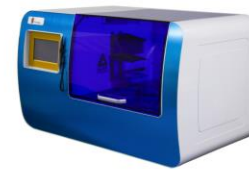


Synthetic DNA Synthesis

D. Lavenier

Principle of DNA data storage

00 → A
01 → C
10 → G
11 → T



A → 00
C → 01
G → 10
T → 11

Coding

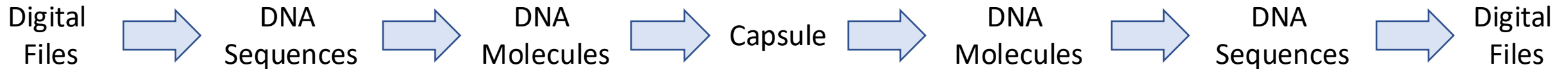
Synthesis

Storage

Retrieval

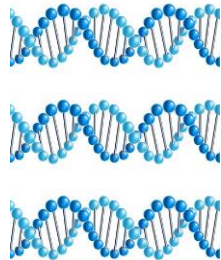
Sequencing

Decoding



010100101
101010011
010001010
001010010

ATGAGGACCG
TTAGGACCAG
GCCATATAAC
GGTTATATAA



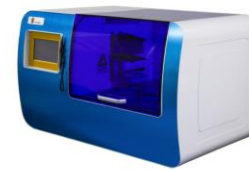
ATGAGGACCG
TTAGGACCAG

010100101
101010011



Principle of DNA data storage

00 → A
01 → C
10 → G
11 → T



A → 00
C → 01
G → 10
T → 11

Coding

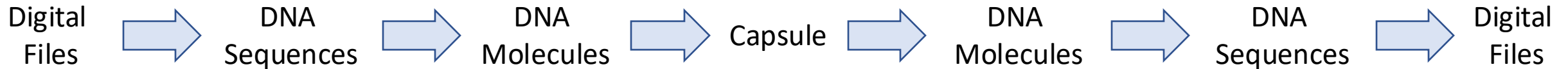
Synthesis

Storage

Retrieval

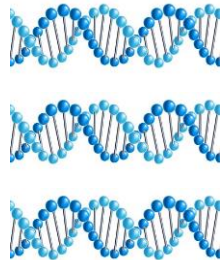
Sequencing

Decoding



010100101
101010011
010001010
001010010

ATGAGGACCG
TTAGGACCAG
GCCATATAAC
GGTTATATAA

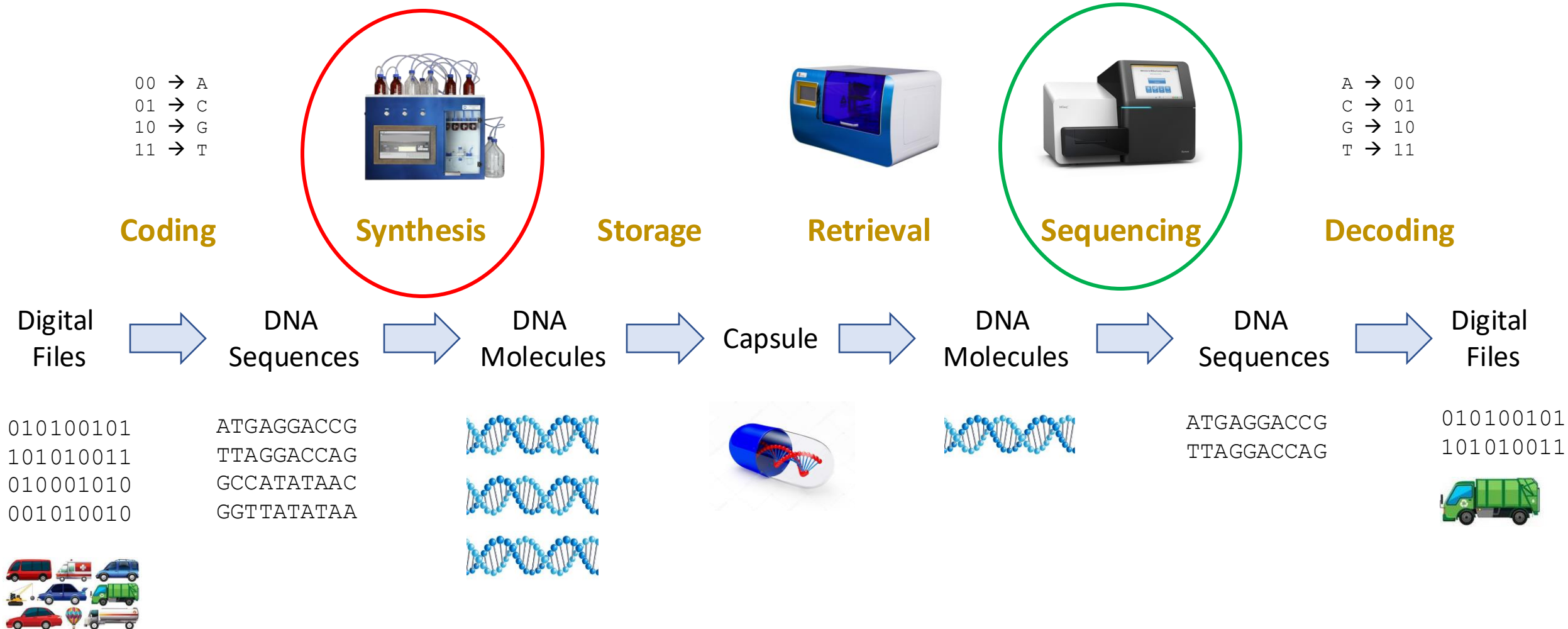


ATGAGGACCG
TTAGGACCAG

010100101
101010011



Principle of DNA data storage



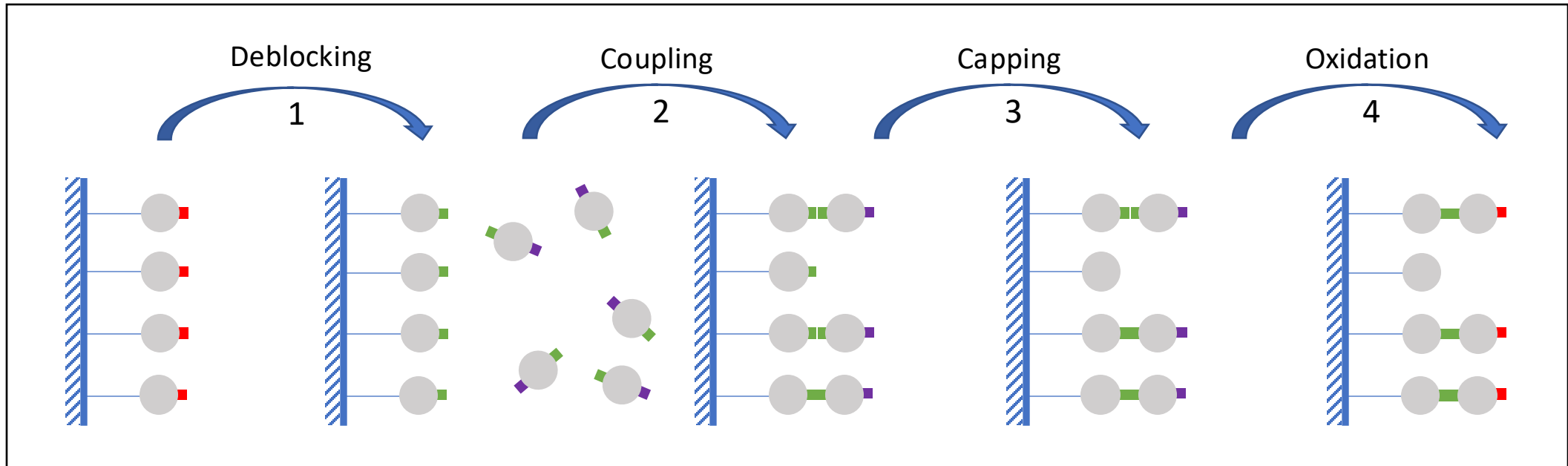
DNA synthesis

Sequential process

Oligonucleotide of size $N \rightarrow N$ cycles

1 cycle $\rightarrow$ 4 steps

1 step $\rightarrow$ a few 10's of sec.

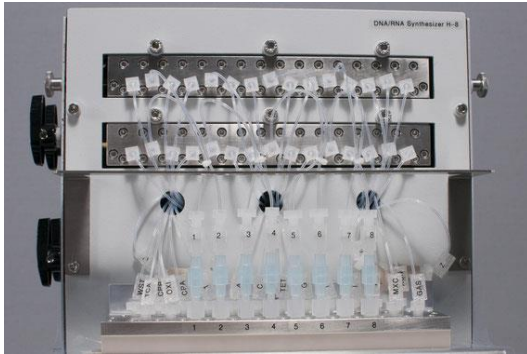


Phosphoramidite synthesis

Synthesis

AGGGACCGGTGACACCCGGTTGGACCGTGA

1 text sequence

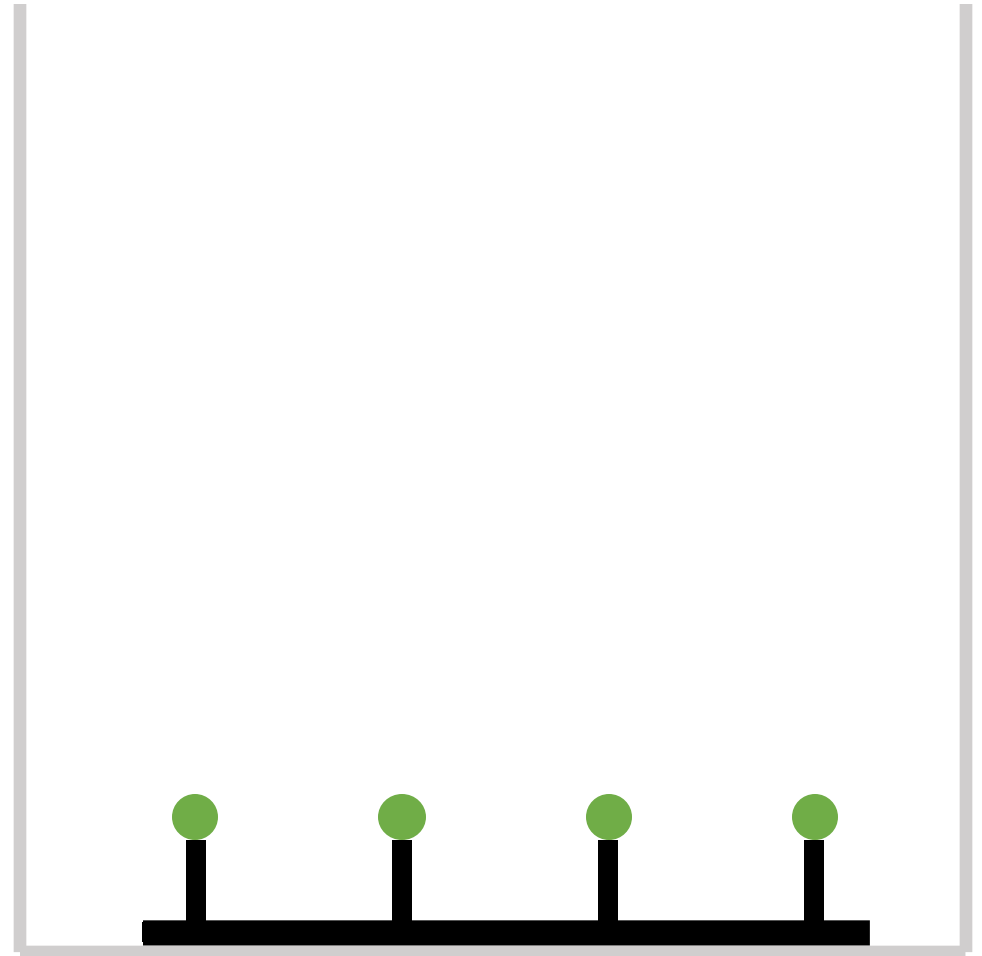


DNA
Synthesizer



Thousands / Million of molecules

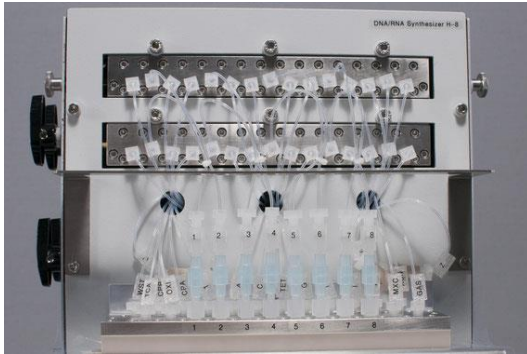
AGTCGTAC...



Synthesis

AGGGACCGGTGACACCCGGTTGGACCGTGA

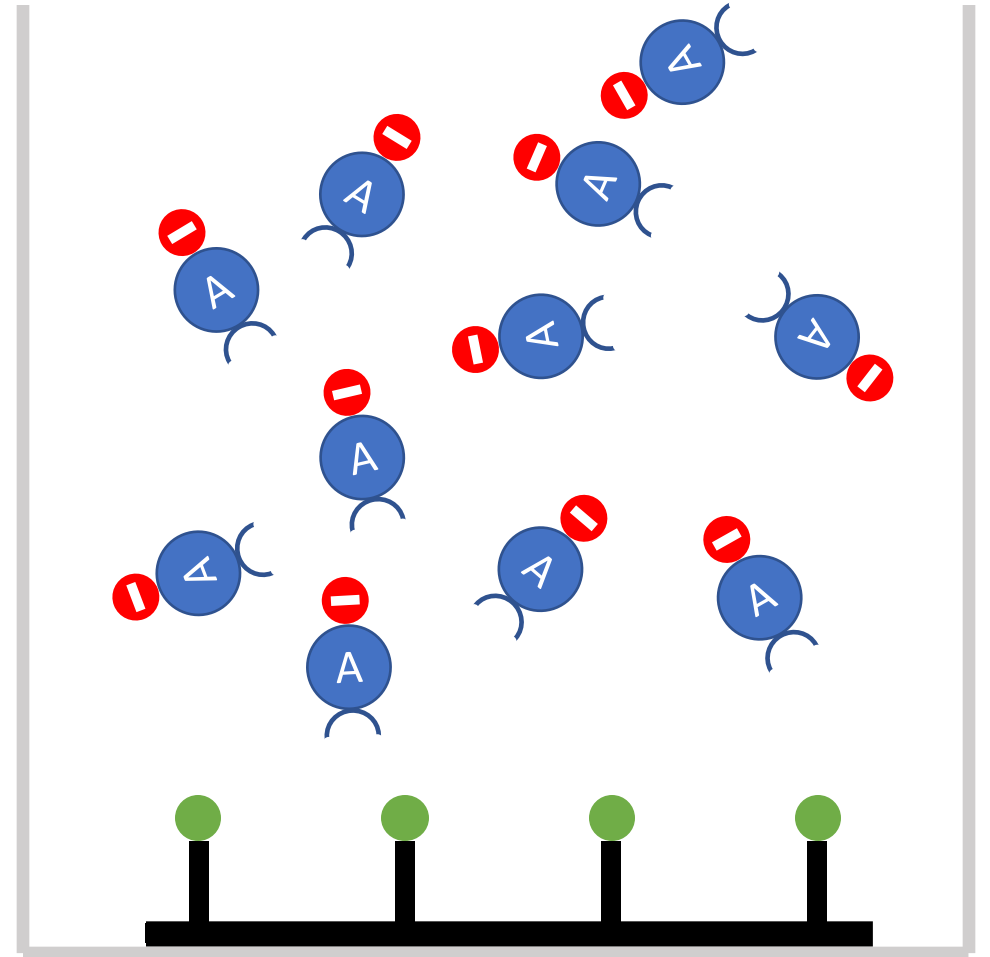
1 text sequence



DNA
Synthesizer



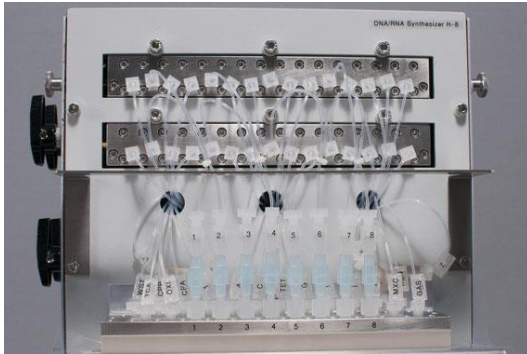
AGTCGTAC...



Synthesis

AGGGACCGGTGACACCCGGTTGGACCGTGA

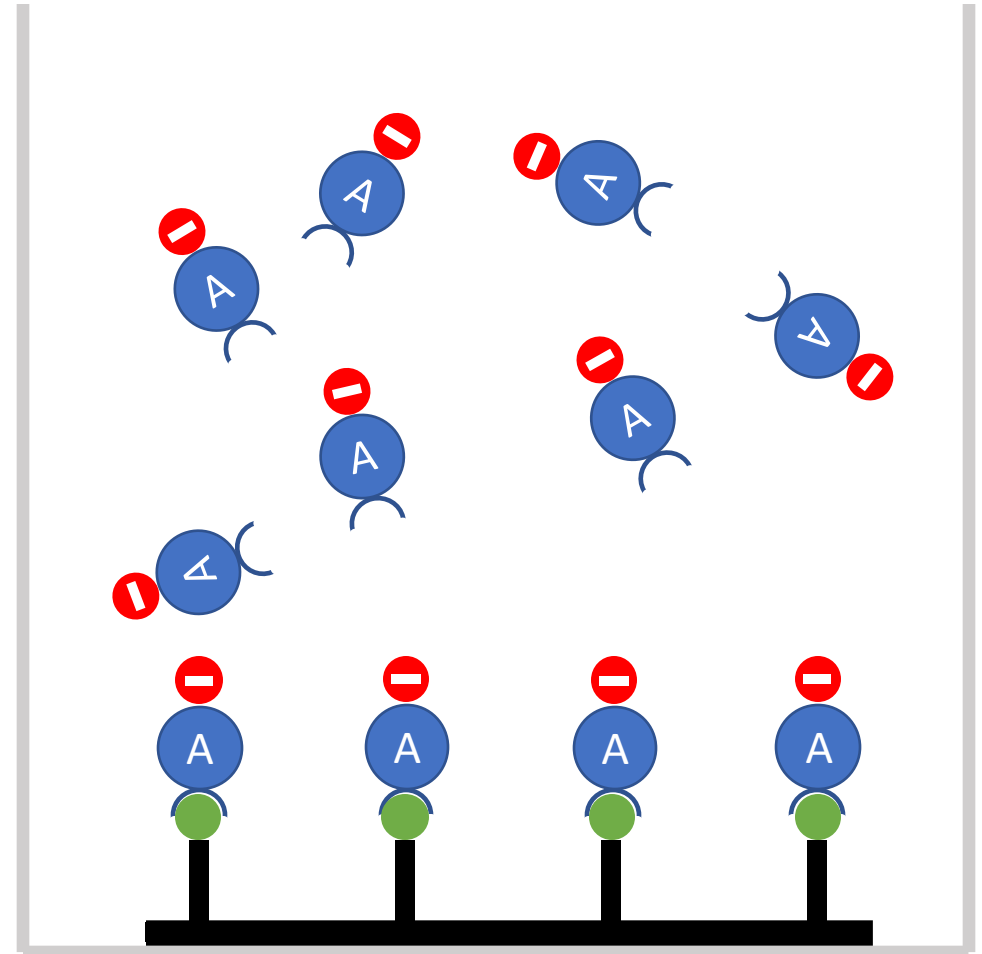
1 text sequence



DNA
Synthesizer



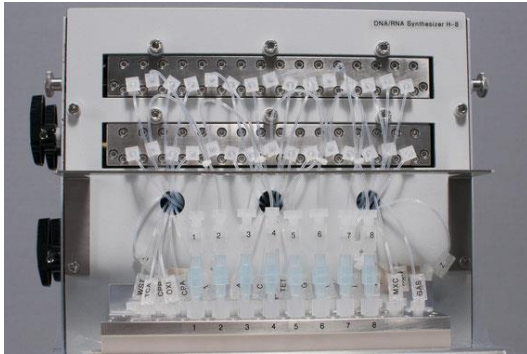
AGTCGTAC...



Synthesis

AGGGACCGGTGACACCCGGTTGGACCGTGA

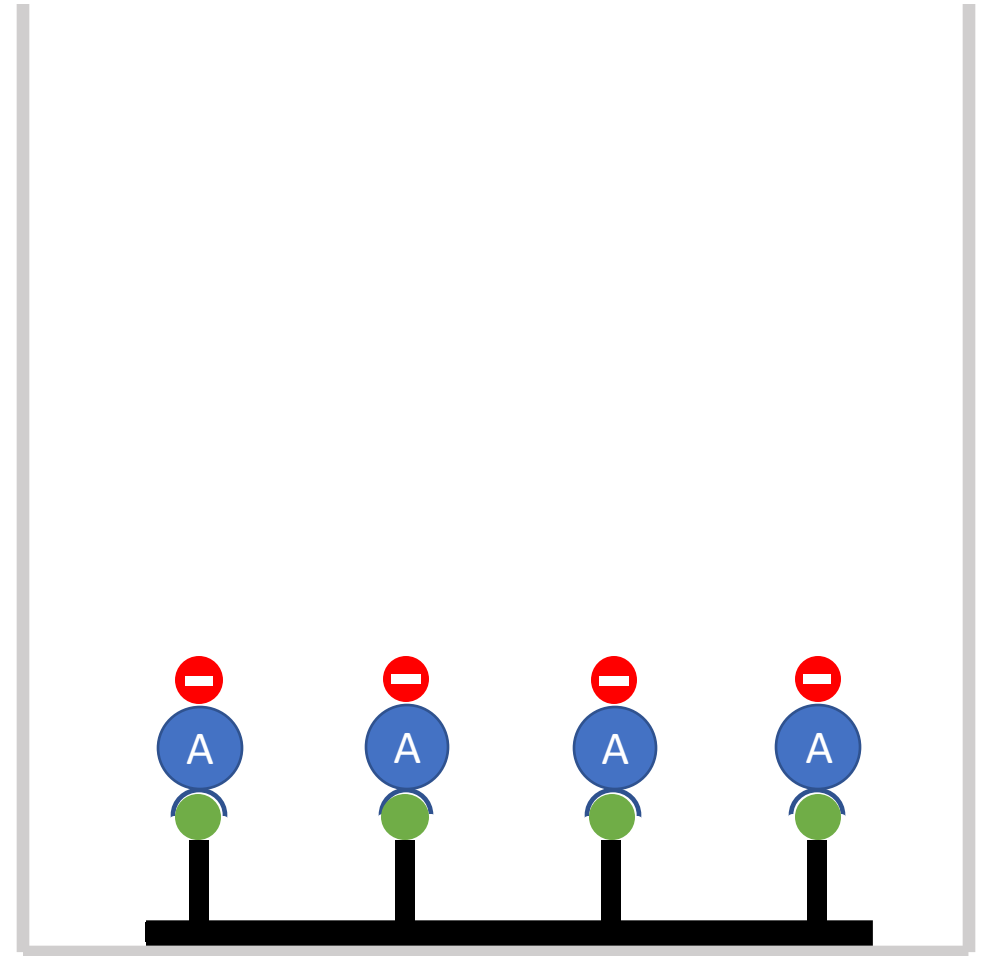
1 text sequence



DNA
Synthesizer



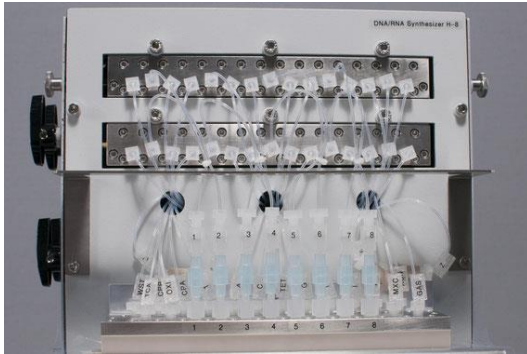
AGTCGTAC...



Synthesis

ATGGACCGGTGACACCCGGTTGGACCGTGA

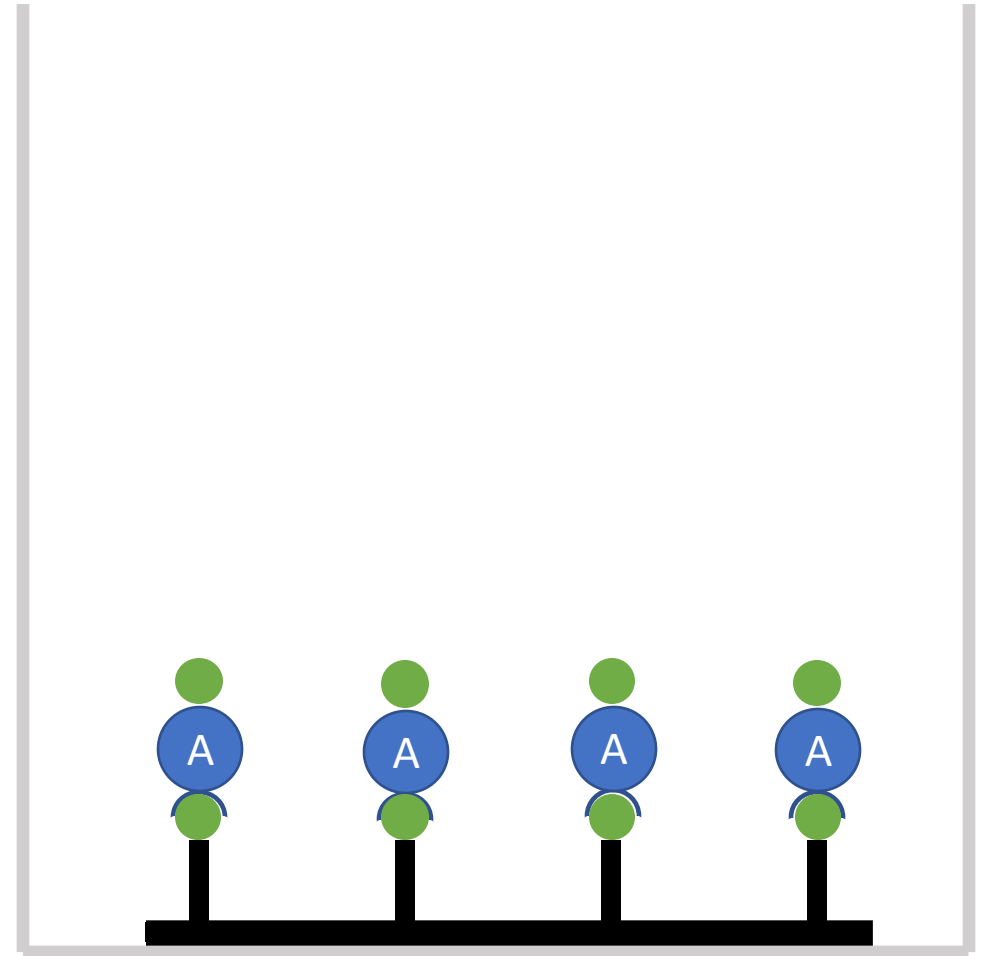
1 text sequence



DNA
Synthesizer



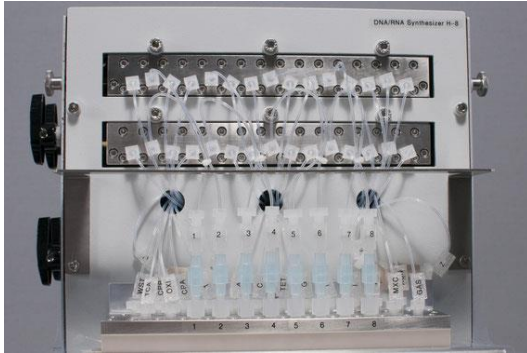
AGTCGTAC...



Synthesis

AGGGACCGGTGACACCCGGTTGGACCGTGA

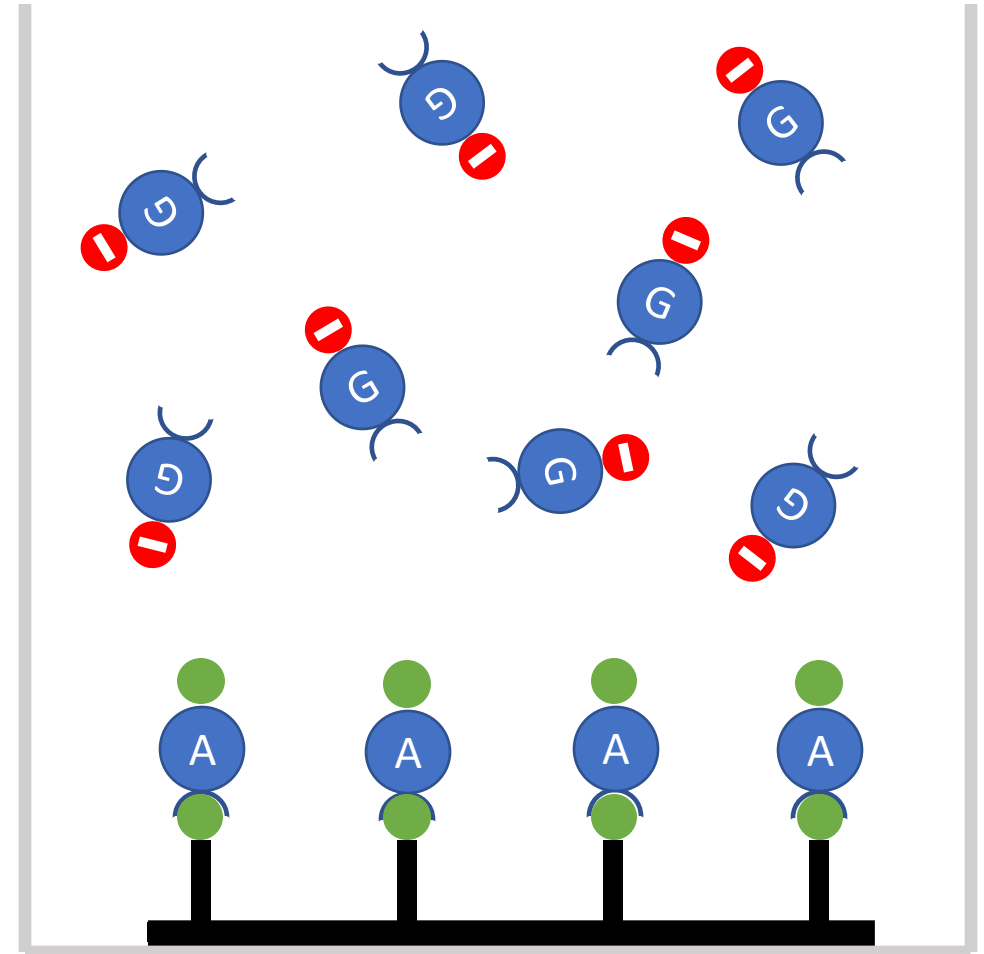
1 text sequence



DNA
Synthesizer



AGTCGTAC...



Synthesis limitation

Synthesizers can only generate short molecules (oligonucleotides)

- oligonucleotide length < 300 nucleotides

Very slow process

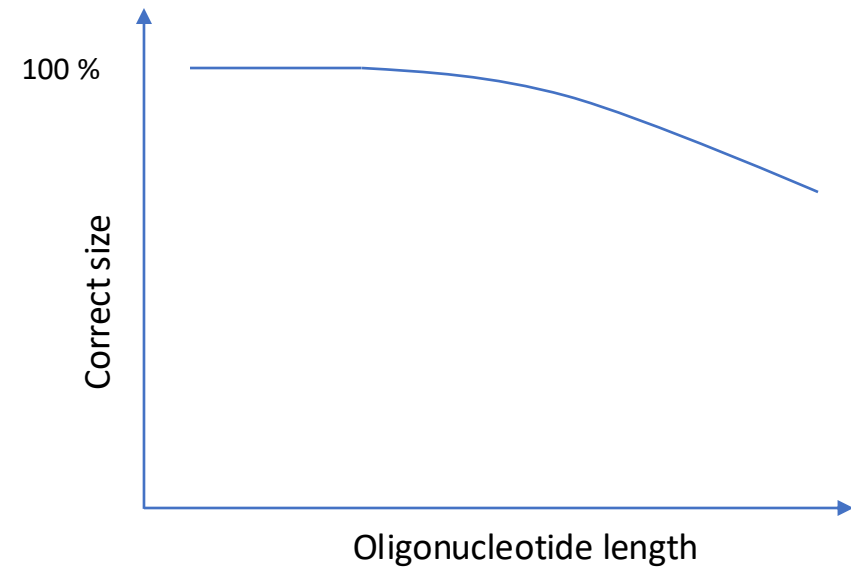
- 1 nucleotide / minute

High cost

- 10-100 € / Kbyte

Errors

- variation on oligonucleotide length



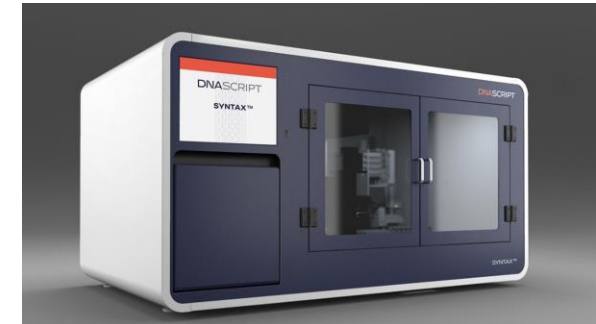
Synthesis perspectives

New technologies

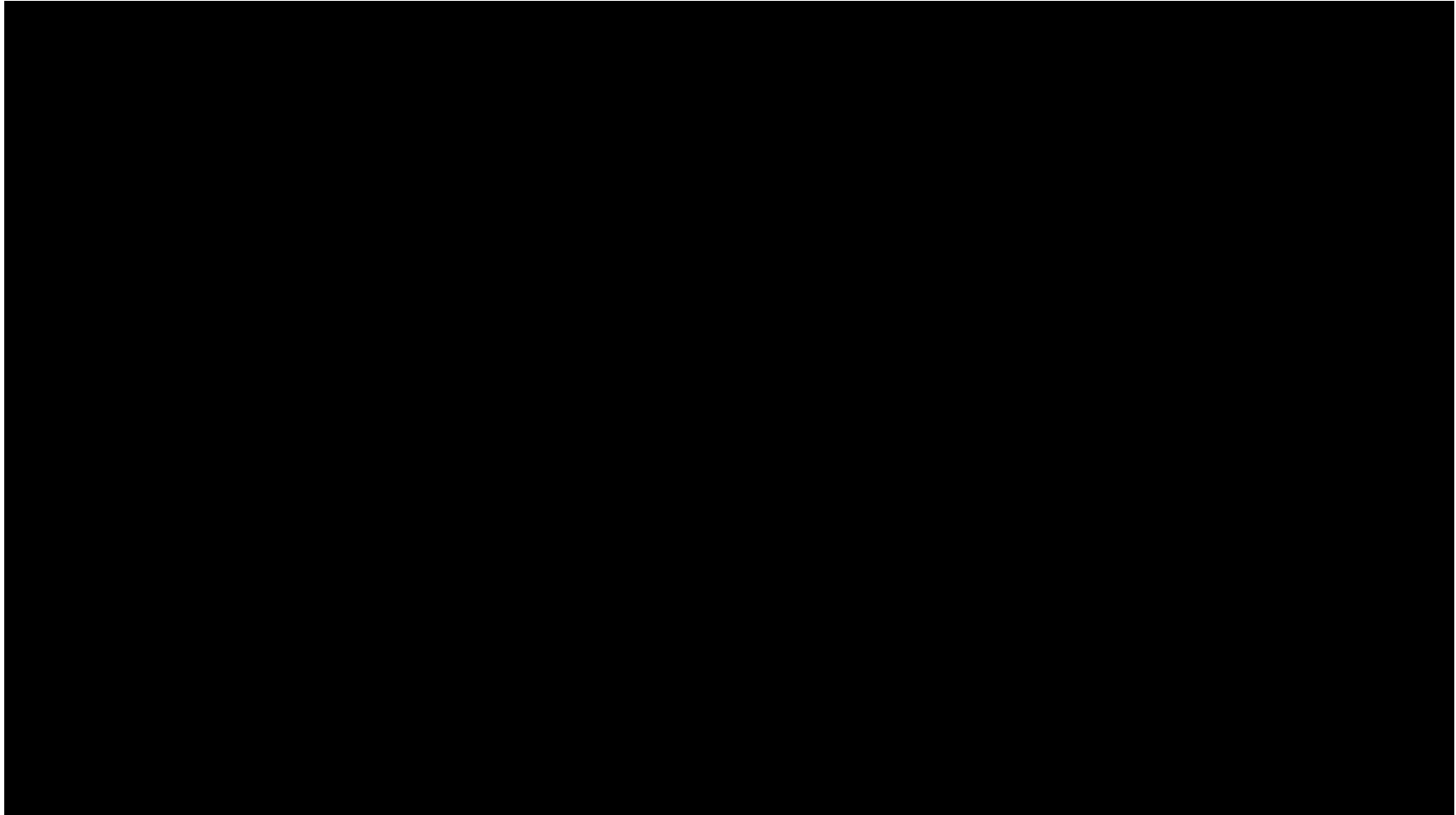
- Enzymatic DNA synthesis
 - exploit the capabilities of the TdT enzyme
 - longer oligonucleotides
 - a little bit faster

Miniaturization & Parallelization

- Roadmap 2025
 - 1 Tbytes
 - 1 day
 - 1000 \$

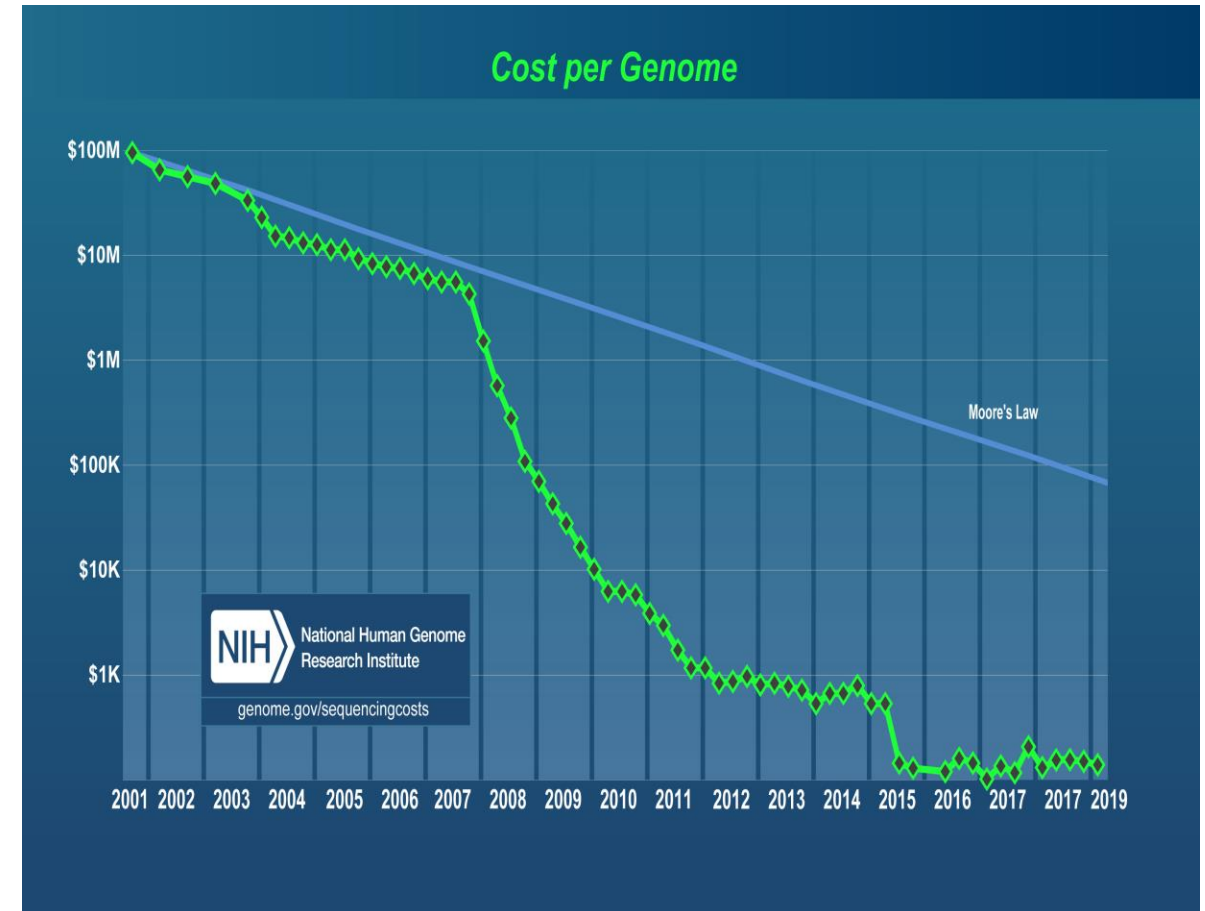


DNA Script = Enzymatic synthesis



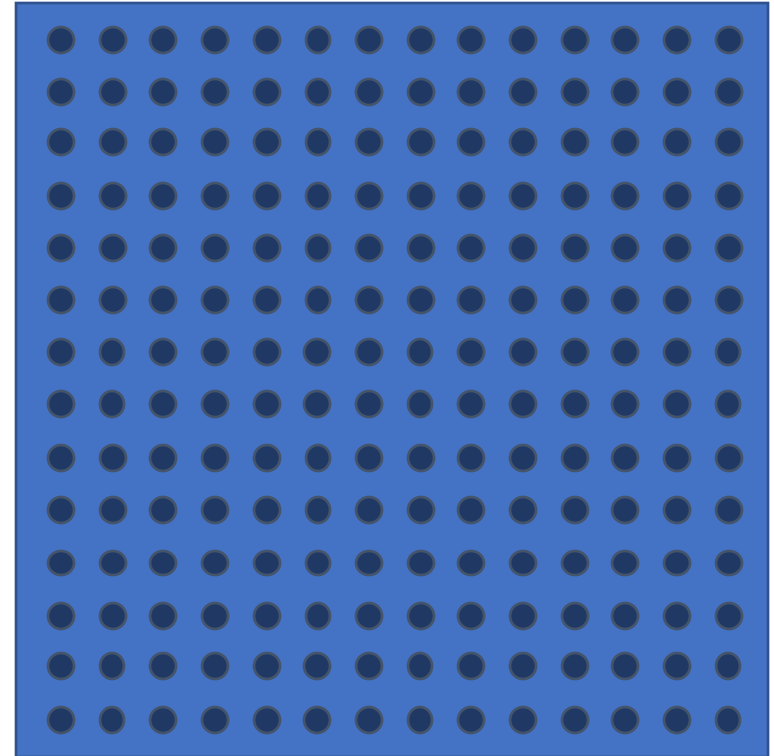
How to improve DNA synthesis ?

- 3 ways to increase throughput and to reduce costs:
 - Speeding-up synthesis
 - Miniaturization
 - Parallelization
- Same strategy which have been followed for sequencing technology 20 years ago



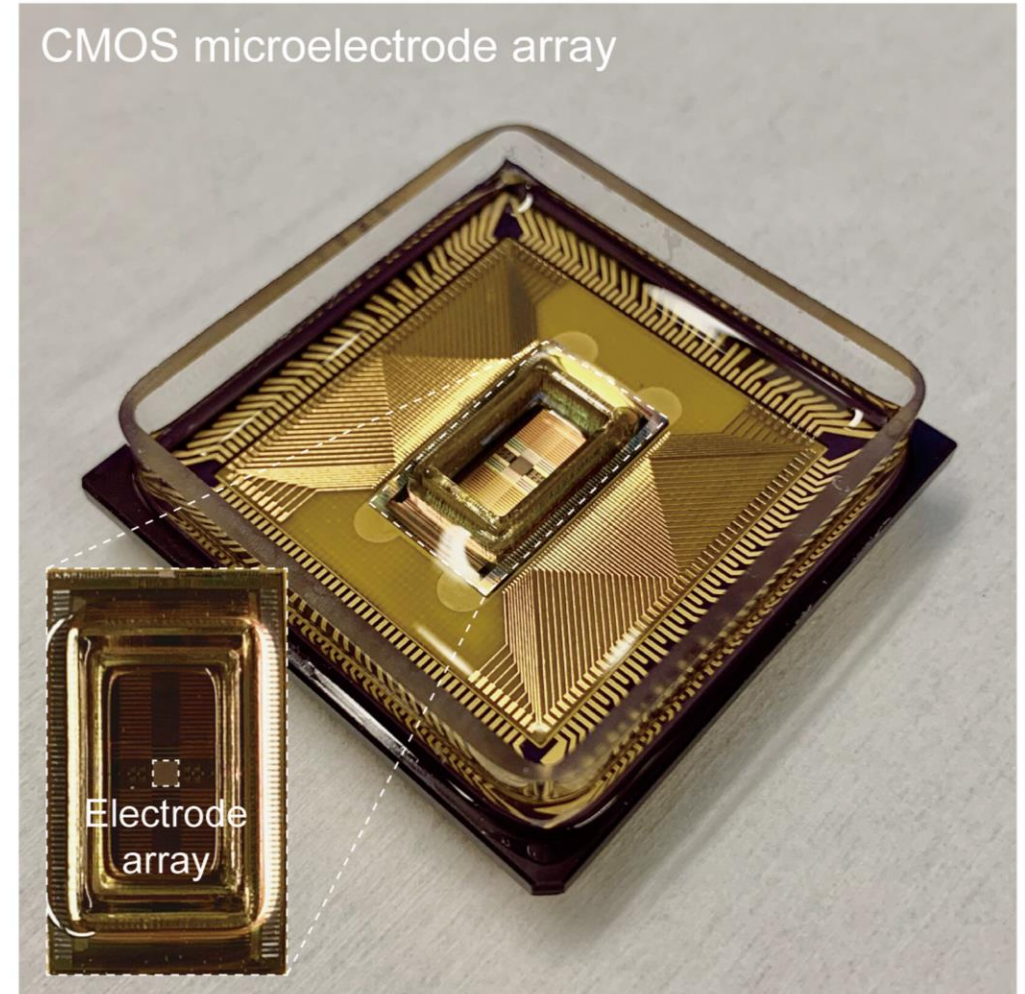
Miniaturization and Parallelization

- Large array of small electrochemical devices patterned on the top of semiconductor chips
- Each device is individually addressed to synthesize a specific pool of oligos
- Different possibilities to control chemical reactions



CMOS microelectrode array

Experimental CMOS chip for DNA synthesis. It features a dense array of microsites capable of locally controlling pH levels, thereby enabling or preventing chemical reactions at each microsite.
Credits: Donhee Ham Research Group / Harvard SEAS."



Acid composition control

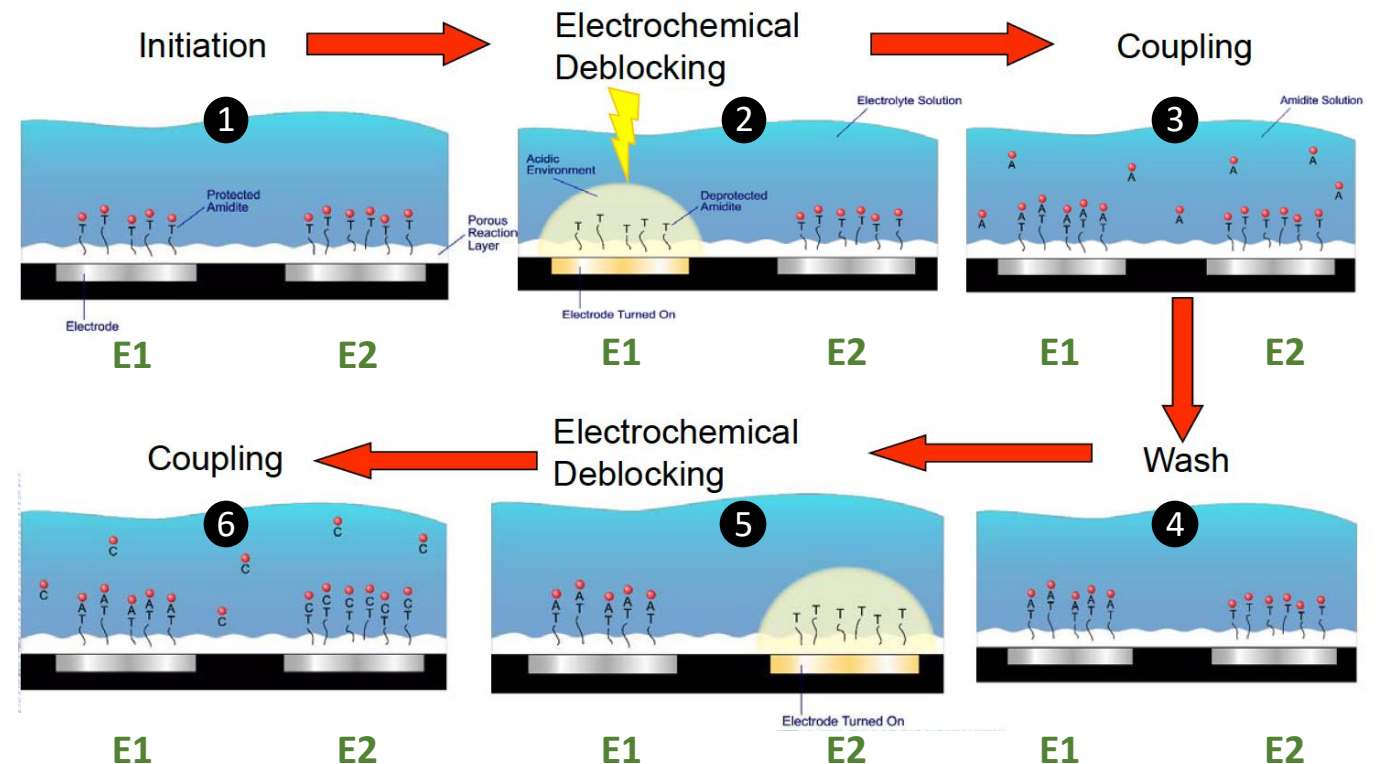
The acid composition is locally modified when electrodes are turned on

Genscript

- Largest chip: 92918 electrodes (diameter: 25 μm)

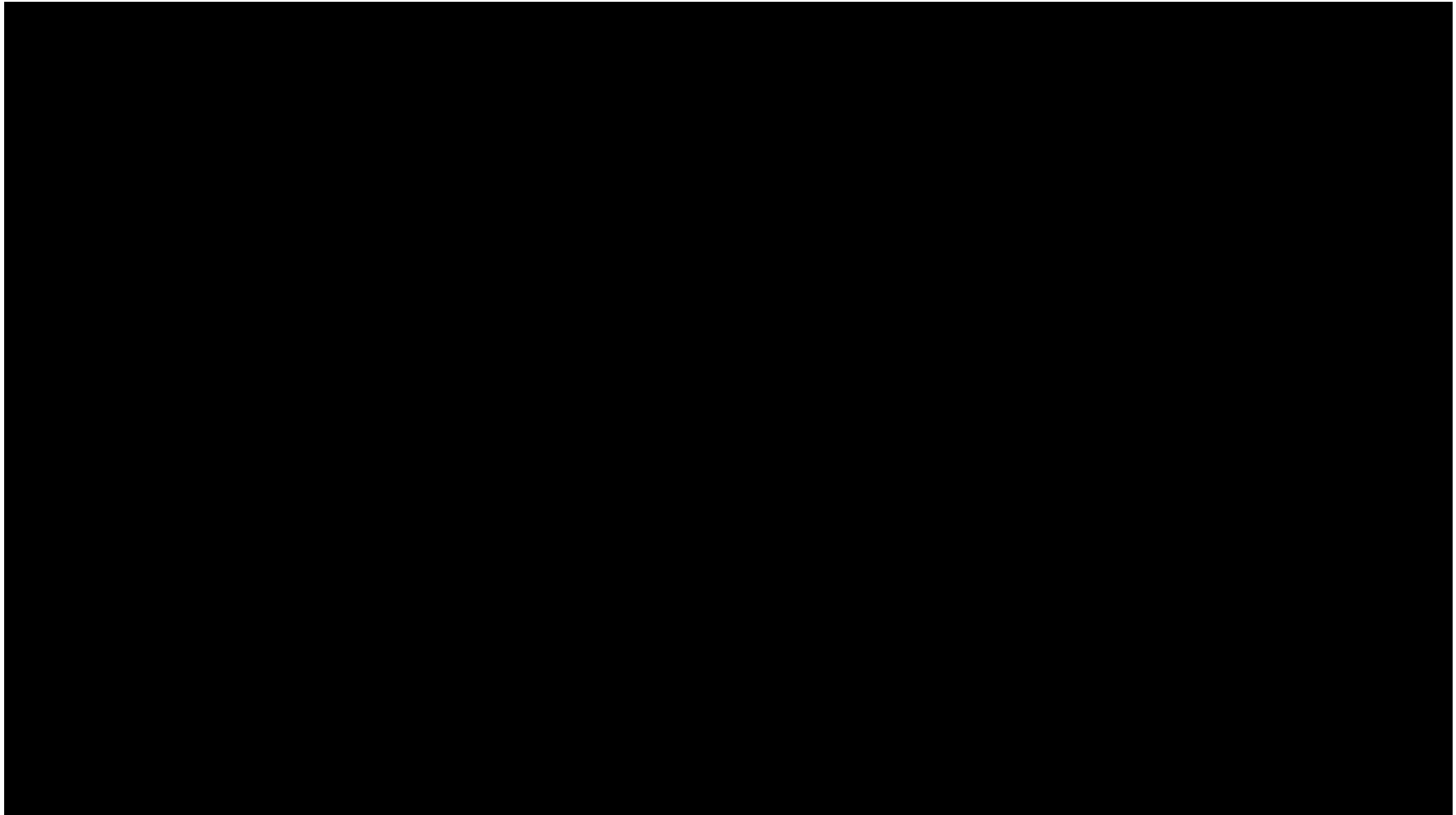
Twist Bioscience

- Roadmap : platform to synthesize 1 billions of oligonucleotides



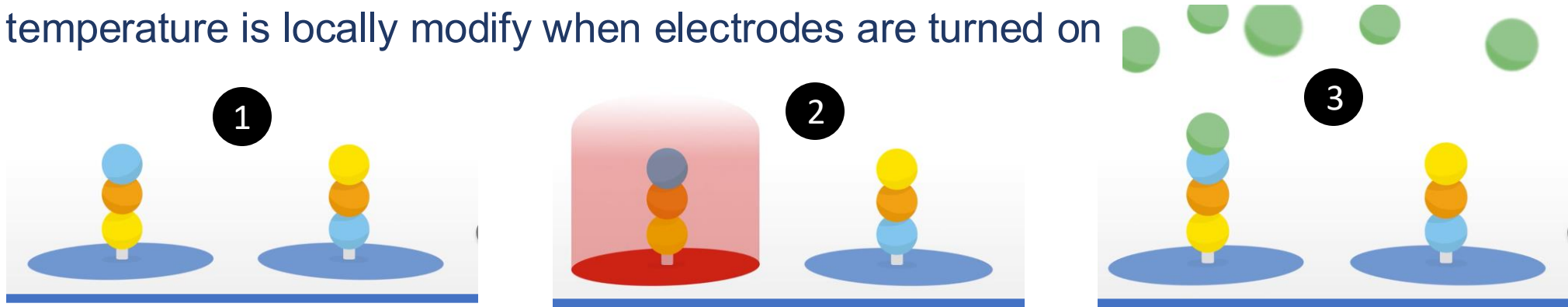
Source: M. Caraballo, *Oligo Pools: Design, Synthesis and Research Applications*, Webinar, 2018

GenScript



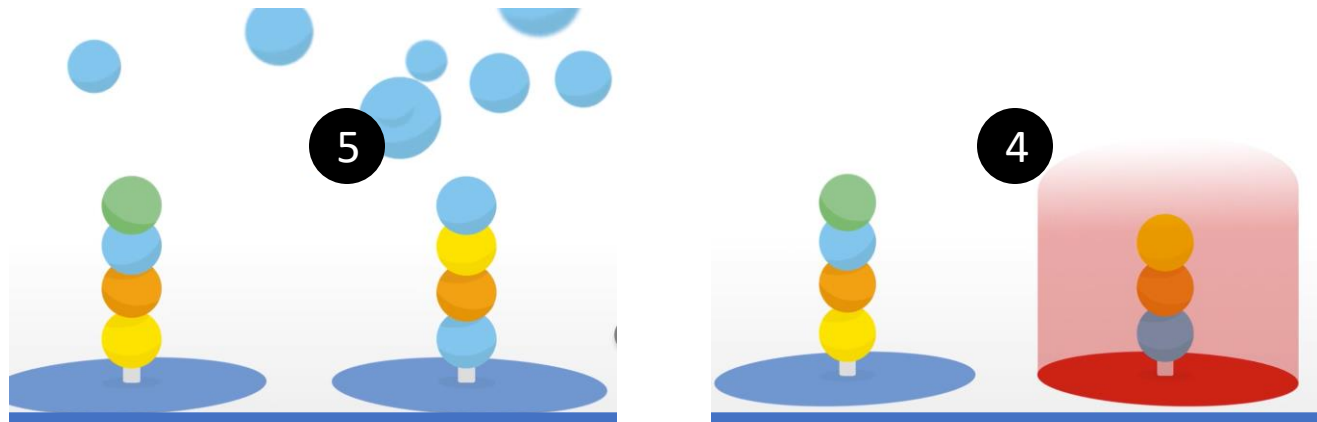
Thermal control

The temperature is locally modify when electrodes are turned on



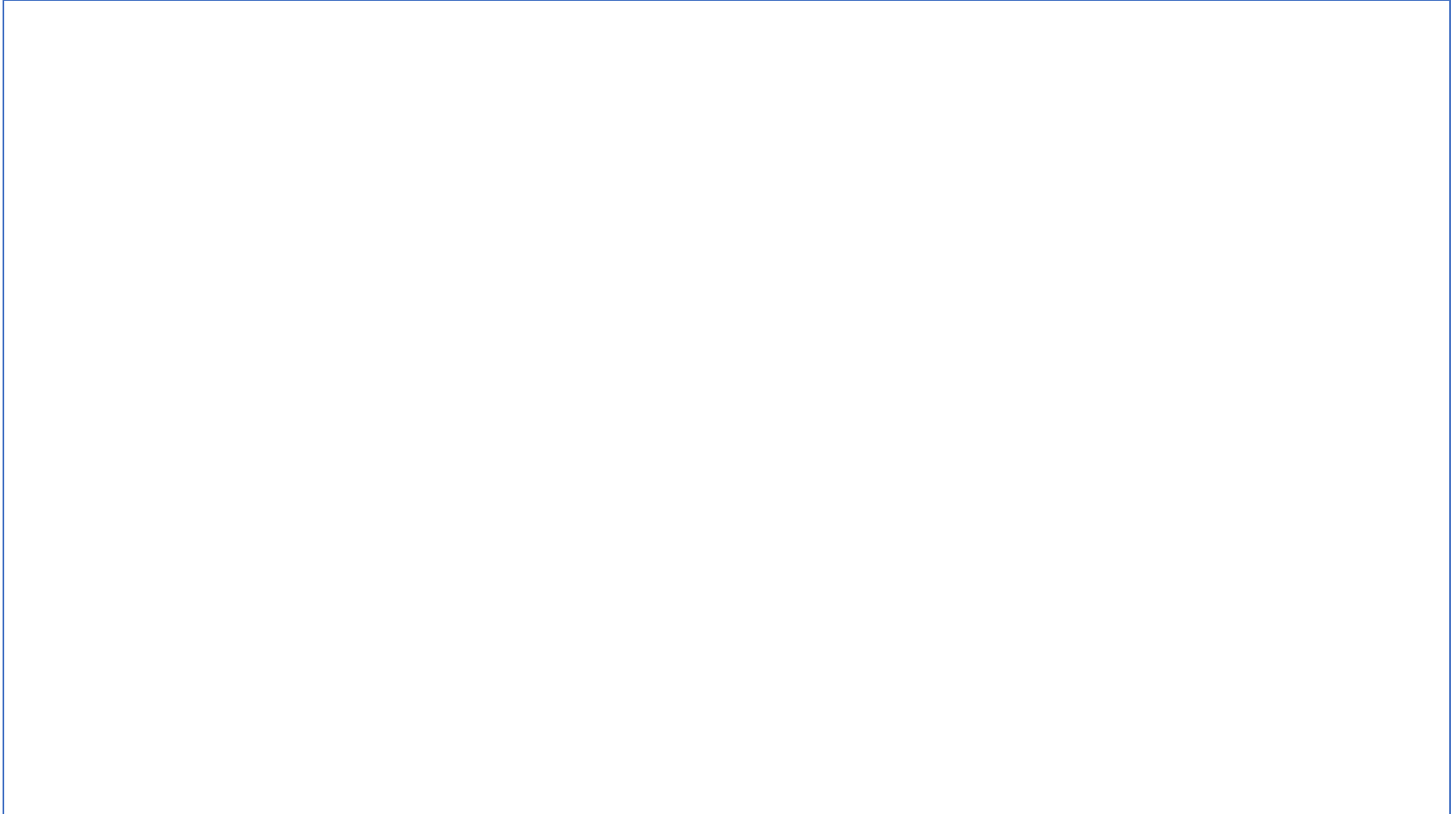
Evonetix company

- Can generate double-stranded DNA molecule
- Compatible with enzymatic synthesis



Source : <https://www.evonetix.com/technology/>

Evonetix



Recap

Synthetic DNA Synthesis

- Sequential process
- Production of oligonucleotides
 - Size < 300 nt
- Slow
- Costly
- Polluting process

Research axe

- Enzymatic synthesis
- Miniaturization / Parallelization
- Achieve faster chemistry

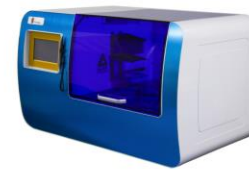
Synthesis Errors

- Oligos of bad size (shorter)
- Missing oligos

Principle of DNA data storage



00 → A
01 → C
10 → G
11 → T



A → 00
C → 01
G → 10
T → 11

Coding

Synthesis

Storage

Retrieval

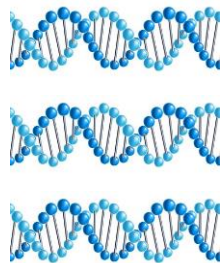
Sequencing

Decoding



010100101
101010011
010001010
001010010

ATGAGGACCG
TTAGGACCAG
GCCATATAAC
GGTTATATAA



ATGAGGACCG
TTAGGACCAG

010100101
101010011

