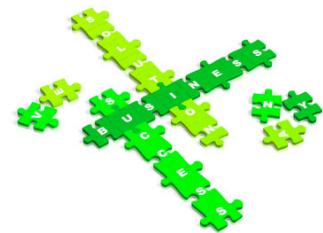


# Next Generation Sequencing – An Overview

Olga Vinnere Pettersson, PhD  
National Genomics Infrastructure hosted by  
ScilifeLab,  
Uppsala Node (UGC)

# Outline:



- 4 slides about history
- NGS technologies
- NGS applications
- NGS sample quality requirements
- Philosophical reflection
- National Genomics Infrastructure – Sweden

# Once upon a time...

- Fredrik Sanger and Alan Coulson  
Chain Termination Sequencing (1977)  
Nobel prize 1980

Principle:

SYNTHESIS of DNA is randomly **TERMINATED** at different points

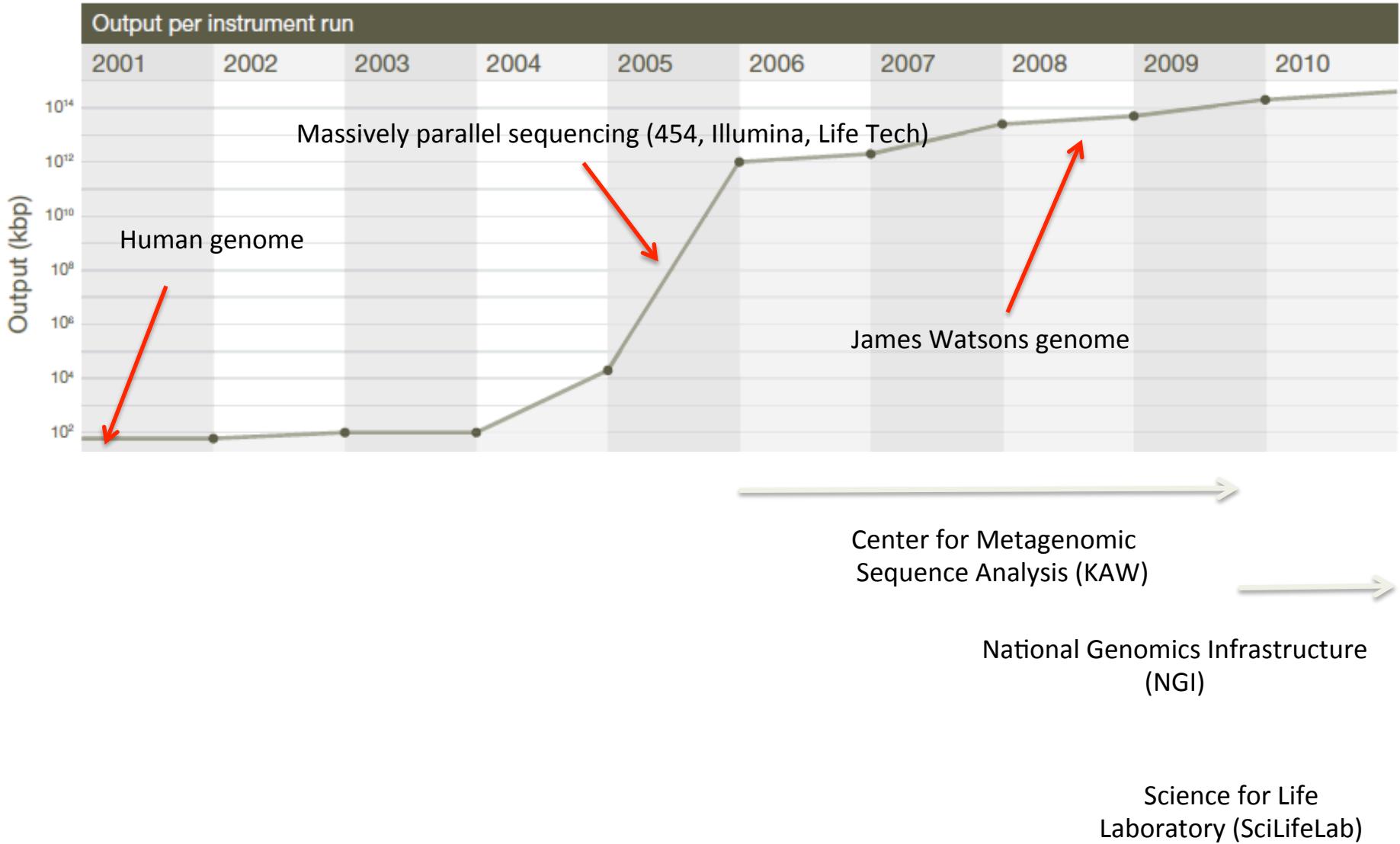
Separation of fragments that are 1 nucleotide different in size



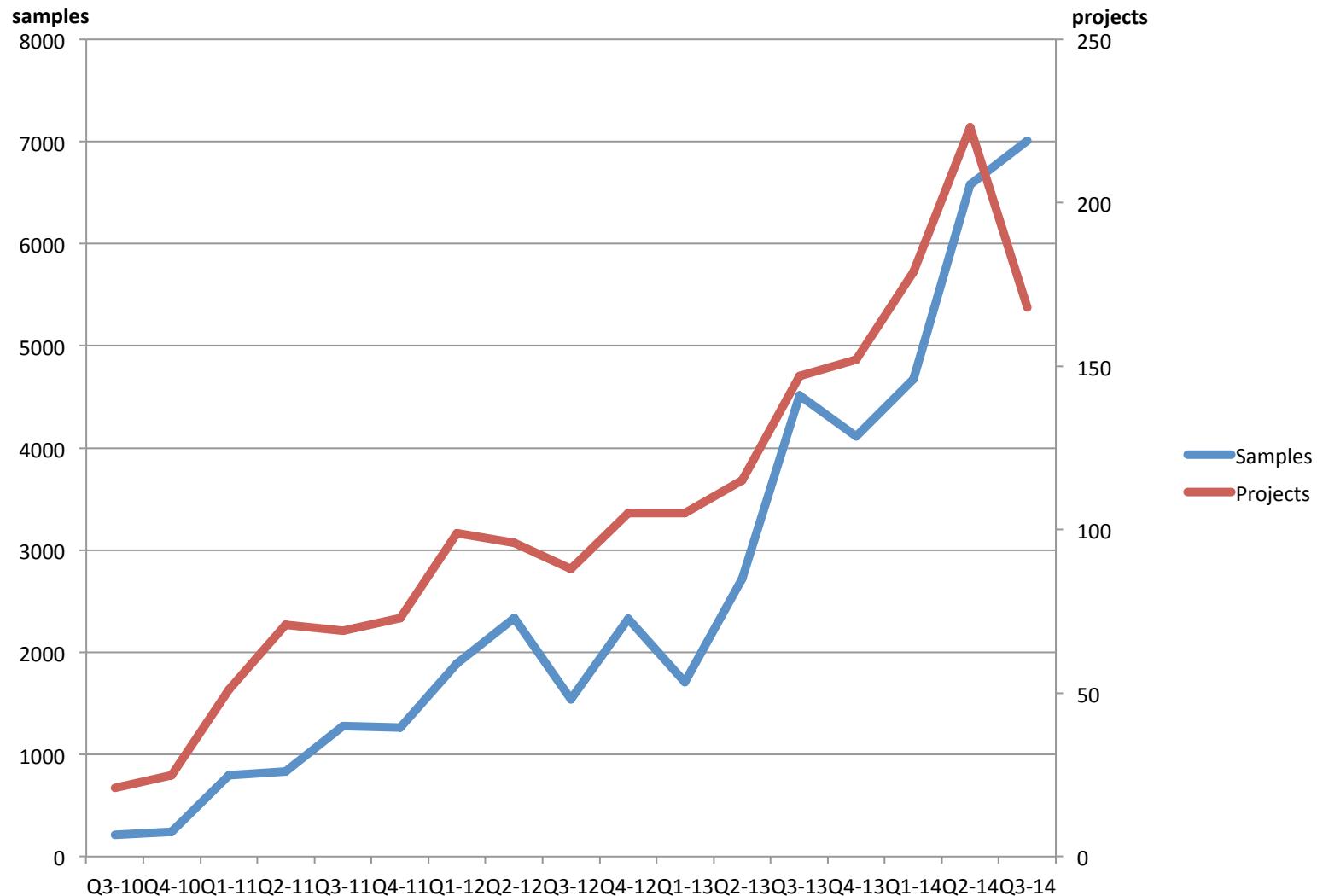
# Sequencing genomes using **Sanger**'s method

- Extract & purify genomic DNA
- Fragmentation
- Make a clone library
- Sequence clones
- Align sequences ( -> contigs -> scaffolds)
- Close the gaps
- Cost/Mb=1000 \$, and it takes TIME

# DNA sequencing revolution - Sweden



# Workload at NGI – Sweden 2010-2014



# NGS technologies

Company	Platform	Amplification	Sequencing method
Roche	454 (until 2016)	emPCR	Pyrosequencing
Illumina	HiSeq, MiSeq NextSeq, X10	Bridge PCR	Synthesis
LifeTechnologies (Thermo Fisher)	Ion Torrent, Ion Proton, S5	emPCR	Synthesis (pH)
Pacific Biosciences	RSII	None	Synthesis (SMRT)
Complete genomics	Nanoballs	None	Ligation
Oxford Nanopore*	MinION GridION	None	Flow

RIP technologies: Helicos, Polonator, SOLiD, 454 etc.

In development: Tunneling currents, nanopores, etc.

# Differences between platforms

- Technology: chemistry + signal detection
- Run times vary from hours to days
- Production range from Mb to Gb
- Read length from <100 bp to > 20 Kbp
- Accuracy per base from 0.1% to 15%
- Cost per base

# Illumina

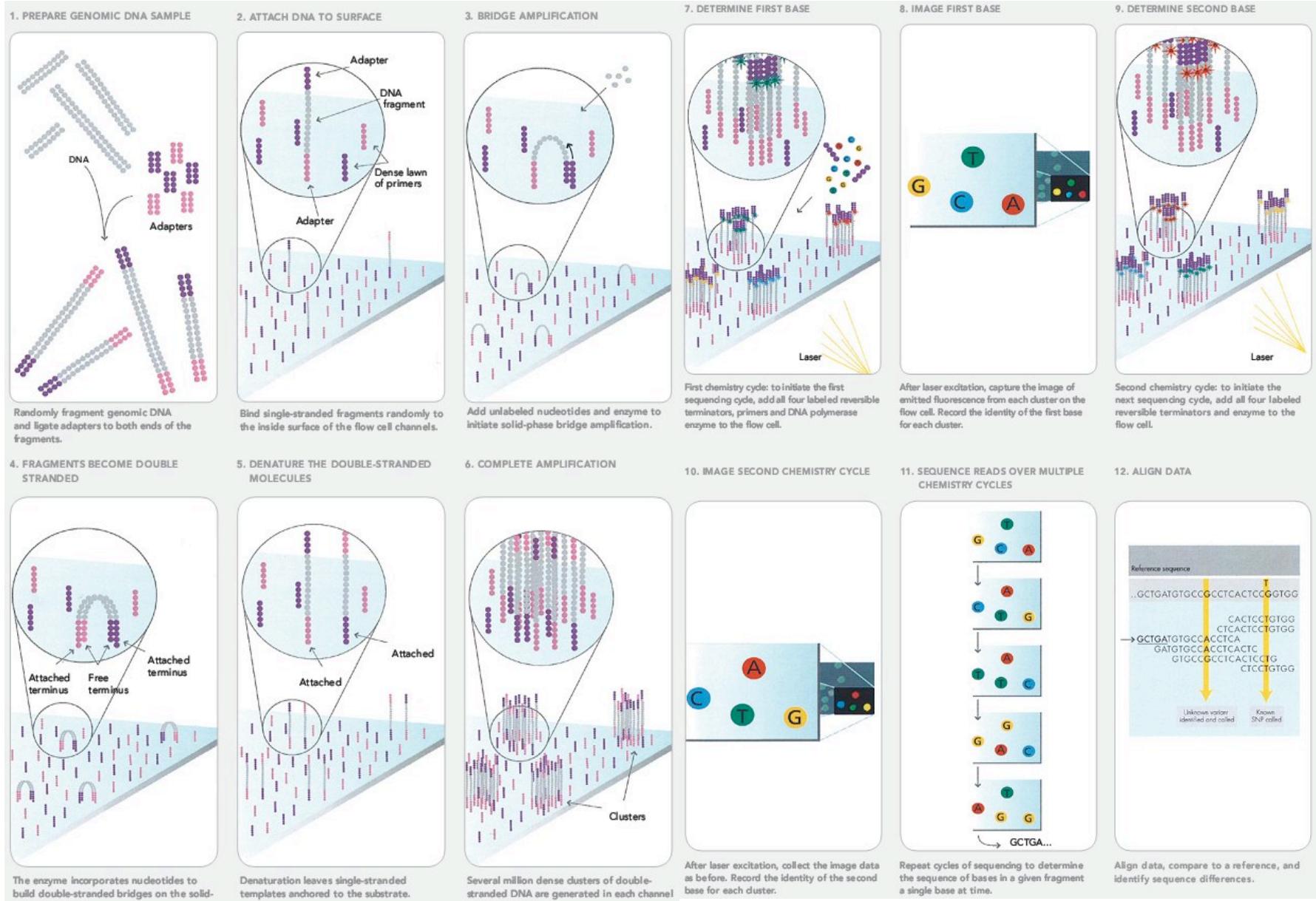
Instrument	Yield and run time	Read Length	Error rate	Error type
HiSeq2500	120 Gb – 600 Gb 27h or standard run	100x100 (250x250)	0.1%	Subst
MiSeq	540 Mb – 15 Gb (4 – 48 hours)	Up to 350x350	0.1%	Subst
HiSeqXten	800 Gb - 1.8 Tb (3 days)	150x150	"	"

## Main applications

- Whole genome, exome and targeted reseq
- Transcriptome analyses
- Methylome and ChIPSeq
- Rapid targeted resequencing (MiSeq)
- Human genome seq (Xten)



# Illumina



# Life Technologies - Ion Torrent & Ion Proton

Chip	Yield - run time	Read Length
314, 316, 318 ( <b>PGM</b> )	0.1 – 1 Gb 3 hrs	200 – 400 bp
P-I ( <b>Proton</b> )	10 Gb 4 hrs	200 bp
520, 530, 540 ( <b>S5</b> )	1 Gb – 10 Gb 3 hrs	400 bp (except 540)

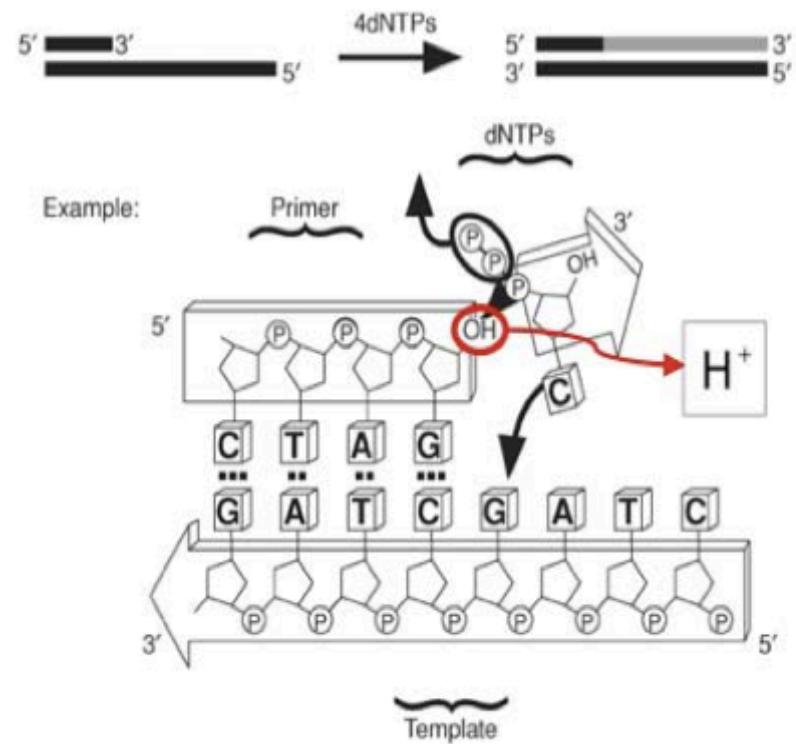
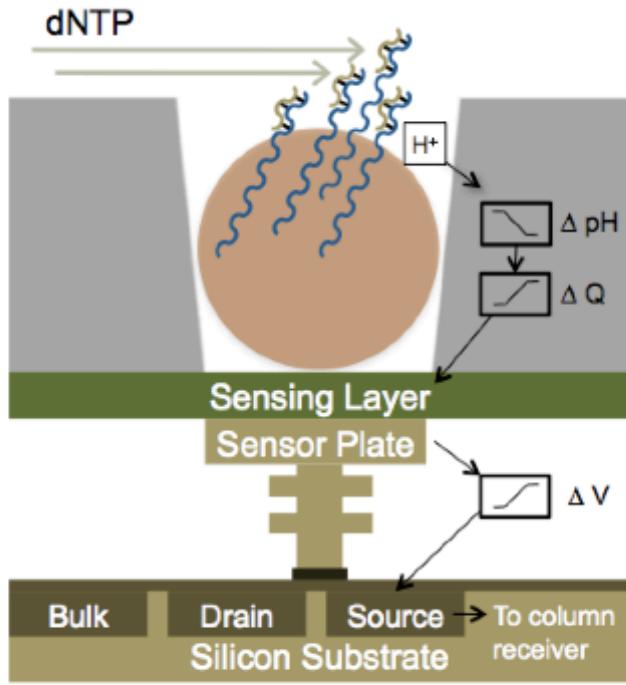
Ion Torrent's PGM



## Main applications

- Microbial and metagenomic sequencing
- Targeted re-sequencing (gene panels)
- Clinical sequencing

# Ion Torrent - H<sup>+</sup> ion-sensitive field effect transistors





314 chip

**10 Mb**

316 chip

**100 Mb**

**200 – 400 bp**

**virus, bacteria, small eukaryote**

318 chip

**1 Gb**

**S5**

**200 bp**

**eukaryote**



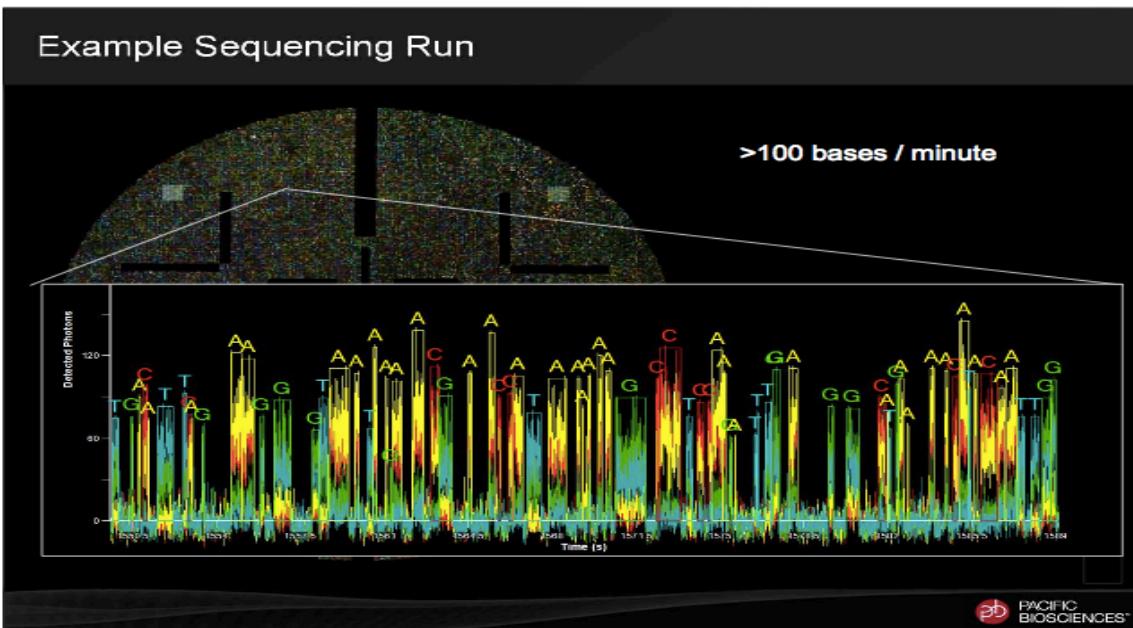
PI chip

**10 Gb**

# PacBio SMRT-technology

Instrument	Yield and run time	Read Length	Error rate	Error type
RS II	250 Mb – 1.3 Gb / 30 - 240 min SMRTCell	250 bp – 30 000 bp (70 000 bp)	15% (on a single passage!)	Insertions, random

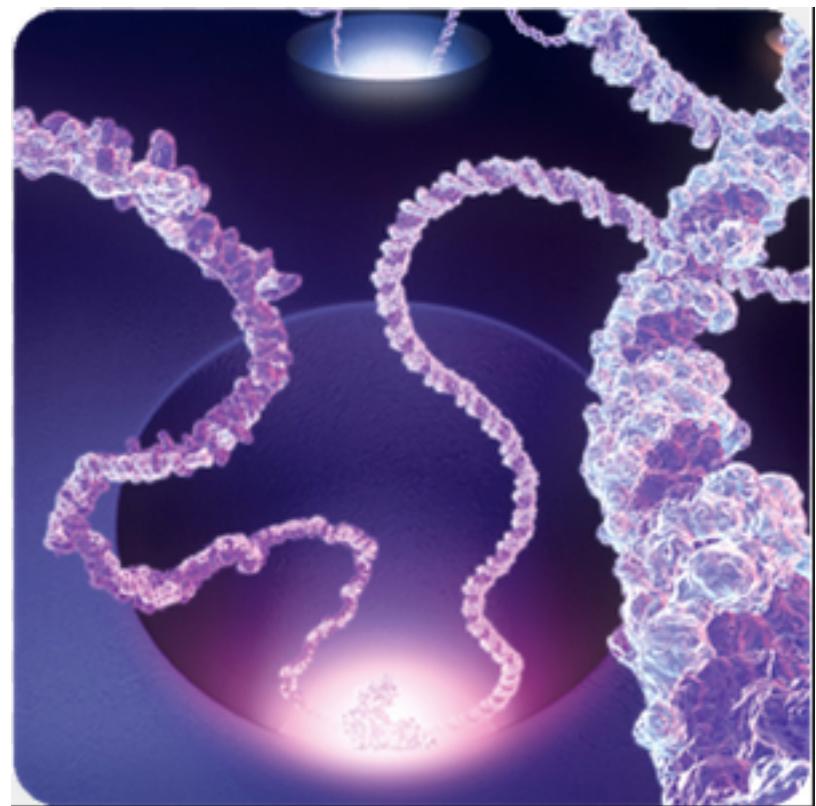
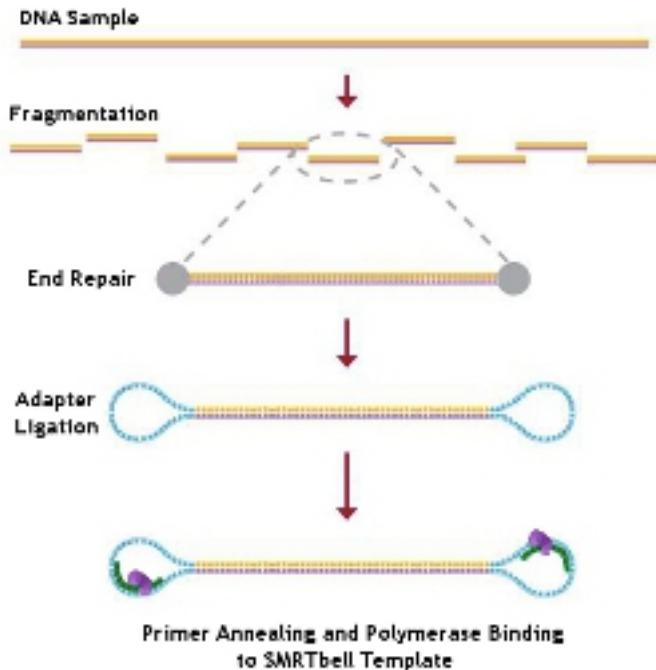
Single-Molecule, Real-Time DNA sequencing



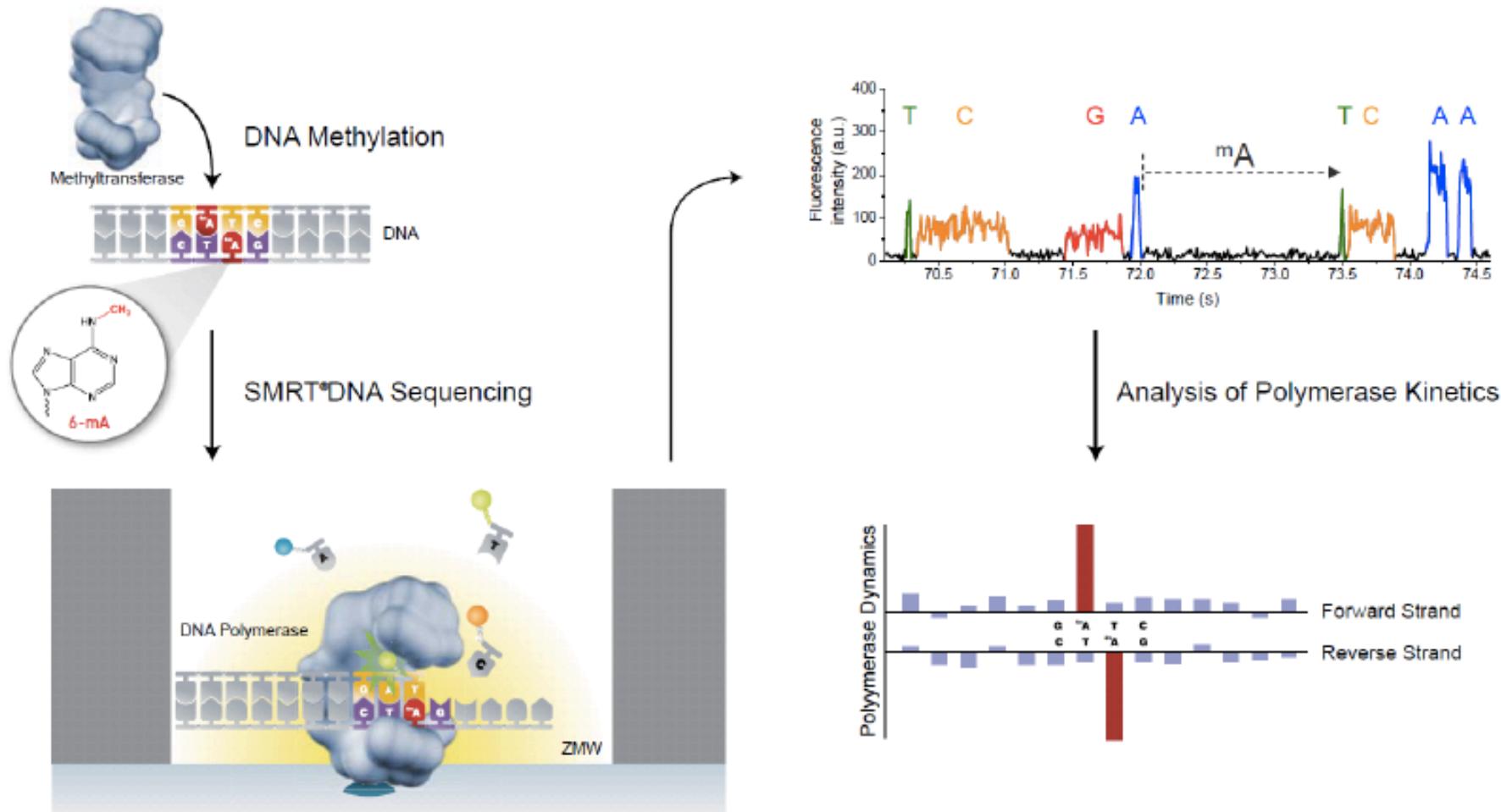
# PacBio SMRT - technology



## Single Molecule Real Time

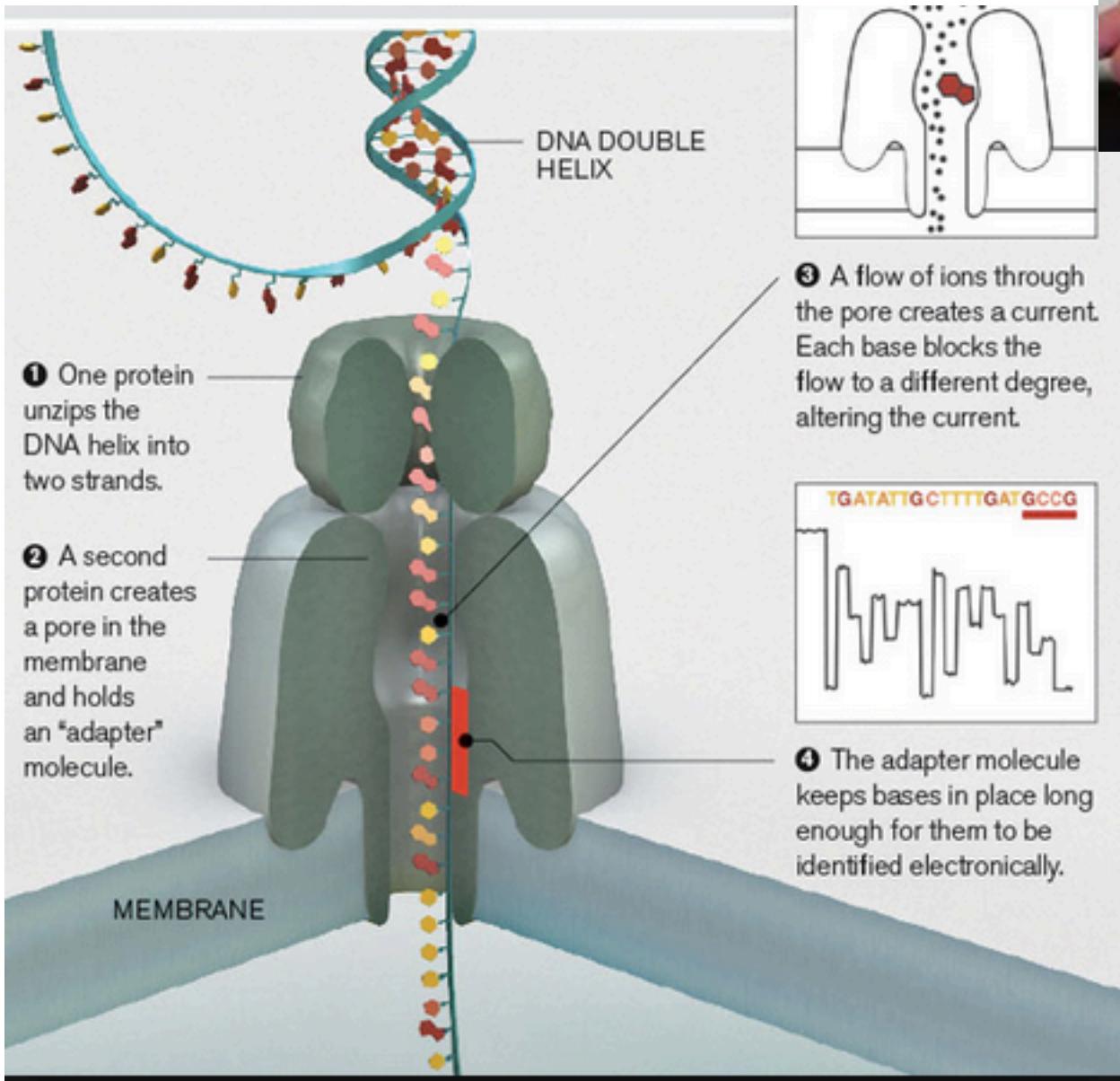


# Base Modification: Discover the Epigenome



Detect base modifications using the kinetics of the polymerization reaction during normal sequencing

# Oxford Nanopore MinION



Reads up to 100k  
1D and 2D reads  
15-40% error rate  
Life time 5 days



# Main types of equipment



Illumina HiSeq  
Illumina Xten  
Illumina MiSeq

Short paired reads  
**HIGH throughput**



Ion Torrent PGM  
Ion Proton  
Ion S5 XL

Short single-end reads  
**FAST throughput**



PacBio RSII

Ultra-long reads  
**FAST throughput**

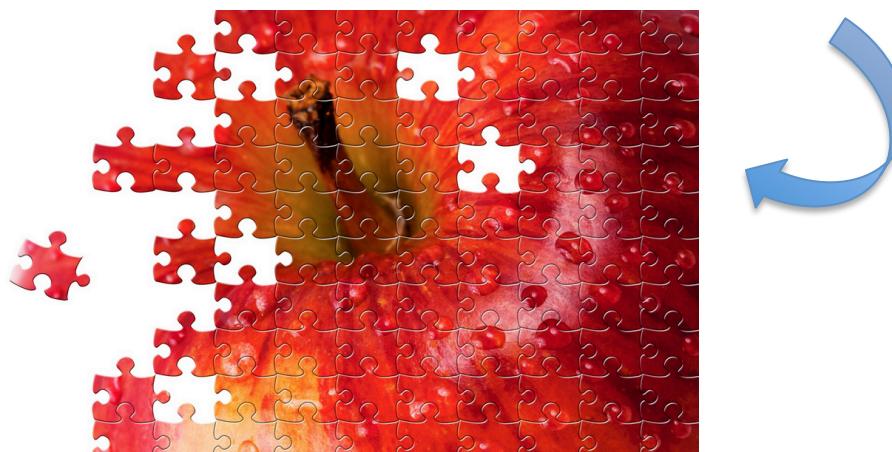
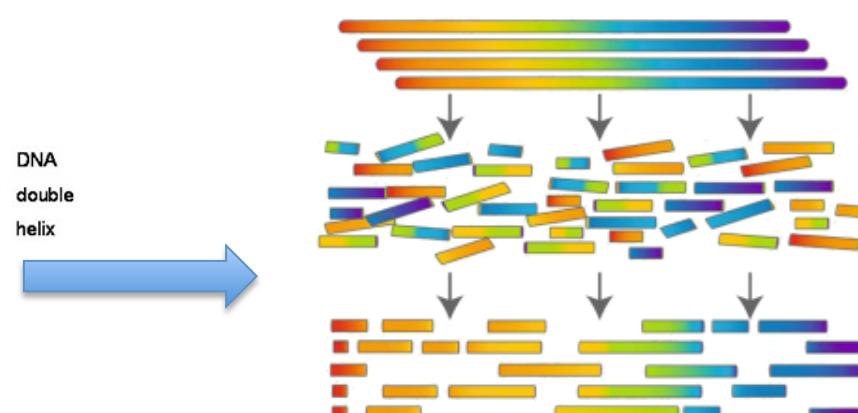
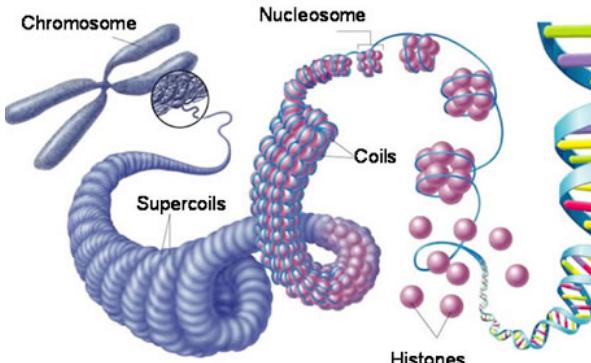
# NGS/MPS applications

- Whole genome sequencing:
  - De novo sequencing
  - Re-sequencing
- Transcriptome sequencing:
  - mRNA-seq
  - miRNA
  - Isoform discovery
- Target re-sequencing
  - Exome
  - Large portions of a genome
  - Gene panels
  - **Amplicons**



# De novo sequencing

- Used to create a reference genome without previous reference





# De novo sequencing:

## Illumina strategy

### Sequencing:

- PE library with 350 bp
- PE library with 600 bp
- MP library with 2 kb
- MP library with 5-8-20 kb

PE: 50-100x, MP 10-15x

## PacBio strategy

### Sequencing:

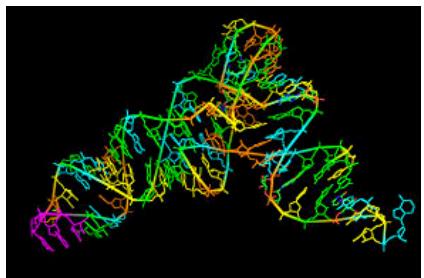
- 10-20 kb library  
50-80x  
(where 30x are reads above 10 kb)

## Analysis:

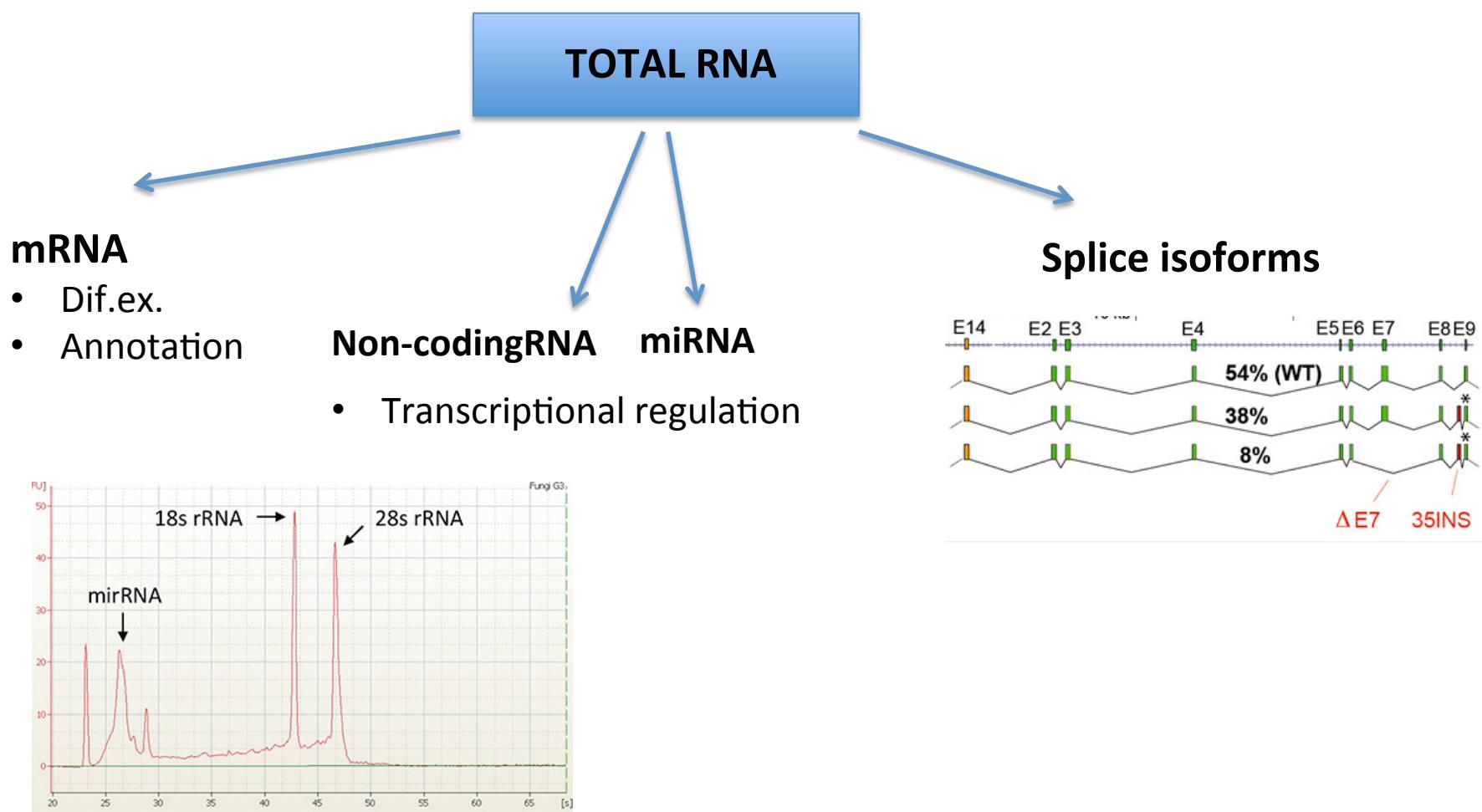
- ALLPATH

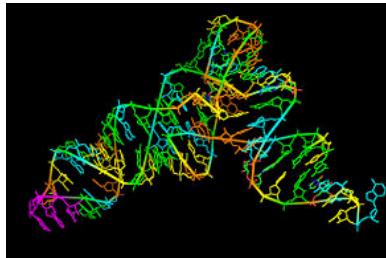
## Analysis:

- HGAP (haploid)
- FALCON (diploid)



# Transcriptome sequencing (RNA-seq)





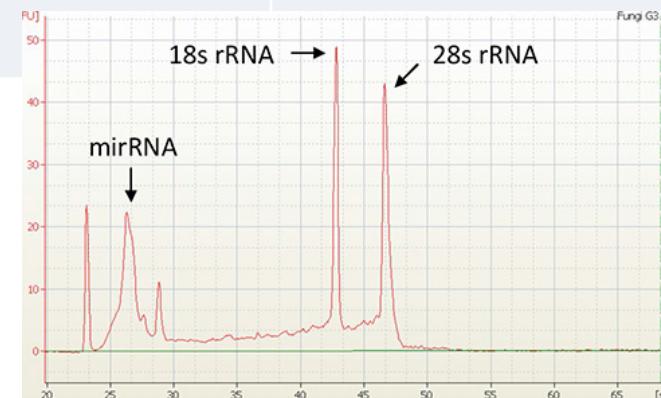
# mRNA: rRNA depletion vs polyA selection

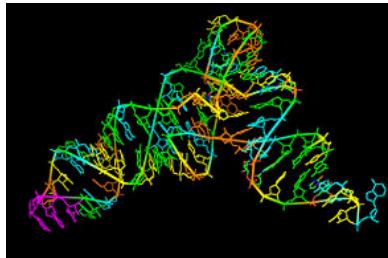
Method	Pros	Cons	Recommended
rRNA depletion	<ul style="list-style-type: none"> <li>Captures on-going transcription</li> <li>Picks up non-coding RNA</li> </ul>	<ul style="list-style-type: none"> <li>Does not get rid of all rRNA</li> <li>Messy Dif.Ex. profile</li> </ul>	20-40 mln reads (single or PE)
polyA selection	<ul style="list-style-type: none"> <li>Gives a clean Dif.Ex. profile</li> </ul>	<ul style="list-style-type: none"> <li>Does not pick non-coding RNA</li> </ul>	5-20 mln reads

Alternative for **human** RNA-seq:

*AmpliSeq Human Transcriptome panel:*

- faster, cheaper, works fine with FFPE
- input: 50 ng **total** RNA
- dif.ex. ONLY



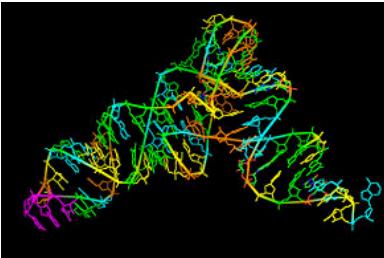


# RNA-seq

## Equipment-related bias

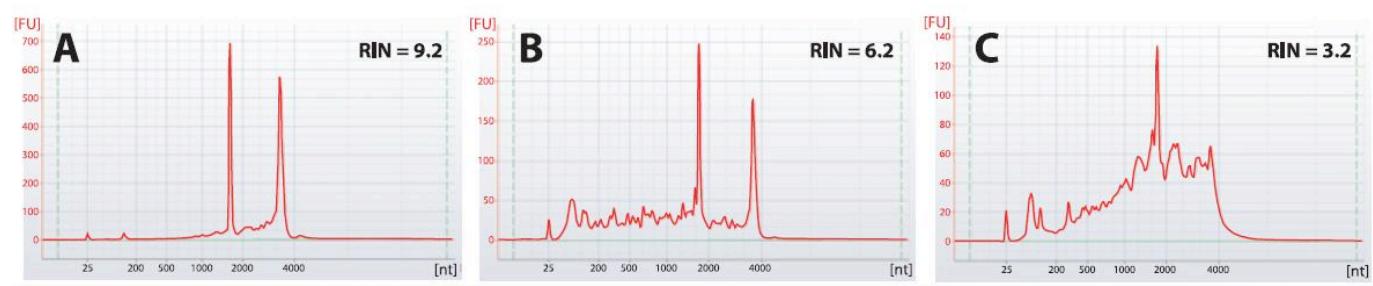
- De novo transcriptome: Illumina **PE only**
- RNA-seq with a good reference:
  - Illumina 50 bp single end for Dif. Ex.
  - Illumina PE for splice information
  - Ion Proton single end in both cases

miRNA: Illumina or IonProton, but stick to the same technology through the project!

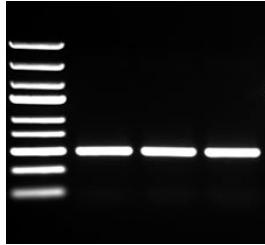


# RNA-seq experimental setup

- mRNA only: any kit
- mRNA **and** miRNA: only specialized kits
- Always use DNase!
- RIN value above 8.



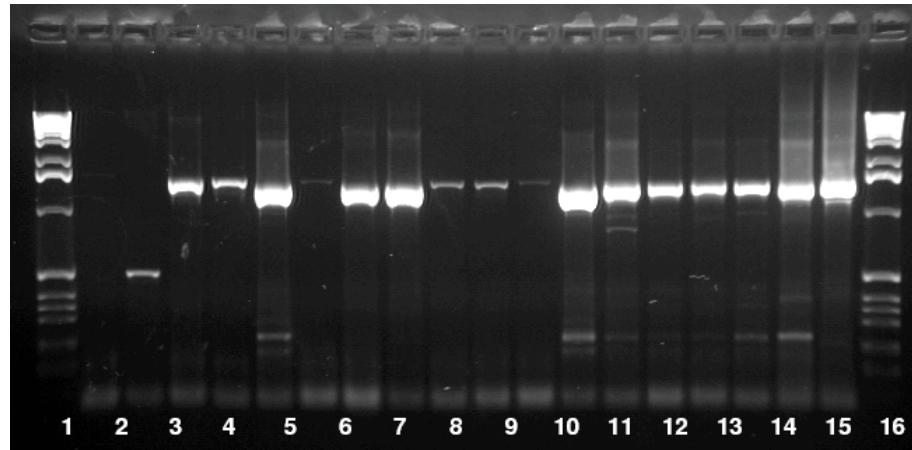
- CONTROL vs experimental conditions
- Biological replicates: 4 strongly recommended



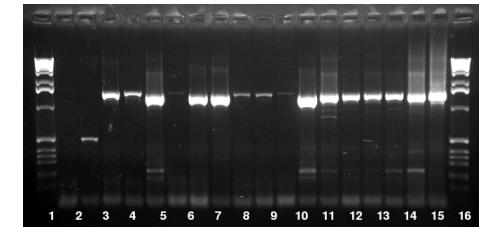
# Amplicon sequencing

Used a lot in metagenomics

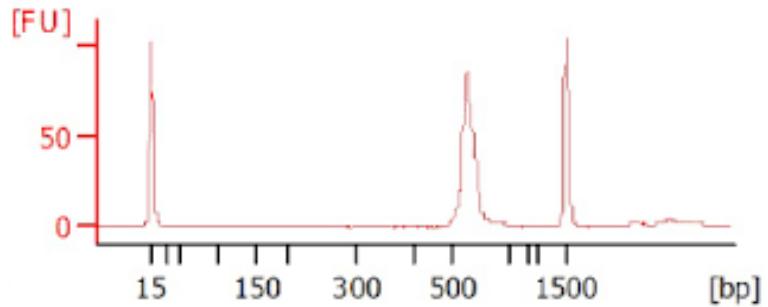
- rRNA genes & spacers (16S, ITS)
- Functional genes
- Genotyping by sequencing



# Amplicon sequencing



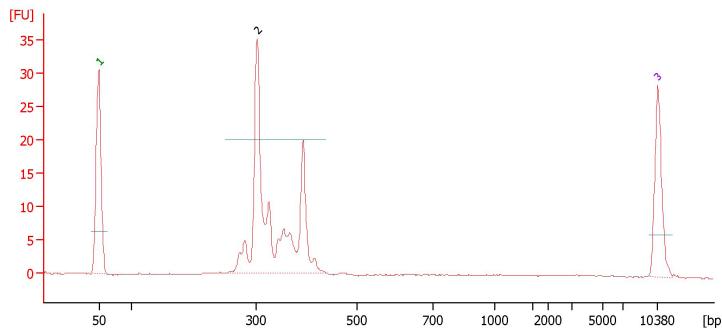
Example 1: tight peak, OK



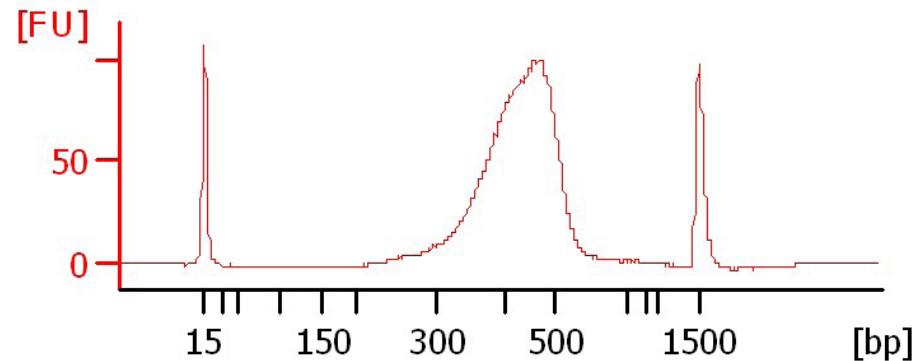
## FOR ANY NGS TECHNOLOGY

Size difference among fragments **must not** exceed 80 bp (or 20% in length)

Reason – preferential amplification of short fragments

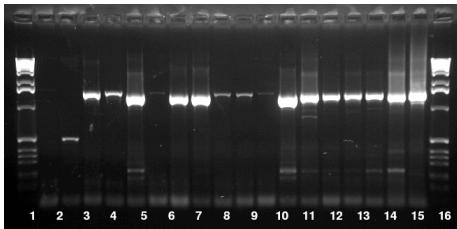


Example 2: several sizes,  
fractionation is needed  
=> we HAVE to make several libraries

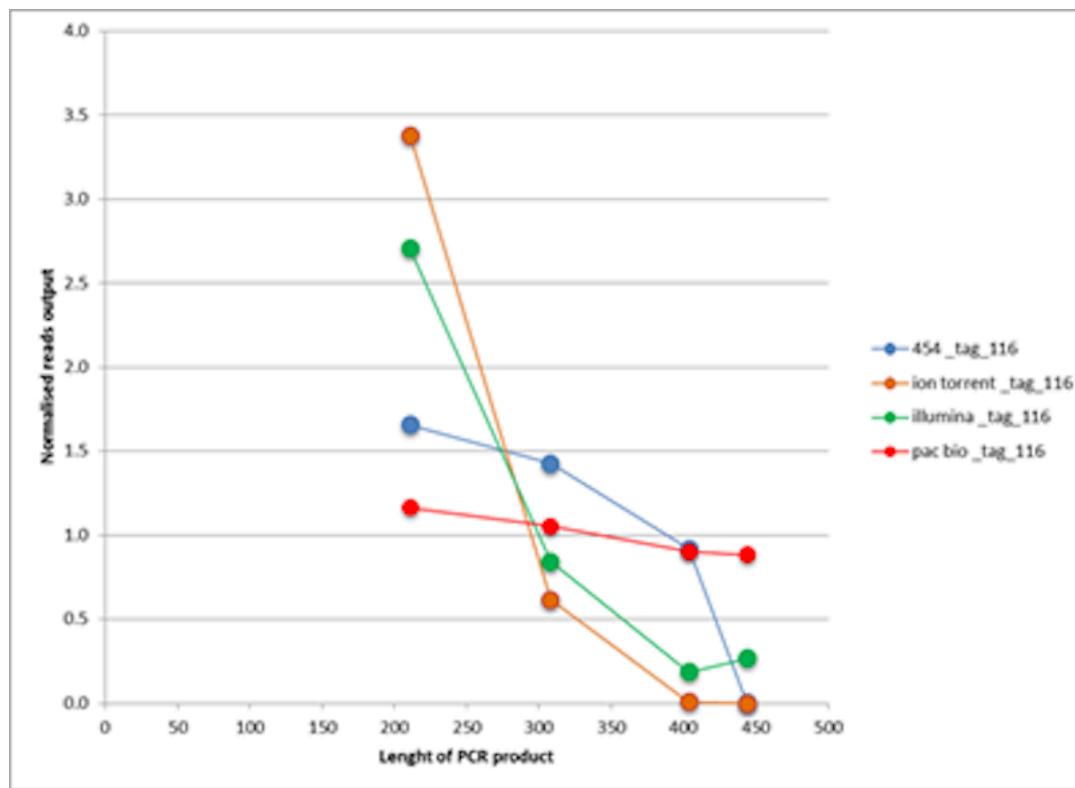


Example 3: broad peak;  
size selection is needed

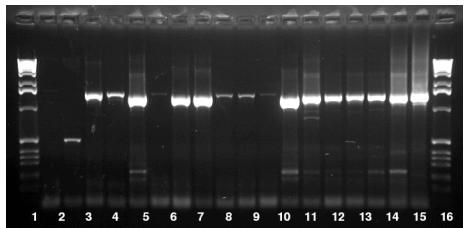
**SIZE MATTERS...**



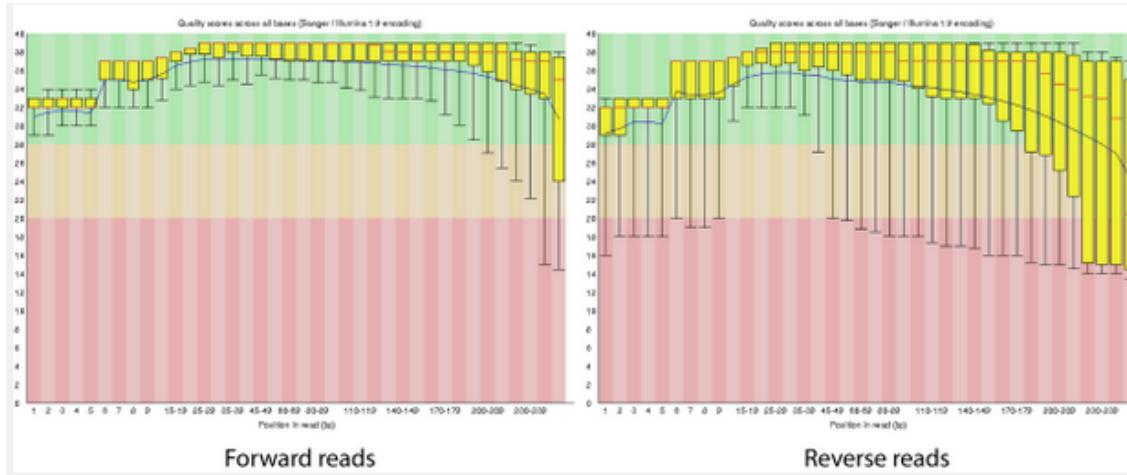
# Size-related bias in amplicon-seq



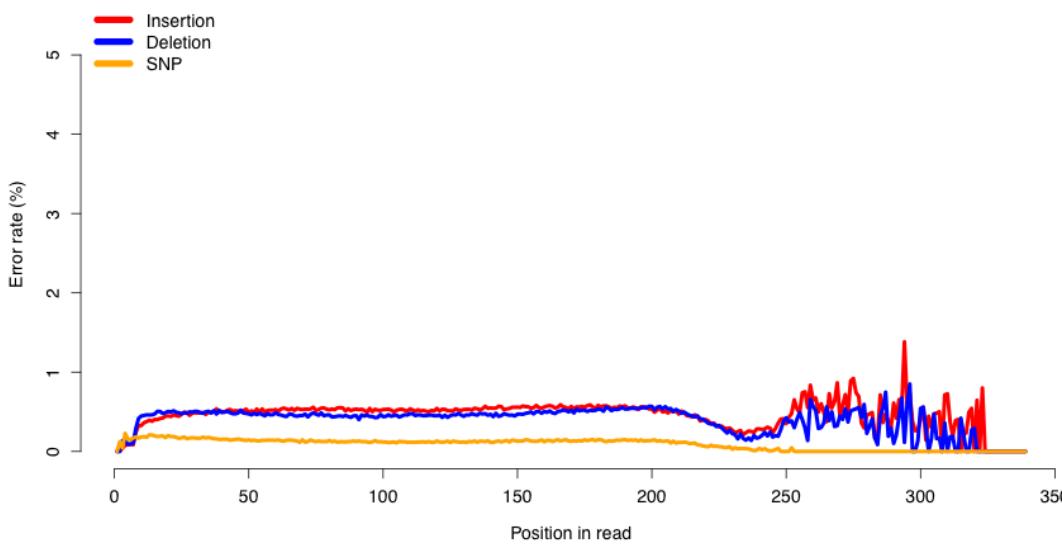
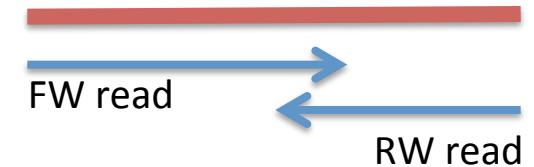
Courtesy Mikael Brandström Durling, Forest Mycology and Pathology, SLU



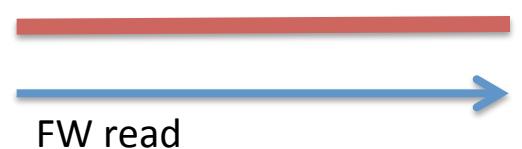
# When you sequence an amplicon...



On MiSeq



On Ion

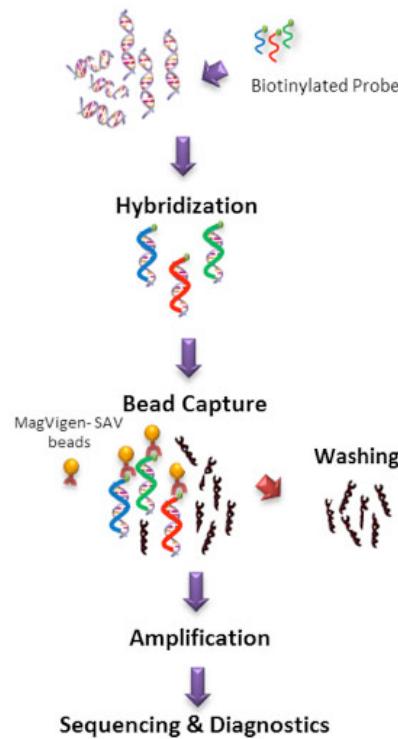


# Sequence capture

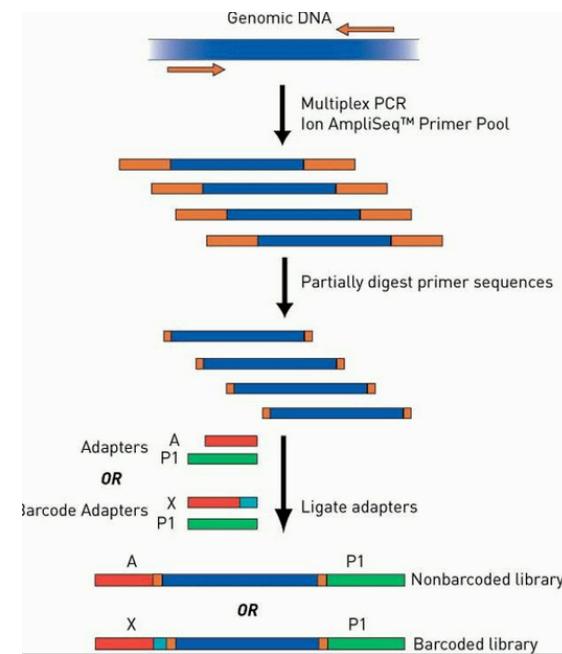
When you are not interested in the entire genome:

- Exome
- Regions of interest
- Genes of interest (gene panels)

## Hybridization-based capture



## PCR-based capture



# Sequence capture: technology choice

- AmpliSeq panels (multiplex PCR) – Ion Only
  - Comprehensive Cancer panel
  - Cancer Hotspot panel
  - AmpliSeq Human Exome, etc
  - AmpliSeq Human Transcriptome
- Hybridization-based: any technology
- Non-multiplex PCR – any technology
  - Short reads (up to 500 bp) – Illumina
  - Medium reads (up to 500 bp) – Ion
  - Long reads (from 500 bp – 20 kb) - PacBio

# Main types of equipment & applications



Illumina HiSeq  
Illumina Xten  
Illumina MiSeq

Short paired reads  
**HIGH throughput**

**Human WGS**  
mRNA and miRNA  
**De novo transcriptome**  
Exome  
ChIP-seq  
Short amplicons  
Methylation



Ion Torrent PGM  
Ion Proton  
Ion S5 XL

Short single-end reads  
**FAST throughput**

mRNA and miRNA  
**Exome**  
ChIP-seq  
Short amplicons  
**Gene panels**  
**Clinical samples**



PacBio RSII

Ultra-long reads  
**FAST throughput**

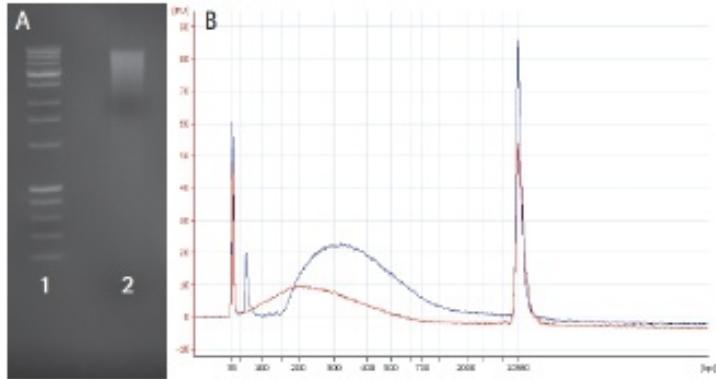
**Long amplicons**  
**Re-sequencing**  
**De novo sequencing**  
Novel isoform discovery  
Fusion transcript analysis  
**Haplotype phasing**  
**Clinical samples**

# What is The BEST?

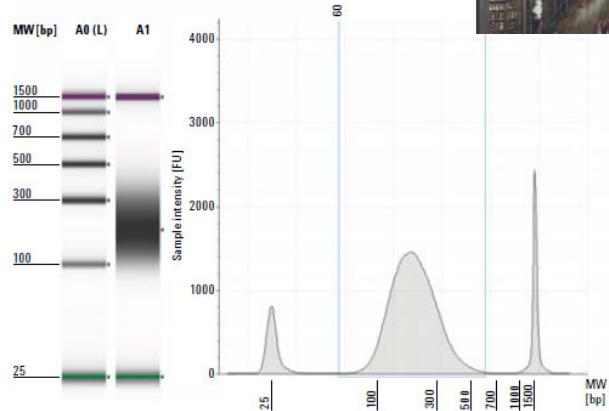


# SAMPLE QUALITY REQUIREMENTS

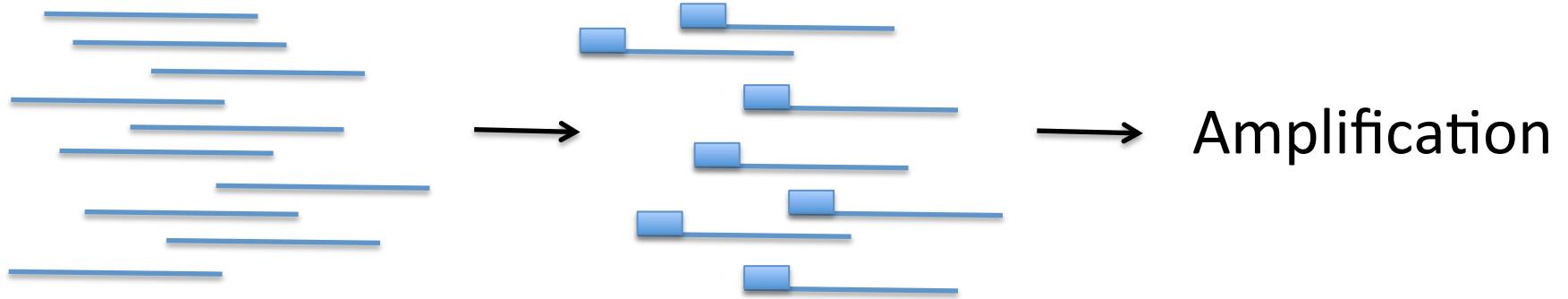
# Making an NGS library



DNA QC – **paramount importance**



Sharing & size selection



Ligation of sequencing adaptors, technology specific

# **Garbage in – garbage out:**

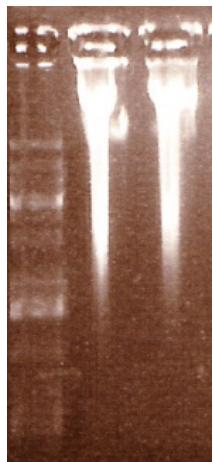
sequencing success to 90% depends  
on the **sample quality**

**Before samples are submitted:**

Send us the gel picture (DNA)  
260/280 and 260/230 readings (DNA)

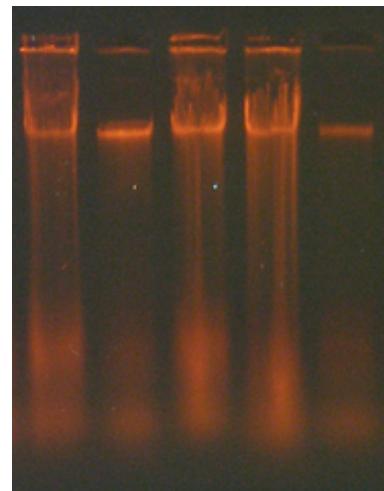
BioAnalyzer readings (RNA)

# Reading gel pictures of genomic DNA



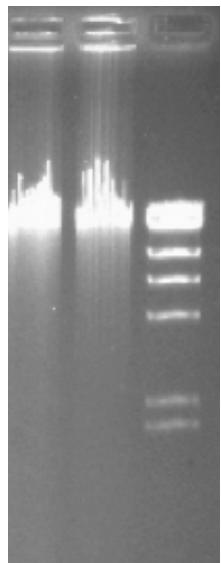
## Protein contamination

- Apply phenol-chloroform

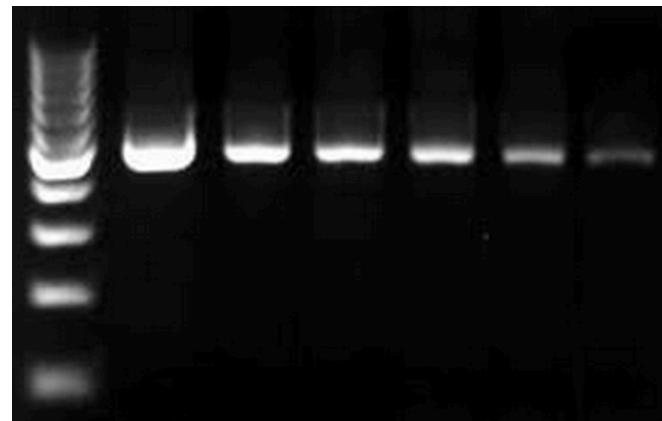


## RNA contamination

- Apply RNase, followed by phenol-chloroform extraction



## Phenol carry-over or overloaded sample?



If unsure, make dilution series.

If problem persists – try MoBio clean-up kit, or re-extract DNA

# What do absorption ratios tell us?

## Pure DNA 260/280: 1.8 – 2.0

< 1.8:

Too little DNA compared to other components of the solution; presence of organic contaminants: proteins and phenol; glycogen - **absorb at 280 nm**.

> 2.0:

High share of RNA.

## Pure DNA 260/230: 2.0 – 2.2

<2.0:

Salt contamination, humic acids, peptides, aromatic compounds, polyphenols, urea, guanidine, thiocyanates (latter three are common kit components) – **absorb at 230 nm**.

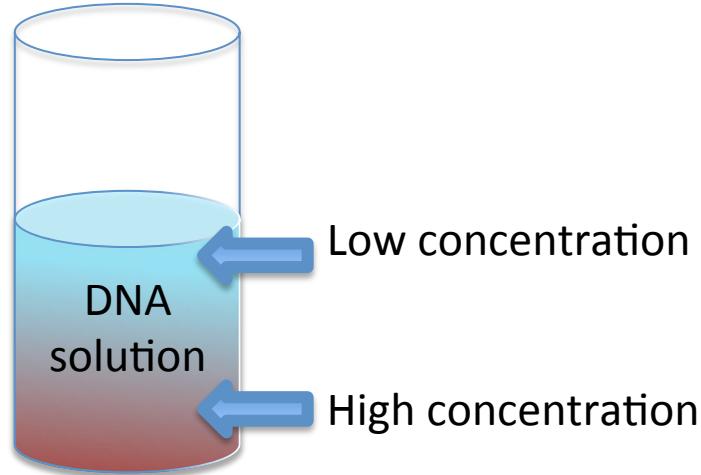
>2.2:

High share of RNA, very high share of phenol, **high turbidity**, dirty instrument, wrong blank.

*Photometrically active contaminants:  
phenol, polyphenols, EDTA, thiocyanate, protein,  
RNA, nucleotides (fragments below 5 bp)*

# How to make a correct measurement

- Thaw DNA completely
- Mix gently (**never vortex!**)
- Put the sample on a thermoblock: 37°C, 15-30 min
- Mix gently
- **Dilute 1:100 (if HMW)**
- Mix gently
- Make a measurement with an appropriate blank
- **NANODROP is Bad.**



# Sample prep: genomic DNA

**Treat DNA as a crystal vase: it is fragile when in solution**

As soon as DNA is released from the cells – use wide-bore tips

Limit pipetting to minimum

Always use RNase!

Never vortex!

Do not heat above 65°C

Reduce amount of freeze-thaw cycles to minimum

Make several aliquotes of the stock DNA



# Sample prep: take home message

PCR-quality sample and

NGS-quality sample

**are two completely different  
things**

# Sample prep: RNA

**mRNA degrades FAST**

Freeze sample or place it in RNA-later within 30 sec (*if possible*)

If going for miRNA seq – chose a correct kit!

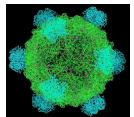
Always treat samples with DNase

Differential expression, miRNA – **RIN value over 8.0**

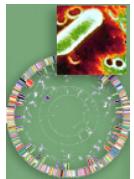
Aim for 4 biological replicates

Let's get philosophical

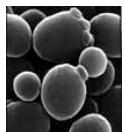
# Since the beginning of Genomics:



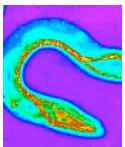
First genome: virus  $\phi$  X 174 - 5 368 bp (1977)



First organism: *Haemophilus influenzae* - 1.5 Mb (1995)



First eukaryote: *Saccharomyces cerevisiae* - 12.4 Mb (1996)



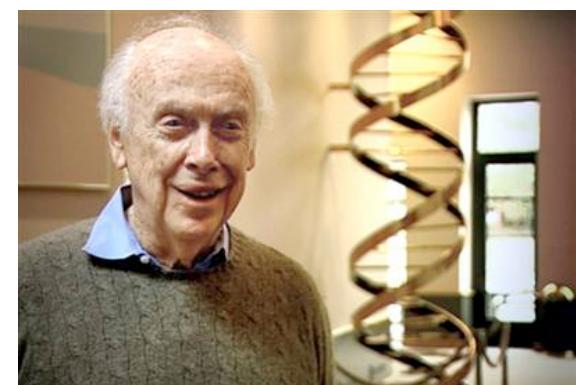
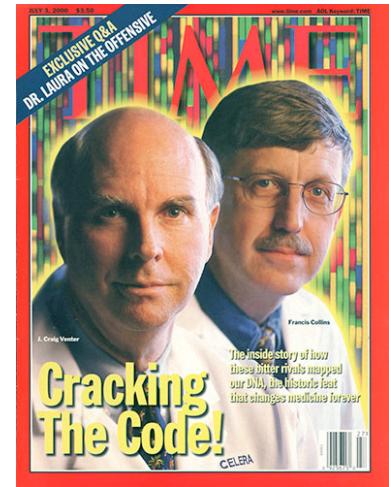
First multicellular organism: *Cenorhabditis elegans* - 100 MB (1998-2002)



First plant: *Arabidopsis thaliana* - 157 Mb (2000)

# ... prices go down

- Human genome project, 2007
  - Genome of Craig Venter costs 70 mln \$
    - Sanger's sequencing
  - Genome of James Watson costs 2 mln \$
    - 454 pyrosequencing
  - Ultimate goal: 1000 \$ / individual  
Almost there! (1200 \$)





# ... paradigm change

- From single genes to complete genomes
- From single transcripts to whole transcriptomes
- From single organisms to complex metagenomic pools
- From model organisms to the species you are studying
- Personal genome = personalized medicine



BEHOLD...THE ANSWER  
TO LIFE, THE UNIVERSE AND EVERYTHING



SORRY... WHAT WAS THE  
QUESTION AGAIN?

HITCHHIKER'S  
GUIDE TO THE  
GALAXY

# ... scientific value diminishes

Science 5 September 1997:  
Vol. 277 no. 5331 pp. 1453-1462  
DOI: 10.1126/science.277.5331.1453

IF 31.6

< Prev | Table of Contents | Next >

ARTICLES

## The Complete Genome Sequence of *Escherichia coli* K-12

Frederick R. Blattner\*, Guy Plunkett III\*, Craig A. Bloch, Nicole T. Perna, Valerie Burland, Monica Riley, Julio Collado-Vides, Jeremy D. Glasner, Christopher K. Rode, George F. Mayhew, Jason Gregor, Nelson Wayne Davis, Heather A. Kirkpatrick, Michael A. Goeden, Debra J. Rose, Bob Mau and Ying Shao

Journal of Biotechnology  
Article in Press, Corrected Proof - Note to users

doi:10.1016/j.jbiotec.2010.12.018 | How to Cite or Link Using DOI

Permissions & Reprints

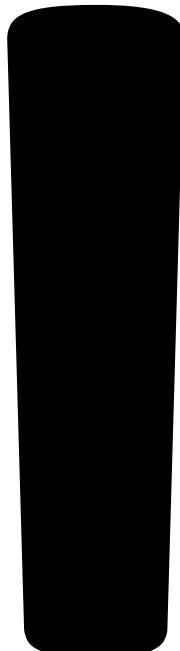
IF 2.9



## The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome

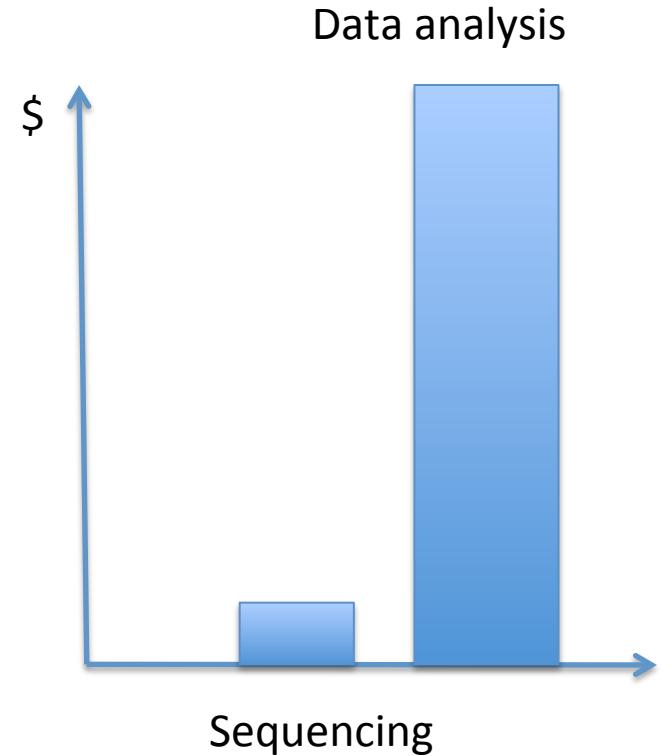
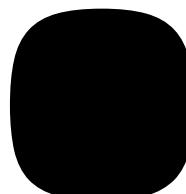
Susanne Schneiker-Bekel<sup>a</sup>, Daniel Wibberg<sup>a</sup>, Thomas Bekel<sup>b</sup>, Jochen Blom<sup>b</sup>, Burkhard Linke<sup>b</sup>, Helko Neuweiler<sup>b</sup>, Michael Stiens<sup>a, c</sup>, Frank-Jörg Vorhölter<sup>a</sup>, Stefan Weidner<sup>a</sup>, Alexander Goesmann<sup>b</sup>, Alfred Pühler<sup>a</sup> and Andreas Schlüter<sup>a</sup>,  

# Main challenge - DATA ANALYSIS and DATA STORAGE



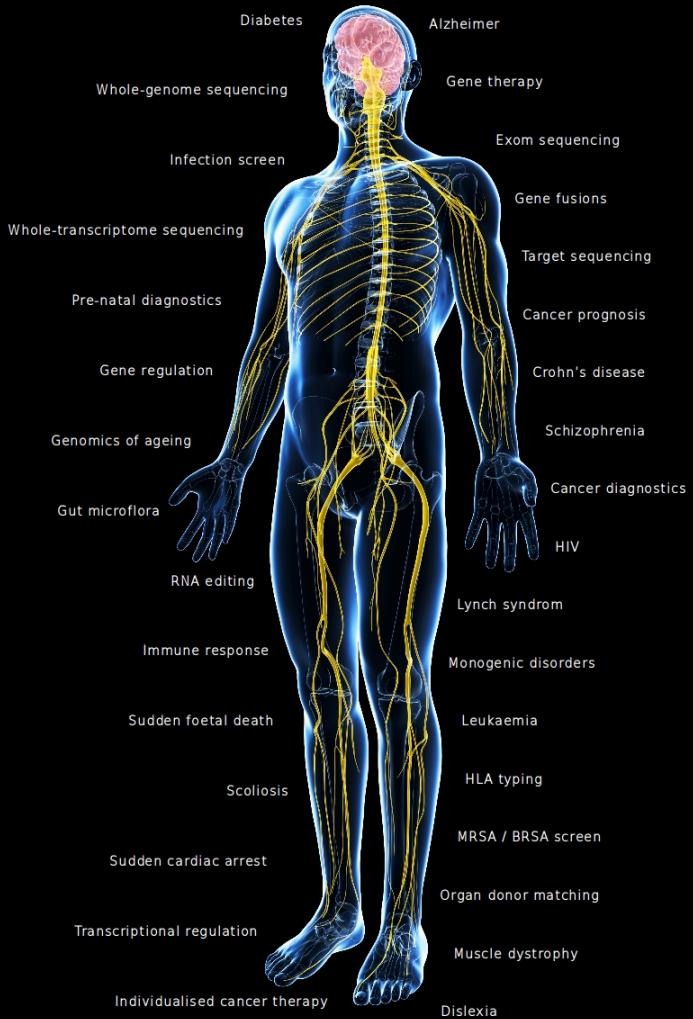
*"If the data problem is not addressed, ABI's SOLiD, 454's GS FLX, Illumina's GAII or any of the other deep sequencing platforms will be destined to sit in their air-conditioned rooms like a Stradivarius without a bow."*

<http://finchtalk.geospiza.com>



=> More bioinformaticians to people!

# What we sequenced at SciLifeLab



165 amplicon, Acinetobacter baumannii, Acrasis kona, Acidotheres javanicus, Actinobacillus succinogenes, African swine fever virus, Agaricomycotina sp., **Alces alces**, **Alligator mississippiensis**, Amphura filiformis, Apis mellifera, Aquila chrysaetus, Arabidopsis thaliana, Arabis alpina, Archaeorhizomycetes finlayi, Arctocephalus gazella, artificial sequences, Arvicolidae amphibus, Ascaris galli, Aspergillus oryzae, Astragalus stephaniae, Atlantic herring, Atlantic salmon, **Avena sativa**, Baccharis breviseta, Baccharis dracunculifolia, Bacteriophages, Balaenoptera musculus, Balaenoptera physalus, Balanus improvisus, Baltic Sea microorganisms, Bathynomus sp., Blodobactrium sp., **Borella burgdorferi**, Borrelia garinii, Bos taurus, Bovine viral diarrhea virus, Brachyspira suauatina, Brassica sp., Brettanomyces naardenensis, Cenocarabditis elegans, Callosobruchus maculatus, Candida intermedia, Candida parapsilosis, Candidatus neoehrlichia mikurensis, **Canis lupus**, Capreolus capreolus, Capsella bursa-pastoris, Capsella grandiflora, Capsella rubella, Ceanothus thyrsiflorus, Cervus dama, Cervus elaphus, Chlidonia submaculata, Clonostachys rosea, Clostridium ultunense, Coelodonta antiquitatis, Collomia crocea, Collomia heterophylla, Coregonus lavaretus, Coronavirus, Corvus corone, Corvus monedula, Crassostrea gigas, Crioceris gryneus, Cryptococcus tephrensis, Cubanola dominicensis, Cyttomegalovirus, Danio rerio, Dafnus glomerata, Deformed wing virus, Dekkera bruxellensis, Dicerorhinus sumatrensis, Dictyostelium discoideum, Diplotaxis erucoides, Diptophorus gymnothorax, Diptostomus longitubus, Drosophila melanogaster, Drosophila paulistorum, Electrophorus electricus, Enterobacter cloacae, Enterococcus faecium, **Equus caballus**, Escherichia coli, Eumecistostomus macrobrachium, Euphorbia lathyrus, Euphorbia peplus, Euplectes afer, Euplectes ardens, Euplectes australis, Euplectes hordeaceus, Euplectes macrourus, Euplectes orix, **Felis catus**, Ficedula albicollis, Ficedula hypoleuca, **Frangula ananassa**, Freshwater microbial communities, Fucus radicans, Fucus vesiculosus, Fumaria sp., Galerucella tanaceti, Geopis magnirostris, Giarla muris, Globodera rostochiensis, Gnetum gnemon, Gnetum luofuense, Gnetum montanum, Gnetum parvifolium, Gnetum pedunculatum, Gymnoscomum semen, Gonzalagunia, Gut microbiota, Hamulus marianae, Heterobasidion abietinum, Heterobasidion annosum, **Hippophae rhamnoides**, **Homo sapiens**, Human immunodeficiency virus, Hyperzia selago, Hymenoscyphus albidus, Hymenoscyphus pseudolepidotus, Idotea baltica, **Influenza A virus**, Klebsiella pneumoniae, Laccaria bicolor, Lactobacillus, Lepidium campestre, **Leptidea sinapis**, Letharia rugosa, Letharia vulpina, Littorina saxatilis, Lycoctonus pyrrhoterus, **Lynx lynx**, Malassezia sympodialis, **Maus sylvensis**, **Mammuthus primigenius**, Marchantia polymorpha, Marine bacteria whole community, **Meliagethes aeneus**, Metagenomes, Methanococcus sp., Metschnikowia andinaensis, Metschnikowia hawalensis, Metschnikowia pulcherrima, Metschnikowia saccharicola, Mixornis galarinus, Moorella thermoacetica, Mus musculus, Mycobacterium malmoense, Mycobacterium marinum, **Mytilus edulis**, Nemertoderma westbladi, Nesoniphantes sp., Neurospora crassa, Neurospora hispaniola, Neurospora intermedia, Neurospora mettenbergii, Neurospora sitophila, Neurospora tetrasperma, Nora Virus, Nothoprocta ornata, Nothoprocta perdicaria, **Notophthalmus viridescens**, Nyctereutes procyonoides, Ogataea pilii, **Oryctolagus cuniculus**, Rana arvalis, Oryzias latipes, **Pacifastacus leniusculus**, Paenibacillus polymyxia, **Panthera leo**, Panthera pardus, Paracardiacia rubra, Parus major, Parus montanus, Paxillus involutus, Penicillium sp., Peridium aciculiferum, Phlomachus pugnax, Phoca sibirica, Phyloscops poliocephalus, Phyloscops tachycinus, Wolbachia persica, Physcomitrella patens, Phytophthora infestans, **Picea abies**, Pieris napi, Pieris rapae, Pinus pinaster, Pinus sylvestris, **Pistacia lentiscus**, Planctomycetes sp., Plasmodiphora brassicae, Plasmoidium falciiparum, Podospora anserina, Polystachya paniculata, Potamogeton pectinatus, Populus maximowiczii, Populus tremula, Populus trichocarpa, Posoqueria sp., Pseudomonas aeruginosa, Pseudomonas brasiliensis, Pseudomonas chlororaphis, Pseudomonas putida, Pteridophora alberti, Ptiloria paradesii, Puccinia striformis, Pythium oligandrum, Quielea quelea, Rangifer tarandus, Rattus rattus, Rhizoctonia sp., Saccharomyces cerevisiae, Salix purpurea, Salix viminalis, Salmonella enterica, Salmonella typhimurium, Salmo salar, Salmo trutta, Schizophyllum commune, Schizosaccharomyces pombe, Scirpus sphaerocephalus, **Semibalanus balanoides**, Setaria digitata, Silene conoidea, Silene latifolia, Silene viscaria, Sindbis virus, Siphocampylus, Siphocampylus retrovirus, Siphoviridae phage, **Stramenopila marinorum**, **Solanum tuberosum**, Sorghum sp., Spirostomus barbatus, Spirotucleus salmonicida, Spirotucleus ventopus, Staphylococcus aureus, Staphylococcus pseudintermedius, Stenommadium sp., Streptococcus pneumoniae, Streptococcus pyogenes, Streptomyces coelicolor, Struthio camelus, Sulfolobus acidocaldarius, **Sus scrofa**, Synthetic *S. scrofa*, Syntrophic actinomycetes schinkii, Taphrina betulinus, Teplidanoaracter acetoxydans, Thamnolla vermicularis, Thelleria parva, Trypanosoma cruzi, Trypanosoma rangeli, **Ursus spelaeus**, Vitex agnus-castus, Yarrowia lipolytica, Zalophus californianus, Zalophus wollebaeki, Zygotypetalum crinitum, Zygosaccharomyces bailii



# SciLifeLab

[TECHNOLOGIES & SERVICES ▾](#)[RESEARCH ▾](#)[EDUCATION ▾](#)[COLLABORATION ▾](#)

Find more information and search for what you need on the page for Technologies & Services

[What is the difference between national and regional facilities?](#)



Search for Technologies & Services

## National facilities

### Affinity Proteomics

Biobank Profiling  
Cell Profiling  
Fluorescence Tissue Profiling  
PLA Proteomics  
Protein and Peptide Arrays  
Tissue Profiling

### Bioimaging

Advanced Light Microscopy  
Fluorescence Correlation Spectroscopy

### Bioinformatics

Bioinformatics Compute and Storage (UPPNEX)  
Bioinformatics Long-term Support (WABI)  
Bioinformatics Short-term Support and Infrastructure (BILS)

### Chemical Biology Consortium Sweden

Laboratories for Chemical Biology Umeå (LCBU)  
The Laboratories for Chemical Biology at Karolinska Institutet (LCBKI)  
Uppsala Drug Optimization and Pharmaceutical Profiling (UDOPP)

### Clinical Diagnostics

Clinical Biomarkers  
Clinical Genomics  
Clinical Sequencing

### Drug Discovery and Development

ADME (Absorption Distribution, Metabolism Excretion) of Therapeutics (UDOPP)  
Biochemical and Cellular Screening  
Biophysical Screening and Characterization  
Human Antibody Therapeutics  
In Vitro and Systems Pharmacology  
Medicinal Chemistry – Hit2Lead  
Medicinal Chemistry – Lead Identification  
Protein Expression and Characterization

### Functional Genomics

Karolinska High Throughput Center (KHTC)

### National Genomics Infrastructure

NGI Stockholm (Genomics Applications)  
NGI Stockholm (Genomics Production)  
NGI Uppsala (SNP&SEQ Technology Platform)  
NGI Uppsala (Uppsala Genome Center)

### Structural Biology

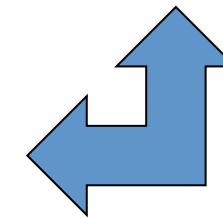
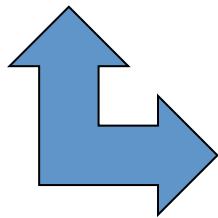
Protein Science Facility

# National Genomics Infrastructure

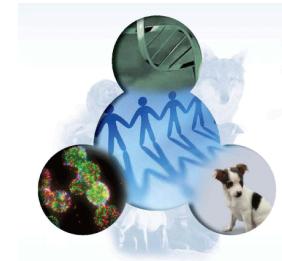
**SciLifeLab, Stockholm**



**SciLifeLab, Uppsala**



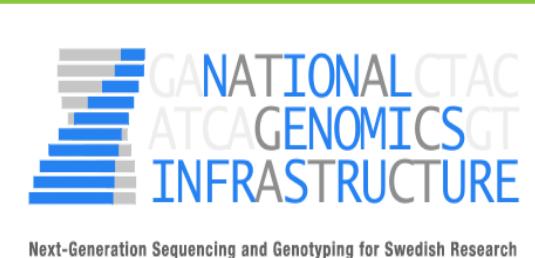
**Uppmax, Uppsala**



NGI-SciLifeLab is one of the most well-equipped  
NGS sites in Europe



- 10 Illumina HiSeq Xten**
- 17 Illumina HiSeq 2000/2500**
- 3 Illumina MiSeq**
- 1 Illumina NextSeq**
- 2 Life Technologies Ion Torrent**
- 6 Life Technologies Ion Proton**
- 2 Pacific Biosciences RSII**
- 2 Sanger ABI3730**
- 1 Argus Whole Genome Map. Syst.**
- 2 Oxford Nanopore Minlon**



# https://portal.scilifelab.se/genomics

- Home
- Newsletters
- Events
- ▼ Services
  - Sequencing
  - Genotyping
  - Bioinformatics
  - Orders
  - Accreditation
- Documents
- FAQ
- Contact us

## User login

Username: \*

Password: \*

Log in

- Create new account
- Request new password

## Order portal for the National Genomics Infrastructure (NGI) hosted by SciLifeLab

This portal is used to describe and submit orders for the services provided by NGI. NGI primarily serves projects run by research groups in Sweden. Projects from other countries are admissible, but have lower priority than Swedish projects. Depending on the queue situation, NGI may decide to decline a non-Swedish project altogether.

### News

- The document "Stockholm node: Sample requirements for genomics projects" has been updated. See the [Documents page](#). [modified 2015-09-14]
- The document "Stockholm node: Price examples for NGI Stockholm (Swedish academia)" has been updated. See the [Documents page](#). [modified 2015-08-17]
- The Human Whole Genome sequencing Toolbox group has a new wiki where the Toolbox work will be documented: <https://wabi-wiki.scilifelab.se/display/SHGATG/>

### Events ([more information](#))

- Course: Perl Programming with Application to Bioinformatics, 2 hp, 12-16 + 19 Oct 2015, Uppsala. [added 2015-09-04]
- Course: RNA-seq, 1 hp, 20-22 Oct 2015, Uppsala. [added 2015-09-04]
- Course: De novo Genome Assembly, 1 hp, 16-17 Nov 2015, Uppsala. [added 2015-09-04]
- Course: Introduction to Bioinformatics using NGS data, 2 hp, 16-20 Nov 2015, Lund (?). [added 2015-09-04]
- Course: Metagenomics, 1 hp, 24-26 Nov 2015, Uppsala. [added 2015-09-04]

### About NGI

NGI is hosted by [Science for Life Laboratory](#) (SciLifeLab). See [the NGI pages](#) at the SciLifeLab web site for more information. The National Genomics Infrastructure (NGI) was launched January 1st 2013. It originates from the Swedish Research Council RFI infrastructure SNISS.

NGI follows the VR RFI guidelines for national infrastructures and is supported by the Swedish Research Council (VR), host universities (KTH, UU), and SciLifeLab. NGI includes facilities established by profs Ulf Gyllensten (Uppsala), Ann-Christine Syvänen (Uppsala) and Joakim Lundeberg (Stockholm). In addition, the Knut and Alice Wallenberg Foundation (KAW) has provided critical capital support for new instruments and computational infrastructure.

# What happens then?

NGI Project coordinators meet twice a week via Skype



Ulrika  
Liljedahl



Ellenor  
Devine

SNP&SEQ, Uppsala node



Mattias  
Ormestad



Beata  
Werne Solenstam

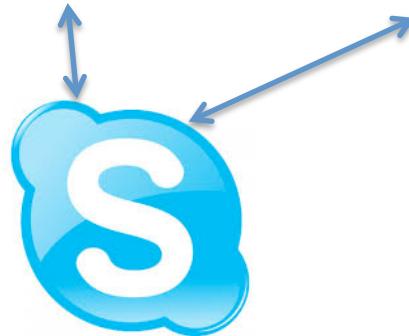
Stockholm Node



Olga  
Vinnere Pettersson  
UGC, Uppsala Node

**Project distribution is based on:**

1. Wish of PI
2. Type of sequencing technology
3. Type of application
4. Queue at technology platforms



Project is then assigned to a certain node and a coordinator contacts the PI

# Project meeting

## What we can help you with:

- Design your experiment based on the scientific question.
- Choose the best suited application for your project.
- Find the most optimal sequencing setup.
- Answer all questions about our technologies and applications, as well as bioinformatics.
- In special cases, we can give extra-support with bioinformatics analysis – development of novel methods and applications

# QUESTIONS?

# Downstream Data Analysis

