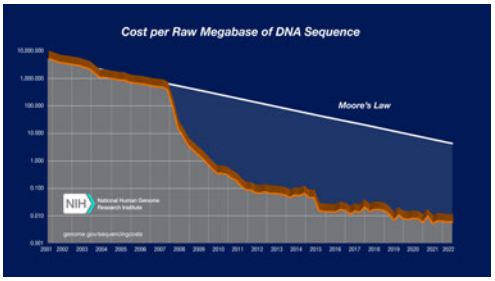
DNA Sequencing Costs: Data

For many years, the National Human Genome Research Institute (NHGRI) has tracked the costs associated with DNA sequencing performed at the sequencing centers funded by the Institute. This information has served as an important benchmark for assessing improvements in DNA sequencing technologies and for establishing the DNA sequencing capacity of the NHGRI Genome Sequencing Program. Here, NHGRI provides an analysis of these data, which gives one view of the remarkable improvements in DNA sequencing technologies and data-production pipelines in recent years.

Overview

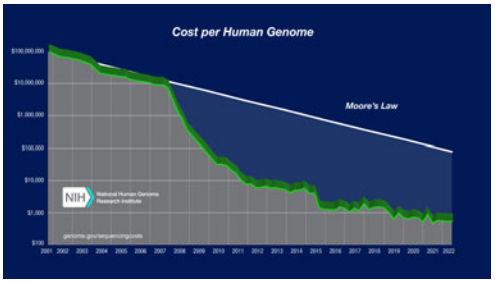
The cost-accounting data presented here are summarized relative to two metrics: (1) "Cost per Megabase of DNA Sequence" - the cost of determining one megabase (Mb; a million bases) of DNA sequence of a specified quality [see below]; (2) "Cost per Genome" - the cost of sequencing a human-sized genome. For each, a graph is provided showing the data since 2001; in addition, the actual numbers reflected by the graphs are provided in a summary table.

NHGRI welcomes people to download these graphs and use them in their presentations and teaching materials. NHGRI plans to update these data on a regular basis. You can view the data in in Excel by downloading the **Sequencing Costs 2022**.



Sequencing cost per megabase - 2022

NeoGenomics
Exhibit 1069 Page 1



Cost per genome data - 2022

To illustrate the nature of the reductions in DNA sequencing costs, each graph also shows hypothetical data reflecting Moore's Law, which describes a long-term trend in the computer hardware industry that involves the doubling of 'compute power' every two years (See: **Moore's Law** [wikipedia.org]). Technology improvements that 'keep up' with Moore's Law are widely regarded to be doing exceedingly well, making it useful for comparison.

In both graphs, note: (1) the use a logarithmic scale on the Y axis; and (2) the sudden and profound out-pacing of Moore's Law beginning in January 2008. The latter represents the time when the sequencing centers transitioned from Sanger-based (dideoxy chain termination sequencing) to 'second generation' (or 'next-generation') DNA sequencing technologies. Additional details about these graphs are provided below.

These data, however, do not capture all of the costs associated with the NHGRI Large-Scale Genome Sequencing Program. The sequencing centers perform a number of additional activities whose costs are not appropriate to include when calculating costs for production-oriented DNA sequencing. In other words, NHGRI makes a distinction between 'production' activities and 'non-production' activities. Production activities are essential to the routine generation of large amounts of quality DNA sequence data that are made available in public databases; the costs associated with production DNA sequencing are summarized here and depicted on the two graphs. Additional information about the other activities performed by the sequencing centers is provided below.

NeoGenomics
Exhibit 1069 Page 2

References

Mardis E. A decade's perspective on DNA sequencing technology. Nature, 470: 198-203. 2011. [**PubMed**]
Metzker M. Sequencing technologies - the next generation. Nature Genetics, 11: 31-46. 2010. [**PubMed**]
Stein L. The case for cloud computing in genome informatics. Genome Biology, 11: 207-213. 2010. [**PubMed**]

Human genome at ten: the sequence explosion. Nature, 464: 670-671. 2010. [**PubMed**]
NHGRI Genome Sequencing Program

How to Cite this Web Page:
Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: **www.genome.gov/sequencingcostsdata**. Accessed [date of access].

Contact

Kris A. Wetterstrand, M.S.
Scientific Liaison to the Director for Extramural Activities
Office of the Director

NeoGenomics
Exhibit 1069 Page 3