

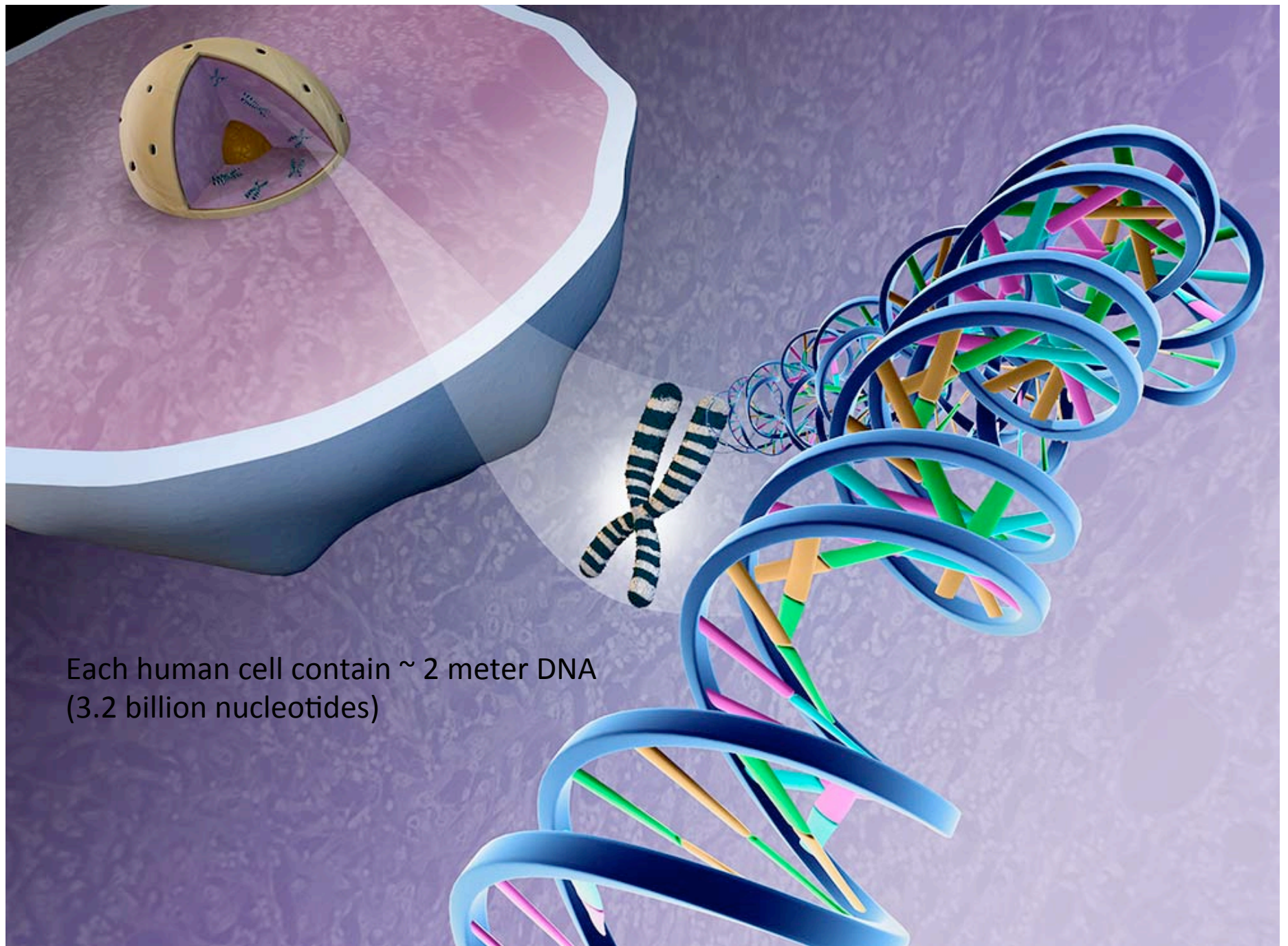


RNA/DNA Library construction

Walter Verweij
walter.verweij@earlham.ac.uk

Me...Born and raised in the Netherlands
During my study I lived 9 months in Portland, USA
PhD on *Petunia hybrida* flower coloration (vacuolar pH regulation), Free
University Amsterdam
Post-doc The Sainsbury Lab in Norwich, UK
Currently Senior Research Scientist at Earlham Institute (TGAC)
Future...moving back to the Netherlands (next month)

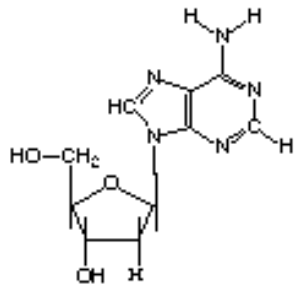




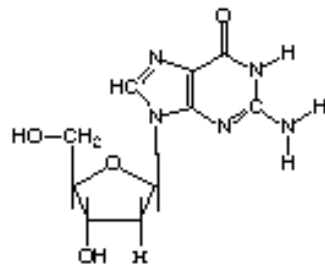
Each human cell contain ~ 2 meter DNA
(3.2 billion nucleotides)



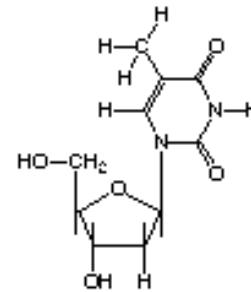
The Nucleotides of DNA



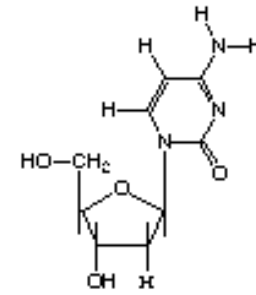
Adenine



Guanosine

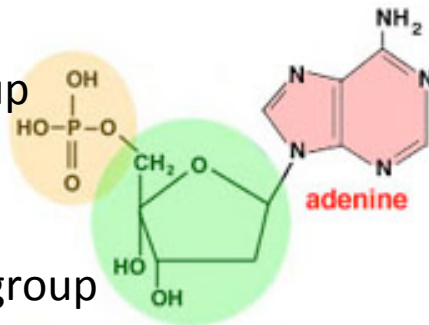


Thymine

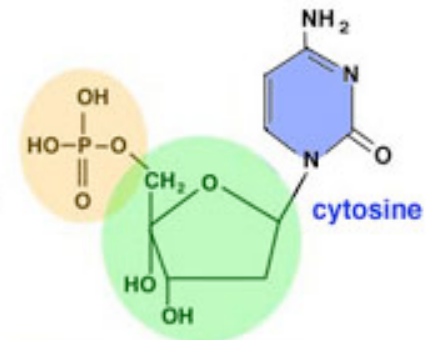
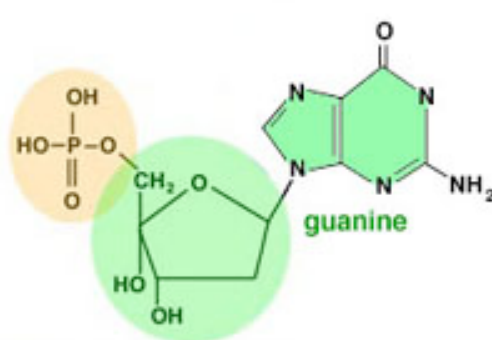
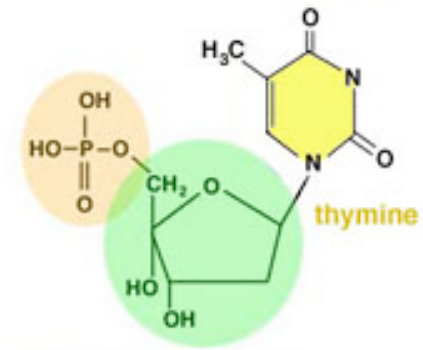


Cytosine

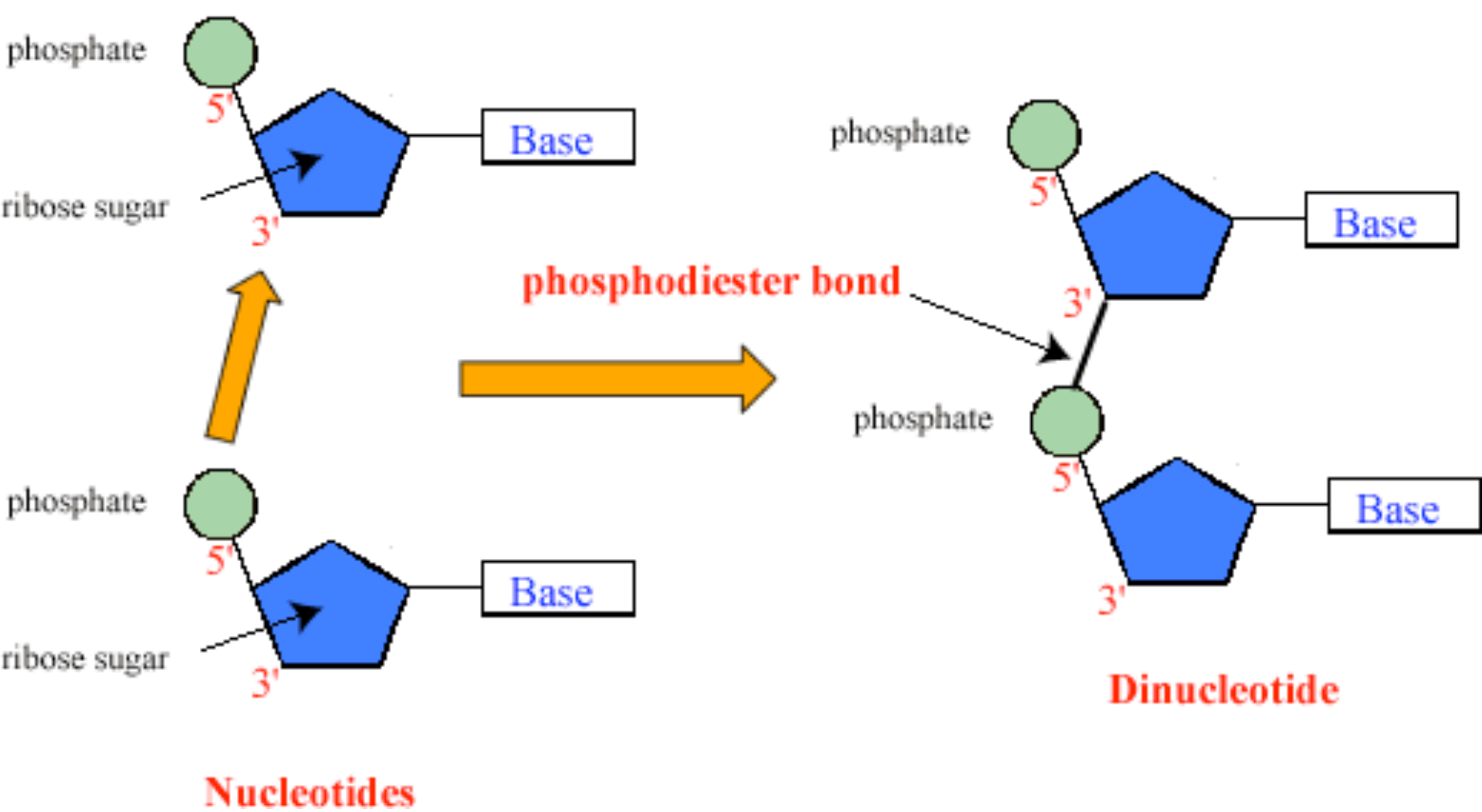
Phosphate group



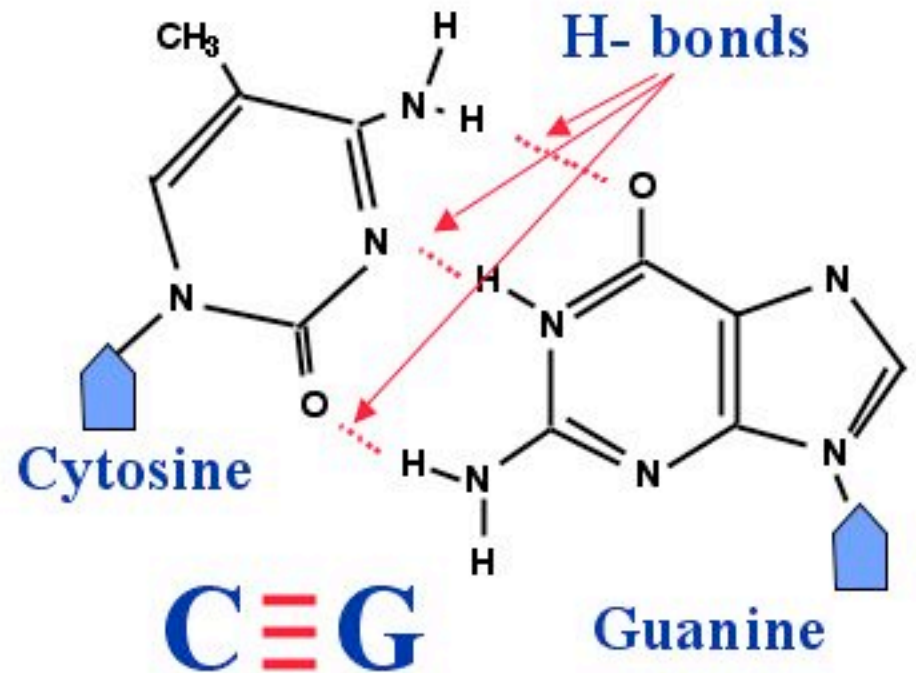
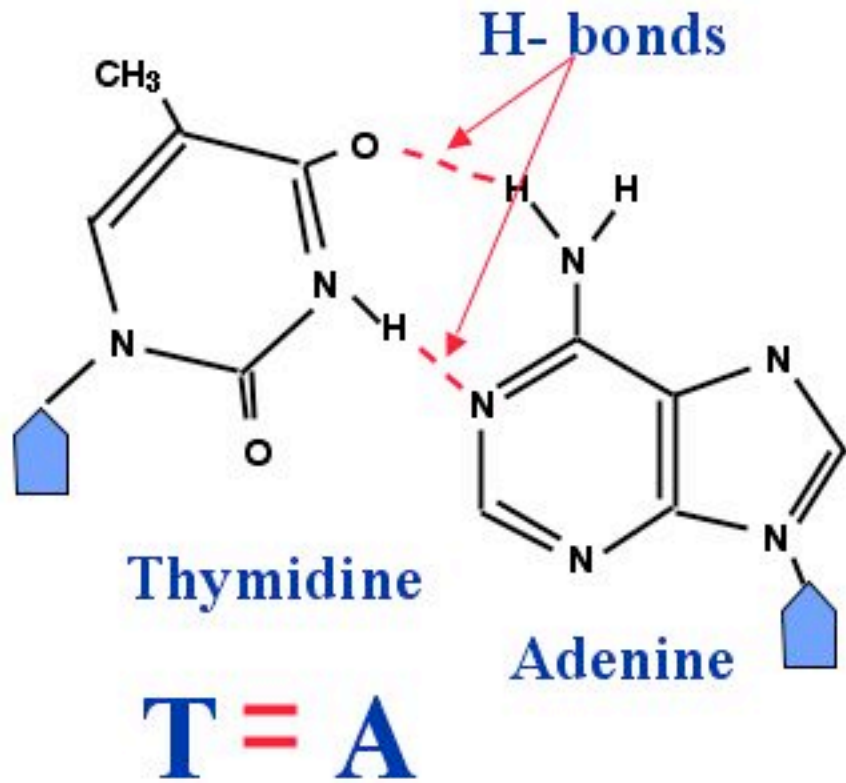
Sugar group



Polynucleotide formation

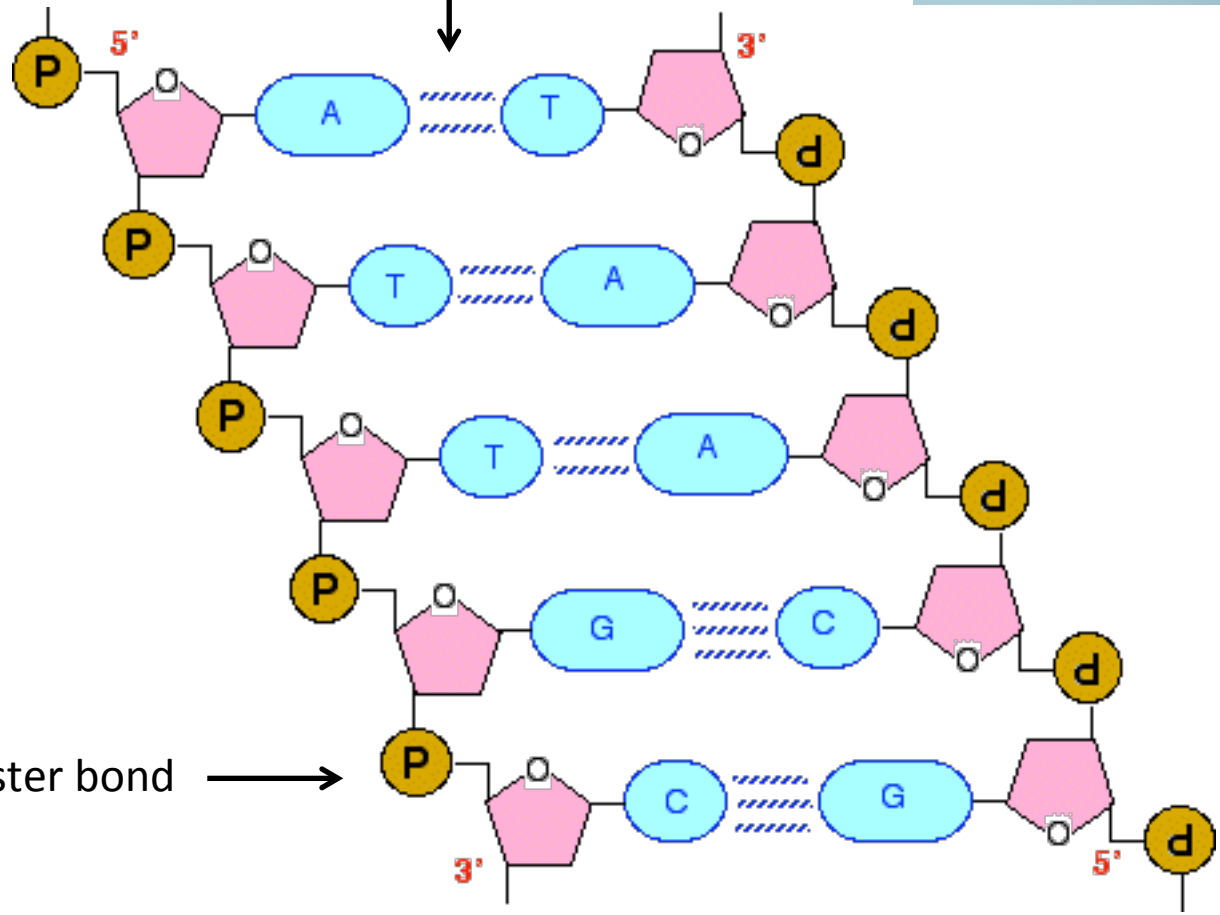


Nucleotide Pairing





Hydrogen bond



phosphodiester bond



E.coli virus Phi174 → 5386 bp
 E.coli → 4 Mb (4 Million bp)
 Candidatus Carsonella ruddii → 160Kb (160.000 bp)
 A. Tumerfaciens → 4Mb (4,674,062 bp)
 Fruitfly → 122 Mb (122,653,977 bp)
 C.Elegans → 100Mb (100,258,171 bp)
 Arabidopsis → 135 Mb
 Human → 3.2 Gb
 Wheat → 17 Gb

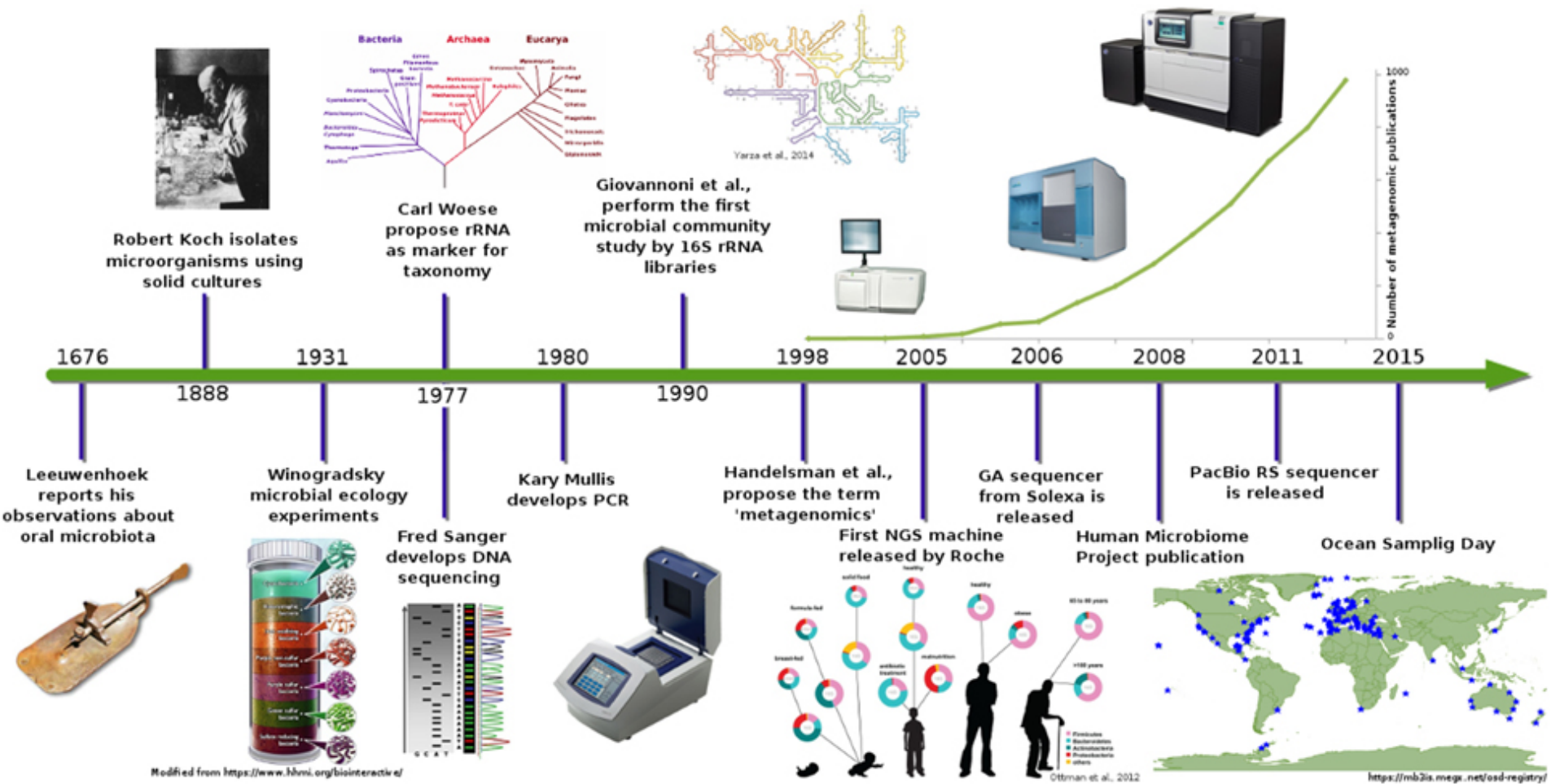
| SPECIES | CHROMOSOMES | GENES | BASE PAIRS |
|---|----------------------|--------------------|---------------------|
| Human (<i>Homo sapiens</i>) | 46 (23 pairs) | 28-35,000 | ~3.1 billion |
| Mouse (<i>Mus musculus</i>) | 40 | 22.5-30,000 | ~2.7 billion |
| Pufferfish (<i>Fugu rubripes</i>) | 44 | ~31,000 | ~365 million |
| Malaria Mosquito (<i>Anopheles gambiae</i>) | 6 | ~14,000 | ~289 million |
| Sea Squirt (<i>Ciona intestinalis</i>) | 28 | ~16,000 | ~160 million |
| Fruit Fly (<i>Drosophila melanogaster</i>) | 8 | ~14,000 | ~137 million |
| Roundworm (<i>C. elegans</i>) | 12 | 19,000 | ~97 million |
| Bacterium (<i>E. coli</i>) | 1* | ~5,000 | ~4.1 million |

*Bacterial chromosomes are **chromonemes**, not true chromosomes.

To understand biology, knowing the DNA sequence of an organism or population of organisms is very useful

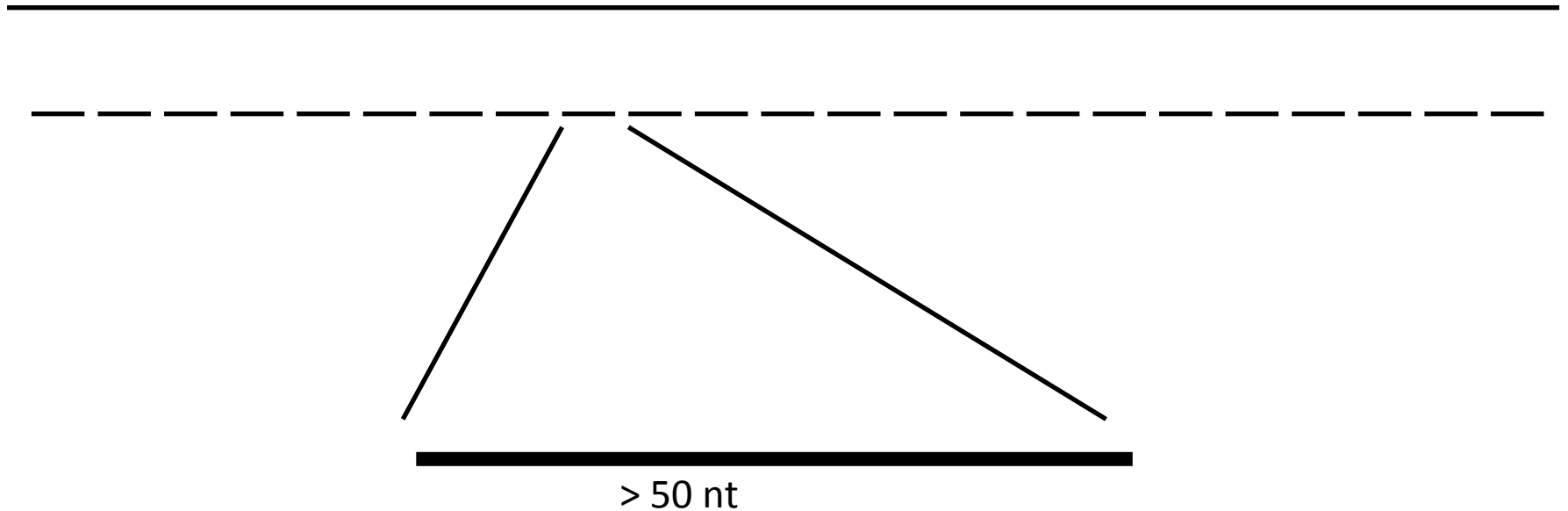
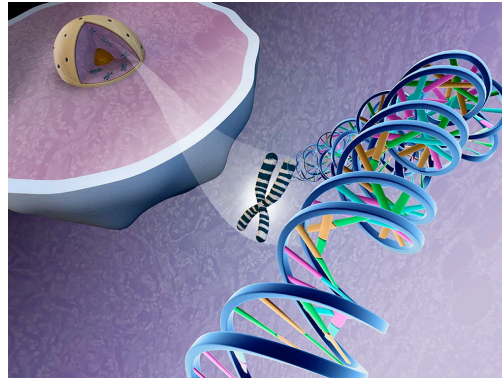
“DNA sequencing is the process of determining the precise order of nucleotides (A,C,G and T) within a DNA molecule. It includes any method or technology that is used to determine the order of the four bases —adenine, guanine, cytosine, and thymine— in a strand of DNA”

Sequence history



Ideal world

There is no technique available to sequence entire genomes/chromosome
We always have to break the DNA in little pieces and determine DNA sequence





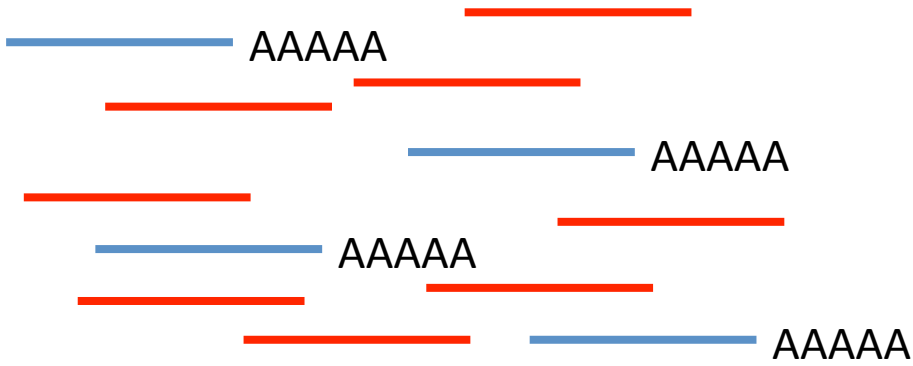
More on this Thursday morning



library preparation

RNA

DNA



sequencing

DNA library prep



Genome



Fragment Genome
50 to 500bp, sonication
or mechanically



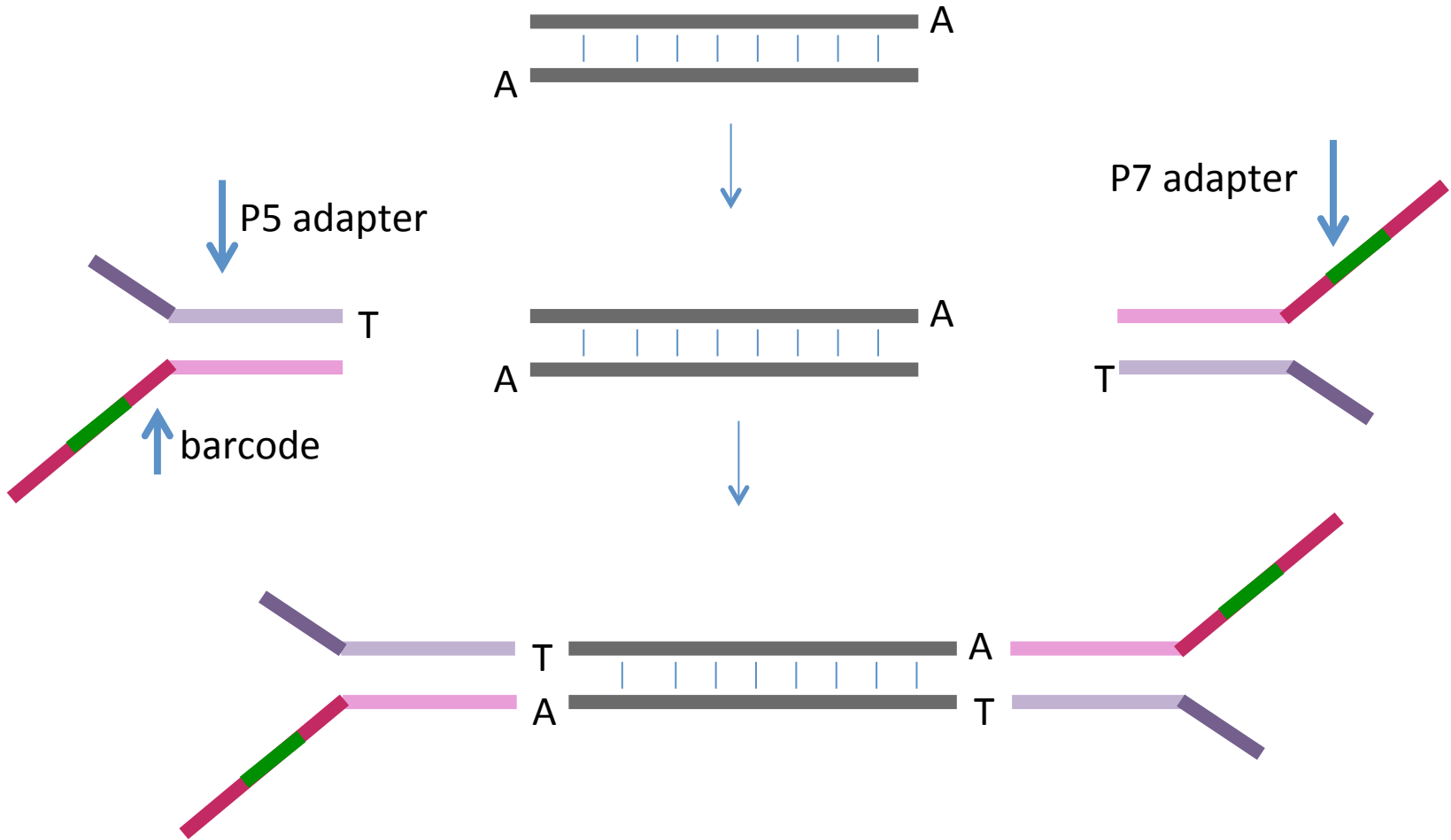
End Repair :DNA ends
are blunt ended



A tailing to the 3'-
end of each DNA
fragment



Illumina library



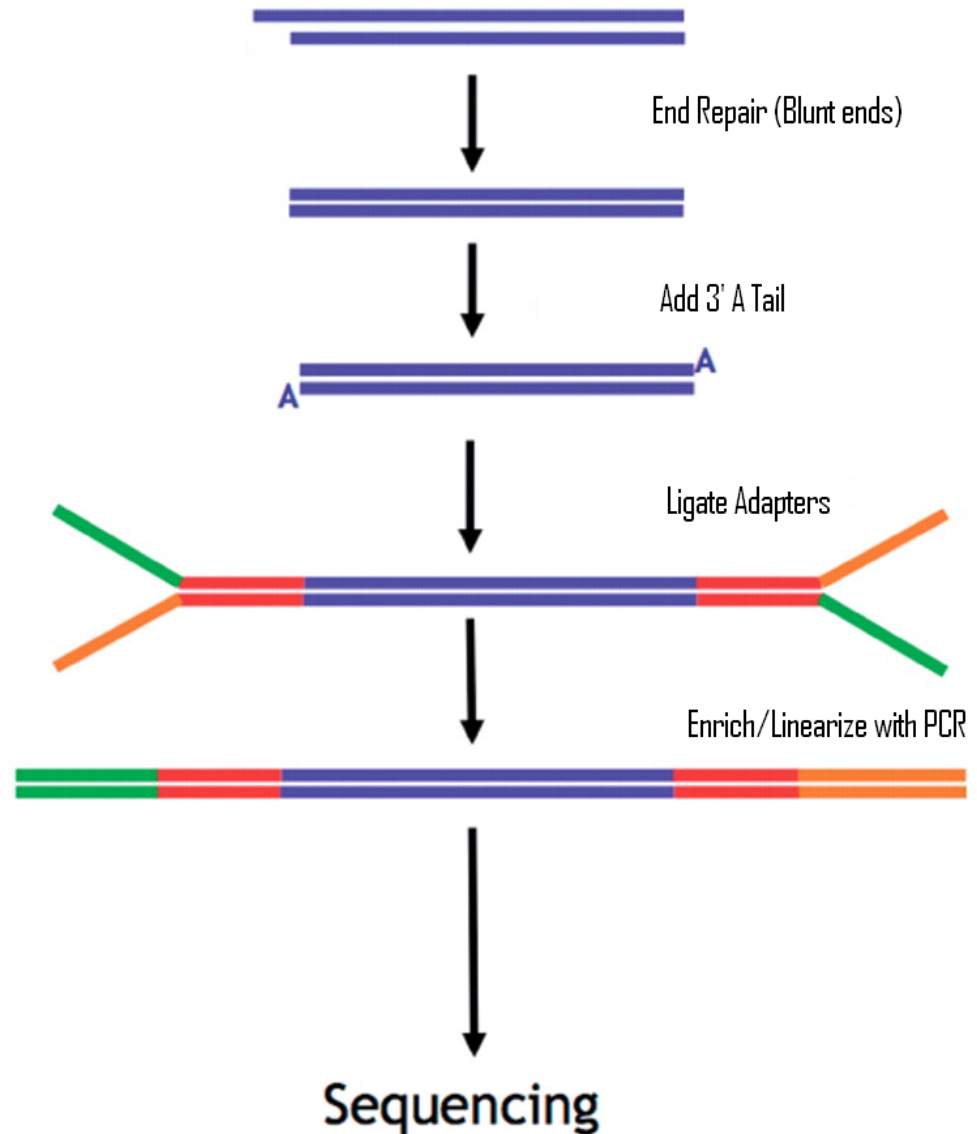
DNA library prep: shearing DNA in small fragments

DNA extraction:

- CTAB
- DNA extraction kit

Shearing DNA

- Sonication
- Mechanically



DNA fragmentation using Tn5 Transposase

A **transposable element (TE or transposon)** is a [DNA sequence](#) that can change its position within a [genome](#),

Transposase is an [enzyme](#) that binds to the end of a [transposon](#) and catalyzes the movement of the [transposon](#) to another part of the genome by a cut and paste mechanism or a replicative transposition mechanism

Tn5 and other transposases are generally inactive. Introduced mutations make the Tn5 transposase very active and consequently, the transposon randomly inserts into the DNA.

Transposomes, adaptor oligos



+

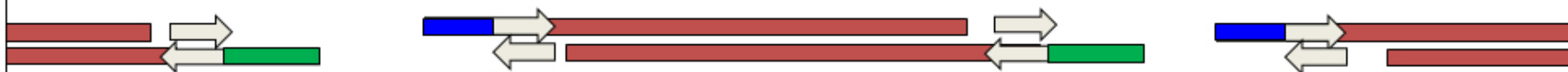


Target dsDNA (gDNA)

Tagmentation (transposition)



Tagmented dsDNA fragments



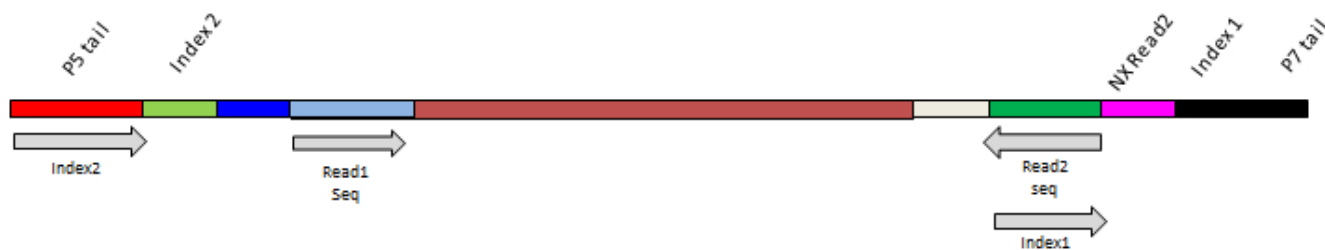
PCR cycle 0: extension to fill 3'



PCR amplification



Sequencing-ready DNA fragments



Sonication

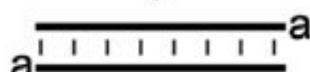
1. Sonication



2. End-repair



3. dA addition



4. Adapter ligation



5. PCR



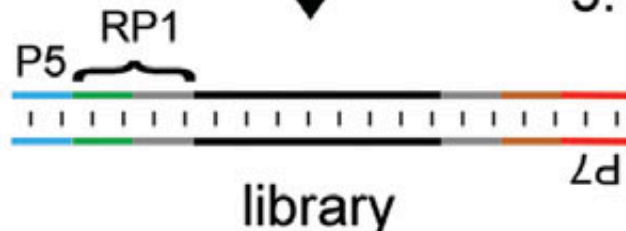
Tn-based prep.



1. Transposition



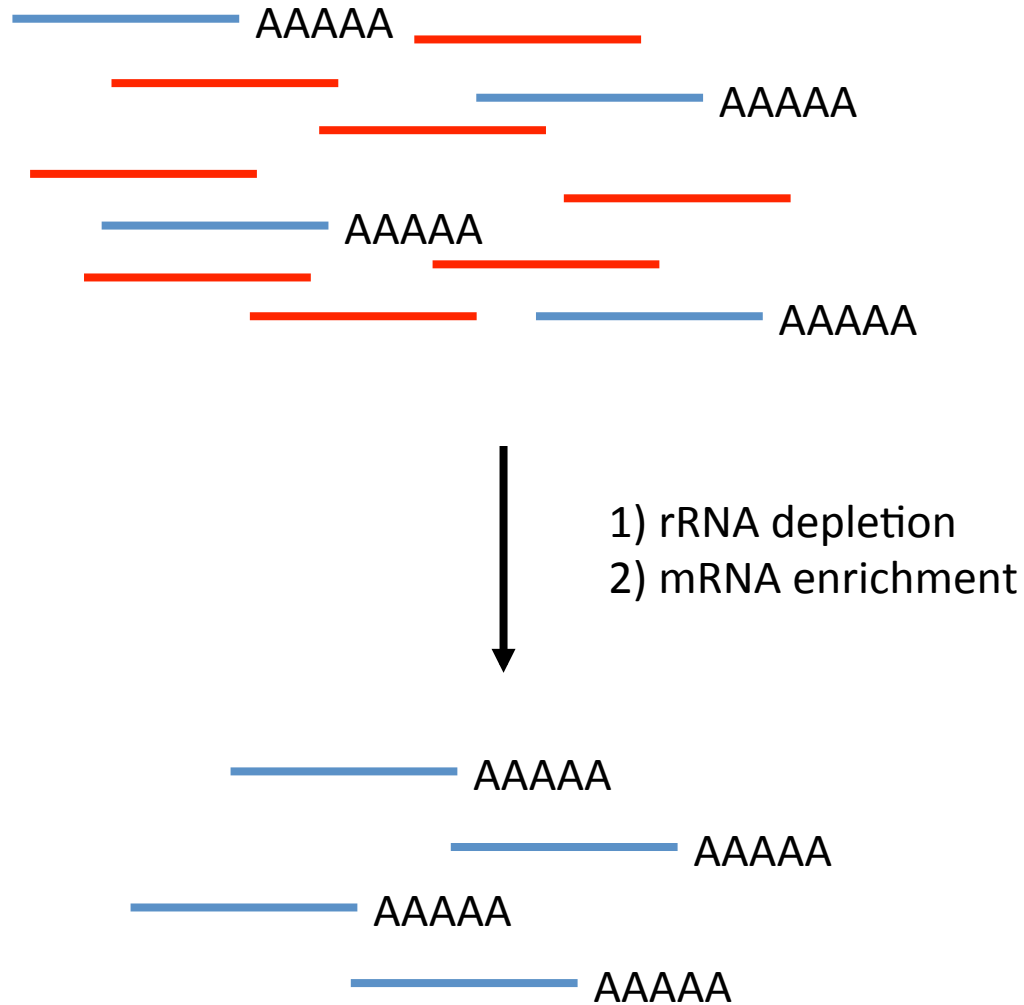
2. Gap fill-in



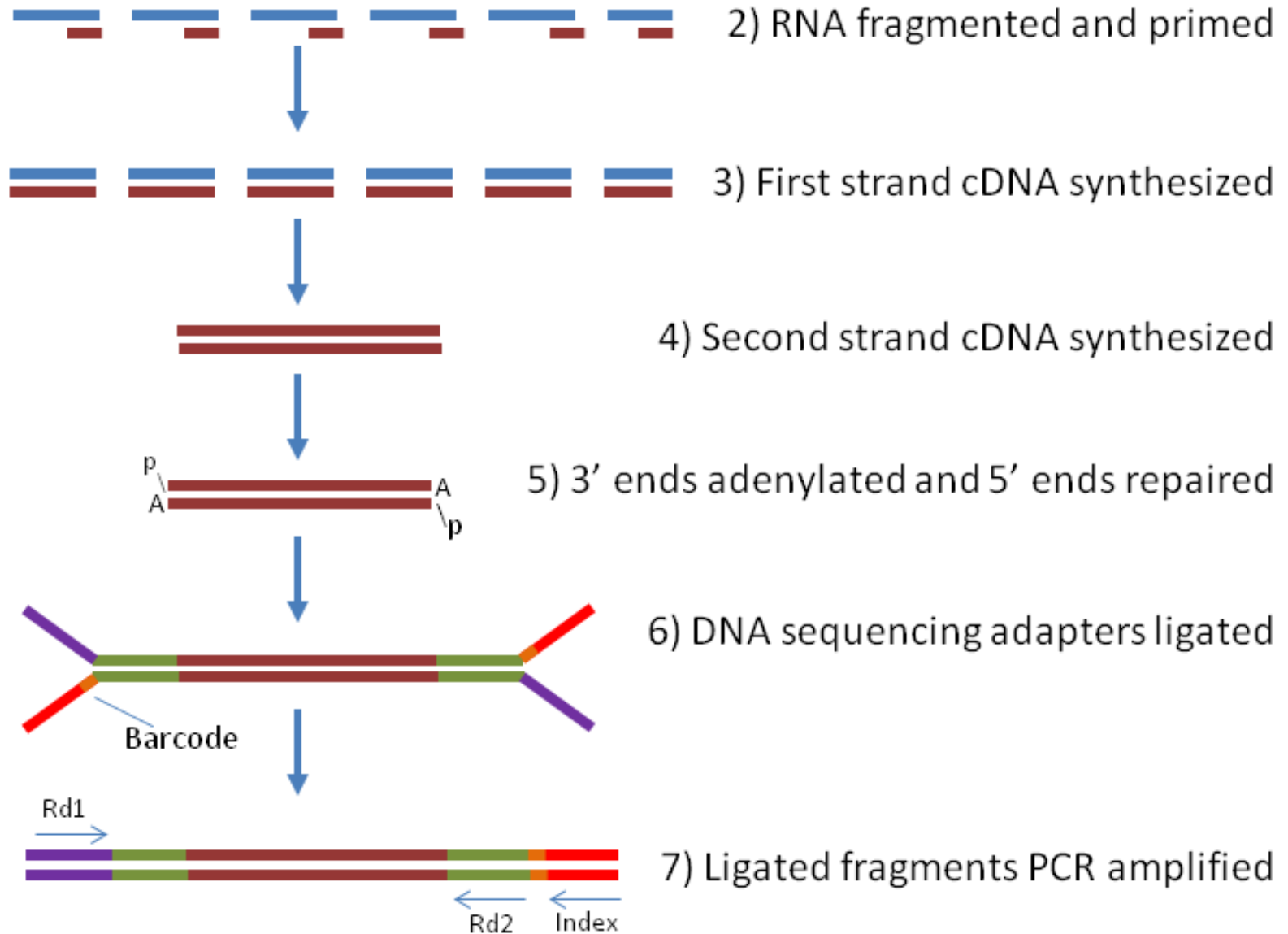
3. PCR

library

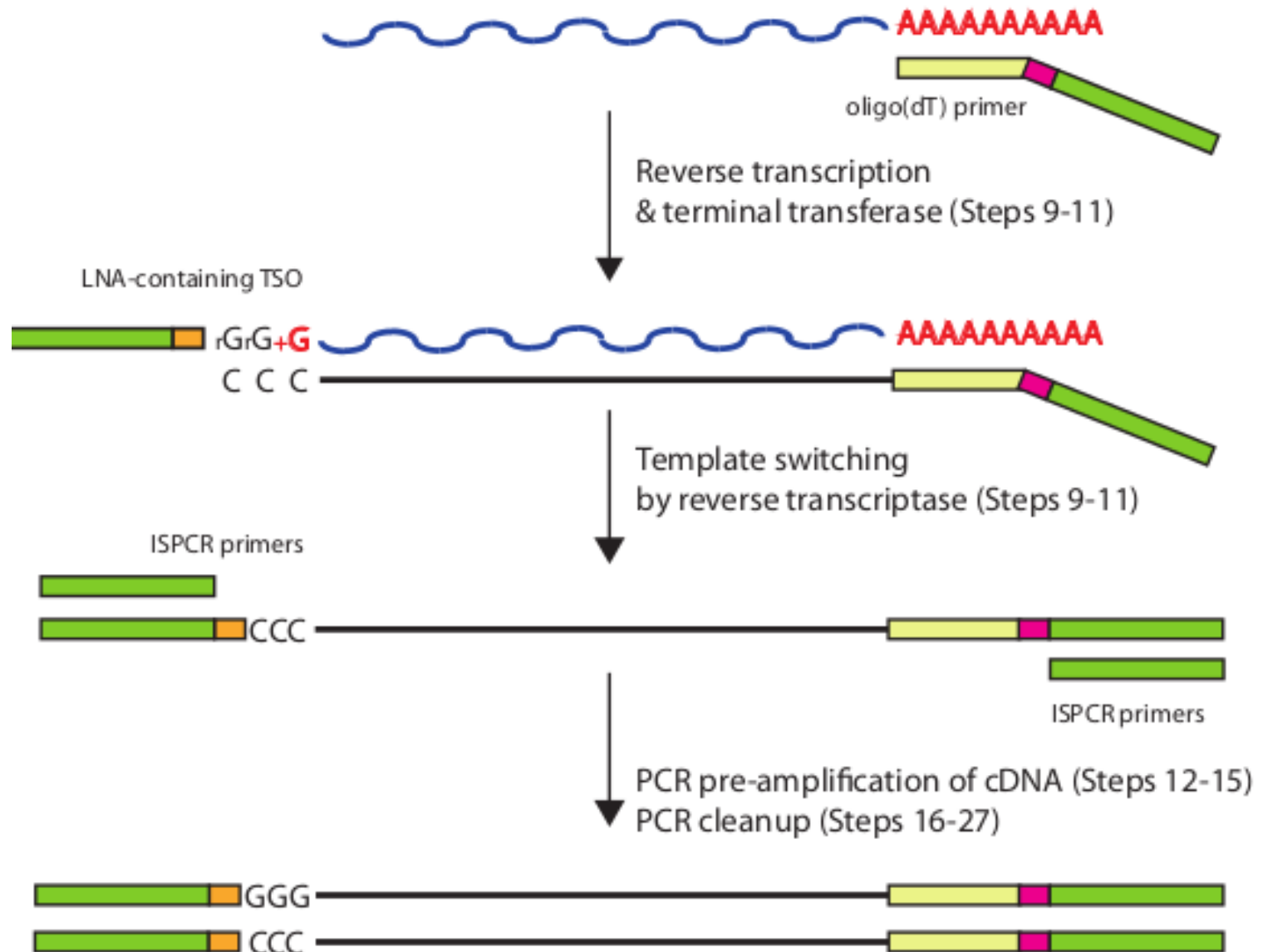
RNA library preparation



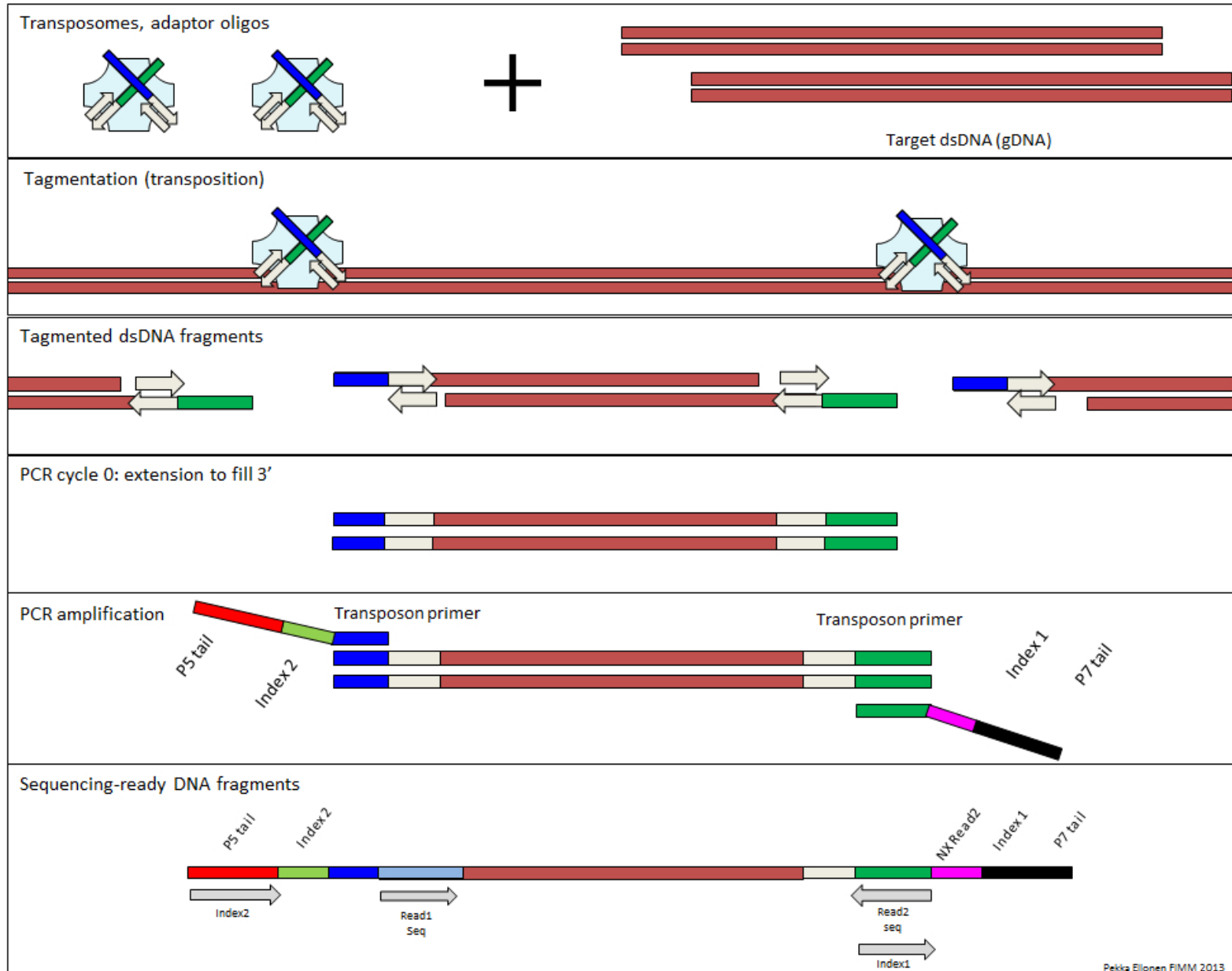
1) rRNA depletion – illumina library prep



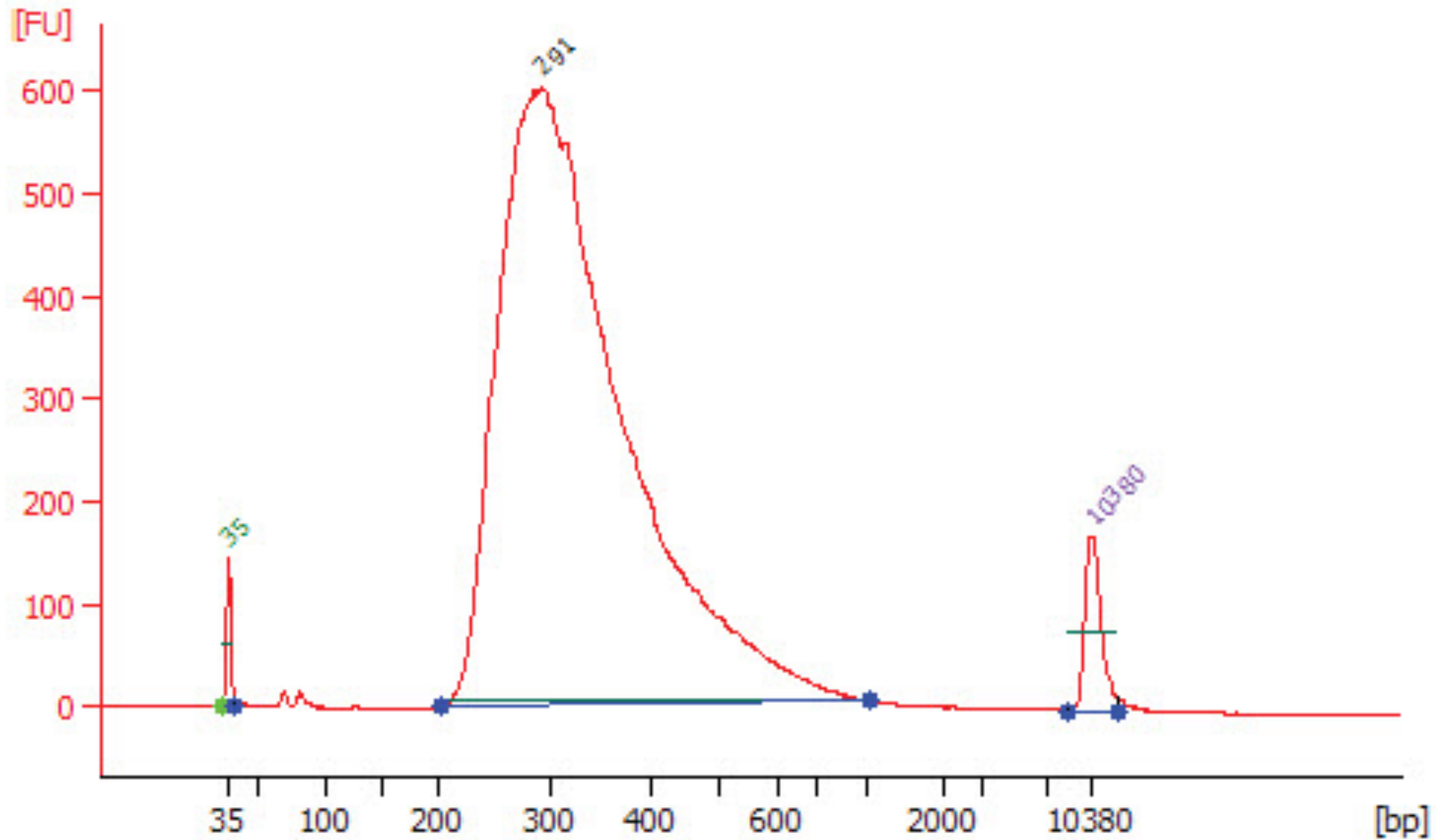
2) mRNA enrichment – SMARTseq2



Nextera library prep today and tomorrow



Bio analyzer trace of successful library construction



Good luck