Silicon meets Biotech:

Building Better Computers by Incorporating Biological Parts

Luis Ceze

University of Washington

Cost per Raw Megabase of DNA Sequence

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Moore's Law

100,000x in 8 years!

$10K
$1K
$100
$10
$1
$0.1

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

NIH National Human Genome Research Institute

genome.gov/sequencingcosts

Table 3: NextSeq Series Specifications

Instrument Configuration
RFID tracking for consumables
Instrument Control Computer (Internal)
Base Unit: Dual Intel Xeon ES-2448L 1.8 GHz CPU
Memory: 96 GB RAM
Hard Drive: 750 GB
Operating System: Windows 7 embedded standard

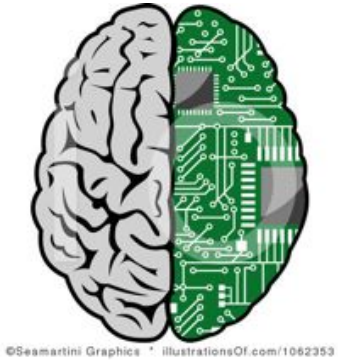
Computer Industry

Helped a **lot**!

Time for payback!

Biotech Industry

Biology
3D Density
Self-assembly
Efficiency
Sensitivity



Silicon
Speed
Engineerability
Integration with Infrastructure

Can sense things silicon can't

Storage

write path

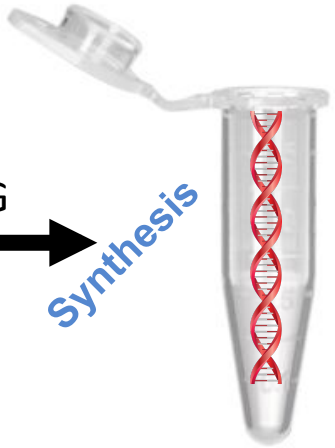
read path

11011101

Encoding

AGCTATCAG

Synthesis



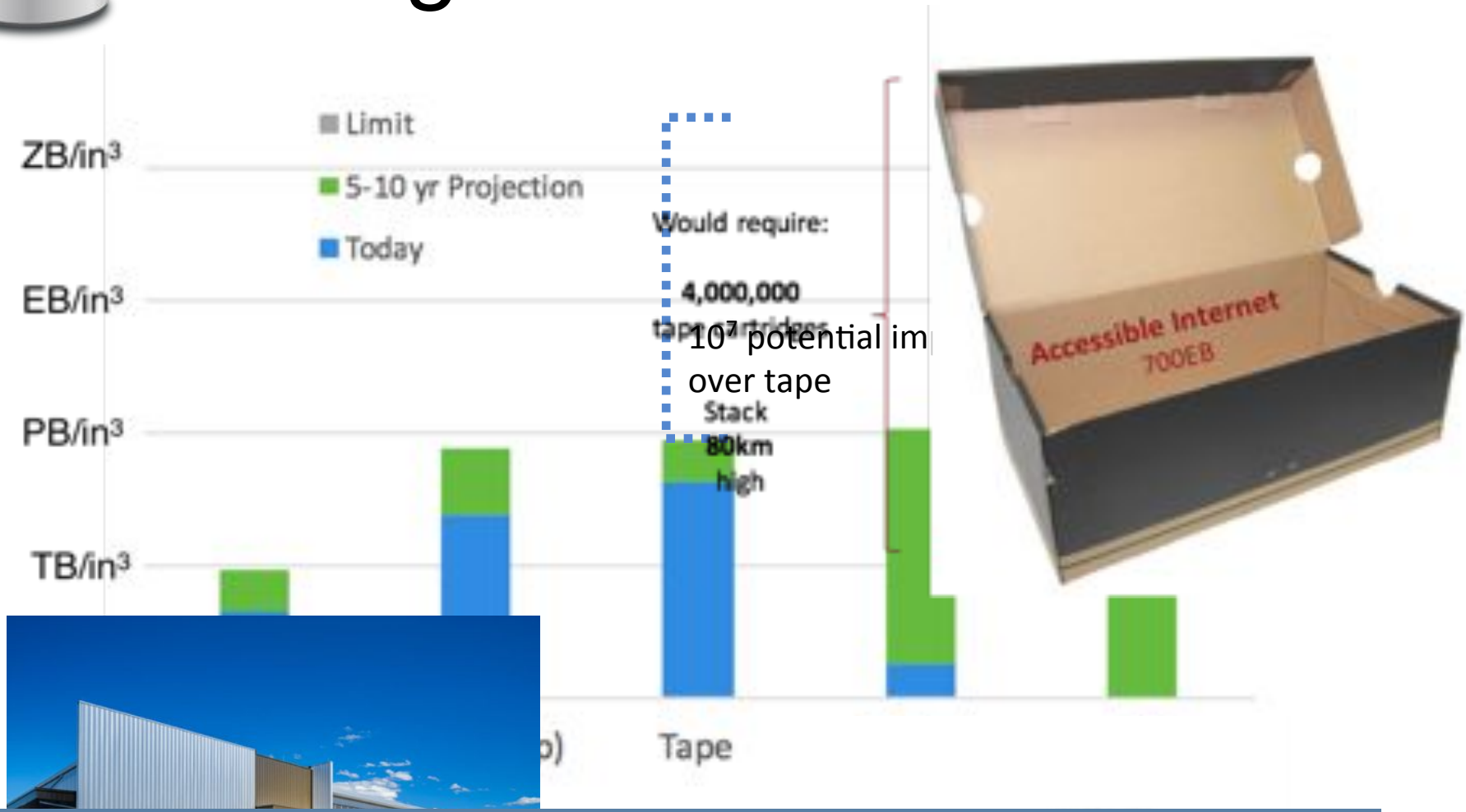
Sequencing

AGCTATCAG

Decoding

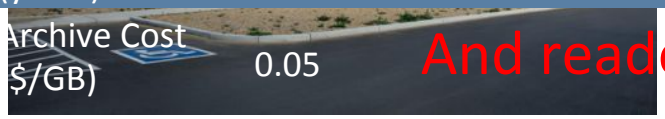
11011101

Storage



10^7 potential im
over tape

Lifetime (years)	3-5	5	10-30	100	**1000+**
Archive Cost ($/GB)	0.05				



And readers never become obsolete!

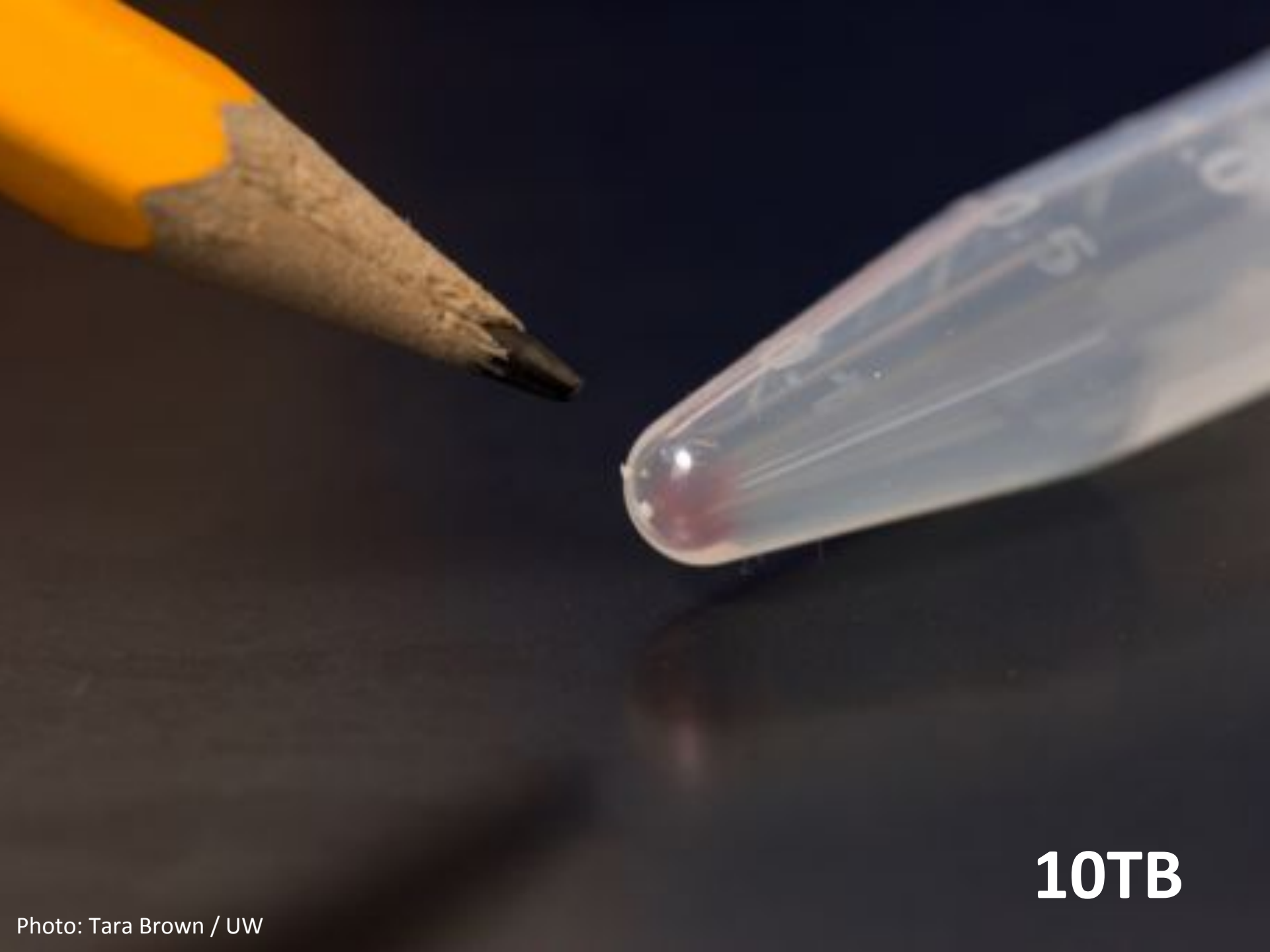


Photo: Tara Brown / UW

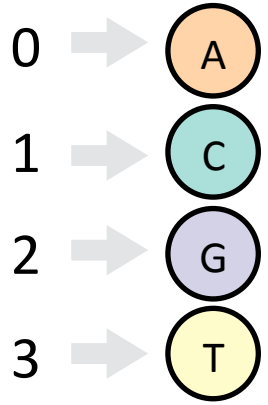
Encoding digital data in DNA

The “direct” way: convert long bit string to base 4

10100011100100011110011111100010110010100101111101…

2 2 0 3 2 1 0 1 3 2 1 3 3 0 1 1 2 1 1 0 2 3 3 1

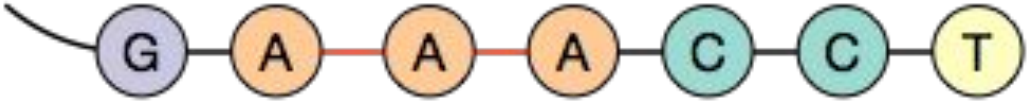
G G A T G C A C T G C T T A C C G C C A G T T C



Direct mapping isn't good enough

The direct mapping approach isn't appropriate for two reasons:

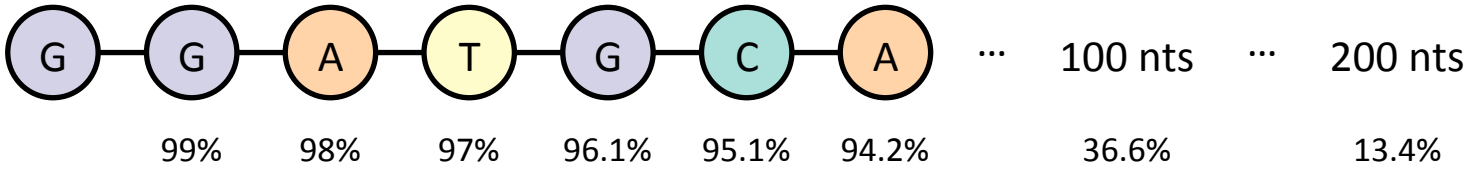
1. Repeated letters are bad:



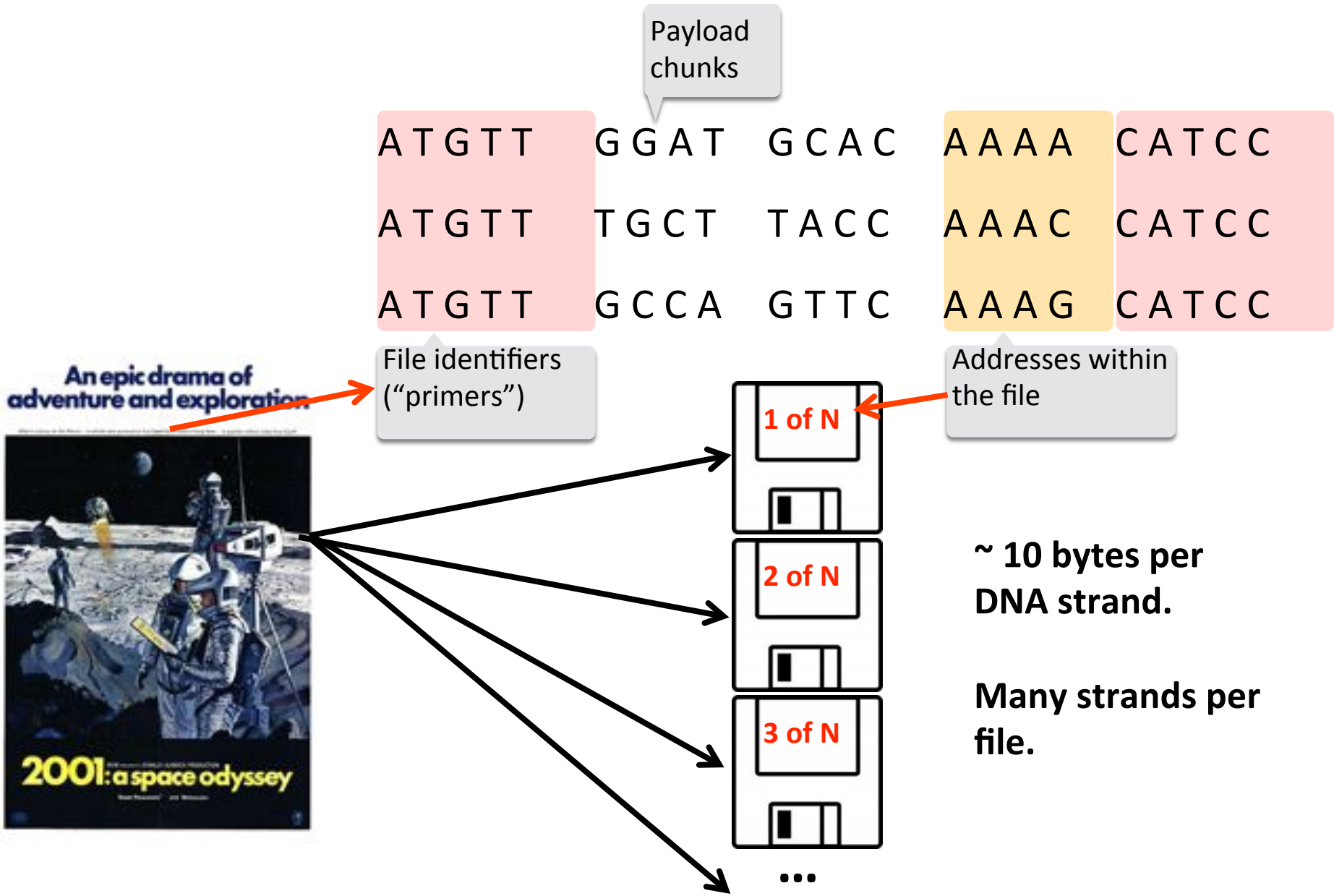
Use base 3 and "rotate" mapping.

2. Synthetic DNA sequences have limited length:

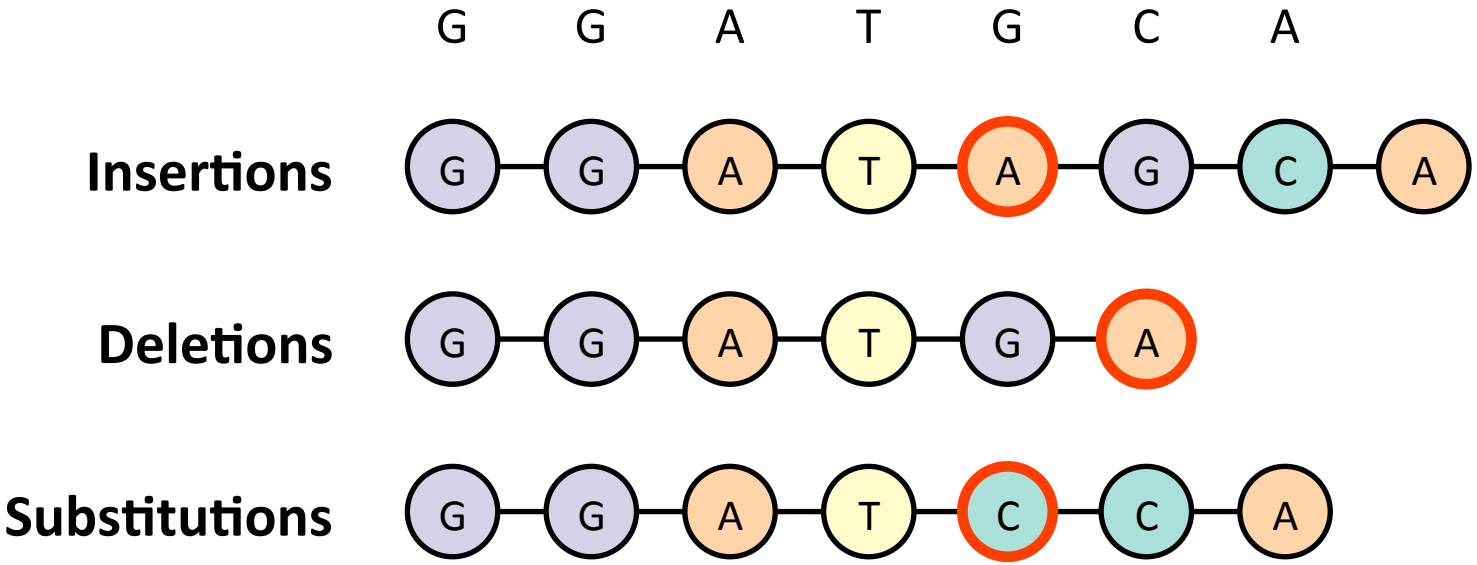
P[Attach] = 99%



Breaking up data into chunks

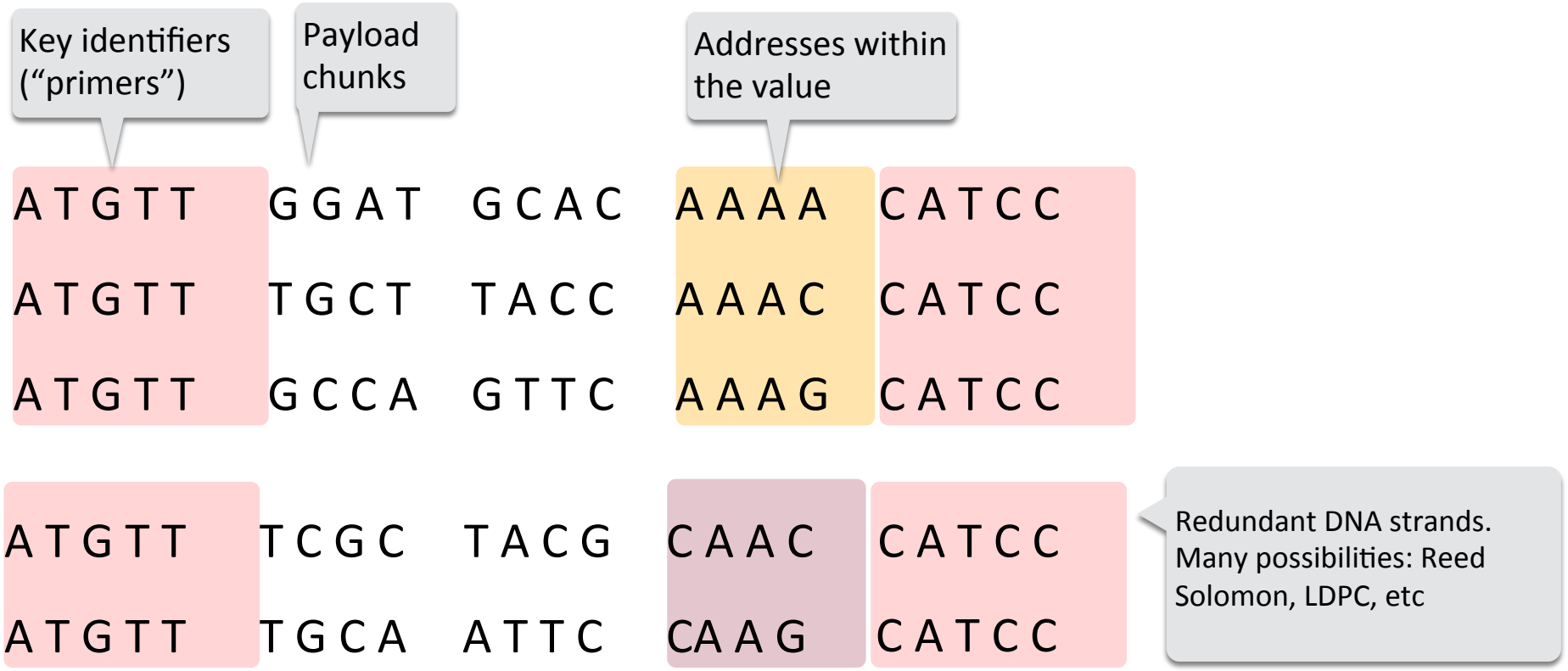


Errors in writing/reading DNA



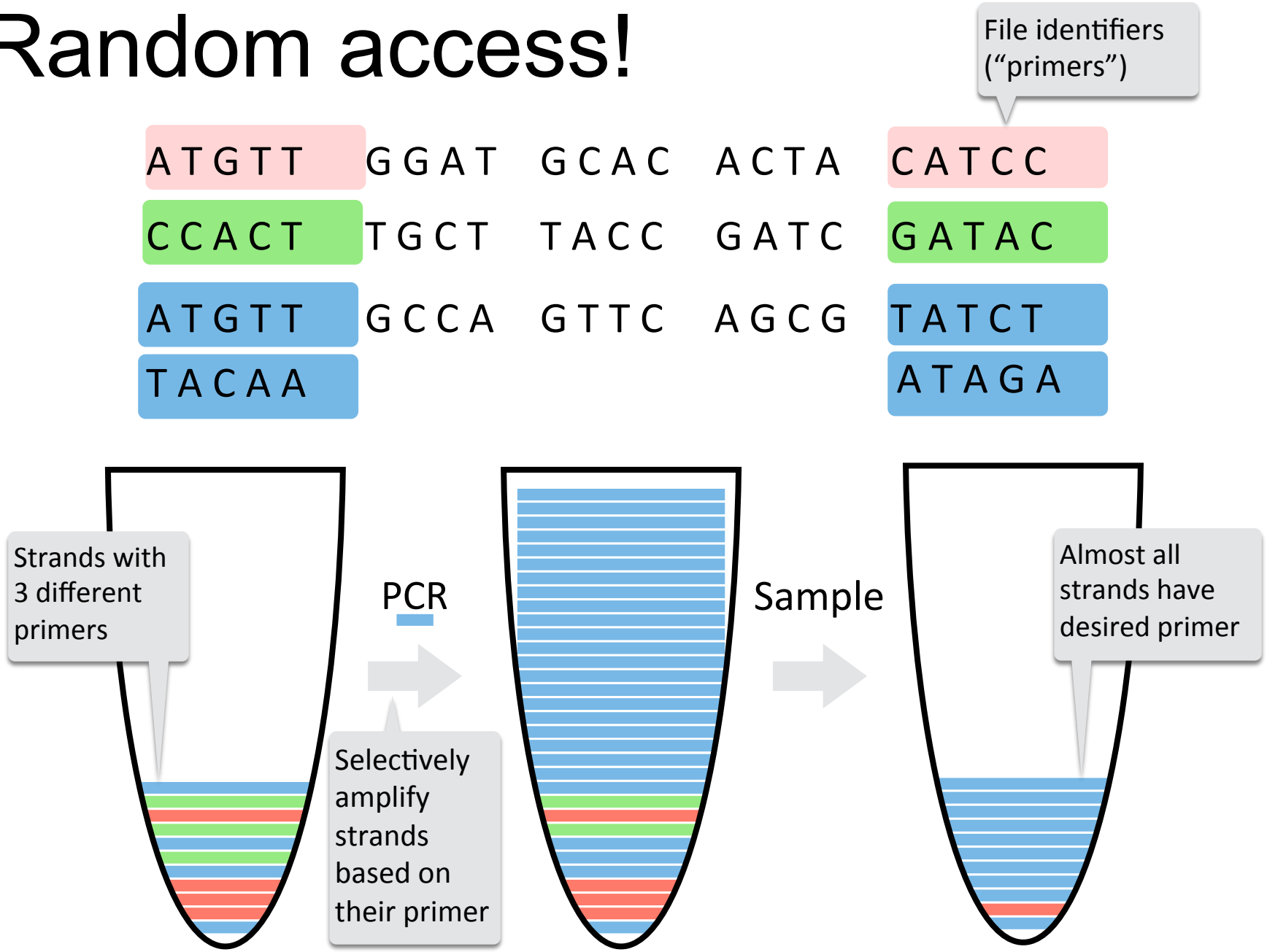
Aggregate error rates ~1%

Adding redundant information



Random access?

Random access!



Experiments

catcatgg

FedEx

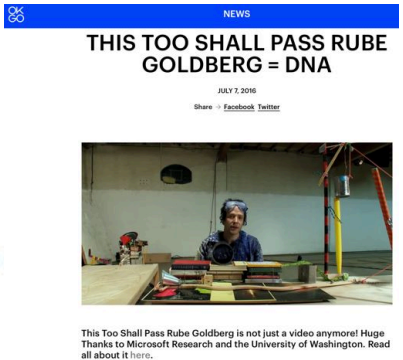
catcatgc

Throughput
10s MB/week

200MB as of July'16
(1.5 B nucleotides)

UNIVERSAL DECLARATION OF HUMAN RIGHTS

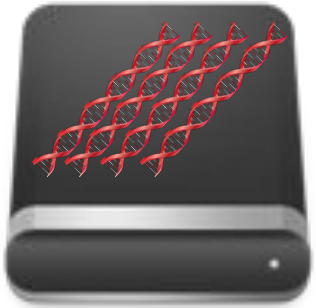
CROP TRUST
SECURING OUR FOOD, FOREVER

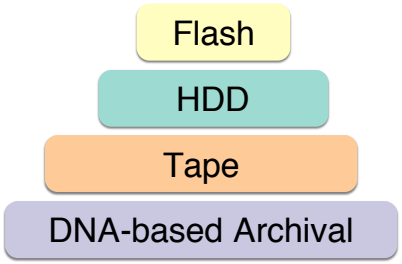


A DNA-based archival storage system

101001101…

[ASPLOS'16]





MISL
Molecular Information Systems Lab

Microsoft Research

Computer architects, coding theorists, biochemist, molecular biologists, fluidics, algorithms, …

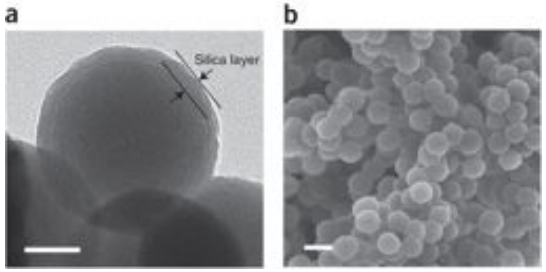
Durable storage by nature

Half-life of DNA in bone: ~521 years.

The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils

Morten E. Allentoft, Matthew Collins, David Harker, James Haile, Charlotte L. Oskam, Marie L. Hale, Paula F. Campos, Jose A. Samaniego, M. Thomas P. Gilbert, Eske Willerslev, Guojie Zhang, R. Paul Scofield, Richard N. Holdaway, Michael Bunce

ETH zürich

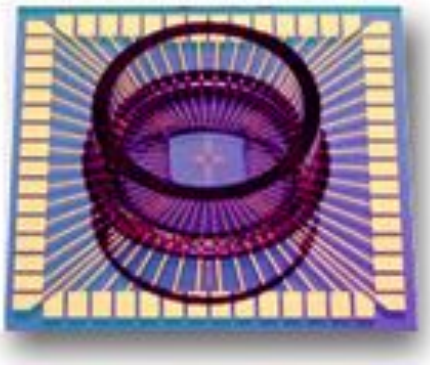


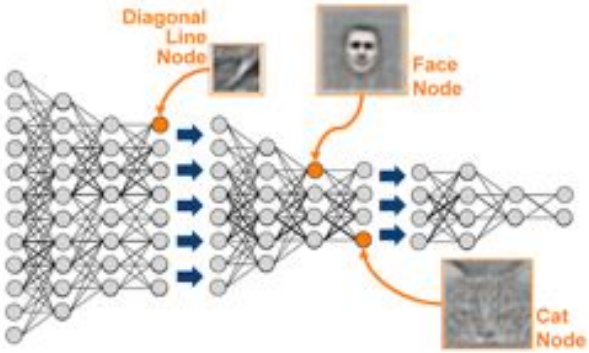
DNA synthetic fossils survive:

Time	Temperature
1 week	70˚C
2,000 years	10˚C
2,000,000 years	-18˚C

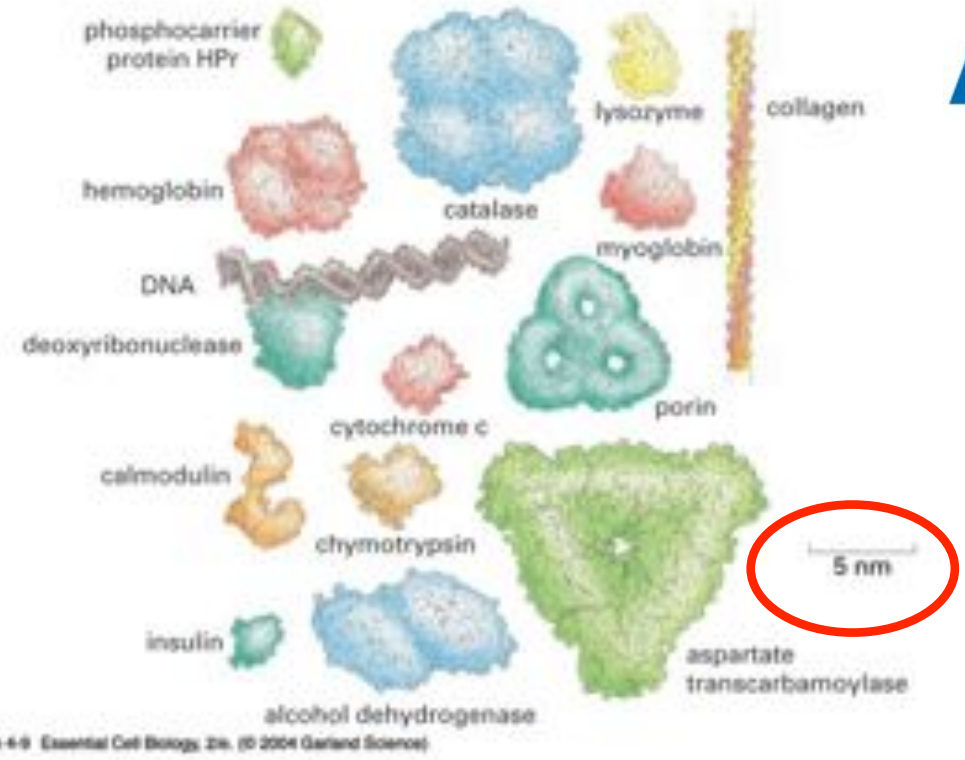
Source: Grass et al. Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes

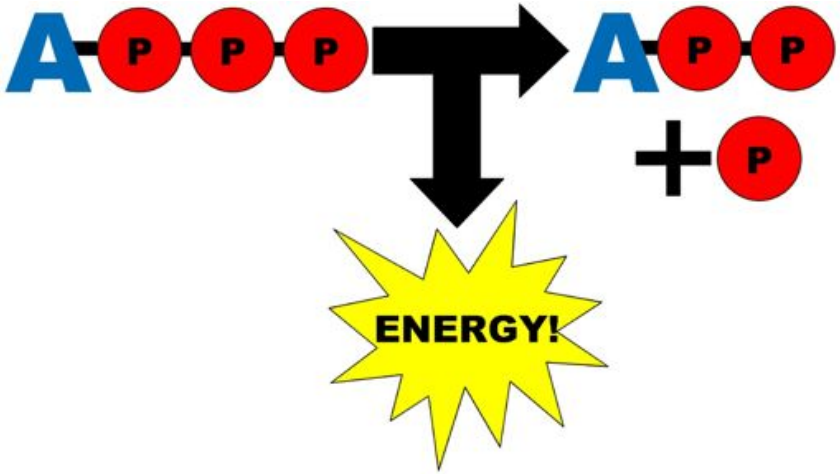
Processing





Proteins are ama





1 ATP molecule = ~ 10^{-21} J

A “proteinistor” would be 10,000x more energy efficient than a CMOS transistor

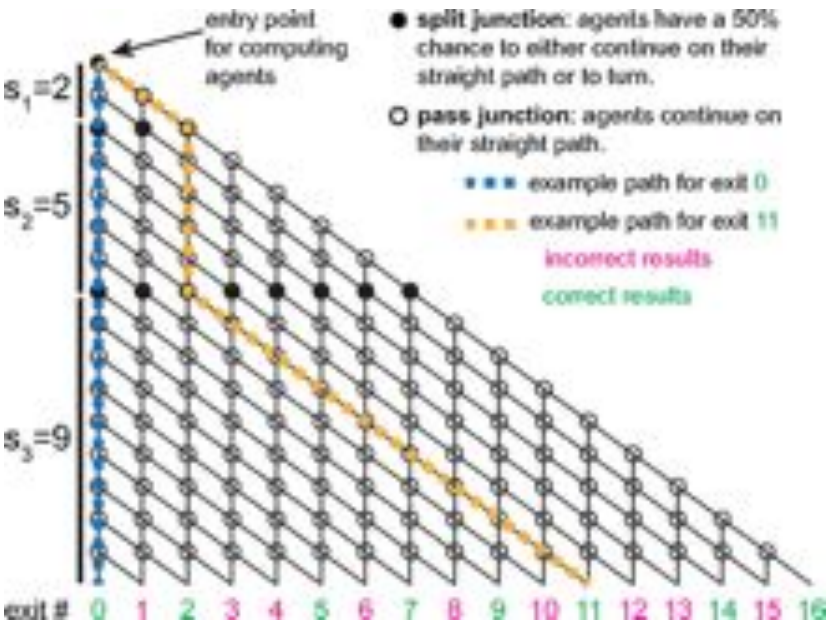
ATP powered “supercomputer”?

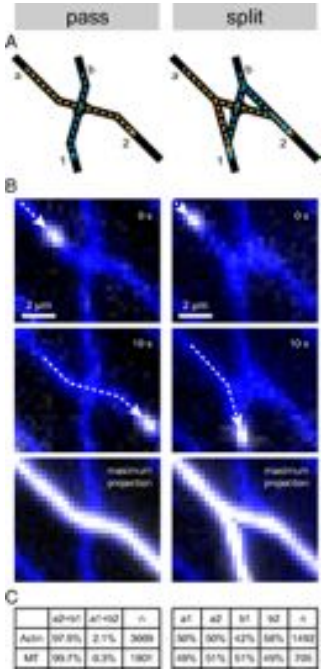
Parallel computation with molecular-motor-propelled agents in nanofabricated networks

Dan V. Nicolau Jr.[a,b,1], Mercy Lard[c,1], Till Korten[d,e,1], Falco C. M. J. M. van Delft[f,2], Malin Persson[g], Elina Bengtsson[g], Alf Månsson[g], Stefan Diez[d,e], Heiner Linke[c,3], and Dan V. Nicolau[h,i,3]

[a]Department of Integrative Biology, University of California, Berkeley, CA 94720-3140; [b]Molecular Sense, Ltd., Wallasey CH44 1AJ, United Kingdom; [c]NanoLund and Solid State Physics, Lund University, S-22100 Lund, Sweden; [d]Center for Molecular Bioengineering (B CUBE) and Center for Advancing Electronics Dresden (cfaed), Technische Universität Dresden, 01069 Dresden, Germany; [e]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; [f]Philips Research (MiPlaza) and Philips Innovation Services, 5656 AE, Eindhoven, The Netherlands; [g]Department of Chemistry and Biomedical Sciences, Linnaeus University, SE-39182 Kalmar, Sweden; [h]Department of Electrical Engineering & Electronics, University of Liverpool, Liverpool L69 3GJ, United Kingdom; and [i]Department of Bioengineering, McGill University, Montreal, QC, Canada H3A 0C3

Edited by Hillel Kugler, Microsoft Research, Cambridge, United Kingdom, and accepted by the Editorial Board January 18, 2016 (received for review June 5, 2015)





Or an ATP battery?

NATURE COMMUNICATIONS | ARTICLE

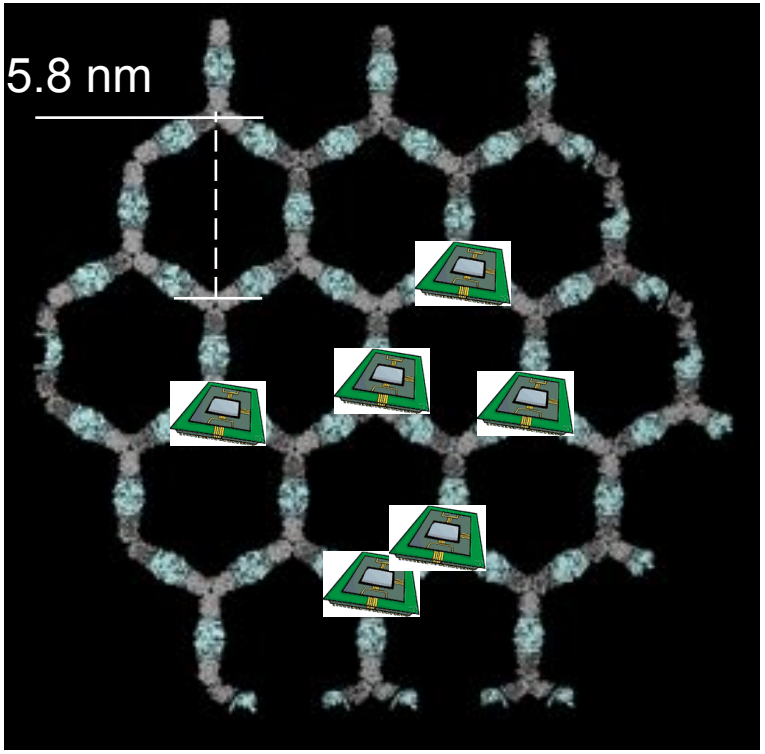
A high-energy-density sugar biobattery based on a synthetic enzymatic pathway

Zhiguang Zhu, Tsz Kin Tam, Fangfang Sun, Chun You & Y. -H. Percival Zhang

10x energy density of li-ion batteries

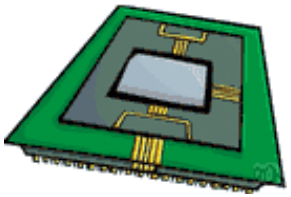
Molecular-Level Self-Assembly *and Reconfiguration*

DNA origami



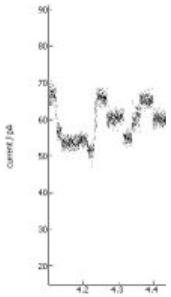
Proteins

[Caltech, Duke]

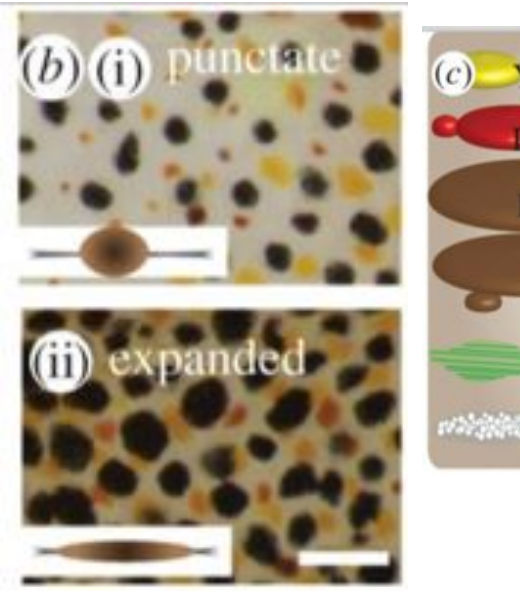


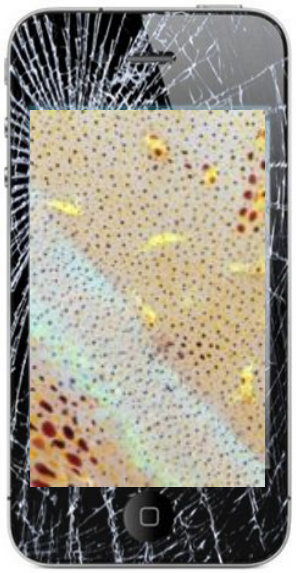
Single-Molecule Sensing

Gu et al. Single molecule sensing by nanopores and nanopore devices. 2009, Analyst.



What about output?





igment
ell
MC
eus
rk

CuttlePhone

Inevitable to go to the molecular/atomic level.

Biology has “invented” many useful parts, not just ideas.

Biotech is quickly developing many useful tools/methods.

Silicon + Biotech has immense opportunities!