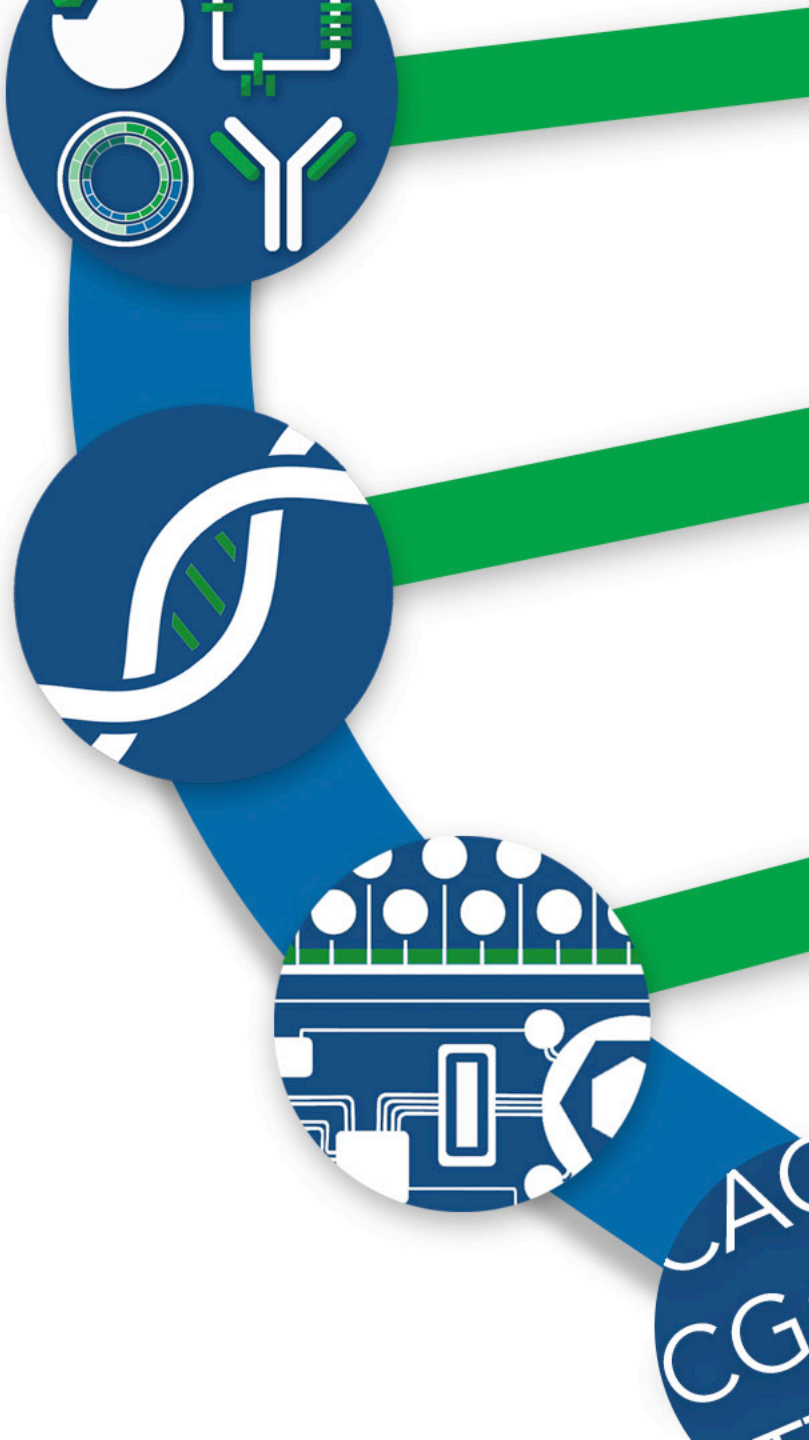
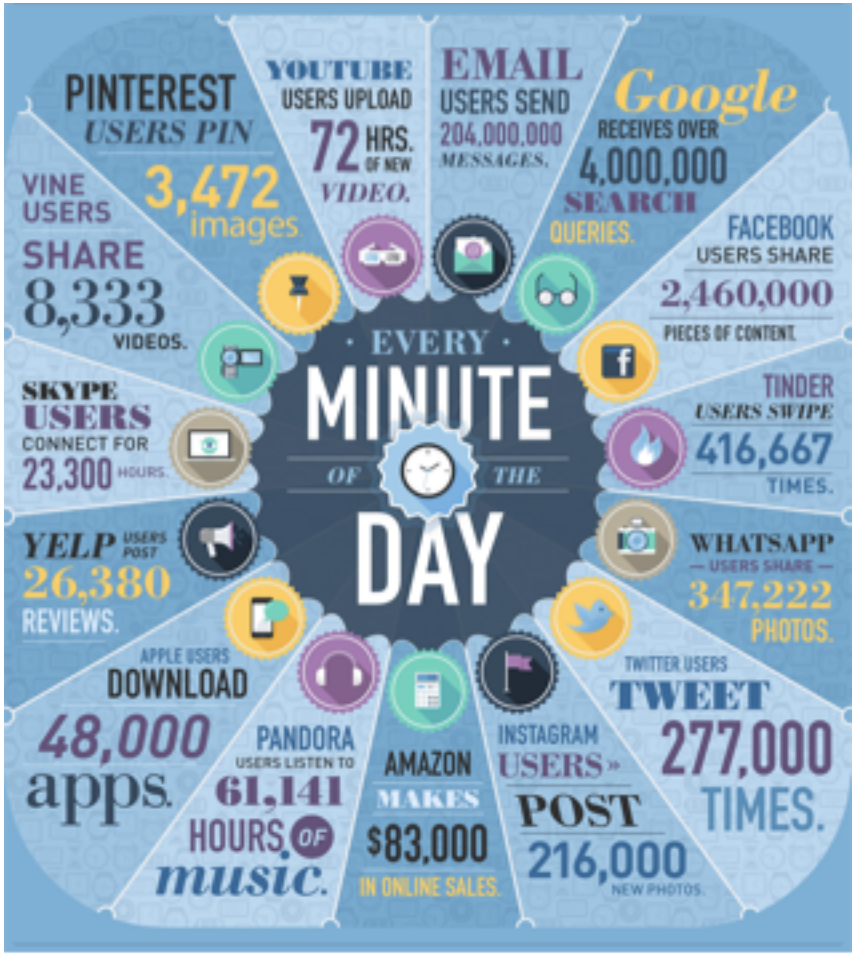
Viability of DNA for Archival Storage

Devin Leake
VP of R&D

Gen9



There is a growing need for storage



Content created by DOMO, "Data Never Sleeps"

- According to the IDC (International Data Corporation)
  - 1.8 zettabytes was generated in 2011
  - By 2020, 40 zettabytes will be generate (or 40 Trillion GB)

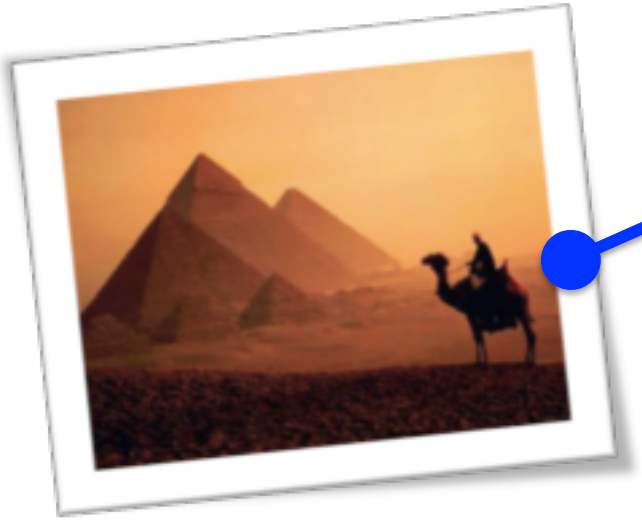
Gen9

Types of storage based on needs

Active Data

Backup Data

Archival Data

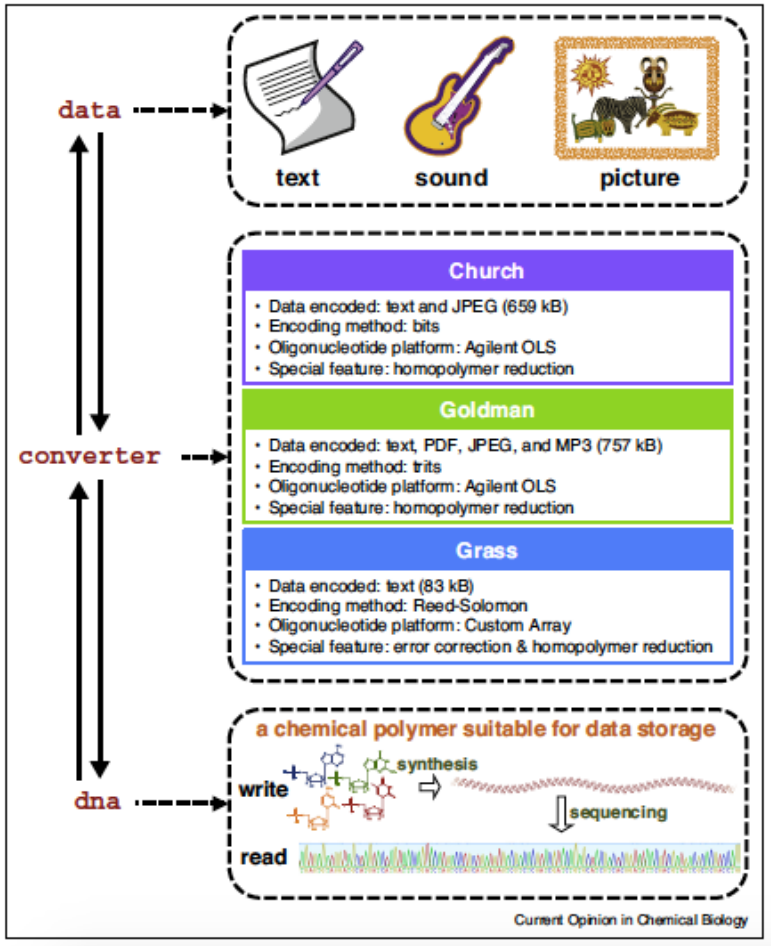


High storage durability

Low total cost of ownership

Gen9

DNA as a medium for archival storage



- Viability of Reed-Solomon error-correcting codes adapted to DNA
- Limited demonstration of storage capacity (less than 1 MB)
- Utility of multiple media formats (ASCI, JPEG, PDF, MP3)

Zakeri and Lu Current Opinion in Chemical Biology 2015

Gen9

High storage durability

- Hardware can fail but storage capability cannot
- Data must be protected from modification and distributed such that data loss is minimized
- DNA has been recovered and analyzed from:
  - Wooly mammoth (20,000 years ago)
  - Neanderthal (40,000 years ago)
  - Bison (60,000 years ago)
  - Pollen samples (~500,000 years ago)

Gen9

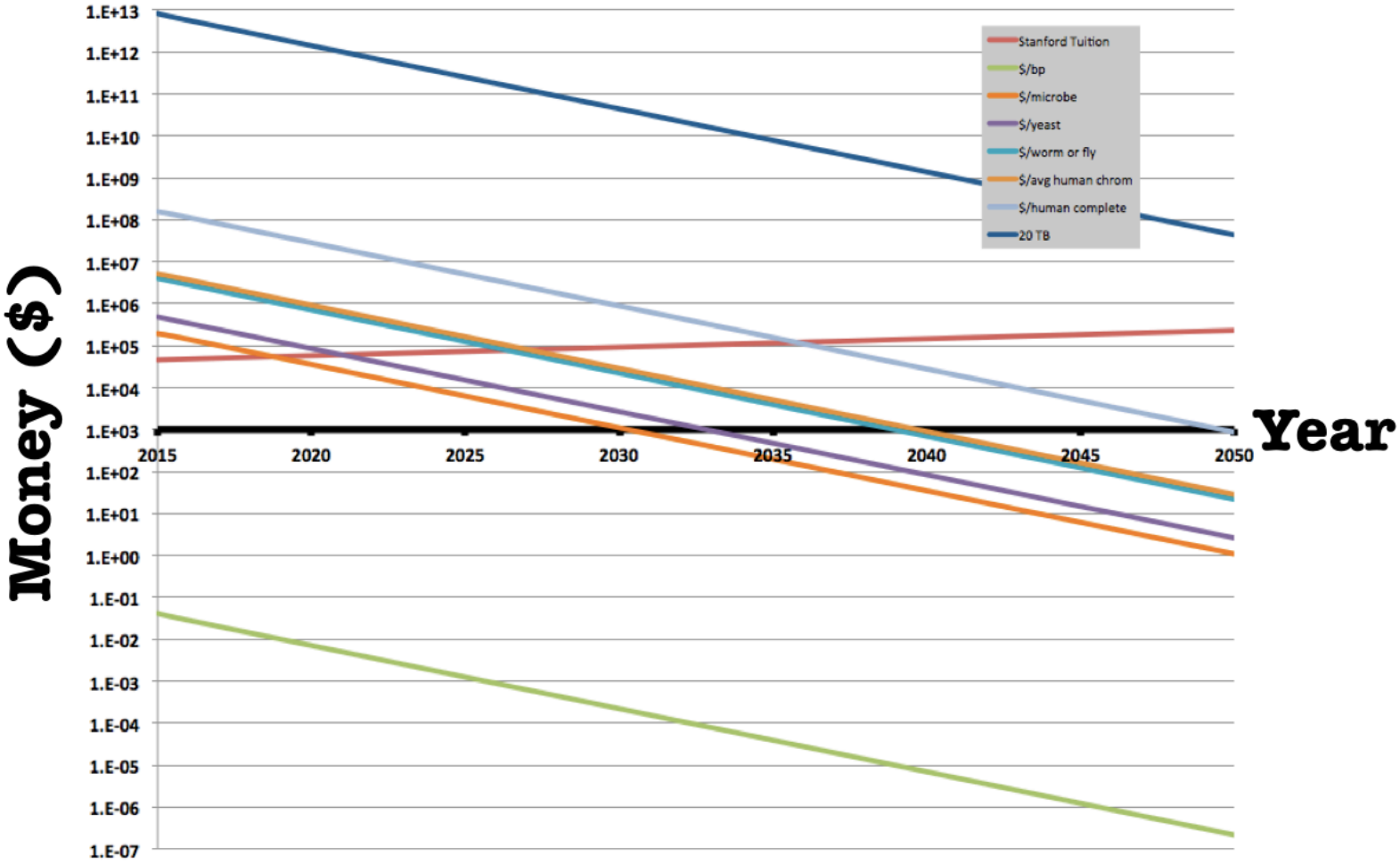
Low total cost of ownership

- Current convention (Grass et.al. Angew. Chem. Ins. Ed. 2015) indicates 10 billion base pairs are need for each 1 GB
- Estimates suggest that a 1.5mL tube filled with DNA could hold approximately a petabyte of data (Goldman et.al. Nature 2013)

➢ Currently, **cost** of DNA writing is **too high** and synthesis **capacity** is **insufficient**

Gen9

DNA synthesis cost and scale



Courtesy of Drew Endy, Gen9/Stanford 2015

Gen9

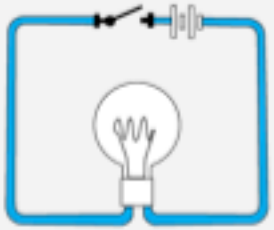
Massive divergence in capacity

Building a scalable BioFab® platform

Joe Jacobson
MIT

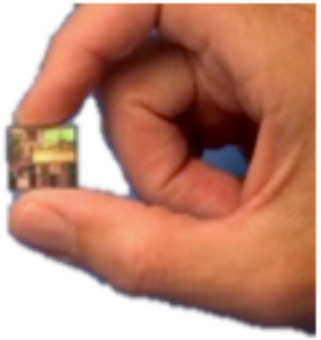
Drew Endy
Stanford

George Church
Harvard



Simple Circuit

Complex Circuit



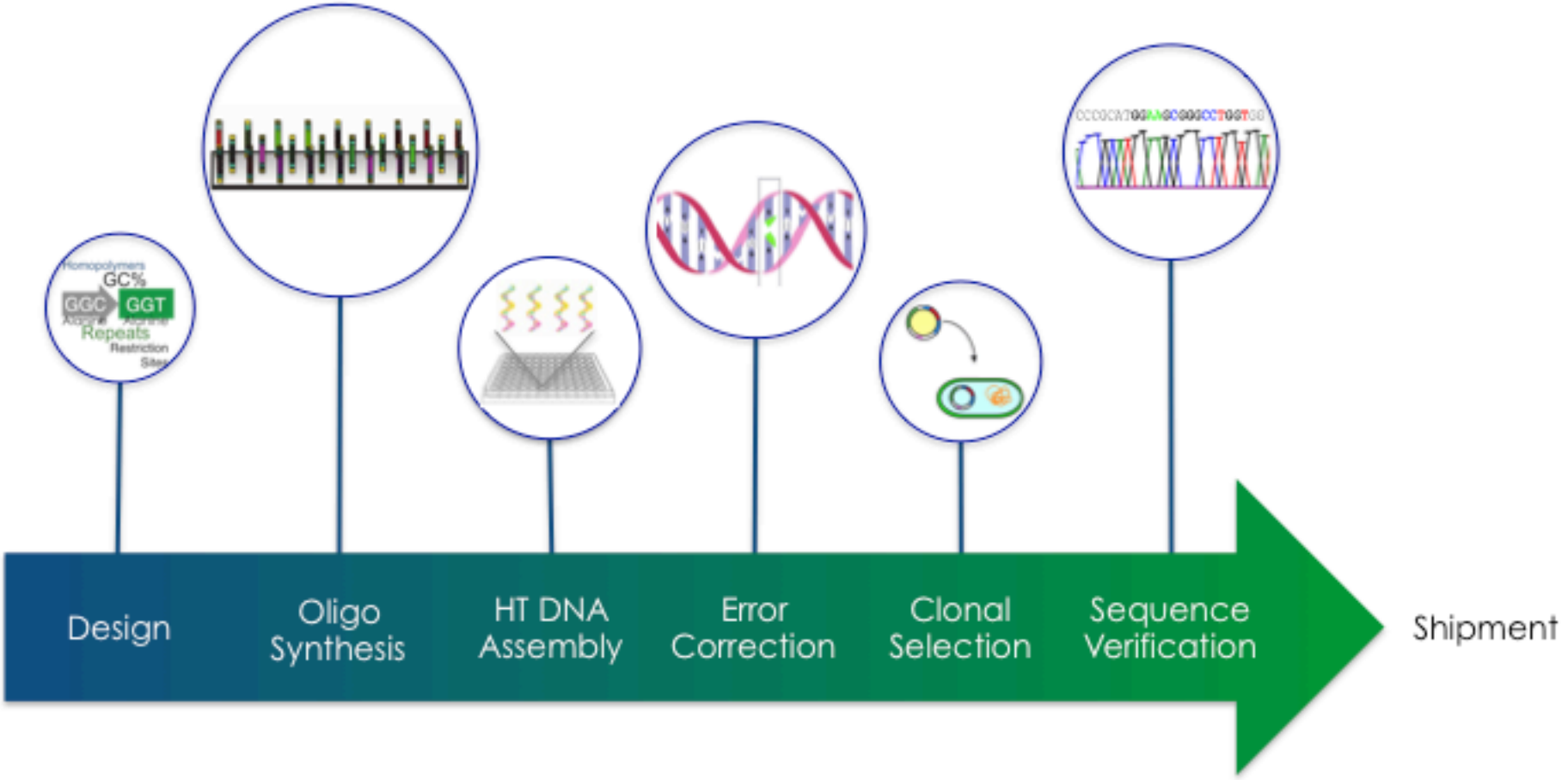
Integrated Circuit

Oligo Synthesis

Gene Synthesis

Gen9 BioFab® Chip

Gen9's BioFab® Platform



Thank You