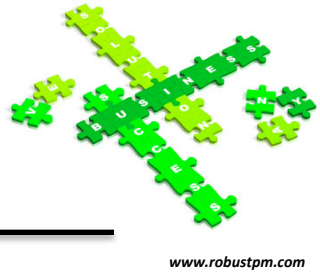
Next Generation Sequencing –

An Overview

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

Version 6.3

Outline

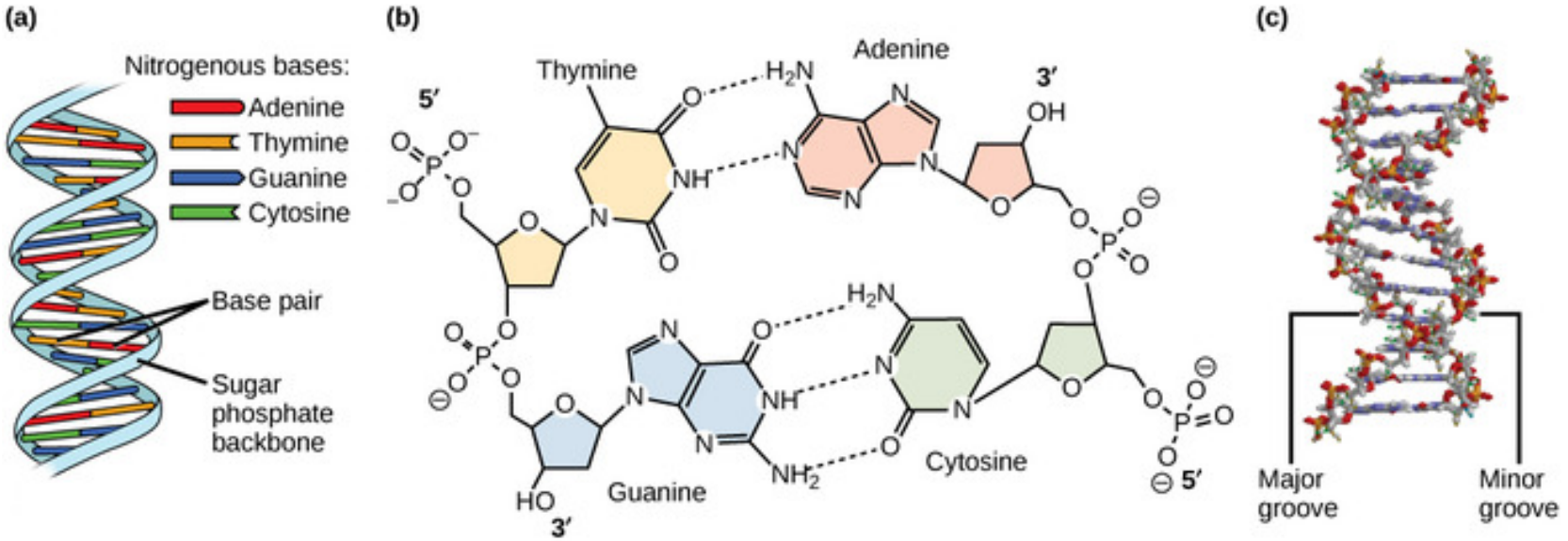


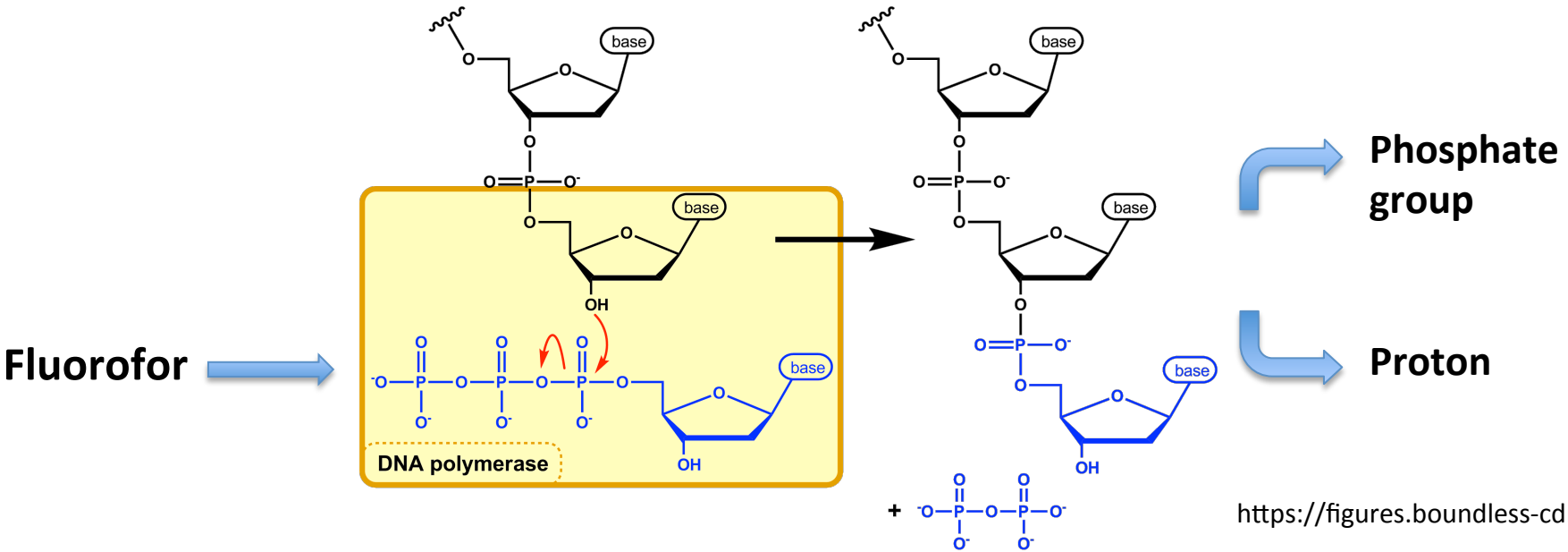
- A bit of history
- NGS technologies

- NGS applications
  - De Novo
  - RNA-seq
  - Targeted enrichment (hybridization & amplicon-Seq)

- National Genomics Infrastructure – Sweden
- Auxiliary technologies (10x Chromium, BioNano)
- Sample prep for NGS

What is sequencing?





https://figures.boundless-cdn.com

Once upon a time…

- Fredrik Sanger and Alan Coulson

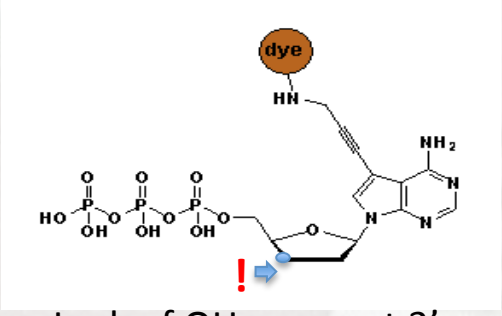
  Chain Termination Sequencing (1977)

  Nobel prize 1980

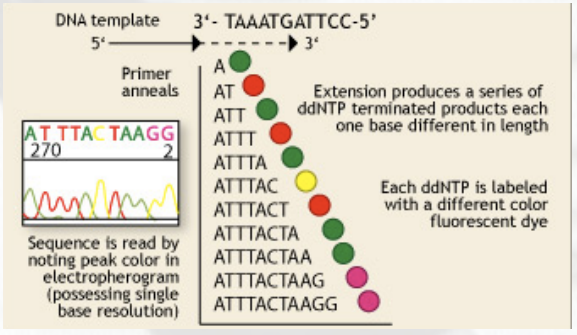
Principle:

SYNTHESIS of DNA is randomly TERMINATED at different points

Separation of fragments that are 1 nucleotide different in size



Lack of OH-group at 3’ position of deoxyribose



1 molecule sequenced at a time = 1 read

Capillary sequencer: 384 reads per run

2006 REVOLUTION

nature
International weekly journal of science

Login Cart

Search go Advanced search

Journal home > Archive > Article > Full Text

Journal content

- Journal home
- Advance online publication
- Current issue
- Nature News
- Archive
- Supplements
- Web focuses
- Podcasts
- Videos
- News Specials

Journal information

- About the journal
- For authors

Article

Nature **437**, 376-380 (15 September 2005) | doi:10.1038/nature03959; Received 6 May 2005; Accepted 10 June 2005; Published online 31 July 2005

There is a Corrigendum (26 January 2006) associated with this document.

There is a Corrigendum (4 May 2006) associated with this document.

Genome sequencing in microfabricated high-density picolitre reactors

Marcel Margulies[1,2], Michael Egholm[1,2], William E. Altman[1], Said Attiya[1], Joel S. Bader[1], Lisa A. Bemben[1], Jan Berka[1], Michael S. Braverman[1], Yi-Ju Chen[1], Zhoutao Chen[1], Scott B. Dewell[1], Lei Du[1], Joseph M. Fierro[1], Xavier V. Gomes[1], Brian C. Godwin[1], Wen He[1], Scott Helgesen[1], Chun He Ho[1], Gerard P. Irzyk[1], Szilveszter C. Jando[1], Maria L. I. Alenquer[1], Thomas P. Jarvie[1], Kshama B. Jirage[1], Jong-Bum Kim[1], James R. Knight[1], Janna R. Lanza[1], John H. Leamon[1], Steven M. Lefkowitz[1], Ming Lei[1], Jing Li[1], Kenton L. Lohman[1], Hong Lu[1], Vinod B. Makhijani[1], Keith E. McDade[1], Michael P. McKenna[1], Eugene W. Myers[2], Elizabeth Nickerson[1], John R. Nobile[1], Ramona Plant[1], Bernard P. Puc[1], Michael T. Ronan[1], George T. Roth[1], Gary J. Sarkis[1], Jan Fredrik Simons[1], John W. Simpson[1], Maithreyan Srinivasan[1], Karrie R. Tartaro[1], Alexander Tomasz[2], Kari A. Vogt[1], Greg A. Volkmer[1], Shally H. Wang[1], Yong Wang[1], Michael P. Weiner[1], Pengguang Yu[1], Richard F. Begley[1] & Jonathan M. Rothberg[1]

subscribe to nature

FULL TEXT

- Readers' Comments
- Subscribe to comments (RSS)
- What is RSS?

- Previous | Next
- Table of contents

Download PDF

View interactive PDF in ReadCube

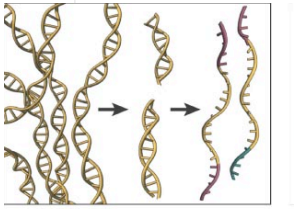
Share this article

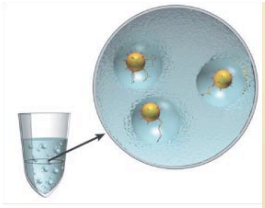
CrossRef lists 376 articles citing this article

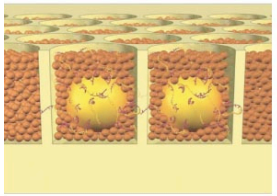
Thousands of molecules sequenced in parallel

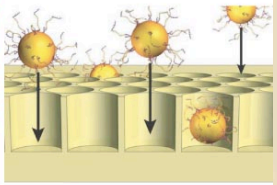
1 mln reads sequenced per run

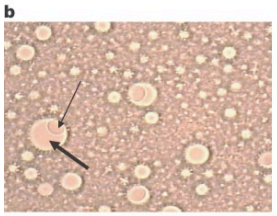
Roche 454 GS FLX

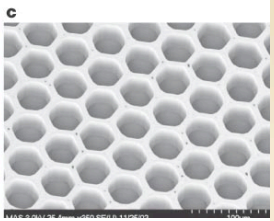


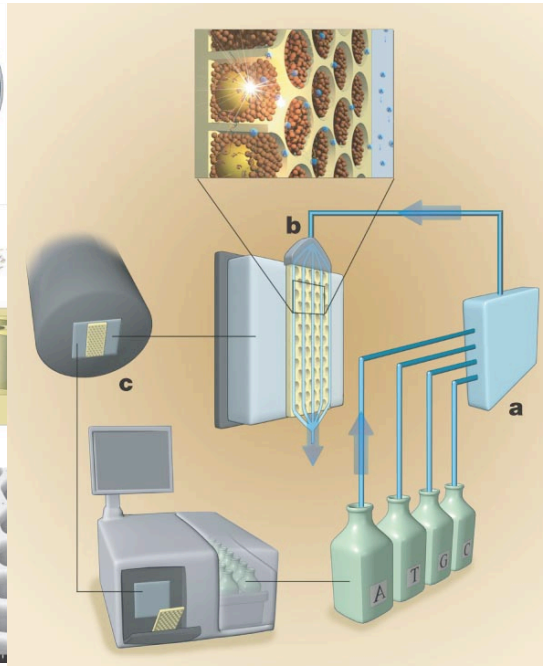












Technologies

Differences between platforms

- Technology: chemistry + signal detection
- Run times vary from hours to days
- Production range from Mb to Gb
- Accuracy per base from 0.1% to 15%
- Cost per base
- Library construction

Read length: from <100 bp to > 20 Kbp

Read length

illumina®
ion torrent
by Thermo Fisher Scientific

PACBIO®

10x GENOMICS™

Oxford NANOPORE Technologies

BIONANO GENOMICS

110
600
10000
50000
100000
300000
1000000

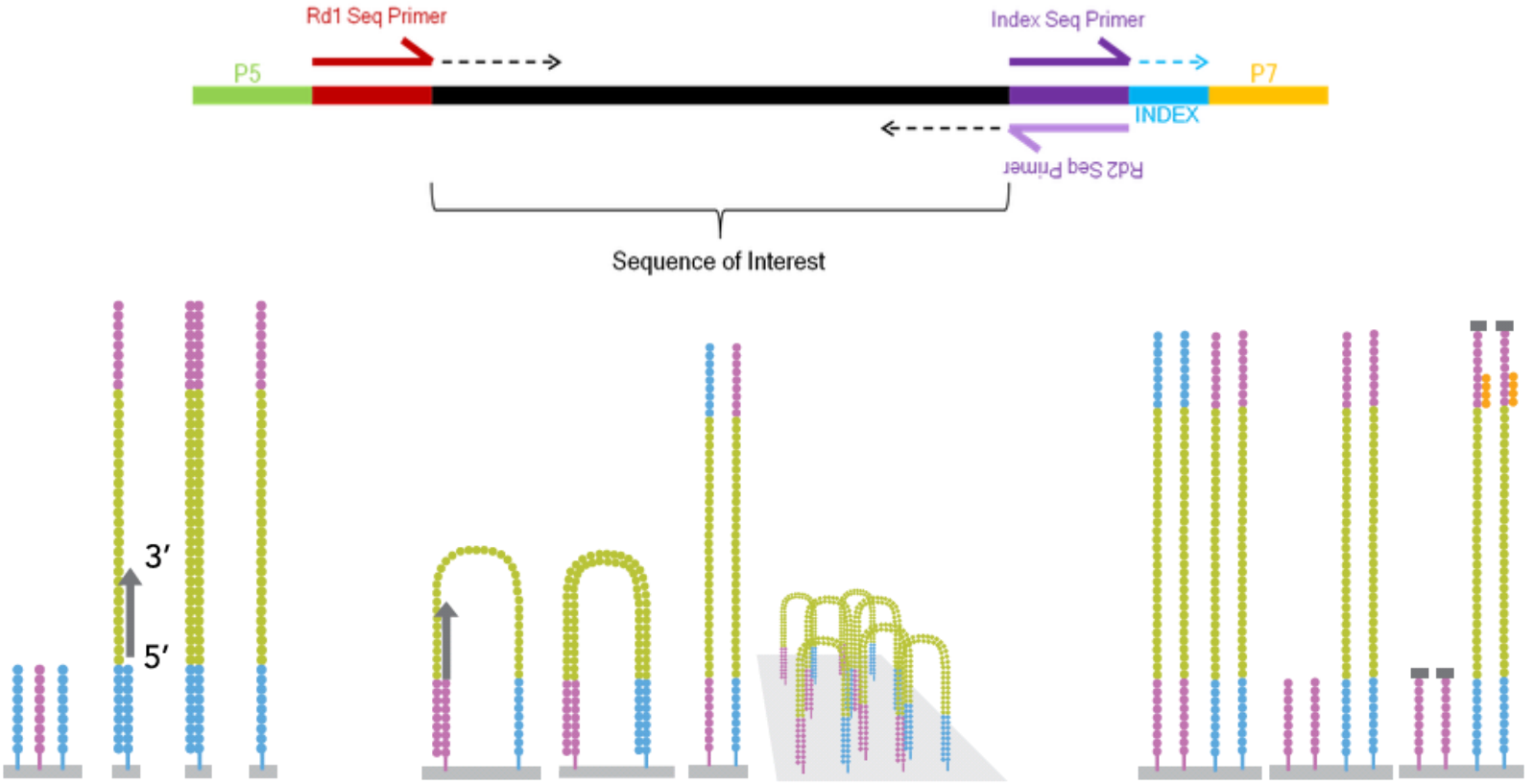
Illumina

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	"	"

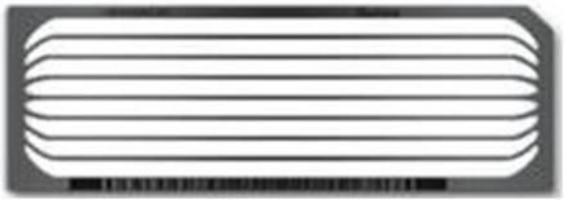
Main applications

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

Illumina: bridge amplification

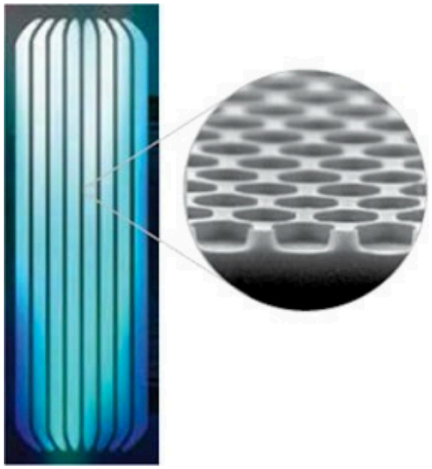


- 200M fragments per lane
- Bridge amplification
- Ends with blocking of free 3’-ends and hybridisation of sequencing primer

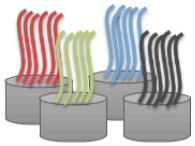


Illumina: ExAmp = black box

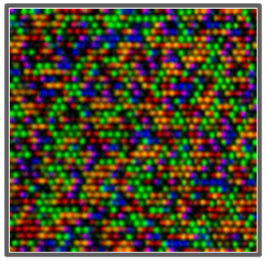
Nanowells on Patterned Flow Cell



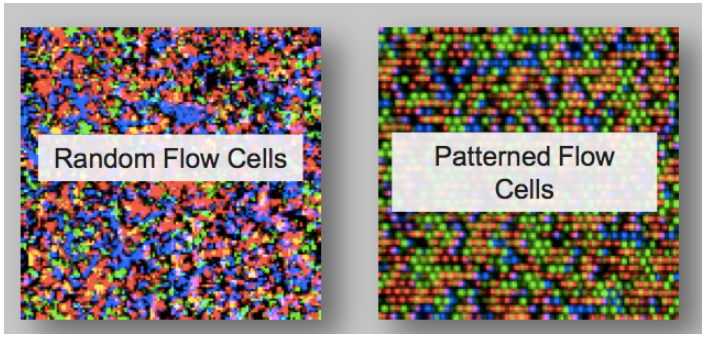
ExAmp on Patterned Flow Cell



Monoclonal wells



Ordered cluster spacing



CSH Cold Spring Harbor Laboratory

bioRχiv beta
THE PREPRINT SERVER FOR BIOLOGY

HOME | A
| CHANN
Search

New Results

Index Switching Causes "Spreading-Of-Signal" Among Multiplexed Samples In Illumina HiSeq 4000 DNA Sequencing

Rahul Sinha, Geoff Stanley, Gunsagar Singh Gulati, Camille Ezran, Kyle Joseph Travaglini, Eric Wei, Charles Kwok Fai Chan, Ahmad N Nabhan, Tianying Su, Rachel Marie Morganti, Stephanie Diana Conley, Hassan Chaib, Kristy Red-Horse, Michael T Longaker, Michael P Snyder, Mark A Krasnow, Irving L Weissman

doi: https://doi.org/10.1101/125724

Affected platforms:

HiSeqXten,
HiSeq 3000 and 4000,
NovaSeq

NovaSeq 6000

NGI acquired 2 instruments in June 2017

Flexible and scalable using multiple flow cell types

Quick and easy operation using RFID labeled reagent cassettes

Onboard clustering and automatic washing minimises hands on time during runs

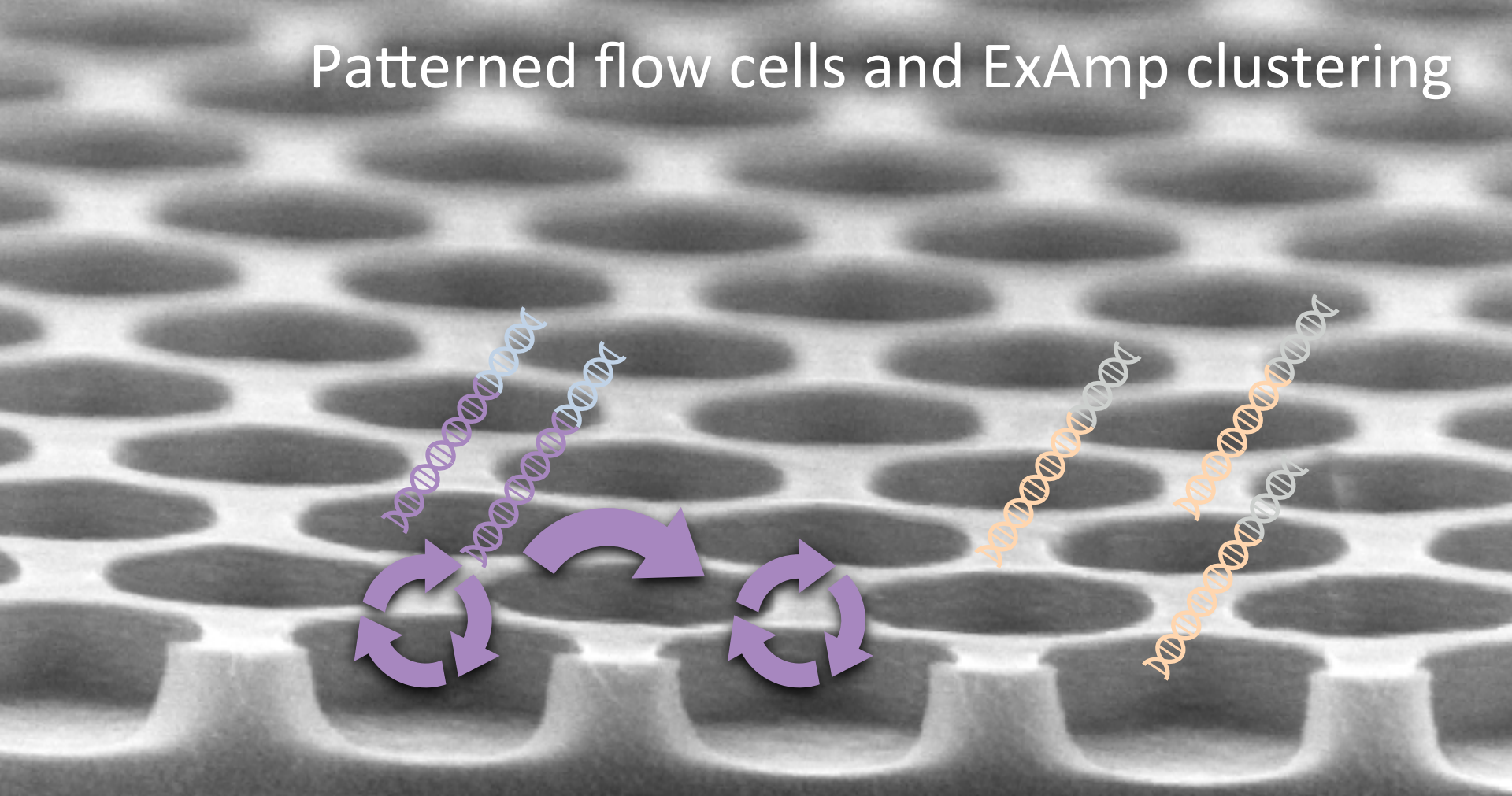
2 color chemistry

T=**Green**
C=**Red**
A=**Green**/**Red**
G=no signal

NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab

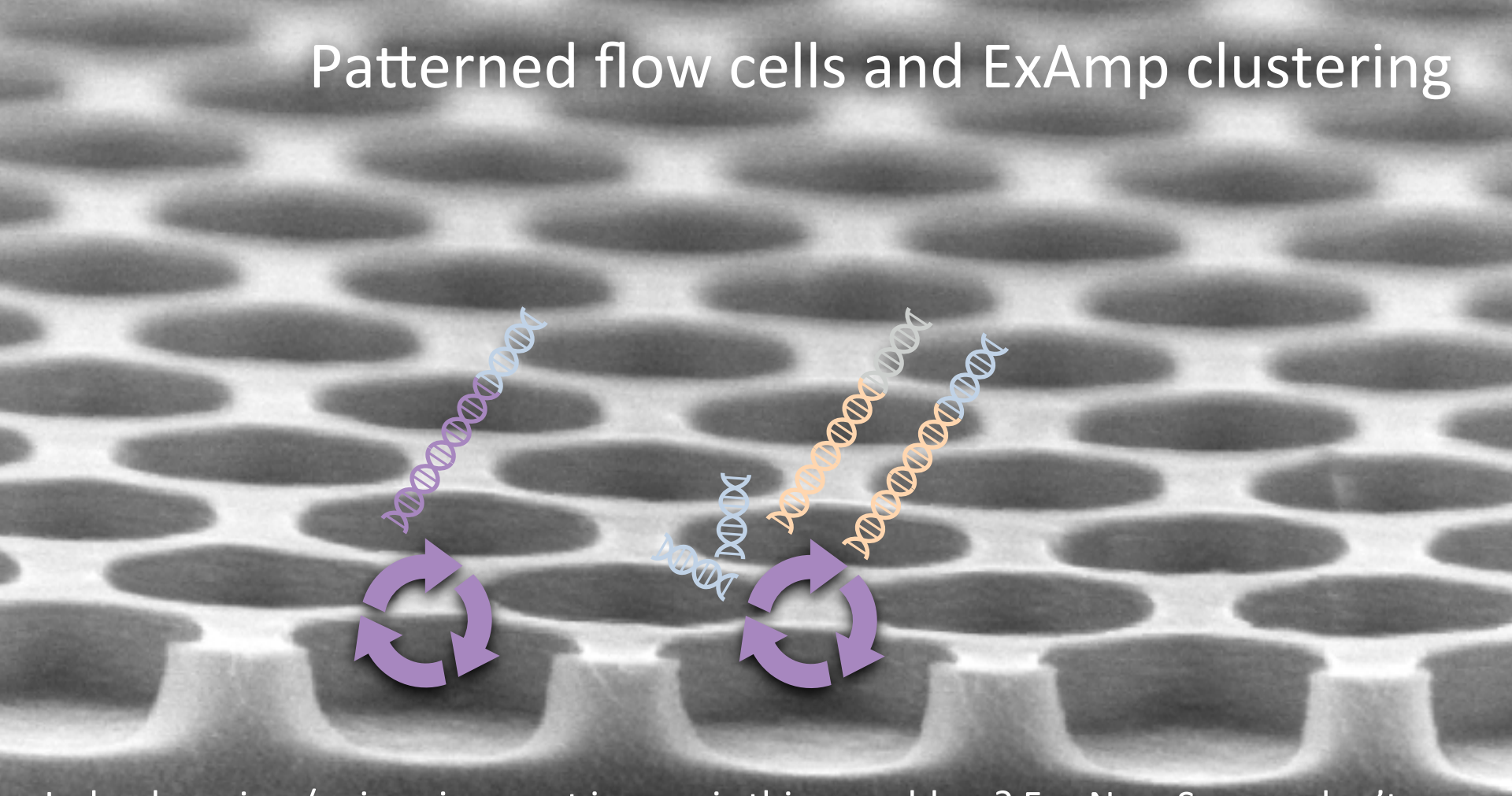
Patterned flow cells and ExAmp clustering



- More densely packed clusters → more data!
- Pre-determined cluster locations → no need for cluster mapping → faster runs!
- Exclusion amplification (isothermal seeding and cluster amplification)
- Technical duplicates / ExAmp duplicates / pad-hopping

Acc.V Spot Magn Det WD 500 nm
7.50 kV 3.0 40000x TLD 4.8 SIS XL.TIF

Patterned flow cells and ExAmp clustering

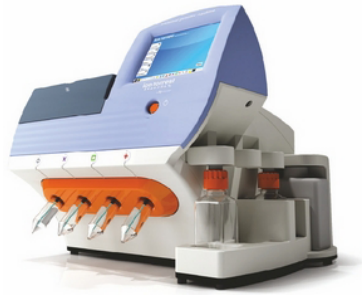


- Index-hopping / misassignment issues, is this a problem? For NovaSeq we don't know yet
- 7 pooled libraries sequenced on HiSeq X with an inline custom barcode on the DNA insert of the library. Show <1% misassigned indexes: https://doi.org/10.1101/179028
- Careful cleaning of library pools, dual indexing

Acc.V Spot Magn Det WD 500 nm
7.50 kV 3.0 40000x TLD 4.8 SIS XL.TIF

Ion

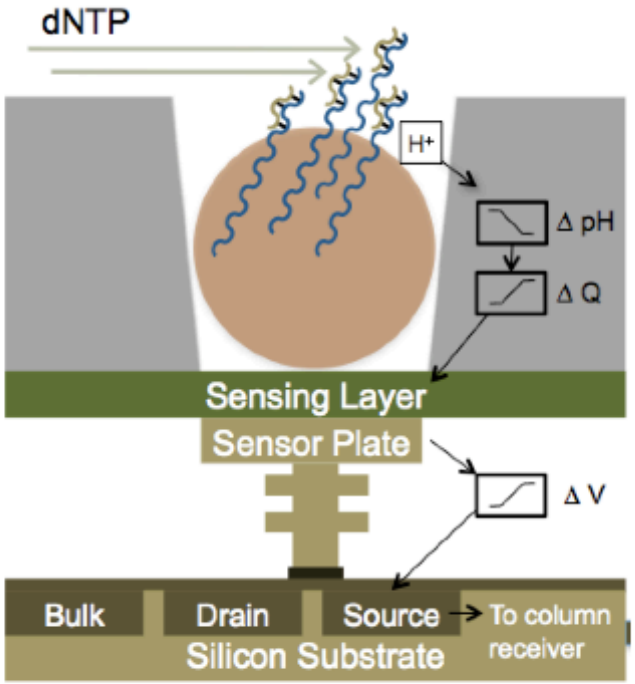
Chip	Yield - run time	Read Length
314, 316, 318 (**PGM**)	0.1 – 1 Gb Gb, 3 hrs	200 – 400 bp
P-I (**Proton**)	10 Gb 4 hrs	200 bp
520, 530, 540, 550 (**S5**)	1 Gb – 13 Gb 3 hrs	200 - 600 bp

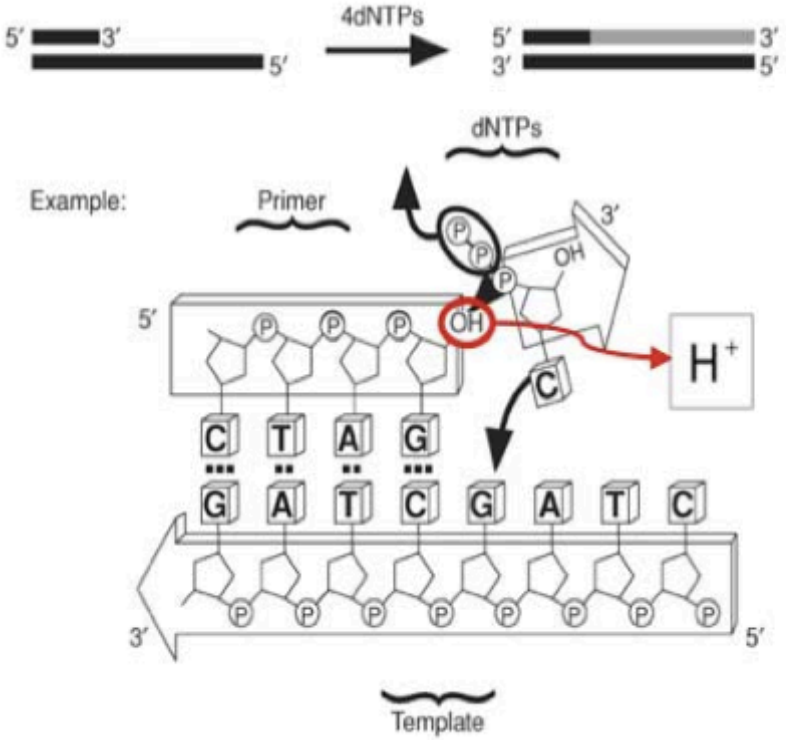


Main applications

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

Ion Torrent - H^+ ion-sensitive field effect transistors

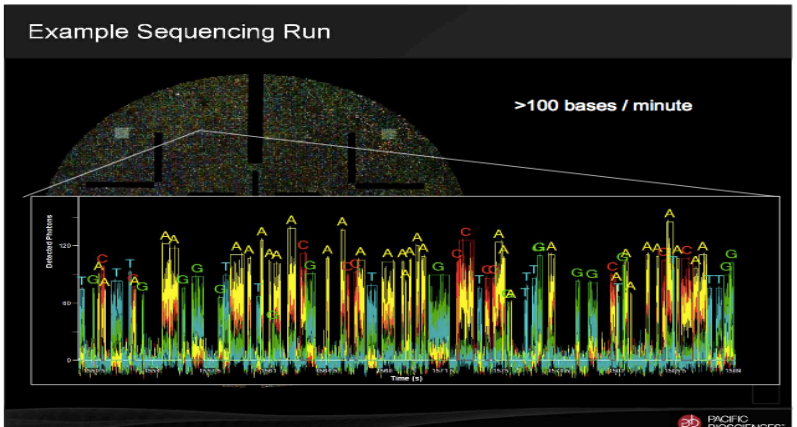




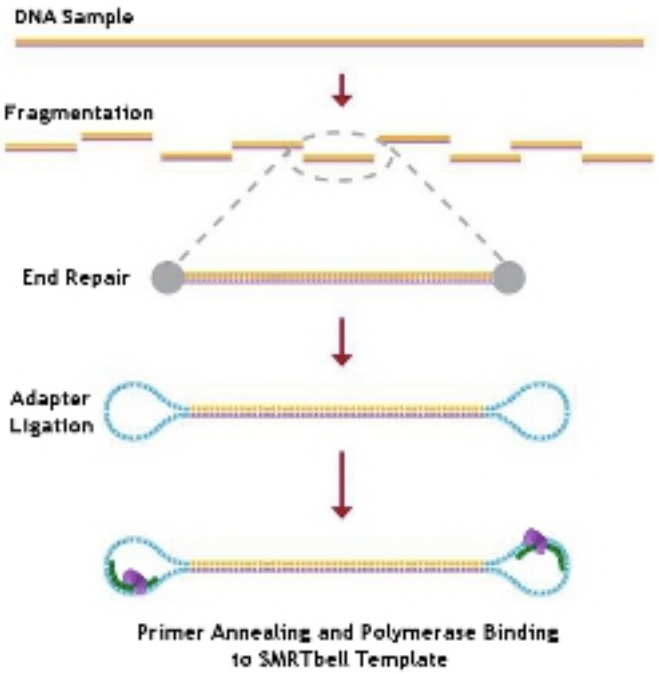
PacBio

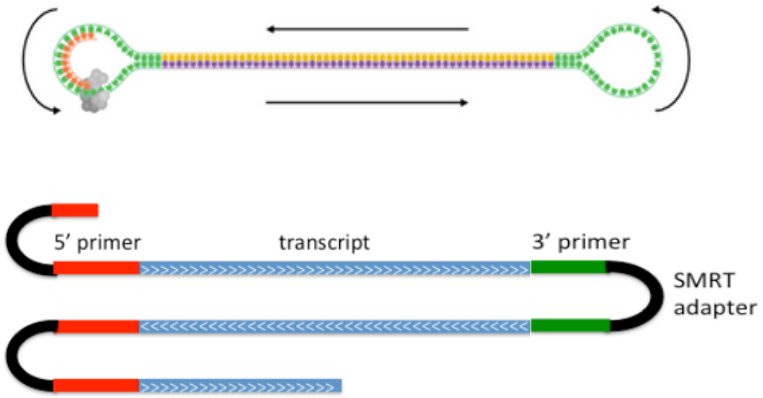
Instrument	Yield/cell and run time	Read Length	Error rate	Error type
RS II	250 Mb – 1.8 Gb 30 - 600 min	250 bp – 60 kb *(78 kb)*	15 % (single pass) 0.0001% (circular consensus)	Indels, **random**
SEQUEL	2-10 Gb 30-600 min	250 bp – 60 kb *(160 kb)*	as RSII	Indels, **random**

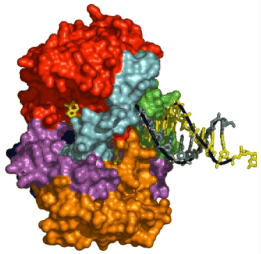
Single-Molecule, Real-Time DNA sequencing

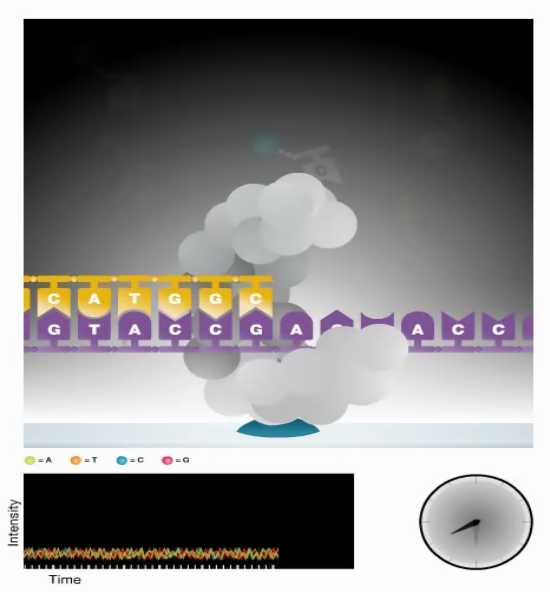


PacBio: SMRT - technology





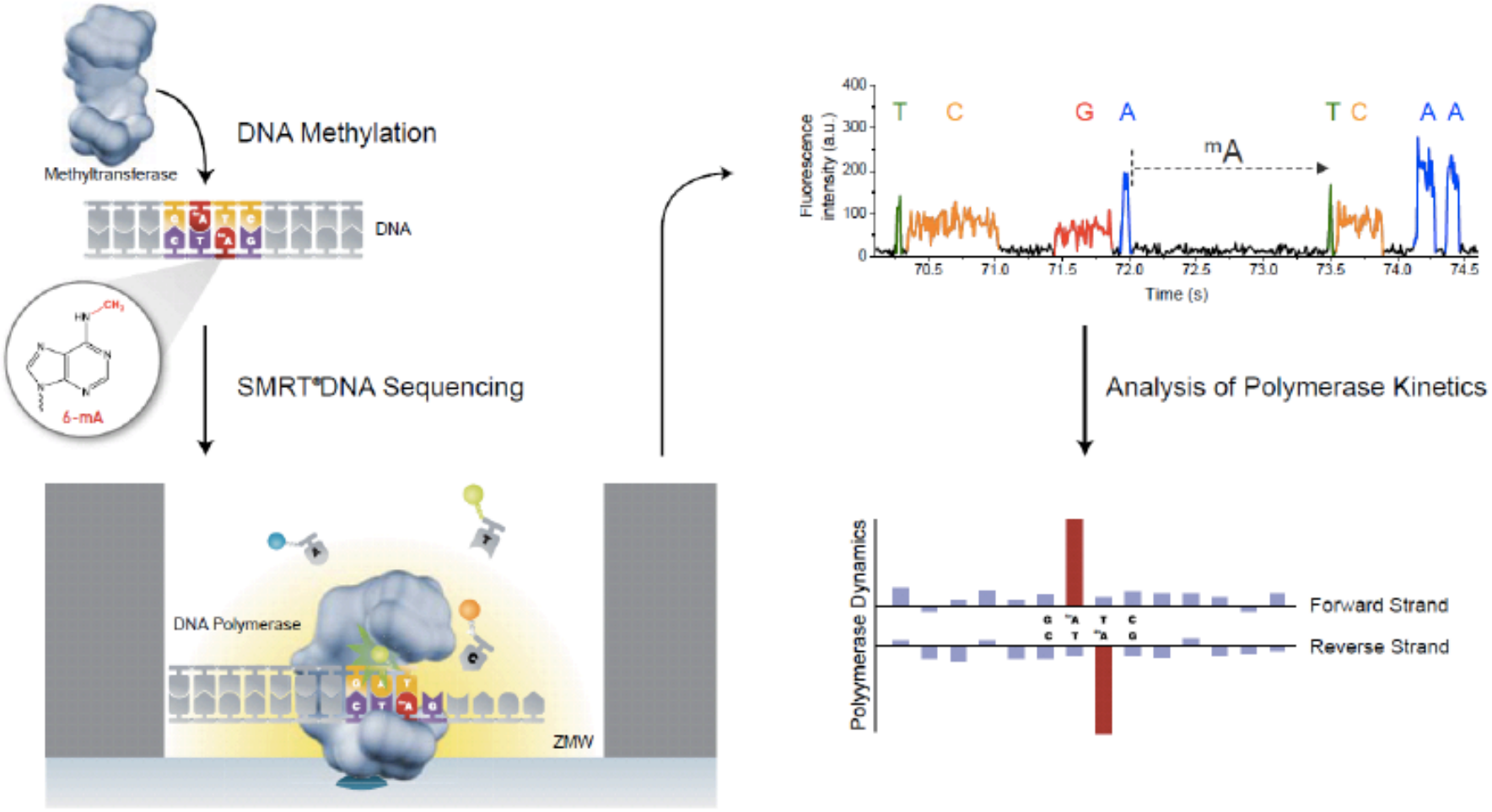




SMRT =

Single Molecule Real Time

Base Modification: Discover the Epigenome



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

SMRT sequencing: common misconceptions

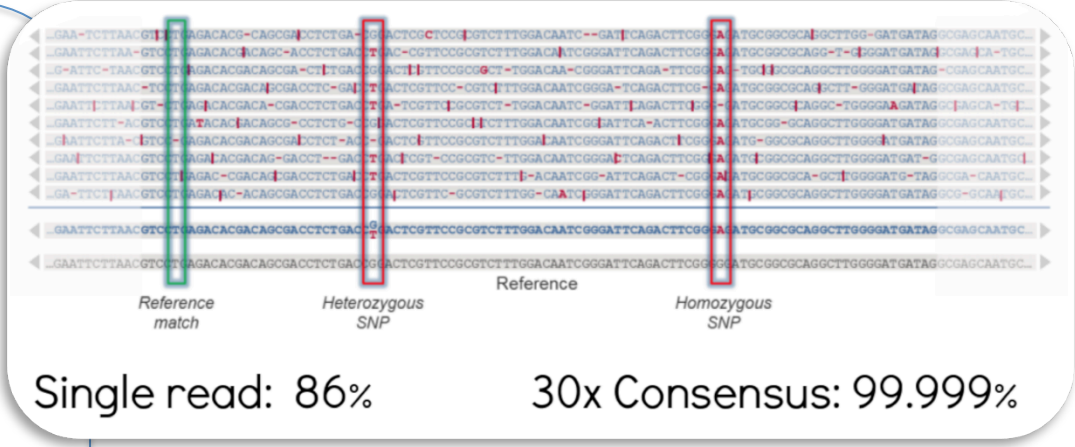
High error rate?

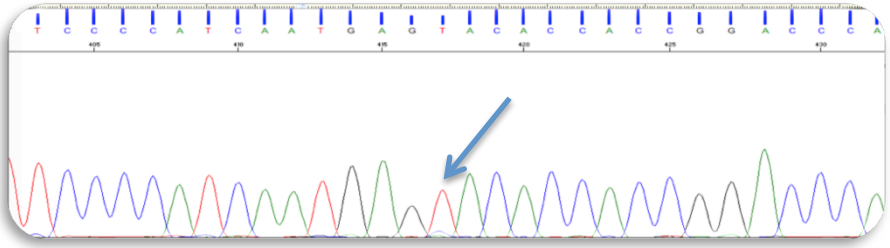
Irrelevant, because errors are random

Depending on coverage

Examples:

- 8 Mb genome, 8 SNPs detected
- 65 kb construct: 100% correct sequence
- Detection of low frequency mutations

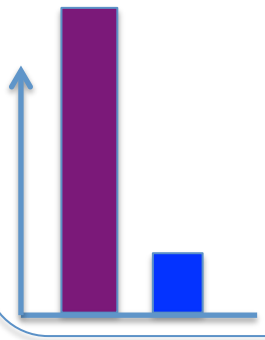




High price?

Not for small genomes

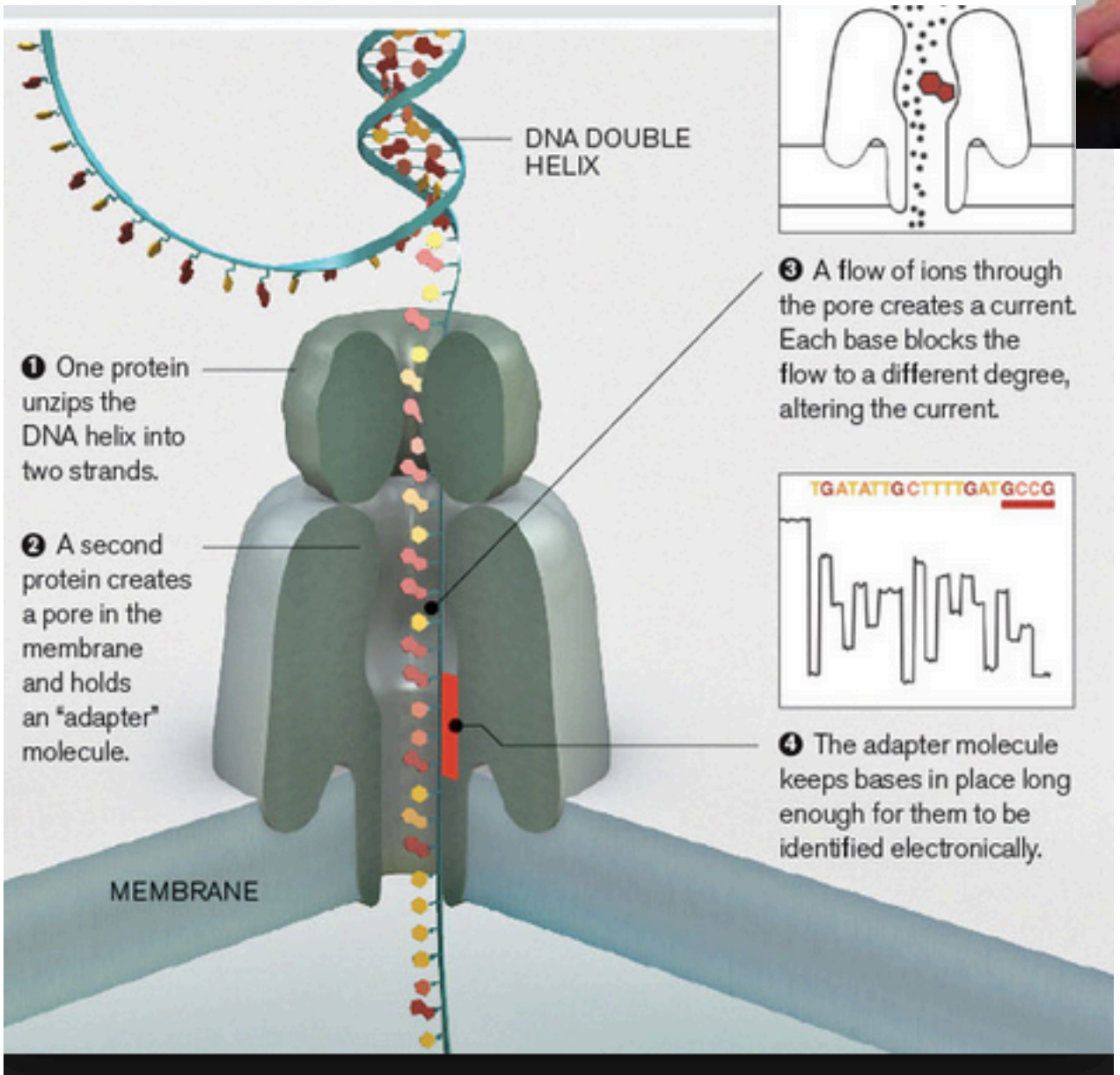
Better assembly quality
Single-molecule reads without PCR-bias



Bioinfo-time to assemble short reads

Bioinfo-time to assemble long reads

Oxford Nanopore MinION



Reads up to 800k
10-15% 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
PacBio Sequel

Ultra-long reads
FAST throughput

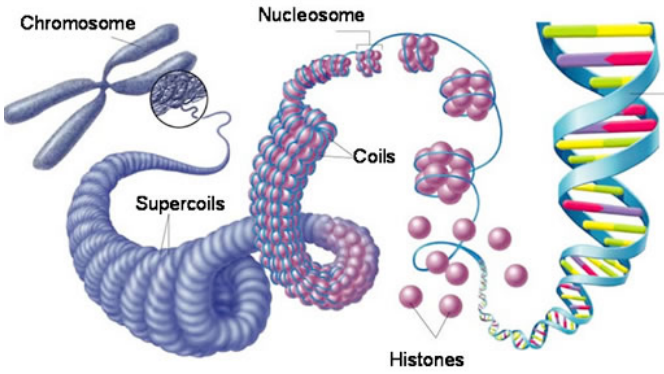
Applications

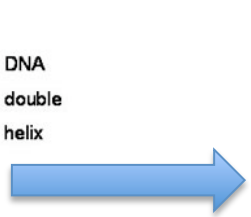
NGS/MPS applications

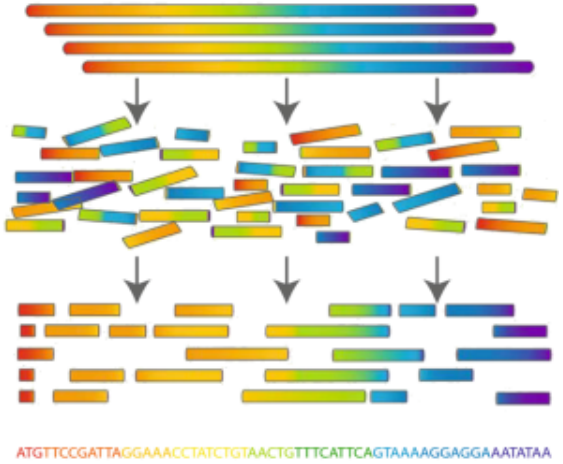
- Whole genome sequencing:
  - De novo sequencing
  - Re-sequencing
- Transcriptome sequencing:
  - **mRNA-seq**
  - **miRNA**
  - Isoform discovery
- Target re-sequencing
  - Exome
  - Large portions of a genome
  - Gene panels
  - **Amplicons**

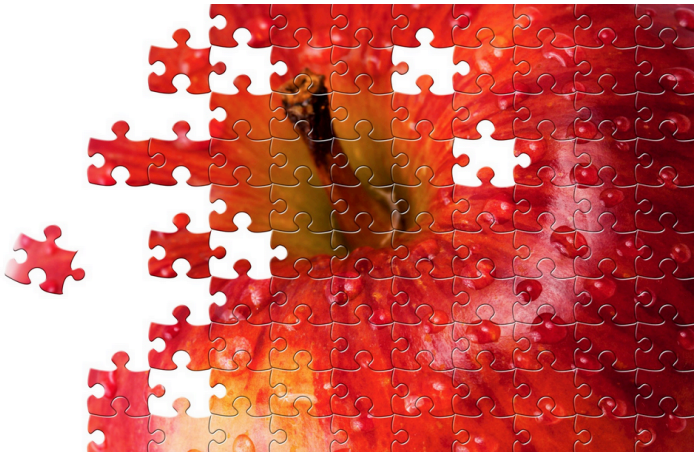
De novo sequencing

- Used to create a reference genome without previous reference

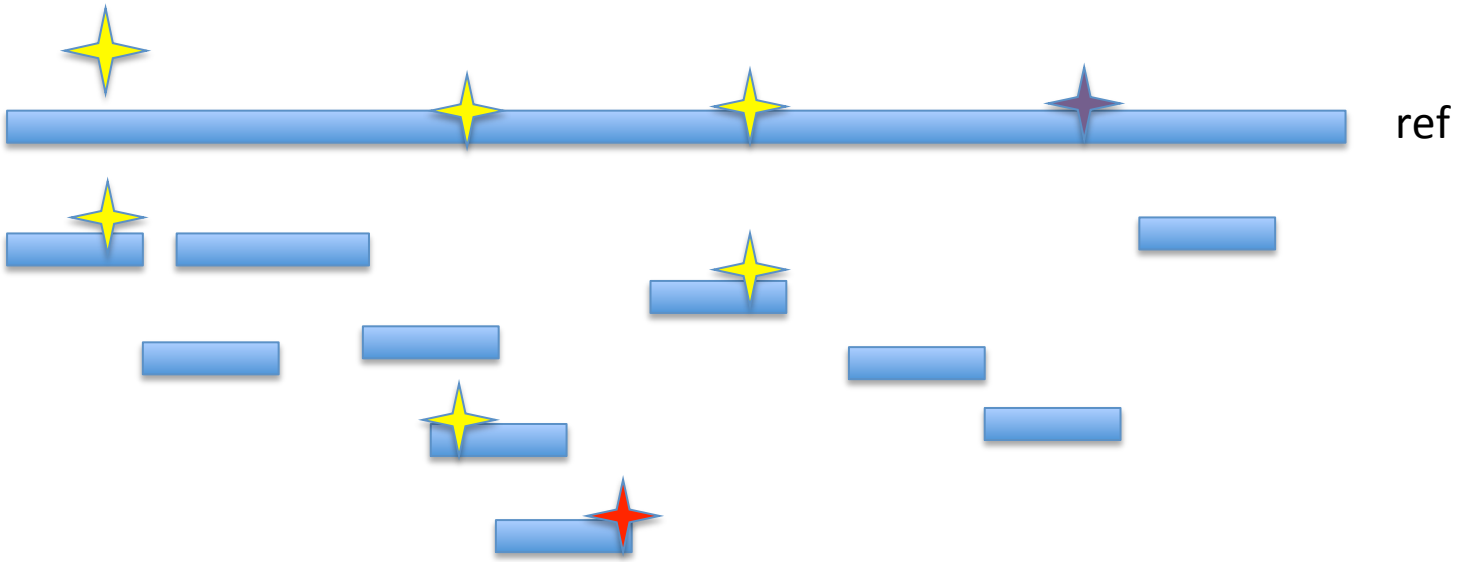








De novo vs re-sequencing



De novo

No bias towards a reference
No template to adapt to

Many contigs
Works best for large-scale events

Re-seq

Finding similarities to a reference
Easier to identify SNPs and minor events
Fewer contigs

Novel events are lost

De novo – do it with long reads!

Science AAAS

Home | News | Journals | Topics | Careers

Science | Science Advances | Science Immunology | Science Robotics | Science Signaling | Science Translational Medicine

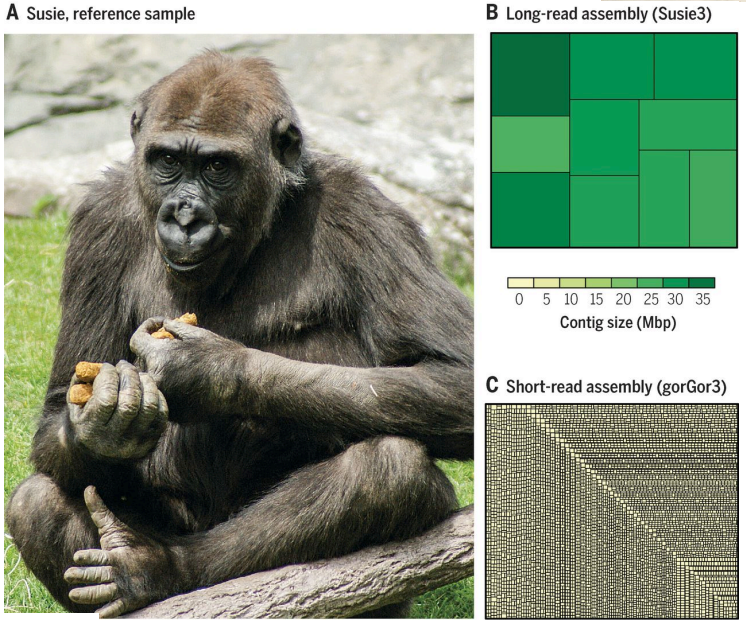
SHARE RESEARCH ARTICLE

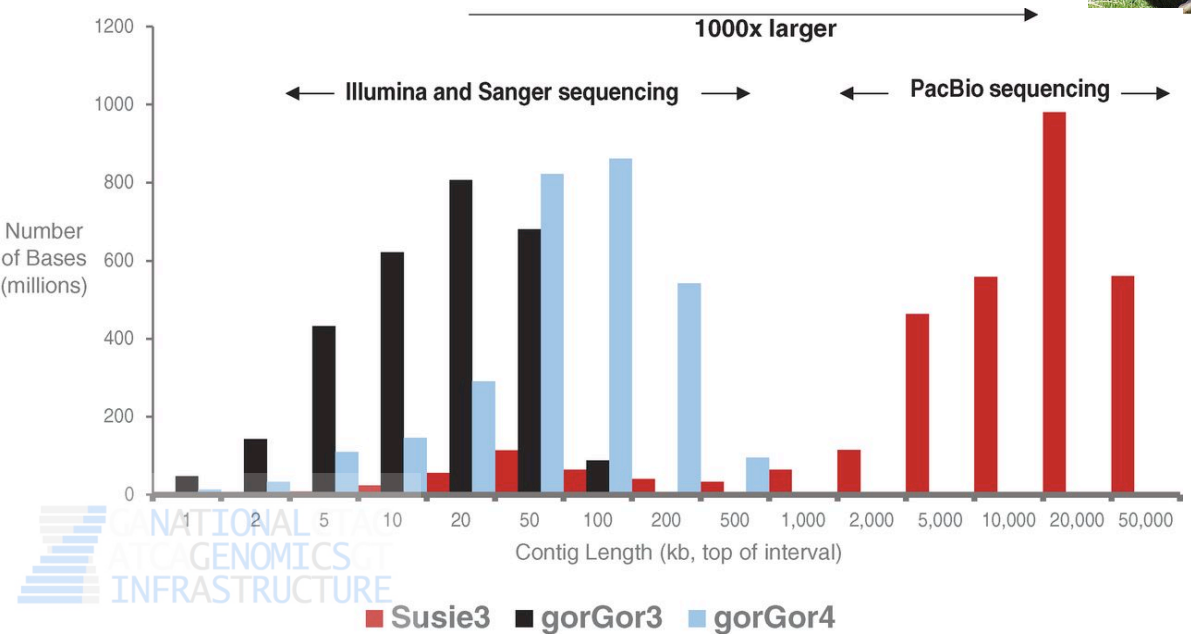
Long-read sequence assembly of the gorilla genome

David Gordon[1,2,*], John Huddleston[1,2,*], Mark J. P. Chaisson[1,*], Christopher M. Hill[1,*], Zev N. Kronenberg[1,*], Katherine ...
+ See all authors and affiliations

Science 01 Apr 2016:
Vol. 352, Issue 6281, aae0344
DOI: 10.1126/science.aae0344

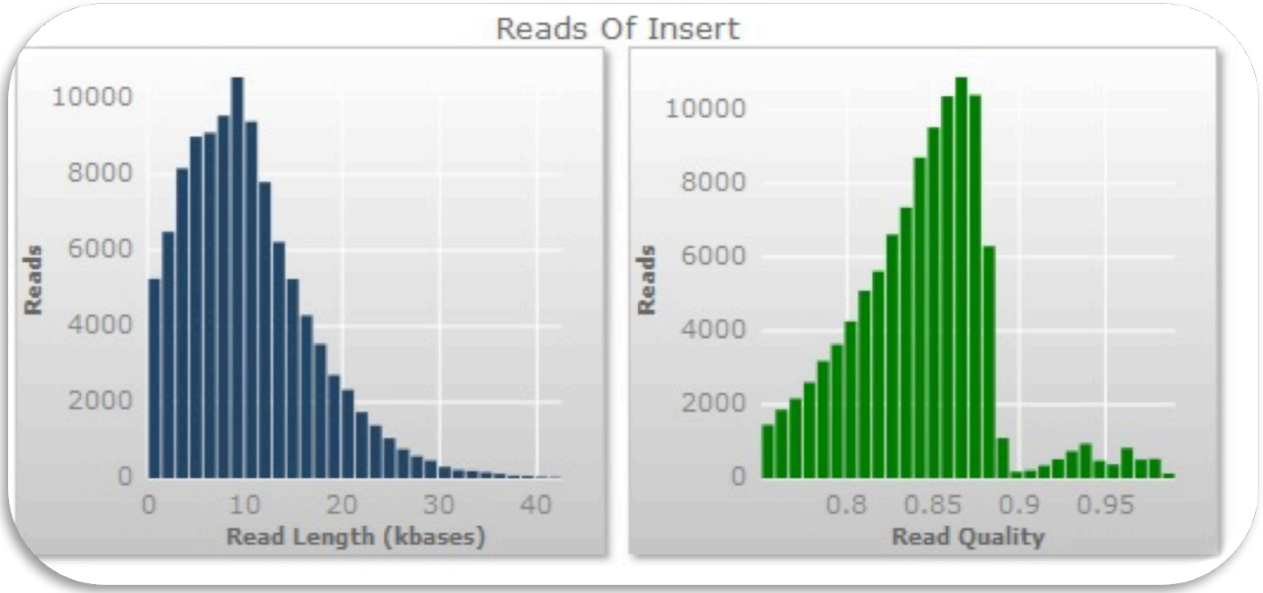
Peer Reviewed
<- see details





TEMPORA MVTANTVR ET NOS MVTAMVR IN ILLIS

Example: de novo PacBio; Crow



Sequencing results

Number of SMRT cells: 70

Total bases per SMRT: 1.39 Gb

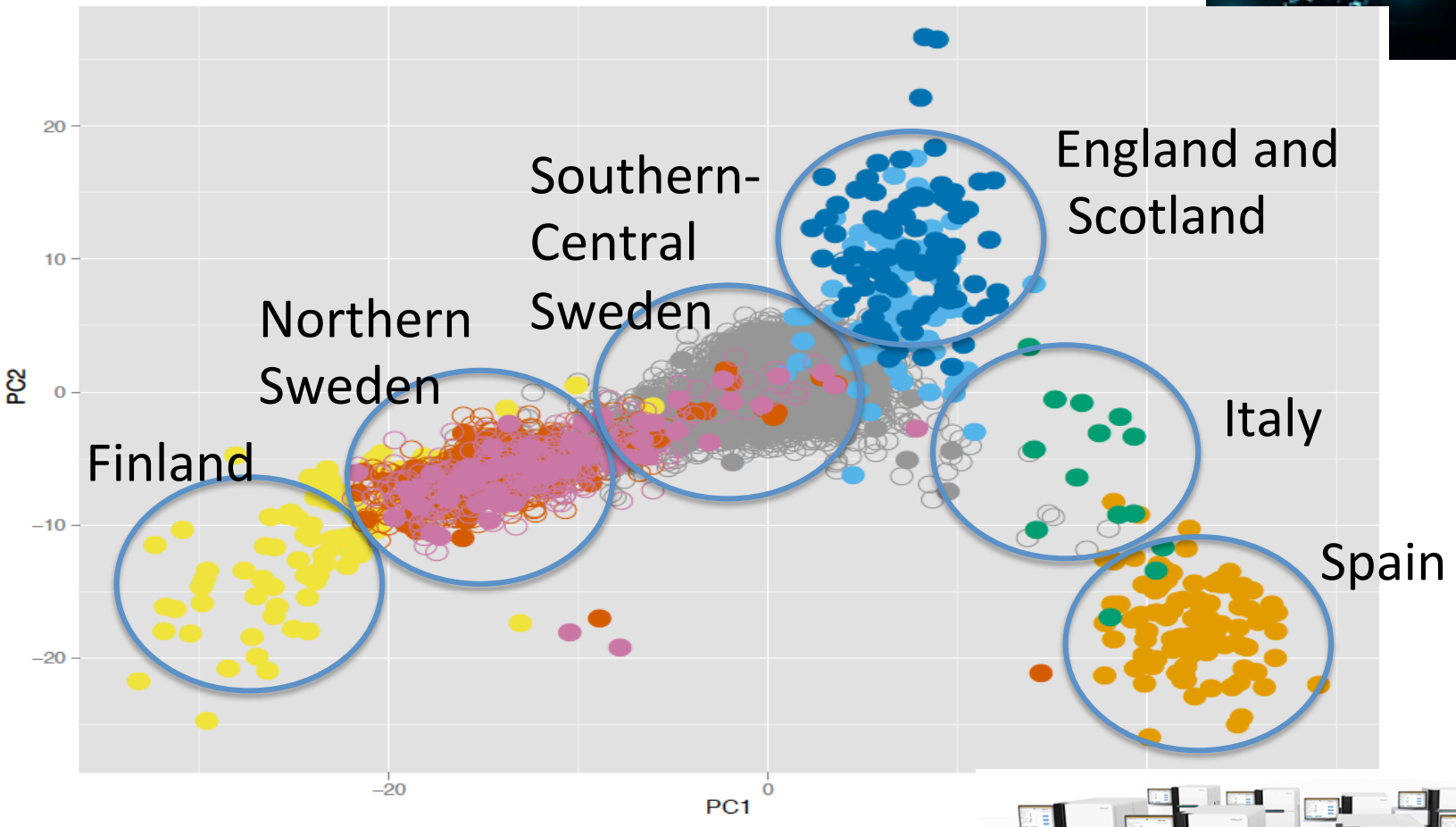
Total reads per SMRT: 106 833

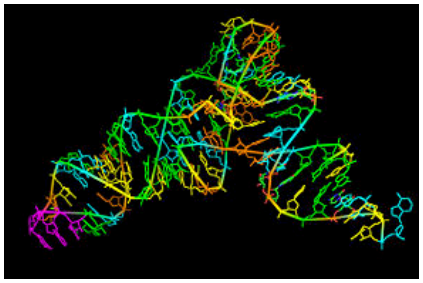
Assembly results, FALCON

	PRIMARY	ALTERNATIVE
N50	8.5 Mb	23 kb
N75	3.9 Mb	18 kb
Nr contigs	4375	2614
Longest contig	36 Mb	121 kb
Total length	1.09 Gb	45 Mb

Re-sequencing

Population studies: Illumina HiSeq is **The Best**





Transcriptome sequencing (RNA-seq)

TOTAL RNA

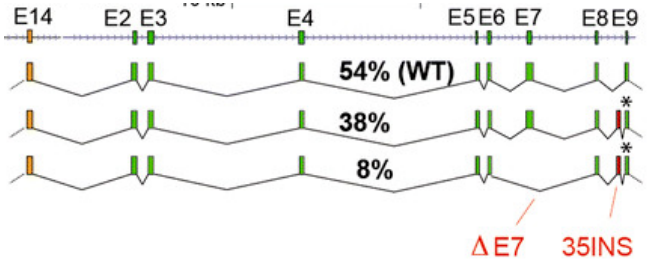
mRNA

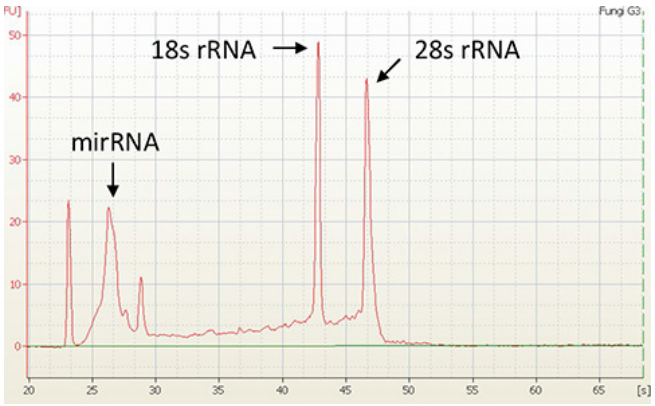
- **Dif.ex.**
- Annotation

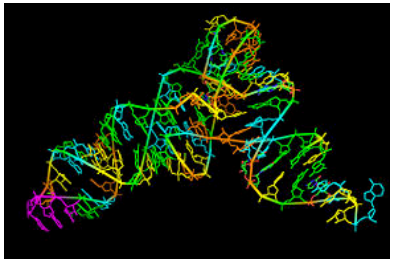
Non-codingRNA **miRNA**

- Transcriptional regulation

Splice isoforms





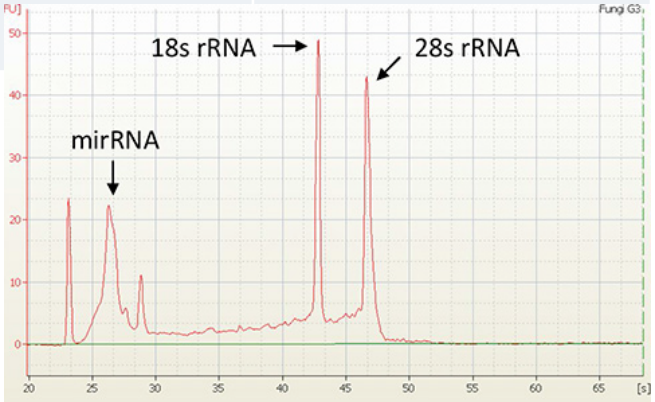


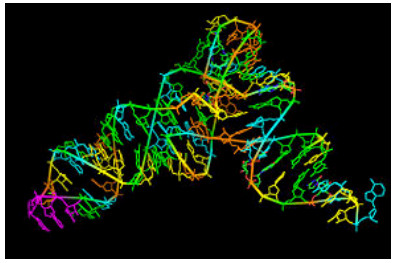
mRNA: rRNA depletion vs polyA selection

Method	Pros	Cons	Recommended
rRNA depletion	• Captures on-going transcription • Picks up non-coding RNA	• Does not get rid of all rRNA • Messy Dif.Ex. profile	20-40 mln reads (single or PE)
polyA selection	• Gives a clean Dif.Ex. profile	• Does not pick non-coding RNA	5-20 mln reads

Alternative for **human** RNA-seq:
AmpliSeq Human Transcriptome panel:

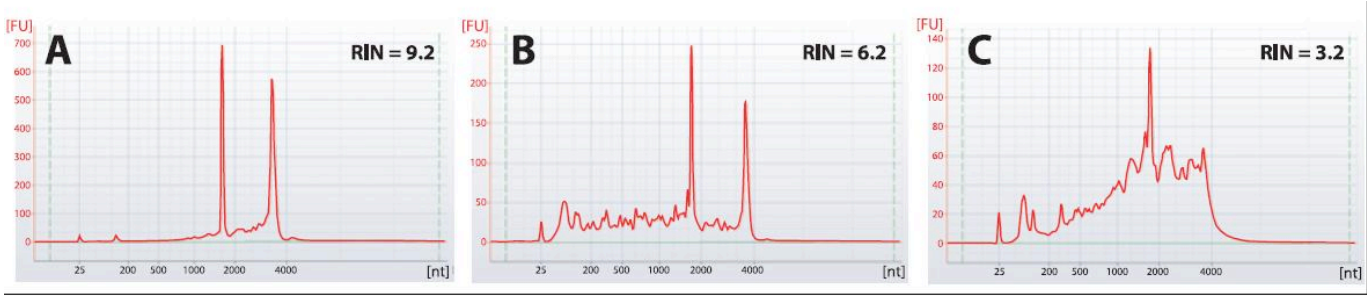
- faster, cheaper, works fine with FFPE
- input: 50 ng **total** RNA
- dif.ex. ONLY



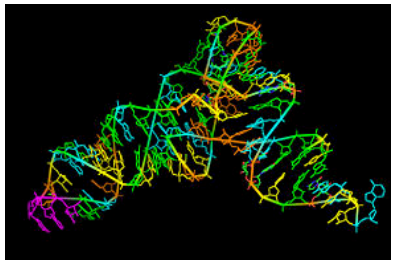


RNA-seq experimental setup

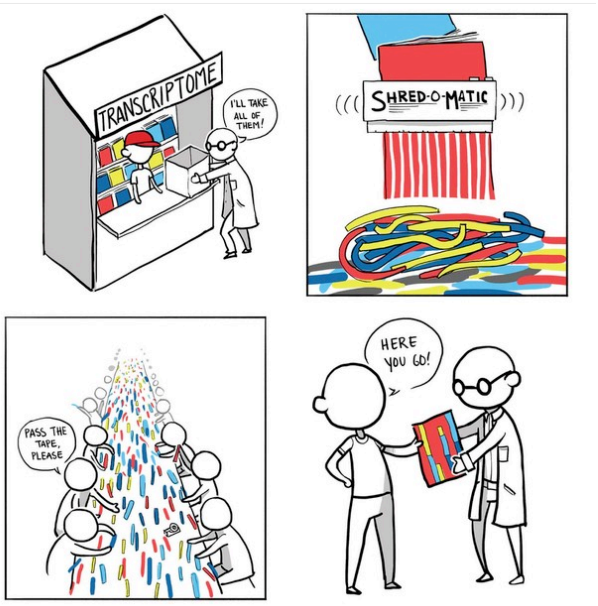
- mRNA only: any kit
- mRNA **and** miRNA: only specialized kits
- Always use DNase!
- RIN value above 8.

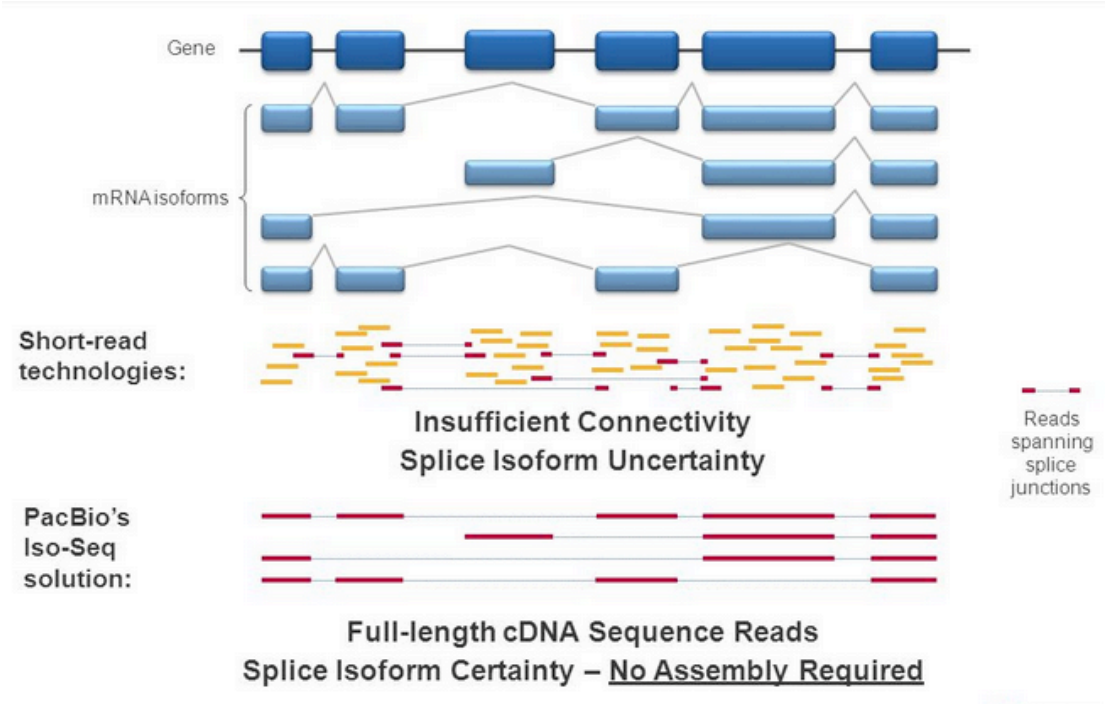


- CONTROL vs experimental conditions
- Biological replicates: 4 strongly recommended



RNA-seq experimental setup





NATURE METHODS | NEWS AND VIEWS

Genomics: the state of the art in RNA-seq analysis

Ian Korf

Nature Methods **10**, 1165–1166 (2013) | doi:10.1038/nmeth.2735
Published online 26 November 2013

PacBio Iso-seq: full-length transcriptome seq

Targeted re-sequencing

Suitable applications for target-seq

- Metagenomics
- Resolving complex regions
- Low frequency mutations
- Human re-sequencing
- Clinical diagnostics
- ....

Approaches

- Hybridization capture (Agilent, NimbleGen, MyBaits)
- PCR (Amplicon sequencing)
  - Long-range
  - Conventional
  - Multiplex
- *Experimental:*
  - *TLA, Samplix, CRISPR-Cas9)*

Example: R&D, sequence capture

Ida Höijer

Modified Agilent SureSelect protocol

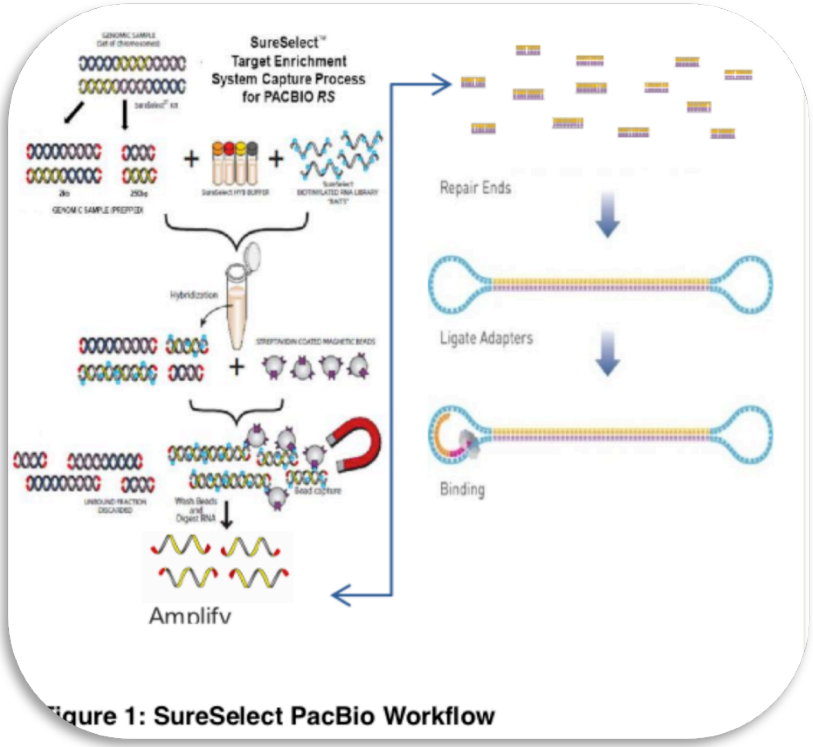
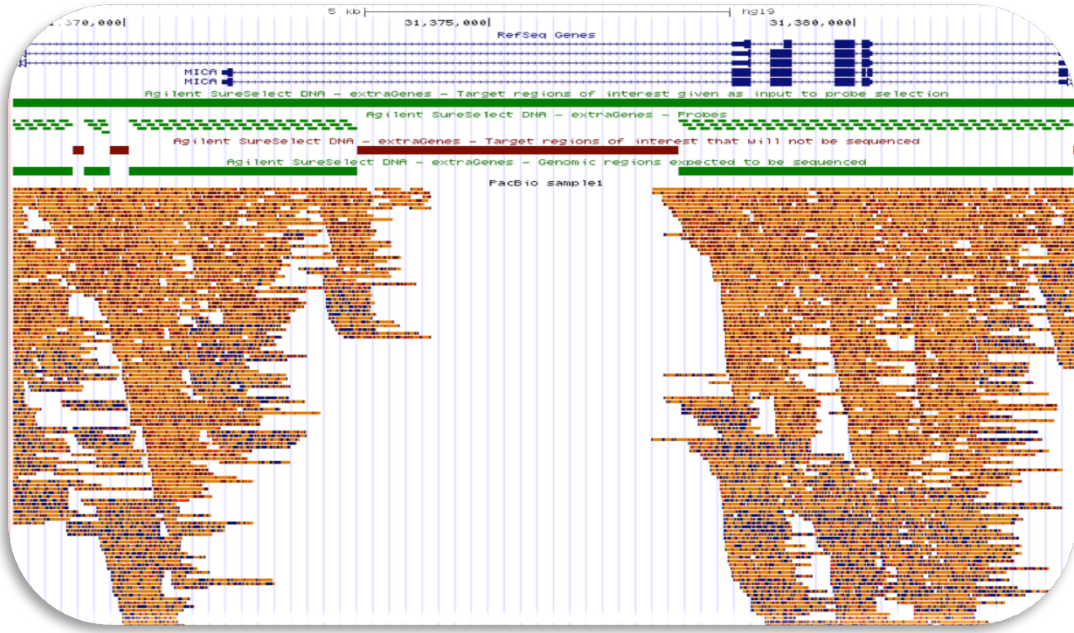
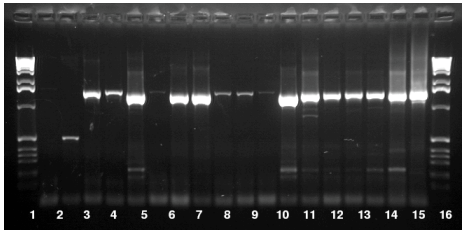


Figure 1: SureSelect PacBio Workflow

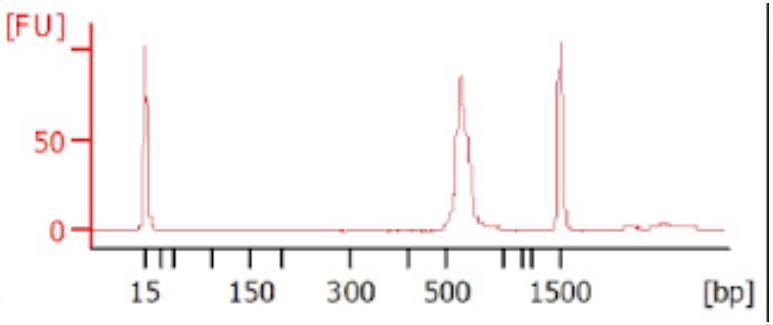


PacBio bridges gaps in Sure Select design
Resolution of gene paralogs and duplications

Amplicon sequencing



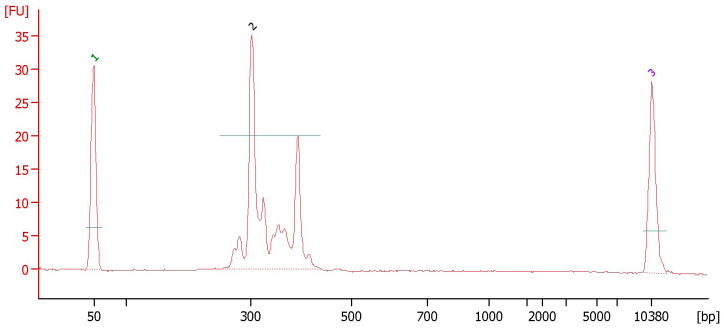
Example 1: tight peak, OK



FOR ANY NGS TECHNOLOGY

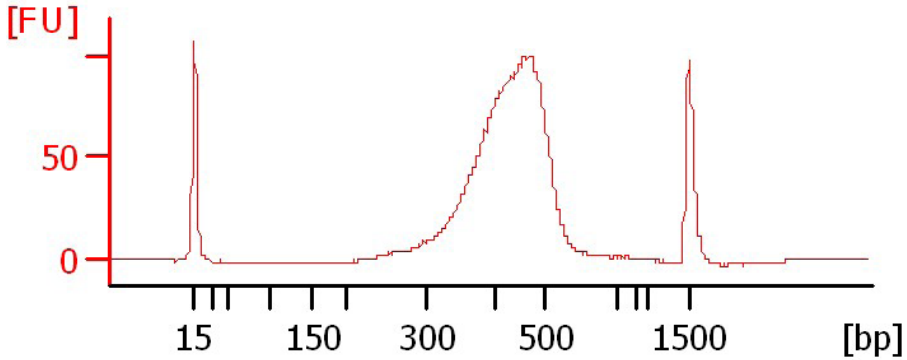
Size difference among fragments **must not** exceed 80 bp (or 20% in length)

Reason – preferential amplification of short fragments



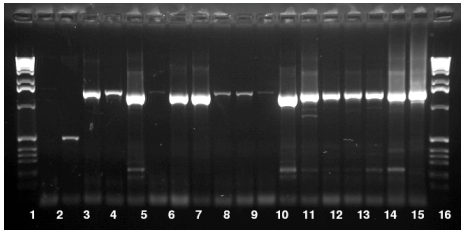
Example 2: several sizes, fractionation is needed

=> we HAVE to make several libraries

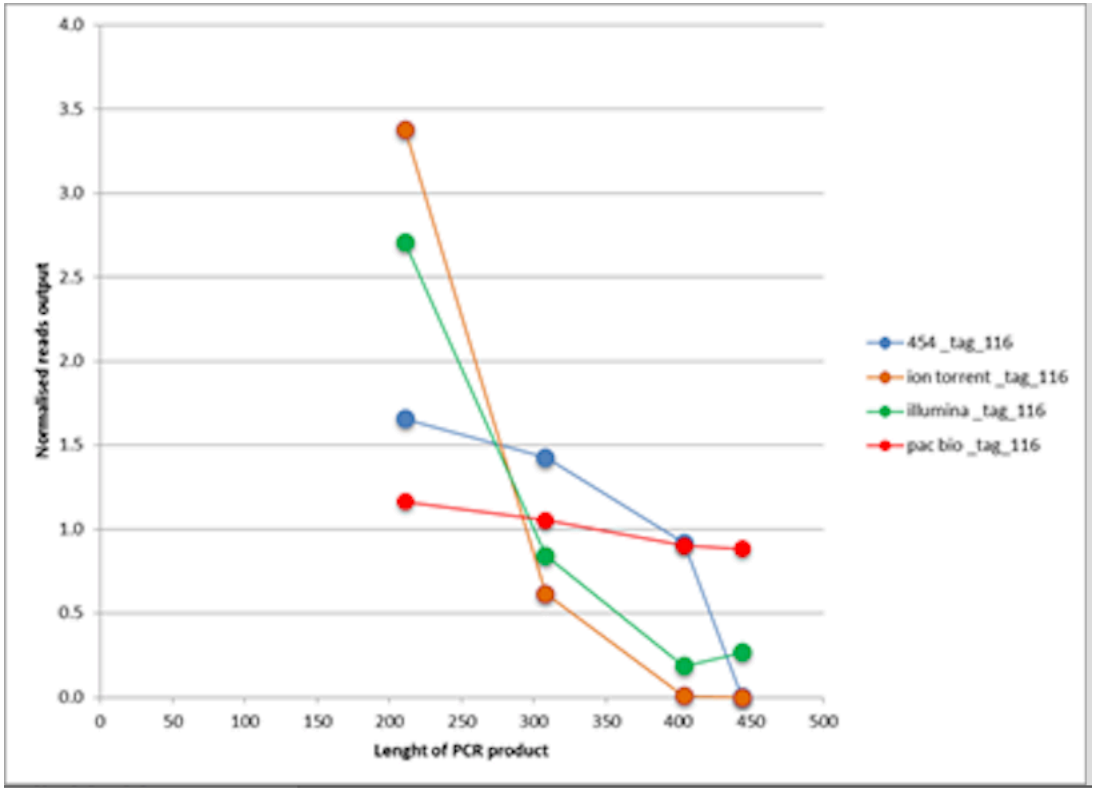


Example 3: broad peak; size selection is needed

SIZE MATTERS…



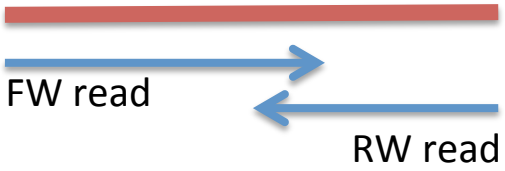
Size-related bias in amplicon-seq



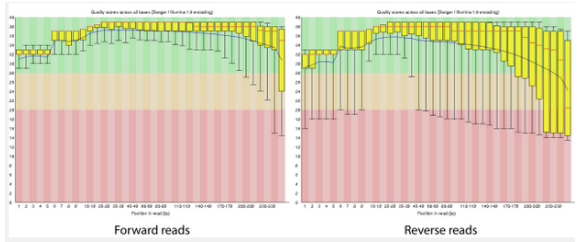
Courtesy Mikael Brandström Durling, Forest Mycology and Pathology, SLU

Amplicon sequencing: Technologies

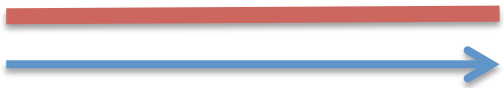
Illumina MiSeq



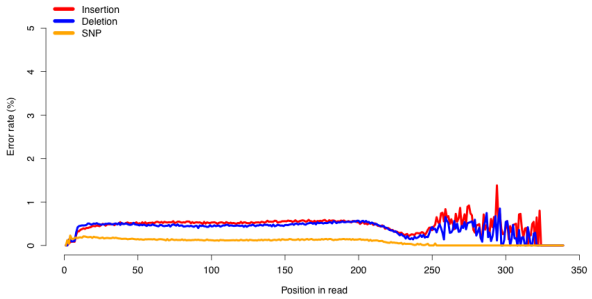
Paired-end reads



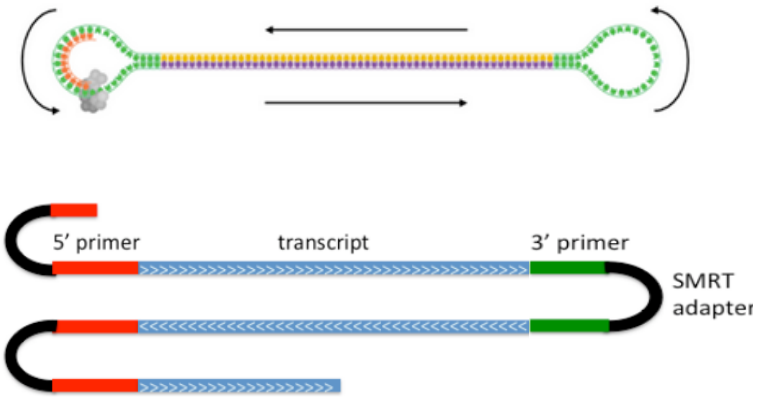
Ion S5XL



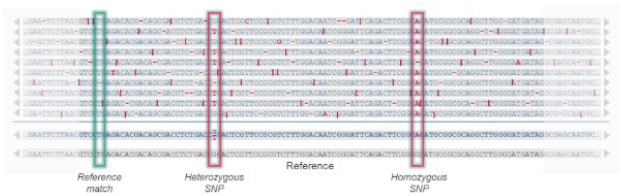
Single-end reads



PacBio RSII



Circular consensus reads

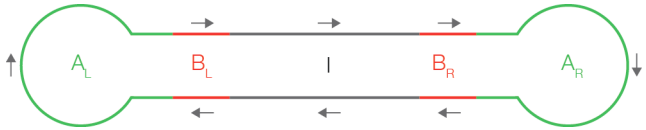


NATIONAL GENOMICS INFRASTRUCTURE

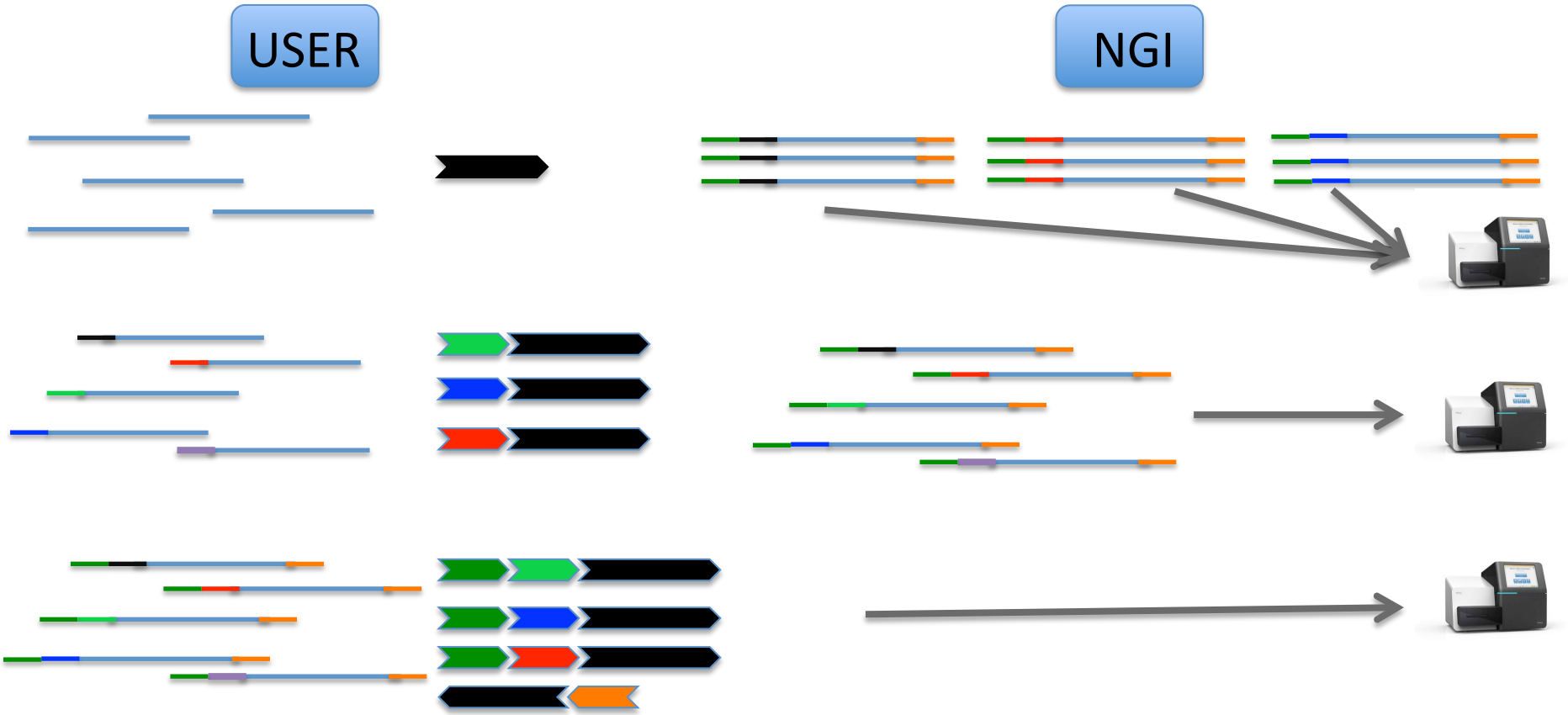
SciLifeLab

Amplicon sequencing: Barcoding strategies

Illumina and Ion



PacBio



Main types of equipment & applications

Illumina HiSeq
NextSeq, X10, MiSeq,
MiniSeq, NovaSeq

Short paired reads
HIGH throughput

Human WGS
Re-sequencing 30x
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

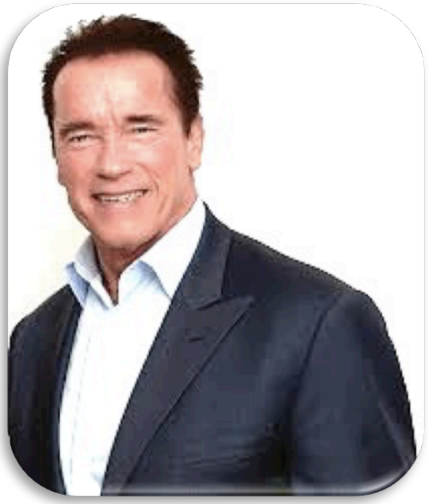
mRNA and miRNA
Exome
ChIP-seq
Short amplicons
Gene panels
Clinical samples

PacBio RSII
SEQUEL

Ultra-long reads
FAST throughput

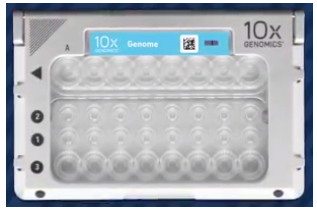
Long amplicons
Re-sequencing
De novo sequencing
Novel isoform discovery
Fusion transcript analysis
Haplotype phasing
Clinical samples

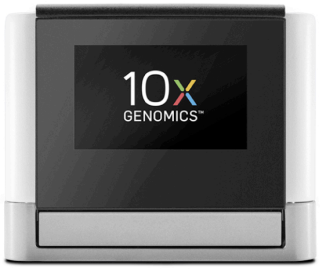


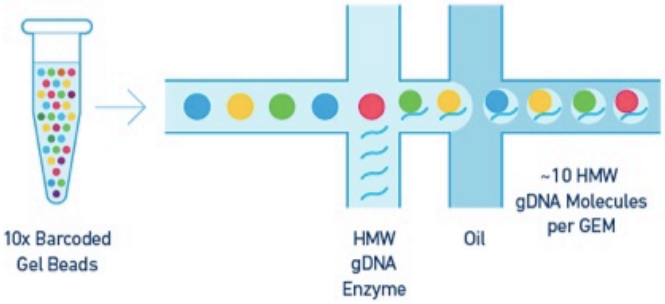



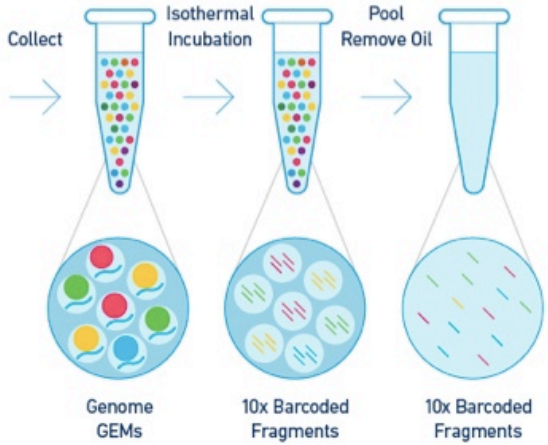

But there is more!

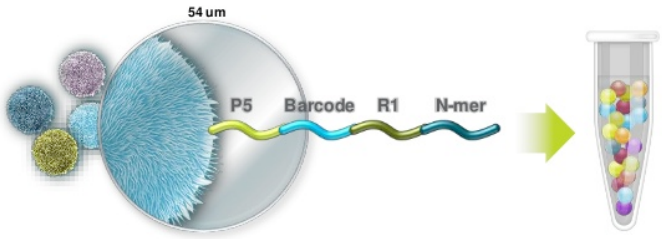
10x Genomics (Chromium)

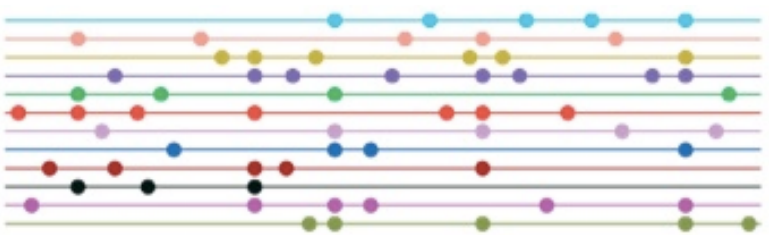












Fragment length: 50 kb – 100+ Kb

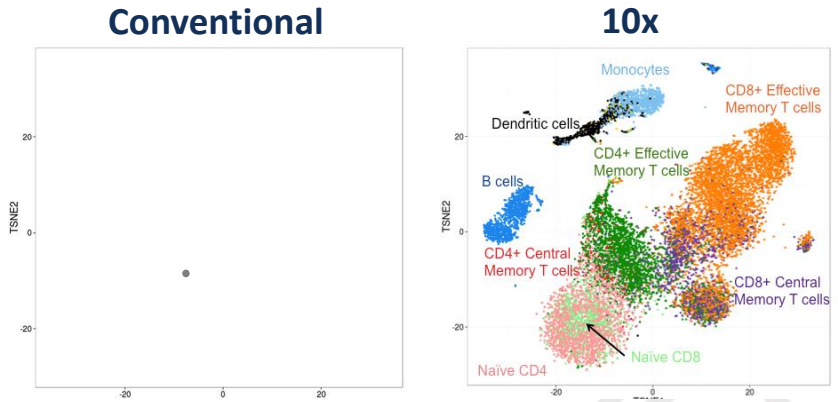
UPPSALA UNIVERSITET

Chromium applications at the NGI

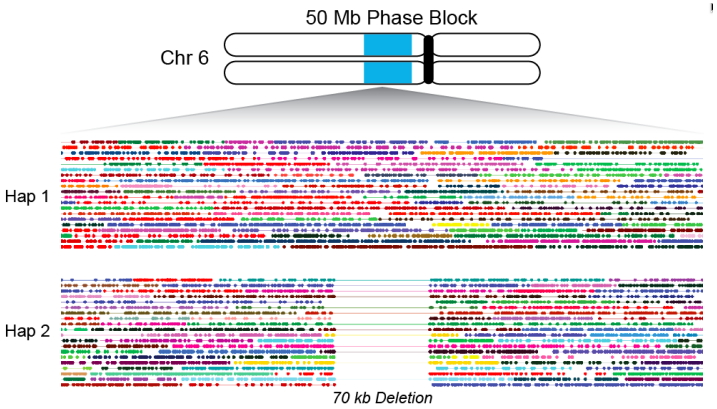
Single Cell 3' RNA:
- Up to 10,000 cells
- Human/mouse
- Fresh/Frozen PBMCs
- Cell lines
- Fractionated cells
- Nuclei (untested)

Chromium Genome:
- Structural variant detection
- Haplotyping
- SNP calling
- De novo assemble:
  - Birds
  - Fish
  - Plants
  - Mammals

Single cell RNA-seq



Phased variant detection



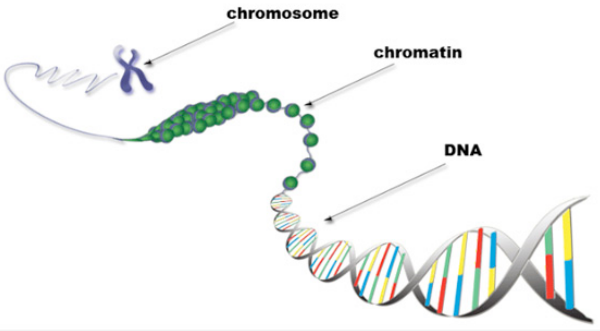
SAMPLE QUALITY REQUIREMENTS

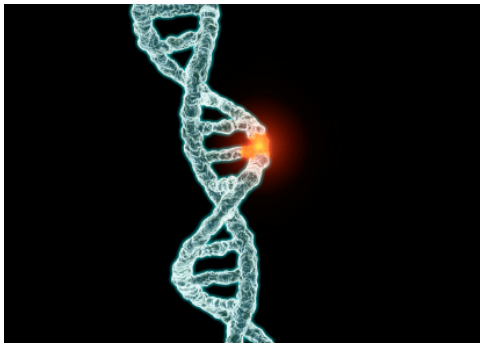
Sample prep: take home message

PCR-quality sample and

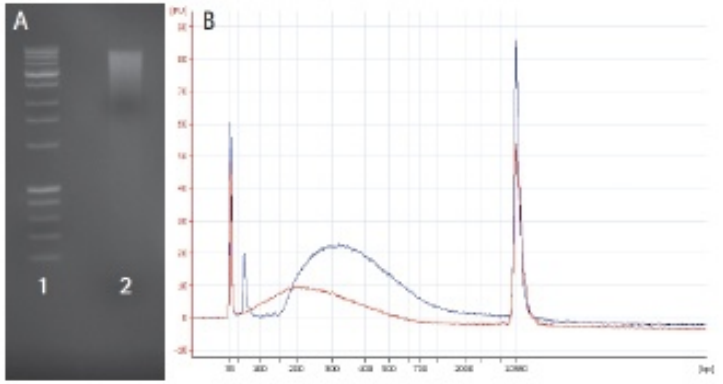
NGS-quality sample

are two completely different things

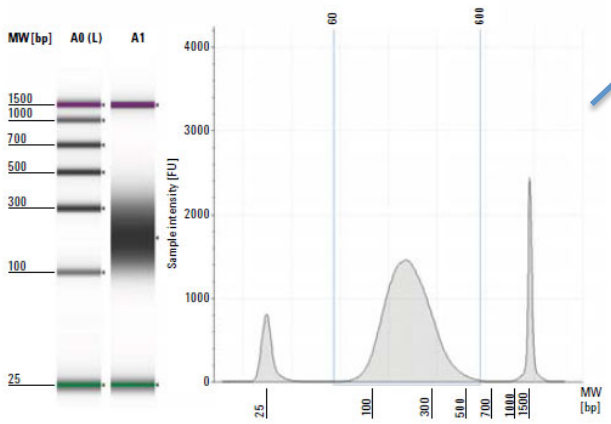




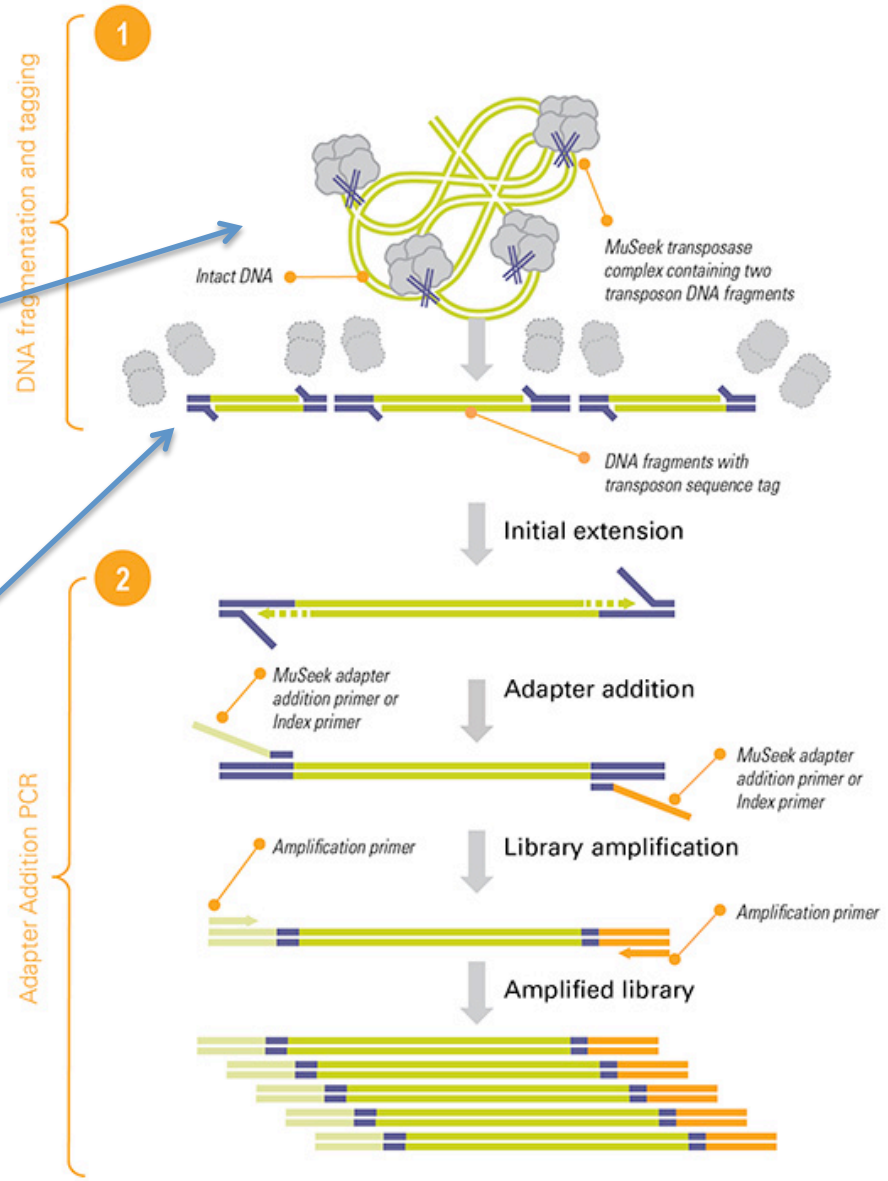
NGS library



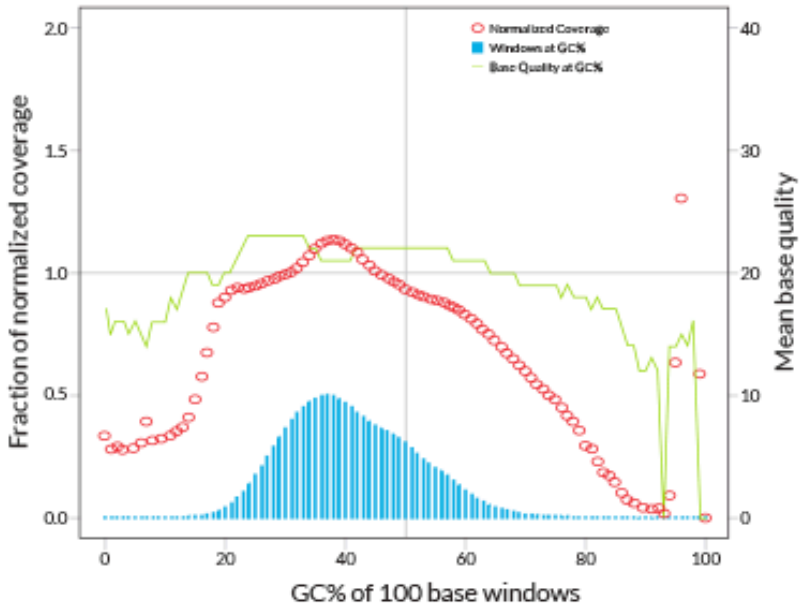
DNA QC – **paramount importance**



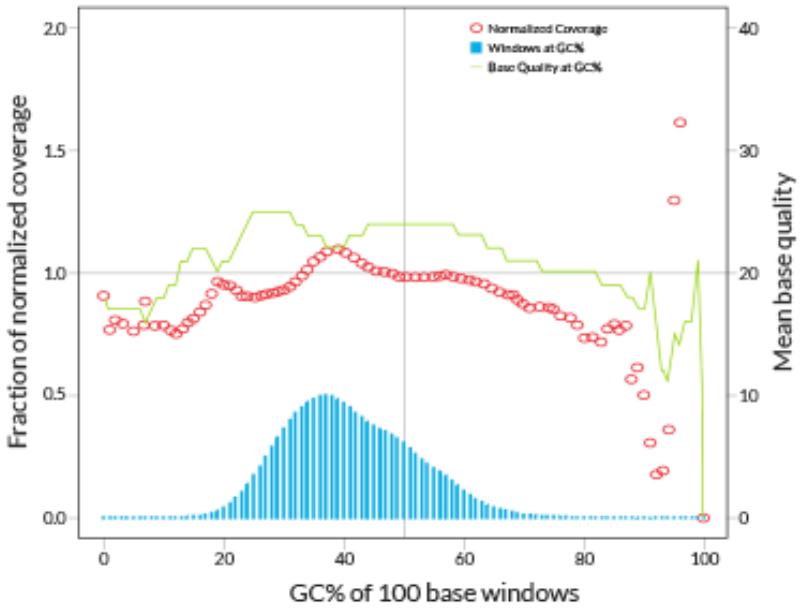
Sharing & size selection



Library complexity



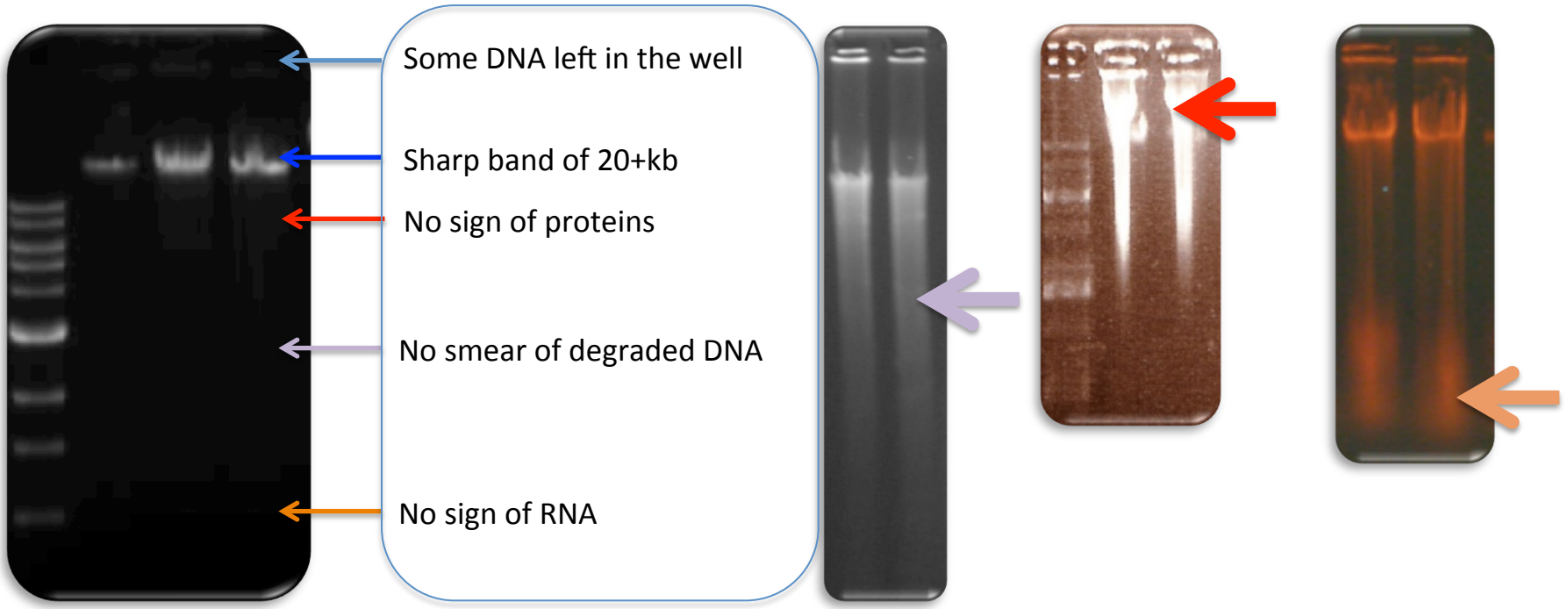
Suboptimal sample



Good sample

(source: https://www.kapabiosystems.com)

DNA quality requirements



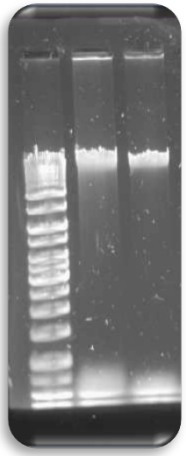
NanoDrop:

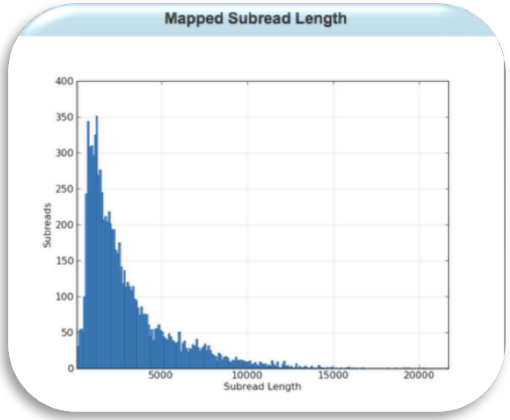
260/280 = 1.8 – 2.0
260/230 = 2.0 – 2.2

Qubit or Picogreen:

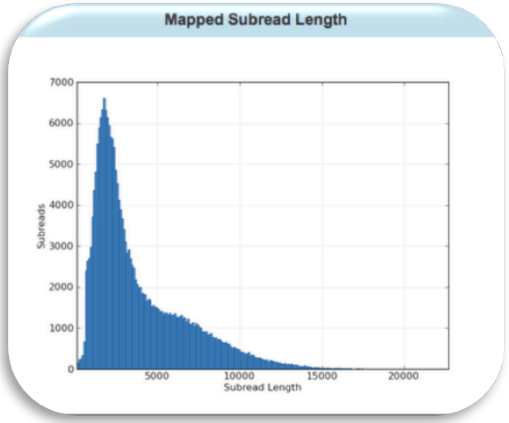
10 kb insert libraries: 3-5 ug
20 kb insert libraries: 10-20 ug

Example:





Polished Contigs	223	Max Contig Length	36,298
N50 Contig Length	2,932	Sum of Contig Lengths	480,087



Polished Contigs	9	Max Contig Length	1,508,929
N50 Contig Length	1,353,702	Sum of Contig Lengths	7,813,244

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:

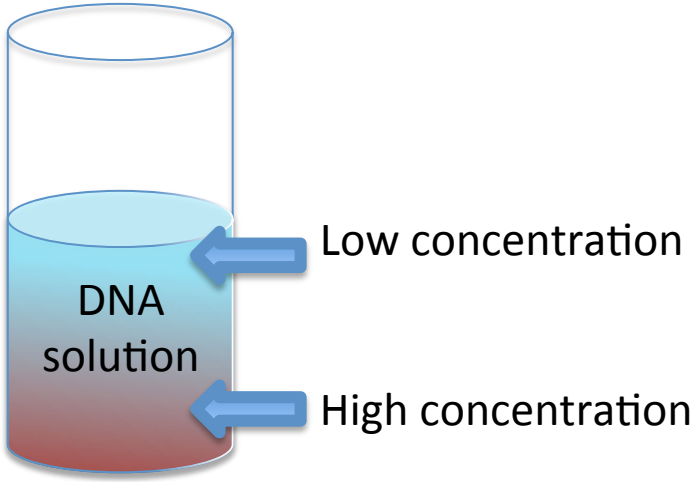
Salt contamination, humic acids, peptides, aromatic compounds, polyphenols, urea, guanidine, thiocyanates (latter three are common kit components) – **absorb at 230 nm**.

>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

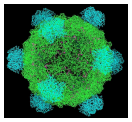


- Thaw DNA completely
- Mix gently (**never vortex!**)
- Put the sample on a thermoblock: 37°C, 15-30 min
- Mix gently
- **Dilute 1:100** (if HMW)
- Mix gently
- Make a measurement with an appropriate blank

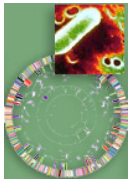
- **NANODROP is Bad**. Point.
- Use Qubit, or PicoGreen.

Let’s get philosophical

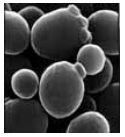
Since the beginning of Genomics:



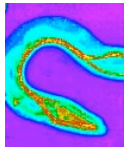
First genome: virus ϕ X 174 - 5 368 bp (1977)



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

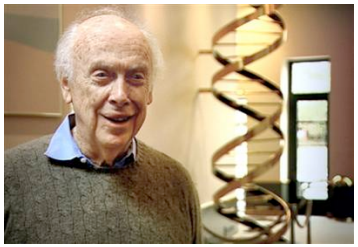
Human genome sequencing:

2004: Genome of Craig Wenter costs 70 mln $

- Sanger's sequencing

2007: Genome of James Watson costs 2 mln $

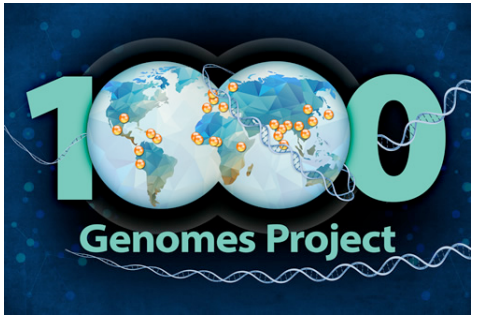
- 454 pyrosequencing

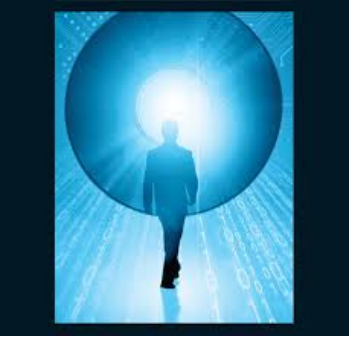


2014: Ultimate goal: 1000 $ / individual

2016: Illumina Xten: Almost there! (1200 $)

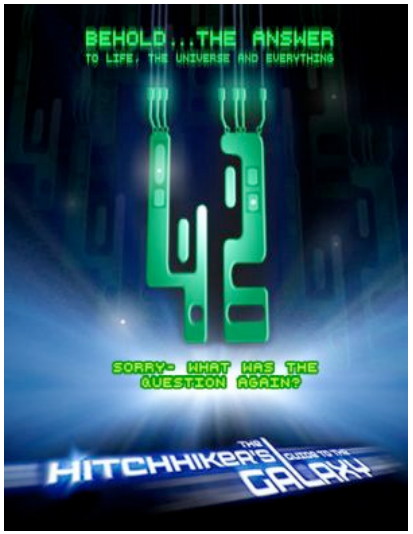
2017: NovaSeq: "Hold my beer…" (100 $)





… paradigm changes

- 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



… 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

IF 2.9

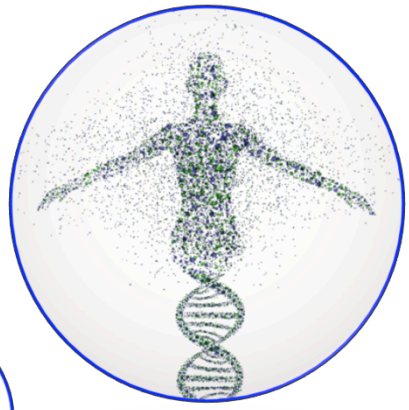
doi:10.1016/j.jbiotec.2010.12.018 | How to Cite or Link Using DOI
Permissions & Reprints

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

Susanne Schneiker-Bekel[a], Daniel Wibberg[a], Thomas Bekel[b], Jochen Blom[b], Burkhard Linke[b], Heiko Neuweger[b], Michael Stiens[a, c], Frank-Jörg Vorhölter[a], Stefan Weidner[a], Alexander Goesmann[b], Alfred Pühler[a] and Andreas Schlüter[a,]

2025 projection: data storage needs

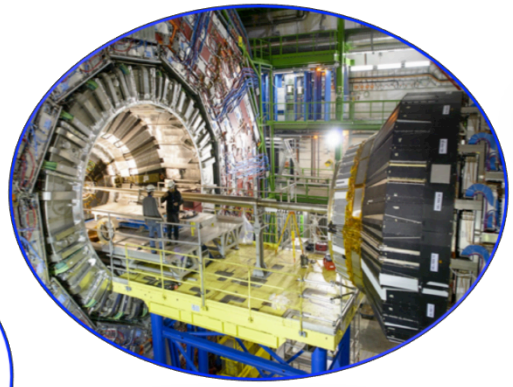
1 petabyte = 10^{15} bytes
1 exabyte = 10^{18} bytes



2-40 exabytes/year

1-2 exabytes/year

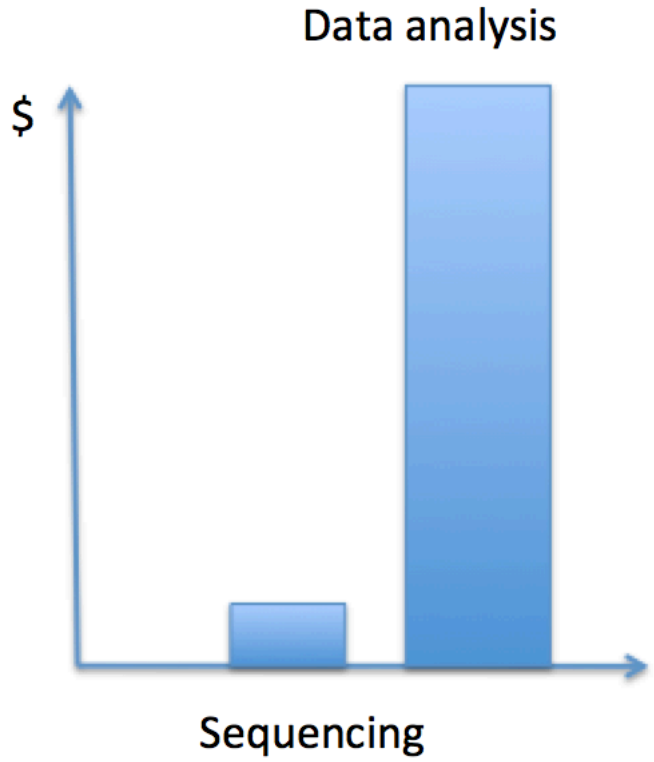
1 exabyte/year

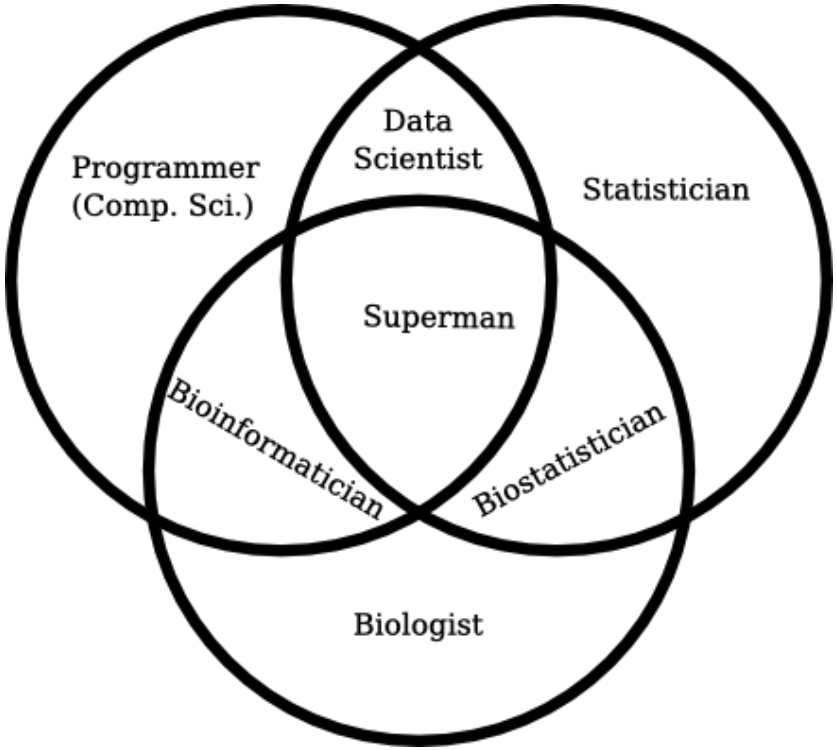


Large Hadron Collider
42 petabytes/year

1-17 petabytes/year

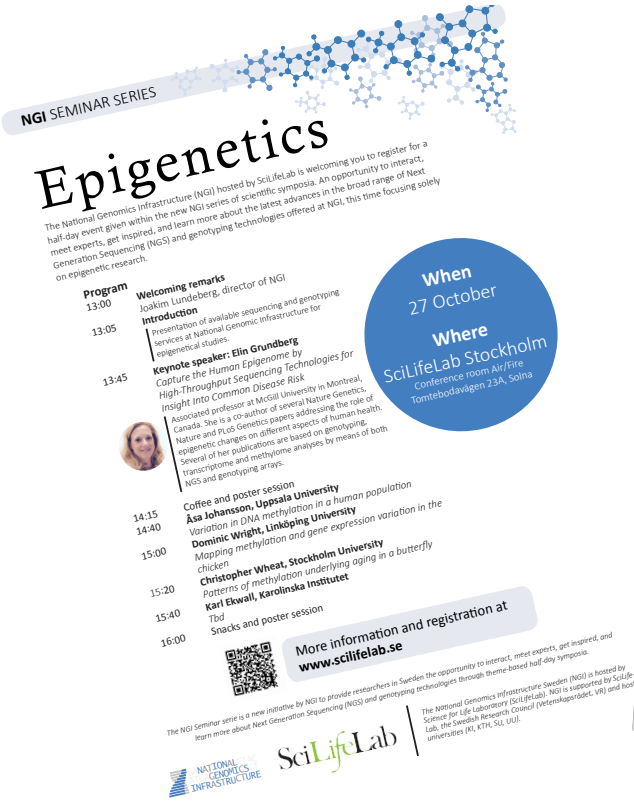
Now, analysis

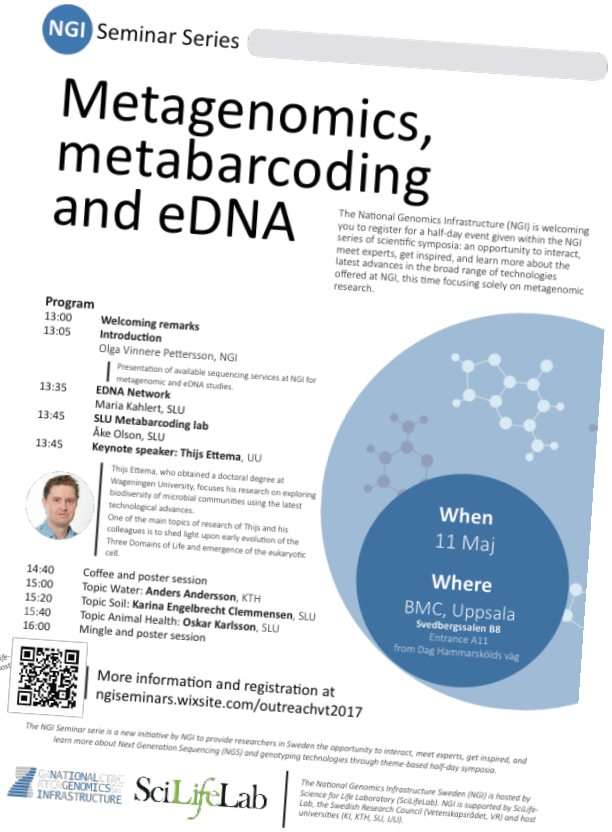


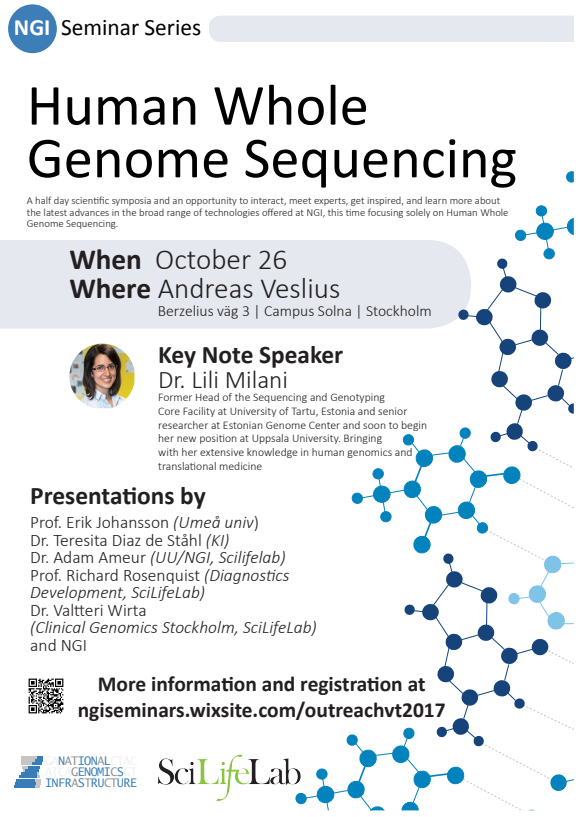


Severe shortage of bioinformaticians

NGI Seminar Series





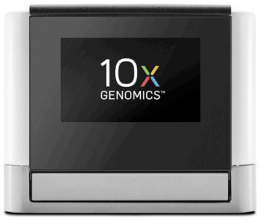


Spring 2018: **De novo sequencing**

Long-Read Sequencing Workshop

BMC Uppsala

Dec 6-7, 2017

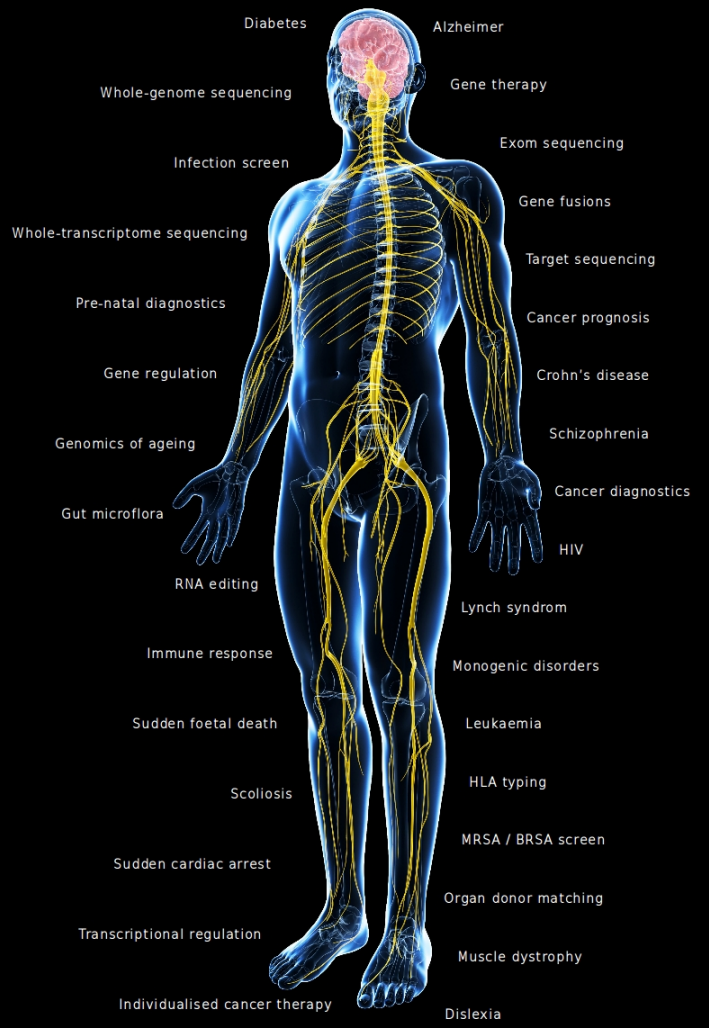


Jason Underwood
Washington University

Robert Sebra
Mount Sinai, NY

Graham Etherington
Earlham Inst. UK

What we sequenced at SciLifeLab



16S amplicon, Acinetobacter baumanii, Acrasis kona, Acridotheres javanicus, Actinobacillus succinogenes, African swine fever virus, Agaricomycotina sp, **Alces alces**, **Alligator mississippiensis**, Amphiura filiformis, Apis mellifera, Aquila chrysaetus, Arabidopsis thaliana, Arabis alpina, Archaeorhizomycetes finlayi, Arctocephalus gazella, artificial sequences, Arvicola amphibius, Ascaridia galli, Aspergillus oryzae, Astrapia stephaniae, Atlantic herring, Atlantic salmon, **Avena sativa**, Baccharis breviseta, Baccharis dracunculifolia, Bacteriophages, Balaenoptera musculus, Balaenoptera physalus, Balanus improvisus, Baltic Sea microorganisms, Bathynomus sp, Bifidobactrium sp, **Borrelia burgdorferi**, Borrelia garinii, Bos taurus, Bovine viral diarrhea virus, Brachyspira suanatina, Brassica sp, Brettanomyces naardensis, Caenorhabditis elegans, Callosobruchus maculatus, Candida intermedia, Candida parapsilosis, Candidatus Neoehrlichia mikurensis, **Canis lupus**, Capreolus capreolus, Capsella bursa-pastoris, Capsella grandiflora, Capsella orientalis, Capsella rubella, Ceanothus thyrsiflorus, Cervus dama, Cervus elaphus, Childla submaculatum, Clonostachys rosea, Clostridium ultunense, Coelodonta antiquitatis, Colias crocea, Collinsia heterophylla, Coregonus lavaretus, Coronavirus, Corvus corone, Corvus monedula, Crassostrea gigas, Cricetulus griseus, Cryptococcus tephrensis, Cubanola dominquensis, Cytomegalovirus, Danio rerio, Datisca glomerata, Deformed wing virus, Dekkera bruxellensis, Dicerorhinus sumatrensis, Dictyostelium discoideum, Diopisthoporus gymnopharyngeus, Diopistoporus longitubus, Drosophila melanogaster, Drosophila paulistorum, Electrophorus electricus, Enterobacter cloacae, Enterococcus faecium, **Equus caballus**, Escherichia coli, Eumecynostomum macrobursalium, Euphorbia lathyris, Euphorbia peplus, Euplectes afer, Euplectes ardens, Euplectes aureus, Euplectes hordeaceus, Euplectes macrourus, Euplectes orix, **Felis catus**, Ficedula albicollis, Ficedula hypoleuca, **Fragaria ananassa**, Freshwater microbial communitites, Fucus radicans, Fucus vesiculosus, Fumaria sp, Galerucella, Gallus gallus, Geospiza magnirostris, Giardia muris, Globodera rostochiensis, Gnetum gnemon, Gnetum luofuense, Gnetum montanum, Gnetum parvifolium, Gnetum pendulum, Gonyostomum semen, Gonzalagunia, Gut microbiota, Hamelia marantha, Heterobasidion abietinum, Heterobasidion annosum, **Hippophae rhamnoides**, **Homo sapiens**, Human Immunodeficiency Virus, Huperzia selago, Hymenoscyphus albidus, Hymenoscyphus pseudoalbidus, Idotea baltica, **Influenza A virus**, Klebsiella pneumoniae, Laccaria bicolor, Lactobacillus, Lepidium campestre, **Leptidea sinapis**, Letharia rugosa, Letharia vulpina, Littorina saxatilis, Lycocorax pyrrhopterus, **Lynx lynx**, Malassezia sympodialis, **Malus domestica**, Malus sylvestris, **Mammuthus primigenius**, Marchantia polymopha, Marine bacteria whole community, **Meligethes aeneus**, Metagenomes, Methanoculleus sp, Metschnikowia andauensis, Metschnikowia hawaiiensis, Metschnikowia lopburiensis, Metschnikowia pulcherrima, Metschnikowia saccharicola, Mixornis gularis, Moorella thermoacetica, Mus musculus, Mycobacterium malmoense, Mycobacterium marinum, **Mytilus edulis**, Nemertoderma westbladi, Nesophontes sp, Neurospora crassa, Neurospora hispaniola, Neurospora intermedia, Neurospora metzenbergii, Neurospora perkinsii, Neurospora sitophila, Neurospora tetrasperma, Nora Virus, Nothoprocta ornata, Nothoprocta perdicaria, **Notophthalmus viridescens**, Nyctereutes procyonoides, Ogataea pini, **Oryctolagus cuniculus**, Rana arvalis, Oryzias latipes, **Pacifastacus leniusculus**, Paenibacillus polymyxa, **Panthera leo**, Panthera pardus, Paradisaea rubra, Parus major, Passer montanus, Paxillus involutus, Penicillium sp, **Perca fluviatilis**, Peridinium aciculiferum, Philomachus pugnax, Phoca sibirica, Phylloscopus collybita, Phylloscopus trochilus, Wolbachia persica, Physcomitrella patens, Phytophthora infestans, **Picea abies**, Pieris napi, Pieris rapae, Pinus pinaster, Pinus sylvestris, **Pisum sativum**, Planctomycetes sp, Plasmodiophora brassicae, Plasmodium falciparum, Podospora anserina, Polystachya paniculata, Pomatoschistus minutus, Populus maximowiczii, Populus tremula, Populus trichocarpa, Posoqueria sp, Pseudomonas aeruginosa, Pseudomonas brassicacearum, Pseudomonas chlororaphis, Pseudomonas putida, Pteridophora alberti, Ptiloris paradiseus, Puccinia striiformis, Pythium oligandrum, Quelea 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, Scrippsiella hangoei, **Semibalanus balanoides**, Setaria digitata, Silene conoidea, Silene latifolia, Silene viscaria, Sindbis virus, Siphocampylus, Siphocampylus retrorsus, Siphoviridae phage, Skeletonema marinoi, **Solanum tuberosum**, Sorghum sp, Spironucleus barkahanus, Spironucleus salmonicida, Spironucleus vortens, Staphylococcus aureus, Staphylococcus pseudintermedius, Stemmadenia sp, Streptococcus pneumoniae, Streptococcus pyogenes, Streptomyces coelicolor, Struthio camelus, Sulfolobus acidocaldarius, **Sus scrofa**, Synthetic DNA, Syntrophaceticus schinkii, Taphrina betulina, Tepidanarobacter aceatoxydans, Thamnolia vermicularis, Theileria parva, Trypanosoma cruzi, Trypanosoma rangeli, **Ursus spelaeus**, Vitex agnus-castus, Yarrowia lipolytica, Zalophus californianus, Zalophus wollebaeki, Zygopetalum 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

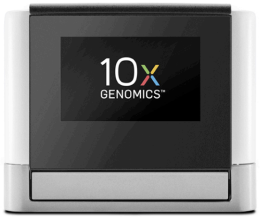
Drug Discovery and Development
ADME (Absorption Distribution, Metabolism Excretion) of Therapeutics (UDOPP)
Biochemical and Cellular Screening
Biophysical Screening and Characterization
Human Antibody Therapeutics
In Vitro and Systems Pharmacology
Medicinal Chemistry – Hit2Lead
Medicinal Chemistry –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

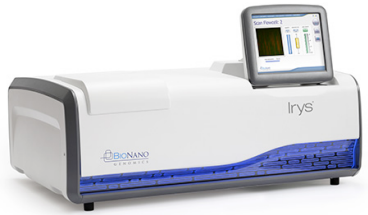
Functional Genomics
Karolinska High Throughput Center (KHTC)

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





NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab

Operational principles of NGI

User community

- Open to all Swedish academic scientists on equal terms.
- Consultation and introduction of new protocols.
- Workshops, courses, etc.

Cost basis

- Academic users of NGI only cover reagent cost.
- Staff salaries at NGI covered by SciLifeLab, VR, and host universities.
- Premises and service contracts covered by SciLifeLab, VR, KAW and host universities.
- Capital equipment covered by KAW, VR, SciLifeLab.

Quality

- Emphasis on data quality and needs of the users.
- Illumina sequencing and genotyping processes accredited by SWEDAC, ISO/IEC 17025
- Ion and PacBio: accreditation during 2017

We are non-profit

We have technology and knowledge

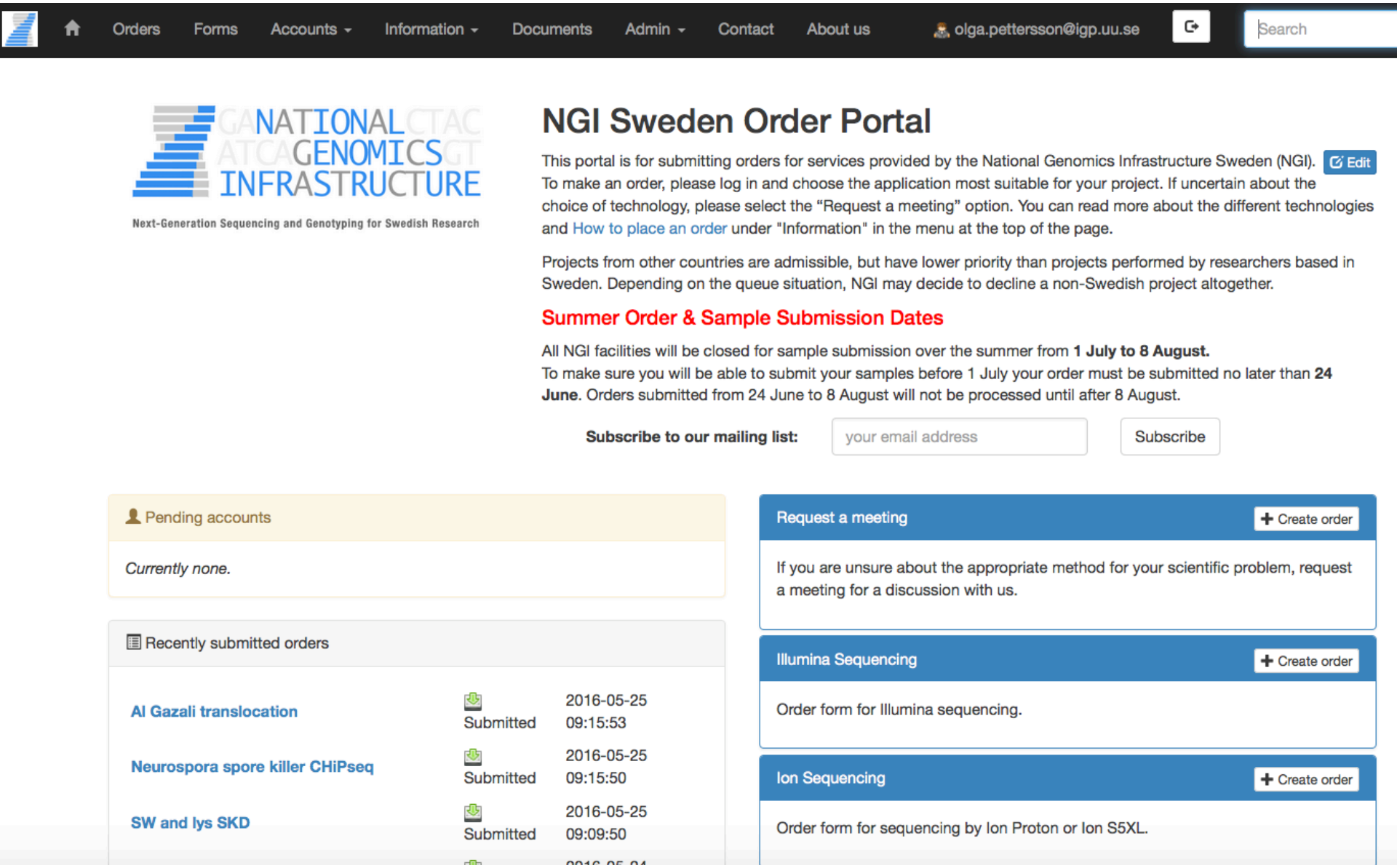
We want to help you to do GREAT research

We do not want co-authorship

Let us help YOU

NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab



https://ngisweden.scilifelab.se/

Contact NGI

Place an order or request a meeting:

https://ngisweden.scilifelab.se/

NGI Stockholm
Illumina



Email: support@ngisweden.se.

Project Coordinators:
Mattias Ormestad
Beata Werne Solnestam
Karin Gillner

NGI Uppsala
Illumina


Email: seq@medsci.uu.se

Project Coordinators:
Ellenor Devine
Johanna Lagensjö

NGI Uppsala
PacBio, Ion

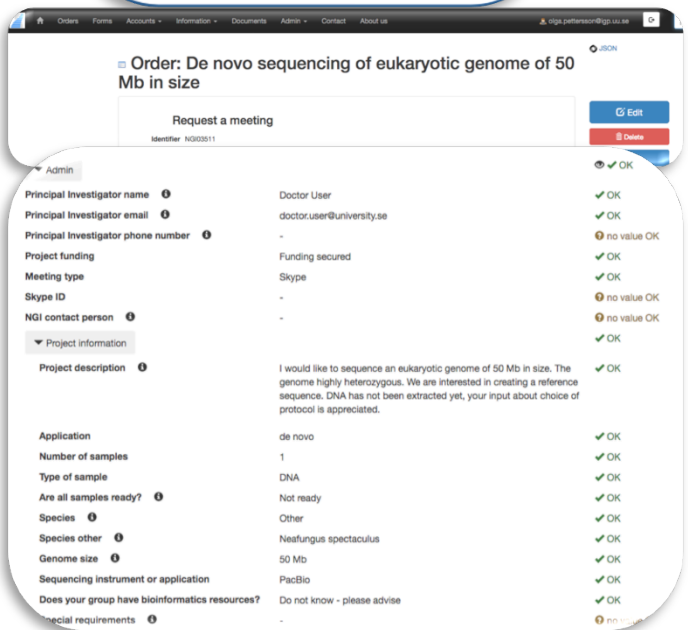

Email:
uppsala_orders@ngisweden.zendesk.com.

Project Coordinators:
Olga Vinnere Pettersson
Susana Häggqvist

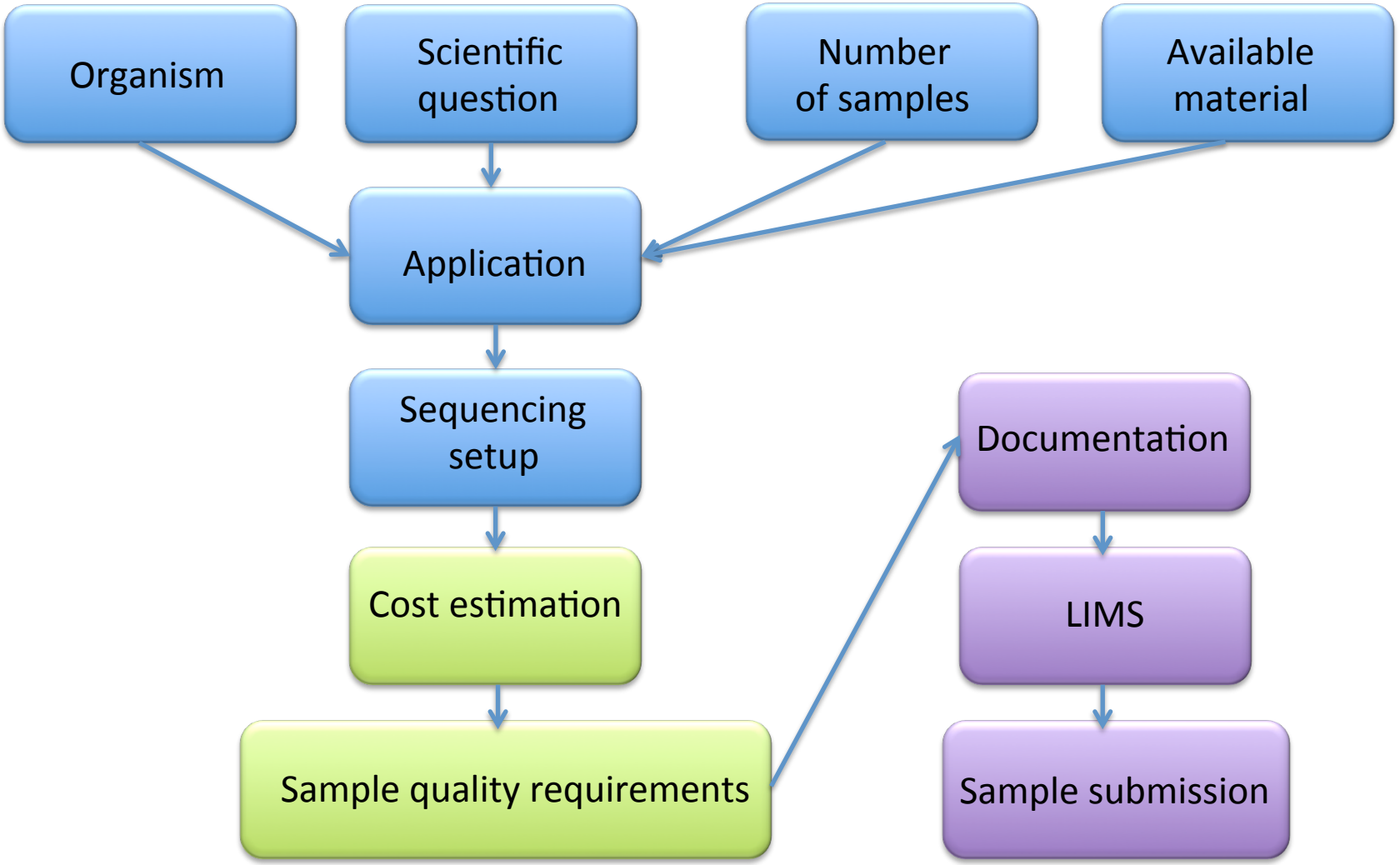
NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab

How does a project go? Project request



Before project starts



Ongoing project

Quality control	Library prep	Quality control	Sequencing	Quality control	Data analysis
Qubit BioAnalyser Fragment Analyser DropSense Pippin Pulse	Shearing Size selection DNA repair SMRT bell construction		Real-time monitoring of run progress		Loading statistics Read length Raw data processing Secondary analysis (PacMan)


Ida Hoijer


Susana Häggqvist


Anna Petri


Nina Williams


Magdalena Andersson


Cecilia Lindau


Sara Olofsson

Thanks for listening! Questions?

support@ngisweden.se

NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab

QUESTIONS?

Pricing

Illumina MiSeq

Ion S5XL

PacBio RSII

Instrument/seq unit	Read length, bp	Mln reads / unit	Library cost, SEK	Sequencing cost, SEK
Illumina MiSeq, Flow cell (FC)	300+300	18	1100	16 000
Illumina HiSeq, Rapid run (FC)	250+250	220	1100	60 000
Ion S5XL				
chip 520	200 – 400 – 600	3	1100	6 500
chip 530	200 – 400 – 600	18	1100	7 300
chip 540	200	80	1100	7 900
PacBio RSII, SMRT cell	250 – 13 000	0,5	1800	3 000

NATIONAL GENOMICS INFRASTRUCTURE

SciLifeLab