Ion Torrent Library Preparation Process using KapaBiosystems Library Quantification Kit significantly increases quantitative sensitivity and improves sample pooling for multiplexed sequencing

David Roberson[1,2], Joseph Boland[1,2], Evan Caporaso[1,2] , Casey Dagnall[1,2], Alexandra McCary[1,2] , Victor Lonsberry[1,2] , Michael Malasky[1,2] ,Timothy Myers[1,2] , Meredith Yeager[1,2], Kevin B. Jacobs[1,2], Amy Hutchinson[1,2] and Stephen Chanock[1,2]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland USA
[2]Core Genotyping Facility, Office of the Director, SAIC-Frederick, National Cancer Institute-Frederick, Frederick, Maryland USA

Abstract

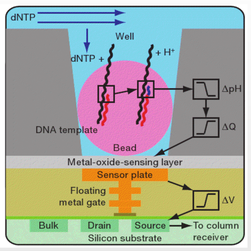
The KAPA Biosystems Library Quantitation Kit for the Ion Torrent Personal Genome Machine (PGM) has helped us achieve a consistent template to sphere ratio and consistent equimolar barcode pooling when compared to other quantitative methods. A correct template to sphere ratio (also known as dilution factor) is essential for a successful sequencing run. If too much template is added during template preparation, the probability of polyclonal IonSpheres is greatly increased. Conversely, if too little template is added the total wells with ISPs during a sequencing run will be less than adequate. Both of these situations will lead to suboptimal sequencing performance on the PGM. Accurate sample quantitation allows us to process more samples per Ion Chip which saves time and direct sample cost. Currently, we have validated 24 custom barcodes that are used with Life's Ion Xpress Plus Kit as well as 96 validated barcodes for use with Fluidigm's Access Array.

We will present data showing the marked differences between the KAPA Biosystems quantitation, Caliper GX quantitation and Agilent Bioanalyzer quantitation. In addition, we will present data from numerous 314 and 316 Ion chip runs highlighting the consistent pooling and uniformity of each set of barcodes.

Introduction

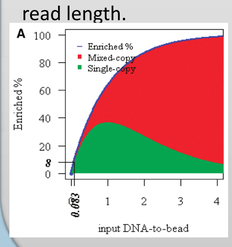
Ion Torrent's PGM is a massively parallel sequencing by synthesis (SBS) instrument. It operates by directly sensing (i.e. no enzymatic cascade or cameras required) the ions produced as DNA polymerase incorporates nucleotides to a sphere bound clonal template.

As with all other SBS technologies template prep is critical to amplify the signal into a range that the measurement device can accurately sense.



The PGM relies on a type of PCR to generate highly templated spheres. If a sphere has many copies of a single template covalently attached to it, the sphere is referred to as clonal. The more templates it has, the more ions will be released during polymerization, and the easier it will be for the base calling software to model the signal. Negative effects such as phasing will be less of an issue.

If a sphere has more than one template molecule attached to it is referred to as polyclonal. The signal modeling software will not be able to determine which template is being copied at any given incorporation event and therefore the entire read will be thrown out. Sometimes, one of the mixed molecules amplifies preferentially and therefore gives off a stronger signal. In this case the mixed signals act as background noise and drastically reduces the sequence quality and read length.

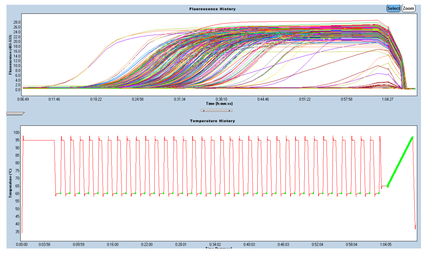


Prior publications have suggested that the ratio of mixed to clonal spheres as a function of input DNA follows a *Poisson* distribution.

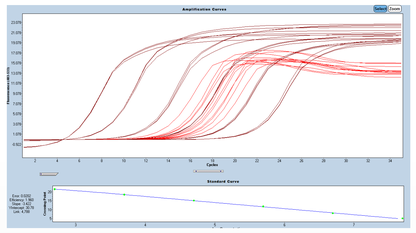
This suggests that if DNA quantitation was accurate enough optimal enrichment can easily be achieved. Therefore it is no longer an issue of trial and error titrations but of DNA quantitation accuracy. This is where quantitative PCR (qPCR) adds significant value.

Experimental Design /Methods

There are two types of qPCR, intercolator-based and probe-based. The probe based method uses a Taqman hydrolysis probe that is complementary to 1 of the ligated sequencing adaptors. The intercolator based probe method uses SYBR Green I which is a fluorescent dye that will bind to all dsDNA in the sample well. Our lab decided to go with the intercolator based ***Kapa Biosystems NGS Library Quant Kits*** because of the lower cost and versatility.



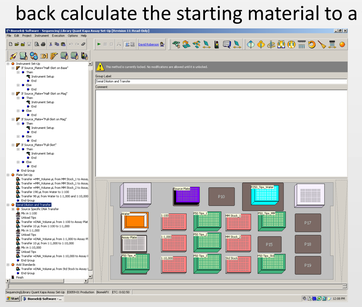
Thermocycling protocol with example data



Standard curve

qPCR is a subset of real time pcr that requires a specialized PCR instrument. Our lab uses a Roche LC480. These types of instruments use a charge-coupled camera in concert with a high intensity xenon lamp to collect fluorescent signals.

In the case of qPCR the fluorescent signal is collected at the end of each elongation step. After each consecutive PCR cycle the signal will build. Post PCR cycling the instruments software (using the 2nd derivative maximum analysis) will determine when the signal crossed a predetermined threshold and thus is able to back calculate the starting material to a very high level of accuracy.

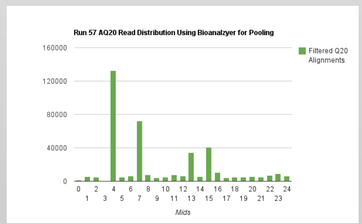


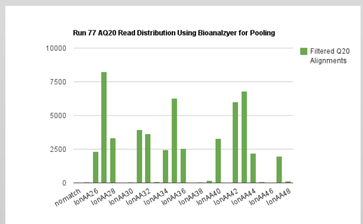
Kapa's Ion PGM NGS Library Quant kit pairs well with the Fluidigm's Access Array technology we have in the lab. The lower throughput of the Agilent Bioanalzyer (11 samples per chip) cannot keep up with our Access Array output.

We have created a robotic method on a Biomek FX that does all prep for the Kapa assay including serial dilutions, master mix addition and sample/standard addition.

Results

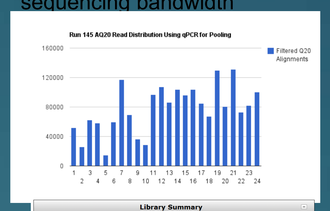
Prior to qPCR implementation at our facility, barcode pooling was either done post library prep in a equi-volume or equi-molar (based on quantitation from Bioanalyzer). We found both methods to be suboptimal.

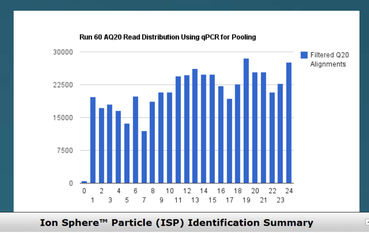


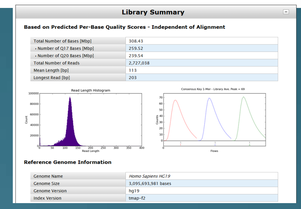


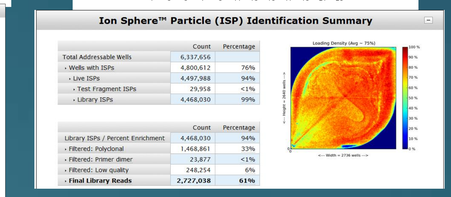
Results cont'd

Kapa Biosystems NGS Library Quant Kit has allowed for a better distribution of sequencing reads and therefore a better use of "sequencing bandwidth"





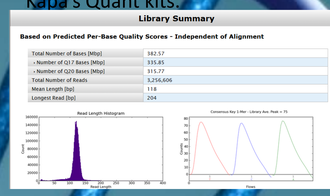


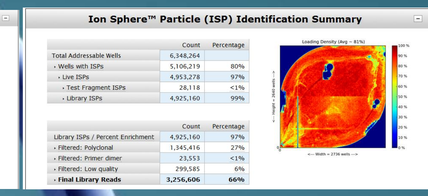


Count	Percentage
Total Addressable Wells	6,337,656	
• Wells with ISPs	4,800,612	76%
• Live ISPs	4,497,988	94%
• Test Fragment ISPs	29,958	<1%
• Library ISPs	4,468,030	99%
Library ISPs / Percent Enrichment	4,468,030	94%
• Filtered: Polyclonal	1,468,861	33%
• Filtered: Primer dimer	23,877	<1%
• Filtered: Low quality	248,254	6%
• **Final Library Reads**	**2,727,038**	**61%**

Most recently we have achieved a high level of success with our sequencing runs. The below run was the highest yielding 100bp library ran on a 316 chip last month (Jan. '12) according to the Ion community leader boards. We attribute our continued success in part to using Kapa's Quant kits.

Week Ending in	314	316	318
February 12, 2012	rk8052zh (71.74Mb)	rk8052zh (466.46Mb)	-
February 5, 2012			
January 2011	rk8052zh (96.06Mb)	pli (596.76Mb)	jkron (924.10Mb)
January 30, 2012	sherke (54.83Mb)	sherke (151.81Mb)	jkron (924.10Mb)
January 23, 2012	rk8052zh (96.06Mb)	tstinear (231.38Mb)	-
January 16, 2012	cindau (30.78Mb)	pli (596.76Mb)	jkron (752.74Mb)
January 9, 2012	tneff (33.48Mb)	RMeize (269.20Mb)	jboland (804.45Mb)
January 2, 2012		jboland (315.76Mb)	-
December 2011	rtewh (103.30Mb)	pli (589.74Mb)	mholder (989.96Mb)
December 25, 2011	leemurphy (56.37Mb)	cover_bio (470.92Mb)	mholder (988.96Mb)
December 18, 2011	cindau (61.97Mb)	donelli (436.20Mb)	jkron (778.66Mb)
December 11, 2011	rtewh (103.30Mb)	pli (589.74Mb)	-
December 4, 2011	cavatorm (78.47Mb)	mzkg (355.13Mb)	-
November 2011	cavatorm (78.47Mb)	jsheffield (350.93Mb)	-
November 27, 2011	-	-	-
November 20, 2011	-	-	-
November 13, 2011	kgould (65.73Mb)	jsheffield (350.93Mb)	-





Count	Percentage
Total Addressable Wells	6,348,264	
• Wells with ISPs	5,106,219	80%
• Live ISPs	4,953,278	97%
• Test Fragment ISPs	28,118	<1%
• Library ISPs	4,925,160	99%
Library ISPs / Percent Enrichment	4,925,160	97%
• Filtered: Polyclonal	1,345,416	27%
• Filtered: Primer dimer	23,553	<1%
• Filtered: Low quality	299,585	6%
• **Final Library Reads**	**3,256,606**	**66%**

Future Development

In the future we believe Kapa Biosystems Quant Kits will have an important role in increasing our throughput by providing a fast (and easily automated) alternative to other NGS quant methods and by removing the need for an amplification step at the end of library preparation. The dynamic range of qPCR is sensitive enough to determine a concentration for libraries that are not detectable on an High Sensitivity chip for Agilent Bioanalyzer 2100 chip or High Sensitively chip for Caliper GXII. These libraries would previously had to be amplified.

Acknowledgements

Eric van der Walt , Kapa Biosystems
John Foskett, Kapa Biosystems
Eric Suh, Kapa Biosystems
Eric Troop, Kapa Biosystems
Andrew Lemire, Life Technologies

References

Rothberg J, Hinz W, Rearick TM, Schultz J, Mileski W, et al. (2011) **An integrated semiconductor device enabling non-optical genome sequencing**. Nature. In press.

Zheng Z, Advani A, Melefors O, Glavas S, Nordstrom H, et al. (2010) **Titration-free massively parallel pyrosequencing using trace amounts of starting material**. Nucleic Acids Res 38: e137.

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.