



Division of Genomic Medicine
University of California Box 0794
San Francisco, CA 94143

Tel: 415-476-3200
Fax: 415-502-0720
email: robert.nussbaum@ucsf.edu

February 14, 2014

Dr. Kellie Kelm
Division of Chemistry and Toxicology Devices
Office of In Vitro Diagnostics and Radiological Health (OIR)
Food and Drug Administration
10903 New Hampshire Avenue WO66, Room 5648
Silver Spring, MD 20993-0002

Re: Pre-Submission for an IDE

Dear Dr. Kellie Kelm:

The following package of documents is for an IDE pre-submission regarding an NIH U19 Research Grant entitled, "Sequencing of Newborn Blood Spot DNA to Improve and Expand Newborn Screening." We are submitting this package to determine whether or not the project requires an IDE from FDA. Enclosed, please find one binder, or hard copy, and two e-copies in the form of two flash drives. The e-copies are an exact duplicate of the paper copy.

The research grant will utilize "next generation" DNA sequencers to clinically validate results of de-identified newborn blood spots previously analyzed with tandem mass spectrometry (MS/MS). It will also determine if DNA sequencing can validate disorders patients have been determined to have, but are not detectable with MS/MS. The ethical, legal and social implications (ELSI) aim will deal with creating a research participants protection framework for newborn sequencing. The ELSI arm of the project will also generate analysis of the legal issues involved with newborn sequencing, work with focus groups to understand public views on the issue and create an ethics and policy board.

As we are clinically validating already known results and working to develop ELSI frameworks to protect patients, we do not believe that our research constitutes a

significant risk. This being the case, we do not believe that our research needs an IDE from FDA, but are seeking your input on this matter. Once the FDA has reviewed the attached packet containing all of the recommended information for an FDA IDE pre-submission package we look forward to answering any questions you have and working with you on this issue.

Sincerely,

A handwritten signature in black ink that reads "Robert L. Nussbaum, MD". The signature is written in a cursive style with a large initial "R" and a stylized "L".

Robert L. Nussbaum, MD, FACP, FACMG,
Chief, Division of Genomic Medicine

Enclosure

PART A – COVER LETTER

PART B – TABLE OF CONTENTS

B

PART C – DEVICE DESCRIPTION

C

AUTOGEN PREP 965 EXTRACTION PROTOCOL	MARKED “MOUSE TAIL 1-MOUSE TAIL 13”
NEXT GENERATION SEQUENCING SOP	MARKED “1-14”
ILLUMINA HISEQ 2500 MANUAL	MARKED “10-204”

PART D – PROPOSED INTENDED USE/INDICATIONS FOR USE

D

PROJECTS 1-3 PROCEDURAL CHARTS	MARKED “PROJECT 1 – PROJECT 3”
<u>RESEARCH PLAN</u>	
OVERVIEW OF THE PROGRAM	MARKED “79-86”
COMPONENT PROJECT 1	MARKED “129-134”
COMPONENT PROJECT 2	MARKED “151-159”
COMPONENT PROJECT 3	MARKED “193-209”

PART E– PREVIOUS DISCUSSIONS OR SUBMISSIONS

E-H

PART F–OVERVIEW OF PRODUCT DEVELOPMENT

E-H

PART G– SPECIFIC QUESTIONS

E-H

PART H– MECHANISM OF FEEDBACK

E-H

Part C – Device Description

The following pages provide detailed information about the extraction and sequencing processes that will be used to carry out the specific aims¹ of our “Sequencing of Newborn Blood Spot DNA to Improve and Expand Newborn Screening” grant.

Extraction

The AutoGenprep 965 Mouse tail (Tissue) DNA protocol includes, “recommended conditions and parameters for blood punch DNA extraction.” The Autogenprep 965 manual is not included in this submission, but can be provided upon request.

Sequencing

The SOP for NGS sequencing follows the extraction protocol as well as the manual for Illumina’s HiSeq 2500, the instrument that will be performing the sequencing.

¹ Project Specific Aims provided in “Research Plan” section on pages marked 129, 151, and 193

AutoGenprep 965 Protocol Note

Mouse tail (Tissue) DNA Protocol

Origin of DNA

Mouse tail and other animal tissues

Cycle Definition

Protocol name: Mouse tail

Sample type: Proteinase K digested homogenates (0.3 ml)
 Optional [Digest] Process automates dispensing of digesting solutions into 96 deep well plates.

Maximum Samples: 384 samples (four of 96 deep well plates) ✓

Processing time: (1) [Extraction] - DNA Extraction
 3.0 hours/192 samples, 4.5 hours/ 384 samples
 (Including 60 min dry up time and excluding incubation time for Proteinase K digestion)
 (2) [Digest] - Optional Process to dispense digesting solutions
 5 min /192 samples, 8 min /384 samples

Yield: Comparable to the manual extraction
 Around 10µg / 10mg mouse tail

Outline of Isolation

Starting from Proteinase K digested mouse tail samples, contaminating protein is removed by the propriety reagent (R3) and genomic DNA is precipitated and washed using alcoholic reagents (Extraction Process). *The protocol has optional (Digest) process to have AutoGenprep 965 dispense Proteinase K solution and detergent solution for Proteinase K digestion reaction.*

Unloading

1. Unload isolated DNA in Stock plates which is on Stock RACK.
2. Discard Sample plates, used pipette tips and reagents remained in the reagent bottles.



Note

Isolated DNA is in the original Stock Plate RACK on the moving rack table

Sample preparation (for mouse tail)

Snip mouse tails (5-10 mm, about 10-15mg) and put them into each well of 96 deep well plates (if you want to digest in 96 deep well plates) or appropriate tubes. Add 150µl of M1 (Tissue Digestion Solution A) containing 0.4mg/ml Proteinase K and 150µl of M2 (Tissue Digestion Solution B). Seal the plates (or tubes) and incubate at 50-65C overnight with consistent shaking.

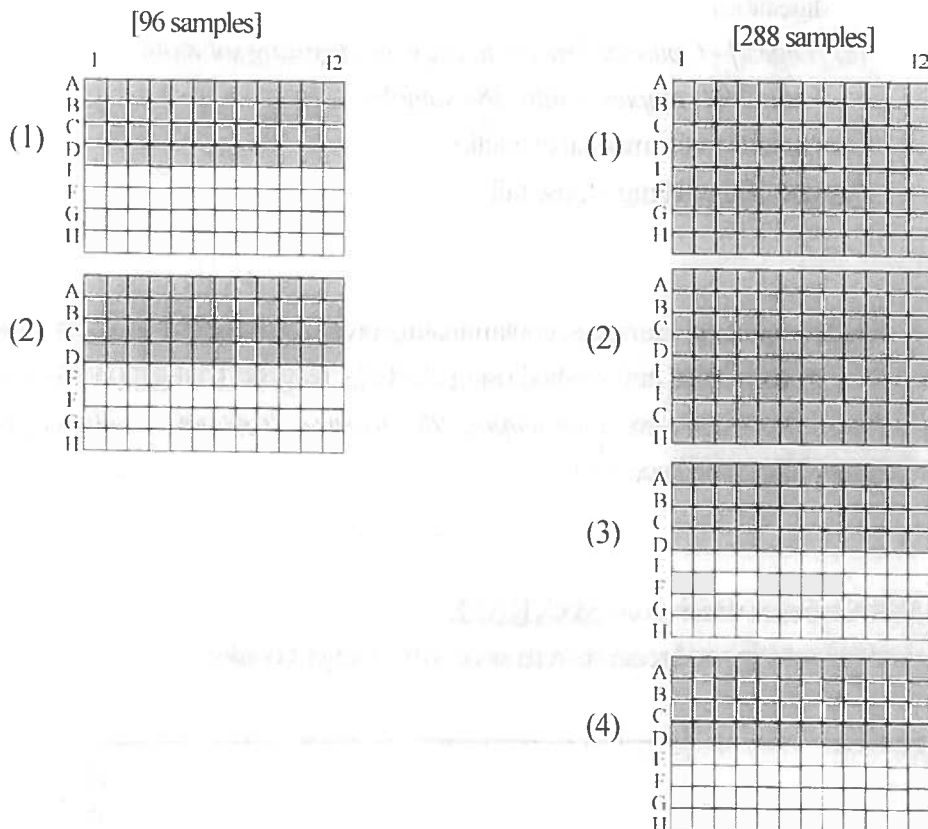
Note

Standard parameters of the Extraction Protocol are optimized for the 0.3 ml digest as the starting volume, and the same 96 deep well plates should be used for both sample loading and DNA recovery. Starting volume more than 0.3 ml can not be recommended.

Loading sample and microplate

Transfer the digests into the wells of 96 deep well plates.

You can process multiples of 24 as the number of samples (minimum 24 and maximum 384 samples). Samples must be loaded in the shadowed wells like following diagram to balance the weights for centrifugation.



1. Set 96 deep well plates with sample on Sample Rack (on the right side)
2. Set fresh 96 deep well plates to recover DNA on the Stock Rack (on the left side)

Set pipette tips

Each sample requires its dedicated pipetting tip (Sample Tip) for the extraction.

Each reagent also requires its own pipetting tip (Reagent Tip).

1. Set sample tips to proper positions of holes of sample tip rack. The lines of tip rack, which correspond to the lines of sample plate with samples, must have the tips. Each line must have twelve (12) tips.
2. Set the tip rack on the moving rack table.
3. Each reagent reservoir use pipette tips which set to dedicated line of reagent tip rack. Refer to the following table for setting line of reagent tip.

All necessary reagents and tips are sold as a complete extraction kit from AutoGen.

Product# AutoGenprep 965 Tissue DNA Extraction Kit

Catalog#AGKT965TD (for up to 3,840 Tissue samples)

Reagents

Following table shows the necessary reagent and their bottle numbers.

Reagent Code	Reservoir No.	Function	Reagent Tip No.
M1 [#] *	Manual (or 1*)	Tissue Digestion	(A*)
Pro K [#]	<i>To be dissolved in M1[#]</i>	Tissue Digestion	-
M2 *	Manual (or 2*)	Tissue Digestion	(B*)
R3	3	Deproteinization	C
R4	4	DNA Precipitation	D
R5/6/7	5, 6 (and 7 ^{**})	DNA Wash	E
R8	8	RNA Digestion	G
R9	9	DNA Resuspension	H

Appropriate amount of Proteinase K should be dissolved into correct volume of M1 at a final concentration of 0.4mg/ml* for the overnight digestion of mouse tail samples.

* Proper Pro K concentration may vary depending on the type of samples.

* If you select Digest process (to have AutoGenprep 965 dispense digestion reagents), M1 and M2 should be prepared onto Reservoir No. 1 and 2 respectively.

** If you select 3 time wash (Standard setting is for 2 time wash), R5/6/7 (DNA Wash) should be prepared onto Reservoir No. 7 also.

Loading reagent

Transfer the necessary amount of reagents into the specially designed reagent reservoirs just before you start run. Use the same reservoir for the same reagents. If you want to use a different reagent from the previous run, discard remaining reagent, wash the reservoir with dd H₂O and rinse it with new reagent you will use before transferring new reagent.

NOTE: Use freshly prepared Proteinase K solution (by dissolving Pro K powder in M1) and M3 (Deproteinization) for every run. Do not top off the Reservoirs which contain the residual reagents from the previous run. These reagents will lose their original activity very quickly.

Reagent Volume for Operation

Reagent Number (Reagent)	Reservoir No.	Volume Sample (ml)	Necessary minimum volume in Reservoir (ml)	Necessary volume (ml) (number of sample)			
				96	192	288	384
M1 Tissue Digestion	Manual (or 1 #)	0.15	25 +	15 (39)	29 (54)	44 (68)	58 (83)
M2 Tissue Digestion	Manual (or 2 #)	0.15	25 +	15 (39)	29 (54)	44 (68)	58 (83)
R3 Deproteinization	3	0.18	25 +	42	60	77	94
R4 DNA Precipitation	4	0.20	25 +	45	64	83	102
R5/6/7 DNA Wash	5	0.45	25 +	69	112	155	198
Reagent 5/6/7 DNA Wash	6	0.45	25 +	69	112	155	198
Reagent 5/6/7 DNA Wash	(7*)	(0.45)	(25) +	(69)	(112)	(155)	(198)
R8 RNA Digestion	8	0.05	25 +	30	35	40	45
R9 DNA Resuspension	9	0.05	25 +	30	35	40	45
Waste fluid volume (ml)	Waste Bottle 1			0	0	0	0
1 time alcohol wash	Waste Bottle 2			114	227	340	454
2 times alcohol wash (default)				157	314	470	627
3 times alcohol wash				200	400	600	800

Necessary volume of R5/6/7 depends on your selection of parameter for DNA wash. Standard parameter is 2 time washes.

* If you select optional Digest process, the required volume in the parentheses of M1 and M2 should be prepared into Reservoir No. 1 and 2 respectively.

Reagent Volume for Operation

Reagent Number (Reagent)	Reservoir No.	Volume Sample (mL)	Necessary Minimum Volume in Reservoir (mL)	Necessary Volume (mL) (Number of Sample) Including Minimum 25mLs						
				24	48	72	96	192	288	384
M1 Tissue Digestion	Manual (or 1)	0.15	25	29	33	36	40	54	69	83
M2 Tissue Digestion	Manual (or 2)	0.15	25	29	33	36	40	54	69	83
R3 Deproteinization	3	0.18	25	30	34	38	43	60	77	94
R4 DNA Precipitation	4	0.20	25	31	35	40	45	64	83	102
R5/R6/R7 DNA Wash	5	0.45	25	36	47	58	69	112	155	198
R5/R6/R7 DNA Wash	6	0.45	25	36	47	58	69	112	155	198
<i>R5/R6/R7* DNA Wash</i>	<i>(7)</i>	<i>(0.45)</i>	<i>(25)</i>	<i>(36)</i>	<i>(47)</i>	<i>(58)</i>	<i>(69)</i>	<i>(112)</i>	<i>(155)</i>	<i>(198)</i>
R8 RNA Digestion	8	0.05	25	27	28	29	30	35	40	45
R9 DNA Resuspension	9	0.05	25	27	28	29	30	35	40	45
Waste Fluid Volume (mL)	Waste Bottle 1			0	0	0	0	0	0	0
1 time alcohol wash	Waste Bottle 2			29	57	86	114	227	340	454
2 times alcohol wash (default)				40	79	118	157	314	470	627
3 times alcohol wash				50	100	150	200	400	600	800

*Necessary volume of R5/R6/R7 depends on your selection of parameters fro DNA wash. Standard parameter is 2 time washes.

If you select optional Digest process, the required volume in the parentheses of M1 and M2 should be prepared into Reservoir No. 1 and 2 respectively.

Start operation

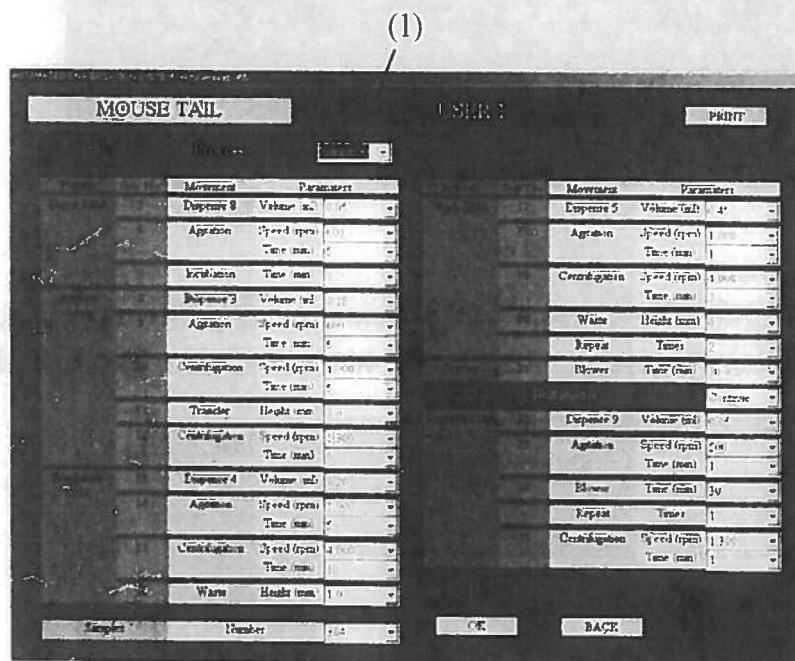
1. Access to Mouse tail DNA protocol

- (1) Select [Mouse tail DNA] protocol from the protocol selection screen (screen 3-2).
- (2) Select the [AUTOMATED RUN] for normal operation from the operation mode selection screen (screen 4). If you want to restart the AutoGenprep 965, select the [RE-START RUN].

2. Input parameter

- (1) Select a parameter file you want to use from the parameter file selection screen (screen 5).
- (2) When you select a parameter file, screen of "Extraction" steps shown below displays. Modify a parameter, if necessary.
- (3) Input sample number

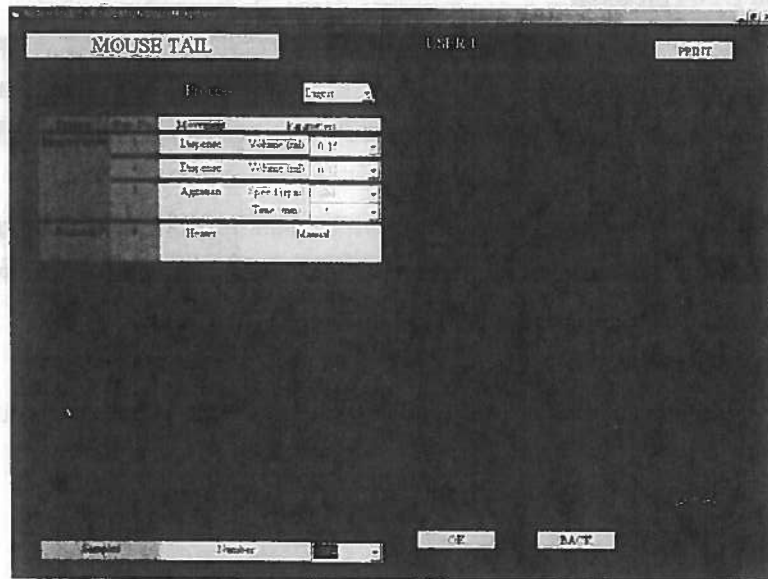
Valid sample number is multiple of 24, minimum 24 and maximum 384.



- (4) If you want to operate from "Digest—the process for adding digestion reagents", first click [▼] marked (1), and then select a process using up/down key on the key board of PC.

(5) When you select “Digest” screen below displays. Input dispensing volumes of M1 and M2.

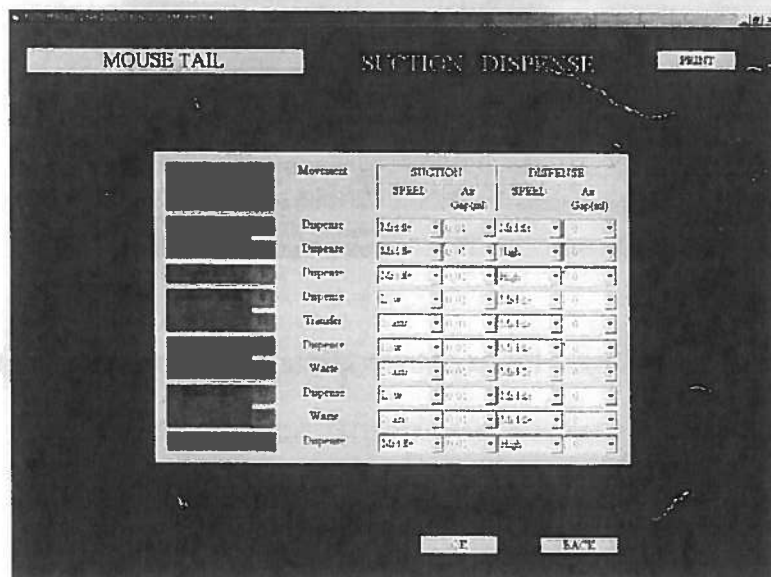
(1)



1. Click [OK] icon to save parameter modifications and enter the next operation step.

4. Dispensing factor adjustment

When you click [SUCTION/DISPENSE] icon in the parameter selection screen, screen shown below displays. Change the factor, if necessary. Refer to page 4-3 of the Operator Manual for details.



Process and Standard Parameters of Mouse tail DNA protocol

1. Digestion process (Optional steps for dispensing tissue digestion reagents)

1/3

Process	Step No.	Movement	Description	Limit of input	Standard
Suspend tissue sample	1	Dispense	Dispense M1 from reservoir No.1	0 ⁺ or 0.02-0.90	0.15 ml
	2	Dispense	Dispense M2 from reservoir No.2	0 ⁺ or 0.02-0.90	0.15 ml
	3	Agitate	Transfer microplate to agitator		
			Agitate	1000,900,800,700,600,500,400,300	500 pm*4
0 ⁺ -20			1 min		
		Transfer microplate to rack table			
Proteinase K digestion	4		Remove 96 deep well plates from the AutoGenprep965, cover the plates with adhesive sheets and incubate at 60-65C for 14-16 hours with sharking in an external incubator.		

Note
Make sure each well of the 96 deep well plates is perfectly covered to avoid evaporation of reagents during incubation for the digestion.

2. Extraction process (DNA extraction)

Digest RNA	5	Dispense	Dispense R8 from reservoir No.8	0 ⁺ or 0.02-0.90	0.05 ml
	6	Agitate	Transfer microplate to agitator		
			Agitate	1000,900,800,700,600,500,400,300	600 rpm*4
			0 ⁺ -20	5 min	
		Transfer microplate to rack table			
	7	Incubate		0 ⁺ -20	0 min
Remove debris	8	Dispense	Dispense R3 from reservoir No.3	0 ⁺ or 0.02-0.90	0.18 ml
	9	Agitate	Transfer microplate to agitator		
			Agitate	1000,900,800,700,600,500,400,300	600rpm*4
			0 ⁺ -20	5 min	
			Transfer microplate to rack table		
	10	Centrifuge	Transfer microplate from rack table to Centrifuge		
Centrifuge to pellet debris			4,000,3,800,3,600,3,400,3,200,3,000,1,300	4,000 rpm	
0 ⁺ -20			5 min		
		Transfer microplate to rack table			
Need 2 operations for complete 4 microplates					

220min
 OBS

Process	Step No.	Movement	Description	Limit of input	Standard	
DBS -1 STEP	11	Liquid transfer	Transfer sample solution from sample microplate to stock microplate using the pipette	-2-30 (Height of tip)	3.0 mm*4 ↘	
	12	Centrifuge	Spin down tube to the bottom of microplate.	4,000,3,800, 3,600, 3,400,3,200, 3,000, 1,300 0 ⁺ -20	1,300rpm 0 min*4	
Precipitate DNA ↓	13	Dispense	Dispense R4 from reservoir No.4	0 ⁺ or 0.02-0.90	0.20 ml	
	14	Agitate	Transfer microplate to agitator			
			Agitate for precipitate DNA	1000,900,800,700, 600,500,400,300	1,000 rpm	
				0 ⁺ -20	5 min	
	15	Centrifuge	Transfer microplate to rack table			
			Transfer microplate to centrifuge			
			Centrifuge to pellet DNA	4,000,3,800, 3,600, 3,400,3,200, 3,000, 1,300 0 ⁺ -20	4,000 rpm 10 min	
			Transfer microplate to rack table			
				Need 2 operation for complete 4 microplates		
	16	Waste	Waste supernatant using pipette	-2-30 (Height of tip):	1.0 mm	
Wash DNA(1) -1	17	Dispense	Dispense R5/6/7 from reservoir No.5, 6 or 7	0 ⁺ or 0.02-0.90	0.45 ml	
	18	Agitate	Transfer microplate to agitator			
			Agitate for wash DNA	1000,900,800,700, 600,500, 400,300 0 ⁺ -20	1,000 rpm 1 min	
			Transfer microplate to rack table			
	19	Centrifuge	Transfer microplate to centrifuge			
			Centrifuge for collect DNA	4,000,3,800, 3,600, 3,400,3,200, 3,000, 1,300 0 ⁺ -20	4,000 rpm 3 min	
			Transfer microplate to rack table			
			Need 2 operation to complete 4 microplates			
	20	Waste	Waste supernatant using pipette	-2-30 (Height of tip)	0 mm	
	Wash DNA(2,3)		Repeat	Specified repeat time of DNA wash(from step 17 to 20) [If you select 2 (3) times wash, R5/6/7 is dispensed from Reservoir No. 5, 6 and 7]	1, 2,3	2

Process	Step No.	Movement	Description	Limit of input	Standard
Dry up	21	Dry-up	Transfer stock microplates to blower		
			Dry-up DNA by blower	0* ¹ -60	60min
			Select dissolving steps Continue: continue dissolving steps Pause: pause the steps until click [START] icon Skip: skip the steps		Continue
Dissolve DNA	22	Dispense	Dispense R9 from reservoir No. 9	0* ¹ or 0.013-0.90* ²	0.05 ml
	23	Agitate	Transfer microplate to Agitator		
			Agitate for dissolve DNA	1000,900,800,700,600,500,400,300	500 rpm
				0* ¹ -20	1 min
	24	Dry-up	Transfer stock microplate to blower		
		Dry-up DNA by using blower	0* ¹ -60	0min	
	-	Repeat time	Specified repeat time from step 25 to 24	1-20	1
	25	Centrifuge	Centrifuge to collect DNA solution	4,000,3,800,3,600,3,400,3,200,3,000,1,300	1,300 rpm
				0* ¹ -20	1 min

*1: When you select <0> or <-> in the instrument skip the step.

*2: More than 0.020ml is recommended as 0.013-0.019 is uneven.

Standard Settings of the adjusting factor of nozzle

Select aspiration and dispensing speeds of reagent dispensing steps from 3 options (High, Middle, or low).
 Aspiration speed of waste fluid and fluid transfer steps has four options (High, Middle, Low and 2-axis).
 When you select the "2-axis" the pipette tip keep moving down while aspirating fluid from wells of microplate.

Input range of air gap of aspiration and dispensing is 0-0.10(ml, allow to input two decimal place)

1/3

	Step No.	Operation	Item	Standard setting
<i>Suspend tissue sample (Optional steps)</i>	1	Dispense M1	Aspiration speed	Middle
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0
	2	Dispense M2	Aspiration speed	Middle
			Air gap after aspiration (ml)	0.01
			Dispensing speed	High
			Air gap after dispensing(ml)	0
Digest RNA	5	Dispense R8	Aspiration speed	Middle
			Air gap after aspiration (ml)	0.01
			Dispensing speed	High
			Air gap after dispensing(ml)	0
Remove debris	8	Dispense R3	Aspiration speed	Low
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0

0.01
 ES
 3/24/11

	Step No.	Operation	Item	Standard setting
Remove debris	11	Transfer supernatant	Aspiration speed	2-axis
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0
Precipitate DNA	13	Dispense R4	Aspiration speed	Low
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0
	16	Discard waste fluid	Aspiration speed	2-axis
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0
Wash DNA	17	Dispensing R5/6/7	Aspiration speed	Low
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0

	Step No.	Operation	Item	Standard setting
Wash DNA	20	Discard waste fluid	Aspiration speed	2-axis
			Air gap after aspiration (ml)	0.01
			Dispensing speed	Middle
			Air gap after dispensing(ml)	0
Dissolve DNA	22	Dispense R9	Aspiration speed	Middle
			Air gap after aspiration (ml)	0.01
			Dispensing speed	High
			Air gap after dispensing(ml)	0

Supplemental Protocol

Recommended Conditions and Parameters for Blood Punch DNA Extraction

A) Sample Preparation

1. Put ^{1 or 2} blood punches into each well of 96 deep-well plate.
2. Dissolve Pro K into M1(Tissue Digestion Solution A) at 0.5mg/ml.
2. Add 100µl of M1 containing Proteinase K to the sample
3. Add 100µl of M2 (Tissue Digestion Solution B)
4. Seal the plate and Incubate at 65C overnight with shaking
5. Unseal and load the plate on the AutoGenprep 965

B) Parameters

Please change the parameters as following table and save the customized parameters file in File#2.

**File#2 – 8 are available for saving the customized file.*

Step.	Movement	Default Value	Recommended Value
10	Centrifuge to pellet debris	4,000,3,800, 3,600, 3,400, 3,200,3,000, 1,300	4,000 rpm
		0 -20	5min
11	Liquid Transfer	-2-30 (Height of tip)	3.0 mm

20 min
1.0mm (1 new)

Sample QC

1. Check sample on gel for sample degradation.
2. Normalize sample to 1ug in at least 30 ul with pico-green assay or nanodrop.
*Pico-green is more accurate than nanodrop.
- a. >1ug is needed for library-prep on robot.
- b. >100ng is needed for manual library-prep.



*

Red arrow shows degraded gDNA.

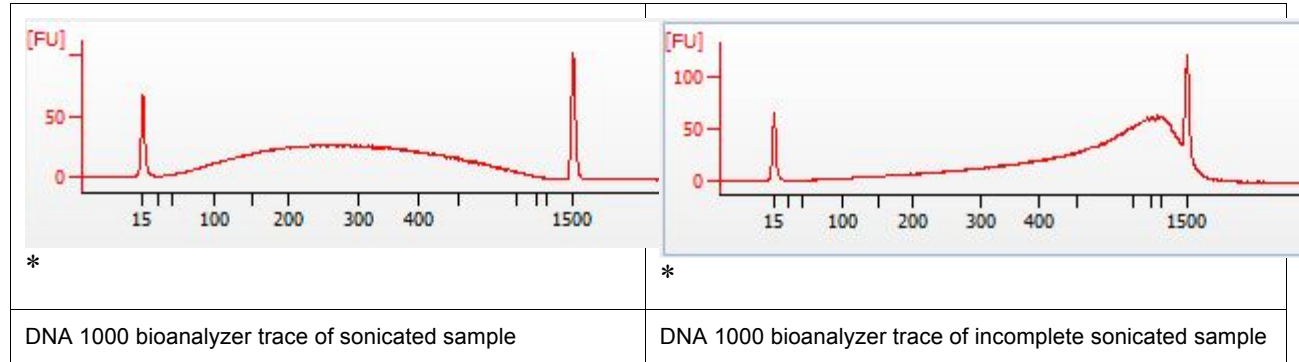
3. If sample looks degraded, replace with non-degraded sample.

Sonication

1. Turn on Covaris S2.
2. Turn on the chiller
3. Open "TruSeq Enrichment" protocol.
- a.
 4. Degas Covaris S2 for 30 minutes.
 5. Set chiller to 2.8 degrees.
. Set the chiller temperature lower to cool water bath faster
 6. Make sure water bath is at 5 to 6 degrees Celsius before starting sonication.
. Adjust chiller temperature to maintain 5 to 6 degrees Celsius for water bath.
 7. Top off sample to at **least 30ul** if it is <30ul using Molecular Biograde Water (MBG).
 8. Pipette sample into covaris micro tube.
. **DO NOT** load more than 120ul into the micro tube.

*sample won't be submerged unless water bath lever is over the cap level

9. Place micro tube into holder.
10. Place micro tube/holder combo into covaris S2 sonicator.
11. Check to make sure water bath level is over sample liquid level
- a. **DO NOT** fill water bath to past the bottom of the micro tube cap
12. Press "Enter" to start sonication.
 - . If the water bath reaches pass 6 degrees, the program will show a warning.
- i. **DO NOT abort** the sonication.
 - *if sonication is aborted, the sonication will be incomplete. There will be a shift in base-pair distribution towards larger fragments.
- ii. Adjust chiller temperature.
- iii. Continue with sonication.



Library Prep (using the KAPA Library Preparation Kit)

KAPA Biosystems-with bead DNA library construction			
DNA input: 50uL at 20ng/uL			
End Repair			
End Repair Master Mix	x1 (uL)		
water	8		
10X KAPA End Repair Buffer	7		
KAPA End Repair Enzyme	5		
	20		
		Add 20uL of the ER master mix to 50uL sonicated DNA.	
		Incubate at 20C for 30 min.	

Clean up					
1		add 120 ul of Ampure XP beads to each 70 ul end repair reaction			
2		mix well, incubate at RT for 15 min.			
3		capture beads on magnet until the liquid is clear. Remove the supernatant.			
4		On the Magnet, add 200 ul of 80% ETOH: incubate at RT for 30 sec, remove the ethanol. Repeat step 4.			
5		allow the beads to dry sufficiently for all ethanol to evaporate.			
A-tailing					
A-tailing master mix	x1				
water	42				
KAPA 10X A-tailing buffer	5				
KAPA A-tailing enzyme	3				
	50				
		add 50 ul of A-tailing master mix to each dried beads carrying the end repaired DNA. mix it and incubate for 30 min at 30C.			
Clean up					
1		To each A-tailing reaction, add 90 ul of 20% PEG 8000/ 2.5M NaCl solution			
2		Mix thoroughly and incubate at RT for 15 min.			
3		capture beads on magnet until the liquid is clear. Remove the supernatant.			
4		On the Magnet, add 200 ul of 80% ETOH: incubate at RT for 30 sec, remove the ethanol. Repeat step 4.			
5		allow the beads to dry sufficiently for all ethanol to evaporate.			
Ligation					
KAPA ligation mix	x1				
water	34				
KAPA 5X ligation buffer	10				
KAPA Ligase	5				
	49				
		add 49 ul of ligation master mix to each reaction. add 1 ul of 50 uM adapter according to the table above. Mix thoroughly to resuspend the beads. incubate at 20C for 15 min.			

Ligation Clean up				
1	To each ligation reaction, add 50 ul of 20% PEG 8000/ 2.5M NaCl solution			
2	Mix thoroughly and incubate at RT for 15 min.			
3	capture beads on magnet until the liquid is clear. Remove the supernatant.			
4	On the Magnet, add 200 ul of 80% ETOH: incubate at RT for 30 sec, remove the ethanol.			
	Repeat step 4.			
5	allow the beads to dry sufficiently for all ethanol to evaporate.			
6	resuspend in 50 ul of elution buffer.			
	repeat clean up one more time, and final elution volume after second clean-up should be 20 ul with water prior to PCR.			
	The ligated samples should now contain no more beads.			
PCR	x1 (ul)			
2X KAPA HiFi Hotstart ready mix	25			
10uM TS-PCR Oligo 1 + 2	5			
	30			
To the 20uL ligated product, directly add 25uL Hotstart ready mix and 5uL oligo (1+2) to each reaction.				
Do not vortex. Gently pipette up and down 5 times.				
		Temp (C')	Time (sec)	
		1	98	45
		2	98	15
		3	60	30
		4	72	30
		5	repeat step 4 times	
		6	72	60
		7	4	forever
Clean up				
1	add 50 ul of Ampure XP beads to each 50 ul PCR product.			
2	mix well, incubate at RT for 15 min.			
3	capture beads on magnet until the liquid is clear. Remove the supernatant.			
4	On the Magnet, add 200 ul of 80% ETOH: incubate at RT for 30 sec, remove the ethanol.			
	Repeat step 4.			
5	allow the beads to dry sufficiently for all ethanol to evaporate.			
6	Add 50 ul of H2O and mix throughly			
7	Incubate at RT for 2 minutes to release the DNA from the beads			

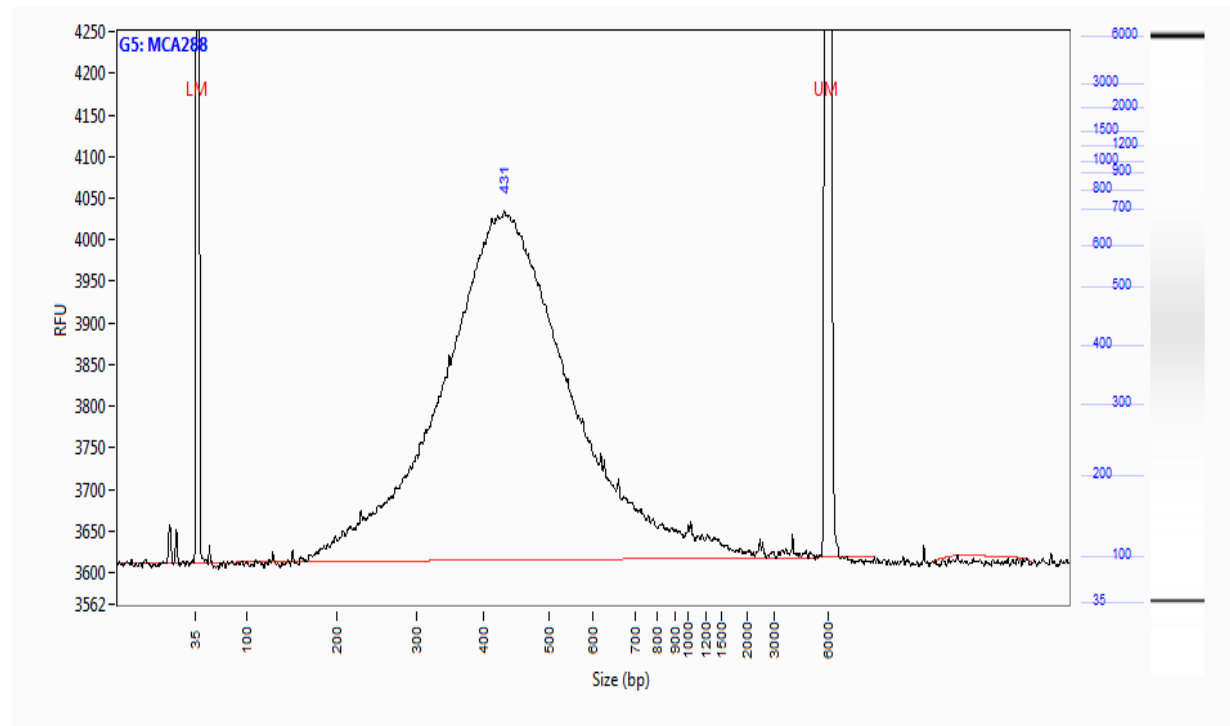
8	Put the plate back on the magnet and wait 3 minutes.
9	Transfer 50uL eluate to new wells.

Library Prep QC

Run the samples on the Fragment Analyzer using the NGS Fragment Analysis High Sensitivity Kit or the DNA 1000 Kit on the Bioanalyzer.

Good libraries should have at least 1ug and should look something like this:

DNA that has been pre-washed should have an average size of about 420-460bp. Without the pre-wash, the average size should be 290 to 370 bp



Pooling

Pool 1000ng of libraries total for each exome capture reaction. If there are 4 samples being pooled, it would be 250ng for each library. Before pooling, make sure each library has a different index for multiplexing.

Exome Capture

Exome capture

5 μ l of 1 mg/ml COT DNA

1 μ l of 1,000 μ M TS-HE Universal Oligo 1 (Forward blocker)

1 μ l of 1,000 μ M barcode blocker oligo mix

1 mg of pooled library DNA

make a hole on the cap of the tube

dry the amplified sample library/COT DNA/TS-HE Oligos in a DNA vacuum concentrator on high heat (60°C).

7.5 μ l of 2X Hybridization Buffer (vial 5)

3 μ l of Hybridization Component A (vial 6)

10.5 μ l total

Vortex for 10 seconds spin down for 10 seconds

Denature at +95°C for 10 minutes in thermocycler

Add 4.5 μ l aliquot of SeqCap EZ Library to the sample

Don't let the temperature drop too much

15 μ l total volume

Incubate in a thermocycler at +47°C for 64 - 72 hours for the hybridization

prepare seq capture wash buffer (RT good for 2 weeks)

10X Wash Buffer I (vial 1) 33 μ l to 297 μ l H₂O (110 μ l) 47°C (220 μ l) RT

10X Wash Buffer II (vial 2) 22 μ l to 198 μ l H₂O RT

10X Wash Buffer III (vial 3) 22 μ l to 198 μ l H₂O RT

10X Stringent Wash Buffer (vial 4) 44 μ l to 396 μ l H₂O 47°C

2.5X Bead Wash Buffer (vial 7) 55 μ l to 495 μ l H₂O RT

Preparing the Streptavidin Dynabeads

Allow the Streptavidin Dynabeads to warm to room temperature for 30 minutes prior to use.

vortex beads for 15 seconds.

100 μ l of beads / each capture

Place the tube in a DynaMag-2 device , remove and discard the liquid

Wash the beads with 200 μ l of 1X Bead Wash Buffer x 2 times

resuspend the beads in 100 μ l of 1X Bead Wash Buffer

Remove and discard the liquid.

Bind the captured DNA immediately the Streptavidin Dynabeads.

Do not allow the Streptavidin Dynabeads to dry out.

Binding and Washing

Transfer the hybridization samples to the Streptavidin Dynabeads

Mix by pipetting up and down 10 times.

Don't let the temperature drop too much

Placing the tubes containing the beads and DNA in an oven set to +47°C for 45 minutes

Mix the samples every 15-minute intervals to ensure that the beads remain in suspension.

Wash with 100 μ l of 1X Wash Buffer I heated to +47°C

Remove and discard the liquid

Wash twice with 200 μ l of 1X Stringent Wash Buffer heated to +47°C.

Incubate at +47°C for 5 minutes.

Wash with 200 μ l of room temperature 1X Wash Buffer I

vortexing for 2 minutes.

Wash with 200 μ l of room temperature 1X Wash Buffer II

vortexing for 1 minute.

Wash with 200 μ l of room temperature 1X Wash Buffer III

vortexing for 30 seconds.

Remove and discard the liquid once clear.

Resuspend the bead-bound captured sample with 50 µl PCR grade water
Store at -20°C.

Amplifying Captured DNA Using LM-PCR

2 tubes of LM-PCR

50 µl Phusion High-Fidelity PCR Master Mix (2x)

2 µl TS-PCR Oligo 1+2 , 100 µM

20 µl captured DNA with bead

28 µl PCR grade water

100 µl total

Step 1: 30 seconds @ +98°C ·

Step 2: 10 seconds @ +98°C ·

Step 3: 30 seconds @ +60°C ·

Step 4: 30 seconds @ +72°C ·

Step 5: Go to Step 2, repeat 17 times ·

Step 6: 5 minutes @ +72°C

·Step 7: Hold @ +4°C Store reactions at +4°C

Cleaning up the Amplified Captured DNA with Qiagen QIAquick PCR Purification Kit

Mix PCR sample with 500 µl Qiagen buffer PBI

Add to QIAquick PCR Purification column.

Spin and discard liquid

Add 750 µl of buffer PE to the column. Centrifuge at 16,000 x g for 1 minute.

Discard the flow-through and Centrifuge the column for an additional minute.

Add 50 µl of buffer EB directly to the column matrix.

Transfer the column to a 1.5 ml microcentrifuge tube. stand for 5 minute.

Centrifuge at 16,000 x g for 1 minute to elute the DNA.

Check the amplified captured DNA on 'Fragment analyzer' with 1:10 dilution

We would like the concentration more than 10 ng/ul.

the yield more than 500 ng

the size between 250-400bp

Exome Capture QC

Measuring Enrichment Using qPCR

4.5 µl of PCR grade water ·

7.5 µl of SYBR Green Master (2X) ·

1 µl of NSC Assay forward + reverse primer (0.6 µM) ·

1 µl ROS dye 1:500 dilution

1 µl amplified captured DNA 5 ng /ul

NSC Assay forward + reverse primer set :

NSC-237

NSC-247

NSC-268

NSC-272

Water as neg control

Ref 103 as pos control

Triplicates / each sample

program:

95°C x 3' 1 cycle

95°C x 10"

60°C x 20" 40 cycles

melting

95°C x 1'

55°C x 30"

95°C x 30" 1 cycle

	ref 103	Enriched sample
NSC 237	31.63	22.42
NSC 237	31.68	22.31
NSC 237	31.56	22.14
NSC 247	28.16	25.65
NSC 247	28.03	26.67
NSC 247	27.94	25.54
NSC 268	34.77	23.9
NSC 268	34.39	23.5
NSC 268	34.2	23.2
NSC 272	35.32	28.28
NSC 272	34.93	28.25
NSC 272	35.05	28.39

Quantitation: qPCR

After library preparation and/or exome capture, the molarities of the samples are determined by quantitative PCR. We will use this information to make further dilutions for denaturation and clustering in cBot.

Before you start library quantification and preparing sample dilutions, take out the KAPA qPCR kit from the -20 freezer to the 4C refrigerator (or cold room) to make sure the master mix is thawed thoroughly.

Prepare sample dilutions

Typically, a 1:10000 dilution is sufficient for the samples to fall within range of the qPCR standards. However, a 1:100000 dilution might be needed, if the sample concentration is too high. Take out more strip tubes, if you have more than 8 samples. You may also use a sterile 96-well PCR plate or single 0.5ml centrifuge tubes. The following instructions are for a 1:10000 dilution, 8 samples.

- 1) Using tweezers, take two 8-well strip tubes with caps.
- 2) Label one strip tube as 1:100 and the other as 1:10000.
- 3) Label the caps to identify which sample is going into which tube.
- 4) Add 198ul of molecular biology grade water to both strip tubes.
- 5) Pulse three times and spin the samples.

- 6) Add 2ul of the samples to the strip tube labeled as 1:100.
- 7) Pulse quickly and spin the strip tube.
- 8) Aspirate 2ul from the 1:100 and dispense into the 1:10000 strip tube.
- 9) Pulse quickly and spin the strip tube.

Alternative

- 1) Using tweezers, take two 8-well strip tubes with caps.
- 2) Label one strip tube as 1:100 and the other as 1:10000.
- 3) Label the caps to identify which sample is going into which tube.
- 4) Add 99ul of molecular biology grade water to both strip tubes.
- 5) Pulse three times and spin the samples.
- 6) Add 1ul of the samples to the strip tube labeled as 1:100.
- 7) Pulse quickly and spin the strip tube.
- 8) Aspirate 1ul from the 1:100 and dispense into the 1:10000 strip tube.
- 9) Pulse quickly and spin the strip tube.

Important Note: Dilutions can be made in different ways. Whichever way you decide to prepare them, it is good practice to pulse and spin after each dilution.

Prepare the quantitative PCR plate

Refer to the quantitative PCR Master Mix excel file in the Q: drive (groupshare). Enter the number of samples you have to determine the total number of reactions. (Note: The samples will be quantitated in triplicates.)

Once you have the recipe, prepare the qPCR master mix. Pulse and spin each tube.

- 1) Have an ice bucket ready, along with the Affy PCR cold block.
- 2) Combine primer mix and water first.
- 3) Then add ROX low.
- 4) Take out a semi-skirted Stratagene PCR plate and placed into the Affy PCR cold block, which should be sitting on ice.
- 5) Determine the layout. (columns or rows). Refer to qPCR template in Agilent software.
- 6) Pulse three times and spin the tube.
- 7) Using the Rainin pipette repeater, aliquot 16ul of master mix into the Stratagene semi-skirted 96-well plate. Either point it forward against the well wall or towards yourself. (Note: When you aliquot samples into the wells, the pipette should be dispensed into the opposite side of the well.)
- 8) Pulse quickly and spin the standards and positive controls.
- 9) Aliquot 4ul of standards and positive controls into their respective wells, changing tips each time. (Note: Standards are in triplicates.)
- 10) Pulse quickly and spin the 1:10000 8-strip tube containing your samples.
- 11) Aliquot 4ul of DNA into their respective wells, changing tips each time. (Note: Samples are in triplicates.)
- 12) Use the strip caps and make sure they are properly sealed into the wells to prevent evaporation.
- 13) Spin the plate, pulse quickly, and spin the plate again.

Important Note: Make sure there are no bubbles in the plate. If so, spin the plate or flick the well to rid them.

- 14) Place the plate in the Stratagene instrument.
- 15) Wipe the top of the caps with Kimwipes.

Important Note: Caps tend to pop open during the qPCR run due to evaporation. To minimize this effect, make sure the caps are properly sealed.

Using the Stratagene instrument software

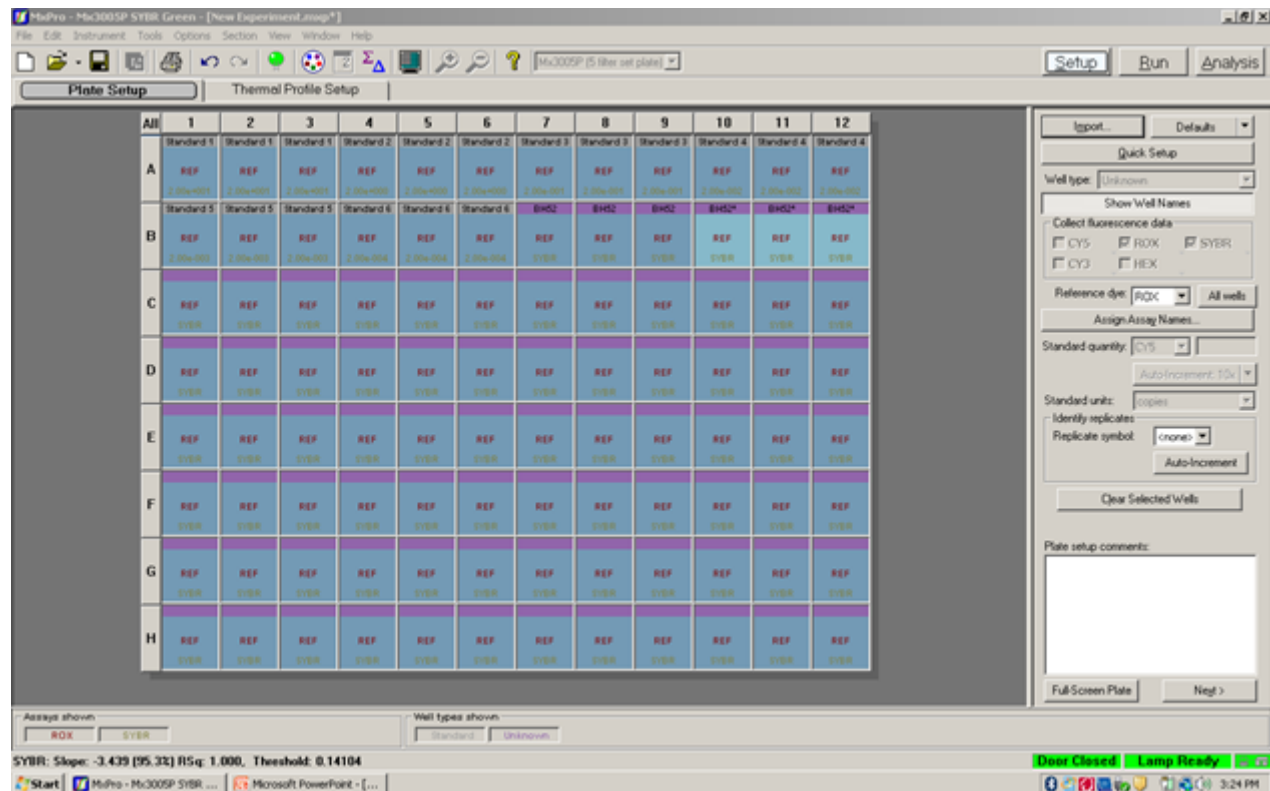
Warm up lamp, if necessary. To turn on lamp, go to

Select new experiment type SYBR Green (with Dissociation Curve).

You can import the plate setup and thermal profile from a template, either rows or columns, depending on your plate layout. The thermal profile details are stated in the KAPA kit instruction manual.

The standards and positive controls should be present.

Important Note: The reference dye is ROX, while the reporter dye is SYBR.



It is up to you to fill in the unknowns.

To label the wells, click assign assay names. All wells are selected by default; if you have less than 96 samples, only select the wells that contain samples to save time.

Analyzing the quantitative PCR results

Click on the Results tab.

There are a few quality checkpoints to ensure the data is reliable and the libraries can proceed further.

This is not in any order of importance:

- Check the efficiency of the standards; it should be around 97% or greater.
- Under area to analyze, click on amplification plots. Look to make sure that the curves or the triplicates are similar; if not, remove the outlier from the graph by deselecting the sample.
- Adjust the threshold, so that all curves are at its exponential growth stage.

Make sure all the unknowns are within the standard curve concentrations; otherwise, the data is useless.

Look at the dissociation curves. Typically, you should only see something similar to this, two peaks representing two types of samples, libraries and standards.

Export the text report to an Excel file.

Calculate the molarity in pM by using the following formula: (refer to the KAPA SYBR qPCR kit manual for more details)

average quantity x dilution factor x (452/average insert size)

The average insert size can be determined by running the Bioanalyzer DNA 1000 or High Sensitivity chip or Fragment Analyzer NGS Fragment kit.

To determine whether the samples proceed with clustering, we compare the molarity of the qPCR to the molarity of either the Bioanalyzer or Fragment Analyzer. If the numbers are reasonably close, then we continue with the next step.

However, if they are not, we would repeat qPCR or Bioanalyzer/Fragment Analyzer.

*Important Note: All samples must be greater than 2nM to proceed. (There is one exception, but that requires an unconventional protocol. This will be explained later.)

Samples that are moving onto the clustering step will be normalized to 2nM with elution buffer.

Clustering

1. Denaturing

- Combine the following:
 - 2 nM Template DNA (10 µl)*(Ensure samples are not cold for better denaturation)
 - 0.1 N NaOH (10 µl)**
- *For molarity lower than 2nM refer to the low concentration protocol in the shared drive under the Hi Seq folder.*
- ** Start the timer as soon as NaOH is added to the first sample.*
- Vortex (3 X for 2 sec. each time) and Centrifuge.
- Incubate for five minutes at room temperature to denature the template into single strands.
- Add 980 µl of pre-chilled HT1 (Hybridization Buffer) to the denatured samples.*
- *In order to give each sample exactly 5 minutes to denature , try adding the HT1 at the same pace as the NaOH was added ,even if it means going beyond the 5 minutes on the timer.*
- Place the denatured DNA template on ice until you are ready to proceed to final dilution.

2. Dilute Denatured DNA

- Dilute the denatured DNA to 10pM by combining the following:
 - 20pM Denatured DNA: 500 µl*
 - Pre-chilled HT1 :500 µl
- Invert several times to mix the template solution and spin briefly.
- Label an eight-tube strip and load 120ul of DNA into each well.**
- * Spin and vortex the DNA briefly BEFORE proceeding with this step.*

*** The sequence for the DNA on strip-tube should match the sequence in which they are shown on the Google doc. performance and queue sheet. Number the tubes 1 to 8.*

3.Loading the samples

- Prewash the c-Bot .
 - Select the appropriate protocol.
 - Make sure the reagents are thawed before they are loaded onto the c-Bot.
 - Follow instructions on the c-Bot to load the flow cell.*
 - At the end load the eight strip-tube (carefully follow the orientation in which the tubes need to be loaded).
- * -When pulling out the flow cell from the storage tube, forceps should only touch the corners avoiding the holes.*
- Wash the flow cell with deionized water .Use Kim-Wipe to wipe it in UNIDIRECTIONAL strokes ,to avoid scratching.*
- Do not touch the surface of the flow cell. Ensure it is completely clean and dry before it is put on the c-Bot.*
- Holes should always face up, when a flowcell is put on the c-Bot.*

Other things to lookout for :

- Make sure none of the sippers are bent or damaged.
- Ensure that all the sippers are parallel to each other and not crossing over.
- Once the run is complete save all the reagents until the flow cell is running in the sequencer
- When taking out the flow cell be very careful since being wet it might stick to the bottom of the manifold and fall down.
- Clean the flow cell properly with ethanol wipe and water , make sure it is dry and dust free.

HiSeq Sequencing

Sequencing

Before proceeding to the reagent preparation and running the instrument makes sure there is enough space on the computer and a sample sheet has been prepared.

Space:

During the run at the end of each lane, the software checks for available disk space. If disk

space drops below the safe threshold, the software pauses the run and a message alerts

you. If you get this message, you need to make disk space available to continue the run.

When sufficient disk space becomes available, the run automatically resumes.

The HiSeq 2500 instrument computer has a storage capacity of 2.7 TB per flow cell, enough space to hold a 209 cycle run if images are not saved to disk. Data from flow cell A is stored on the D: drive, while data from flow cell B is stored on the E: drive.

Sample sheet: Create a sample sheet for all the samples to be loaded for the run. The different sections must be filled according to the sample information that is accompanying the samples to be processed.

Reagents:

1. For High output (SR50 and PE100) runs:

Pull out reagent 3 and reagent 7 bottles from the freezer to thaw in a water bath overnight at 4 C.

2. On the sequencing day, for high output and rapid runs: remove all the reagents from the freezer and thaw them at 4 degrees.

(For the High output PE100 run pull out the ICB reagent left over from last run and use it first. Split the ICB into two separate 47 ml aliquots each. Reserve one for the spike in.)

3. If performing a multiplexed run, reagent 18 (HP3) must be prepared following the directions in the protocol.

Preparing the Instrument:

1. The instrument should have undergone a maintenance wash PRIOR to the new sequencing run.

2. The vacuum gaskets should be replaced prior to each run.

3. Enter all the run parameters using the HCS interface on the HiSeq computer. For this step:

First base report should be confirmed.

Control if any should be specified.

Under Menu->Tools->Options all the proper drives should appear.

4. Cycle numbers should be entered according to the Run being performed.

5. Thoroughly review the entire information before proceeding with the running of the instrument.

Running the Instrument:

1. Load all the required reagents.

2. Slot the waste lines(only line 5 and 6 for Rapid flow cells) into individual tubes before performing the next step.

3. Perform a water pump on a USED flow cell to confirm proper flow.

4. Ensure the levels of waste in the individual tubes are at level with each other.

5. Next the priming step is performed and again the levels in the tube must be checked for evenness.

6. Remove the used flow cell. Clean the stage with Ethanol wipes. Dry it with a Kim wipe.

7. Place the clustered flow cell onto the stage and engage the vacuum.

8. Pump SBS reagent through to confirm proper flow and check the volumes in the waste.

9. Slot the waste lines back into the waste.(for Rapid it is only line 5 and 6, rest are in clean water bottles next to the instrument)

10. Close the door.

11. Hit start to begin sequencing.



HiSeq[®] 2500 System User Guide

FOR RESEARCH USE ONLY

ILLUMINA PROPRIETARY

Part # 15035786 Rev. A
October 2012



This document and its contents are proprietary to Illumina, Inc. and its affiliates ("Illumina"), and are intended solely for the contractual use of its customer in connection with the use of the product(s) described herein and for no other purpose. This document and its contents shall not be used or distributed for any other purpose and/or otherwise communicated, disclosed, or reproduced in any way whatsoever without the prior written consent of Illumina. Illumina does not convey any license under its patent, trademark, copyright, or common-law rights nor similar rights of any third parties by this document.

The Software is licensed to you under the terms and conditions of the Illumina Sequencing Software License Agreement in a separate document. If you do not agree to the terms and conditions therein, Illumina does not license the Software to you, and you should not use or install the Software

The instructions in this document must be strictly and explicitly followed by qualified and properly trained personnel in order to ensure the proper and safe use of the product(s) described herein. All of the contents of this document must be fully read and understood prior to using such product(s).

FAILURE TO COMPLETELY READ AND EXPLICITLY FOLLOW ALL OF THE INSTRUCTIONS CONTAINED HEREIN MAY RESULT IN DAMAGE TO THE PRODUCT(S), INJURY TO PERSONS, INCLUDING TO USERS OR OTHERS, AND DAMAGE TO OTHER PROPERTY.

ILLUMINA DOES NOT ASSUME ANY LIABILITY ARISING OUT OF THE IMPROPER USE OF THE PRODUCT(S) DESCRIBED HEREIN (INCLUDING PARTS THEREOF OR SOFTWARE) OR ANY USE OF SUCH PRODUCT(S) OUTSIDE THE SCOPE OF THE EXPRESS WRITTEN LICENSES OR PERMISSIONS GRANTED BY ILLUMINA IN CONNECTION WITH CUSTOMER'S ACQUISITION OF SUCH PRODUCT(S).

FOR RESEARCH USE ONLY

© 2012 Illumina, Inc. All rights reserved.

Illumina, IlluminaDx, BaseSpace, BeadArray, BeadXpress, cBot, CPro, DASL, DesignStudio, Eco, GAIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NuPCR, SeqMonitor, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.

This software contains the SeqAn Library, which is licensed to Illumina and distributed under the following license:

Copyright © 2010, Knut Reinert, FU Berlin, All rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- 1 Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- 2 Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- 3 Neither the name of the FU Berlin or Knut Reinert nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Revision History

Part #	Rev.	Date	Description of Change
15035786	A	October 2012	Initial release.

Table of Contents

Revision History	iii
Table of Contents	v
Chapter 1 Overview	1
Introduction	2
Delivery and Installation	4
HiSeq 2500 Configuration and Components	6
Single-Indexed Sequencing Overview	13
Dual-Indexed Sequencing Overview	16
Sequencing Consumables	20
User-Supplied Consumables	26
Starting the HiSeq 2500	27
Chapter 2 HiSeq Software	29
Introduction	30
HiSeq Control Software Interface	32
Run Configuration Screens	38
Pre-Run Setup Screens	50
Initiate Run Screens	54
Run Overview Screen	56
Stopping and Resuming a Run	59
Pausing a Run	63
Run Folders	64
Sample Sheet Overview	66
Available Disk Space	68
Switching Sequencing Modes	69
Switching to High Output Mode	70
Switching to Rapid Run Mode	73
Chapter 3 Performing a High Output Sequencing Run	75

Introduction	76
High Output Sequencing Workflow	77
Preparing SBS Reagents	79
Preparing Indexing Reagents	84
Entering Run Parameters	88
Loading and Priming Reagents	93
Loading a Flow Cell	103
Monitoring the Run	110
Preparing Reagents for Read 2	112
Loading Reagents for Read 2	117
Chapter 4 Performing a Rapid Sequencing Run	121
Introduction	122
Rapid Sequencing Workflow	124
Preparing SBS and Cluster Reagents	126
Preparing Indexing Reagents	134
Checking Pre-Run Fluidics and Run Parameters	135
Loading and Priming Reagents	140
Loading a Flow Cell	151
Monitoring the Run	158
Chapter 5 Post-Run Procedures	161
Introduction	162
Unloading and Weighing Reagents	163
Performing an Instrument Maintenance Wash	164
Performing a Water Wash	169
Idling the Instrument	172
Shutting Down the Instrument	173
Chapter 6 Real Time Analysis	175
Real Time Analysis Overview	176
Monitoring Runtime Statistics	178
Real Time Analysis Output Data	180
Chapter 7 Troubleshooting	185

Introduction	186
Possible Run Setup Problems	187
Performing a Fluidics Check	188
Repeating Read 2 Primer Hybridization	189
If BaseSpace is Unavailable	190
 Index	 191
 Technical Assistance	 195

Overview

Introduction	2
Delivery and Installation	4
HiSeq 2500 Configuration and Components	6
Single-Indexed Sequencing Overview	13
Dual-Indexed Sequencing Overview	16
Sequencing Consumables	20
User-Supplied Consumables	26
Starting the HiSeq 2500	27



Introduction

The Illumina HiSeq® System combines innovative engineering with proven SBS technology to set new standards in output, simplicity, and cost-effectiveness.

The HiSeq includes the following features:

- ▶ **Dual Surface Imaging**—The HiSeq uses a four camera epifluorescence system with cutting-edge scanning technology to enable dual surface imaging.
- ▶ **High-Capacity Reagent Chiller**—The reagent compartment is a high-capacity chiller that holds enough reagents the entire sequencing run.
- ▶ **Integrated Fluidics for Paired-End Runs**—Integrated paired-end fluidics provide reagents from the reagent compartment to the flow cell for Read 2 resynthesis and for indexed sequencing.
- ▶ **Interface Control Options**—The instrument software interface provides options for setting up a run and operating the instrument using the touch screen monitor or the integrated keyboard.
- ▶ **Real Time Analysis**—The instrument software performs real time data analysis on the instrument computer, which allows you to monitor important quality metrics during the run and save valuable downstream analysis time. Downstream analysis of sequencing data can be performed with Illumina and third-party software applications, on IlluminaCompute computing and storage appliances, in the BaseSpace genomics cloud solution, or on a custom infrastructure.
- ▶ **BaseSpace Connectivity**—The HiScan System now features an option to send instrument health and sequencing data to the BaseSpace genomics cloud solution in real time to streamline both instrument quality control and analysis. Real time monitoring of runs from anywhere, including Illumina tech support, enables very fast troubleshooting. BaseSpace facilitates collaboration by enabling you to share results instantly with anyone anywhere in the world. Free alignment and variant calling and the soon to be launched BaseSpace app store provide many easy to use workflows that tailor analysis for diverse biological applications.
- ▶ **Dual Flow Cells**—The HiScan System is equipped with a dual flow cell system. You can sequence a single flow cell or two flow cells with different read lengths simultaneously, and start and stop experiments independently.

Audience and Purpose

This guide contains information about the HiSeq 2500. It provides an overview of instrument components and software features, and full instructions for performing a sequencing run.

This guide is for laboratory personnel and other individuals responsible for the following:

- ▶ Operating the instrument
- ▶ Performing sequencing runs
- ▶ Performing instrument and component maintenance
- ▶ Training personnel on the instrument

Delivery and Installation

An Illumina-appointed service provider will deliver the HiSeq 2500 System, uncrate the system, and place it on the lab bench. The space and bench must be ready in advance of delivery. For more information, see Lab Bench Guidelines in the *HiSeq Site Prep Guide*, Part # 15006407.



CAUTION

Only Illumina-authorized personnel should uncrate, install, or move the HiSeq 2500. Mishandling of the instrument can impact the alignment or damage instrument components.

An Illumina representative will install and align the instrument. If you plan to connect the instrument to a data management system or remote network location, please have the path for data storage selected prior to the date of installation. This allows your Illumina representative to test the data transfer process during installation. For more information, see Network Considerations in the *HiSeq Site Prep Guide*, Part # 15006407.



CAUTION

After your Illumina representative has installed and aligned the HiSeq 2500, *do not* relocate the instrument. Moving the instrument improperly can impact the optical alignment and compromise data integrity. If you need to relocate the HiSeq 2500, contact your Illumina representative.

Crate Dimensions and Contents

The HiSeq 2500 and its components are shipped in three crates. Use the following dimensions to determine the minimum door width required to accommodate the shipping crates.

- ▶ Crate #1 contains the HiSeq 2500 instrument.
- ▶ Crate #2 contains one uninterruptible power supply (Japan, North America, or International version).
- ▶ Crate #3 contains the following items:
 - Workstation PC, monitor, and monitor arm
 - Waste bottle and reagent racks
 - Instrument panels
 - Power cords—One domestic power cord or three international power cords (UK, EU, and China)

Table 1 Crated Dimensions

Measurement	Crate #1: Instrument	Crate #2: UPS	Crate #3: Workstation, Covers, and Power Cords
Width	164 cm (64.5 in.)	77.5 cm (30.5 in.)	104 cm (41 in.)
Height	121 cm (47.5 in.)	51 cm (20 in.)	99.1 cm (39 in.)
Depth	104 cm (41 in.)	61 cm (24 in.)	81.3 cm (32 in.)
Weight	312 kg (688 lbs.)	80.3 kg (177 lbs.)	125 kg (275 lbs.)

HiSeq 2500 Configuration and Components

System Configuration

The HiSeq is shipped with a dedicated workstation that is customized to the latest system requirements. The computer workstation is a dedicated subsystem of the instrument and is not intended for use nor supported as a general-purpose workstation.

The HiSeq system comprises the HiSeq instrument, which includes a monitor, a dedicated workstation that is customized to the latest system requirements, a barcode scanner, a keyboard, a mouse, and a universal power supply.

Components

- ▶ **Reagent Compartment**—Contains racks that hold reagents for sequencing runs, or wash solution for instrument washes.
- ▶ **Flow Cell Compartment**—Contains the vacuum-controlled flow cell stage, which holds the flow cell in place during sequencing runs.
- ▶ **Fluidics Compartment**—Contains fluidics pumps that deliver reagents to the flow cell, and then to the waste container.
- ▶ **Sample Loading Station** – In the flow cell compartment, contains the sippers that enable the introduction of the sample for the Rapid Run mode of operation. The sample will need to be introduced using 1.5/1.7mL microtubes from the approved list specified in *User-Supplied Consumables* on page 26.



NOTE

Tubes with screw caps cannot be used in the Template Loading Station.

- ▶ **Enclosed Optics Module**—Contains optical components that enable dual surface imaging of the flow cell, imaging A, C, G, and T at the same time using epifluorescence. The excitation laser beam passes through the objective and the fluorescence is collected through the same objective.
- ▶ **Status Bar**—Uses three colors to indicate instrument status. Blue indicates the instrument is running, orange indicates the instrument needs attention, and green indicates that the instrument is ready to begin the next run.

Figure 1 HiSeq 2500 Main Compartments



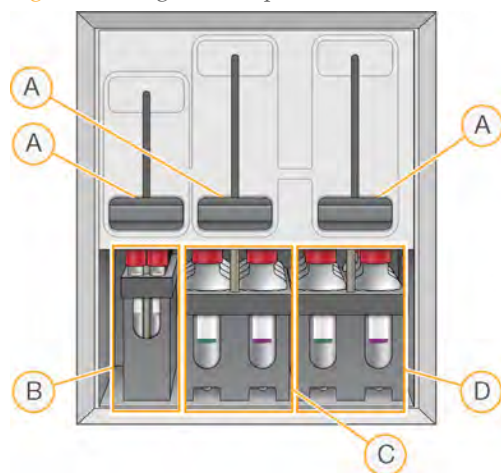
- A Optics Module
- B Flow Cell Compartment
- C Fluidics Compartment
- D Status Bar
- E Reagent Compartment

To set up a sequencing run on the HiSeq 2500, you need to access the reagent compartment and the flow cell compartment.

Reagent Compartment

The reagent compartment is a high-capacity reagent chiller that holds three reagent racks: two for SBS reagents and one for cluster and paired-end reagents. Sipper handles lower the sippers into the reagent bottles.

Figure 2 Reagent Compartment



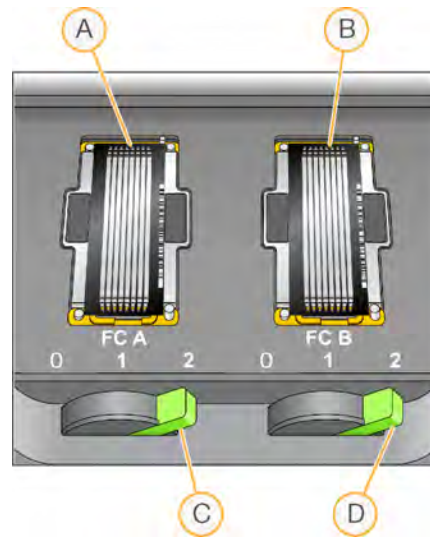
- A Sipper Handles
- B Reagent Rack for Clustering, Paired-End, or Indexing Reagents
- C Reagent Rack for SBS Reagents, Flow Cell A
- D Reagent Rack for SBS Reagents, Flow Cell B

- ▶ **SBS Reagent Racks**—SBS reagent racks hold 250 ml conical bottles. The reagent rack for flow cell A is located in the center position, and the rack for flow cell B is located in the far right position. Each reagent rack has numbered positions that correspond to connections on the fluidics pump.
- ▶ **Cluster, Paired-End, and Index Reagent Rack**—The cluster and paired-end reagent rack, located to the left of racks A and B, has two rows of numbered positions that hold 15 ml conical tubes containing cluster and paired-end reagents and indexing reagents. The left row is for flow cell A, and the right row is for flow cell B.
- ▶ **Reagent Chiller**—The reagent chiller is designed to maintain reagent temperature at 4°C. The reagent chiller is not intended to cool down reagents; therefore, reagents must be cooled before they are loaded onto the instrument.

Flow Cell Compartment

The flow cell compartment houses the flow cell stage, the thermal stations, the vacuum system, and the fluidics connections to each flow cell.

Figure 3 Flow Cell Stage With High Output Flow Cells



- A Flow Cell A
- B Flow Cell B
- C Flow Cell Lever A
- D Flow Cell Lever B

The flow cells are seated on the flow cell stage, which moves in and out of the optics module. For the dual flow cell system in the HiSeq 2500, flow cell A is on the left, and flow cell B is on the right. The flow cell is positioned on the flow cell holder with the inlet and outlet ports facing down, and held in place by a vacuum. The vacuum is controlled by an illuminated flow cell lever in front of each flow cell. The flow cell lever turns green when the vacuum seal is secure.

You cannot open the flow cell compartment if the flow cell stage is not in forward-most position. The flow cell stage is controlled by the HiSeq Control Software. The software will position the flow cell stage in the forward-most position when it prompts you to load or remove a flow cell.

Fuses

Only Illumina Field Service Engineers are qualified to replace the internal fuses. The power entry module includes two input fuses on the high voltage input lines. These

fuses are size 5x20 and are rated for 10 Amps, 250 Vac, slo blow.

Power Cords

The HiSeq 2500 comes with an international standard IEC 60320 C13 receptacle and is shipped with a region-specific power cord.

Hazardous voltages are removed from the instrument only when the power cord is disconnected from the AC power source.

To obtain equivalent receptacles or power cords that comply with local standards, consult a third-party supplier such as Interpower Corporation (www.interpower.com).



CAUTION

Never use an extension cord to connect the instrument to a power supply.

Heat Output

Heat output for a one HiSeq 2500 instrument with personal computer (PC) and monitor is shown below.

Table 2 Heat Output Per Instrument

Component	Measured Power (Watts)	Thermal Output (Btu/h)
Instrument	750	2,600
PC and Monitor	750	1,400
Totals for Instrument, PC, and Monitor	1,500	4,000



CAUTION

Do not operate the HiSeq with any of the panels removed. Do not touch the temperature station in the imaging compartment. The peltier-effect heater used in the stage area is normally controlled between ambient room temperature (22°C) and 85°C. Exposure to temperatures at the upper end of this range could result in burns.

Laser Radiation



The HiSeq is a Class 1 Laser Product that contains two Class 4 lasers. Class 4 lasers present an eye hazard from direct and diffuse reflections. Avoid eye or skin exposure to direct or reflected Class 4 laser radiation. Class 4 lasers can cause combustion of flammable materials and produce serious skin burns and injury from direct exposure.

Do not operate the instrument with any of the panels removed. When the front door of the instrument is open, the laser beam is blocked via safety interlock switches. If you operate the instrument with any panels removed, you risk exposure to direct or reflected laser light.

Figure 4 Class 4 Laser Warning







Barcode Scanner Laser Precautions

A handheld barcode scanner that contains a Class 2 laser is included with the HiSeq sequencing system. The barcode scanner is a Class 2 laser product. Do not stare into the visible-light beam of the barcode scanner.

Compliance and Regulatory Markings

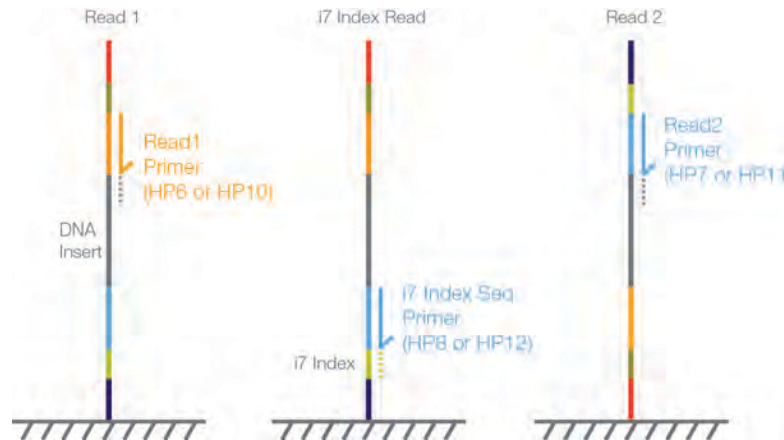
The HiSeq is labeled with the following compliance and regulatory markings.

Marking	Meaning
	<p>This label assures that the product meets requirements of all relevant EU Directives.</p>
	<p>This label assures that the product meets Chinese ROHS regulatory requirements.</p>
	<p>This label assures that the product is compliant with safety standards and has been tested and certified by a third-party organization.</p>
	<p>This label assures that the product meets WEEE no waste directives. Return the instrument to Illumina for disposal.</p>

Single-Indexed Sequencing Overview

Single-indexed sequencing includes one Index Read following Read 1. The details of the workflow will depend on whether you are running High Output or Rapid Run mode.

Figure 5 Single-Indexed Sequencing



Read 1: High Output

Read 1 follows the standard Read 1 sequencing protocol using reagents provided in the TruSeq SBS Kit. The Read 1 sequencing primer is annealed to the template strand during the cluster generation process on the cBot.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP10**, which is used on the cBot in place of HP6. For more information, see the *cBot User Guide*, Part # 15006165.
- ▶ When sequencing single-indexed TruSeq LT/v2 libraries or dual-indexed TruSeq HT libraries, you can use either HP6 or HP10.

Read 1: Rapid Run

Read 1 preparation is performed entirely on the HiSeq instrument, including sample loading, cluster generation and sequencing primer annealing. The Read 1 sequencing primer HP10, included in TruSeq Rapid Cluster Generation Kits, is compatible with

Nextera, TruSeq LT/v2, dual-indexed TruSeq HT libraries, TruSeq Custom Amplicon, and TruSeq Amplicon-Cancer Panel.

Index Read Preparation: High Output

The Read 1 product is removed and the Index 1 (i7) sequencing primer is annealed to the same template strand, producing the Index 1 (i7) Read.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the appropriate sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP12**, which is used in place of HP8.
- ▶ When sequencing single-indexed TruSeq LT/v2 libraries or dual-indexed TruSeq HT libraries, you can use either HP8 or HP12.

Index Read Preparation: Rapid Run

The Read 1 product is removed and the Index 1 (i7) sequencing primer HP12 is annealed to the same template strand producing the Index 1 (i7) Read. HP12 is included in the TruSeq Rapid Cluster Generation Kits and is compatible with dual-indexed Nextera, TruSeq LT/v2, dual-indexed TruSeq HT libraries, TruSeq Custom Amplicon, and TruSeq Amplicon-Cancer Panel.

Index 1 (i7) Read

Following Index Read preparation, the Index 1 (i7) Read is performed.

- ▶ When sequencing single-indexed libraries, the Index Read consists of seven cycles of sequencing.
- ▶ When sequencing dual-indexed libraries (Nextera or TruSeq HT), the Index Read consists of eight cycles of sequencing.

Read 2 Resynthesis

The Index Read product is removed and the original template strand is used to regenerate the complementary strand. Then, the original template strand is removed to allow hybridization of the Read 2 sequencing primer.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP11**, which is used in place of HP7.

- ▶ When sequencing single-indexed TruSeq LT/v2 libraries or dual-indexed TruSeq HT libraries, you can use either HP7 or HP11.
- ▶ For Rapid Run mode, the Read 2 sequencing primer HP11 is contained within the TruSeq Rapid Cluster Generation Kits and is compatible with dual-indexed Nextera, TruSeq LT/v2 and dual-indexed TruSeq HT libraries.

Read 2

Read 2 follows the standard paired-end sequencing protocol using reagents provided in the TruSeq SBS Kit.

Dual-Indexed Sequencing Overview



NOTE

Performing dual-indexed sequencing runs requires HiSeq Control Software (HCS) v2.0.

Dual-indexed sequencing includes two Index Reads following Read 1.

Figure 6 Dual-Indexed Single-Read Sequencing

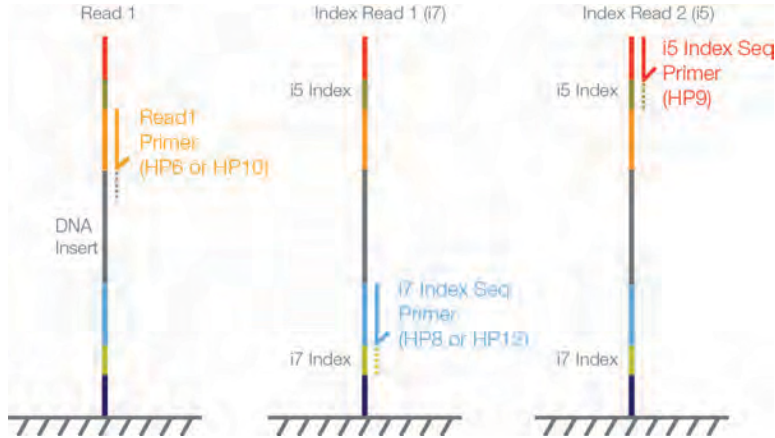
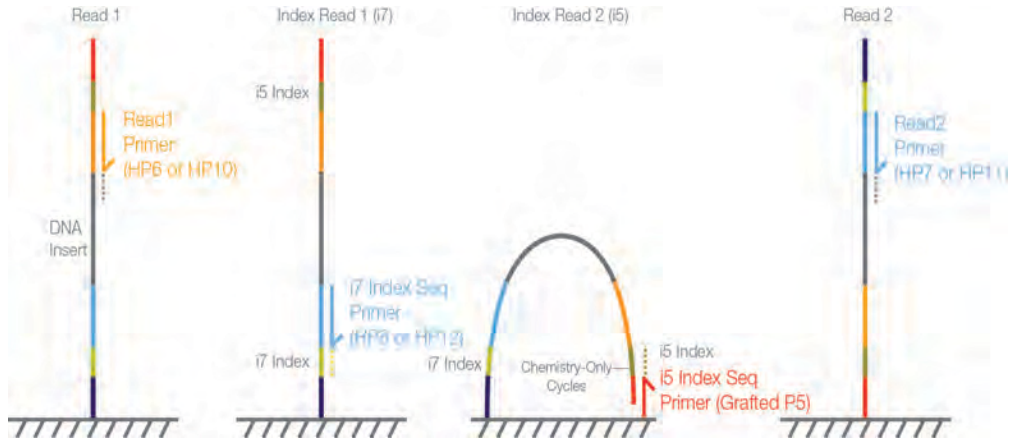


Figure 7 Dual-Indexed Paired-End Sequencing



Read 1: High Output

Read 1 follows the standard Read 1 sequencing protocol using reagents provided in the TruSeq SBS Kit. The Read 1 sequencing primer is annealed to the template strand during the cluster generation process on the cBot.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the appropriate sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP10**, which is used on the cBot in place of HP6. For more information, see the *cBot User Guide*, Part # 15006165.
- ▶ When sequencing dual-indexed TruSeq HT libraries, you can use either HP6 or HP10.

Read 1: Rapid Run

Read 1 preparation is performed entirely on the HiSeq instrument, including sample loading, cluster generation and sequencing primer annealing. The Read 1 sequencing primer HP10, included in TruSeq Rapid Cluster Generation kits, is compatible with Nextera, TruSeq LT/v2, dual-indexed TruSeq HT libraries, TruSeq Custom Amplicon, and TruSeq Amplicon-Cancer Panel.

Index Read Preparation: High Output

The Read 1 product is removed and the Index 1 (i7) sequencing primer is annealed to the same template strand, producing the Index 1 (i7) Read.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP12**, which is used in place of HP8.
- ▶ When sequencing dual-indexed TruSeq HT libraries, you can use either HP8 or HP12.

Index Read Preparation: Rapid Run

The Read 1 product is removed and the Index 1 (i7) sequencing primer HP12 is annealed to the same template strand producing the Index 1 (i7) Read. HP12 is included in the TruSeq Rapid Cluster Generation Kits and is compatible with dual-indexed Nextera, TruSeq LT/v2, dual-indexed TruSeq HT libraries, TruSeq Custom Amplicon, and TruSeq Amplicon-Cancer Panel.

Index 1 (i7) Read

Following Index Read preparation, the Index 1 (i7) Read is performed. When sequencing dual-indexed libraries (Nextera or TruSeq HT), the Index Read consists of eight cycles of sequencing.

Index 2 (i5) Read

The process for Index 2 (i5) Read is different for runs performed on single-read flow cells than for runs performed on paired-end flow cells:

- ▶ **Single-Read Flow Cells**—The Index 1 (i7) Read product is removed and the Index 2 (i5) Sequencing Primer Mix, **HP9**, is annealed to the same template strand. HP9 is provided in the TruSeq Dual Index Sequencing Primer Box and required for dual-indexed runs on single-read flow cells regardless of the library type (Nextera or TruSeq HT). The run proceeds through eight cycles of sequencing.
- ▶ **Paired-End Flow Cells**—The Index 1 (i7) Read product is removed. The Resynthesis Mix, **RMX**, deprotects the grafted P5 primer on the surface of the flow cell and the template anneals to the grafted P5 primer. The run proceeds through an additional seven chemistry-only cycles (no imaging occurs), followed by eight cycles of sequencing.
- ▶ For Rapid Run mode, index 2 (i5) sequencing primer HP9 is required for single read flow cells only and is included in the TruSeq Rapid SR Cluster Generation Kit.

Read 2 Resynthesis

The Index Read product is removed and the original template strand is used to regenerate the complementary strand. Then, the original template strand is removed to allow hybridization of the Read 2 sequencing primer.

- ▶ When sequencing dual-indexed Nextera libraries, you must use the sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box, **HP11**, which is used in place of HP7.
- ▶ When sequencing dual-indexed TruSeq HT libraries, you can use either HP7 or HP11.
- ▶ For Rapid Run mode, HP11 is provided in the TruSeq Rapid PE Cluster Generation Kit and compatible with all libraries.

Read 2

Read 2 follows the standard paired-end sequencing protocol using reagents provided in the TruSeq SBS Kit.

Dual-Indexed Libraries

Dual-indexed libraries are libraries prepared with adapters containing two indices:

- ▶ Index 1 (i7), an eight-base index adjacent to the P7 strand
- ▶ Index 2 (i5), an eight-base index adjacent to the P5 strand

During sample preparation, up to 12 unique Index 1 sequences are arrayed across the columns of a 96-well plate and up to eight unique Index 2 sequences are arrayed down the rows, which create up to 96 uniquely dual-indexed adapters.

Supported Dual-Indexed Libraries

Dual-indexed sequencing on the HiSeq 2500 is supported for libraries prepared with the following Illumina sample prep kits:

- ▶ TruSeq HT Sample Prep Kits
- ▶ Nextera DNA Sample Prep Kits
- ▶ Nextera XT Sample Prep Kits
- ▶ TruSeq Custom Amplicon
- ▶ TruSeq Amplicon-Cancer Panel

Sequencing Consumables

Illumina-supplied consumables are required for sequencing on the HiSeq 2500. The kits that you need depend on the type of run you perform and the type of libraries you are sequencing.

The following table lists the type of reagents required for a High Output run:

Run Type	Library Type	Compatible Consumables
All runs	Any library type	TruSeq SBS Kit
Indexed sequencing run	Any library type	TruSeq Cluster Kit, SR or PE
Paired-end sequencing run	Any library type	TruSeq PE Cluster Kit
All run types	Nextera libraries	TruSeq Dual Index Sequencing Primer Box
Dual-indexed single-read run Single-read flow cell	TruSeq HT libraries	TruSeq Dual Index Sequencing Primer Box

The following table lists the type of reagents required for a Rapid Run:

Run Type	Library Type	Compatible Consumables
All runs	Any library type	TruSeq Rapid SBS Kit
All sequencing runs	Any library type	TruSeq Rapid Duo cBot Sample Loading Kit
All dual/single indexed single read runs	Any library type	TruSeq Rapid SR Cluster Kit
All dual/single indexed paired-end sequencing run	Any library type	TruSeq Rapid PE Cluster Kit

For more information, see *Sequencing Primers By Run Type* on page 24.

SBS Reagents

High Output Mode

Sequencing reagents are provided in the TruSeq SBS Kit. All sequencing runs require either one 200-cycle TruSeq SBS Kit, or one or more 50-cycle TruSeq SBS Kits depending on the total number of cycles in the run.

- ▶ One 200-cycle kit provides sufficient reagents for up to 209 cycles.
- ▶ One 50-cycle kit provides sufficient reagents for up to 58 cycles.

Kit	Catalog #
TruSeq SBS Kit v3 - HS (200 Cycles)	Catalog # FC-401-3001
TruSeq SBS Kit v3 - HS (50 Cycles)	Catalog # FC-401-3002

Rapid Run Mode

Sequencing reagents are provided in the TruSeq Rapid SBS Kit. All sequencing runs require either one 200-cycle TruSeq Rapid SBS Kit, or one or more 50-cycle TruSeq Rapid SBS Kits depending on the total number of cycles in the run.

- ▶ One 200-cycle kit provides sufficient reagents for up to 225 cycles.
- ▶ One 50-cycle kit provides sufficient reagents for up to 75 cycles.

Kit	Catalog #
TruSeq Rapid SBS Kit - HS (200 Cycles)	Catalog # FC-402-4001
TruSeq Rapid SBS Kit - HS (50 Cycles)	Catalog # FC-402-4002

Indexing Reagents

TruSeq Cluster Kits (SR or PE) contain reagents for single indexed libraries only. TruSeq Rapid Cluster Kits (SR or PE) contain reagents for both single and dual indexed libraries. One kit provides sufficient reagents for sequencing one flow cell.

Kit	Catalog #
TruSeq SR Cluster Kit v3 (cBot - HS)	Catalog # GD-401-3001
TruSeq PE Cluster Kit v3 (cBot - HS)	Catalog # PE-401-3001
TruSeq Rapid SR Cluster Kit	Catalog # GD-402-4001
TruSeq Rapid PE Cluster Kit	Catalog # PE-402-4001

**NOTE**

For all High Output sequencing runs using Nextera libraries and for dual-indexed runs using TruSeq HT libraries on a single-read flow cell, additional sequencing primers are required for the Index Reads. For more information, see *Sequencing Primers for Dual-Indexed Libraries* on page 23.

For Rapid Run mode, all required sequencing primers for all the libraries listed are in the above kits.

Paired-End Reagents

All paired-end sequencing runs require paired-end reagents provided in the TruSeq PE Cluster Kit and TruSeq Rapid Cluster Kits. One kit provides sufficient reagents for sequencing one paired-end flow cell.

Kit	Catalog #
TruSeq PE Cluster Kit v3 (cBot - HS)	Catalog # PE-401-3001
TruSeq Rapid PE Cluster Kit	Catalog # PE-402-4001

**NOTE**

Dual-indexed Nextera libraries require the Read 2 sequencing primer provided in the TruSeq Dual Index Sequencing Primer Box. For more information, see *Sequencing Primers for Dual-Indexed Libraries* on page 23.

Sequencing Accessories

High Output

The TruSeq Cluster Kit (cBot - HS) for High Output mode contains an accessories kit, which includes the following components used for setting up a High Output sequencing run on the HiSeq.

- ▶ **ICB Instrument Tube**—One 250 ml bottle used for splitting and preparing ICB.
- ▶ **Flow Cell Gaskets**—Four flow cell gaskets to be installed prior to performing an instrument maintenance wash.
- ▶ **Funnel Caps**—Eight funnel caps used when loading SBS reagents.

Rapid Run

- ▶ **Flow Cell Gaskets**—Four flow cell gaskets, 2 for the back and 2 for the front.
- ▶ **Funnel Caps**—8 funnel caps will arrive as part of the cluster kit.

**NOTE**

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

Sequencing Primers for Dual-Indexed Libraries

Dual-indexed Nextera libraries and dual-indexed TruSeq HT libraries have different sequencing primer requirements:

- ▶ **Nextera libraries**—Nextera libraries require additional sequencing primers provided in the TruSeq Dual Index Sequencing Primer Box. These primers are used throughout the sequencing process, including Read 1 (applied during cluster generation), any Index Reads, and Read 2.
 - Sequencing runs on single-read flow cells require the sequencing primers provided in the single-read kit.
 - Sequencing runs on paired-end flow cells require the sequencing primers provided paired-end kit, even if you plan to perform a single-read run on a paired-end flow cell.
- ▶ **TruSeq HT libraries**—When sequencing on a paired-end flow cell, TruSeq HT libraries can use the sequencing primers provided in the TruSeq Cluster Kit throughout the sequencing process. However, when performing a dual-indexed run on a single-read flow cell, TruSeq HT libraries require the Index 2 Read sequencing primer (HP9) from the Dual Index Sequencing Primer Box (Single-Read).

**NOTE**

Sequencing primers provided in the TruSeq Dual Index Sequencing Primer Box are compatible with all TruSeq and Nextera libraries. If your flow cell contains a combination of TruSeq and Nextera libraries in different lanes, use the sequencing primers in the TruSeq Dual Index Sequencing Primer Box for the entire run.

One TruSeq Dual Index Sequencing Primer Box provides sufficient reagents for sequencing one flow cell.

Kit Name	Catalog #
TruSeq Dual Index Sequencing Primer Box (Single Read)	Catalog # FC-121-1003
TruSeq Dual Index Sequencing Primer Box (Paired End)	Catalog # PE-121-1003

Dual Index Sequencing Primer Box Contents, Single Read

Reagent Name	Storage	Description
HP10	-15° to -25°C	Read 1 Sequencing Primer Mix
HP12	-15° to -25°C	Index 1 (i7) Sequencing Primer Mix
HP9	-15° to -25°C	Index 2 (i5) Sequencing Primer Mix

Dual Index Sequencing Primer Box Contents, Paired End

Reagent Name	Storage	Description
HP10	-15° to -25°C	Read 1 Sequencing Primer Mix
HP12	-15° to -25°C	Index 1 (i7) Sequencing Primer Mix
HP11	-15° to -25°C	Read 2 Sequencing Primer Mix

Sequencing Primers By Run Type

The following table lists required sequencing primers by library type and flow cell type.

Run Type	Read 1 (cBot)	Index 1 (i7)	Index 2 (i5)	Read 2
Nextera libraries Single-Read Flow Cell	HP10	HP12	HP9	--
Nextera libraries Paired-End Flow Cell	HP10	HP12	Grafted P5 on flow cell	HP11
TruSeq HT libraries Single-Read Flow Cell	HP6 or HP10	HP8 or HP12	HP9	--
TruSeq HT libraries Paired-End Flow Cell	HP6 or HP10	HP8 or HP12	Grafted P5 on flow cell	HP7 or HP11
TruSeq LT/v2 libraries Single-Read Flow Cell or Paired-End Flow Cell	HP6 or HP10	HP8 or HP12	--	HP7 or HP11
TruSeq Small RNA libraries Single-Read Flow Cell or Paired-End Flow Cell	HP6 or HP10	HP8 or HP12	--	HP7 or HP11



NOTE

TruSeq HT libraries are dual-indexed libraries. TruSeq LT libraries are single-indexed libraries and include libraries prepared with TruSeq v2 kits.

Version Compatibility and Requirements

For best performance and run results, always use compatible versions of sequencing consumables and software.

Run Type	HiSeq Control Software
Dual-Indexed Sequencing	HCS v2.0 or later

Reagent Kit Version	HiSeq Control Software
TruSeq Cluster Kit v3 (cBot - HS) with Flow Cell v3	HCS v2.0 or later
TruSeq SBS Kit v3 - HS	HCS v2.0 or later
TruSeq Dual Index Sequencing Primer Box (SR or PE)	HCS v2.0 or later
TruSeq Rapid Cluster Kit (cBot - HS) with Rapid Flow Cell	HCS v2.0 or later
TruSeq Rapid SBS Kit - HS	HCS v2.0 or later

For more information about version compatibility, visit the Illumina website at <http://www.illumina.com/VersionCompatibility>.

User-Supplied Consumables

Consumables

Consumable	Supplier	Purpose
1 N NaOH	General lab supplier	Instrument maintenance wash.
Alcohol wipes, 70% Isopropyl or Ethanol, 70%	VWR, catalog # 15648-981 General lab supplier	Cleaning the flow cell and flow cell stage.
Centrifuge tubes, 250 ml	Corning, catalog # 430776	Instrument maintenance wash and water wash.
Conical tubes, 15 ml	Corning, catalog # 430052	Collecting and measuring waste volumes.
Disposable gloves, powder-free	General lab supplier	General use.
Lab tissue, low-lint	VWR, catalog # 21905-026	Cleaning the flow cell holder.
Lens paper, 4 x 6 in.	VWR, catalog # 52846-001	Cleaning the flow cell.
Water, laboratory-grade, 18 M Ohm	Millipore	SBS reagent rack, position 2. Water wash.

Microtubes

The following microtubes are acceptable for use with the sample loading station for Rapid Runs.

Consumable	Supplier	Purpose
MicroCentrifuge Tube, 1.5mL	VWR, catalog #20170-038	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.7mL, capless	VWR, catalog #20170-575	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.5mL	VWR, catalog #20170-650	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.5mL	VWR, catalog #89000-028	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.5mL	AXYGEN, catalog #MCT-150-C	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.7mL	AXYGEN, catalog #MCT-175-C	Sample(template) loading and washing the template fluidic line.
MicroCentrifuge Tube, 1.7mL	Sorenson BioScience, catalog #16070	Sample(template) loading and washing the template fluidic line.

Starting the HiSeq 2500

Use the following instructions to start the HiSeq 2500 and instrument computer.

- 1 Start the instrument control computer.
- 2 Log on to the operating system using the default user name and password. Wait until it has completely loaded.
 - User name: sbsuser
 - Password: sbs123


If the default values do not work, consult your IT personnel for the user name and password for your site.

- 3 Switch the main power switch to the ON position. If you are facing the front of the instrument, the power switch is located on the left side.
- 4 Wait for the instrument drive called “DoNotEject” to initialize. A window opens when the drive is initialized. Close the window.



NOTE

Never eject the “DoNotEject” flash drive located inside the instrument chassis, or modify the files on it. This drive contains hardware configuration files and initializes whenever the instrument is turned on.

- 5 To ensure adequate disk space, archive the data on the instrument computer from previous runs to a network location.
- 6 Open the HiSeq Control Software (HCS) using the shortcut icon on the computer desktop. HCS takes a few minutes to initialize. When the software is initialized, the Welcome screen opens and the initialization icon  appears on the bottom left of the screen.

[This page intentionally left blank]

HiSeq Software

Introduction	30
HiSeq Control Software Interface	32
Run Configuration Screens	38
Pre-Run Setup Screens	50
Initiate Run Screens	54
Run Overview Screen	56
Stopping and Resuming a Run	59
Pausing a Run	63
Run Folders	64
Sample Sheet Overview	66
Available Disk Space	68
Switching Sequencing Modes	69
Switching to High Output Mode	70
Switching to Rapid Run Mode	73



Introduction

Three software applications are installed on the instrument computer:

- ▶ **HiSeq Control Software**—The HiSeq Control Software (HCS) interface guides you through the steps to specify run parameters, and load the flow cell and reagents prior to starting the run. During the run, HCS provides a summary of quality statistics throughout the run.
- ▶ **Real Time Analysis**—The integrated primary analysis software, Real Time Analysis (RTA), performs base calling and assigns a quality score to each base for each cycle. For more information, see *Real Time Analysis Overview* on page 176.
- ▶ **Sequencing Analysis Viewer**—The Sequencing Analysis Viewer (SAV) provides detailed quality statistics and updated metrics as the run progresses. For more information, see *Sequencing Analysis Viewer* on page 178.

What's New in HCS v2.0

HCS v2.0 introduces the following features:

- ▶ Support for dual mode High Output and Rapid Run on instruments that support both modes.
- ▶ New gasket change workflow.
- ▶ Improvement and enforcement of wash workflows:
 - Wash timers that require a wash of the system prior to a sequencing run.
 - Water wash every 24 hours from the completion of the last water wash.
 - Maintenance wash every 240 hours from the completion of the last maintenance wash.
 - Maintenance wash is considered as a water wash in terms of the timers and precedence.
 - The software will prompt the user to run post run washes when runs are completed.
 - Rapid Run mode will cycle through a water wash unless the system is in need of a maintenance wash.
 - High Output mode workflow includes a maintenance wash.
 - When you halt a run, HCS will prompt for a wash if it appears that the user does not intend to resume the run (Complete for Run).
 - Halting a run allows the resumption of the run without a wash of the system.

- Starting a new run requires a wash.
- New BCL zipping and Q scores binning to reduce the size of the data footprint.
- ▶ Optional pre run fluidics check.
- ▶ Support for BaseSpace: With BaseSpace, you will need a sample sheet.
- ▶ On instrument cluster generation for Rapid Run mode.
- ▶ Optional bcl compression and qScore binning as required for BaseSpace.

To download a copy of HCS v2.0, go to the HiSeq 2500 support page on <http://www.illumina.com>, and click the Downloads tab. A MyIllumina login is required.

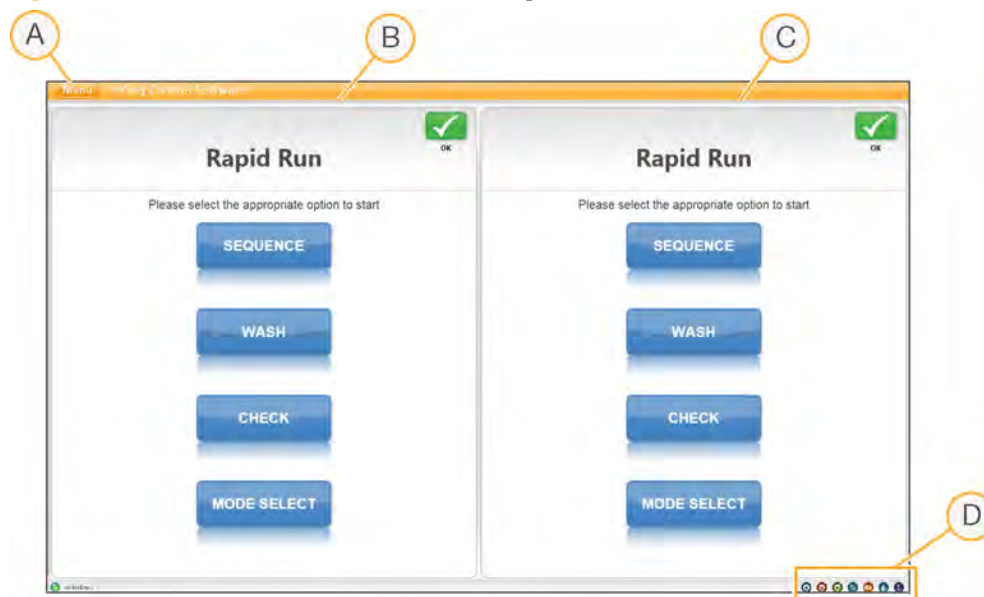
HiSeq Control Software Interface

The HiSeq Control Software (HCS) interface opens to run configuration screen. From here you can select the RAPID RUN or HIGH OUTPUT run configuration. When you change the mode, you will have to perform a gasket change and maintenance wash. HiSeq Control Software (HCS) will guide you through the process. Please be sure to allow sufficient time for this.

Once you have selected a mode and completed any required wash steps, the Welcome screen opens for the mode you have selected. Both runs must be in the same mode. The software returns to the Welcome screen of the given mode when a run is complete.

The Welcome screen is split into two panels, one for flow cell A and one for flow cell B. You can set up a run for flow cell A and flow cell B in parallel using the software interface.

Figure 8 HCS Interface, Welcome Screen for Rapid Run Mode



- A Welcome Screen Menu Button
- B Interface Panel for Flow Cell A
- C Interface Panel for Flow Cell B

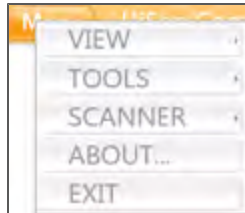
D Activity Icons

Welcome Screen Menu

The Welcome screen menu button, located in the upper-left corner of the Welcome screen, provides the following options:

- ▶ **View**—Provides options to view the interface in full screen or in a window, or to minimize the interface.
- ▶ **Tools**—Provides access to the Options window and Show Log file:
 - **Options**—From the Options window, you can define the run Id template, default folder locations, and the LIMS server, username, and password. See *Menu Options Window* on page 33.
 - **Show Log File**—The Show Log file is a text file used for troubleshooting.
- ▶ **Scanner**—Activates the command to manually initialize the software.
- ▶ **About**—Provides information about instrument hardware, software versions, and technical support contact information.
- ▶ **Exit**—Closes the HCS interface.

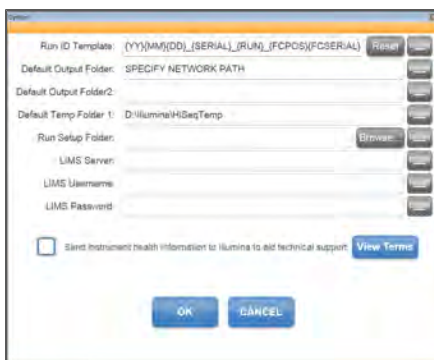
Figure 9 Menu Button



Menu Options Window

From the Options window, you can define the run Id template, default folder locations, and the LIMS server, username, and password.

Figure 10 Menu Options Window



Run ID Template — The name used to name a run folder

Default Output Folder — The default location for flow cell A runs to be saved. This can be changed on a per run basis

Default Output Folder2 — The default location for flow cell B runs to be saved. This can be changed on a per run basis

Default Temp Folder1 — The location to which temporary files can be written during a run

Run Setup Folder — Where the HiSeq will pick up LIMS sample forms

LIMS Server — The server name for interactions with supported Illumina LIMS

LIMS Username — The LIMS username HiSeq will use when authenticating to Illumina LIMS

LIMS Password — The LIMS password HiSeq will use when authenticating to Illumina LIMS

Send instrument health information to Illumina to aid technical support — Gives permission to the instrument to send information to BaseSpace for each run. Any and all information will remain confidential. Illumina recommends that you keep this feature enabled and will never share your information. Run sample data will be sent only if you specifically indicate on the Storage screen that you want to send run sample data. See *Storage* on page 40 for more information.

Start Screen Options

Four options on the Welcome screen are Sequence, Wash, Check, and Mode.

- ▶ **Sequence**—Select **Sequence** to begin the steps to setup a sequencing run. From the Sequence button, you can start a new run or resume an existing run.

Figure 11 Sequence Command Options



- **New Run**—The software guides you through setting run parameters, loading and priming reagents, loading the flow cell, performing fluidics checks, and starting the run.
 - **Resume Run**—The software guides you through selecting the existing run folder and setting parameters for resuming the run.
- ▶ **Wash**—Select **Wash** to initiate an instrument wash. From the Wash button, you can start a Water Wash or a Maintenance Wash.

Figure 12 Wash Command Options



- **Water Wash**—The water wash flushes water through the system. A water wash is recommended before you start a new run if the instrument has been idle for one day or more. For more information, see *Performing a Water Wash* on page 169.
 - **Maintenance Wash**—The maintenance wash flushes water, then NaOH, and then water through the system. This wash is recommended after each sequencing run in High Output mode and every 10 days in Rapid Run Mode. For more information, see *Performing an Instrument Maintenance Wash* on page 164.



NOTE

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

- ▶ **Check**—Select **Check** to open the fluidics check screen and confirm proper flow during instrument installation or fluidics troubleshooting. For more information, see *Performing a Fluidics Check* on page 188.

- ▶ **Mode Select**—Select **Mode Select** to return to the run configuration screen.

Activity Indicators






The Welcome screen contains a series of icons in the lower-right corner of the screen that indicate which activity the instrument is performing. From left to right, they represent the X, Y, and Z motors, electronics functionality, the camera, the fluidics system, and processing functions.

Figure 13 Activity Indicators



Status Icons

Status icons are located in the upper-right corner of each screen, and notify you of any change in conditions during run setup or during the run.

Status Icon	Status Name	Description
	Status OK	No change. System is normal.
	Information	Information only. No action is required
	Attention	Information that might require attention.
	Warning	Warnings do not stop a run, but might require action before proceeding.
	Error	Errors usually stop a run and generally require action before proceeding with the run.

When a change in condition occurs, the icon changes and blinks to alert you. If this happens, select the icon to open the status window, which contains a general description of the condition. Select **Acknowledge** to accept the message and **Close** to close the dialog box.

You can filter the types of messages that appear in the status window by selecting the icons along the top margin of the window. Selecting an icon toggles the condition to show or to hide.



NOTE

For more information about known issues related to errors and warnings, see the release notes for the version of instrument control software you are using.

Run Configuration Screens

When you select **Sequence | New Run** from the Welcome screen, the software opens the run setup screens that guide you through the run setup steps, starting with the screens on the Run Configuration tab.

Each run configuration screen contains drop-down lists, checkboxes, or text fields. Use the hand-held barcode scanner to scan the flow cell or reagent kit ID, or enter the ID using the touch screen keyboard. The keyboard icon is located to the right of text fields.



When you have finished entering the required information on each screen, the Next button becomes active. Select **Next** to move to the next screen or select **Back** to return to the previous screen. At any time during the run setup steps, you can select **Cancel** to abort the run setup and return to the Welcome screen.

Run configuration screens are different for Rapid Run mode than for High Output mode.

- ▶ High Output includes Storage, Flowcell-Set-up, Advanced, Recipe, Reagents, and Review.
- ▶ Rapid Run includes Pre Run Fluidics Check, Storage, Flowcell-Set-up, Advanced, Recipe, Sample Sheet, Reagents, and Review.

Gasket Change

When you change the run mode from the Welcome screen (Rapid Run to High Output or vice versa), the software opens the Gasket Change screens that guide you through a maintenance wash.

- 1 **Load Gasket** — You are prompted to load a new gasket.
- 2 **Load Wash Reagents** — You are prompted to load wash reagents.
- 3 **Load Wash Flow Cell** — You are prompted to load a non-sequencing flow cell and enter a wash flow cell ID.
- 4 **Fluidics Check** — You are prompted to select a solution. Use default values for Volume, Aspiration Rate, and Dispense Rate.
- 5 **Wash** — The instrument performs a wash.
- 6 **Load Solvent** — You are prompted to load NaOH in all positions for washing.

- 7 **Solvent Wash** — The instrument performs a solvent wash.
- 8 **Load Wash Reagents** — You are prompted to load wash reagents.
- 9 **Wash** — The instrument performs a wash.

Maintenance Wash

If required, HCS will prompt you to run a maintenance wash before a run. In all cases, a maintenance wash is recommended after every High Output run and every 10 days in Rapid Run mode.

- 1 **Load Wash Reagents** — You are prompted to load wash reagents.
- 2 **Load Wash Flow Cell** — You are prompted to load a non-sequencing flow cell and enter a wash flow cell ID.
- 3 **Fluidics Check** — You are prompted to select a solution. Use default values for Volume, Aspiration Rate, and Dispense Rate.
- 4 **Wash** — The instrument performs a wash.
- 5 **Load Solvent** — You are prompted to load NaOH in all positions for washing.
- 6 **Solvent Wash** — The instrument performs a solvent wash.
- 7 **Load Wash Reagents** — You are prompted to load wash reagents.
- 8 **Wash** — The instrument performs a wash.

Volume Check

The optional pre run fluidics check is a feature of only Rapid Run mode. It enables you to ensure that fluidics lines are clear and pumps are performing to calibrated specification before you start a run. The fluidics check has four steps:

- ▶ **Load Reagents** — You are prompted to load reagents.
- ▶ **Load Wash Flow Cell** — You are prompted to load a non-sequencing flow cell and enter a wash flow cell ID.
- ▶ **Fluidics Check** — You are prompted to select a solution. Use default values for Volume, Aspiration Rate, and Dispense Rate.
- ▶ **Wash** — The instrument performs a fluidics check.

Storage

The Storage screen enables you to define where you run data will be output and stored.

Figure 14 Storage Screen in High Output Mode

Please enter "Password".

Connect to BaseSpace Storage and analysis Run Monitoring Only

Enter your MyIllumina account

Email:

Password:

If you do not have an account, please go to <https://my.illumina.com/sign-up>.
A sample sheet is required for BaseSpace runs.

Save to an output folder:

Zip BCL files Bin QCcores (≈36% reduction in storage space)

Save Auxiliary Files:

- ▶ **Connect to BaseSpace**—When you select this option you will be prompted to enter your MyIllumina account information. Zip BCL files (below) will be selected by default. Illumina recommends that you also save files locally. To save files locally, select **Save to an output folder** and enter a path.
- ▶ **Storage and analysis**—This option enables the HiSeq to send run data as well as system health information to BaseSpace.
- ▶ **Run Monitoring Only**—This option sends only system health information to BaseSpace.



NOTE

For every run, your HiSeq 2500 will attempt to send at least general run metrics to BaseSpace. Metrics are sent anonymously, associated with the HiSeq 2500 from which they came, not to any particular user. Metrics include interoperation files (that SAV uses) and global files such as RunInfo.xml and RunParameters.xml. No data files are sent unless you specify for the system to do so.

Permissions for this feature are enabled by default, and Illumina recommends that you leave them enabled. To disable permissions, select **Menu | Tools | Options**, and then unselect **Send system health information to Illumina to aid technical support**. You can also view terms in the Options dialog box.

If permissions for this feature are not enabled, HCS will prompt you to accept Terms and Conditions. If you do not accept the terms and conditions,

you will not be able to use BaseSpace. If you are not using BaseSpace, go back to the previous screen and ensure that the BaseSpace option is de-select (otherwise the Next button will not be enabled).

- ▶ **Save to an output folder**—Enter the path to the run folder, which is the destination for all output files. Illumina recommends that you specify a network location for the run folder. For more information, see *Run Folders* on page 64.
- ▶ **Zip BCL files**—Zipping BCL files reduces the footprint of disc space needed for the run. If you are using BaseSpace, files are zipped by default.
- ▶ **Bin QScores**—Binning Q Scores reduces the footprint of disc space needed for the run. This is done by grouping Q Scores together over a wider range of values. For technical details regarding how data quality score resolution can be reduced without sacrificing score accuracy or standard analysis and variant calling performance, please find the white paper on www.Illumina.com titled, *Reducing Whole-Genome Data Storage Footprint*.
- ▶ **Save Auxiliary Files**—This option enables you to define which image files will be saved.
 - **Save All Thumbnails**—Thumbnail images are a combination of eight images from each swath, combined in one thumbnail image. A swath is a column of tiles, and there are three swaths per lane per surface. Save All Thumbnails saves swath thumbnails and tile thumbnails.
 - **Save Tile Thumbnails**—This option saves only tile thumbnails. Tile thumbnails show just the tile rather than the whole swath.

Flow Cell Set-up

The Flow Cell Set-up screen is the run configuration screen with which you record information about your run, starting with the flow cell ID.

Figure 15 Flow Cell Set-up Screen for a High Output Flow Cell



NOTE

It is critical that you scan the flow cell barcode or accurately enter the flow cell ID when you set up your run. The software uses the flow cell ID to determine flow cell type and component compatibility.

- ▶ **Flow Cell ID**—Enter the flow cell ID of the new flow cell. You can enter the flow cell ID using the barcode scanner, the keyboard, or the touch screen keyboard. Use the barcode scanner to ensure accuracy; otherwise, carefully enter the flow cell ID.
- ▶ **Flow Cell Type**—The appropriate flow cell is automatically selected based on the flow cell ID. The software uses this information to confirm flow cell compatibility with instrument software and run components. For more information, see *Recipe Screen* on page 43.
- ▶ **Experiment**—Enter an experiment name of your choosing. This name appears at the top of each interface screen to help identify the sequencing run in progress.
- ▶ **User Name**—Enter the name of the person setting up the run.

Advanced

The Advanced screen includes settings for first base confirmation and indicating alignment lanes.

Figure 16 Advanced Screen for a High Output Flow Cell

- ▶ **Confirm First Base**—Select the checkbox to receive a first base confirmation dialog box that provides the option to abort the run at this point or to continue. Selecting this option requires that you confirm the results before the run can continue. If you do not select the checkbox, the first base report is generated as usual and the run continues without showing the first base report confirmation dialog box. The first base report may be accessed at any time during the run from the run folder. For more information, see *First Base Report Confirmation* on page 57.
- ▶ **Control Lane**—If applicable, select the lane on your flow cell containing the control library from the drop-down list. Otherwise, select **None**.
- ▶ **Align to PhiX**—This allows for real-time reporting of error rates for PhiX control lanes or PhiX spike-in lanes. By default, all lanes are selected for alignment by Real Time Analysis (RTA). Clear the checkbox for any lanes that do not have PhiX added to the sample, so that RTA does not align and calculate error rates for that lane, resulting in arbitrarily high error rates. For more information, see *Using a PhiX Control for HiSeq Sequencing Runs*, which is available on the HiSeq support page on the Illumina website at <http://www.illumina.com>.
- ▶ **Keep Intensity Files**—Select the checkbox to save cluster intensity files (*.cif) for later reanalysis, troubleshooting, or custom processing. Saving intensity files is not required for on-instrument analysis. Enabling this option significantly increases the size of the data output folder. By default, this option is not selected.

Recipe Screen

From the information that you enter on the Recipe screen, the software creates a recipe specific to the run. The recipe is automatically copied to the run folder and saved in the

Recipe subfolder. The saved recipe includes only the number of cycles for each read and index; it does not save the options for the control lane, PhiX alignment, images, or cifs.

Figure 17 Recipe Screen in for a High Output Flow Cell with Dual Index Selected

Please check the remaining parameters and select "Next" to continue.

Index Type: No Index Dual Index Single Index Custom

Flowcell Format: Single Read Paired End

Read 1: 8 Index 1 (I7): 8 Index 2 (I5): 1 Read 2: 8

Cycles: [8] [8] [1] [8]

SBS: TruSeq SBS Kit v3

Index: TruSeq Dual Index Sequencing Primer Box

PE Turnaround: []

Existing Recipe

Index Type

- ▶ Select **No Index** to perform a non-indexed run.
- ▶ Select **Single Index** to perform a run with one Index Read immediately following Read 1. The read length of the Index Read depends on the kit you select from the **Index** drop-down list. Single-indexed libraries require seven cycles and dual-indexed libraries require eight cycles.
- ▶ Select **Dual Index** to sequence dual-indexed libraries with two eight-cycle Index Reads immediately following Read 1. The second Index Read includes seven chemistry-only cycles prior to eight cycles of imaging.
- ▶ Select **Custom** to perform a single-indexed or dual-indexed run with a custom number of cycles for each Index Read.

Flow Cell Format

In High Output mode, this option appears when you select Dual Index or Custom as an index type. Select a **Flow Cell Format: Single Read** or **Paired End**.



NOTE

A dual-indexed paired-end run must be performed on a paired-end flow cell. A dual-indexed single-read run can be performed on either a single-read flow cell or a paired-end flow cell with recipe adjustments. Your selection in HCS, paired-end or single-read flow cell, will determine the recipe. Always

use the appropriate dual-index sequencing primers box according to your flow cell type.

Cycles

- ▶ **Cycles**—Enter the number of cycles you plan to sequence for each read including Read 1, the Index Reads, and Read 2.



NOTE

Always add one cycle to the total cycles of Read 1 or Read 2. For example, to perform 100 cycles for Read 1, enter 101. To perform 50 cycles for Read 1, enter 51.

The number of cycles performed in a read is one more cycle than the number of cycles analyzed. The last cycle is required for phasing and prephasing calculations.

- For paired-end runs, enter the **same** number of cycles for Read 2 as you entered for Read 1; Or, if you selected the **Custom** indexing option, enter the number of cycles for each Index Read (Read lengths do not need to be identical).



NOTE

If you selected **Single Index** or **Dual Index**, the number of indexing cycles are entered automatically.

SBS

With HCS v2.0 and later, compatible kit versions for SBS, Index, and PE turnaround chemistry are entered by default based on the flow cell ID you entered on the Scan screen.

The following table lists the default kit name as listed on the Recipe screen for each flow cell type:

Flow Cell Type	SBS	Index	Clustering and/or PE Turnaround
HiSeq Flow Cell v3	TruSeq SBS Kit v3	TruSeq Multiplex Sequencing Primer Box (<i>single-indexed libraries</i>) or TruSeq Dual Index Sequencing Primer Box (<i>dual-indexed libraries</i>)	TruSeq PE Cluster Kit v3 - cBot - HS

Flow Cell Type	SBS	Index	Clustering and/or PE Turnaround
Rapid Flow Cell	TruSeq Rapid SBS Kit	--	TruSeq Rapid PE Cluster Kit or TruSeq Rapid SR Cluster Kit

**NOTE**

The TruSeq Multiplex Sequencing Primer Box is part of the TruSeq Cluster Kit v3. Reagents provided in these kits can be used with any High Output flow cell type. Ensure that you use High Output reagents with High Output runs and Rapid Run reagents with Rapid Runs.

Index

This is the Illumina kit name that is being used for the indexing reagents.

PE Turnaround/Cluster in Rapid Run Mode

This is the Illumina kit name that is being used for the PE reagents.

Existing Recipe

If you choose to use a custom recipe for your run, enter the path and recipe name. Otherwise, leave this checkbox blank. Illumina recommends that you use the recipe wizard to build your recipe based on the information you enter on the Recipe screen. For more information, see *Recipe Screen* on page 43.

Sample Sheet Screen**NOTE**

To view indexing results in Sequencing Analysis Viewer (SAV), it is essential that you load a sample sheet at the start of your run. SAV v1.8 requires a sample sheet to properly report indexing information. For more information, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

High Output Mode

In the **Sample Sheet** field, enter the path of a valid sample sheet for your run. When the run begins, the software copies the sample sheet to the root level of the run folder.

Unless you are using BaseSpace, a sample sheet is optional. For indexed runs, the sample sheet is required if you want indices to be reported in SAV. If you are running BaseSpace, a sample sheet is required. For more information about sample sheets, see *Sample Sheet Overview* on page 66.

Figure 18 Sample Sheet Screen for a High Output Flow Cell



Please select a sample sheet if you want indexes to be reported.

Sample Sheet:

If a sample sheet is associated with your run, index-specific information can be viewed in Sequencing Analysis Viewer (SAV) v1.8.4 (or later). For more information, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

Rapid Run Mode

In the **Sample Sheet** field, enter the path of a valid sample sheet for your run. When the run begins, the software copies the sample sheet to the root level of the run folder. Unless you are using BaseSpace, a sample sheet is optional. For indexed runs, the sample sheet is required if you want indices to be reported in SAV. If you are running BaseSpace, a sample sheet is required. For more information about sample sheets, see *Sample Sheet Overview* on page 66.

Rapid Run mode does not support sample sheets and offers preparation options that are not available in High Output mode:

- ▶ Select **On-Board Cluster Generation** when you want to perform all cluster generation on the HiSeq. If you select On-Board cluster generation, each Rapid Run flow cell that is to be sequenced can contain only one sample in both lanes.
- ▶ Select **Template Hybridization on cBot** if you have performed template hybridization on the cBot. If you perform template hybridization on the cBot, each lane of a Rapid Run flow cell can have a unique sample.

Reagents

The Reagents screen records information about the reagents used for your run.

Figure 19 Reagents Screen in High Output Mode

- ▶ **Reagent Kit ID**—Enter the SBS reagent kit ID from either box 1 or box 2 of the SBS kit using the barcode scanner, the keyboard, or the touch screen keyboard.
- ▶ **Indexing Reagent Kit ID**—Enter the indexing reagent kit ID using the barcode scanner, the keyboard, or the touch screen keyboard. The Indexing Reagent Kit ID is the barcode of the reagent kit that is being used for an indexed run.
- ▶ **SBS Kit**—Select the SBS Kit you are using for your run, either 200 cycles, 50 cycles, or a the number of cycles that you want to define. Your selection automatically populates the SBS Cycles remaining according the number of cycles supported by that kit.
- ▶ **Cycles Remaining**—This field is populated automatically depending on the SBS kit you are using. If you are using a partial kit, enter the number of cycles reagents are expected to last. The software counts down the number of cycles entered. When the cycles are low, the software prompts you to load fresh reagents. For more information, see *Changing Reagents During a Run* on page 63.



NOTE

If you are starting a new run using less than a full reagent kit, make sure that you can be present when new reagents are needed.

- ▶ **Prime SBS Reagents**—Select this checkbox to prime reagents before you start a run. Always prime reagents before loading a new flow cell. If you are performing on-board cluster generation, this is not an option.

Review Screen

The Review screen lists the information you have entered for your run. Select **Back** if you need to change any entries.

Pre-Run Setup Screens

After you have entered the required information in the run configuration screens, the pre-run setup screens open. Pre-run setup screens include: Load Reagents, Load Prime Flow Cell, Prime Fluidics Check, and Prime. The Load Prime Flow Cell, Prime Fluidics Check, and Prime screens are available in only High Output mode and Rapid Run mode when you have performed template hybridization on the cBot.

Load Reagents Screen

The Load Reagents screen prompts you to load SBS reagents. If you specified a number of cycles for indexing, the software prompts you to load indexing reagents on the paired-end rack.

After Read 1 of a paired-end run in High Output mode, the software pauses so that you can load paired-end reagents. Paired-end reagents must be loaded prior to Read 2 resynthesis.

Figure 20 Load Reagents Screen for a High Output Flow Cell



NOTE

To prepare for the post-run rinse at the end of a sequencing run, it is important that you load 25 ml of PW1 or laboratory-grade water in position 2 when you load SBS reagents at the beginning of your run.

The post-run rinse does *not* replace the post-run instrument wash.

Load Prime Flow Cell Screen

This screen is available in High Output mode and Rapid Run mode when you have performed template hybridization on the cBot.

The Load Prime Flow Cell screen prompts you to load a *used* flow cell for the priming step. After loading a used flow cell, confirm that the vacuum is engaged.



NOTE

Illumina recommends that you use the flow cell from a previous run for priming reagents on a subsequent run or for a post-run maintenance wash.

Figure 21 Load Prime Flow Cell Screen for High Output Flow Cells



Prime Fluidics Check Screen

This screen is available in High Output mode and Rapid Run mode when you have performed template hybridization on the cBot.

The Prime Fluidics Check screen prompts you to confirm proper flow before priming reagents. Select **Pump** to start the fluidics check.

You can pause or cancel the fluidics check from this screen:

- ▶ Select **Pause** to temporarily pause the fluidics check. The Pause button changes to the Resume button.
- ▶ Select **Resume** to continue.

- ▶ Select **Stop** to cancel the fluidics check.

Figure 22 Prime Fluidics Check Screen



Prime

This screen is available in High Output mode and Rapid Run mode when you have performed template hybridization on the cBot.

From the Prime screen you can monitor the priming step in progress. Select **Start Prime** to begin, or select **Skip Prime** if you have already primed reagents.

Figure 23 Prime Screen



You can pause or cancel the priming step from this screen:

- ▶ Select **Pause** to temporarily pause the priming step. The Pause button changes to the Resume button.
- ▶ Select **Resume** to continue.
- ▶ Select **Stop** to stop priming.
- ▶ Select **Cancel** to abort the run setup and return to the Welcome screen.

When the priming step is complete, the initiate run screens open automatically.

Initiate Run Screens

After you have loaded reagents and primed reagents, the software opens the initiate run screens. Initiate run screens include Load Sequencing Flow Cell, Fluidics Check, and Close Door.

Load Sequencing Flow Cell Screen

The Load Sequencing Flow Cell screen illustrates the steps to load the clustered flow cell and engage the vacuum. When the vacuum is successfully engaged, the system automatically selects the Vacuum Engaged checkbox.

In Rapid Run mode, enter a Flow Cell ID.

Figure 24 Load Seq Flow Cell Screen for a High Output Flow Cell

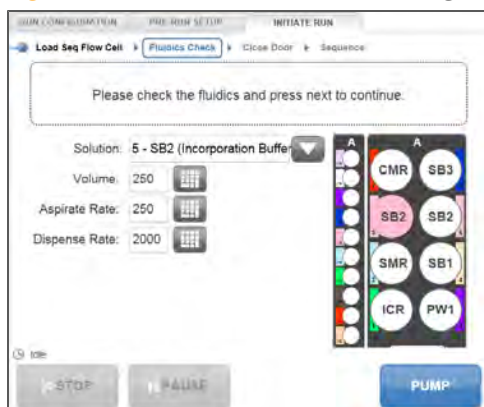


Fluidics Check Screen

The Fluidics Check screen prompts you to confirm proper flow prior to starting the sequencing run.

SB2 *must* be selected for a fluidics check on a clustered flow cell. Use the default volume, aspirate rate, and dispense rate. Alternate between position 5 and position 6 if you need to repeat the fluidics check.

Figure 25 Fluidics Check Screen for a High Output Flow Cell



Close Door Screen

The Close Door screen prompts you to close the flow cell compartment door. The system confirms that the vacuum is engaged and that the flow cell compartment door is closed. The checkboxes are selected automatically when the system is ready.

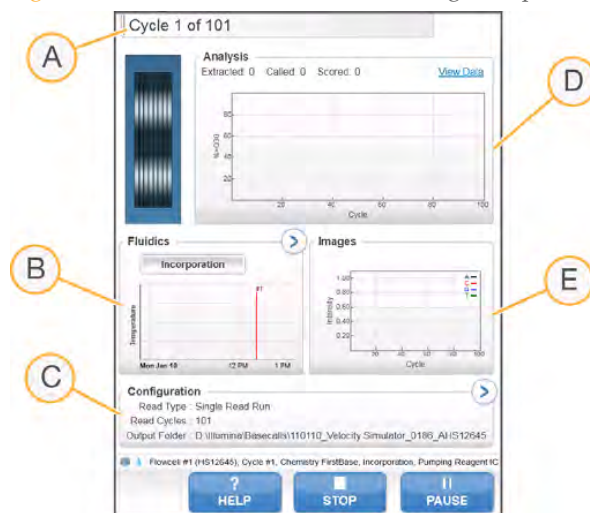
Figure 26 Close Door Screen



Run Overview Screen

The Run Overview screen provides an overview of run progress, fluidics status, imaging status, and intensities by cycle.

Figure 27 Run Overview Screen for a High Output Flow Cell

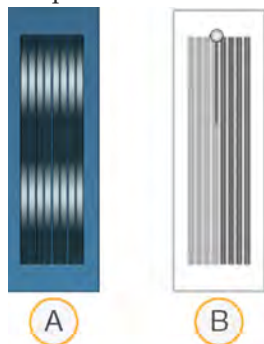


- A Run Progress
- B Fluidics and Temperature
- C Run Configuration
- D Analysis
- E Images

This screen contains the following sections to monitor your run:

- ▶ **Run Progress**—Provides a visual display of run progress and lists the number of cycles completed. Run progress is indicated by a visual progress bar and an image of a flow cell. The progress bar at the top of the screen shows the number of cycles completed and the flow cell image shows which lanes have been imaged for that cycle.
 - Light gray indicates that the lane has been imaged; dark gray indicates that the lane has not been imaged.

- Wavy lines across the flow cell indicate that reagents are pumping through the flow cell.
- Changing background colors indicate temperature changes during the run. Blue indicates cooler temperatures, while orange and red colors indicate warmer temperatures.



- A Flow Cell Image (of High Output flow cell): Fluidics in Progress
 B Flow Cell Image (of High Output flow cell): Imaging in Progress

- ▶ **Configuration**—Lists the read type, output folder, and run parameters for this run. Select the arrow button to expand the view.
- ▶ **Analysis**—Shows a graph of intensities by cycle.
- ▶ **Fluidics**—Shows the current temperature, protocol step, and reagent. Select the arrow button to expand the view.
- ▶ **Images**—Shows a graph of intensities per cycle while the flow cell is in chemistry and thumbnails while the flow cell is imaging.

First Base Report Confirmation

The first base report is an *.htm file that provides early information about cluster density, focus quality, and intensities of the flow cell.

If you selected **First Base Confirmation** during the run setup, the first base report confirmation dialog box opens automatically after the completion of cycle 1. For more information, see *Advanced* on page 42.

Sequencing Analysis Viewer

The Sequencing Analysis Viewer (SAV) is an application that allows you to view important sequencing metrics generated by Real Time Analysis (RTA) in plots, graphs, and tables. SAV automatically opens to the analysis page after image analysis begins.

Statistics are updated as soon as new analysis information is available. Select **Refresh** at any time during the run to view updated metrics as the run progresses.

SAV can be viewed from a remote location by installing SAV on another computer with access to the network location of the output run folder. You can download a copy of SAV from the Illumina website. A MyIllumina login is required.

For a detailed description of available metrics and installation instructions, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

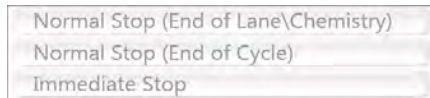
Stopping and Resuming a Run

You can stop a run from the Run Overview screen. You might stop a run if the run was set up incorrectly, if the data quality is bad, or if there is a hardware error. If you plan to resume your run, make sure that you select the appropriate normal stop options in conjunction with one of the RTA options that allow the run resume.

Stop Option	RTA Option	Able to Resume?
Normal Stop (End of Lane\Chemistry)	Keep As Is	Yes. The run resumes at the next chemistry or imaging command.
	Complete For Run	No. The run cannot be resumed.
	Complete For Read	Yes. The run resumes at the beginning of the next read.
Normal Stop (End of Cycle)	Keep As Is	Yes. The run resumes at the next cycle.
	Complete For Run	No. The run cannot be resumed.
	Complete For Read	Yes. The run resumes at the beginning of the next read.
Immediate Stop	No option	No.

- 1 From the Run Overview screen, select **Stop**. The stop menu opens.

Figure 28 Stop the Run



- 2 Select one of the following stop options:
 - **Normal Stop (End of Lane\Chemistry)**—Stops the run only after the current chemistry or imaging command is complete, and then places the flow cell in a safe state.
 - **Normal Stop (End of Cycle)**—Stops the run after the current cycle is complete, and then places the flow cell in a safe state.
 - **Immediate Stop**—Stops the run without completing the current operation and does *not* place the flow cell in safe state. You cannot resume a run that was stopped using the immediate stop option.
- 3 Select one of the following RTA options:

- **Keep As Is**—The run is stopped without any modifications to RTA. You can resume the run where you left off.
 - **Complete For Run**—RTA is stopped. The run info, run parameters, and recipe files are updated to reflect total cycles as the last cycle completed. Then RTA restarts to complete analysis for the run up to the point the run was stopped. You will *not* be able to resume the run.
 - **Complete For Read**—RTA is stopped. The run info, run parameters, and recipe files are updated to trim the length of the current read to the last cycle completed. Subsequent reads are not affected. Then RTA restarts to complete analysis for the current read. You can resume the run at the beginning of the next read.
- 4 After the run is stopped, select **Return to Start** on the Run Overview screen. The Welcome screen opens.

Resuming a Stopped Run

You can resume a stopped run if the run was stopped safely using a normal stop option in conjunction with one of the RTA options that allow the run to resume.

- 1 From the Welcome screen, select **Sequence**, and then select **Resume Run**. The Resume screen opens.
- 2 Select the run folder for the run you want to resume from the drop-down list.



NOTE

The software resumes the run at the point where the run was stopped, and defaults to the correct setup on the Resume screen.

- 3 Confirm the settings on the Resume screen or select the appropriate point in the run to resume the run.
 - **Resume At** lists the read or point in the run to resume.
 - **Start At Cycle** lists the cycle to resume.

Figure 29 Points to Resume the Run

**CAUTION**

Illumina does not recommend resuming a run at the point of paired-end turnaround, other than for Read 2 primer rehybridization. For more information, see *Repeating Read 2 Primer Hybridization* on page 189.

- 4 Confirm the settings on the Resume screen or select the imaging and chemistry commands to resume. For more information, see *Example Settings for Resuming a Run* on page 61.
- 5 Select **Next** to proceed. The software guides you through the remaining run setup steps.

Example Settings for Resuming a Run

If you stopped a run after imaging lane 1 at cycle 23 of a 101-cycle read, the software automatically sets up the resume run settings for Read #1 at cycle 23. You should see the following settings on the Resume screen:

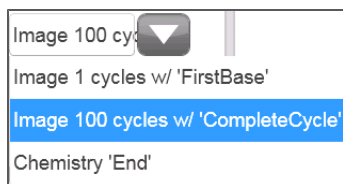
- ▶ **Resume At:** Read #1
- ▶ **Start At Cycle:** 23

Figure 30 Example of Resume at Cycle 23



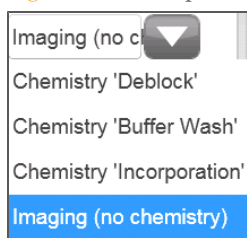
For a 101-cycle read, there is one cycle of **FirstBase** followed by 100 cycles of **CompleteCycle**. At cycle 23, you are within 100 cycles of **CompleteCycle**. **Image 100 cycles w/ 'CompleteCycle'** should automatically be selected.

Figure 31 Examples of Imaging Commands



Because the run in this example was stopped during an imaging step, **Imaging (no chemistry)** should automatically be selected.

Figure 32 Examples of Chemistry Commands



Pausing a Run

You can pause a run from the Run Overview screen. You might pause a run to check run components, such as reagent volumes, before proceeding with the run. Under normal procedures, you don't need to use this feature. The software pauses the run automatically before Read 2 in High Output mode to allow you to load reagents.

- 1 From the Run Overview screen, select **Pause**. The pause menu opens.

Figure 33 Pause Options



- 2 Select **Normal Pause**.
- 3 Select **Yes** to confirm the pause command. The software completes the current chemistry or imaging command and places the flow cell in a safe state.
- 4 Select **Resume** when you are ready to resume the run.

Changing Reagents During a Run

If you started your run with a partial volume of reagents, use the Change Reagents feature to pause your run and replenish reagents. There is no need to pause a run to load fresh reagent before Read 2; the software pauses the run automatically at that point.

- 1 From the Run Overview screen, select **Pause**. The pause menu opens.
- 2 Select **Change Reagents**.
- 3 Select **Yes** to confirm the pause command. The software completes the current chemistry or imaging command and places the flow cell in a safe state.
- 4 On the Reagent screen enter the following reagent parameters:
 - Reagent kit ID for new reagents you are loading.
 - The number of cycles the reagents are expected to last.



NOTE

The priming checkbox is disabled on the Change Reagents screen. Priming is not required.

- 5 Select **Next** to proceed to loading reagents.

Run Folders

Each sequencing run generates a run folder that contains data files and log files.

Run Folder Name and Path

During run setup, the software prompts you to enter the path for the run folder. By default, the folder is named in the following format:

YYMMDD_<Workstation Name>_<Run Number>_<Flow Cell ID>

Example: 110114_SN106_0716_A90095ACXX

The run number increments by one each time you perform a run on a given instrument. The flow cell ID entered during the run setup steps appends to the end of the run folder name.

The run folder is written to the output path specified on the Scan screen during run setup. The temporary run folder for flow cell A is written to the D: drive and the temporary run folder for flow cell B is written to the E: drive.

Contents of Run Folders

Subfolder	Key Files/Folders	Description
[Root level]	First_Base_Report.htm	File with basic metrics for the first cycle.
	IALog.txt	Log file for image analysis events.
	Sample Sheet (optional)	If you plan to use CASAVA, copy the sample sheet to the root level of the run folder. For more information, see the <i>CASAVA User Guide</i> , Part # 15011196.
	RunInfo.xml	Identifies high-level run information.
	runParameters.xml	Provides the setup parameters for the run.
Config		Contains a copy of the software configuration parameters used for the run.

Subfolder	Key Files/Folders	Description
Data	Status.htm	Contains run and analysis-specific metrics generated during the run.
	Intensities folder	Contains image analysis results. This folder is saved on the network folder under <Run Folder>\Data, and is given the same name as the run folder on the instrument computer.
	CopyLog.txt	Log file for copy events.
	Log.txt	Log file for image analysis and base calling.
Data\ Intensities	BaseCalls folder	Contains RTA base calling results.
	config.xml	Contains run and image analysis-specific information.
	RTAConfiguration.xml	Configuration file for RTA in *.xml format.
Logs	Recipe_<date>_<time>.log	Contains the run log information. Whenever the log file gets too big (approximately 5 MB), HCS starts a new log file.
Recipe	*.xml	When you start a run, a copy of the recipe is moved to the Recipe folder. If you stop the run and resume it later, this recipe restarts where the run was stopped.
Thumbnail_Images	thumbnail images (*.jpg files)	Contains thumbnails of nine locations from each tile, which can be used to troubleshoot a run.

Sample Sheet Overview

The sample sheet is a user-generated file in *.csv format that stores information about a sequencing run. The sample sheet is copied into the run folder and later used for analysis.

Using a sample sheet is optional. If you choose to use a sample sheet, create one *before* you start the run.



NOTE

To view indexing results in Sequencing Analysis Viewer (SAV), it is essential that you load a sample sheet at the start of your run. SAV v1.8 requires a sample sheet to report indexing information. See the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

You can create your sample sheet in a variety of software applications such as Excel or Notepad, or use the Illumina Experiment Manager.

Use the following guidelines when creating your sample sheet:

- ▶ All index sequences are specified in the Index field. If you are doing dual-indexed reads, multiple indices must be separated by a hyphen: ACCAGTAA–GGACATGA.
- ▶ If a sample does not contain an index sequence, leave the Index field blank.
- ▶ For a sample sheet to be valid, every cycle in an Index Read for every lane *must* have at least one base in the red channel (A,C) and one base in the green channel (G,T) to ensure proper image processing. An invalid sample sheet will prevent proper secondary analysis.

Illumina Experiment Manager

The Illumina Experiment Manager is a wizard-based application that guides you through the steps to create your sample sheet.

Using the Illumina Experiment Manager not only reduces syntax errors, but also prompts you to provide information for sections that apply to your sample type and analysis workflow. It provides a feature for recording parameters for your sample plate, such as sample ID, project name, dual indices, and barcode information. Using the Illumina Experiment Manager, you can import sample plate parameters to your sample sheet.

The Illumina Experiment Manager can be run on any Windows platform. You can download a copy from the Illumina website at www.illumina.com. Go to the HiSeq 2500 support page and click **Downloads**. A MyIllumina login is required.

For more information, see the *Illumina Experiment Manager User Guide*, Part # 15031335.

Available Disk Space

During the run at the end of imaging each lane, the software checks available disk space on the local D: and E: drives. The network location is not checked at this time. If disk space drops below the safe threshold, the software pauses the run and places the flow cell in a safe state.



NOTE

Illumina recommends that you specify a network location for the output run folder. The output run folder is specified on the Scan screen. For more information, see *Flow Cell Set-up* on page 41.

If the disk space becomes low, you need to make disk space available to continue the run. When sufficient disk space becomes available, the run automatically resumes.

Switching Sequencing Modes

When you switch between High Output mode to Rapid Run mode and vice versa, a maintenance wash and gasket change are required before you start a new run. If the instrument has been left idle for 10 days or more prior to switching between modes, a maintenance wash will be required before you set up a run, regardless of the mode you have selected. HiSeq Control Software (HCS) will guide you through the process.

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

To change from the current mode, use the mode select option on the HiSeq Control Software (HCS) software main page. You can change modes only when there are no active runs on the system. The mode select function is disabled while a sequencing run is in progress. All flow cells need to complete all mode change procedures in order for the system to properly enter the new mode. Only runs of the same mode type can be performed simultaneously. You cannot run one Rapid Run and one High Output flow cell at the same time.

If you are switching to Rapid Run from High Output, a maintenance wash will have been carried out at the end of the High Output run, and only a water wash will be required. You can use clean, Milli-Q grade (18 M Ω) water in place of PW1 for system washes.

The expected fluidics output volumes are the specified volumes +/- 10%.

Switching to High Output Mode

Once a Rapid Run is finished, you will perform the required post-run water wash (or maintenance wash if after 10 days). The software will automatically guide you through this.

When switching modes, both flow cells must fully complete the mode switch washes. If the mode wash is not carried out for one or both of the flow cells, the software will not let you switch back to the previous mode without completing the full mode wash.

If one position is to be left idle after a mode switch, the mode switch wash can be carried out either at the time of switching to the current mode or just before switching back to the previous mode.

If one of the positions is canceled during the mode switch wash, you will not be able to switch back to the previous mode without completing the mode wash on the canceled position.

If a maintenance wash is required after the previous Rapid Run is complete and you want to switch to High Output mode, to avoid additional washing you can cancel the Rapid Run maintenance wash and switch immediately to High Output mode. However, you will still need to complete the necessary gasket change, Rapid Run maintenance wash, and High Output maintenance wash after you choose to switch to High Output mode.



CAUTION

A gasket change and a Rapid Run maintenance wash followed by a High Output maintenance wash (to condition the gasket) are required when you change the mode to High Output.

To switch to High Output mode, navigate to the Home Screen and select **MODE SELECT**. This takes you to the Run Configuration screen. Select **HIGH OUTPUT**. A warning message will appear to remind you that the gaskets should be changed and that a full Rapid Run maintenance wash should be carried out using a Rapid flow cell before the software switches to High Output mode. Once in High Output mode, a full High Output maintenance wash should be performed.

Table 3 Maintenance Wash Volumes (volumes are +/- 10%)

	Run Mode	
Flow Cell Type	Rapid (2 lane)	High Output (8 lane)
Flow Cell Gasket	10 port	8 or 10 port
Reagents	PW1 > 1 N NaOH > PW1	PW1 > 1 N NaOH > PW1
Expected Per Flow Cell Volumes (ml)	19 + 9.5 + 19 = 47.5	72 + 36 + 72 = 180
Time (minutes)	13 + 6 + 13 = 32	60 + 30 + 60 = 150
Total Mode Change Time	~ 3 hours	

After selecting **OK**, the software will automatically prompt you to change the gaskets prior to the Rapid Run maintenance wash only. Place waste lines 4 and 5 into a 50 ml tube to before starting the wash.

**NOTE**

If you cancel out of this Rapid Run maintenance wash, the software will still be in Rapid Run mode. If you want to remain in Rapid Run mode, a water wash is required before starting a new Rapid Run. If you still want to switch to High Output mode, the required Rapid Run maintenance wash and High Output maintenance wash will need to be completed.

Once in High Output mode, a full High Output maintenance wash should be performed to condition the gasket. The software will automatically guide you through this. During the fluidics check, ensure that all 8 syringes are at least 75% full of water before you start the wash. A small amount of air is not a problem as the air will disperse during the maintenance wash, Gaskets do not need to be changed prior to the High Output maintenance wash. Select whether or not you want to wash the PE positions. Place waste lines 1-8 into an appropriate container to measure the volumes used.

**NOTE**

If you cancel out of the wash once the Rapid Run maintenance wash is complete or at any point once the software is in High Output mode (displayed at top of screen), you will remain in High Output mode. A High

Output maintenance wash will need to be completed prior to starting a run in High Output mode.

Once all washes are complete, you will be ready to start a new High Output run.

Switching to Rapid Run Mode

Once a High Output run is completed, change the gaskets and perform the required post-run maintenance wash. The software will automatically guide you through this. A High Output maintenance wash takes approximately 2 hours and 30 minutes, and the volume should be 180 ml +/- 10% per lane. You will perform this wash prior to starting the switching process.

To switch to Rapid Run mode, navigate to the Welcome Screen, and select **MODE SELECT**. This takes you to the Mode Select screen. Select **RAPID RUN**.



NOTE

If you cancel out of this wash, the software will not let you switch modes until the wash is completed.

A warning message will appear telling you that a maintenance wash will need to be performed before a new run can be started in Rapid Run mode.

Table 4 Maintenance Wash Volumes (volumes are +/- 10%)

	Run Mode
Flow Cell Type	Rapid (2 lane)
Flow Cell Gasket	10 port
Reagents	PW1 > 1 N NaOH > PW1
Expected Per Flow Cell Volumes (ml)	$19 + 9.5 + 19 = 47.5$
Time (minutes)	$13 + 6 + 13 = 32$
Total Mode Change Time	~ 35 minutes

The software automatically guides you through a Rapid Run maintenance wash using a Rapid flow cell. Place waste lines 4 and 5 into a 50 ml tube. The wash will take ~30 minutes.



NOTE

You cannot start a new Rapid Run in Rapid Run mode until the required water wash is completed.

Once the wash is finished, the software automatically brings up the welcome screen. Select **SEQUENCE** and then **NEW RUN** to start a Rapid Run.

Performing a High Output Sequencing Run

Introduction	76
High Output Sequencing Workflow	77
Preparing SBS Reagents	79
Preparing Indexing Reagents	84
Entering Run Parameters	88
Loading and Priming Reagents	93
Loading a Flow Cell	103
Monitoring the Run	110
Preparing Reagents for Read 2	112
Loading Reagents for Read 2	117



Introduction

To perform a High Output sequencing run on the HiSeq 2500, you first prepare reagents for your run and then set up the run using the HCS interface. Run setup steps include: entering run parameters, loading and priming reagents, loading the High Output flow cell, and performing a fluidics check.

- ▶ Reagents for Read 1 and the Index Reads are loaded during run setup prior to starting the run.
- ▶ Reagents for Read 2 resynthesis are loaded after the completion of Read 1 and the Index Reads.
- ▶ Illumina recommends that you prepare and load fresh ICB for Read 2. For more information, see *Preparing ICB* on page 81.

Types of High Output Runs

You can perform six types of High Output runs on the HiSeq 2500:

Run Type	Read 1 Cycles	Index 1 (i7) Read Cycles	Index 2 (i5) Read Cycles	Read 2 Cycles	Total Cycles
Single-Read, Non-Indexed	≤ 101	--	--	--	≤ 101
Single-Read, Single-Indexed	≤ 101	7 * 8 **	--	--	≤ 108 * ≤ 109 **
Single-Read, Dual-Indexed	≤ 101	8	8	--	≤ 117
Paired-End, Non-Indexed	≤ 101	--	--	≤ 101	≤ 202
Paired-End, Single-Indexed	≤ 101	7 * 8 **	--	≤ 101	≤ 209 * ≤ 210 **
Paired-End, Dual-Indexed	≤ 101	8	7 + 8 ***	≤ 101	≤ 225

* Number of cycles for single-indexed libraries

** Number of cycles for dual-indexed libraries

*** Index 2 Read of a paired-end dual-indexed run includes seven additional chemistry-only cycles

For more information, see *Single-Indexed Sequencing Overview* on page 13 and *Dual-Indexed Sequencing Overview* on page 16.

High Output Sequencing Workflow



Prepare SBS reagents. This includes thawing and inverting all reagents.



For indexed runs:

Prepare indexing reagents. Make sure that you prepare the appropriate sequencing primers for your library and type of run.



Using the HiSeq Control Software (HCS), enter parameters for your run.



When prompted, load SBS reagents for Read 1 and Read 2, except ICB (recommended).

For indexed runs:

Load indexing reagents. For dual-indexed runs, load reagents for both Index Reads.



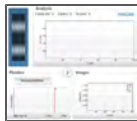
With a used flow cell on the instrument, confirm proper flow as prompted by the HCS user interface.

Prime SBS reagents and measure priming waste.

Note: For Rapid Run mode you do not need to load the used flow cell as priming is completed within recipe once the run has started.



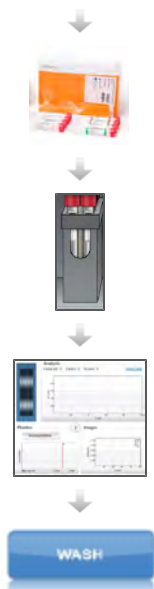
Load a flow cell that was previously clustered on the cBot. Confirm proper flow as prompted by the HCS user interface.



Start your sequencing run. After cycle 1, inspect the first base report (optional setting), and then continue Read 1.

For indexed runs:

Following Read 1, the run proceeds to the Index Read (single-indexed runs) or two Index Reads (dual-indexed runs).



For paired-end runs:

Prepare paired-end reagents for Read 2 Resynthesis.
Prepare fresh ICB for Read 2 (recommended).

For paired-end runs:

Load paired-end reagents onto the paired-end rack.
Load fresh ICB in position 1 (recommended).

For paired-end runs:

Continue the run. The software automatically primes paired-end reagents and performs Read 2 resynthesis and Read 2.

When the run is complete, unload and weigh reagents.
Perform an instrument maintenance wash.

Approximate Run Duration

The following table lists the approximate duration of a High Output sequencing run based on read length and version of instrument software.

Run Format	HCS v2.0 or later with Flow Cell v3
36	Approx. 2 days
2 x 50	Approx. 5 days
2 x 100	Approx. 10 days
2 x 150 Cycles	N/A

Preparing SBS Reagents

The first step in setting up a run is to prepare SBS reagents. Reagent preparation requires overnight thawing at 2° to 8°C, or about 90 minutes in a room temperature deionized water bath. When completely thawed, reagents are ready for use. For High Output mode, Incorporation Mix requires preparation before use.



NOTE

TruSeq SBS v3 reagents enable an alternative workflow of loading all SBS reagents at the start of a paired-end sequencing run for both Read 1 and Read 2. Although this alternative SBS workflow provides acceptable sequencing performance in Read 2, Illumina recommends that you prepare and load fresh ICB at the beginning of Read 2 for optimal performance.

SBS Kits for Dual-Indexed Sequencing

One 200-cycle SBS Kit provides sufficient reagents for up to 209 cycles of sequencing. For runs greater than a total of 209 cycles, and to maximize reagent use for other read lengths, you can use multiple 50-cycle kits.

The following table lists the number of kits required for dual-indexed runs. For dual-indexed paired-end runs, seven chemistry-only cycles are calculated in the total number of cycles.

Run Type	Run Cycles	Total Cycles	Kit Option 1 SBS Kits (200-cycle)	Kit Option 2 SBS Kits (50-cycle)
Dual-Indexed Paired-End	101+8+(7)+8+101	225	-	4*
	93+8+(7)+8+93	209	1	4*
	76+8+(7)+8+76	175	1	3
	51+8+(7)+8+51	125	1	3
	36+8+(7)+8+36	95	1	2
Dual-Indexed Single-Read	101+8+8	117	1	2
	76+8+8	92	1	2
	51+8+8	67	1	2
	36+8+8	52	1	1

*Fill the container of SRE and SB3 to the top. No replenishing is required.

For more information, see *Dual-Indexed Sequencing Overview* on page 16.

Thaw SBS Reagents

- 1 Remove SRE and CMR from -15° to -25°C storage, thaw, and invert according to the instructions in the following table:

Reagent Name	SBS Kit (200 Cycles)	SBS Kit (50 Cycles)
SRE	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for 90 minutes. Invert 5 times to mix. Leave on ice until ready to use.	Thaw in a room temperature water bath for one hour.
CMR	Thaw at 2° to 8°C for 16 hours or in room temperature water bath for 90 minutes. Invert 5 times to mix. Leave on ice, away from IMM and SRM until ready to use.	Thaw in a <i>separate</i> room temperature water bath for one hour.

- 2 When thawed, set SRE and CMR aside on ice until you are ready to load reagents.
- 3 Remove LFN36 from -15° to -25°C storage and thaw in a beaker containing room temperature deionized water for 20 minutes.

Reagent Name	SBS Kit (200 Cycles)	SBS Kit (50 Cycles)
LFN	Thaw two tubes to prepare ICB for Read 1. Thaw the other two tubes when you prepare ICB for Read 2 or a subsequent run. Alternatively, thaw four tubes to prepare ICB for both Read 1 and Read 2.	Thaw one tube. One tube of LFN prepares ICB for 58 cycles of sequencing.

- 4 When thawed, set LFN36 aside on ice until you are ready to prepare ICB.
- 5 Leave EDP in -15° to -25°C storage until you are ready to prepare ICB.
- 6 Use ICB, SB1, SB2, and SB3 directly from 2° to 8°C storage.
- 7 Record the lot numbers of each reagent on the lab tracking form.

Preparing ICB

Prepare ICB according to your type of run and the size of kit you are using:

- ▶ For non-indexed or single-indexed runs using a 200-cycle kit, see *Prepare ICB for Read 1 (200-Cycle Kit)* on page 81 and *Prepare ICB for Read 2 (200 Cycle Kit)* on page 116.
- ▶ For non-indexed or single-indexed runs using a 50-cycle kit, see *Prepare ICB for 58 Cycles (50-Cycle Kit)* on page 81.
- ▶ For dual-indexed runs, see *Prepare ICB for Dual-Indexed Runs* on page 82.
- ▶ For non-indexed or single-indexed runs using the alternative workflow for SBS v3 reagents, see *Prepare ICB for Read 1 and Read 2 (200-Cycle Kit)* on page 83.



NOTE

When you prepare ICB, gently invert the bottle of ICB to mix. For best performance, make sure that the temperature of ICB remains below 8°C.

Prepare ICB for Read 1 (200-Cycle Kit)

Use the following instructions to prepare ICB for a non-indexed or single-indexed run up to 2 x 101 cycles. For dual-indexed runs, see *Prepare ICB for Dual-Indexed Runs* on page 82.

- 1 Aliquot 47 ml of ICB-200 into a spare 250 ml bottle to result in two bottles of ICB.
- 2 Store the spare bottle containing 47 ml of ICB at 2° to 8°C storage until you are ready to prepare and load reagents for Read 2.
- 3 Add the contents of two tubes of LFN to the original bottle of ICB-200.
- 4 Rinse each tube of LFN with ICB-200 to make sure that all LFN is transferred.
- 5 Aliquot 1.1 ml of EDP-200 into the ICB/LFN mix.
- 6 Return the unused portion of EDP-200 to -15° to -25°C storage.
- 7 Cap the bottle and *gently* invert several times to mix.
- 8 Set aside on ice until you are ready to load reagents onto the instrument.

Prepare ICB for 58 Cycles (50-Cycle Kit)

Use the following instructions to prepare ICB for up to 58 cycles of sequencing.

- 1 Add the contents of one tube of LFN to the bottle of ICB-50.
- 2 Rinse the tube of LFN with ICB-50 to make sure that all LFN is transferred.
- 3 Add the contents of one tube of EDP-50 into the ICB/LFN mix.
- 4 Rinse the tube of EDP-50 with ICB/LFN mix to make sure that all EDP is transferred.
- 5 Cap the bottle and *gently* invert several times to mix.
- 6 Set aside on ice until you are ready to load reagents onto the instrument.

Prepare ICB for Dual-Indexed Runs

Illumina recommends the following procedure to calculate the volume of ICB required for a dual-indexed run.

For more information about the number of SBS Kits required for a dual-indexed run based on the read length, see *SBS Kits for Dual-Indexed Sequencing* on page 79.

- 1 At the beginning of your run, aliquot ICB, LFN, and EDP using the following volumes for every ten cycles of sequencing to be performed. Do **not** mix ICB, LFN, and EDP.

Reagent	Volume per Ten Cycles	Storage
ICB	4.7 ml	2° to 8°C
LFN	0.7 ml	-15° to -25°C
EDP	0.11 ml	-15° to -25°C

For example, a 101-cycle dual-indexed paired-end run has a total of 225 cycles:

- 124 cycles (101 + 8 + 7 + 8) for Read 1 and the Index Reads
- 101 cycles for Read 2

Read 1 and Index Reads (124 Cycles)	Read 2 (101 Cycles)
124 cycles / 10 cycles = 12.4	101 cycles / 10 cycles = 10
• ICB: 12.4 x 4.7 ml = 58.3 ml	• ICB: 10 x 4.7 ml = 47 ml
• LFN: 12.4 x 0.7 ml = 8.7 ml	• LFN: 10 x 0.7 ml = 7 ml
• EDP: 12.4 x 0.11 ml = 1.36 ml	• EDP: 10 x 0.11 ml = 1.1 ml

You can use the same calculation for any dual-indexed run based on the total number of cycles in Read 1 and the Index Reads, and the total number of cycles in Read 2. For dual-indexed paired-end runs, include the seven chemistry-only cycles in the Index 2 Read.

- 2 Store each volume of ICB, LFN, and EDP at the proper storage, as needed. For paired-end runs, store the aliquots calculated for Read 2 until you are ready to prepare fresh ICB for Read 2.
- 3 Prepare ICB, LFN, and EDP as follows:
 - a Add LFN to the bottle of ICB.
 - b Rinse each tube of LFN with ICB to make sure that all LFN is transferred.
 - c Add EDP to the ICB/LFN mix.
 - d Rinse the tube of EDP with ICB/LFN mix to make sure that all EDP is transferred.
 - e Cap the bottle and *gently* invert several times to mix.
 - f Set aside on ice until you are ready to load reagents onto the instrument.

Prepare ICB for Read 1 and Read 2 (200-Cycle Kit)

If you plan to use the alternative workflow and load ICB for Read 1 and Read 2 at the start of your run, use the following instructions for runs up to 209 total cycles for single indexing runs.

- 1 Add the contents of four tubes of LFN to the bottle of ICB-200.
- 2 Rinse each tube of LFN with ICB-200 to make sure that all LFN is transferred.
- 3 Add the contents of one tube of EDP-200 into the ICB/LFN mix.
- 4 Rinse the tube of EDP-200 with ICB/LFN mix to make sure that all EDP is transferred
- 5 Cap the bottle and *gently* invert several times to mix.
- 6 Set aside on ice until you are ready to load reagents onto the instrument.

Preparing Indexing Reagents

Indexed sequencing requires indexing reagents based on the type of libraries you are sequencing and the number of Index Reads that you perform.

- ▶ All library types require HP3 and HT2 provided in the TruSeq Cluster Kit.
- ▶ Dual-indexed Nextera libraries require additional sequencing primers provided in the TruSeq Dual Index Sequencing Primer Box. Use the single-read kit when sequencing on a single-read flow cell, and the paired-end kit when sequencing on a paired-end flow cell.
- ▶ Dual-indexed TruSeq HT libraries require HP9 (Index 2 Read primer) from the TruSeq Dual Index Sequencing Primer Box (Single Read) when performing a dual-indexed run on a single-read flow cell.



NOTE

The sequencing primers provided in the TruSeq Dual Index Sequencing Primer Box are compatible with non-indexed libraries and single-indexed libraries. If your flow cell contains a combination of non-indexed, single-indexed, and dual-indexed libraries, use the sequencing primers in the TruSeq Dual Index Sequencing Primer Box.

For more information, see *Sequencing Primers By Run Type* on page 24.

Library Type and Flow Cell Type	Index 1 (i7)	Index 2 (i5)
Nextera libraries Single-Read Flow Cell	HP12	HP9
Nextera libraries Paired-End Flow Cell	HP12	RMX
TruSeq HT libraries Single-Read Flow Cell	HP8 or HP12	HP9
TruSeq HT libraries Paired-End Flow Cell	HP8 or HP12	RMX
TruSeq LT/v2 libraries Single-Read Flow Cell or Paired-End Flow Cell	HP8 or HP12	--
TruSeq Small RNA libraries Single-Read Flow Cell or Paired-End Flow Cell	HP8 or HP12	--

Indexing reagents take less than 30 minutes to prepare.

Table 5 Indexing Reagents

Reagents	Kit Name
HP3	TruSeq PE Cluster Kit (Indexing Reagents Box)
HT2	TruSeq PE Cluster Kit (Indexing Reagents Box)
HP8 or HP12	TruSeq PE Cluster Kit (Indexing Reagents Box)
HP12	TruSeq Dual Index Sequencing Primer Box
HP9	TruSeq Dual Index Sequencing Primer Box
RMX	TruSeq PE Cluster Kit (PE Reagents Box)


NOTE

The sequencing primers provided in the TruSeq Dual Index Sequencing Primer Box are compatible with non-indexed libraries and single-indexed libraries. If your flow cell contains a combination of non-indexed, single-indexed, and dual-indexed libraries, use the sequencing primers in the TruSeq Dual Index Sequencing Primer Box.


NOTE

HP8, provided in the TruSeq Cluster Kit, is not compatible with dual-indexed Nextera libraries. You must use **HP12** from the TruSeq Dual Index Sequencing Primer Box.

Thaw Indexing Reagents

- 1 Remove the reagents from -15° to -25°C storage. Confirm that you are using the appropriate sequencing primers for your library type and run type.
- 2 Thaw reagents in a beaker filled with room temperature deionized water for about 20 minutes or until reagents are fully thawed.

Prepare HT2

- 1 Invert the tube of HT2 five times to mix the reagent.
- 2 Label the tube of HT2 **Reagent #19**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare HP3 for Single-Indexed Runs

- 1 Invert the tube of HP3 five times to mix the reagent, and then pulse centrifuge.
- 2 Transfer 3,325 μ l of PW1 into a 15 ml Sarstedt conical tube and add 175 μ l of HP3.
- 3 Invert the tube five times to mix the reagent.
- 4 Label the tube of diluted HP3 **Reagent #18**.
- 5 Centrifuge at 1,000 rpm for one minute.
- 6 Set aside at room temperature.

Prepare HP3 for Dual-Indexed Runs



NOTE

In order to provide sufficient HP3 for both Index Reads of a dual-indexed run, the preparation instructions for HP3 differ from instructions for single-indexed runs.

- 1 Invert the tube of HP3 five times to mix the reagent, and then pulse centrifuge.
- 2 Transfer 3,800 μ l of PW1 into a 15 ml Sarstedt conical tube and add 200 μ l of HP3.
- 3 Invert the tube five times to mix the reagent.
- 4 Label the tube of diluted HP3 **Reagent #18**.
- 5 Centrifuge at 1,000 rpm for one minute.
- 6 Set aside at room temperature.

Prepare HP8

- 1 Invert the tube of HP8 five times to mix the reagent.
- 2 Briefly pulse centrifuge to collect droplets.
- 3 Label the tube of HP8 **Reagent #17**.
- 4 Set aside at room temperature.

Prepare HP12

You must use HP12 for sequencing dual-indexed Nextera libraries.

- 1 Invert the tube of HP12 five times to mix the reagent.
- 2 Briefly pulse centrifuge to collect droplets.
- 3 Label the tube of HP12 **Reagent #17**.
- 4 Set aside at room temperature.

Prepare HP9

You must use HP9 for dual-indexed runs on a single-read flow cell.

- 1 Invert the tube of HP9 five times to mix the reagent.
- 2 Briefly pulse centrifuge to collect droplets.
- 3 Label the tube of HP9 **Reagent #16**.
- 4 Set aside at room temperature.

Prepare RMX

Prepare RMX for dual-indexed runs on a paired-end flow cell.

- 1 Invert the tube of RMX five times to mix the reagent.
- 2 Label the tube of RMX **Reagent #10**.
- 3 Briefly centrifuge the reagent at 1,000 rpm for up to one minute. Alternatively, gently tap or flick the tube to make sure that reagent droplets collect at the bottom of the tube.



NOTE
Do not vortex RMX.

- 4 Set aside on ice.

Entering Run Parameters

The HiSeq Control Software (HCS) guides you through the run setup steps. From the HCS Welcome screen, select **Sequence | New Run**. The Scan screen opens.

Storage Screen

- 1 Select **Use BaseSpace for storage and analysis** if you want to save your data to a BaseSpace account. If you select this option you will be prompted to enter a My Illumina account email and password.



NOTE

For every run, your HiSeq 2500 will attempt to send general run metrics to BaseSpace. These metrics include interoperation files (that SAV uses) and global files such as RunInfo.xml and RunParameters.xml. No data files are ever sent, and all information is sent anonymously, associated with the HiSeq 2500 from which it came, not to any particular user. This feature is enabled by default and Illumina recommends that you leave it enabled. To disable this feature, select **Menu | Tools | Options**, and then unselect **Send system health information to Illumina to aid technical support**. You can also view terms in the Options dialog box.

- 2 In **Save to an output folder**, browse to and select the path of your run folder.
- 3 Select **Zip BCL files** to reduce the footprint of disc space needed for the run. If you are using BaseSpace, this option is selected by default.



NOTE

If you are using BaseSpace or an updated version of CASAVA for data analysis, zipping BCL files is mandatory.

- 4 Select **Bin QScores** to reduce the footprint of disc space needed for the run.
- 5 Select a setting from the **Save Auxiliary Files** drop-down list:
 - Select **Save All Thumbnails** to save tile level and full thumbnails.
 - Select **Save Tile Thumbnails** to save tile level thumbnails only.
- 6 Select **Next**.

Flow Cell Set-up Screen



NOTE

It is critical that you scan the flow cell barcode or accurately enter the flow cell ID when you set up your run. The software uses the flow cell ID to determine flow cell type and reagent compatibility.

- 1 Scan the flow cell barcode or enter the flow cell ID (barcode number) of the flow cell to be sequenced.
- 2 Confirm the flow cell type, which is automatically selected based on the flow cell ID.
- 3 Enter the experiment name and user name.
- 4 Select the lane containing the control, if applicable. Otherwise, select **None**.
- 5 Select **Next**.

Advanced Screen

- 1 Select the **Confirm First Base** checkbox to generate a first base report confirmation.
- 2 Select a setting from the **Control Lane** drop-down list.
- 3 Select or deselect **Align to PhiX** for each lane of the flow cell.
- 4 Select **Keep Intensity Files** for later reanalysis or custom processing (optional).
- 5 Select **Next**.

Recipe Screen

- 1 Select one of the following indexing options:
 - **No Index**—Performs a non-indexed single-read or paired-end run.
 - **Single Index**—Performs a single-read or paired-end run with one Index Read (Index 1 (i7) Read).
 - **Dual Index**—Performs a single-read or paired-end run with two Index Reads (Index 1 (i7) Read and Index 2 (i5) Read).
 - **Custom**—Performs a single-read or paired-end run with a custom number of cycles for the Index Reads.

- 2 Select a flow cell format, either **Single Read** or **Paired End**. This option appears in High Output mode only if you have selected Dual Index or Custom.
- 3 Enter the number of cycles you plan to sequence for Read 1 and Read 2, if applicable.

If you selected the **Custom** indexing option, enter the number of cycles you plan to sequence for the Index Reads.



NOTE

It is possible to perform a dual-indexed single-read run using a paired-end flow cell containing paired-end libraries. To do so, set up a *dual-indexed paired-end run* and enter zero for the number of cycles in Read 2. This will stop the run after the Index 2 (i5) Read.

- 4 Confirm the chemistry selections for your run:
 - a **SBS**—This setting is determined by the flow cell ID.
 - b **Index**—This setting is determined by the indexing option selected. However, you must specify the kit name for single-indexed runs to make sure that the correct number of indexing cycles are applied to your run.
 - **For non-indexed or single-indexed libraries**—Select **Multiplexing Primer Box**. This sets up your run to perform a seven-cycle Index Read.
 - **For dual-indexed libraries**—Select **TruSeq Dual Index Sequencing Primer Box**. This sets up your run to perform an eight-cycle Index Read.



NOTE

Select TruSeq Dual Index Sequencing Primer Box setting for sequencing any dual-indexed libraries (Nextera or TruSeq HT) regardless of which sequencing primers you will use for your run. Selecting any other setting will result in less than eight-cycle Index Read.

For more information about which sequencing primers to use for your run, see *Sequencing Primers By Run Type* on page 24.

- c **PE turnaround chemistry**—This setting is determined by the flow cell ID.
- 5 Select the **Use Existing Recipe** checkbox only if you want to use a custom recipe. Otherwise, allow the software to create your recipe based on information you provide on the run setup screens.

Sample Sheet Screen



NOTE

To view indexing results in Sequencing Analysis Viewer (SAV), it is essential that you load a sample sheet at the start of your run. SAV v1.8 requires a sample sheet to properly report indexing information. For more information, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

- 1 Enter the path or browse to a valid sample sheet for your run. For indexed runs, the sample sheet is required if you want indices to be reported in SAV. A sample sheet is also required if you are using BaseSpace.
- 2 Select **Next**.

Reagents Screen

- 1 Scan or enter the Reagent Kit ID. You can enter the kit ID from either box 1 or box 2.
- 2 For indexed runs, scan or enter the Indexing Reagent Kit ID.
- 3 Select the SBS reagent kit you are using for your run.
 - Select **200 Cycles** for a 200 cycle kit; select **50 Cycles** for a 50 cycle kit.
 - Select **Used Kit** for a partial kit, and then enter the number of SBS cycles remaining, or the number of cycles that reagents are expected to last.



NOTE

If you are combining four 50-cycle SBS kits for a dual-indexed run, select **Used Kit** instead of 50 Cycles to correctly set the cycle counter.

In High Output mode, the counter defaults to 225 cycles, which you can change according to the number of cycles in your run.

In Rapid Run mode, the counter defaults to 325 cycles, which you can change according to the number of cycles in your run.

- 4 For indexed runs, scan or enter the Indexing Reagent Kit ID.
- 5 For paired-end runs, scan or enter the PE or SR Reagent Kit ID.
- 6 The Cycles Remaining field is populated automatically depending on the SBS kit you are using. If you are using a partial kit, enter the number of cycles reagents are expected to last. The software counts down the number of cycles entered. When the

cycles are low, the software prompts you to load fresh reagents. For more information, see *Changing Reagents During a Run* on page 63.



NOTE

If you are starting a new run using less than a full reagent kit, make sure that you can be present when new reagents are needed.

- ▶ Select **Prime SBS Reagents** to prime reagents before you start a run. Always prime reagents before loading a new flow cell.
- 7 Select **Next**.

Review Screen

Review your run information. If it is correct, select **Next**.

Select **Back** if you need to change any entries.

Loading and Priming Reagents

The next step is to load sequencing reagents. The software interface guides you through the process.

Load SBS Reagents

Make sure that SBS reagents are thoroughly thawed and ready to load onto the instrument.

ILLUMINA-SUPPLIED CONSUMABLES

- ▶ Eight funnel caps (Provided in the TruSeq Cluster Kit)

USER-SUPPLIED CONSUMABLES

- ▶ One 250 ml bottle (Corning, catalog # 430776)

Procedure

- 1 Record the weight of each reagent on the lab tracking form.



NOTE

Weighing reagents before and after a sequencing run confirms proper reagent delivery.

- 2 Open the reagent compartment door.
- 3 Raise the sippers for the sequencing reagent rack using the following motion:
 - a Pull the sipper handle towards you.
 - b Push the sipper handle up while pulling it towards you.
 - c Release the sipper handle into the slot on the top end of the groove. You should feel the sipper handle rest securely into the slot.
- 4 Slide the reagent rack out of the reagent compartment.
- 5 Place each reagent bottle onto the rack in the associated numbered position, making sure that the conical end of the bottle rests in the indentation on the base of the rack.

Table 6 SBS Reagent Positions

Position	Reagent	Description
1	ICB	Incorporation Mix
2	PW1 (25 ml)	Wash Buffer
3	SRE	Scan Mix Reagent
4	SBS Buffer 1 (SB1)	High Salt Buffer
5	SBS Buffer 2 (SB2)	Incorporation Wash Buffer
6	SBS Buffer 2 (SB2)	Incorporation Wash Buffer
7	CMR	Cleavage Mix Reagent
8	SBS Buffer 3 (SB3)	Cleavage Buffer

- 6 Load the template into the sample loading station.
- 7 Add 25 ml of PW1 or laboratory-grade water to the bottle in position 2.



NOTE

HCS v2.0 and later includes an automatic post-run rinse following the completion of the sequencing run. To prepare for the post-run rinse, it is important that you load 25 ml of PW1 or laboratory-grade water in position 2 when you load SBS reagents at the beginning of your run.

The post-run rinse does *not* replace the post-run instrument wash.

- 8 Remove the cap from each reagent bottle and replace it with a funnel cap.



CAUTION

After handling the bottle of CRM, discard your gloves and replace them with a new pair.

- 9 Slide the reagent rack into the reagent compartment, aligning the rack with the raised guide on the floor of the compartment.
- 10 Lower the sippers into the sequencing reagent bottles as follows:
 - a Pull the sipper handle towards you.
 - b Lower the sipper handle while pulling it towards you.
 - c Visually inspect the sippers to ensure that they do not bend as they lower into the funnel caps.
 - d Release the sipper handle into the slot on the bottom end of the groove. You should feel the sipper handle rest securely into the slot.
- 11 Select the checkbox labeled **PW1 (25mL) loaded in Position 2**.
- 12 Do one of the following:

- If you are performing an indexed run, select **Next**. The software prompts you to load indexing reagents.
- If you are performing a non-indexed run, close the reagent compartment door and select **Next**. The software proceeds to priming reagents.

Load Indexing Reagents

If you are performing an indexed run, load indexing reagents before starting the run.

- 1 Before beginning, record the weight of each reagent on the lab tracking form.
- 2 Raise the sippers for the paired-end reagent rack using the following motion:
 - a Pull the handle towards you.
 - b Push the handle up while pulling it towards you.
 - c Release the handle into the slot on the top end of the groove. You should feel the handle rest securely into the slot.
- 3 Slide the reagent rack out of the reagent compartment using the rack handle.
- 4 Place each reagent tube onto the rack in the associated numbered position depending on the type of run you are performing.

Table 7 Single-Indexed Run

Position	Reagent	Description
17	HP12	Index 1 (i7) Sequencing Primer Mix
18	HP10	R1 Primer
19	--	Not Used

Table 8 Dual-Indexed Run on a Single-Read Flow Cell

Position	Reagent	Description
16	HP9	Index 2 (i5) SR Sequencing Primer Mix
17	HP12	Index 1 (i7) Sequencing Primer Mix
18	HP10	R1 Primer
19	--	Not Used

Table 9 Dual-Indexed Run on a Paired-End Flow Cell

Position	Reagent	Description
10	RMX	Resynthesis Mix
16	HP11	R2 Primer
17	HP12	Index 1 (i7) Sequencing Primer Mix
18	HP10	Denaturation Solution
19	--	Not Used



NOTE

Place tubes filled with deionized water into rack positions that are not used.

- 5 Remove the caps from each reagent tube.
- 6 Slide the reagent rack into the reagent compartment, aligning the rack with the raised guide on the floor of the compartment.
- 7 Lower the sippers into the tubes on the paired-end reagent rack as follows:
 - a Pull the handle towards you.
 - b Lower the handle while pulling it towards you.
 - c Visually inspect the sippers to make sure that they do not bend as they lower into the tubes.
 - d Release the handle into the slot on the bottom end of the groove. You should feel the handle rest securely into the slot.
- 8 Close the reagent compartment door.
- 9 Select **Next** to proceed to *Priming Reagents* on page 145.

Priming Reagents



NOTE

You will need to prime reagents only if you have done template hybridization on the cBot for a Rapid Run flow cell.

Perform the following steps to prime reagents:

- ▶ *Clean the Flow Cell Holder* on page 146
- ▶ *Load the Used Flow Cell* on page 98
- ▶ *Confirm Proper Flow* on page 100
- ▶ *Prime Reagents* on page 101

User-Supplied Consumables

- ▶ Lens cleaning tissue
- ▶ 70% ethanol or alcohol wipes
- ▶ Low-lint lab tissue
- ▶ Used flow cell

Clean the Flow Cell Holder

- 1 Open the flow cell compartment door.

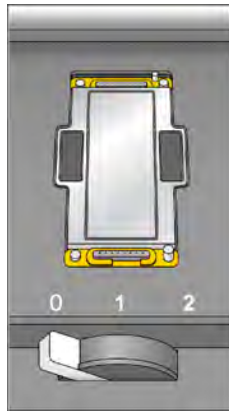


CAUTION

Do not place fluids on the flow cell compartment door or on the flow cell stage when the door is open. Spills in this area can damage the instrument.

- 2 Make sure that the flow cell lever is in the OFF position.

Figure 34 Flow Cell Lever in Position 0



- 3 Put on a new pair of powder-free latex gloves.
- 4 If the flow cell from a previous run is present, remove it and set aside. It can be used to confirm proper flow prior to loading the clustered flow cell.
- 5 Using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol, carefully wipe the surface of the flow cell holder until it is completely clean.

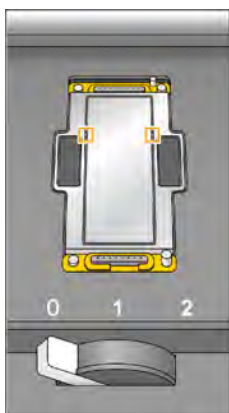


CAUTION

Do not allow alcohol to drip into the vacuum holes or around the manifolds. Use a low-lint lab tissue to dry the stage, if necessary.

- 6 Visually inspect the flow cell holder to make sure that it is free of lint and the vacuum holes are free of obstructions.

Figure 35 Vacuum Hole Locations



Load the Used Flow Cell

- 1 Rinse the used flow cell with laboratory-grade water. Dry it with lens cleaning tissue or lint-free tissue.
- 2 Clean the flow cell using alcohol wipes and lens cleaning tissue.



NOTE

Do not remove or replace the manifold gaskets at this time.

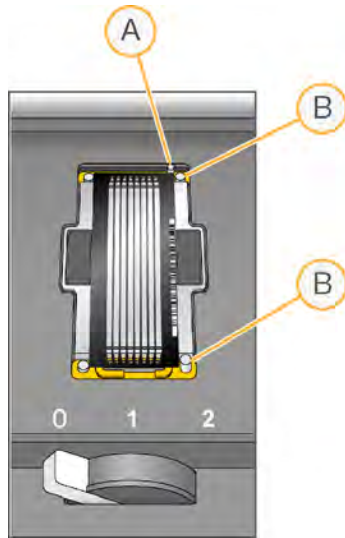
- 3 Place the used flow cell on the flow cell holder with the inlet and outlet ports facing *down* and the barcode on the right. The arrow on the left edge of the flow cell indicates flow direction, and should point towards the instrument.
- 4 Gently slide the flow cell towards the top and right guide pins until it stops.



NOTE

Remove your hand from the flow cell before engaging the vacuum switch. This prevents any possible alignment drift over time.

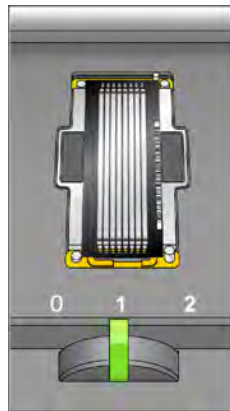
Figure 36 Flow Cell Positioned Against Top and Right Guide Pins.



- A Top Guide Pin
- B Right Guide Pins

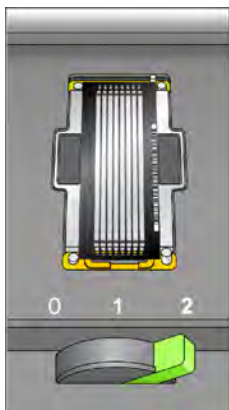
- 5 Slowly move the flow cell lever to position 1. This engages the vacuum and secures the flow cell into position. When the flow cell lever is green, the vacuum is engaged.

Figure 37 Flow Cell Lever in Position 1



- Wait for about five seconds, and then slowly move the flow cell lever to position 2 (far-right). When the flow cell lever is solid green, the manifolds are in position and the flow cell is ready for use.

Figure 38 Flow Cell Lever in Position 2



- Make sure that the **Vacuum Engaged** checkbox is selected on the load prime flow cell screen, and then select **Next**.

Confirm Proper Flow

After the used flow cell is loaded, check for proper flow. Checking for proper flow confirms that the flow cell and gaskets are properly installed and the manifold is engaged.

- Select solution 2 (laboratory-grade water) from the drop-down list.



CAUTION

You can use water to confirm proper flow on a used flow cell only. Never use water to confirm proper flow on a clustered flow cell.

- Confirm the following default values:
 - Volume: **250**
 - Aspirate Rate: **250**
 - Dispense Rate: **2000**
- Select **Pump**.

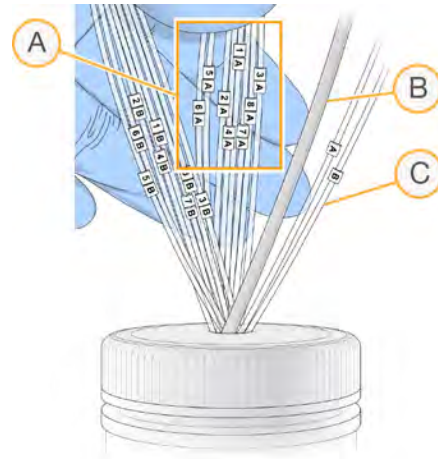
- 4 Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.

If you see excessive bubbles, check the gaskets for obstructions, reduce the aspirate rate to 100, and pump another 250 μ l of water to the flow cell. If problems persist, remove the flow cell, repeat the cleaning steps, and reload the flow cell.

Prime Reagents

Loosen and remove the eight lines of waste tubing for the appropriate flow cell from the waste container. Make sure that you do not include the eight lines for the opposite flow cell or the lines for the condensation pump or the paired-end priming pumps

Figure 39 Waste Lines for a HiSeq 2500



- A Flow Cell Waste Lines for Reagent Positions 1–8 on High Output Flow Cell
- B Condensation Pump Tubing (Do not remove)
- C Paired-End Priming Pump Tubing (Do not remove)

- 1 Place each waste tubing into an empty 15 ml tube, one line per tube. Priming waste is collected and measured after the priming step.
- 2 Select **Next**. The priming screen opens and the priming step starts automatically. You can monitor the progress of the priming step from the priming screen.

- 3 When the priming step is complete, measure the collected priming waste and confirm that the volume is 2 ml from each reagent position primed for a total of **14 ml**. Record the results on the lab tracking form.



NOTE

The color of priming waste might be brown in appearance. This is normal. The change in waste color is not toxic and does not impact run performance.

- 4 Return the waste tubing to the waste container before proceeding.
- 5 Select **Next**. You are ready to load the clustered flow cell.

Loading a Flow Cell

The next step is to remove the used flow cell and load the flow cell that you want to sequence.

Perform the following steps to load the flow cell:

- ▶ *Remove the Used Flow Cell* on page 103
- ▶ *Clean the Flow Cell Holder* on page 152
- ▶ *Clean the Flow Cell* on page 153
- ▶ *Load the Flow Cell* on page 106
- ▶ *Confirm Proper Flow* on page 109

User-Supplied Consumables

- ▶ Lens cleaning tissue
- ▶ 70% ethanol or alcohol wipes
- ▶ Low-lint lab tissue



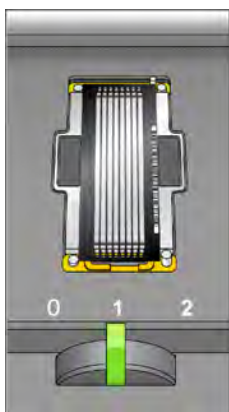
NOTE

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

Remove the Used Flow Cell

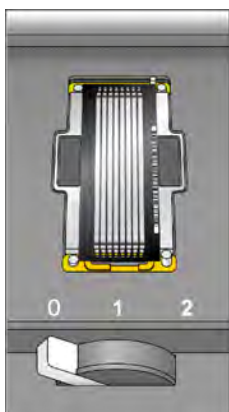
- 1 Open the flow cell compartment door.

Figure 40 Flow Cell Lever in Position 1



- 2 Slowly move the flow cell lever to position 1 to disengage the manifolds.
- 3 Slowly move the flow cell lever to position 0 to disengage the vacuum seal and release the flow cell.

Figure 41 Flow Cell Lever in Position 0



- 4 Lift the used flow cell from the flow cell holder.

Clean the Flow Cell Holder

- 1 Put on a new pair of powder-free latex gloves.

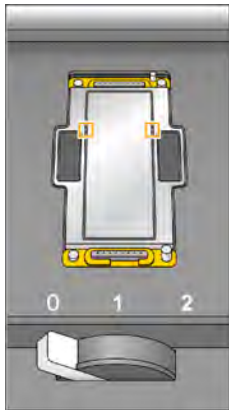
- Using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol, carefully wipe the surface of the flow cell holder until it is completely clean.



CAUTION

Do not allow alcohol to drip into the vacuum holes or around the manifolds. Use a low-lint lab tissue to dry the stage, if necessary.

Figure 42 Inspect Vacuum Holes



- Visually inspect the flow cell holder to make sure that it is free of lint and the vacuum holes are free of obstructions.

Clean the Flow Cell

- Rinse the flow cell with laboratory-grade water and dry it with a lens cleaning tissue.
- Fold an alcohol wipe to approximately the size of the flow cell.
- Hold the edges of the clustered flow cell with two fingers. Make sure that the inlet and outlet ports are facing *up*.
- Wipe off each side of the flow cell with a single sweeping motion. Repeat, refolding the alcohol wipe with each pass, until the flow cell is completely clean.
- Dry the flow cell using a dry lens cleaning tissue.



CAUTION

If you clean the flow cell while it is lying on the bench top, you could easily apply too much pressure and break the flow cell. Illumina recommends cleaning the flow cell while holding the edges between your fingers.

- 6 Protect the flow cell from dust until you are ready to load it onto the instrument.

Load the Flow Cell



NOTE

Do not replace the manifold gaskets at this time. Leave the gaskets that were installed prior to the instrument maintenance wash in place when loading the flow cell. You will replace manifold gaskets after the current sequencing run is completed, prior to the instrument maintenance wash.

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

Do not perform more than one sequencing run on the same set of manifold gaskets .

- 1 Place the flow cell on the flow cell holder with the inlet and outlet ports facing *down* and the barcode on the right. The arrow on the left edge of the flow cell indicates flow direction, and should point towards the instrument.

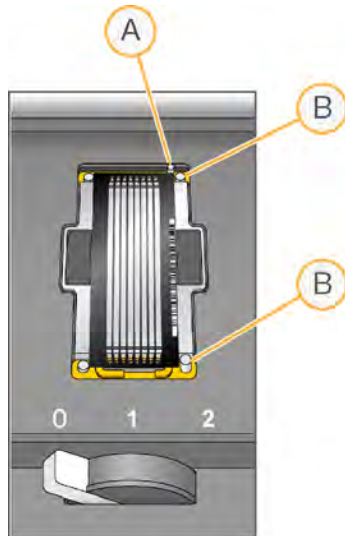


NOTE

Flow Cell v3 and the Rapid Run Flow Cell have a mechanically keyed corner, which provides a visual orientation for loading the flow cell. Install the Flow Cell so that the keyed corner is on the output end of the flow cell facing towards the instrument, and on the left side of the flow cell by lane 1.

- 2 Gently slide the flow cell towards the top and right guide pins until it stops.

Figure 43 Flow Cell Positioned Against Top and Right Guide Pins



- A Top Guide Pin
- B Right Guide Pins

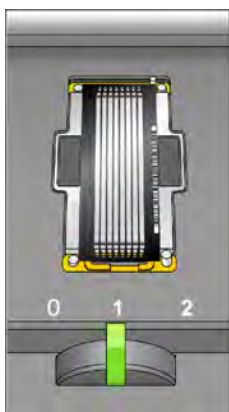


NOTE

Remove your hand from the flow cell before engaging the vacuum switch. This prevents possible alignment drift over time.

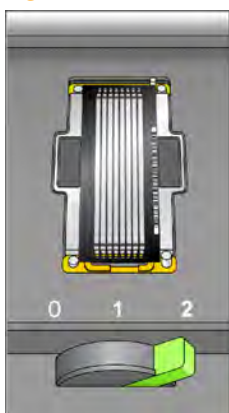
- 3 Slowly move the flow cell lever to position 1. This engages the vacuum and secures the flow cell into position. When the flow cell lever is green, the vacuum is engaged.

Figure 44 Flow Cell Lever in Position 1



- 4 Wait for about five seconds, and then slowly move the flow cell lever to position 2. When the flow cell lever is solid green, the manifolds are in position and the flow cell is ready for use.

Figure 45 Flow Cell Lever in Position 2



- 5 Make sure that the **Vacuum Engaged** checkbox is selected on the load sequencing flow cell screen.

Confirm Proper Flow

After loading the flow cell, check for proper flow using the fluidics check screen. Checking for proper flow confirms that the flow cell and gaskets are properly installed and the manifold is engaged.

- 1 Select solution 5 (SB2) from the drop-down list.
- 2 Enter the following default values:
 - Volume: **25**
 - Aspirate Rate: **250**
 - Dispense Rate: **2000**
- 3 Select **Pump**.
- 4 Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.
If you see excessive bubbles, check the manifold gaskets for obstructions, and repeat the process.
 - a Select solution 6 (SB2) to avoid depleting SB2 from position 5.
 - b Reduce the aspirate rate to 100, and pump another 250 μ l of SB2 to the flow cell.
- 5 After you have confirmed proper flow, select **Next** to proceed. Make sure that the flow cell lever is green, and then close the flow cell compartment door.
- 6 Confirm that the **Vacuum Engaged** and **Door Closed** checkboxes are selected, and then select **Next**.
- 7 Select **Start** to start the sequencing run.

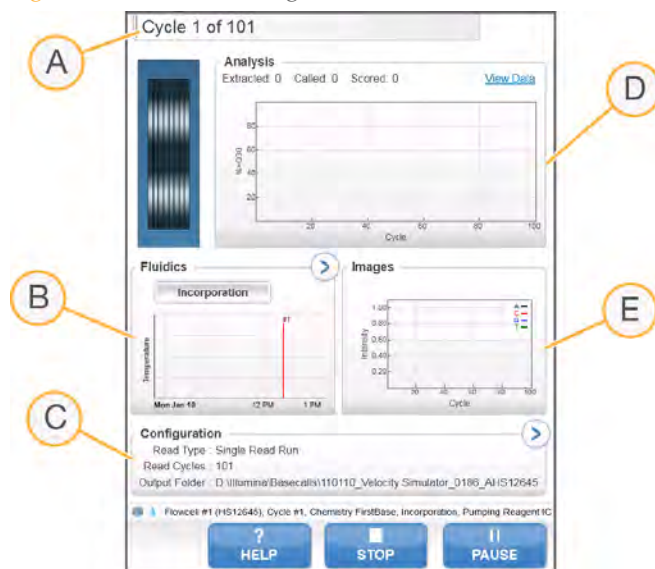
Monitoring the Run

Use the run overview screen to monitor on-instrument analysis, fluidics, and imaging.

- ▶ **Progress Bar**—Use the progress bar to monitor how many cycles have been completed.
- ▶ **Flow Cell Image**—Use the flow cell image to monitor which lanes have been imaged for that cycle.
- ▶ **Analysis Graph**—Use the analysis graph to monitor quality scores by cycle.
- ▶ **Images Graph**—Use the images graph to monitor intensities by cycle.
- ▶ **Fluidics Graph**—Use the arrow button to expand the fluidics section and monitor chemistry steps in the protocol.

In Rapid Run mode, HCS will automatically prompt you to perform a water wash when sequencing is finished. See *Performing a Water Wash* on page 169.

Figure 46 Run Monitoring Screen Overview



- A Run Progress
- B Fluidics and Temperature
- C Run Configuration
- D Analysis

E Images



NOTE

You might notice an intermittent clicking sound coming from the right side of the instrument during operation. This is normal. The clicking sound indicates that the refrigeration condensation pump is working.

First Base Report

The first base report confirmation dialog box opens automatically after cycle one is complete, if you selected **Confirm First Base** on the scan screen during run setup. The run will pause at this step.

If you did not select the checkbox, the run continues after first base incorporation without prompting you to review the first base report. For more information, see *Advanced* on page 42.

- 1 Review the first base report from the first base report confirmation dialog box or by opening `First_Base_Report.htm` from the root level of the run folder.
- 2 If the results are satisfactory, select **Continue** to continue your run. If results are not satisfactory, abort the run.

Sequencing Analysis Viewer

You can monitor your run using the Sequencing Analysis Viewer, which launches automatically after image analysis begins. For more information, see *Sequencing Analysis Viewer* on page 178.

For a detailed description of available metrics and installation instructions, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

Preparing Reagents for Read 2

Prior to the completion of Read 1 and any Index Reads, prepare reagents for Read 2 resynthesis and fresh ICB for Read 2. Reagents take less than 30 minutes to prepare.



NOTE

For optimal performance, Illumina recommends that you prepare fresh ICB (Incorporation Mix) for Read 2.

Prepare Paired-End Reagents

Paired-end sequencing requires paired-end reagents provided in the TruSeq Cluster Kit. Dual-indexed Nextera libraries require HP11 (Read 2 sequencing primer) provided in the TruSeq Dual Index Sequencing Primer Box.

Library Type	Read 2
Nextera libraries	HP11
TruSeq HT libraries	HP7 or HP11
TruSeq LT/v2 libraries	HP7 or HP11
TruSeq Small RNA libraries	HP7 or HP11

Paired-end reagents take less than 30 minutes to prepare.



NOTE

HP7, provided in the TruSeq PE Cluster Kit, is not compatible with dual-indexed Nextera libraries. You must use **HP11** from the TruSeq Dual Index Sequencing Primer Box (PE).

Table 10 Paired-End Reagents

Reagents	Kit Name
RMX	TruSeq PE Cluster Kit (PE Reagents Box) Note, You do not need to prepare RMX for dual-indexed paired-end runs. RMX is used during the Index 2 (i5) Read of a dual-indexed paired-end run and is not used again during Read 2 resynthesis. In non-indexed paired-end runs, RMX is required and used during the paired-end turn.
LMX2	TruSeq PE Cluster Kit (PE Reagents Box)
BMX	TruSeq PE Cluster Kit (PE Reagents Box)

Reagents	Kit Name
AMX2	TruSeq PE Cluster Kit (PE Reagents Box)
APM2	TruSeq PE Cluster Kit (PE Reagents Box)
AT2	TruSeq PE Cluster Kit (PE Reagents Box)
HP7 or HP11	TruSeq PE Cluster Kit (PE Reagents Box)
HP11	TruSeq Dual Index Sequencing Primer Box
HP3	TruSeq PE Cluster Kit (PE Reagents Box)
HT2	TruSeq PE Cluster Kit (PE Reagents Box)



WARNING

This set of reagents contains formamide, an aliphatic amide that is a probable reproductive toxin. Personal injury can occur through inhalation, ingestion, skin contact, and eye contact.

Dispose of containers and any unused contents in accordance with the governmental safety standards for your region.

For more information, see the MSDS for this kit on www.illumina.com.

Thaw Reagents

- 1 Remove the reagents from -15° to -25°C storage. Confirm that you are using the appropriate sequencing primer for your library type and run type.
- 2 Thaw reagents in a beaker filled with room temperature deionized water for about 20 minutes, or until reagents are fully thawed.
- 3 After reagents have completely thawed, place LMX2, BMX, AMX2, and RMX on ice.

Prepare RMX

- 1 Invert the tube of RMX five times to mix the reagent.
- 2 Label the tube of RMX **Reagent #10**.
- 3 Briefly centrifuge the reagent at 1,000 rpm for one minute or less. Alternatively, gently tap or flick the tube to make sure that reagent droplets collect at the bottom of the tube.



NOTE

Do not vortex RMX.

- 4 Set aside on ice.

Prepare LMX2

- 1 Invert the tube of LMX2 five times to mix the reagent.
- 2 Label the tube of LMX2 **Reagent #11**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside on ice.

Prepare BMX

- 1 Invert the tube of BMX five times to mix the reagent.
- 2 Label the tube of BMX **Reagent #12**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside on ice.

Prepare AMX2

- 1 Invert the tube of AMX2 five times to mix the reagent.
- 2 Label the tube of AMX2 **Reagent #13**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside on ice.

Prepare APM2

- 1 Invert the tube of APM2 five times to mix the reagent.
- 2 Label the tube of APM2 **Reagent #14**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare AT2

- 1 Invert the tube of AT2 five times to mix the reagent.
- 2 Label the tube of AT2 **Reagent #15**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare HP7

- 1 Invert the tube of HP7 five times to mix the reagent.
- 2 Label the tube of HP7 **Reagent #16**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare HP11

You must use HP11 for sequencing dual-indexed Nextera libraries.

- 1 Invert the tube of HP11 five times to mix the reagent.
- 2 Label the tube of HP11 **Reagent #16**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare HP3

- 1 Invert the tube of HP3 five times to mix the reagent, and then pulse centrifuge.
- 2 Transfer 2.85 ml of PW1 into a 15 ml Sarstedt conical tube and add 150 μ l of HP3.
- 3 Invert the tube five times to mix the reagent.
- 4 Label the conical tube of diluted HP3 **Reagent #18**.
- 5 Centrifuge at 1,000 rpm for one minute.
- 6 Set aside at room temperature.

Prepare HT2

- 1 Invert the tube of HT2 five times to mix the reagent.
- 2 Label the tube of HT2 **Reagent #19**.
- 3 Centrifuge the reagent at 1,000 rpm for one minute.
- 4 Set aside at room temperature.

Prepare ICB for Read 2 (200 Cycle Kit)

Use the following instructions for non-indexed or single-indexed runs. For dual indexed runs, prepare ICB using the volumes that you set aside for Read 2. For more information, see *Prepare ICB for Dual-Indexed Runs* on page 82.

- 1 Add the contents of two tubes of LFN to the bottle containing 47 ml of ICB-200.
- 2 Rinse each tube of LFN with ICB-200 to make sure that all LFN is transferred.
- 3 Aliquot 1.1 ml of EDP-200 into the ICB/LFN mix.
- 4 Return any unused portion of EDP-200 to -15° to -25°C storage.
- 5 Cap the bottle and invert several times to mix.
- 6 Set aside on ice until you are ready to load reagents onto the instrument.

Loading Reagents for Read 2

After completion of Read 1 and any Index Reads, the software prompts you to load paired-end reagents for Read 2 resynthesis and freshly-prepared ICB for Read 2.

Load Paired-End Reagents

- 1 Record the weight of each reagent on the lab tracking form.
- 2 Raise the sippers for the paired-end reagent rack using the following motion:
 - a Pull the handle towards you.
 - b Push the handle up while pulling it towards you.
 - c Release the handle into the slot on the top end of the groove. You should feel the handle rest securely into the slot.
- 3 Slide the reagent rack out of the reagent compartment using the rack handle.
- 4 Place each of the reagent tubes onto the rack in the associated numbered positions depending on the type of run you are performing.

Table 11 Paired-End Reagent Positions

Position	Reagent	Description
10	RMX	Resynthesis Mix
11	LMX2	Linearization Mix 2
12	BMX	Blocking Mix
13	AMX2	Amplification Mix 2
14	APM2	AMX2 Premix
15	AT2	100% Formamide
16	HP7 or HP11	Read 2 Sequencing Primer
18	HP3	Denaturation Solution
19	HT2	Wash Buffer



NOTE

You do not need to load RMX for dual-indexed paired-end runs. RMX is used during the Index 2 (i5) Read and is not used again during Read 2 resynthesis.

- 5 Remove the caps from each reagent tube.

- 6 Slide the reagent rack into the reagent compartment, aligning the rack with the raised guide on the floor of the compartment.
- 7 Lower the sippers into the paired-end reagent tubes as follows:
 - a Pull the handle towards you.
 - b Lower the handle while pulling it towards you.
 - c Visually inspect the sippers to make sure that they do not bend as they lower into the tubes.
 - d Release the handle into the slot on the bottom end of the groove. You should feel the handle rest securely into the slot.

Load ICB for Read 2

- 1 Record the weight of the reagent on the lab tracking form.
- 2 Raise the sippers for the sequencing reagent rack using the following motion:
 - a Pull the sipper handle towards you.
 - b Push the sipper handle up while pulling it towards you.
 - c Release the sipper handle into the slot on top end of the groove. You should feel the sipper handle rest securely into the slot.
- 3 Slide the reagent rack out of the reagent compartment.
- 4 Remove the existing ICB reagent bottle from position 1 of the reagent rack and remove the funnel cap from the bottle.
- 5 Place the funnel cap on the new bottle of ICB and load the bottle in position 1, making sure that the conical end of the bottle rests in the indentation on the base of the rack.
- 6 Slide the reagent rack into the reagent compartment, aligning the rack with the raised guide on the floor of the compartment.
- 7 Lower the sippers into the sequencing reagent bottles as follows:
 - a Pull the sipper handle towards you.
 - b Lower the sipper handle while pulling it towards you.
 - c Visually inspect the sippers to ensure that they do not bend as they lower into the funnel caps.

- d Release the sipper handle into the slot on the bottom end of the groove. You should feel the sipper handle rest securely into the slot.
- 8 Close the reagent compartment door, and then select **Next** to resume the run.

[This page intentionally left blank]

Performing a Rapid Sequencing Run

Introduction	122
Rapid Sequencing Workflow	124
Preparing SBS and Cluster Reagents	126
Preparing Indexing Reagents	134
Checking Pre-Run Fluidics and Run Parameters	135
Loading and Priming Reagents	140
Loading a Flow Cell	151
Monitoring the Run	158



Introduction

To perform a Rapid sequencing run on the HiSeq 2500, you first prepare reagents for your run and then set up the run using the HCS interface. Run setup steps include: entering run parameters, loading reagents, loading the Rapid Run flow cell, and performing a fluidics check. If you have preformed template hybridization on the cBot, you will also need to prime reagents.

When loading the same low-diversity sample into both Rapid flow cell lanes, take into account that you will not have a balanced control lane for matrix and phasing estimations. This can result in poor quality data. Illumina recommends spiking in PhiX at 30% to 50% to increase base diversity. Alternatively, perform template hybridization on the cBot with a balanced sample in one lane and set this as the control lane on the HiSeq.

For Rapid Run mode, all SBS, clustering, PE and indexing reagents are loaded at the start of the run.

Types of Rapid Runs

You can perform six types of Rapid runs on the HiSeq 2500:

Run Type	Read 1 Cycles	Index 1 (i7) Read Cycles	Index 2 (i5) Read Cycles	Read 2 Cycles	Total Cycles
Single-Read, Non-Indexed	≤ 151	--	--	--	≤ 151
Single-Read, Single-Indexed	≤ 151	7 * 8 **	--	--	≤ 158 * ≤ 159 **
Single-Read, Dual-Indexed	≤ 151	8	8	--	≤ 167
Paired-End, Non-Indexed	≤ 151	--	--	≤ 151	≤ 302
Paired-End, Single-Indexed	≤ 151	7 * 8 **	--	≤ 151	≤ 309 * ≤ 310 **
Paired-End, Dual-Indexed	≤ 151	8	7 + 8 ***	≤ 151	≤ 325

* Number of cycles for single-indexed libraries

** Number of cycles for dual-indexed libraries

*** Index 2 Read of a paired-end dual-indexed run includes seven additional chemistry-only cycles

**NOTE**

If you are running more than 100 cycles per read, you will need to combine multiple kits. See *SBS Kits for Dual-Indexed Sequencing* on page 126 and *SBS Kits for Non-Indexed Sequencing* on page 128 for details.

For more information, see *Single-Indexed Sequencing Overview* on page 13 and *Dual-Indexed Sequencing Overview* on page 16.

Rapid Sequencing Workflow



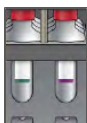
Prepare reagents and library sample. This includes thawing and inverting all reagents and diluting your sample to the relevant concentration.

Note that your workflow will vary depending upon whether you perform template hybridization on the cBot as opposed to on the HiSeq. You will find workflow differences noted in the following sections:

- *Introduction* on page 122
- *Priming Reagents* on page 145
- *Preparing SBS and Cluster Reagents* on page 126
- *Denature Library for On-Instrument Clustering* on page 131
- *Load SBS and Clustering Reagents and Template* on page 140
- *Loading a Flow Cell* on page 151
- *Confirm Proper Flow* on page 157



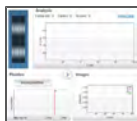
Using the HiSeq Control Software (HCS), perform a pre-run fluidics check and enter parameters for your run.



If you are performing on-board cluster generation, you will load clustering, SBS, PE turn, indexing reagents, and the library sample at start of run. If you have performed template hybridization on the cBot, you will need to load the reagents at the start of the run, and the library sample will not be needed.



If you are performing a template hybridization on the cBot, clustering reagents on the HiSeq will need to be primed through a used flow cell. If you are performing on-board cluster generation, all clustering and sequencing reagents will be primed through the new flow cell once the run has started.



Start your sequencing run. After cycle 1, inspect the first base report (optional setting), and then continue Read 1.

The sequencing run will continue through the PE turn and read 2 with no intervention necessary.



When the run is complete, unload and weigh reagents. You will be prompted to carry out a post run water wash. After the post run water wash, the instrument is ready to use. If a maintenance wash is required, the software will prompt you to perform one. This occurs every 10 days or when switching modes.

Approximate Run Duration

The following table lists the approximate duration of a Rapid sequencing run based on read length and version of instrument software.

Run Format	HCS v2.0 in Rapid Run Mode on instruments with a serial number greater than 700895*
36	Approx. 7 hours
2 x 50	Approx. 16 hours
2 x 100	Approx. 27 hours
2 x 150 Cycles	Approx. 40 hours

*Rapid Runs will take longer on instruments with a serial number less than 700895.

Preparing SBS and Cluster Reagents



NOTE

Regardless of whether you are performing on-board cluster generation or cBot template hybridization, the same clustering and SBS reagents will be prepared as described in this section.

If you are performing template hybridization on the cBot, you can skip the section titled, Denature Library for On-Instrument Clustering.

When running in Rapid mode, the HiSeq draws all sequencing primers from a single tube; consequently, only a single tube of sequencing primer can be used in each primer hybridization step on the HiSeq in Rapid Run mode. Therefore the same sequencing primer must be used for both lanes of a Rapid flow cell.

The first step in setting up a run is to prepare SBS reagents. Reagent preparation requires overnight thawing at 2° to 8°C, or about 90 minutes in a room temperature deionized water bath. When completely thawed, reagents are ready for use. For Rapid Run mode, the Incorporation Mix does not require preparation apart from other reagents.

SBS Kits for Dual-Indexed Sequencing

One 200-cycle SBS Kit provides sufficient reagents for up to 225 cycles of sequencing, and one 50-cycle SBS Kit provides sufficient reagents for up to 75 cycles of sequencing. For runs greater than a total of 225 cycles, and to maximize reagent use for other read lengths, you can use multiple 50-cycle kits. For a 2x151-cycle run Illumina recommends that you use either 1.5x200-cycle kits or 1x200-cycle + 2x50-cycle kits

The following table lists the number of kits required for dual-indexed runs. For dual-indexed paired-end runs, seven chemistry-only cycles are calculated in the total number of cycles.

Table 12 Kit Combinations for Indexed Runs

Run Type	Run Cycles	Total Cycles	Kit Combination 1		Kit Combination 2		Kit Combination 3		Kit Combination 4	
			200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits
Dual-Indexed Paired-End	151+8+(7)+8+151	325	2	-	1.5	-	1	2	-	4
	101+8+(7)+8+101	225	1	-	-	3	-	-	-	-
	93+8+(7)+8+93	209	1	-	-	3	-	-	-	-
	76+8+(7)+8+76	175	1	-	-	2	-	-	-	-
	51+8+(7)+8+51	125	-	2	-	-	-	-	-	-
	36+8+(7)+8+36	95	-	2	-	-	-	-	-	-
Dual-Indexed Single-Read	151+8+8	167	-	2	-	-	-	-	-	-
	101+8+8	117	-	2	-	-	-	-	-	-
	93+8+8	109	-	2	-	-	-	-	-	-
	76+8+8	92	-	2	-	-	-	-	-	-
	51+8+8	67	-	1	-	-	-	-	-	-
	36+8+8	52	-	1	-	-	-	-	-	-


NOTE

If you are combining kits for a 2x151 run (with or without indexing), fill each USB bottle to 200 ml. No replenishing is required. The remaining USB can be stored, tightly capped at 2° to 8°C, for up to 7 days.

Discard any other remaining SBS reagents once the run is complete; do not store for future use.

For more information, see *Dual-Indexed Sequencing Overview* on page 16.

SBS Kits for Non-Indexed Sequencing

One 200-cycle SBS Kit provides sufficient reagents for up to 225 cycles of sequencing, and one 50-cycle SBS Kit provides sufficient reagents for up to 75 cycles of sequencing. For runs greater than a total of 225 cycles, and to maximize reagent use for other read lengths, you can use multiple 50-cycle kits. For a 2x151-cycle run Illumina recommends that you use either 1.5x200-cycle kits or 1x200-cycle + 2x50-cycle kits

The following table lists the number of kits required for non-indexed runs.

Table 13 Kit Combinations for Non-Indexed Runs

Run Type	Run Cycles	Total Cycles	Kit Combination 1		Kit Combination 2		Kit Combination 3		Kit Combination 4	
			200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits	200-cycle kits	50-cycle kits
Paired-End	151+151	302	2	-	1.5	-	1	1	-	4
	101+101	202	1	-	-	3	-	-	-	-
	93+93	186	1	-	-	3	-	-	-	-
	76+76	152	1	-	-	2	-	-	-	-
	51+51	102	-	2	-	-	-	-	-	-
	36+36	72	-	1	-	-	-	-	-	-
Single-Read	151	151	1	-	-	2	-	-	-	-
	101	101	-	2	-	-	-	-	-	-
	93	93	-	2	-	-	-	-	-	-
	76	76	-	2	-	-	-	-	-	-
	51	51	-	1	-	-	-	-	-	-
	36	36	-	1	-	-	-	-	-	-



NOTE

If you are combining kits for a 2x151 run (with or without indexing), fill each USB bottle to 200 ml. No replenishing is required. The remaining USB can be

stored, tightly capped at 2° to 8°C, for up to 7 days.
 Discard any other remaining SBS reagents once the run is complete; do not store for future use.

Thaw SBS Reagents



NOTE

Be sure to thaw CRM separately from the other SBS reagents.

- 1 Remove IMM, SRM, CRM, and USB from -15° to -25°C storage, thaw, and invert according to the instructions in the following table:



NOTE

If USB has been stored at 2° to 8°C, it can be used directly from storage after being inverted 5 times to mix. Leave at room temperature until ready to use.

Reagent Name	SBS Kit (200 Cycles)	SBS Kit (50 Cycles)
IMM	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for 90 minutes. Invert 5 times to mix. Leave on ice until ready to use.	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for one hour. Invert 5 times to mix. Leave on ice until ready to use.
SRM	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for 90 minutes. Invert 5 times to mix. Leave on ice until ready to use.	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for one hour. Invert 5 times to mix. Leave on ice until ready to use.
CRM	Thaw at 2° to 8°C for 16 hours or in room temperature water bath for 90 minutes. Invert 5 times to mix. Leave on ice, away from IMM and SRM until ready to use.	Thaw at 2° to 8°C for 16 hours or in a <i>separate</i> room temperature water bath for one hour. Invert 5 times to mix. Leave on ice, away from IMM and SRM until ready to use.
USB (two bottles per flow cell)	Thaw at 2° to 8°C for 16 hours or in room temperature water bath for 90 minutes. Invert 5 times to mix. Leave a room temperature until ready to use.	Thaw at 2° to 8°C for 16 hours or in a room temperature water bath for one hour. Invert 5 times to mix. Leave a room temperature until ready to use.

- 2 Ensure that reagents are completely thawed before use.
- 3 Record the lot numbers of each reagent on the lab tracking form.

Thaw Clustering Reagents

- 1 Remove the following reagents from -15° to -25°C storage, thaw, and invert according to the instructions in the table:

Reagent	TruSeq Rapid PE Cluster Kit	TruSeq Rapid SR Cluster Kit
FRM	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.	--
FLM2	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.	--
FLM1	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.	--
FLS	--	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.
AMS	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.	Thaw in room temperature water bath for one hour. Invert five times to mix. Store on ice until needed.
FPM	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.
FDR	Thaw in room temperature water bath for one hour. Invert five times to mix.	Thaw in room temperature water bath for one hour. Invert five times to mix.

Reagent	TruSeq Rapid PE Cluster Kit	TruSeq Rapid SR Cluster Kit
	Store at room temperature until needed.	Store at room temperature until needed.
HP9	--	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.
HP10	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.
HP11	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.	--
HP12	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.	Thaw in room temperature water bath for one hour. Invert five times to mix. Store at room temperature until needed.

- 2 Ensure that reagents are completely thawed before use.

Denature Library for On-Instrument Clustering



NOTE

If you are performing the template hybridization on the cBot, you can skip this section.

This step is only for performing the template hybridization step on the HiSeq 2500.

Perform the following steps to denature the template DNA with 0.1 N NaOH to a final DNA concentration of 20 pM.



NOTE

If your application requires higher than a 20 pM final concentration of your library, make sure that your concentration of NaOH is not higher than 0.05 N in the denaturation solution and not more than 0.001 N (1 mM) in the final solution after diluting with HT1.

Higher concentrations of NaOH in the loaded library will inhibit library hybridization and decrease cluster density.

Illumina Supplied Consumables

- ▶ HT1

User Supplied Consumables

- ▶ 1N NaOH
- ▶ Microcentrifuge tube. See *Load SBS and Clustering Reagents and Template* on page 140.

Procedure

- 1 Combine the following volumes of template DNA and 0.1 N NaOH in a microcentrifuge tube:
 - 2 nM template DNA (10 μ l)
 - 0.1 N NaOH (10 μ l)
- 2 Vortex briefly to mix the template solution.
- 3 Centrifuge the template solution to 280 xg for one minute.
- 4 Incubate for five minutes at room temperature to denature the template into single strands.
- 5 Transfer denatured template (Temp) to a tube containing 980 μ l pre-chilled HT1 (Hybridization buffer).

Diluting Denatured DNA

Perform the following steps to dilute the denatured DNA with pre-chilled HT1.

- 1 Dilute the denatured DNA in a 1.5 or 1.7 ml microcentrifuge tube, to the desired concentration in a total volume of 420 μ l, according to the following examples:

Required Concentration	20 pM template (μ l)	HT1 (μ l)
2 pM	42	378

Required Concentration	20 pM template (μl)	HT1 (μl)
6 pM	126	294
10 pM	210	210
12 pM	252	168
16 pM	336	84
20 pM	420	0

- 2 Invert several times to mix, and then pulse centrifuge the template solution.
- 3 Place the denatured template DNA on ice until you are ready to load it onto the instrument. See *Load SBS and Clustering Reagents and Template* on page 140 for template loading details.

Preparing Indexing Reagents

A separate indexing kit is not necessary for the Rapid Run mode. All rapid run indexing reagents are supplied in the clustering PE and SR kits.

Thaw Indexing Reagents

- 1 Thaw indexing reagents (HP9 - i5 indexing primer, HP12 - i7 indexing primer, FRM) at 2° to 8°C for 16 hours or in a room temperature water bath for one hour.
- 2 Invert 5 times to mix.
- 3 Pulse centrifuge to collect droplets.
- 4 Keep reagents on ice until they are ready to be used.
- 5 At the start of the run, load onto the instrument at the same time as all the other SBS and clustering reagents from the kit.

Checking Pre-Run Fluidics and Run Parameters

The HiSeq Control Software (HCS) guides you through the run setup steps. Ensure that you are in Rapid Run mode.



CAUTION

Place waste lines 1, 2, 3, 6, 7, and 8 for the appropriate position in a 1 liter container filled with deionized water. If these lines are not placed in deionized water, damage to the reagent pumps can occur.

From the HCS Welcome screen, select **Sequence | New Run**. The Pre Run Fluidics Check screen opens. Illumina recommends that you perform the fluidics check to ensure proper functioning of your instrument before starting a run.

Volume Check Screen

For Rapid Run mode only, the Volume Check screen enables you to perform a pre-run fluidics check for volume.

If you select Yes to perform the fluidics check at the start of the run, HiSeq Control Software (HCS) will prompt you to load water in all positions, including the sample loading station, and load a used flow cell. After fluidics are checked the through the used flow cell, the water wash will start (and will take ~12 minutes). When the check is complete, the software will advance to the next screen.

If you select No to skip the fluidics check, the software will advance directly to the next screen.

- 1 Load laboratory grade water onto all 8 SBS positions, 10 Clustering positions, and the sample position for the relevant flow cell. Ensure that a used Rapid Run flow cell is loaded on the instrument, and carry out a flow check.



NOTE

Before you perform a fluidics check, place lines 4 and 5 into empty 50 ml tubes. The expected delivery volume per lane per flow cell is 9.5 ml.

Storage Screen

- 1 Select **Use BaseSpace for storage and analysis** if you want to save your data to a BaseSpace account. If you select this option you will be prompted to enter a My Illumina account email and password.



NOTE

For every run, your HiSeq 2500 will attempt to send general run metrics to BaseSpace. These metrics include interoperation files (that SAV uses) and global files such as RunInfo.xml and RunParameters.xml. No data files are ever sent, and all information is sent anonymously, associated with the HiSeq 2500 from which it came, not to any particular user. This feature is enabled by default and Illumina recommends that you leave it enabled. To disable this feature, select **Menu | Tools | Options**, and then unselect **Send system health information to Illumina to aid technical support**. You can also view terms in the Options dialog box.

- 2 In **Save to an output folder**, browse to and select the path of your run folder.
- 3 Select **Zip BCL files** to reduce the footprint of disc space needed for the run. If you are using BaseSpace, this option is selected by default.



NOTE

If you are using BaseSpace or an updated version of CASAVA for data analysis, zipping BCL files is mandatory.

- 4 Select **Bin QScores** to reduce the footprint of disc space needed for the run.
- 5 Select a setting from the **Save Auxiliary Files** drop-down list:
 - Select **Save All Thumbnails** to save tile level and full thumbnails.
 - Select **Save Tile Thumbnails** to save tile level thumbnails only.
- 6 Select **Next**.

Flow Cell Set-up Screen



NOTE

It is critical that you scan the flow cell barcode or accurately enter the flow cell ID when you set up your run. The software uses the flow cell ID to determine flow cell type and reagent compatibility.

- 1 Scan the flow cell barcode or enter the flow cell ID (barcode number) of the flow cell to be sequenced.
- 2 Confirm the flow cell type, which is automatically selected based on the flow cell ID.
- 3 Enter the experiment name and user name.
- 4 Select the lane containing the control, if applicable. Otherwise, select **None**.

- 5 Select **Next**.

Advanced Screen

- 1 Select the **Confirm First Base** checkbox to generate a first base report confirmation.
- 2 Select a setting from the **Control Lane** drop-down list.
- 3 Select or deselect **Align to PhiX** for each lane of the flow cell.
- 4 Select **Keep Intensity Files** for later reanalysis or custom processing (optional).
- 5 Select **Next**.

Recipe Screen

- 1 Select one of the following indexing options:
 - **No Index**—Performs a non-indexed single-read or paired-end run.
 - **Single Index**—Performs a single-read or paired-end run with one Index Read (Index 1 (i7) Read).
 - **Dual Index**—Performs a single-read or paired-end run with two Index Reads (Index 1 (i7) Read and Index 2 (i5) Read).
 - **Custom**—Performs a single-read or paired-end run with a custom number of cycles for the Index Reads.
- 2 Select a flow cell format, either **Single Read** or **Paired End**.
- 3 Enter the number of cycles you plan to sequence for Read 1 and Read 2, if applicable.
If you selected the **Custom** indexing option, enter the number of cycles you plan to sequence for the Index Reads.



NOTE

If you choose to run a custom recipe, the number of cycles is determined by that recipe, and the number of cycles option in the software will be grayed out when the recipe is loaded. You can enter up to 325 cycles.



NOTE

It is possible to perform a dual-indexed single-read run using a paired-end flow cell containing paired-end libraries. To do so, set up a *dual-indexed paired-end run* and enter zero for the number of cycles in Read 2. This will stop the run after the Index 2 (i5) Read.

- 4 Select the **Use Existing Recipe** checkbox only if you want to use a custom recipe. Otherwise, allow the software to create your recipe based on information you provide on the run setup screens.

Sample Sheet Screen



NOTE

To view indexing results in Sequencing Analysis Viewer (SAV) or to use BaseSpace, it is essential that you load a sample sheet at the start of your run. SAV v1.8 requires a sample sheet to properly report indexing information. For more information, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

- 1 Enter the path or browse to a valid sample sheet for your run. For indexed runs, the sample sheet is required if you want indices to be reported in SAV. A sample sheet is also required if you are using BaseSpace.
- 2 If you want to perform on-board cluster generation on the HiSeq, select **On-Board Cluster Generation**. Or, if you have already performed template hybridization on the cBot, select **Template Hybridization on cBot**.
- 3 Select **Next**.

Reagents Screen

- 1 Scan or enter the Reagent Kit ID. You can enter the kit ID from either box 1 or box 2.
- 2 For indexed runs, scan or enter the Indexing Reagent Kit ID.
- 3 Select the SBS reagent kit you are using for your run.
 - Select **200 Cycles** for a 200 cycle kit; select **50 Cycles** for a 50 cycle kit.



NOTE

If you are combining four 50-cycle SBS kits for a dual-indexed run, select **Used Kit** instead of 50 Cycles to correctly set the cycle counter.

- 4 For paired-end runs, scan or enter the PE or SR Reagent Kit ID.
- 5 The Cycles Remaining field is populated automatically depending on the SBS kit you are using. If you are using a partial kit, enter the number of cycles reagents are expected to last. The software counts down the number of cycles entered. When the cycles are low, the software prompts you to load fresh reagents. For more information, see *Changing Reagents During a Run* on page 63.

**NOTE**

If you are starting a new run using less than a full reagent kit, make sure that you can be present when new reagents are needed.

- 6 If you selected Template Hybridization on cBot in the Sample Sheet window, select **Prime SBS Reagents** to prime reagents before you start a run. If you are performing on-board cluster generation, do not prime SBS Reagents.
- 7 Select **Next**.

Review Screen

Review your run information. If it is correct, select **Next**.

Select **Back** if you need to change any entries.

Loading and Priming Reagents

The next step is to load sequencing reagents. The software interface guides you through the process.

Load SBS and Clustering Reagents and Template

Make sure that SBS reagents are thoroughly thawed and ready to load onto the instrument.

Illumina-Supplied Consumables

- ▶ Eight funnel caps (Provided in the SBS Kit)

User-Supplied Consumables

- ▶ One 250 ml bottle (Corning, catalog # 430776)
- ▶ One flip-top eppendorf tube per flow cell (1.5 ml or 1.7 ml). The following eppendorfs are acceptable for use with the HiSeq 2500. Do not use tubes with screw caps.

Consumable	Supplier	Purpose
MicroCentrifuge Tube, 1.5mL	VWR, catalog #20170-038	Sample(template) loading and washing the template fluidics line.
MicroCentrifuge Tube, 1.7mL, capless	VWR, catalog #20170-575	
MicroCentrifuge Tube, 1.5mL	VWR, catalog #20170-650	
MicroCentrifuge Tube, 1.5mL	VWR, catalog #89000-028	
MicroCentrifuge Tube, 1.5mL	AXYGEN, catalog #MCT-150-C	
MicroCentrifuge Tube, 1.7mL	AXYGEN, catalog #MCT-175-C	
MicroCentrifuge Tube, 1.7mL	Sorenson BioScience, catalog #16070	

Procedure

- 1 Record the weight of each reagent on the lab tracking form.

**NOTE**

Weighing reagents before and after a sequencing run confirms proper reagent delivery.

- 2 Open the reagent compartment door.
- 3 Raise the sippers for the sequencing reagent rack using the following motion:
 - a Pull the sipper handle towards you.
 - b Push the sipper handle up while pulling it towards you.
 - c Release the sipper handle into the slot on the top end of the groove. You should feel the sipper handle rest securely into the slot.
- 4 Slide the reagent rack out of the reagent compartment.
- 5 Place each reagent bottle onto the rack in the associated numbered position as shown on the Reagents screen, making sure that the conical end of the bottle rests in the indentation on the base of the rack.

**NOTE**

Load laboratory-grade water into positions 2, 4, 8, 10, 11, and 12 for single read flow cell and positions 2, 4, 8, and 19 for a paired end flow cell to prevent the lines from drying out.

Table 14 Single Read Flow Cell Positions

Position	Reagent	Description
1	IMM	Incorporation Master Mix
2	PW1 (25ml)	Wash Buffer
3	SRM	Scan Reagent Master Mix
4	PW1 (25ml)	Wash Buffer
5	USB	Universal Sequencing Buffer
6	USB	Universal Sequencing Buffer
7	CRM	Cleavage Reagent Master Mix
8	PW1 (25 ml)	Wash Buffer
10	PW1 (10 ml)	Wash Buffer
11	PW1 (10 ml)	Wash Buffer

Position	Reagent	Description
12	PW1 (10 ml)	Wash Buffer
13	AMS	Fast Amplification Mix
14	FPM	Fast Premix
15	FDR	Fast Denaturation Reagent
16	HP9	i5 Index Primer
17	HP12	i7 Index Primer
18	HP10	Read 1 Primer
19	FLS	Fast Linearisation Solution
Sample Loading Station Template	Template	420 μ l in 1.5 or 1.7 ml microcentrifuge tube at required concentration

Table 15 Paired End Flow Cell Positions

Position	Reagent	Description
1	IMM	Incorporation Master Mix
2	PW1 (25 ml)	Wash Buffer
3	SRM	Scan Reagent Master Mix
4	PW1 (25 ml)	Wash Buffer
5	USB	Universal Sequencing Buffer
6	USB	Universal Sequencing Buffer
7	CRM	Cleavage Reagent Master Mix
8	PW1 (25 ml)	Wash Buffer
10	FRM	Fast Resynthesis Mix
11	FLM2	Fast Linearisation Mix 2 (read 2)
12	FLM1	Fast Linearisation Mix 1 (read 1)
13	AMS	Fast Amplification Mix
14	FPM	Fast Premix
15	FDR	Fast Denaturation Reagent
16	HP11	Read 2 Primer
17	HP12	i7 Index Primer
18	HP10	Read 1 Primer
19	PW1 (10 ml)	Wash Buffer
Sample Loading Station Template	Template	420 μ l in 1.5 or 1.7 ml microcentrifuge tube at required concentration

Loading the Template



NOTE

If you used the alternative, cBot template hybridization recipe, you will not need to load the template. For details about template hybridization on the cBot, see the *cBot User Guide*, Part # 15006165.

- 1 The software will prompt you to load the template at the same time as the other reagents. Load the template into the sample loading station. When loading the same low-diversity sample into both Rapid flow cell lanes, take into account that you will not have a balanced control lane for matrix and phasing estimations. This can result in poor quality data. Illumina recommends spiking in PhiX at 30% - 50% to increase base diversity. Alternatively, perform template hybridization on the cBot with a balanced sample in one lane and set it as the control lane on the HiSeq.

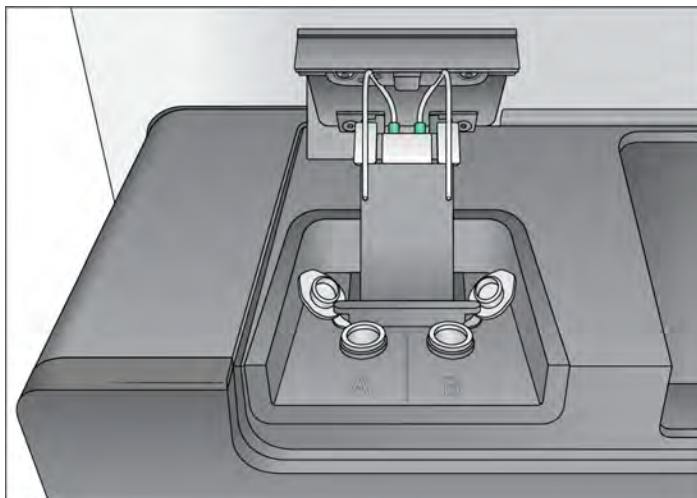


CAUTION

It is important to add no more than the required volume of template. When the sequencing run is finished successfully, the system will perform a flush-back wash adding an additional 500 μ l of water back into the template tube.

- a Ensure that the flow cell compartment door is open.
- b Lift the door of the sample loading station on the left side of the flow cell stage. As you lift the door, the sippers, which are attached to the door, also lift up.
- c Remove the eppendorf containing water from the appropriate positions (A and B) and replace it with the eppendorfs containing exactly 420 μ l of diluted library template per eppendorf. Holding the lids away from you, push them under the bar behind the tubes (to hold them out of the way of the sippers).

Figure 47 Sample Loading Station With the Door Open



NOTE

The liquid remaining in the tube following the run is highly diluted; it should be discarded immediately and not retained for further use.

- d Slowly close the sample loading station door; ensure that the sippers are properly aligned so that they enter the eppendorf tubes when the lid is closed. (Ensure that the sample loading station door is closed before closing the flow cell compartment door.)

Loading the SBS and Clustering Reagents

- 1 Add 25 ml of PW1 or laboratory-grade water to the bottles in positions 2, 4, and 8.



NOTE

HCS v2.0 and later includes an automatic post-run rinse following the completion of the sequencing run.

The post-run rinse does *not* replace the post-run instrument wash.

- 2 Add 10 ml of PW1 or laboratory-grade water to 15 ml conical tubes in positions 10, 11, and 12 for a single read flow cell or position 19 for a paired end flow cell.
- 3 Remove the cap from each reagent bottle and replace it with a funnel cap. Replace the cap on the bottle of CRM last, and then replace your gloves.

**CAUTION**

After handling the bottle of CRM, discard your gloves and replace them with a new pair.

- 4 Slide the reagent racks into the reagent compartment, aligning the racks with the raised guide on the floor of the compartment.
- 5 Lower the sippers into the sequencing reagent bottles as follows:
 - a Pull the sipper handle towards you.
 - b Lower the sipper handle while pulling it towards you.
 - c Visually inspect the sippers to ensure that they do not bend as they lower into the funnel caps.
 - d Release the sipper handle into the slot on the bottom end of the groove. You should feel the sipper handle rest securely into the slot.
- 6 Select the checkbox labeled **PW1 (25 ml) loaded in Position 2**.
- 7 Select **Next**.

Priming Reagents

**NOTE**

You will need to prime reagents only if you have done template hybridization on the cBot for a Rapid Run flow cell.

Perform the following steps to prime reagents:

- ▶ *Clean the Flow Cell Holder* on page 146
- ▶ *Load the Used Flow Cell* on page 98
- ▶ *Confirm Proper Flow* on page 100
- ▶ *Prime Reagents* on page 101

User-Supplied Consumables

- ▶ Lens cleaning tissue
- ▶ 70% ethanol or alcohol wipes
- ▶ Low-lint lab tissue
- ▶ Used flow cell

Clean the Flow Cell Holder

- 1 Open the flow cell compartment door.

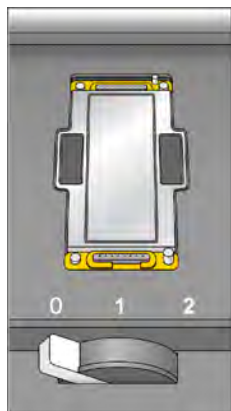


CAUTION

Do not place fluids on the flow cell compartment door or on the flow cell stage when the door is open. Spills in this area can damage the instrument.

- 2 Make sure that the flow cell lever is in the OFF position.

Figure 48 Flow Cell Lever in Position 0



- 3 Put on a new pair of powder-free latex gloves.
- 4 If the flow cell from a previous run is present, remove it and set aside. It can be used to confirm proper flow prior to loading the clustered flow cell.
- 5 Using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol, carefully wipe the surface of the flow cell holder until it is completely clean.

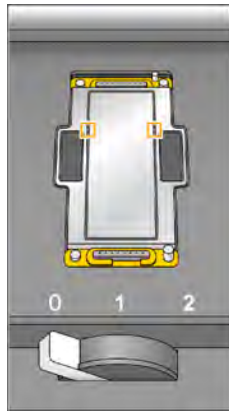


CAUTION

Do not allow alcohol to drip into the vacuum holes or around the manifolds. Use a low-lint lab tissue to dry the stage, if necessary.

- 6 Visually inspect the flow cell holder to make sure that it is free of lint and the vacuum holes are free of obstructions.

Figure 49 Vacuum Hole Locations



Load the Used Flow Cell

- 1 Rinse the used flow cell with laboratory-grade water. Dry it with lens cleaning tissue or lint-free tissue.
- 2 Clean the flow cell using alcohol wipes and lens cleaning tissue.



NOTE

Do not remove or replace the manifold gaskets at this time.

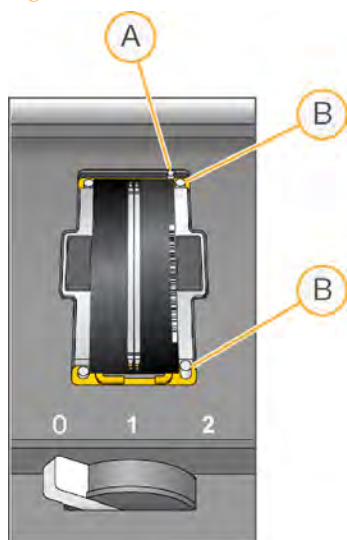
- 3 Place the used flow cell on the flow cell holder with the inlet and outlet ports facing *down* and the barcode on the right. The arrow on the left edge of the flow cell indicates flow direction, and should point towards the instrument.
- 4 Gently slide the flow cell towards the top and right guide pins until it stops.



NOTE

Remove your hand from the flow cell before engaging the vacuum switch. This prevents any possible alignment drift over time.

Figure 50 Flow Cell Positioned Against Top and Right Guide Pins.



- A Top Guide Pin
- B Right Guide Pins

- 5 Slowly move the flow cell lever to position 1. This engages the vacuum and secures the flow cell into position. When the flow cell lever is green, the vacuum is engaged.

Figure 51 Flow Cell Lever in Position 1



- Wait for about five seconds, and then slowly move the flow cell lever to position 2 (far-right). When the flow cell lever is solid green, the manifolds are in position and the flow cell is ready for use.

Figure 52 Flow Cell Lever in Position 2



- Make sure that the **Vacuum Engaged** checkbox is selected on the load prime flow cell screen, and then select **Next**.

Confirm Proper Flow

After the used flow cell is loaded, check for proper flow. Checking for proper flow confirms that the flow cell and gaskets are properly installed and the manifold is engaged.

- Select solution 2 (laboratory-grade water) from the drop-down list.



CAUTION

You can use water to confirm proper flow on a used flow cell only. Never use water to confirm proper flow on a clustered flow cell.

- Confirm the following default values:
 - Volume: **250**
 - Aspirate Rate: **1500**
 - Dispense Rate: **2000**
- Select **Pump**.

- 4 Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.
If you see excessive bubbles, check the gaskets for obstructions, reduce the aspirate rate to 1000, and pump another 250 μ l of water to the flow cell. If problems persist, remove the flow cell, repeat the cleaning steps, and reload the flow cell.

Prime Reagents

- 1 Loosen and remove waste lines 4 and 5 for the appropriate flow cell from the waste container. Make sure that you do not include the eight lines for the opposite flow cell or the lines for the condensation pump or the paired-end priming pumps.
- 2 Place lines 4 and 5 into two 15 ml tubes. Place the remaining lines (1, 2, 3, 6, 7 and 8) into a container with laboratory-grade water. Priming waste is collected and measured after the priming step.
- 3 Select **Next**. The priming screen opens and the priming step starts automatically. You can monitor the progress of the priming step from the priming screen.
- 4 When the priming step is complete, measure the collected priming waste and confirm that the volume is 2.5 ml +/- 10%, 500 μ l per reagent, per lane. Record the results on the lab tracking form.



NOTE

The color of priming waste might be brown in appearance. This is normal. The change in waste color is not toxic and does not impact run performance.

- 5 Return the waste tubing to the waste container before proceeding. Leave waste lines 1, 2, 3, 6, 7 and 8 in container with laboratory-grade water
- 6 Select **Next**. You are ready to load the hybridized flow cell.

Loading a Flow Cell



NOTE

If you have performed template hybridization on the cBot, you will be loading a hybridized flow cell. If you are performing template hybridization on the HiSeq, you will load a new flow cell.

The next step is to remove the used flow cell and load the flow cell that you want to sequence.

If you want to have individually addressable lanes on the 2-lane Rapid Run flow cell, you will use the cBot to carry out template hybridization and first extension; the rest of the clustering process is carried out on the HiSeq.

Perform the following steps to load the flow cell:

- ▶ *Remove the Used Flow Cell* on page 103
- ▶ *Clean the Flow Cell Holder* on page 152
- ▶ *Clean the Flow Cell* on page 153
- ▶ *Load the Flow Cell* on page 106
- ▶ *Confirm Proper Flow* on page 109

User-Supplied Consumables

- ▶ Lens cleaning tissue
- ▶ 70% ethanol or alcohol wipes
- ▶ Low-lint lab tissue



NOTE

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

Remove the Used Flow Cell

- 1 Open the flow cell compartment door.

Figure 53 Flow Cell Lever in Position 1



- 2 Slowly move the flow cell lever to position 1 to disengage the manifolds.
- 3 Slowly move the flow cell lever to position 0 to disengage the vacuum seal and release the flow cell.

Figure 54 Flow Cell Lever in Position 0



- 4 Lift the used flow cell from the flow cell holder.

Clean the Flow Cell Holder

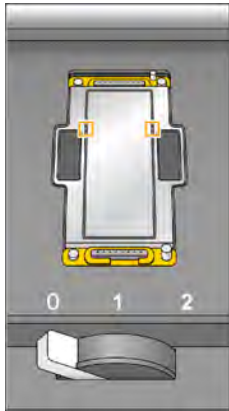
- 1 Put on a new pair of powder-free latex gloves.

- Using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol, carefully wipe the surface of the flow cell holder until it is completely clean.

**CAUTION**

Do not allow alcohol to drip into the vacuum holes or around the manifolds. Use a low-lint lab tissue to dry the stage, if necessary.

Figure 55 Inspect Vacuum Holes



- Visually inspect the flow cell holder to make sure that it is free of lint and the vacuum holes are free of obstructions.

Clean the Flow Cell

- Rinse the flow cell with laboratory-grade water and dry it with a lens cleaning tissue.
- Fold an alcohol wipe to approximately the size of the flow cell.
- Hold the edges of the clustered flow cell with two fingers. Make sure that the inlet and outlet ports are facing *up*.
- Wipe off each side of the flow cell with a single sweeping motion. Repeat, refolding the alcohol wipe with each pass, until the flow cell is completely clean.
- Dry the flow cell using a dry lens cleaning tissue.



CAUTION

If you clean the flow cell while it is lying on the bench top, you could easily apply too much pressure and break the flow cell. Illumina recommends cleaning the flow cell while holding the edges between your fingers.

- 6 Protect the flow cell from dust until you are ready to load it onto the instrument.

Load the Flow Cell



NOTE

Do not replace the manifold gaskets at this time. Leave the gaskets that were installed prior to the instrument maintenance wash in place when loading the flow cell. You will replace manifold gaskets after the current sequencing run is completed, prior to the instrument maintenance wash.

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

The manifold must be changed for Rapid Run mode every 10 days, prior to the instrument maintenance wash.

- 1 Place the flow cell on the flow cell holder with the inlet and outlet ports facing *down* and the barcode on the right. The arrow on the left edge of the flow cell indicates flow direction, and should point towards the instrument.

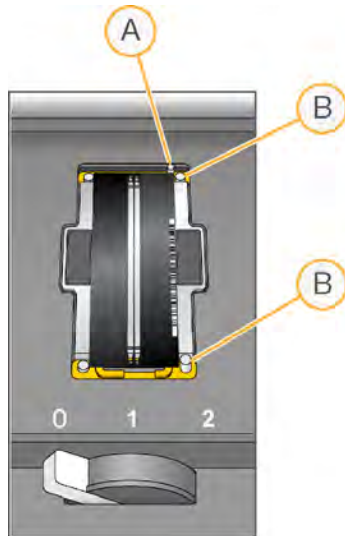


NOTE

Flow Cell v3 and the Rapid Run Flow Cell have a mechanically keyed corner, which provides a visual orientation for loading the flow cell. Install the Flow Cell so that the keyed corner is on the output end of the flow cell facing towards the instrument, and on the left side of the flow cell by lane 1.

- 2 Gently slide the flow cell towards the top and right guide pins until it stops.

Figure 56 Flow Cell Positioned Against Top and Right Guide Pins



- A Top Guide Pin
- B Right Guide Pins



NOTE

Remove your hand from the flow cell before engaging the vacuum switch. This prevents possible alignment drift over time.

- 3 Slowly move the flow cell lever to position 1. This engages the vacuum and secures the flow cell into position. When the flow cell lever is green, the vacuum is engaged.

Figure 57 Flow Cell Lever in Position 1



- 4 Wait for about five seconds, and then slowly move the flow cell lever to position 2. When the flow cell lever is solid green, the manifolds are in position and the flow cell is ready for use.

Figure 58 Flow Cell Lever in Position 2



- 5 Make sure that the **Vacuum Engaged** checkbox is selected on the load sequencing flow cell screen.

Confirm Proper Flow



NOTE

If you are performing template hybridization in the cBot, you will need to load a used flow for reagent priming and the fluidics check. You will also still need to check fluidics through the hybridized flow cell.

After loading the flow cell, check for proper flow using the fluidics check screen. Checking for proper flow confirms that the flow cell and gaskets are properly installed and the manifold is engaged.



NOTE

Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

- 1 Select solution 5 (USB) from the drop-down list.
- 2 Ensure that the following default values are entered:
 - Volume: **250**
 - Aspirate Rate: **1500**
 - Dispense Rate: **2000**
- 3 Select **Pump**.
- 4 Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.

If you see excessive bubbles, check the manifold gaskets for obstructions, and repeat the process.

 - a Select solution 6 (USB) to avoid depleting USB from position 5.
 - b Reduce the aspirate rate to 1000, and pump another 250 μ l of USB to the flow cell.
- 5 After you have confirmed proper flow, select **Next** to proceed. Make sure that the flow cell lever is green, and then close the flow cell compartment door.
- 6 Confirm that the **Vacuum Engaged** and **Door Closed** checkboxes are selected, and then select **Next**.
- 7 Select **Start** to start the sequencing run.

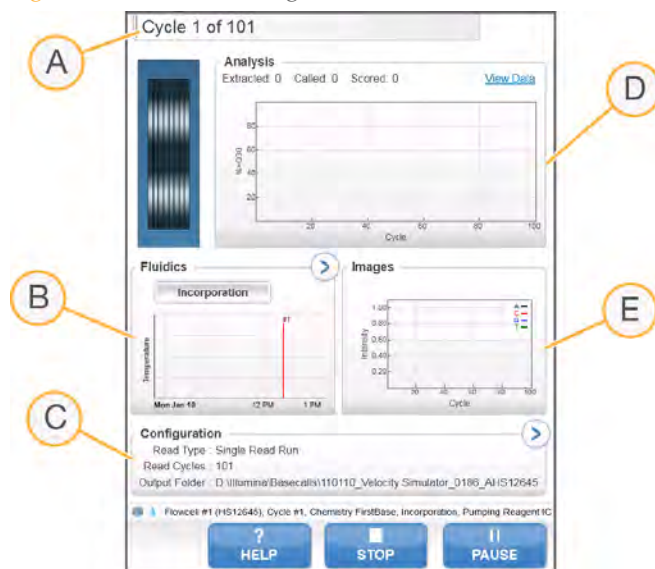
Monitoring the Run

Use the run overview screen to monitor on-instrument analysis, fluidics, and imaging.

- ▶ **Progress Bar**—Use the progress bar to monitor how many cycles have been completed.
- ▶ **Flow Cell Image**—Use the flow cell image to monitor which lanes have been imaged for that cycle.
- ▶ **Analysis Graph**—Use the analysis graph to monitor quality scores by cycle.
- ▶ **Images Graph**—Use the images graph to monitor intensities by cycle.
- ▶ **Fluidics Graph**—Use the arrow button to expand the fluidics section and monitor chemistry steps in the protocol.

In Rapid Run mode, HCS will automatically prompt you to perform a water wash when sequencing is finished. See *Performing a Water Wash* on page 169.

Figure 59 Run Monitoring Screen Overview



- A Run Progress
- B Fluidics and Temperature
- C Run Configuration
- D Analysis

E Images



NOTE

You might notice an intermittent clicking sound coming from the right side of the instrument during operation. This is normal. The clicking sound indicates that the refrigeration condensation pump is working.

First Base Report

The first base report confirmation dialog box opens automatically after cycle one is complete, if you selected **Confirm First Base** on the scan screen during run setup. The run will pause at this step.

If you did not select the checkbox, the run continues after first base incorporation without prompting you to review the first base report. For more information, see *Advanced* on page 42.

- 1 Review the first base report from the first base report confirmation dialog box or by opening `First_Base_Report.htm` from the root level of the run folder.
- 2 If the results are satisfactory, select **Continue** to continue your run. If results are not satisfactory, abort the run.

Sequencing Analysis Viewer

You can monitor your run using the Sequencing Analysis Viewer, which launches automatically after image analysis begins. For more information, see *Sequencing Analysis Viewer* on page 178.

For a detailed description of available metrics and installation instructions, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

[This page intentionally left blank]

Post-Run Procedures

Introduction	162
Unloading and Weighing Reagents	163
Performing an Instrument Maintenance Wash	164
Performing a Water Wash	169
Idling the Instrument	172
Shutting Down the Instrument	173



Introduction

This section and its subsections pertain to both Rapid Run and High Output mode runs.

Post-run procedures include the removal and weighing of reagents, and an instrument wash.

Regular instrument washes ensure continued performance in the following ways:

- ▶ Flushes any remaining reagents and sample from the fluidics lines and sippers
- ▶ Prevents salt accumulation and crystallization in the fluidics lines and sippers
- ▶ Prevents cross-contamination from the previous run, including cross-contamination from Rapid to High Output runs and vice versa.



NOTE

A water wash will be performed at the end of each Rapid Run to wash the system and check fluidics. HCS will not allow you to start a new run unless a water wash or required maintenance wash has been performed within the last 24 hours. Every 10 days, HCS will prompt you to perform a maintenance wash without regard to when you last performed a maintenance wash.

Unloading and Weighing Reagents

- 1 Open the reagent compartment door.
- 2 Raise the sippers for the reagent rack and clustering rack you are loading using the following motion:
 - a Pull the sipper handle towards you.
 - b Push the sipper handle up while pulling it towards you.
 - c Release the sipper handle into the slot on the top end of the groove. You should feel the sipper handle rest securely into the slot.
- 3 Slide the reagent rack out of the reagent compartment using the rack handle.
- 4 Remove each bottle from the reagent rack and record the weight on the lab tracking form.
- 5 (Rapid Run mode) Remove tubes from sample loading station.



NOTE

The color of sequencing reagent waste might be brown in appearance. You also might see a slight discoloration on the funnel caps or in any spills. The change in waste color is not toxic and does not impact run performance.



WARNING

A component in the clustering kit contains formamide, an aliphatic amide that is a probable reproductive toxin. Personal injury can occur through inhalation, ingestion, skin contact, and eye contact.

Dispose of containers and any unused contents in accordance with the governmental safety standards for your region.

For more information, see the MSDS for this kit, at <http://www.illumina.com/msds>.

Performing an Instrument Maintenance Wash

To ensure optimal instrument performance, perform an instrument maintenance wash every 10 days, after each sequencing run, and prior to switching between Rapid Run mode and High Output mode. Change the gasket at the time of the maintenance wash. Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold. The maintenance wash consists of three wash steps:

- ▶ **First water wash**—Flushes the system with laboratory-grade water.
- ▶ **NaOH wash**—Washes the system with NaOH.
- ▶ **Final water wash**—Flushes the system with laboratory-grade water to remove NaOH.

Positions	Approximate Run Time Per Step
Eight SBS positions	20 minutes
Eight SBS positions and ten cluster positions	60 minutes
(Rapid Run mode) Eight SBS positions, ten cluster positions, and 1 sample loading position for each flow cell	10 minutes

User-Supplied Consumables

- ▶ Ethanol wipes
- ▶ 24 bottles, 250 ml (Corning, catalog # 430776), eight bottles per wash step
- ▶ 30 tubes, 15 ml (Corning, catalog # 430052), 10 tubes per wash step
- ▶ Laboratory-grade water
- ▶ 1 N NaOH
- ▶ 3 eppendorfs per flow cell for sample loading station wash steps

First Water Wash

- 1 From the HCS Welcome screen select **Wash | Maintenance**.
- 2 (High Output mode) Select **Yes** to wash PE reagent positions. Otherwise, select **No**. Select **Next** to proceed.
- 3 Load the instrument with laboratory-grade water as follows:

- a Fill eight 250 ml bottles with more laboratory-grade water than there were reagents present at the beginning of the run to ensure that the sippers are rinsed.
 - b (Rapid Run mode) Fill eppendorfs with 1 ml laboratory-grade water
 - c If you are washing clusters, load ten 15 ml tubes with 10 ml laboratory-grade water.
 - d Load the bottles and tubes onto the instrument in the appropriate reagent rack.
- 4 Remove the flow cell from the flow cell stage, and set it aside. You will reload this flow cell before starting the instrument wash.



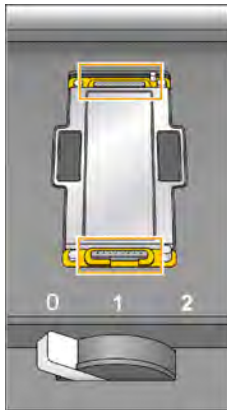
NOTE

Gaskets should not be changed during the NaOH and second wash steps; they should remain installed until the sequencing run is complete.

Always install new manifold gaskets before performing an instrument maintenance wash. Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

- 5 Ensure you have protective gloves on, and remove the front and back gaskets. Do so by applying light pressure with your finger to one side of the gasket. The other side will pop up. Use tweezers to grasp and remove each gasket.

Figure 60 Remove Used Manifold Gaskets



- 6 Position two new gaskets in the slots on the front end and back end of the flow cell holder. Regardless of which mode you are running, you will always have a 10 port gasket in the front manifold and an 8 port gasket in the back manifold.

- 7 Load the flow cell that you removed to install the new gaskets. Select **Next**.
- 8 Perform a fluidics check:
 - a Select solution 2 from the drop-down list.
 - b Enter the following default values:
 - Volume: **250**
 - (High Output mode) Aspirate Rate: **250**
 - (Rapid Run mode) Aspirate Rate: **1500**
 - Dispense Rate: **2000**
 - c Select **Pump**.
 - d Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.
- 9 (High Output mode) Bundle the eight lines of waste tubing with parafilm, making sure to keep all of the ends even. Place the bundled tube ends into a 250 ml bottle. (Rapid Run mode) Place the ends of tubes 4 and 5 into the waste bottle. Place the end of all of the other tubes into separate 15 ml tubes to measure the volumes for each lane for each step.
- 10 Select **Next**.
- 11 When the wash is complete, measure the delivered volume.



NOTE

The priming pump is washed during the wash steps. Keep the waste from the priming pump separate.

The following tables list the wash volumes from the eight lines of waste tubing. They do not include the priming pump wash volume.

Positions	Delivered Volume
Eight SBS positions	32 ml
Eight SBS positions and ten clusters	72 ml
(Rapid Run mode) One sample position, eight SBS reagents, 10 cluster reagents. SBS and cluster positions will be washed every time as they will always be used for every run.	9.5 ml per lane

NaOH Wash



CAUTION

For safety purposes, you should wear protective eyewear when working with large volumes of NaOH.

- 1 Load the instrument with 1 N NaOH as follows:
 - a Load eight 250 ml bottles with 5 ml of NaOH.
 - b If you are washing cluster positions, load ten 15 ml tubes with 5 ml NaOH.
 - c (Rapid Run mode) Fill eppendorfs with 1 ml 1N NaOH
- 2 Load the bottles and tubes onto the instrument in the appropriate reagent rack.
- 3 Select **Next** to start the NaOH wash.
- 4 When the NaOH wash is complete, measure the delivered volume.

Positions	Delivered Volume
Eight SBS positions	16 ml
Eight SBS positions and ten cluster positions	36 ml
(Rapid Run mode) One sample position, eight SBS reagents, ten cluster positions. SBS and cluster positions will be washed every time as they will always be used for every run.	4.75 ml per lane

Final Water Wash

- 1 Load the instrument with laboratory-grade water as follows:
 - a Load eight 250 ml bottles with at least 20 ml of laboratory-grade water.
 - b If you are washing cluster positions, load ten 15 ml tubes with at least 10 ml laboratory-grade water.
 - c Fill eppendorfs with 1 ml laboratory-grade water.



CAUTION

Do not reuse the same water or wash bottles that you used for the first wash step. The water from the first wash step may be contaminated with reagents that were present on the sippers.

- 2 Load the bottles and tubes onto the instrument in the appropriate reagent rack.
- 3 Select **Next** to start the final water wash.

- When the final water wash is complete, measure the delivered volume.

Positions	Delivered Volume
Eight SBS positions	32 ml
Eight SBS positions and ten cluster positions	72 ml
(Rapid Run mode) One sample position, eight SBS reagents, ten cluster positions. SBS and cluster positions will be washed every time as they will always be used for every run.	9.5 ml per lane

- Unwrap the waste lines and return them to the waste container.

Performing a Water Wash

Perform a water wash if the instrument has been idle for one day or more before you begin a new sequencing run.

For Rapid Run mode, you will not need to navigate to the home screen before starting the post-run wash as HCS will automatically prompt you to do so. Volume for the Rapid Run post run wash is greater than for the standard water wash.



NOTE

HCS will prompt you to perform a water wash after each Rapid run.

HCS will not allow you start a new run if the instrument has been idle for more than 24 hours.

Before the start of every run, be sure to carry out the recommended fluidics check, as you are prompted by the software, to ensure fluidics are working properly.

Illumina-Supplied Consumables

- ▶ Laboratory-grade water

User-Supplied Consumables

- ▶ Eight 250 ml bottles (Corning, catalog # 430776)
- ▶ Ten 15 ml conical tubes (Corning, catalog # 430052)
- ▶ One eppendorf flip-top tube for each flow cell

Procedure

- 1 From the HCS Welcome screen select **Wash | Water**.
- 2 (High Output mode) Select **Yes** to wash cluster positions. Otherwise, select **No**. Select **Next** to proceed.
- 3 Load the instrument with laboratory-grade water as follows:
 - a Fill eight 250 ml bottles with 20 ml laboratory-grade water.
 - b If you are washing cluster positions, load ten 15 ml tubes with 10 ml laboratory-grade water.
 - c (Rapid Run mode) Fill the eppendorf tube with 1 ml laboratory-grade water.



NOTE

Manifold gaskets do not require replacement prior to a water wash.

- 4 Make sure that a used flow cell is loaded. Load a used flow cell, if necessary. Select **Next**.



NOTE

Illumina recommends a fluidics check if you loaded a used flow cell for the wash.

- 5 If you loaded a used flow cell in the previous step, perform a fluidics check:
 - a Select solution 2 from the drop-down list.
 - b Enter the following default values:
 - Volume: **250**
 - (High Output mode) Aspirate Rate: **250**
 - (Rapid Run mode) Aspirate Rate: **1500**
 - Dispense Rate: **2000**
 - c Select **Pump**.
 - d Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.
- 6 Bundle all of the lines of waste tubing with parafilm, making sure to keep all of the ends even.

High Output: Place the bundled tube ends into a 250 ml bottle.

Rapid Run: Place the waste line from each lane into a separate 15 ml tube.
- 7 Select **Next** to start the water wash.

Positions	Approximate Run Time
Eight SBS positions	20 minutes
Eight SBS positions and ten paired-end positions	60 minutes
(Rapid Run mode) Eight SBS positions, ten cluster positions, and 1 sample loading position for each flow cell	10 minutes per SBS cluster

- 8 When the wash is complete, measure the delivered volume. Record the delivery volume on the lab tracking form.

Positions	Delivered Volume
Eight SBS positions	32 ml
Eight SBS positions and ten cluster positions	72 ml
(Rapid Run mode) Eight SBS reagents, ten cluster positions, and 1 sample loading position. SBS and cluster positions will be washed every time as they will always be used for every run.	9.5 ml per lane

- 9 Return the waste tubing to the waste bottle.

Idling the Instrument

If the instrument will not be used within the next ten days, prepare it to sit idle. If the instrument will not be used for more than ten days, see *Shutting Down the Instrument* on page 173.

- 1 Perform a complete maintenance wash to fully flush the system. For more information, see *Performing an Instrument Maintenance Wash* on page 164.
- 2 Leave the flow cell on the flow cell stage with the flow cell lever in position 2. This leaves the manifolds in the raised position.
- 3 (High Output mode) Load 10 ml of laboratory-grade water in each reagent position in the reagent racks, and then lower the sippers into the water.
(Rapid Run mode) Load 10 ml of laboratory-grade water in each reagent position and 1 ml of laboratory-grade water in the sample loading station position in the reagent racks, and then lower the sippers into the water.
- 4 Leave the instrument on.
- 5 When you are ready to use the instrument for another run, perform a water wash. For more information, see *Performing a Water Wash* on page 169.

Shutting Down the Instrument

Turn off the instrument only if it will not be used for more than ten days. If the instrument will be used within the next ten days, see *Idling the Instrument* on page 172. If you turn off the instrument, you do not need to turn off the PC.

Use the following procedure to safely prepare fluidics and shut down the system.

- 1 Perform a complete maintenance wash to fully flush the system. For more information, see *Performing an Instrument Maintenance Wash* on page 164.
- 2 Remove the flow cell(s) from the flow cell stage.
- 3 Using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol, carefully wipe the surface of the flow cell holder until it is completely clean.



CAUTION

Do not allow alcohol to drip into the vacuum holes or around the manifolds. Use a low-lint lab tissue to dry the stage, if necessary.

- 4 (High Output mode) Load 10 ml of laboratory-grade water in each reagent position in the reagent racks, and then lower the sippers into the water.
(Rapid Run mode) Load 10 ml of laboratory-grade water in each reagent position and 1 ml laboratory-grade water in each sample loading station position in the reagent racks, and then lower the sippers into the water.
- 5 Turn off the instrument.
- 6 When you are ready to restart the instrument, load water in all reagent positions, turn on the instrument, and perform a water wash. For more information, see *Performing a Water Wash* on page 169.

[This page intentionally left blank]

Real Time Analysis

Real Time Analysis Overview	176
Monitoring Runtime Statistics	178
Real Time Analysis Output Data	180



Real Time Analysis Overview

The integrated primary analysis software, Real Time Analysis (RTA), performs image analysis and base calling, and assigns a quality score to each base for each cycle. The analysis is performed during the chemistry and imaging cycles of a sequencing run, which saves downstream analysis time and allows you to quickly decide whether or not your run is progressing as expected.

Real Time Analysis (RTA) runs automatically on the instrument computer, and is configured through the HCS interface at the beginning of a run.



NOTE

If for any reason RTA stops functioning, the software automatically resumes analysis during the next cycle at the appropriate point on the flow cell. Do not restart the application manually.

Data Transfer

The output from a sequencing run is a set of quality-scored base call (*.bcl) files, which are generated from the raw image files and contain the base calls per cycle. By default, images are deleted from the instrument computer after image analysis. The raw image data are not needed for downstream analysis.

If for any reason image analysis falls too far behind, the software stops sequencing and places the flow cell in a safe state. Sequencing resumes when image analysis issues are resolved.

Basecalling_Netcopy_complete

Data transfer is complete when the Basecalling_Netcopy_complete files are present in the root level of the run folder. There should be one generic file named Basecalling_Netcopy_complete.txt, plus one file for each read.

Analysis of RTA Data

Both bcl2fastq and CASAVA v1.8's configureBclToFastq.pl script have the capability to convert bcl files to fastq files and demultiplex indexed data simultaneously. However, CASAVA 1.8's configureBclToFastq.pl script is not compatible with zipped bcl files. Bcl2fastq will handle bcl conversion and demultiplexing of both zipped and unzipped bcl files.

CASAVA v1.8.2 or later or bcl2fastq 1.8.3 or later are required to analyze RTA data output from dual-indexed sequencing runs. For more information, see the *CASAVA v1.8 User Guide*, Part # 15011196.



NOTE

CASAVA v1.8.2 or later is required to analyze RTA data output from dual-indexed sequencing runs.

Monitoring Runtime Statistics

Real Time Analysis automatically generates quality metrics when image analysis begins. Not all metrics are available at the early cycles because some processes need multiple cycles to generate data. The following summary describes points in the run when metrics are available:

- ▶ **Image Analysis**—Begins after cycle 4. The software first needs to generate a template with cluster locations. At cycle 4, the software catches up with the image analysis.



NOTE

Because the software catches up with the image analysis during cycle 4, cycle 4 takes significantly longer to finish. The software waits for the analysis to end before proceeding with the next cycle. This behavior is normal.

- ▶ **Phasing Estimates and Base Calls**—The software generates phasing/prephasing estimates and base calls after cycle 12.
- ▶ **Quality Scores**—The software generates quality scoring after cycle 25. The software determines which of the reads pass the quality filter, applying the Chastity formula over the first 25 bases. At least 24 of the first 25 cycles must have a Chastity value >0.6 . Reads may pass the quality filter if only one base in the first 25 has a lower Chastity value. Quality scores will be reported only for reads passing the quality filter. See RTA Theory of Operations for more details.
- ▶ **Error Rates**—If a PhiX control was included, RTA generates error rates after cycle 25.
- ▶ **In-Line Controls**—If you prepared your samples with the TruSeq sample prep kits, RTA will perform in-line control alignment at cycle 52 of each read or the end of the read if the read length is less than 52 cycles.
- ▶ **Index Count**—After the last index read is completed, RTA will report the index count per lane if a sample sheet was provided.

Sequencing Analysis Viewer

The Sequencing Analysis Viewer (SAV) is an application that allows you to view important sequencing metrics generated by Real Time Analysis (RTA) in plots, graphs, and tables. SAV automatically opens to the analysis page after image analysis begins.

Statistics are updated as soon as new analysis information is available. Select **Refresh** at any time during the run to view updated metrics as the run progresses.

SAV can be viewed from a remote location by installing SAV on another computer with access to the network location of the output run folder. You can download a copy of SAV from the Illumina website. A MyIllumina login is required.

For a detailed description of available metrics and installation instructions, see the *Sequencing Analysis Viewer User Guide*, Part # 15020619.

Analysis Page

The analysis page reports the progress of the analysis, and provides access to the most important runtime statistics.

Statistics are updated as soon as new analysis information is available. The runtime statistics can be viewed from any computer with access to the experiment run folder. Navigate to the <RunFolder>\Data folder and open the Status file. The Status file is transferred to the Data subfolder in the analysis folder at the end of each cycle that is processed.



NOTE

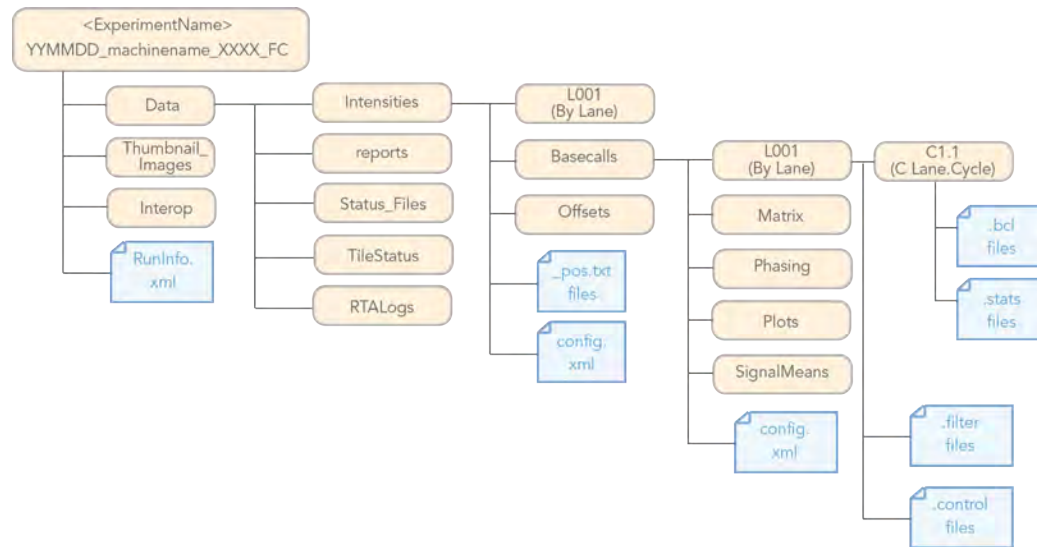
Your browser might warn you about active content. If so, click the warning in the information bar and allow blocked content.

Real Time Analysis Output Data

RTA output results are saved to the network location that you specified during the run setup steps.

All analysis results are saved in an analysis output folder that is given the same name as the run folder on the instrument computer. The following figure illustrates the organization of the network analysis folder, including the location of the most important output files (blue boxes).

Figure 61 Real Time Analysis Folder



Analysis Results

Table 16 Contents of Network Analysis Folder

Key File	Subfolder	Description
*.bcl files	AnalysisFolder\Data\ Intensities\Basecalls\ L00X\CX.X	Each *.bcl file contains RTA base calling and base quality scoring results for one cycle, one tile. These files are the input into CASAVA v1.8.2 or the BCL Converter. For more information, see <i>Main Sequence Output Files</i> on page 183.
*.cif files	AnalysisFolder\Data\ Intensities\L00X\CX.X	Each binary *.cif file contains RTA image analysis results for one cycle, one tile. These files should be used in the OLB analysis software if you want to perform BUSTARD base calling; for a description, see the <i>Off-line Basecaller Software User Guide</i> .
.clocs/.locs files	AnalysisFolder\Data\Intensities\L00	These files contain the x, y position for every cluster. The *.clocs file is a compressed version of the *.locs file.
*.filter files	AnalysisFolder\Data\ Intensities\Basecalls	*.filter files contain filter results per tile. These files are used by CASAVA v1.8.2 or later.
*.jpg files	AnalysisFolder\ Thumbnail_Images\ L00X\CX.X	HCS can generate thumbnail images of nine locations from each tile. The thumbnails are generated for each cycle and base, and can be used to troubleshoot a run. For more information, see <i>Thumbnail Images</i> on page 184.
*.stats files	AnalysisFolder\Data\ Intensities\Basecalls\ L00X\CX.X	*.stats files contain RTA base calling statistics for a given cycle/tile. These files are used by CASAVA v1.8.2 or later.

Key File	Subfolder	Description
*_pos.txt files	AnalysisFolder\Data\ Intensities\	Reports the cluster coordinates. There is one *_pos.txt file for each tile. These files should be used in the OLB analysis software if you want to perform BUSTARD base calling.
config.xml	AnalysisFolder\Data\ Intensities\ AnalysisFolder\Data\ Intensities\Basecalls\ 	Contains meta-information about the run, image analysis, and base calling.
CopyLog.txt	AnalysisFolder\Data	All copy events during analysis are captured in the CopyLog file.
ErrorLog.txt	AnalysisFolder\Data	Any error during analysis is captured in the ErrorLog file. If there are no errors, no ErrorLog.txt file is generated.
RunInfo.xml	AnalysisFolder\ 	Identifies the boundaries of the reads (including indexing reads), the Flowcell ID (if available), and the quality table selected for run.
SampleSheet.csv	AnalysisFolder\ AnalysisFolder\Data\ Intensities\Basecalls\ 	Contains sample information, including indexing information. RTA copies the Sample Sheet file in the root folder of the run to the Basecalls folder.

Tile Numbering

The tile name is a four-digit number encoding the position on the flow cell.

- ▶ The first digit encodes surface:
 - 1 is for top
 - 2 is for bottom
- ▶ The second digit encodes the swath:
 - 1 is for the first swath
 - 2 is for the second swath
 - 3 is for the third swath (only on Flow Cell v3)

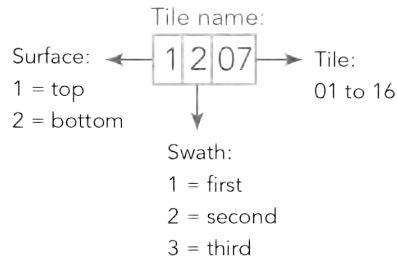


NOTE

A swath is a column of tiles. There are two swaths per lane per surface on Flow Cell v3 there are three swaths per lane per surface.

- ▶ The last two digits encode the tile, 01 through 16 (with HCS v2.0). Tile numbering starts with 01 at the output end of the flow cell through 16 at the input end.

Figure 62 Tile Numbering



This example indicates a tile from the top surface of the flow cell, the second swath, and the seventh tile.

Main Sequence Output Files

The following output file types contain primary analysis data generated by Real Time Analysis:

Base Call Files

A base call (*.bcl) file contains base calls and quality scores for one cycle and one tile. They can be used for later analysis off-instrument.



NOTE

If you are using CASAVA v1.7 or earlier, the *.bcl files need to be converted to the _qseq.txt format in OLB or the BCL Converter to enable downstream analysis. For more information, see the *BCL Converter User Guide* or the *Off-line Basecaller Software User Guide*.

This step is not necessary with CASAVA v1.8.2 or later.

Stats Files

The Stats file is a binary file containing base calling statistics for clusters passing filter. These files are used by the BCL Converter and CASAVA v1.8.

Filter Files

The filter (*.filter) files are binary files containing filter results. These files are used by the BCL Converter and CASAVA v1.8.

Control Files

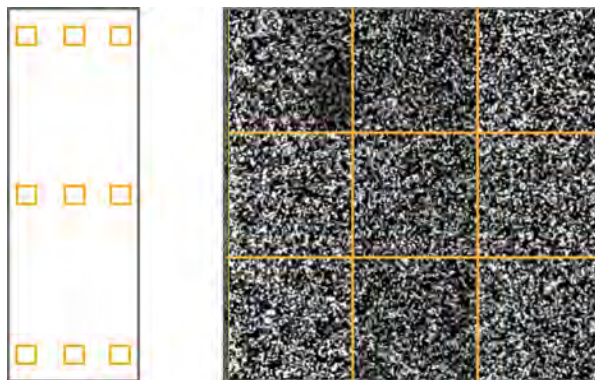
The control (*.control) files are binary files containing control results. These files are used by the BCL Converter and CASAVA v1.8.

Thumbnail Images

You can configure HCS to generate thumbnail images in *.jpg format. Thumbnail images are generated for each cycle and base.

HCS collects images from nine sections of a tile. These nine images are combined in one thumbnail image and can be used to troubleshoot a run. Thumbnail images are not suitable for image analysis.

Figure 63 Thumbnail Image



Troubleshooting

Introduction	186
Possible Run Setup Problems	187
Performing a Fluidics Check	188
Repeating Read 2 Primer Hybridization	189
If BaseSpace is Unavailable	190



Introduction

This section describes what you possible solutions if you encounter a problem during a sequencing run and how to perform a fluidics check from the Welcome screen.

Possible Run Setup Problems

Problem	Possible Cause	Action
The software did not initialize.	The software was unable to initialize internal hardware devices.	<p>Close the error message and then re-launch the instrument software.</p> <p>If the problem persists, restart the instrument computer. If you are going to restart the computer, you must first shutdown the instrument to ensure that the DoNotEject drive is recognized correctly.</p> <p>If the problem persists after restarting the instrument computer, shut down the instrument, wait a minimum of 60 seconds, and then restart the instrument.</p>
Flow cell lever is orange.	<p>The flow cell did not seat properly.</p> <p>The vacuum did not seal.</p> <p>Manifolds did not raise.</p>	<p>Remove the flow cell and repeat the cleaning steps.</p> <p>Make sure that the gaskets are present and well-seated.</p> <p>Reload the flow cell.</p>
Poor fluid delivery.	Potential bubbles in the system.	<p>Reposition the flow cell and confirm that the holes are facing <i>down</i>.</p> <p>Look for white precipitate around the gaskets. If precipitate is present, replace the gaskets. Always replace gaskets before an instrument maintenance wash.</p> <p>Confirm that the sipper assemblies are fully lowered and each sipper is in contact with the reagents.</p>

Performing a Fluidics Check

The Check button on the Welcome screen allows you to perform a fluidics check when you are not setting up a run. Use this option during instrument installation and fluidics troubleshooting.

- 1 Load a used flow cell onto the instrument. For more information, see *Load the Used Flow Cell* on page 98.
- 2 Load eight 250 ml bottles with wash solution or laboratory-grade water, and load the bottles onto the corresponding reagent rack. Load the rack onto the instrument.
- 3 Select **Check** on the Welcome screen.
- 4 Select solution 5 (SB2) from the drop-down list. If you are performing a fluidics check with a used flow cell installed, you can select solution 2, which is water.
- 5 Enter the following default values:
 - Volume: **250**
 - (High Output mode) Aspirate Rate: **250**
 - (Rapid Run mode) Aspirate Rate: **1500**
 - Dispense Rate: **2000**
- 6 Select **Pump**. If you need to pause the fluidics check, select **Pause**.
- 7 Visually inspect the flow cell for bubbles passing through the lanes and leaks near the manifolds.

If you see excessive bubbles, check the manifold gaskets for obstructions, reduce the aspirate rate to 100, and pump another 250 μ l of water to the flow cell.

Repeating Read 2 Primer Hybridization

It is possible to repeat Read 2 primer hybridization on the instrument; however, additional reagents are required. Contact Illumina Technical Support for further evaluation on whether a Read 2 rehybridization is appropriate.

For more information, visit the HiSeq support page on the Illumina website, and click **Troubleshooting**. A MyIllumina login is required.

If BaseSpace is Unavailable

If BaseSpace is not available, please open Windows Services and ensure that Illumina BaseSpace Broker is started; if it is not started, restart it. If services are running and you are still unable to access BaseSpace, please contact Illumina technical support.

A

- activity indicators 36
- analysis folder 180
- analysis page 58, 179
- analysis results 181, 183

B

- base calling results 64
- bcl files 181, 183

C

- cif files 181
- clicking sound 110, 158
- clocs files 181
- components
 - flow cell compartment 6, 8, 96, 145
 - fluidics compartment 6
 - optics module 6
 - reagent compartment 6-7, 163
- config.xml 64, 182
- consumables
 - Illumina sequencing kits 20
 - user supplied 26
 - version compatibility 25
- control files 183
- customer support 195
- cycles in a read 45

D

- data transfer 176
- disk space, checking 68
- documentation 195

- dual-indexed libraries 19
 - primers for 23-24
- dual indexing
 - on paired-end flow cell 89, 137
 - overview 16
 - settings on recipe screen 89, 137

F

- filter files 181, 183
- first base incorporation
 - confirmation 42, 57, 111, 135, 159
 - report location 42, 64, 135
- flow 104, 152
- flow cell
 - cleaning 105, 153
 - flow cell ID 41, 88, 135
 - lever 98, 106, 147, 154
 - loading 96, 103, 145, 151
 - troubleshooting 187
 - version compatibility 25
- flow cell compartment 6, 8, 96, 145
- Flow Cell Set-up Screen 41
- fluidics check 188
- fluidics compartment 6
- fluidics, troubleshooting 187
- funnel caps 22

H

- help, technical 195
- HiSeq Control Software 30, 32, 68

I

- icons
 - errors and warnings 36
 - status alert 36
- idling the instrument 172
- images
 - saving images samples 40
 - thumbnails 40
- indexed sequencing
 - reagents, loading 95
- indexing
 - reagents 84
- initialize software 33
- instrument wash 162, 169

L

- loading reagents 93, 117, 140, 163
- locs files 181

M

- maintenance wash 162, 164
- manifold gaskets 22, 164
- monitoring 110, 158
 - cluster intensities 110, 158
 - image samples 110, 158
 - run fluidics 110, 158

N

- Nextera libraries 19
- normal stop 59

O

- optics module 6

P

- paired-end
 - reagents 112
- paired-end sequencing
 - reagents, loading 117

- pos.txt files 182
- post-run procedures 162
- primer rehybridization 189
- primer requirements 24
- priming
 - preparation 101, 150
 - waste volume 101, 150

Q

- quality scores
 - generating 178

R

- Read 2 hybridization 189
- read length 45
- reagent compartment 6-7, 163
- reagent positions
 - indexing 95
 - paired end 117
- reagents
 - funnel caps 22
 - loading 95, 117, 163
 - loading, ICB 118
 - loading, SBS 93, 140
 - paired end 117
 - priming 101, 150
 - Reagents screen 48
 - recording kit ID 48
 - sequencing 163
 - version compatibility 25
- Reagents screen 48
- Real Time Analysis 30, 176
 - analysis folder 180
 - analysis page 58, 179
 - analysis results 181
 - data transfer 176
 - runtime statistics 178
 - status page 179
 - with stopped run 59
- recipe folder 64
- recipe screen 88, 135

- reports
 - first base incorporation 42, 64, 111, 135, 159
 - run folder contents 64
- resume run 34
- review run setup screen 88, 135
- Review screen 49
- run duration 78, 125
- run folders 64
- run overview screen 110, 158
- Run Overview screen 56
- RunInfo.xml 64, 182
- S**
- sample preparation
 - kits for dual-indexing 19
- sample sheet
 - editor 66
 - in run folder 64
 - overview 66
 - Sample Sheet screen 47
 - uploading 47
- Sample Sheet screen 47
- SAV 58, 111, 159, 179
- scan screen 88, 135
- Sequencing Analysis Viewer 30, 58, 111, 159, 179
- sequencing consumables 20
- sequencing primers 24
- show log file 33
- single indexing
 - overview 13
- software
 - about, version 33
 - Experiment Manager 66
 - HiSeq Control 32, 68
 - HiSeq Control Software 30
 - initialization 27
 - initialize manually 33
 - Real Time Analysis 30
 - RTA 176
 - run duration 78, 125
 - Sequencing Analysis Viewer 30
 - show log file 33
 - troubleshooting 187
 - version compatibility 25
 - view full screen 33
- sound, clicking 110, 158
- stats files 181, 183
- status alert icon 36
- status page 110, 158
- status.xml 64
- stopping a run 59
- Storage
 - BaseSpace 40
 - Storage screen 40
 - swath 40
- T**
- technical assistance 195
- tile numbering 182
- troubleshooting 187
- TruSeq HT libraries 19
- turning off instrument 173
- U**
- user-supplied consumables 26
- V**
- version compatibility
 - consumables 25
 - software 25
- W**
- washes
 - consumables 164, 169
 - delivered volumes 164, 169
 - maintenance wash 34, 162, 164
 - manifold gaskets 164
 - recommendations 162
 - water wash 34, 162, 169
 - water wash 162

Welcome screen 32
workflow
 run duration 78, 125

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 17 Illumina General Contact Information

Illumina Website	www.illumina.com
Email	techsupport@illumina.com

Table 18 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Italy	800.874909
Austria	0800.296575	Netherlands	0800.0223859
Belgium	0800.81102	Norway	800.16836
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000

MSDSs

Material safety data sheets (MSDSs) are available on the Illumina website at www.illumina.com/msds.

Product Documentation

Additional product documentation in PDF is available for download from the Illumina website. Go to www.illumina.com/support and select a product. A MyIllumina login is required. To register for a MyIllumina account, please visit my.illumina.com/Account/Register.

Part D – Proposed Intended Use/Indications for Use

Projects 1 and 2 of our research aim to clinically validate whether or not the DNA sequencing protocols we have described in Part C, Device Descriptions, will produce the same results as tandem mass spectrometry (MS/MS), the current technology, does for newborn blood spots. We also hope to determine if DNA sequencing will be able to detect certain Primary Immunodeficiencies (PID) that are not currently discoverable via MS/MS. We will utilize the blood spots of already diagnosed PID patients to see if DNA sequencing can clinically validate what MS/MS could not from the blood spots. Project 3 aims to set up Human Research Protection Protocols for the DNA sequencing of neonates for the entire University of California system. In addition, project 3 will analyze the legal, ethical and policy issues associated with DNA sequencing of newborns. The following pages contain procedural charts for the three projects. These are followed by the overall research plan for this grant to provide an in depth understanding of the project. The pages marked 129, 151 and 193 indicate the specific aims of each project followed by a more detailed discussion of each project's goals.

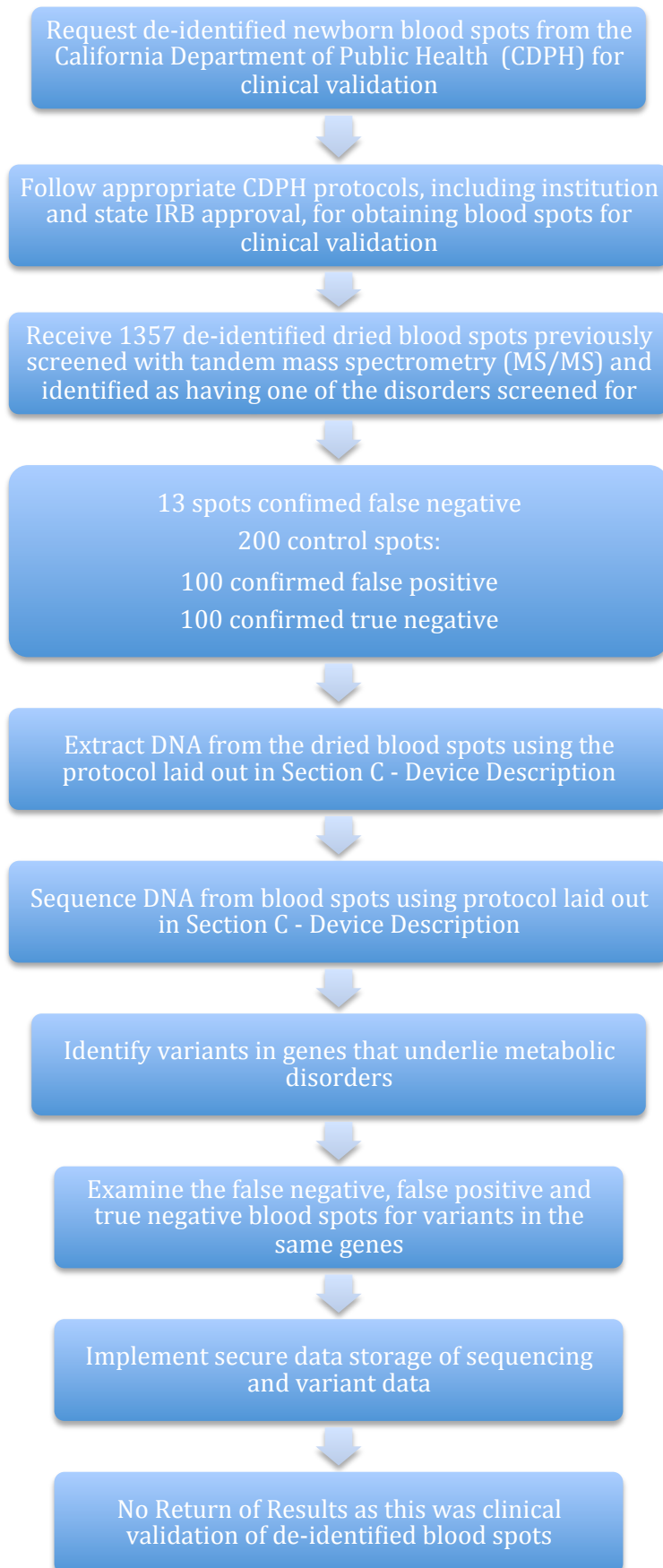
Sequencing of Newborn Blood Spot DNA to Improve and Expand Newborn Screening

University of California, San Francisco

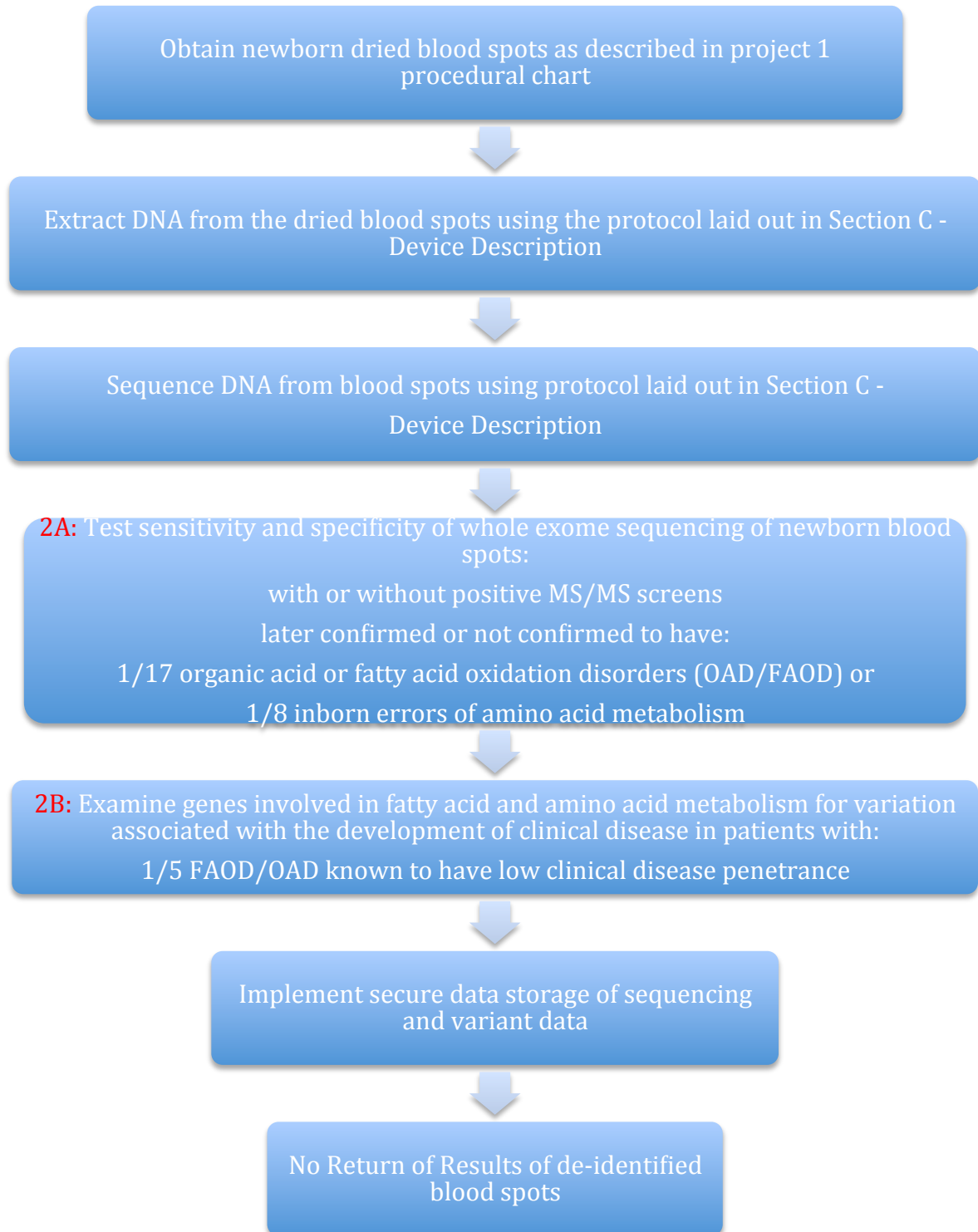
Principal Investigator: Robert Nussbaum, M.D

Projects 1-3

Project 1: Whole Exome Sequencing and Analysis of Variants in Newborn Blood Spots Relevant to Metabolic Disorders and Primary Immunodeficiency



**Project 2A and 2B - Whole Exome Sequencing as a Newborn Screening Tool for Metabolic and Immunological Disorders
in Conjunction with
Project 1 - Whole Exome Sequencing and Analysis of Variants in Newborn Blood Spots Relevant to Metabolic Disorders and Primary Immunodeficiency**



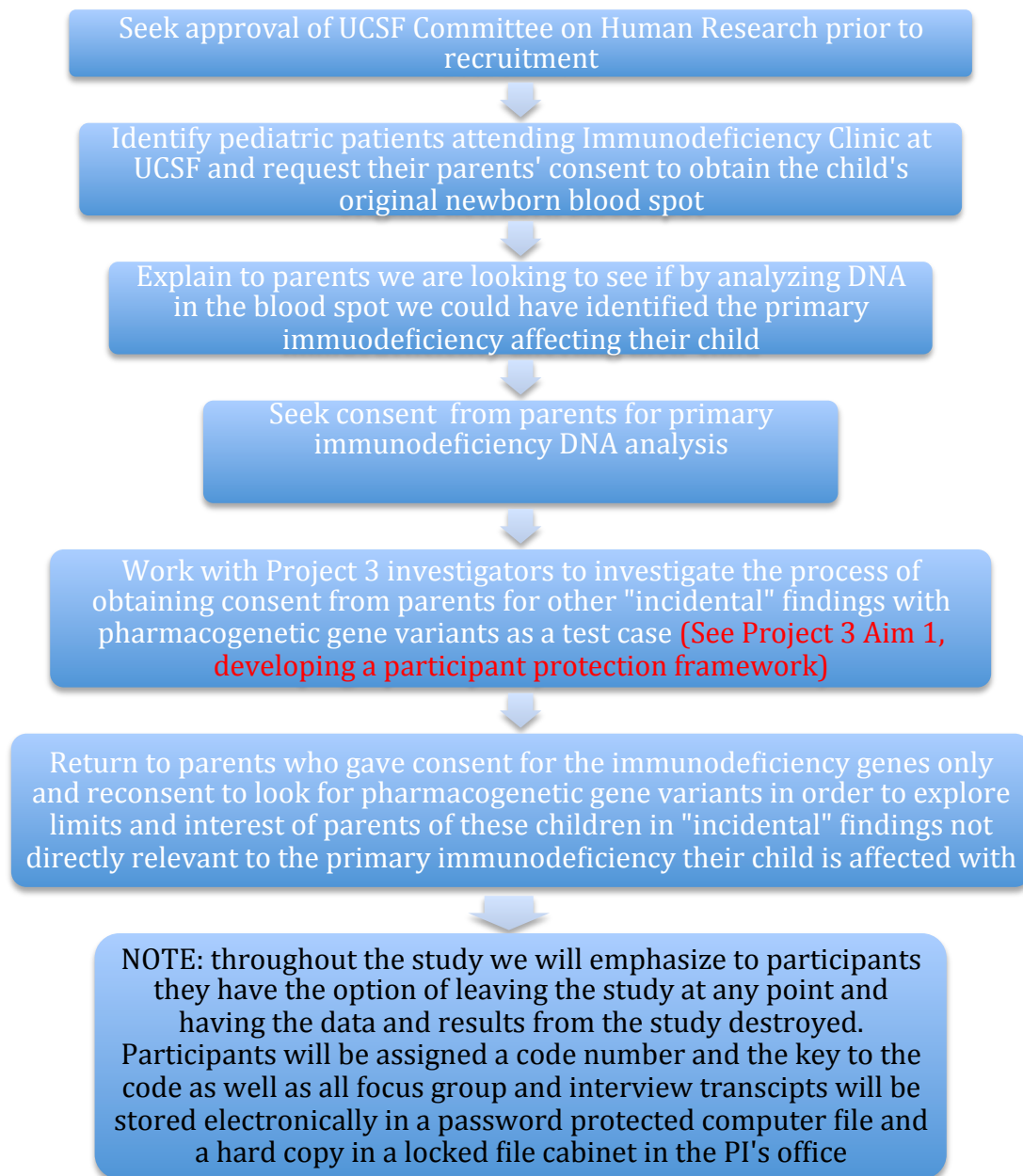
Project 2c - Examination of Variants in Selected Immunodeficiency and Pharmacogenetic Genes Obtained by Whole Exome Sequencing of Newborn Blood Spots Obtained from Patients Seen in Immunodeficiency Clinic for TREC+ Primary Immunodeficiency

in Conjunction with

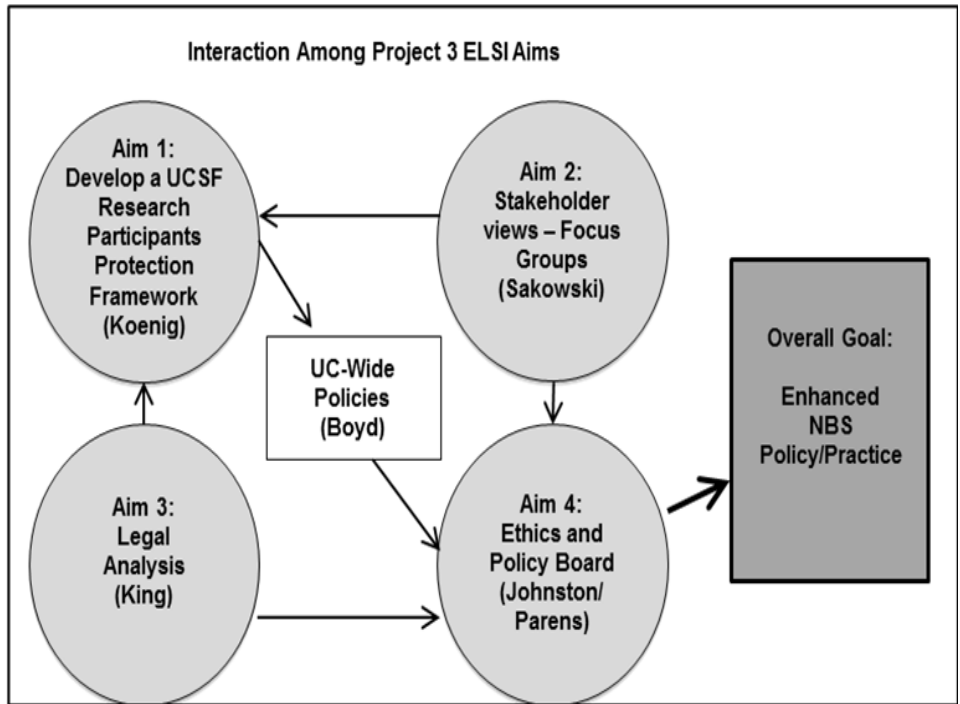
Project 1 - Whole Exome Sequencing and Analysis of Variants in Newborn Blood Spots Relevant to Metabolic Disorders and Primary Immunodeficiency

and

Project 3 - ELSI Implications of Research Related to DNA-Based Analysis Associated with Newborn Screening



Project 3 – ELSI Implications of Research Related to DNA-Based Analysis Associated with Newborn Screening



Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

OVERVIEW OF THE PROGRAM – SPECIFIC AIMS

Newborn screening (NBS) began in the United State nearly 50 years ago, and has become an essential public health program in all 50 states. It began with screening tests based on phenotypic assays specific to one or a few individual disorders but has expanded and evolved substantially over time. With tandem mass spectrometry¹, one test could detect as many as 48 different disorders, some of which were of uncertain clinical significance, were poorly understood, or had no effective treatment. Another major change occurred when genotyping was introduced for cystic fibrosis screening. The falling cost of whole genome/ exome sequencing has now created a new opportunity to ask whether sequencing of the genome might serve as a method of cost-effective newborn screening for every and any condition, in parallel, at the genotype level. We will address certain critical questions in the application of this technology to NBS as follows:

Project 1. Whole Exome Sequencing and Analysis of Variants in Newborn Blood Spots Relevant to Metabolic Disorders and Primary Immunodeficiencies (Analysis and Assembly of Genomic Datasets)

Carry out Whole Exome Sequencing (WES) from 1620 newborn blood spots of patients carefully selected to drive the research questions in Projects 2 and 3. The blood spots and associated data will be obtained from the California Department of Public Health Newborn Screening Program. The exome sequence will be analyzed for variants that will be characterized as to their frequency and pathogenicity. The researchers performing the variant analysis will be blinded to the phenotypes for all 1463 samples analyzed.

Project 1 is directly responsive to the RFA stipulation that the proposal address the “Acquisition and analysis of genomic datasets that expand considerably the scale of data available for analysis in the newborn period”

Project 2 Testing Whole Exome Sequencing as a Newborn Screening Tool for Metabolic and Immunological Disorders (Clinical Research of Disorders Identifiable through Newborn Screening)

Aim 2a Compare sensitivity and specificity of mutation data with biochemical testing by tandem mass spectrometry (MS/MS) in patients known to have one of 25 metabolic disorders for which MS/MS had been used for screening, i.e. fatty acid oxidation, organic acid, and amino acid disorders.

Aim 2b Combine WES with clinical outcomes data to identify gene variants in pathways associated with fatty acid oxidation that predict which children with fatty acid oxidation disorders are at greater risk for metabolic decompensation, thereby enhancing the clinical value of the WES data.

Aim 2c Apply WES to identify mutations in genes responsible for primary immunodeficiencies that warrant early diagnosis and treatment similar to what is provided for severe combined immunodeficiency (SCID) but are not detected by the current T-Cell receptor excision circle assay used for SCID screening.

Project 2 is directly responsive to the RFA stipulation that the proposal address “For disorders currently screened for in newborns, how can genomic sequencing replicate or augment known newborn screening results” and that the proposal include “Clinical research that will advance understanding of specific disorders identifiable via newborn screening through promising new DNA-based analysis” as well as address “What knowledge about conditions not currently screened for in newborns could genomic sequencing of newborns provide”

Project 3: ELSI Analysis of Whole Genome Analysis (WGA)-Enabled NBS

Aim 1 We will develop a participant protection framework for conducting WGA during the neonatal period, as an adjunct to the standard NBS blood spot.

Aim 2 We will determine the views, perspectives, and value preferences of key stakeholders about the potential expansion of NBS; methods will include focus groups with prospective parents, including parents of children with immunodeficiency (2a) and focus groups with pediatricians and obstetricians/gynecologists (2b).

Aim 3 In collaboration with the UC Hastings Consortium on Law, Science and Health Policy, we will identify the legal and constitutional issues for using WGA, and for incorporating PGx into NBS programs.

Aim 4 Informed by Aims 1-3, and in collaboration with a “Policy Advisory Board” convened by The Hastings Center, we will develop and disseminate policy recommendations for expanded NBS programs based on WGA. Project 3 utilizes the sequences generated from the patient cohorts in Projects 1 and 2 and is directly responsive to the RFA stipulations that the proposal address “Research related to the ethical, legal and social implications (ELSI) of the possible implementation of genomic sequencing of newborns”.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

SIGNIFICANCE

Newborn screening (NBS) began in the United States nearly 50 years ago, and although it has expanded and evolved substantially in this time, the decision as to which conditions to screen for has continued to be based on the following principles developed by the World Health Organization (WHO) ¹:

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole
10. Case-finding should be a continuing process and not a “once and for all” project.

Until 1997, public health departments applied very stringent WHO criteria and only used screening tests based on biochemical or other phenotypic assays specific to one or a few individual disorders in NBS programs. A major change occurred in 1997 when tandem mass spectrometry (MS/MS) was first used in a statewide newborn screening (NBS) program in North Carolina². The use of MS/MS represented a major advance in newborn screening because, for the first time, a technology was applied that could with one test detect at reasonable cost as many as 48 different disorders (See Table 1), some of which were very rare, or were of uncertain clinical significance, or had natural histories that were poorly understood, or for which no effective treatment was available.

Another major change occurred over the past decade with cystic fibrosis screening. For the first time, states began to introduce NBS based on genotype, although not as a stand-alone test but as a component of a tiered testing protocol that begins with biochemical phenotype (immunoreactive trypsin) followed by genotyping. For cystic fibrosis (CF) screening, genotype is determined for a few dozen common mutations that are well accepted as being pathogenic and complete sequencing of the *CFTR* gene is not used, thereby avoiding the issue of assessing pathogenicity of variants found from sequencing. Nonetheless, CF screening opened the door to the possibility that NBS need not be a purely phenotype-based form of testing.

Screening by MS/MS is a highly successful model of how to screen for many disorders simultaneously³. The falling cost of whole genome analysis (WGA) either by whole genome sequencing and its more limited but less costly surrogate, whole exome sequencing, has now created a new opportunity to consider whether WGA could serve as a single test for all newborn screening for any and all relevant disorders. Massively parallel sequencing of the entire genome might ultimately be the most effective form of newborn screening for all conditions that meet the WHO conditions 1-4 and 7-10 above, particularly those for which no disease-specific, phenotype-driven test is available.

The ethical and legal issues surrounding using a whole genome/exome approach for NBS are of course extremely important, given that one of the core principles for NBS, listed as #6 by the WHO, is that “the test should be acceptable to the population.” It is not clear that the general population would welcome a state-mandated test that was as intrusive as a whole exome or whole genome sequence. Thus, Project 3 is very important to the overall goals of this Program.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

MULTIPLE PI LEADERSHIP PLAN

Rationale and Justification. This U19 Project, “Exome Sequencing of Newborn Blood Spots to Improve and Expand Newborn Screening”, will use a Multiple PI Leadership strategy to effectively evaluate several applications of genome sequencing as a method of newborn screening at the genotype level and address important ethical and legal issues around genome sequencing in newborn screening. This type of ambitious project requires a multi-disciplinary team of experienced investigators with skill sets in genome sequencing, genetic epidemiology, statistical genetics and bioinformatics, as well as in ethical, legal and social implications (ELSI) research. To address this broad spectrum of scope, the Multi-PIs of this project appropriately reflect expertise in those disciplines: **Robert Nussbaum** (genomics, medical genetics), **Pui Kwok** (genomics, analysis of next-generation sequencing) and **Barbara Koenig** (anthropology, bioethics). By jointly-leading from their respective strengths, the design of the research plan is strengthened and the likelihood for advancement and success of the proposed projects are heightened beyond what any of the leaders could deliver independently or in a hierarchical approach. Each of the Multi-PIs carries responsibility to: (1) lead in a key function on the project, (2) provide centralized oversight that the talent and resources of their institute are engaged for the success of this project, and (3) foster successful crosstalk and collaboration between all major disciplines and collaborating sites within the study.

While responsibility for the different discreet tasks resides with each of the Multi-PIs, the execution of the responsibilities is affected through the Components (**Figure 1**). The Components will operate under the leadership of the noted PI, with co-Leadership responsibility to be shared by a designated individual member of a specific Component. This ensures timely execution of tasks, enhanced availability to the Executive Committee, a shared ownership for dispersed administrative responsibilities, and offers leadership development opportunities.

Administrative and Scientific Responsibilities. **Dr. Nussbaum** and UCSF are responsible for all major administrative tasks of the project. Dr. Nussbaum has a complete and dedicated team of administrators and grant specialists for this task. Plans, practices, and relationship maturation for bringing the blood spot samples from the California Department of Health will be guided by Dr. Nussbaum and operationally supported by a Clinical Coordinator; UCSF will house the overall study biorepository of spots and DNA samples. Dr. Nussbaum will be working closely with **Drs. Fred Lorey** and **Robert Currier** at the California Department of Public Health to identify and retrieve the appropriate samples and data. **Dr. Neil Risch** will be in charge of genetic epidemiology-related activities, which include: standardized case and control accrual protocols, standardized data elements, standardized instruments for questionnaire interview and medical chart review, and will contribute to study design and data analysis. **Dr. Steven Brenner** will be in charge of all elements of study design, data processing and statistical analysis. Dr. Brenner has a dedicated team at UC Berkeley for these tasks, including a statistician/statistical programmer and a research assistant. **Dr. Sean Mooney** and his group at the Buck Institute will work with Dr. Brenner, applying his extensive expertise in the pathogenicity and functional assessment of DNA variants by bioinformatics and *in silico* tools. The necessary IT infrastructure is already in place at UCSF, with dedicated secure servers and full access to the needed statistical software. **Dr. Joseph Shieh** will oversee the design and execution of the integrated fatty acid oxidation project, while working closely with **Drs. Nussbaum** and **Packman** as a collaborator. **Dr. Jennifer Puck** will provide her expertise in the identification of appropriate patients with primary immunodeficiencies and work with **Drs. Kwok, Brenner and Mooney** in assessing candidate variants for primary immunodeficiencies. Quality control and quality assurance of the primary sequence data are the responsibilities of **Dr. Pui Kwok** with additional assistance from Mr. Paul Tang in the Institute for Human Genetics Genomics Core for data analysis. Team members at UCSF are responsible for all web-based secure database development for study data entry and tracking of biospecimens.

Governance and Organizational Structure (Figure 1). Drs. Nussbaum, Kwok and Koenig, with substantial input from Drs. Brenner, Mooney, Shieh, Puck, Sakowski and King, are responsible for developing and communicating the Vision, specifying Experimental Design, and providing Supervision. The PIs will interface through weekly conferences (via tele- or web-based conference calls) to update on progress, problems, and advances, as well as to coordinate the monthly **Investigator Committee** meetings.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

The Investigator Committee is composed of the Multi-PIs and Team members representing the required spectrum of disciplines (Table 1). All scientific and operational assessments will be made by the Investigator Committee monthly meetings as a team. The Committee will also guide the development of Scientific Presentations and Manuscripts together with the input from others involved in the project within each of the listed disciplines. Execution of the recommended work from the Investigator Committee is directed by the Multi-PI leadership team as delivered by the efforts of the overall team. The cross-disciplinary Components interface through their own monthly **Component Meetings** where data is presented, and the interpretation and implications of the findings are processed. Additionally, two annual program meetings will be held to discuss all progress, problems, and advances in study accrual, data analysis and interpretation, scientific manuscripts and grants.

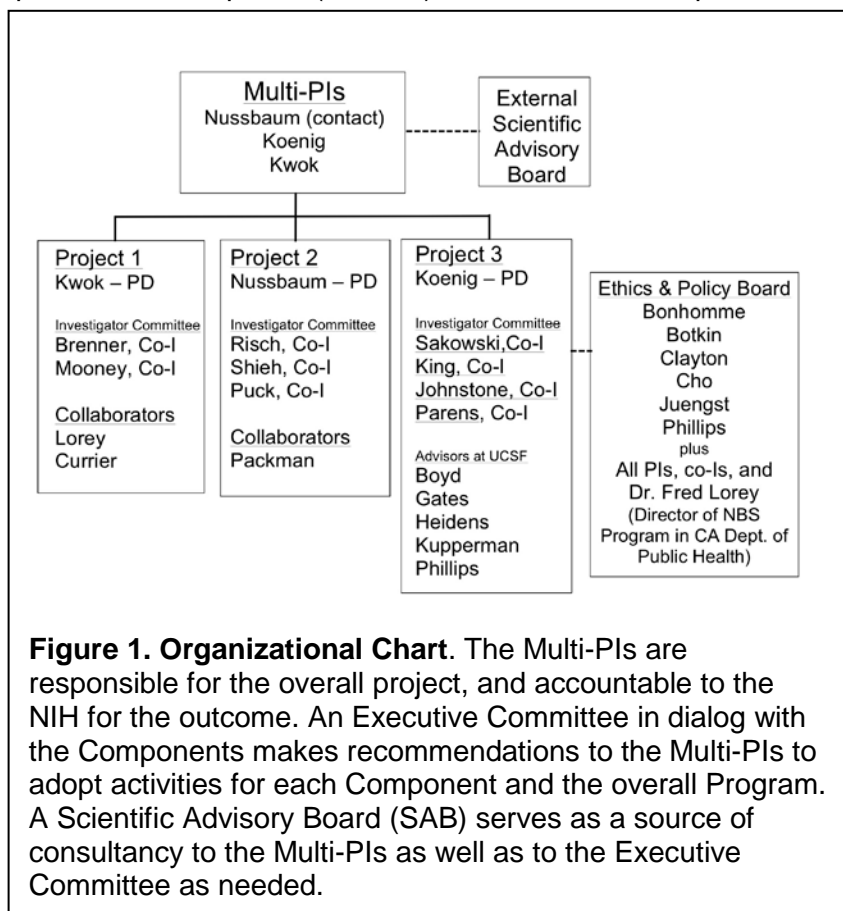


Figure 1. Organizational Chart. The Multi-PIs are responsible for the overall project, and accountable to the NIH for the outcome. An Executive Committee in dialog with the Components makes recommendations to the Multi-PIs to adopt activities for each Component and the overall Program. A Scientific Advisory Board (SAB) serves as a source of consultancy to the Multi-PIs as well as to the Executive Committee as needed.

expertise, will meet with the Multi-PIs and Investigator Committee and other key personnel. The advisors will be available on an *ad hoc* basis. The members of the ESAB will be asked to critically review details of study design and methods, study progress and interpretation of study results as they become available. The responsibilities of the ESAB are to engage with the research team on a quarterly basis in year 01 and biannually thereafter; this will be affected by tele-conference/web-conference and once a year at a US-based, in-person meeting. If necessary, the ESAB may be asked to make recommendations to resolve scientific conflicts (see below). In addition, Project 3 will specifically seek the guidance and input of an Ethics and Policy Board, convened by the Hastings Center, charged with developing an ethical framework for determining which variants revealed by WGA should be included (or excluded) as part of a state-run newborn screening program, identifying the barriers to implementation and factors likely to promote appropriate use.

Procedures for Resolving Conflicts. In the unlikely event that a conflict arises in the scientific trajectory of the project, the Multi-PIs will make every effort to resolve it amongst themselves. If outside counsel is warranted, the expertise of the SAB will be utilized to make a recommendation back to the Multi-PIs. Recommendations for resolution of conflicts related to administrative or operational issues will be made by the Program Managers and business infrastructure at the conflicted institutes; only in the improbable event of a complete impasse or a complete inability to resolve an issue will the contact PI (Dr. Nussbaum) then resolve administrative or operational issues since these ultimately impact performance towards the aims of the project.

Policies for Communication, Data Sharing, Publication, and Intellectual Property. UCSF utilizes bridge telecommunication and web-based tools from WebEx and Global Crossing for webcasts and dial-in conference calls, respectively; these facilitate connectivity for voice and visual exchanges. UCSF utilizes DropBox from Google with which all data, files, spreadsheets, presentations, and communications that do not contain protected health information related to specific projects are hosted AND accessible to pre-approved participants. All data developed in this project is to be openly available to the other members (Key Personnel)

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

of the project via their unique individual logins and passwords. Preparation of presentations, reports, posters and manuscripts developed from the proposed research will be co-authored by the scientific leadership, including the members of the multi-PI team and other professional research partners as appropriate; technical support and other research support will be assigned as either co-authors or acknowledged depending on the scientific contribution by the personnel in accord with international convention.

Table 1. Team Members

Robert Nussbaum	Multi-PI (contact), PD Project 2	Human Genetics and Genomics; Clinical and Biochemical Genetics
Pui Kwok	Multi-PI, and PD Project 1	Genomics, Next Generation Sequencing
Barbara Koenig	Multi-PI, and PD Project 3	Ethical and Social Implications of Genetics
Steven Brenner	Co-Investigator Project 1, Investigator Committee	Bioinformatics
Sean Mooney	Co-Investigator Project 1, Investigator Committee	Bioinformatics
Fred Lorey	Collaborator, CA Dept. of Public Health	Newborn Screening
Robert Currier	Collaborator, CA Dept. of Public Health	Epidemiology and Newborn Screening
Jennifer Puck	Co-Investigator, Project 2 Investigator Committee	Pediatrics and Immunology
Neil Risch	Co-Investigator Project 2, Investigator Committee	Genetic Epidemiology and Biostatistics; Human Genetics
Joseph Shieh	Co-Investigator Project 2, Investigator Committee	Human Genetics and Genomics; Clinical and Biochemical Genetics
Seymour Packman	Collaborator, Project 2	Clinical and Biochemical Genetics
Julie Sakowski	Co-Investigator Project 3, Investigator Committee	Genetics and Economics, Public Policy
Jaime King	Co-Investigator Project 3, Investigator Committee	Genetics and the Law
Josie Johnstone	Co-Investigator Project 3 (Hastings Center)	Genetics and the Law
Erik Parens	Co-Investigator Project 3 (Hastings Center)	Philosophy

Inventions conceived and rendered to practice in the course of this project will be disclosed as intellectual property in accord with each institute's policies on intellectual property; the institute which employs the individual making the dominant portion of the discovery has right of first refusal to develop the invention, after which institutes employing other co-inventors will have the option to exercise their right of refusal. Other co-inventors will be listed on the institute's invention disclosure form. Development of the invention (patent filing, licensing, trademarking, sale, etc.) is at the discretion of the inventor's institute, but the terms under which development is pursued or declined, and how the other inventors participate in any commercialization, if any, will be negotiated by the multi-PI team and the respective institutions prior to formal launch of the project.

Budget Issues. This is a multi-disciplinary project that relies heavily on the dedicated expertise of each of the members of the multi-PI team; salary support for this effort is a necessary component of the project's

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

budget, as is salary support for other dedicated professionals within the different teams and dedicated research staff as necessary. Travel expenses to present the findings at scientific meetings as well as publication costs will be a secondary source of adjustments to the budget if necessary. Financial information on the status of the expenditures will be securely and privately held for access by the members of the multi-PI team and the necessary individuals from their Grants and Contracts offices at their respective institutions.

INNOVATION

Although whole exome sequencing is not a new technology, combining the whole exome approach to sequencing with the very comprehensive biochemical and clinical data available through the California NBS program is a novel “marriage” of new and powerful technology with the largest newborn screening program in the United States. In addition, our innovative multi-disciplinary approach brings together experts in genetics, newborn screening, metabolic disorders, genomic analysis, bioinformatics, genetic epidemiology, public health law, and social implications of genomics research together into a single research endeavor.

Obtaining WES data in NBS offers additional potential advances in terms of clinical prediction by making use of all the derived sequence data, as compared to a limited mutation panel, as is conventionally done in CF screening, for example. Hence, another innovative element in this application is that we will attempt to characterize mutations at the primary disease locus in terms of mild versus severe pathogenicity, as well as search for modifier variants at other loci that predict clinical severity. If successful, these results can have a substantial impact on our ability to provide clinical care in the FAODs.

Our proposal also tests whether we can expand the capacity of NBS to identify children at risk from forms of immunodeficiency that are not detected by the T-cell receptor excision circle (TREC) assay for neonatal T-cell lymphopenia and for which there are many highly specialized and expensive tests particular to each of the immunodeficiencies that would need to be run to screen successfully for all of these disorders. Given that these children will have many heterogeneous disorders, there is no single analyte or laboratory phenotype that can be used for this screening and, therefore, the application of WES is the only uniform method available. Such a novel newborn screening method could serve as a model for many other disorders for which no single analyte or laboratory phenotype exists.

Project 3 is an innovative coordination of ethical and legal scholarship to address the broader questions raised by the falling cost of whole genome/exome sequencing and the serious discussion of whether and how sequence information from a newborn’s blood spot might be used for clinical care. Since NBS is an individual’s first point of contact with the healthcare system and is applied to everyone, the newborn period has been proposed as an opportune time for initiating a widespread, lifelong program of personalized medicine based on genome sequencing. There are complex technical, economic, clinical and ethical issues that must be dealt with first before any such program could be initiated.

Our proposed focus on a particular case example—pharmacogenetic (PGx) variants predicting response to drugs commonly used in childhood—is innovative. To our knowledge no NBS programs or clinical testing programs currently return PGx information that is relevant to the well-being of healthy infants or children, those who may be exposed to drugs used routinely such as codeine or medications for asthma or seizures. In addition, our plans to create a research participant protection program in collaboration with our regulatory and compliance colleagues at UCSF, and then expand that work throughout the UC system, are also highly innovative.

APPROACH

Interdependent Associations between Component Projects 1, 2 and 3 (Figure 2)

Project 1 is the foundational project in which all of the extensive exome sequence data will be generated and lists of exon and splice site variants prepared. A detailed description of the sequencing, analysis, and informatics tools to be used to store and analyze the data are outlined in the Approach Section of the Component Project 1.

The relationship between genotype and clinical phenotype (development of disease) is captured by the concepts of sensitivity, specificity and positive and negative predictive values. These are key issues to be

Program Director/Principal Investigator (Last, First, Middle):

Nussbaum, Robert L.

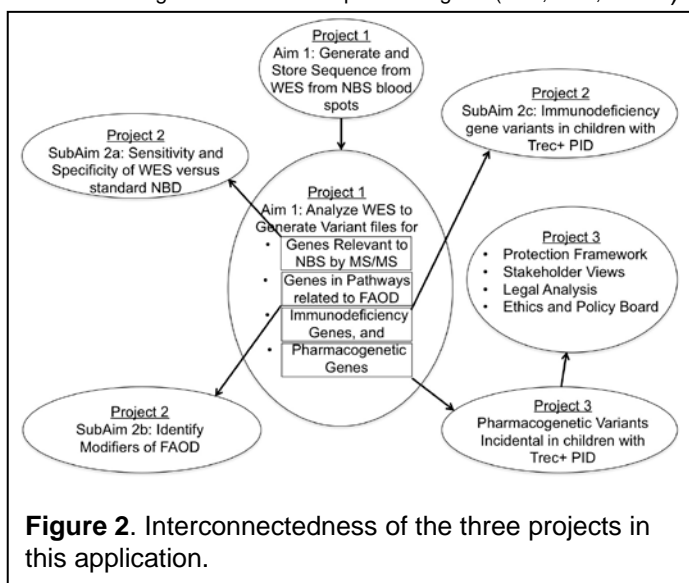


Figure 2. Interconnectedness of the three projects in this application.

from whole exome sequencing and applied retrospectively to blood spots of patients known to have a variety of metabolic disorders will work to identify these patients. This project will answer questions of sensitivity and specificity for whole exome sequencing in comparison to NBS by MS/MS for metabolic disorders. If sequencing can be shown to be comparable to current MS/MS methods, it will drive the field of newborn screening by allowing a single test to do everything the MS/MS is doing now but in addition would provide much additional information of susceptibility for many other types of diseases for which we cannot yet screen, such as non-SCID primary immunodeficiencies.

A second critical question, Question B in the RFA, asks, “*What knowledge about conditions not currently screened for in newborns could genomic sequencing of newborns provide?*” This question is addressed in Aim 2c of Project 2 of this application. We cannot currently screen for non-SCID primary immunodeficiency disorders, yet there is clinical value in knowing a child is affected by one of these non-SCID PIDs for all the reasons that SCID screening has been implemented using the TREC assay. We are looking to show in a retrospective study that whole exome sequencing of newborns already known to have a non-SCID PID would have allowed that child to be identified and treated from birth.

A third question, Question C in the RFA, asks, “*What additional clinical information could be learned from genomic sequencing relevant to the clinical care of newborns?*” This question is addressed by Aim 2b of Project 2 that seeks to identify genetic variation responsible for the variable phenotype of fatty acid oxidation disorders. (FAODs). If we can identify such main effects and modifiers, our ability to predict which patients with one of these FAODs will need treatment or is at increased risk for life-threatening metabolic decompensation may be greatly improved.

Project 3 will address a key question of how to approach the legal issues that arise when genetic variants that are not the direct reason for the sequencing, so-called “off-target” or “incidental” or “secondary findings” are obtained in the course of using a massively parallel genome-wide technology in a legally mandated program which up until now has been much more narrowly focused on ~50 disorders for which there is general consensus of the value of the screening. Project 3 will utilize the clinical information from Project 2 and genome sequence information from Project 1.

considered when contemplating using a genotype-based screen derived from whole exome sequencing to replace a biochemical phenotype-driven screen. Project 2 will combine clinical information from the California NBS program’s database of follow-up information as well as patient records from the Primary Immunodeficiency Clinic/Jeffrey Modell Center at UCSF with the genotype information obtained in Project 1 to address critical Questions A, B and C as outlined in the RFA.

Question A in the RFA asks “*For disorders currently screened for in newborns, how can genomic sequencing replicate or augment (e.g. make more accurate, comprehensive or inexpensive) known newborn screening results?*” This question is addressed in Aim 2a of Project 2 of this proposal in which we will determine how well a genotype-based screen derived

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Overview References

References Cited

1. Wilson, J. M. G. & G. Jungner, G. *Principles and Practice of Screening for Disease*. (World Health Organization, 1968).
2. Frazier, D. M. *et al.* The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. *J Inherit Metab Dis* **29**, 76-85, doi:10.1007/s10545-006-0228-9 (2006).
3. Sweetman, L. Newborn screening by tandem mass spectrometry (MS-MS). *Clin Chem* **42**, 345-346 (1996).

COMPONENT PROJECT 1**Whole Exome Sequencing and Analysis of Variants in Newborn Blood Spots Relevant to Metabolic Disorders and Primary Immunodeficiencies****SPECIFIC AIMS**

Whole exome sequencing focuses on the important ~2% of the genome that contains the coding regions and immediate peri-exonic splice sites in the genome. Deep exonic sequencing utilizes massively parallel “next generation sequencing”, which is accurate for single nucleotide polymorphisms, somewhat less so for small insertions and deletions, and provides only scant information about large insertions or deletions. Furthermore, although it is less costly than whole genome sequencing, the variable efficiency with which various exonic fragments are captured means whole exome coverage is not complete and only provides information for ~90-95% of the coding genome. Nonetheless, whole exome sequence is a cost-effective method for identifying variants genome-wide in a critical subfraction of the genome.

Accurate identification of variants is only the first step in utilizing whole exome sequencing. Deducing the functional consequences of variants is an even greater challenge, especially for missense variants. Large exome sequencing projects have demonstrated that nearly 85% of all variants are rare and are unevenly distributed among different ethnic groups. Most rare variants have no functional annotation and interpreting them requires a variety of imperfect *in silico* methods to assess whether a variant is deleterious based on such parameters as how frequent the allele is in the population and whether the affected residue is conserved among orthologs and paralogs.

The California Department of Public Health Newborn Screening (NBS) Program holds a vast inventory of dried blood spots as well as up to 5 years of clinical follow-up information. We have identified a set of 1463 newborn blood spots with linked follow-up data as a resource upon which to base the research questions in Projects 2 and 3. In support of Project 2 SubAims 2a and 2b, we will carry out Whole Exome Sequencing (WES) from 1413 dried blood spots from individuals who either screened positive or negative, and were later confirmed either to have, or not to have, one of a group of 25 metabolic diseases for which the state uses tandem mass spectrometry (MS/MS) for screening. In support of Project 2 SubAim 2c, we will carry out WES of DNA extracted from 50 NBS dried blood spots from children followed in Immunodeficiency Clinic with one of the many primary immunodeficiencies that would not have been detected by the recently implemented T-cell excision circle (TREC) screen for neonatal lymphopenia. Specifically, we propose to:

1. Perform whole exome sequencing on each of 1357 NBS blood spots from patients identified by tandem mass spectrometry and confirmed as having one of 25 amino acid, organic acid and fatty acid oxidation disorders for which MS/MS was used for screening. Identify variants in genes that underlie the metabolic disorders. Perform WES on 13 samples from patients who were found to be false negative NBS as well as 200 control patients without metabolic disorders, 100 of whom had a false positive screen and 100 had a true negative screen, and examine them for variants in the same genes.
2. Perform whole exome sequencing on each of 50 NBS blood spots from patients identified in the UCSF Immunology Clinic with TREC+ primary immunodeficiencies and identify variants in genes that could have identified these patients as having a primary immunodeficiency at birth.
3. Implement secure data storage of sequencing and variant data.
4. Analyze the WES of the of 50 NBS blood spots from patients identified in the UCSF Immunology Clinic for alleles at one of the 9 non-HLA genes considered by expert opinion to be of substantial clinical utility by the Clinical Pharmacogenetics Implementation Consortium (CPIC) of the Pharmacogenomics Research Network.

SIGNIFICANCE

The precedents of screening for many disorders in parallel with MS/MS and for having a test that is genotype-based, as is done with cystic fibrosis, lead naturally to the question of whether massively parallel sequencing of the entire genome might ultimately be the most effective form of newborn screening for all disorders, including many disorders for which no disease-specific, phenotype-driven test is available. The falling cost of whole genome sequencing and its more limited but less costly surrogate, whole exome sequencing, has now created a new opportunity to study massively parallel sequencing of the genome as a method of newborn screening at the genotype level.

Whole exome sequencing focuses on only a fraction, albeit an important fraction, of the genome, the ~1% that contains the coding regions and immediate peri-exonic splice sites¹. Deep exonic sequencing utilizes massively parallel “next generation sequencing”, which is highly accurate for single nucleotide polymorphisms, less so for small insertions and deletions, and provides only scant information about large insertions or deletions. One might argue that the ENCODE project identifies non-coding regions of the genome that should also be examined, for example the 19.4% of the genome showing open chromatin or transcription-factor binding across all cell lines studied by the ENCODE investigators¹. We chose not to pursue whole genome sequencing because (1) it is more efficient to concentrate on what whole exome sequencing provides, given we are dealing with a trade-off between efficiency/cost *versus* completeness, (2) ENCODE regions are a minimum of 20 times larger, many fold more variable, and less susceptible to analytical programs and database analysis to assess the functional and pathogenic impact of variation, and (3) we are operating under the assumption, based on precedents, that these metabolic disorders are Mendelian disorders caused mostly by coding or splice site mutations of large impact, which whole exome sequencing will detect. Furthermore, although it is less costly than whole genome sequencing, the variable efficiency with which various exonic fragments are captured means whole exome coverage is not complete and only provides information for ~97.5% of the coding genome of our genes of interest (see below). Nonetheless, whole exome sequence is a cost-effective method for identifying variants genome-wide in a critical subfraction of the genome. Ultimately, with further reductions in price, WGS will likely replace WES in clinical application, but for the present and near future it is still cost prohibitive for widespread use.

INNOVATION

Although whole exome sequencing is not a new technology, combining the whole exome approach to sequencing with the very comprehensive biochemical and clinical data available through the California NBS program is a novel “marriage” of new and powerful technology with the largest newborn screening program in the United States. It will allow a direct comparison of the sensitivity and specificity of newborn screening by traditional testing *versus* whole exome sequencing.

APPROACH

We are very fortunate to have Drs. **Fred Lorey** and **Robert Currier** from the California Department of Public Health as “in-kind” collaborators. **Dr. Lorey** is Chief of the California Genetic Disease Screening Program and a distinguished figure on the national scene in newborn screening. He joined the Genetic Disease Branch, California Department of Health Services, in 1989 and has served with distinction in the state agency for over 20 years. Since 2003 he has been Member, American College of Medical Genetics/Health and Human Services Agency Newborn Screening Diagnosis and Follow-up Advisory Group. Since 1994, **Dr. Currier** is a statistician in charge of evaluation of the California Genetic Disease Screening Program within the California Department of Public Health. In this capacity he has been responsible for both statistical monitoring of the existing newborn and prenatal screening programs and development and implementation of new screening modalities. In 2011, he was appointed to the Steering Committee of the Newborn Screening Translational Research Network (NBSTRN). Drs. Lorey and Currier will work with us to identify blood spots, linked clinical data, and assure the integrity and completeness of the samples and data we will be obtaining from the state.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

We will choose ~1600 newborn blood spots from babies born in California between 2006 and 2011. These spots will be as follows:

- i. 1357 spots positive by tandem mass spectrometry (MS/MS) screening for one of 25 metabolic disorders and then confirmed to have the disorder (True positive screens). These constitute nearly the entire number of all infants born between 2006 and 2011 with positive newborn screen by MS/MS who were later confirmed to have either a FAOD, an OAD, or an AAD by biochemical testing and have up to 5 years of postnatal follow-up.
- ii. 100 spots chosen at random that were positive by tandem mass spectrometry (MS/MS) for one of 25 metabolic disorders but then confirmed not to have the disorder (False positive screens)
- iii. 13 spots negative by tandem mass spectrometry (MS/MS) for one of 25 metabolic disorders but then confirmed during follow-up to have the disorder (False negative screens)
- iv. 100 spots chosen at random that are negative by tandem mass spectrometry (MS/MS) for one of 25 metabolic disorders and then confirmed not to have the disorder (True negative controls)

In addition, we will choose 50 spots from patients known to have a primary non-SCID immunodeficiency (PID) ascertained through clinical assessment at the UCSF PID clinic. We therefore plan to sequence a minimum of $1413+13=1463$ blood spots for Project 1 over the course of the grant period.

DNA will be extracted from each dried blood spot (about 14mm) as follows. The spot will be punched six times with a 3.2mm paper punch. Six 3.2mm punches are placed into each of 6 wells of an Autogen DNA purification robotics system plate for DNA extraction. Our pilot experiments indicate the 6 punches provides ~900 ng, which is 4-5 times more DNA than the minimum required for whole exome sequence. Reprecipitation with ethanol gives high quality DNA with A260/A80 ratio = ~1.9-2.0 that has been successfully used to make libraries suitable for exon capture (Nimblegen).

The sequences of the standard Roche-Nimblegen oligonucleotide reagents used for exon capture were examined for coverage of all 316 exons (not just the coding exons) contained within all splice variants of transcripts from the 15 genes most closely associated with FAODs. The capture reagents contain elements corresponding to 308 out of 316, or 97.5%, of the exons for these genes. Of the 8 exons with no non-zero coverage, 5 were from a single gene, *ACADVL*. To avoid missing even the 2.5% of the exons not well captured, we will supplement the Nimblegen kit with additional oligonucleotides chosen to fill in for the missing exons.

The goal of doing whole exome sequencing is to identify, with great accuracy and efficiency, all DNA sequence variants found in all exonic regions of each subject. Thus, target enrichment has to be done one sample at a time, with great care to make sure that the entire exome is enriched greatly and evenly. The genomic DNA is sheared (by a precision sonicator) and assembled into a library with the adaptors containing barcodes. Size-selected DNA fragments are incubated with biotinylated RNA oligonucleotides corresponding to the entire human exome for 24 hours. The RNA bait-DNA hybrids are then selected out of the complex mixture by incubation with streptavidin-labeled magnetic beads. After the beads have been washed, the RNA baits are digested so that the only remaining nucleotides are the targeted DNA species of interest. After 14 cycles of DNA amplification, the captured sample is loaded onto the sequencer.

With a HiSeq2500 Sequencer from Illumina and using a new chemistry and a new analytical software package, up to 200 million reads can be obtained in each lane of the flow cell. At 100 bp per read and paired-end sequencing (sequencing both ends of each DNA fragment), the amount of DNA sequence generated for each lane is 2 reads X 100 bp/read/cluster X 200,000,000 clusters/lane. Assuming 80% of the reads passing the quality filter, there will be 37.2 Gb of high quality sequence/lane. Since the targeted DNA sequence for the exome is ~51 Mb, each lane has the capacity to sequence the exome >180 times even when 4 barcoded samples are sequenced on each lane. At >180X coverage, one can ensure that >80% of the bases in the targeted regions have a read depth of at least 20X (to account for the Poisson distribution of DNA fragments captured) and provide sufficient data for accurate DNA sequence variant discovery. At 4 samples per lane and 8 lanes per flowcell, and 2 flowcells per run, each sequencing run for the paired-end 100 bp reads takes 11 days. In other words, we can sequence 64 samples per run (equivalent to ~6 samples per day). While the sequencing protocol is routine and automated, it is important to place an optimal number of DNA fragments

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

onto each lane. We have extensive experience with mixing different samples for multiplexing and loading the correct amount of DNA to generate accurate sequence for variant detection.

This sequencing work will be done in the UCSF Institute for Human Genetics Genomics Core, which has extensive experience having performed and analyzed hundreds of whole exome sequences to date.

Sequence analysis is actually one of the most challenging tasks in this project. The first step of the exome analysis is base calling, including identification of small insertions and deletions (indels). The raw reads are aligned against the reference human genome using software such as BWA² or Bowtie2³, and sorted and indexed in a BAM file for processing by tools such as Illumina's CASAVA, Picard [<http://picard.sourceforge.net>] and GATK^{4,5} to reduce artifacts introduced by sequencing errors and to improve call accuracy. Manipulations include, for example, marking PCR duplicates, performing local realignment around known indel locations, and recalibrating scores based on covariates such as position in read and sequencing chemistry. Variant calls are then generated, currently using GATK. We then apply further evaluation (e.g., VQSR from GATK), and compare the sequence against our own local database of exomes from the UCSF Genomics Center to filter likely experimental artifacts. The resulting VCF file is the basis for further analyses.

We next will apply a battery of quality assessment metrics, including average coverage, total number of variants, number of novel variants, and transition/transversion ratio. Once convinced of the quality of base calls, we annotate the variants to enable prioritization of those likely to be associated with the disease phenotype. Variants will be identified, mapped, filtered and evaluated for functional significance. We tag variants according to whether they have been observed in the thousand genomes project⁶, in dbSNP or in the NHLBI GO Exome Sequencing Project (ESP) (<https://esp.gs.washington.edu/drupal/>) (noting if they are labeled as rare or clinically relevant). Variants are also annotated to indicate where they fall relative to the exon capture system (those not tagged are far richer in false positives), name and location within known genes (e.g., 3' UTR), defined coding and splicing impact (e.g., synonymous; GT/AC splice site change); and functional annotation for region, gene, and variant (e.g., OMIM⁷, HGMD (<http://hgmd.org/>), GeneTests⁸, and locus specific databases⁹). We have access to the commercial Ingenuity Systems Pathway Analysis database (Ingenuity Systems, Inc), which enables basic pathway analysis and access to their proprietary knowledgebase. Finally, we add information about predicted phenotypic pathogenicity (e.g., 'pathogenic' calls for non-synonymous amino acid changes from software prediction programs SNAP¹⁰, Align-Grantham Variation Grantham Deviation, multivariate analysis of protein polymorphisms [MAPP], MutPred, PolyPhen-2.1, Sorting Intolerant From Tolerant (SIFT), and Xvar¹¹ and probable impact on splice enhancers or other conserved motifs. Our selection of methods are directly informed by Dr. Brenner's leadership of the Critical Assessment of Genome Interpretation (CAGI) project [<http://genomeinterpretation.org>] that directly evaluates the ability of methods to predict the phenotypic impact of genetic variation.

With the large datasets we will be generating, a high-performance computer infrastructure and an effective and up-to-date bioinformatics pipeline that must be ready to ensure the generation of useful results. Furthermore, the data needs to be stored in a secure environment. Data for this project will be stored and analyzed in UCSF's secure MyResearch environment. This HIPAA compliant desktop environment is hosted on servers housed at the UCSF Data Center on Minnesota Street. The MyResearch environment is hosted on six Dell PowerEdge R710s and Five EqualLogic PS6100E SAN, which are located inside the locked rack. There are two layers of physically redundant Cisco firewalls that protect the servers and SAN. The MyResearch environment utilizes VM Ware View Virtual Desktop, which must be logged into using UCSF Active Directory credentials. The servers are locked inside a rack with a combination lock. The rack is located in a data center secured by two sets of locked doors. All sequence analysis will be carried out in a blinded fashion with regards to the phenotype of the individual whose sequence is being analyzed.

Data will be stored in the Principal Investigator's group network folder in the remote MyResearch environment where only the research team members are able to view the data sets and access is audited. This folder is physically located in a data store on the SAN in the locked rack. Network traffic between MyResearch and the UCSF campus network traverses a SSL VPN tunnel in encrypted format.

Once Dr. Koenig and her colleagues have completed Aim 1 of Project 3, we will have in hand a well-vetted consent procedure consistent with the participant protection framework for conducting whole genome/whole

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

exome sequencing during the neonatal period. At this point, we will expand the consent process used with the primary immunodeficiency patients to include requesting consent to examine other “incidental” or “secondary” variants not directly relevant to their immunodeficiency. In particular, we will analyze the WES of the 50 children from Immunodeficiency Clinic for alleles at 9 non-HLA genes considered to be of substantial clinical utility by the highest level of expert opinion as considered by the Clinical Pharmacogenetics Implementation Consortium (CPIC) of the Pharmacogenomics Research Network¹². These genes and the drugs whose metabolism are most affected are shown in Table 1.

Gene	Drugs Affected
CYP2C19	Clopidogrel
CYP2C9	Warfarin
CYP2D6	Codeine, tramadol, risperidone, tricyclic antidepressants, selective serotonin reuptake inhibitors
DPYD	5FU/capecitabine
G6PD	Rasburicase, Septra
SLCO1B1	Statins
TPMT	Thiopurines
UGT1A1	Irinotecan
VKORC1	Warfarin

Some of the important variants in these genes are not exonic single nucleotide variants that would be most readily targeted by the Roche-Nimblegen exon capture reagents. Instead, they occur in promoter regions, such as the *VKORC1* promoter variant (c.-1639G>A) and the insertion/deletion of (TA)₆/(TA)₇ in *UGT1A1* (*UGT1A1**28), or they are gene duplications, such as

the ultra-rapid metabolizing allele of *CYP2D6*¹³. Obtaining sequence for these variants using WES is problematic. First, a variant in a promoter as far as 1639 upstream from an exon, or a variant that consists of insertions/deletions of a simple dinucleotide repeat, or a variant caused by an entire gene duplication, are not well-captured or detected. Instead, we plan to target these particular alleles using direct PCR and sequencing or MLPA.

The variants in these genes obtained from WES, supplemented by the specific allele testing for variants not well detected by WES, will be used to generate genotypes from which pharmacogenetic phenotype, such as rapid or slow metabolizer or pharmacodynamics sensitivity can be inferred¹³. We will not always know phase and, therefore, it may occasionally be difficult to assign a pharmacogenetic phenotype.

Limitations

Because the capture probes focus only on exons, and the enrichment is not uniform, copy number changes are hard to detect reliably. That being said, in regions where the capture is uniform, deletions can be detected reliably when the read depth is good. Duplications are more difficult to detect. With paired end reads, one can occasionally detect the breakpoints but most of the “hybrid reads” are artifacts that do not map back to the genome and are disregarded. For that reason, we do not expect to have copy number assessment of large-scale indels (>100 bp) using exome sequence data because current methods are insufficiently robust, but we may incorporate this in the future.

Base calling algorithms are evolving rapidly, and we submit each new one to extensive validation and regression tests. We also use our own approaches alongside those provided by companies (e.g., Illumina) to better understand the error spectrum and hone it to be appropriate to this application.

Timeline: See Project 2 Timeline; the Timeline for Projects 1 and 2 have been combined given the highly integrated nature of these two projects.

REFERENCES (Component Project 1)

1. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
2. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).
3. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
4. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).
5. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-498, doi:10.1038/ng.806 (2011).
6. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-1073, doi:10.1038/nature09534 (2010).
7. OMIM; <http://www.ncbi.nlm.nih.gov/omim> (Johns Hopkins University (Baltimore, MD)).
8. GeneTest website; <http://www.ncbi.nlm.nih.gov/sites/GeneTests/?db=GeneTests> (NCBI)
9. Mitropoulou, C., Webb, A. J., Mitropoulos, K., Brookes, A. J. & Patrinos, G. P. Locus-specific database domain and data content analysis: evolution and content maturation toward clinical use. *Hum Mutat* **31**, 1109-1116, doi:10.1002/humu.21332 (2010).
10. Bromberg, Y. & Rost, B. SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res* **35**, 3823-3835, doi:10.1093/nar/gkm238 (2007).
11. Thompson, B. A. *et al.* Calibration of Multiple In Silico Tools for Predicting Pathogenicity of Mismatch Repair Gene Missense Substitutions. *Hum Mutat*, doi:10.1002/humu.22214 (2012).
12. Relling, M. V. & Klein, T. E. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther* **89**, 464-467, doi:10.1038/clpt.2010.279 (2011).
13. Whirl-Carrillo, M. *et al.* Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther* **92**, 414-417, doi:10.1038/clpt.2012.96 (2012).

COMPONENT PROJECT 2**Whole Exome Sequencing as a Newborn Screening Tool for Metabolic and Immunological Disorders****SPECIFIC AIMS**

Expanded newborn screening by tandem mass spectrometry (MS/MS) began in California as a pilot in 2002 and was instituted in 2005. The mass spectrometer is used to measure carnitine, acylcarnitines, and certain amino acids chosen for their positive predictive value for organic acid disorders (OAD), fatty acid oxidation disorders (FAOD) and amino acid disorders (AAD). Normal ranges are arbitrary and the cut-offs are generally set at mean+4 standard deviations. As with any screening test, a “positive newborn screen” must be confirmed by more specific biochemical testing before a diagnosis is made.

All newborns that screen positive by MS/MS and other NBS tests in California are given an initial appointment to see an appropriate specialist by the state Department of Public Health (DPH) follow-up coordinators. The coordinators also gather follow-up data, record the confirmed diagnoses, and collect annual summaries of clinical symptoms for 5 years. Furthermore, specialists are required by state law to inform the California DPH whenever they see a patient with a disorder for which NBS is performed; thus if there were a negative screen and months later a patient is diagnosed with one of the screened disorders, the specialist must inform the DPH and the DPH must record the child as a false negative or missed case. Thus, the California DPH has an accurate record of all true positive and false negative screens. Among ~3,100,000 newborns screened between 2006-2011, 1083 true positives for fatty acid oxidation disorders (FAOD) and organic acid disorders (OAD) and 274 true positives for amino acid disorders (AAD) were found and were followed for up to 5 years. In addition, the state learned of thirteen patients with FAOD and AAD that were “missed” by screening. These patients provide an excellent opportunity to measure the sensitivity and specificity of whole exome sequencing for 25 different FAOD, OAD and AAD currently screened for by tandem mass spectrometry.

Another issue facing NBS by MS/MS is the variability of the incidence of clinically significant, disease-causing abnormalities in individuals who screen positive and are later confirmed with some of the disorders for which MS/MS is used for screening, particularly FAODs and at least one OAD (LCHADD, MCADD, VLCADD, SCADD and 3MCC1D)¹. Many, but not all of these patients are asymptomatic. There is currently no method available that will predict which patients with these disorders will suffer metabolic decompensation and/or neurological dysfunction and which will not. The California Department of Public Health has follow-up data on which newborns become symptomatic and which do, thereby offering an opportunity to perform an association study between metabolic decompensation and modifier gene variants.

A third clinical research problem we will address is primary immunodeficiency. The most recent NBS test introduced in California is the TREC (T-cell receptor excision circle) assay for neonatal lymphopenia to screen for severe combined immunodeficiency (SCID). SCID, however, comprises only a small fraction of the total spectrum of all primary immunodeficiencies (PIDs). Yet, just as with SCID, there is obvious benefit in diagnosing the non-SCID PIDs soon after birth, before infections occur, in order to optimize survival and freedom from medical complications. Unfortunately, for the vast majority of these disorders, the only phenotype-based tests are expensive and different for the different disorders, making NBS for these disorders cumbersome. The PIDs are just one example of how a large number of disorders for which early intervention could benefit infant health might be approached through whole genome or whole exome sequencing in lieu of a battery of specific newborn screening tests.

Project 2 has three Aims:

Aim 2a Test the sensitivity and specificity of whole exome sequencing of newborn blood spots from newborns with or without positive MS/MS screens and later either confirmed or not confirmed to have one of 17 organic acid or fatty acid oxidation disorders (OAD/FAOD) or one of 8 inborn errors of amino acid metabolism.

Aim 2b Examine genes involved in fatty acid and amino acid metabolism for variation associated with the development of clinical disease in patients with one of five fatty acid oxidation and organic acid disorders well known to have low clinical disease penetrance.

Aim 2c Test the sensitivity of whole exome sequencing of newborn blood spots from patients ascertained through UCSF immunodeficiency clinic with an immunodeficiency that would not have been identified through current newborn screening based on the absence of T cell receptor excision circles (TRECs).

SIGNIFICANCE

The California NBS program is an outstanding public health program that has provided newborn screening, case-finding, and follow-up to the newborn citizens of California and their parents for decades. They have an excellent track record of collecting and inventorying NBS blood spots and ascertaining nearly complete clinical outcomes information through specialty clinics throughout California. Out of ~3,100,000 newborns screened between 2006-2011, disease-specific screening tests identified approximately 3600 babies who were subsequently confirmed to have hypothyroidism, galactosemia, biotinidase deficiency, cystic fibrosis, hemoglobinopathies, and congenital adrenal hyperplasia.

Expanded newborn screening by MS/MS began as a pilot in 2002 and was instituted in 2005. The mass spectrometer is used to measure carnitine, acylcarnitines, and certain amino acids chosen for their positive predictive value for organic acid disorders (OAD), fatty acid oxidation disorders (FAOD) and amino acid disorders (AAD). Normal ranges are arbitrary and the cut-offs are generally set at mean+4 standard deviations. As with any screening test, a “positive newborn screen” must be confirmed by more specific biochemical testing before a diagnosis is made.

In California, all newborns that screen positive by MS/MS and other NBS tests are reported initially to the newborn’s primary care provider and are given an initial appointment to see an appropriate specialist by the state Department of Public Health (DPH) follow-up coordinators. The coordinators also gather follow-up data, record the confirmed diagnoses, and collect annual summaries of clinical symptoms for 5 years. Furthermore, specialists are required by state law to inform the California DPH whenever they see a patient with a disorder for which NBS is performed; thus if there were a negative screen and months later a patient is diagnosed with one of the screened disorders, the specialist must inform the DPH and the DPH must record the child as a false negative or missed case. The California DPH specifically asks for hospital admissions and emergency department visits that are related to the “management of the disorder”. They also ask about the reason for the hospital admission, but this data is not always provided. Thus, the California DPH can provide an accurate list of all patients who have been confirmed to have a FAOD, OAD and AAD born since 2006 in the state, provide their newborn screening spots, and provide basic clinical follow-up for the first 5 years of life.

The positive predictive value of many DNA variants for causing disease is unknown. The significance of this study is that it will allow a careful assessment of the sensitivity and specificity of WES in a well-characterized set of patients for whom NBS screening results and clinical follow-up are available

Aim 2a

Between 2006 and 2011, California identified 1083 true test positives for FAOD and OAD out of ~8,100 initial positives and 274 true positives for AAD out of ~5,200 positive for amino acid disorders (AAD). The Program also recorded 13 false test negative patients who had had negative newborn screens but were later found to have a an FAOD or AAD. The total number of true-positive patients = true test positives + false test negatives = 1357+13=1370.

Despite the power and broad applicability of MS/MS for screening, using MS/MS to screen for genetic FAODs, OADs and AADs presents certain difficulties. Carnitine and acylcarnitine levels are critical analytes used in MS/MS to screen for carnitine transporter defects as well as OADs and FAODs. Abnormalities may arise due to deficiency of the carnitine transporter, prematurity², medications^{3,4} or other metabolic disorders, thereby reducing the specificity of NBS based on carnitine levels. Mothers who are carnitine deficient may have newborns with false positive results on biochemical screening because their newborns may appear deficient prior to significant dietary intake⁵. False negative NBS for carnitine deficiency also occurs⁶.

If WES were to supplant conventional NBS for nearly every NBS disorder as a cost efficient initial screen, it would be very important to determine how it compares to the conventional NBS in terms of sensitivity and specificity. Focusing only on the metabolic disorders in this proposal, the results from California’s MS/MS screening between 2006-2011 for FOAD, OAD, and AADs gives a

$$\text{sensitivity} = (\text{test-pos true-pos patients})/\text{alltrue-pos patients} = 1357/1370 = 0.99$$

and

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

specificity = (test-neg true-neg patients)/alltrue-negative patients = $\sim 3100000/3107000 = 0.9977$.

The high specificity meant that only about 8100 MS/MS screen positives needed to be followed up by biochemical analysis among the 3,100,000 newborns screened. Based on criteria described below, we propose to use different screening thresholds for declaring mutations pathogenic, and compare the sensitivity and specificity for each threshold. From those results, we can draw conclusions regarding the appropriate threshold to be adopted using WES for NBS.

Aim 2b

Another issue facing NBS by MS/MS is the variability of the incidence of clinically significant, disease-causing abnormalities in individuals who screen positive by MS/MS and are later confirmed to have a metabolic derangement with highly variable penetrance¹. This is particularly the case for FAODs and certain OADs.

- At one end of the spectrum are carnitine-acylcarnitine-translocase and CPT I deficiencies, which have high clinical penetrance and generally present early in life.
- At the other end of the spectrum is short chain acyl-CoA dehydrogenase deficiency (SCADD) and 3-methylcrotonyl-CoA carboxylase 1 deficiency (3MCC1D). In these conditions, the majority of newborns and infants have no clinical symptoms but a few percent may develop life threatening hypoglycemia and metabolic acidosis. SCADD is often referred to as a “non-disease”.

In between are:

- Isolated long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency (LCHAD) deficiency is treatable, but many patients suffer episodic hypoketotic hypoglycemia and rhabdomyolysis.
- Medium chain acyl-CoA dehydrogenase deficiency (MCADD) is particularly problematic. With treatment, the risk of death in MCADD is low and long-term outcome good. However, if untreated, 90% of patients are asymptomatic while 10% can suffer serious complications, even death, in childhood but are generally asymptomatic after age 6. These data suggest that 90% of patients being treated for MCADD could forgo any treatment without suffering any sequelae while the remaining 10% are at greatly increased risk, but we have no way of determining which patients do or do not need the treatment¹.
- VLCADD is another FAOD that causes early infantile death, but the condition is generally treatable. Nonetheless, approximately 10–20% of VLCADD patients found through newborn screening and started on treatment suffer episodic rhabdomyolysis while the remaining 80-90% never become symptomatic.

There is currently no method available that will predict which LCHADD, MCADD, VLCADD, SCADD and 3MCC1D patients will suffer metabolic decompensation and/or neurological dysfunction and which will not. WES offers an opportunity to examine the pathogenic severity of mutations at the primary locus as well as variation in other genes that might modify the phenotype and could be useful for predicting which patients with one of these disorders is at greater risk for metabolic decompensation and, therefore, requires more intensive therapy and follow-up.

Aim 2c

The most recent NBS test introduced in California is for severe combined immunodeficiency (SCID). SCID is a rare, but life threatening inherited disorder in which infants appear healthy at birth, but lack adaptive immunity provided by T and B lymphocytes. SCID can be caused by defects in any of 14 known, and additional unknown genes that impair lymphocyte development. SCID affected infants acquire serious infections early in life and do not survive unless they receive immune-reconstituting treatment, such as a hematopoietic cell transplant (HCT), enzyme replacement or gene therapy. Diagnosing SCID soon after birth, before infections occur, optimizes survival and freedom from medical complications. With no recognized family history in >80% of SCID cases, diagnosis of pre-symptomatic SCID requires population-based screening for T lymphopenia⁷. An effective screening test for neonatal lymphopenia using newborn blood spots was developed that is based on quantitating *T cell receptor excision circles* (TRECs), which are present in newly formed T cells, but essentially absent in the blood of infants with SCID, in whom T cell maturation is impaired⁸. A favorable evidence-based

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

review by the DHHS Secretary's Advisory Committee on Heritable Disorders of Newborns and Children in 2010 led to the recommendation to add SCID screening to the universal newborn screening panel in all states. Statewide screening for SCID with the TREC assay began in Wisconsin and is now underway in Massachusetts, California, New York, and additional states. By April, 2012, in >1.7 million screens, 29 infants with SCID were diagnosed, giving an incidence of ~1/60,000 infants (PIDTC, unpublished).

SCID comprises only a small fraction of the total spectrum of all primary immunodeficiencies (PIDs)^{9,10}. More than 100 genes are known in which mutations cause other forms of PID (See **Table 2**). Yet, the same arguments used for SCID concerning the value of early diagnosis soon after birth, before infections occur, in order to optimize survival and freedom from medical complications also apply to these non-SCID PIDs. Unfortunately, for the vast majority of these disorders, only cumbersome phenotype-based tests exist to make NBS possible. The UCSF Pediatric Primary Immunodeficiency Clinic is following >50 patients with non-SCID immunodeficiencies who could be approached for recruitment. Dried blood spots have been stored by the CA Dept. of Public Health (CDPH) since 1982 and are a unique resource, providing an opportunity to investigate whether these non-SCID immunodeficiencies could have been identified as part of a NBS screening program based on whole exome sequencing.

The PIDs are just one example of how a large number of disorders for which early intervention could benefit infant health might be approached through whole genome analysis in lieu of a battery of specific newborn screening tests.

INNOVATION

Using the whole exome approach outline in Project 1, combined with the large dataset available through the California Newborn Screening Program, Project 2 will allow a novel analysis of patients with MS/MS results for conditions that have poor genotype/phenotype correlation. This project also expands the capacity of NBS to identify children at risk from forms of immunodeficiency that are not severe combined immunodeficiency and for which there is currently no known newborn screening available. Given that these children will have many heterogeneous disorders, there is no single analyte or laboratory phenotype that can be used for this screening and, therefore, the application of WES is the only method available. Such a novel newborn screening method could serve as a model for many other disorders for which no single analyte or laboratory phenotype exists.

APPROACH

We are fortunate to have **Dr. Seymour Packman** as a consultant on Project 2. Dr. Packman has a distinguished career spanning more than three decades in research and clinical care of patients with metabolic disorders. He established a broadly based and productive program of clinical investigation and genetics education, embodied in the Neurometabolic Program and the Lysosomal Disease Center at UCSF. Clinical investigations under these programs have contributed to the genetics, biology, and treatment of diverse heritable metabolic disorders, including phenylketonuria, galactosemia, organic acidemias, and to the indications and outcomes of liver transplant for inborn errors of metabolism with uncertain natural history. He was among the first to delineate the two major forms of biotin-responsive multiple carboxylase deficiency, based on biochemical parameters at the cellular level. In this work, he contributed in a novel direction to our understanding of the role of acetyl-CoA carboxylase deficiency in the pathophysiology of these organic acidemias. He is the recipient of a Special Recognition Award from the March of Dimes Birth Defects Foundation for his work.

Aim 2a

We will identify four groups of NBS blood spots:

1357 true positive NBS spots by MS/MS with confirmed diagnoses of FAOD, OAD or AOD

13 false negative NBS spots by MS/MS with confirmed diagnoses of FAOD, OAD or AOD

100 false positive NBS spots by MS/MS with confirmation that FAOD, OAD or AOD is NOT present

100 true negative NBS spots by MS/MS with confirmation that FAOD, OAD or AOD is NOT present

The derived exome sequence for each individual will undergo base calling, mapping, quality assessment,

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

and variant annotation as described in Component Project 1. We will test each variant for consistency with the known inheritance pattern for each of the diseases for which NBS is applied, i.e., a recessive compound heterozygote must have two mutations in the same gene in the sample. In applying such tests, we are careful to account for base calling limitations, realizing that many bases may not be robustly called in every individual; this is very important to avoid excluding real disease variants. The determination of whether compound heterozygous or homozygous mutations are present in genes relevant to the 25 disorders screened by MS/MS will be made by individuals blinded to the NBS and confirmed diagnosis status of each sample.

		WES+	WES-
Confirmed Diagnosis +	NBS+ 1357	a	1357-a
	NBS- 13	b	13-b
Confirmed Diagnosis -	NBS+ 100	100-c	c
	NBS- 100	100-d	d

Whether or not a spot is considered "positive" by whole exome sequencing (WES+) or negative by whole exome sequencing (WES-) will depend on what we set as our threshold for considering any variant as pathogenic. We will calculate the values of a,b,c and d in four ways:

I) Very conservatively: Homozygosity or compound heterozygosity for two mutations that are known from public databases to cause the disease

II) Moderately conservatively: Homozygosity or compound heterozygosity for two mutations from Group I above or are likely highly deleterious (such as premature terminations, frame-shift and canonical splice site mutations)

III) Less conservatively: Homozygosity or compound heterozygosity for two mutations from Groups I and II above or are novel missense mutations judged to be damaging and, therefore, likely pathogenic by in silico methods, such as PolyPhen or SIFT.

IV) Least conservative (Most Aggressive): Homozygosity or compound heterozygosity for two mutations from Groups I and II and III above or novel mutations outside the canonical splice regions that might still interfere with normal splicing

We will then calculate sensitivity and specificity of whole exome sequencing by determining the values of a,b,c and d in a blinded fashion as follows: Sensitivity of WES = $a+b/1370$ and Specificity of WES = $(c+d)/200$. We will also examine the 13 patients whose NBS was falsely negative to see if WES would have picked up the relevant defective gene. Comparing sensitivity and specificity for each threshold will provide some guidance regarding appropriate thresholds that could be used in practice. For example, to obtain sufficient sensitivity, level IV above may be required; however, as a consequence, the specificity may drop, leading to a larger number of individuals requiring follow up testing by biochemical analysis to exclude false positives.

Aim2b

In this Aim, we are seeking genetic factors underlying phenotypic variation in four different FAOD diseases: LCHADD, MCADD, VLCADD and SCADD, and in one organic aciduria, deficiency of 3-methylcrotonyl-CoA carboxylase involved in leucine metabolism (3MCC1 deficiency or 3MCC1D). For each one of these diseases, we will classify individuals, within each disease entity, as "symptomatic" and "non-symptomatic" in a dichotomous fashion, based on whether they had made any emergency room visits because of their metabolic abnormality.

In the five year follow up of patients between infancy and age 5 identified through the California NBS program, the frequency of episodes of metabolic decompensation, as documented by emergency room visits or number of hospital stays reported by age 1 year, is shown below in Table 1.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Disease	Number of Patients Followed	Reports at 1 year of age	
		ER Visits	Hospitalizations
VLCADD	56	3	7
LCHADD	22	3	3
MCADD	170	11	11
SCADD	123	6	6
3MCC1D	94	3	3

There are several possibilities regarding genetic variation influencing disease penetrance. For a broad range of biochemical genetic disorders, such as enzymopathies, residual functional activity of the enzyme is highly predictive of disease course. For example, in Gaucher disease, a complete absence of enzyme activity is associated with an early onset and severe course, while a small level of residual enzyme activity associates with a later age of onset and milder course, including completely asymptomatic cases. Therefore, our first hypothesis is that clinical severity is dictated by degree of residual enzyme activity. We propose to grade mutations as either type 1 - complete loss of activity, for example represented by frameshift, truncation or splice mutations, versus type 2 - partial loss of activity, for example represented by certain amino acid substitutions. For the latter, we will still grade the likely severity of loss of function based on the criteria described above. Analysis will focus on comparing the frequency of mutations in category (1) versus category (2) between cases and controls, as defined above.

Our second line of attack will focus on common genetic variation outside the primary disease locus. Here we will perform a genome-wide association study, comparing allele frequencies between cases and controls. We will prioritize other genes in the FAO pathway or other relevant pathways. Because we are testing only exonic variants, the number of tests performed will be considerably less than in a typical GWAS, and so we can employ a somewhat less stringent significance level than the conventional 5×10^{-8} ; the P-value will be determined by the number of SNPs tested. We will focus on SNPs with minor allele frequency of 5% or greater. Here we hypothesize that common variants in pathway genes determine clinical outcome.

Our third hypothesis is that rare variants in candidate pathways determine clinical outcome, and hence we propose to examine rare variants in an association analysis with clinical severity. Here, a mutational load or burden test will be employed, where for each gene in the appropriate pathway we will compare the mutational burden between cases and controls. Again, P-values will be determined by the number of tests performed. While a genome-wide approach may be attractive for both the GWAS and rare variant approaches, we are mindful of the sample sizes we will be dealing with, and hence reduced power by requiring more stringent P-values for larger scale testing. In any event, both the GWAS and rare variant analyses may need to be considered preliminary, requiring further replication and validation.

For analysis of hypothesis 1, we will cluster the genotypes at the causal locus into subgroups based on type: two type 1 mutations (described above) versus one type 1 and one type 2 mutation versus 2 type 2 mutations. To do so, we need a characterization of predicted total or partial loss of function based on mutation. While some mutations will be straightforward to characterize as non-functional (e.g., frameshift truncations), others, such as amino acid substitutions, will be more complex. For these cases, we will apply an analysis of features used to train machine learning methods that predict deleterious or functional variants using human allele data (i.e., PolyPhen or our MutPred method). We will also include features that describe location in gene and location relative to domains in the protein product. These features will be ranked based on the Area Under the Curve of a Receiver Operator Characteristics plot (ROCAUC). Based on experience with other metabolic disorders such as enzymopathies, we will then cluster genotypes into two groups – those with two type 1 mutations, who presumably have no functional gene activity, and those with one or two type 2 mutations, who presumably have some low level of residual gene activity. We will then compare the frequency of the two different genotype groups between the cases and controls by Fisher exact test. Because the analysis involves a single test with a directional alternative hypothesis, we will employ a one-sided test with a P-value of .05 for significance.

For analysis of hypothesis 2, we will perform a SNP association study using the computer program Plink. Genes relevant to the FAOD pathway will be ranked using several candidate gene prioritization approaches including GeneMANIA (for example, see Table 3 in the Appendix) and GENIE. GeneMANIA will evaluate

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

candidates using a semi-supervised approach with known FAOD genes as input and will classify based on its functional network. The score is a similarity score that is reported by the GeneMANIA algorithm. They are not probabilities, but just values that are used to rank the genes. The ranking itself reflects how close the genes are to the input set of genes (the known FAO genes) based on the GeneMANIA algorithm. The first one is closest, with respect to the biological networks that GeneMANIA uses (protein-protein interaction networks, co-expression networks, etc.) then the second in the list and so on. There is no specific cut-off as the scores are not p-values. GENIE will evaluate a text mining approach that will identify genes most closely associated with FAOD concepts in PubMed abstracts. The two resulting lists will be averaged to come up with a final ranked list for evaluation. The number of genes to be included in the final analysis is limited due to statistical power issues related to multiple testing. Therefore, we propose to test at most 100 SNPs due to power considerations, as described below. For these association analyses, we propose to use a 2-sided test with a P-value of .0005, which allows for 100 independent tests by a Bonferroni correction.

For analysis of hypothesis 3, we will form the union of all the mutations in each gene and perform a "burden" test¹¹. The P-value to use would be 0.05 divided by the number of genes tested. For the burden test, we will consider an individual to be in an "at risk" group if they are homozygous or compound heterozygous for mutations considered to be dysfunctional. Evaluation of functionality will be based on application of PolyPhen, MutPred and SNAP. In addition, we will build and evaluate a new supervised classification approach for FAOD mutations using MutPred and PolyPhen attributes, highly ranked genes from the previously described pathway analysis, and trained with neutral and disease-causing missense substitutions present only in those genes. Classification will be performed using Random Forests and will be evaluated using Leave One Out (LOO) cross-validation. The resulting performance will be visualized using the ROCAUC and will be compared to a similarly computed MutPred ROCAUC not trained on the FAOD gene variants.

Power Calculations: Regardless of which of the 3 hypotheses we are testing, we can create two groups defined as "genotype present" or "genotype absent". For hypothesis 1, as described above, those with two predicted non-functional alleles would be considered "genotype present" and the remaining subjects with at least one reduced function allele as "genotype absent". For hypothesis 2, we would consider individuals carrying at least one variant in one of the genes on the candidate gene prioritization list formed by the intersection between GENIE and GeneMania lists as "genotype present" (i.e. dominant model). For hypothesis 3, we would consider any mutation in one of the genes on the candidate gene prioritization list as "genotype present" (dominant model). We will also examine a recessive model for hypotheses 2 and 3, whereby an individual must possess two deleterious alleles in a gene to be included in the "genotype present" group.

All power calculations are based on a Fisher Exact test for a 2x2 contingency table, as follows:

	Symptomatic	Not symptomatic
Genotype present	x	y
Genotype absent	a-x	b-y

Where a+b is the total number of patients being followed, a is the number of patients with a diagnosis with an emergency room visit or hospitalization for metabolic decompensation, b is the number of patients carrying this diagnosis who never went to the emergency room or hospital during follow-up, x is the number of symptomatic patients who also had "genotype present," and y is the number of asymptomatic patients who are "genotype present."

For hypothesis 1, we assume a one-sided Fisher Exact test with a P-value of 0.05. We calculate the difference in frequency of the predisposing genotype group ("genotype present") between symptomatic and asymptomatic individuals ("cases" versus "controls") necessary to obtain 80% power given the number of cases and controls for each disease. For hypotheses 2 and 3, we perform a similar power analysis, but in this case use a P-value of 0.0005 for a 2-sided test, assuming we perform 100 independent tests. The results are in the Table 2 below. Given the sample sizes, the required group frequency difference for hypothesis 1 varies across disorders from 0.42 for MCAD to 0.78 for LCHAD. These frequency differences are not at all unreasonable to expect if level of residual gene function is highly predictive of clinical symptomatology. For hypotheses 2 and 3, because of multiple testing, the required group frequency difference is larger but still manageable (except for LCHAD where the same size is too small to obtain statistical significance).

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Table 2. Required group frequency difference to obtain 80% power to reject the null hypothesis in favor of hypothesis 1, 2 or 3.		
Disease	Hypothesis 1	Hypothesis 2 and 3
VLCAD	.53	.78
LCHAD	.78	NA
MCAD	.42	.62
SCAD	.54	.78
3MCC1	.64	.93

Limitations

We recognize that the number of samples from symptomatic patients is small, which would reduce our power to detect associations. We are proposing this analysis as a pilot since we recognize the power is low and there is no replication sample. However, if we can find any evidence for association, it might justify a larger national consortium of data sets to increase numbers and collect a replication sample.

Aim 2b only addresses the issue of why some patients with a FAOD develop symptoms during the first few years of life. We will not be capturing information on which patients remained totally well throughout infancy and childhood and only developed symptoms later in adult life.

It is clear that dietary management, maintaining calorie support, avoiding fasting and metabolic stress, are important factors in preventing decompensation for fatty acid disorders. It is distinctly possible that the genetic contribution to metabolic decompensation is rather small and most of the symptomatic patients became sick because of environmental factors. In this case, we would not be likely to find any modifying genotype.

Aim 2c

We will apply WES to identify mutations in genes responsible for immunodeficiencies that are unlikely to be detected by the current T-Cell receptor excision circle assay (i.e. are TREC+) but nonetheless warrant early diagnosis and treatment similar to what is currently provided for severe combined immunodeficiency. These are primarily non-SCID PIDs but also includes late onset SCID patients whose T-cell lymphopenia is not marked enough at birth to give a positive TREC test. The derived exome sequence for each of 50 individuals with PID will undergo base calling, mapping, quality assessment, and variant annotation as described in Project 1. As discussed above, we will enroll infants and children with TREC+ immunodeficiencies (PID). We will obtain informed consent for each of these patients and if minors, their parents, and obtain their newborn blood spot for DNA extraction and whole exome sequencing. We will perform standard TREC assay to confirm the TREC-based screen would have missed the patient as birth. We will recruit 10 patients per year for whole exome sequencing. The sequence will undergo base calling, mapping, quality assessment, and variant annotation as described in Project 1. We will test each variant for consistency with the known inheritance pattern for each of the diseases for which NBS is applied, i.e. a recessive compound heterozygote must have two mutations in the same gene in the sample. In applying such tests, we are careful to account for base calling limitations, realizing that many bases may not be robustly called in every individual; this is very important to avoid excluding real disease variants. The determination of whether compound heterozygous or homozygous mutations are present in the 110 genes relevant to the non-SCID PID disorders (see Supplementary Table 2) will be made by individuals blinded to the NBS and confirmed diagnosis status of each sample.

Timeline: Project 1 will continue to generate sequence and prepare lists of variants throughout the 5 year period. We will choose blood spots from the children with fatty acid oxidation disorders on whom there is clinical follow-up in the first two years so the analysis for modifiers in Aim 2b can be done. We will also recruit and sequence the 50 children with primary immunodeficiencies in the first 2 years so the pharmacogenetic data will be available to Project 3. Project 2, Aim 2a, will require all of the sequencing be completed in order to measure sensitivity and specificity, although some preliminary results should be generated in years 4 and 5.

	Year 1	Year 2	Year 3	Year 4	Year 5
Project 1 Generate WES and Analyze for Variants	X	X	X	X	X
Project 2 Aim 2a				X	X
Project 2 Aim 2b	X	X	X		
Project 2 Aim 2c	X	X			

REFERENCES (Component Project 2)

1. Wilcken, B. Fatty acid oxidation disorders: outcome and long-term prognosis. *J Inherit Metab Dis* **33**, 501-506, doi:10.1007/s10545-009-9001-1 (2010).
2. Shah, M. D. & Shah, S. R. Nutrient deficiencies in the premature infant. *Pediatr Clin North Am* **56**, 1069-1083, doi:10.1016/j.pcl.2009.08.001 (2009).
3. Makino, Y., Sugiura, T., Itop, T., Sugiyama, N. & Koyama, N. Carnitine-associated encephalopathy caused by long-term treatment with an antibiotic containing pivalic acid. *Pediatrics* **120**, e739-741 (2007).
4. Rasmussen, J., Nielsen, O. W., Lund, A. M., Kober, L. & Djurhuus, H. Primary carnitine deficiency and pivalic acid exposure causing encephalopathy and fatal cardiac events. *J Inherit Metab Dis*, epub ahead of print (2012).
5. Leydiker, K. B. *et al.* Maternal medium-chain acyl-CoA dehydrogenase deficiency identified by newborn screening. *Mol Genet Metab* **103**, 92-95, doi:10.1016/j.ymgme.2011.01.011 (2011).
6. Lee, N. C. *et al.* Diagnoses of newborns and mothers with carnitine uptake defects through newborn screening. *Mol Genet Metab* **100**, 46-50, doi:10.1016/j.ymgme.2009.12.015 (2010).
7. Puck, J. M. The case for newborn screening for severe combined immunodeficiency and related disorders. *Ann N Y Acad Sci* **1246**, 108-117, doi:10.1111/j.1749-6632.2011.06346.x (2011).
8. Puck, J. M. Laboratory technology for population-based screening for severe combined immunodeficiency in neonates: the winner is T-cell receptor excision circles. *J Allergy Clin Immunol* **129**, 607-616, doi:10.1016/j.jaci.2012.01.032 (2012).
9. Al-Herz, W. & Notarangelo, L. D. Classification of primary immunodeficiency disorders: one-fits-all does not help anymore. *Clin Immunol* **144**, 24-25, doi:10.1016/j.clim.2012.05.003 (2012).
10. Al-Herz, W. *et al.* Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. *Front Immunol* **2**, 54, doi:10.3389/fimmu.2011.00054 (2011).
11. Lee, S. *et al.* Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* **91**, 224-237, doi:10.1016/j.ajhg.2012.06.007 (2012).

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

COMPONENT PROJECT 3

ELSI implications of Research Related to DNA-based Analysis Associated with Newborn Screening SPECIFIC AIMS

The history of newborn screening (NBS), the first concrete application of modern genetic knowledge to public health, has been intimately tied to technological innovation. Expanding genetic knowledge and technology continues to transform the scope and breadth of state-mandated screening programs, from testing for an expanding list of Mendelian single gene disorders, through the advent of tandem mass spectrometry, and most recently, into the era of rapid, inexpensive whole genome analysis. Screening programs were initially structured around a simple premise: if identified early, inborn errors of metabolism could be ameliorated, dramatically changing clinical outcomes. **The clear benefit justified a public health approach built on universal, population-wide screening.** Parents were rarely informed and certainly not asked to provide informed consent, which was seen as unnecessary in the face of clear public health benefit.

The public health paradigm came into question when tandem mass spectrometry enabled screening for many conditions that could not be readily or fully treated; states responded variously, some adding multiple conditions to NBS panels, some taking a more conservative approach. The utilization of whole genome analysis (WGA) to replace NBS, or to improve its accuracy, will radically transform that calculus. **Cheap and accurate WGA of newborns will reveal unanticipated findings, including predictive risk of late-onset disorders, multiple variants of unknown significance, incidental information of potential benefit to the health of parents or other family members, as well as information pertinent to the parents' reproductive decision making.** Some have argued that harmful or unwanted information could simply be "masked" in WGA. Even if accepted, such a solution would require careful discussion about which gene variants should be returned, and how. Inevitably, clinicians and public health professionals will be challenged to identify variants revealed by WGA that meet traditional NBS criteria. **We will illuminate that discussion by focusing on one likely category of variants as a case study: pharmacogenomic (PGx) predictors of drug response.**

Joining together empirical, legal, and policy (normative) studies, we will pursue the following aims:

1) Through a collaboration with the UCSF Human Research Protection Program, we will develop a participant protection framework for conducting whole genome/whole exome sequencing during the neonatal period, as an adjunct to the standard NBS blood spot. Families experiencing severe primary immunodeficiency who are offered WGA of their child's NBS blood spots (Project 2, Aim 2c) will serve as the exemplar for discussion. The framework will include:

- Procedures for parental informed consent
- Strategies for determining and then reporting actionable incidental findings (including PGx)
- Approaches to findings that affect the personal health or reproductive decisions of parents
- Policies for managing late-onset conditions identified in the neonate

Best practices developed as part of this research would then be shared throughout the University of California (UC) system via the UC BRAID efforts to establish harmonization among human subjects protection procedures at the 5 UC campuses that include academic health science centers.

2) Using PGx variants that predict response to drugs commonly used in childhood as a case example, we will determine the views, perspectives, and value preferences of key stakeholders about the potential expansion of newborn screening programs with the advent of WGA. Empirical methods will include:

2a focus groups with pregnant women/prospective parents from throughout California and parents of children with severe primary immunodeficiency who are offered WGA of NBS blood spots

2b focus groups with pediatricians and obstetricians/gynecologists

3) In collaboration with the UCSF/UC Hastings Consortium on Law, Science and Health Policy, identify the legal and constitutional issues for using WGA, and for incorporating PGx into public health NBS programs.

3a Analyze the constitutional foundations of newborn screening programs and the reach of police power.

3b Conduct a fifty state survey of state laws and public health department regulations that govern existing newborn screening panels to identify the factors that determine inclusion.

4) Informed by Aims 1-3, and in collaboration with a "Policy Advisory Board" convened by The Hastings Center (New York), develop and disseminate policy recommendations for expanded NBS programs based on WGA. Using PGx variants as one exemplar, the Board will develop an ethical framework for determining which variants revealed by WGA should be included (or excluded) as part of a state-run newborn screening program, identifying the barriers to implementation and factors likely to promote appropriate use.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

SIGNIFICANCE

Technological advances are increasing the ability to analyze large amounts of genomic information for diagnosis, risk assessment and treatment decision-making quickly and accurately. Whole genome analysis (WGA) by whole exome or whole genome sequencing will become more widely available as costs continue to decline and new service providers come on the market.¹⁻⁴ As WGA technologies become more accessible, opportunities to integrate genomic information into public health and clinical care will expand.^{1,5}

The long held promise of genomic medicine is the potential to increase quality and efficiency of care, that WGA can “transform screening and diagnostic testing.”^{1,6,7,10} For that potential to be reached, the appropriate use of genomic technologies must be specified. There are concerns that the usefulness and clinical utility of WGA may be “oversold.”^{1,11}

We need evidence and critical reflection to manage expectations and guide the appropriate adoption of WGA. For genomic information to be incorporated into public health and clinical settings in a fashion that maximizes public benefit, we must anticipate the ethical, legal and social issues that may arise.¹² One pending clinical application of WGA is newborn screening (NBS).^{13,14} Although the uptake, perceived value, impact on behaviors, ethical and social consequences of genomic profile testing provided directly to the consumer have been studied, the translation of WGS into public health programs like NBS is not well understood.¹⁵⁻¹⁷ This study will address that gap by examining ethical, legal, and social implications of the application of WGA to newborn screening.

Existing public health newborn screening programs vary by state, but with current tandem mass spectrometry technologies, states include about 29 of 60 possible conditions in their panels.^{14,18-20} WGA provides the means to greatly expand the conditions screened for, not to mention the potential to reveal “incidental” or “secondary” findings.²¹ Studies have explored parents’ attitudes towards expansion of newborn screening panels.²²⁻²⁴ Although population screening for disorders with proven treatments was widely supported, attitudes toward screening for untreatable conditions or conditions that may present later in life were mixed and more controversial. Concerns about potential ethical and social issues associated with genetic profiling in general and NBS in particular, include informed consent, confidentiality, potential for discrimination, reporting of incidental or non-actionable results, false negative and false positive results, and over diagnosis^{13,14,22,25-27}

One potential benefit of using WGA in NBS is the capability to screen for additional disorders/conditions. **This project will use pharmacogenomic (PGx) testing as a case study to explore the ethical, legal and policy ramifications of using WGA to expand state NBS panels.** Pharmacogenomics is the measurement of genetic variants associated with drug response. Drug treatments in many cases are ineffective: either patients do not adequately respond to the treatment or they experience adverse drug reactions.⁹ In addition to the incalculable costs associated with the use of ineffective treatments, adverse drug reactions or events (ADR/ADEs) are estimated to cost well over \$100 billion and cause 100,000 deaths per year in the United States^{9,28-30} It has been estimated that 4.7 % of hospitalizations are the result of an ADE³¹ and that 6.5% of inpatients suffer an ADE during their hospitalization.³¹

ADR/ADEs in pediatrics, although less well studied, may be even more profound. There is wide variation in drug response in pediatric patients leading to opportunities for overdosing or underdosing causing toxicity or loss of efficacy³². Many ADEs occur on the first use of a drug. Medication side effects or accidental overdoses account for over half a million pediatric outpatient visits per year, particularly in children aged 4 and under.³¹ It is estimated that 4.3% of admissions to pediatric hospitals in children less than 2 years old are due to ADEs. Overall, risk of ADRs is higher in infants.³³ Up to 16% of all outpatient prescriptions are associated with an ADE. Lack of efficacy and ADEs are frequently cited as reasons for non-adherence to medications in children. In pediatric psychiatry, side effects or lack of efficacy are a major cause of non-adherence.³⁴

Some ADEs are fatal. There is increasing evidence documenting a genetic impact on treatment response.

Pharmacogenomics (PGx): The use of genetic information on drug transporters, drug metabolizing enzymes, and drug receptors to facilitate optimal choice of drugs and dosages.⁹

Although drug response is influenced by a variety of factors including environment, diet, age, lifestyle and state of health and interactions with other drugs; germ line genetic variants – mainly in genes encoding drug metabolizing enzymes, drug transporters, drug targets and human-leukocyte antigen (HLA) are reported to be responsible for many of the observed inter-individual differences in drug efficacy, the risk for ADRs, or both.⁸

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Genetic factors explain 20–95% of variation in drug response.²⁹ For example, cytochrome P450 enzymes (such as CYP2D6) are involved in the primary metabolism of commonly-used drugs such as codeine. If an infant or child is an ultra-rapid metabolizer, morphine toxicity, including mortality, may result. (For breast-feeding infants, the mother's PGx status is also a factor.) There are other examples of critical PGx variants in child health. Carbamazepine, a commonly used anti-seizure medication, causes severe Stevens-Johnson Syndrome in individuals with particular HLA variants. Some variants occur more frequently in certain geographic populations, raising concerns about targeting by race/ethnicity.

The classic demonstration of the critical importance of PGx variants involves thiopurine methyltransferase (TPMT), which catalyzes the S-methylation of thiopurine drugs such as 6-mercaptopurine and azathioprine. Before testing became routine, children with leukemia often had severe ADEs following chemotherapy, including death. Clearly, in the situation of cancer therapy, specific testing would be conducted at the point of care, a result from a NBS panel would have little use. **However, for healthy children exposed to common drugs, PGx information may well rise to the level of clinical actionability. Parents and clinicians may desire such information. If variants are revealed during WGA-enabled NBS, should they be reported?**

In summary, non-targeted genomic profiling applications such as WGA are becoming clinically available before we fully understand how they should be integrated into clinical practice, or what impact they will have on the delivery of care, costs and outcomes. It's unknown which technologies should be implemented or how to most effectively and efficiently integrate them into existing practices and programs. These questions become even more critical when screening populations such as newborns and their families, who could live with the ramifications of WGA information their entire lives.^{20,22,25,35}

INNOVATION

Understanding of stakeholder perspectives is an essential first step in explaining how emerging technologies are disseminated, implemented and used.³⁶⁻³⁸ The aims of this proposed study contribute toward filling the gap in our understanding of the translation of WGA into public health NBS programs by: 1) assessing the current legal environment, 2) evaluating policies in place for the protection of participants in genetic research, and 3) ascertaining the views of key stakeholders. Joined together, the information generated by this project will aid the development of robust policy recommendations. We have a clear plan for dissemination, including extensive use of new media.

Our proposed focus on a particular case example—PGx variants predicting response to drugs commonly used in childhood—is innovative. To our knowledge no NBS programs or clinical testing programs currently return PGx information that is relevant to the well-being of healthy infants or children, those who may be exposed to drugs used routinely such as codeine or medications for asthma or seizures. **This focus will allow us to be concrete when gathering empirical data** from stakeholders about what constitutes clinical utility. It also aligns with other national efforts in translational genomics, such as the NHGRI eMERGE network, which have concluded that selected PGx variants are actionable, and thus most appropriate to integrate into clinical practice—and the medical record—at this time. In addition, **our plans to create a research participant protection program in collaboration with our regulatory and compliance colleagues at UCSF, and then expand that work throughout the UC system are also highly innovative.** Other innovative features of the proposal are noted aim by aim below.

APPROACH

Strengths of the Project 3 Team and Prior Studies: Project Director (PD) **Barbara Koenig, a medical anthropologist**, has considerable experience leading interdisciplinary bioethics research endeavors linked to genetic testing policy. Her current and past work focuses on “return of results” of unexpected or incidental genetic findings,^{21,39} problems that are directly relevant to the proposed project on NBS. She is currently funded by NIH (R01CA154517) to investigate—using mixed ELSI methods—return of clinically actionable genetic findings to family members of research participants. As part of the eMERGE network (U01HG004599), she participated in ELSI investigations linked to ongoing scientific investigations, a model similar to the proposed U19 consortium network.⁴⁰⁻⁴² At UCSF, Koenig oversees the Clinical and Translational Science Institute's “Research Ethics Consultation Services,” and chairs the “Responsible Conduct of Research Committee.” Koenig has also participated in national policy development, having served on the Secretary's Advisory Committee on Genetic Testing, which recommended expanded oversight of the introduction of novel tests. Currently, she is a member of the NHGRI Return of Results Consortium, established to guide NIH-supported research using deep sequencing technologies in research and in clinical care.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Co-Investigator **Julie Sakowski is Associate Professor in the Department of Clinical Pharmacy and the Executive Director of the UCSF Center for Translation and Policy Research on Personalized Medicine (TRANSPERS)**, which provides a unique resource for our assessment of PGx variants and newborn screening. Sakowski offers the perspective of a health economist. She has extensive experience assessing stakeholder attitudes and preferences using a variety of methods, including focus groups. Sakowski's work on the policy and economic evaluation of pharmacogenomics, genetic testing, and implementation of emerging technologies is directly relevant to the proposed project.^{43,44}

Co-Investigator **Jaime King is the Associate Director of the UCSF/UC Hastings Consortium on Law, Science and Health Policy**. She has conducted an extensive analysis of the ethical, legal, regulatory, and policy dilemmas raised by non-invasive prenatal genetic testing, with a focus on the implications of WGA prior to birth, including policy frameworks for regulating genetic tests.^{45,46} She has also conducted a 50 state survey of informed consent laws, demonstrating her methodological expertise.^{47,48} In the context of noninvasive prenatal genetic testing, King developed a new model for informing the family about prenatal genetic testing using WGA.⁴⁶ Her most recent work focuses on the tensions between individual reproductive autonomy and public health goals, examining the constitutional implications of advances in perinatal genetic testing. The legal issues raised by noninvasive prenatal whole genome sequencing overlap considerably with those in expanded NBS.^{49,50}

Our project team is enhanced by collaboration with the Hastings Center (Garrison, NY). Josephine Johnston is a lawyer and bioethicist whose work has been focused on controversies surrounding the use of biomedical technologies, including genetic testing. She is currently PI of a March of Dimes Foundation-funded project to investigate the contribution of fertility treatments to the multiple birth rate and recommend changes to practice and policy to reduce the risks associated with multiple births. Previously she was Co-I of a NIMH-funded project to investigate the controversies surrounding the use of medications in children with emotional and behavioral disturbances, and related domains of pediatric psychiatry.

Erik Parens, trained in philosophical analysis, has investigated contentious issues involving child health. He has been the PI on 3 NIH-funded projects: on the controversies surrounding the disabilities rights critique of prenatal genetic testing; on the behavioral genetics research; and most recently on the use of psychopharmacology to treat emotional and behavioral disturbances in children. His work considers the impact of advances in genetics on our concepts of health and disease, the rights and responsibilities of parents making decisions for and about children, and the impact that health and social policy, as well as cultural attitudes, have on clinical and parental decision-making. Relevant to expanded NBS is his book *Wrestling with Behavioral Genetics: Science, Ethics, and Public Conversation*.⁵¹ He also has considerable experience developing fair and nuanced analyses of the ethical and social implications of new biomedical technologies that take account of public input and perspectives, an essential issue in NBS.

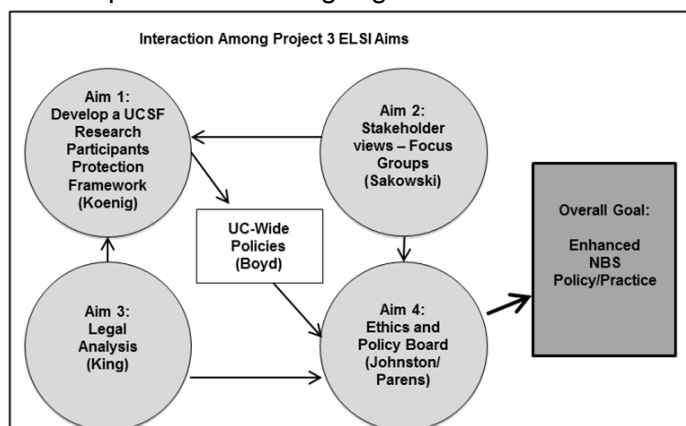
Hastings Center NBS Policy Expertise: The Hastings Center conducted a large study of newborn screening policy, culminating in several policy reports and an edited volume, *Ethics and Newborn Genetic Screening: New Technologies, New Challenges*.⁵² Funded by NHGRI (R01HG002579) and led by Thomas H. Murray, this study included in-depth ethical analysis in four critical areas of newborn screening: fairness in the distribution of the costs and benefits of screening; information, consent, and privacy; consultation and decision-making; and race, ethnicity, and socioeconomic status. Their analysis pointed out the limitations of existing programs, offering standards for best practice. **Johnston and Parens are experienced in conducting interdisciplinary bioethics research on contested and complex ELSI issues, including identifying the various perspectives and expertise needed to understand a complex issue, convening interdisciplinary teams, creating agendas that unpack and explore the various issues at play, leading group discussion and debate on these issues, identifying areas of consensus and ongoing debate, and synthesizing what is learned** at these meetings as well as through an analysis of the literature, to create readable and relevant scholarship containing, where appropriate, specific policy and practice recommendations.

The **UCSF Center for Translational and Policy Research on Personalized Medicine (TRANSPERS)** examines the translation of genomic technologies, including PGx variants into clinical care and health policy. Based in the UCSF School of Pharmacy, the center conducts interconnected projects on utilization, preferences, cost-effectiveness, evidence development, diverse populations, and health policies. TRANSPERS addresses new and evolving challenges in building and analyzing the evidence base for PGx by focusing on clinical applications of emerging genomic technologies such as WGA.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Several UCSF and University of California (UC)-wide activities also support this application. Specifically, UCSF's Clinical and Translational Science Institute (CTSI), a member of the Clinical and Translational Science Awards (CTSA) consortium, includes a dedicated focus on enhanced clinical research methods, including human subject protection. CTSI, which takes a broad-based approach to improving clinical research, has a goal of harmonizing human subjects protection across UCSF and the UC system. Elizabeth Boyd, Associate Vice Chancellor, Ethics and Compliance, oversees UCSF's Human Subjects Protection Program (HRPP). A sociologist, she is also the Program Director for the Regulatory Knowledge and Support Program in CTSI. Boyd, in cooperation with the director of the UCSF Human Subjects Protection Program (HSPP), John Helden, will participate in efforts to establish policies governing the use of WGA in neonates and children. Inclusion of Boyd and Helden will assure translation of our policy efforts into practice.

Another major strength of this proposal lies in its ability to leverage the power of the UC system wide "Biomedical Research, Acceleration, Integration & Development" program (UC BRAID), whose leaders recognized that shared UC governance facilitates development of a strong regional research network. (UC BRAID is described in "Resources.") These five national CTSA awardees created UC BRAID to enable coordination, standardization and systematic innovation in clinical and translational research. The UC Office of the President (UCOP) provides financial and in kind support to BRAID initiatives, including establishment of reciprocal IRB agreements across the five campuses. IRB harmonization demonstrates UC-wide cooperation in creating standardized human subjects protection practices; the commitment to addressing informed consent and related policies fills a major gap in local and regional policy.



Aim 1: Through a collaboration with the UCSF Human Research Protection Program (HRPP), develop a participant protection framework for conducting whole genome/whole exome sequencing during the neonatal period.

Overall approach: With the UCSF HRPP, we will develop a participant protection framework for conducting whole genome/whole exome sequencing during the neonatal or early childhood period when such testing is done as an adjunct to (or replacement of) the standard newborn screening blood spot. Families with a child experiencing severe primary immunodeficiency who are offered WGA of their child's blood spot (as part of Project 2 Aim 2c) will serve as the exemplar for discussion. Currently, UCSF has no standard policies and procedures for using WGA with neonates or children. **Offering WGA to families of parents with immunodeficiency will generate many serious questions and concerns, including data sharing, and how to manage uncertain information or findings that cascade through the family.**

We anticipate that the framework we recommend will include:

- Procedures for parental informed consent or child assent for those over age seven
- Strategies for determining and then reporting actionable incidental findings (including PGx)
- Approaches to findings that affect the personal health or reproductive decisions of parents
- Policies for managing late-onset conditions identified in the neonate or child
- Consideration of NIH data-sharing policies in NBS research using WGA

Best practices developed locally at UCSF as part of this research would then be shared throughout the UC system via the UC BRAID efforts to establish harmonization among human research protection procedures at five UC campuses that include academic health science centers.

Participants: Aim 1 will involve the UCSF ELSI and clinical investigators (Koenig, Sakowski, King, Nussbaum, Puck) in collaboration with the HRPP leadership and staff. Additional stakeholders, such as representatives of patient advocacy groups devoted to children with immunodeficiency, or disability rights advocates, will be included as needed. We will work directly with the UCSF HRPP Policy Group. (Committee membership is described under "Resources.") This committee meets quarterly and serves as an advisory group to Associate Vice Chancellor, Ethics and Compliance, Elizabeth Boyd. It is a standing committee composed of the Chairs

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

and Vice-Chairs of each of the four UCSF IRB panels, Associate Vice Chair, Ethics and Compliance, the Associate Vice Chancellor, Research, the HRPP Director and Associate Directors, UCSF Legal Counsel, the Deputy Associate Chief of Staff for Clinical Research at the VA, and other UCSF faculty members as appointed by the Associate Vice Chancellor, Ethics and Compliance. Others may be invited to or request to attend a particular meeting to provide additional expertise or perspectives.

The UCSF HRPP group develops and recommends new policies for the protection of human subjects and human research at UCSF. Policies will be developed using an iterative approach, informed by the emerging empirical and conceptual literature. Revised data sharing, informed consent and return of results policies will be presented to the HRPP leadership group, reviewed, and revised again. Before policies are finalized and promulgated, they are reviewed and approved by UCSF Legal Counsel with input from UC General Counsel as needed.

Procedures for Specific Aim 1: During the first quarter of the project, PD Koenig will meet with HRPP leadership to finalize implementation plans, including identification of a subcommittee (see Heldens letter of support). The group will review existing processes and procedures for genetic testing and genomic analysis at UCSF and in the UC system. The group will also consider the special informed consent issues of retesting NBS blood spots (reporting of incidental findings, late onset disorders/protecting the right “not to know,” information of reproductive concern to parents, *etc.*). The issue of “actionability” will be modeled through the discussion of specific PGx findings as an exemplar. The group will be informed of recommendations from the NHGRI “Return of Results” Consortium. We will also make available bibliographic resources developed for the legal analysis in Aim 3 and the policy work in Aim 4. During this time, the research assistant will identify practices at other academic institutions and recommended practices by professional organizations.

In addition, results from the Aim 2 focus groups and Aim 3 legal analysis, as well as other emerging evidence about participant and family preferences for return of results emerging from ongoing NHGRI WGA consortia, will inform the development of specific UCSF policies. We will develop and implement policies and procedures to create an exemplary system for requesting participation in WGA for neonates and children at UCSF. As a final step, Vice Chancellor Boyd will present findings to UC BRAID for consideration across the UC system (see Boyd letter of support).

Advantages and Innovation for Aim 1—Policy Translation: Geographic, structural, and cultural diversity of the five UC medical campuses has resulted in individual IRBs that operate in an uncoordinated fashion, approving a wide range of consent documents that emphasize legalistic concepts such as compensation for subject injury and liability. Rarely is the authentic voice of the potential research participants included in the participant protection framework. The current processes are variable, administratively inefficient, and may introduce added cost and delay the onset of clinical and translational research. Our project is unique in its attempt **to implement through policy** best practices for a participant protection process that is ethical, legal and meaningful to participants. This approach will serve as a regional—all UC campuses are subject to state NBS statutes—and potentially a national model.

This project is innovative in its ability to leverage state-wide resources (through UC BRAID harmonization goals) to implement a single set of policies and procedures across multiple universities to enhance WGA research. It also incorporates significant and meaningful empirical data from stakeholders (prospective parents, families with immunodeficiency, pediatricians, and OB/GYNs) throughout the process.

Potential Challenges and Limitations: The primary challenge is to gain consensus among the various stakeholders in the UCSF HRPP and harmonization of participant protection procedures among diverse campuses. We mitigate this issue in several ways: first, the UCSF HRPP and all five campuses are motivated to find common solutions to the challenge of WGA (see Boyd and Heldens letters of support). Second, UC BRAID leadership has already been working to bridge IRB procedural differences among campuses. Finally, the inclusion of empirical data, legal analysis, and interactions with the national return of results consortium (through PD Koenig’s participation) will ensure that a full range voices and perspectives are consistently heard in policy discussions.

Expected Outcomes: We expect to develop specific recommendations that will translate first into UCSF practice and second into system-wide policies and procedures. Those policies will result in a harmonized, coordinated, ethical approach to WGA in neonates and children across UC campuses.

Aim 2: Using PGx variants that predict response to drugs commonly used in childhood as one case example, we will determine the views, perspectives, and value preferences of key stakeholders about the potential expansion of newborn screening programs with the advent of WGA.

To meet Aim 2, we will conduct focus groups with (2a) pregnant women/prospective parents without prior experience of genetic disease, and parents of children with severe primary immunodeficiency who are offered WGA of NBS blood spots, along with (2b) pediatricians and obstetricians/gynecologists.

Overall Approach: The objective of this aim is to gather information on potential stakeholders' understanding of the implications of expanded NBS panels based on WGA, comprising preferences for the conduct of NBS research, including informed consent; information sharing within the family; data sharing/protection of information; thresholds of actionability, and opinions about how screening information using genomic technologies might best be incorporated into clinical practices. There have been several studies that have examined preferences toward expansion of NBS panels and attitudes toward genetic testing of newborns.^{22,23,26,38,53,54} However, **there is little information on stakeholder views of using WGS in public health NBS programs and scant evidence on public and physician attitudes toward routine sharing of information from expanded NBS, such as PGx data predicting drug response during early childhood. This study will fill that gap.**

Aim 2 will compile qualitative, empirical information on stakeholder views of WGA in newborn screening and preferences about the potential implications of expanding NBS to include return of potentially actionable PGx variants. We will use PGx as one exemplar to probe participants views about implementation of an expanded NBS panel made possible because of new technologies. This case will be discussed in the context of current NBS practice, which includes screening for rare diseases, such as those where diagnosis resolves uncertainty and ends diagnostic odysseys, but effective treatments are not currently available. This study will gather data through focus groups with two key stakeholder groups: parents of children undergoing NBS and physicians. Focus groups will be used because they enable the gathering of a large amount of information from groups of people in relatively short periods of time.⁵⁵ In addition, they allow for interaction among participants to clarify views and identify areas of agreement/disagreement. Drs. Sakowski and/or Koenig will lead all the groups.

Participants and Rationale: Focus group participants will include: 1) prospective parents contacted during the third trimester of pregnancy, and who will thus be experiencing NBS in the near future, and 2) the parents of children who have been identified, as part of Project 2, Aim2c, as having severe primary immunodeficiency, and who thus are offered WGA of newborn blood spots as a means of improving NBS technology. Physician focus group participants will include both OB/GYNs and pediatricians—the clinicians whose practices might be most immediately affected by the availability of expanded NBS information. Understanding stakeholder knowledge and views of the technology and its ethical implications is a critical first step to anticipate decisions to adopt and implement a technology.

Specific Procedures—Parents: We will conduct five focus groups with parents in which we will explore participants' knowledge of NBS (what they know and how they know it); their level of interest in expanded NBS information made possible by WGA; and their hesitations or concerns about using WGA, including PGx information for NBS as part of a mandated public health program. Existing studies have assessed attitudes toward the use of genetic testing in the clinical setting and expansion of newborn screening. The themes and topics identified in those studies will serve as the starting point of discussion for our focus groups. Factors previously identified to influence attitudes about PGx and genetic testing generally include: analytic validity and potential benefits of genomic diagnosis; testing process and what happens to test information (return of results, privacy, etc.).^{56,57,58} In addition, these focus groups will help provide a baseline understanding of participants' knowledge of newborn screening that will provide insights into stakeholder awareness of current practices and identify opportunities for future communication, education and outreach about NBS. The open-ended nature of the questions in focus groups will enable unanticipated topics to arise and to be explored.⁵⁵

Recruitment: **We will conduct four focus groups with prospective parents** (n=32-40) (pregnant women and partners, when appropriate) in their third trimester of pregnancy and **one focus group with parents of children diagnosed with severe immunodeficiency** (n=8-10). To increase generalizability to the California population, participants for the four prospective parent focus groups will be recruited from the obstetrical clinics at UCSF, UCSF East Bay clinics (Oakland, CA), University of California Davis, and University of California Los Angeles. Subjects will be recruited so that focus groups mirror the race/ethnicity, age, and income distribution

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

of the California newborn population. We will also seek to include a mix of NBS-naive new parents and parents with previous children who have prior experience with the NBS process. Parents of children with immunodeficiency will be selected at random from UCSF clinic patients identified by Co-I Puck.

Recruitment of patients will be conducted on site via posting of notices, announcements during pre-natal and Lamaze classes, and referrals from participating physicians. Assistance with recruitment will be provided by the UCSF Department of Obstetrics, Gynecology, and Reproductive Sciences (see Kuppermann letter of support). Parents of infants and children with immunodeficiency will be recruited at UCSF with the assistance of Co-I Puck. We will also work with UCSF's CTSI Participant Recruitment Services consultants to develop an optimal recruitment strategy and materials. Informed consent will be obtained immediately prior to the focus group. Each focus group will consist of 8-10 participants and last approximately 120 minutes. Meals will be provided during the focus groups and participants will be compensated for their time and expenses.

Research Questions: The focus group discussion guides will be designed with the assistance of the full investigator team and the Ethics & Policy Advisory Board. Potential research questions are in the following table:

- What is the public's understanding of and familiarity with current NBS practices and policies in California?
- Do parents have concerns about research use of NBS samples as technology advances?
- Are parents interested in results of expanded NBS, including PGx variants?
- If so, how would they like PGx information reported?
- What are parent's concerns about the use of genomic information from their children? (e.g., collection of genetic information by the state; privacy; reproductive freedom)
- What notification or consent processes, if any, would parents prefer?
- What information would parents prefer "not to know"?
- How would parents like genomic information to be reported? (e.g., primary physician, directly)
- Do parents with children diagnosed with a severe disease differ in their views of NBS and expanded NBS panels from other parents?

Specific Procedures—Physicians: We will conduct three focus groups to explore: physicians' familiarity with current NBS processes, their attitudes toward using WGA for NBS, what changes would be needed to NBS processes should WGA be adopted, the clinical impact of expansion of NBS to include reporting of PGx variants, and how PGx variants could be implemented in health care delivery practices. We will conduct one group with pediatricians, one with OB/GYNs, and one mixed group. Each focus group will consist of 8-10 participants (n=24–30, half pediatricians, half OB/GYNs) and last approximately 120 minutes.

Recruitment: Recruitment of physicians will be conducted via email or phone call to physicians practicing in clinics in the San Francisco Bay Area and the nearby central valley, in order to gain a mix of perspectives and experience, including urban/rural and academic/non-academic settings. Recruitment of her OB/GYN colleagues in California will be facilitated by Dr. Elena Gates (see letter of support) and pediatricians by Co-I Puck. Informed consent will be obtained immediately prior to the focus group. Meals will be provided during the focus groups and participants will be compensated for their time and expenses.

Research Questions: Focus group discussion guides will be designed with the assistance of the full investigator team and the Ethics & Policy Advisory Board. Potential research questions are in the following table:

- What are California physicians' understanding of and familiarity with current state NBS practices and policies?
- What are physician concerns about the use of WGA-informed NBS information for research and clinical care?
- What role do physicians believe they should play in NBS notification or consent processes?
- What tools would physicians need to effectively communicate PGx information to their patients?
- How important is clinical utility in determining return of results in WGA; does PGx information meet that bar?
- How could PGx be incorporated into existing work flows and decision support tools?
- How should information be managed within the family (e.g., potential reproductive implications, reporting of late onset conditions)?

Analysis and Interpretation of Focus Group Data: Focus groups will be audio-recorded and fully transcribed. We will use standard qualitative data analysis techniques based in grounded theory for analysis and interpretation of the focus group data.^{59,60} Grounded theory is an analytic approach commonly used in the social sciences to analyze textual data.⁶¹ This iterative process of inductive analysis involves examining and re-examining the facts and meanings contained in the data to develop successively more refined ideas about domains of interest. Transcripts will be analyzed with the aid of a qualitative data analysis software program (such as N-Vivo or Atlas-ti).⁶² Two researchers will independently code and analyze the data. The PD, Co-I Sakowski, and the RA will meet regularly to discuss coding and come to consensus on any discrepancies. Initially, coding and analyses will proceed according to the conceptual framework for this study; subsequently the

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

coding scheme will evolve so as to incorporate unanticipated themes and findings. We will develop coding memos to summarize interim findings, and as the basis for discussion with the rest of the research team.⁶⁰

Advantages and Innovation—Empirical Data: Formative focus groups provide highly valid qualitative information that is useful for exploring new topics that are poorly understood, such as the early dissemination and use of emerging technologies in research and clinical practice. The data gathered will inform policy development and provide useful information for later, larger scale studies and surveys.

Challenges and Limitations: If certain groups of informants prove difficult to recruit into focus groups, for example the parents of ill children with a rare disease who might not be able to assemble as a group, we will, if necessary, substitute semi-structured stakeholder telephone interviews with individual parents.

Expected Outcomes: We will compile qualitative, empirical information on stakeholder views of WGA in newborn screening and preferences about 1) the potential implications of expanding NBS to include actionable variants, and, 2) human subjects protection practices. A manuscript describing the focus group findings will be published as a free standing article in the special supplement to the *Hastings Center Report* summarizing the findings of our full project. The information gathered will enable us to explore social and ethical implications of the use of WGA in expanded NBS programs. The anticipated outcome of the focus groups will be a list of the key themes and examples of questions to be explored and considered when developing policy recommendations.

Aim 3: In collaboration with the UCSF/UC Hastings Consortium on Law, Science and Health Policy, identify the legal and constitutional issues surrounding the use of WGA to incorporate PGx into public health NBS programs.

Even if incorporating WGA information into newborn screening protocols is feasible, acceptable for stakeholders, and desirable as a public health mandate, the method of incorporation must be both constitutional and consistent with state laws. State newborn screening programs screen over 4 million newborns each year for genetic and metabolic disorders.⁶³ While all fifty states have a newborn screening program, the programs vary widely in the conditions currently screened for, as well as the requirements for adding a gene variant to the newborn screening panel.^{64,65} States also differ in whether participation in newborn screening is mandatory, opt-out, or entirely voluntary. The ease with which testing for PGx variants can be incorporated into state newborn screening protocols will depend on specific features of each PGx variant, as well as the state newborn screening laws.

From a constitutional perspective, **the states derive their ability to impose mandatory newborn screening programs from the Supreme Court decision in *Jacobson v. Massachusetts*, which established the states' ability to impinge the autonomy of individuals in certain situations to protect the health of the public.**⁶⁶ This ability arises from the police power granted to the states by the Constitution. State laws often grant public health authorities the power to compel individuals to submit to medical testing in instances where testing will benefit the individual or may prevent the spread of disease. However, **compulsory medical testing can also compromise an individual's autonomy and invade his or her privacy, and therefore cannot constitutionally be required without sufficient justification.**

At the state level, in 2002, the Maternal Child Health Bureau of the Health Resources and Services Administration commissioned the American College of Medical Genetics (ACMG) to gather evidence on the effectiveness of newborn screening, delineate the best evidence for screening for specified conditions, and to develop recommendations for a uniform newborn screening panel. ACMG conducted its evaluation and issued its report, but since that time a uniform panel has not been adopted.⁶⁵ We aim to update and expand upon ACMG's examination of state programs to develop a unified framework for incorporating PGx variants into NBS programs. However, any proposed expansion of the variants included in a newborn screening panel necessitates an examination of the specific characteristics of the condition being tested for, the bounds of the state police power to compel individuals to subject their newborns to genetic testing, and the existing state laws that govern the state's NBS program. In addition, **the use of WGA raises significant constitutional questions regarding the ability of the state to place limits on the information returned to parents, due to First Amendment protections for speech, which must be accounted for in this context as well.**

Overall Approach Aim 3

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

In Aim 3, **we will conduct legal research and analyze relevant constitutional precedent, state and federal case law, public health department regulations, and academic literature that govern existing newborn screening programs to inform the development of our framework for incorporating PGx variants into NBS programs.** Our legal analysis will begin with a broad examination of the constitutional requirements for any mandatory public health initiative and then move into a more granular investigation of each state's NBS statute and its implementing regulations. This narrowing approach will enable us to estimate the minimum constitutional requirements for incorporating testing for a specific gene variant with PGx significance in a mandatory NBS program and then explore the implications of those requirements for specific state programs with mandatory, opt-out, and voluntary participation. We will next use this information to inform the development of our framework for determining which PGx variants to include in mandatory newborn screening protocols and which would only be included in voluntary testing programs.

Procedures for Specific Aim 3

We propose three legal research strategies:

a. **Conduct legal research and analysis regarding the constitutional foundations of newborn screening programs and the reach of the police power.**

Research will be conducted on the relevant constitutional precedents, state and federal case law, and academic literature via various electronic databases including LexisNexis, WestLaw, Social Science Research Network, and HeinOnline regarding the use of the state police power to mandate medical procedures for public health purposes and the First Amendment protections afforded to content-based speech. This research will include an in-depth examination of individuals' Fourteenth Amendment rights to prevent the state from interfering with their bodily integrity, as well as their First Amendment right to access information discovered during WGA. **Professor King's legal team will analyze this material and synthesize it into a white paper describing the constitutional boundaries of mandatory newborn genetic screening tests, the factors that would help determine whether a test for a particular PGx variant would be eligible for inclusion into a mandatory newborn screening panel,** and what limits, if any, the state can place on access to information uncovered by WGA testing.

b. **Conduct a fifty state survey of state laws and public health department regulations that govern existing newborn screening panels to determine the factors that determine inclusion.**

For each of the fifty state newborn screening programs, we will examine the state statutes identified by the National Conference of State Legislatures and administrative regulations that govern which conditions can be tested for through the program in an effort to identify the range of factors states deem relevant to adding new conditions to newborn screening panels.⁶¹ As part of this subpart, we will also examine the protocol for administering newborn screening tests, specifically whether they are required of all newborns: what exceptions, if any, exist; whether consent is obtained; and whether newborn screening is discussed with before or after delivery. Finally, we will analyze regulations surrounding optional genetic testing panels offered in the prenatal or neonatal context for conditions not included in the newborn screening panels to determine the characteristics of conditions and factors that states consider appropriate for neonatal testing, but not mandatory testing. Research will be conducted through electronic databases, such as LexisNexis, WestLaw, Social Science Research Network, and HeinOnline, as well as contacting state public health departments directly in instances where the protocol for expanding or modifying state newborn screening panels was left up to the discretion of the administration. All conversations with state public health department officials will be taped for further analysis. Professor King's legal team will analyze our research findings and draft a white paper that identifies the relevant factors for both incorporation of a particular genetic test into a NBS panel and inclusion in a mandatory, opt-out, or voluntary screening panel.

c. **Incorporate these legal factors into the development of the framework for determining which PGx variants to include in state-run newborn screening programs.**

We will present our findings from the research above and an initial legal framework at the first meeting of the Ethics and Policy Advisory Board during project year 2. Following the interdisciplinary critique, we will refine our analysis. The final **legal analysis will be presented as a review article to be published in two formats:** 1) as a free standing article in the special supplement to the Hastings Center Report summarizing the findings of our full project, and, 2) as a free-standing law review article.

Advantages and Innovation—Legal Analysis: Legal analysis of both the broad constitutional requirements of the First and Fourteenth Amendments as well as the more specific state statutes and public health depart-

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

ment regulations provides a comprehensive examination of the legal restrictions on NBS programs. While this specific examination focuses on the incorporation of PGx variants, the analysis provides a framework that is relevant to regulating of WGA more broadly, and implementing other public health genomics initiatives.

Challenges and Limitations: First and Fourteenth Amendment jurisprudence is often unsettled and ill-equipped to address technological advances and the societal change. As a result, our constitutional legal analysis will be largely normative. However, nearly all legal scholarship on the constitutionality of developing technologies faces similar issues, and as a result has the opportunity to influence future Supreme Court decisions.

Expected Outcomes: We will produce two white papers for presentation to the Ethics and Policy Advisory Board. After comments and suggestions from the Advisory Board, two papers will be drafted for publication – one for the *Hastings Center Report* summarizing our findings and our framework, and a second more in depth law review article.

Aim 4: Informed by Aims 1-3, and in collaboration with an “Ethics and Policy Advisory Board” convened by the Hastings Center, develop and disseminate policy recommendations for expanded NBS programs using WGA.

Overview: Using PGx variants as one exemplar of potentially “actionable” findings revealed by expanded NBS, the Board will develop an ethical framework for determining which variants revealed by WGA should be included (or excluded) as part of a state-run newborn screening program, identifying the barriers to implementation and factors likely to promote appropriate use. The Board will include key scientists and clinicians from Projects 1 and 2, as well as the full team of Project 3 ELSI investigators. Formed during Year 1, the Board will hold two face-to-face meetings on the UCSF campus (Years 2 and 4). A “white paper” summarizing Board deliberations, along with Project 3 Aims 1, 2, and 3 summary papers will be published as a special supplement of *The Hastings Center Report*, and disseminated through well-established procedures for maximum policy impact.

Project Leadership: The Hastings Center team, including Josephine Johnston and Erik Parens, will lead the Ethics & Policy Advisory Board, in close collaboration with Project 3 Director Koenig and PI Nussbaum. They will convene **the Board described in the table below, which will convene quarterly by teleconference, and in person in Years 2 and 4:**

<u>Ethics & Policy Advisory Board</u>		
Natasha Bonhomme	Genetic Alliance	Patient Advocacy/Maternal-Child Health
Jeff Botkin, MD	University of Utah	Pediatrics/Genetics/Human Subjects Protection
Ellen Clayton, JD, MD	Vanderbilt University	Law/Pediatrics/Bioethics
Mildred Cho, PhD	Stanford Ctr. for Biomedical Ethics	Biology/Genetics Policy/Bioethics
Josephine Johnston, JD	Hastings Center	Law/Bioethics
Eric Juengst, PhD	University of North Carolina	Philosophy/Bioethics
Jaime King, JD, PhD	UC Hastings College of Law	Law/Public Policy
Barbara Koenig, PhD	UCSF	Medical Anthropology/Bioethics/Empirical Bioethics Methods
Pui-Yan Kwok, MD, PhD	UCSF	Human Genetics/Genotyping Technology Development
Frederick Lorey, PhD	California Department of Public Health NBS program	Biological Anthropology/Human Genetics
Robert Nussbaum, MD	UCSF	Medical Genetics
Erik Parens, PhD	Hastings Center	Philosophy/Bioethics
Kathryn Phillips, PhD	UCSF	Health Economics/Policy/Pharmacy
Jennifer Puck, MD	UCSF	Pediatrics/Genetics/Infectious Disease
Neil Risch, PhD	UCSF	Genetic Epidemiology/Human Genetics
Julie Sakowski, PhD	UCSF	Health Economics/Policy/Pharmacy

Searchable Bibliography: To support the project group’s analysis, Johnston will, beginning in year 1, work with the ELSI team to create a searchable bibliographic database of medical, scientific, ethical, legal and policy literature relevant to expansion of NBS and the use of PGx results in children. This database will be available to all members of the ELSI project, and will inform the project group’s work throughout the project. The data base will highlight: 1) the expansion of NBS; 2) the family consequences of identifying genetic variants in children, and 3) the identification of PGx variant information. Relevant works will be identified systematically using computerized literature searches (PubMed, PsychINFO, Philosopher’s Index, Sociological Abstracts, Google-Scholar, and LexisNexis) and by tracking citation chains in relevant papers, books, and reports. Search strategies will be modified based on initial results and searches will continue until no new relevant results are identified. Sources will be catalogued in a searchable database, which will be housed on a web-based platform, so

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

that it is available to all members of the U19 team. It will inform the project Ethics & Policy Advisory Board's work throughout the project and aid the planning of the two project meetings by Johnston, Parens and Koenig.

Project Meetings: In consultation with the full ELSI team—Parens, Koenig, Sakowski, and King—Hastings Center PI Johnston will lead and oversee all planning required for the project meetings in years 2 and 4, including identification and articulation of meeting goals, selection of invited speakers, creation of detailed meeting agendas, and assignment of responsibility for presentations at the meetings and/or for leading discussions on particular issues. In addition to Ethics & Policy Advisory Board members, who will all attend the project meetings, we anticipate inviting 2 additional speakers to each meeting to address issues that cannot be adequately covered by the group or to represent perspectives not represented on the Board (for instance, a pediatric PGx expert, and the perspective of disability rights).

The main purpose of project meeting 1 is to identify and discuss the ELSI issues relevant to expanded NBS, using the case example of PGx variants that predict response to drugs commonly used in childhood. At this meeting, we will also hear reports from Aims 1, 2, and 3—the UCSF efforts to develop a participant protection framework, the empirical studies of stakeholder views, and the legal analysis—and will offer comments and advice. The main purpose of project meeting 2 is to discuss the final report in detail, including the draft of the Board's policy recommendations. We will also hear updates from Aims 1, 2, and 3. Johnston and Parens will chair the project meetings and moderate the discussion. They will ensure that the meetings are faithful to their goals and agendas, and that discussion and debate are productive and, where possible, conclusive. Meeting presentations and discussion will be recorded and transcribed by the Center's Research Assistant. The Hastings team will then analyze the transcript to identify themes, specific concerns, and specific suggestions. Identified themes, concerns and suggestions will be further studied and analyzed after the meetings, and used to plan the final report and draft the Ethics and Policy Advisory Board's policy recommendations.

The Hastings Center has organized this kind of project meeting for over 4 decades, and has found that it allows complex issues to be analyzed from different perspectives—legal, medical, biological, philosophical, socioeconomic—in order to find common ground. (Additional details about Hastings Center can be found in "Resources.") Two facts about this kind of meeting—that it is interdisciplinary and that it entails face-to-face interaction⁶⁷—make it especially well-suited to analyzing complex issues of interpretation and for producing clear understandings. It is an improvement over scholars or researchers working in isolation or relying only on the published literature, and attendees learn more than at an academic conference because the partial view of each participant, limited by their disciplinary perspective and their experience, is challenged and thoroughly discussed, rendering everyone's understanding of the issues more full and complex. The NIH Roadmap supports interdisciplinary research,⁶⁸ and studies of interdisciplinary interaction support the efficacy of the Hastings Center method.⁶⁹ Klein and other leading authorities on interdisciplinary projects and processes have written about the intellectual power that a multi-perspective approach brings to the production, critique, and dissemination of knowledge.⁷⁰⁻⁷³

Final Report: Johnston and Parens will work closely with Project 3 Director Koenig and overall PI Nussbaum to plan and draft the project's final report. The report will consist of a summary white paper containing policy recommendations, plus 4-5 papers which together will address the ELSI issues raised by the advent of NBS expanded through WGA. The case study of PGx variants will inform the reporting. The white paper will lay out the ELSI challenges and include the Ethics & Policy Advisory Board's full recommendations. The remaining papers will be commissioned by Johnston, Parens and Koenig from project members. They will deal with specific issues, such as the constitutional analysis and the focus group findings. Since the final report will aim to identify and address a range of ELSI issues, the papers will be commissioned to build on one another and present a comprehensive analysis of the issues, including representing divergent views.

Subject to peer review, **the final report will be included in the *Hastings Center Report* for publication as a free-standing supplement, mailed with the *Hastings Center Report* and made available from the Report's web site on the Wiley Online Library, and also available independently in hard copy and as a pdf from Hastings' own web site.** *Hastings Center Report* supplements, known as Special Reports, are peer reviewed and indexed in Pubmed and other databases. Readership includes roughly 2,000 institutional and individual subscribers, plus roughly 3,000 libraries that will receive *Hastings Center Report* content through a Wiley collection, and additional readers through database licensing arrangements Wiley has entered into with EBSCO and JSTOR. Access in the developing world will be possible through the International Network for the

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Availability of Scientific Publications (INASP) and the Health InterNetwork Access to Research Initiative (HINARI)—estimated at around 3,500 libraries.

Expected Outcomes—Dissemination: **The Hastings Center will disseminate the final report to policymakers and journalists through its Public Affairs and Communications Department.** Created in 2007 following a capacity building grant from the Ford Foundation, this department is now a pioneer in bioethics communication, reaching thousands of journalists, policymakers, leaders of professional organizations, and other opinion-leaders each year and increasingly engaging with members of the public.

The Hastings Center will create a project-specific website, which will reproduce the final report in HTML format. This kind of web-based dissemination is far superior to simply posting a pdf of the published final report because it is optimized for search engines and can return usage metrics more sensitive than simple download statistics. It is also highly readable (on electronic devices) and is visually compelling. In addition to the final report text and images, the website will also contain information about the full U19 project and all Project 3 aims, as well as background on newborn screening, a glossary, and a searchable bibliography. An example is a website created for the Hastings Center project on child psychiatry.⁷⁴ The Hastings’s Center’s “New Media Director” will work on the website in year 4 and it will be launched contemporaneous with publication of the final report in year 5. In addition, **the Center will develop a series of press releases about the final report**, which will be sent to its database of more than 4,000 health and health policy journalists (newspapers, magazines, TV, radio, and leading online publications), as well as relevant professional organizations. In addition, the Hastings Center’s Washington, D.C. liaison officer will meet personally with the staff of selected federal policy-makers to bring them the report and discuss its findings. Since newborn screening is an important state policy issue, we will identify and notify relevant state policy makers and public health departments. Other dissemination activities will be discussed with the Ethics & Policy Advisory Board, for example, we might inform members of the U.S. Secretary’s Advisory Committee on Heritable Disorders and Genetic Diseases in Newborns and Children, or work with The Genetic Alliance to make our report available through their public education campaign, “Baby’s First Test.”

The project and its final report will also be publicized through the Center’s existing web platforms. The Center’s main website will include a description of the project and a link to its website and the electronic version of the final report. We will also post commentaries on *Bioethics Forum*, the blog of the *Hastings Center Report* discussing ethical, legal and policy issues addressed by the project. News about the project will also be disseminated to varied audiences via the Center’s existing Twitter feed, which has almost 3,000 followers including journalists from the *New York Times*, *Wall Street Journal*, *Nature*, *Reuters*, *Scientific American*, and other major news outlets, as well as Hastings’s Facebook account, which has almost 1,500 followers.

Advantages and Limitations of Our Approach: Our project design is flexible, builds from a strong team, allows corrections, and establishes robust, interdisciplinary collaborations. The policy advisory board cannot represent all perspectives, but we can add guest speakers as needed.

Timeline of Project 3 Activities

Aims	Activities	Year 01				Year 02				Year 03				Year 04				Year 05			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Aim 1 Protection	Meet with HRPP	X	X	X	X	X	X														
	Draft guidance						X	X	X												
Aim 2 Focus Groups	Develop study guides				X	X															
	Recruit/ conduct groups						X	X	X												
	Analyze data									X	X										
Aim 3 Legal	Legal analysis & review	X	X	X	X	X	X	X	X												
Aim 4 Eth- ics and Policy Ad- visory	Teleconference calls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Meeting preparation				X	X								X	X						
	Board meetings						X									X					
	Synthesize findings							X	X								X	X			
All Aims	ELSI team meetings/calls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Write-up and dissemination								X	X	X	X	X	X	X	X	X	X	X	X	X

Component Project 3 References

1. Brunham LR, Hayden MR. Medicine. Whole-genome sequencing: the new standard of care? *Science*. Jun 1 2012;336(6085):1112-1113.
2. Drmanac R. The advent of personal genome sequencing. *Genetics in Medicine*. March 2011;13(3):188-190.
3. Biesecker LG. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project. *Genet Med*. 2012.
4. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends in Genetics*. 2008;24(3):133-141.
5. Welch JS, Link DC. Genomics of AML: clinical applications of next-generation sequencing. *Hematology Am Soc Hematol Educ Program*. 2011;2011:30-35.
6. Collins FS, Guttmacher AE. Genetics moves into the medical mainstream. *JAMA*. Nov 14 2001;286(18):2322-2324.
7. Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. *Genet Med*. Jun 2011;13(6):499-504.
8. Wei CY, Michael Lee MT, Chen YT. Pharmacogenomics of adverse drug reactions: implementing personalized medicine. *Hum Mol Genet*. Oct 15 2012;21(R1):R58-65.
9. Ingelman-Sundberg M. Pharmacogenetics: an opportunity for a safer and more efficient pharmacotherapy. *J Intern Med*. Sep 2001;250(3):186-200.
10. Guttmacher AE, McGuire AL, Ponder B, Stefansson K. Personalized genomic information: preparing for the future of genetic medicine. *Nature Reviews: Genetics*. February 2010;11:161-165.
11. David G. Genetics and Genomics to the Clinic: A Long Road ahead. *Cell*. 2011;147(1):17-19.
12. Caulfield T, McGuire AL, Cho M, Buchanan JA, Burgess MM, Danilczyk U, Diaz CM, Fryer-Edwards K, Green SK, Hodosh MA, Juengst ET, Kaye J, Kedes L, Knoppers BM, Lemmens T, Meslin EM, Murphy J, Nussbaum RL, Otlowski M, Pullman D, Ray PN, Sugarman J, Timmons M. Research Ethics Recommendations for Whole-Genome Research: Consensus Statement. *PLoS Biol*. 2008;6(3):e73.
13. Almond B. Genetic profiling of newborns: ethical and social issues. *Nat Rev Genet*. Jan 2006;7(1):67-71.
14. Burke W, Tarini B, Press NA, Evans JP. Genetic screening. *Epidemiol Rev*. Jul 2011;33(1):148-164.
15. Bloss CS, Schork NJ, Topol EJ. Effect of Direct-to-Consumer Genomewide Profiling to Assess Disease Risk. *New England Journal of Medicine*. 2011;364(6):524-534.
16. Caulfield T, McGuire AL. Direct-to-Consumer Genetic Testing: Perceptions, Problems, and Policy Responses. *Annual Review of Medicine*. 2012;63(1):23-33.
17. McGuire AL, Burke W. An unwelcome side effect of direct-to-consumer personal genome testing: raiding the medical commons. *JAMA*. Dec 10 2008;300(22):2669-2671.
18. Bodurtha J, Strauss J. Genomics and perinatal care. *The New England Journal of Medicine*. January 5 2012;366(1):64-73.
19. Howell RR. Quality improvement of newborn screening in real time. *Genetics in Medicine*. March 2011;13(3):205.
20. Campbell E, Ross LF. Parental Attitudes and Beliefs regarding the Genetic Testing of Children. *Public Health Genomics*. 2005;8(2):94-102.
21. Wolf SM, Crock BN, Van Ness B, Lawrenz F, Kahn JP, Beskow LM, Cho MK, Christman MF, Green RC, Hall R, Illes J, Keane M, Knoppers BM, Koenig BA, Kohane IS, Leroy B, Maschke KJ, McGeveran W, Ossorio P, Parker LS, Petersen GM, Richardson HS, Scott JA, Terry SF, Wilfond BS, Wolf WA. Managing incidental findings and research results in genomic research involving biobanks and archived data sets. *Genet Med*. Apr 2012;14(4):361-384.
22. Hasegawa LE, Fergus KA, Ojeda N, Au SM. Parental attitudes toward ethical and social issues surrounding the expansion of newborn screening using new technologies. *Public Health Genomics*. 2011;14(4-5):298-306.
23. Lipstein EA, Nabi E, Perrin JM, Luff D, Browning MF, Kuhlthau KA. Parents' decision-making in newborn screening: opinions, choices, and information needs. *Pediatrics*. Oct 2010;126(4):696-704.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

24. Plass AM, van El CG, Pieters T, Cornel MC. Neonatal screening for treatable and untreatable disorders: prospective parents' opinions. *Pediatrics*. Jan 2010;125(1):e99-106.
25. Goldenberg AJ, Sharp RR. The ethical hazards and programmatic challenges of genomic newborn screening. *JAMA*. Feb 1 2012;307(5):461-462.
26. Miller FA, Hayeems RZ, Bombard Y, Little J, Carroll JC, Wilson B, Allanson J, Paynter M, Bytautas JP, Christensen R, Chakraborty P. Clinical obligations and public health programmes: healthcare provider reasoning about managing the incidental results of newborn screening. *J Med Ethics*. Oct 2009;35(10):626-634.
27. Tarini BA, Goldenberg AJ. Ethical Issues with Newborn Screening in the Genomics Era. *Annu Rev Genomics Hum Genet*. May 1 2012.
28. Kitzmiller JP, Groen DK, Phelps MA, Sadee W. Pharmacogenomic testing: relevance in medical practice: why drugs work in some patients but not in others. *Cleve Clin J Med*. Apr 2011;78(4):243-257.
29. Ross CJ, Katzov H, Carleton B, Hayden MR. Pharmacogenomics and its implications for autoimmune disease. *J Autoimmun*. Mar-May 2007;28(2-3):122-128.
30. Loo TT, Ross CJ, Sistonen J, Visscher H, Madadi P, Koren G, Hayden MR, Carleton BC. Pharmacogenomics and active surveillance for serious adverse drug reactions in children. *Pharmacogenomics*. Sep 2010;11(9):1269-1285.
31. Bourgeois FT, Mandl KD, Valim C, Shannon MW. Pediatric adverse drug events in the outpatient setting: an 11-year national analysis. *Pediatrics*. Oct 2009;124(4):e744-750.
32. Husain A, Loehle JA, Hein DW. Clinical pharmacogenetics in pediatric patients. *Pharmacogenomics*. Oct 2007;8(10):1403-1411.
33. Becker ML, Leeder JS. Identifying genomic and developmental causes of adverse drug reactions in children. *Pharmacogenomics*. Nov 2010;11(11):1591-1602.
34. Wall CA, Oldenkamp C, C. S. Safety and efficacy pharmacogenomics in pediatric psychopharmacology. *Primary Psychiatry*. May 2010;17(5):53-58.
35. Schmidt JL, Castellanos-Brown K, Childress S, Bonhomme N, Oktay JS, Terry SF, Kyler P, Davidoff A, Greene C. The impact of false-positive newborn screening results on families: a qualitative study. *Genet Med*. Jan 2012;14(1):76-80.
36. Carlifjord S, Lindberg M, Bendtsen P, Nilsen P, Andersson A. Key factors influencing adoption of an innovation in primary health care: a qualitative study based on implementation theory. *BMC Family Practice*. 2010;11(1):60.
37. Rogers EM. *Diffusion of innovations*. 5th ed. New York: Free Press; 2003.
38. Etchegary H, Dicks E, Green J, Hodgkinson K, Pullman D, Parfrey P. Interest in newborn genetic testing: a survey of prospective parents and the general public. *Genet Test Mol Biomarkers*. May 2012;16(5):353-358.
39. Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, Fletcher JG, Georgieff MK, Hammerschmidt D, Hudson K, Illes J, Kapur V, Keane MA, Koenig BA, Leroy BS, McFarland EG, Paradise J, Parker LS, Terry SF, Van Ness B, Wilfond BS. Managing incidental findings in human subjects research: analysis and recommendations. *J Law Med Ethics*. Summer 2008;36(2):219-248, 211.
40. Clayton EW, Smith M, Fullerton SM, Burke W, McCarty CA, Koenig BA, McGuire AL, Beskow LM, Dressler L, Lemke AA, Ramos EM, Rodriguez LL. Confronting real time ethical, legal, and social issues in the Electronic Medical Records and Genomics (eMERGE) Consortium. *Genet Med*. Oct 2010;12(10):616-620.
41. Fullerton SM, Wolf WA, Brothers KB, Clayton EW, Crawford DC, Denny JC, Greenland P, Koenig BA, Leppig KA, Lindor NM, McCarty CA, McGuire AL, McPeck Hinz ER, Mirel DB, Ramos EM, Ritchie MD, Smith ME, Waudby CJ, Burke W, Jarvik GP. Return of individual research results from genome-wide association studies: experience of the Electronic Medical Records and Genomics (eMERGE) Network. *Genet Med*. Apr 2012;14(4):424-431.
42. McGuire AL, Basford M, Dressler LG, Fullerton SM, Koenig BA, Li R, McCarty CA, Ramos E, Smith ME, Somkin CP, Waudby C, Wolf WA, Clayton EW. Ethical and practical challenges of sharing data from genome-wide association studies: the eMERGE Consortium experience. *Genome Res*. Jul 2011;21(7):1001-1007.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

43. Brody AA, Barnes K, Ruble C, Sakowski J. Evidence-based practice councils: potential path to staff nurse empowerment and leadership growth. *J Nurs Adm.* Jan 2012;42(1):28-33.
44. Sakowski JA, Kahn JG, Kronick RG, Newman JM, Luft HS. Peering into the black box: billing and insurance activities in a medical group. *Health Aff (Millwood).* Jul-Aug 2009;28(4):w544-554.
45. King JS. Predicting probability: regulating the future of preimplantation genetic screening. *Yale J. Health Policy, Law & Ethics.* 2008;8.
46. King JS. And genetic testing for all..the coming revolution in non-invasive prenatal genetic testing. *Rutgers Law Journal.* 2011;42(3).
47. King JS, Moulton BW. Rethinking informed consent: the case for shared medical decision-making. *Am J Law Med.* 2006;32(4):429-501.
48. Moulton B, King JS. Aligning ethics with medical decision-making: the quest for informed patient choice. *J Law Med Ethics.* Spring 2010;38(1):85-97.
49. King JS. Genetic tests: Politics and fetal diagnostics collide. *Nature.* Nov 1 2012;491(7422):33-34.
50. King JS. Not this child: constitutional questions in regulating noninvasive prenatal genetic testing and selective abortion. *U.C.L.A. L. Rev.* 2012;60:2.
51. Parens E, Chapman AR, Press N. *Wrestling with behavioral genetics : science, ethics, and public conversation.* Baltimore: Johns Hopkins University Press; 2006.
52. Baily MA, Murray TH. *Ethics and newborn genetic screening : new technologies, new challenges.* Baltimore: Johns Hopkins University Press; 2009.
53. Etchegary H, Dicks E, Hodgkinson K, Pullman D, Green J, Parfey P. Public attitudes about genetic testing in the newborn period. *J Obstet Gynecol Neonatal Nurs.* Mar 2012;41(2):191-200.
54. Miller FA, Hayeems RZ, Carroll JC, Wilson B, Little J, Allanson J, Bytautas JP, Paynter M, Christensen R, Chakraborty P. Consent for newborn screening: the attitudes of health care providers. *Public Health Genomics.* 2010;13(3):181-190.
55. Berg BL. *Qualitative Research Methods for the Social Sciences.* Boston, MA: Pearson; 2007.
56. Issa AM, Tufail W, Hutchinson J, Tenorio J, Baliga MP. Assessing Patient Readiness for the Clinical Adoption of Personalized Medicine. *Public Health Genomics.* 2009;12(3):163-169.
57. Tabor HK, Stock J, Brazg T, McMillin MJ, Dent KM, Yu JH, Shendure J, Bamshad MJ. Informed consent for whole genome sequencing: A qualitative analysis of participant expectations and perceptions of risks, benefits, and harms. *Am J Med Genet A.* Jun 2012;158A(6):1310-1319.
58. Issa AM, Hutchinson JF, Tufail W, Fletcher E, Ajike R, Tenorio J. Provision of personalized genomic diagnostic technologies for breast and colorectal cancer: an analysis of patient needs, expectations and priorities. *Personalized Medicine.* 2011/07/01 2011;8(4):401-411.
59. Miles MS, Huberman AM. *Qualitative data analysis: An expended source book (2nd Edition).* Thousand Oaks, CA: Sage Publications; 1994.
60. Strauss AL. *Qualitative analysis for social scientists.* New York: Cambridge University Press; 1987.
61. Glaser BG, Strauss AL. *The discovery of grounded theory: Strategies for qualitative research.* Chicago, IL: Aldine; 1967.
62. Weitzman EA, Miles MB. *Computer programs for qualitative data analysis: a software sourcebook.* Thousand Oaks, CA Sage Publications; 1995.
63. Newborn genetic and metabolic screening laws. <http://www.ncsl.org/issues-research/health/newborn-genetic-and-metabolic-screening-laws.aspx>.
64. National newborn screening status report. <http://genes-r-us.uthscsa.edu/sites/genes-r-us/files/nbsdisorders.pdf>, 2012.
65. *Newborn screening: toward a uniform screening panel and system.* 16.
66. *Jacobson v. Massachusetts,* (197 US 11 1905).
67. Nardi BA, Whittaker S. The place of face-to-face communication in distributed work. In: Hinds P, Kiesler S, eds. *Distributed work.* Cambridge: The MIT Press; 2002.
68. Research teams of the future. <https://commonfund.nih.gov/researchteams/>.
69. Boix-Mansilla V, Gardner H. *Assessing interdisciplinary work at the frontier: An empirical exploration of "symptoms of quality"*2004. 26.
70. Klein JT. *Interdisciplinarity : history, theory, and practice.* Detroit: Wayne State University Press; 1990.
71. Moran J. *Interdisciplinarity.* London ; New York: Routledge; 2002.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

72. Nowotny H, Scott P, Gibbons M. *Re-thinking science. Knowledge and the public in an age of uncertainty*. Cambridge: Polity Press; 2001.
73. Turner BS. The interdisciplinary curriculum: from social medicine to postmodernism. *Sociology of Health & Illness*. 1990;12(1):1-23.
74. Parens E, Johnston J. Troubled children: Diagnosing, treating, & attending to context. *Hastings Center Report* 2011; 41:<http://childpsychiatry.thehastingscenter.org/>.

Program Director/Principal Investigator (Last, First, Middle): Nussbaum, Robert L.

Resource Sharing Plan

A. Data Sharing Plan

For Projects 1 and 2, California law considers blood specimens and any individual data derived from blood specimens to be confidential information subject to strict protections. These confidentiality protections require that such material be protected by administrative, physical and technical safeguards; be returned or destroyed when no longer needed for the approved research activity for which they were provided; and that they shall not be reused, disclosed to any other person or entity, or used in any manner not approved in the research protocol which was submitted with the request for these materials. Thus, California law precludes any researcher from sharing blood specimens or uploading individual data derived from these blood specimens into dbGap or any other genomic data repository.

For the studies involving sequencing of whole exomes from primary immunodeficiency patients, patients will be ascertained through the clinic. We will seek informed consent to examine all genes relevant to immunodeficiency. We will also examine their genomes for variants in any one of the genes contained in the peer-reviewed, updated, evidence-based guidelines for gene/drug pairs identified by the Clinical Pharmacogenetic Implementation Consortium as particularly clinically relevant {Relling, 2011 #150}. We will seek consent to upload these data to dbGAP. The applicants acknowledge that controlled access to individual genotype, phenotype, and exposure data will be available for approved research purposes through the Executive Committee, and that interested researchers will be required to complete a "Data Use Certification". We expect that this certification will include the core elements of NIH Data Use Certification Agreements.

For the studies in Project 3, the data generated through the proposed research are not amenable to centralized electronic data dissemination and will be shared in traditional manners, e.g., via publication and public presentation of the data at meetings. The products from all aims will be shared fully, via the dissemination plans outlined in the description of Aim 4. The bibliographic resources from Aims 3 and 4 will be made available via the project website. Focus group data collected as part of Aim 2 will be made available, in deidentified format, by request to the investigators.

B. Sharing Model Organisms: N/A; No model organisms will be generated under the scope of this proposal.

C. Genome-Wide Association Studies: N/A; No genome wide association data will be available through this application.

D. Data Management

MyResearch

Data for this project will be stored and analyzed in UCSF's Secure MyResearch environment. This HIPAA compliant desktop environment is hosted on servers housed at the UCSF Data Center on Minnesota Street, San Francisco, CA. The MyResearch environment is hosted on six Dell PowerEdge R710s and Five EqualLogic PS6100E SAN, which are located inside the locked rack. There are two layers of physical redundant Cisco firewalls that protect the servers and SAN. The MyResearch environment utilizes VM Ware View Virtual Desktop, which must be logged into using UCSF Active Directory credentials. The servers are locked inside a rack locked with a combination lock. The rack is located in a data center secured by two sets of locked doors.

Data set will be stored in the Principal Investigator's group network folder in the remote MyResearch environment where only the research team members are able to view the data sets and this access is audited. This folder is physically located in a data store on the SAN in the locked rack. Network traffic between MyResearch and the UCSF campus network traverses a SSL VPN tunnel in encrypted format.

Analysis tools are hosted in the environment contiguous to data storage.

There is a tape back-up. Every day the MyResearch team backs-up the incremental changes to the data. Every week they create a full back-up that is stored off-site. A monthly copy of the full back-up is retained off-site for 18 months.

Part E - Previous Discussions of Submissions

10/28/13 – Dr. Robert Nussbaum, principal investigator of the UCSF Newborn Screening (NBS) grant, and Dr. Kellie Kelm of the Division of Chemistry and Toxicology at FDA had a phone call to discuss the question of whether or not the NBS grant recipients would need to submit an IDE to FDA.

1/28/14 – Dr. Kellie Kelm of FDA; Dr. Tiina Urv of NIH; Sara Kalia, Harvard's NBS project manager; Cristina Kapustij, UCSF's project manager; and Dr. Laura Milko, UNC's project manager, had a conference call to discuss specific questions about what information should be included in the IDE pre-submission form.

Part F – Overview of Product Development

Due to the nascent state of our grant research, we have not generated any data nor have we produced any publications yet.

Please note, nothing from this research is intended to lead to commercialization of a product.

Part G - Specific Questions

1. As can be seen in our specific aims, neither project 1 nor project 2 will be returning any results. To our team this does not seem to constitute a significant risk. Does FDA consider any components of these projects as creating a significant risk?
2. Will our research require an IDE from FDA?
3. We have provided you a current version of an evolving protocol. If our protocols are modified during the course of our research do we need to contact FDA? If yes, how do we do this?

Part H – Mechanism of Feedback

Our preferred method of feedback is a teleconference involving Principal Investigator, Dr. Robert Nussbaum, and Program Manager, Cristina Kapustij.

Dr. Robert Nussbaum

UCSF Institute for Human Genetics
513 Parnassus Avenue
HSE 901D
San Francisco, CA 94143
Tel: (415) 476-3200
Email: nussbaumr@humgen.ucsf.edu

Cristina Kapustij

UCSF Institute for Human Genetics

513 Parnassus Avenue

HSE 901D

San Francisco, CA 94143

Tel: (415) 502-2859

Email: kapustijc@humgen.ucsf.edu