Challenges of sequencing the 16S rRNA gene with the MiSeq Platform. Prior to the release of the Real Time Analysis software v. 1.17.28 and MiSeq Control Software v. 2.2.0, the Illumina platform was challenged by samples that had low genetic diversity due to the need to evenly excite both light channels on each cycle. Because of the low genetic diversity of the 16S rRNA gene fragment, this limitation of the Illumina platforms was particularly problematic. For example, within the V4 region of our mock community, 81 of the 250 sites were homogeneous for the G+T or A+C channels in the forward direction while in the reverse direction 51 sites were homogeneous. There were three commonly used solutions to solve this problem, which include reducing the amount of DNA being sequenced, increasing the amount of random control DNA that is spiked into the sample, and using a pre-defined calibration file based on the sequencing a high diversity sample. Commercially prepared libraries of the PhiX phage genome were commonly used for this application and Illumina recommends that between 40 and 50% of the DNA should be PhiX. We performed 7 MiSeq runs with varying amounts of input DNA and PhiX and with and without using the pre-defined calibration file using older versions of RTA and MCS. For each of these runs we assessed the error rates using the consensus sequence (**Tables S1 and S2**).

Table S1. Summary of operating conditions for various MiSeq sequencing runs and machine reported quality metrics. Shaded cells indicate the seven runs that were performed prior to the development of RTA and MCS software, which significantly reduced the amount of PhiX required to obtain high quality sequence data.

Run	RTA version	MCS version	[DNA] (pM)	Cluster density (K/mm^2)	% PhiX	Correction	%>= Q30	Reported machine error (%)	16S rRNA pairs (10^6)
121203	1.16.18	2.0.5	9	743	49.6	Adaptive	57.0	2.94	5.6
121205	1.16.18	2.0.5	8	607	59.9	Adaptive	64.5	1.37	3.9
121207	1.16.18	2.0.5	6	567	59.2	Adaptive	63.3	1.50	3.7
130125	1.17.22	2.1.13	4	192	82.8	Adaptive	94.4	0.61	0.4
130211	1.17.22	2.1.13	4	387	68.9	Hardcoded	92.0	0.59	1.9
130220	1.17.22	2.1.13	10	702	58.8	Hardcoded	93.9	0.87	4.5
130306	1.17.22	2.1.13	10	781	42.3	Hardcoded	78.9	0.96	4.6
130401	1.17.28	2.2.0	10	1313	3.8	Adaptive	63.7	1.31	12.6
130403	1.17.28	2.2.0	5	1094	3.2	Adaptive	70.5	0.89	12.4
130417	1.17.28	2.2.0	3	839	6.2	Adaptive	74.6	0.92	10.5
130422	1.17.28	2.2.0	1.8	690	8.0	Adaptive	80.1	0.65	9.0

Table S2. Summary of the error rates and number of observed OTUs for the sequencing runs described in Table S1. Shaded cells indicate the seven runs that were performed prior to the development of RTA and MCS software, which significantly reduced the amount of PhiX required to obtain high quality sequence data.

Region	Run	Error rates (%)			% Reads remaining from Basic	Average number of OTUs[A]				
		Basic	ΔQ=6	Pre-cluster		Mock[B]	Mock[C]	Soil	Mouse	Human
V34	121203	2.10	2.50	2.34	0.5	ND	ND	ND	ND	ND
	121205	1.95	0.65	0.51	1.5	ND	ND	ND	ND	ND
	121207	1.85	0.73	0.57	1.0	ND	ND	1,233.6	ND	ND
	130125	0.30	0.11	0.04	64.9	ND	ND	1,101.0	ND	151.0
	130211	0.54	0.24	0.11	57.5	36.3	47.4	1,078.1	154.4	164.3
	130220	0.78	0.23	0.11	41.8	34.9	49.5	1,050.2	157.9	154.3
	130306	0.96	0.26	0.13	27.5	36.5	53.0	1,102.1	175.2	138.7
	130401	2.14	0.37	0.21	10.3	26.9	49.6	1,110.6	175.1	187.5
	130403	1.30	0.26	0.12	27.6	31.1	47.8	1,095.8	158.2	164.1
	130417	1.12	0.24	0.10	27.9	35.1	52.2	1,038.6	ND	142.6
	130422	0.91	0.29	0.17	47.5	41.4	54.3	1,053.0	162.5	145.1
V4	121203	0.36	0.07	0.01	68.1	21.9	33.2	1,273.9	135.7	123.5
	121205	0.29	0.05	0.01	76.5	21.7	33.1	1,278.5	135.4	121.4
	121207	0.25	0.04	0.01	79.6	21.5	32.3	1,271.9	135.6	121.8
	130125	0.07	0.03	0.01	97.1	22.1	29.5	1,279.7	136.0	124.2
	130211	0.09	0.03	0.01	94.7	21.8	32.4	1,274.8	134.2	122.3
	130220	0.20	0.04	0.01	85.6	21.4	33.6	1,274.3	139.7	122.8
	130306	0.47	0.05	0.01	70.7	23.4	45.0	1,251.2	141.1	120.8
	130401	1.08	0.06	0.01	44.2	23.5	43.4	1,248.2	136.1	115.4
	130403	0.67	0.05	0.01	60.1	23.5	40.9	1,261.8	133.9	117.8
	130417	0.40	0.05	0.01	69.3	22.8	37.5	1,257.3	135.5	117.2
	130422	0.28	0.05	0.01	78.4	23.2	37.2	1,256.2	132.8	117.3
V45	121203	2.42	0.57	0.31	11.1	64.5	85.2	1,181.1	121.5	138.0
	121205	2.20	0.72	0.50	32.9	144.8	202.6	1,304.2	153.0	243.8
	121207	1.74	0.57	0.33	34.9	69.4	101.9	1,211.2	142.1	181.2
	130125	0.47	0.35	0.17	93.1	50.6	60.1	1,301.5	162.2	191.0
	130211	0.84	0.51	0.30	79.5	85.1	120.0	1,374.1	174.5	193.5
	130220	1.78	0.74	0.53	57.5	159.1	217.0	1,538.8	204.5	255.4
	130306	2.62	0.85	0.63	33.7	209.3	275.1	1,582.4	265.8	337.9
	130401	4.60	0.87	0.64	13.5	191.9	271.7	1,462.8	198.0	312.9
	130403	3.31	0.79	0.56	32.3	180.4	246.1	1,519.7	213.3	324.0
	130417	2.38	0.66	0.43	36.5	110.8	158.3	ND	180.4	242.1
	130422	1.67	0.58	0.36	56.4	98.0	131.6	1,403.3	186.3	227.2

A   The average number of OTUs is based on rarefaction of each sample to 5,000 sequences per sample; cells labeled as ND reflect samples that did not have at least one replicate with more than 5,000 sequences.

B   Number of OTUs in the mock community when all chimeras were removed; in the absence of chimeras and sequencing errors, there should be 20 OTUs for all three regions.

C   Number of OTUs in the mock community when chimeras were removed using UCHIME.

	Standard	Author	James Kozich
	Operating	Effective Date	20-May-13
	Procedure	Site	Ann Arbor
Title: 16S Sequencing with the Illumina MiSeq Personal Sequencer			
Detailed Title	Library Generation, QC, Sequencing, and Data Quality Analysis Procedures		
Original Date	25-Mar-13		
Additional Authors	Patrick Schloss		
Version	2.0		
Last Updated	20-May-13		
Updated By	James Kozich		

1.0) Aim/Rationale of Protocol

- The Purpose of this protocol is to define the steps and costs in both material and man hours for the preparation and sequencing of 16S libraries using the Illumina MiSeq in the laboratory of Dr. Patrick Schloss Ph.D., Department of Microbiology and Immunology.

2.0) Introduction and Workflow

2.1) Introduction

- The Illumina MiSeq Personal Sequencer can produce 2 x 250 paired-end reads and up to 15 Gb of data in a single run. Dual indexing of library samples allows up to 384 samples to be run simultaneously. The instrument is capable of producing in excess of 15 million pairs of reads, however, for low diversity runs about 12 million reads can be expected. A wide range of applications is possible including 16S analysis, metagenomics, genome sequencing, transcriptomics, and RNA sequencing.

- There are several steps in preparing samples for sequencing on the MiSeq. Broadly, these include library generation and indexing, quality control, normalization and pooling, quantification, sequencing, run quality assessment, and data export.

2.2) 16S Prep Workflow

1. Samples must be arrayed in 96 well plate format. If possible, leave two wells on each plate open for controls. Plate used should be compatible with Eppendorf epMotion 5075.

2. The investigator will provide the technician with the names of the project, experiment, a one-sentence description of the experiment, a name for each plate submitted (up to 4 96 well plates per run), the names and groupings of each sample.

3. A Sample Plate will then be created for each plate using Illumina Experiment Manager. Sample Plates will then be used to create a Sample Sheet. This sheet serves as the set of run parameters and indexing scheme used by the MiSeq for the run.

4. PCR of 16S samples with Schloss lab indices (up to 94 samples can be processed per plate). Each plate will contain a negative control (water) and a positive control (mock community).

5. A subset of 12 samples from each plate will undergo electrophoresis on a 1% agarose gel to ensure amplification proceeded normally.

6. Library clean up and normalization performed using the Invitrogen SequalPrep Plate Normalization Kit.

7. Samples from all plates are pooled.

8. Library QC includes quantification using a KAPA Biosystems Q-PCR kit cat# KK4824, and obtaining a Bioanalyzer trace using the Agilent Technologies HS DNA kit cat# 5067-4626.

9. Library enters the Sequencing Workflow.

2.3) Sequencing Workflow

1. The technician will review all documentation including the sample sheet. The sample sheet will be transferred to the instrument. The Technician will consult with the investigator to choose an appropriate amount of DNA to load for optimal cluster density and ensure optimum data quality/quantity.

2. The reagent cartridge will be thawed in a water bath per the MiSeq User Guide.

3. Unless otherwise specified, dilution and loading will follow the steps outlined in the document: Preparing DNA Libraries for Sequencing on the MiSeq (15039740 B)

4. Any custom primers required (including those for 16S) will be mixed with the native primers on the reagent cartridge.

5. The technician will load the sample sheet, flow cell, reagent cartridge, PR2 bottle, and an empty waste bottle onto the MiSeq and start the run. A 500 cycle run takes approx. 44 hours.

6. The technician will monitor the run using Illumina Sequence Analysis Viewer.

7. The data will be stored on the Schloss Lab NAS drive and made available to the investigator following run completion. Unless otherwise requested, default output will be fastq files.

8. Technician will perform a post run wash and any required maintenance of instrument.

3.0) Safety and Waste Disposal

- The Schloss Lab Chemical Hygiene Plan should be followed at all times.

- Standard PPE (nitrile gloves, safety glasses, and lab coat) should be used at all times.

- Each reagent cartridge contains a small amount of formamide and must be disposed of in an appropriate container following the run. Liquid waste from a run must also be disposed of as hazardous due to the formamide content.

- All other safety concerns may be addressed to the lab chemical safety officer.

4.0) Consumables

Reagent/Kit	Catalog #	Price
MiSeq® Reagent Kit v2 (500 cycle)	MS-102-2003	$910.80
PhiX Control Kit v3	FC-110-3001	$132.00
16s Index Primers (40 total)	IDT	$1180.00
16s Sequencing Primers (Read 1, Index, Read 2)	IDT	41.71
AccuPrime™ Pfx SuperMix 200rxns	12344-040	$331.28
Agilent High Sensitivity DNA Kit	5067-4626	$434.70
Library Quantification Kit - Illumina/Universal	KK4824	$575.00
SequalPrep™ Normalization Plate (96) Kit	A10510-01	$453.60
epT.I.P.S. Motion 1-50 µL Reloads 24 racks of 96 tips	30014421	$190.00
Tip One Refill Wafers 200µL elongated graduated	NC9549602	$39.24
Tip One Refill Wafers 1000µL graduated	1111-2721	$25.25
Fisher 1N NaOH 1L	SS266-1	$46.32
TWIN.TEC 96 Well Plate Skirted Blue	E951020460	$99.23

5.0) Run Costs

For 384 sample run	PCR and Indexing	Cleanup Pooling & Normalization	Library QC	Sequencing	Totals	Total Cost with Man Hours	Duration (days)
16S Reagents	$588	$265	$138	$941	$1932	$2,232	4.5 days
16S Man hours	4	3	4	4	15		

6.0) Method(s)

6.1) Published Protocols

- The following methods and references are used in the workflows above.
  - MiSeq User Guide Rev. H
  - Preparing DNA Libraries for Sequencing on the MiSeq (15039740 B)
  - Kapa Biosystems Q-PCR Library Quantification Kit Illumina
  - Accuprime Pfx Super Mix
  - Agilent High Sensitivity DNA Kit Guide
  - SequalPrep Normalization Plate (96) Kit

6.2) Working Protocols

- A Detailed protocol for the 16S workflow can be found in Appendix A.

7.0) Data Analysis

7.1) During Run

- The Technician will monitor the run using Illumina Sequence Analysis Viewer.
- To asses run quality, the technician will review the following data: cluster density, clusters passing filter, data yield, Q30 Score, intensity, corrected intensity, % base composition, % aligned, reads, reads passing filter, and indices identified following index reads.
- The tolerances for the preceding data are run type dependent. The investigator will consult the technician prior to the run to determine these values.

7.2) Post Run

- The default repository for run data is the "runs" folder on the Schloss Lab NAS Drive.
- Investigators wishing for personal copies of fastq files must provide the technician with storage media.
- The investigator will determine whether sufficient coverage and quality have been achieved or if there is a need for re-sequencing, and notify the technician immediately.
- It is recommended that investigators use mothur for fastq analysis.

8.0) Change Control

New Version	Date	Change	Reason for Change
1.0	3/18/2013	SOP Created	N/A
2.0	5/16/2013	Note about software update added to Appendix B. Updated linked Illumina Documents to current versions section 2.3 and 6.1. Removed Ampure link section 6.1. Updated consumables to reflect new prices and added pcr plates section 4.0. Adjusted man-hours calculations and run cost section 5.0. Eliminate Hardcoding language section 2.3.5. Corrected nM to µM Appendix A Initial Set Up lines 1 and 2, added reference to Appendix B, and eliminated hardcoding language substituting with current guidelines line 7. Under Appendix A PCR Program changes penultimate step to 72C 10:00. Appendix A Sequencing removed former line 3, which was a check to ensure hardcoding was in use; line 9 edited to reflect current NaOH practice; line 11 edited to reflect changes resulting from software update. Section 7.1 bullet 2 added reads and reads passing filter. Section 7.2 bullet 3 added note about technician notification. Appendix A Cleanup Normalization and Pooling line 8 changed to create pool from each plate. Use 5µl instead of 3µl. Appendix A Library QC line 1 eliminated 1:20 and 1:40 dilutions, changed to "each" pool; line 2j adjusted to load 4 plate pools on Bioanalyzer chip, line 3ki added "from Bioanalyzer", and added line 4. Appendix A Run Monitoring line 2 eliminated hardcoding language and adjusted expected parameters reflecting lower PhiX. Removed % base composition target.	MiSeq software update, general corrections, new price quotes.

9.0) Signature Block

Written By:_____ Date: _____

Reviewed and Approved By:_____ Date: _____

Appendix A

Detailed 16S Protocol

Initial Set up

1. Reconstitute indexed primers to 100 μM. See Appendix D for primer design.
2. Prepare 100µl 10 μM aliquots of indexed primers.
3. Array aliquots into four 96 well plates. Use the following scheme:
   a. A701 – A712 with A501 – A508
   b. A701 – A712 with B501 – B508
   c. B701 – B712 with B501 – B508
   d. B701 – B712 with A501 – A508
4. Extract template DNA and array in 96 well format leaving two wells open. (One for a negative water control and another for the positive Mock Community control)
5. Using Illumina Experiment Manager, create a sample plate for each 96 well plate of template. Choose indexes that correspond to one of the four index pair plates above. See Appendix C for instruction on creating a custom assay in IEM.
6. Using Illumina Experiment Manager, create a sample sheet for the run. Ensure that index choices are compatible with one another and there is sufficient diversity in the index reads so as to activate both light channels every cycle.
7. The MiSeq requires base diversity on every cycle. 16S is a low diversity library. With MiSeq software v2.2, 16S libraries can be loaded with 5% PhiX. Additionally, other high diversity samples such as metagenomes can be run simultaneously. This requires manually editing the sample sheet. Older software versions required "hardcoding" the matrix and phasing/pre-phasing values. See Appendix B.

PCR

Note: These steps may be performed using an epMotion or similar automated pipetting system.

1. Dispense 17 µl of Accuprime Pfx Supermix into each well of a new 96 well plate.
2. Using a multichannel pipette, transfer 1 µl of template DNA per well to the corresponding well on the PCR plate.
3. Using a multichannel pipette, transfer 2 µl of each paired set of index primers to the corresponding well on the PCR plate. Be sure to follow the layout chosen in the sample sheet.
4. Add 1 µl of PCR grade dH$_2$O to the negative control well, and 1 µl of Mock Community at a 1:3 dilution to the positive control well.
5. Repeat for up to four 96 well plates. Seal plates, vortex briefly and spin down contents.
6. Place in thermocycler.

PCR Program

Use the following program:

```
        95C  2:00
--------30 cycles--------
        95C  00:20
        55C  00:15
        72C  5:00
--------------------------
        72C  10:00
        4C   end
```

Gel Electrophoresis

1. A random row of 12 should be selected from each PCR plate and run on a gel to confirm success of the PCR.
2. Use 2 µl of sample, 4 µl of loading dye in a 1% agarose gel.
3. Run at 100v for 30 minutes alongside a standard ladder.
4. Photograph gel under UV. Check to be sure there is a band for every well.

Cleanup, Normalization, and Pooling

Use the SequalPrep Normalization Plate Kit

1. Transfer 18 µl of PCR product from PCR plate to corresponding well on the normalization plate.
2. Add 18 µl of Binding Buffer. Mix by pipetting, sealing, vortexing, and spinning briefly.
3. Incubate at room temperature for 60 minutes. Note: can incubate overnight if needed. Extra time does not improve results.
4. Aspirate the liquid from the wells. Do not scrape the sides.
5. Add 50 µl of Wash Buffer and pipette up and down twice, then aspirate immediately. Ensure there is no residual wash buffer in any wells.
6. Add 20 µl of Elution Buffer. Mix by pipetting up and down 5 times. Seal, vortex, and spin briefly.
7. Incubate at room temperature for 5 minutes.
8. Create a pool from each plate. Take 5 µl of each well to pool. The use of an empty 96 well plate may facilitate the use of multichannel pipettes.
9. Freeze the remaining sample for later use.

Library QC

1. Prepare the following dilutions of each pooled library in PCR grade H_2O:
   a. 1:1
   b. 1:10
   c. 1:1000 (dilute in several steps for better results)
   d. 1:2000
   e. 1:4000
2. Agilent Bioanalyzer Trace
   a. Prepare Gel-Dye mix if not already prepared.
   b. Let reagents equilibrate to room temperature.

c. Turn Bioanalyzer on and load 350 µl of dH$_2$O onto electrode cleanser and place in analyzer for 5 minutes.

d. Open a high sensitivity chip and place on the priming station. Base plate should be a position "C" and syringe clip should be at lowest position.

e. Load 9.0 µl of gel-dye mix to position 12 market with a large "G". Ensure the syringe plunger is at 1.0 ml and close the station. Press plunger until it is held by clip.

f. Wait for exactly 60 seconds then release the plunger clip. Wait an additional 5 seconds, then slowly pull the plunger back to the 1.0 ml position.

g. Open the priming station. Pipette 9.0 µl of gel-dye mix into the other wells marked "G" in positions 4,8,and 16.

h. Pipette 5.0 µl of marker to all wells excluding the right column. (No marker positions 4,8,12, and 16)

i. Load 1 µl of ladder into position 15 marked by the ladder symbol.

j. Pipette 1 µl of each of dilutions a – b above. Top row Plate 1 Pool 1:1 x 1, 1:10 x 2. Second row Plate 2 pool 1:1 x 1, 1:10 x 2. Third row Plate 3 Pool 1:1 x 1, 1:10 x 2. Bottom row Plate 4 Pool 1:1 x 1, 1:10 x 2.

k. Place chip in the designated vortex for 1 minute, then transfer chip to the Bioanalyzer.

l. Open the 2100 Expert software and select the HS DNA Assay. Enter sample names/dilutions for each of the test wells. Click Start.

m. Print .pdf when run finishes.

3. Kapa Q-PCR Library Quantification

a. Before Q-PCR reaction setup, add 1 ml Primer Premix (10X) to the 5 ml bottle of KAPA SYBR® FAST Q-PCR Master Mix (2X) and mix by vortexing for 10 sec. Record the date of Primer Premix addition on the KAPA SYBR® FAST Q-PCR Master Mix bottle.

b. Reaction can be either 10 µl or 20 µl. A 10 µl reaction volume is recommended.

c. Prepare a 96 well Q-PCR plate compatible with the real time thermocycler. There are six standards. Each should be run in triplicate. Each pool at each dilution should be run in triplicate.

d. For 10 µl reaction volume dispense 6 µl of master mix into each well needed.

e. Pipette 4 µl of standards and library dilutions into appropriate wells. Mix by pipetting. Vortex and spin optional.

f. Place plate in thermocycler. Start control software

g. Program the following cycle
   i. Initial Activation   95 °C   5 minutes
   ii. 35 cycles
      1. Denaturation   95 °C   30 seconds
      2. Annealing   60 °C   45 seconds
      3. If library fragment size exceeds 700bp, extend annealing step to 90 seconds.
   iii. Perform melt curve to check for primer/adaptor dimer

h. Assign wells and group replicates.

i. Enter values for standards
   i. Std. 1   20pM
   ii. Std. 2   2pM
   iii. Std. 3   0.2pM
   iv. Std. 4   0.02pM
   v. Std. 5   0.002pM
   vi. Std. 6   0.0002pM

vii. Note: The concentrations provided here are for the DNA Standards as supplied in the kit, and are NOT the concentrations in the reactions. Provided that the volume of template added to each reaction is the same for Standards and for library samples (i.e. 4 µl in each case), there is no need to account for these volumes when calculating the concentrations of library samples, nor should one need to calculate the concentration of template in the reaction.
j. Run program
k. To calculate library concentration use the following formula:
   i. Average x (452/Avg fragment length from bioanalyzer) x dilution factor
   ii. Use the average of the triplicate data points corresponding to the most concentrated library DNA dilution that falls within the dynamic range of the DNA Standards to calculate the concentration of the undiluted library.
   iii. Do not include outliers in calculation. If there is more than one outlier in a group, the assay must be repeated.
4. Create normalized pools from each plate by diluting to the concentration of the least concentrated plate. Create a single final pool by adding equal amounts of each post qpcr normalized pool. Final pool must be >10µl in total volume. 40-80µl is ideal.

Sequencing

1. Remove a 500 cycle reagent cartridge from the -20 °C freezer. Place in room temperature water bath for one hour. Place HT1 buffer tube in 4 °C fridge.
2. Copy sample sheet to sample sheet folder on MiSeq. Rename sample sheet to match barcode of reagent cartridge.
3. When the reagent cartridge has thawed, dry bottom with paper towel. Invert the cartridge repeatedly to check each well is thawed. This also serves to mix the reagents. Place in hood.
4. Thaw library, PhiX, and sequencing primers. Check to make sure HT1 is thawed.
5. Place 3 µl of the Read 1 Sequencing Primer(s) into a clean PCR tube. Repeat in separate tubes for the Index Primer(s) and Read 2 Sequencing Primer(s).
6. Using a 1000 µl pipette tip, break the foil over wells 12, 13, 14, and 17.
7. Use an extra long 100 µl tip with the pipettor set on 75 µl to transfer the 3 µl of Read 1 Sequencing Primer to the bottom of well 12 and pipette 10X to mix. Repeat this process spiking the Index Primer into well 13 and the Read 2 Sequencing Primer into well 14.
8. Prepare a fresh dilution of 0.2N NaOH.
9. To a 1.5ml tube add 10 µl of library, and 10 µl of 0.2N NaOH. To a separate tube add 2 µl PhiX, 3 µl PCR grade water, and 5 µl of 0.2N NaOH. Vortex both tubes to mix and spin for 1 minute at 400rcf. Note: NaOH concentration on the flow cell must remain under 0.001N. Adjusting the concentration of the NaOH used to denature the DNA to 0.1N may be necessary.
10. Allow the tubes to incubate at room temperature for 5 minutes. Immediately add 980 µl of HT1 to the library tube, and 990 µl HT1 to the PhiX tube.
11. Use HT1 to dilute both the library and PhiX to 10pM. For a 5% PhiX run, combine 950 µl of 3.5pM Library and 50 µl PhiX in a final tube. Vortex. Load 600 µl of this solution into well 17 on the reagent cartridge. See example below:
    a. (1.45 nM library x 10 µl) + (0.2N NaOH x 10 µl) + 980 µl HT1 = 14.5pM Lib, 0.002N NaOH
    b. (14.5pM lib x 241.38 µl) + 758.62 µl HT1 = 3.5pM lib, 0.00048N NaOH
    c. [(10nM PhiX x 2 µl) + 3 µl H_2O] + (0.2N NaOH x 5 µl) + 990 µl HT1 = 20pM PhiX, 0.001N NaOH
    d. (20pM PhiX x 175 µl) + 825 µl HT1 = 3.5pM PhiX, 0.0000175N NaOH

      e.   (3.5pM Lib x 950 µl) + (3.5M PhiX x 50 µl) = solution loaded

      f.   Solution loaded is 3.5pM overall with a 3.325pM Library concentration, 0.175pM PhiX concentration, and 0.0000457N NaOH

13. Set reagent cartridge aside. Unbox flow cell and PR2 bottle.
14. Thoroughly rinse the flow cell with Milli-Q water. Carefully dry by blotting with lint free wipes (Kimwipes). Give special attention to the edges and points of intersection between the glass and plastic.
15. Wet a new wipe with 100% alcohol and wipe the glass on both sides avoiding the rubber intake ports.
16. Visually inspect the flow cell to ensure there are no blemishes, particles, or fibers on the glass.
17. Follow on screen instructions and load the flow cell, reagent cartridge, and PR2 bottle. Empty and replace the waste bottle.
18. Ensure the machine recognizes the correct sample sheet and the run parameters are correct.
19. Wait for the MiSeq to perform its pre-run checks, and press start.

Run Monitoring

1. The run should be monitored periodically using Illumina Sequence Analysis Viewer.
2. Ideal parameters for a 95% 16S run:
      a.   Cluster density 700-800k/mm2
      b.   >85% clusters passing filter
      c.   5% aligned (amount of PhiX)
      d.   No spikes in corrected intensity plot
      e.   All indices identified following index reads
      f.   Final >Q30 score of >70%

Final Steps

1. Perform post run wash.
2. Dispose of liquid waste in appropriate hazardous jug and reagent cartridge in hazardous bucket.
3. When MiSeq Reporter finishes, copy the fastq files from the analysis folder to the run folder on the NAS drive.
4. Perform maintenance wash if required.
5. Confirm with investigator that data are of sufficient quality and quantity.

Appendix B

Note: The need for hardcoding has been eliminated with the 3/25/2013 release of the MiSeq 2.2 software update package. The changes calculate phasing/pre-phasing values dynamically for each cycle. See the release notes here: MiSeq_v2.2_Software_Release_Notes

Hardcoding

Background

   When sequencing low diversity samples like 16S libraries, the MiSeq can generate unusually high phasing/pre-phasing values. Since the MiSeq is made to run a wide variety of sample types, it attempts to determine these values in the first 12 cycles of each read. High values may lead to a known RTA software glitch called a corrected intensity spike. This manifests as a massive increase in corrected intensity values either in one large spike or in a pattern of exponential growth in amplitude over the course of a read. This causes the MiSeq to either fail outright or generate extremely low Q30 scores and error rates in the 80-100% range. Hardcoding provides the matrix and phasing/pre-phasing values in advance based on an ideal run. In most cases this eliminates the corrected intensity spike issue. Note however, hardcoding is not an officially supported customization.

Preparation

   The procedure requires a paired end, non-indexed PhiX (or other balanced sample) run to generate the values that will be hardcoded. It is important to ensure that this run is done with the same reagent kit version and software version as your low-diversity samples, and if any major service is done on the instrument, the PhiX run must be performed again, and the values updated. The values are generated in the first 12 cycles, so running 250 cycles in each direction in not necessary. Twenty cycles in each direction should be more than sufficient. Illumina technical support can provide instructions and two versions of the MiSeqConfiguration.xml file. One version is the original, the other contains the framework for hardcoding the values from your PhiX run.

Procedure

   The files from Illumina contain a series of instructions, which should be followed with a few alterations.
1. In the C:\Illumina\RTA\Configs directory, create a new folder called "alternate configurations". Copy the standard_MiSeqconfiguration.xml and lowdiversity_MiSeqconfiguration.xml from Illumina to this folder. Copies of the matrix and phasing/pre-phasing value files from the PhiX run can be stored here for future reference.
2. The low diversity file is not initially set up for a paired end, dual indexed run. It must be modified. Open the file with the text editor of your choice. Copy the entire block for index read including the </ArrayofFloat> lines at the beginning and end. Insert a new line between the index read section and the read 2 section. Paste the new section here.
3. Additionally, a line must be added to the phasing and pre-phasing sections.
4. Use the values from read one for both index reads.

Note: The MiSeqconfiguration.xml file must be swapped back to the original if a non-low diversity run is to be performed.

Appendix C

Adding An Assay To Illumina Experiment Manager

Introduction

      Illumina Experiment Manager is used to generate sample plates and sheets. A new assay must be added to the system to efficiently prepare sample sheets for 16S sequencing. Not only does this eliminate the need to manually assemble a sample sheet for 16S runs, but allows the user to retain use of the IEM index analysis feature. This ensures the indices selected for a particular run have sufficient diversity on a cycle-by-cycle basis and will successfully demultiplex.

Procedure

1. Open C:\program files\Illumina\Illumina Experiment Manager\Sample Prep Kits
2. Copy the file Nextera.txt, and rename the file Schloss.txt
3. Open in a text editor
4. Under [Name] change to Schloss
5. Under [PlateExtension] change to Schloss
6. Under [I7], clear the Nextera indices and paste in the SA701 through SB712 index names, and the REVERSE COMPLIMENT of the primer index sequences.
7. Under [I5], remove the Nextera indices and paste in the SA501 through SB512 index names and index sequences with no alteration (NOT the reverse compliment).
8. Under [DefaultLayout_SingleIndex] and [DefaultLayout_DualIndex] each well of a 96 well plate is listed with a corresponding Nextera index name. These must be replaced with Schloss index names. It is recommended that a text editor with column select capability be used to leave the index name number unchanged (i.e. 701, 702, etc.), while replacing the character "N" for Nextera with "SA" for Schloss A.
9. Save and close.
10. Return to the IEM folder and open the Applications folder. Open and edit the file GenerateFASTQ.txt.
11. Add a line at to the bottom with the text "Schloss"

Appendix D

Primer design

Overall design considerations
- The sequencing primers must have a melting temperature near 65°C. This can be achieved by altering the pad sequence
- The index sequences must balance the number of bases at each position. The index sequences listed here have a 25% ATGC composition at each site. If you are going to cherry pick indices from the list, make sure that you have even representation.

Generic PCR primer design:
```
AATGATACGGCGACCACCGAGATCTACAC <i5><pad><link><16Sf>              VX.N5??
CAAGCAGAAGACGGCATACGAGAT <i7><pad><link><16Sr>                   VX.N7??
```

Generic read 1 primer design
```
<pad><link><16Sf>                                                VX.read1
```

Generic read 2 primer design
```
<pad><link><16Sr>                                                VX.read2
```

Generic index read primer design
```
Reverse complement of (<pad><link><16Sr>)                        VX.p7_index
```


The listed sequences in the generic design, above, are the adapter sequences to allow annealing of the amplicons to the flow cell. The i5 and i7 sequences are the 8-nt index sequences. The pad is a 10-nt sequence to boost the sequencing primer melting temperatures. The link is a 2-nt sequence that is anti-complementary to the known sequences. The 16Sf and 16Sr are the gene specific primer sequences. Primers are purchased from IDT with no special purification. This system should work for any other region of the 16S rRNA gene or any other gene. The only thing to change would be the 16Sf/16Sr sequences and confirm that when combined with the pad sequence that the melting temperature is near 65°C.

16Sf
```
V3:    CCTACGGGAGGCAGCAG
V4:    GTGCCAGCMGCCGCGGTAA
```


16Sr
```
V4:    GGACTACHVGGGTWTCTAAT
V5:    CCCGTCAATTCMTTTRAGT
```


Link:
```
V4f:   GT
V4r:   CC
V3f:   GG
V5r:   GG
```

Pad:
```
Forward:     TATGGTAATT
Reverse:     AGTCAGTCAG
```


i5
```
SA501 ATCGTACG
SA502 ACTATCTG
SA503 TAGCGAGT
SA504 CTGCGTGT
SA505 TCATCGAG
SA506 CGTGAGTG
SA507 GGATATCT
SA508 GACACCGT

SB501 CTACTATA
SB502 CGTTACTA
SB503 AGAGTCAC
SB504 TACGAGAC
SB505 ACGTCTCG
SB506 TCGACGAG
SB507 GATCGTGT
SB508 GTCAGATA

SC501 ACGACGTG
SC502 ATATACAC
SC503 CGTCGCTA
SC504 CTAGAGCT
SC505 GCTCTAGT
SC506 GACACTGA
SC507 TGCGTACG
SC508 TAGTGTAG

SD501 AAGCAGCA
SD502 ACGCGTGA
SD503 CGATCTAC
SD504 TGCGTCAC
SD505 GTCTAGTG
SD506 CTAGTATG
SD507 GATAGCGT
SD508 TCTACACT
```


i7
```
SA701 AACTCTCG
SA702 ACTATGTC
SA703 AGTAGCGT
SA704 CAGTGAGT
SA705 CGTACTCA
SA706 CTACGCAG
SA707 GGAGACTA
SA708 GTCGCTCG
SA709 GTCGTAGT
SA710 TAGCAGAC
SA711 TCATAGAC
SA712 TCGCTATA
```

```
SB701 AAGTCGAG
SB702 ATACTTCG
SB703 AGCTGCTA
SB704 CATAGAGA
SB705 CGTAGATC
SB706 CTCGTTAC
SB707 GCGCACGT
SB708 GGTACTAT
SB709 GTATACGC
SB710 TACGAGCA
SB711 TCAGCGTT
SB712 TCGCTACG

SC701 ACCTACTG
SC702 AGCGCTAT
SC703 AGTCTAGA
SC704 CATGAGGA
SC705 CTAGCTCG
SC706 CTCTAGAG
SC707 GAGCTCAT
SC708 GGTATGCT
SC709 GTATGACG
SC710 TAGACTGA
SC711 TCACGATG
SC712 TCGAGCTC

SD701 ACCTAGTA
SD702 ACGTACGT
SD703 ATATCGCG
SD704 CACGATAG
SD705 CGTATCGC
SD706 CTGCGACT
SD707 GCTGTAAC
SD708 GGACGTTA
SD709 GGTCGTAG
SD710 TAAGTCTC
SD711 TACACAGT
SD712 TTGACGCA
```

Primers used to amplify 384 samples using the V4 region:
```
v4.SA501   AATGATACGGCGACCACCGAGATCTACACATCGTACGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA502   AATGATACGGCGACCACCGAGATCTACACACTATCTGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA503   AATGATACGGCGACCACCGAGATCTACACTAGCGAGTTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA504   AATGATACGGCGACCACCGAGATCTACACCTGCGTGTTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA505   AATGATACGGCGACCACCGAGATCTACACTCATCGAGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA506   AATGATACGGCGACCACCGAGATCTACACCGTGAGTGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA507   AATGATACGGCGACCACCGAGATCTACACGGATATCTTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SA508   AATGATACGGCGACCACCGAGATCTACACGACACCGTTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB501   AATGATACGGCGACCACCGAGATCTACACCTACTATATATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB502   AATGATACGGCGACCACCGAGATCTACACCGTTACTATATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB503   AATGATACGGCGACCACCGAGATCTACACAGAGTCACTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB504   AATGATACGGCGACCACCGAGATCTACACTACGAGACTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB505   AATGATACGGCGACCACCGAGATCTACACACGTCTCGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB506   AATGATACGGCGACCACCGAGATCTACACTCGACGAGTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB507   AATGATACGGCGACCACCGAGATCTACACGATCGTGTTATGGTAATTGTGTGCCAGCMGCCGCGGTAA
v4.SB508   AATGATACGGCGACCACCGAGATCTACACGTCAGATATATGGTAATTGTGTGCCAGCMGCCGCGGTAA
```

```
v4.SA701    CAAGCAGAAGACGGCATACGAGATAACTCTCGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA702    CAAGCAGAAGACGGCATACGAGATACTATGTCAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA703    CAAGCAGAAGACGGCATACGAGATAGTAGCGTAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA704    CAAGCAGAAGACGGCATACGAGATCAGTGAGTAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA705    CAAGCAGAAGACGGCATACGAGATCGTACTCAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA706    CAAGCAGAAGACGGCATACGAGATCTACGCAGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA707    CAAGCAGAAGACGGCATACGAGATGGAGACTAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA708    CAAGCAGAAGACGGCATACGAGATGTCGCTCGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA709    CAAGCAGAAGACGGCATACGAGATGTCGTAGTAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA710    CAAGCAGAAGACGGCATACGAGATTAGCAGACAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA711    CAAGCAGAAGACGGCATACGAGATTCATAGACAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SA712    CAAGCAGAAGACGGCATACGAGATTCGCTATAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB701    CAAGCAGAAGACGGCATACGAGATAAGTCGAGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB702    CAAGCAGAAGACGGCATACGAGATATACTTCGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB703    CAAGCAGAAGACGGCATACGAGATAGCTGCTAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB704    CAAGCAGAAGACGGCATACGAGATCATAGAGAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB705    CAAGCAGAAGACGGCATACGAGATCGTAGATCAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB706    CAAGCAGAAGACGGCATACGAGATCTCGTTACAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB707    CAAGCAGAAGACGGCATACGAGATGCGCACGTAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB708    CAAGCAGAAGACGGCATACGAGATGGTACTATAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB709    CAAGCAGAAGACGGCATACGAGATGTATACGCAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB710    CAAGCAGAAGACGGCATACGAGATTACGAGCAAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB711    CAAGCAGAAGACGGCATACGAGATTCAGCGTTAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
v4.SB712    CAAGCAGAAGACGGCATACGAGATTCGCTACGAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
```

Read 1 primer for V4 region
```
TATGGTAATTGTGTGCCAGCMGCCGCGGTAA
```

Read 2 primer for V4 region
```
AGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
```

Index primer for V4 region
```
ATTAGAWACCCBDGTAGTCCGGCTGACTGACT
```