Davide Barbagallo,

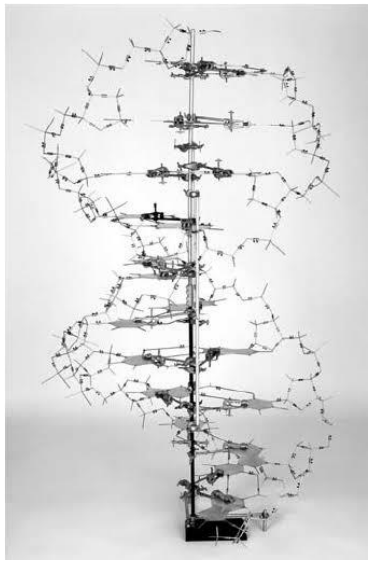
Accademia Gioenia di Catania

1 Dicembre 2023

“Le molteplici sfide della materia oscura del genoma umano”

Milestone 1

April 1953



Human Genome Project
(official start – October 1990)

- 1961: Niremberg, Khorana and coll. cracked the «code for life»
- 1968: The Nobel Prize in Physiology or Medicine 1968 was awarded jointly to Robert W. Holley, Har Gobind Khorana and Marshall W. Nirenberg "for their interpretation of the genetic code and its function in protein synthesis“
- 1972: Paul Berg and coll. publish the first paper on a recombinant DNA molecule. It’s the beginning of genetic engineering
- 1975 (and following 1977): Frederick Sanger and coll. develop the “dideoxys’ method” for sequencing DNA. The genome of ΦX174 virus is sequenced.
- 1977: Allan Maxam and Walter Gilbert develop a DNA sequencing technique complementary to that published by Sanger
- 1980: The Nobel Prize in Chemistry 1980 was divided, one half awarded to Paul Berg "for his fundamental studies of the biochemistry of nucleic acids, with particular regard to recombinant-DNA", the other half jointly to Walter Gilbert and Frederick Sanger "for their contributions concerning the determination of base sequences in nucleic acids”
- 1980: David Botstein e coll, obtained the first Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms
- 1981: Anderson and coll. sequenced for the first time the 16,569 bp of the human mtDNA
- 1983: Kary Mullis invents the Polymerase Chain Reaction (in 1993 he will receive Nobel prize in Chemistry together with Michael Smith)

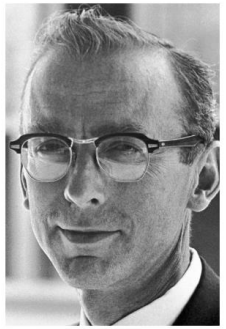


Photo from the Nobel Foundation archive.
Robert W. Holley
Prize share: 1/3

Photo from the Nobel Foundation archive.
Har Gobind Khorana
Prize share: 1/3

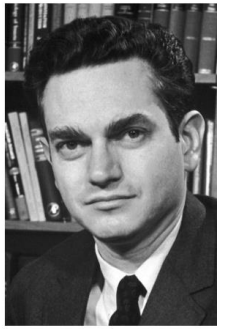


Photo from the Nobel Foundation archive.
Marshall W. Nirenberg
Prize share: 1/3

Photo from the Nobel Foundation archive.
Paul Berg
Prize share: 1/2

Photo from the Nobel Foundation archive.
Walter Gilbert
Prize share: 1/4

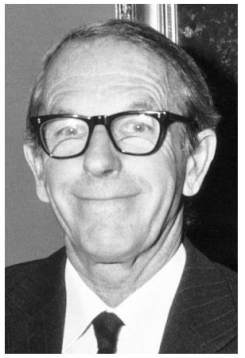


Photo from the Nobel Foundation archive.
Frederick Sanger
Prize share: 1/4

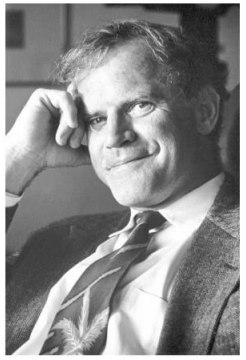


Photo from the Nobel Foundation archive.
Kary B. Mullis
Prize share: 1/2

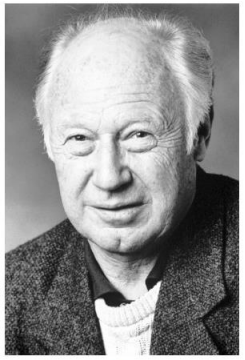


Photo from the Nobel Foundation archive.
Michael Smith
Prize share: 1/2

Cedalion on Orion's shoulders.
1658 painting by Nicolas Poussin

“Hiroshima meeting”, March 1984 and further “The Alta Summit”, December 1984

Charles DeLisi, Professor of Science and Engineering at Boston University. Director of the Office of Health at DOE (1985)

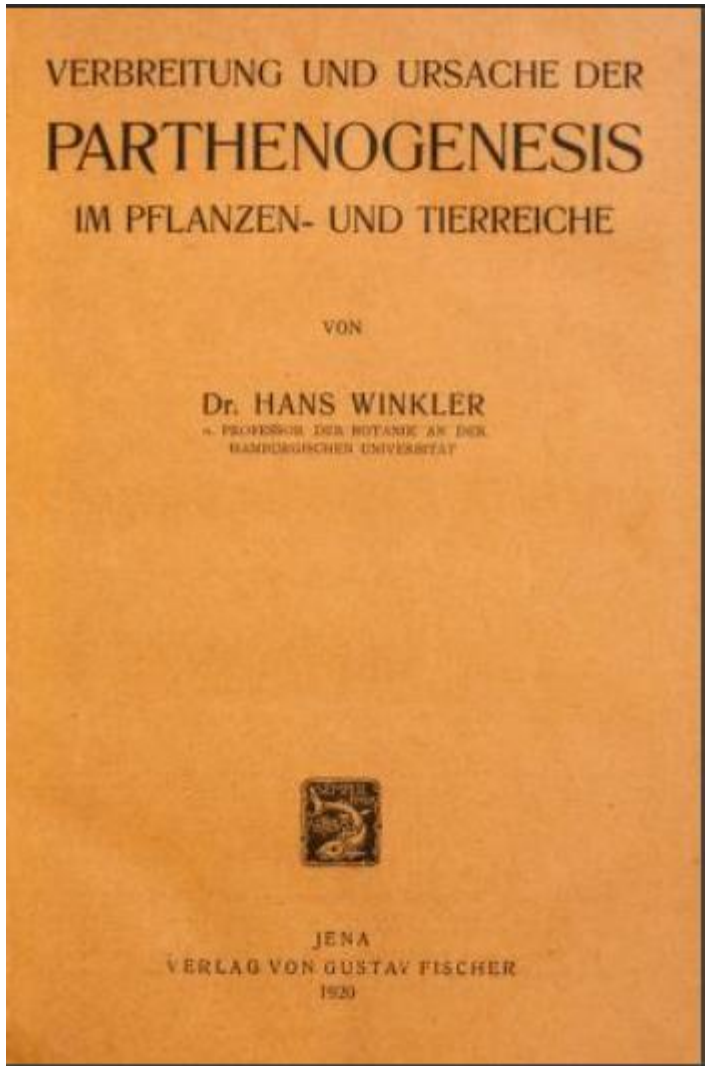
The Alta meeting is thus the bridge from DOE's traditional interest in detection of mutations to DeLisi's push for a Human Genome Initiative, and provides one of several historical links between genome projects and another massive technical undertaking of the 20th century—the Manhattan project.

The origin of term «Genome»

Ich schlage vor, für den haploiden Chromosomensatz, der im Verein mit dem zugehörigen Protoplasma die materielle Grundlage der systematischen Einheit darstellt den Ausdruck: das Genom zu verwenden ... (Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche. Winkler, Hans. Jena, G. Fischer, 1920).

This may be translated as: "I propose the expression ***Genom*** *for the haploid chromosome set, which, together with the pertinent protoplasm, specifies the material foundations of the species ..."*

From the Commentary: *The Scientist* 15[7]:8, Apr. 2, 2001



Letter from Robert Sinsheimer to the President of UCSC: the Human Genome Project (HGP) is proposed for the first time

November 19, 1984

President David Pierpont Gardner
Office of the President
714 University Hall
Berkeley, California 94720

Dear David:

Let me expand a bit on our brief discussion at the Regents' Meeting on Friday.

If the "Hoffmans" firmly intend to withdraw from the TMT project, then I have another project that we might propose to them. It is an opportunity to play a major role in a historically unique event - the sequencing of the human genome.

A genome is the complete set of DNA instructions for the making of a species. The human genome is the complete set of DNA instructions for a human being. We know that the haploid human genome is composed of a sequence of some three billion nucleotide pairs (3×10^9).

A few months ago, I posed to our biologists the question: could the human genome now be sequenced, with extant technique, and in a reasonable time (10 years)? If so, what scale of effort would be required? (Obviously, I had made a guess as to the answer.)

Their reply is enclosed. It can be done. We would need a building in which to house the Institute formed to carry out the project (cost approximately $25 million), and we would need an operating budget of some $5 million/year (in current dollars). Not at all extraordinary.

Clearly, the human genome will be sequenced. It will be done, once and for all time, providing a permanent and priceless addition to our knowledge.

In addition to satisfying our scientific curiosity, this knowledge will provide deep insight into other questions of interest. It will have major medical implications: we know that literally thousands of human ailments have genetic bases, in whole or part.

This knowledge will also have highly significant evolutionary implications. The biological differences between homo sapiens and the chimpanzee are certainly due to the changes and rearrangements in the genomes of each as they have diverged from that of our common ancestor. To understand these changes will surely illuminate the ancient human quest to know what we are and where we came from.

The enterprise could be known as the Hoffman Project and, of course, the building could be named the Hoffman Laboratory or Institute. If we had the building and equipment, I feel quite confident we could obtain the operating funds from government and/or private sources.

Needless to say, should the "Hoffmans" not be interested in this project, I will intend to look elsewhere for funding.

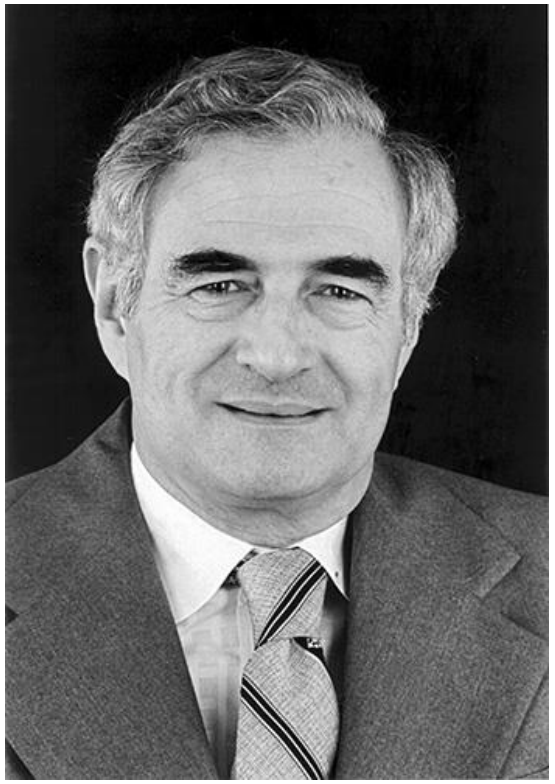
Sincerely yours,

Robert L. Sinsheimer
Chancellor

cc: Vice President Frazer
Vice Chancellor Moldave

enclosure

Figure 1. Meeting organizers: Sinsheimer, Edgar, Lugwig, Noller. Permission to use photograph: David Haussler, UC Santa Cruz.



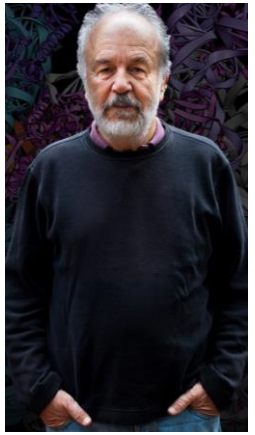
Robert L. Sinsheimer (1920 – 2017)
Chancellor emeritus at the UCSC

Answer of the biologists Bob Edgar, Harry Noller e Bob Ludwig to the questions of Robert (Bob) Sinsheimer

We see this as a noble and inspiring enterprise. In some respects, like the journeys to the moon, it is simply a "tour de force"; it does not necessarily follow that knowledge of the nucleotide sequence of the human genome will provide deep insights into the physical nature of man. Nevertheless, we are confident that this project will provide an integrating focus for all efforts to use DNA cloning techniques in the study of human genetics. The ordered library of cloned DNA that must be produced to allow the genome to be sequenced will itself be of great value to all human genetics researchers. The project will also provide an impetus for improvements in these techniques, techniques that already have revolutionized the nature of biological research in all areas, from biochemistry to evolution, and promise to have wide-ranging applications in agriculture and medicine.

The scale of this project is indeed awesome. The largest genome so far completely sequenced, the Epstein-Barr virus, is 1.72×10^5 nucleotides in length (Baer et al., 1984) ; the human genome is twenty thousand times larger. However, the technology to carry out this project is already at hand and much of it is routine, and so can be scaled up using bio-engineering procedures. Ultimately we envision that the actual sequence-gathering steps will be automated and computerized.

Robert Edgar, † professor emeritus of biology at UC Santa Cruz



Harry F. Noller, director of the University of California, Santa Cruz's Center for the Molecular Biology of RNA. Noller has just received the 2017 Breakthrough Prize in Life Sciences for his discoveries about the ribosome *Science* magazine (Sept. 24, 1999), the tiny structure of the cell that Noller calls the “mothership of life.”

Bob Ludwig, † professor emeritus of MCD Biology at UC Santa Cruz

of the human body, forming the structure and determining the function of all the body's cells.

Roughly 100,000 genes and some associated information about when they should be turned on and off, all adding up to about three billion pairs of nucleotides, make up the genome.

Inside the nucleus of just about every human cell, the genome takes the form of a yard-long thread of DNA coiled up and packed together with some proteins in 46 squiggly-shaped chromosomes, half coming from the mother and half from the father.

The DNA book could probably tell scientists, if they knew how to read it, how to build a human body.

Errors in the sequence of nucleotides can create too much or not enough of a protein, causing the 3,000 known genetic diseases. The instructions contained in the sequence also determine much about how babies develop, how adults age, and who inherits bodily conditions making them susceptible to ailments that aren't strictly hereditary, such as heart disease.

"The common diseases all have inherited components," says Sir Walter. "This is applicable to all medical problems."

Progress in Understanding Disease

Although sequencing a human genome would help in the understanding of all human disease, it would not stamp out disease. Scientists need to understand many other steps in the building process between DNA and the human body. Scientists understaad how DNA is translated into proteins, but they don't understand how the proteins fold into their final form as compact molecules.

In the case of some genetic diseases, scientists have found the place on the genes where the disease originates, or at least a marker near the genetic error, but that is sometimes all they know. For example, researchers working at eliminating sickle cell anemia, a hereditary disease that predominantly affects blacks, know where on the genes the disease originates but they do not completely understand why the genetic defect causes the sickle-shaped red blood cells of the disease.

Uncertainty about the value of the information that would be obtained from a complete genome sequence is at the center of the debate over what resources should be devoted to the effort. Despite the uncertainty, the idea has already began to attract money.

The Department of Energy, which sponsored one of the first conferences about

Continued on Following Page

A GENETIC PRIMER

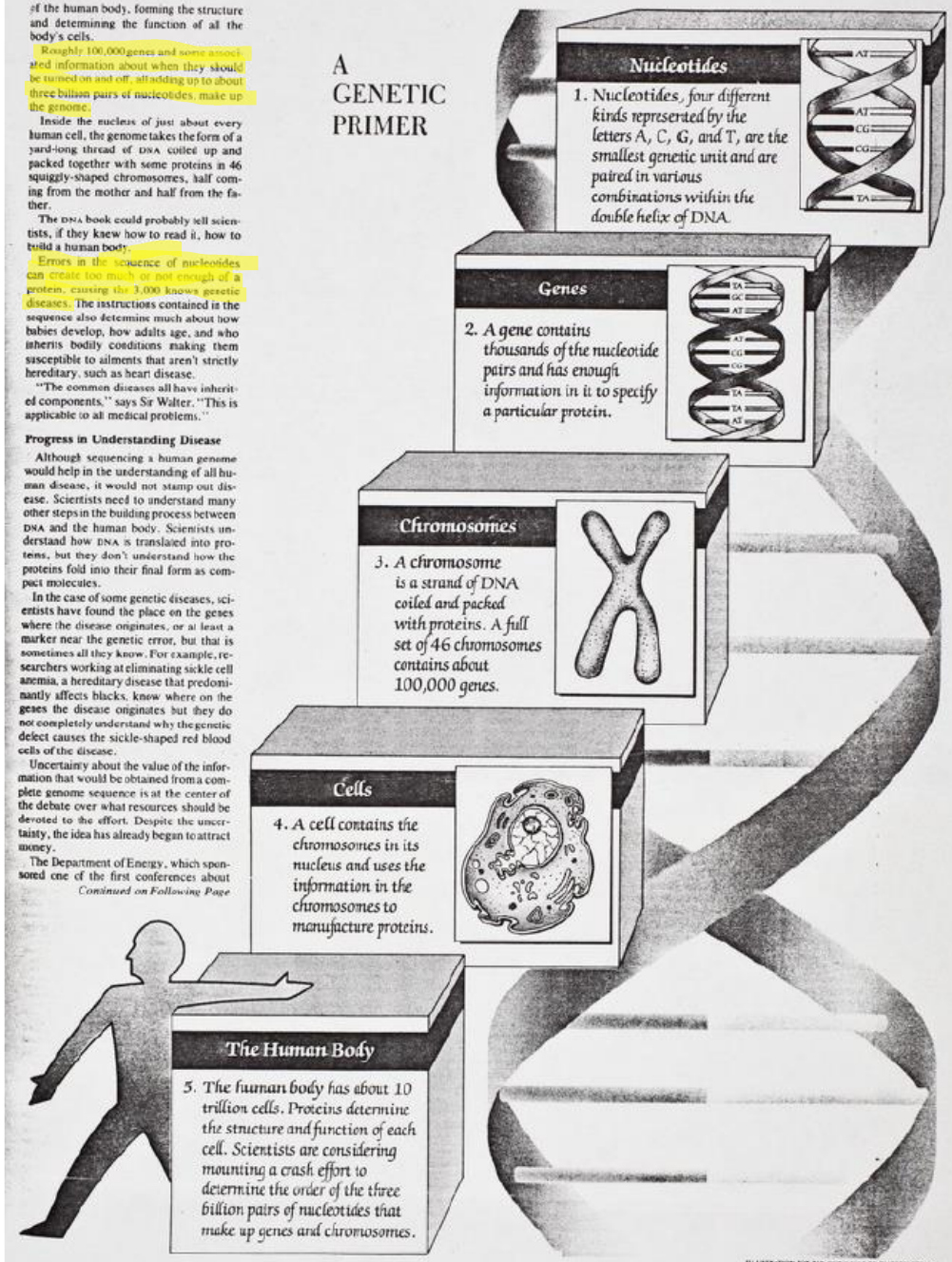


ILLUSTRATION FOR THE CHRONICLE BY WARREN JENSEN

Renato Dulbecco's perspective (1986) (President of the Salk Institute)

Perspective

A Turning Point in Cancer Research: Sequencing the Human Genome

RENATO DULBECCO

Science

A turning point in cancer research: sequencing the human genome

R Dulbecco

Science **231** (4742), 1055-1056.
DOI: 10.1126/science.3945817

GENOME SEQUENCING WORKSHOP

MARCH 3 & 4, 1986

SANTA FE, NEW MEXICO

SPONSOR

DOE

OFFICE OF HEALTH AND

ENVIRONMENTAL RESEARCH

“The National Research Council (NRC) of the National Academy of Sciences… met several times during 1987, and issued a report in February 1988 that argued strongly for a broader human genome project. The committee included several members who were initially skeptical of genome proposals, such as David Botstein, Shirley Tilghman, and Leon Rosenberg. The chairman, Bruce Alberts, had written a 1983 editorial highly critical of big science in Biology”

Origins of the human genome project
James Dewey Watson, Robert Mullan Cook-Deegan
The FASEB Journal, 1991

TOP, BOTTOM RIGHT: COLD SPRING HARBOR LAB. LIBRARY & ARCHIVES; BOTTOM LEFT: ERIC GREEN

1989, Pre-HGP Banbury meeting. Francis Collins & James Watson in top row.

1990, Washington University School of medicine. McDonald Dixon & Eric Green.

1997, HGP meeting at Cold Spring Harbor. E Green, R Myers, J Witkowski & R Gibbs.

The HGP officially started in 1990 jointly by NIH and DOE and became a truly international project with many participating countries. James Watson headed the NIH HGP and David Galas the DOE HGP project. In 1993 they were replaced by Francis Collins (NIH) and Aristides Patrino (DOE).

"What more powerful form of study of mankind could there be than to read our own instruction book" – Francis S. Collins, Director, NIH.

"Along with Bach's music, Shakespeare's sonnets, and the Apollo Space Program, the Human Genome Project (HGP) is one of those achievements of the human spirit that makes me proud to be a human" – Richard Dawkins, British ethologist & evolutionary biologist.

$3 billion - 15 years (predicted time, originally)

Main challanges and opportunities

- High costs
- Concerns on sequencing and assembling technologies (in relationship with time and repetitive sequences of the genome, respectively)
- Interdisciplinariety (needs of expertise in the field of biology, chemistry, phisics, computational analysis) and transnationality of the project (an opportunity and a risk at the same time)
- Data analysis and technical issue for the storage of the data
- Parallel sequencing of genomes of model organisms
- Ethical and legal implications (3% of the budget reserved for Ethical, Legal and Social Implications – ELSI Research Program)

The Buffalo News/Sunday, March 23, 1997

WANTED

20 Volunteers

to participate in the

Human Genome Project

a very large international scientific research effort.

The goal is to decode the human hereditary information (*human blueprint*) that determines all individual traits inherited from parents. The outcome of the project will have tremendous impact on future progress of medical science and lead to improved diagnosis and treatment of hereditary diseases.

Volunteers will receive information about the project from the Clinical Genetics Service at Roswell Park, and sign a consent form before participating.

No personal information will be maintained or transferred.

Volunteers will provide a one-time donation of a small blood specimen. A small monetary reimbursement will be provided to the participants for their time and effort.

Individuals must be at least 18 years of age.
Persons who have undergone chemotherapy are not eligible.

ROSWELL PARK CANCER INSTITUTE

For more information please contact the
Clinical Genetics Service
845-5720 (9:00 am - 3:00 pm)
March 24 - 26, 1997

The project researchers used a thoughtful process to recruit volunteers, acquire their informed consent, and collect their blood samples. Most of the human genome sequence generated by the Human Genome Project came from blood donors in Buffalo, New York; specifically, 93% from 11 donors, and 70% from one donor.

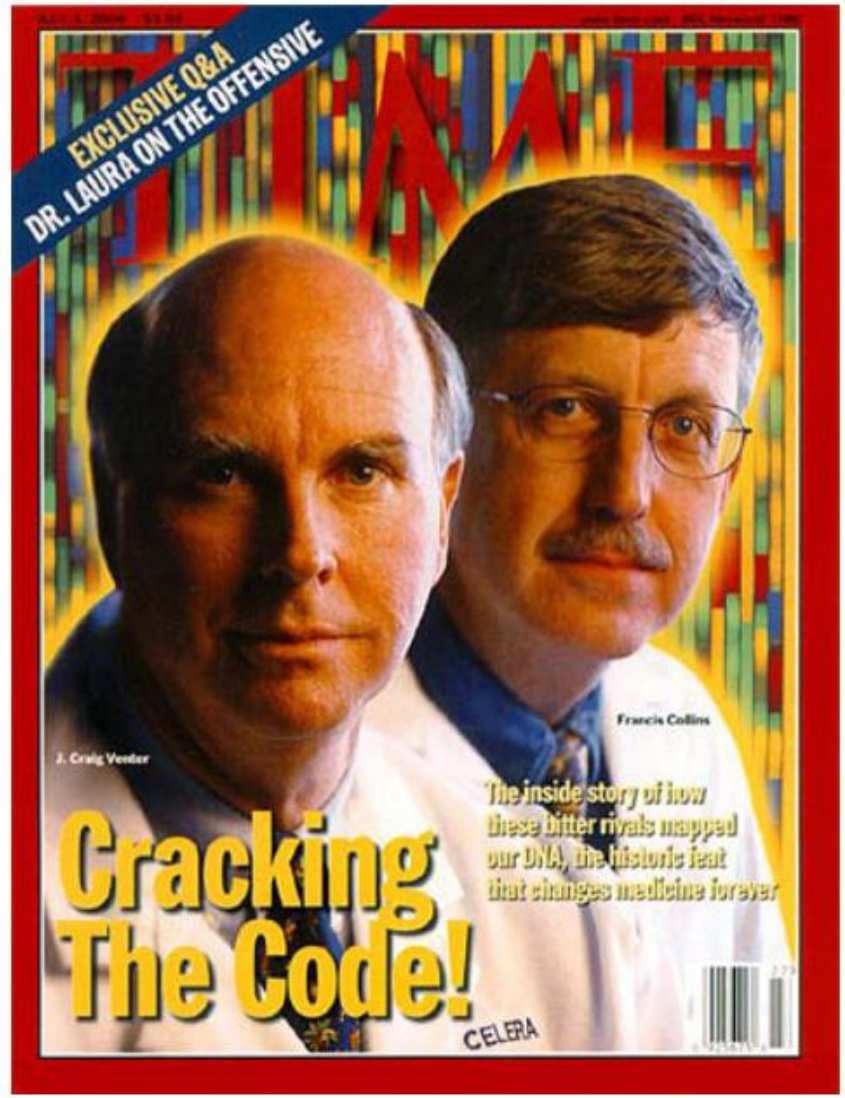
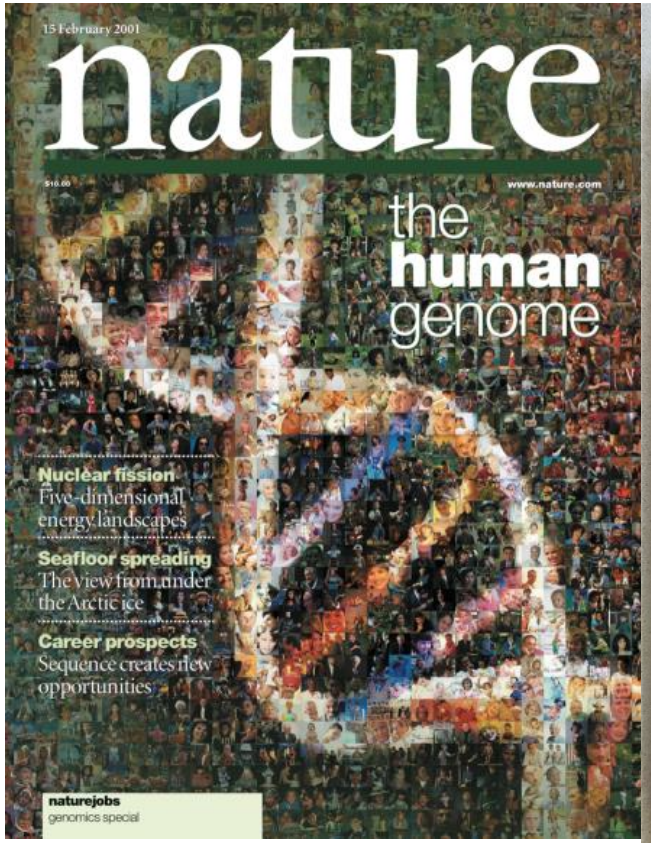


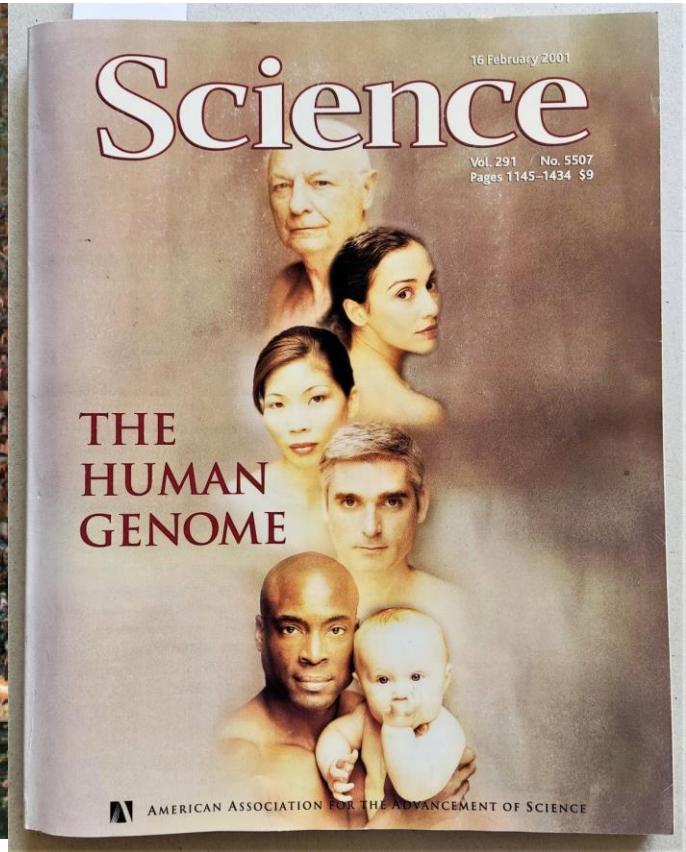
Figure 1: Time magazine cover showing Craig Venter head of Celera Genomics who led the private HGP and Francis Collins, head of NIH, who led the Public HGP.

June, 26th 2000: public announcement of the completion of the first draft of the human genome sequence

J. Craig Ventor, President Bill Clinton, and Francis Collins at the White House, June 26, 2000.

Robert Waterston, M.D., Ph.D., at the 2001 press conference announcing the publication describing the draft sequence of the human genome generated by the Human Genome Project (NHGRI Photo Archive).



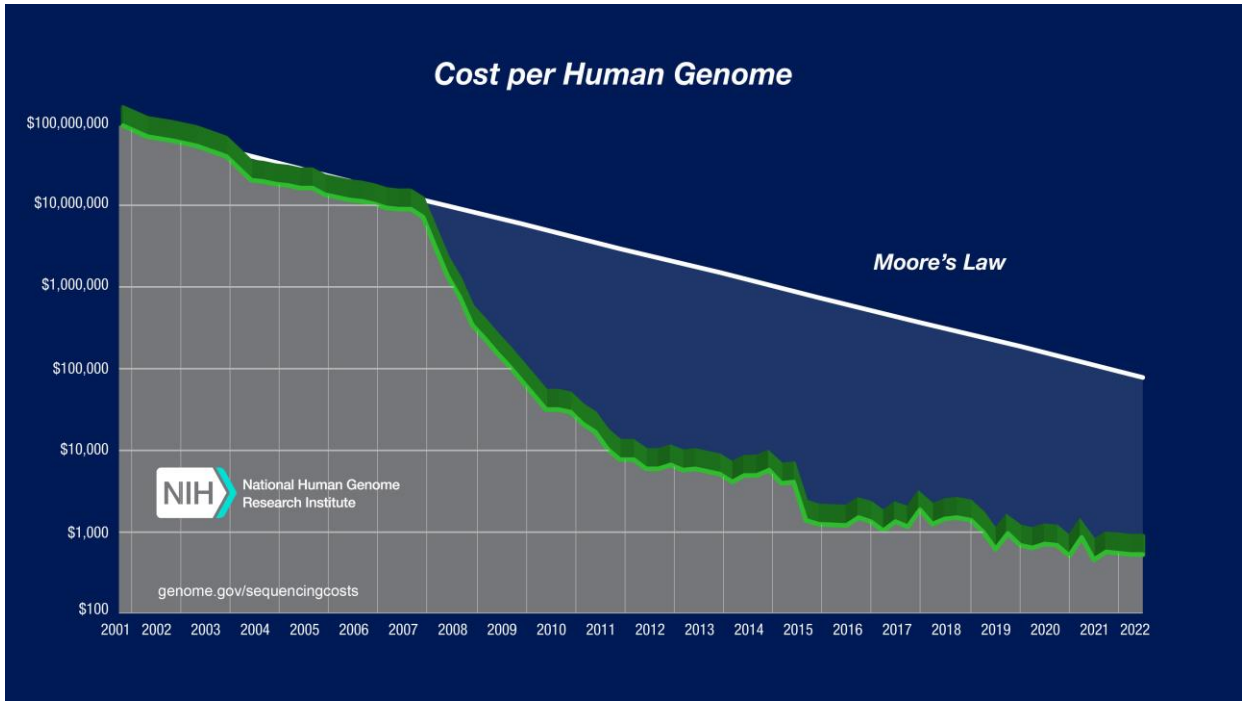


A summary of the main results

- «Total human genome size expected aroung 3.2 Gb»
- «Estimate of the euchromatic genome size of 2.9 Gb»
- «Variation in the distribution of several features (genes, transposable elements, GC content, recombination rate, etc…) within the genome landscape»
- «More than 1.4 million single nucleotide polymorphisms (SNPs) identified»
- «Number of protein-coding genes less than expected (about 30,000-40,000 against the 100,000 genes predicted at the beginning of the project)»
- «The total number of human protein-genes only twice as many as in worm or fly BUT more alternative splicing observed in humans»
- «Presence of vertebrate-specific DNA sequences involved in the synthesis of motif and domains typical in vertebrates»
- «The mutation rate twice as high in male as in female meiosis: most mutation occurs in males»
- «About 88% of the human genome represented in the draft»

Main challanges coming from these results

- Individual variation of the sequence: the need to obtain personalized genome sequences
- Reducing the costs of sequencing
- Develop technologies leading to long-read sequences, covering the gaps
- Understand the meaning of the so called «junk DNA»

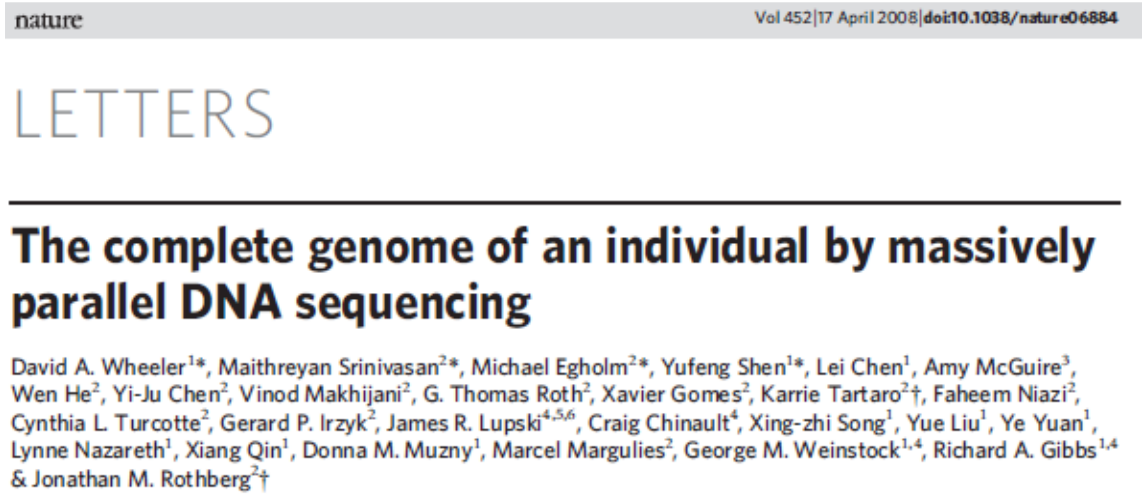


Dr. Eric Green, the director of the National Human Genome Research Institute, recalled that "the first genome cost us about a billion dollars … Now when we sequence a person's genome, it's less than $1000, so that's a million-fold reduction."

Using technology originally acquired in the US, the Chinese gene giant BGI Group says it will make genome sequencing cheaper than ever, breaking the $100 barrier for the first time.

A fall in the cost of DNA sequencing from $1,000,000,000 to $100 over 20 years would imply a compound rate of decline of 6.5 percent a month. (Adjusting for the time price puts the compound rate of decline at 7.13 percent per month.) Moore's law indicates that prices of computing decline at 2.85 percent a month. So, the cost of DNA sequencing per genome may amount to the fastest price decline in history.

- In April 2003, the consortium announced that it had generated an essentially complete human genome sequence, which was significantly improved from the draft sequence. Specifically, it accounted for 92% of the human genome and less than 400 gaps; it was also more accurate.



nature Vol 452|17 April 2008|doi:10.1038/nature06884

LETTERS

The complete genome of an individual by massively parallel DNA sequencing

David A. Wheeler[1]*, Maithreyan Srinivasan[2]*, Michael Egholm[2]*, Yufeng Shen[1]*, Lei Chen[1], Amy McGuire[3], Wen He[2], Yi-Ju Chen[2], Vinod Makhijani[2], G. Thomas Roth[2], Xavier Gomes[2], Karrie Tartaro[2]†, Faheem Niazi[2], Cynthia L. Turcotte[2], Gerard P. Irzyk[2], James R. Lupski[4,5,6], Craig Chinault[4], Xing-zhi Song[1], Yue Liu[1], Ye Yuan[1], Lynne Nazareth[1], Xiang Qin[1], Donna M. Muzny[1], Marcel Margulies[2], George M. Weinstock[1,4], Richard A. Gibbs[1,4] & Jonathan M. Rothberg[2]†

17 April 2008 –
James Watson's genome sequenced

H3 Africa (Mulderr et al., Pharmac. Pers. Med. 2018)

The Genome Asia 100K project (Wall et al., Nature 2019)

Indi Genomes (Jain et al., NAR 2021)

- On March 31, 2022, the Telomere-to-Telomere (T2T) consortium announced that had filled in the remaining gaps and produced the first truly complete human genome sequence.

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

SPECIAL SECTION COMPLETING THE HUMAN GENOME

RESEARCH ARTICLE

HUMAN GENOMICS

The complete sequence of a human genome

Nurk et al., Science 376, 44–53 (2022) 1 April 2022

Since its initial release in 2000, the human reference genome has covered only the euchromatic fraction of the genome, leaving important heterochromatic regions unfinished. Addressing the remaining 8% of the genome, the Telomere-to-Telomere (T2T) Consortium presents a complete 3.055 billion–base pair sequence of a human genome, T2T-CHM13, that includes gapless assemblies for all chromosomes except Y, corrects errors in the prior references, and introduces nearly 200 million base pairs of sequence containing 1956 gene predictions, 99 of which are predicted to be protein coding. The completed regions include all centromeric satellite arrays, recent segmental duplications, and the short arms of all five acrocentric chromosomes, unlocking these complex regions of the genome to variational and functional studies.

Workshop Summary

Workshop on the Comprehensive Extraction of Biological Information from Genomic Sequence

Bethesda, Md.

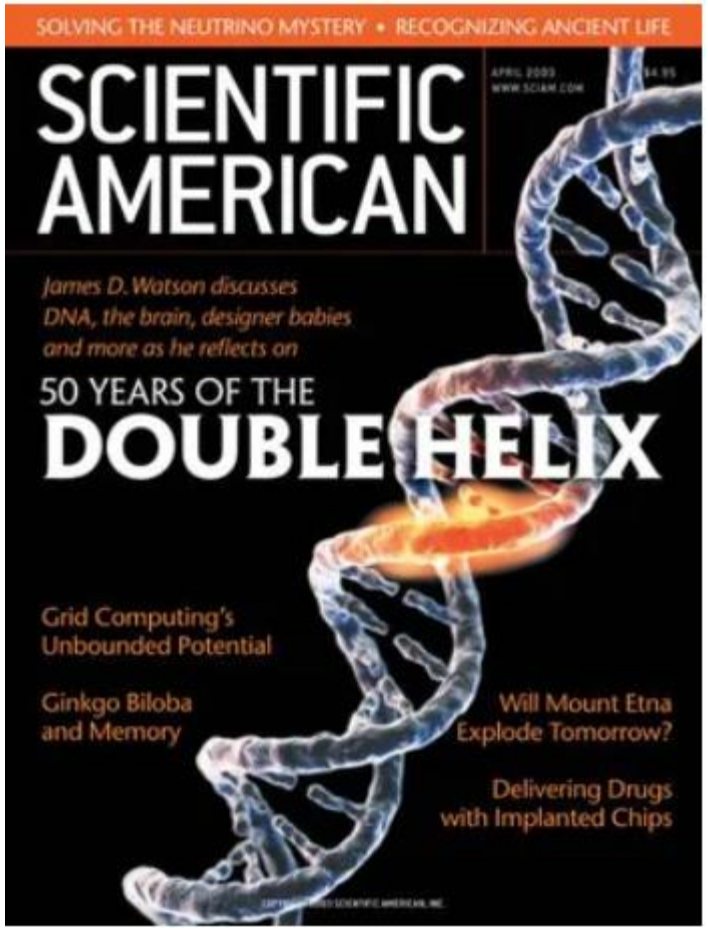
July 23 - 24, 2002

- *Initial Inventory of Functional Elements to Identify*
- *Initial List of Technologies/Approaches that Could be Utilized*
- *Process for Selection of Genomic Targets*
- *Criteria for Participation*
- *Organizational Issues*
- *Data Management*
- *Data Release*
- *Other Issues*

https://www.genome.gov/10005115/encode-workshop-summary

A Conversation with James D. Watson (Interview by John Rennie, EiC of Sc. American at that time)

APRIL 1, 2003



SA: "In a century, we went from rediscovering Mendel's laws and identifying chromosomes as agents of heredity to having the human genome largely worked out. Finding the double helix drops neatly in the middle of that span. How much, with respect to DNA, is left for us to do? Are there still great discoveries to be made, or is it all just filling in details?"

JW: "The major problem, I think, is chromatin [the dynamic complex of DNA and histone proteins that makes up chromosomes]. What determines whether a given piece of DNA along the chromosome is functioning, since it's covered with the histones? You can inherit something beyond the DNA sequence. That's where the real excitement of genetics is now. And it seems to be moving pretty fast. You don't really want to make a guess, but I'd guess that over these next 10 years, the field will be pretty played out. A lot of very good people are working on it. We have the tools. At some stage, the basic principles of genetics will be known in terms of gene functioning, and then we'll be able to apply that more to problems such as how the brain works."

Back to 2003

3 phases planned as the beginning of the ENCODE project:

1) Pilot study: approx. 30 Mb of the human genome (1%) divided into 44 discrete regions

2) Technical Skills Improvement Phase

3) Production phase (extension of the acquired technical capabilities to the remaining 99% of the genome)

A summary of the main results

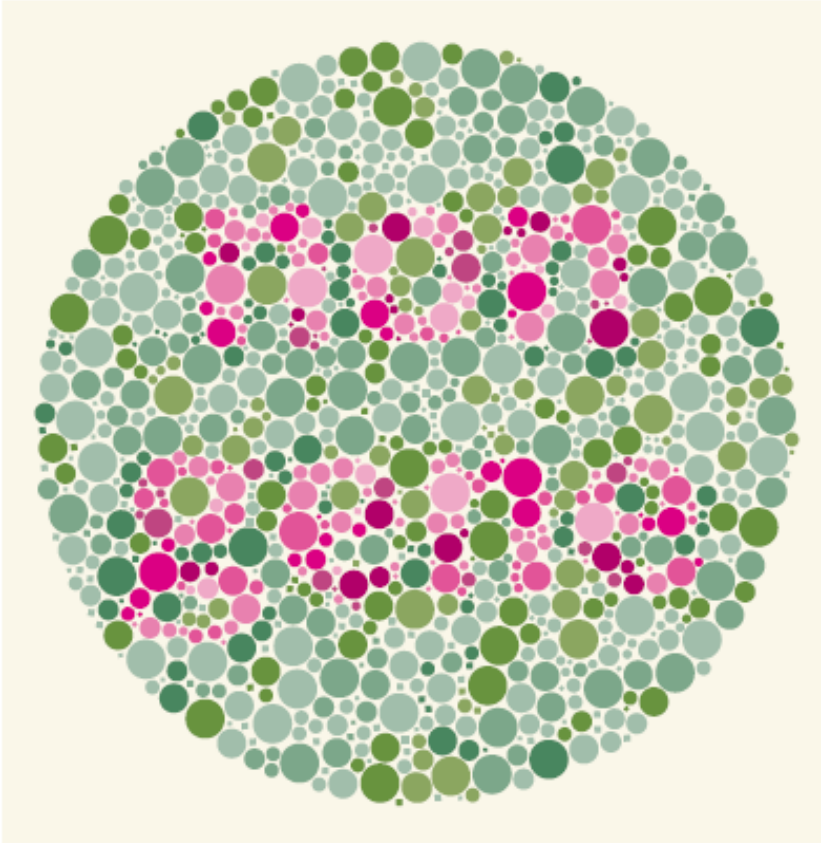
GENOMICS

Encyclopaedia of humble DNA

John M. Greally

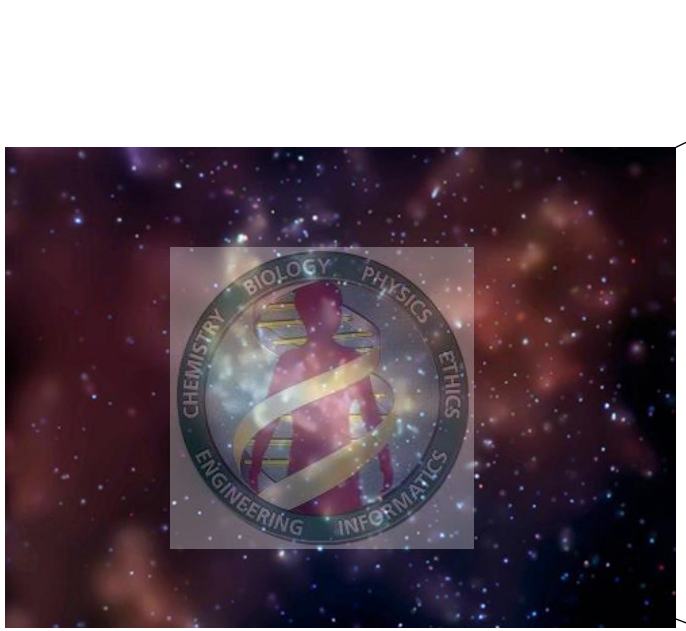
Researchers of the ENCODE consortium have analysed 1% of the human genome. Their findings bring us a step closer to understanding the role of the vast amount of obscure DNA that does not function as genes.

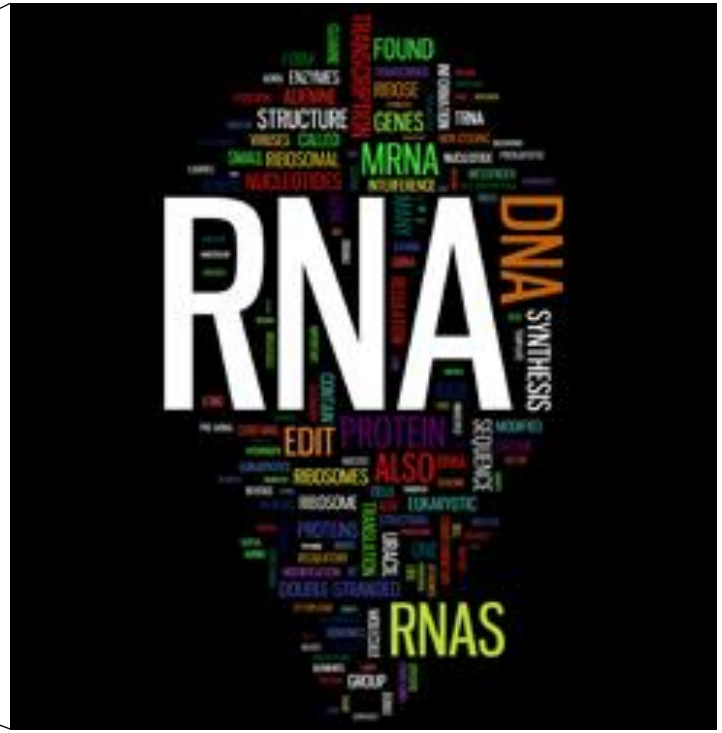
- Many more sections of the genome are transcribed into RNA than had previously been recognized
- Identified where within the chromatin certain histones are marked by chemical modifications, and identified positions at which transcription-regulatory proteins were binding to the DNA (as Watson said…the real issue is chromatin!)
- Evidence of regulatory functions for sequences at transcription start sites, as expected, but also at other sites in the DNA



Ishihara's test
(From: NATURE|Vol 447|14 June 2007)

The dark matter of the genome and non-coding RNA





Just 1.2% of the 3 billion bases of our genome is really «translated» into proteins!

Notwithstanding this, most of the remaining part of the genome (about 75%) is trancribed in RNA, even if not translated into proteins

Biophysicist Alexander Rich first mixed two differing strands of RNA in 1956 to discover the RNA double helix.

A NEW TWO STRANDED HELICAL STRUCTURE: POLYADENYLIC ACID AND POLYURIDYLIC ACID

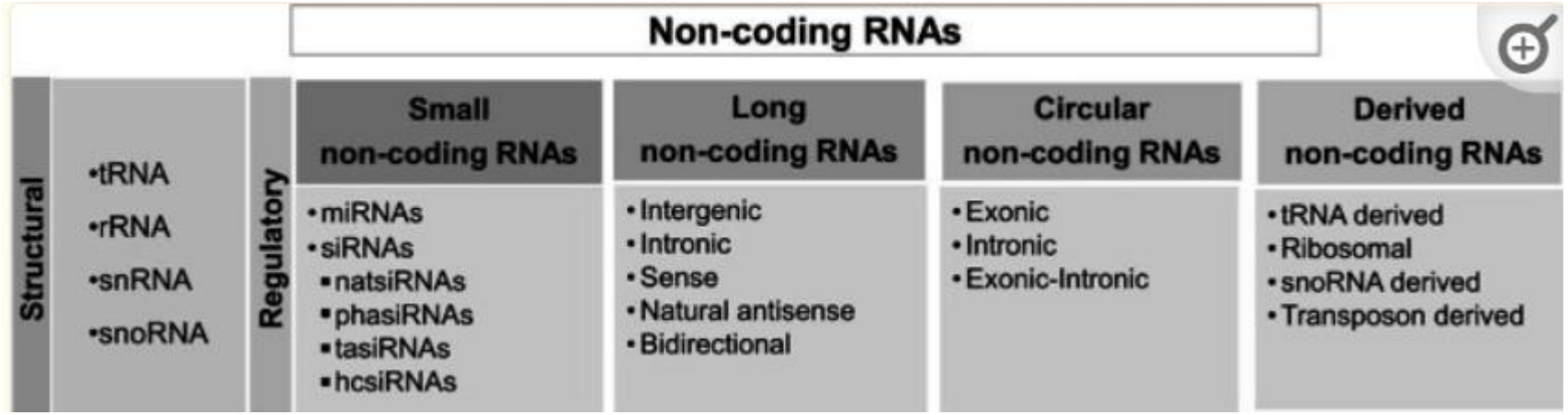
Alexander Rich and David R. Davies

Cite this: *J. Am. Chem. Soc.* 1956, 78, 14, 3548–3549
Publication Date: July 1, 1956
https://doi.org/10.1021/ja01595a086

These results show for the first time that it is possible for the ribonucleic acid (RNA) backbone to assume a configuration not unlike that found in DNA, using the same complementarity in the base pairs. This implies that there may exist a form of the RNA molecule similar to that of DNA and that this could be the form in which RNA carries out its implied molecular duplication in the plant and smaller animal viruses.

Finally, we would like to point out that this method for forming a two-stranded helical molecule by simply mixing two substances can be used for a variety of studies directed toward an understanding of the formation of helical molecules utilizing specific interactions.

The era of transcriptomics



From: Funct Integr Genomics. 2021; 21(3-4): 313–330

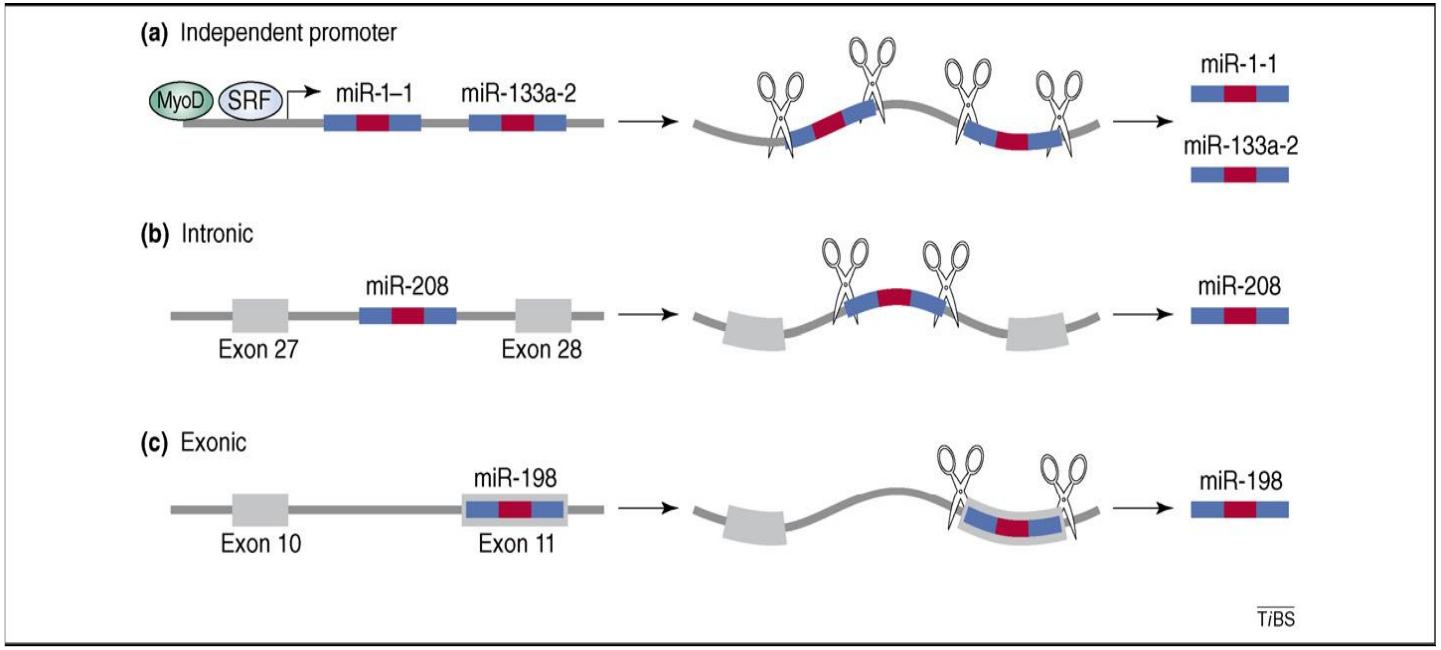
Micro RNAs

The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14

R C Lee [1], R L Feinbaum, V Ambros

Affiliations + expand

PMID: 8252621 DOI: 10.1016/0092-8674(93)90529-y



The Nobel Prize in Physiology or Medicine 2006

Photo: L. Cicero

Andrew Z. Fire

Prize share: 1/2

Photo: J. Mottern

Craig C. Mello

Prize share: 1/2

Letter | Published: 19 February 1998

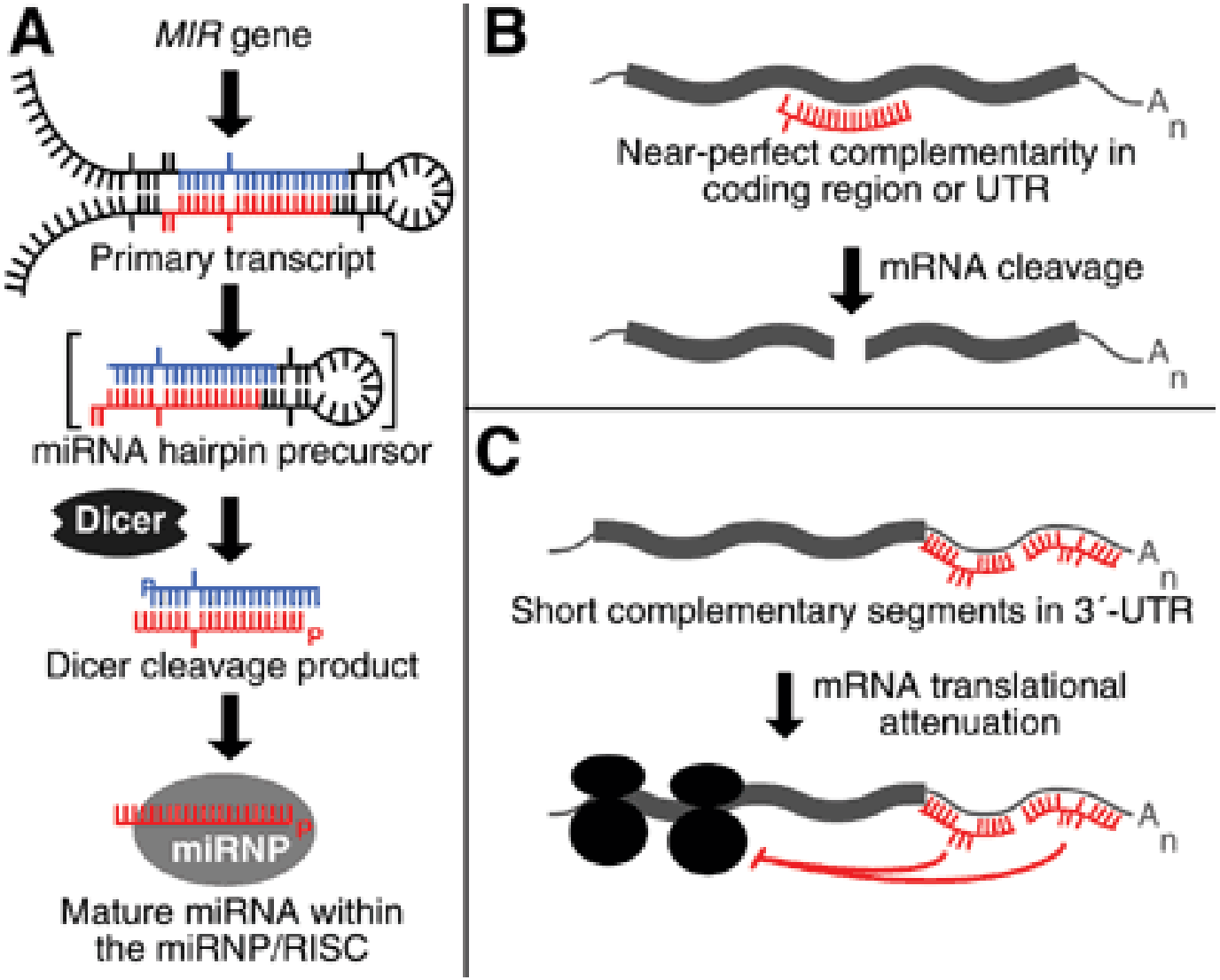
Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*

Andrew Fire ✉, SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver & Craig C. Mello

Nature **391**, 806–811 (1998) | Cite this article

« Whatever their target, the mechanisms underlying RNA interference probably exist for a biological purpose. Genetic interference by dsRNA could be used by the organism for physiological gene silencing. Likewise, the ability of dsRNA to work at a distance from the site of injection, and particularly to move into both germline and muscle cells, suggests that there is an effective RNA-transport mechanism in *C. elegans*. » □

miRNAs are negative regulators of gene expression



- **Each miRNA can target up to 200 different mRNAs (quite 1/3 of the whole transcriptome is regulated my miRNAs)**
- **Each mRNA can be targeted by several miRNAs (sinergistic control)**

Courtesy of Bonnie Bartel and David P. Bartel Reprinted from *Plant Physiol,* 132:1-9, June 2003.

miRNAs in cancer (e.g.: CRC, GBM, neuroblastoma, uveal melanoma) – may act as oncogenes or tumor suppressors

circSMARCA5 Is an Upstream Regulator of the Expression of miR-126-3p, miR-515-5p, and Their mRNA Targets, *Insulin-like Growth Factor Binding Protein 2* (*IGFBP2*) and *NRAS Proto-Oncogene, GTPase* (*NRAS*) in Glioblastoma.

Merulla AE, Stella M, Barbagallo C, Battaglia R, Caponnetto A, Broggi G, Altieri R, Certo F, Caltabiano R, Ragusa M, Barbagallo GMV, Di Pietro C, Purrello M, Barbagallo D.

Int J Mol Sci. 2022 Nov 8;23(22):13676. doi: 10.3390/ijms232213676.

Dysregulated miR-671-5p / CDR1-AS / CDR1 / VSNL1 axis is involved in glioblastoma multiforme.

Barbagallo D, Condorelli A, Ragusa M, Salito L, Sammito M, Banelli B, Caltabiano R, Barbagallo G, Zappalà A, Battaglia R, Cirnigliaro M, Lanzafame S, Vasquez E, Parenti R, Cicirata F, Di Pietro C, Romani M, Purrello M.

Oncotarget. 2016 Jan 26;7(4):4746-59. doi: 10.18632/oncotarget.6621.

miRNA profiling in **vitreous** humor, vitreal exosomes and serum from uveal melanoma patients: Pathological and diagnostic implications.

Ragusa M, Barbagallo C, Statello L, Caltabiano R, Russo A, Puzzo L, Avitabile T, Longo A, Toro MD, Barbagallo D, Valadi H, Di Pietro C, Purrello M, Reibaldi M.

Cancer Biol Ther. 2015;16(9):1387-96. doi: 10.1080/15384047.2015.1046021. Epub 2015 May 7.

Specific alterations of the **microRNA transcriptome** and global network structure in colorectal cancer after treatment with MAPK/ERK inhibitors.

Ragusa M, Statello L, Maugeri M, Majorana A, Barbagallo D, Salito L, Sammito M, Santonocito M, Angelica R, Cavallaro A, Scalia M, Caltabiano R, Privitera G, Biondi A, Di Vita M, Cappellani A, Vasquez E, Lanzafame S, Tendi E, Celeste S, Di Pietro C, Basile F, Purrello M.

J Mol Med (Berl). 2012 Dec;90(12):1421-38. doi: 10.1007/s00109-012-0918-8. Epub 2012 Jun 4.

miRNAs may regulate woman fertility at different levels (e.g.: oocytes, follicular fluid, granulosa cells)

Down-regulation of long non-coding RNAs in reproductive aging and analysis of the lncRNA-**miRNA**-mRNA networks in human cumulus cells.

Caponnetto A, Battaglia R, Ferrara C, Vento ME, Borzì P, Paradiso M, Scollo P, Purrello M, Longobardi S, D'Hooghe T, Valerio D, **Di Pietro C**; Italian Society of Embryology, Reproduction, Research (SIERR).

J Assist Reprod Genet. 2022 Apr;39(4):919-931. doi: 10.1007/s10815-022-02446-8. Epub 2022 Mar 5.

Ovarian aging increases small extracellular vesicle $CD81^+$ release in human **follicular fluid** and influences miRNA profiles.

Battaglia R, Musumeci P, Ragusa M, Barbagallo D, Scalia M, Zimbone M, Lo Faro JM, Borzì P, Scollo P, Purrello M, Vento EM, **Di Pietro C.**

Aging (Albany NY). 2020 Jun 17;12(12):12324-12341. doi: 10.18632/aging.103441. Epub 2020 Jun 17.

Non-coding RNAs in the Ovarian Follicle.

Battaglia R, Vento ME, Borzì P, Ragusa M, Barbagallo D, Arena D, Purrello M, **Di Pietro C.**

Front Genet. 2017 May 12;8:57. doi: 10.3389/fgene.2017.00057. eCollection 2017.

MicroRNAs Are Stored in Human MII **Oocyte** and Their Expression Profile Changes in Reproductive Aging.

Battaglia R, Vento ME, Ragusa M, Barbagallo D, La Ferlita A, Di Emidio G, Borzí P, Artini PG, Scollo P, Tatone C, Purrello M, **Di Pietro C.**

Biol Reprod. 2016 Dec;95(6):131. doi: 10.1095/biolreprod.116.142711. Epub 2016 Nov 9.

miRNAs in SARS-CoV-2

Competing endogenous **RNA** network mediated by circ_3205 in SARS-CoV-2 infected cells.

Barbagallo D, Palermo CI, Barbagallo C, Battaglia R, Caponnetto A, Spina V, Ragusa M, Di Pietro C, Scalia G, Purrello M.

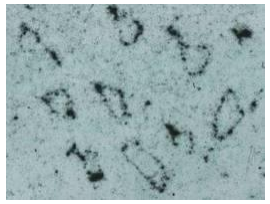
Cell Mol Life Sci. 2022 Jan 17;79(2):75. doi: 10.1007/s00018-021-04119-8.

MicroRNA-Mediated Regulation of the Virus Cycle and Pathogenesis in the SARS-CoV-2 Disease.

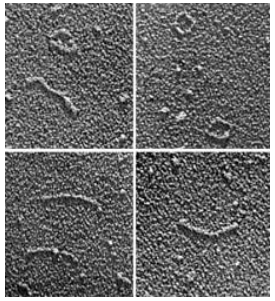
Battaglia R, Alonzo R, Pennisi C, Caponnetto A, Ferrara C, Stella M, Barbagallo C, Barbagallo D, Ragusa M, Purrello M, **Di Pietro C.**

Int J Mol Sci. 2021 Dec 7;22(24):13192. doi: 10.3390/ijms222413192.

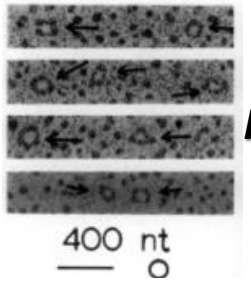
Circular RNAs



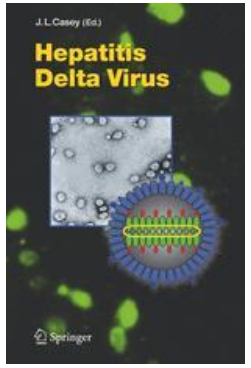
Viroids (1976)



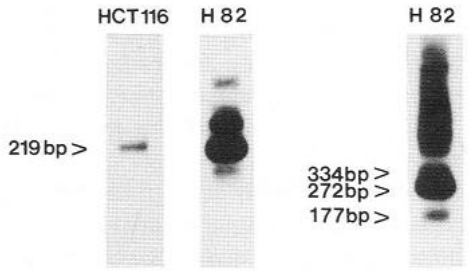
Cytoplasmic eucaryotic circular RNAs (1979)



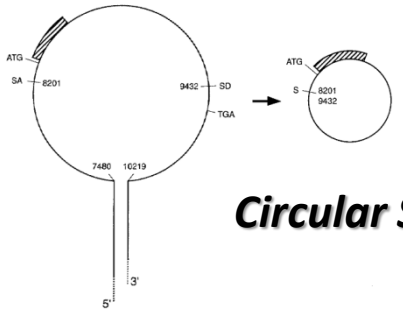
Circular RNAs from the Intervening sequence of rRNA (1981)



Hepatitis δ virus (1986)



«Scrambled exons» in human DCC gene (1991)



Circular Sry transcripts (1993)

RNA. 2013 Feb;19(2):141-57. doi: 10.1261/rna.035667.112. Epub 2012 Dec 18.

Circular RNAs are abundant, conserved, and associated with ALU repeats

William R Jeck [1], Jessica A Sorrentino, Kai Wang, Michael K Slevin, Christin E Burd, Jinze Liu, William F Marzluff, Norman E Sharpless

Nature. 2013 Mar 21;495(7441):384-8. doi: 10.1038/nature11993. Epub 2013 Feb 27.

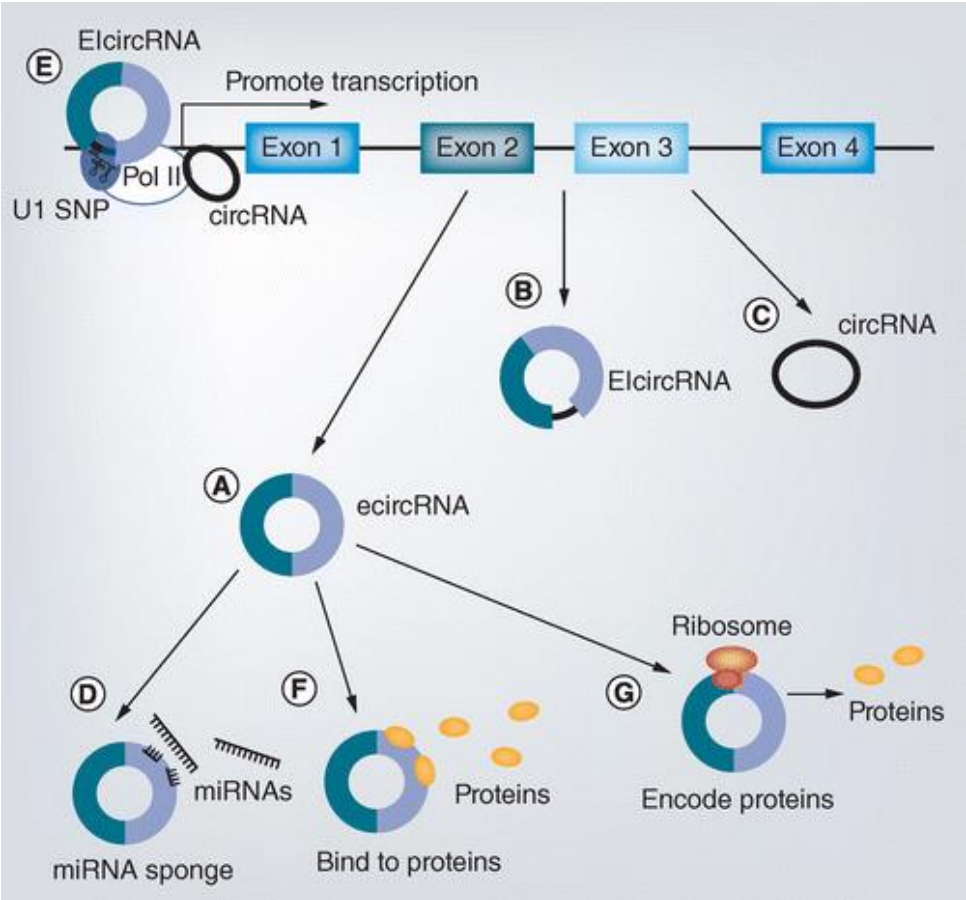
Natural RNA circles function as efficient microRNA sponges

Thomas B Hansen [1], Trine I Jensen, Bettina H Clausen, Jesper B Bramsen, Bente Finsen, Christian K Damgaard, Jørgen Kjems

Nature. 2013 Mar 21;495(7441):333-8. doi: 10.1038/nature11928. Epub 2013 Feb 27.

Circular RNAs are a large class of animal RNAs with regulatory potency

Sebastian Memczak [1], Marvin Jens, Antigoni Elefsinioti, Francesca Torti, Janna Krueger, Agnieszka Rybak, Luisa Maier, Sebastian D Mackowiak, Lea H Gregersen, Mathias Munschauer, Alexander Loewer, Ulrike Ziebold, Markus Landthaler, Christine Kocks, Ferdinand le Noble, Nikolaus Rajewsky



CircSMARCA5 Regulates **VEGFA** mRNA Splicing and Angiogenesis in Glioblastoma Multiforme Through the Binding of SRSF1.

Barbagallo D, Caponnetto A, Brex D, Mirabella F, Barbagallo C, Lauretta G, Morrone A, Certo F, Broggi G, Caltabiano R, Barbagallo GM, Spina-Purrello V, Ragusa M, Di Pietro C, Hansen TB, Purrello M.

Cancers (Basel). 2019 Feb 7;11(2):194. doi: 10.3390/cancers11020194.

Noncoding RNA therapeutics — challenges and potential solutions

Melanie Winkle, Sherien M. El-Daly, Muller Fabbri & George A. Calin

Nature Reviews Drug Discovery **20**, 629–651 (2021) | Cite this article

nature biomedical engineering

Article

https://doi.org/10.1038/s41551-023-01081-7

Profiling of repetitive RNA sequences in the blood plasma of patients with cancer

Received: 22 August 2022

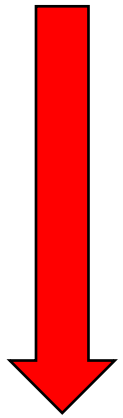
Accepted: 26 July 2023

Published online: 31 August 2023

Roman E. Reggiardo[1], Sreelakshmi Velandi Maroli[2], Vikas Peddu[1], Andrew E. Davidson[1], Alexander Hill[1], Erin LaMontagne[1], Yassmin Al Aaraj[3], Miten Jain[1,8,9], Stephen Y. Chan[3] & Daniel H. Kim[1,4,5,6,7]

Concluding remarks

Regularoty ncRNAs  Epigenetic regulators – What is their relationship with the genome?



Back to the (personalized) genomes

Grazie a tutti per l’attenzione!

“We shall not cease from exploration. And the end of all our exploring will be to arrive where we started, and know the place for the first time”. DT. S. Eliot