Cluster Generation

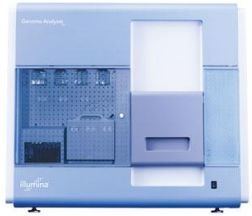
Module 2: Overview

BROAD INSTITUTE illumina

Sequencing Workflow

Sample Preparation

Cluster Generation

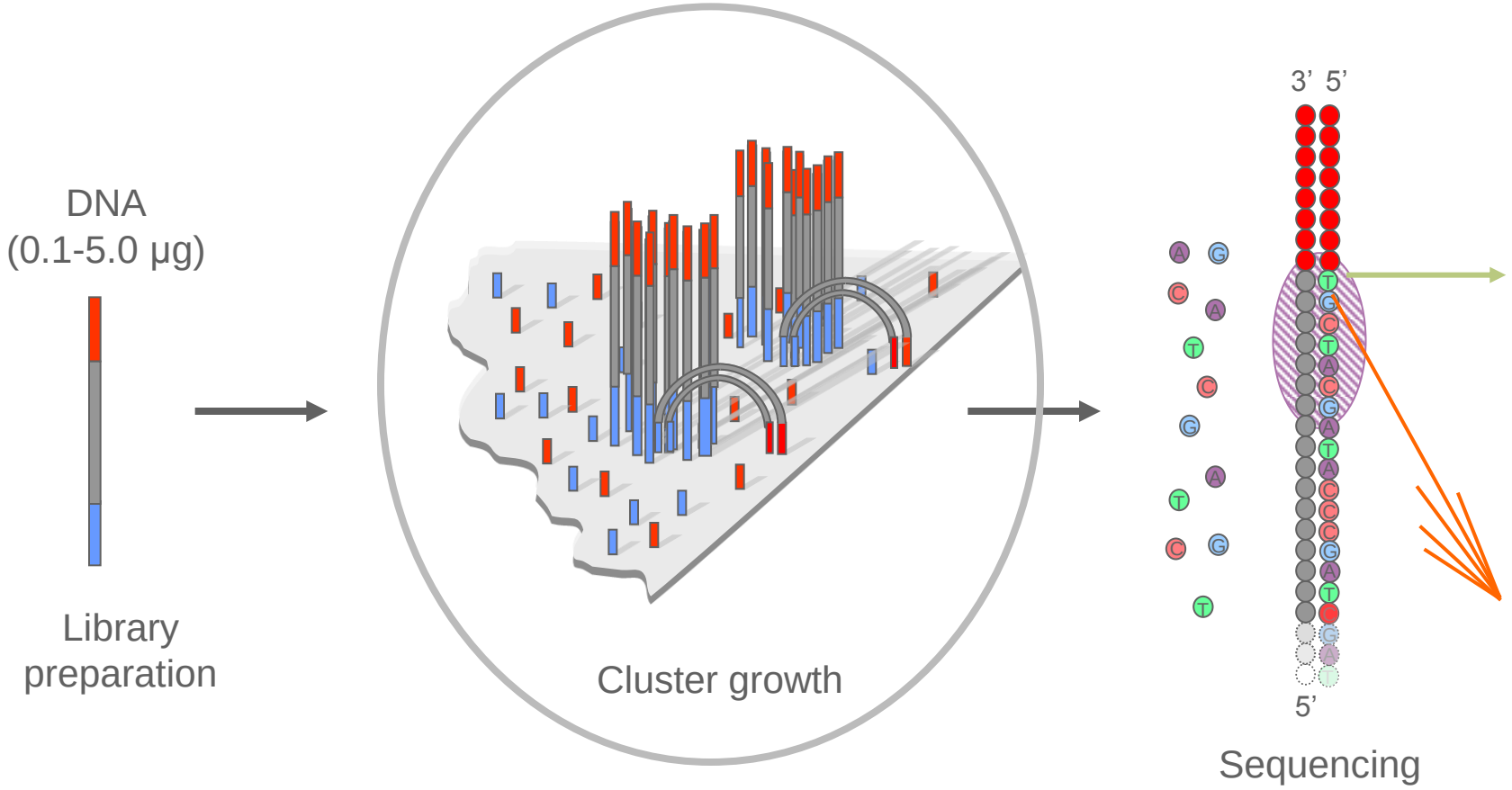


Sequencing

Data Analysis

BROAD INSTITUTE illumina

Cluster Generation



BROAD INSTITUTE illumina®

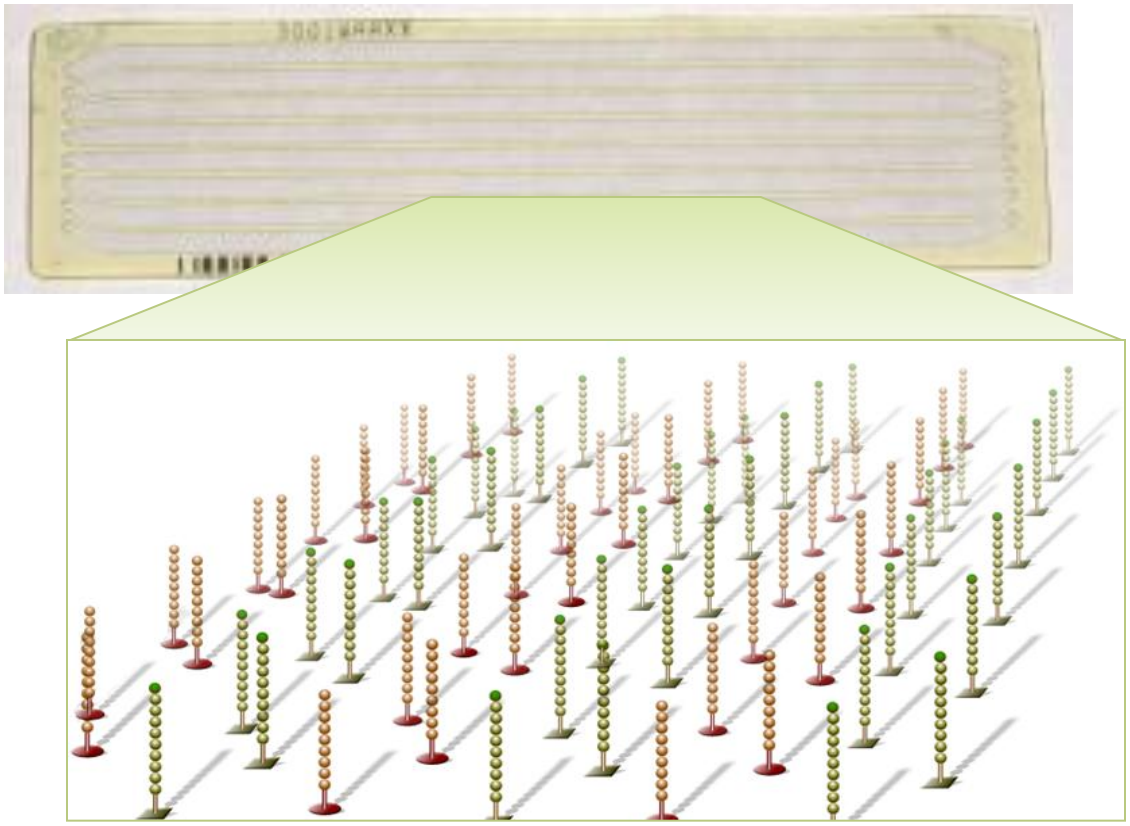
Cluster Generation

Cluster Generation

- Cluster Generation turns libraries into clonal clusters on a flow cell
- All library preparation protocols result in dsDNA
  - Each sample is treated identically
  - Samples are denatured and hybridized to flow cell
  - Captured sample DNA is used as template for second strand synthesis
  - Second strand is amplified into a clonal cluster
  - The cluster is linearized and active sites are blocked
  - Added sequencing primer provides a site for sequencing by synthesis
- Cluster Station/cBot delivers fluidics and controls temperature

BROAD INSTITUTE illumina

Flow Cell



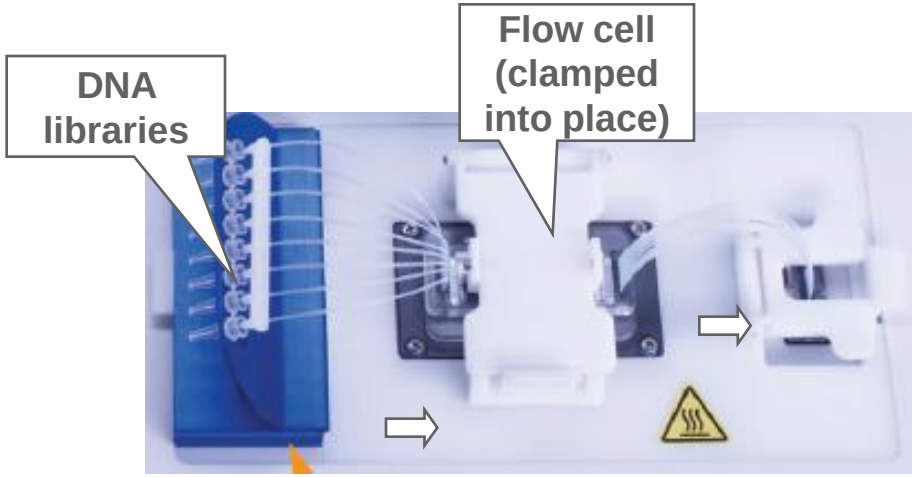
Surface of flow cell coated with a lawn of oligo pairs

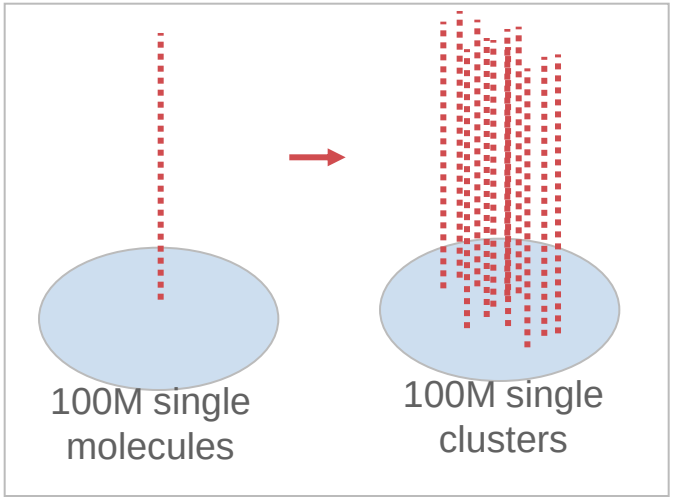
- Simple workflow
- Clusters are generated in a contained environment (no need for clean rooms)
- Sequencing performed on the clusters on the flow cell

BROAD INSTITUTE illumina

Cluster Station/cBot

- Aspirates DNA samples into flow cell
- Automates the formation of amplified clonal clusters from the DNA single molecules





BROAD INSTITUTE illumina

Cluster Generation Workflow

Prepare Reagents

↓

Open and Run a Recipe

↓

Follow Recipe Prompts to Load Flow Cell

↓

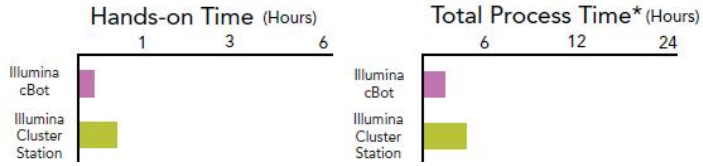
Follow Recipe Prompts to Load Reagents

↓

Cluster Generation Steps
Hybridization, Amplification, Linearization, Blocking, Primer Hybridization

↓

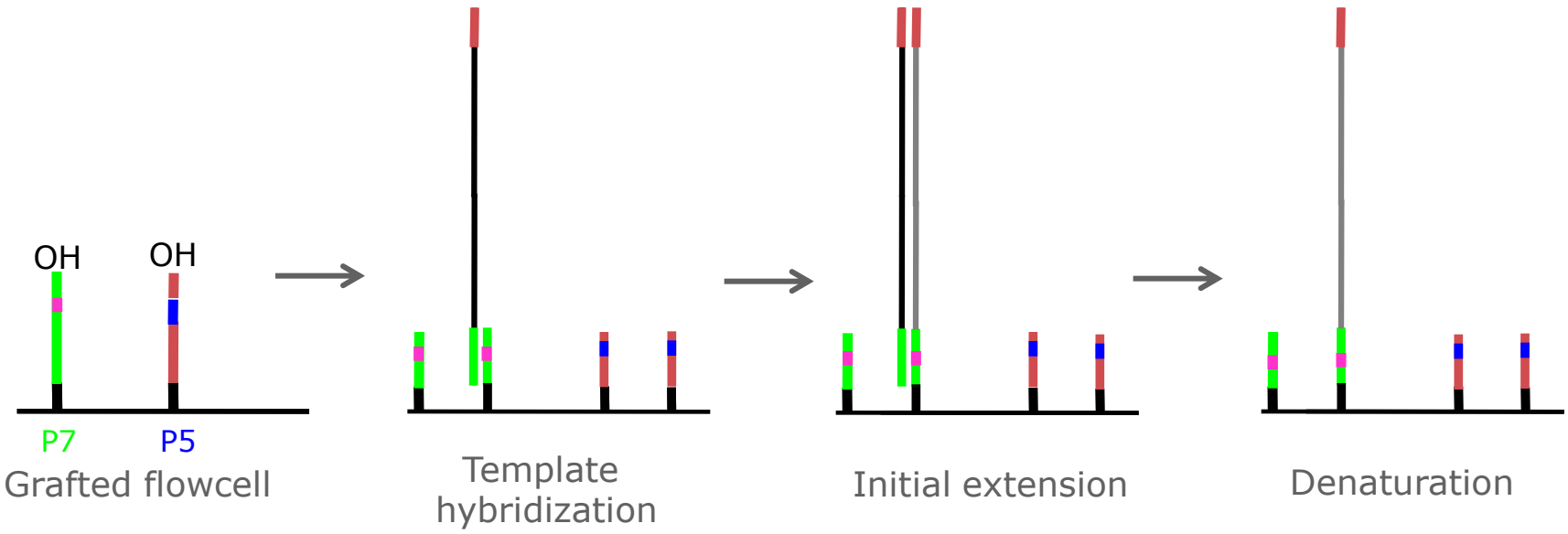
Prepare to Sequence Flow Cell



BROAD INSTITUTE illumina®

Cluster Generation (v4 Chemistry)

Hybridize Fragment & Extend

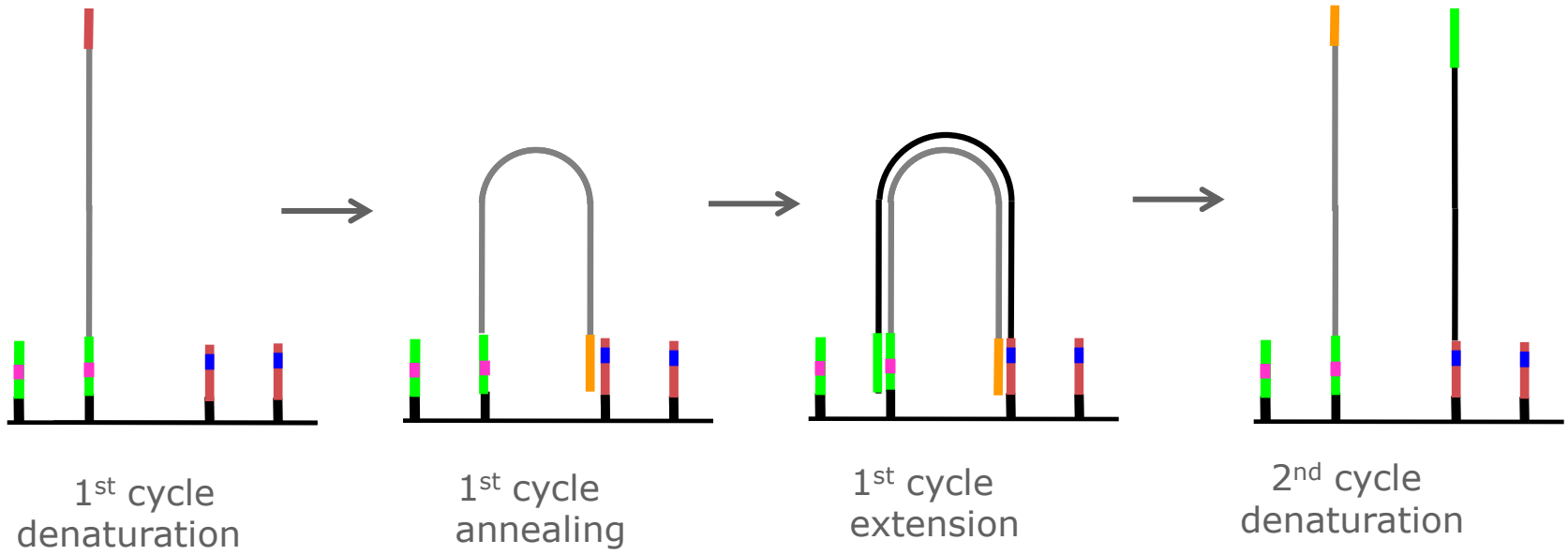


Hyb Manifold	Denatured Library	Amp Premix AMP1	NaOH Wash
Strip Tubes	Temp Ramp: 96-40°C	Phusion HFE 90 sec	Wash Buffer HT2
Hyb Buffer HT1	Wash Buffer HT2	Temp Ramp: 20°C	

BROAD INSTITUTE illumina

Cluster Generation

Bridge Amplification



AMP Manifold

Ramp to 60°C

Formamide AT1

Amp Premix APM1

Amp Mix AMX1

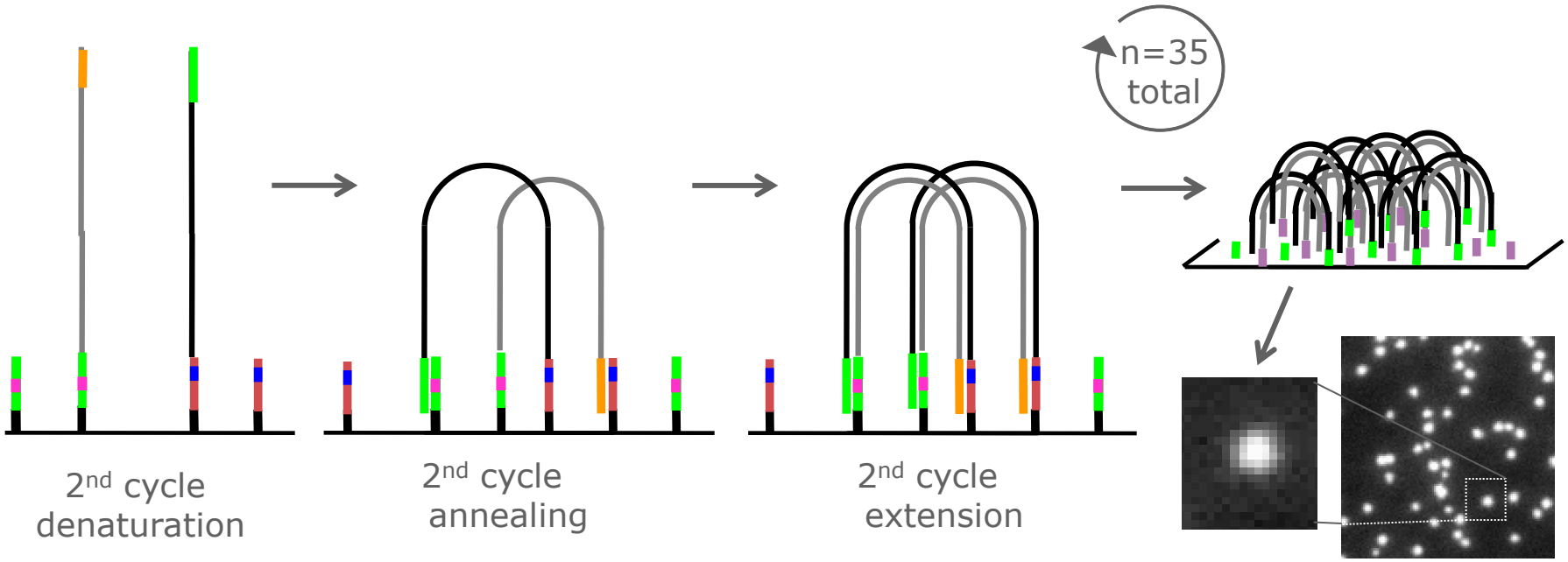
Contains BST pol & nucleotides

Formamide AT1

BROAD INSTITUTE illumina

Cluster Generation

Bridge Amplification



Formamide AT1

Amp Premix APM1

Amp Mix AMX1

Contains BST pol & nucleotides

35 Cycles

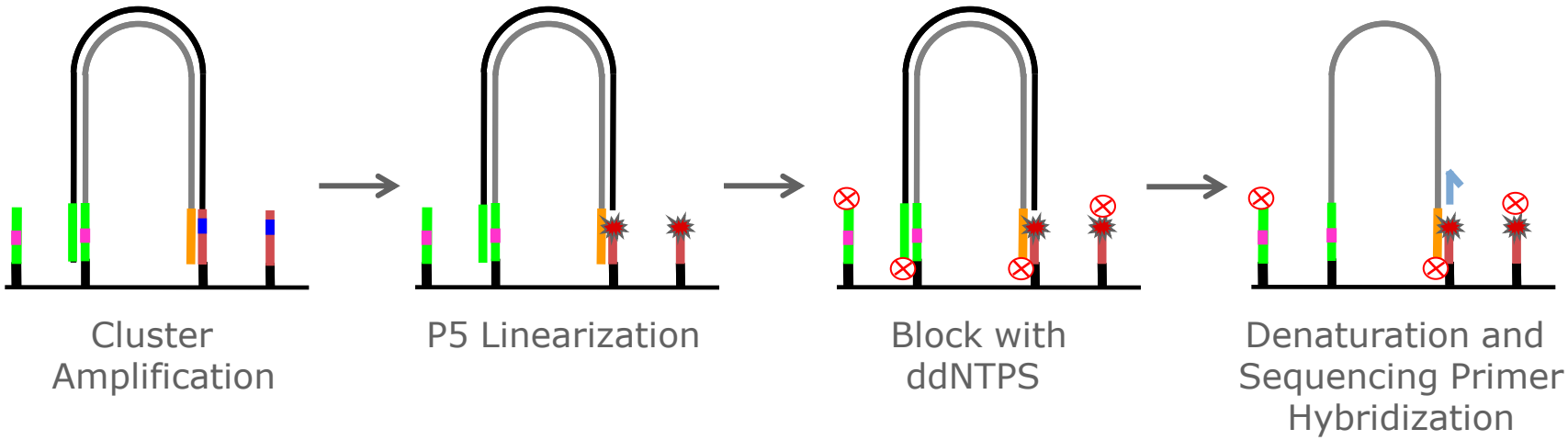
Wash Buffer HT2

Hyb Buffer HT1

BROAD INSTITUTE illumina

Cluster Generation

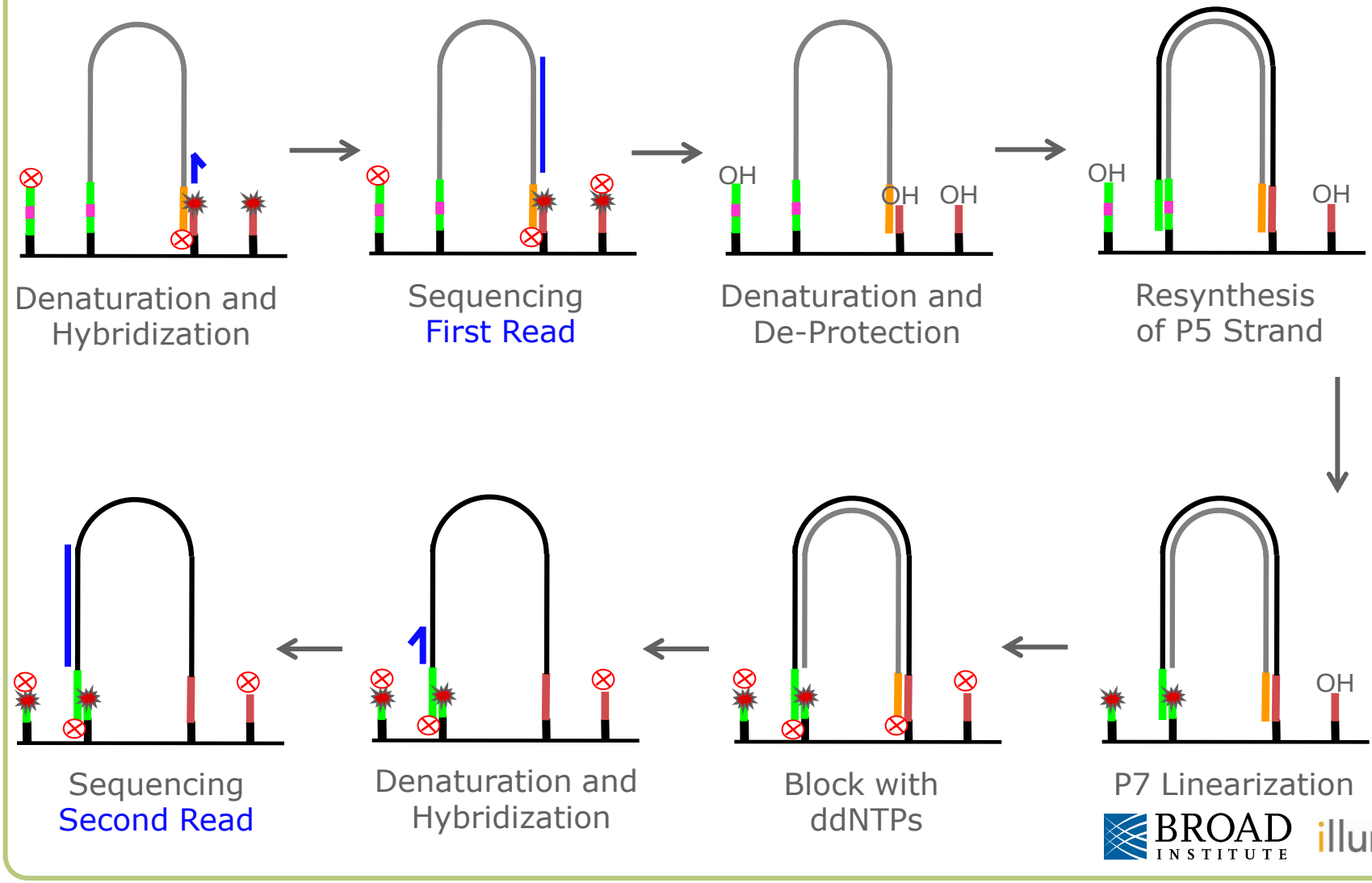
Linearization, Blocking & Sequencing Primer Hybridization



AMP Manifold	PE Linearization LMX1	Blocking Mix BMX	0.1N NaOH
Hyb Buffer HT1	Ramp 37.9°C, 30 min	38°C, 30 min	Seq. Primer
Temp 20°C	Temp Ramp: 20°C	60°C, 15 min	60°C, 5 min
	Wash Buffer HT2	20°C, HT2, HT1 Washes	20°C, HT2, HT1 Washes

Cluster Generation: Read 2 Turnaround

Resynthesis of the second strand



Best Practices

Cluster Generation

- Store reagents as recommended
  - 9 month shelf life from date of manufacture
  - Track lot numbers and barcodes
- Prepare all reagents fresh before use
  - Mix reagents well
- Use proper recipes and remove unnecessary recipes
- Check fluidics delivery both visually and by weight of reagent delivered
- Ensure that Cluster Station is washed regularly

BROAD INSTITUTE illumina

Cluster Generation

Best Practices from The Broad Institute

BROAD INSTITUTE illumina

Cluster Generation

QPCR

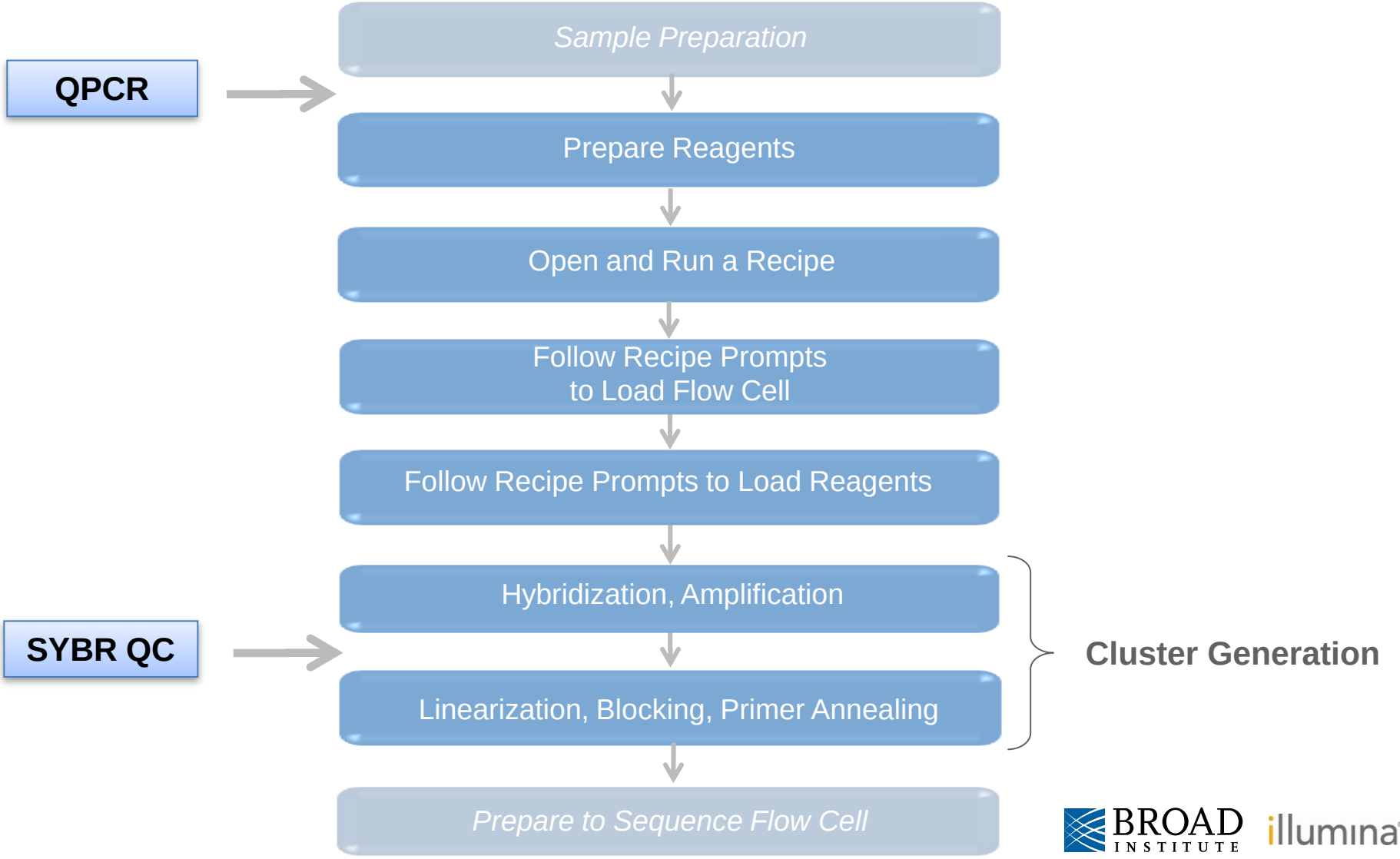
Cluster Amplification

SYBR QC

Linearization & Blocking / Primer Annealing (LB/PA)

BROAD INSTITUTE illumina

Cluster Generation Workflow -- Broad

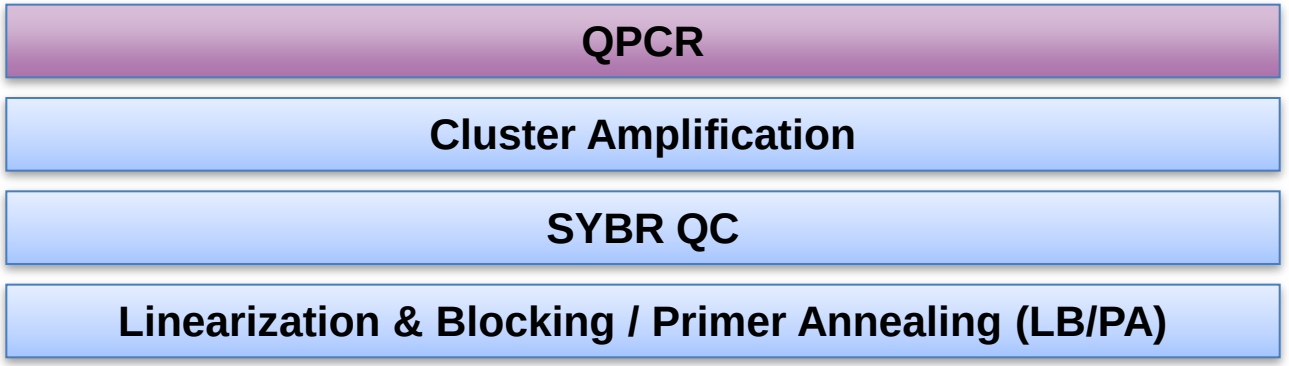


Cluster Generation Overview

- qPCR: Obtaining optimal cluster density
- Addition of Process Controls
- SYBR QC: Ensure successful amplification before continuing
- Lab: Various process details can be monitored and standardized

Quantitative PCR (qPCR)

- **GOAL:** Quantitate amplifiable library fragments in order to obtain optimal cluster density without performing a titration flow cell.

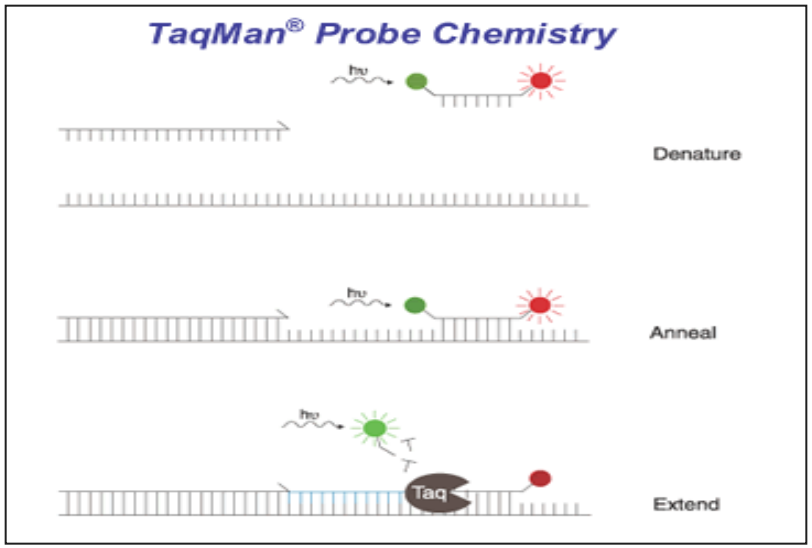


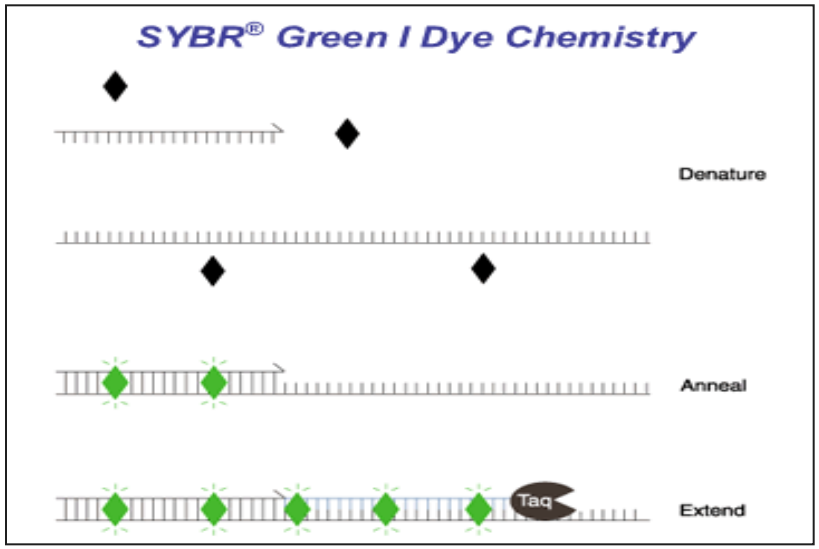
BROAD INSTITUTE illumina

Quantitative PCR (qPCR) Overview

Quantitative Real-Time Polymerase Chain Reaction

- Fluorescence is detected and measured at each cycle
  - Increases proportional to PCR product
- Library's fluorescence is used to determine its concentration



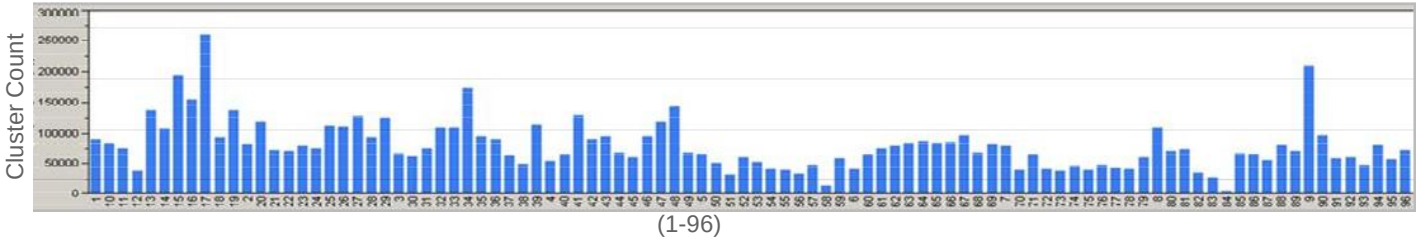


BROAD INSTITUTE illumina®

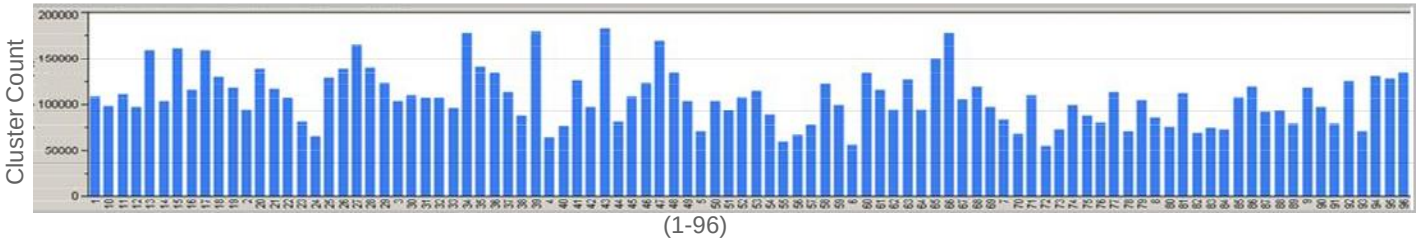
Quantitative PCR (qPCR) Accuracy

- 96 barcoded (indexed) libraries were quantified, normalized, and pooled. qPCR produced more accurate quantitation results, leading to more consistent cluster counts among the libraries.

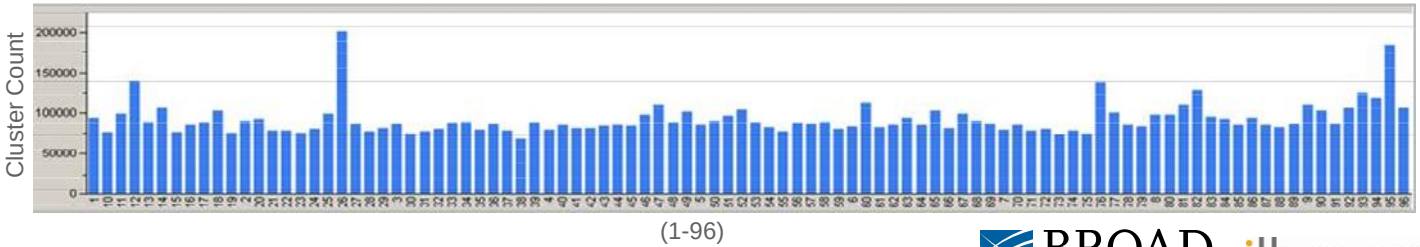
Quant Method:
None (volume)
Fold Difference: 60X
% CV: 50.8%



Quant Method:
Pico
Fold Difference: 3.3X
% CV: 27.4%



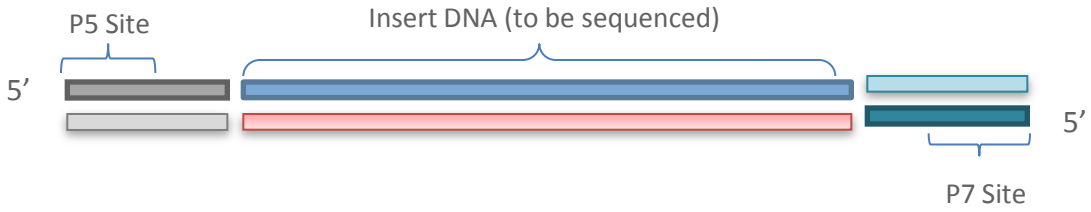
Quant Method:
Automated qPCR
Fold Difference: 3.0X
% CV: 21.4%

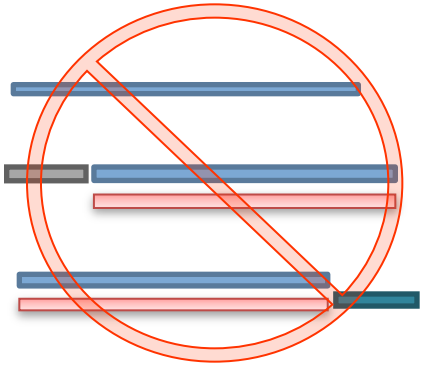


BROAD INSTITUTE illumina

Quantitative PCR (qPCR)

- qPCR is designed to quantify only cluster-forming fragments in the sample

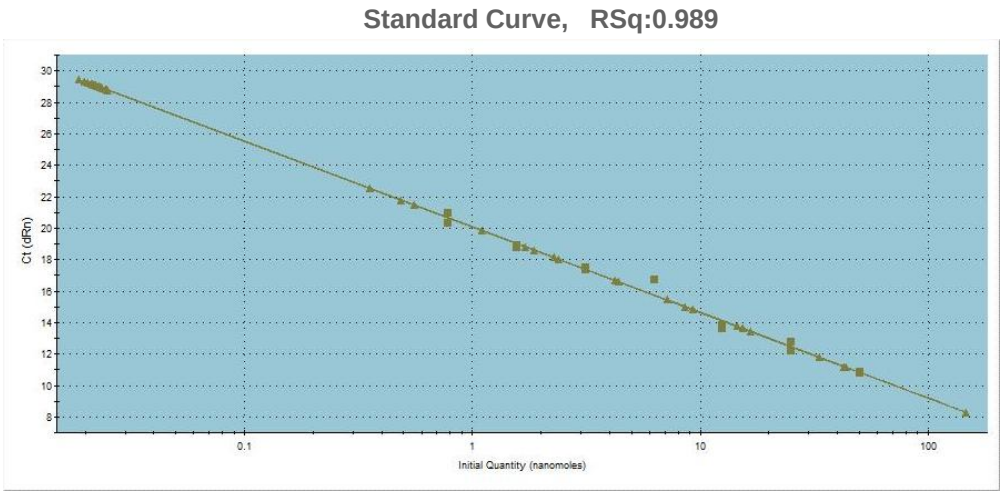




  - Uses P5 & P7 primers to mimic amplification on the flow cell
  - Primers only amplify library fragments with proper adapters

Quantitative PCR (qPCR) Standards

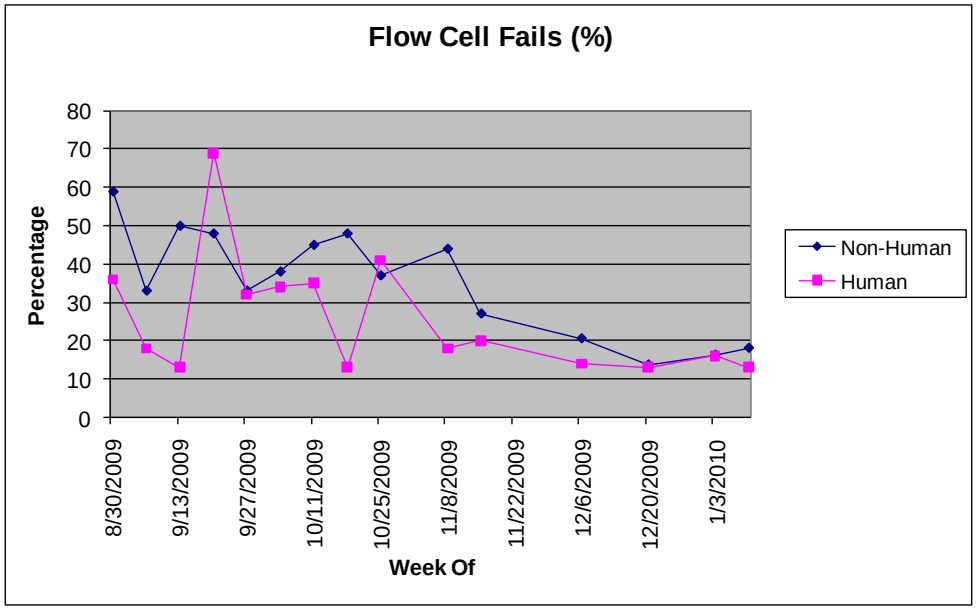
- Using high-quality standards is critical
- Standards' fragment size is critical
  - Ideally, a different set of standards would be used for each library size
  - Library concentration is calculated using size ratio of a monotemplate standard to library* *(*size determined by Agilent assay)*
- We have had success using two types of standards
  - a) Monotemplate (from KAPA)
  - b) Well-characterized genomic library with predictable cluster-forming behavior (PhiX from Illumina)



BROAD INSTITUTE illumina®

KAPA Library Quantification

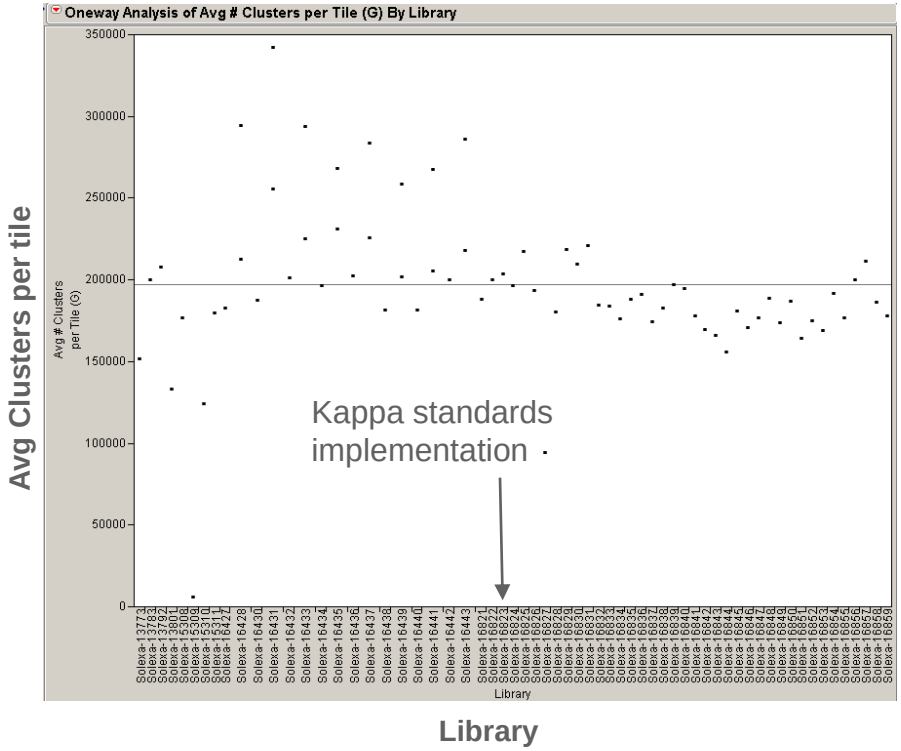
- qPCR-based assay designed specifically for Illumina libraries
- 452bp monotemplate standards
- Multiple different samples 384 well plate = 96 Samples (triplicate)
  - Quants are adjusted to library size: (**452**/average library size)*Quant)
- Reduce flow cell rework rate due to high or low densities
  - (failed flow cell has 3 or more failed lanes)



BROAD INSTITUTE illumina

KAPA Library Quantification

- Reduction in Variabilty
  - KAPA standard qPCR vs. in-house standard qPCR
- Low Yielding Library
  - Pico unable to detect DNA
  - KAPA quant: .3nM
    - Successfully sequenced



Quantitative PCR (qPCR) Standards

- Commercial standards work well at large scale (pre-QCd)
- In-house standards work well at small scale if:
  - Standard sets should be made in large batches
  - QC: Test each new batch against the last batch
  - Ensures accuracy and reproducibility
  - Stored at -80°C

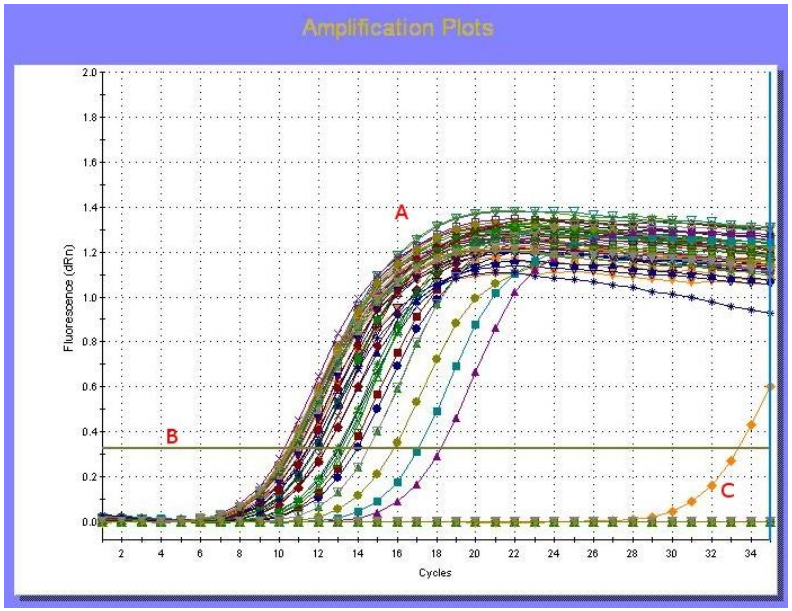
BROAD INSTITUTE illumina

Quantitative PCR (qPCR) Specs

Enrichment Curves

- Flat for at least 4 cycles
- Should increase, then plateau

A. Enrichment Curve
B. Threshold
C. NTC (No Template Control) Enrichment Curve



Threshold

- Should be above background
- Should be below plateau
- Automatically generated

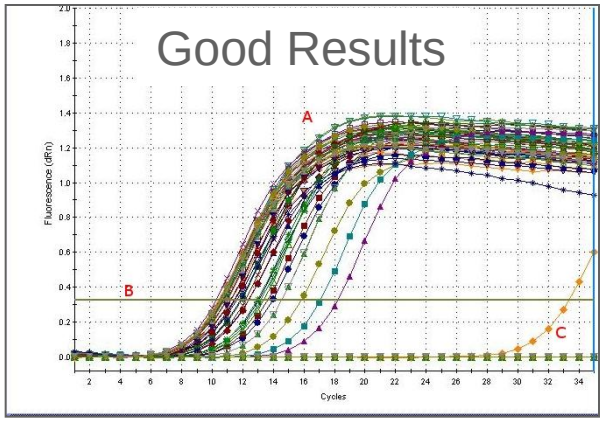
No-Template (Negative)Control

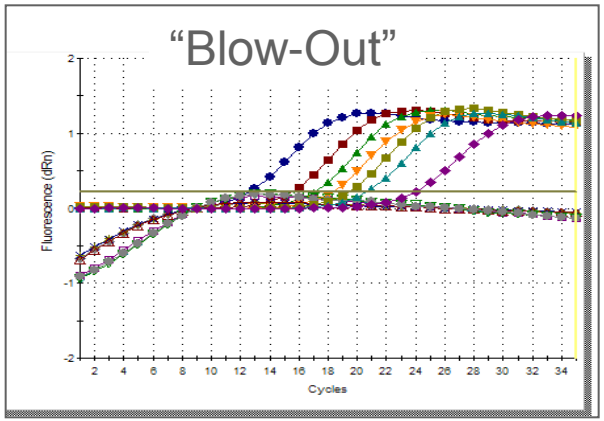
- Displays low level of enrichment, mostly due to primer dimers
- C_t (Cycle threshold) should be at least 10 cycles after template samples.

BROAD INSTITUTE illumina

Quantitative PCR (qPCR) Sample Dilutions

- qPCR requires very low amounts of input DNA
  - "Blow-outs" occur if samples are too concentrated





- Two options for dilution of library
  - Dilute to ~10nM (based on fragment size (Agilent) and rough quant (Pico or Agilent))
  - Perform qPCR on several serial dilutions and ignore blow-outs

BROAD INSTITUTE illumina

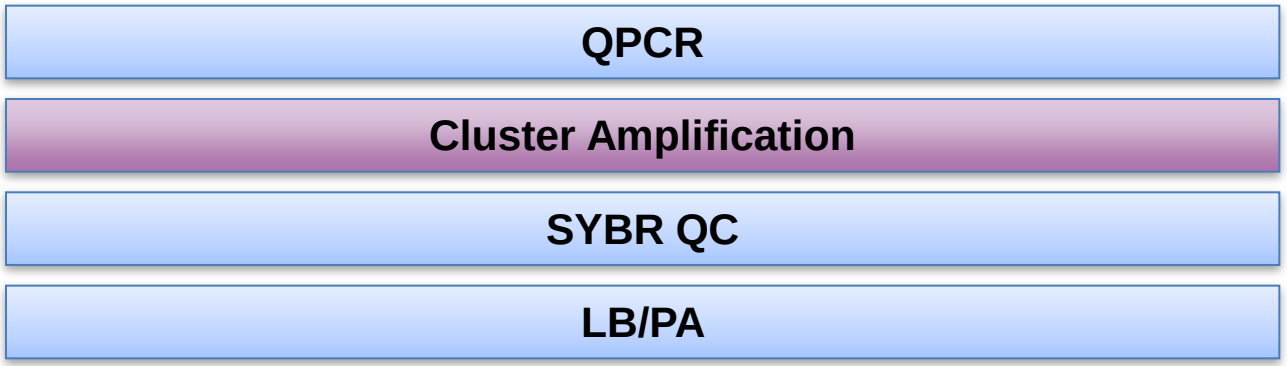
Quantitative PCR (qPCR) Best Practices

- THOROUGH MIXING IS CRITICAL!
  - Sample dilutions (vortex and spin)
  - Vortex mastermix (even though it contains enzyme)
  - Place paper between plate and vortex to keep plate clean
  - Always spin plates before removing caps (to avoid cross-contamination or sample loss)
- UNIFORMITY
  - Use automated deck whenever possible
  - Designate a qPCR user to reduce operator-to-operator variability
  - Make mastermix in a large tube
  - If using a multichannel pipettor, check for uniform volume between tips
  - With any pipettor, check for (and be able to recognize) correct volume levels in tips

BROAD INSTITUTE illumina

Control Samples

- **GOAL:** Detect normal and aberrant variation by including a control sample during cluster generation/amplification



BROAD INSTITUTE illumina

Choosing a Control Sample

- We currently use one of two types of control samples

Phi X Lane	Internal Controls (ICs)
Create a good crosstalk matrix for use in base-biased samples	Regain “real estate” by adding a small amount to each experimental lane (1/100 reads instead of 1/8 reads)
Well characterized small genome	Synthetic sequences require no error-inducing PCR enrichment
~ 50% GC	~ 50% GC
Simple data analysis	Easily identified monotemplates*
Simple workflow (treat as library)	Control in every lane

**Using Broad-designed pipeline software*

BROAD INSTITUTE illumina

Internal Controls (IC) Design

- Randomly-generated, unique synthetic sequences don't align to any genomes (easy to identify)
- The Internal Controls pool contains 4 monotemplate sequences:
  - Pool contains equal base composition at every cycle.
  - Easily-identified, known sequences allow accurate calculation of error rate at any cycle*.

CACGTCACGTGTACGTAGTACGTAC…
TCTAGAGTACAGTGACGAGTGACGA…
GTGCAGTGCGCACATATCACATGCT…
AGATCTCATATCGTCGCTCGTCATG…

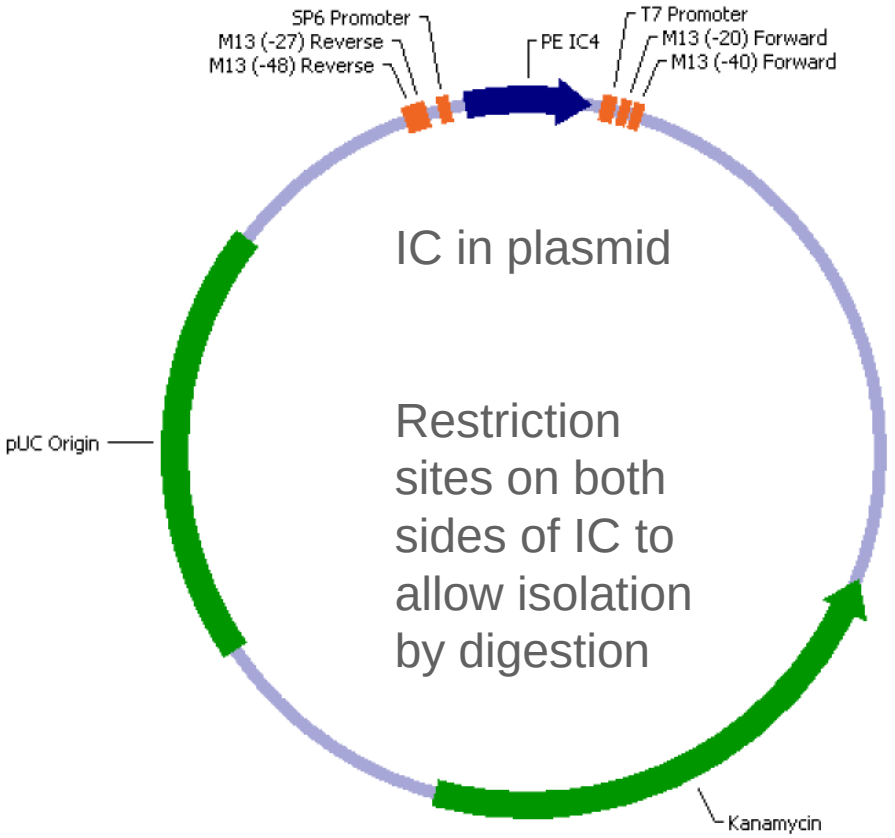
- ICs add a level of lane-to-lane control
  - Control fragments are subject to identical chemistry, imaging, and analysis as each lane's experimental sample.

*Covered in more detail in the Working with Data Module

BROAD INSTITUTE illumina

Internal Controls (IC) Preparation

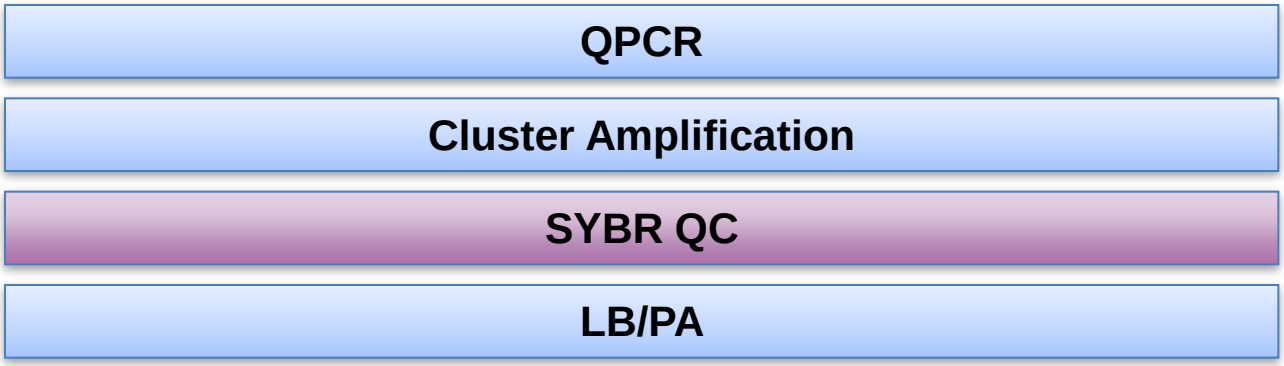
- Cellular reproduction of plasmid (more accurate than PCR amplification)
- IC fragment with final adapter sequence are flanked by unique restriction sites in plasmid
- ICs with adaptors are digested/purified from plasmid
- After final purification, all 4 IC's are quantified (KAPA qPCR) and pooled equimolarly to give even representation of each base at each cycle



BROAD INSTITUTE illumina

SYBR Green Cluster Generation QC

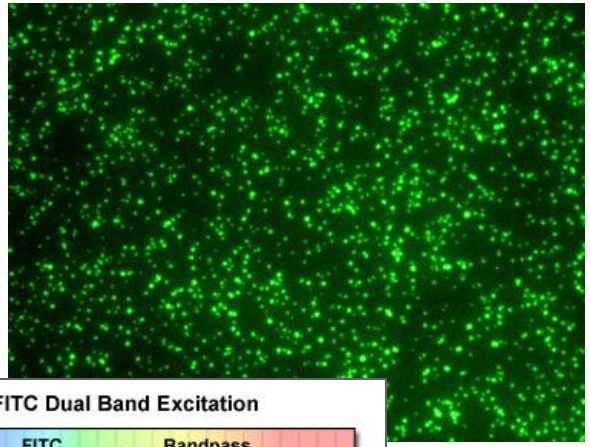
- **GOAL: Visually confirm successful cluster generation and optimal density before continuing with LBPA and sequencing**

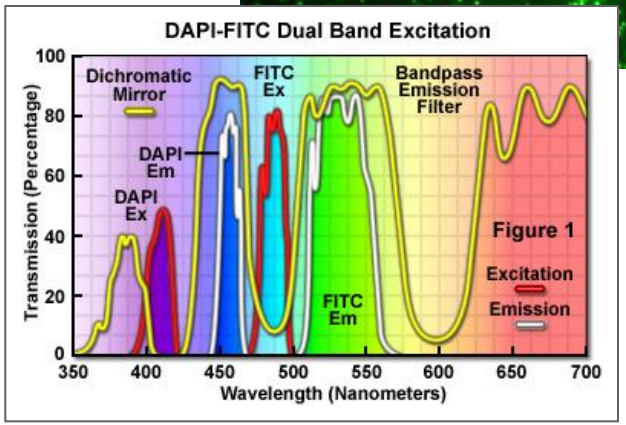


BROAD INSTITUTE illumina

SYBR Green Cluster Generation QC

- SYBR Green dye is run through flow cell after cluster generation is complete.
  - SYBR intercalates into dsDNA
  - May be less mutagenic than ethidium bromide
- Visible clusters using standard microscope
  - Excited by Xenon bulb*
  - Fluoresces @ 500-560λ (green)
  - Visible through FITC filter
  - Clusters visible at 40X magnification
- Will see more in lab module

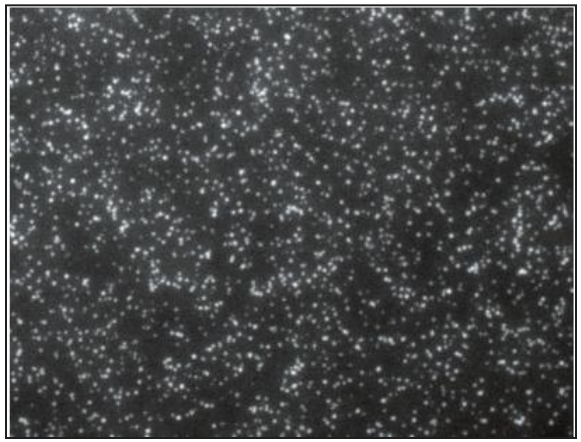




* Xenon bulb must be changed every 350 hours to prevent explosion risk

BROAD INSTITUTE illumina

SYBR Green Cluster Generation QC



Sparse

Good

Dense

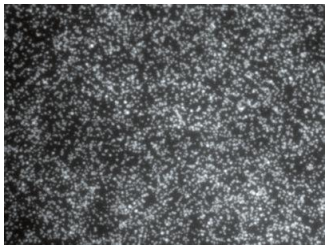
*1.6 RTA

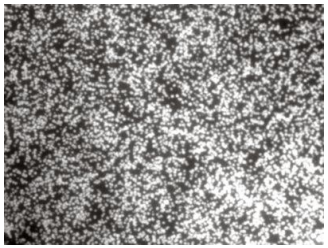
- Visual QC of flow cells: Verify successful amplification and cluster count
  - Too Sparse: Loss of valuable real estate on flow cell
  - Too Dense: Analysis problems
- High density limits ability to accurately count clusters at SYBR step
- Afterward, flow cell can be stored for up to ~1 month

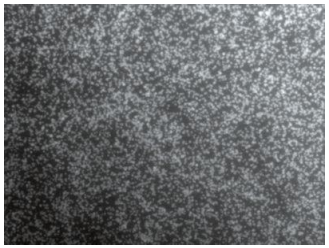
BROAD INSTITUTE illumina

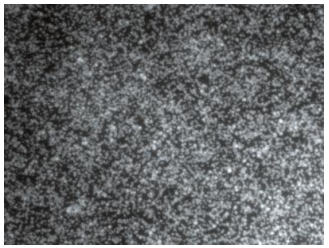
SYBR Green Cluster Generation QC

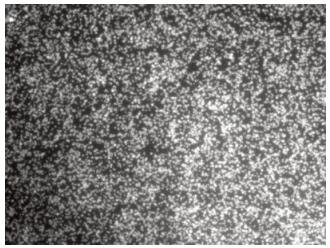
SYBR image → Estimated Cluster #

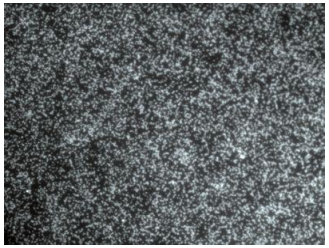

100322

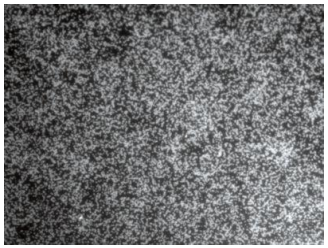

110493

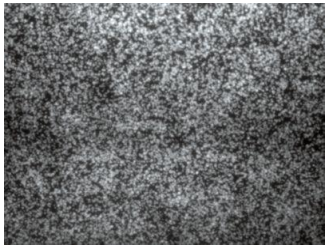

121380

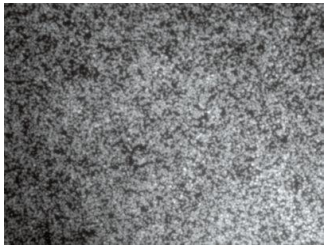

132532

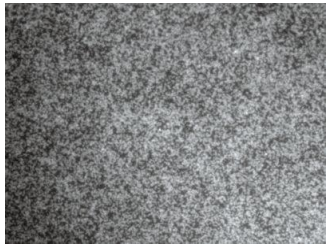

140432


148136


160834


169500


180580


190794

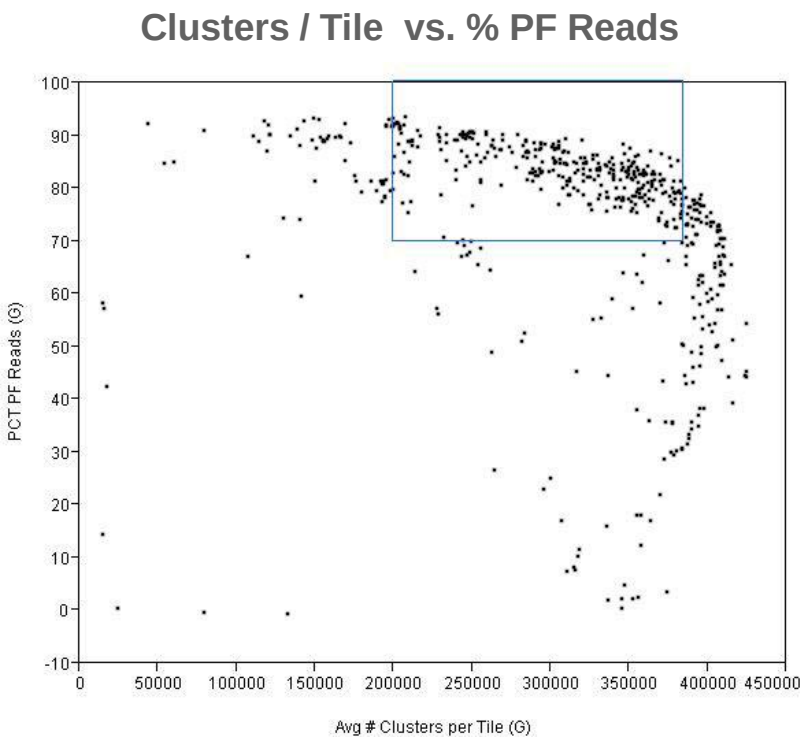

319000


360000

BROAD INSTITUTE illumina

Cluster Densities (1.6 RTA)

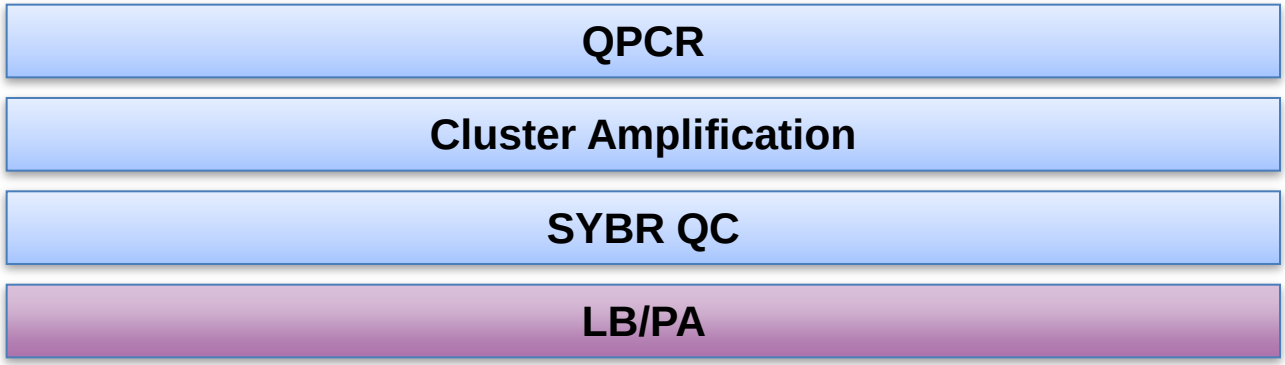
- 1.6 RTA is able to identify many more clusters than previous pipelines
- SYBR QC images are so dense that it is difficult to differentiate between 350,000 and 400,000 clusters/tile
- QPCR accuracy, along with historical data, allow us to pinpoint optimal loading density



BROAD INSTITUTE illumina

LBPA: Linearization, Blocking, Primer Annealing

- Prepare clusters for sequencing
- Use same amplification manifold as in cluster generation
  - Use wash flowcell to run water through manifold
- Ensure proper flow of reagent using Hybridization Buffer
- Immediately place on sequencer



BROAD INSTITUTE illumina

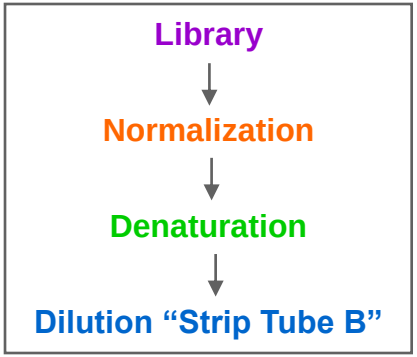
Cluster Generation – Summary

- QPCR
  - Using high quality standards is critical
  - Determines loading concentration for optimal cluster density
- Control Samples
  - Phi X – simple workflow and good for use on flow cells with base biased samples
  - Internal Controls – Control in every lane,
- SYBR QC
  - Determine number of clusters amplified prior to placing on sequencer

BROAD INSTITUTE illumina

Lab Best Practices – Pipetting Variability

- Experiment: Compare variation in cluster counts between lanes containing **the same E.coli genomic DNA sample** (split off at various points in the process):

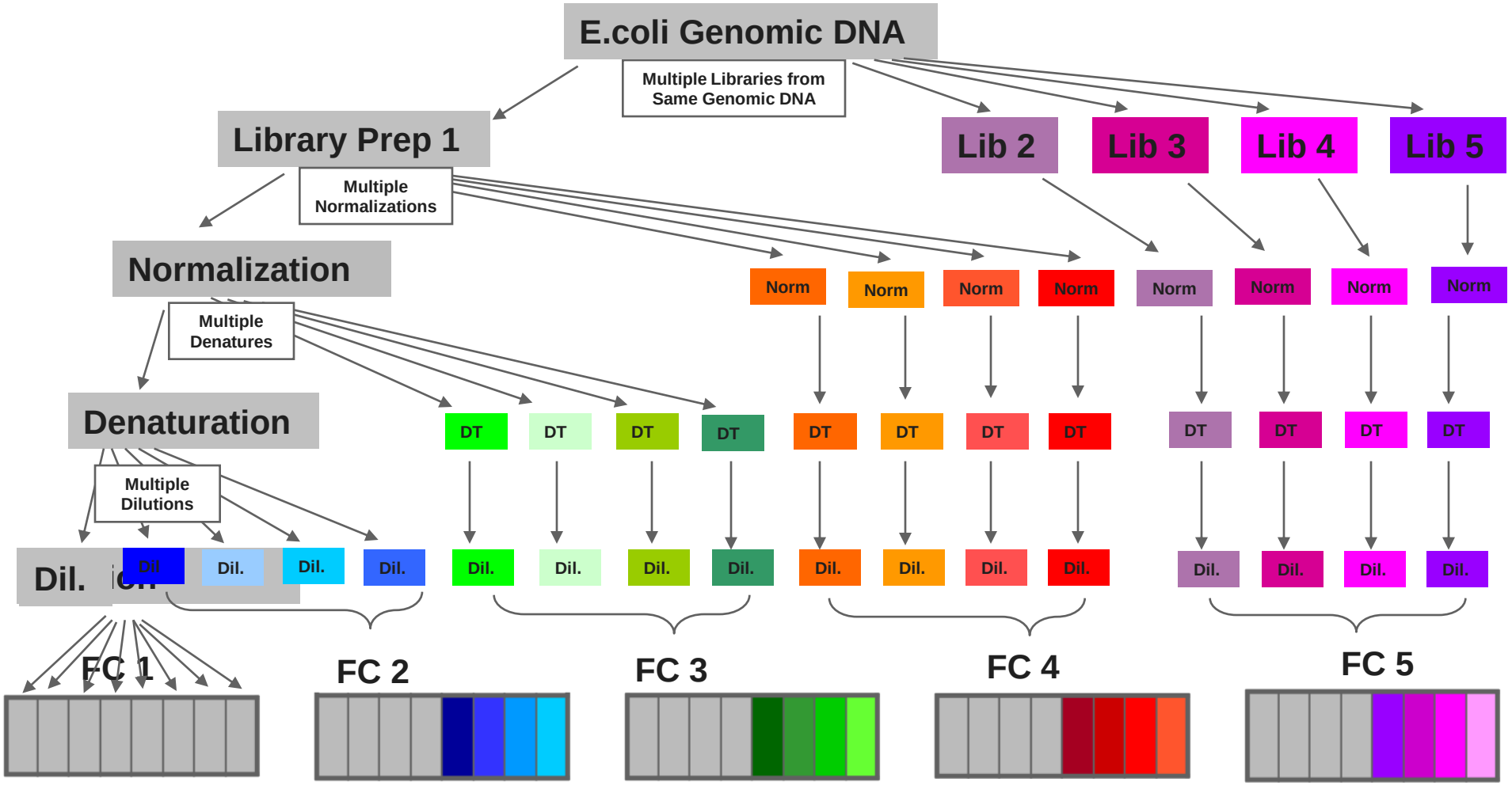


  - Replicate libraries
  - Same library, replicate Normalizations
  - Same library, same normalization, replicate denaturations
  - Same library, same normalization, same denaturation, replicate dilutions
  - Same library, same normalization, same denaturation, same dilution

- Problem confirmed: Small-volume pipetting introduces significant variation at each phase between library prep and cluster generation. Inserting variation earlier in the process increases amount of final variation.

BROAD INSTITUTE illumina

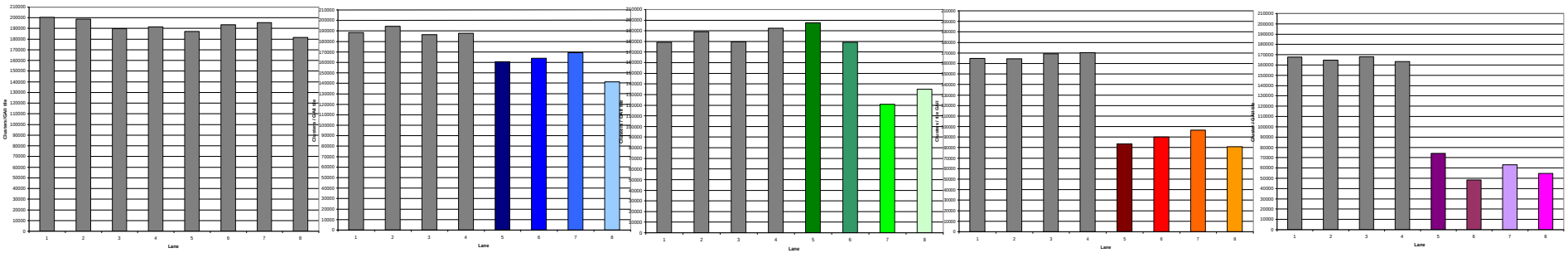
Lab Best Practices – Pipetting Variability

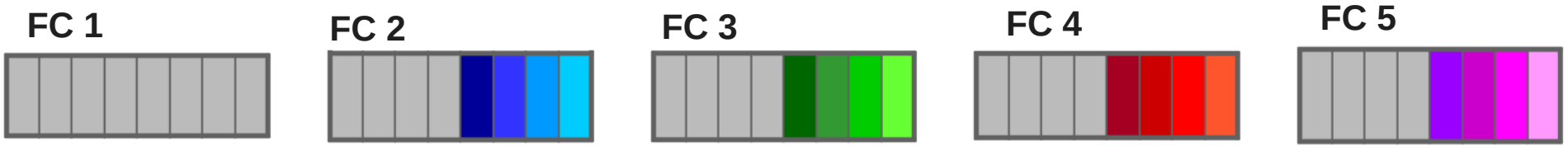


Lab Best Practices – Pipetting Variability

Same Library	Same Library	Same Library	Same Library	**REPLICATE Libraries**
Same Normalization	Same Normalization	Same Normalization	**REPLICATE Normalizations**	
Same Denaturation	Same Denaturation	**REPLICATE Denaturations**		
Same Dilution (Aliquotted into different lanes)	**REPLICATE Dilutions**			

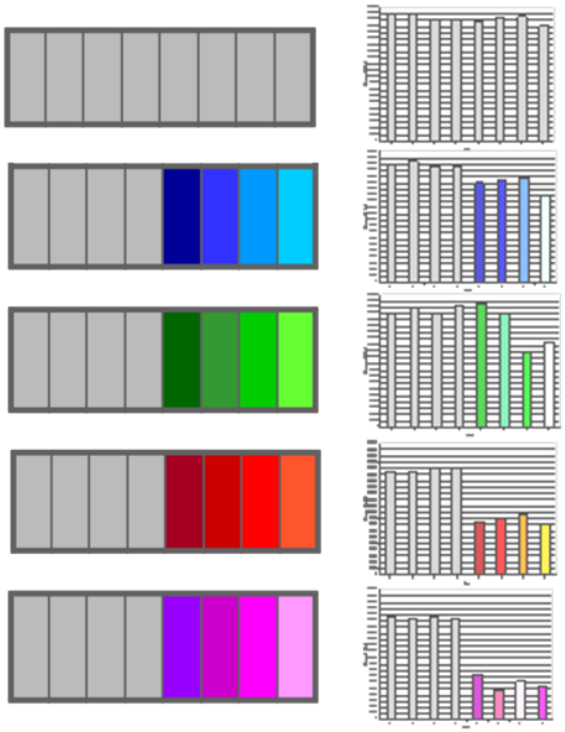
Cluster Count By Lane





BROAD INSTITUTE illumina®

Lab Best Practices – Pipetting Variability



CV across 8 lanes: 3.2%

CV across 8 lanes: 10.5%
CV across 4 lanes: 7.7%
Split @ Dilution

CV across 8 lanes: 16.4%
CV across 4 lanes: 22.9%
Split @ Denature

CV across 8 lanes: 33.5%
CV across 4 lanes: 8.0%
Split @ Normalization

CV across 5 lanes: 50.5%
CV across 4 lanes: 18.6%
Split @ Library Prep

- Inserting variation earlier in the process between library prep and cluster generation increases final variation in cluster counts.
- **To reduce small-volume pipetting over time:**
  - **Use automated qPCR to determine optimal loading concentrations**
  - **Use automated denaturation**
  - **Store denatured libraries for re-use**

BROAD INSTITUTE illumina

Lab Best Practices – Cluster Station Use

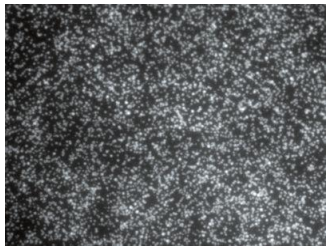
- Reagent Preparation
- Flow Problems
  - Line clogs, Manifold gaskets, Watch Prime, Wiping Ports
- Cluster Station Maintenance
  - Routine Decon Wash
- Tracking – manifold re-use
- Thermal Profile
  - Peltier block failures
- Lab Bench Organization

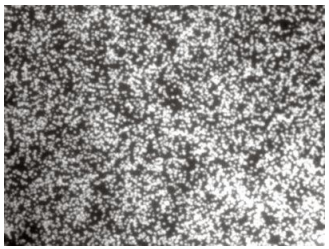
- Reagent and waste volumes are monitored and tracked for future troubleshooting

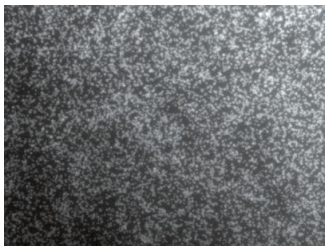
BROAD INSTITUTE illumina

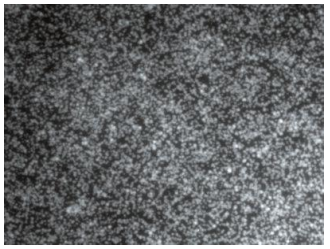
Lab Best Practices – SYBR QC

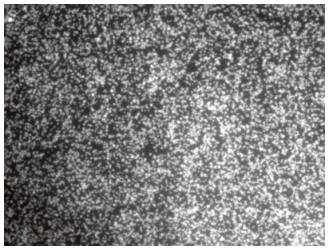
- Visualizing clusters on microscope
- SYBR image → Estimated Cluster # (Pass/Fail with trained eye)

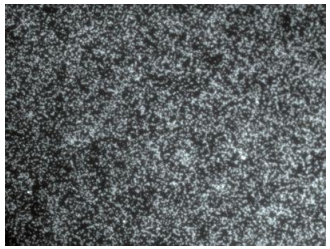

100322

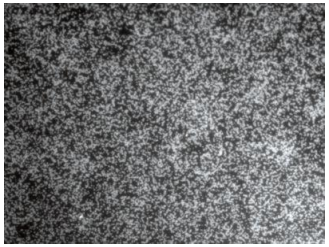

110493

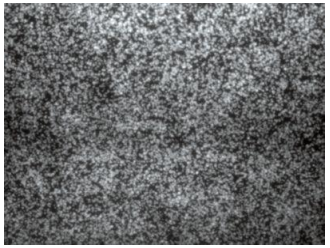

121380

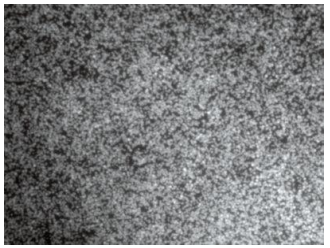

132532

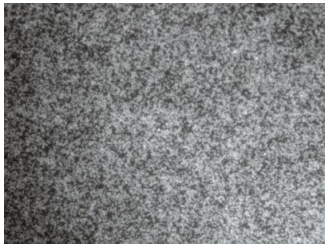

140432


148136


160834


169500


180580


190794


319000


360000

BROAD INSTITUTE illumina