454 Sample Prep / Workflow

at the

BioMedical Genomics Center (BMGC)

University of Minnesota

Sushmita Singh

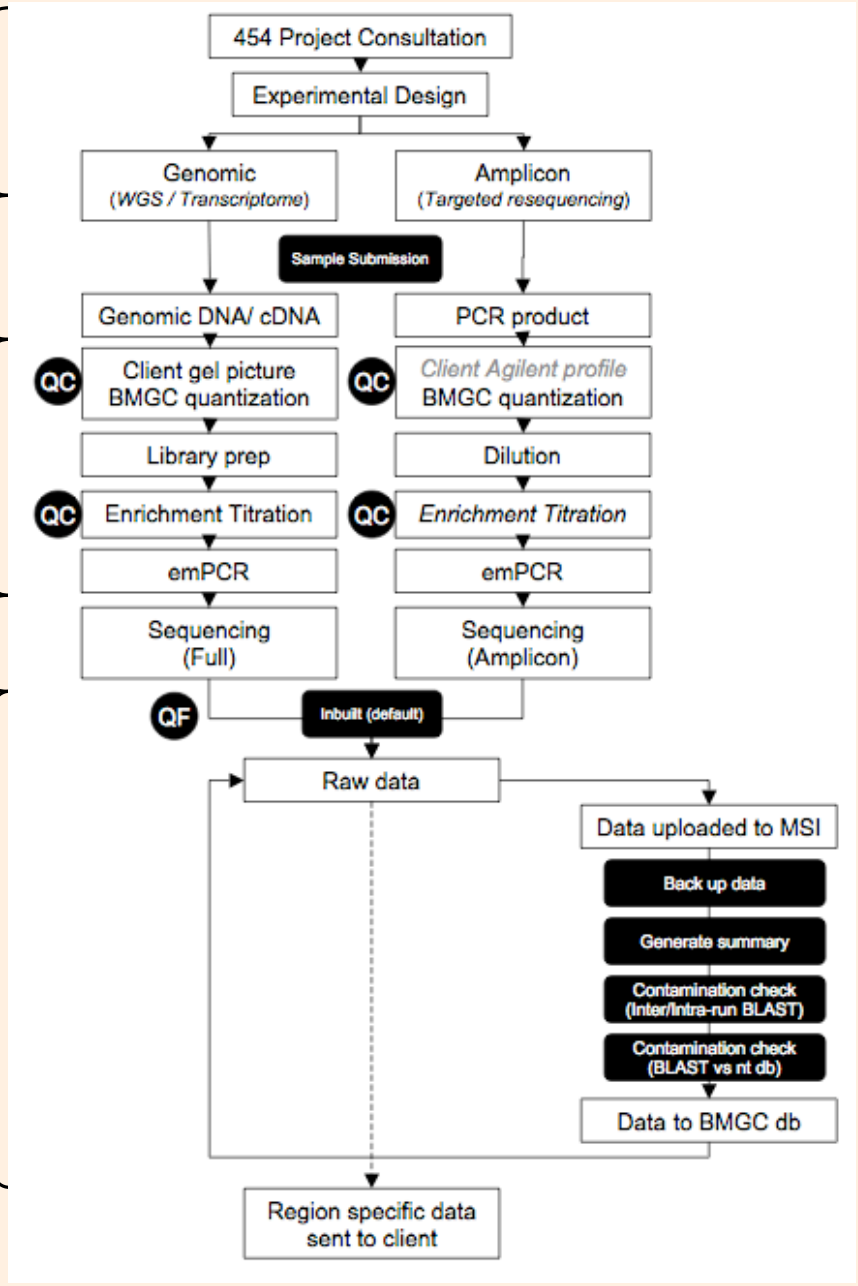
1 **Consultation**

2 **Sample Prep (client)**

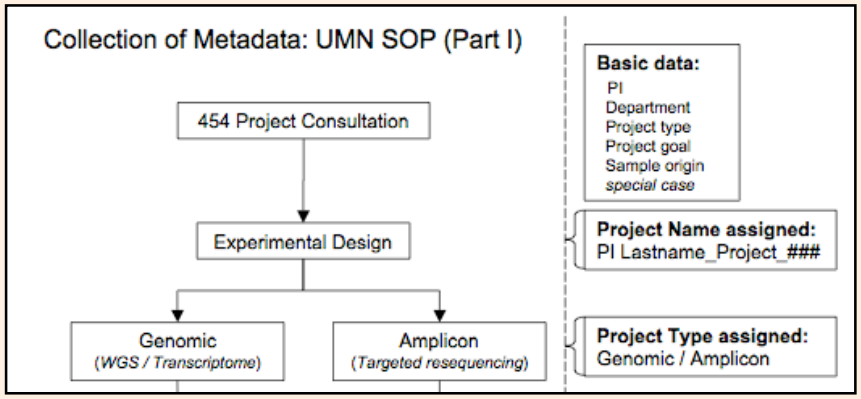
3 **Sample Prep (Core)**

Sequencing

4 **Data QC**



1 Consultation



Based on application-specific project goals and various other factors, help client develop suitable experimental design,

- sequencing type: genomic / amplicon
- coverage: what size regions / plates

UNIVERSITY OF MINNESOTA

BioMedical Genomics Center
DNA Sequencing and Analysis Facility
Rm. 20 Snyder Hall
1475 Gortner Ave
St. Paul, MN 55108

Phone: 612.625.7736
Fax: 612.624.2785
Email: agac@umn.edu

BMGC 454 Read Metrics

Sequencing Kit	Number of Regions per Gasket	Read Length (bases)	Number of Reads (*per* region)	Number of Bases (*per* region)	Total Number of Reads (all regions)	Total Number of Bases (all regions)
GSFLX Titanium XLR70	2	400	375,000	150,000,000	750,000	300,000,000
	4	400	120,000	48,000,000	480,000	192,000,000
	8	400	60,000	24,000,000	480,000	192,000,000
	16	400	20,000	8,000,000	320,000	128,000,000
GSFLX Standard LR70	2	250	200,000	50,000,000	400,000	100,000,000
	4	250	64,000	16,000,000	256,000	64,000,000
	8	250	27,500	6,875,000	220,000	55,000,000
	16	250	11,250	2,812,500	180,000	45,000,000
GSFLX Standard SR70	2	100	200,000	20,000,000	400,000	40,000,000
	4	100	65,000	6,500,000	260,000	26,000,000
	8	100	27,500	2,750,000	220,000	22,000,000
	16	100	11,250	1,125,000	180,000	18,000,000

UNIVERSITY OF MINNESOTA

BioMedical Genomics Center
DNA Sequencing and Analysis Facility
Rm. 20 Snyder Hall
1475 Gortner Ave
St. Paul, MN 55108

Phone: 612.625.7736
Fax: 612.624.2785
Email: agac@umn.edu

GS FLX/Titanium a la carte pricing

Library Preparation	Cost
Standard Chemistry (FLX)	
1st library	$800.00
Each additional library created at the same time	$400.00
Pico Green assay + dilutions (per sample)	$5.00
DNA 7500 or Pico Agilent chip	$30.00
Titanium	
1st library	$1,000.00
Each additional library created at the same time	$750.00
Paired End Library (3kb, 8kb, 20kb)	$1,800.00

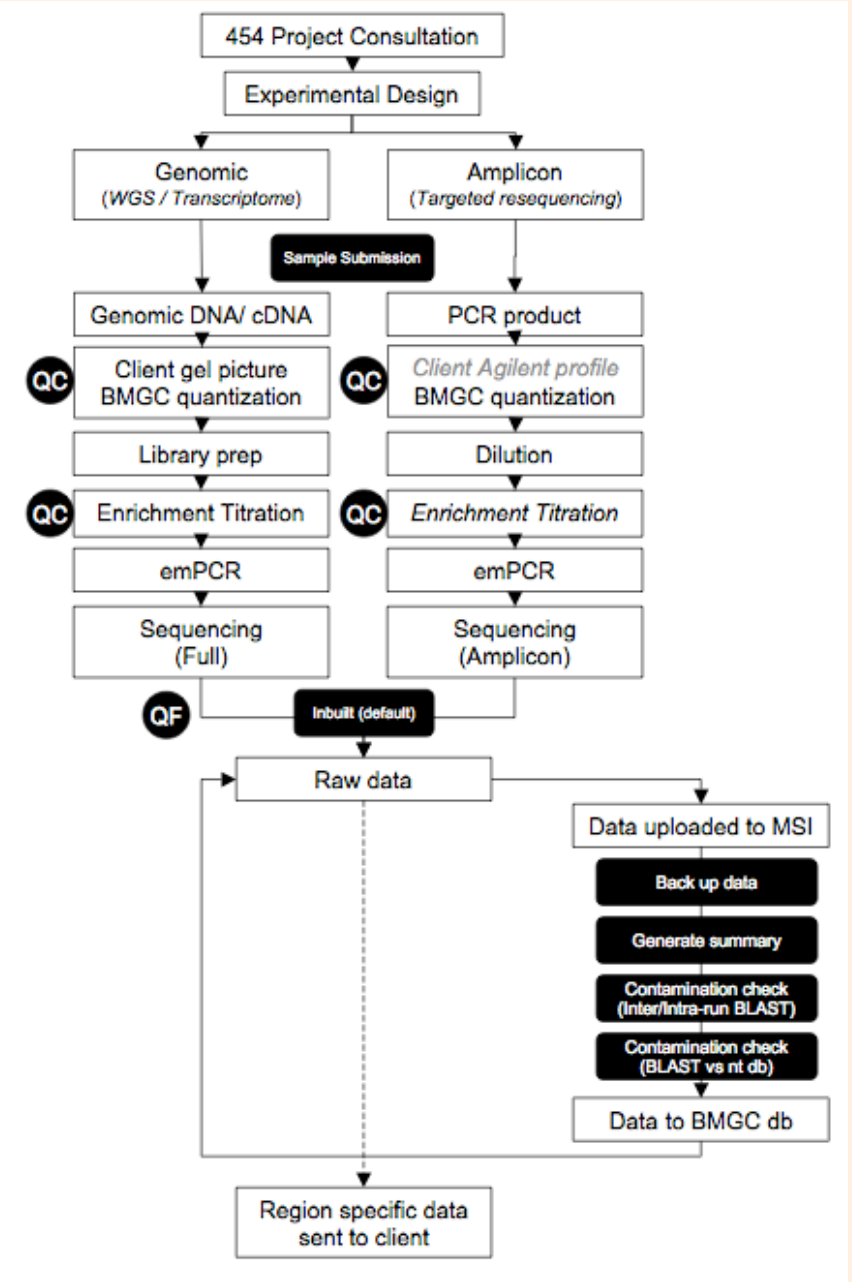
Emulsion Titration	Cost
Standard Chemistry (FLX)	
1st emulsion titration	$500.00
Each additional emulsion titration done at the same time	$250.00
Titanium	
1st emulsion titration	$600.00
Each additional emulsion titration done at the same time	$300.00

Bulk emPCR	Cost
Standard Chemistry (FLX)	$2,000.00
Titanium	$2,700.00

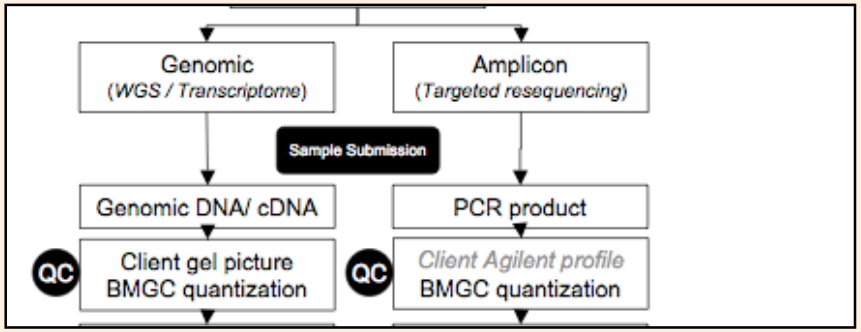
2 Sample Prep (client)

Provide detailed information on sample prep and sample submission requirements,
- during consultation
- providing information through brochures
- through links on BMGC website

- Genomic DNA
- cDNA (total RNA)
- Amplicon (PCR product)



2 Sample Prep (client)



Client is provided with detailed sample submission specifications in terms of quality and quantity

- At time of submission client is required to submit a sample submission form which queries user on information on sample and basic protocol used

Check:

- Gel picture
- OD values
- Quantization data

BMGC: Quantifies sample using Nanodrop spec analysis and PicoGreen™ assay (*these numbers are used in core*)

Genomic DNA

UNIVERSITY OF MINNESOTA

BIOMEDICAL GENOMICS CENTER
DNA SEQUENCING AND ANALYSIS FACILITY
1475 Gortner Avenue,
123 Snyder Hall, Room 20
St. Paul, MN 55108

Tel: (612) 625-7736
Fax: (612) 624-2785
www.bmgc.umn.edu

GS FLX/Titanium WGS Sample Description Form
Please fill out this document and include it with your samples

Name:	Ph:	Email:
Full Name of PI:		

SAMPLE RUN SPECIFICATIONS

DNA Type:
□ gDNA
□ cDNA
□ Plasmid
□ Cosmid / Fosmid
□ BAC
□ PCR Product
□ low mol.weight DNA [70-500 bp] *(please provide an explanation of content in 'remarks' section below)*
□ Other *(please provide an explanation of content in 'remarks' section below)*

SAMPLE REQUIREMENTS

At minimum, the DNA sample must be double-stranded, purified, non-degraded (>1.5kb in length, 70-500bp for low molecular weight) with an $OD_{260/280}$ of ≥1.8
For high molecular weight DNA, 5-10 µg of input DNA at a concentration of ≥ 300 ng/uL is requested (*please note that spectrophotometric analysis usually overestimates DNA conc.*)
For low molecular weight DNA (70-500bp): 3 µg at a conc. ≥ 50 ng/uL is requested
For other sample types please contact the BMGC Sequencing and Analysis Facility for input DNA requirements.
Samples can be provided in water or TE buffer (TE buffer is preferred)

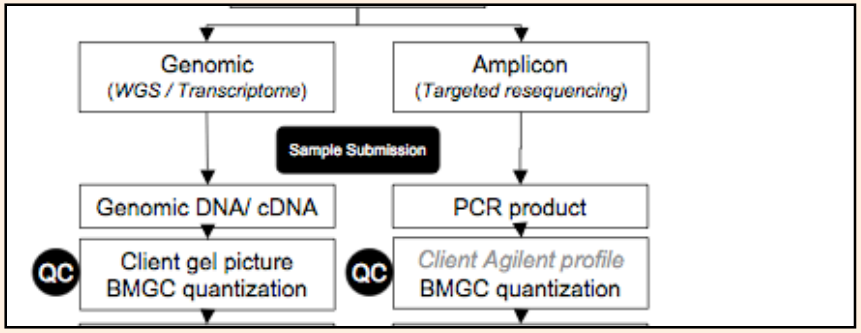
Have samples been purified? □ Yes □ No
Method used for sample prep: _____
Sample Buffer (TE preferred): _____

By what method did you quantify your sample/s?
□ Spectrophotometer
Did you use a Nanodrop? □ Yes □ No
□ Fluorescent dye based assays

Remarks: _____

Run Type:
□ GSFLX Standard chemistry (*can be run using LR70, SR70, LR25 kits*)
□ GSFLX Titanium Chemistry (*cannot be used with SR70 or LR25 kits*)

2 Sample Prep (client)



Client is provided with detailed sample submission specifications in terms of quality and quantity

- At time of submission client will be required to submit a sample submission form which queries user on information on sample and basic protocol used

Check:

- Quantization data
- Agilent profile (gel picture)

BMGC: Ethanol precipitate out RNA, quantify sample using Nanodrop spec analysis and run Agilent analysis to check quality

cDNA (now total RNA)

UNIVERSITY OF MINNESOTA

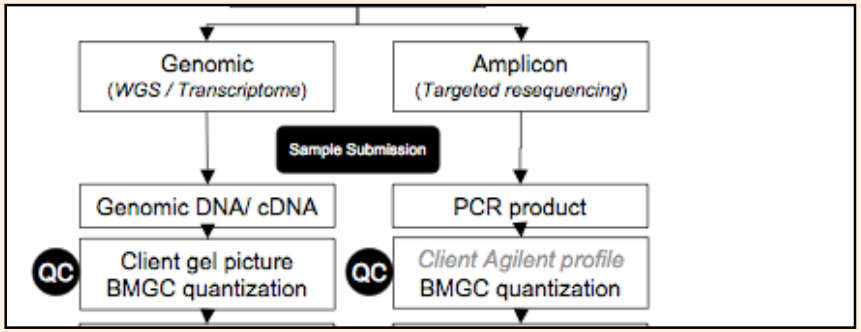
BIOMEDICAL GENOMICS CENTER
DNA SEQUENCING AND ANALYSIS FACILITY

1475 Gortner Avenue,
123 Snyder Hall, Room 20
St. Paul, MN 55108

Tel: (612) 625-7736
Fax: (612) 624-2785
www.bmgc.umn.edu

Under Construction!

2 Sample Prep (client)



Client is provided with detailed sample preparation and submission specifications

- At time of submission client is required to submit a sample submission form which queries user on information on sample

Check:

- Quantization data
- Agilent profile, if available

BMGC: Quantifies sample using PicoGreen assay and runs an Agilent profile

Amplicon (PCR product)

UNIVERSITY OF MINNESOTA

BIOMEDICAL GENOMICS CENTER
DNA SEQUENCING AND ANALYSIS FACILITY
1475 Gortner Avenue,
123 Snyder Hall, Room 20
St. Paul, MN 55108

Tel: (612) 625-7736
Fax: (612) 624-2785
www.bmgc.umn.edu

GS FLX Amplicon Sample Description Form
Please fill out this document and include it with your samples

Name:	Ph:	Email:

SAMPLE RUN SPECIFICATIONS

DNA Type:
□ Amplicon □ Other *(please specify)*: _____

Amplicon size designed:
□ 100bp (SR70 run)
□ 250bp (LR70 run)

Amplification primer preference:
□ Fusion primerA
□ Fusion primerB
□ Fusion primerA&B (for bi-directional sequencing)

Enter # of lanes and size desired:
(On a Large PTP plate)
___ of 2 region
___ of 4 region
___ of 8 region
___ of 16 region

SAMPLE REQUIREMENTS

Sample quality:
Please ensure that the DNA sample is highly purified and provide **only ONE** of the following recordings/information of your sample:

1. An Agilent profile (DNA1000 chip) of 1 µL of each sample. Clearly indicate sample identity in each profile OR
2. Spectrophotometric quantization of your samples.
   (*please note: spectrophotometric analysis usually overestimates DNA concentration so we highly recommend quantifying your samples with a fluorescent dye assay.*
   Our recommendation -Invitrogen Pico Quant-iT kit)

Sample specifications:
Please send us (or drop off) 100 µl of final mix to be amplified and sequenced in a clearly labeled microfuge tube at the BMGC Sequencing Facility.

► Samples can be provided in water or TE buffer (although TE is preferred)
Have samples been purified? □ Yes □ No
Method used: _____
Sample Buffer: _____

Special instructions: If you are 'pooling' different barcoded amplicons together then please quantify the DNA from all separate PCR reactions and combine equimolar amounts, mix and submit in one tube.

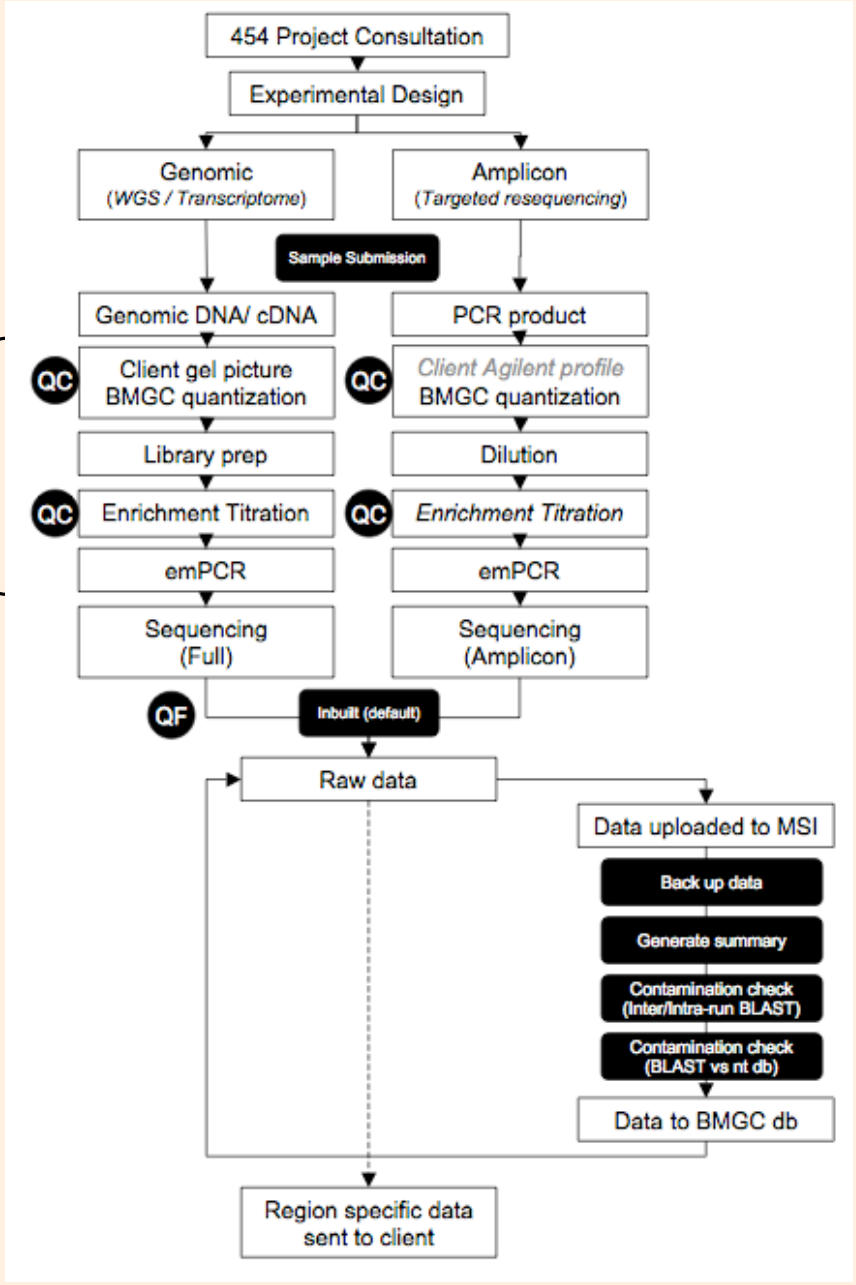
Please enter 'Sample' and 'Billing' information on the next page

3 Sample Prep (Core)

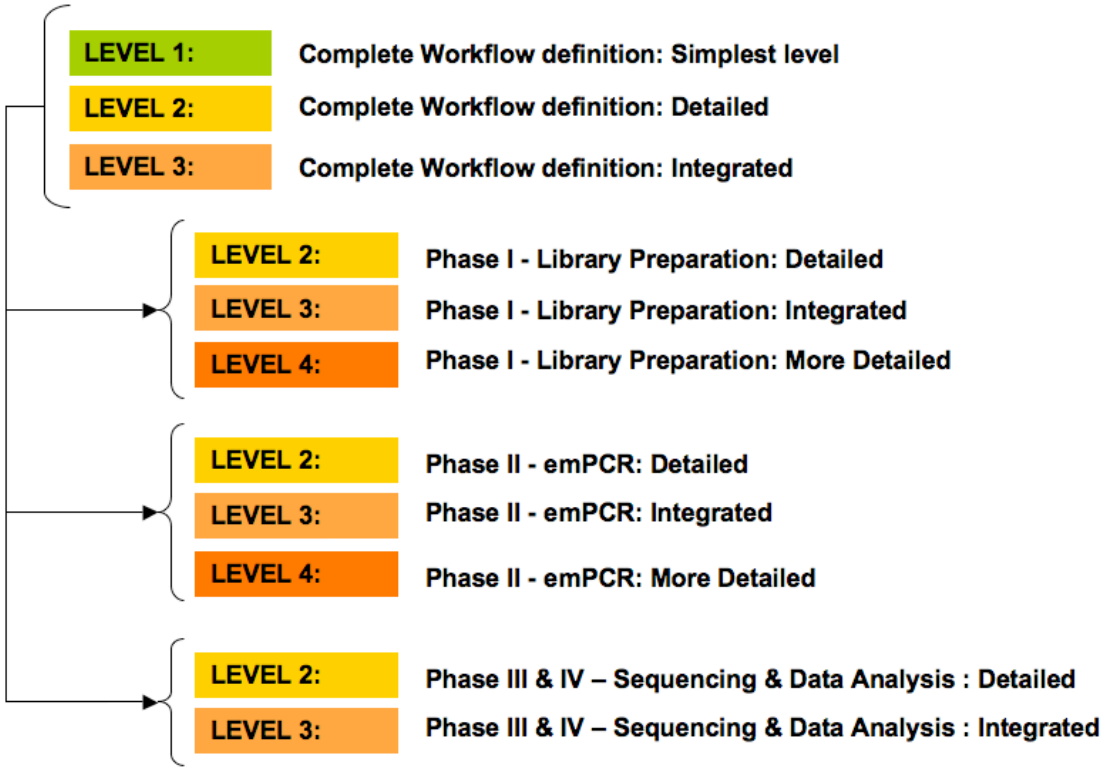
For the most part, vendor specified protocols are followed - A detailed review of the protocol was conducted and a workflow generated to,

- optimize efficiency
- minimize human error
- record detailed information

- Genomic DNA
- cDNA (total RNA)
- Amplicon (PCR product)

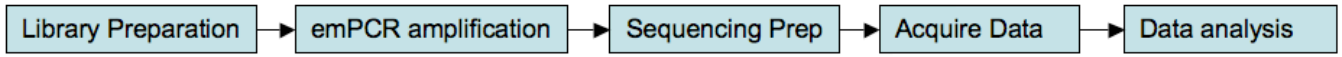


Workflow definition: 454 Sequencing



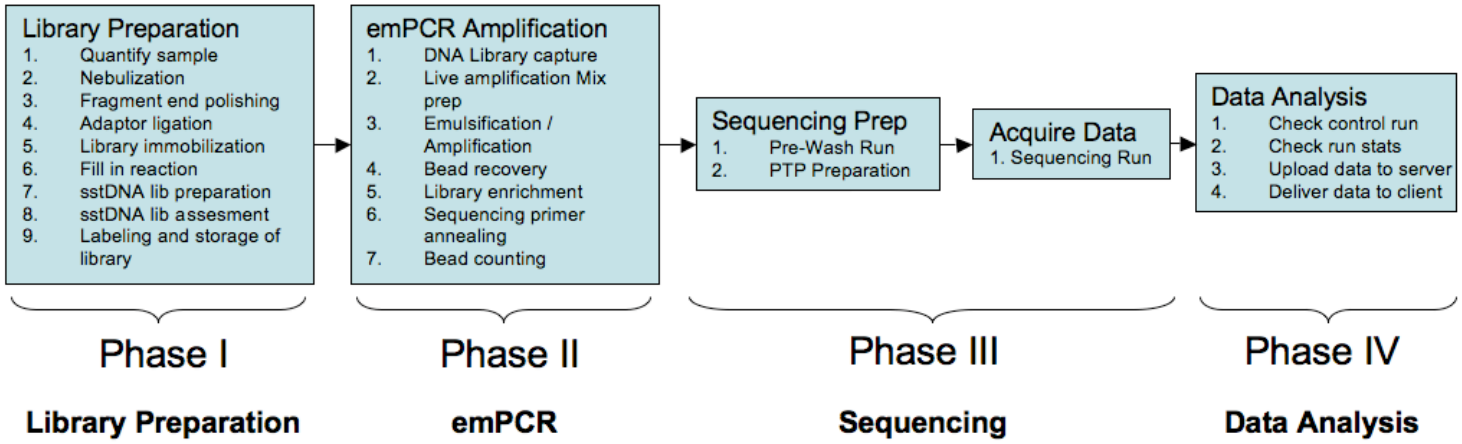
Complete Workflow definition: Simplest level

LEVEL 1:



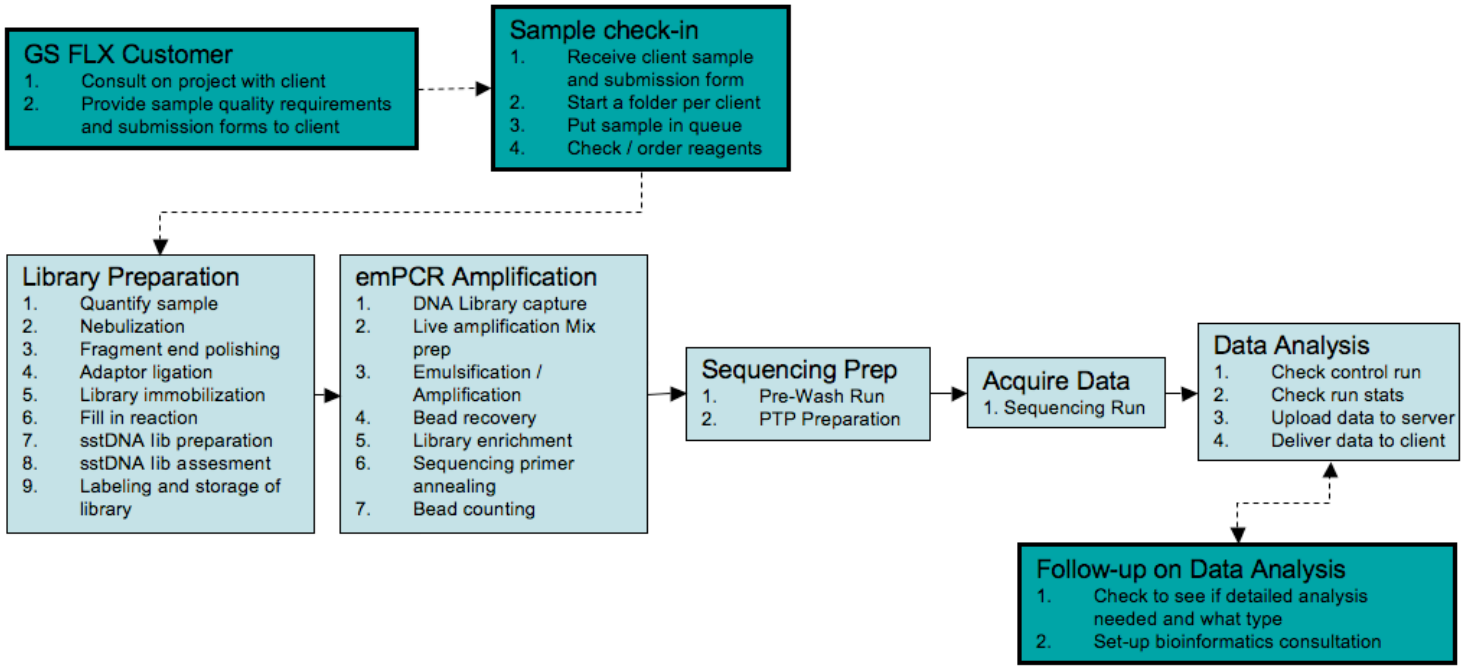
Complete Workflow definition: Detailed

LEVEL 2:



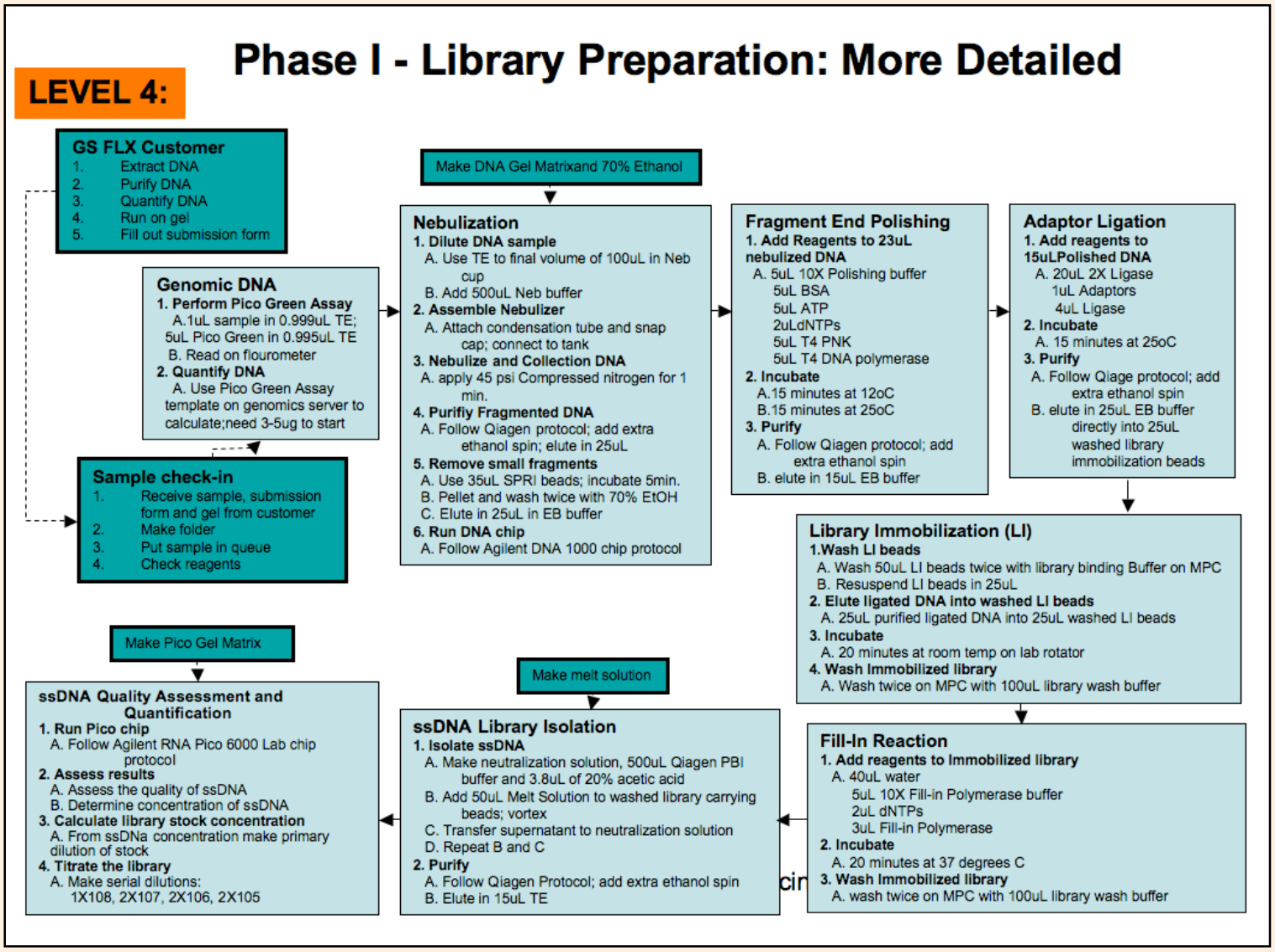
Complete Workflow definition: Integrated

LEVEL 3:



Phase I - Library Preparation: More Detailed

LEVEL 4:

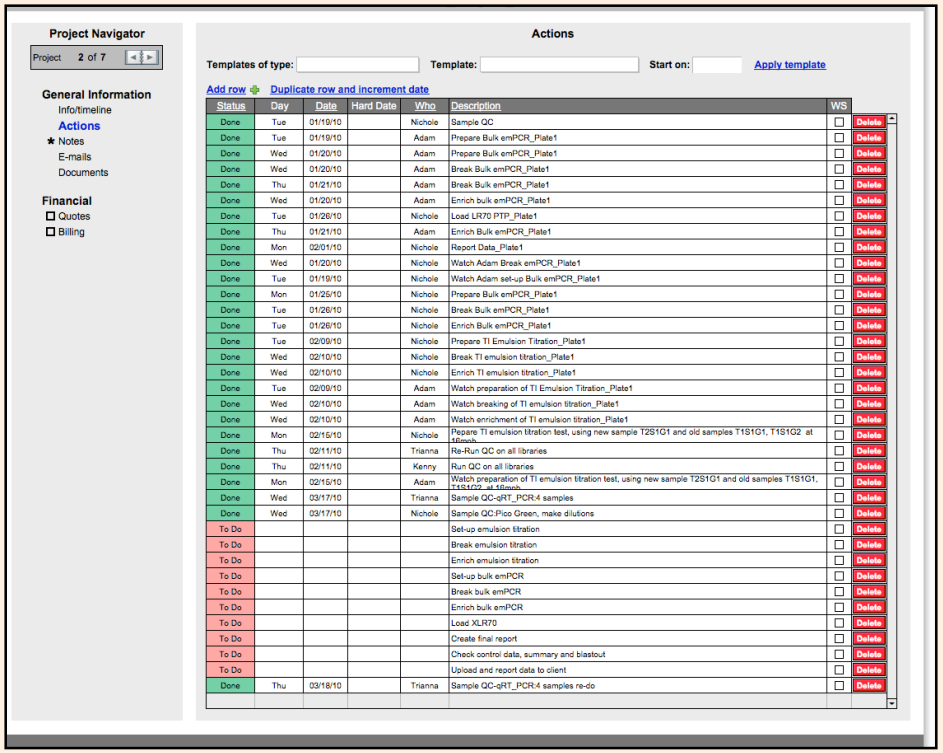


Preliminary LIMS system (FM Pro)

Project list:

Project Name	Project Type	Status	Title	Start Date	Due Date	End Date	Billable	Billing Status	Payment Status
						Total =	$0		
Bowser_Project_002	GS-FLX	In Progress	Bowser_Project_002	01/19/10				To Be Billed	To Be Invoiced
Isaacson_Project_007	GS-FLX	In Progress	Isaacson amplicon sequencing: Five Plates	01/19/10				To Be Billed	To Be Invoiced
Johnson_Project_008	GS-FLX	In Progress	Johnson_Project_008	02/19/10				To Be Billed	To Be Invoiced
Muehlbauer_Project_002	GS-FLX	In Progress	Muehlbauer GS-FLX transcriptomic analysis, attempt #2: Roche labeling and TItanium rapid library.	02/03/10				To Be Billed	To Be Invoiced
Murtaugh_Project_003	GS-FLX	In Progress	Murtaugh_Project_003	02/10/10				To Be Billed	To Be Invoiced
Sadowsky_Project_003	GS-FLX	In Progress	Sadowsky_Project_003	03/15/10				To Be Billed	To Be Invoiced
Sreevatsan_Project_005	GS-FLX	In Progress	Sreevatsan_005_16 lane	11/09/09			$0.00	To Be Billed	To Be Invoiced

Project detail:



A very basic LIMS system in place

Looking into possibilities with commercial NGS data handling software for future tech adaptations

Genomic DNA Sample Prep: QC & *Library Preparation*

Genomic DNA received

↓

Project opened

↓

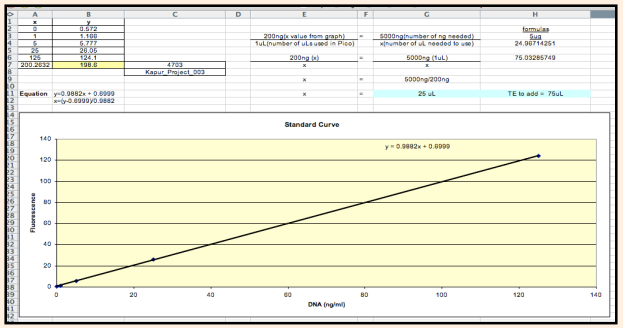
Submissions checked for given specifications

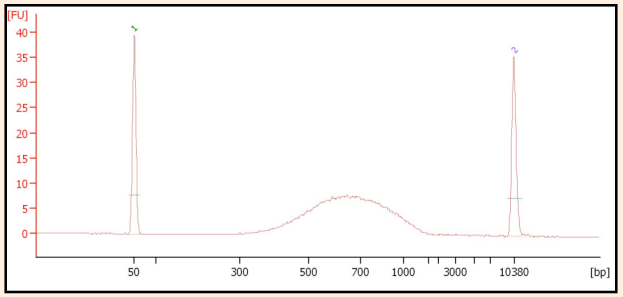
↓

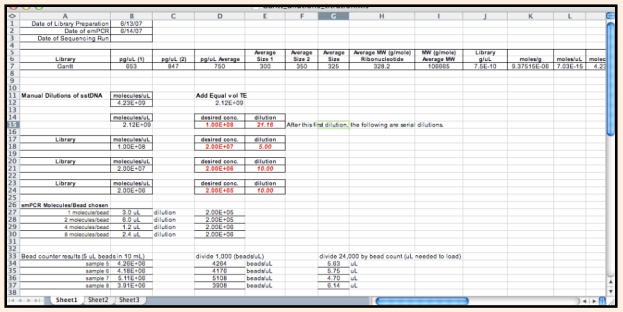
Samples quantified at core

↓

Library prep initiated using vendor specified protocols







Library Preparation

1. Quantify sample
2. Nebulization
3. Fragment end polishing
4. Adaptor ligation
5. Library immobilization
6. Fill in reaction
7. sstDNA lib preparation
8. sstDNA lib assessment
9. Labeling and storage of library

- Rapid library prep method
- MID library prep protocol

cDNA Sample Prep: QC & *Library Preparation*

Total RNA received

↓

Project opened

↓

Submissions checked for given specifications

↓

Samples quantified at core

↓

Total RNA quality checked using Agilent RNA Nano

↓

mRNA isolated using a commercial kit

↓

Library prep initiated using vendor specified protocols

Library Preparation

1. Quantify sample
2. RNA fragmentation
3. lib assessment
4. Labeling and storage of library

Total RNA QC

Project: ***Date:***

Enter Nanodrop Readings:

Sample ID	Concentration (ng/ul)	260/280	260/230
Example	500	2	1.7

The OD A260/A280 ratio should be very close to 2.0 (cutoff: 1.8) indicative of RNA purity. Pure RNA has a ratio of 2.1
The OD A260/A230 ratio should be very close to 2.0 (cutoff: 1.7) lower numbers indicative of contamination with chaotropic salts, phenol or protein

Dilution for Agilent Analysis:

(optimal range for Agilent RNA Nano assays is 25-500 ng/ul)

Enter desired final volume (ul) here: 2 *Note: default is 2ul, minimun required to heat and run a sample; you can set it to have reasonable pipetting volumes*

Sample ID	For 100ng/ul (ul)	Vol of H20 (ul)	Dilution Factor
Example	0.40	1.60	5
	#DIV/0!	#DIV/0!	#DIV/0!
	#DIV/0!	#DIV/0!	#DIV/0!
	#DIV/0!	#DIV/0!	#DIV/0!

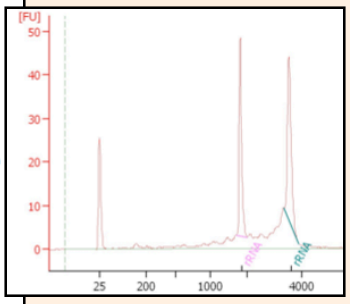
Agilent Readings:

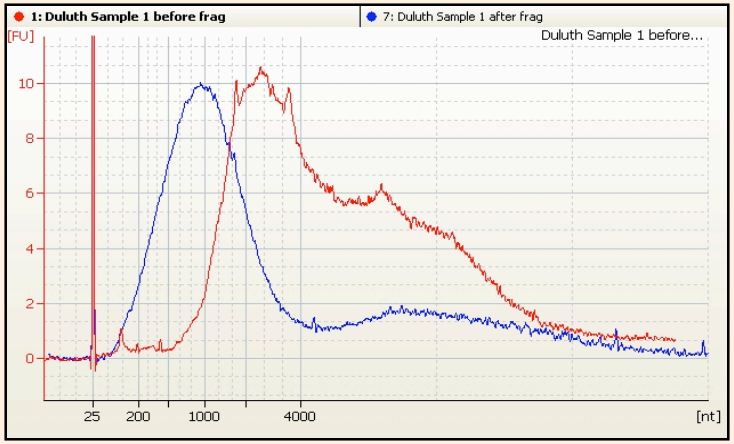
Assay:	Eukayotic	Total RNA	Nano
Sample ID	Concentration (ng/ul)	rRNA Ratio (28S/18S)	RIN
Example	223	1.7	9

The rRNA ratio should ideally be 2.0 or more (cutoff: 1.0)
The RNA Integrity Number or RIN should be close to 10 (cutoff: 7.0)

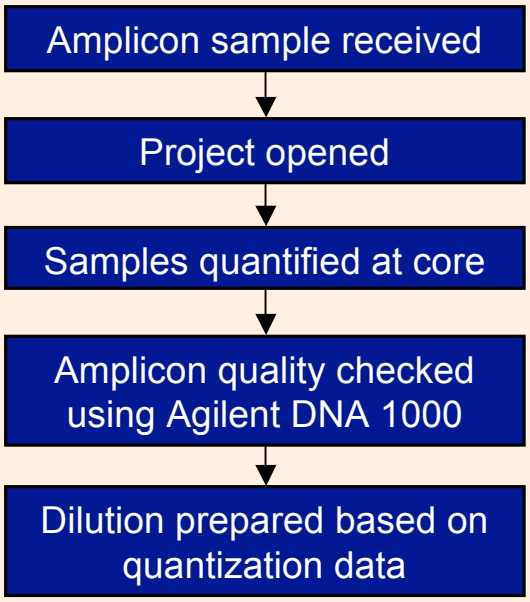
Please note:
The concentrations obtained from Agilent are of the diluted sample
so if you want the 'original' concentration, multiply by the dilution factor

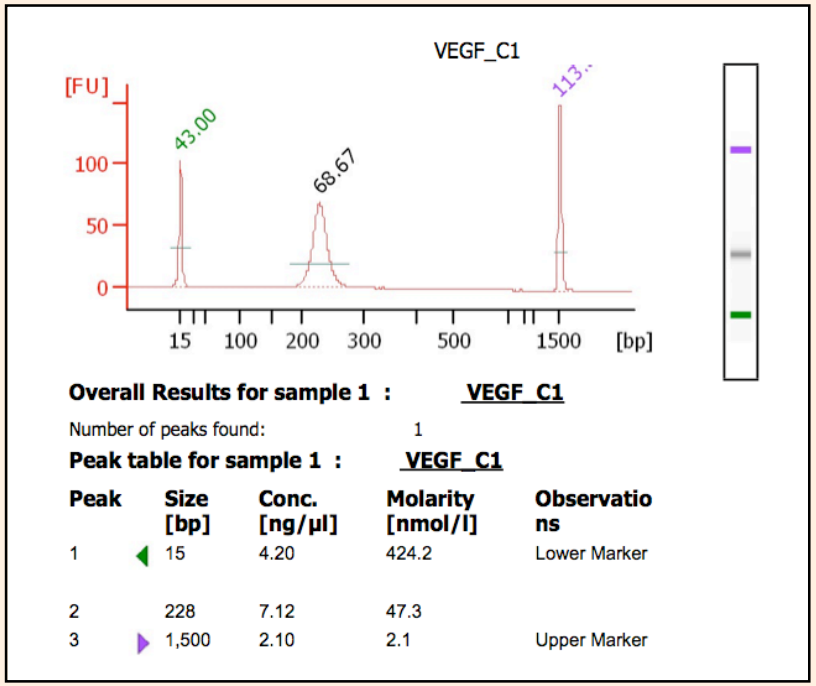
Total RNA QC





Amplicon Sample Prep: *QC and Dilution*





Enter sample names in column 1 in following Table:
Enter library concentration in ng/ul measured by flurometry in column 2 of following Table:
Enter average/ expected amplicon length in bp in column 3:

Sample	Lib Conc (ng/ul)	Amplicon Size in bp	Molecule/ul	First Dilution	1st Dilution (Column 5)	Final Vol in ul	Vol of TE	Desired Conc	2nd Dilution (Column 8)	Final Vol in ul	Vol of H2O
CRD 454 A-Agilent	10.41	216	4.42E+10	2.00E+08	1.1	250	248.9	2.00E+05	1.00	1000	999.00
CRD 454 B -Agilent	9.94	216	4.22E+10	2.00E+08	1.2	250	248.8	2.00E+05	1.00	1000	999.00
CRD 454 A- Pico	13.90	216	5.90E+10	2.00E+08	1.4	400	398.6	2.00E+05	1.00	1000	999.00
CRD 454 B -Pico	12.14	216	5.15E+10	2.00E+08	1.6	400	398.4	2.00E+05	1.00	1000	999.00

All samples: *enrichment titration and emPCR*

A bead enrichment titration assay conducted ascertain the best ‘molecules per bead’ (mpb) ratio for each library to ensure maximal output from emPCR reaction

(Bead stock is at 10,000 beads/ul)

600,000 beads for sst and Paired end
450,000 beads for Amplicon

Enter Coulter Counter (CC) Reading:

Sample	CC reading	Beads/ul	Total beads have	% Enriched
CRD 454 A 2mpb	2.51E+06	2512	*251,200*	56%
CRD 454 A 4mpb	3.52E+06	3524	*352,400*	78%
CRD 454 A 8mpb	3.85E+06	3852	*385,200*	86%
CRD 454 A 16mpb	5.23E+06	5232	*523,200*	116%
CRD 454 A .5 mpb	1.41E+06	1408	*140,800*	31%
CRD 454 A 1 mpb	1.91E+06	1908	*190,800*	42%
CRD 454 B .5 mpb	1.30E+06	1296	*129,600*	29%
CRD 454 B 1mpb	2.69E+06	2692	*269,200*	60%

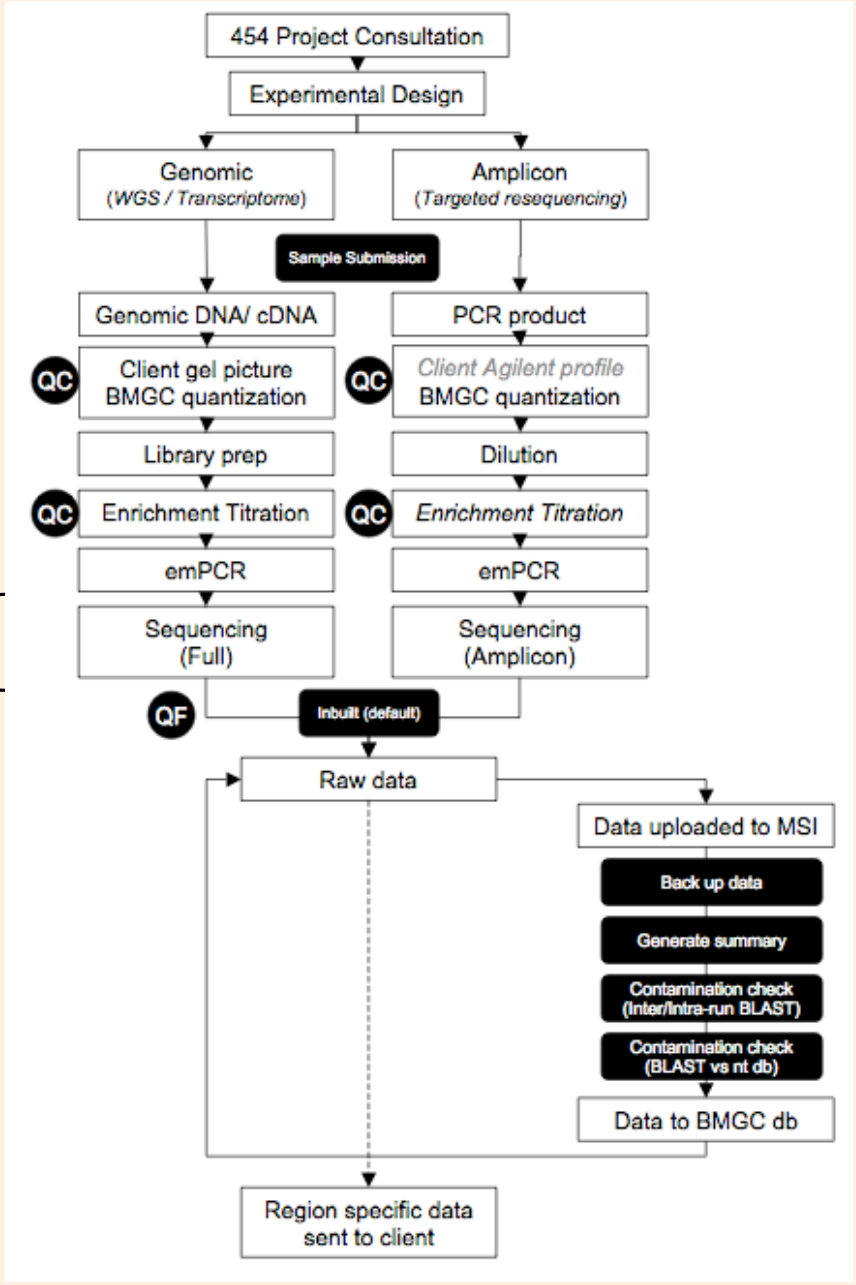
mpb	Bead recovery	%Enrichment
0.5	140800	31%
1	190800	42%
2	251200	56%
4	352400	78%
8	385200	86%
16	523200	116%

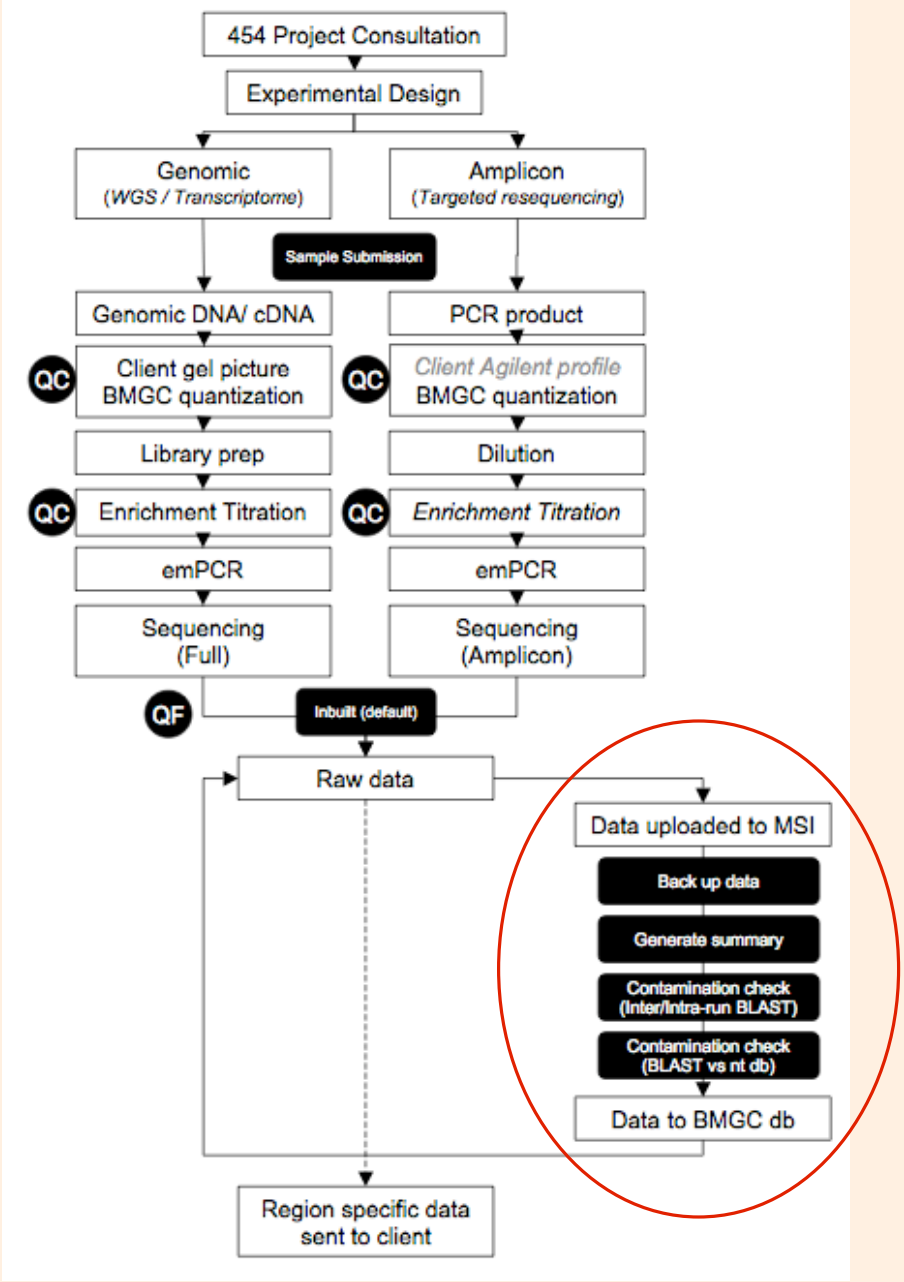
Default filtering parameters used (*reanalysis with variant filter parameter used in special cases only*)

Control sequences generated are pretty good indicators of the sequencing run quality

Sequencing

Close attention paid to machine functioning by checking for any leakage etc. Maintenance washes are conducted frequently and volume tests are conducted periodically



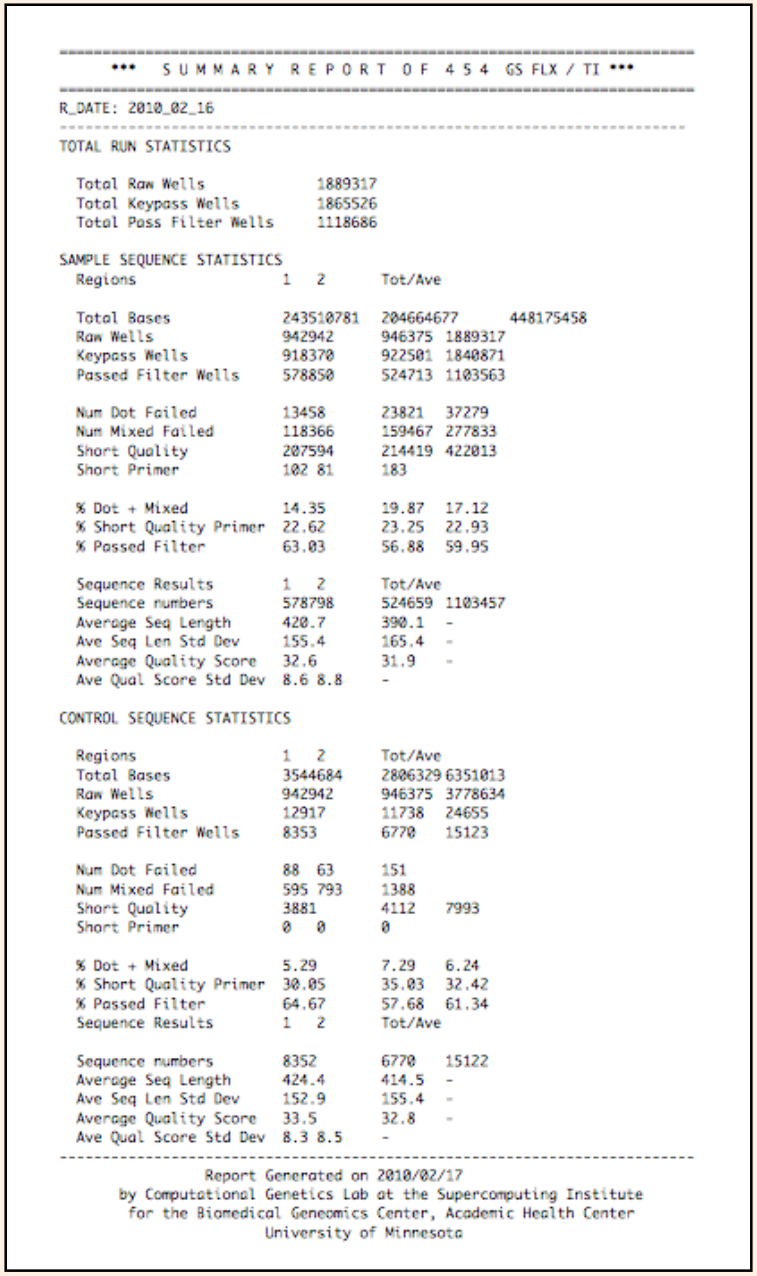


4 **Data QC**

Summary file generated used to evaluate run

↓

Contamination checks for the run conducted to detect cross-contamination



```
*** SUMMARY REPORT OF 454 GS FLX / TI ***

R_DATE: 2010_02_16

TOTAL RUN STATISTICS

  Total Raw Wells              1889317
  Total Keypass Wells          1865526
  Total Pass Filter Wells      1118686

SAMPLE SEQUENCE STATISTICS
  Regions                  1   2       Tot/Ave

  Total Bases              243510781   204664677      448175458
  Raw Wells                942942      946375  1889317
  Keypass Wells            918370      922501  1840871
  Passed Filter Wells      578850      524713  1103563

  Num Dot Failed           13458       23821   37279
  Num Mixed Failed         118366      159467  277833
  Short Quality            207594      214419  422013
  Short Primer             102 81      183

  % Dot + Mixed            14.35       19.87   17.12
  % Short Quality Primer   22.62       23.25   22.93
  % Passed Filter          63.03       56.88   59.95

  Sequence Results         1   2       Tot/Ave
  Sequence numbers         578798      524659  1103457
  Average Seq Length       420.7       390.1   -
  Ave Seq Len Std Dev      155.4       165.4   -
  Average Quality Score    32.6        31.9    -
  Ave Qual Score Std Dev   8.6 8.8     -

CONTROL SEQUENCE STATISTICS

  Regions                  1   2       Tot/Ave
  Total Bases              3544684     2806329 6351013
  Raw Wells                942942      946375  3778634
  Keypass Wells            12917       11738   24655
  Passed Filter Wells      8353        6770    15123

  Num Dot Failed           88  63      151
  Num Mixed Failed         595 793     1388
  Short Quality            3881        4112    7993
  Short Primer             0   0       0

  % Dot + Mixed            5.29        7.29    6.24
  % Short Quality Primer   30.05       35.03   32.42
  % Passed Filter          64.67       57.68   61.34
  Sequence Results         1   2       Tot/Ave

  Sequence numbers         8352        6770    15122
  Average Seq Length       424.4       414.5   -
  Ave Seq Len Std Dev      152.9       155.4   -
  Average Quality Score    33.5        32.8    -
  Ave Qual Score Std Dev   8.3 8.5     -

                Report Generated on 2010/02/17
        by Computational Genetics Lab at the Supercomputing Institute
          for the Biomedical Genomics Center, Academic Health Center
                        University of Minnesota
```

Summary Report

Check:

- Raw well numbers
- Keypass numbers
- Passed filter wells
- For Control:
  - % Dot + Mixed
  - % Short quality primer
- For Sample:
  - % Dot + Mixed
  - % Short quality primer

We are currently in the process of obtaining an Illumina system

Let the games begin….. again!….

Thank you for your attention!