



**NRC-CMRC**

*Plant  
Biotechnology  
Institute*

# **Sample Preparation and Sequencing for 454 GS FLX ROCHE**

Janet Condie, Christine Sidebottom,  
Carling Tallon and Inge Roewer

January 20, 2011



National Research  
Council Canada

Conseil national  
de recherches Canada

**Canada**

**NRC-CNRC**

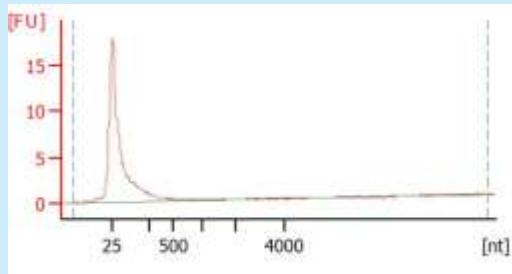
*Plant  
Biotechnology  
Institute*

# 454 Genome Sequencer since February 2008

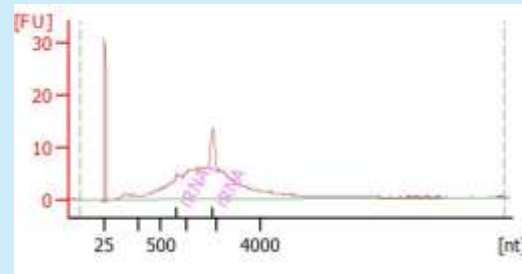


# Sample Submission

- Contact us to obtain information and set-up a contract with PBI.
- Principal Investigator obtains access to **LIMS** for submitting the request(s) with all Sample and Library relevant information including concentration and contract number.
- Please submit **purified, high quality** gDNA, cDNA, Amplicons or enriched mRNA at the **appropriate** concentration.
- Verification of concentration and quality via Fluorometer, RNA or DNA BioAnalyzer CHIPs and/or agarose gel.



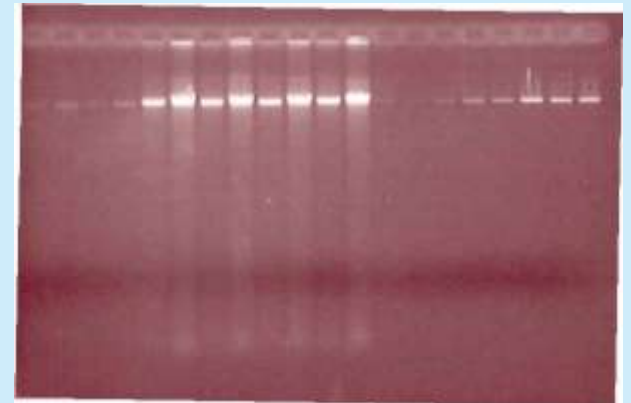
Degraded mRNA Sample, not usable



acceptable mRNA Sample

RNA Concentration: 1,573 pg/ $\mu$ l

rRNA Contamination: 6.8 %



# Sample dependent Library Constructions for different Applications

## Paired End Libraries, 3 kb to 40 kb Span

- 5 µg to 50 µg high molecular ds DNA
- Hydroshear and End Repair
- Circularization Adaptor Ligation
- Library Span Size Selection
- DNA Circularization
- Nebulization of circ. DNA →
- End Repair
- Library Adaptor (A/B) Ligation
- Library Size Selection, Cleaning and Isolation
- Quality and Quantity Verification
- Dilution for emPCR

## Shotgun Library from gDNA or cDNA

- Minimum of 500 ng ds DNA
- Nebulization of gDNA or cDNA
- End Repair
- MID or (A/B) Adaptor Ligation
- Library Size Selection, Cleaning and Isolation
- Quality and Quantity Verification →
- Dilution (and Pools) for emPCR

## Amplicon/ PCR Library

.....  
.....

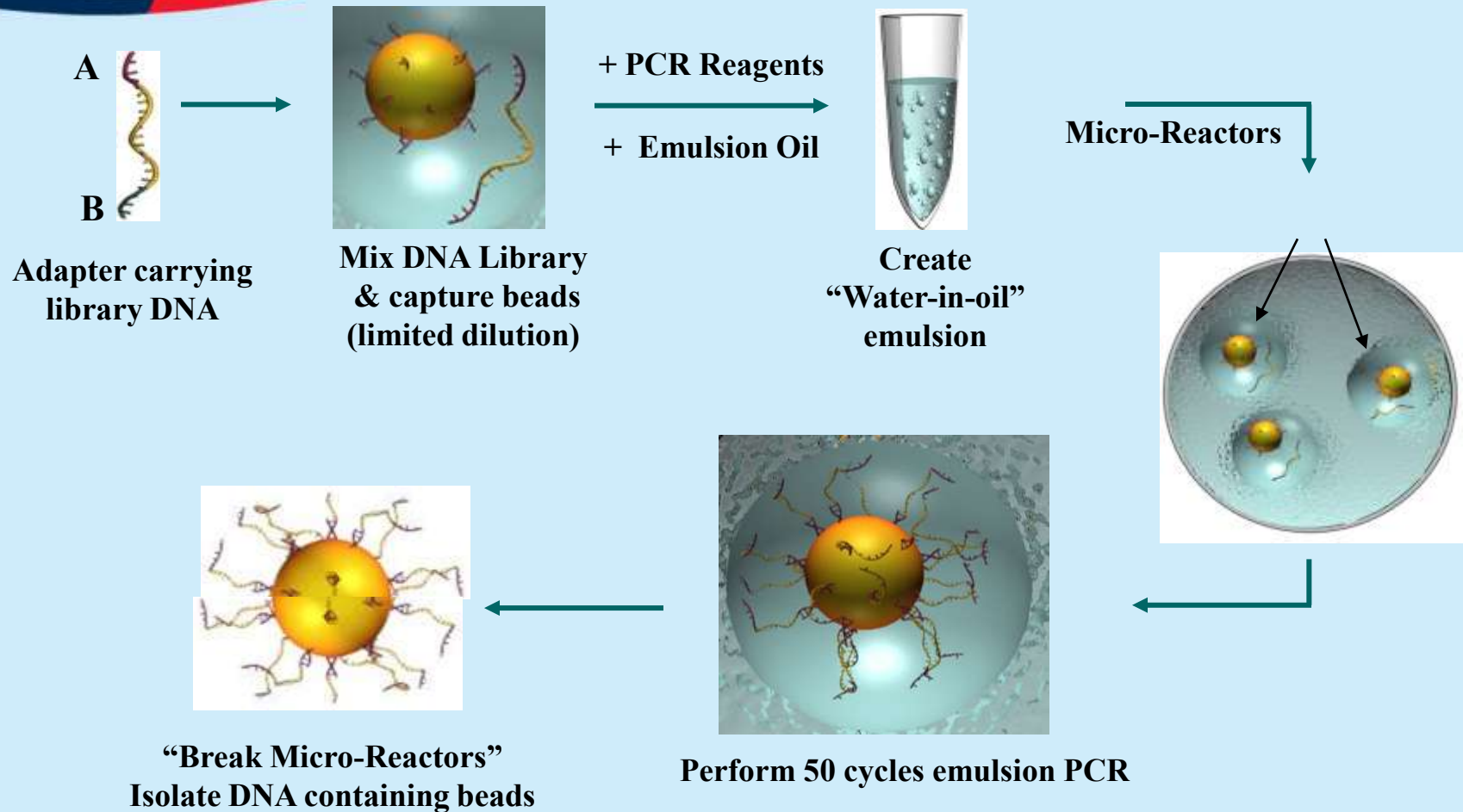
## MID-Tagged Amplicons

.....

## Roche cDNA Rapid Library Construction Kit using Purified Enriched mRNA

- Minimum of 400 ng good quality, **enriched** mRNA
- Fragmentation followed by 1. strand Synthesis with **Random** Primers (less issues with polyA)
- 2. strand Synthesis to produce cDNA
- End Repair
- ..... , **etc**

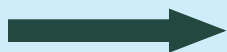
# Emulsion Based Clonal Amplification



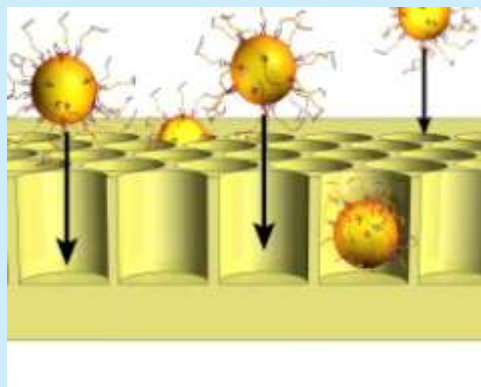
- Generation of millions of clonally amplified sequencing templates on each bead
- No cloning and colony picking

# Load DNA Beads into the PicoTitre™ Plate

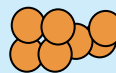
Load DNA beads  
(20 µm) into Plate



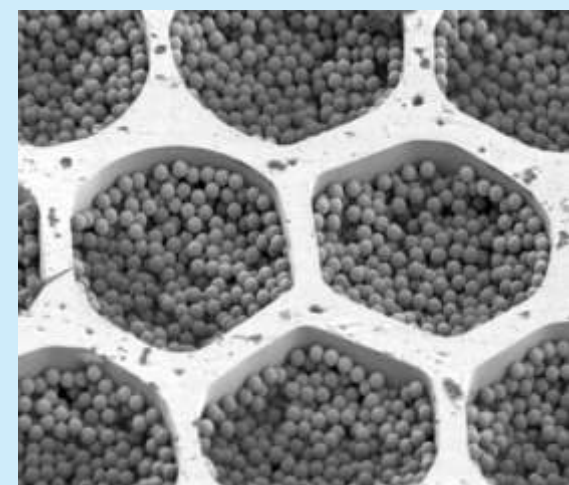
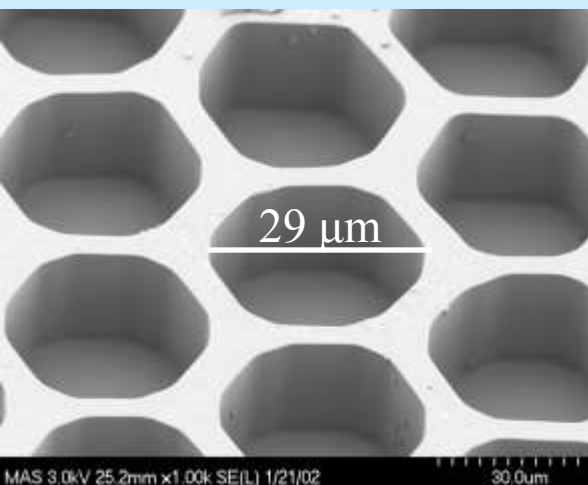
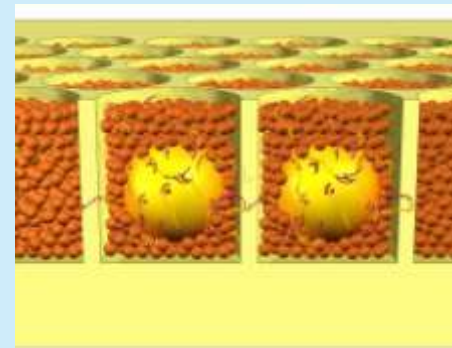
Centrifuge Step



Load Enzyme  
Beads



Centrifuge Step

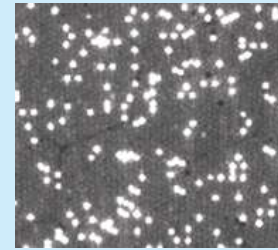


Wells are made of Fiber Optics with a thin layer of Titanium

# Sequencing-By-Synthesis

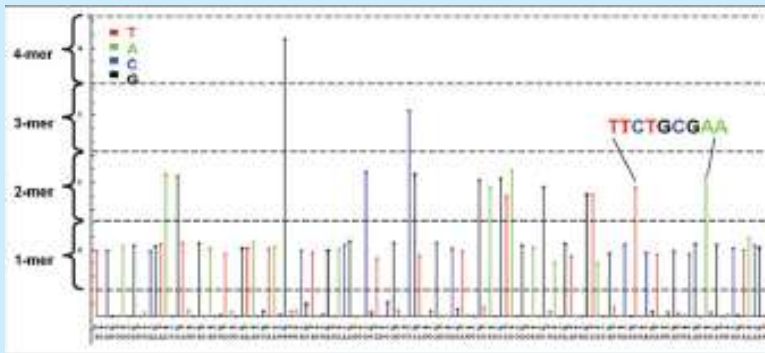
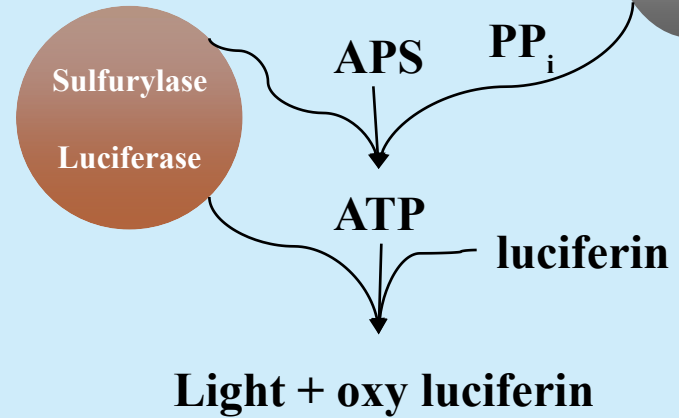
- Simultaneous sequencing in hundreds of thousands of picoliter-size wells
- Pyrophosphate signal generation

DNA Capture Bead  
Containing Millions of  
Copies of a Single  
Clonal Fragment



A A T C G G C A T G C T A A A A G T C A

T  
Anneal Primer



# Selection of Regions on PicoTitre-Plate (PTP)

Depending on Library type, Genome size, Sample Pool, required Depth and Coverage, the PTP can be divided in different size regions.

Each Library type has a specific Key recognized by the Software.

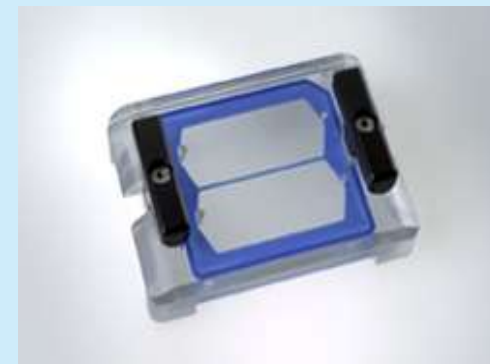
How many Beads to load --- How many Reads --- How many Bases to expect

<b>Loading Region Size</b>	<b><u>Number of DNA Library Beads to Load per Region</u></b>	<b><u>Reads / Region (x 1000)</u></b>	<b><u>Bases/ Region (Mbp)</u></b>	<b><u>Bases per full plate (Mbp)</u></b>
2 x Large (30x60 mm)	2,000,000 (x2)	450-650	180-280	360-560
4 x Medium (14x43 mm)	790,000 (x4)	160-250	60-110	240-440
8 x M/S (5.4x60 mm)	340,000 (x8)	80-120	30-55	240-440
16 x Small (2x53 mm)	125,000 (x16)	25-40	10-20	160-320

# Post-Run Metrics

GACT (Library)	Region		Total
	1	2	
Raw Wells	1,133,257	1,109,490	2,242,747
Key Pass Wells	1,081,719	1,071,226	2,152,945
Passed Filter Wells	838,603	827,523	1,666,126
Total Bases	343,060,830	343,675,731	686,736,561
Length Average	409.11	415.32	412.18
Length Std Deviation	114.18	111.73	
Longest Reads Length	787	724	787
Shortest Reads Length	40	40	40
Median Reads Length	433.0	439.0	437.0

GACT (Library)		Region		Total
		1	2	
Raw Wells		1,133,257	1,109,490	2,242,747
Key Pass Wells		1,081,719	1,071,226	2,152,945
Failed	Dot	10,532	13,743	24,275
	Mixed	75,282	84,157	159,439
	Short Quality	157,287	145,793	303,080
	Short Primer	15	10	25
Passed Filter Wells		838,603	827,523	1,666,126
% Dot + Mixed		7.93	9.14	8.53
% Short		14.54	13.61	14.08
% Passed Filter		77.53	77.25	77.39



GACT (Library)	Region				Total
	1	2	3	4	
Raw Wells	385,448	398,997	420,598		1,205,043
Key Pass Wells	346,633	363,728	369,049		1,079,410
Passed Filter Wells	257,178	262,079	235,927		755,184
Total Bases	92,936,334	84,809,045	83,329,199		261,074,578
Length Average	361.47	323.67	353.36		345.71
Length Std Deviation	133.99	145.15	127.46		
Longest Reads Length	1,118	693	670		1,118
Shortest Reads Length	40	40	40		40
Median Reads Length	395.0	351.0	381.0		376.0

TCAG (Library)		Region				Total
		1	2	3	4	
Raw Wells					438,775	438,775
Key Pass Wells					419,733	419,733
Passed Filter Wells					184,924	184,924
Total Bases					61,724,223	61,724,223
Length Average					334.02	333.78
Length Std Deviation					125.45	
Longest Reads Length					712	712
Shortest Reads Length					40	40
Median Reads Length					363.0	363.0



GACT (Library)		Region																Total
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Raw Wells		40,182	27,874	52,709	56,803	63,664	53,460	23,021	9,646	54,434	58,551	40,840	63,761	29,859	31,508	36,906	40,470	683,688
Key Pass Wells		37,375	24,623	49,721	53,880	61,575	50,961	20,476	7,845	52,337	56,400	38,420	61,762	27,675	28,938	34,333	25,564	631,885
Failed	Dot	1,249	600	845	6,312	5,077	2,122	635	161	589	789	4,046	593	384	542	329	4,243	28,516
	Mixed	1,928	995	2,123	14,376	9,189	7,722	1,286	178	3,376	2,550	7,352	2,159	413	552	1,031	16,101	71,331
	Short Quality	5,000	3,047	5,020	6,706	10,622	7,484	1,625	449	7,029	7,183	3,278	6,922	2,539	2,897	3,612	1,611	75,024
	Short Primer	27	61	37	103	63	90	18	6	20	13	24	21	28	44	15	21	591
Passed Filter Wells		29,171	19,920	41,696	26,383	36,624	33,543	16,912	7,051	41,323	45,865	23,720	52,067	24,311	24,903	29,346	3,588	456,423
% Dot + Mixed		8.50	6.48	5.97	38.40	23.17	19.32	9.38	4.32	7.58	5.92	29.67	4.46	2.88	3.78	3.96	79.58	15.80
% Short		13.45	12.62	10.17	12.64	17.35	14.86	8.02	5.80	13.47	12.76	8.59	11.24	9.28	10.16	10.56	6.38	11.97
% Passed Filter		78.05	80.90	83.86	48.97	59.48	65.82	82.59	89.88	78.96	81.32	61.74	84.30	87.84	86.06	85.47	14.04	72.23



- **Postrun Data retrieval via LIMS**

Like for Illumina PI has access to LIMS to download 454 data and PI has access to **Newbler Software** for data assembly

- **Future Upgrades**

Release of 1 KB Titanium Chemistry Kits for 445 GS FLX with **average Read Length of 800 bases** in 2011

- **Thank you**