Primer Walking in the Finishing Lab at the Broad Institute

Xiaohong Liu, Mostafa Benamara, Tashi Lokyitsang, Charles Matthews, Tamrat Negash, Thu Nguyen, Nicole Allen, Daniel Bessette, Michael FitzGerald
The Broad Institute, Cambridge, MA 02141

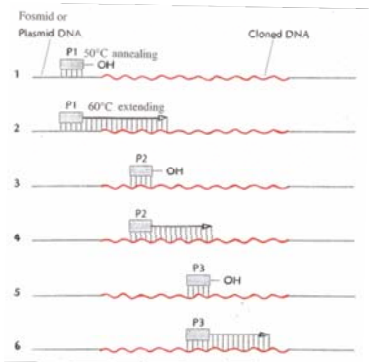
OVERVIEW

Primer walking at the broad is an effective strategy to sequence gaps or resolve small regions of difficult sequence. The DNA of interest may be a plasmid (4kb) or a fosmid (40kb) insert. The initial sequencing is performed by Production Sequencing from each end using universal primers. In order to completely sequence the region of interest, Computer Finishers analyze the initial data, identify the directions of sequence, and design custom oligonucleotide primers near the end of the known sequence. The first sequencing reaction is carried out and, based on the reads found, a second primer can be designed off of the new sequence to extend farther into the gap. Multiple rounds of this can be repeated until the gap is filled or questionable sequence can be resolved.

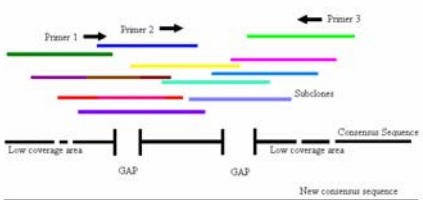
In the Finishing Lab, plasmid or fosmid 384-well glycerol plates are pulled out and specific clones are inoculated from their initial positions into a single 96-well growth plate (medium with chloramphenicol). Cloned DNA is isolated from the cells using a modified alkaline lysis procedure which has been specially adapted for high-throughput DNA purification. A Pico Green fluorescence DNA quantization assay is done to determine DNA concentration and the template is normalized. Custom primers are ordered automatically in a 96-well plate format and three chemistries (TB, GB, GA) can be chosen for sequencing. Primer Walk reads in the Finishing Lab have, on average, an NHGRI pass rate of ~75% and average Q20 around 500.

PRIMER WALKING SEQUENCING PRINCIPLE

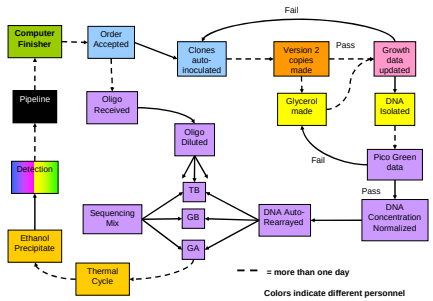
In our Finishing Group, we are interesting pOT (insert 4kb), pEpiFOS-5 (insert 40kb) and pJAN (insert 10kb) primer walk sequencing. If the gap is small, one primer walking process could be done, otherwise more primers will be designed for the future sequencing.



Both forward and reverse directions could be chosen to design the primers to generate over lapping sequence until the low coverage area or gaps are filled.



PRIMER WALKING PROCESS IN THE LAB



Pull glycerol plates:
Retrieve plasmid or fosmid subclone glycerol plates from -80°C freezers according to SQUID layout

Auto-inoculation on a Beckman MultiPROBE:
Specific clones are inoculated into a 96-well plate with medium/chloramphenicol and incubated at 37°C for 17 hours at 300rpm

Verify bacterial growth and make Version 2 copies:
To preserve glycerol stocks, clones with multiple orders are inoculated once, grown up, then reinoculated into multiple wells of a new deep well. We call this a 'Version 2' plate

Isolate template DNA using a modified alkaline lysis procedure:

1) Centrifuge cells
2) Resuspend pellets in 150uL Qiagen P1
3) Lyse cells for 5 minutes in 150uL Qiagen P2
4) Neutralize in 150uL Qiagen P3. Centrifuge to separate the supernatant and cell debris
5) Transfer the supernatant into a Whatman filter by Hydra 96 and spin the supernatant to clear plate
6) Precipitate DNA by 280uL of 99% Isopropanol and centrifugation
7) Wash DNA pellets two times with 200uL 70% cold EtOH

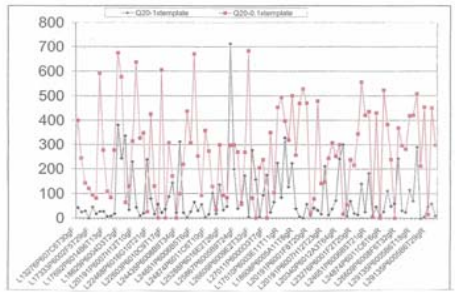
PicoGreen Fluorescence DNA Quantitation Assay
(ultra sensitive fluorescent nucleic acid stain used for quantitating double stranded DNA in solution)
The plasmid or fosmid DNA is generally 100-300ng/uL or 20-100ng/uL, respectively, in 50uL

Plate Name:	FTEM9522
	pOT
Date:	51407
Operator:	xliu

Total DNA - ng/assay well — *assay ng/μL * 120*

	1	2	3	4	5	6	7	8	9	10	11	12
A	97	199	85	146	317	230	225	203	142	79	206	164
B	575	313	178	156	361	420	323	247	342	219	367	281
C	8	326	273	242	422	497	219	202	337	333	199	154
D	125	283	203	298	184	187	248	312	244	232	242	2

DNA is generated in high concentration, a test of 300~400ng/uL (1xtemplate) vs. 30~40ng/uL (0.1xtemplate) sequencing is shown below. The average Q20 was 200 for the higher concentration vs. 400 for the lower. We now normalize our template concentration to 40 ng/uL.



DNA normalization is carried out by first pasting PicoGreen data into a spreadsheet formatted to drive a MultiPROBE. Dilution calculations are made automatically within the sheet.

FINAL CONCENTRATION ATTEMPTED (ng/uL)	4
ELUTION VOLUME PER COPY PREPPED (uL)	5
# of COPIES PREPPED	3

Copy and paste green area to notepad and save it as: FTEM9365.nrm

#FILENAME : FTEM9365.nrm							
Water=Water=96							
Dest_1=FTEM9365=96							
Position	#SPWellLocation	Sheet	DPWell	Dwell	Vol. Water (uL)	Wash	Final concentration
Water	A1	Dest_1	A1	370	Y	40	139
Water	A2	Dest_1	A2	412	Y	40	150
Water	A3	Dest_1	A3	436	Y	40	156
Water	A4	Dest_1	A4	450	Y	46	185
Water	A5	Dest_1	A5	405	Y	40	148
Water	A6	Dest_1	A6	450	Y	51	203
Water	A7	Dest_1	A7	124	Y	40	73
Water	A8	Dest_1	A8	9	Y	40	42
Water	A9	Dest_1	A9	316	Y	40	124
Water	A10	Dest_1	A10	313	Y	40	124
Water	A11	Dest_1	A11	69	Y	40	58
Water	A12	Dest_1	A12	409	Y	40	149
Water	B1	Dest_1	B1	131	Y	40	75
Water	B2	Dest_1	B2	301	Y	40	120
Water	B3	Dest_1	B3	450	Y	64	257
Water	B4	Dest_1	B4	155	Y	40	81
Water	B5	Dest_1	B5	450	Y	76	305

A Beckman MultiPROBE is used to normalize template concentration and rearray DNA into a sequencing plate.

Custom primers are delivered at a lab standard of 20uM and must be diluted to appropriate concentration before sequencing for this workflow. 14uL oligo is diluted in a 100uL final volume to get 2.8uM to use in a sequencing reaction.

For Sequencing, any of three chemistries (TB, GB, GA) can be used for sequencing. 5.5uL DNA template and 2.0uL oligo are transferred

FINP-TB + 2.5 uL/well

TB (regular Big Dye):
5x v3.1 Big Dye buffer: 1.5uL
v3.1 Big Dye terminator sequencing mix: 1.0uL

FINP-GB + 2.5 uL/well

GB (4:1 dGTP: Big Dye):
5x v3.1 Big Dye buffer: 1.5uL
v3.1 Big Dye terminator sequencing mix: 0.8uL
dGTP sequencing mix : 0.2uL

FINP-GA + 9.4 uL/well

GA (dGTP with additive A):
5x v3.0 buffer: 1.9uL
dGTP sequencing mix: 2.4uL
additive A: 5.1uL (2M 1:1 L-Proline: 4-Methylmorpholine N-Oxide)

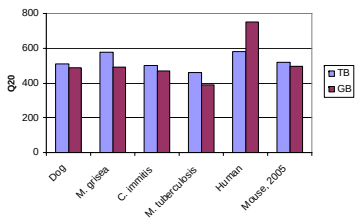
Finish up: Thermal cycling, ethanol precipitation, and EDTA resuspension for detection

Cycling	Denaturing	Annealing	Extending	Cycles	End
pOT or pJAN	96°C for 1 min then 96°C for 30 sec	50°C for 15 sec	60°C for 4 min	40	4°C for ∞
pEpiFOS-5	95°C for 30 sec	50°C for 15 sec	60°C for 2 min 30 sec	75	4°C for ∞

Ethanol cleanup	TB	GB	GA	Centrifuge	Wash	Centrifuge
Plasmid (pOT & pJAN)	28uL 85%	28uL 85%	19uL 95%	1920g 45 min	38uL 73%	1920g 20 min
Fosmid	22uL 83%	22uL 73%	19uL 95%	1920g 45 min	38uL 73%	1920g 20 min

RESULTS AND DISCUSSION

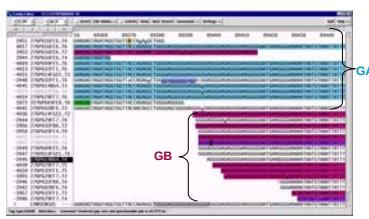
NHGRI Average Q20

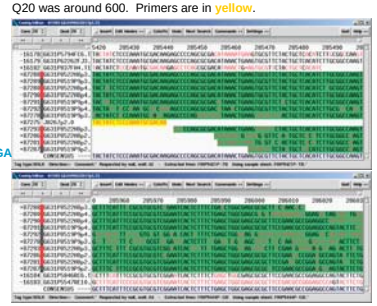


We generate about 8,000 Primer Walk reads per month.
The NHGRI average pass rate for all organisms is ~75% and the NHGRI average Q20 for TB and GB is around 500 (see left).

TB (regular BigDye) chemistry is the workhorse and adequate for may gaps. TB gives longer average read lengths while GA and GB are more suited to handle difficult sequence regions: SSRs (simple sequence repeats), GC-rich, secondary structures.

Below is a successful primer walk with TB and GB chemistries. GB successfully read through the area, which contained >50% GC. Q20 was around 600. Primers are in yellow.





GA special chemistry may work better for GC-rich and homopolymeric runs. Above, GA reads (in blue) crossed the gap. GB reads (in red) did not and prematurely terminate here.

All walk chemistry typically stop at extremely GC-rich regions (78%). Transposon insertions or small insert shatter libraries can sometimes interrupt or break up these stretches enough for complete sequencing.