Center for Surveillance, Epidemiology, and Laboratory Services

Next Generation Sequencing and the CDC/APHL NGS Quality Initiative

Christin Hanigan (APHL)
Genevieve Langley (CDC/Booz Allen Hamilton)
Sarah Mason St. Prix (CDC/ Booz Allen Hamilton)
Diego Arambula (CDC)

September 16th, 2021

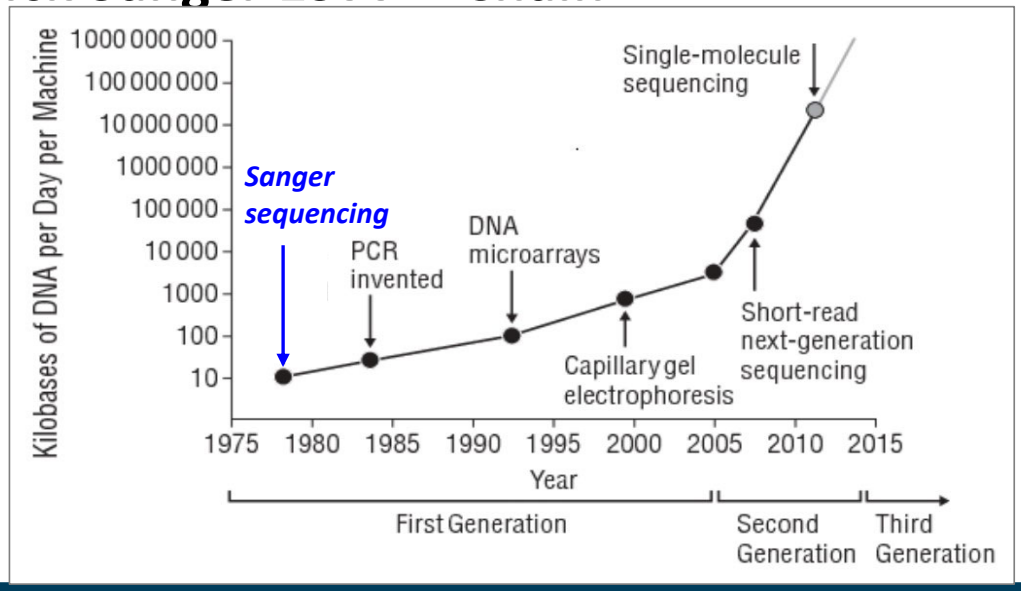
APHL

CDC

U.S. Department of Health and Human Services
Centers for Disease Control and Prevention

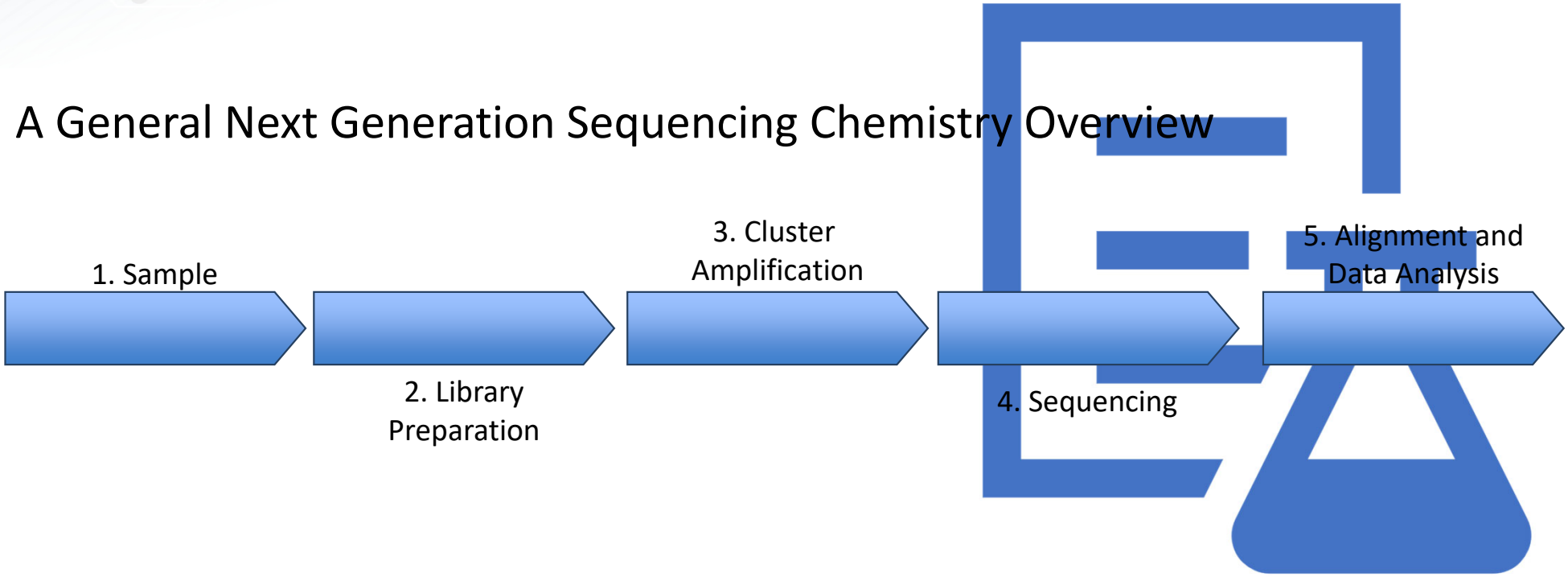
A Brief History of Sequencing

- **Oswald Theodore Avery in 1944** isolated DNA as the material of which genes and chromosomes are made.
- **Watson & Crick in 1953** determined DNA structure was a double helix.
- **First Generation Sequencing: Frederick Sanger 1977 – Chain Termination**
  - Incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication
  - Most widely used sequencing method for 40 years
  - Still used for small scale projects and validation of next generation sequencing (NGS) results



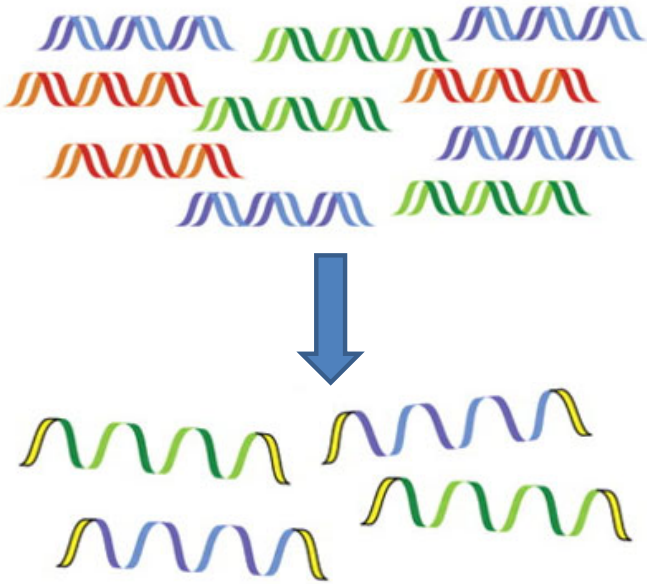
Next Generation Sequencing Workflow and Chemistry

A General Next Generation Sequencing Chemistry Overview



Next Generation Sequencing Library Prep

Genomic DNA or complementary DNA (cDNA) is sheared into fragments and size selected, then separated into single strands.



FRAGMENTS OF 200-300 BASE PAIR (bp)

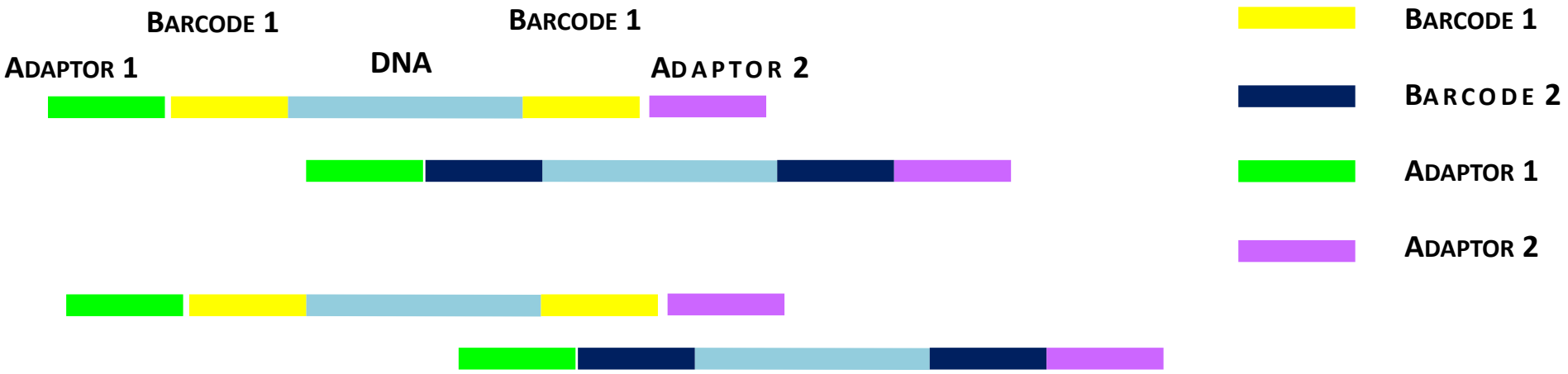
Barcodes and universal adaptor sequences are ligated to the target pool of DNA.

https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

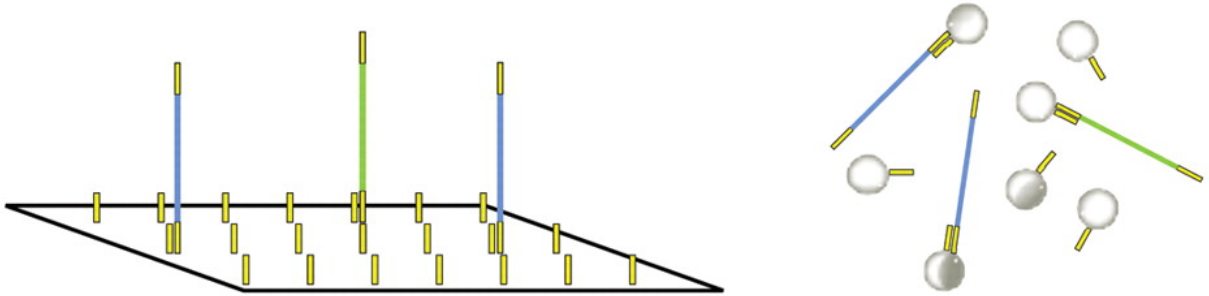
Next Generation Sequencing- Library Prep

Adapters and barcodes are ligated to the DNA

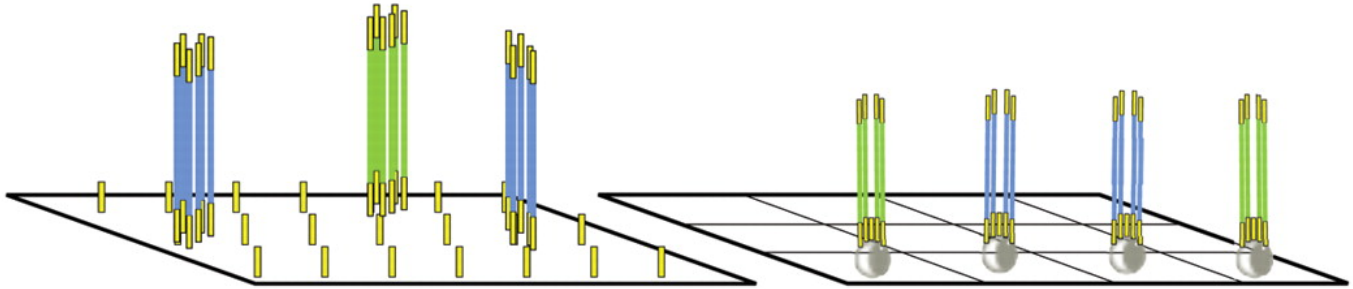
- Adapters are short DNA oligos that contain the primer sites that the sequencer needs to generate the sequencing read.
- Barcodes allow for multiple samples to be in the same sequencing run.



Next Generation Sequencing – Cluster Amplification



The DNA fragments are washed over an array or incubated with microscopic beads such that one DNA molecule is anchored by its adaptor on a single bead or away from other fragments on an array.

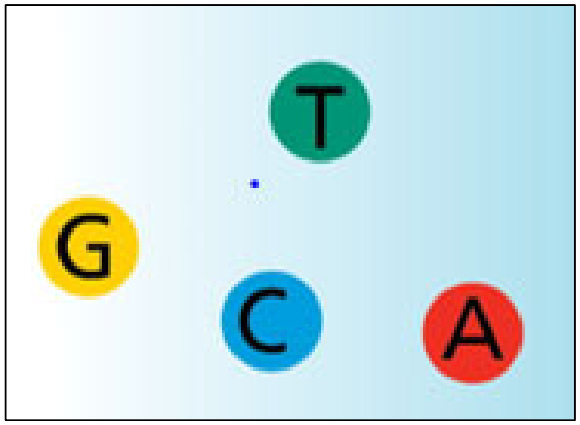


DNA is amplified with the net result that clusters of cloned fragments are fixed in distinct areas of the array or on separate beads.

Next Generation Sequencing

- Primers and fluorescently-labelled terminators are added to the flowcell.
- The DNA polymerase then binds to the primer and adds the first fluorescently-labelled terminator to the new DNA strand.
- Lasers are passed over the flowcell to activate the fluorescent label on the nucleotide base. This fluorescence is detected by a camera and recorded.
- Each base (G, A, T, C) fluoresces a different color.
- Process continues until millions of clusters have been sequenced.

GATCTCCAGTAC



Quality Management Considerations for NGS

Next Generation Sequencing (NGS) is being **broadly implemented** across clinical as well as state and local public health laboratories.

Incorporating **quality management practices** early in the process aids in the efficient transition of LDTs from R&D to results reporting.

Standardized quality practices and protocols can **help ensure** test results are consistent, accurate, and adhere to regulatory requirements where applicable.

Maximize **resources** and **facilitate** sharing of data, equipment, and protocols.

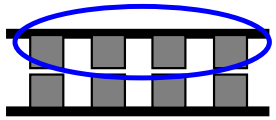
There is a **high risk** associated with the NGS process because resulting data becomes part of the public record and ultimately may be used for research by many people or linked to other data sets.

Quality Management Considerations for NGS Compared to Other Molecular Methods

PCR Quality Control

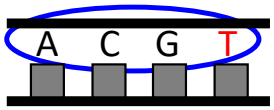
A small selection qPCR QC considerations.

PCR



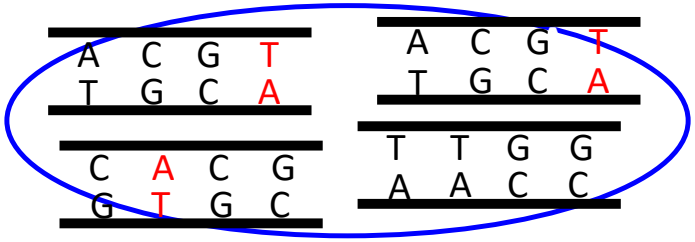
Analysis of sequences (including variants) for specific genomic locations

Sangar

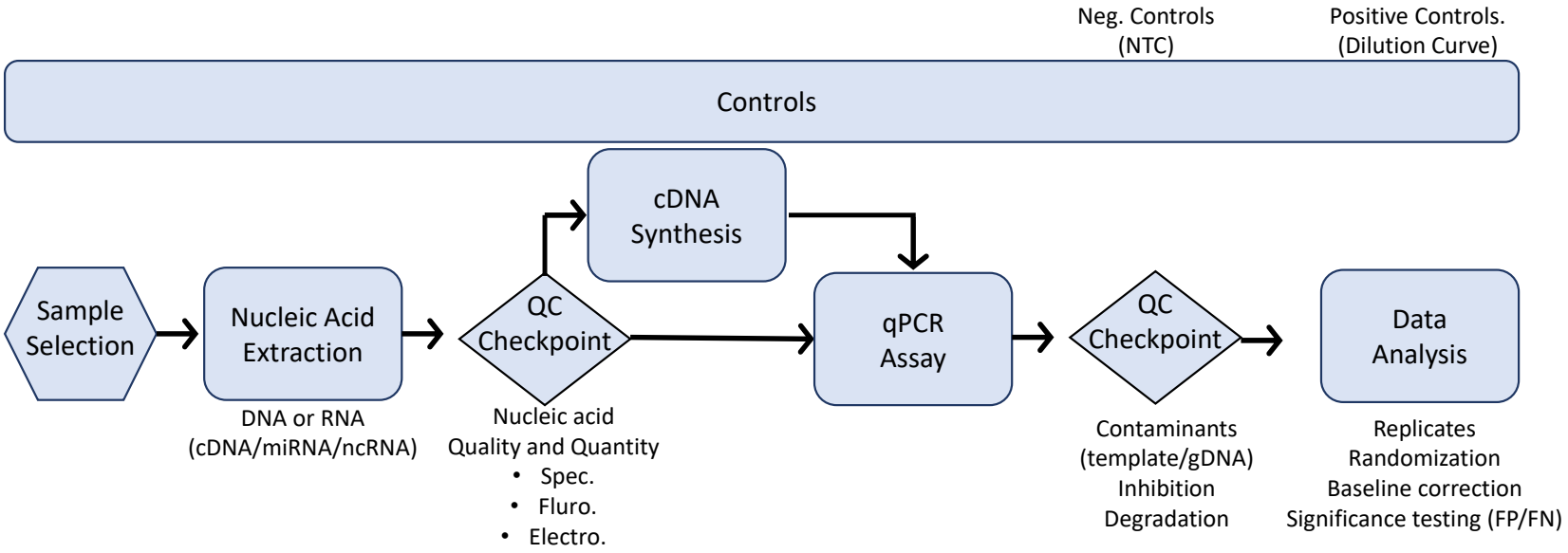


Detect and interrogate of sequences (including variants) of specific genes

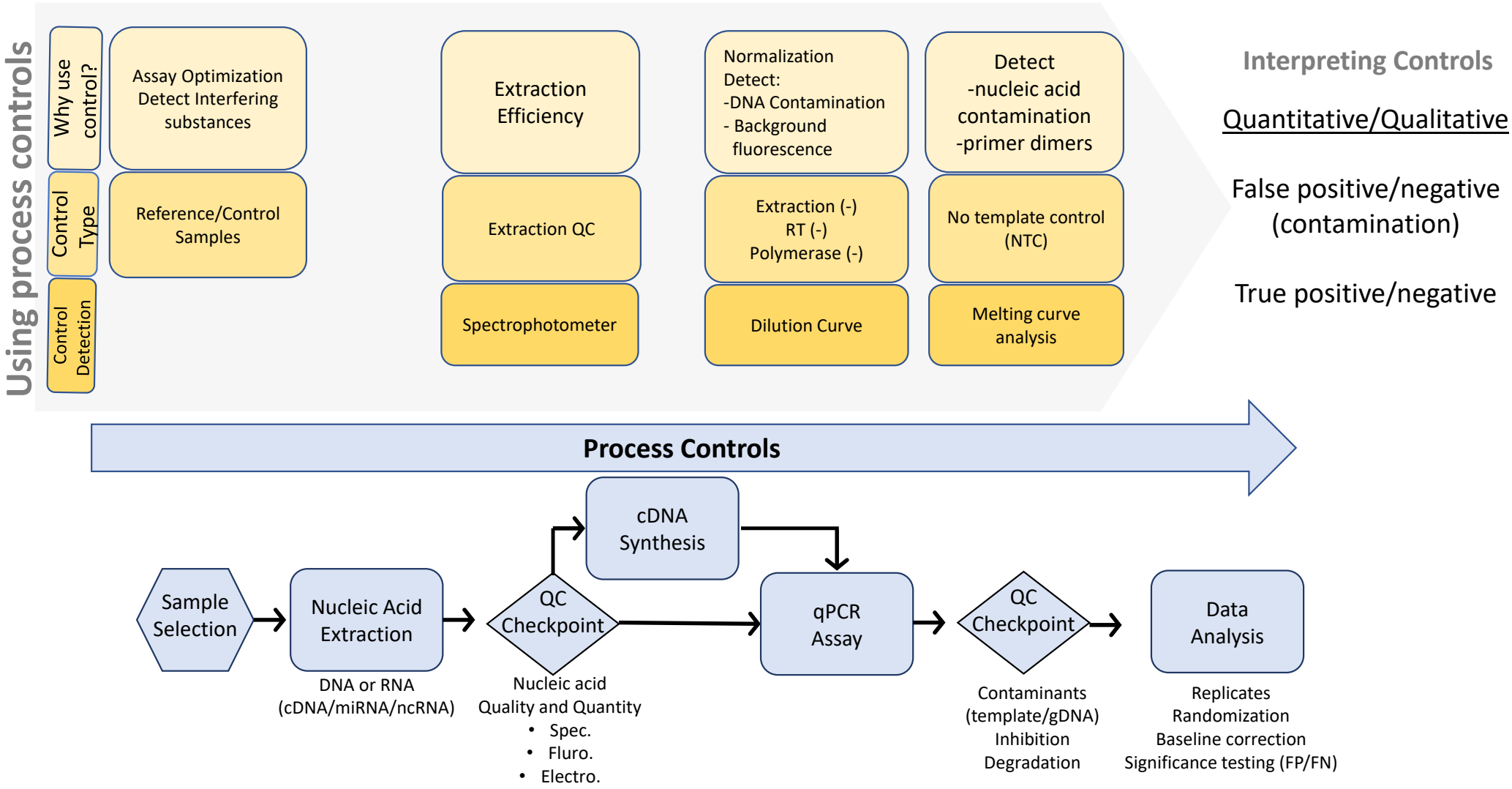
NGS



Detect and interrogate sequences (including variants) of hundreds to thousands of genes



PCR QC considerations – positive and negative controls



Quality Management Considerations for NGS

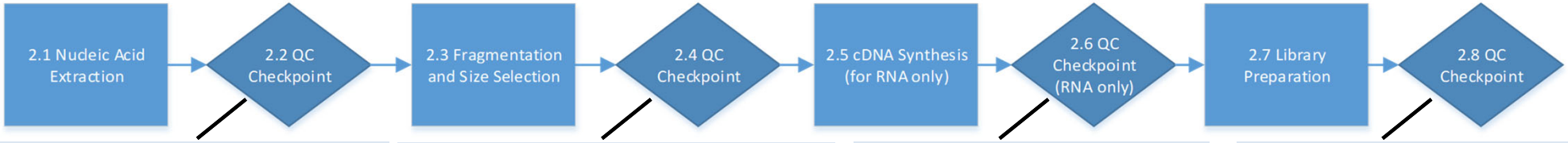
NGS Quality Control

Quality Management Considerations for NGS

NGS Library Prep QC

QC Checkpoints (Library Prep)

QC checkpoints occur between laboratory processes to ensure each step is completed correctly, with high confidence, and to generate quality data metrics that are informative for downstream bioinformatics processes.



Post-Extraction Nucleic Acid

Quantify purity and concentration to ensure sufficient nucleic acid (dsDNA, RNA, or cDNA) and mostly free of contamination.

- Fluorescent measurements
- Spectrophotometric measurement of absorbance ratio (A_{260}/A_{280})

Fragmentation and Size Selection

Information on quality of fragmentation and uniformity. Inefficient sizing can waste capacity and reduce Bioinformatics assembly accuracy.

- Electrophoretic instrument - gel band should reveal a single peak with no tailing and excessive broadening.

cDNA Synthesis

Purity and concentration should be quantified after extraction to ensure successful extraction, as well as prior to use if not used immediately.

- Fluorescent measurements using dye to measure dsDNA.

Library Preparation

Quantify libraries prior to pooling and loading into sequencer.

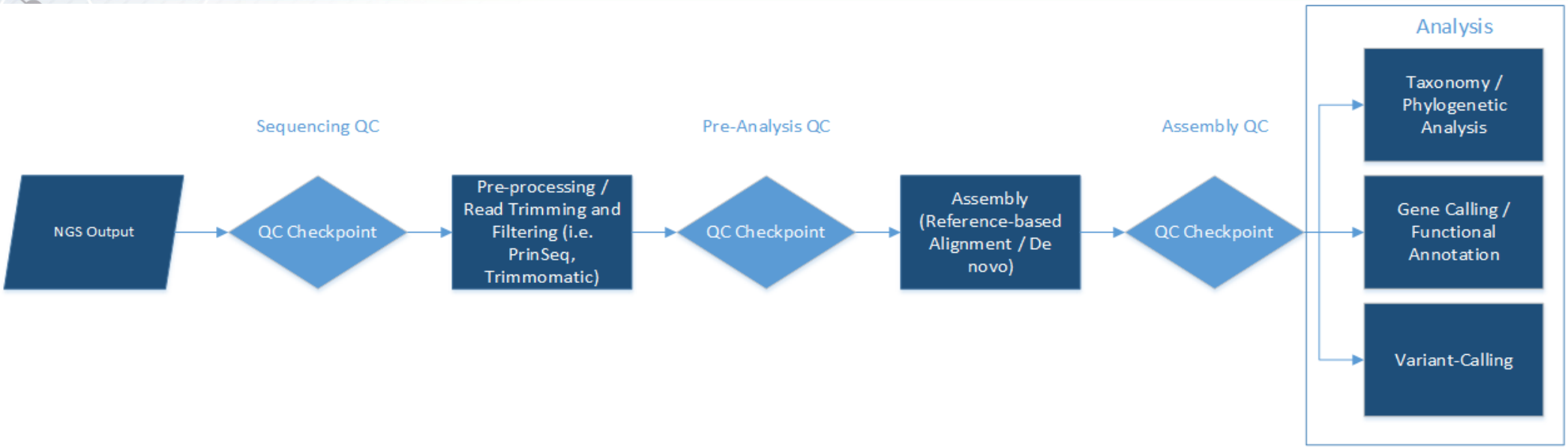
- Optimum cluster density.
- Fluorometric measurement or using real-time PCR (e.g., KAPA qPCR

NGS QC Guidance for Illumina Workflows, CDC NGS Quality Workgroup

Quality Management Considerations for NGS

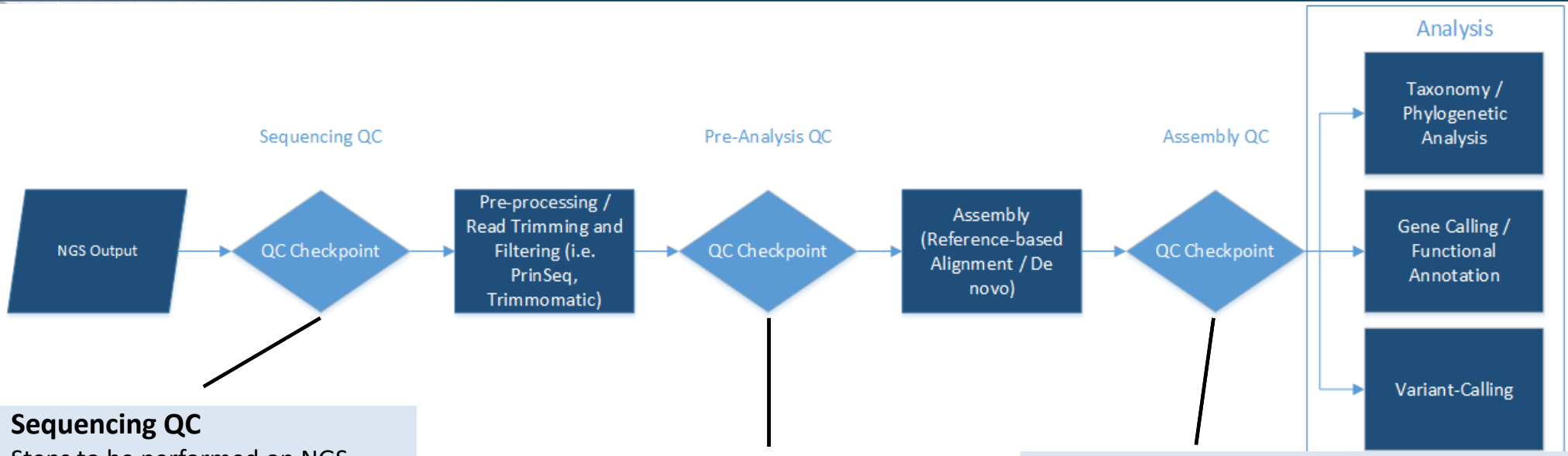
Bioinformatics QC

QC Checkpoints (Bioinformatics)



QC checkpoints are necessary at several stages of bioinformatics analysis including evaluation of run metrics, filtering of raw sequences, alignment/assembly, and characterization stages. These steps ensure the sequence data meets standards for analysis, allows removal of low-quality reads, and reduces false negatives and positives.

QC Checkpoints (Bioinformatics)



Sequencing QC
Steps to be performed on NGS data after it comes off the sequencing instrument and before Pre-Analysis QC.

Pre-Analysis QC
This stage of QC comes after FASTQ file generation.

Alignment and Assembly
Overlapping reads are aligned to create contigs. Assembly joins reads that overlap into contigs. This is controlled by establishing minimum coverage, **N50**, **L50** and **minimum length of contigs** cutoff values. N50 describes a contig length whereas L50 describes a number of contigs.

NGS Sequencing Controls Overview

Control Type	Frequency	What does it control
External controls	Each run	Test system
QC checkpoints	Each sample on each run	Sample
Bioinformatic controls - Sequencing, Pre-analysis, Assembly - Archived sequence data - In silico controls	- Each sample on each run - After pipeline modifications - As needed during analysis	Sample and test system

External Controls

- **What it is**
  - Positive and negative control strains prepared and analyzed alongside test samples.
- **Examples**
  - Previously characterized isolates with and without the marker of interest
  - Precision strain(s) from method validation
- **What does it monitor?**
  - variation over time mitigating process drift
- **What does pass/fail mean?**

  Pass - indicates the test system is working. No further action required.

  Fail - may indicate the test system is not working as it should. Immediate action for failed controls can be to repeat as test samples on the next sequencing run.
- **Frequency**

  Included on each run
- **Benefit**

  Over time, QC metrics from external controls can provide supporting data for the development of an Individualized Quality Control Plan (IQCP) where labs may define the frequency or purpose for using external controls.

QC Checkpoints

- **What it is**
  - QC checkpoints generate data metrics between procedures or processes in the NGS test system that are informative for ensuring each step was completed correctly with high confidence. QC checkpoints are within both the laboratory and bioinformatic (wet and dry) workflows. QC checkpoint data metrics are recorded and evaluated for each individual test sample and external control.
- **Examples**
  - DNA concentration/purity
  - Library insert size
  - Cluster density/Q30
  - Number of contigs/minimum coverage
- **What does it monitor?**
  - Determines if test sample and external controls are of acceptable quality to continue through the process
- **What does pass/fail mean?**
  - Pass – the test sample/external control met the acceptance criteria and may proceed to the next step.
  - Fail - the test sample/external control did not meet the acceptance criteria and the result should be evaluated before the sample can proceed to the next step.
- **Frequency**
  - Recorded for each sample on each run
- **Benefit**
  - Data metrics from QC checkpoints provide real-time evaluation of each sample as it moves through NGS test system/high complexity test thus saving resources for the end user.

Bioinformatic Controls

- **What it is**
  - Bioinformatic controls are steps to ensure the sequence data meets standards for analysis. Bioinformatic controls may also be used to verify analysis or the performance of pipelines.
- **Examples**
  - Filtering/trimming of low-quality reads and index sequences
  - Per base quality scores (Q30)
  - Assembly quality scores (N50, L50)
  - Coverage quality (uniformity and depth)
  - In-silico positive and negative controls
  - Raw reads, FASTQ, assemblies
- **What does it monitor?**
  - Sequence quality (sample level performance) and pipeline integrity (test system performance)
- **What does pass/fail mean?**
  - Sample level - ensures only samples with good quality sequence data pass through the test system.
  - Test system -ensures pipelines are reliable and analysis is accurate/reproducible.
- **Frequency**
  - Pre-analysis bioinformatic controls (i.e. filtering/trimming, assessment of quality scores) are used for each sample on each run. Archived sequence data are used after pipeline upgrades before placing the pipeline back into service. In silico controls may be used during analysis to provide confidence in the interpretation of results.
- **Benefit**
  - Bioinformatic controls generate quality data metrics to ensure each computational process is completed correctly and provide a standard for interpretation or verification of output (e.g. NGS results).

The CDC/APHL NGS Quality Initiative

Purpose of this collaborative effort is to **harmonize quality standards** for NGS across public health and provide laboratories with confidence in reported results.

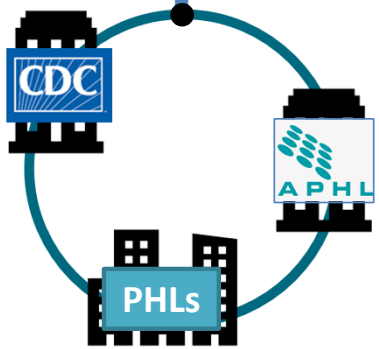
The NGS Quality Initiative is moving forward a coordinated and comprehensive plan to develop and implement an NGS-focused QMS to **ensure high quality, reliable data**.

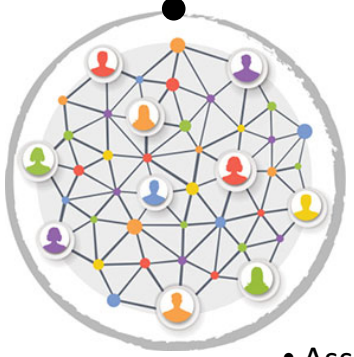
The initiative supports NGS activities to **meet multiple requirements**, including CLIA as a prerequisite for use of this transformative technology in public health laboratory diagnostic activities.

CDC and APHL are coordinating and leveraging existing resources to **create a "toolkit"** to prevent duplication of efforts, increase efficiency, and save costs across internal and external laboratories.

Developing a Quality Management System for Assuring Sequencing Quality

Organizations partner to harmonize quality standards to ensure access to high quality, reliable data



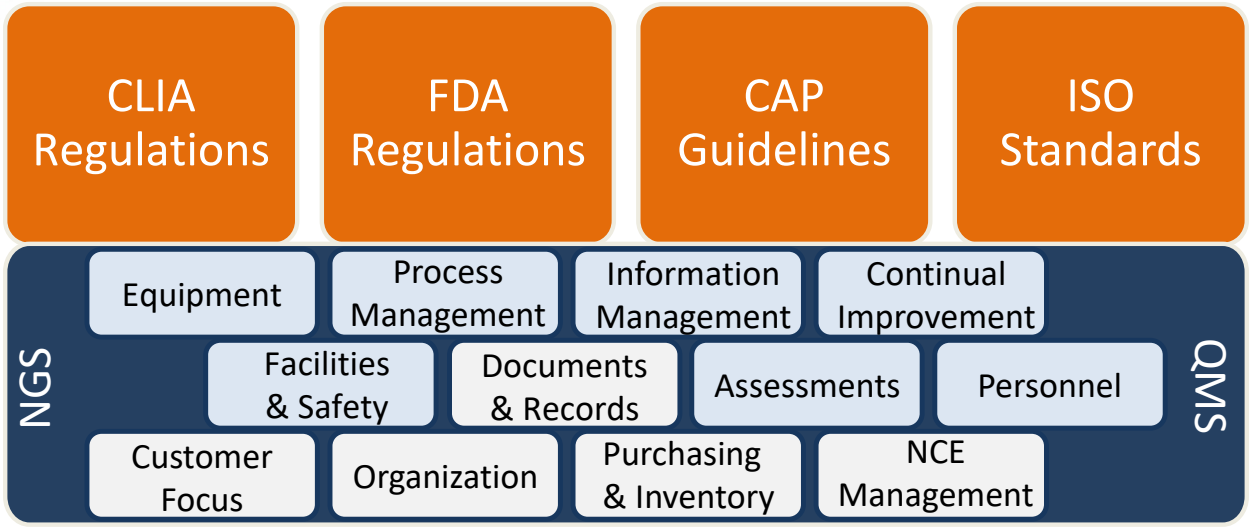


- Centers for Disease Control and Prevention
  - CSELS Division of Laboratory Systems
  - Office of Deputy Director for Infectious Diseases
  - NGS Quality Workgroup
  - NGS Review Board

- Association of Public Health Laboratories (APHL)
  - NGS Subcommittee
  - Laboratory Systems and Standards Committee

- PulseNet Steering Committee

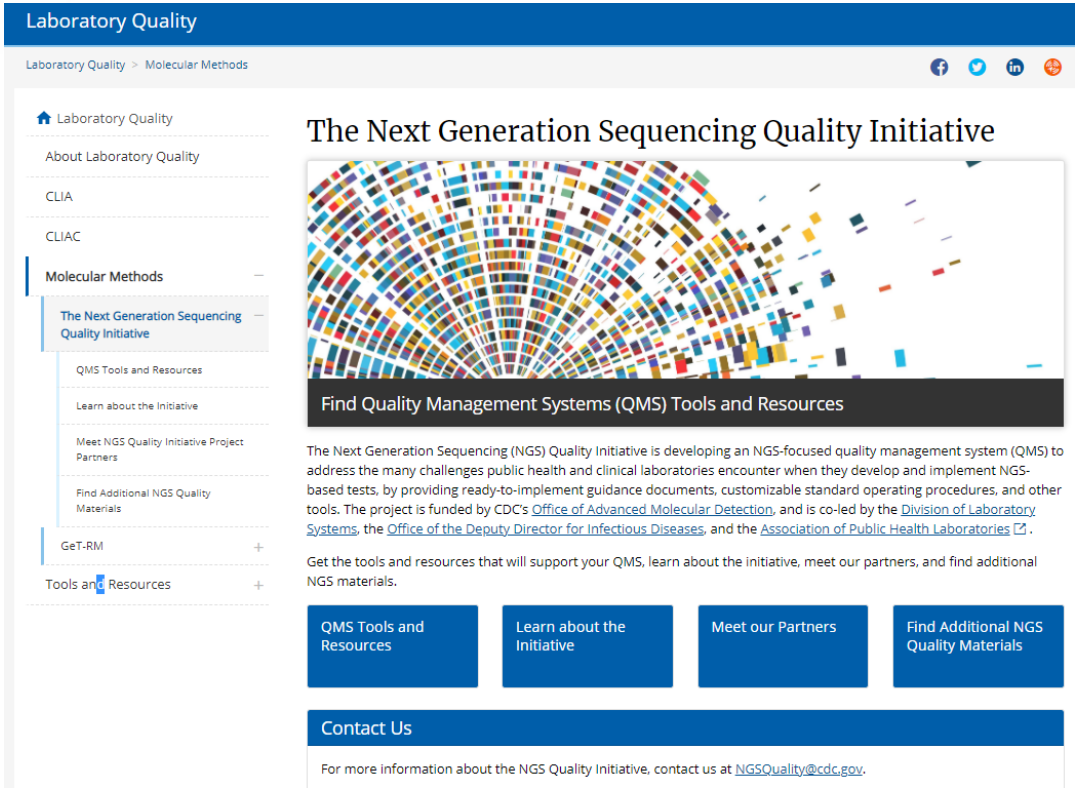
NGS-Focused Quality Management System



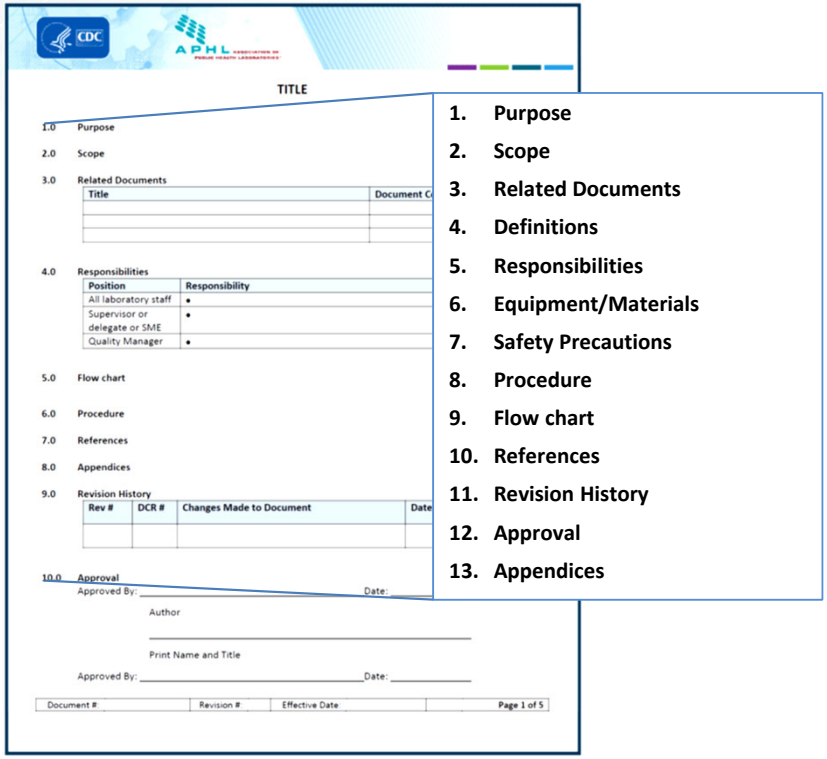
The NGS Quality Initiative's Quality Management System (QMS):

- based on Clinical and Laboratory Standards Institute (CLSI)
- uses 12 building blocks known as the ***Quality System Essentials (QSEs)***

The CDC and APHL NGS Quality Initiative



NGS Quality Initiative Products



~85 ready to implement

Customizable products include:

- Standard Operating Procedures
- Guidance Documents
- Flowcharts
- Templates
- Forms
- Tools
- Logs
- Instructional Videos

NGS Quality Initiative Guidance and Tools: https://www.cdc.gov/labquality/ngs-quality-initiative.html

NGS Quality Initiative Quality Management System (QMS) Assessment Tool

Overview and Purpose

- Intended to guide the assessment of compliance with QMS recommendation
- Used the Clinical Laboratory Standards Institute Framework's 12 Quality System Essentials (QSE)
- Objectives:
  - Catalog essential elements for NGS QMS
  - Monitor ongoing compliance
  - Develop chart reports
  - General and NGS-specific QMS focus

Utilizing the QMS Assessment Tool

- CDC Tools and Resources webpage under "Assessments"
  - https://www.cdc.gov/labquality/qms-tools-and-resources.html
  - Excel workbook with embedded formulas and charts
- Tabs in the workbook align with recommended QSE components
- Each tab lists a series of questions related to the QSE
- QSE may be divided into two sections: "General" and "NGS"
- Each question has a weighted score
- A compliance rating is auto-generated based on the selection of "complaint" or "partially compliant" or "non-compliant" in the period
- Each QSE rating can range from 0-100%
- Workbook may be edited to add or revise questions

NGS Quality Initiative QMS Assessment Tool

- Charted average weighted score
  - Green = Compliance
  - Yellow = Partial compliance
  - Red = Non-compliance
- Display options:
  - **General and NGS**
  - General Only
  - NGS Only
- Export as a PDF for record retention

Example: Personnel Management

Maximum sub-element value

	A	B	C	D	E	F	G
1	Max Value Reference	Quality System Essential	Baseline Rating QMS	P1 Performance			P1 Rating
2				Compliant	Partially Compliant	Non-Compliant	
3		**QSE: Personnel Management**					
4		**General**					
5	33.33	Is the effectiveness of training evaluated and documented?		○	◉	○	16.67
6	33.33	Does the lab maintain records of the following documents: job descriptions, training records, and continuing education records?		◉	○	○	33.33
7	33.33	Is personnel competency documented?		◉	○	○	33.33
8		**Subtotal**	**100**	2	1	0	83.33
9		**NGS related**					
10	25	Has personnel received Human Subjects, HIPPA, scientific integrity training?		◉	○	○	25
11	25	Is personnel competence on NGS methods assessed at least annually?		◉	○	○	25
12	25	Is NGS training documented and competence assessed after training?		◉	○	○	25
13	25	Are personnel encouaged to engage in NGS continual learning activities?		○	◉	○	12.5
14		**Subtotal**	**100**	3	1	0	87.5
15		**Total**	100%	5	2	0	85%
16							

... | Organization | Customer Focus | Facilities and Safety Mgmt | **Personnel Mgmt** | Supply and Inventory

General weighted score

NGS weighted score

Overall score

Example: Personnel Management

B	C	D	E	F	G
Quality System Essential	Baseline Rating QMS	P1 Performance			P1 Rating
		Compliant	Partially Compliant	Non-Compliant	
QSE: Personnel Management					
General					
Is the effectiveness of training evaluated and documented?		○	◉	○	16.67
Does the lab maintain records of the following documents: job descriptions, training records, and continuing education records?		◉	○	○	33.33
Is personnel competency documented?		◉	○	○	33.33
Subtotal	**100**	2	1	0	83.33
NGS related					
Has personnel received Human Subjects, HIPPA, scientific integrity training?		◉	○	○	25
Is personnel competence on NGS methods assessed at least annually?		◉	○	○	25
Is NGS training documented and competence assessed after training?		◉	○	○	25
Are personnel encouaged to engage in NGS continual learning activities?		○	◉	○	12.5
Subtotal	**100**	3	1	0	87.5
Total	100%	5	2	0	85%

)rganization | Customer Focus | Facilities and Safety Mgmt | **Personnel Mgmt** | Supply and Inventory

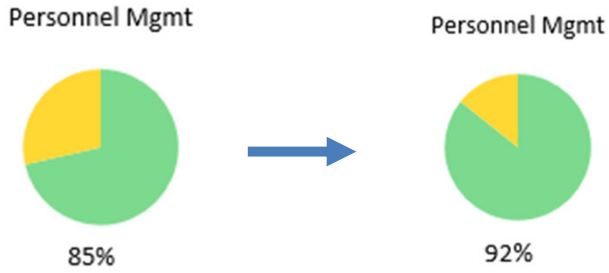
Element for further development
- Performance
- Knowledge
- Blind Samples

Element for further development
- List of available resources
- Work plan for professional development

Example: Personnel Management

- Ongoing compliance assessment for personnel management
  - Between period 1 (P1) and period 2 (P2) the lab addressed existing gaps
    - Progressive work plan for NGS continuing education

	NGS related	
25	Has personnel received Human Subjects, HIPPA, scientific integrity training?	
25	Is personnel competence on NGS methods assessed at least annually?	
25	Is NGS training documented and competence assessed after training?	
25	Are personnel encouaged to engage in NGS continual learning activities?	
	Subtotal	**100**



NGS Quality Initiative QMS
Inventory Management Tool

Ordering

- View product and supplier details, including cost
- Search by product description or item/catalog number

Receiving

- Capture expiration date, SDS, CoC/CoA, QC*
- Lot notes*
- Location
- Adding to inventory or rejecting

Warnings

- Critical levels
- Expiration dates*

Tracking

- Entering inventory observations
- Lot levels
- Active inventory
- Storage locations

Metrics

- Burn rates
- Shipment times
- Expiration

*compliance critical data or process

NGS Quality Initiative Products and Resources
QMS Tools and Resources | CDC

QSE: Process Management
- Validation template and SOP
- Individualized QC Plan
- Bioinformatics QC
- Illumina workflow QC guidance
- Nanopore workflow QC guidance

QSE: Information Management
- Bioinformatics Pipeline Documentation Management
- Containerization
- Sequence Data Storage and Retention

QSE: Personnel
- Training forms for library preparation and instrument run
- Training forms for bioinformaticians
- Competency Assessments (Laboratory and Bioinformaticians)
- Trainer designation form

QSE: Equipment
- Illumina, Ion, Nanopore technologies
  - Maintenance and installation forms/SOPs

QSE: Assessments
- NGS/QMS Assessment Tool

QSE: Continual Improvement
- NGS Key Performance Indicators SOP and Form

QSE: Purchasing and Inventory
- Inventory Management Tool and User Guide

QSE: Facility and Safety
- Waste Disposal SOPs

Support the NGS Quality Initiative

Interested in participating?

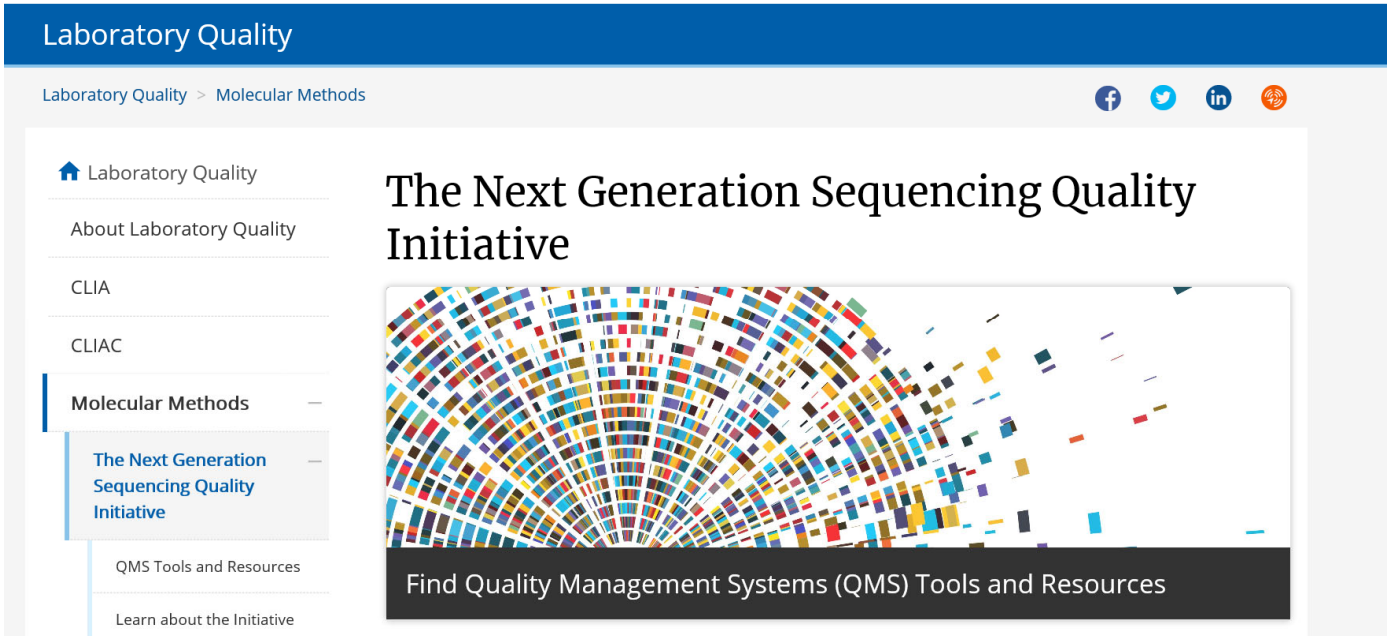
SME input

Meet ~1-3 times per month

Product development
in shared team space

Contact:

- NGSquality@cdc.gov



Thank You

Questions: NGSQuality@cdc.gov

Acknowledgements

- APHL
- CDC Office of Advanced Molecular Detection
- NGS Quality Initiative - Technical Coordinating Committee
- Technical Coordinating Committee workgroups
- CDC Division of Laboratory Systems
  - Heather Stang
- CDC Office of the Deputy Director for Infectious Diseases
  - Atis Muhlenbachs
  - Rebecca Hutchins

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 www.cdc.gov

Use of trade names is for identification only and does not imply endorsement by U.S. Centers for Disease Control and Prevention.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of Centers for Disease Control and Prevention.