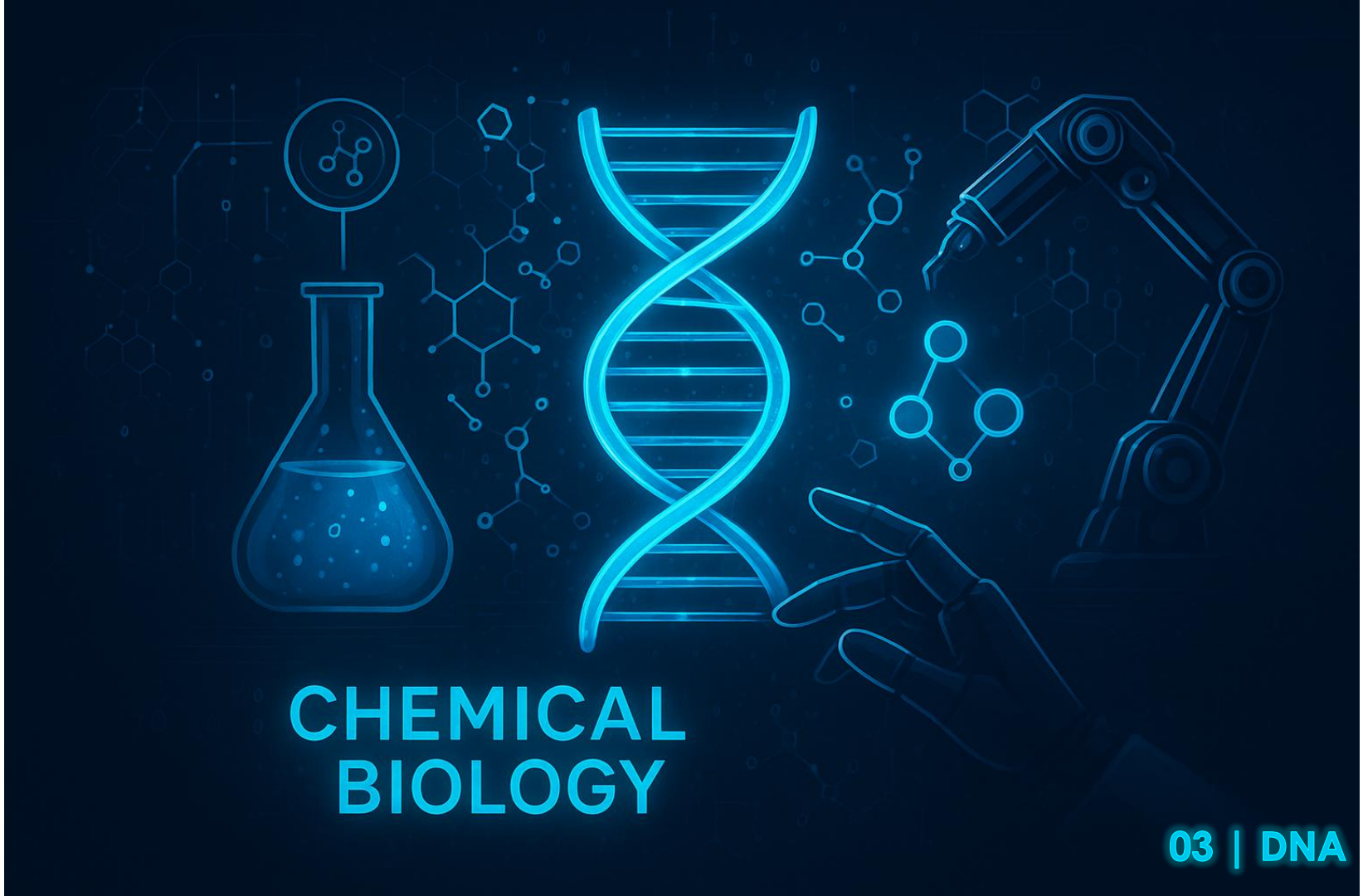


Prof. Dr. Markus Jeschek, LSAM

General Course Information

- Moodle: https://go.epfl.ch/CH-313
  - Lecture slides (evening before the lecture)
  - Distributed presentation topics (assignments)
  - Forum (for questions and announcements)

- Examination (written, graded, detailed information will follow)
- Contact:
  - Moodle forum (for questions)
  - markus.jeschek@epfl.ch

- **“Concepts over details!”**
- **Interact! Ask! Discuss! Anytime!**

Prof. Dr. Markus Jeschek, LSAM

Group Presentations

- Critical discussion of primary literature
- Illustrative examples for topics from the lecture

- Why?
  - Repetition of core concepts, techniques etc.
  - Presentation skills and critical discussion of research
  - Insight into current research topics

- How?
  - Two students per group
  - Assignments distributed one week before delivery of presentation (via Moodle)
  - **Send slides: markus.jeschek@epfl.ch (Mon evening before presentation)**
  - **15 min presentation (both group members should present!)** + Q&A

Tipps for Group Presentations

- Rough structure
  - Short intro on general topic
  - Main presentation according to assignment
  - Brief outlook incl. points of criticism/open questions/personal opinion as kick-starter for the discussion
- Everybody should participate in the discussion, incl. constructive(!) feedback on presentation style
- Questionnaires with different points, feedback by peers
- Typical assignment:
  - You will receive a certain topic including a related publication
  - Introduce the topic using the publication
  - present the motivation behind the research, methodology, key results (not every graph!)
  - Additional questions will be provided hinting towards central points
  - Be encouraged to look/present beyond the questions and the provided paper

Group Presentations – Schedule

#	Name1	Name2	Presentation on…	Assignment on…
1	Wininger Quentin	Jeremy	**Sep 23, 2025**	Sep 16, 2025
2	Ema	Ariane	**Sep 30, 2025**	Sep 23, 2025
3	Benjamin	Matthieu	**Oct 7, 2025**	Sep 30, 2025
4	Ivana	Ipek	**Oct 14, 2025**	Oct 7, 2025
5	Mridhula	Elodie	**Oct 28, 2025**	Oct 14, 2025
6	Abigail	Robin	**Nov 4, 2025**	Oct 28, 2025
7	Andrea	Florian	**Nov 11, 2025**	Nov 4, 2025
8	Melodie	?	**Nov 18, 2025**	Nov 11, 2025
9	Bastien	**Siolène (E-Mail)**	**Nov 25, 2025**	Nov 18, 2025
10	Nicole	Maria	**Dec 2, 2025**	Nov 25, 2025
11	Eva	?	**Dec 9, 2025**	Dec 2, 2025

Course Topics – Overview

- Week 1 | Introduction + DNA
- Week 2 | DNA
- **Week 3 | DNA**
- Week 4 | DNA/RNA
- Week 5 | Protein/Enzymes
- Week 6 | Enzymes
- Week 7 | Enzymes
- Week 8 | Membranes
- Week 9 | Metabolism
- Week 10 | Metabolism
- Week 11 | Engineering
- Week 12 | Engineering
- Week 13 | Engineering
- Week 14 | LSAM Intro + Exam Preparation

[tentative schedule]

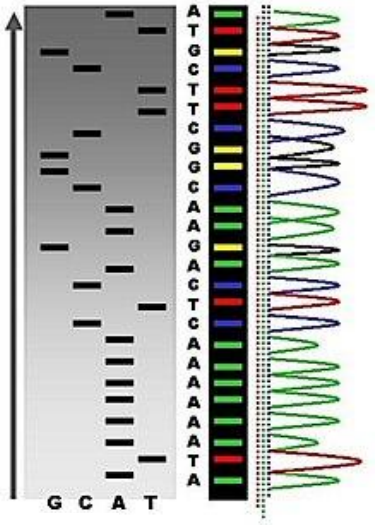
DNA (“read”)

DNA Sequencing – Overview

- Analytical process to determine the sequence of nucleotides (nucleobases) in a DNA molecule

- Key technology in molecular biology/biotech (“era of genomics“)
  - early methods: 1970s
  - wider availability: 1980s/1990s
  - “next-generation“ sequencing: late 1990s/early 2000s, ongoing!

- Selected applications
  - molecular biology
  - evolutionary biology
  - epidemiology, virology
  - quality control (DNA synthesis, cloning etc.)
  - diagnostics, forensics
  - etc. etc.



Second generation sequencer: 454 GS20

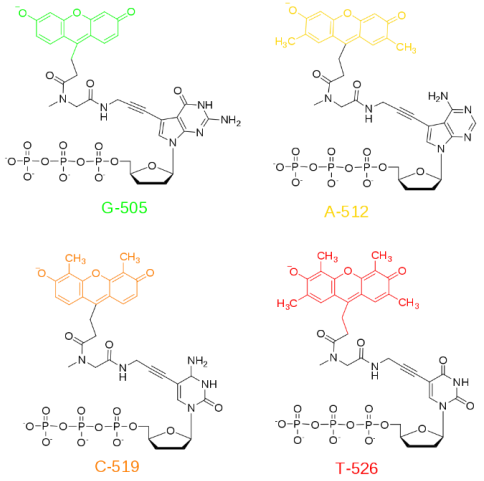
Prof. Dr. Markus Jeschek, LSAM

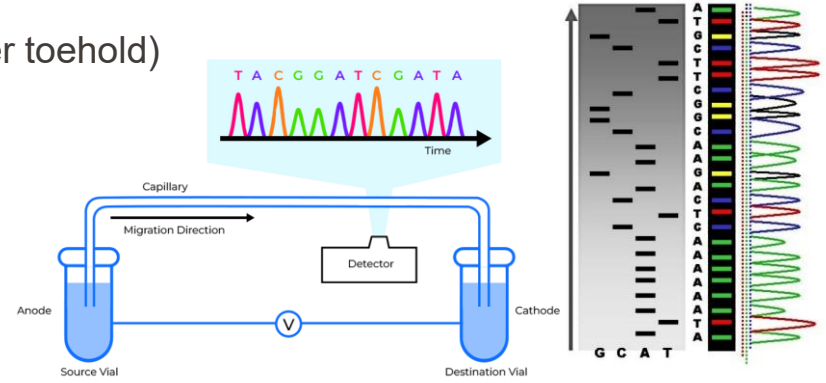
Sanger Sequencing – Pushing the Price

- Fluorescently labelled ddNTPs (“dye-terminator sequencing“)
  - → no radioactive labels/autoradiography
  - → one instead of four reactions
- Capillary electrophoresis
- Automation, miniaturization etc.
- **600 – 1200 bp, < 4 CHF per sample, <24 h result delivery**
- high accuracy (> 99.9% per positions)
- Limitations:
  - weak signal in the beginning (~50 bp from primer toehold)
  - length limit
  - repetitive sequences

Q: What could be reasons for the length limitation?

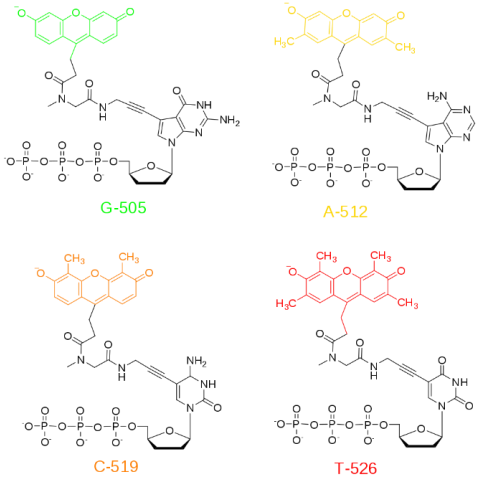
How can one sequence fragments larger than ~ 1kbp?

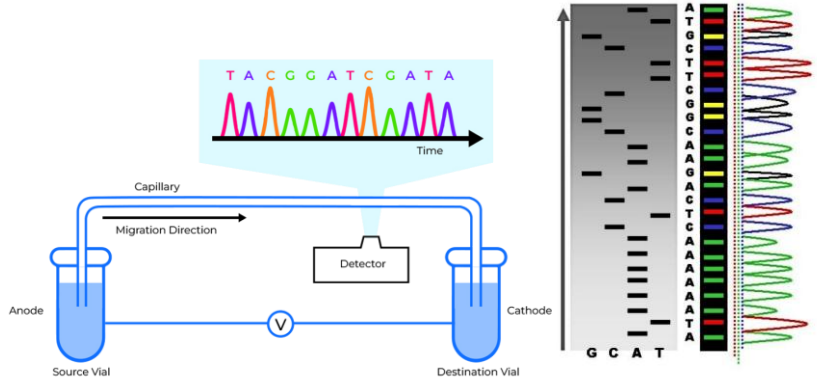




Question from last lecture

- Where exactly does the length limitation/decreased signal strength for longer sequences in Sanger sequencing come from?

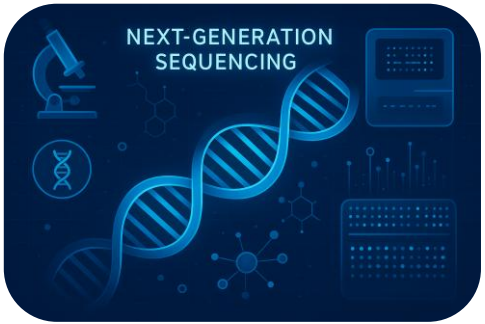




Prof. Dr. Markus Jeschek, LSAM

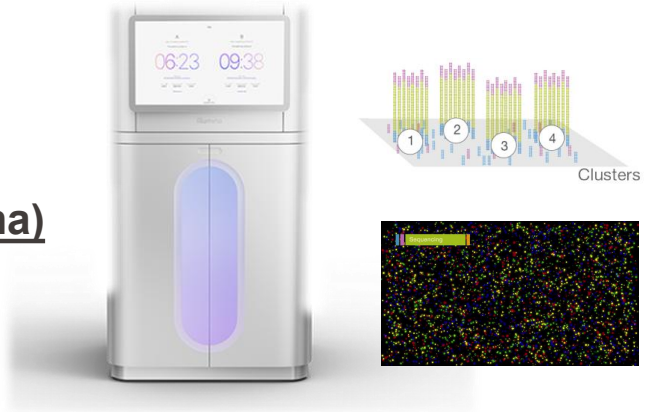
Next-generation sequencing (NGS)

- “Massively parallel sequencing“, “2nd/3rd generation sequencing“, “deep sequencing”, “high-throughput sequencing”…
- Different “recent” technologies (from late 1990s/early 2000s)
- application scope (selection)
  - whole-genome sequencing (“shotgut” approach)
  - diagnostics
  - transcriptomics, amplicon sequencing (libraries!)
  - single-cell sequencing approaches
  - etc.
- **key technology for modern day biotech and medicine**

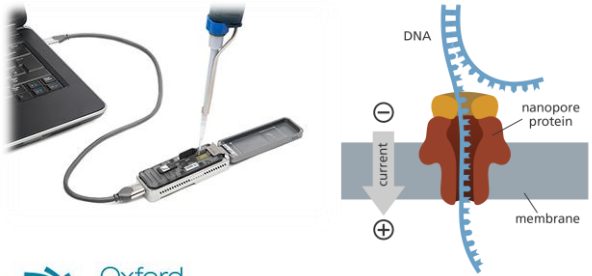


NGS – Main Methods (Overview)

- "Second generation" methods
  - **(1) Pyrosequencing (Roche 454)**
  - **(2) Sequencing by synthesis (Illumina)**
  - (3) Sequencing by ligation
  - (4) Ion semiconductor sequencing



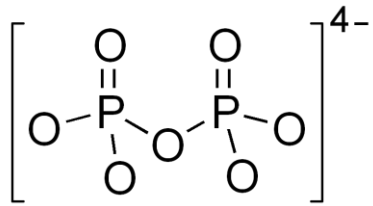
- "Third generation" methods (single molecule)
  - **(1) Single-molecule real-time sequencing (SMRT; Pacific Biosciences)**
  - **(2) Nanopore sequencing (MinION etc.; Oxford Nanopore Technologies)**

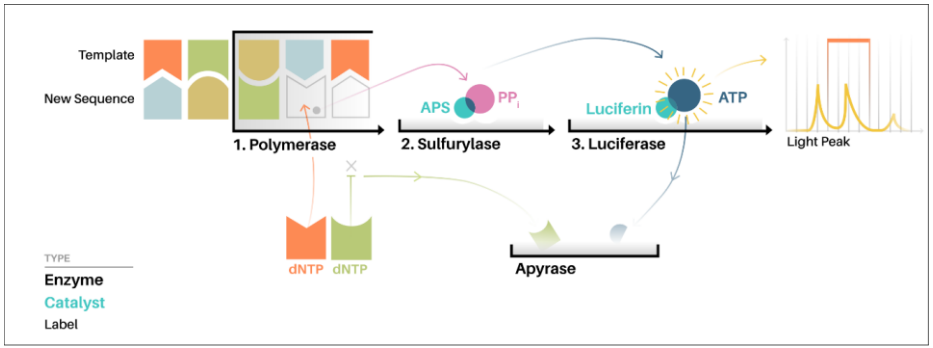


Prof. Dr. Markus Jeschek, LSAM

Pyrosequencing

- **Sequencing by synthesis** (following Sanger approach)
- dNTP incorporation detected via enzyme-coupled assay
  - (1) **DNA polymerase** adds one nt releasing pyrophosphate (PPi)
  - (2) **sulfurylase** converts adenosine 5´-phosphosulfate (APS) and PPi to ATP
  - (3) ATP used by **luciferase** to convert luciferin under light emission
  - (4) unincorporated dNTPs/ATP degraded by **apyrase**





$(NA)_n$ + Nucleotide —Polymerase→ $(NA)_{n+1}$ + PPi

PPi + APS —ATP sulfurylase→ ATP + SO_4^{2-}

ATP + Luciferin + O_2 —Luciferase→ AMP + PPi + Oxyluciferin + CO_2 + Light

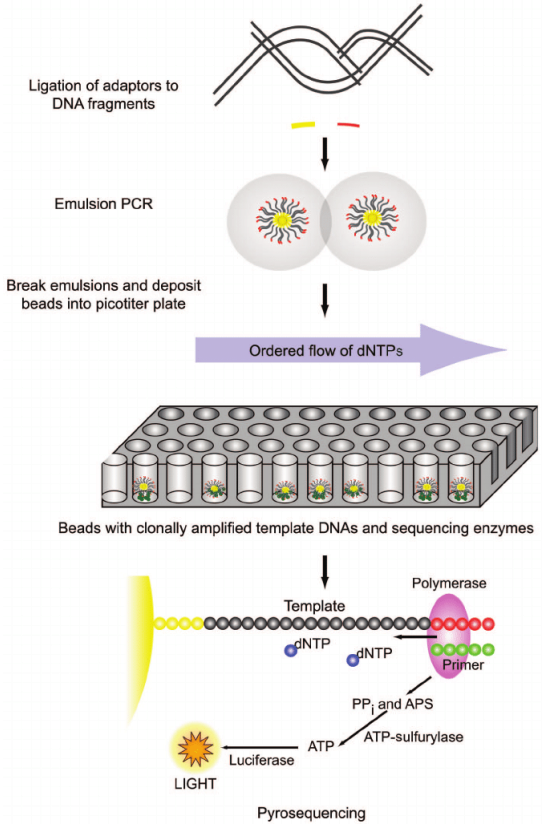
Prof. Dr. Markus Jeschek, LSAM

Pyrosequencing – 454 Life Sci. (later: Roche)

- 1st benchtop HTP sequencer
- Target DNA ligated to "**adaptors**"
- Single-stranded fragments captured on beads (**one per bead!**)
- Emulsion PCR to amplify fragments **monoclonally**
- Decorated beads loaded into picotiter plate (**Poisson distribution!**)
- ~10^6 Wells = reaction vessels (~29 µm diameter)
- dNTPs added one by one

- Limitations:
  - no termination → multiple incorporations, homopolymer read errors
  - expensive
- discontinued 2013

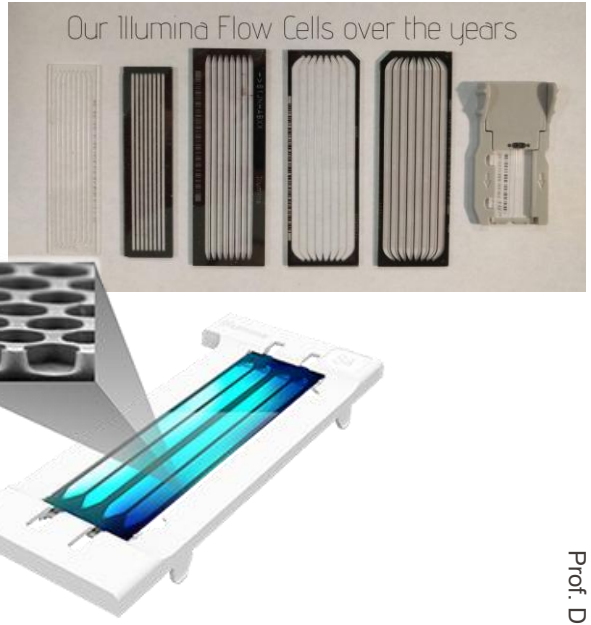
Q: Why is the amplification by PCR needed?
What is "monoclonality" and why is it needed?
What is a Poisson distribution?

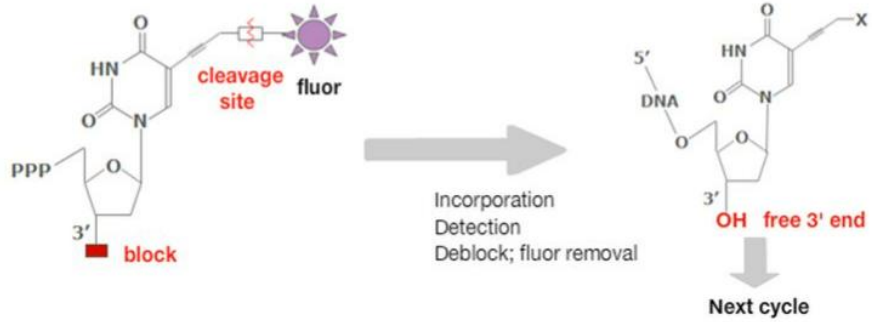


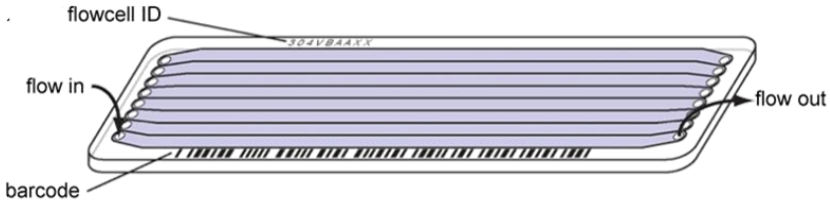
Prof. Dr. Markus Jeschek, LSAM

Sequencing by Synthesis (Illumina)

- "Extension of Sanger principle"
- Most widely used NGS technology today
- Sequencing on chip surface ("flow cells")
- **Fluorescently labelled, "blocked" dNTPs**
- **Reversible termination**: blocked dNTPs converted into dNTPs to continue elongation after each nucleotide/"cycle"

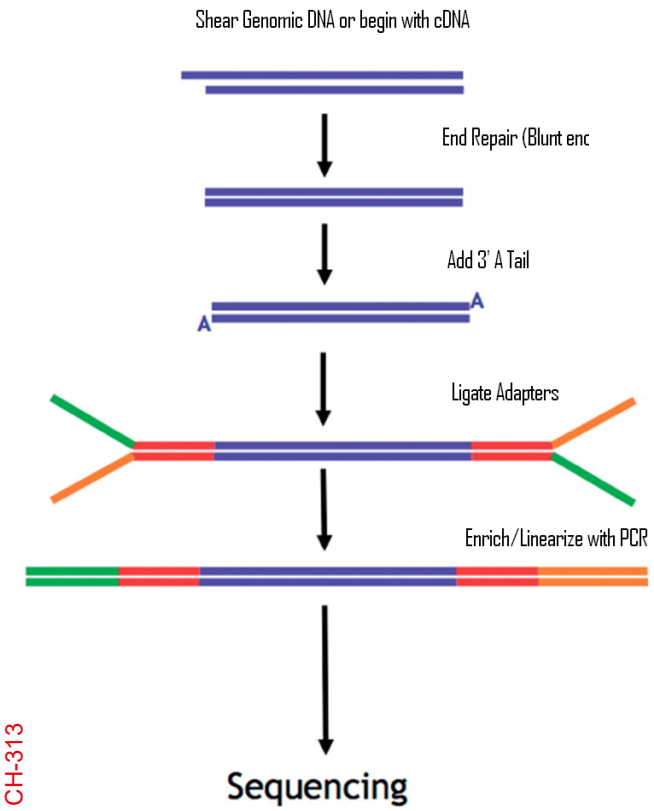


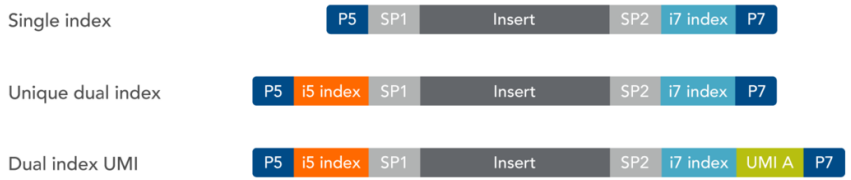




Prof. Dr. Markus Jeschek, LSAM

Illumina Seq. – Sample preparation (1)



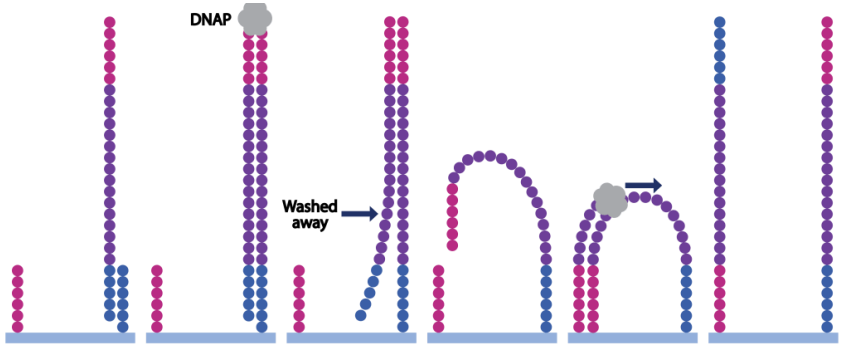


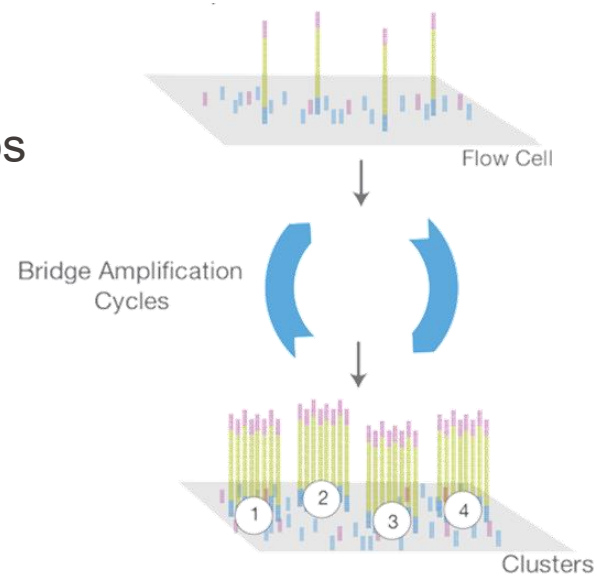
- **Flow cell binding sequence:** Platform-specific sequences for library binding to instrument
- **Sequencing primer sites:** Binding sites for general sequencing primers
- **Sample indexes:** Short sequences specific to a given sample library
- **Molecular index/barcode:** Short sequence used to uniquely tag each molecule in a given sample library
- **Insert:** Target DNA or RNA fragment from a given sample library

Prof. Dr. Markus Jeschek, LSAM

Illumina Seq. – Cluster Amplification (2)

- Isothermal DNA amplification on flow cell surface
- Via “bridges“ using two types of immobilized oligos
- Resulting clusters:
  - **monoclonal** (~1000 copies per cluster)
  - several hundred thousand per mm^2
  - defined by 2D-coordinates on flow cell
  - reverse strands washed off before sequencing



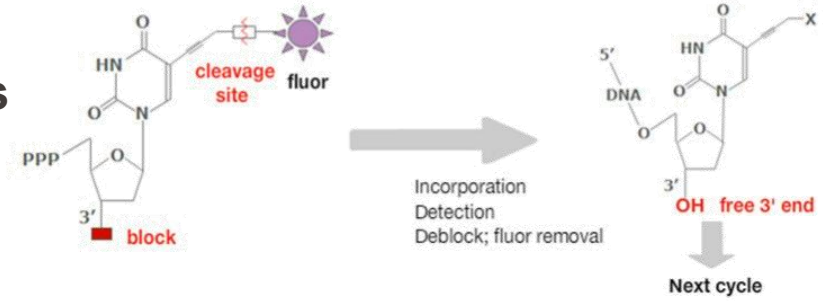


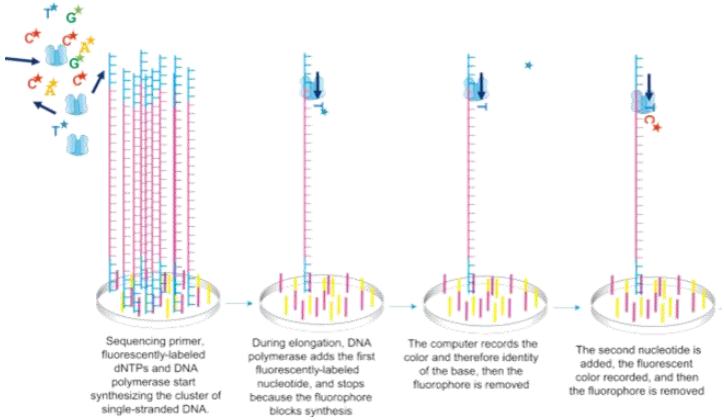
Library is loaded into a flow cell and the fragments are hybridized to the flow cell surface. Each bound fragment is amplified into a clonal cluster through bridge amplification.

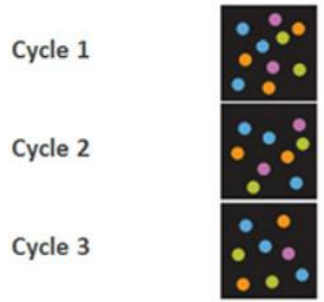
Prof. Dr. Markus Jeschek, LSAM

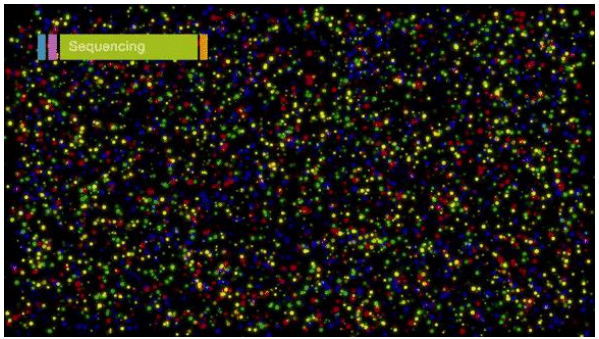
Illumina Seq. – Sequencing (3)

- DNA polymerase extends sequencing primer
- **Fluorescent, reversible termination dNTPs**
- Cycle (once per nt/position)
  - (1) elongation by one nucleotide
  - (2) imaging
  - (3) de-blocking+removal of fluorophore





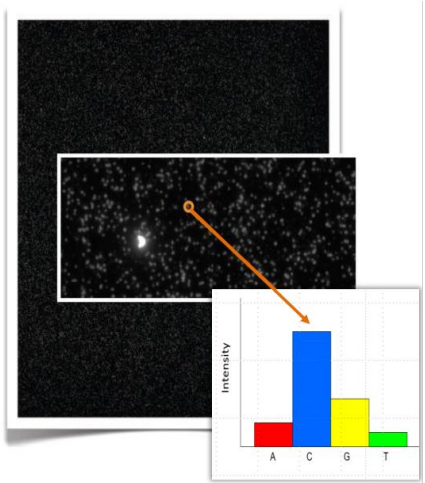




Prof. Dr. Markus Jeschek, LSAM

Illumina Seq. – Data Processing (4)

- Images analysed in real time (each cycle+channel)
  - “base calling” and quality scores obtained for each cluster
  - images deleted
  - FASTQ file produced
- FASTQ files
  - line 1: unique read ID (instrument/flow cell ID, lane, cluster coordinates, index, PE: forward or reverse read etc.)
  - **line 2: sequence**
  - line 4: quality score (“Phred“, ASCII code)




```
@HISEQ:126:H14YJADXX:1:1101:1118:2101 1:N:0:ATCACG
CTCCATAGTCAGAAACTTCAGCATGACAGTACCTCATGCTGCATCAGGTGATCATGAAAAGATTAC
+
@@?ADDDD?ADHDIIIIIIIEIIIGEFHC<?FH4C9E9BGAFIGH<DG9BD?@DGGEGHHG<DCBB
```

Phred quality scores are logarithmically linked to error probabilities

Phred Quality Score	Probability of incorrect base call	Base call accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%
60	1 in 1,000,000	99.9999%

$$\text{Phred} = -10 \log_{10} p$$

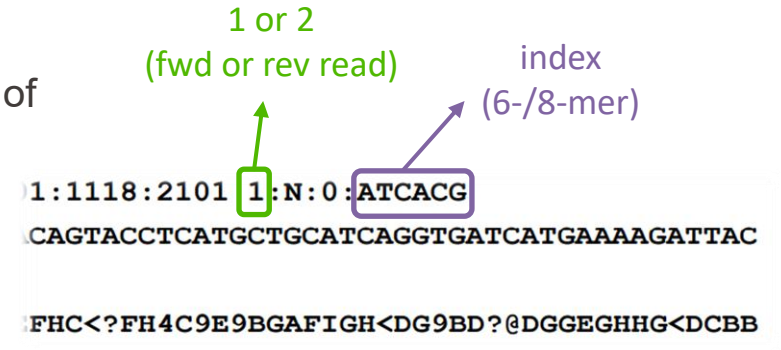
p = Probability call is incorrect

https://www.youtube.com/watch?v=fCd6B5HRaZ8

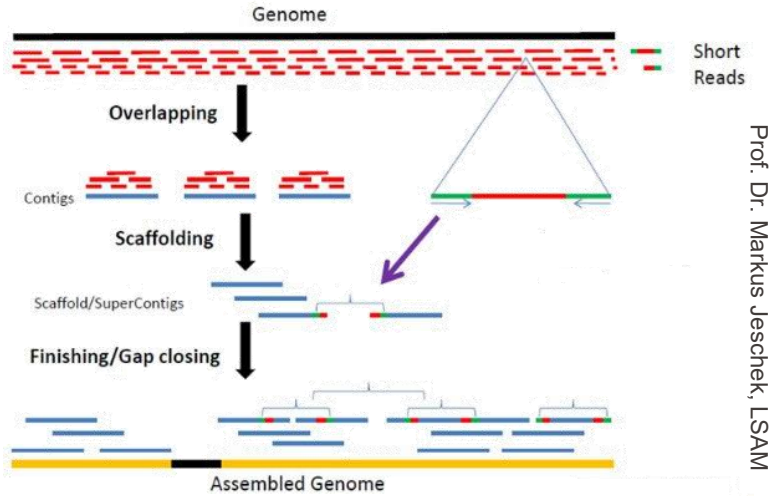
Prof. Dr. Markus Jeschek, LSAM

Illumina Seq. – Data Analysis (5)

- From FASTQ files
  - optional: de-multiplexing of samples (deconvolution of index combinations)
  - optional: pairing of reads (paired-end reading)



- E.g. for whole-genome assembly…
  - short reads “stitched“ together in silico via overlaps
  - contigs → scaffolds → entire assembly
  - “oversampling“ for statistical coverage (30-200x)
  - mapping to reference genomes



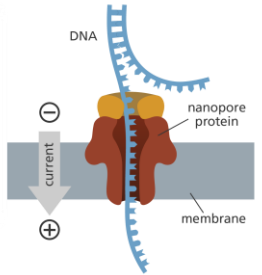
- Different procedures for each application (amplicon sequencing etc.)

NGS – Main Methods (Overview)

- "Second generation" methods
  - (1) Pyrosequencing (Roche 454)
  - (2) Sequencing by synthesis (Illumina)
  - (3) Sequencing by ligation
  - (4) Ion semiconductor sequencing

illumina®

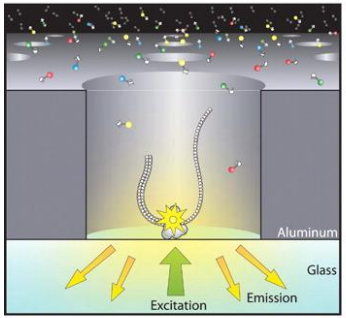
- "Third generation" methods (single molecule)
  - **(1) Single-molecule real-time sequencing (SMRT; Pacific Biosciences)**
  - **(2) Nanopore sequencing (MinION etc.; Oxford Nanopore Technologies)**

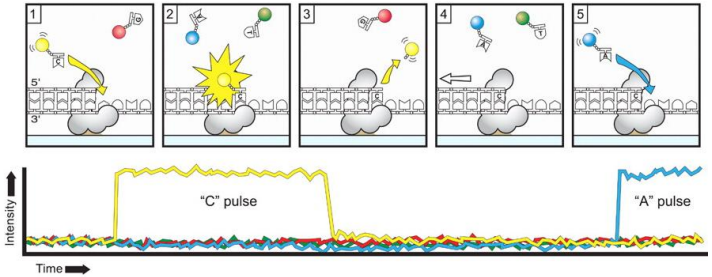


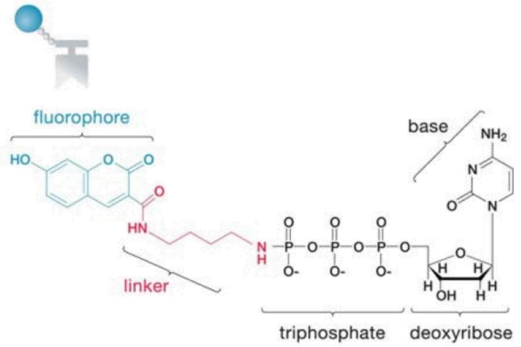
Oxford NANOPORE Technologies

Single-molecule Real-time (SMRT) Sequencing

- Pacific Biosciences (PacBio)
- **Zero-mode waveguides (ZMWs)** contain…
  - single DNA polymerase molecule, immobilized
  - single molecule of DNA template
  - “phospholinked” nucleotides
  - reaction vessel: (70 nm diameter, 100 nm depth)
- Principle: labelled nucleotides incorporated by DNA pol. → labels are held a little longer in the ZMW than average diffusion → detected as a “flash“



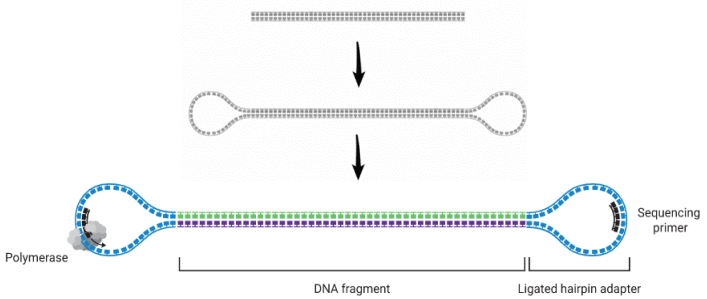


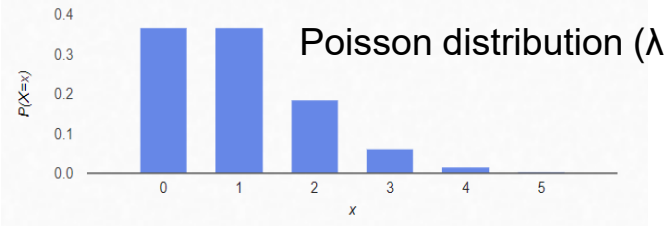


Prof. Dr. Markus Jeschek, LSAM

Single-molecule Real-time (SMRT) Sequencing

- Brief procedure:
  - linear DNA fragments ligated to bell-shaped adapters
  - DNA loaded on SMRT cell (Poisson distribution!)
  - Reading is performed in a circle (multiple times)
- Long-read technology (30 kb average, good for repetitive sequences!)
- Up to 4,000,000 reads
- High error rates (5-15% per base)
- Fast (<< 24 h)
- Some modified bases can be directly detected (label-free!)
  - mainly methylated bases (e.g. N^6-methyladenine, N^4-methylcytosine)
  - epigenetics!

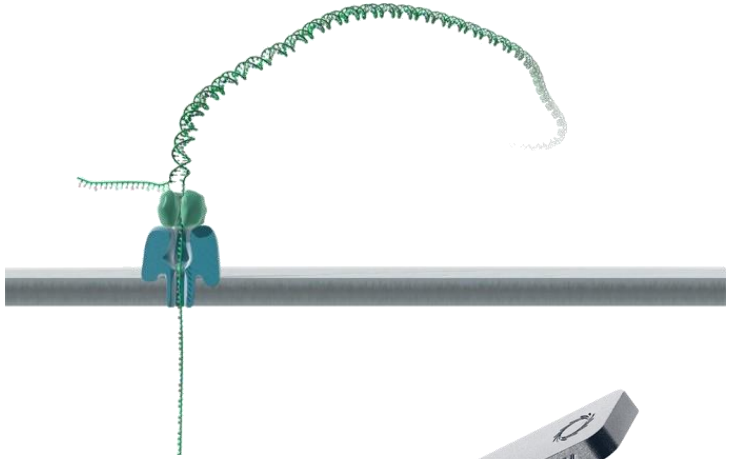


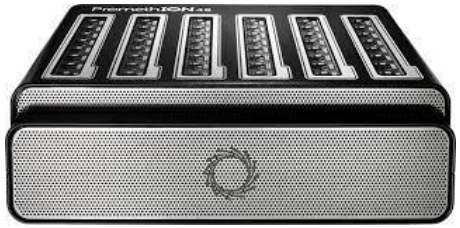


Prof. Dr. Markus Jeschek, LSAM

Nanopore Sequencing

- Oxford Nanopore Technologies (MinION, PromethION etc.)
- Label-free, single-molecule technique
- Portable equipment, real-time results
- Long-read technology (> 4Mbp successfully demonstrated)
- 512 – 2,675 pores per flow cell, repeated passage possible
- Error rates of 3-8% (lately improving drastically)
- Modified bases and other molecules (RNA, proteins) can be directly “sequenced“

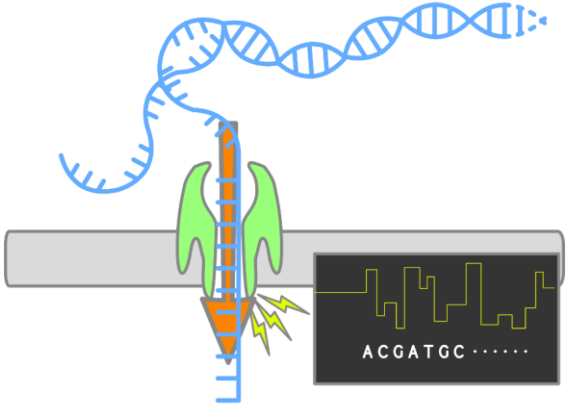


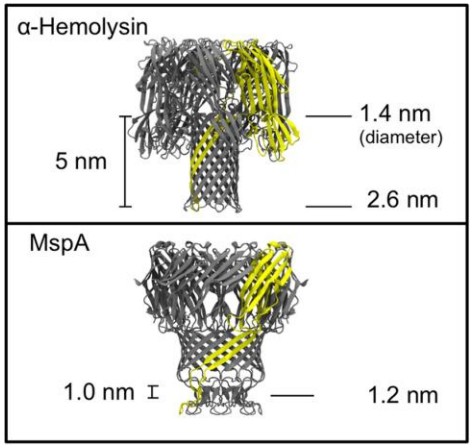


Prof. Dr. Markus Jeschek, LSAM

Nanopore Sequencing - Principle

- single-stranded DNA/RNA molecules are “pushed“ through nanopore via processive enzyme (e.g. DNA helicase)
- pore embedded in membrane and surrounded by electrolyte
- electric field across the membrane → electrophoretic motion of ions through pore
- if a larger molecule (e.g. DNA strand) occupies pore, ion flux is disrupted (detectable by voltage change in real time)
- voltage changes are specific for base/molecules





- pores: biological (α-hemolysin, MspA) or solid-state (metal, metal alloy)
- mostly synthetic membranes

https://www.youtube.com/watch?v=E9-Rm5AoZGw

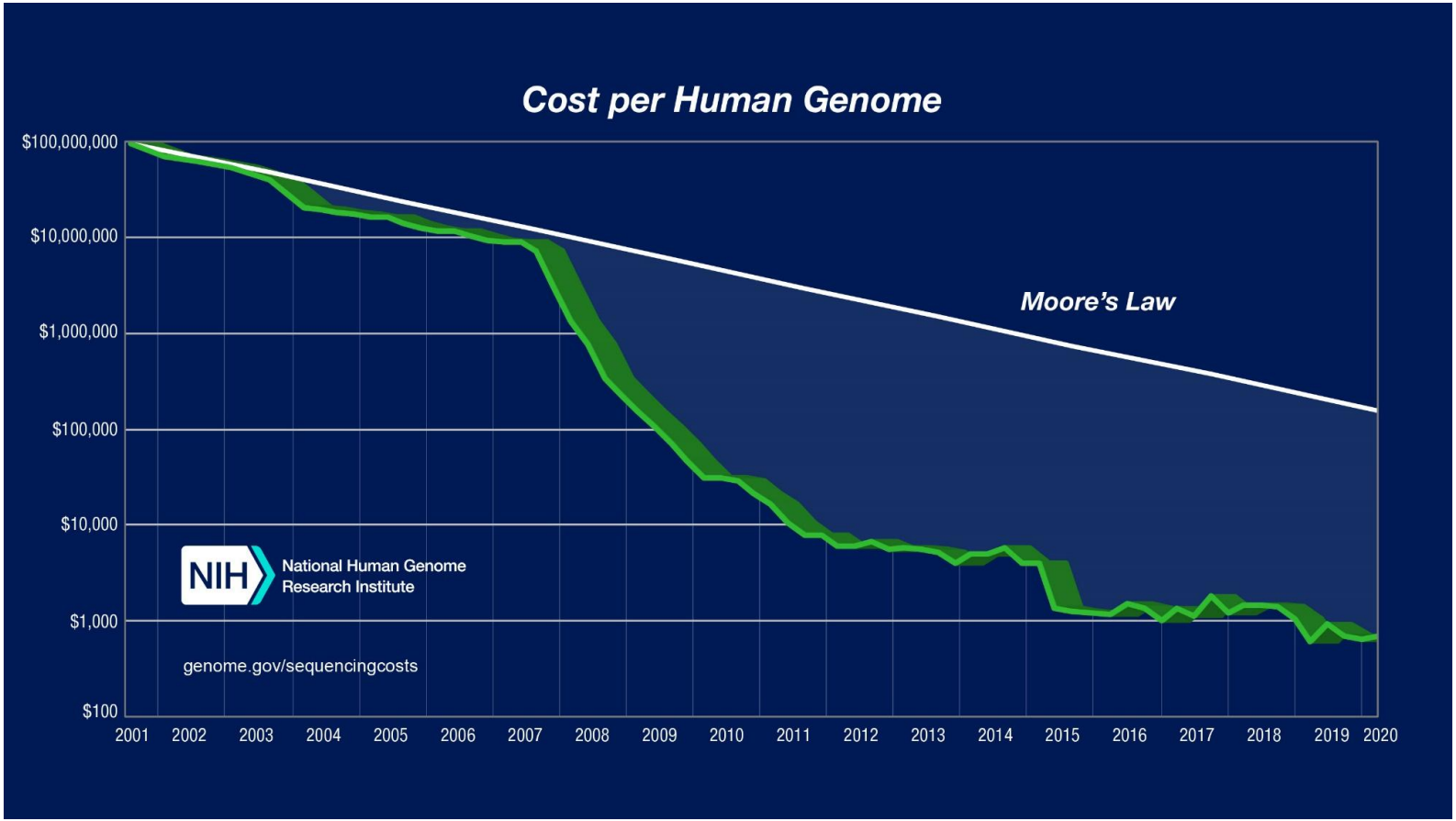
Prof. Dr. Markus Jeschek, LSAM

NGS – Comparison of Main Methods (approximate numbers)

Method	max. read length [bp]	error rate [%]	max. reads per run	time per run [h]	single molecule?	costs per Gb [US-$]	Remarks
Pyrosequencing (Roche)	700	0.1-1	1M	24	no	10k	discontinued, expensive, homopolymer errors
Sequencing by synthesis (Illumina)	50-600	**0.1-1**	**52B**	4-48	no	**2-150**	expensive equipment, cheap Gb price, low error rates
Ion semiconductor (Ion Torrent)	600	~0.5	80M	**2**	no	50-1000	cheap equipment, very fast, homopolymer errors
SMRT sequencing (PacBio)	**30k-100k**	5-15	4M	0.5-20	**yes**	5-50	expensive equipment, long reads, fast, methylation
Nanopore Sequencing (Oxford)	**> 4000k**	3-8	dep. on length (~few 100k)	72	**yes**	5-100	handheld, cheap equipment, longest reads, other molecules
Sanger	1200	**0.01**	1	0.2-3	no	2-3M	gold standard, low throughput

Prof. Dr. Markus Jeschek, LSAM

DNA Sequencing – Costs



Prof. Dr. Markus Jeschek, LSAM

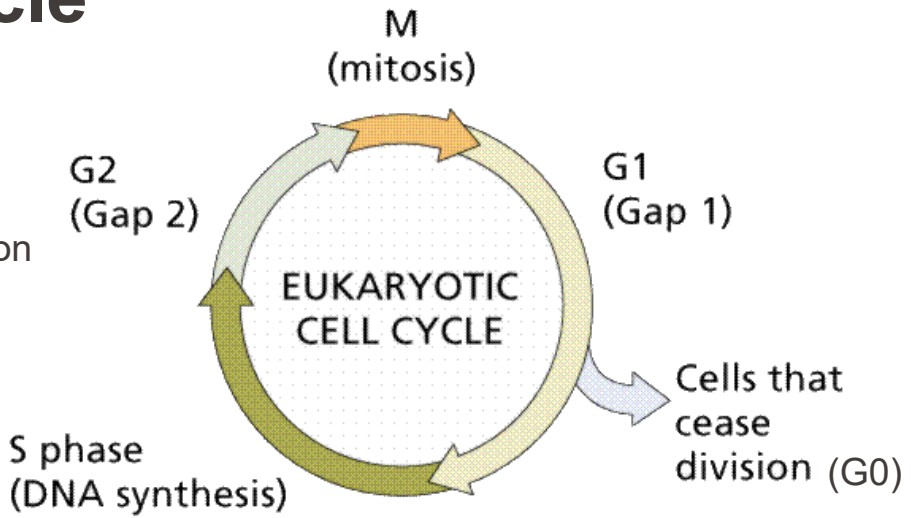
DNA

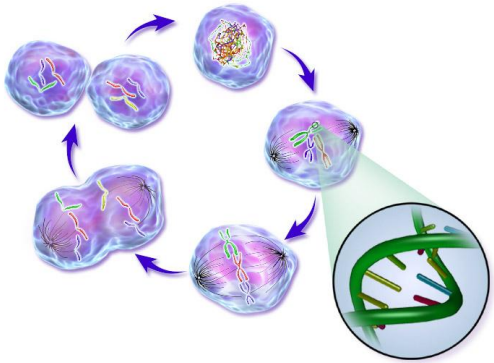
("How to interrogate and manipulate DNA in cells?")

Prof. Dr. Markus Jeschek, LSAM

The Eukaryotic Cell Cycle

- Four main phases
  - **G1** (Gap 1)
    - "normal" functions, preparations for division
  - **S** (Synthesis)
    - DNA replication
  - **G2** (Gap 2)
    - continued growth
  - **M** (Mitotic phase)
    - mitosis (nuclear division) & cytokinesis (cytoplasmic division)



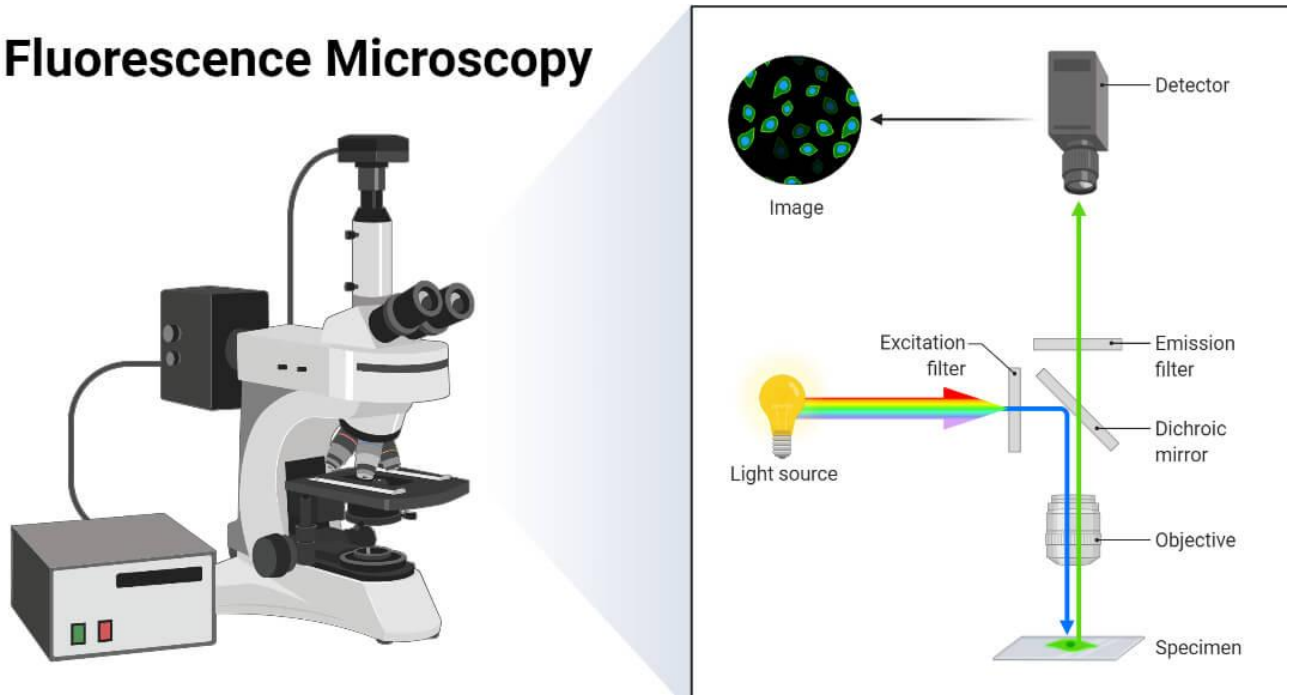


- Checkpoints for transition between phases (e.g. G1/S, G2/M)
- Interphase = G1+S+G2
- G0 = resting phase

RECAP

Prof. Dr. Markus Jeschek, LSAM

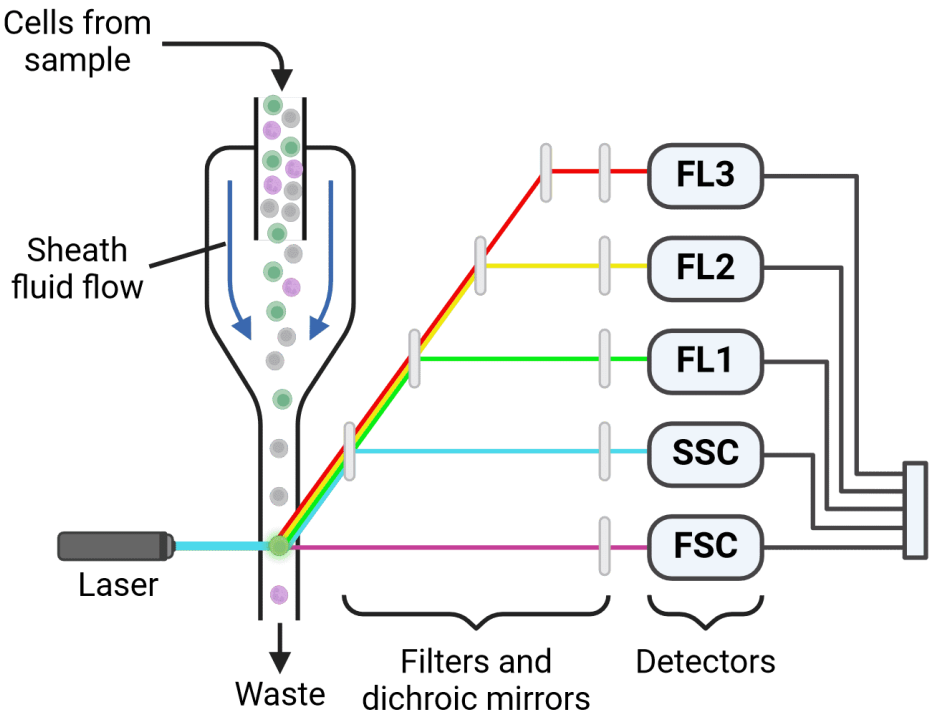
Fluorescence Microscopy

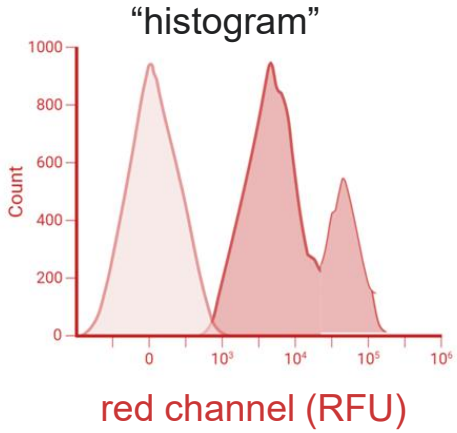


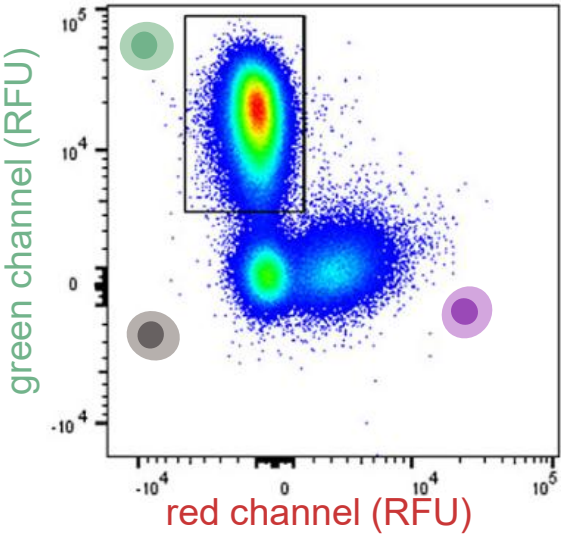
Prof. Dr. Markus Jeschek, LSAM

RECAP

Flow Cytometry





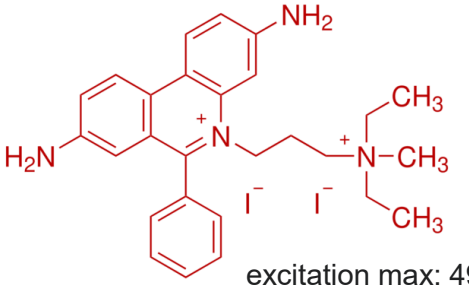


RECAP

Prof. Dr. Markus Jeschek, LSAM

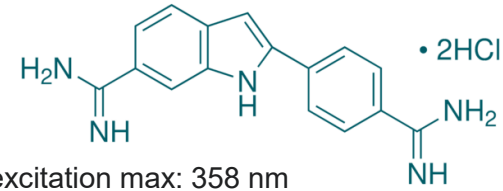
DNA Intercalators

- Propidium iodide (PI)
  - stains DNA (and RNA) intercalating between bases
  - membrane impermeable
  - “dead cell stain”



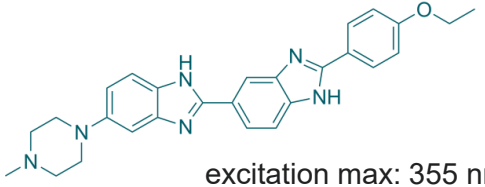
excitation max: 493 nm
emission max: 636 nm (red)

- DAPI (4′,6-diamidino-2-phenylindole)
  - stains dsDNA intercalating at AT-rich regions
  - binds minor groove
  - membrane permeable



excitation max: 358 nm
emission max: 461 nm (blue)

- Hoechst 33342
  - stains dsDNA intercalating at AT-rich regions
  - binds minor groove
  - membrane permeable

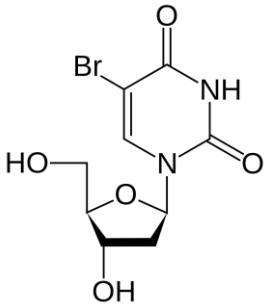


excitation max: 355 nm
emission max: 465 nm (blue)

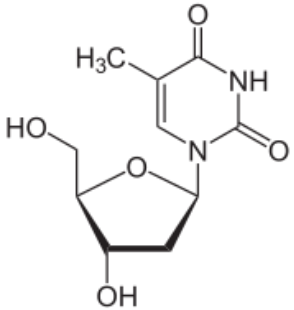
Prof. Dr. Markus Jeschek, LSAM

Nucleoside Analogues – Labelling

- BrdU (5-bromo-2'-deoxyuridine, “broxuridine”)
  - incorporates in DNA instead of T
  - detectable with anti-BrdU antibodies (fixed samples!)
  - cancerogenic
  - *in vivo* use possible

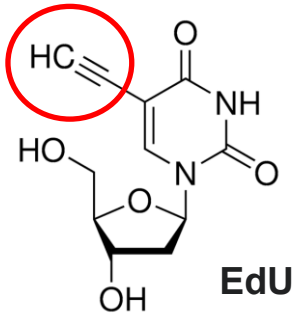


BrdU



thymidine

- EdU (5-ethynyl-2’-deoxyuridine)
  - incorporates in DNA instead of T
  - labelling with “clickable” dyes
  - no denaturing required!
  - DNA damage via interstrand crosslinking



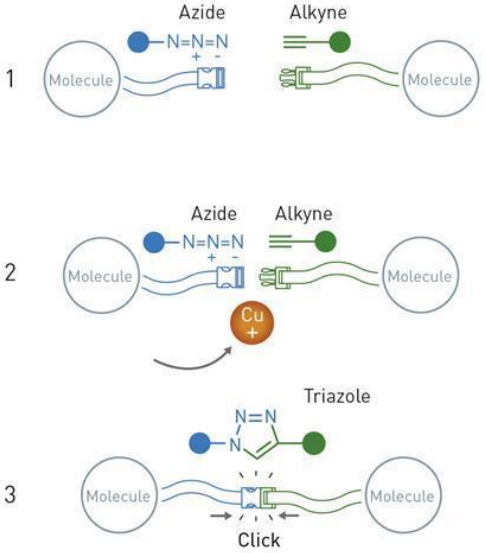
EdU

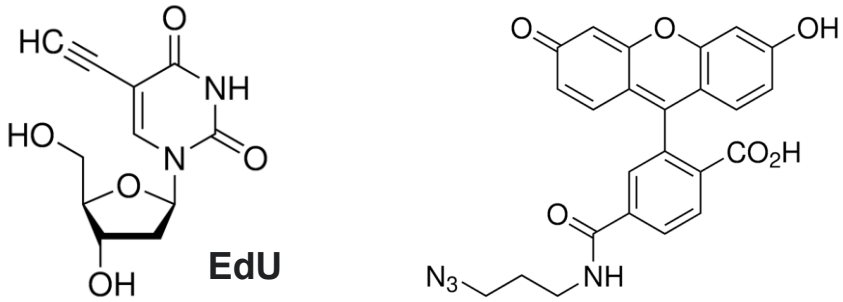
Click Chemistry

- Nobel Prize Chemistry 2022: "…for the development of click chemistry and biorthogonal chemistry"
  - labelling/modification through azides and alkynes
  - minimal invasive for biological systems!
  - simple, modular
  - e.g. copper(I)-catalyzed cycloaddition (CuAAC)
  - e.g. strain-promoted cycloaddition (SPAAC)

Source: © Nobel Prize Outreach/Niklas Elmehed

The 2022 chemistry laureates (left to right) Carolyn Bertozzi, Morten Meldal and Barry Sharpless

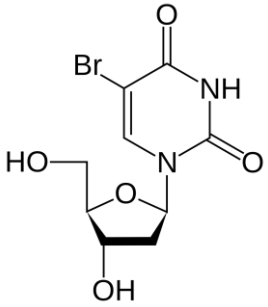




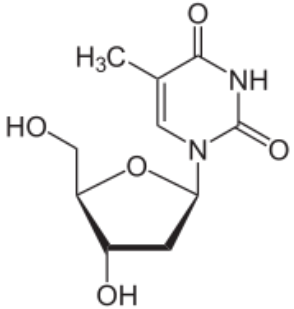
Prof. Dr. Markus Jeschek, LSAM

Nucleoside Analogues – Labelling

- BrdU (5-bromo-2'-deoxyuridine, “broxuridine”)
  - incorporates in DNA instead of T
  - detectable with anti-BrdU antibodies (fixed samples!)
  - cancerogenic
  - *in vivo* use possible

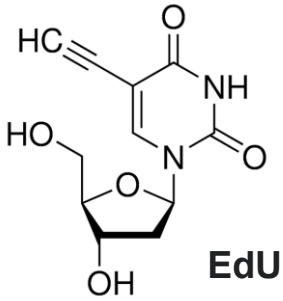


BrdU



thymidine

- EdU (5-ethynyl-2’-deoxyuridine)
  - incorporates in DNA instead of T
  - labelling with “clickable” dyes
  - no denaturing required!
  - DNA damage via interstrand crosslinking

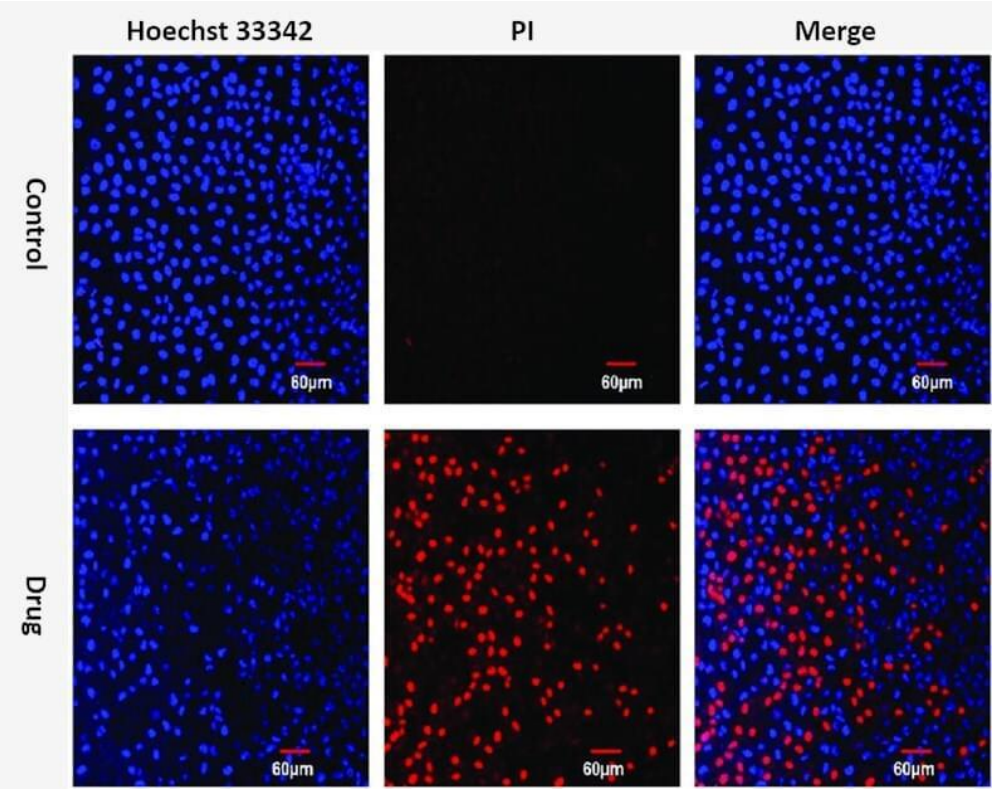


EdU

Q: In which phase of the cell cycle do these analogues act?

Prof. Dr. Markus Jeschek, LSAM

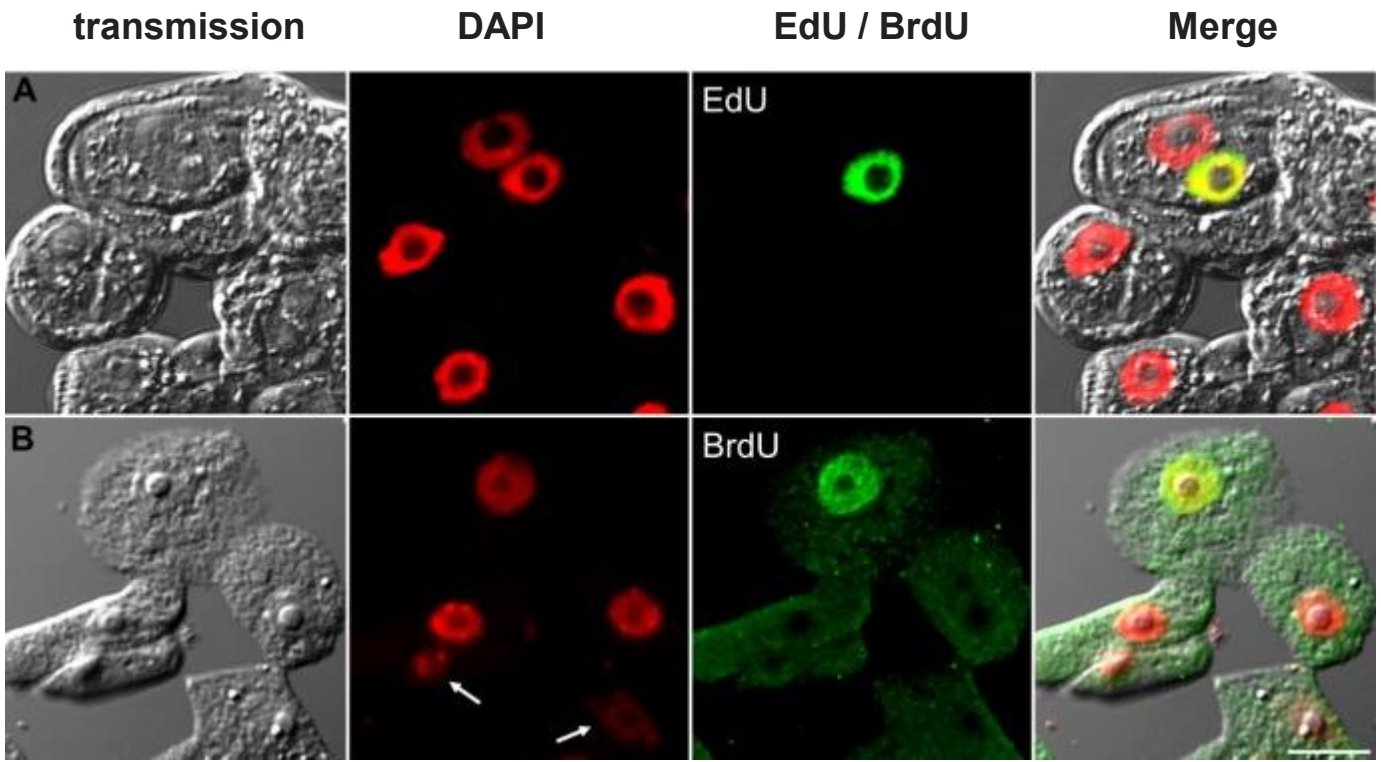
Labelling – Examples



Q: What does the drug do here?

Prof. Dr. Markus Jeschek, LSAM

Labelling – Examples



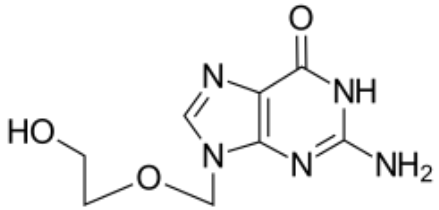
Q: What can you say about the green cells?

How would these samples look in flow cytometry (2D plot)?

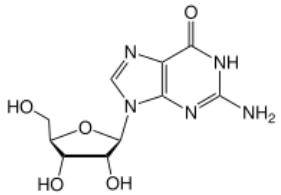
Prof. Dr. Markus Jeschek, LSAM

Nucleoside Analogues – Therapeutics

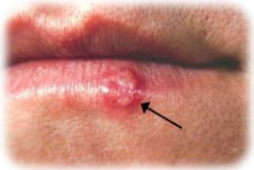
- Aciclovir (Zovirax)
  - antiviral medication
  - herpes simplex virus (amongst others)
  - inhibits viral DNA polymerase



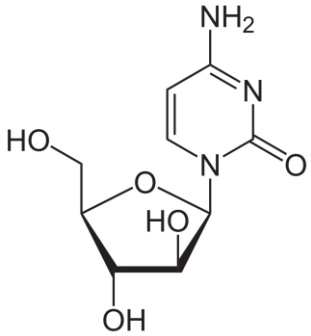
acyclovir



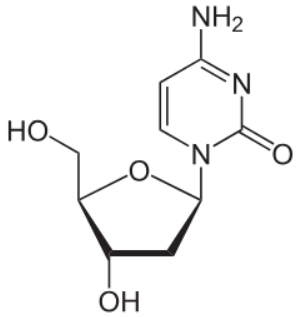
guanosine



- Cytarabine (cytosine arabinoside)
  - chemotherapeutic medication
  - different forms of leukemia
  - incoporates in DNA, DNA damage (in S phase)
  - inhibits DNA and RNA polymerases



cytosine arabinoside



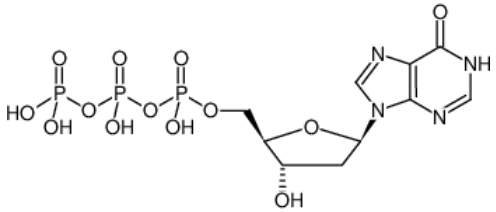
deoxycytidine

Q: What could be a second mode of action for acyclovir?

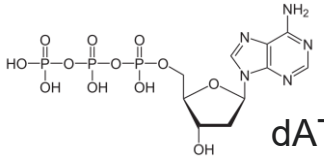
What must these drugs not do to be of medical use?

Nucleotide Analogues for Mutagenesis

- Deoxyinosine-triphosphate (dITP)
  - pairs with A, G or C
  - mostly A→G / G→A mutations

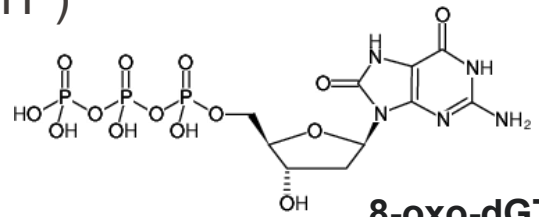


dITP

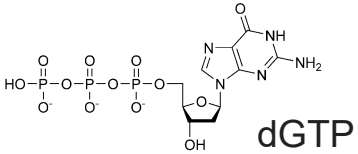


dATP

- 8-oxo-2'-deoxyguanosine (8-oxo-dGTP)
  - pairs with C and A
  - various mutations (mainly G→T/A)



8-oxo-dGTP



dGTP

- Used in error-prone PCR (epPCR)

→ will be discussed in “Engineering” part of the lecture

Prof. Dr. Markus Jeschek, LSAM