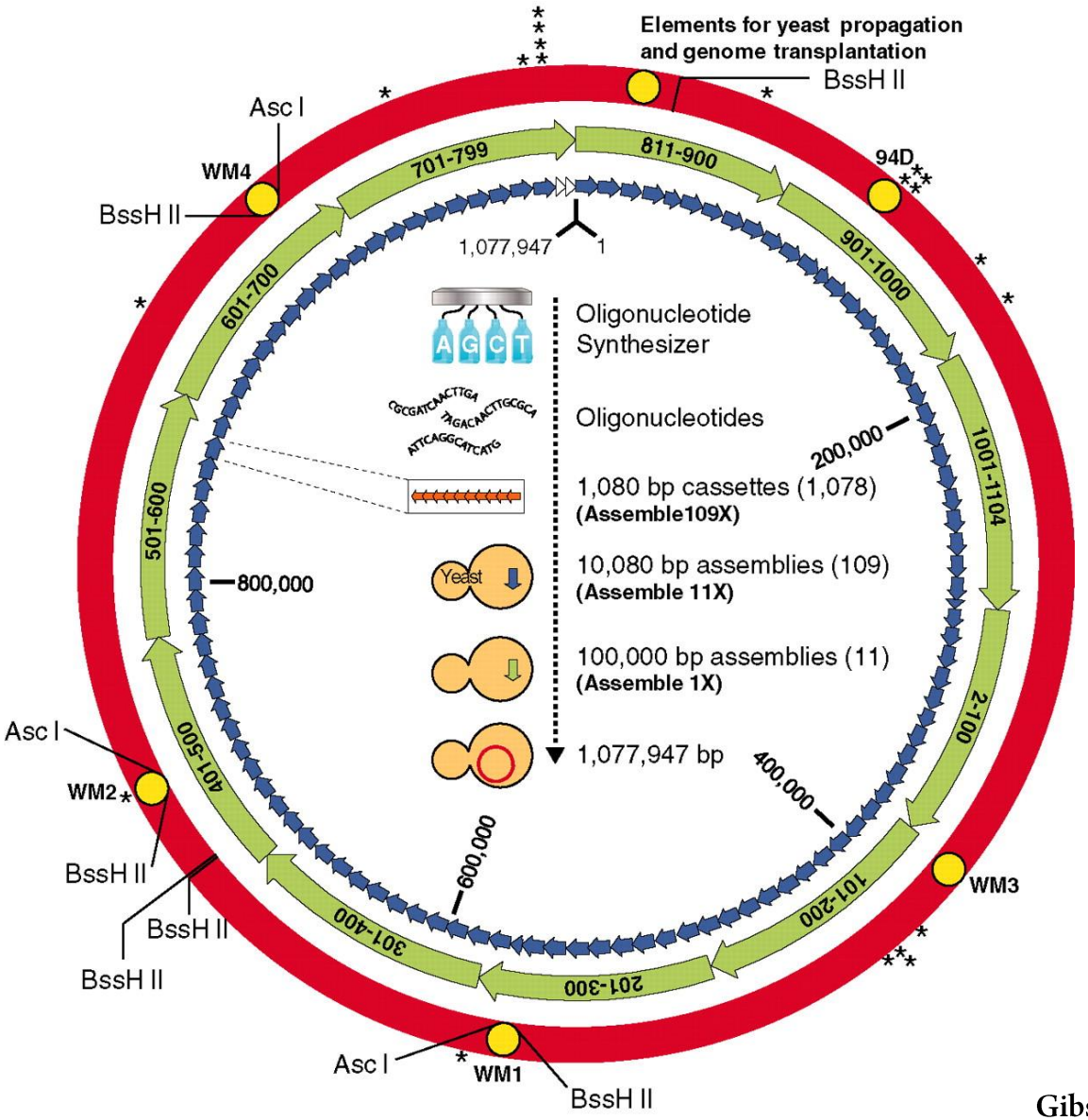
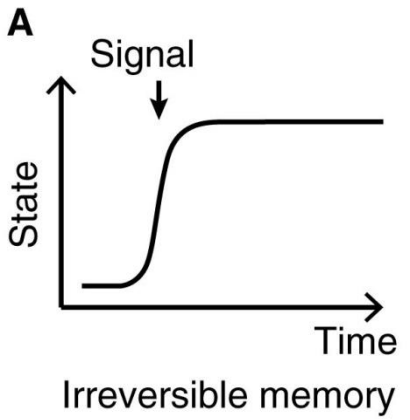
SYNTHETIC BIOLOGY



Gibson DG et al., Science 2010

SYNTHETIC BIOLOGY FOR *DE NOVO* ASSEMBLY OF MEMORY DEVICES – GLOSSARY

(Non-)volatile memory – synthetic memory circuits (do not) need active cellular processes to maintain their state

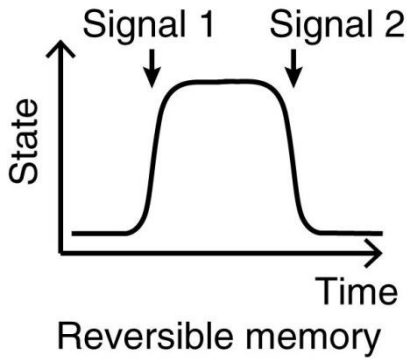


volatile memory > e.g. transcription-based

non-volatile m. > recombination-based

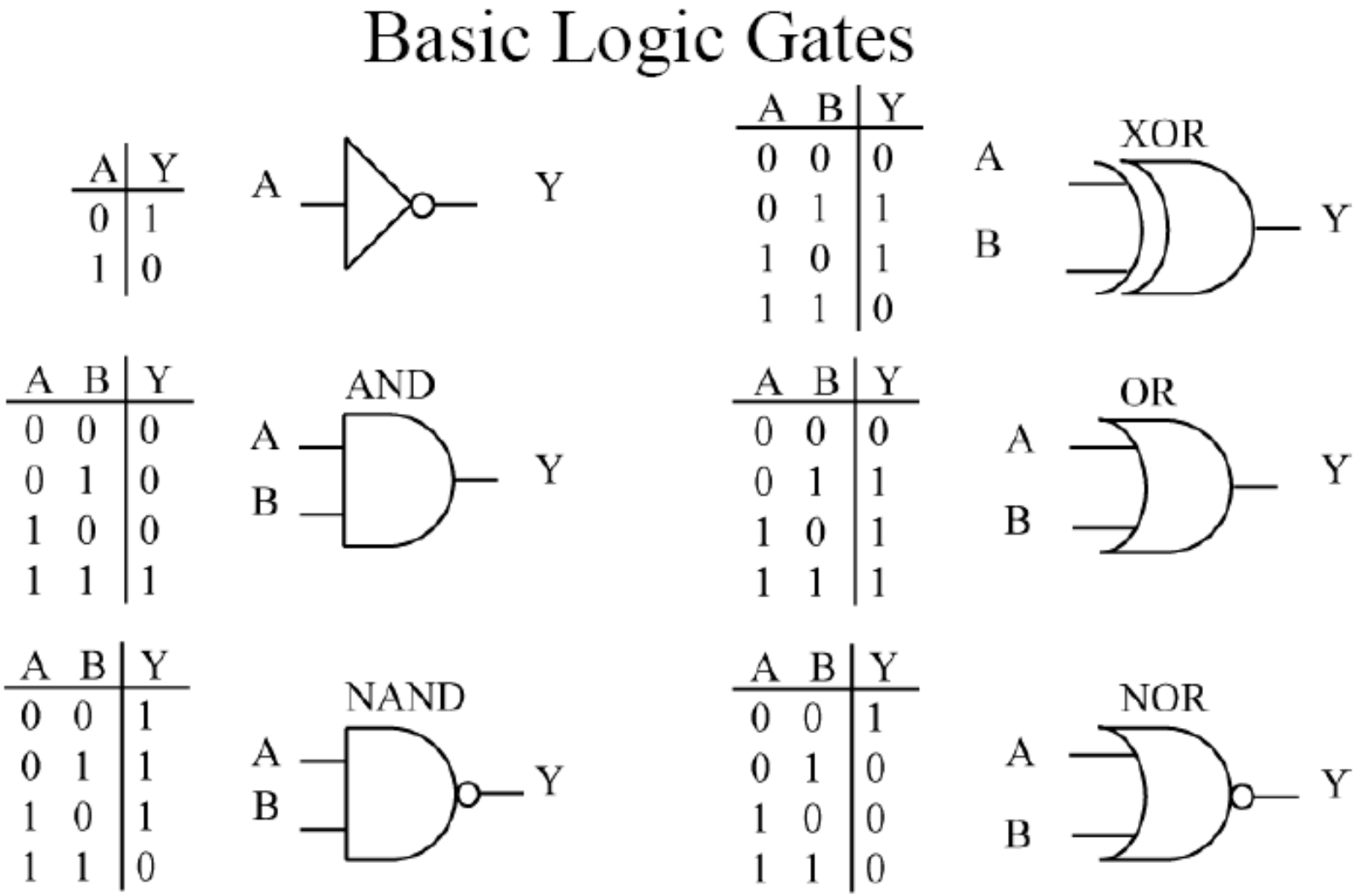
volatile memories are bistable

stochastic switching between states should be rare



Inniss MC & Silver PA, Curr Biol 2013

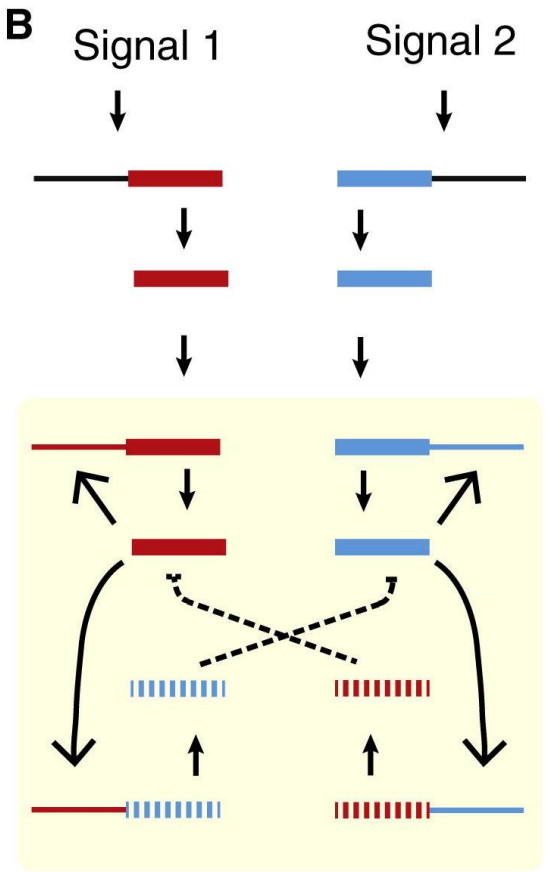
LOGIC GATES



http://www.kaituanui.com/elect-logic-1.htm

GENETICALLY ENCODED OPERATORS AND REGULATORS *IN VITRO*

In vitro



Nucleic acid hybridization

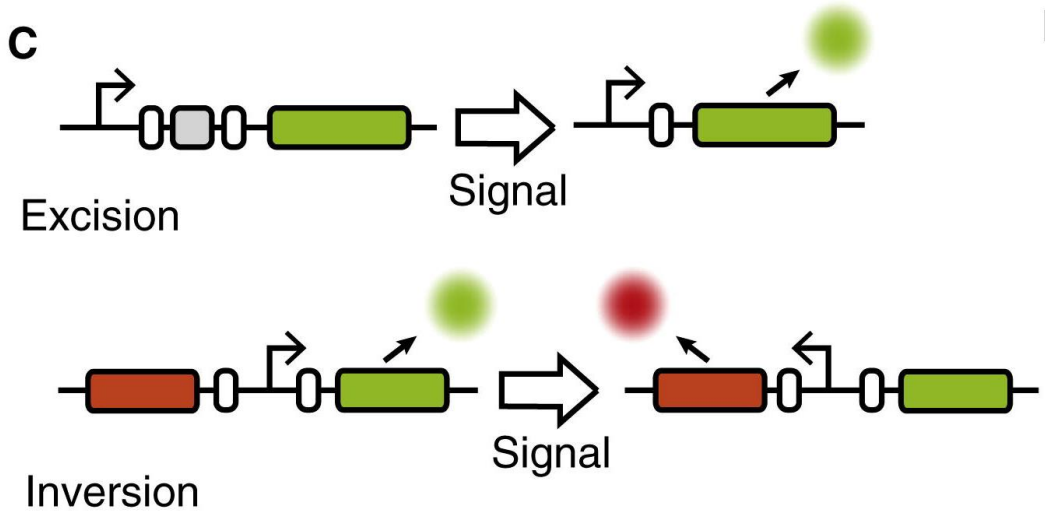
In vitro **memory circuits are typically composed of interlocking negative and positive feedback loops that form a bistable core**

Inniss MC & Silver PA, Curr Biol 2013

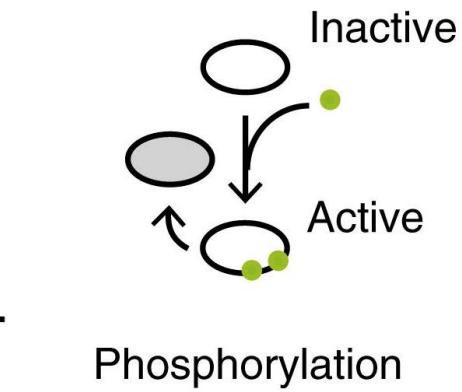
GENETICALLY ENCODED OPERATORS AND REGULATORS *IN VIVO*

In vivo

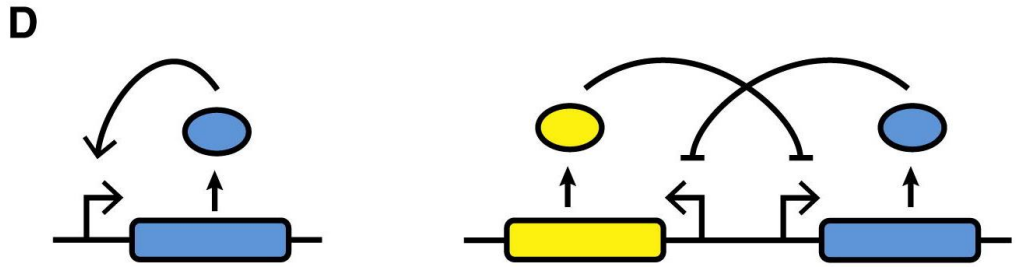
RECOMBINATION-BASED



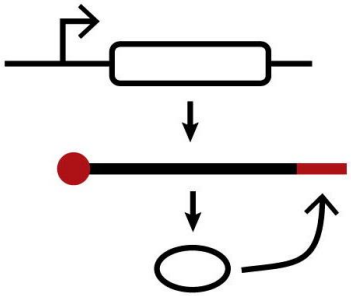
PROTEIN PHOSPHORYLATION



TRANSCRIPTIONAL FEEDBACK LOOPS

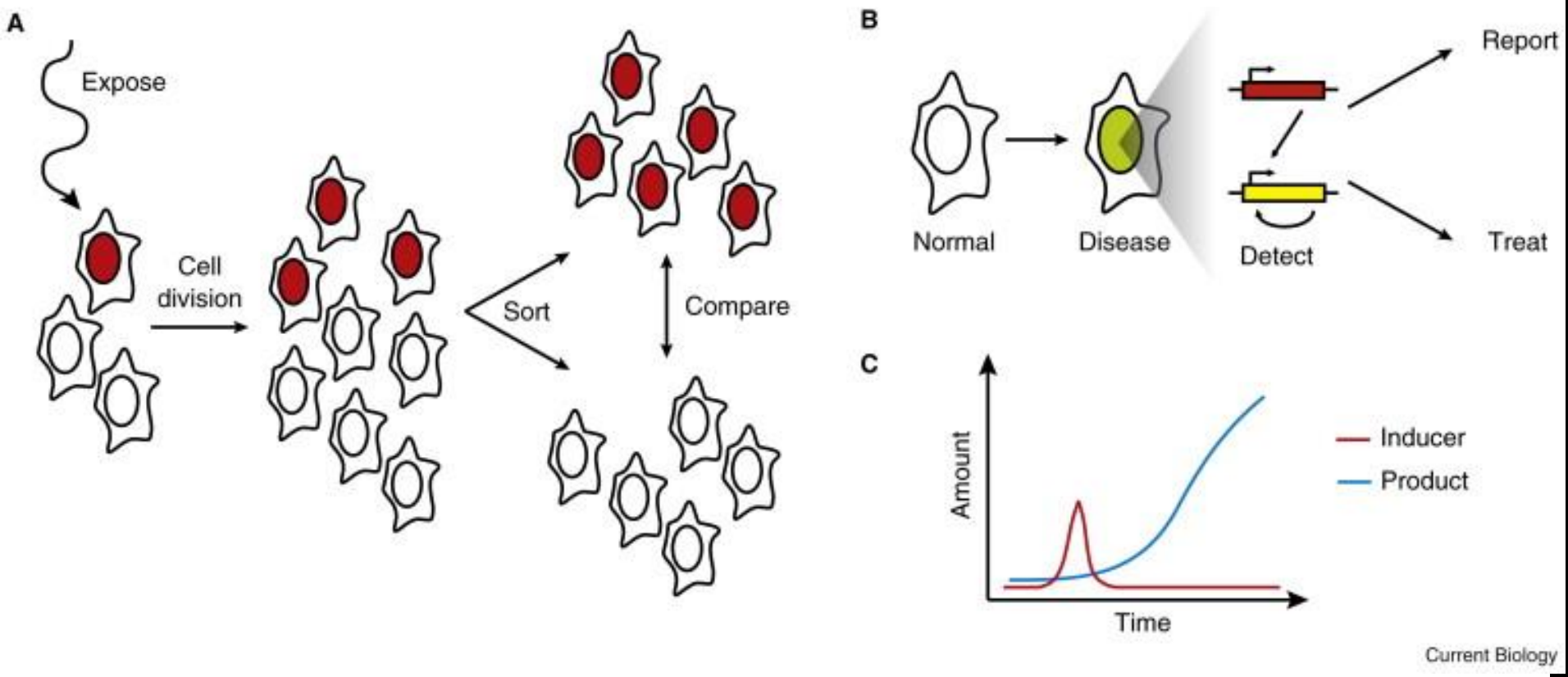


RNA EDITING



Inniss MC & Silver PA, Curr Biol 2013

APPLICATIONS OF SYNTHETIC BIOLOGY



Inniss MC & Silver PA, Curr Biol 2013

PAPER #1 – KOTULA JW ET AL., PNAS 2014

«PROGRAMMABLE BACTERIA DETECT AND RECORD AN ENVIRONMENTAL SIGNAL IN THE MAMMALIAN GUT"

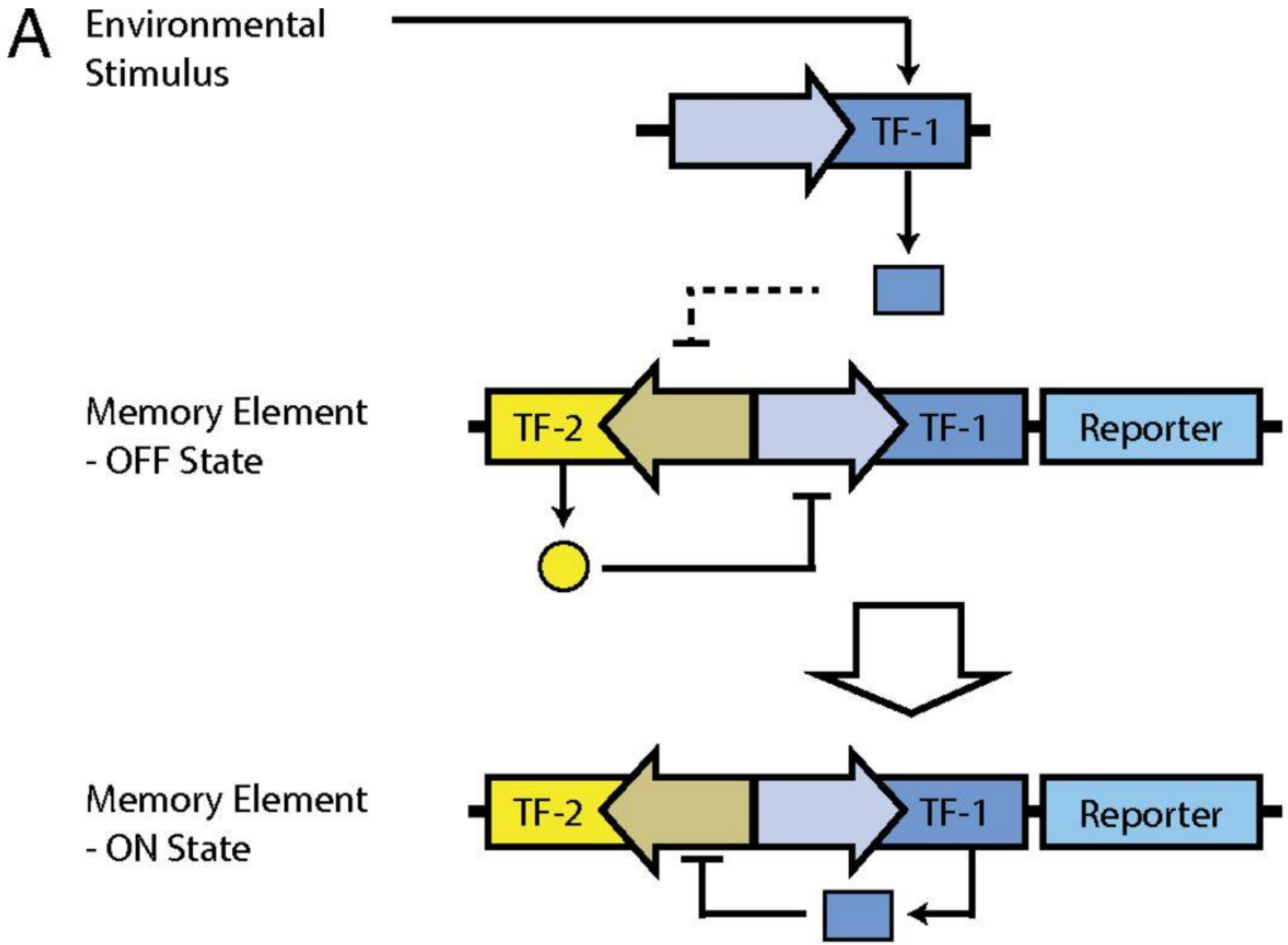
REPROGRAMMING THE MICROBIOME

The Scientific American

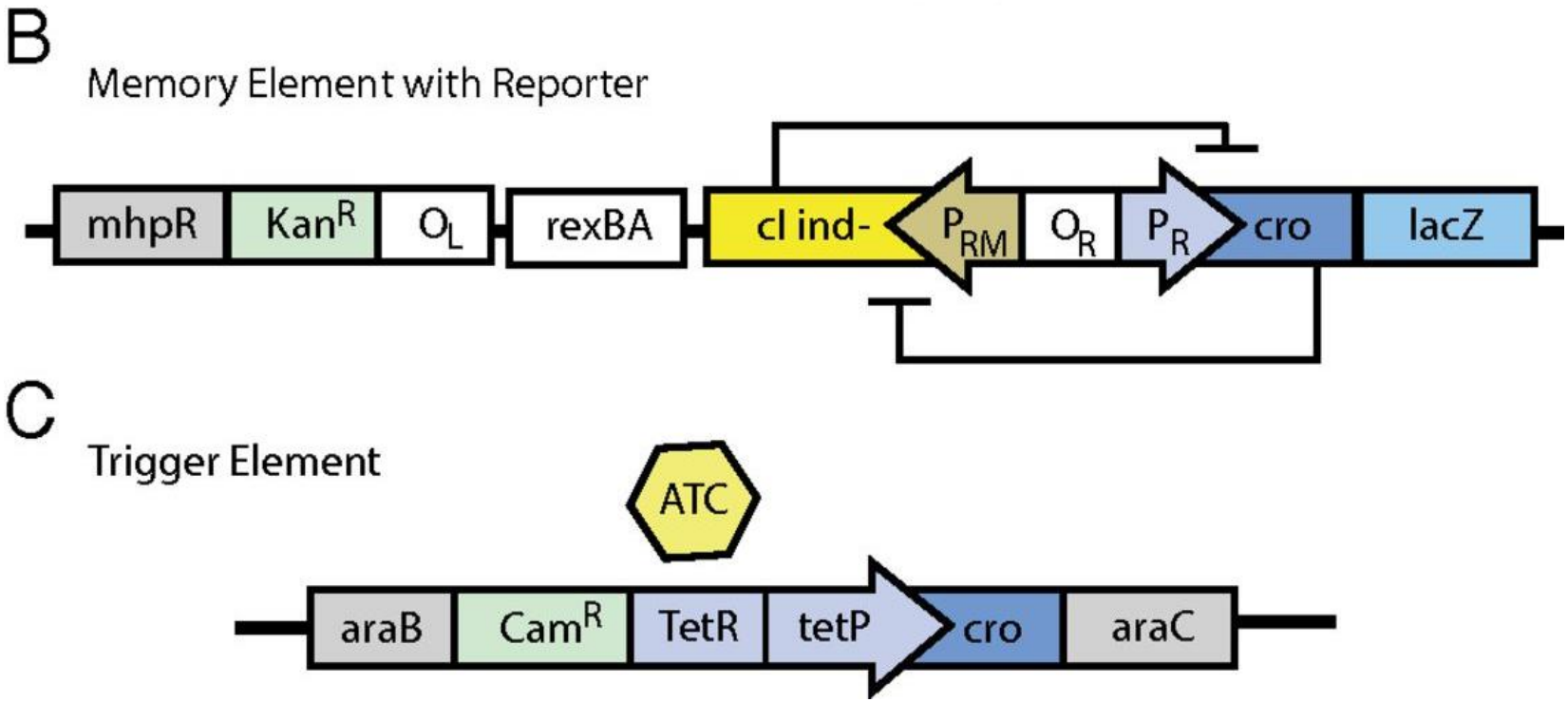
MEMORY ANALYSIS OF E. COLI

- **Modify E. coli with memory capabilities , transfer into mice, record events during antibiotic exposure**
- **Prerequisites**
  - inital «nonmemory» state should be highly stable, only failing as a result of mutation
  - «memory» state should also be highly stable
  - Chromosomal integration instead of plasmids to minimize loss of genetic information
  - Integrated elements should not impose high fitness burden on the host

TRIGGER/REPORTER SYSTEM WITH A TOGGLE SWITCH



cI/Cro//tetP-Cro SYSTEM FROM BACTERIOPHAGE LAMBDA

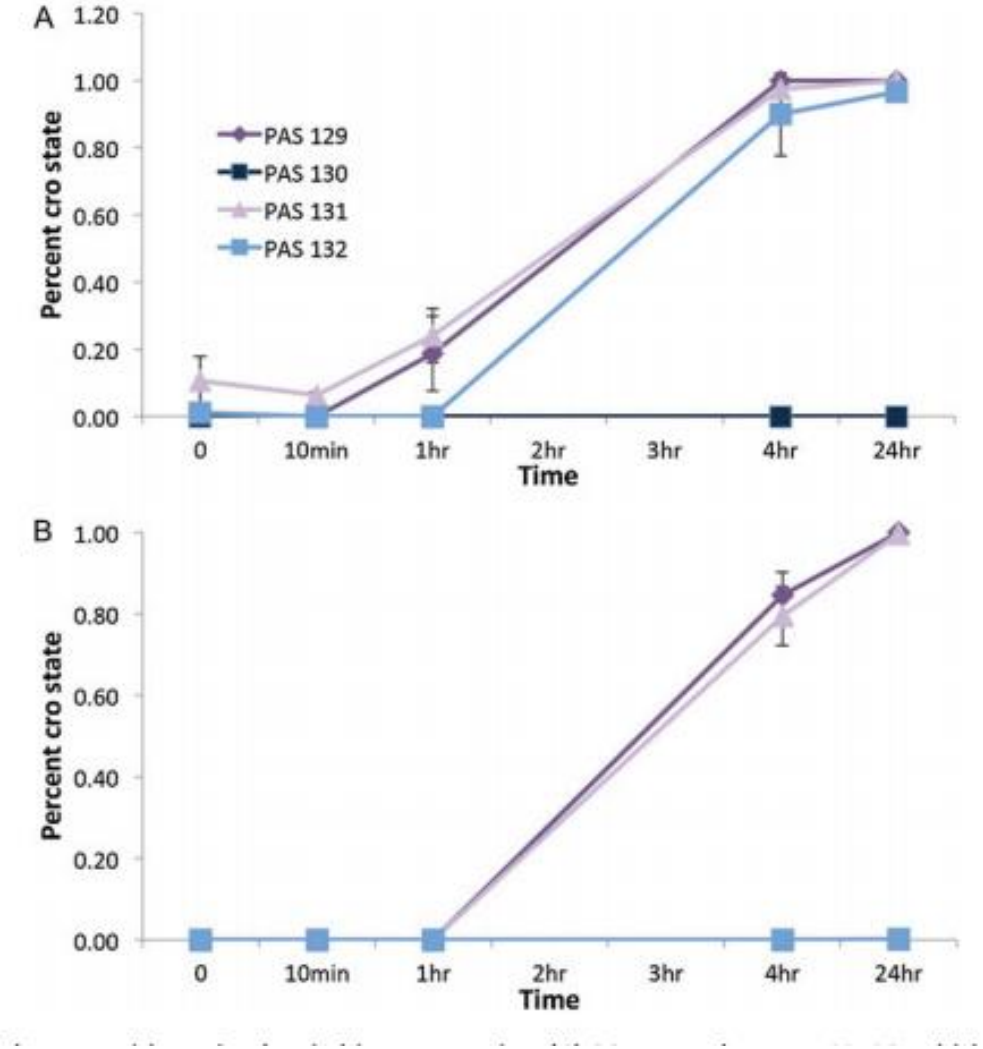


- repressor state of induction-deficient cI ind- lysogen only fails due to mutation
- little burden on bacterial host since only 100-200 monomers of cI ind- are present per cell, and, if activated < 1000 molecules of Cro
- Tn10 tetracycline repressor is particularly sensitive to anhydrotetracycline (ATC) >> 100 ng/mL ATC will cause full derepression of the promoter without inhibiting growth of tetracycline-sensitive E.coli

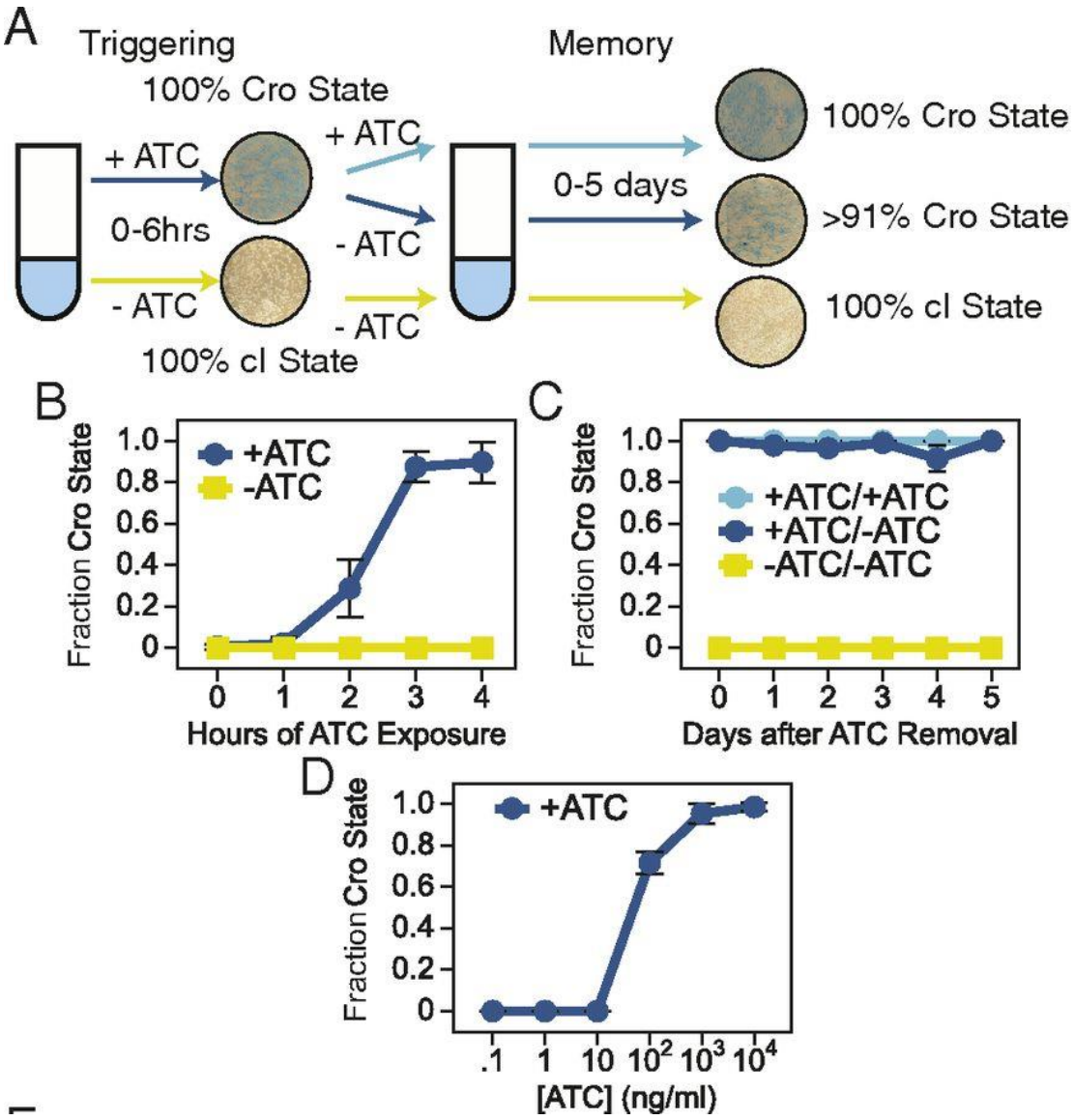
SELECTING THE OPTIMAL MUTANT

ATC

42°C

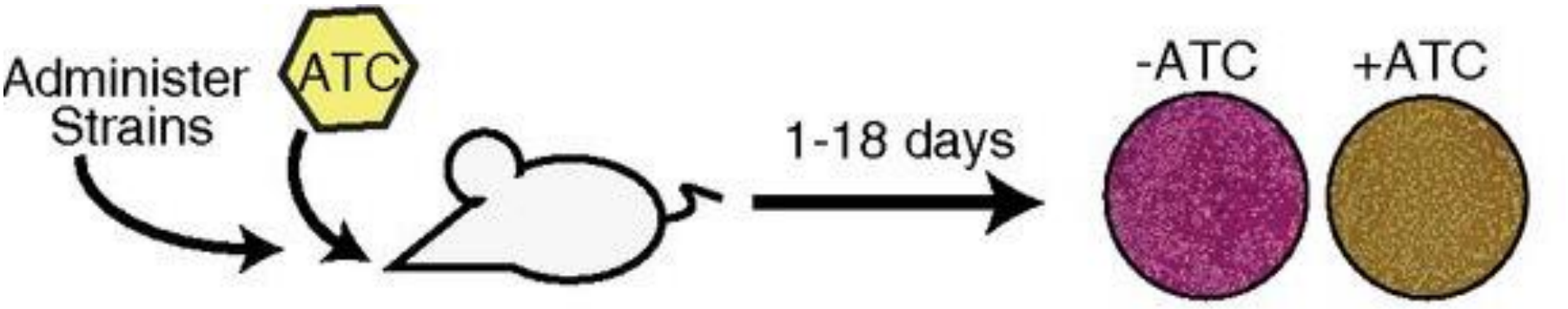


ATC EXPOSURE LEADS TO CI>CRO SWITCH FOR >5D *IN VITRO*

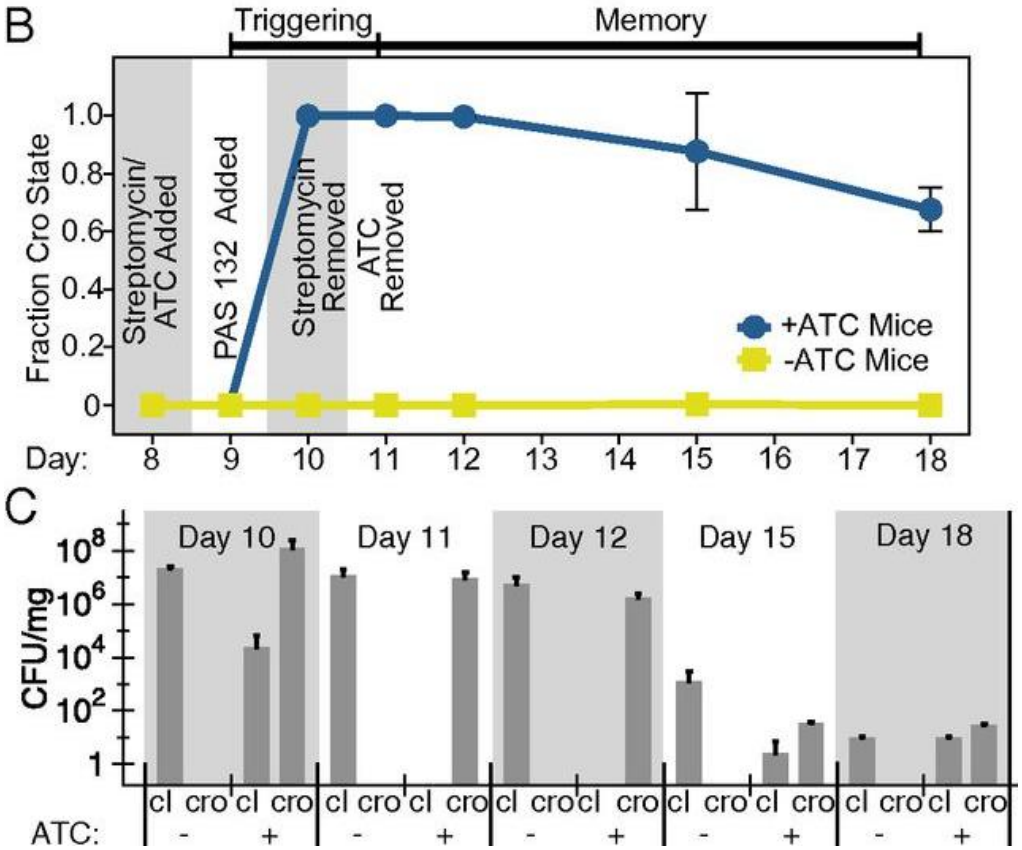


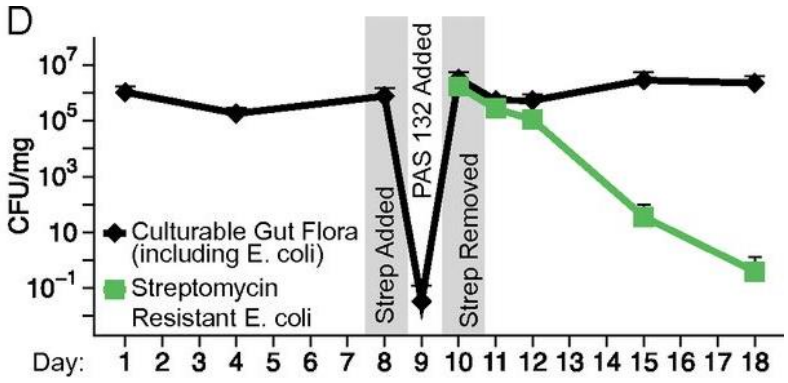
DETECTION OF ATC IN THE MAMMALIAN GUT

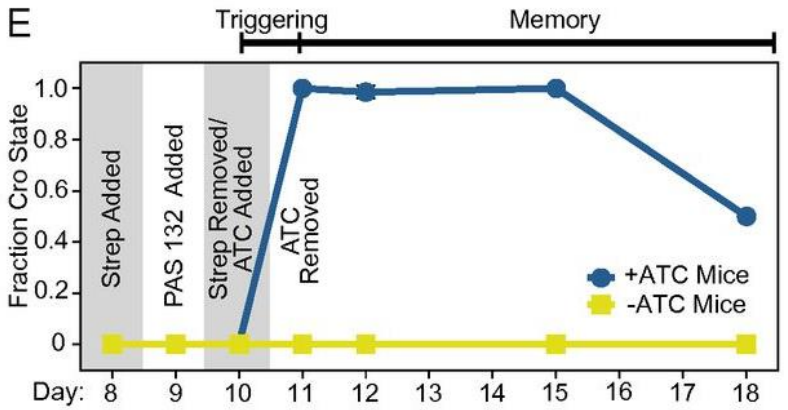
- **PAS132: mutated clone chosen for analysis**
  - Contains a mutation in rspL, conferring resistance to Streptomycin
- **Colonization of mice with E. coli^{PAS132}**
  - Female Balb/c mice received 10^7 bacteria per animal via oral gavage while receiving streptomycin in the drinking water (+/- ATC)



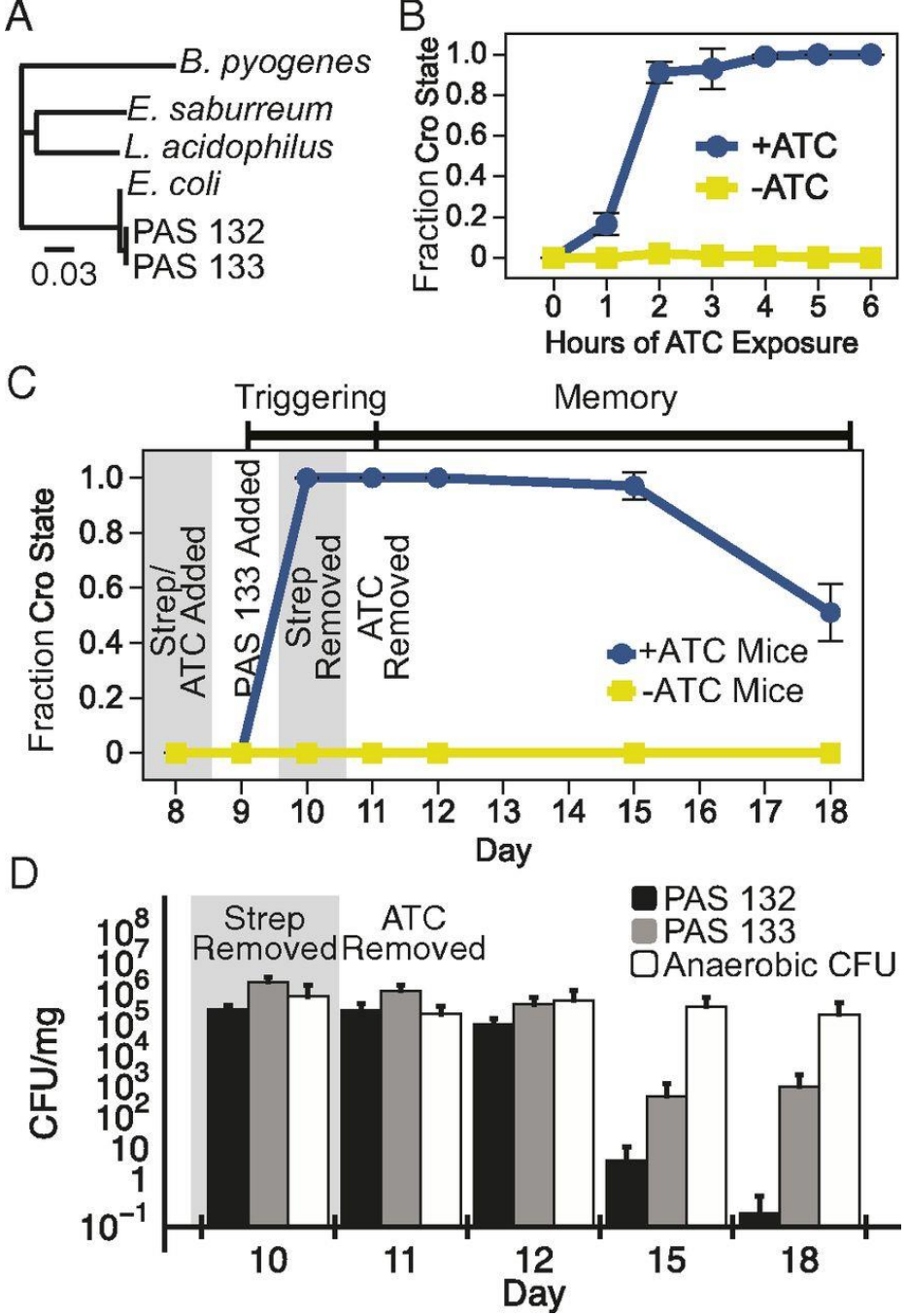
ATC-INDUCED LASTING CI>CRO SWITCHES *IN VIVO*







ATC-INDUCED LASTING CI>CRO SWITCHES *IN VIVO* IN AN UNCHARACTERIZED (E.COLI 16S RIBOSOME CONTAINING) COLIFORM BACTERIUM

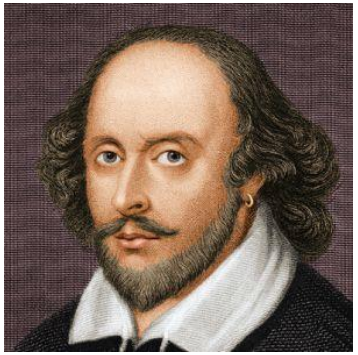


CONCLUSIONS

- **Stable (& bistable) expression system in 2 different E.coli (one as yet uncharacterized) strains to record exposure to antibiotic treatment**
- **Lasting recordings over hours to days (and multiple cell cycles) with no significant fitness burden and no observed «spontaneous», stochastic switching of states**
- **<> FLP based recombination systems are non-volatile and may be leaky due to constitutive low expression of FLP recombinase**
- **Stable cI/Cro system could be completely transformed into E. coli and may further have stabilizing function on the transcripts**
- **May lead to design of efficient, probiotic bacteria for diagnostics or treatment**

PAPER #2 – GOLDMAN N ET AL., NATURE 2014

"TOWARDS PRACTICAL, HIGH-CAPACITY, LOW-MAINTENANCE INFORMATION STORAGE IN SYNTHESIZED DNA"

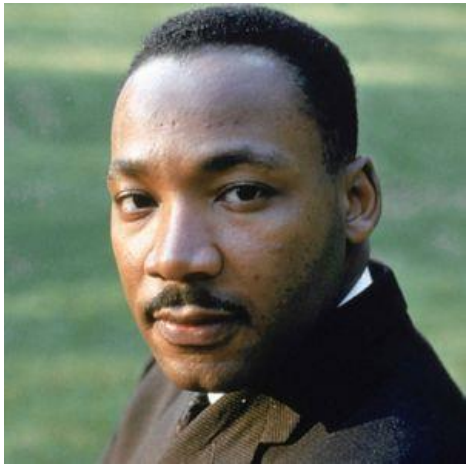


All 154 of Shakespeare's sonetts (ASCII «.txt»)

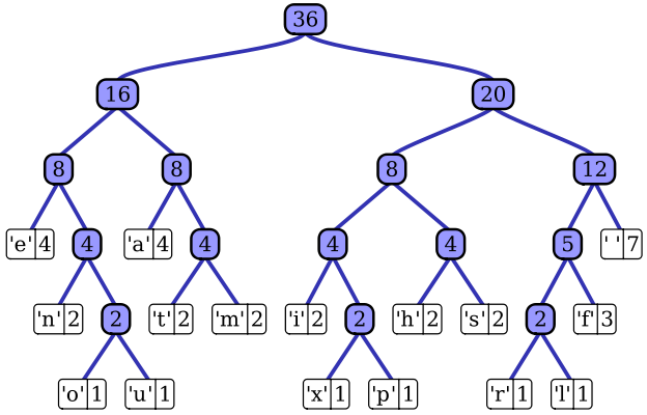
Watson&Crick, Nature 1953 («watsoncrick.pdf»)

>> 757'051 bytes

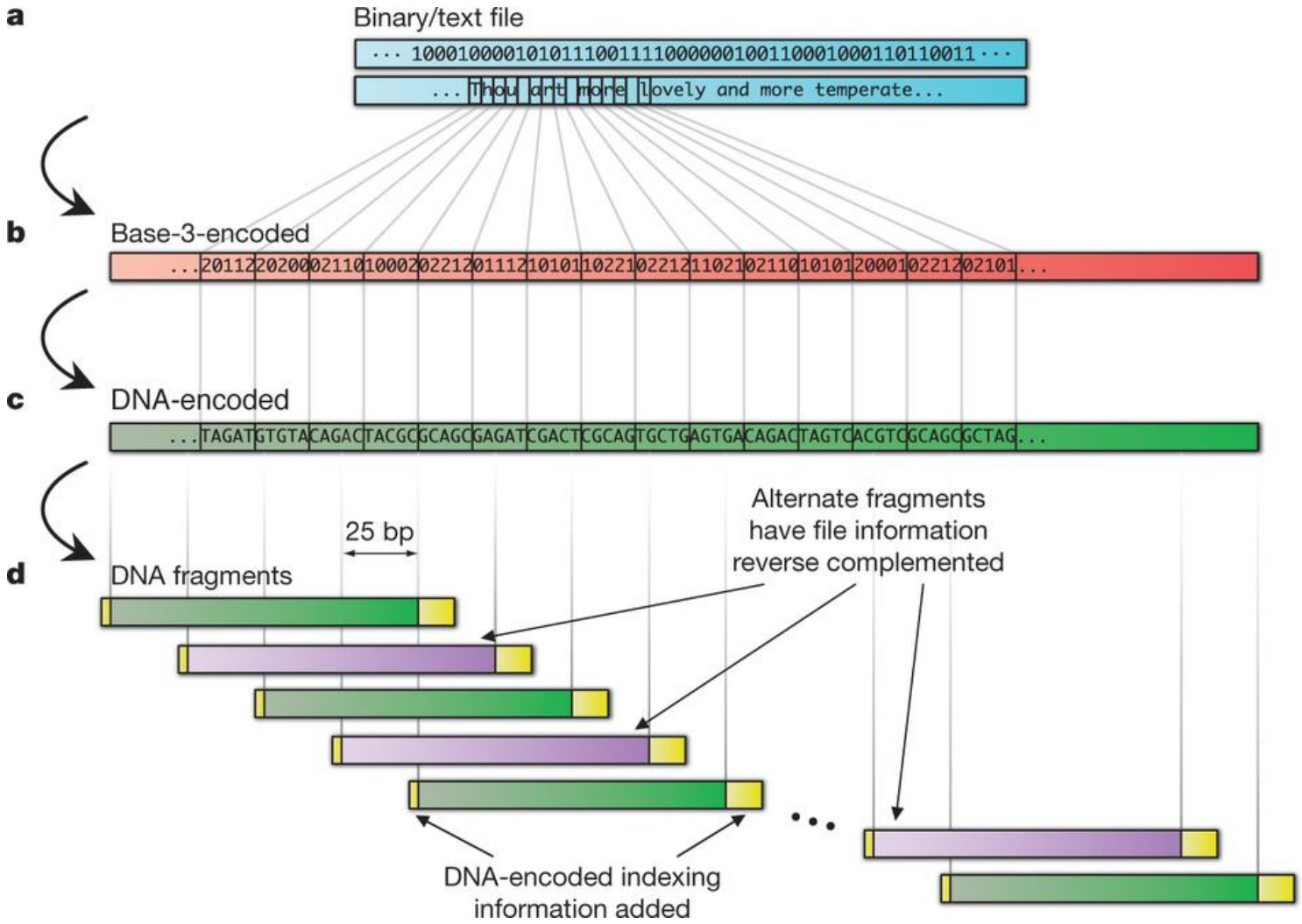
«EMBL-EBI.jpg»



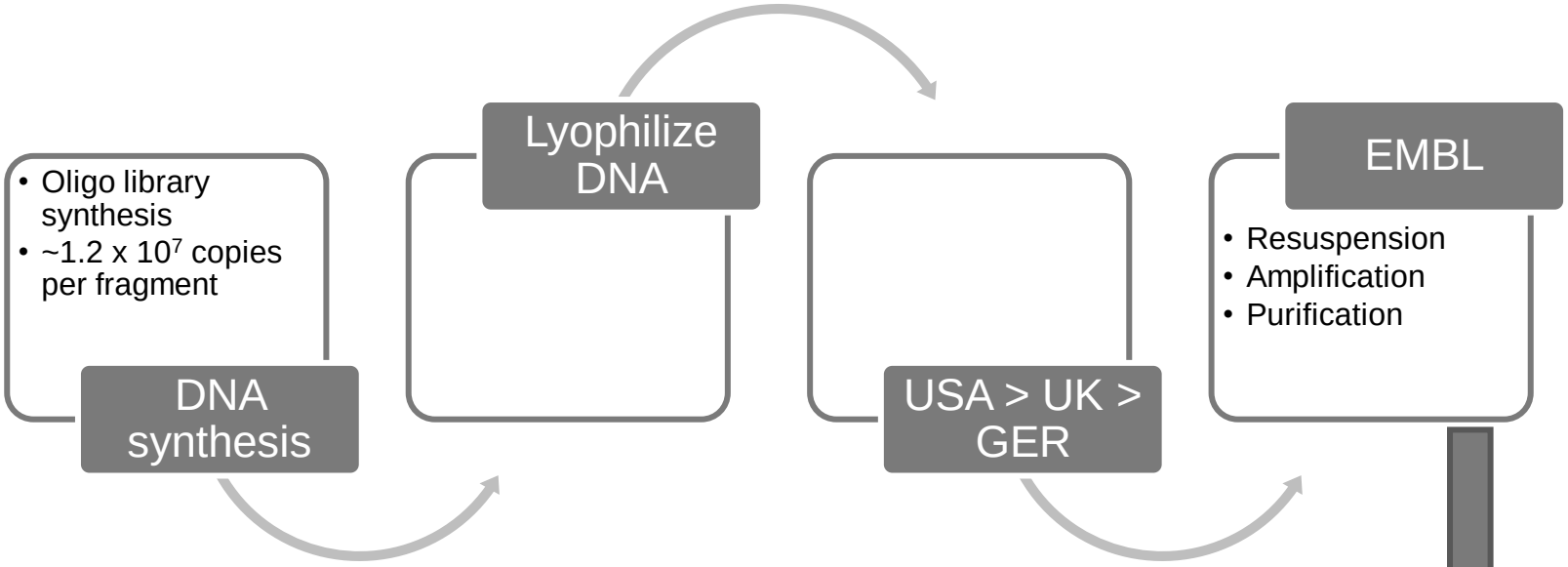
26s, «ihaveadream.mp3»



Huffmann code (ASCII «.txt»)



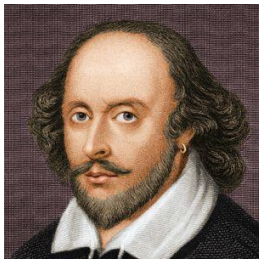
- 4-fold redundancy per DNA fragment
- Reverse complements of each fragment to minimize data loss
- total of 153,335 strings of DNA, each comprising 117 nt (100 nt data + 17 nt barcoding)
- No homopolymers

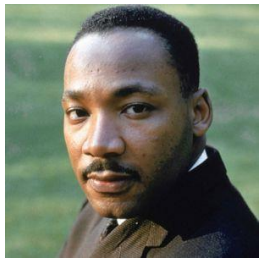


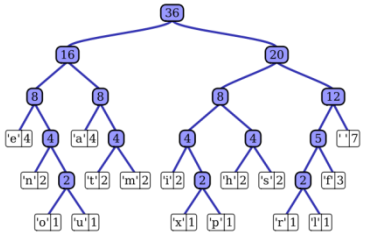
Sequencing on
Illumina HighSeq
In «paired-end» mode

79.6 x 10^6 read-pairs of 104 bases in length

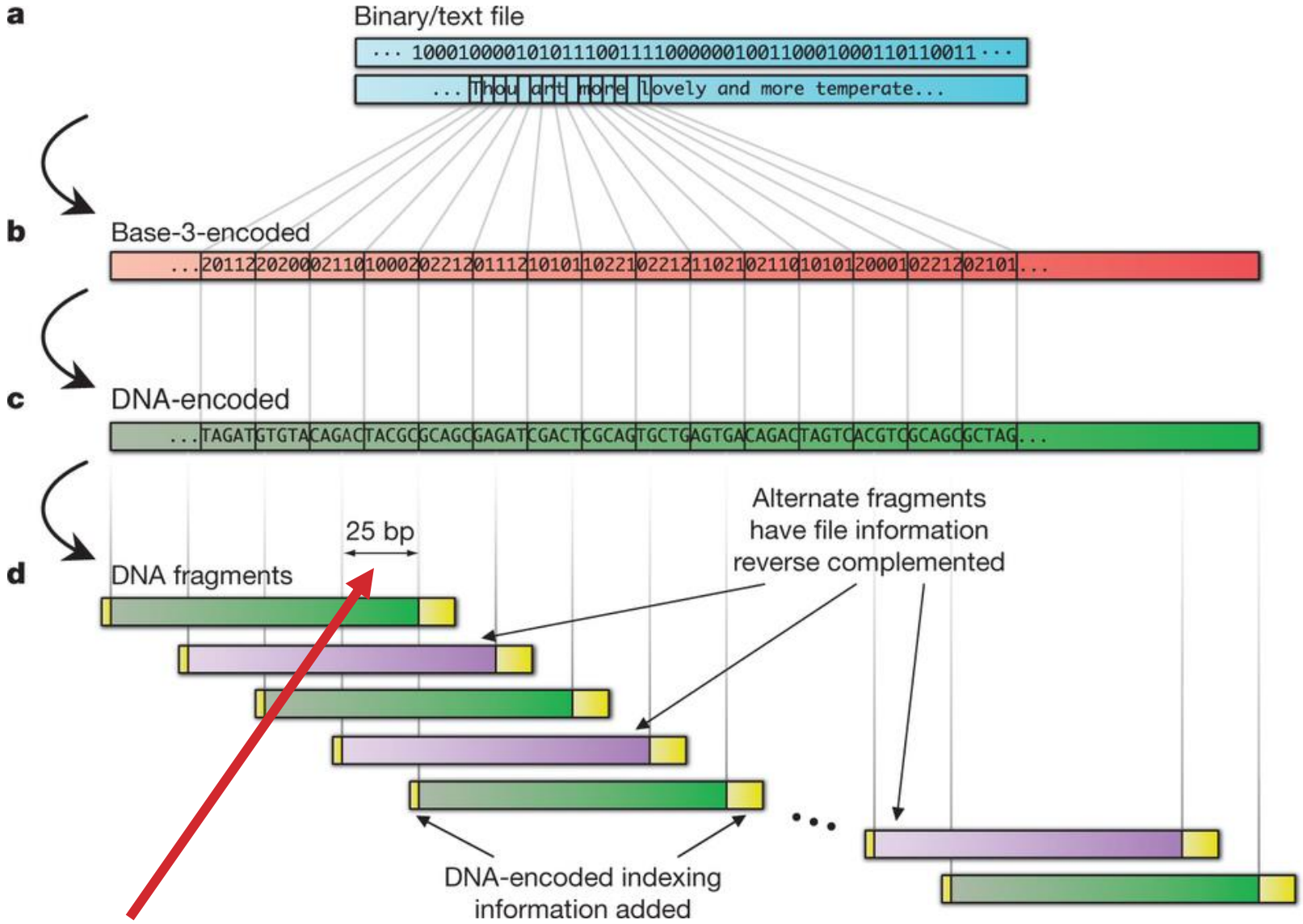
full-length reconstruction *in silico*





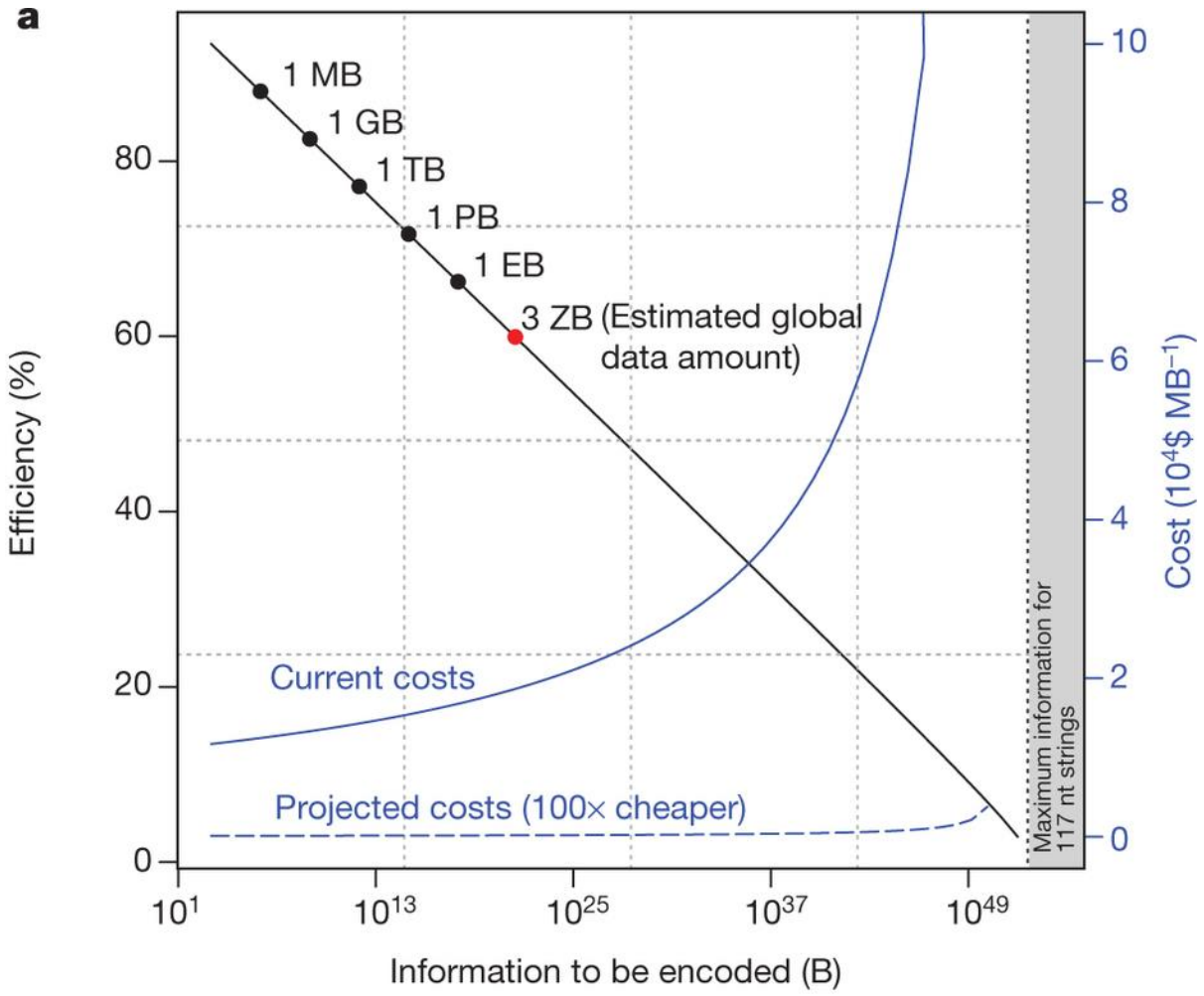


TROUBLESHOOTING



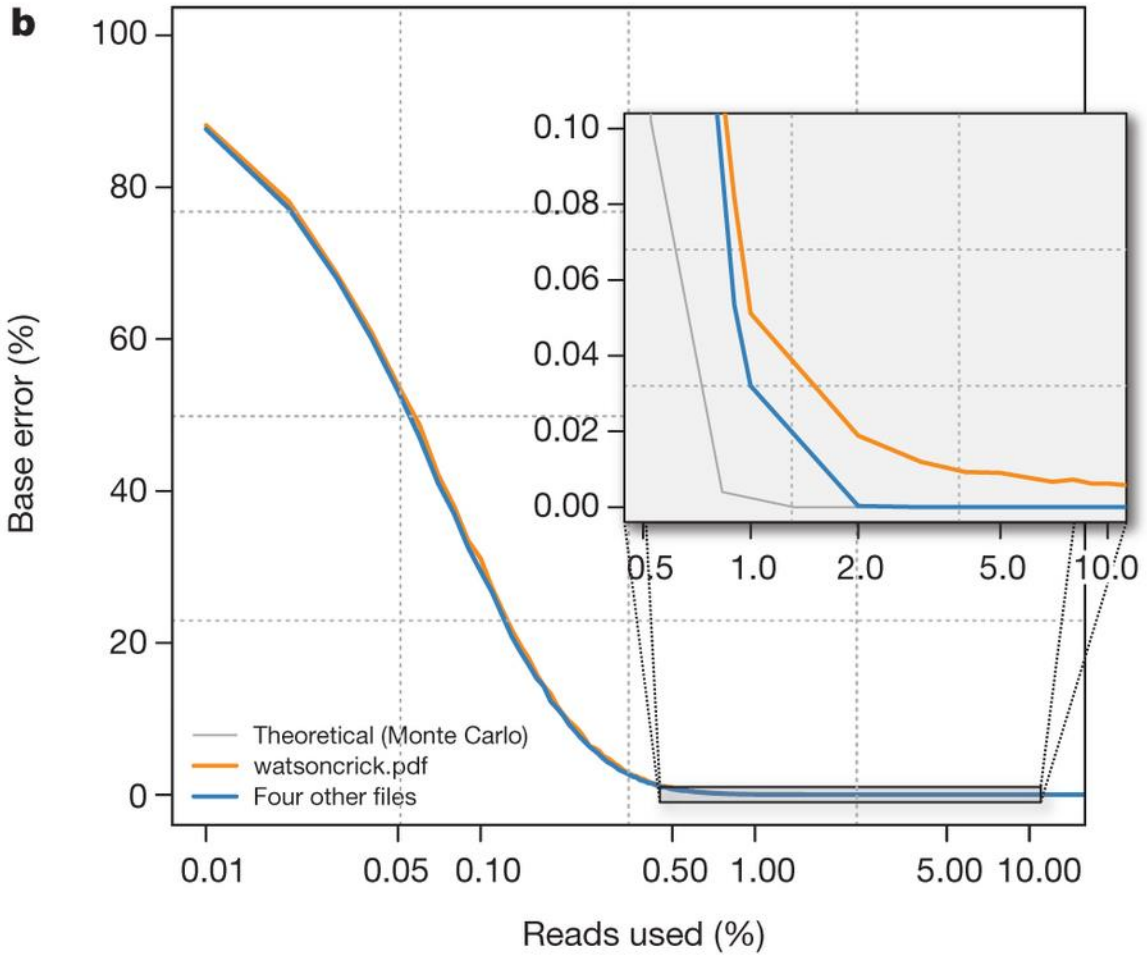
2x25bp could not be assigned > manual filling through neighbor comparison > «100% correctly aligned sequences»

CONCLUSIONS & OUTLOOK – ENCODING LARGER DATA VOLUMES



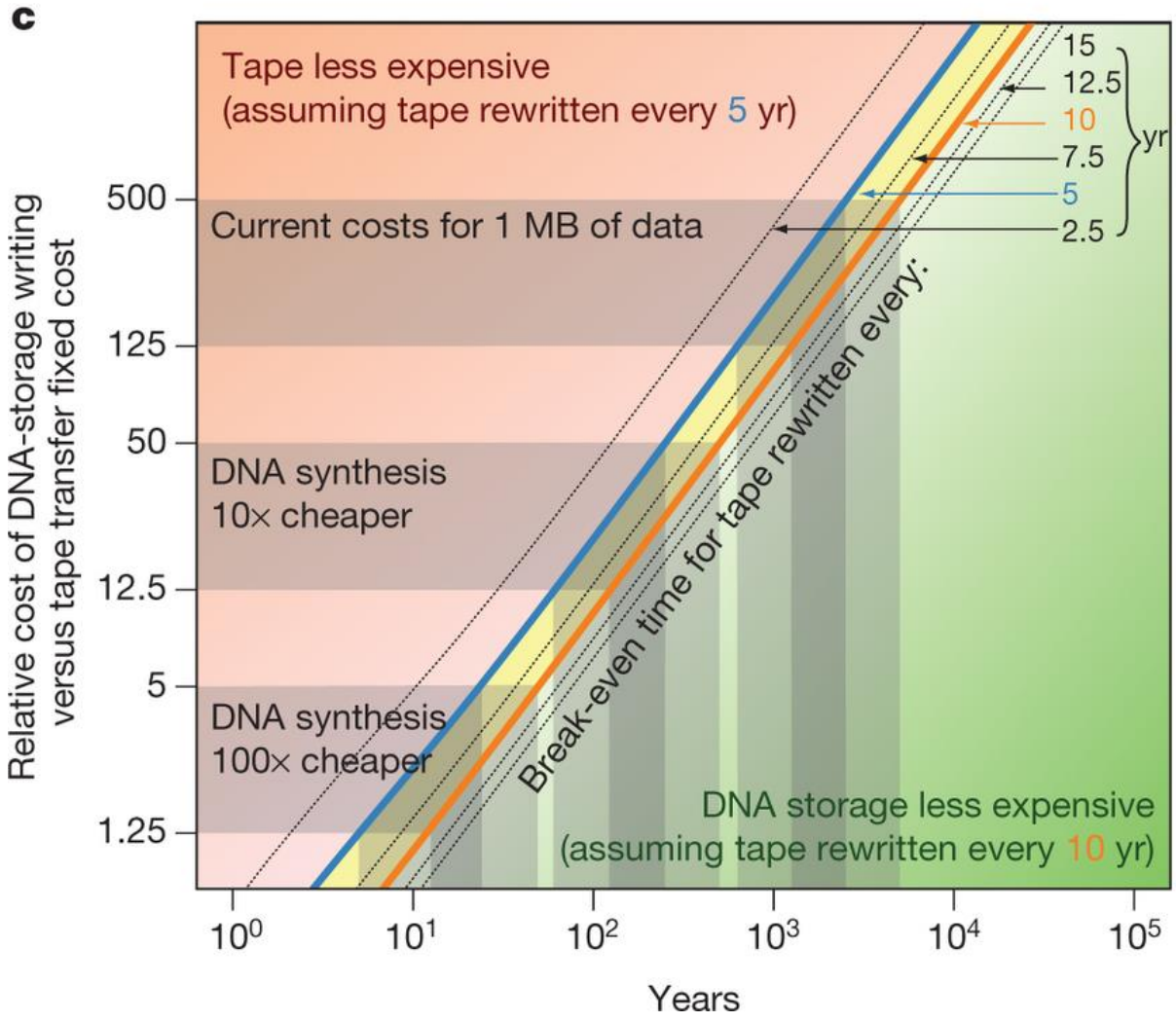
Indexing DNA grows logarithmically to the encodable data, while total amount of synthetized DNA for coding grows sub-linearly

CONCLUSIONS & OUTLOOK – READ ERRORS AS A FUNCTION OF COVERAGE



- Error is zero when ≥2% of the reads are used - taken from the originally obtained 79.6×10^6 read-pairs of all but "watsoncrick.pdf"
- Manual correction of «watsoncrick.pdf» leads to a minimal error rate of 0.0036%

CONCLUSIONS & OUTLOOK – COST EFFECTIVENESS



CONCLUSIONS & OUTLOOK – COST EFFECTIVENESS

- **Costs can further be decreased by lower sequencing coverage (now mean sequencing coverage ~ 1’308x)**

CASTOR

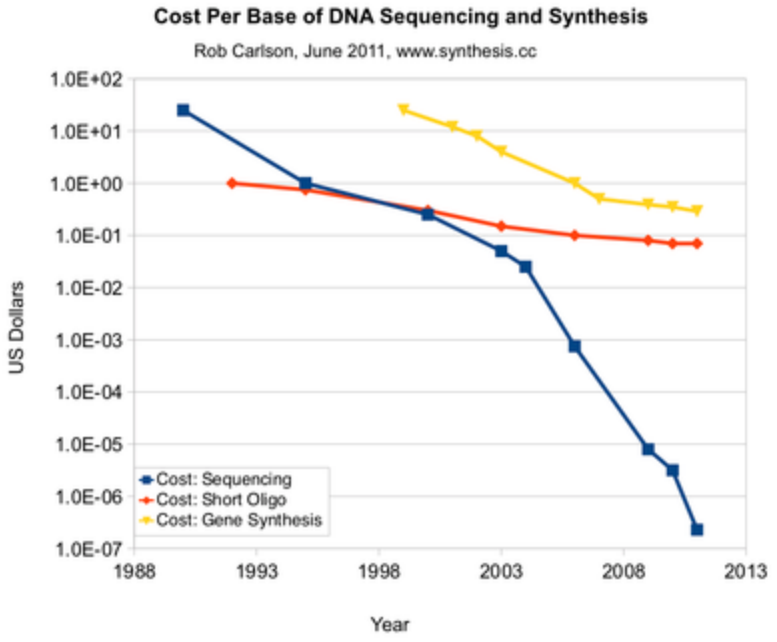
CERN Advanced STORage manager

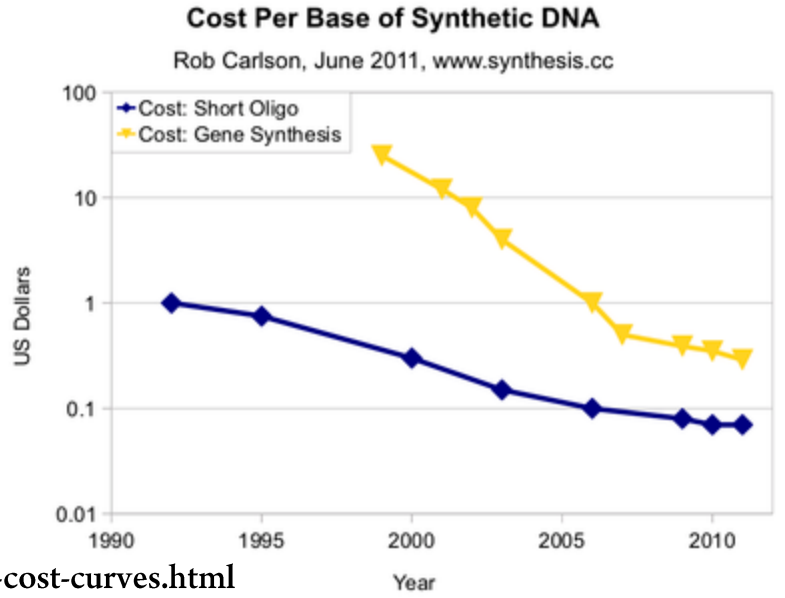
~ 80 PB (80x1000 TB) of data as of writing of Paper, growth of ~15 PB per year

- 10% is maintained on disk, CASTOR migrates regularly between tapes
- Access to old data decreases considerably after 2-3 years after aquisition
- *Current costs ~ 12’400$ per MB for coding and ~220$ per MB for decoding*

CONCLUSIONS & OUTLOOK

- With current synthesis & reading costs as well as current coding scheme, DNA-based storage may be cost-effective for archives of several megabytes within a ~600-5'000 yr period
- If costs decrease by 1 log, ~50-500 yr period
- If costs decrease by 2 logs, < 50 yrs
- ~2.2 PB of information per gram DNA
- long-term storage seems to be evolutionary proven

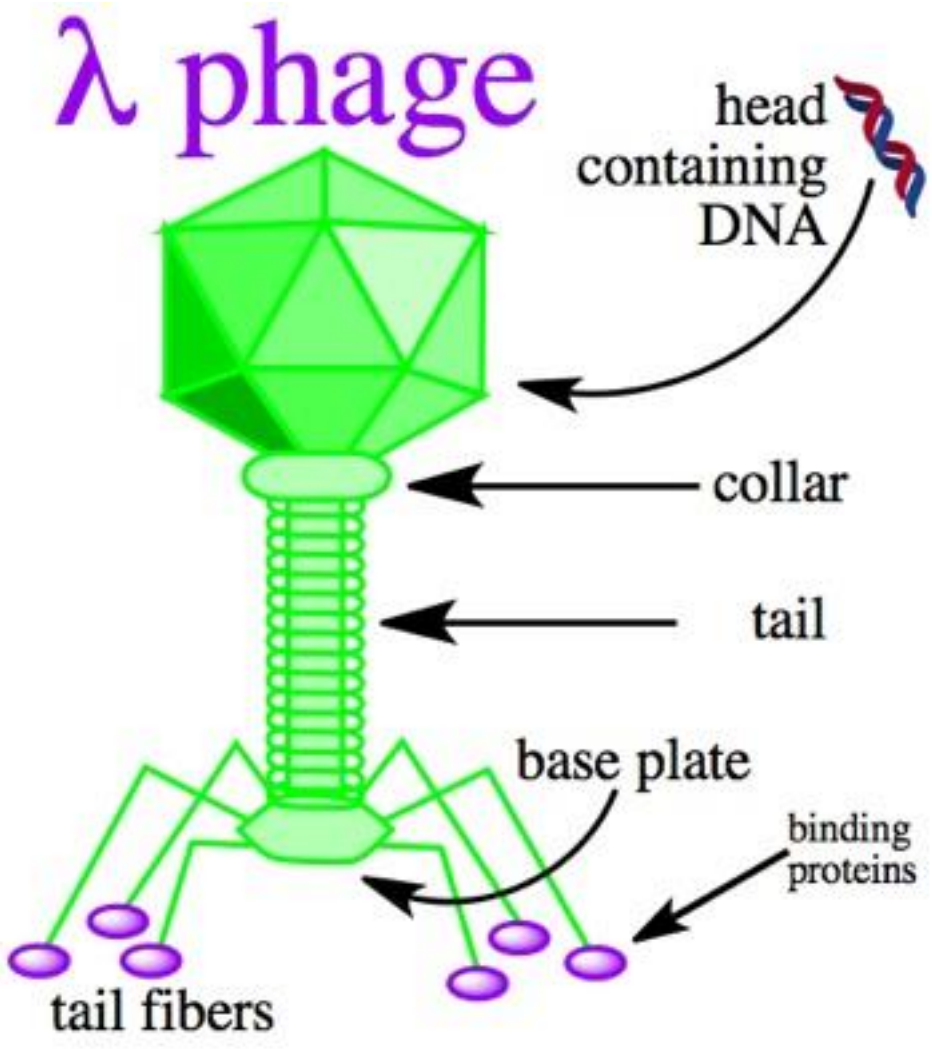




http://www.synthesis.cc/2011/06/new-cost-curves.html

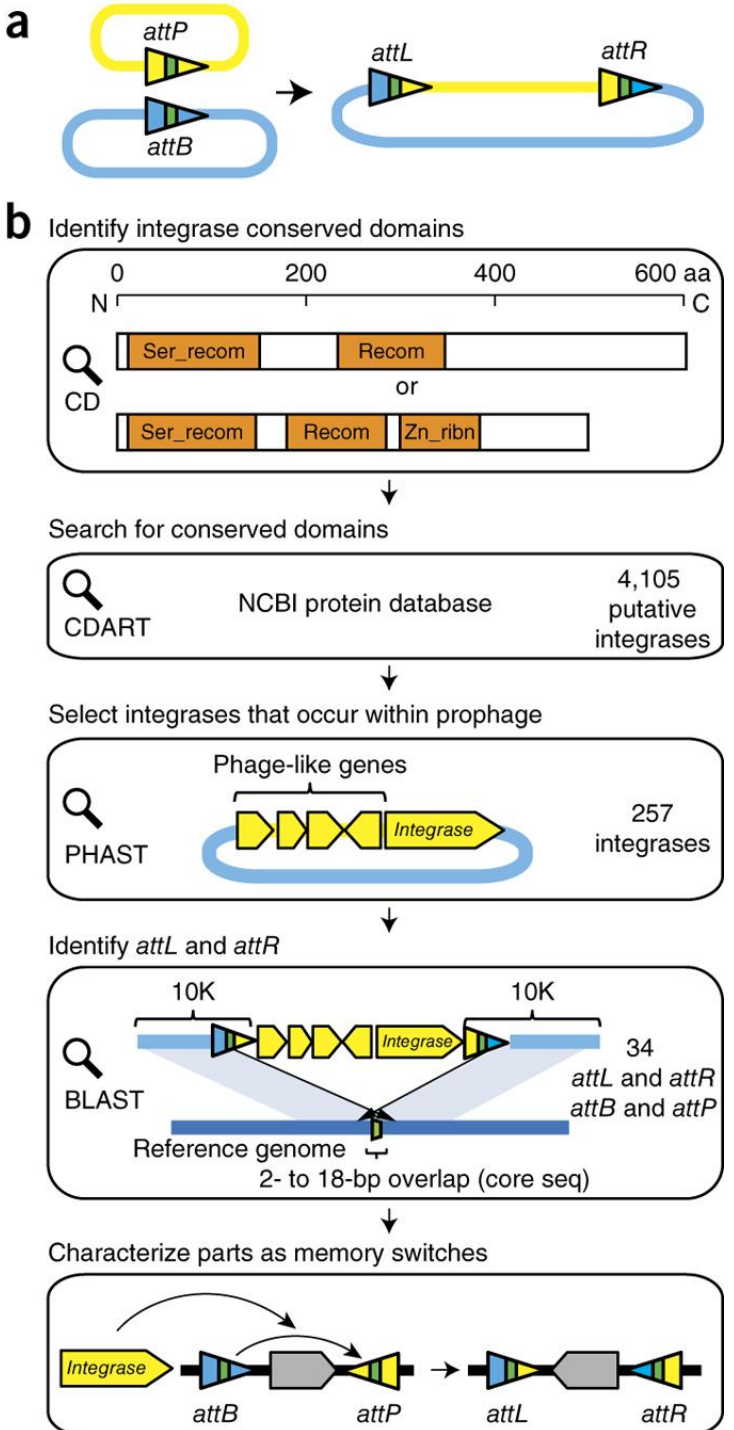
PAPER #3 – YANG ET AL., NATURE METHODS 2014

"PERMANENT GENETIC MEMORY WITH >1-BYTE CAPACITY"

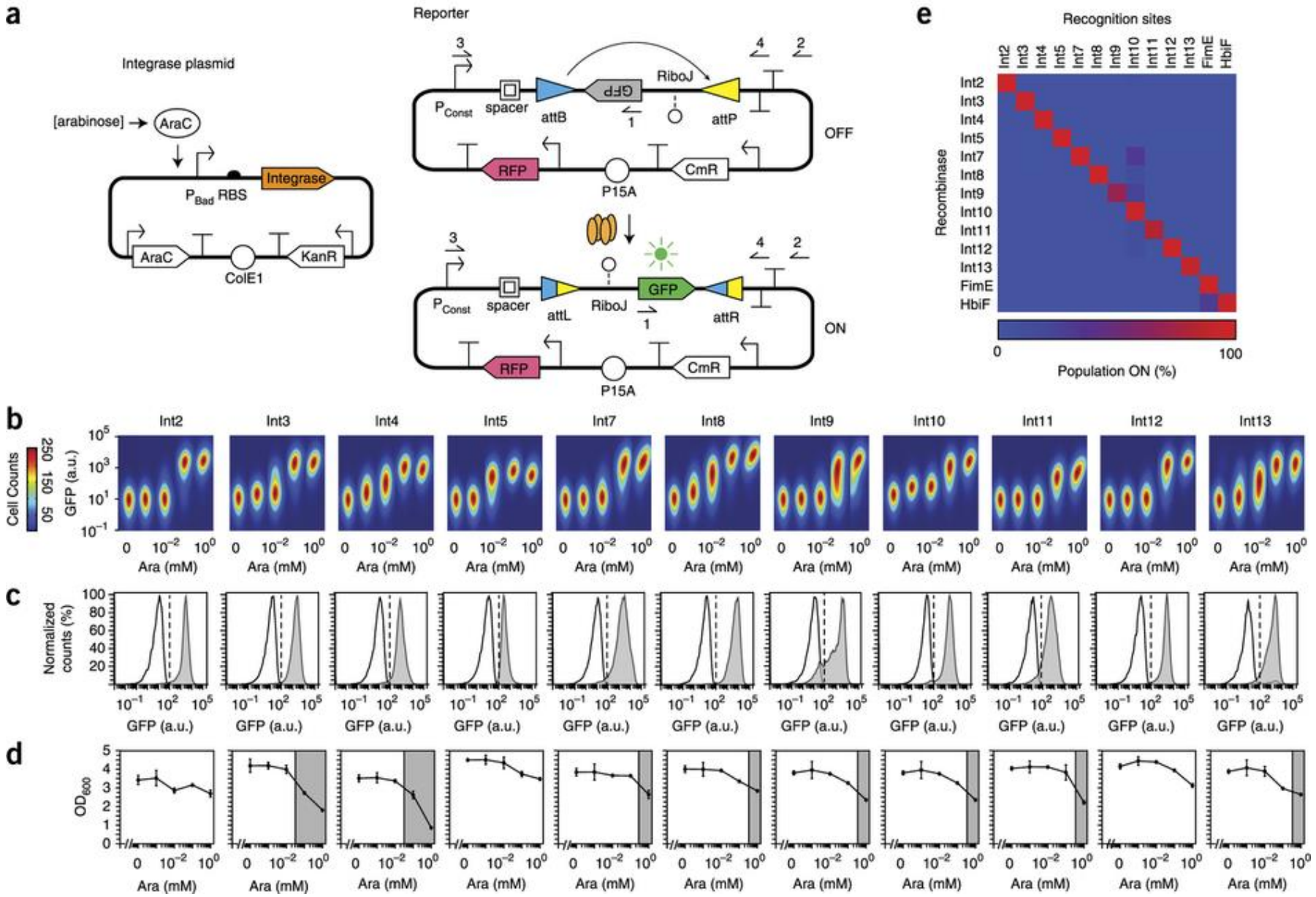


LSTP
«large serine-type phage» integrases

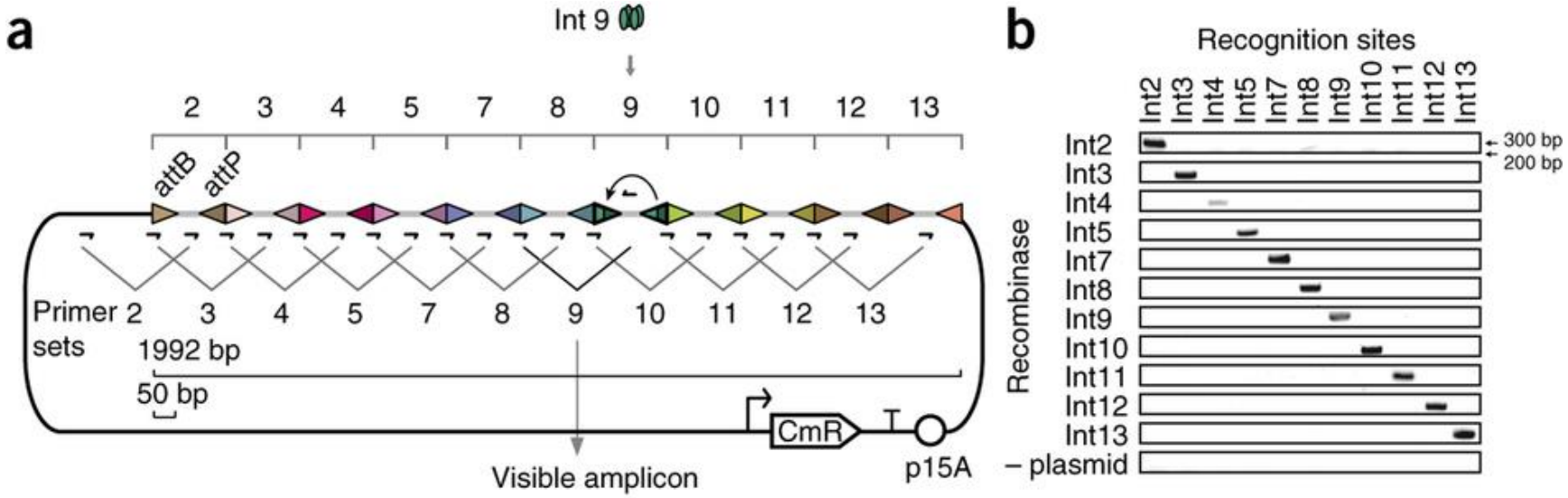
finding LSTP integrases



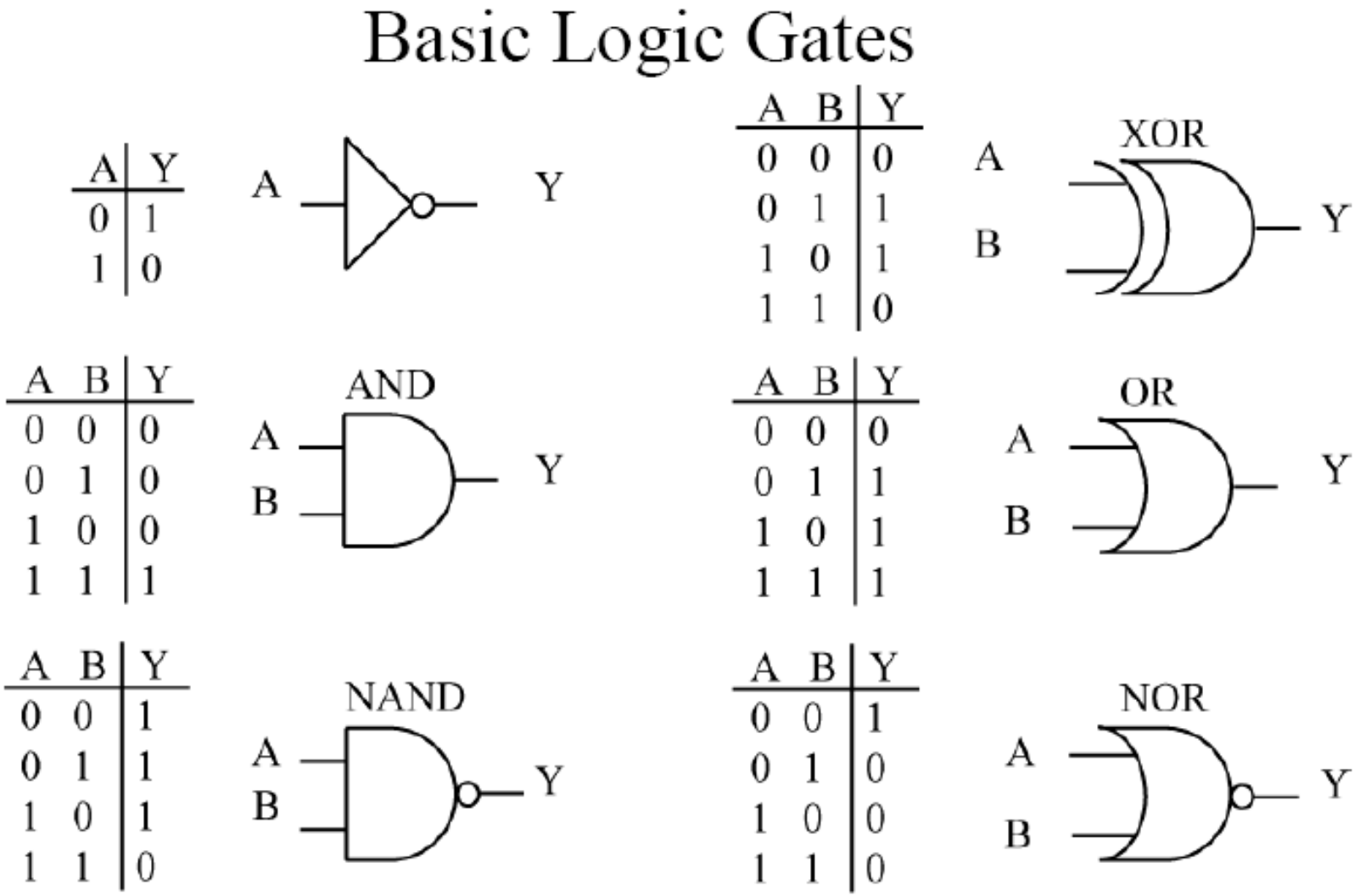
CHARACTERIZATION OF MEMORY SWITCHES AND ORTHOGONALITY



CREATING A MEMORY ARRAY WITH 11 ORTHOGONAL INTEGRASES FOR RECORDING OF 2^11 = 2048 DIFFERENT STATES (> 1 BYTE)

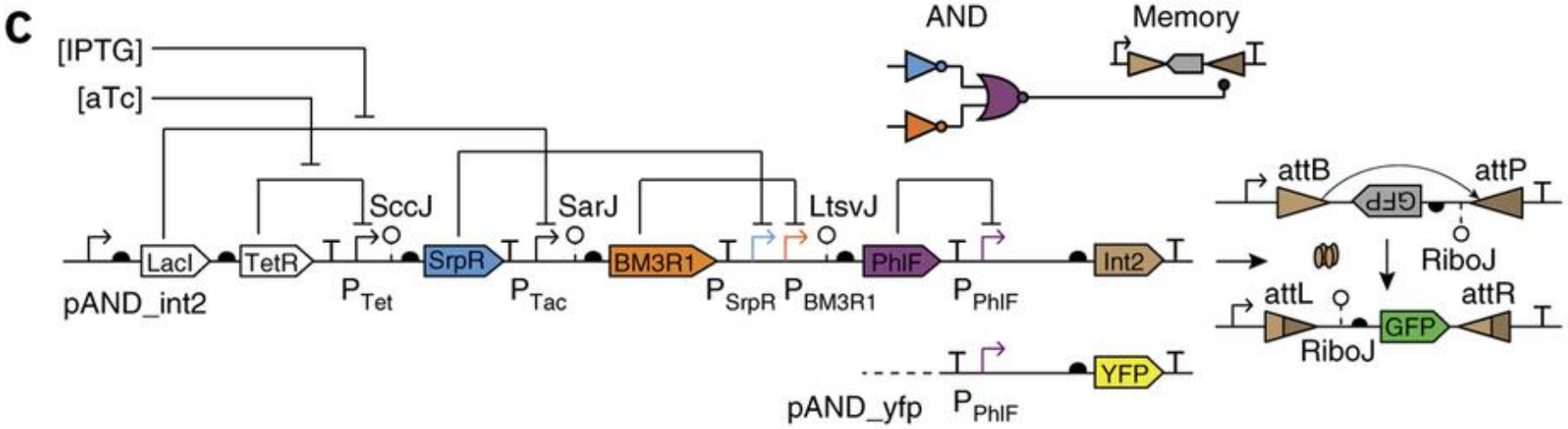


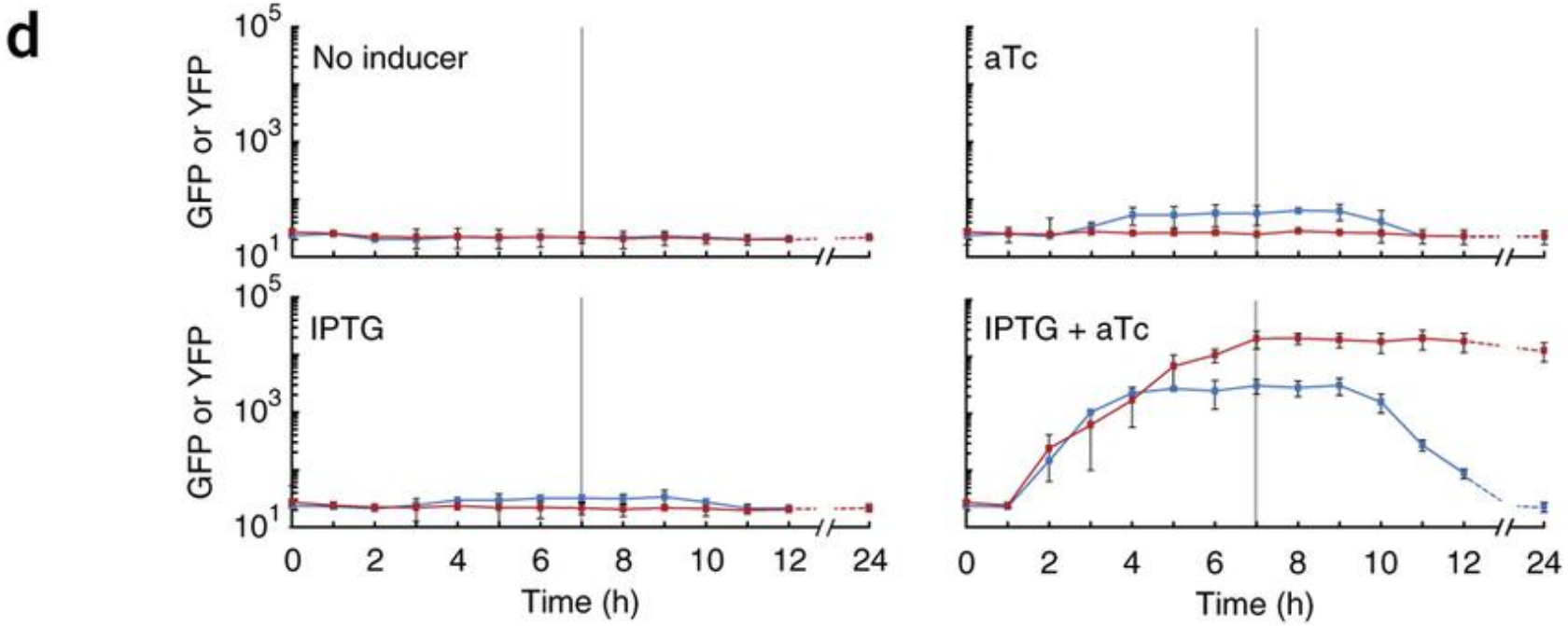
LOGIC GATES



http://www.kaituanui.com/elect-logic-1.htm

LOGIC AND-GATE





CIRCUITS OF MULTIPLE RECOMBINASES