

Solexa Sequencing Library Quality Control Analysis by Quantitative Real-Time PCR

This protocol describes a method to estimate the *relative* expected performance of your libraries in sequencing on the Illumina Genome Analyzer II (i.e. Solexa) platform. Performance estimates are normalized relative to an existing library of similar construction that you have already verified to give adequate data yields. If you do not have access to such a library, contact us about providing you with a sample of one.

We perform this assay on an MJR DNA Engine Opticon 2, but virtually any real-time PCR instrument could be substituted. With 96-well block designs we recommend avoiding the wells at the edge of the block, since the thermal parameters of these wells often show excessive variance. This still leaves 60 usable wells, which will allow you to analyze up to 12 libraries in a single run.

We use iQ SYBR Green Supermix (cat. #170-8882) from BioRad, but you can probably substitute any good-quality commercial SYBR Green master mix without affecting assay performance.

PCR primer sequences for single read libraries are as follows:

AdapA 5'-CAAGCAGAAGACGGCATAACG-3'

AdapB 5'-AATGATACGGCGACCACCGA-3'

Remember to pipette very carefully, since the quality of the results depends on it.

1. Make one >18 μ l aliquot of the reference library at 0.625 nM final concentration. For example, combine 2 μ l of 10 nM DNA + 30 μ l nuclease-free water to yield 32 μ l of 0.625 nM DNA. Divide into three equal aliquots for serial dilution in triplicate (below).
2. Prepare triplicate 8-step serial dilutions of the reference library by combining 3 μ l DNA with 3 μ l nuclease-free water at each step. Mix each dilution thoroughly before proceeding to the next. Upon completion you should have triplicate aliquots (each \geq 3 μ l) at 0.625, 0.3125, 0.15625, 0.078125, 0.0390625, 0.01953125, 0.009765625 and .0048828125 nM final concentrations. A multi-channel pipette capable of accurately delivering 3 μ l volumes will speed this process.
3. Prepare triplicate \geq 3 μ l aliquots of your unknown libraries at 0.05 nM final concentration. This should place them somewhere near the middle of the standard curve.
4. Prepare enough PCR master mix for three wells per unknown library plus 24 wells for the reference library plus 5% overage. For five unknown libraries (15 + 24 wells) and 25 μ l reactions the recipe would be as follows:

<u>1X</u>	<u>39X (+5%)</u>	
12.5 μ l	511.9 μ l	iQ SYBR Green Supermix
3 μ l	122.9 μ l	1 μ M (each) AdapA/AdapB primer mix

7.5 μ l 307.1 μ l nuclease-free water

5. Aliquot 23 μ l PCR master mix into the appropriate reaction locations. The use of multichannel pipettors can speed this step and the next. For five libraries in a standard 96-well PCR plate you could use the following layout:

<u>Library</u>	<u>Replicate 1</u>	<u>Replicate 2</u>	<u>Replicate 3</u>
Reference	wells B3-B10	wells C3-C10	wells D3-D10
Library 1	well E4	well E5	well E6
Library 2	well E7	well E8	well E9
Library 3	well F4	well F5	well F6
Library 4	well F7	well F8	well F9
Library 5	well G4	well G5	well G6

6. Aliquot 2 μ l of template DNA (either reference or unknown) into the appropriate wells, pipetting up and down several times to mix. Seal wells with optical-grade lids or adhesive film. We use Excel Scientific ThermalSeal RT 2 mil adhesive film, available from ISC BioExpress (cat. #T-2417-8). Spin briefly to distribute contents to the bottoms of the wells.
7. Thermal cycle reactions as follows:

95°C for 3 minutes
then 26 cycles of:
95°C for 30 seconds
60°C for 30 seconds
72°C for 1 minute
then 72°C for 10 minutes

8. After cycling, select appropriate background correction parameters and set the threshold. On the Opticon 2 instrument we use global minimum baseline subtraction, suppressing pre-threshold data as necessary, and set the threshold at one standard deviation of the first cycle data. But you can vary these parameters considerably without affecting the overall results much, so long as

you select a threshold that lies in the zone of exponential amplification (the linear portion of the log-transformed fluorescence plots).

9. When you are satisfied with the analysis parameters, export the C_t values for all wells, saving them in plain text format. Import them into Excel and cut and paste them into the appropriate cells of an analysis spreadsheet. A blank sample spreadsheet is available.
10. Excel analysis consists of plotting the average C_t for each standard versus the log of the template amount (fmoles) in each reaction. When plotted in this way, the dilution series for the reference library should yield a straight line. The linear regression feature of Excel plots the trendline and displays the R^2 value on the chart (which should be quite close to 1).
11. Interpreting results: Using the best-fit line equation the spreadsheet back-calculates the starting concentration of each unknown library as *normalized by PCR performance relative to the reference library*. PCR performance is assumed to be a suitable proxy for relative performance of libraries during cluster amplification. The normalized concentration estimates given by the spreadsheet reflect both starting nucleic acid amount and performance as a PCR template, so they can't be expected to agree with measurements of DNA concentration given by NanoDrop, PicoGreen, Bioanalyzer, etc. However the normalized concentrations are useful in determining how much library DNA to use for cluster generation.

Since the true concentration of the template DNA of your unknown libraries was 0.05 nM, any normalized concentration estimate below 0.05 nM means the library can be expected to yield fewer clusters than the reference library from the same molar amount of starting material. If the normalized concentration estimate is greater than 0.05 nM the library should yield more clusters than the reference. For example, if your unknown library has a normalized concentration of 0.1 nM then it is expected to yield 1/5 the number of clusters as the reference library per picomole added. If the normalized concentration is 1.0 nM the library would be expected to yield twice the number of colonies as the reference library per picomole.

Since the amount of reference library DNA needed to yield an ideal number of clusters is known, the simplest way to determine how much DNA to use for an unknown library is to use the same number of picomoles as you would for the reference library, but prepared *using the normalized (i.e. spreadsheet-calculated) concentrations*. For example, suppose your reference library gives ideal results when 3.5 pmol are loaded on the cluster station and your unknown library is 250 nM according to the NanoDrop. You dilute your unknown library 5000-fold to get a 0.05 nM sample for the q-PCR assay, and the spreadsheet calculates a *normalized* concentration of 0.025 nM. Accounting for the dilution, the *normalized* concentration of your unknown library is $5000 \times 0.025 = 125$ nM, so to load a *normalized* 3.5 pmol you will need 28 ul of your unknown library. This should yield the same number of clusters as 3.5 pmol of the reference library.

Sample Results

	A	B	C	D	E	F	G	H	I										
1		Actual			Ct														
	Ref. Lib. Dilution Number	Starting Quantity (fmoles)	Log of Starting Quantity	Reference	Library 1	Library 2	Library 3	Library 4	Library 5										
2																			
3	1	1.25	0.09691001	2.63															
4	2	0.625	-0.20411998	3.45															
5	3	0.3125	-0.50514998	4.30															
6	4	0.15625	-0.80617997	5.44															
7	--	0.1	-1		6.10	8.63	7.00	7.28	6.90										
8	5	0.078125	-1.10720997	6.74															
9	6	0.0390625	-1.40823997	7.89															
10	7	0.01953125	-1.70926996	9.01															
11	8	0.009765625	-2.01029996	10.13															
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27			slope:	-3.6516163	intercept:	2.706933736													
28																			
29			Back-calculated																
			Log of Starting Quantity	Starting Quantity (fmoles)	Starting conc. relative to ref. lib. (nM)														
30	Sample	Ct																	
31	Library 1	6.10	-0.9283	0.1180	0.0590														
32	Library 2	8.63	-1.6211	0.0239	0.0120														
33	Library 3	7.00	-1.1747	0.0669	0.0334														
34	Library 4	7.28	-1.2533	0.0558	0.0279														
35	Library 5	6.90	-1.1492	0.0709	0.0355														