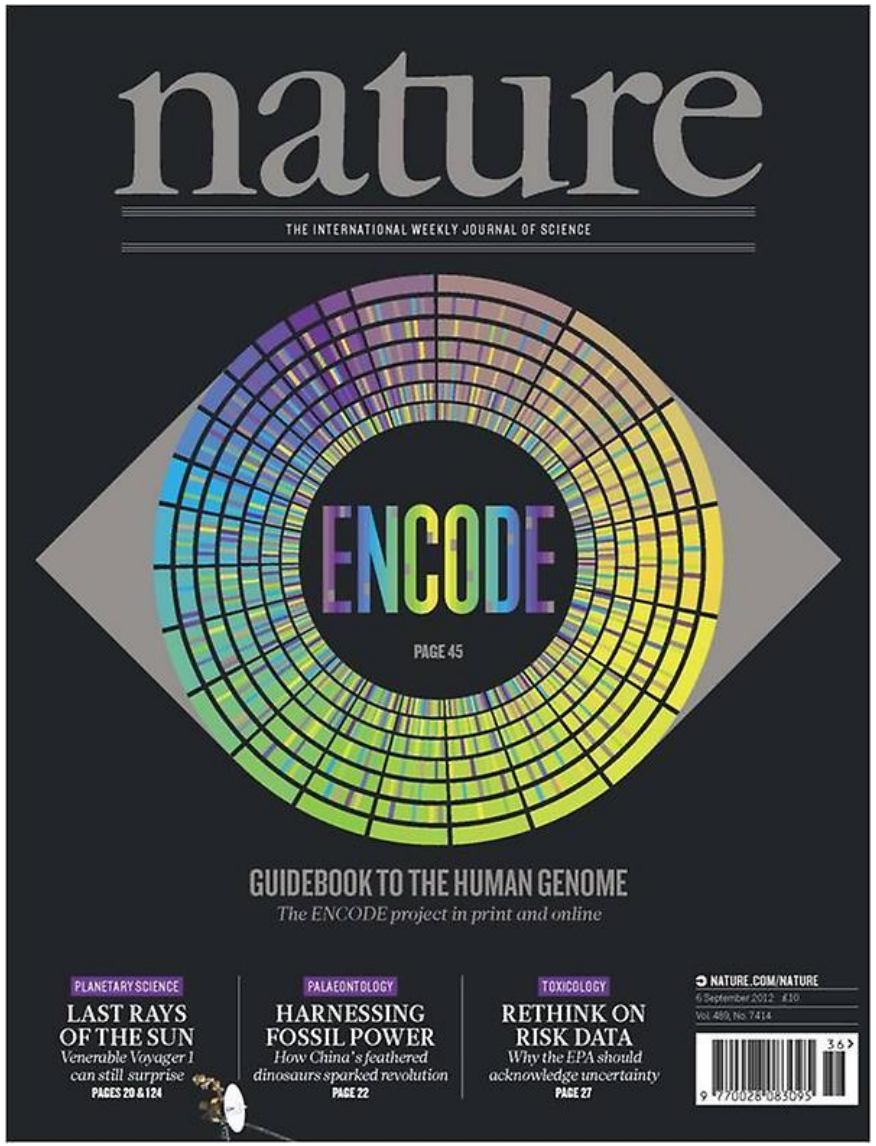
estrategia de secuenciación: ATAC-SEQ

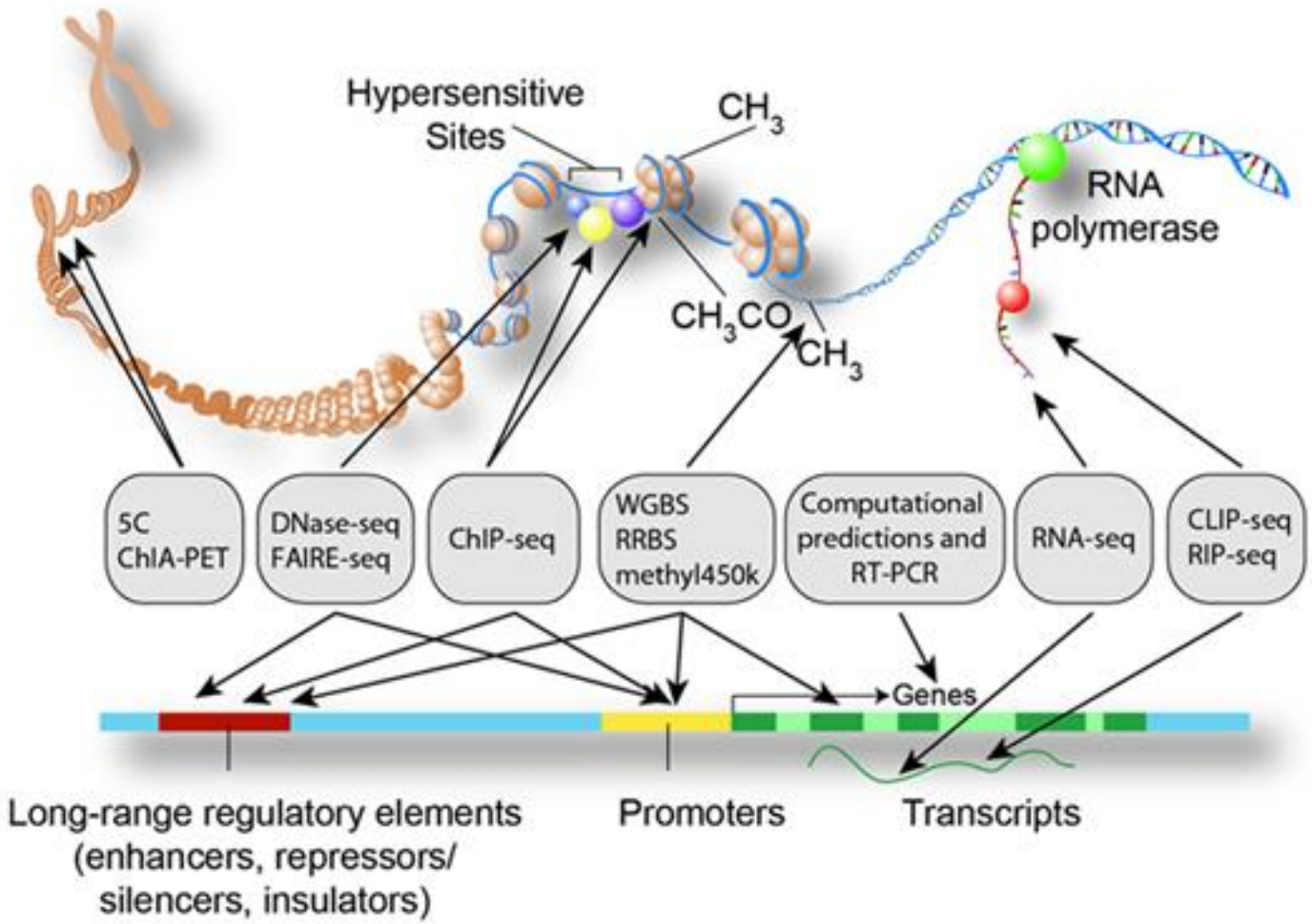


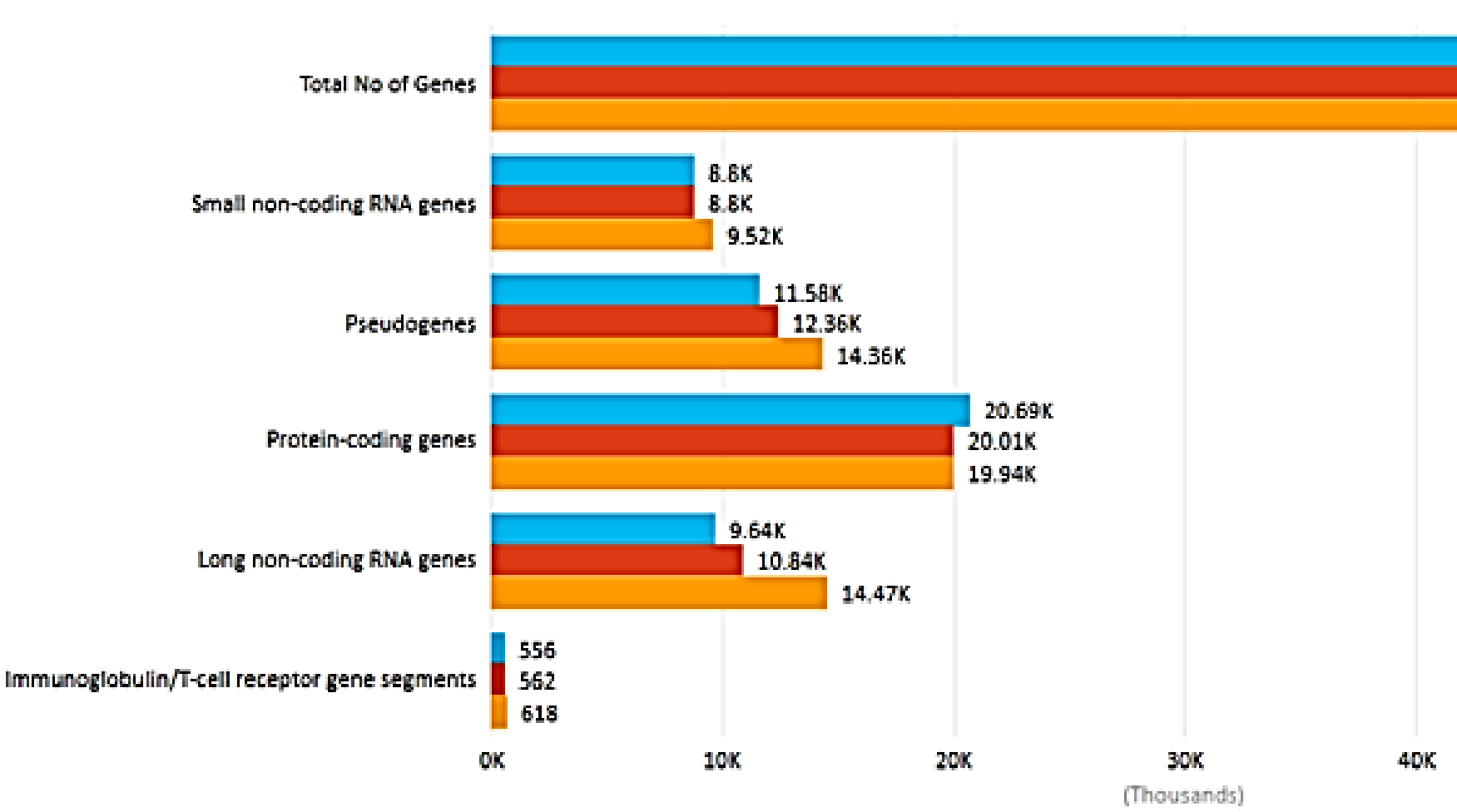


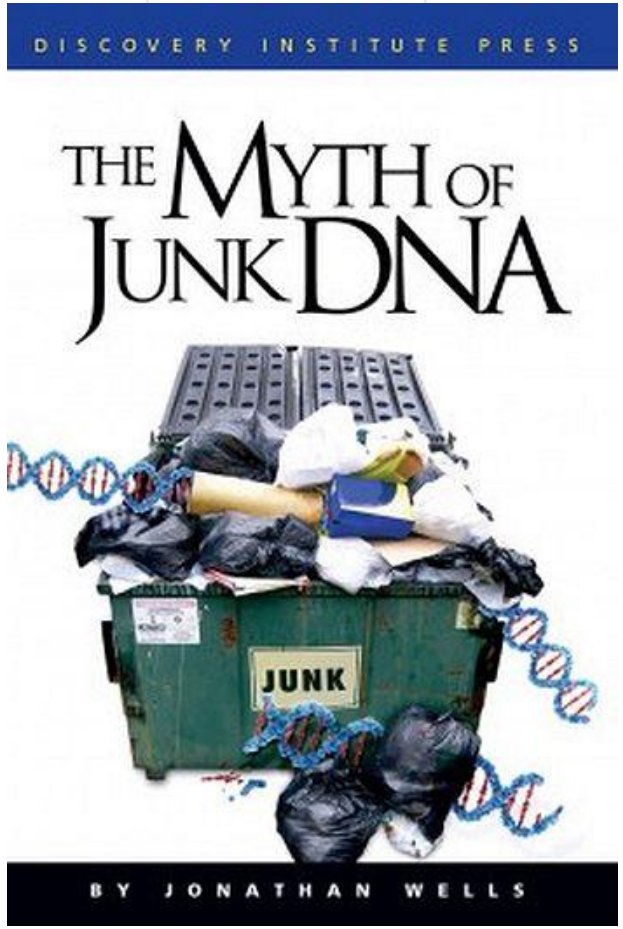
https://doi.org/10.1186/s41065-019-0105-9

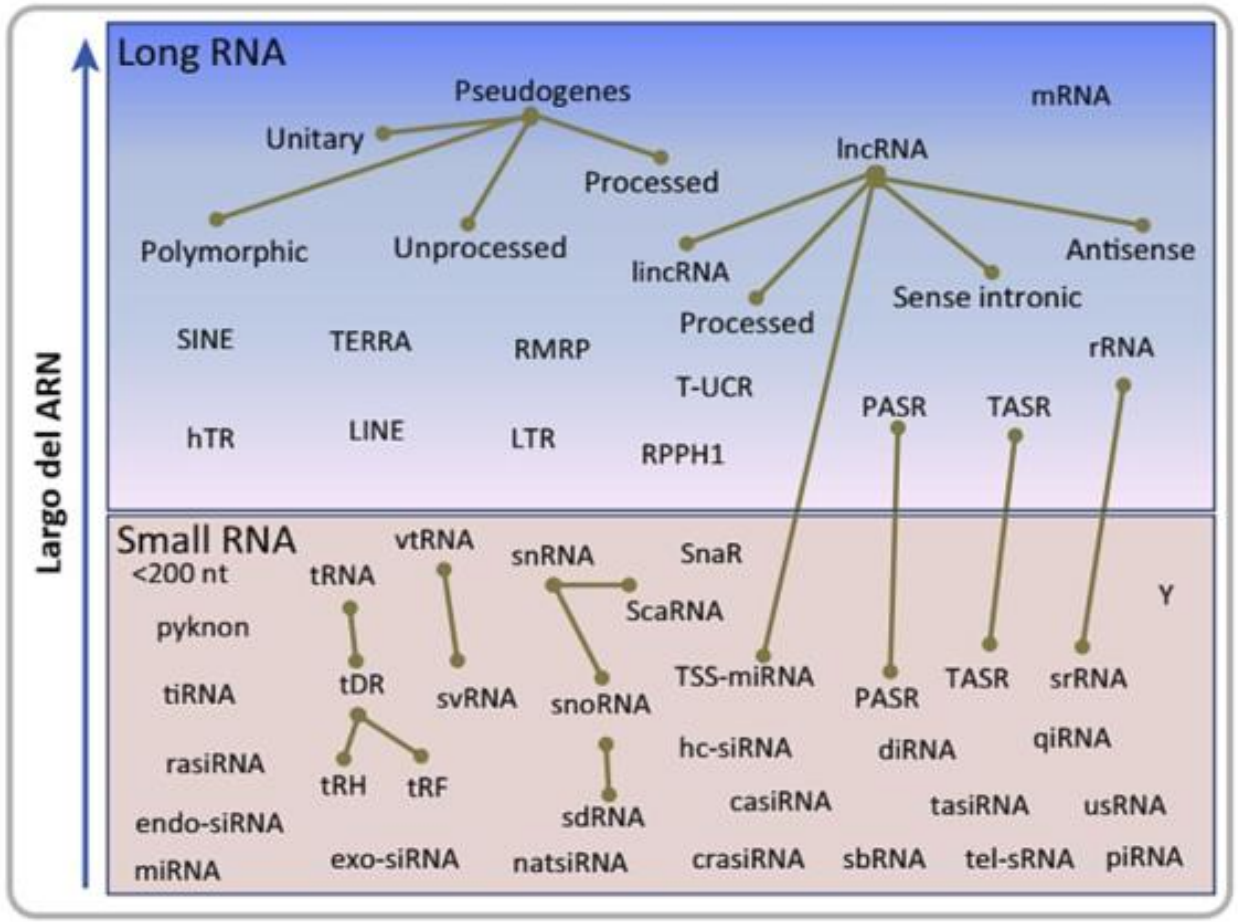
Proyecto ENCODE 2003-2016.






Version
v07
v10
v20
Total No of Genes
51.08K
52.38K
58.69K
Small non-coding RNA genes
8.6K
8.6K
9.52K
Pseudogenes
11.58K
12.36K
14.36K
Protein-coding genes
20.69K
20.01K
19.94K
Long non-coding RNA genes
9.64K
10.84K
14.47K
Immunoglobulin/T-cell receptor gene segments
556
562
618
0K
10K
20K
30K
40K
(Thousands)


DISCOVERY INSTITUTE PRESS
THE MYTH OF JUNK DNA
JUNK
BY JONATHAN WELLS



Gloss and Dinger *Experimental & Molecular Medicine* (2018) 50:97
DOI 10.1038/s12276-018-0087-0

Experimental & Molecular Medicine

REVIEW ARTICLE Open Access

Realizing the significance of noncoding functionality in clinical genomics

Brian S. Gloss[1,2] and Marcel E. Dinger[1,2,3]

Non-Coding Loss-of-Function Variation in Human Genomes

Zachary Zappala[a] Stephen B. Montgomery[a, b]

Departments of [a]Genetics and [b]Pathology, Stanford University, Stanford, Calif., USA

Trends in Genetics

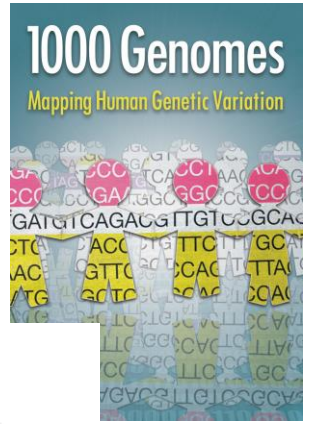
CellPress

Review

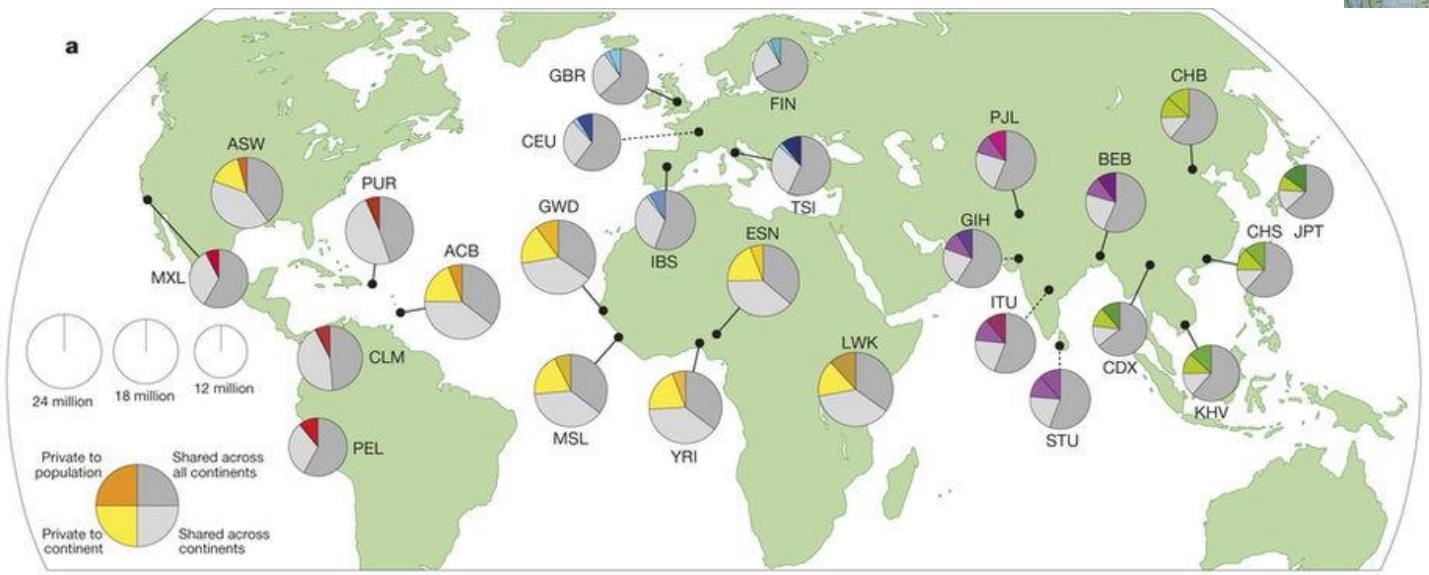
The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome

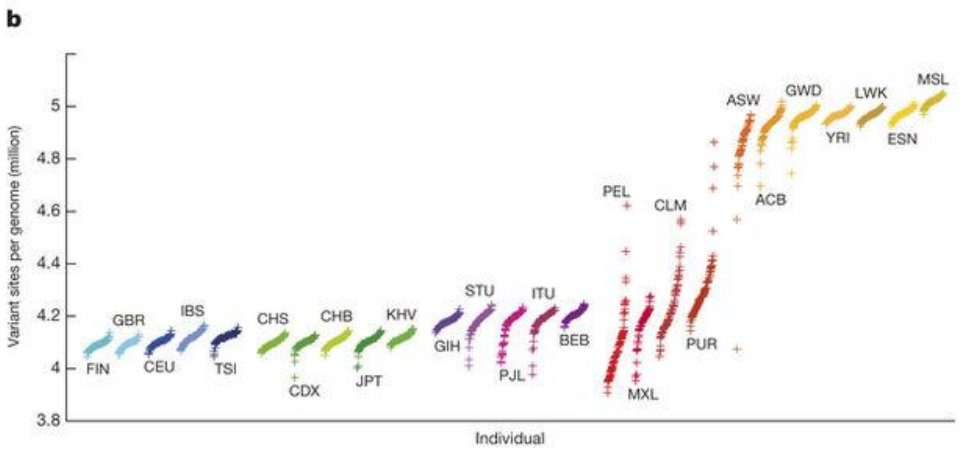
Ira W. Deveson,[1,2] Simon A. Hardwick,[1,3] Tim R. Mercer,[1,3] and John S. Mattick[1,2,3,*]

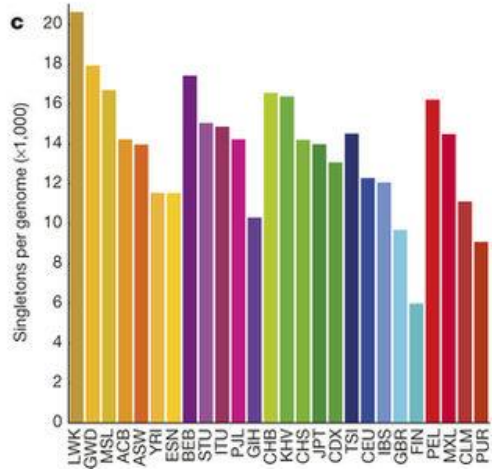
Aplicaciones: estudio de variantes.



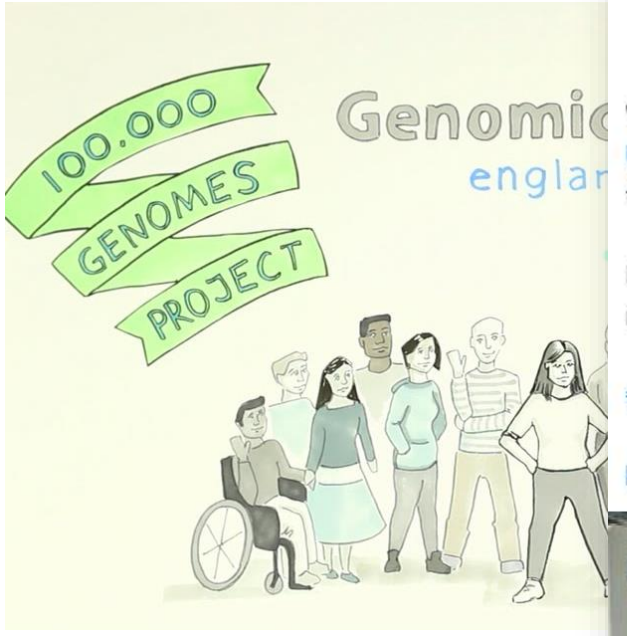
- El Proyecto 1000 genomas es una iniciativa para analizar el material genético de **mil personas** en todo el mundo, con la finalidad de obtener una base de datos que permita estudiar la variabilidad genética humana.
- Según los resultados obtenidos cada individuo sano tiene en su genoma por término medio alrededor de 150 anomalías que provocan la finalización prematura de proteínas y otras 30 implicadas en la aparición de enfermedades raras.
- También se ha confirmado el origen único de todos los humanos actuales, en un antepasado común de hace entre 150 000 y 200 000 años.







Aplicaciones: estudio de variantes.



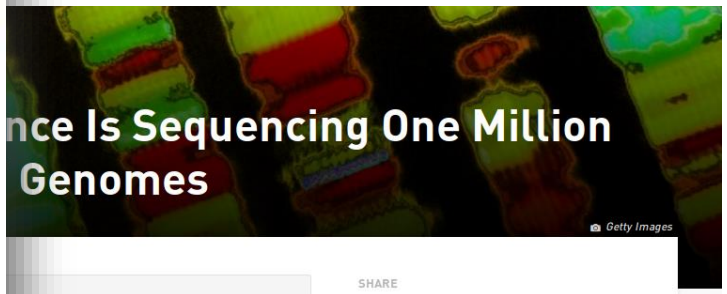
Wellcome Sanger Institute
45,668 followers
35m •
+ Follow

We have sequenced over 243,000 human genomes in a record 3.5 years, for the UK Biobank project. The data will be available, together with health information, for researchers studying diseases including cancer, Alzheimer's and diabetes.

In our latest blog, some of the key people at the Wellcome Sanger Institute reflect on what it took to deliver sequencing at such scale.

#genomes #scale #teamwork

https://lnkd.in/eSd5zKvG

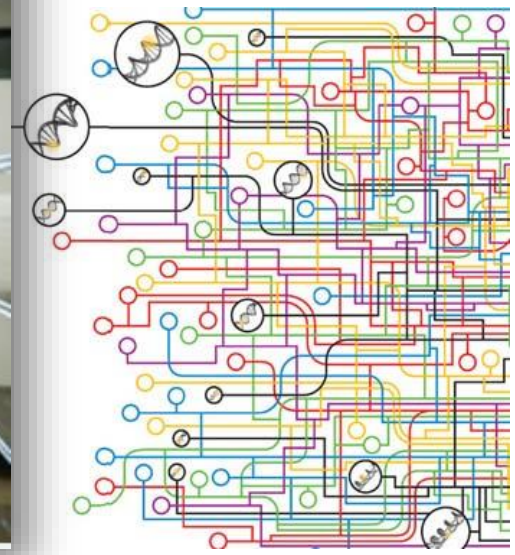


Medicine Big Data Nanjing Center will
s project is an effort to identify
ealth disorders.

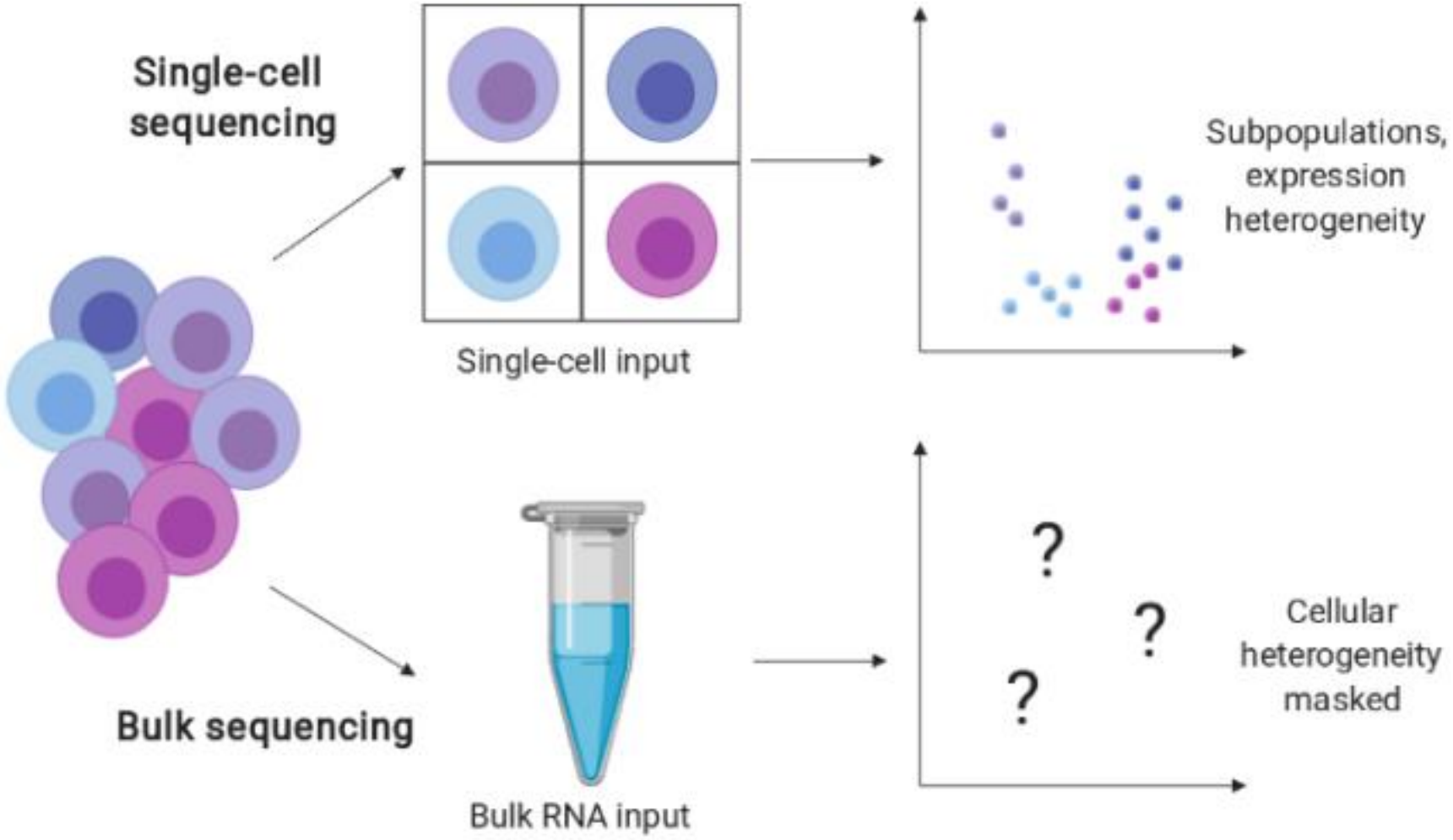
SHARE

WRITTEN BY

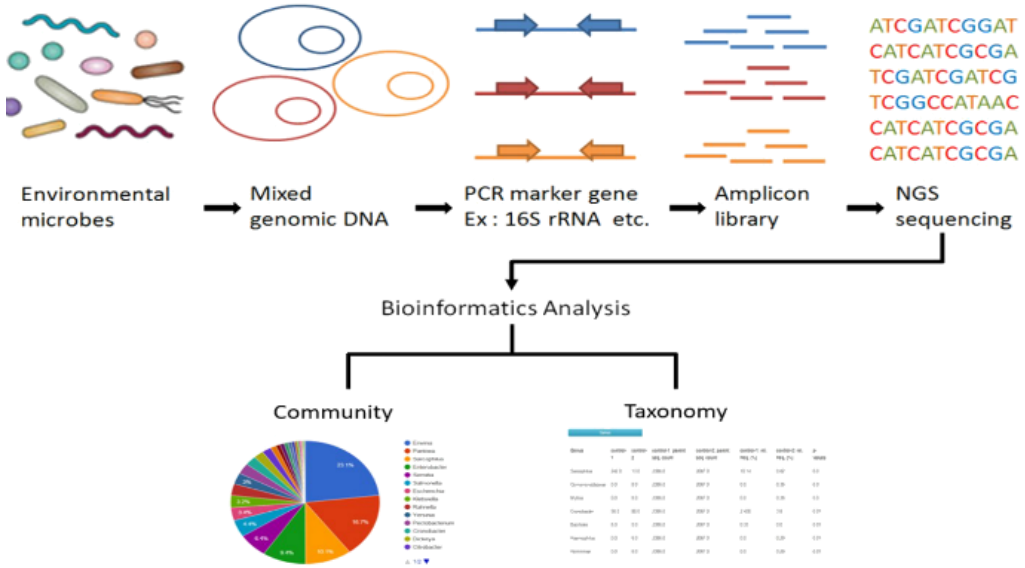
Claudia Geib

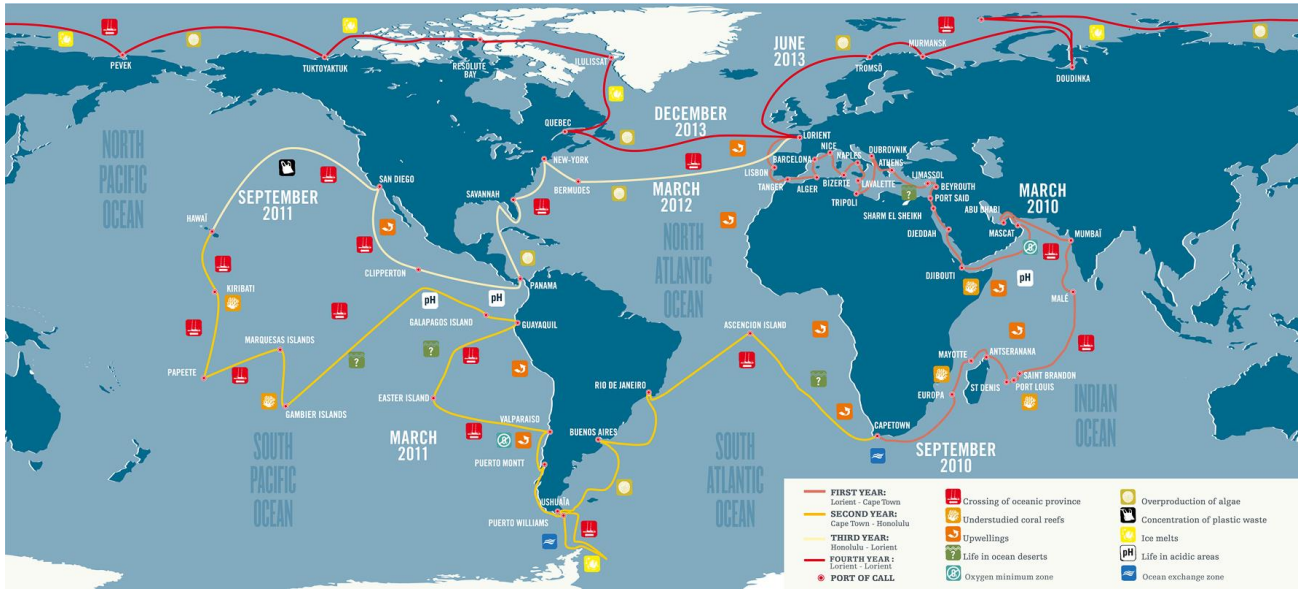


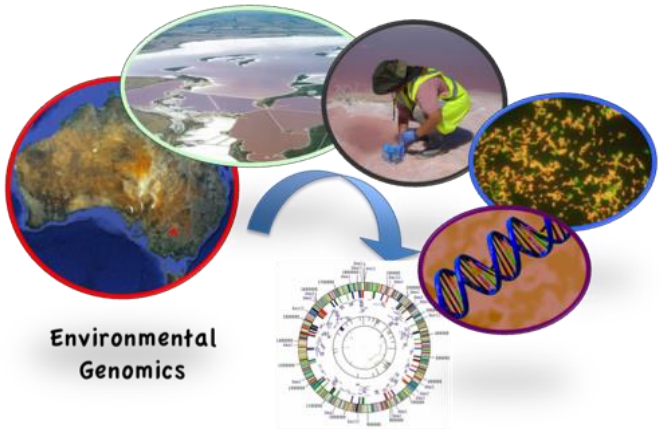
estrategia de secuenciación: Single cell SEQ

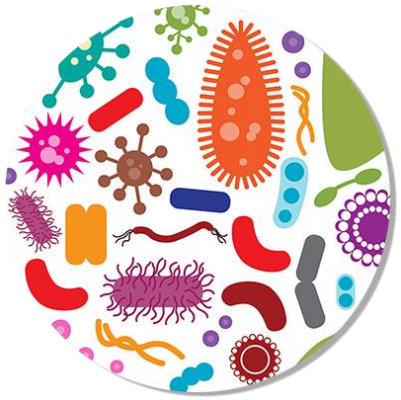


Aplicaciones: metagenómica ambiental

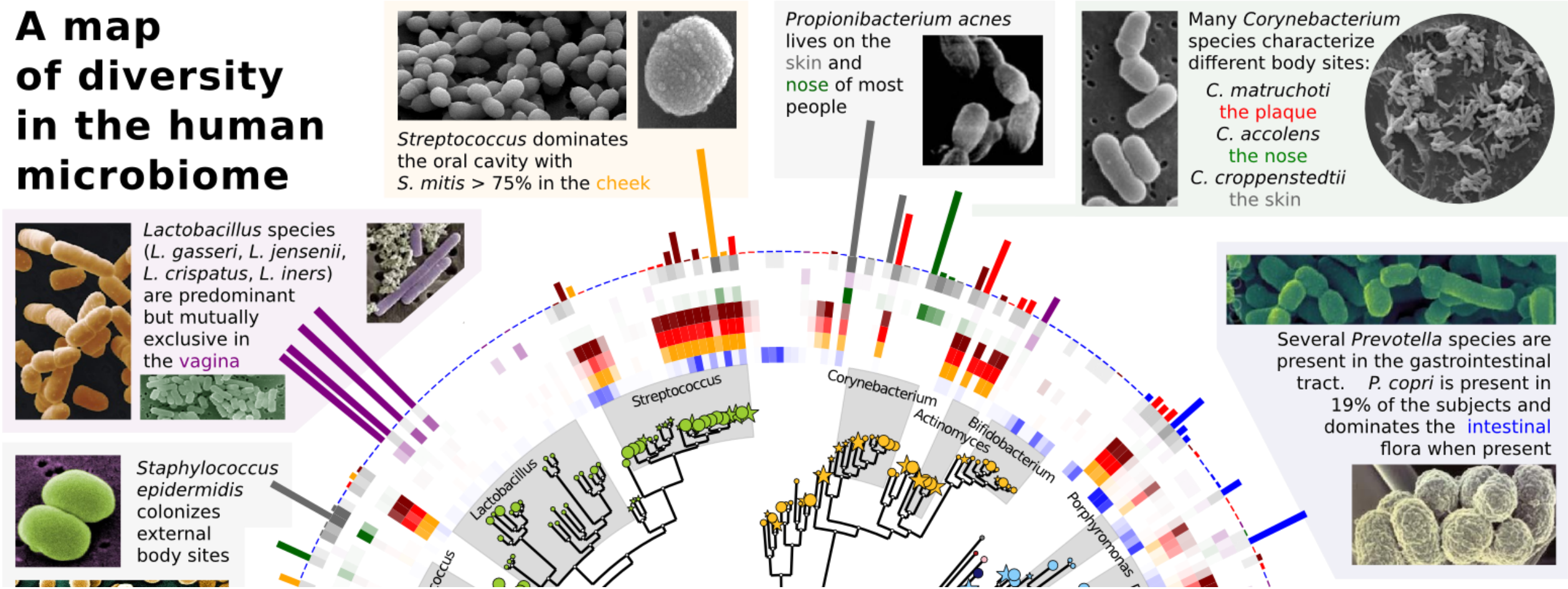








Aplicaciones: ¿Con quiénes convivimos?



Aplicaciones: abarcando la complejidad de la vida

Earth BioGenome Project: Sequencing life for the future of life

Harris A. Lewin, Gene E. Robinson, W. John Kress, William J. Baker, Jonathan Coddington, Keith A. Crandall, Richard Durbin, Scott V. Edwards, Félix Forest, M. Thomas P. Gilbert, Melissa M. Goldstein, Igor V. Grigoriev, Kevin J. Hackett, David Haussler, Erich D. Jarvis, Warren E. Johnson, Aristides Patrinos, Stephen Richards, Juan Carlos Castilla-Rubio, Marie-Anne van Sluys, Pamela S. Soltis, Xun Xu, Huanming Yang and Guojie Zhang

PNAS April 23, 2018. 201720115; published ahead of print April 23, 2018. https://doi.org/10.1073/pnas.1720115115
Add to Cart ($10)

Edited by John C. Avise, University of California, Irvine, CA, and approved March 15, 2018 (received for review January 6, 2018)

Article | Figures & SI | Authors & Info | PDF

Abstract

Increasing our understanding of Earth's biodiversity and responsibly stewarding its resources are among the most crucial scientific and social challenges of the new millennium. These challenges require fundamental new knowledge of the organization, evolution, functions, and interactions among millions of the planet's organisms. Herein, we present a perspective on the Earth BioGenome Project (EBP), a moonshot for biology that aims to sequence, catalog, and characterize the genomes of all of Earth's eukaryotic biodiversity over a period of 10 years. The outcomes of the EBP will inform a broad range of major issues facing humanity, such as the impact of climate change on biodiversity, the conservation of endangered species and ecosystems, and the preservation and enhancement of ecosystem services. We describe hurdles that the project faces, including data-sharing policies that ensure a permanent, freely available resource for future scientific discovery while respecting access and benefit sharing guidelines of the Nagoya Protocol. We also describe scientific and organizational challenges in executing such an ambitious project, and the structure proposed to achieve the project's goals. The far-reaching potential benefits of creating an open digital repository of genomic information for life on Earth can be realized only by a coordinated international effort.

EARTH BIOGENOME PROJECT

Aplicaciones: abarcando la complejidad de la vida

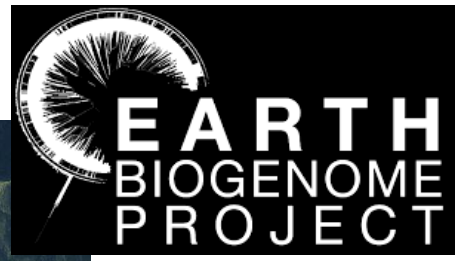
Earth BioGenome Project: Sequencing life for the future of life

Harris A. Lewin, Gene E. Robinson, W. John Kress, William J. Baker, Jonathan Coddington, Keith A. Crandall, Richard Durbin, Scott V. Edwards, Félix Forest, M. Thomas P. Gilbert, Melissa M. Goldstein, Igor V. Grigoriev, Kevin J. Hackett, David Haussler, Erich D. Jarvis, Warren E. Johnson, Aristides Patrinos, Stephen Richards, Juan Carlos Castilla-Rubio, Marie-Anne van Sluys, Pamela S. Soltis, Xun Xu, Huanming Yang and Guojie Zhang

PNAS April 23, 2018. 201720115; published ahead of print April 23, 2018. https://doi.org/10.1073/pnas.1720115115
Add to Cart ($10)

Edited by John C. Avise, University of California, Irvine, CA, and approved March 15, 2018 (received for review January 6, 2018)

Article



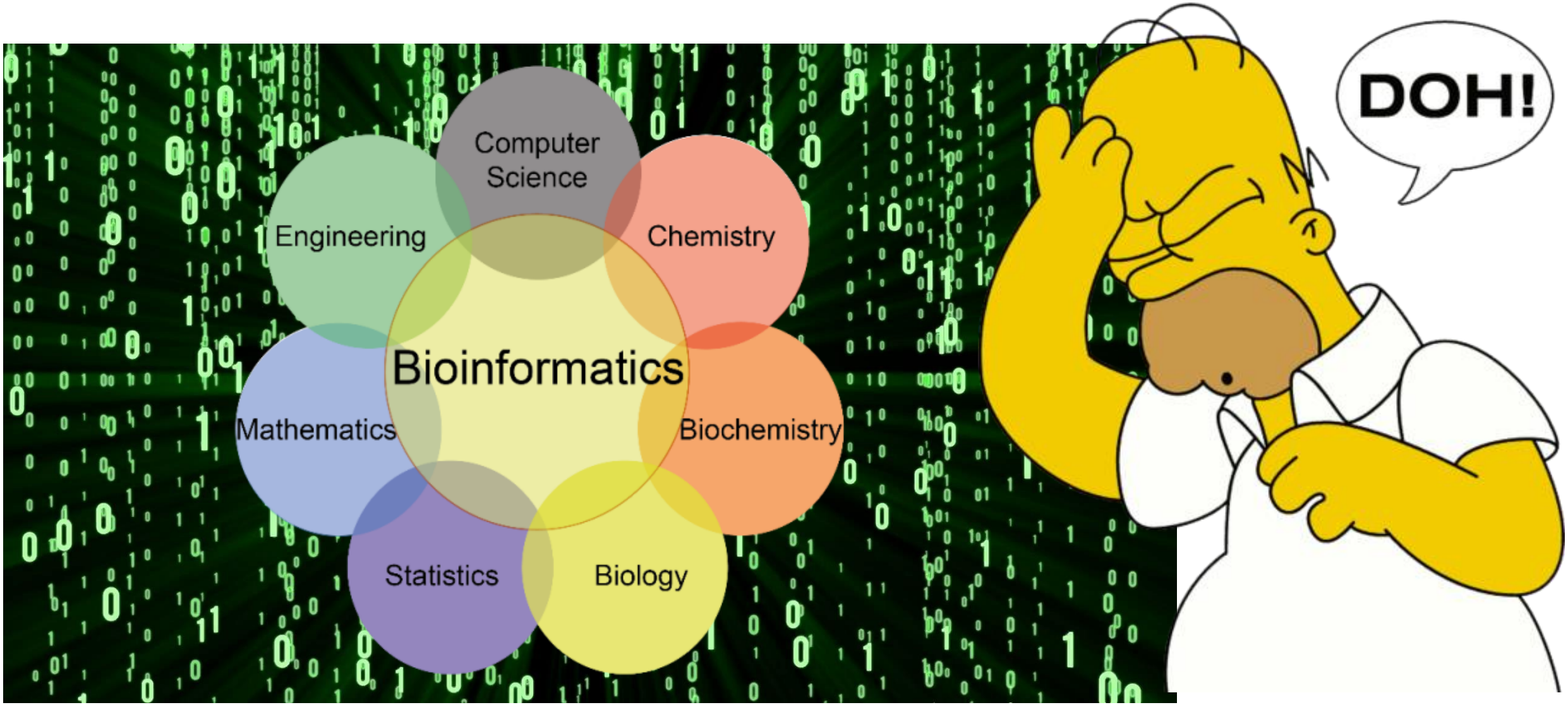
El proyecto tiene como objetivo secuenciar y anotar las aproximadamente 1,5 millones de especies eucariotas conocidas.

Se proyecta que el proyecto costará 4.700 millones de dólares

Abstract

Increasing our und
are among the mo
challenges require
interactions among millions of the planet's organisms. Herein, we present a perspective on the Earth BioGenome Project (EBP), a moonshot for biology that aims to sequence, catalog, and characterize the genomes of all of Earth's eukaryotic biodiversity over a period of 10 years. The outcomes of the EBP will inform a broad range of major issues facing humanity, such as the impact of climate change on biodiversity, the conservation of endangered species and ecosystems, and the preservation and enhancement of ecosystem services. We describe hurdles that the project faces, including data-sharing policies that ensure a permanent, freely available resource for future scientific discovery while respecting access and benefit sharing guidelines of the Nagoya Protocol. We also describe scientific and organizational challenges in executing such an ambitious project, and the structure proposed to achieve the project's goals. The far-reaching potential benefits of creating an open digital repository of genomic information for life on Earth can be realized only by a coordinated international effort.

Análisis de datos: Bioinformática

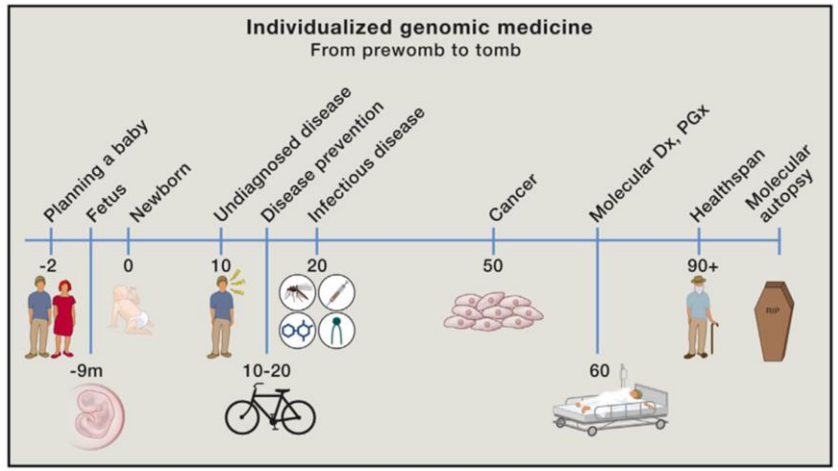


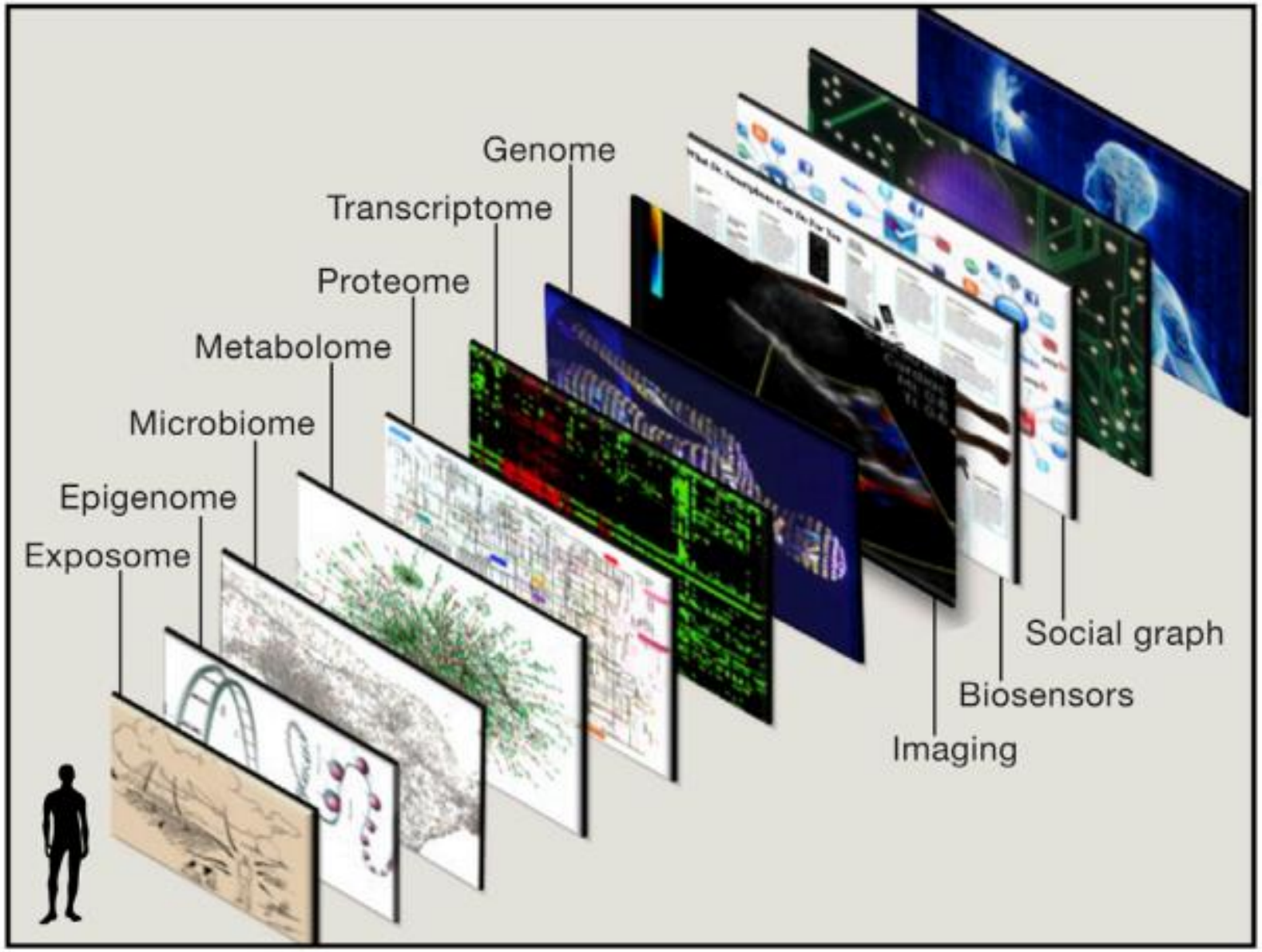
Leading Edge
Review

Cell

Individualized Medicine from Prewomb to Tomb

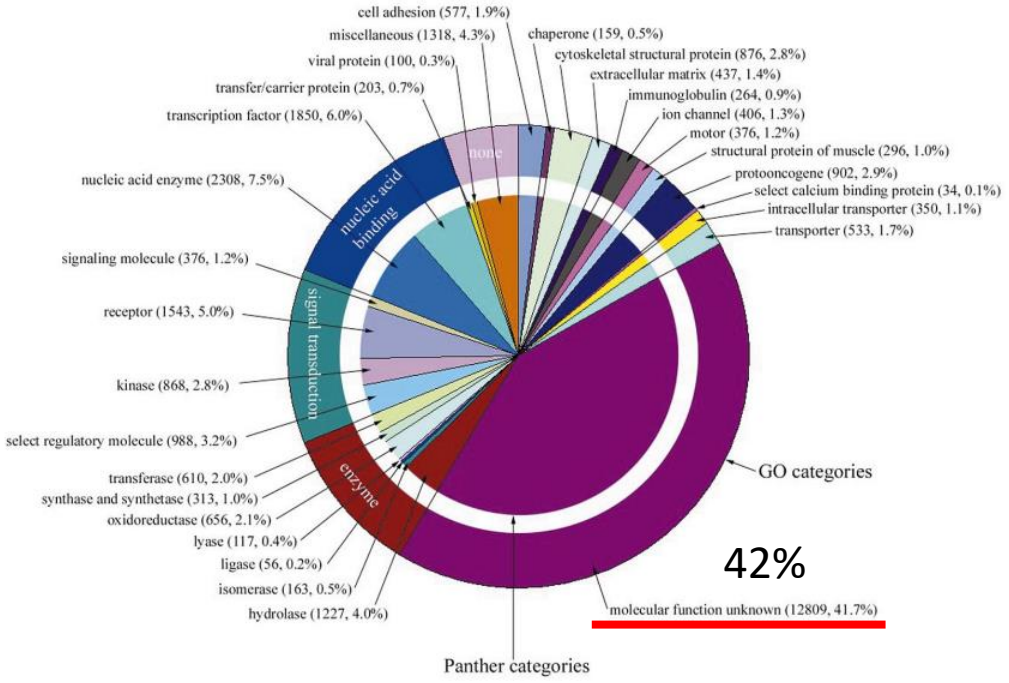
Eric J. Topol[1,*]
[1]The Scripps Translational Science Institute, The Scripps Research Institute and Scripps Health
*Correspondence: etopol@scripps.edu
http://dx.doi.org/10.1016/j.cell.2014.02.012





Cuántos genes sin función definida existen en el genoma humano?

2001



2019

Hidden in plain sight: what remains to be discovered in the eukaryotic proteome?

Valerie Wood[1,2], Antonia Lock[3], Midori A. Harris[1,2], Kim Rutherford[1,2], Jürg Bähler[3] and Stephen G. Oliver[1,2]

Alrededor del 20% de las proteínas incluso en organismos modelo bien estudiados no poseen descripciones informativas de sus funciones biológicas. En el humano 3117 (15% aprox.)

"Esta es la alegría de la ciencia y la investigación: el trabajo nunca está completo. Cada avance abre nuevas perspectivas de oportunidades con la frecuente sensación de que lo mejor aún está por venir."

Eric D. Green is director of the National Human Genome Research Institute (NHGRI) at the National Institutes of Health

Muchas gracias por su atención