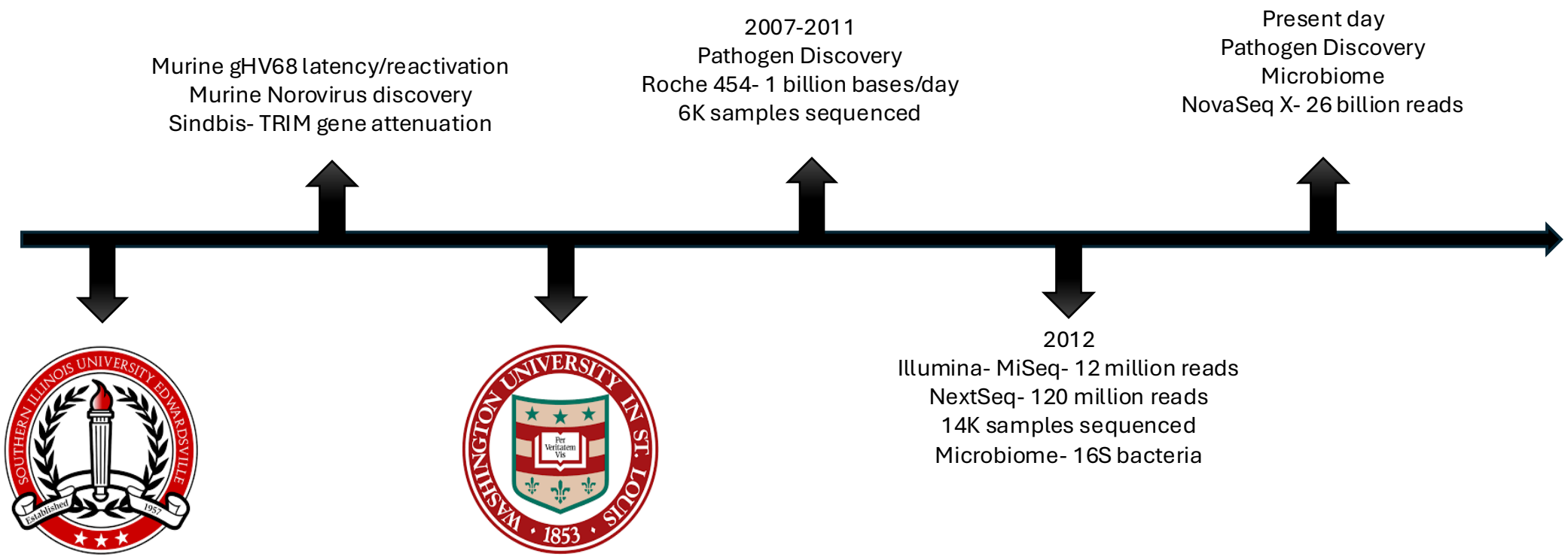
Targeted Virus Genomics Sample Preparation

Lindsay Droit

Washington University School of Medicine

St. Louis, Missouri

Career path


Murine gHV68 latency/reactivation
Murine Norovirus discovery
Sindbis- TRIM gene attenuation
2007-2011
Pathogen Discovery
Roche 454- 1 billion bases/day
6K samples sequenced
Present day
Pathogen Discovery
Microbiome
NovaSeq X- 26 billion reads
2012
Illumina- MiSeq- 12 million reads
NextSeq- 120 million reads
14K samples sequenced
Microbiome- 16S bacteria
SOUTHERN ILLINOIS UNIVERSITY EDWARDSVILLE
Established
1957
WASHINGTON UNIVERSITY IN ST. LOUIS
1853
Per Veritatem Vis

Research interests

Women's health in developing countries

IMBOKODO

FRESH

Females Rising through Education, Support & Health

Pathogen Discovery

Marsabit, Kenya

Capacity building and training/teaching

Durban, South Africa

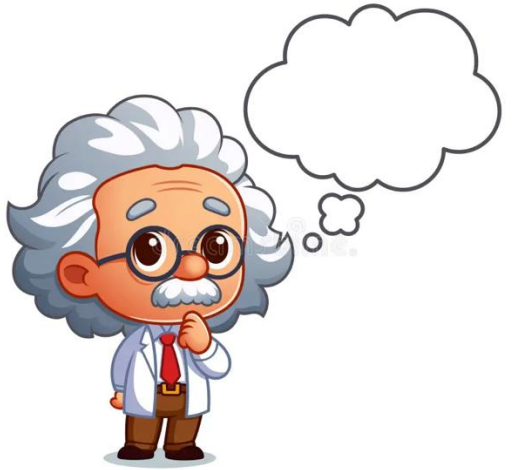
Kathmandu, Nepal

Addis Ababa, Ethiopia

Source: HVTN

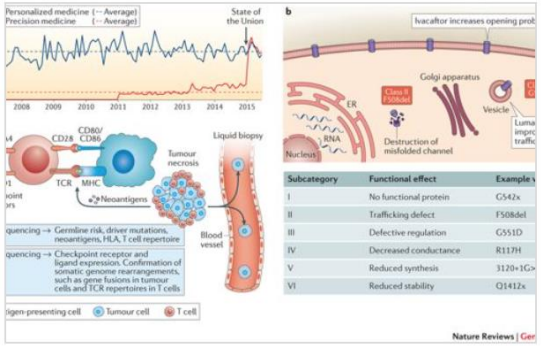
Questions to ask

- Hypothesis
- Resources
  - What is the project budget
  - Is there access to cohort/samples
  - Who will collect samples and how will they be transported/stored
- Sample requirements
  - How many samples should the study include
  - How much specimen is available for each sample
  - Does my experiment require replicates
  - Statistics- How many reads per sample are needed for analysis
- Equipment and sequencing access
  - Is there access to equipment needed to perform experiments
  - What sequencing platforms are available
  - What sequencing reagents are available



Next Gen Sequencing applications

Nature Reviews Genetics | Review Article

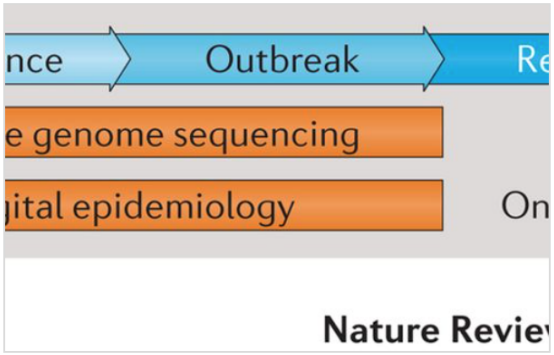


Towards precision medicine

Precision medicine is a strategy for tailoring clinical decision making to the underlying genetic causes of disease. This Review describes how, despite the straightforward overall principles of precision medicine, adopting it responsibly into clinical practice will require many technical and conceptual hurdles to be overcome. Such challenges include optimized sequencing strategies, clinically focused bioinformatics pipelines and reliable metrics for the disease causality of genetic variants. show less

Euan A. Ashley

Nature Reviews Genetics | Review Article

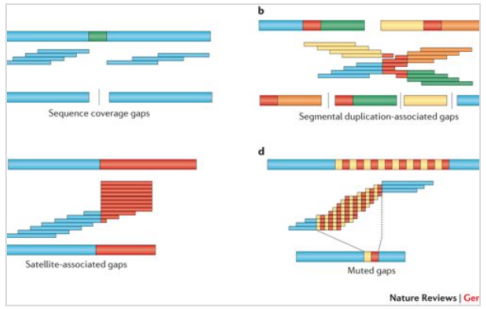


Towards a genomics-informed, real-time, global pathogen surveillance system

Next-generation sequencing has the potential to support public health surveillance systems to improve the early detection of emerging infectious diseases. This Review delineates the role of genomics in rapid outbreak response and the challenges that need to be tackled for genomics-informed pathogen surveillance to become a global reality. show less

Jennifer L. Gardy & Nicholas J. Loman

Nature Reviews Genetics | Review Article

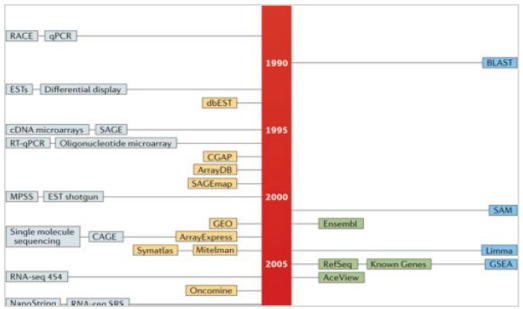


Genetic variation and the *de novo* assembly of human genomes

The wealth of existing and emerging DNA-sequencing data provides an opportunity for a comprehensive understanding of human genetic variation, including the discovery of disease-causing variants. This Review describes how the limitations of current reference-genome assemblies confound the characterization of genetic variation and how this can be mitigated by important advances in algorithms and sequencing technology that facilitate the *de novo* assembly of genomes. show less

Mark J. P. Chaisson, Richard K. Wilson & Evan E. Eichler

Nature Reviews Genetics | Review Article



Cancer transcriptome profiling at the juncture of clinical translation

Although cancer genome sequencing is becoming routine in cancer research, cancer transcriptome profiling through methods such as RNA sequencing (RNA-seq) provides information not only on mutations but also on their functional cellular consequences. This Review discusses how technical and analytical advances in cancer transcriptomics have provided various clinically valuable insights into gene expression signatures, driver gene prioritization, cancer microenvironments, immuno-oncology and prognostic biomarkers. show less

Marcin Cieślik & Arul M. Chinnaiyan

Virome sequencing applications

Article

Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease

Jason M. Norman,[1,10] Scott A. Handley,[1,10] Megan T. Baldridge,[1] Lindsay Droit,[1] Catherine Y. Liu,[1] Brian C. Keller,[1,2] Amal Kambal,[1] Cynthia L. Monaco,[1,2] Guoyan Zhao,[1,3] Phillip Fleshner,[4] Thaddeus S. Stappenbeck,[1] Dermot P.B. McGovern,[5] Ali Keshavarzian,[6] Ece A. Mutlu,[6] Jenny Sauk,[7] Dirk Gevers,[8] Ramnik J. Xavier,[7,8] David Wang,[1,3] Miles Parkes,[9] and Herbert W. Virgin[1,*]

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA
[2]Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA
[3]Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110, USA
[4]Division of Colorectal Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[5]The F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute; Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[6]Department of Medicine, Division of Digestive Diseases and Nutrition, Rush University Medical Center, Chicago, IL 60612, USA
[7]Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[8]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[9]Division of Gastroenterology Addenbrooke's Hospital and Department of Medicine, University of Cambridge, Cambridge CB2 0QQ, UK
[10]Co-first author
*Correspondence: virgin@wustl.edu
http://dx.doi.org/10.1016/j.cell.2015.01.002

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES
August 2021 Volume 10 Issue 31 10.1128/mra.00699-21
https://doi.org/10.1128/mra.00699-21

Divergent Enteroviruses from Macaques with Chronic Diarrhea

Kathie A. Mihindukulasuriya[a,b], Lindsay Droit[a,b], Margaret H. Gilbert[c], Peter J. Didier[d,e], Anne Paredes[a,b], Scott A. Handley[a,b], Rudolf P. Bohm[c], David Wang[a,b]

[a]Department of Pathology, Washington University School of Medicine, St. Louis, Missouri, USA
[b]Department of Immunology, Washington University School of Medicine, St. Louis, Missouri, USA
[c]Tulane University Office of Research, Institutional Animal Care and Use Committee, Covington, Louisiana, USA
[d]Division of Veterinary Medicine, Tulane National Primate Research Center, Covington, Louisiana, USA
[e]Division of Comparative Pathology, Tulane National Primate Research Center, Covington, Louisiana, USA

ELSEVIER

Virology
Volume 582, May 2023, Pages 83-89

Isolation of a rhesus calicivirus that can replicate in human cells

Tianyu Gan [a], Lindsay Droit [b], Susan Vernon [a], Dan H. Barouch [c], David Wang [a b]

[a] Department of Molecular Microbiology, School of Medicine, Washington University in St. Louis, St. Louis, MO, 63110, USA
[b] Department of Pathology & Immunology, School of Medicine, Washington University in St. Louis, St. Louis, MO, 63110, USA
[c] Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA, 02115, USA

RESEARCH ARTICLE | BIOLOGICAL SCIENCES

Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children

Guoyan Zhao, Tommi Vatanen, Lindsay Droit, Arnold Park, Aleksandar D. Kostic, Tiffany W. Poon, Hera Vlamakis, Heli Siljander, Taina Härkönen, Anu-Maaria Hämäläinen, Aleksandr Peet, Vallo Tillmann, Jorma Ilonen, David Wang, Mikael Knip, Ramnik J. Xavier, and Herbert W. Virgin -13 Authors Info & Affiliations

Contributed by Herbert W. Virgin, June 7, 2017 (sent for review April 17, 2017; reviewed by Mya Breitbart and Julie A. Segre)

July 10, 2017 | 114 (30) E6166-E6175 | https://doi.org/10.1073/pnas.1706359114

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

Microbiol Resour Announc. 2024 Oct 21;13(11):e00768-24. doi: 10.1128/mra.00768-24

Human immunodeficiency virus-1 genome from patient with fever, Nepal

Eans Tara Tuladhar [1,2], Bimal Sharma Chalise [3], Binod Khadka [1], Mamta Tamang [1], Jenish Neupane [3], Shankar Poudel [3], Lindsay Droit [4], Kathie A Mihindukulasuriya [4], Annie Elong Ngono [5], Yuba Nidhi Basaula [3], Sujan Shresta [5], David Wang [4], Krishna Das Manandhar [1,✉]

Cell Host Microbe. 2022 Jan 12;30(1):110–123.e5. doi: 10.1016/j.chom.2021.12.002

Enteric virome negatively affects seroconversion following oral rotavirus vaccination in a longitudinally sampled cohort of Ghanaian infants

Andrew HyoungJin Kim [1,2], George Armah [5], Francis Dennis [5], Leran Wang [2,3], Rachel Rodgers [2,4], Lindsay Droit [3], Megan T Baldridge [1,2,6], Scott A Handley [2,3], Vanessa C Harris [7,8,9,*]

Nat Med. Author manuscript; available in PMC: 2016 Jan 12.

Published in final edited form as: Nat Med. 2015 Sep 14;21(10):1228–1234. doi: 10.1038/nm.3950

Early life dynamics of the human gut virome and bacterial microbiome in infants

Efrem S Lim [1,2], Yanjiao Zhou [3,4], Guoyan Zhao [1], Irma K Bauer [3], Lindsay Droit [1,2], I Malick Ndao [3], Barbara B Warner [3], Phillip I Tarr [1,3], David Wang [1,2], Lori R Holtz [3]

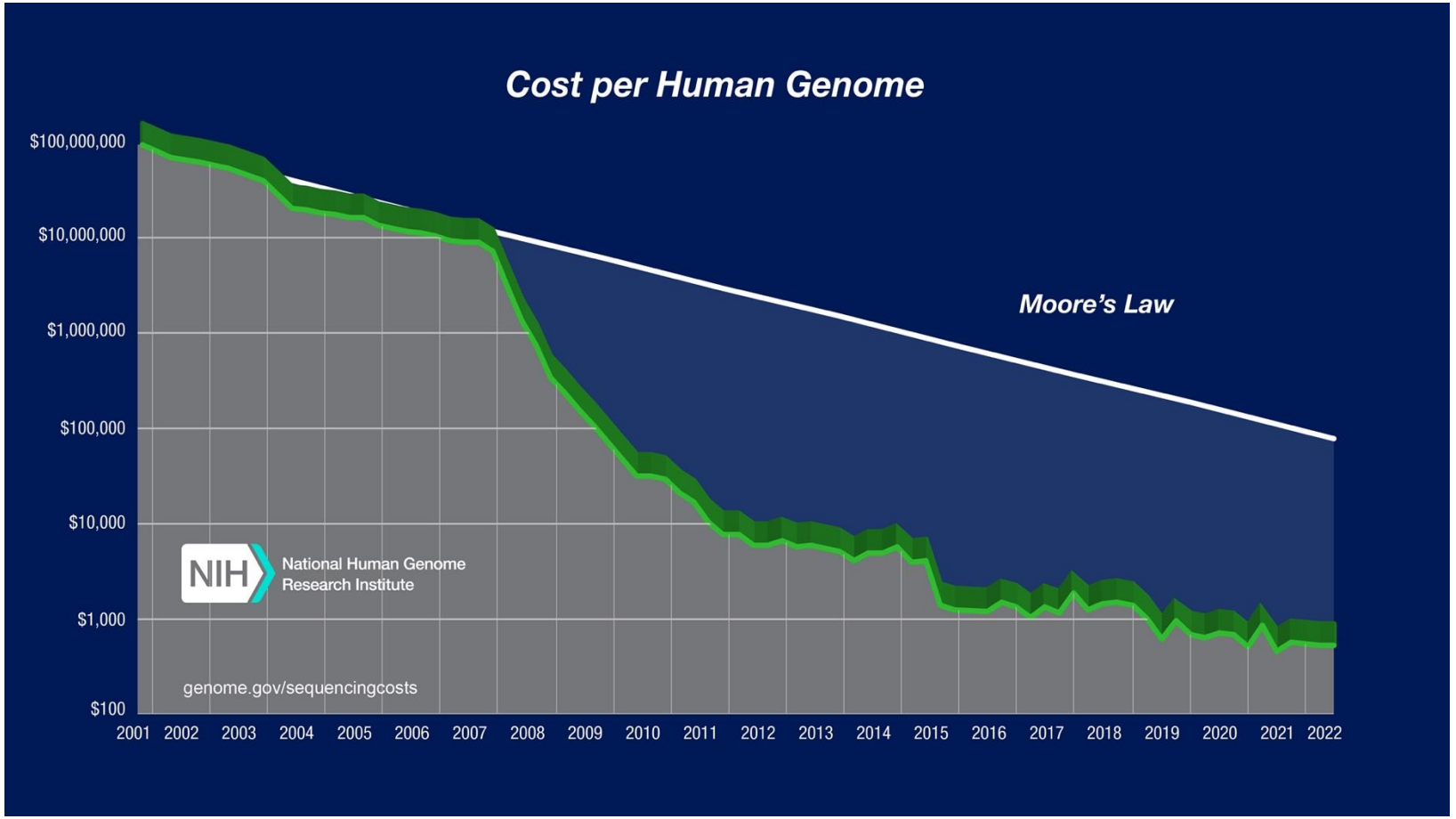
Outline

- Sequencing history ⟵
- Current sequencing technologies
- Preparing samples for **N**ext **G**eneration **S**equencing
  - Techniques for virus like particle enrichment
  - Library preparation

Sequencing history

- First Generation
  - Sanger Sequencing (1977)
    - Foundation for DNA sequencing
    - Short reads
    - High labor and cost
- Second Generation (NGS)
  - 454 Sequencing (2005)
    - First Next Gen Sequencing
    - Massively parallel pyrosequencing technique
  - Illumina (Solexa)Sequencing (2007)
    - Sequencing by synthesis
      - Genome Analyzer 2007-2009
      - HiSeq 2010
      - MiSeq 2011
      - NextSeq 2014
      - MiniSeq 2016
      - NovaSeq 2017
- Third Generation
  - Oxford Nanopore
  - Single molecule
    - MinIon

Sequencing costs



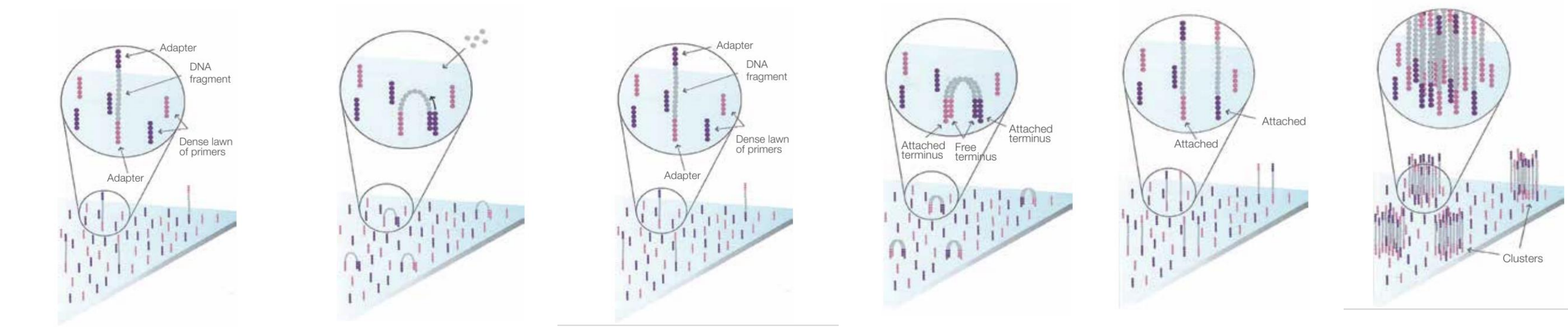
Outline

- Sequencing history
- Current sequencing technologies ⟵
- Preparing samples for **N**ext **G**eneration **S**equencing
  - Techniques for virus like particle enrichment
  - Library preparation

From Sample to Sequencing

1. Nucleic acid extraction (VLP enrichment)
2. Library preparation
3. Sequencing ⟵
   - Illumina
   - Oxford Nanopore
4. Data analysis

Illumina sequencing technology- cluster generation



1.A flow cell is coated with oligonucleotides

2.DNA fragments are added to the flow cell and attach to the oligonucleotides

3.DNA polymerase replicates the DNA fragments

4.The double-stranded DNA fragments are denatured and the original strands are washed away

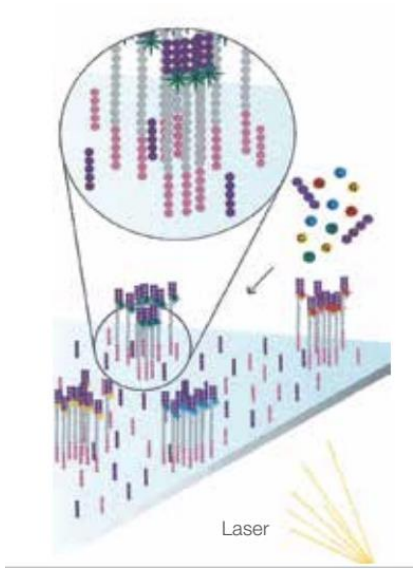
5.The remaining strands fold over and attach to the other oligonucleotides on the flow cell

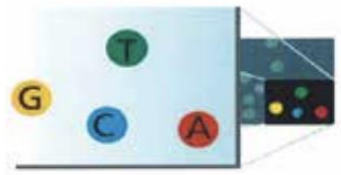
6.DNA polymerase replicates the folded strands, forming a bridge

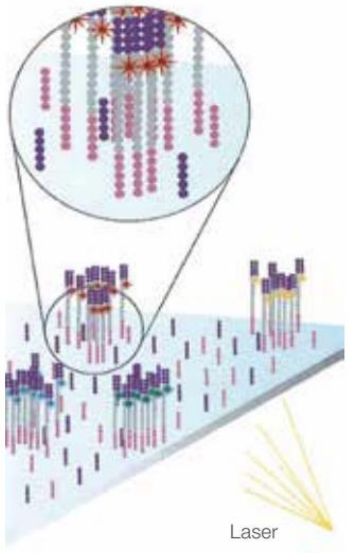
7.The bridge is denatured, creating two single-stranded copies of the DNA

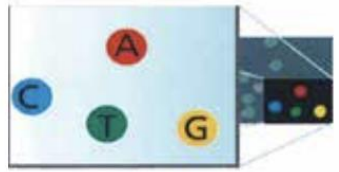
8.Repeat the process until enough DNA has been amplified

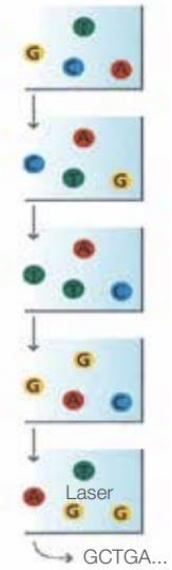
Illumina sequencing technology- sequencing by synthesis











SBS- method for sequencing DNA that detects bases as they are added to a growing DNA strand

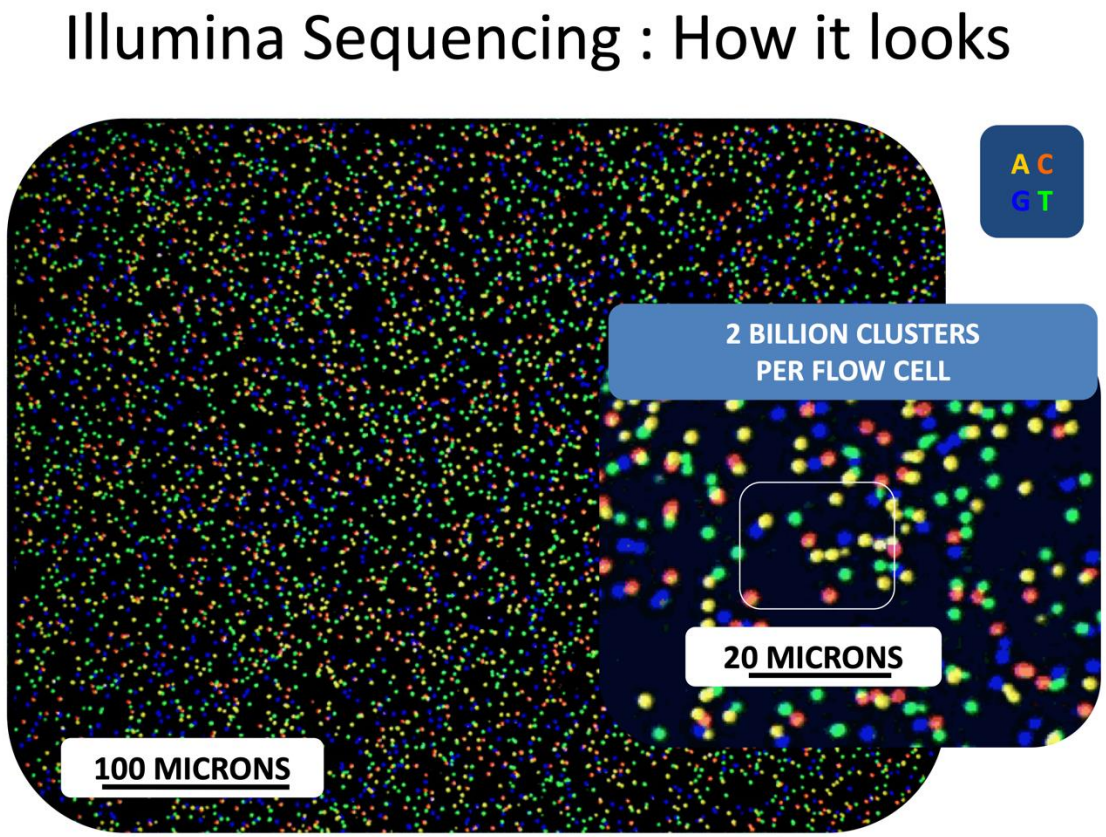
1.A DNA polymerase incorporates nucleotides into a complementary DNA strand.

2.A fluorescently-labeled nucleotide is added to the nucleic acid chain.

3.The fluorescent dye is imaged to identify the base.

4.The dye is enzymatically cleaved to allow the next nucleotide to be incorporated.

Illumina sequencing technology- base calling



2 channel SBS sequencing chemistry:

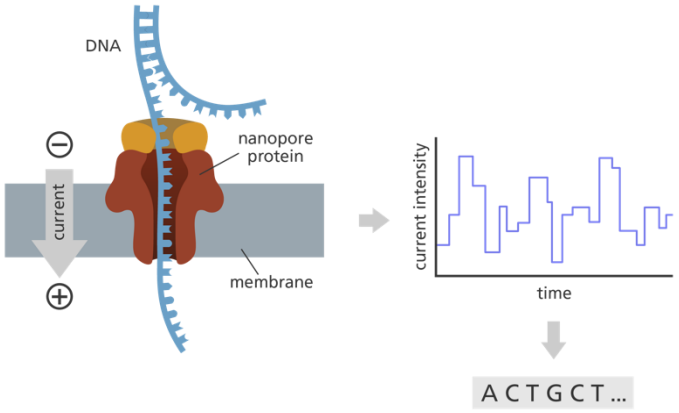
- T- labeled with green fluorophores
- C- labeled with red fluorophores
- A- red and green fluorophores
- G- permanently dark

Illumina sequencing video

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

Oxford Nanopore sequencing technology

- Nanopore sequencing
- Detects changes in electrical current as a single strand of DNA passes through a tiny protein pore “nanopore”
- Each different nucleotide causes a unique current fluctuation
- Advantage- longer reads / Disadvantage- lower accuracy, no redundancy, large amount of high-quality DNA are needed for input



Oxford Nanopore sequencing video

https://www.youtube.com/watch?v=RcP85JHLmnI

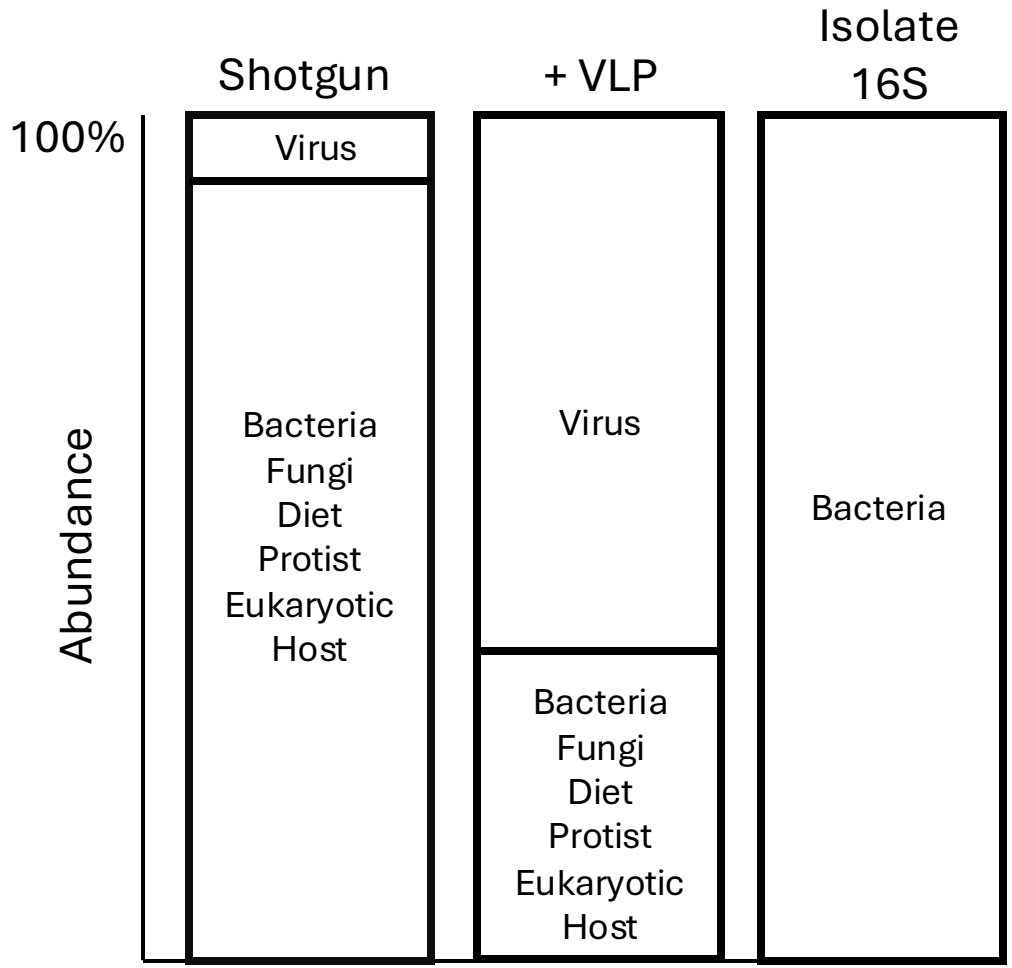
Outline

- Sequencing history
- Current sequencing technologies
- Preparing samples for **N**ext **G**eneration **S**equencing
  - Techniques for virus like particle enrichment
  - Library preparation

From Sample to Sequencing

1. Nucleic acid extraction (VLP enrichment) ⟵
2. Library preparation
3. Sequencing
4. Data analysis

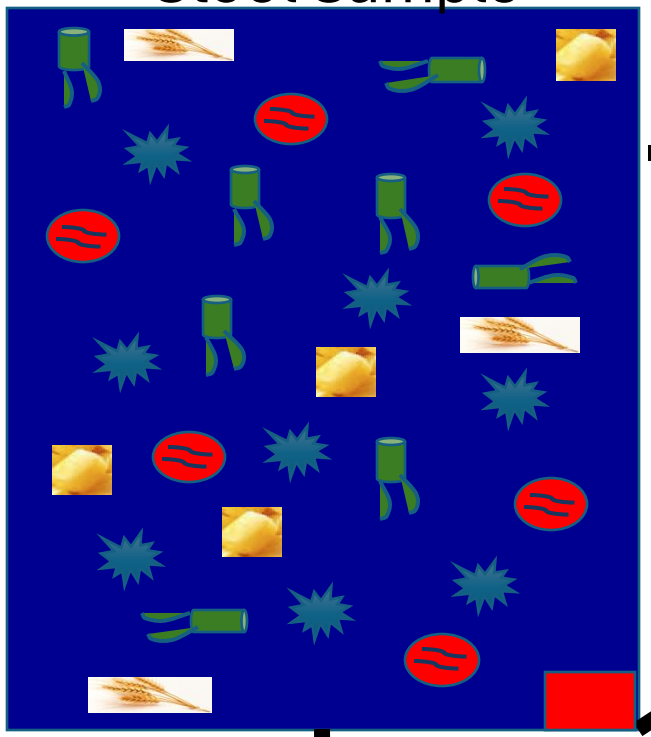
Why VLP enrichment



Abundance- what is your question and how will you maximize the number of sequences you are specifically interested in

VLP enrichment

Stool Sample



Amplicon Sequencing

16S/18S/ITS
-DNA
-targeted amplicons

Enriched Metagenome- Virome

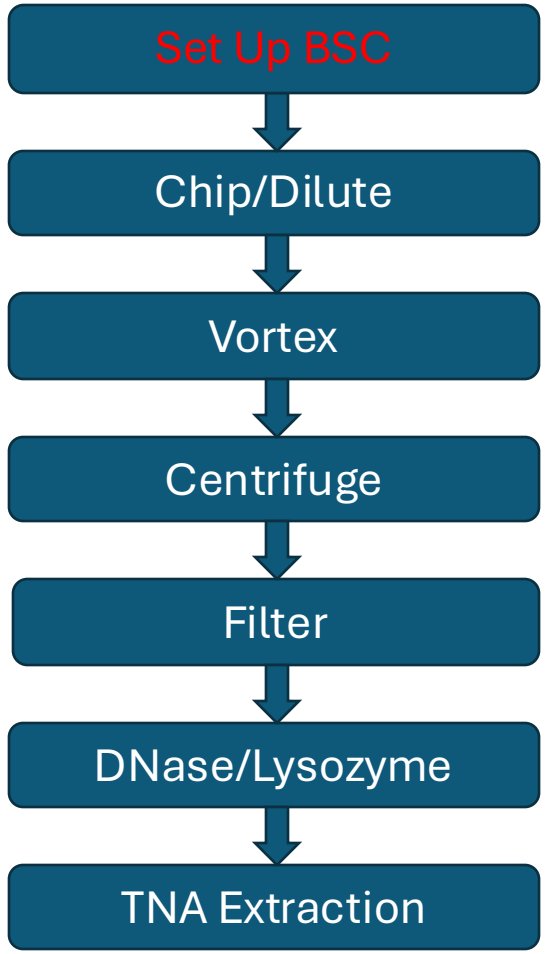
Virus Like Particle (VLP)
-DNA, RNA
-remove bacteria
-remove encapsulated DNA

Shotgun Sequencing

Preparing samples for virome sequencing

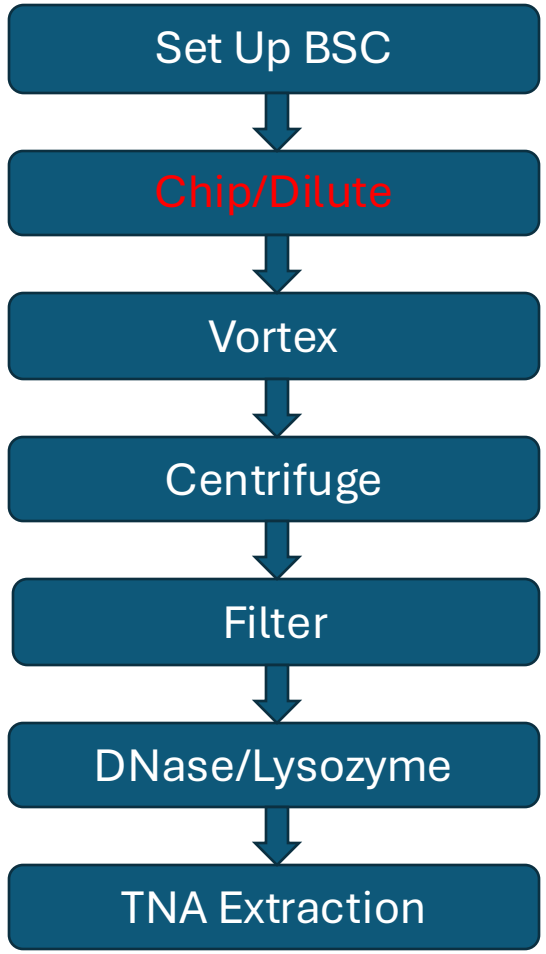
1. Virus Like Particle (VLP) Enrichment and Total Nucleic Acid Extraction
2. Reverse Transcription, Second Strand Synthesis and PCR Amplification
3. Library Construction

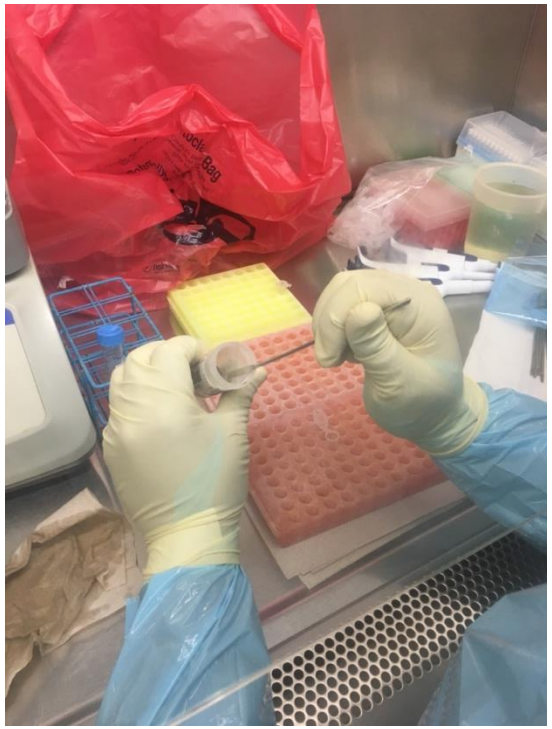
Set up biosafety cabinet



- Samples are handled at BSL2+ - gown and double glove
- Decon with 10% bleach solution, 70% ethanol, and UV for 30 minutes
- Waste collected in biohazard bags and autoclaved

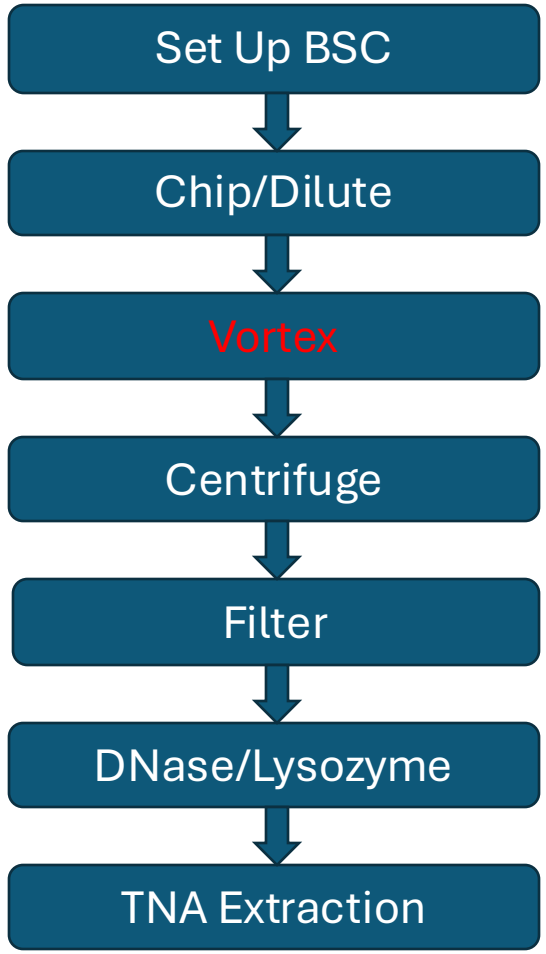
Chip/Dilute to obtain desired input





- Keep sample frozen
- Chip ~200mg of stool
- Add SM Buffer (NaCl, Tris, MgSO4)

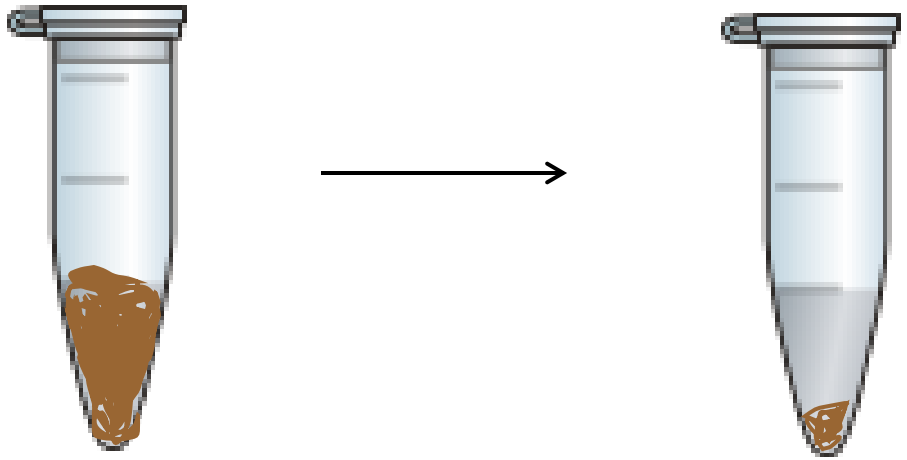
Homogenize to break up stool material



5 minutes

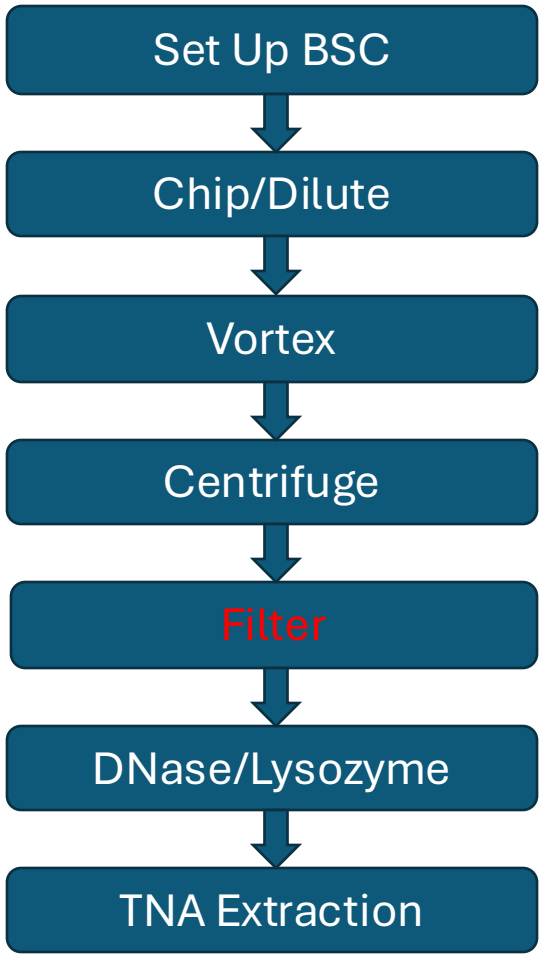
Centrifuge to pellet stool particles

- Set Up BSC
- Chip/Dilute
- Vortex
- Centrifuge
- Filter
- DNase/Lysozyme
- TNA Extraction



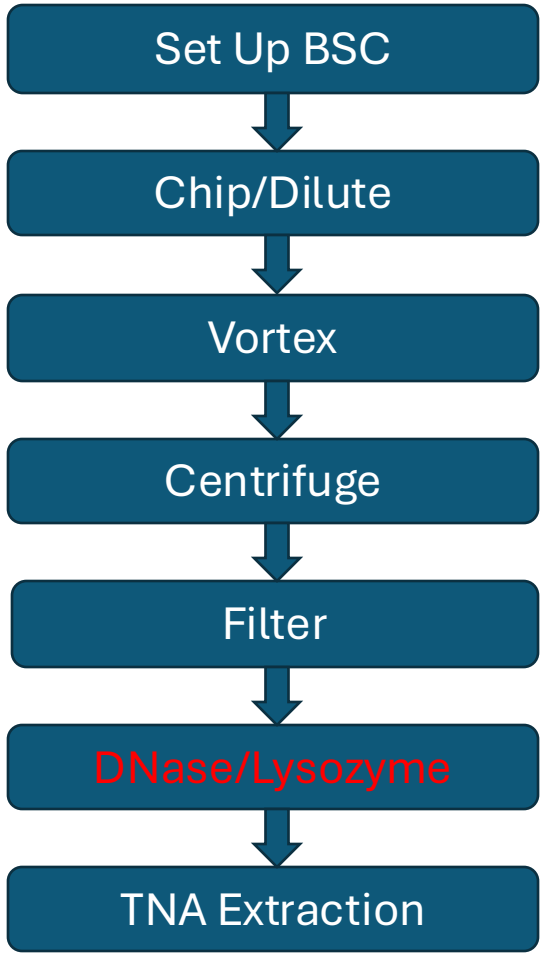
7,000G for 10 minutes

Filter to remove bacteria



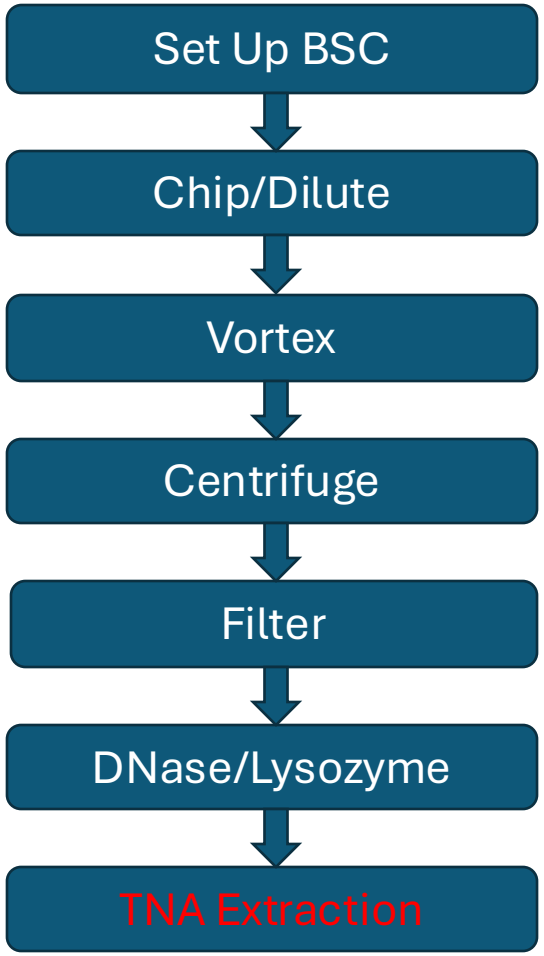
.45u filter

Non encapsulated DNA removal



	Per 800ul sample	12 +1 =13 samples
Turbo DNase buffer	108 ul	1,404
TurboDNaseI (2U/ul)	20 ul	260
Baseline zero (1U/ul)	4 ul	52
Lysozyme (10mg/ml)	80 ul	1,040
H2O	68 ul	884
	280ul	

Total Nucleic Acid extraction



Roche

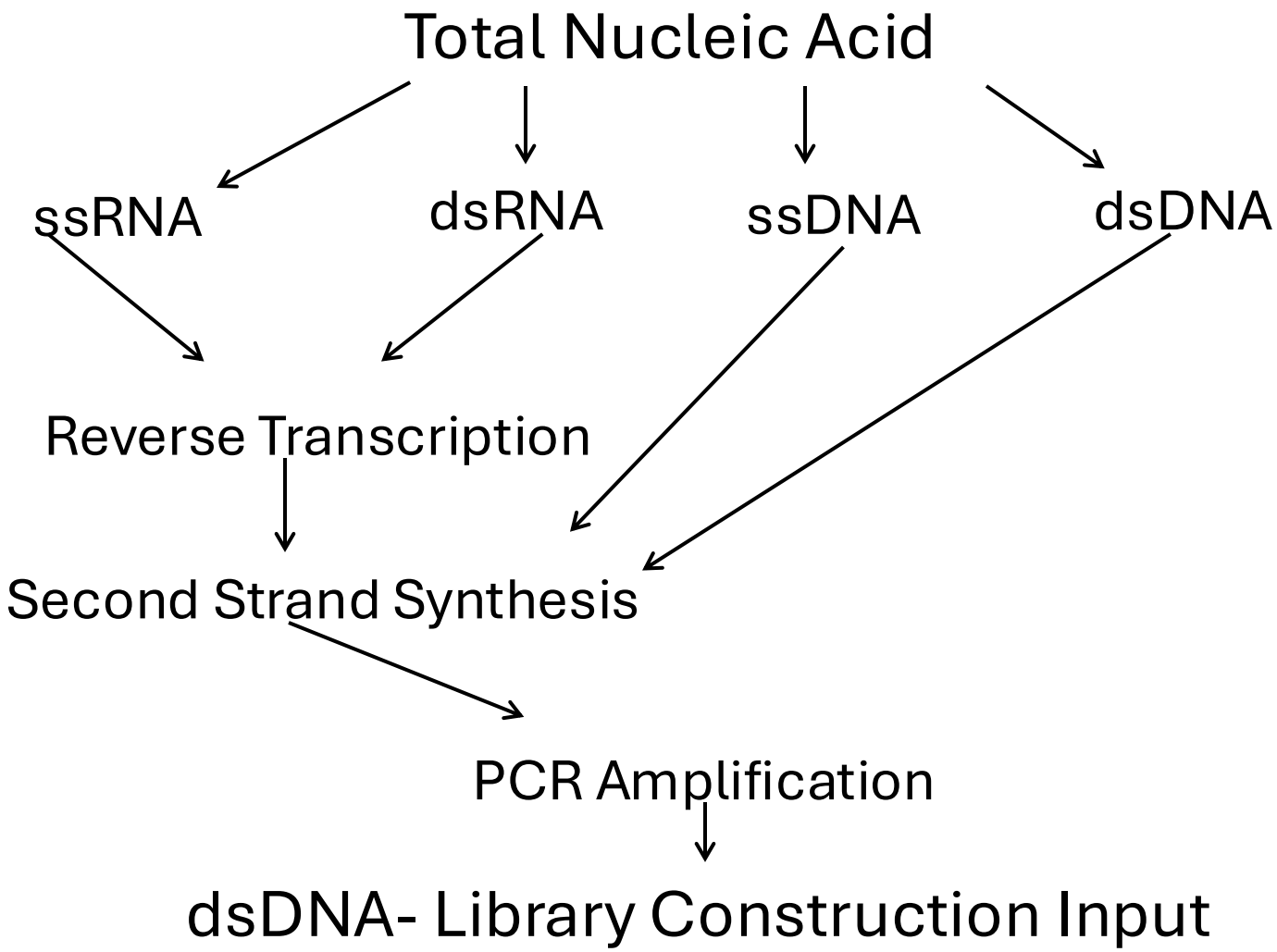
MagNA Pure 24 System

- Extract both DNA and RNA
- Automated systems
- Manual kits- Qiagen DNeasy

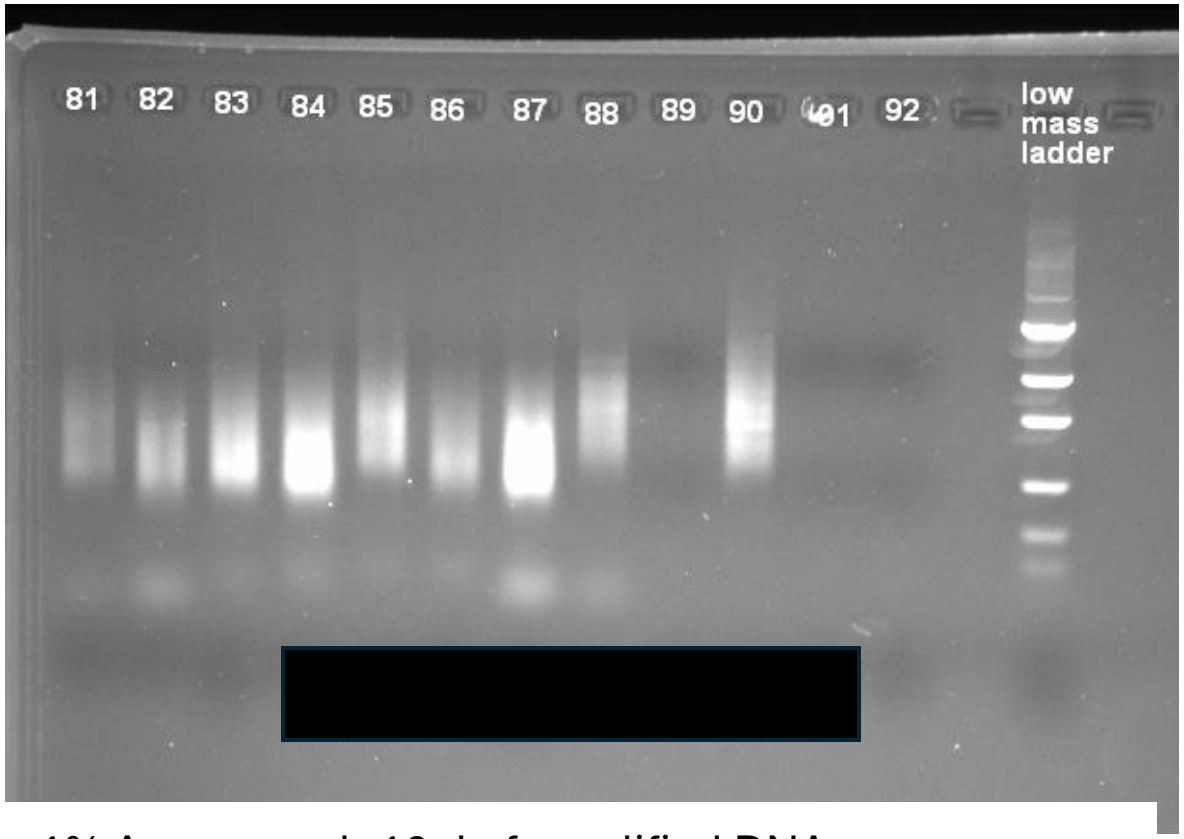
Preparing samples for virome sequencing

1. Virus Like Particle (VLP) Enrichment and Total Nucleic Acid Extraction
2. Reverse Transcription, Second Strand Synthesis and PCR Amplification
3. Library Construction

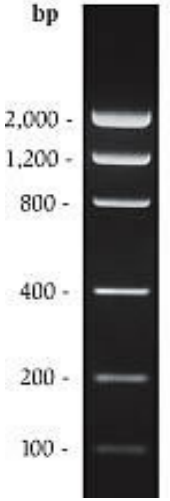
Convert RNA and ssDNA to dsDNA



Post PCR amplification DNA visualization



Invitrogen Low Mass Ladder



- 1% Agarose gel- 10ul of amplified DNA
- Expected smear 200bp-1kb
- Lanes 81-88 samples, lane 90 + control, lanes 89, 91 and 92 - controls
- Failure rate averages 8-10%

From Sample to Sequencing

- Nucleic acid extraction (VLP enrichment)
- Library preparation 
- Sequencing
- Data analysis

Preparing samples for virome sequencing

1. Virus Like Particle (VLP) Enrichment and Total Nucleic Acid Extraction
2. Reverse Transcription, Second Strand Synthesis, and PCR Amplification
3. Library Construction
   - Mechanical shearing (TruSeq)
   - Transposomes to cleave (Nextera)

Basic steps of library preparation

1. Fragmentation
2. End Repair
3. Addition of adapters
   - Index sequence (barcodes for pooling multiple samples)
   - Sequencing primer binding site
   - Amplification primers
   - Flow cell binding sequence- allows the sequence to bind to the flow cell
4. PCR amplification

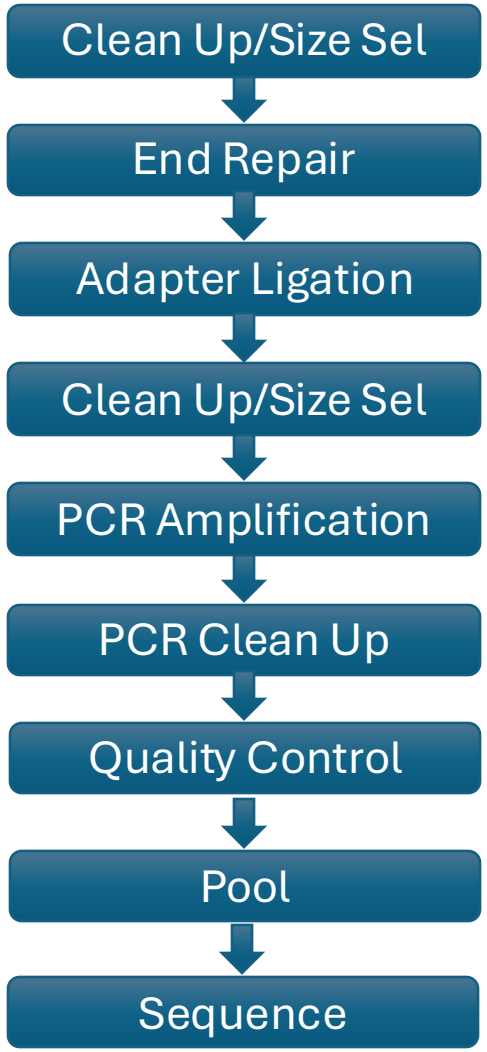
Many options for library preparation kits

SureSelectXT Reagent Kits

Preparing samples for virome sequencing

1. Virus Like Particle (VLP) Enrichment and Total Nucleic Acid Extraction
2. Reverse Transcription, Second Strand Synthesis, and PCR Amplification
3. Library Construction
   - Mechanical shearing (TruSeq/NEB Next DNA Library prep)
   - Transposomes to cleave (Nextera)

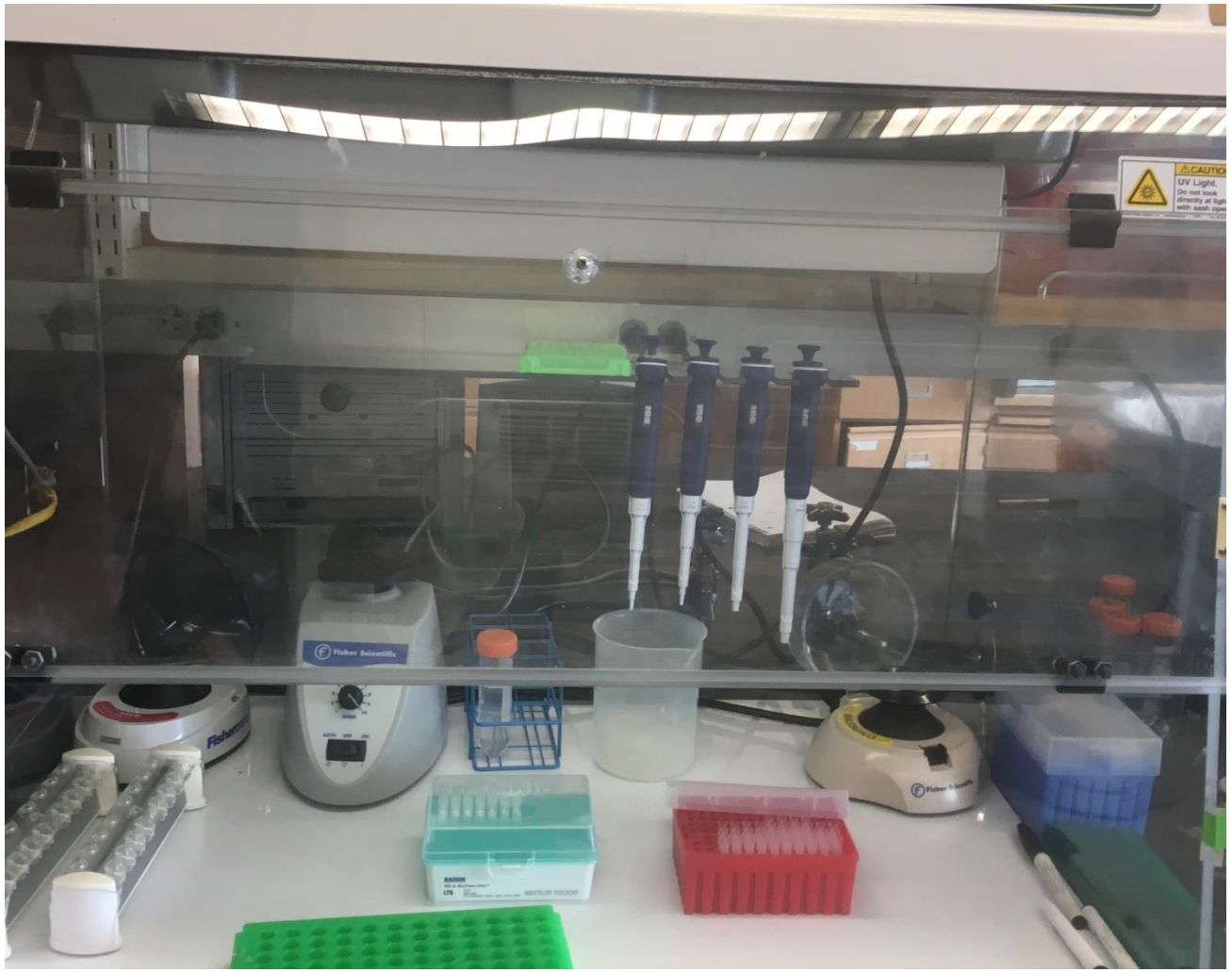
New England Biolabs NEBNext DNA Library Kit



PCR workstation

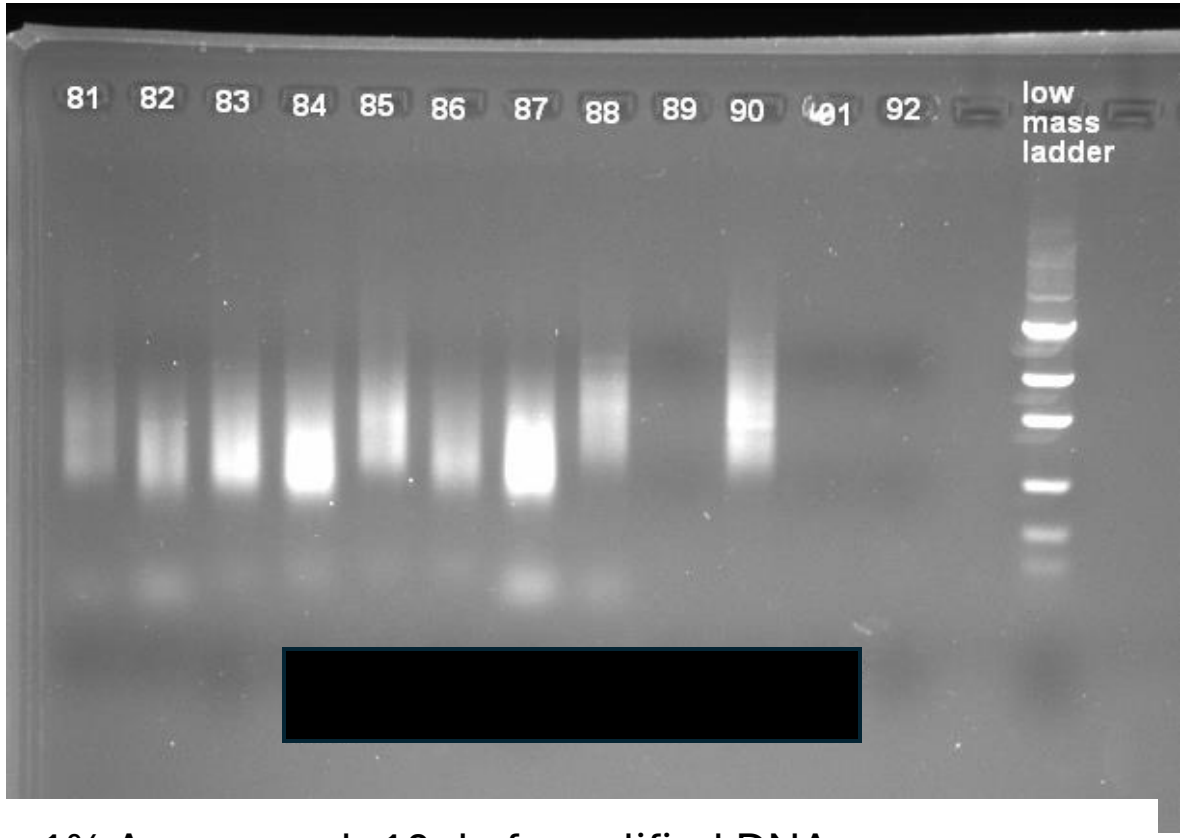
PCR workstation qualities:

- Sterile work area
- Vertical Laminar air flow
- Hepa filtration system
- Built in UV



Post PCR amplification visualization

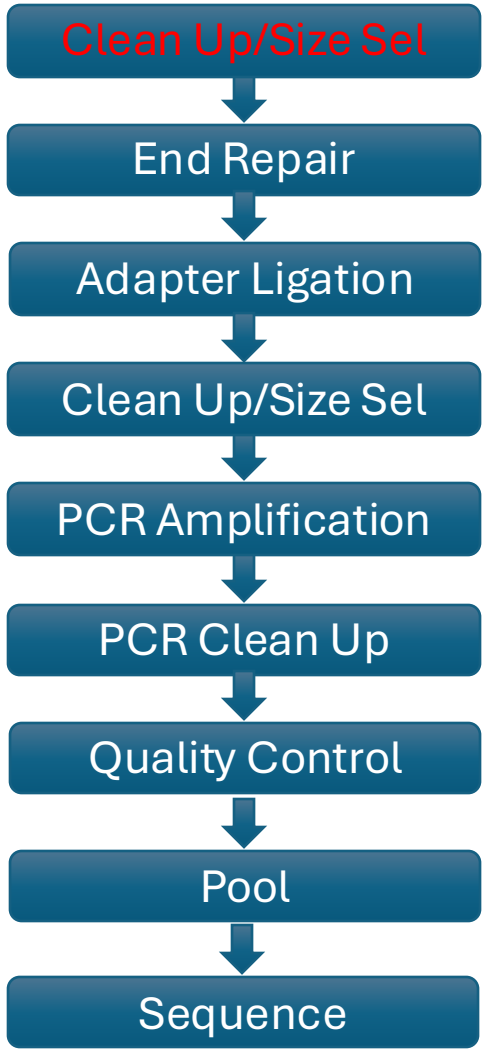
- Clean Up/Size Sel
- End Repair
- Adapter Ligation
- Clean Up/Size Sel
- PCR Amplification
- PCR Clean Up
- Quality Control
- Pool
- Sequence

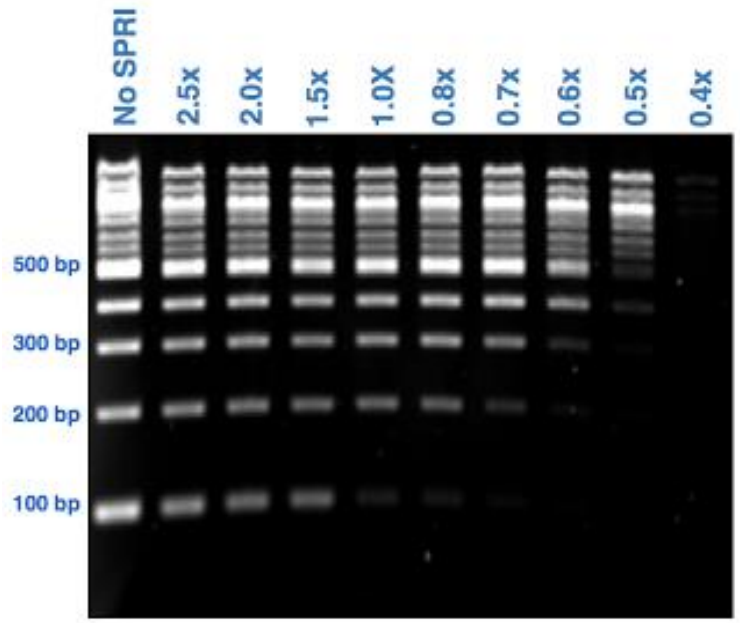


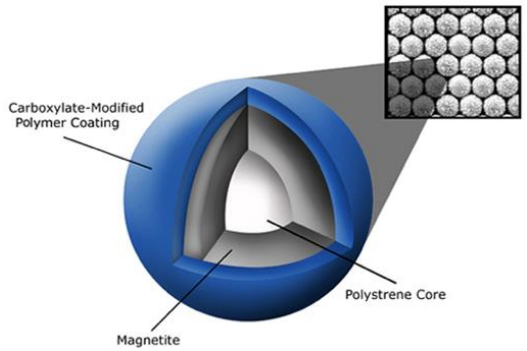
Invitrogen Low Mass Ladder

- 1% Agarose gel- 10ul of amplified DNA
- Expected smear 200bp-1kb
- Lanes 81-88 samples, lane 90 + control, lanes 89, 91 and 92 - controls
- Failure rate averages 8-10%

Size selection- Beckman Coulter AmPure bead





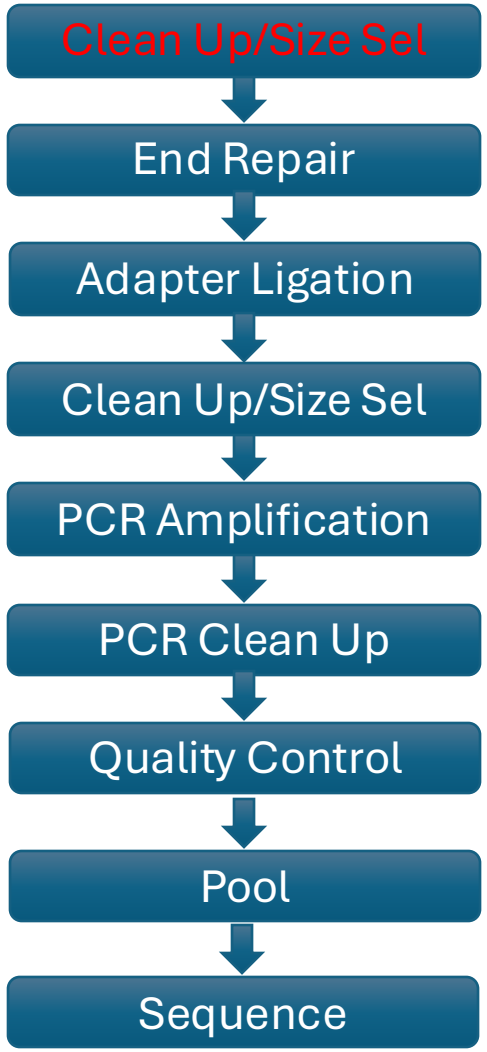


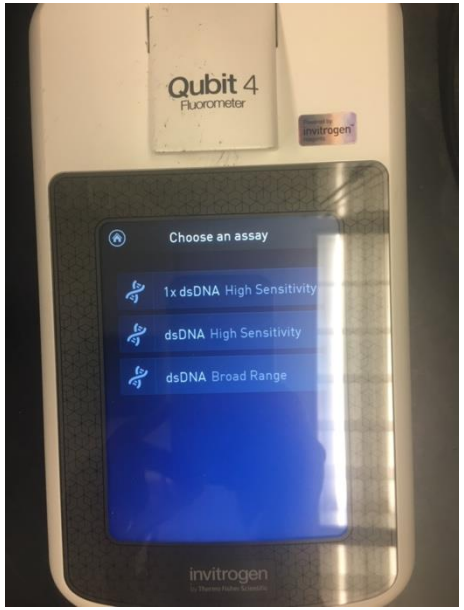
- SPRI Bead (Solid Phase Reversible Immobilization)
- Uses Paramagentic beads to selectively bind nucleic acid by size
- PEG (polyethylene glycol) causes the negatively charged DNA to bind to the carboxyl molecules on bead surface
- Lower the ratio of SPRI:DNA= larger final fragments at elution

Beckman Coulter SPRI bead video

https://youtu.be/zGV0SjCe0CU

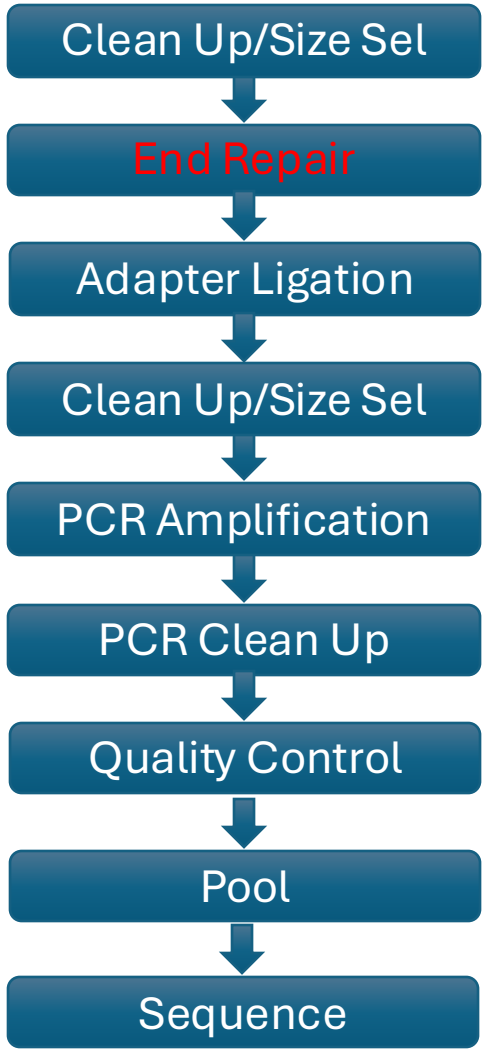
Sample quantification

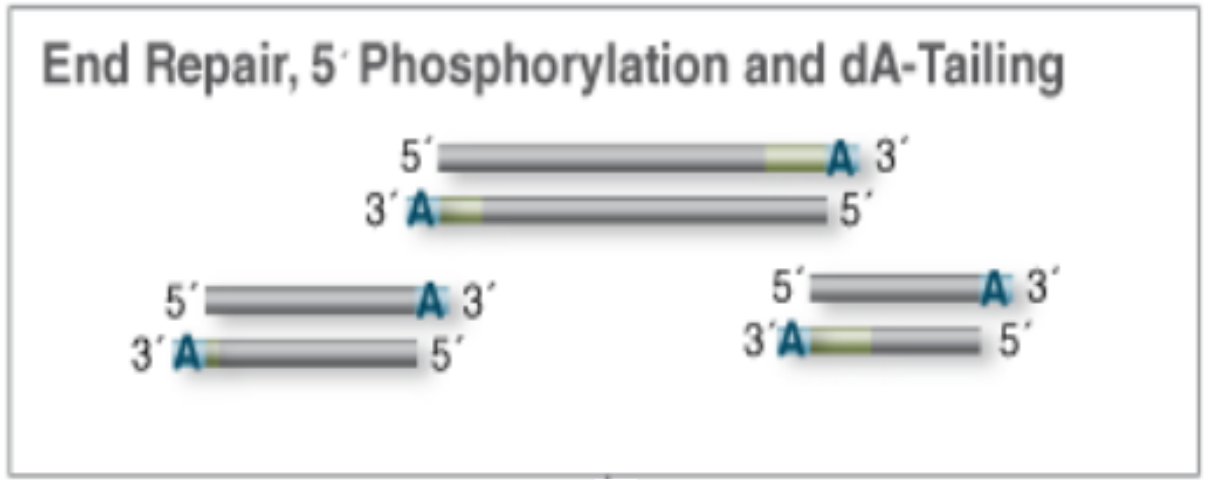




- Library input DNA 20-100ng
- Can go as low as 1ng
- Knowing input is critical for downstream steps- adapter concentration and PCR amplification cycle number

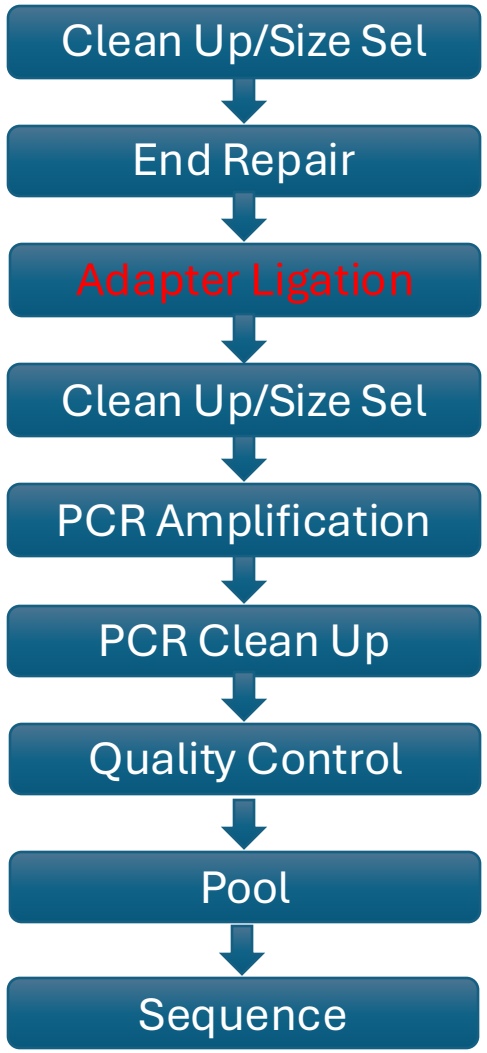
End Repair

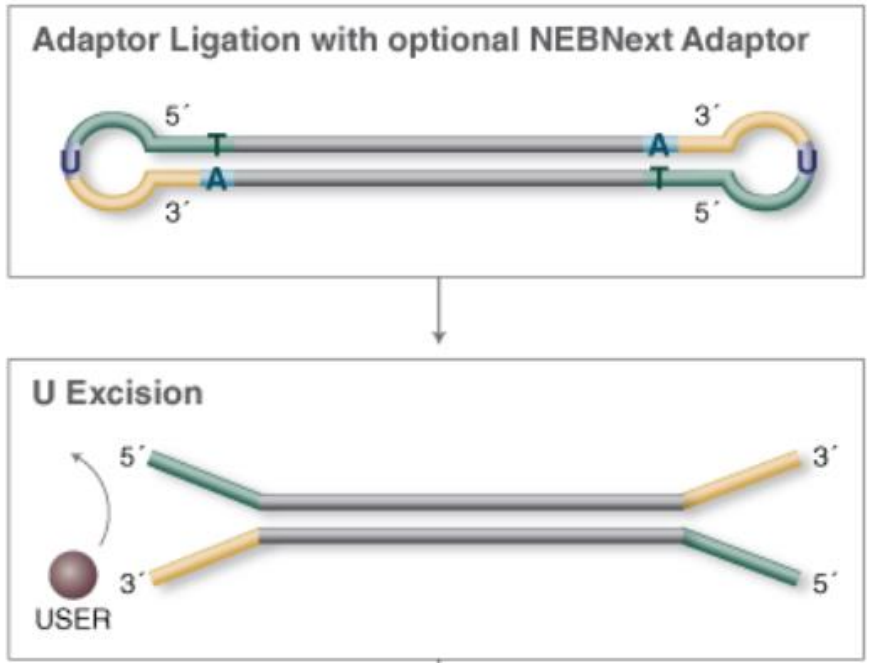


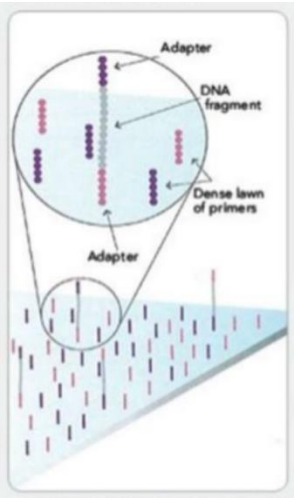


- Strands are blunted and phosphorylated
- Adding an A to 3’ ends

Adapter Ligation

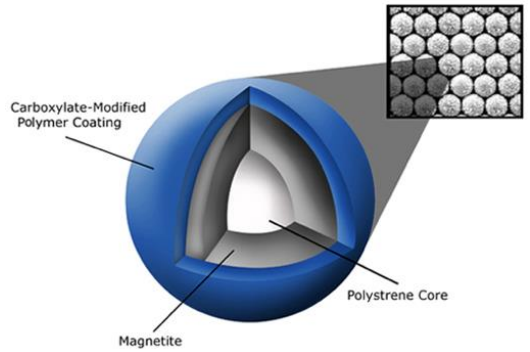






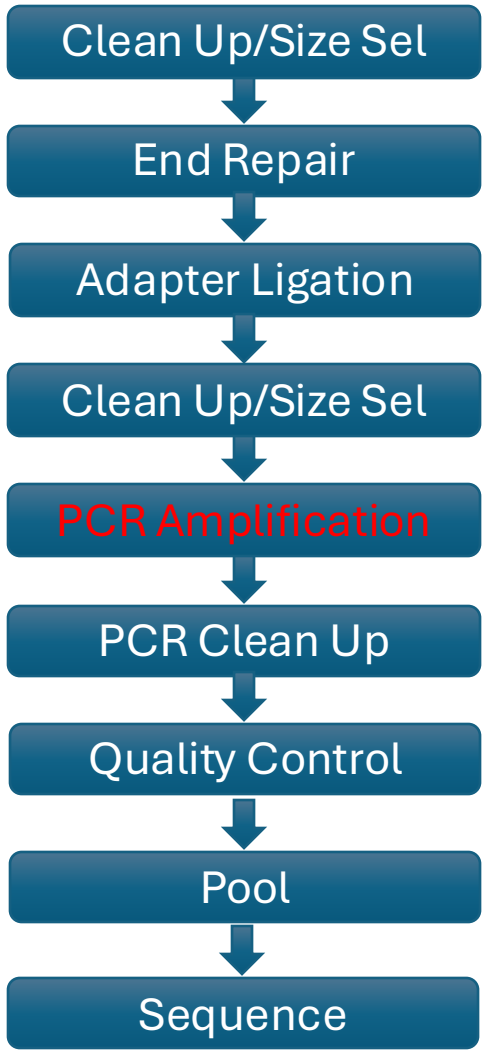
- Adapters with single T overhang ligated on the end repair dA fragment
- Amount of adapter is critical
- User enzyme used to cleave hairpin loop

Clean up- post adapter ligation

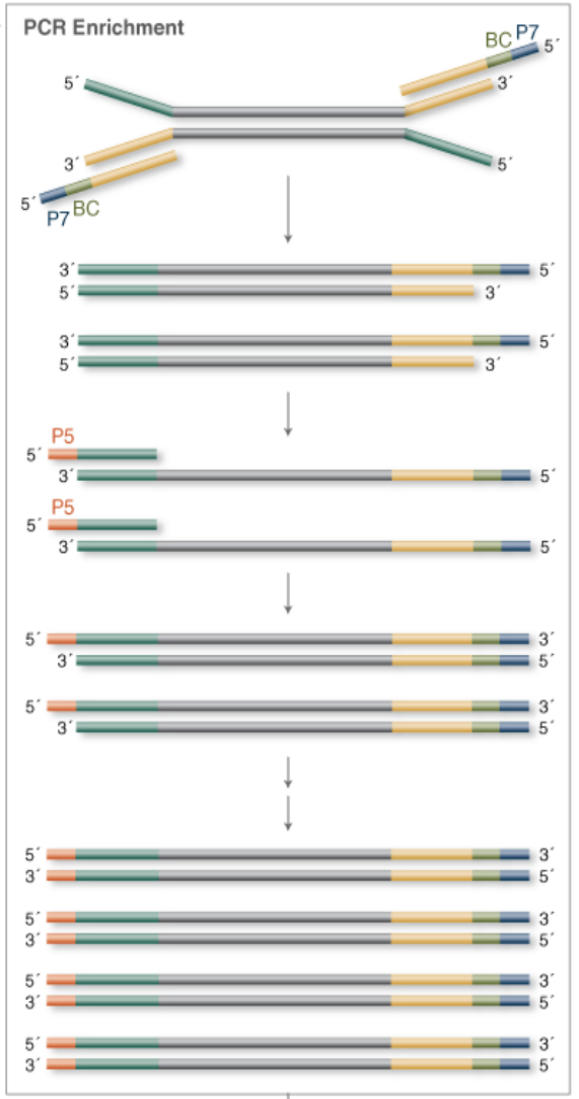


- Size Selection 400-600bp
- Remove unused ligation reaction components, adapter dimers, and concatemers

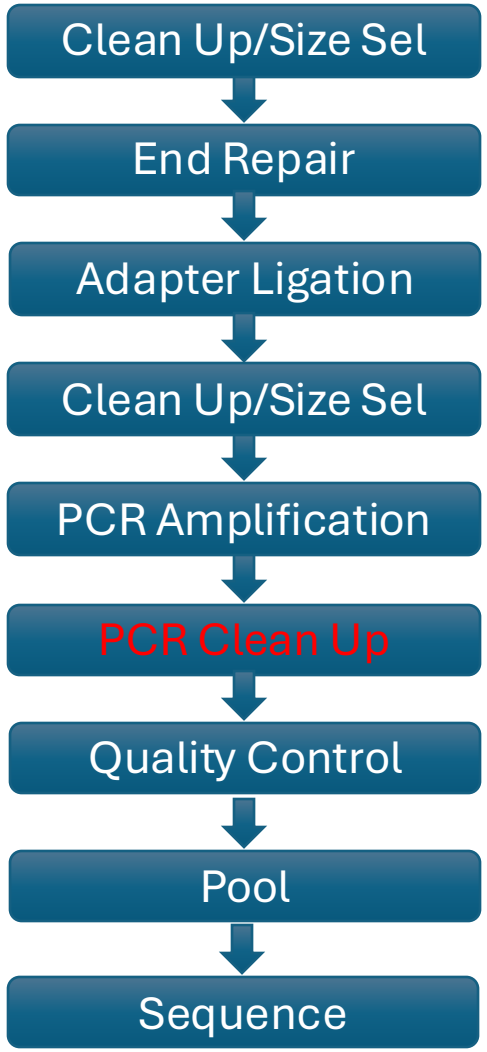
Library Amplification by PCR

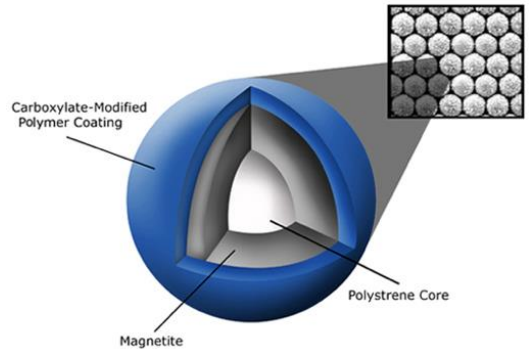


- Select for libraries with adapters on each end
- Increase the amount of library
- Indexes can be added for multiplexing- 24 unique indexes



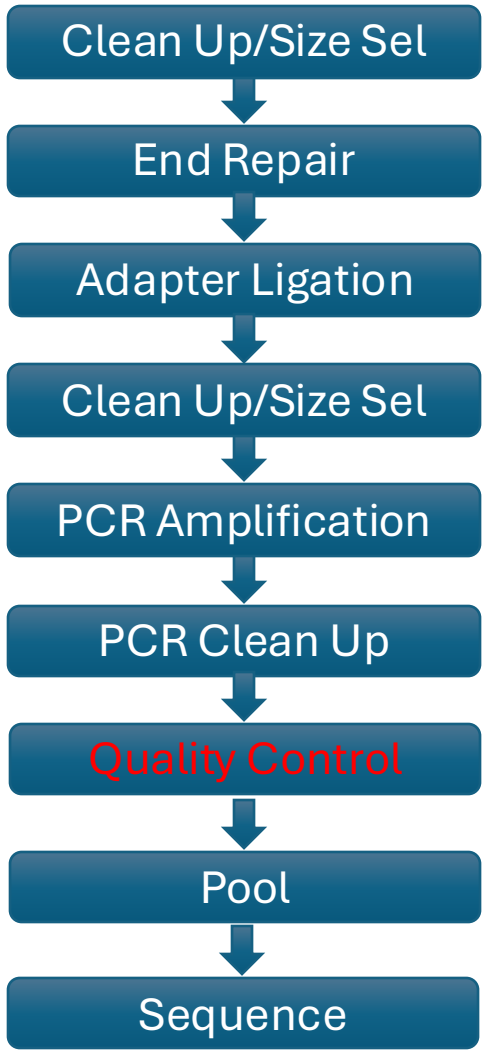
PCR clean up

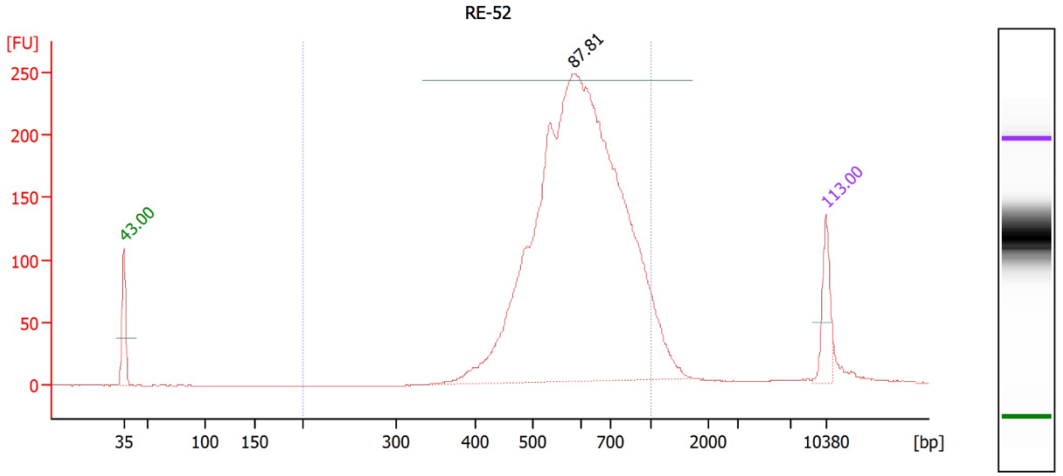




- Remove free barcodes, nucleotides
- Remove adapter dimers

Quality control





Overall Results for sample 9 : RE-52

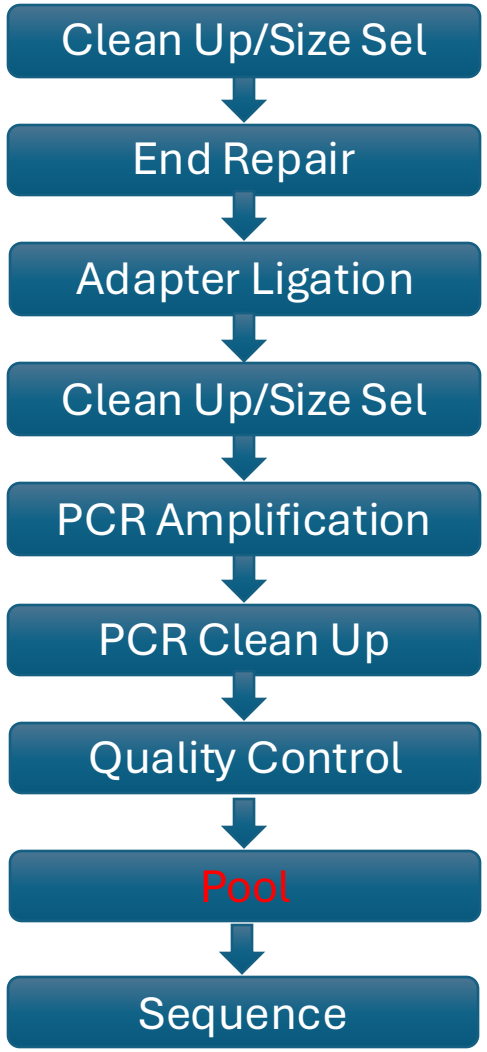
Number of peaks found:	1	Corr. Area 1:	2,893.5
Noise:	0.3		

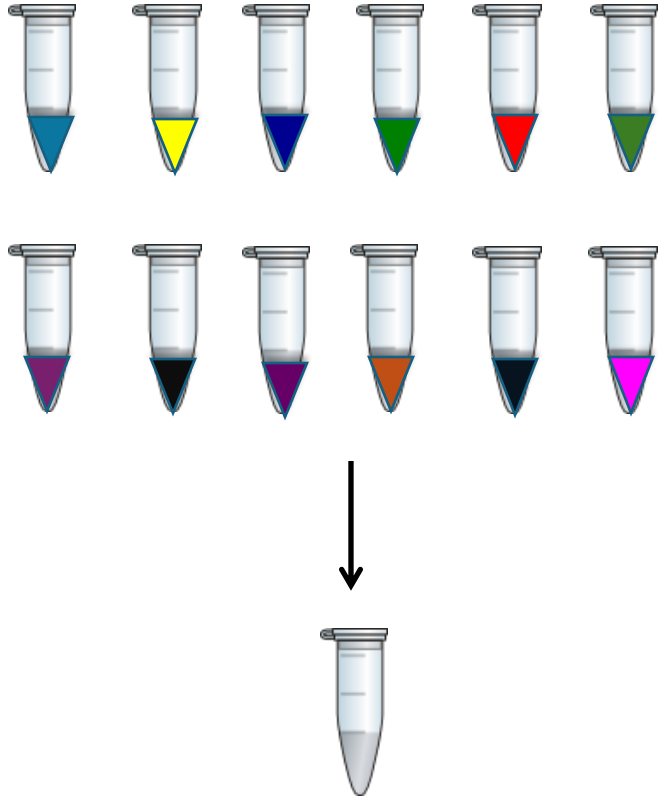
Peak table for sample 9 : RE-52

Peak	Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	35	125.00	5,411.3	Lower Marker
2	585	2,754.89	7,140.6	
3	10,380	75.00	10.9	Upper Marker

- Agilent Bioanalzyer 2100
- Microfluidics platform for sizing and quantification
- Tapestation

Pool final libraries





- Individual Barcode for multiplexing
- Pool equal molar concentration
- Sequencing Core requires 20ul at 2-10nM

Sequence

Clean Up/Size Sel

↓

End Repair

↓

Adapter Ligation

↓

Clean Up/Size Sel

↓

PCR Amplification

↓

PCR Clean Up

↓

Quality Control

↓

Pool

↓

Sequence

- MiSeq
- Paired end 2X250
- Target 1 million reads/sample

Illumina Sequencing Platforms

Key specifications	iSeq 100 System	MiniSeq System	MiSeq System	MiSeq i100 Series[a] (New)	NextSeq 550 System	NextSeq 1000 and 2000 Systems
Max output per flow cell	1.2 Gb [b]	7.5 Gb [c]	15 Gb [d]	30 Gb [a]	120 Gb [c]	540 Gb [e]
Run time (range)[e]	~9.5–19 hr	~5–24 hr	~5.5–56 hr	~4–15.5 hr	~11–29 hr	~8–44 hr
Max reads per run (single reads)	4M [ab]	25M [c]	25M [d]	100M [a]	400M [c]	1.8B [e]
Max read length	2 × 150 bp	2 × 150 bp	2 × 300 bp	2 × 300 bp	2 × 150 bp	2 × 300 bp

Key specifications	NextSeq 1000 and 2000 Systems	NovaSeq 6000 System	NovaSeq X Series
Max output per flow cell	540 Gb[a]	3 Tb[b]	8 Tb[c]
Run time (range)[d]	~8–44 hr	~13–44 hr	~17–48 hr
Max reads per run (single reads)	1.8B[a]	10B (single flow cell)[b] 20B (dual flow cells)[b]	26B (single flow cell)[c] 52B (dual flow cells)[c,f]
Max read length	2 × 300 bp	2 × 250 bp	2 × 150 bp

Cost per gigabase	$262	$54	$40-$63	$10-$35	$2

Preparing samples for virome sequencing

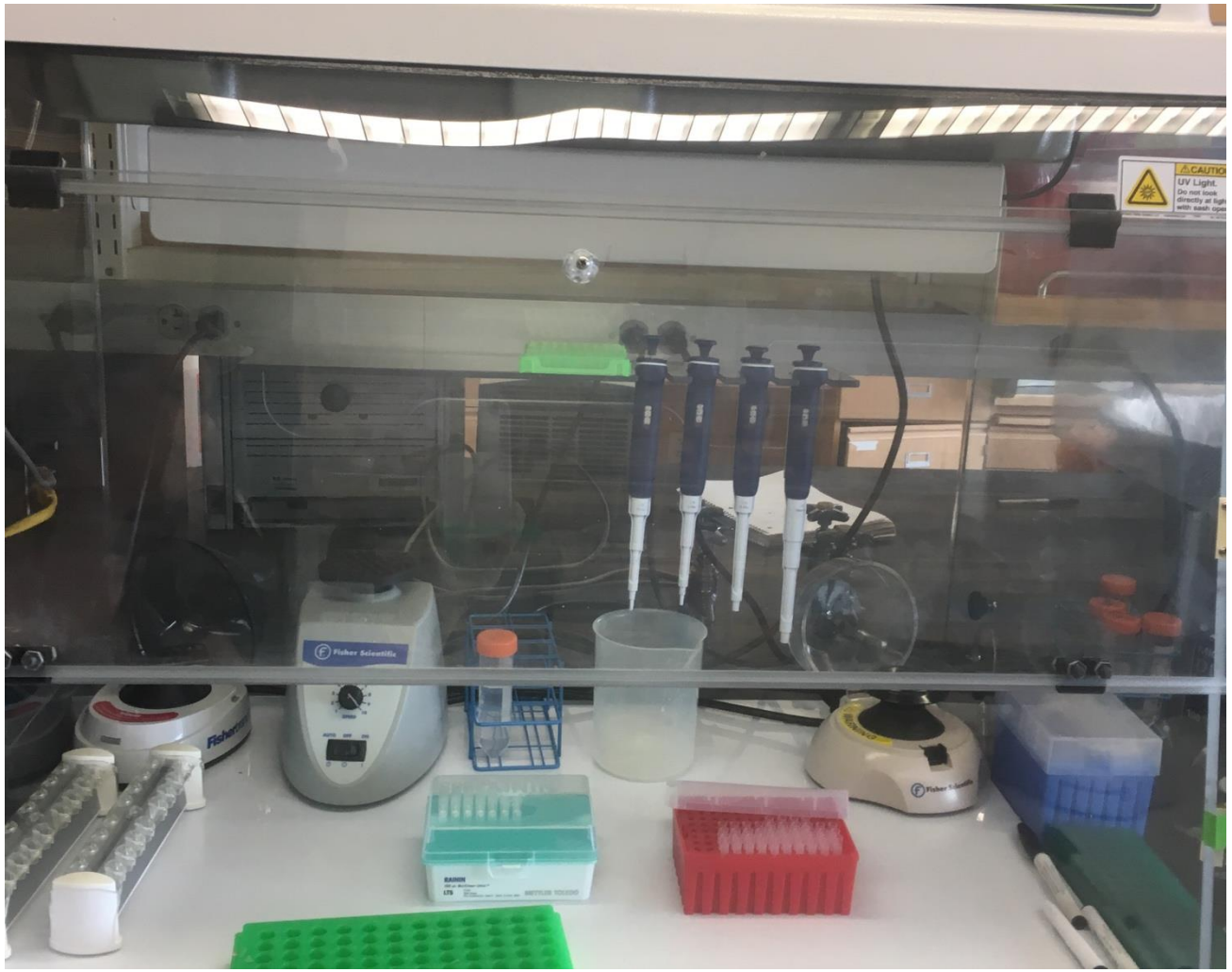
1. Virus Like Particle (VLP) Enrichment and Total Nucleic Acid Extraction
2. Reverse Transcription, Second Strand Synthesis, and PCR Amplification
3. Library Construction
   - mechanical fragmentation (TruSeq/NEB Next DNA Library prep)
   - transposon fragmentation (Nextera/Illumina DNA Library prep) 

Illumina DNA Library Prep

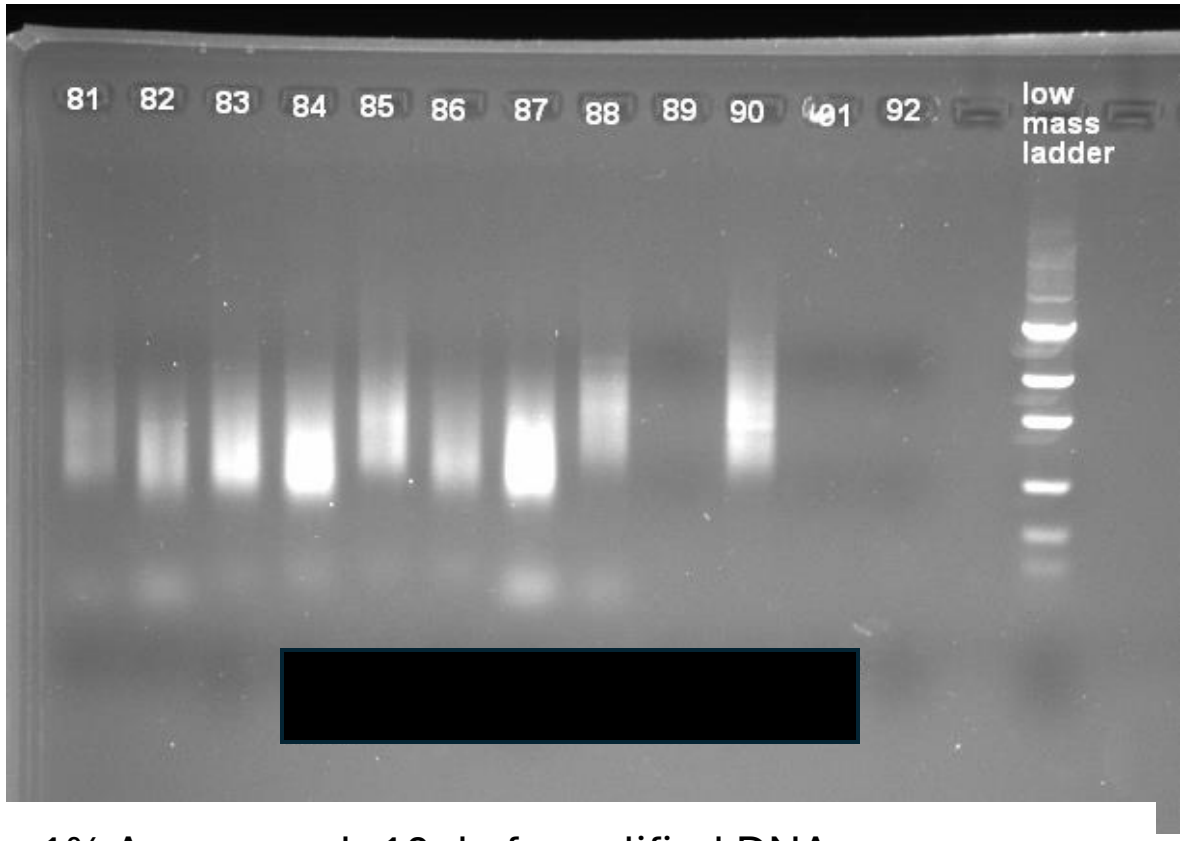
PCR workstation

PCR workstation qualities:

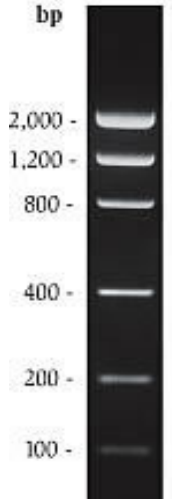
- Sterile work area
- Vertical Laminar air flow
- Hepa filtration system
- Built in UV



Post PCR amplification visualization

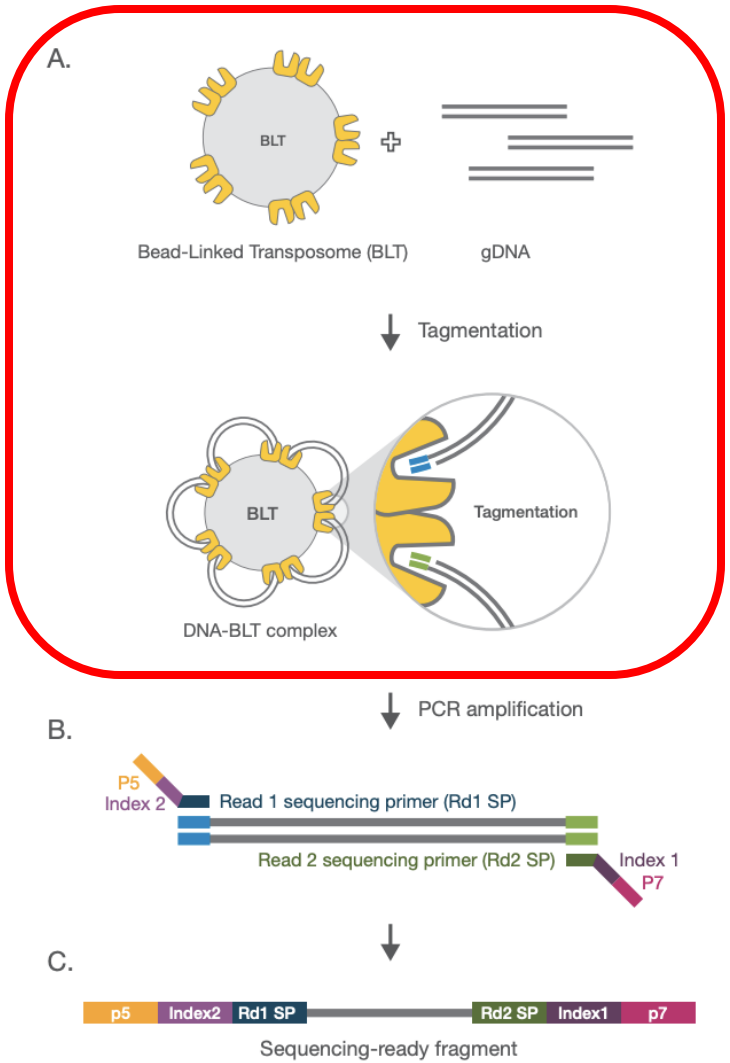


Invitrogen Low Mass Ladder

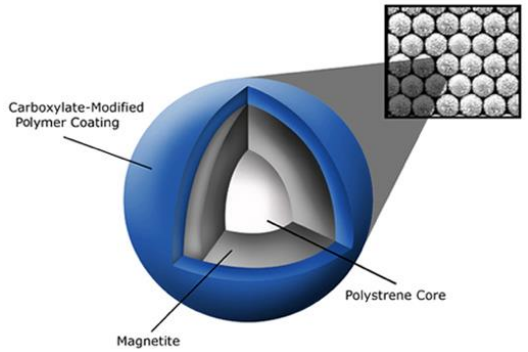


- 1% Agarose gel- 10ul of amplified DNA
- Expected smear 200bp-1kb
- Lanes 81-88 samples, lane 90 + control, lanes 89, 91 and 92 - controls
- Failure rate averages 8-10%

Tagmentation: fragmentation and adapter ligation



Post tagmentation clean up



- Washes the adapter-tagged DNA on the BLT before PCR amplification

PCR amplification

Clean Up

Tagmentation

Post Tag Clean Up

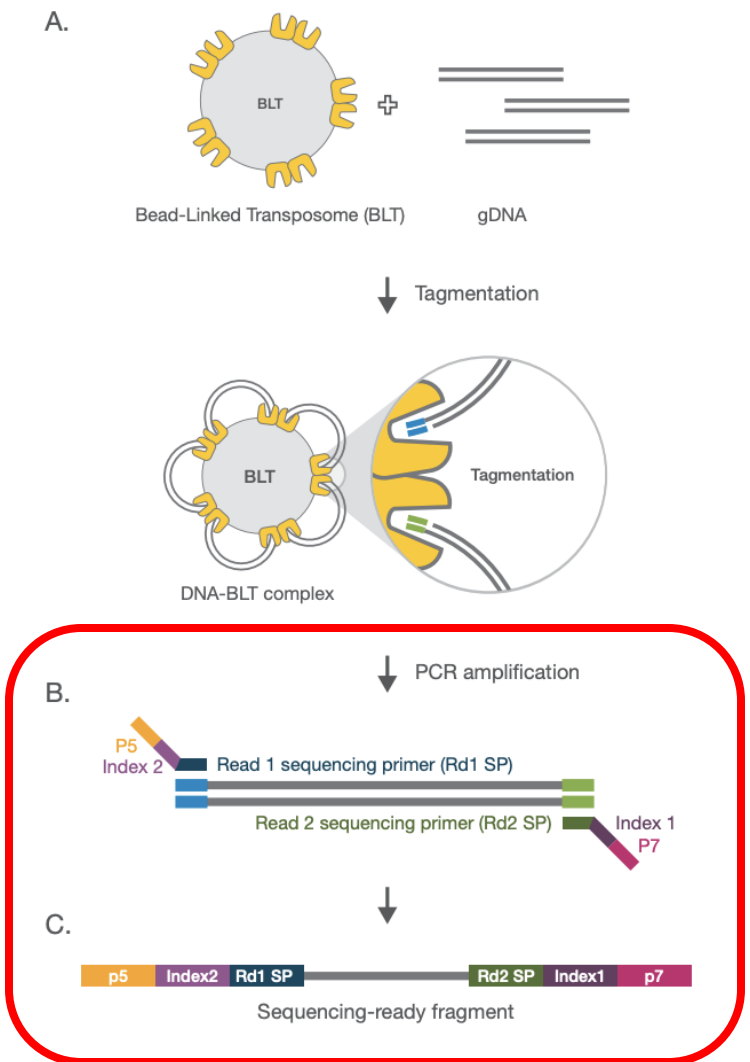
PCR Amplification

PCR Clean Up

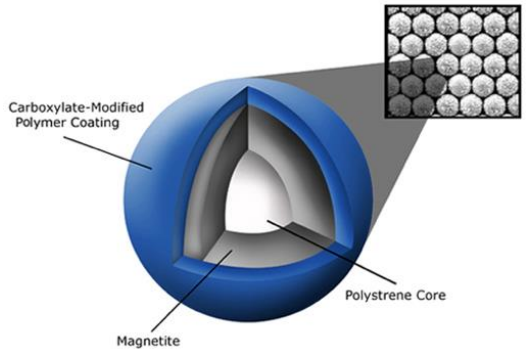
Quality Control

Pool

Sequence

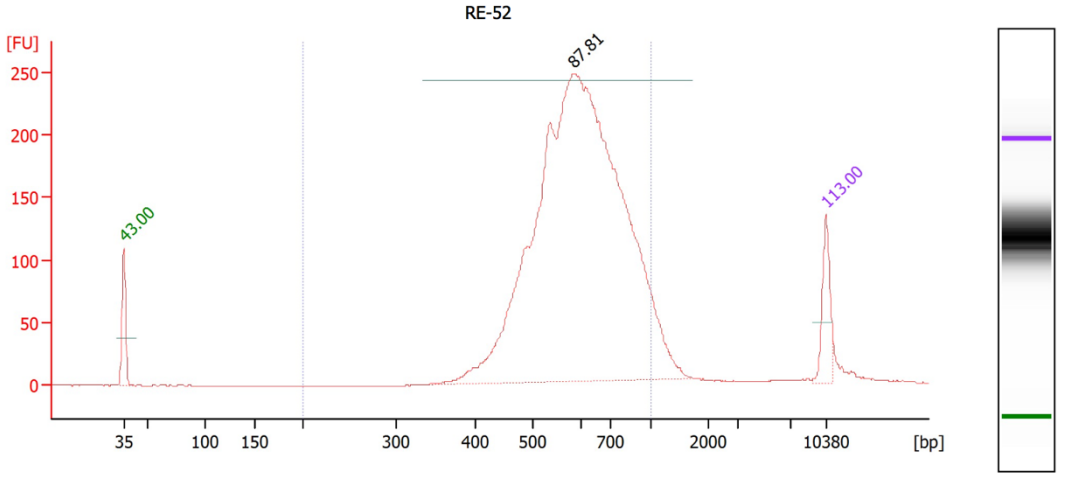


PCR amplification clean up



- Remove free barcodes, nucleotides
- Remove adapter dimers

Quality control



Overall Results for sample 9 : RE-52

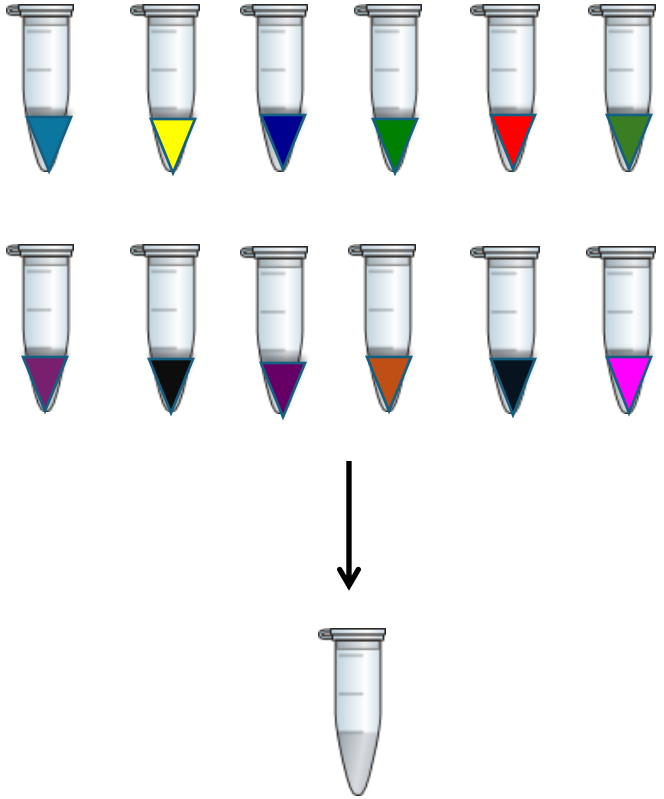
Number of peaks found:	1	Corr. Area 1:	2,893.5
Noise:	0.3		

Peak table for sample 9 : RE-52

Peak	Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	35	125.00	5,411.3	Lower Marker
2	585	2,754.89	7,140.6	
3	10,380	75.00	10.9	Upper Marker

- Agilent Bioanalzyer 2100
- Microfluidics platform for sizing and quantification
- Tapestation

Pool final libraries



- Individual Barcode for multiplexing
- Pool equal molar concentration
- Sequencing Core requires 20ul at 2-10nM

Sequence

Clean Up

Tagmentation

Post Tag Clean Up

PCR Amplification

PCR Clean Up

Quality Control

Pool

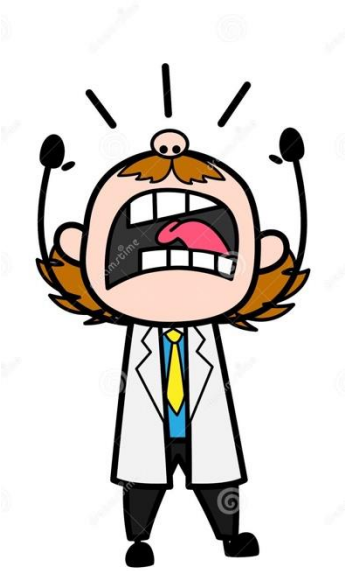
Sequence

- NovaSeq X
- Paired end 2X150
- Target 2 million reads/sample

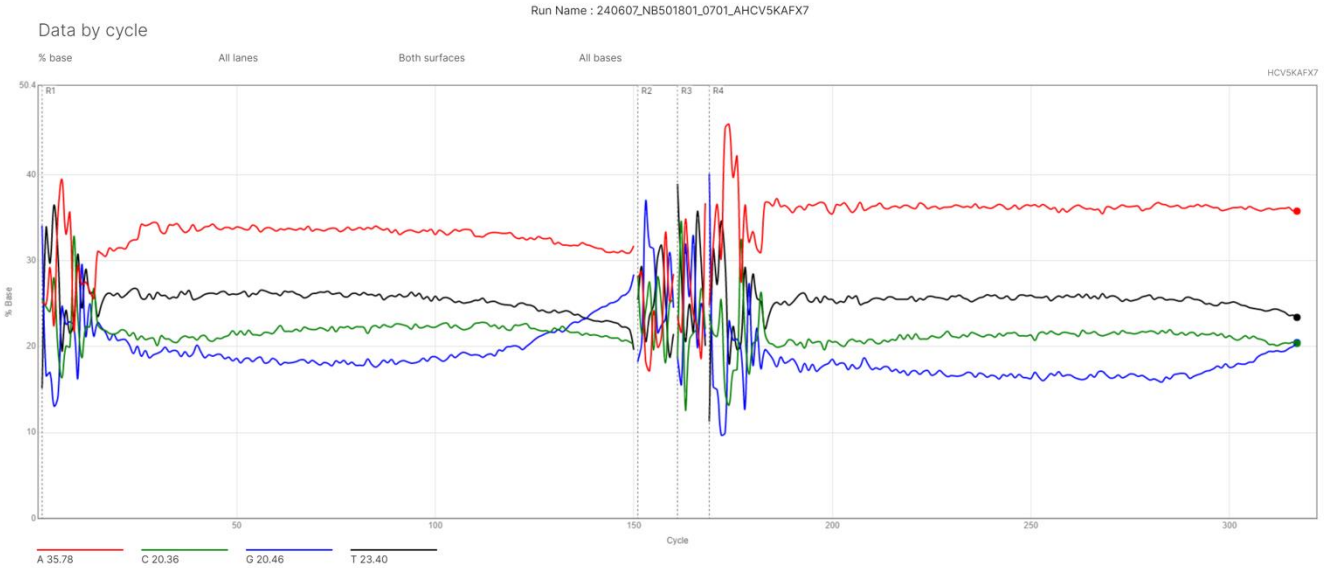
What can go wrong

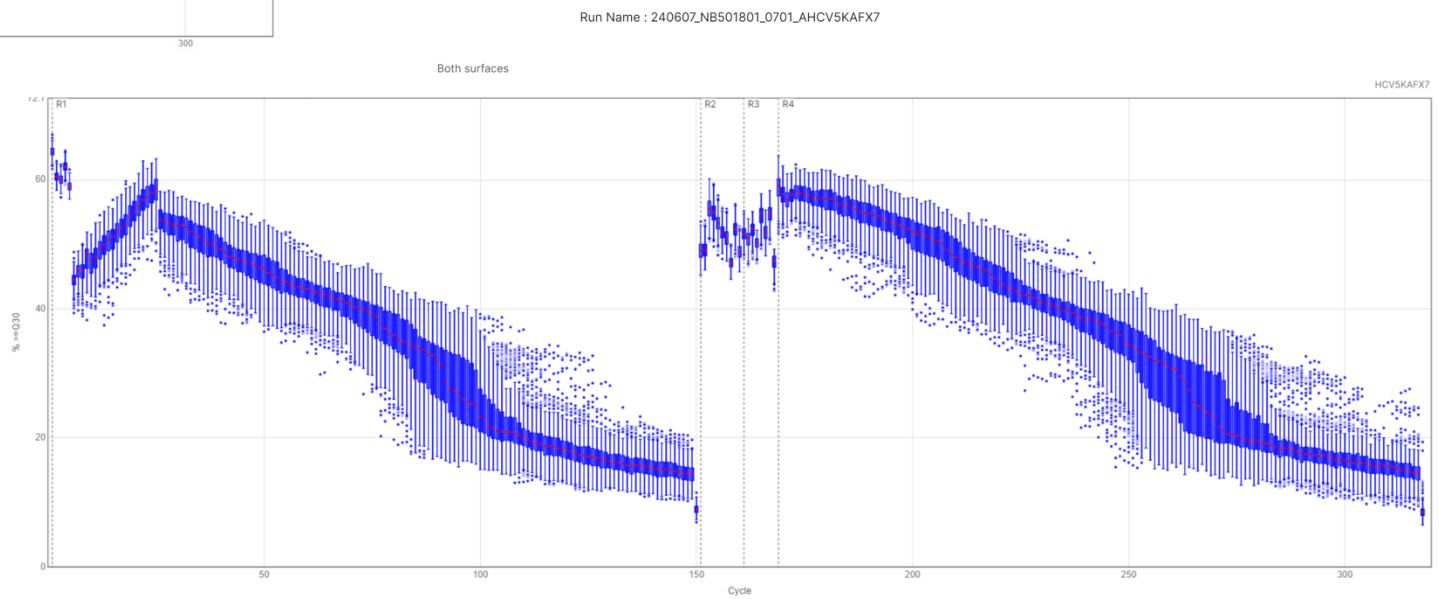
- Sequencing
  - Small fragments preferentially bind to flow cells
  - Uneven base composition
  - Flow cell is overloaded
  - Read through to adapters
  - Read through beyond adapters
  - Regions with high GC content are underrepresented due to limitations of sequencing chemistry
  - Base calling errors at the end of reads
  - Homopolymers
    - Regions containing long stretches of the same base

- Library preparation
  - Poor DNA quality for input
  - Libraries are not the correct length
  - Library artifacts

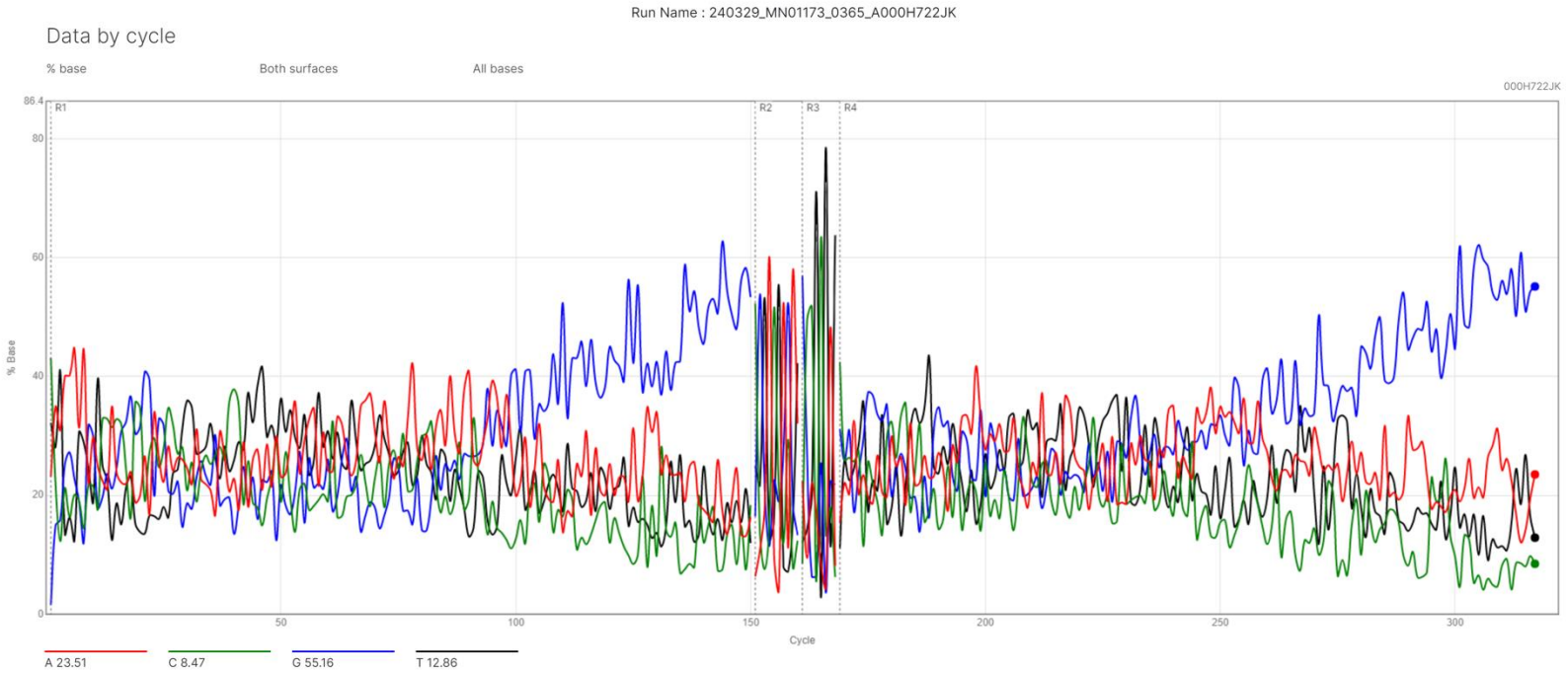


Uneven base composition

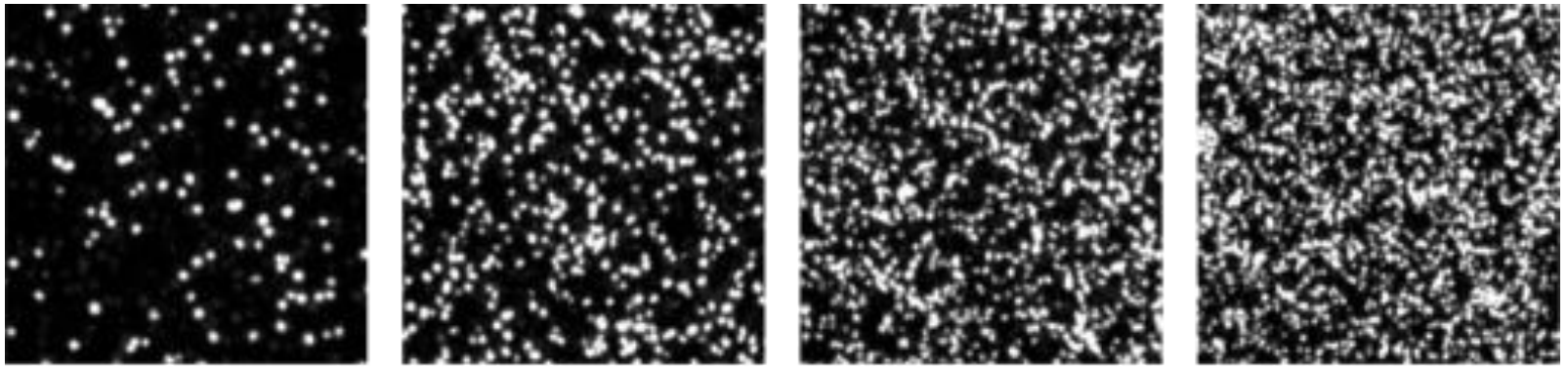


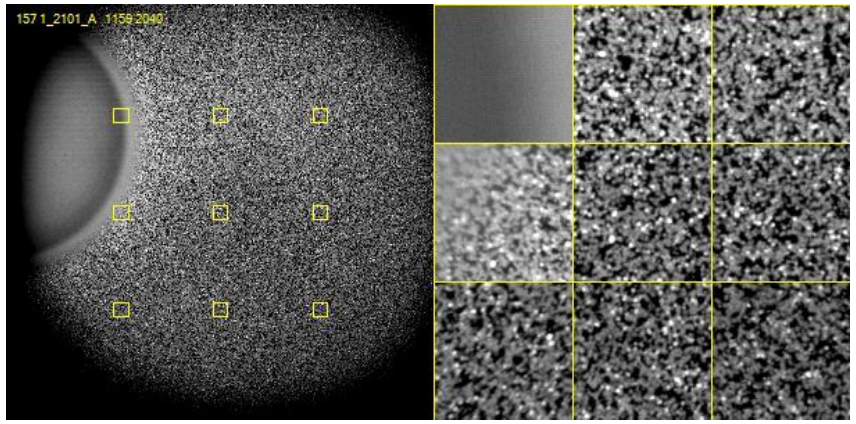


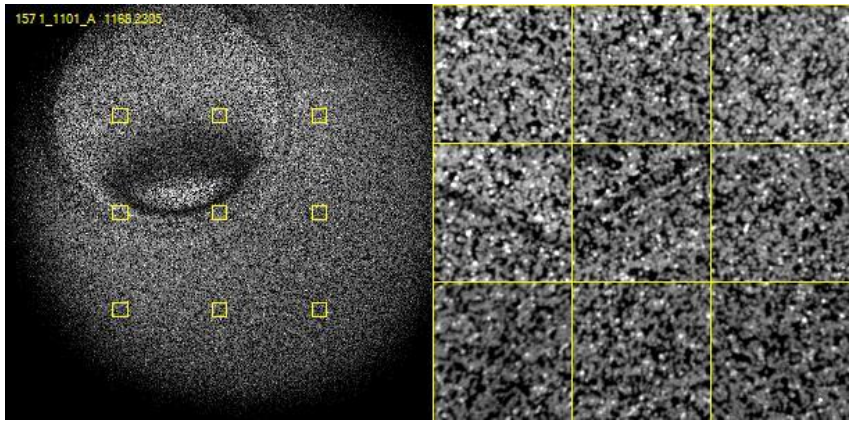
Library fragments are too short



Cluster Issues

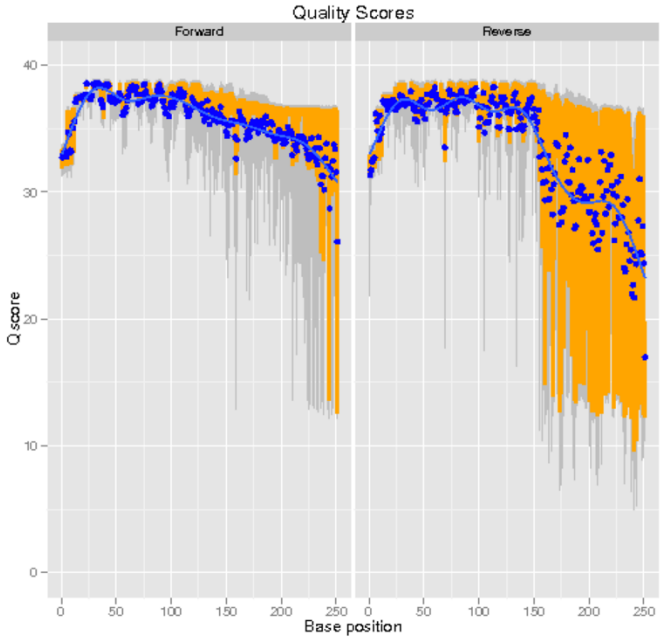




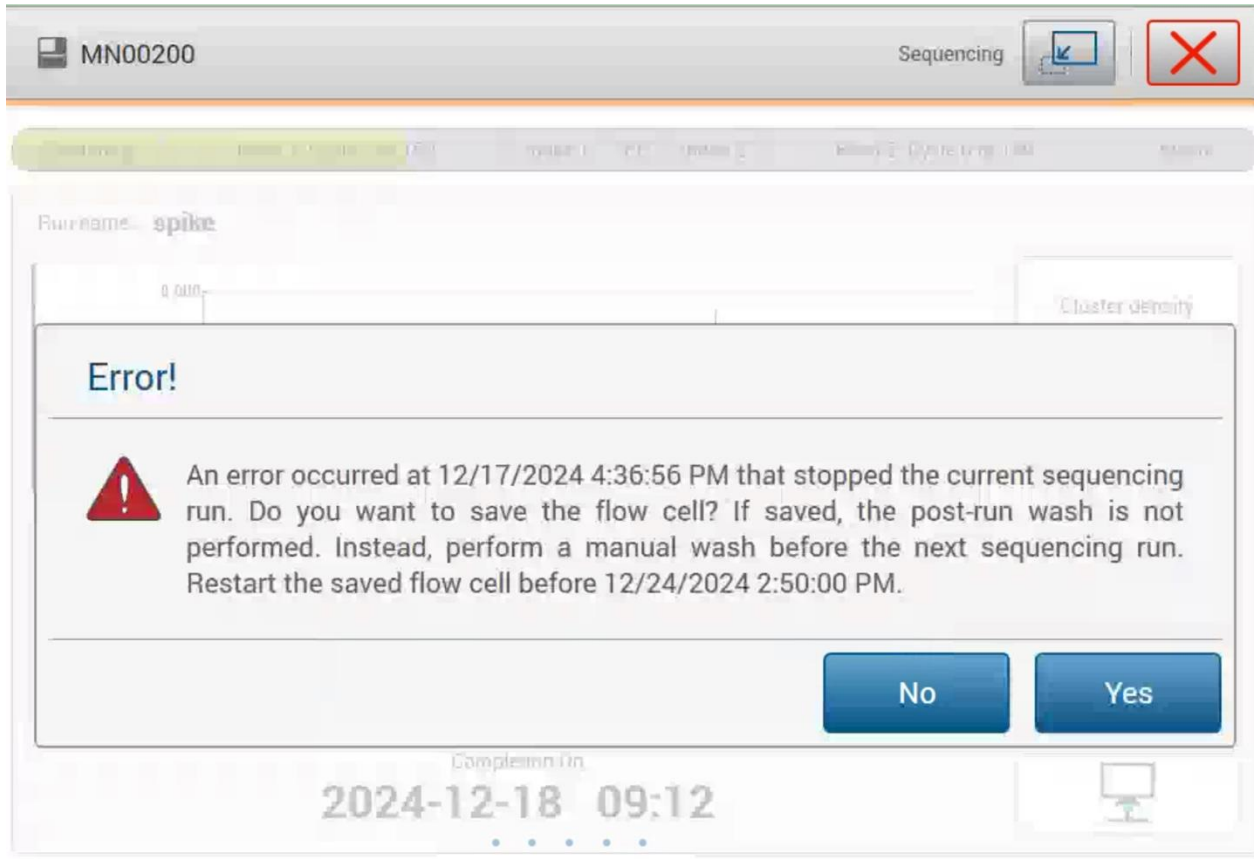


Reagent issues

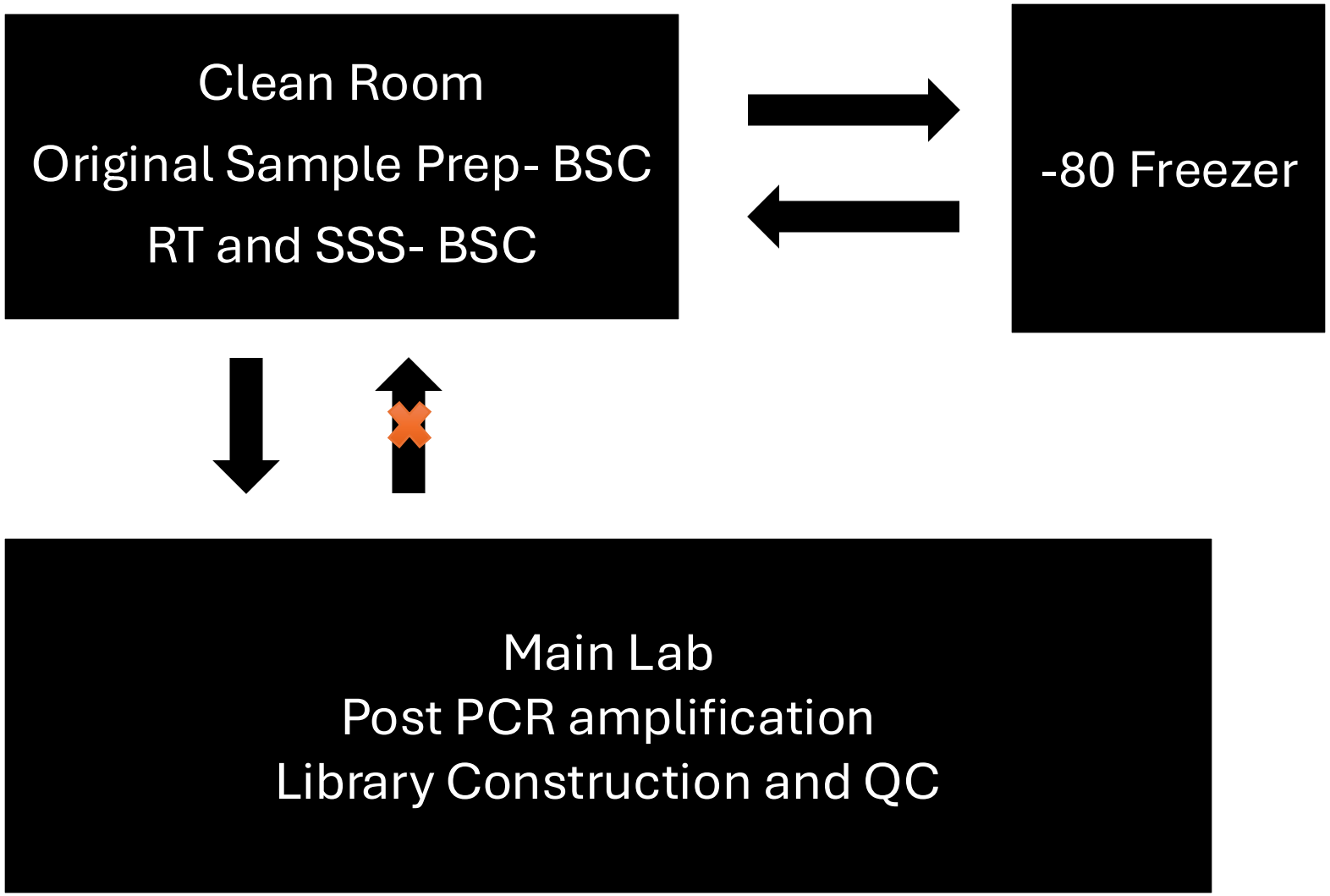
Why do quality scores drop towards the end of a read?



Run failure


MN00200
Sequencing
Error!
An error occurred at 12/17/2024 4:36:56 PM that stopped the current sequencing run. Do you want to save the flow cell? If saved, the post-run wash is not performed. Instead, perform a manual wash before the next sequencing run. Restart the saved flow cell before 12/24/2024 2:50:00 PM.
No
Yes
2024-12-18 09:12

Laboratory layout



Acknowledgements

Handley Lab
Scott Handley
Leran Wang
Kathie Mihindukulasuriya
Luis Chica
Dhoha Abid
Megan Johnson
Martina Moore

Dave Wang

Kwon Lab
Doug Kwon
Joseph Elsherbini
Sarah Eisa
Cameron Reitan

Center for Genome Sciences
Jessica Hoisington-Lopez
MariaLynn Crosby