Il punto di svolta



The turning point

Il Progetto Genoma Umano



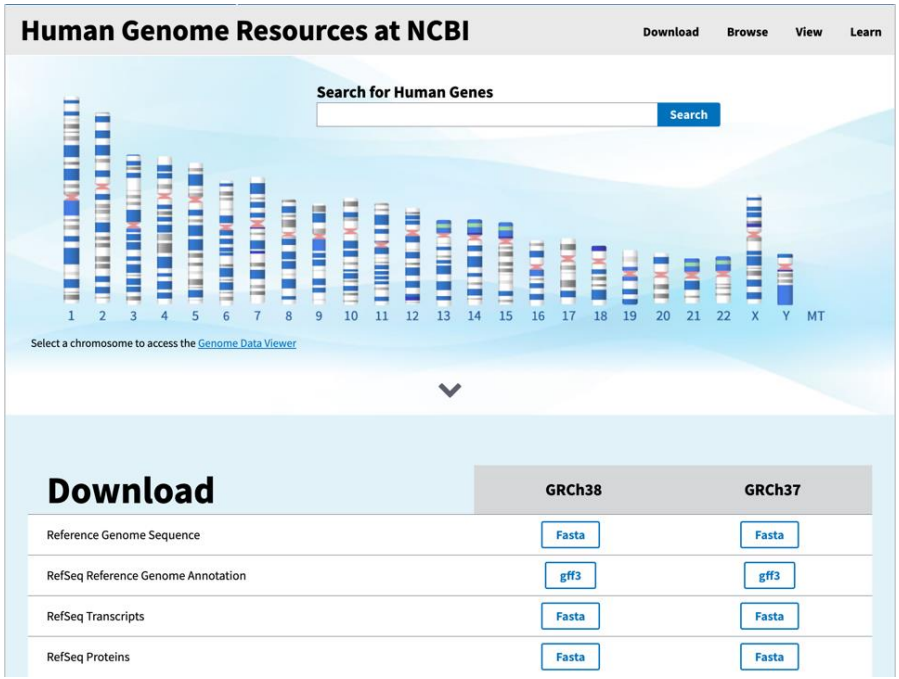
- ✓ Iniziato negli Stati Uniti nel **1990**, finanziato per tre miliardi di dollari dal Department of Energy e dal National Institutes of Health degli USA.
- ✓ Il **Consorzio Internazionale** è formato da ricercatori americani, inglesi, francesi, tedeschi, cinesi, indiani e giapponesi.
- ✓ La durata prevista inizialmente è di **quindici anni**.
- ✓ Il **26 giugno 2000** il presidente degli Stati Uniti, Bill Clinton, insieme a Francis Collins e Craig Venter, **annunciano il completamento del primo *draft* del genoma umano**, che contiene tra I 30.000 e i 40.000 geni e 3,1 miliardi di basi circa.



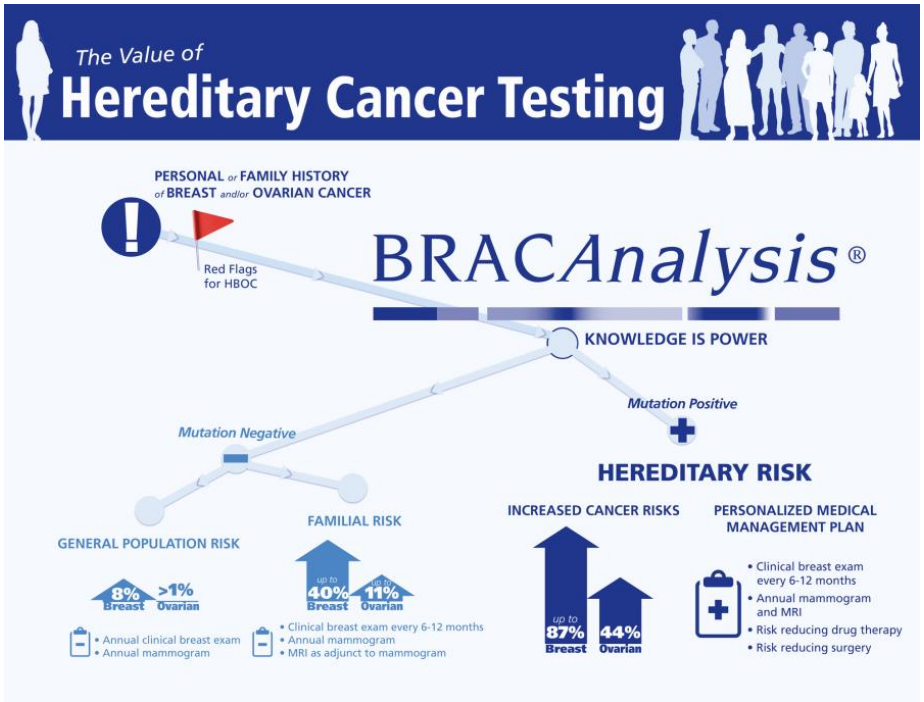
Il Progetto Genoma Umano

- ✓ Nel Febbraio del 2001 vengono pubblicati i risultati delle analisi del genoma umano, **sulle riviste Nature e Science**. Il lavoro del consorzio internazionale esce su Nature, quello della Celera Genomics su Science, 1 giorno dopo.
- ✓ Sia i dati del Consorzio Internazionale che quelli della Celera Genomics erano stati resi disponibili già durante il progetto.
- ✓ I dati vengono resi pubblici a tutti **senza alcuna restrizione** attraverso portali web o tramite il database GeneBank.



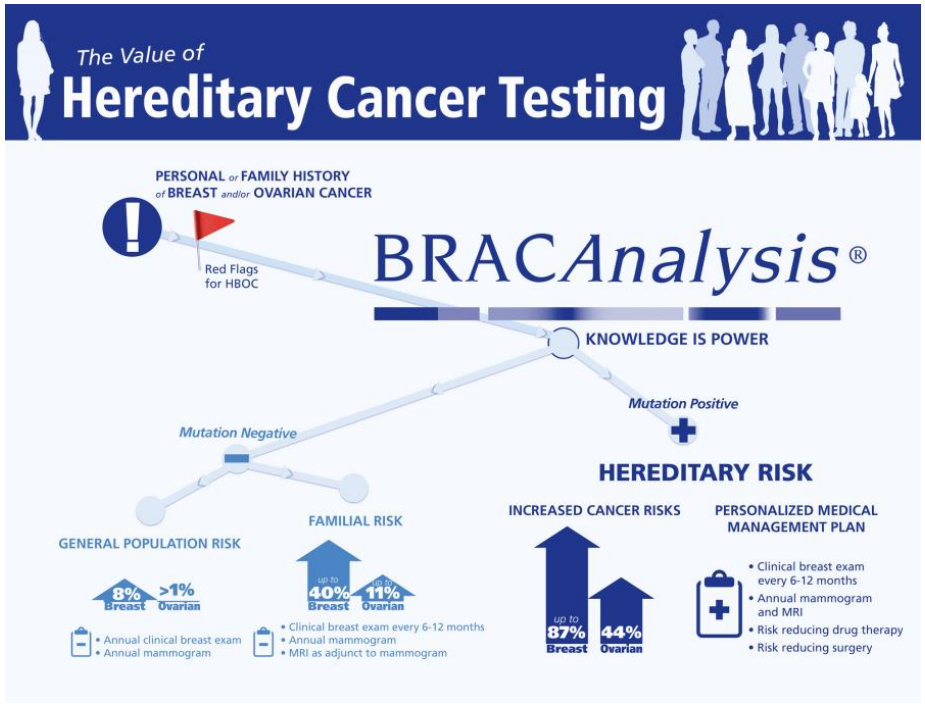


IL caso della Myriad Genetics Inc.



- La Myriad è una company americana che sviluppa tool di **diagnostica molecolare**
- Fu fondata nel 1991 da Mark Skolnick che grazie ad un nuovo sistema di mappatura genica identifica e clona il gene **BRCA1**, noto per il suo ruolo di oncosoppressore in tumore al seno e alle ovaie. Se mutato le pazienti hanno un rischio maggiore di sviluppare forme tumorali.
- **Nel 1994 la Myriad brevetta il gene BRCA1** e impedisce a qualsiasi altra company di sviluppare test diagnostici per l'identificazione di mutazioni.
- La Myriad avrà per molti anni il monopolio dei test sul gene BRCA, lo venderà a 4.000$ raggiungendo un profitto di 57 milioni $ fino a giugno 2012, data in cui il brevetto viene dichiarato non valido.

IL caso della Myriad Genetics Inc.



Molti ricercatori insieme ad un gruppo di pazienti fanno causa alla Myriad.

Nell'Ottobre del 2012 la Corte Suprema dichiara:

« *laws of nature, natural phenomena, and abstract ideas are not patentable… A naturally occurring DNA segment is a product of nature and not patent eligible merely because it has been isolated.*»

Immediatamente dopo la sentenza che invalida i brevetti della Myriad, altre società inizieranno a offrire test sui geni BRCA1/2 a basso costo.

Open genome:
il contributo di Sulston e Waterson

- ✓ A metà degli anni '80, **John Sulston e Bob Waterston** cominciano a collaborare e nel 1998 completano il genoma dell'organismo modello *C. elegans,* un progetto satellite del progetto genoma umano

- ✓ Entrambi grandi **sostenitori della scienza come bene pubblico**, rendono disponibili i dati e i risultati dei loro studi, prima della pubblicazione.

- ✓ Sono i promotori del primo **Bermuda Meeting nel 1996,** in cui per la prima volta si discutono le **politiche di condivisione dei dati del progetto Genoma**



Open genome: Bermuda Principles

Nel febbraio 1996, i leader del progetto Genoma Umano si incontrano alle Bermuda all'International Strategy Meeting on Human Genome Sequencing.

Principio 1: tutti i dati devono essere resi disponibili gratuitamente alla comunità scientifica;

Principio 2: i dati devono essere rilasciati entro **24 ore** dalla loro generazione.

Obiettivo: incoraggiare la ricerca, promuovere un rapido sviluppo in campo biomedico e massimizzare i benefici del Progetto Genoma Umano per la società.



Questi principi rivoluzioneranno il modo di fare scienza in ambito biologico, e diventano la base fondante di tutti i grandi progetti genomici internazionali successivi.

Open genome:
il contributo di Sulston e Waterson

John Edward Sulston e Robert Hugh Waterston ricevono **nel 2000 la medaglia della Genetics Society of America** per i loro eccezionali contributi sul sequenziamento del genoma, per la loro integrità e per la loro notevole visione nell'ambito delle scienze della vita.



John Edward Sulston **nel 2002 vince il Nobel per la medicina** per gli studi compiuti sulla regolazione genica cellulare e sull'apoptosi.



Gli anni successivi

HapMap

- Comincia nel 2003
- Catalogare la variabilità genetica umana
- Consorzio Internazionale di 7 nazioni
- Approccio mirato
- **Dati liberamente accessibili**

The Cancer Genome Atlas

- Comincia nel 2005
- Caratterizzare da un punto di vista molecolare diversi tipi di tumori
- Progetto Americano
- **Dati liberamente accessibili**

1000 Genomes

- Comincia nel 2008
- Catalogare la variabilità genetica umana
- Consorzio Internazionale 8 Nazioni
- Sequenziamento dell'intero genoma di 1000 individui da diversi gruppi etnici
- **Dati liberamente accessibili**

Il progetto Genoma Umano è stata la pietra fondante su cui si sono sviluppate tutte le tecnologie omiche successive:

- ✓ Genomica
- ✓ Trascrittomica
- ✓ Epigenomia

Spirito collaborativo e per la **politica di condivisione**



Gli anni successive

International Cancer Genome Consortium (ICGC)

- Iniziato nel 2007
- Ombrello che racchiude tutti i progetti genomici per evitare ridondanze e promuovere degli standard
- Consorzio Internazionale
- **Dati liberamente accessibili**

Human microbiome

- Comincia nel 2007
- Catalogare le comunità microbiche umane e il loro ruolo nelle malattie
- Finanziato NIH
- **Dati liberamente accessibili**

100.000 genomes

- Annunciato nel 2012
- Far entrare la genomica nella pratica clinica
- Progetto UK
- Sequenziamento genomico
- Dati liberamente accessibili **previa approvazione**

Human Cell Atlas

- Comincia nel 2016
- Caratterizzare tutte le cellule del corpo umano utilizzando tecniche ad altissima risoluzione
- Consorzio Internazionale
- **Dati liberamente accessibili**





Minimum information about a microarray experiment (MIAME)—toward standards for microarray data

Alvis Brazma[1], Pascal Hingamp[2], John Quackenbush[3], Gavin Sherlock[4], Paul Spellman[5], Chris Stoeckert[6], John Aach[7], Wilhelm Ansorge[8], Catherine A. Ball[4], Helen C. Causton[9], Terry Gaasterland[10], Patrick Glenisson[11], Frank C.P. Holstege[12], Irene F. Kim[4], Victor Markowitz[13], John C. Matese[4], Helen Parkinson[1], Alan Robinson[1], Ugis Sarkans[1], Steffen Schulze-Kremer[14], Jason Stewart[15], Ronald Taylor[16], Jaak Vilo[1] & Martin Vingron[17]

Microarray analysis has become a widely used tool for the generation of gene expression data on a genomic scale. Although many significant results have been derived from microarray studies, one limitation has been the lack of standards for presenting and exchanging such data. Here we present a proposal, the Minimum Information About a Microarray Experiment (MIAME), that describes the minimum information required to ensure that microarray data can be easily interpreted and that results derived from its analysis can be independently verified. The ultimate goal of this work is to establish a standard for recording and reporting microarray-based gene expression data, which will in turn facilitate the establishment of databases and public repositories and enable the development of data analysis tools. With respect to MIAME, we concentrate on defining the content and structure of the necessary information rather than the technical format for capturing it.

BIOINFORMATICS

Editorial

AN OPEN LETTER TO THE SCIENTIFIC JOURNALS

One of the underlying principles of scientific publication in peer-reviewed journals has been the requirement that the authors make available the data and materials necessary for a reader to reproduce the experiment or analysis and to determine whether the data support the conclusions of the authors. In many instances, such as DNA sequence or protein structure data, this has evolved into the requirement that the data underlying each published report be deposited in an appropriate international database. For microarray experiments, simply defining the appropriate data has been a challenge, because the large quantity of data generated in each experiment and the typical complexity of the ancillary information needed to interpret the results are unlike anything that has yet faced the biological research community. Databases to hold microarray data and the tools to annotate them properly are under development. As an interim solution, this guide describes the types of data that are necessary to reproduce and interpret a microarray experiment. It should go without stating that this information is only of value as long as data are available, and that every effort should be made to provide stable access to published data until such time as it is available from a public database.

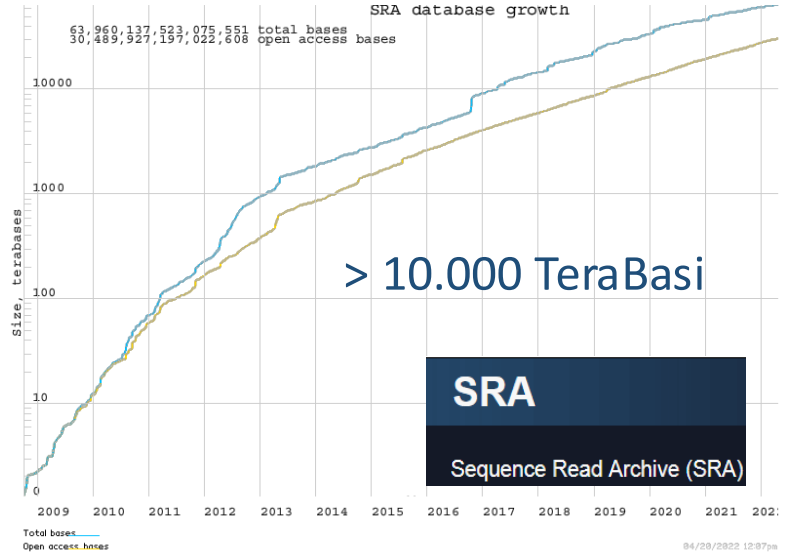
La potenza è nulla senza il controllo

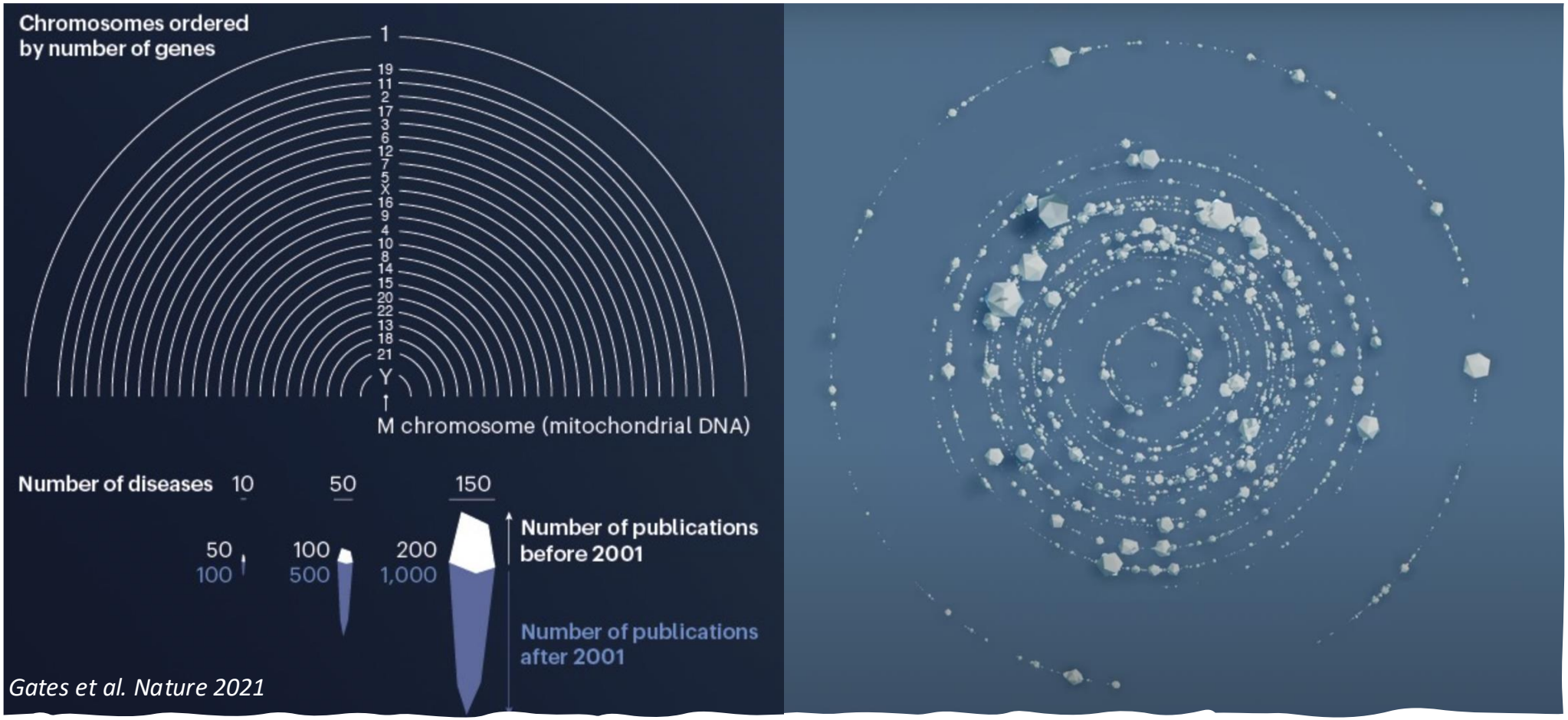
- Avere la disponibilità del dato senza avere le informazioni necessarie su come interpretarlo è rischioso e inutile.
- Nascono gruppi di lavoro per definire gli **standard e le linee guida** per sottomettere i dati (trascrittomica) in modo che sia tracciato l'intero esperimento, dal campione al dato finale.
- Le linee guida vengono recepite **dai database di riferimento mondiale.**
- Le principali riviste scientifiche appoggiano la politica **open data** sottoscrivendo un accordo per cui ogni lavoro scientifico deve aver i dati sottomessi in database pubblici prima della fase di referaggio.

La dimensione dei Database aumenta

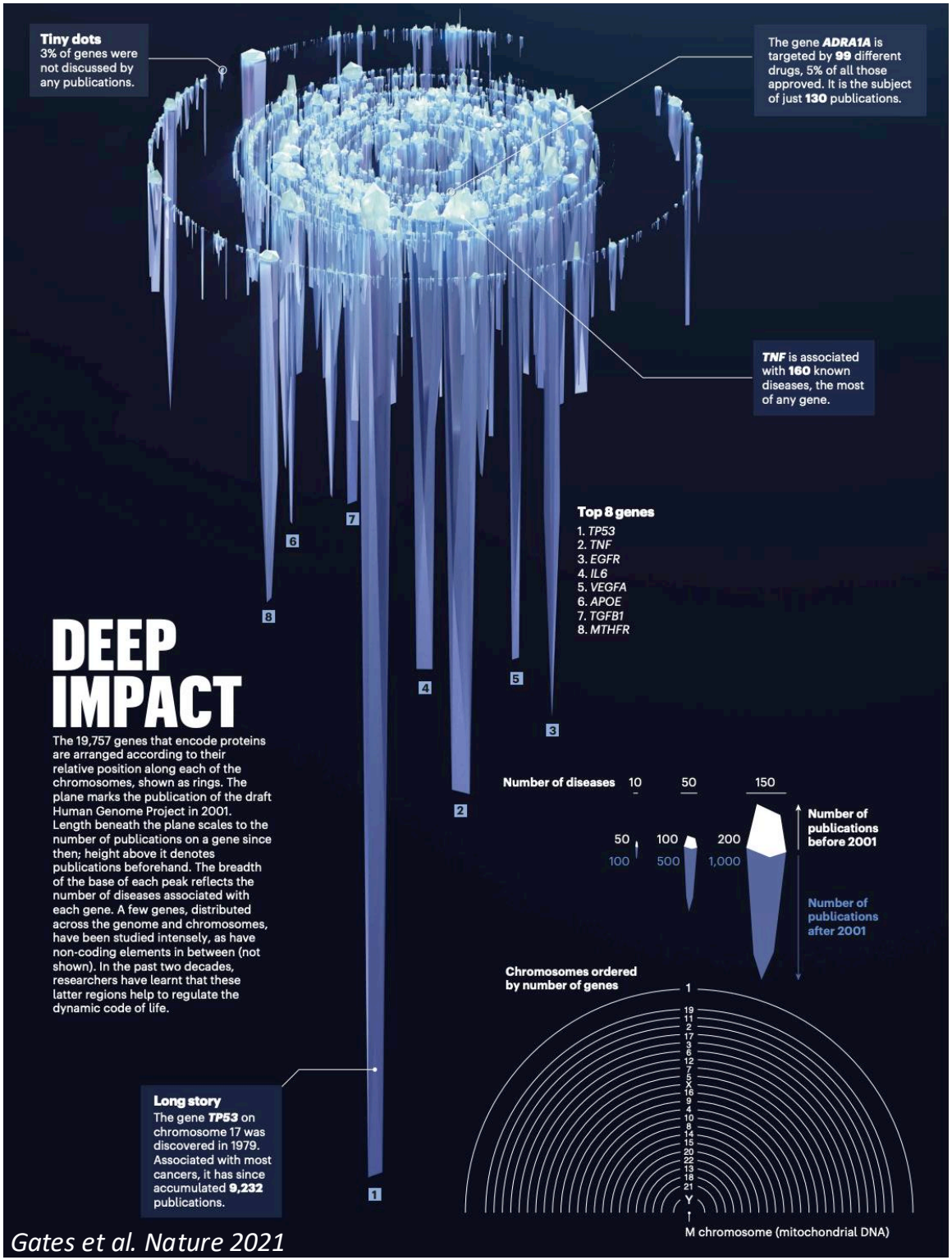
- Il numero di pubblicazioni scientifiche basate su dati omici cresce enormemente
- Questa crescita si riflette sulle dimensioni dei database di riferimento mondiale.







Impatto dei Progetto Genoma



Gates et al. Nature 2021

Comment

A wealth of discovery built on the Human Genome Project – by the numbers

Alexander J. Gates, Deisy Morselli Gysi, Manolis Kellis & Albert-László Barabási

Il Progetto genoma ha permesso lo sviluppo di nuove tecnologie, e lo studio di nuovi geni,

ma la maggior parte degli studi riguardano solo l'1% dei geni.

Il maggior contributo è stato diretto allo studio di geni non codificanti e alle regioni regolative, coinvolte nei meccanismi di regolazione genica

La «chiusura» del Genoma

- La prima bozza conteneva il 92% dell'informazione.
- 150 Mb (8%) del genoma non erano state risolte, parti del genoma distribuite in tutti i cromosomi, specialmente in regioni vicine ai centromeri e ai telomeri, regioni essenziali per moltissimi processi fondamentali della cellula.
- Il **1° Aprile 2022** il Consorzio Internazionale Telomere2Telomere (T2T) ha pubblicato su Science la "chiusura" del Genoma Umano grazie alle nuove tecnologie in grado di fare sequenze molto lunghe.



Data and materials availability: The T2T-CHM13 and T2T-HG002-ChrX assemblies generated by this study are archived under National Center for Biotechnology Information (NCBI) GenBank accession numbers GCA_009914755 and CP086568, respectively. CHM13hTERT cells were obtained for research use through a material transfer agreement with U. Surti and the University of Pittsburgh. The raw sequencing data were described in prior studies and are summarized in table S1. For convenience, links to the sequence data and genome browsers are also available from https://github.com/marbl/CHM13. Supplementary data for fig. S39 and the string graph construction code are archived at Zenodo (*61*) and also https://github.com/snurk/sg_sandbox.

Una storia lunga 22 anni