ddRADseq protocol

McDaniel Lab - University of Florida
Adam Payton

(Revised May 2015)

This ddRADseq protocol is a variation on the Peterson 2012 protocol. Like all ddRadseq protocol the major steps/processes are all the same but like many labs we have made some changes to optimize the procedure for our projects, in our case for large numbers of samples while maintaining an attractive cost/benefit ratio.

Please familiarize yourself with the Peterson et. al 2012 ddRADseq paper, as well as the protocol they provide as it contains more background information and explanations of restriction enzyme selection and adaptor design, which we do not discuss here. We assume you have a solid understanding of the methodological steps. The McDaniel lab uses the restriction enzymes EcoRI and MseI, all barcodes are on the EcoRI adaptor and thus to utilize this protocol EcoRI must be one of the restriction enzymes used.

It is possible to use this protocol for both single end and paired-end Illumina sequencing on any of the common Illumina platforms (HiSeq, NextSeq, MiSeq).

Disclaimer: This protocol will generate massive quantities of data, ie. several hundred million sequence reads with stochastic variation in both numbers of reads per individual and in read depth at any given locus. To successfully use these data you need to run the pipeline for processing and analyzing the data. If you are unsure how to handle >40gb files containing hundreds of millions of sequence reads or the analysis of these data, this method will not be useful to you.

Primary reagents used in this protocol:

EcoRI-HF	(NEB, R3101 20,000 units/ml)
MseI	(NEB, R0525 10,000 units/ml)
T4-DNA ligase	(NEB, M0202 400,000 units/ml)
ATP	(NEB, P0756 10mM)
dNTP	(NEB, N0447 8um each nucleotide)
Q5 high fidelity DNA polymerase	(NEB, M0491 2,000 units/ml)
OneTaq standard PCR mastermix	(NEB, M0486)
Wizard SV Gel and PCR clean-up system	(Promega, A9281)

Note: access to a Pippin ELF or Pippin Blue precision electrophoresis machine is essential to the success of this protocol. All other necessary equipment should be easily found in a standard molecular biology lab.

Adaptor Hybridization:

Individual complementary single stranded oligos are hybridized in this step to form double stranded adaptors which will contain an overhanging enzyme cut-site sequence (a "stick-end"), the associated barcode, and a priming site for the Illumina flowcell and sequencing primer (Fig. 1). This hybridization is preformed in a high salt environment using a thermocycler to control the hybridization process.

10x hybridization buffer = 100mM Tris-HCL ph 8, 500mM NaCl, 10mM EDTA

For 100ml:

10mL	1M Tris-Hcl ph8
10mL	5M NaCl
0.3722g	EDTA
80mL	Filtered Water

To make a 10um stock of hybridized (double-stranded) adaptor combine:

10ul	adaptor_1 (100um)
10ul	adaptor_1.b (100um)
10ul	10x annealing buffer
70ul	water

Repeat for all adaptors needed, mix well either by pipetting or using a plate vortexer, place in a thermocycler and incubate at 97.5°C for 2.5 minutes then cool at a rate not greater than 3°C per minute until the solution reaches 21C. Hold at 4°C.

The hybridization needs to be performed on all 5' and 3' adaptor pairs (EcoRI and MseI). This will create a 10um stock solution of hybridized adaptors which will serve as a stock solution that will be diluted depending on the concentration of adaptor that is needed given the amount of DNA present in each individual (discussed in more detail in adaptor ligation section). Once this hybridization has been done these can be stored frozen for long periods of time.

For a detailed explanation of adaptors see Peterson et al 2012.

Enzyme Digestion

DNA is digested by both restriction enzymes in a single reaction. The enzymes in this protocol are compatible for double digest reactions, ie they use the same reaction buffer and have the same optimal operating temperature. Use New England Biolab's "Double Digest Finder" if other enzymes are used to ensure enzymes are compatible.

DNA should be at a concentration of at least 20ng/ul. However regardless of the concentration of DNA all samples should be normalized so all are at a similar concentration. Normalization is important because the quantity of DNA used in a given sample is highly correlated with the number of reads and loci that a sample will generate when sequenced.

Unequal quantities of starting DNA will result in unequal number of reads and loci between samples.

1) Place 6ul of each sample's DNA in the well of a PCR plate, store on ice.

2) Prepare Master Mix 1:
For each reaction:

0.9ul	Cutsmart 10x buffer
0.28ul	EcoRI enzyme
0.12ul	MseI enzyme
1.7ul	H2O

Mix well, centrifuge, and store on ice. Due to the small volumes used and the viscous nature of the glycerol the enzymes are stored in we recommend making at least a 130% excess of master mix 1 to accommodate multiple rounds of pipetting.

3) Add 3ul of Master Mix 1 to each sample's DNA.

4) The total reaction volume is now 9ul. Seal samples, vortex, centrifuge, and incubate at 37C for 8 hours on a thermocycler with a heated lid set to 50C.

Adaptor Ligation

During this step double stranded adaptors are ligated onto the sticky ends resulting from the enzyme digest (Fig. 1). The EcoRI adaptors contain a unique barcode sequence and the Mse1 adaptor is universal and the same across all samples (unless paired end sequencing is desired and you have Mse1 adaptors that contain spacer sequences).

Thaw the double stranded EcoRI and MseI adaptors made previously.
Calculate the quantity of adaptor necessary given the quantity of starting DNA, the frequency of the enzyme cutsites in the genome, and the desired adaptor excess. This can easily be done using the excel spreadsheet provided by Peterson 2012 (INSERT LINK HERE). Use the stock adaptors to make dilutions so that your calculated molar concentration of adaptor is contained within 1ul of solution, ie. we only want to add 1ul of the adaptor to the digested DNA.

Perform all steps with samples on ice.

1) Add 1ul of the diluted EcoRI adaptor directly to the digested DNA
1b) If performing paired end sequencing add 1ul of the appropriate barcoded Mse1 adaptor. If performing single read sequencing add the MseI adaptor to MM2.
2) Prepare Master Mix 2:
For each reaction:

0.4ul	Cutsmart buffer 10x
1.3ul	ATP 10um
0.2ul	T4 Ligase enzyme
0.1ul	H2O
1.0ul	MseI adaptor (add to MM2 if preforming single read sequencing)

3) Add 3ul of MM2 (or 2ul of MM2 if preforming paired end sequencing) to the digested DNA. Due to the small volumes used and the viscous nature of the glycerol the enzymes are stored in we recommend making at least a 130% excess of master mix 2 to accommodate multiple rounds of pipetting.

4) The total reaction volume is now 14ul. Seal the samples, vortex, centrifuge, and incubate at 16C for 6 hours on a thermocycler with a heated lid set to 50C.

After ligation the samples can be stored frozen for a few weeks if needed.

PCR 1

This step will be used to test the success of library construction of each individual sample (Fig. 4). This PCR reaction uses primers that contain the Illumina flow cell binding sequences, thus allowing the incorporation of these sequences onto the ends of the DNA fragments (Fig. 1). Again this PCR is just a test to verify if the library had both adaptors successfully ligated to their respective ends of the fragment, this PCR product will not be used for any Illumina sequencing. This PCR can be preformed with inexpensive non-high-fidelity Taq.

For each reaction:

8ul	NEB One-Taq 2x Master Mix
0.8ul	10um F and R Illumina primers (premixed)
6.2ul	H2O
1ul	DNA Restriction/ligation product

Run PCR using the following conditions on a thermocycler with a heated lid:
94C for 2min then 20 cycles of (94C for 30sec, 60C for 30sec, 68C for 45sec) 4-10C hold.

Agarose Gel

Run the products of PCR1 for each sample on a dense (1.4-2%) agarose gel with a 100bp ladder (Fig. 4). Successful amplification will result in a smear of fragments typically ranging from ~50bp to ~700bp. Failed libraries will produce amplification products of only the primers and adaptors creating a smear of fragments typically ranging from ~50bp to ~200bp. Note the fragment range that produces the highest density of fragments, this can be helpful when determining the fragment range to capture during the size selection.

Once the gel is successfully run there is no need to keep the PCR products.

Sample cleanup and Size Selection

1) Combine 3-5ul of the restriction ligation product from every sample that produced successful PCR amplification in PCR1 into a single tube.

2) It is now necessary to clean up the pooled sample to remove the digestion and ligation enzymes and buffers prior to size selection. This step also helps concentrate the sample as a small volume is necessary for loading the gel in size selection.

Estimate the quantity of DNA present in your pooled sample based, often this can be done based on the average quantity of DNA used in the initial digestion, plus the quantity of adaptor added.

Using a PCR cleanup spin column follow the manufactures directions for DNA cleanup being absolutely sure the binding capacity of the column is not exceeded by your sample. If the sample contains more DNA than a single column can handle use multiple columns and pool the final elutions.

NOTE: We have the following steps performed by our CORE genomic facility: concentration of restriction/ligation products, Pippin size selection, and Bioanalyzer evaluation.

Concentrate/precipitate the DNA so that it can be loaded into either a Pippin ELF or Pippin Blue instrument for precise size selection, agarose gel based size selection is not sufficiently accurate for RADseq and we strongly discourage its use. We recommend a Pippin ELF as it allows for fractioning of a sample so you generate elutions ranging from, for example, 250-400, 400-550, 550-700. This allows you the flexibility to choose which fractions to use or to combine after you have run the second PCR (Fig. 1 & 2).

The fragment range chosen will depend on several criteria:
Are you running paired end or single end sequencing.
The length of the read you are generating during sequencing.
The size of your organism's genome.
Where the highest density of fragments were in PCR1.

Remember that you are size selecting fragments that have already had the adaptors ligated on but not the flow cell binding sequences, these are added during PCR2 after size selection.

We also recommend that you run the size selected fractions on a Bioanalyzer, in order to get a precise concentrations estimate of each fractions as well as an evaluation of how successful the size selection was (Fig. 1 & 2).

PCR2

PCR2 is run using the size selected fraction(s) and is necessary to incorporate the flowcell biding sequence onto the size-selected barcoded DNA fragments (fig. 1).
In order to help mitigate stochastic variation associated with PCR amplification we recommend running ~20 separate PCR reactions on the size selected fraction(s) and then combining those PCR products for sequencing.

This PCR should be run using high-fidelity Taq, such as NEB's Q5.

For each reaction:

3ul	Q5 buffer 5x
0.3ul	DNTP
1ul	Primers F&R combined 5um
0.15ul	Q5 Taq
9.55-7.55ul	H20 (depending on the volume of DNA added)
1-3ul	Size selected DNA (volume dependent on concentration)

These reactions have a final volume of 15ul, adjust components proportionally if larger volumes are desired.

Run the PCR with the following conditions on a thermocycler with a heated lid:
98C for 30 sec, 8-12 cycles of (98C for 15sec, 60C for 30sec, 72C for 30sec), hold 4-10C.
Denaturing and extension temperatures are based on the recommendations from NEB when using Q5 DNA polymerase.

The number of PCR cycles is deliberately low, experimentation may be necessary to determine the best number of cycles for a give sample. Generally 10 cycles is sufficient.

Agarose Gel

Run 2ul of the products of PCR2 on a 1.4-2% agarose gel with a 100bp ladder. Successful amplification should produce a faint smear of the fragment range that was targeted during size selection. If your PCR product is bright, consider running fewer cycles, alternatively if it is very faint or visibly non-existent increase the number of cycles

Pool all successful PCR2 reactions into a single tube. At this point the sample is delivered to the sequencing core facility explaining the sample is in need DNA concentration, a 1.5X Ampure purification to remove the unincorporated primers, and other associated library validations prior to sequencing (eg. Quantification, qPCR, ect).

References:

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLoS ONE 7(5): e37135. doi:10.1371/journal.pone.0037135

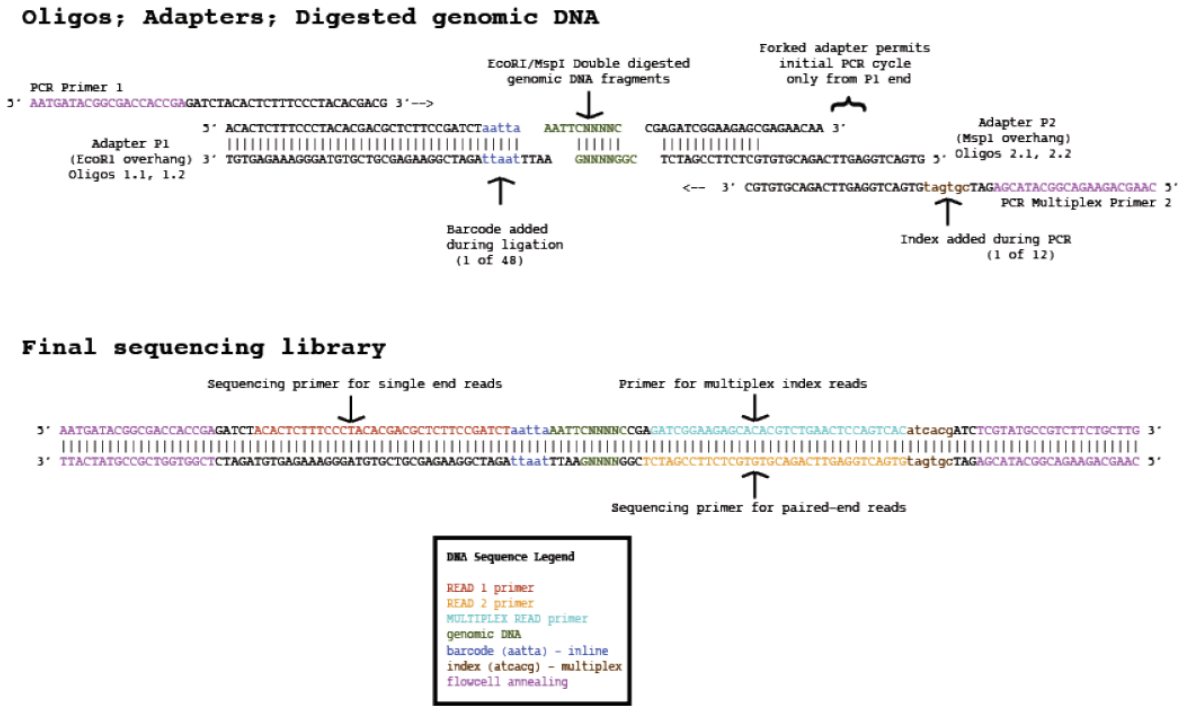


Fig. 1. From Peterson 2012. Schematic showing how the adaptor sequences, the Illumina primers, and the digested DNA are assembled into a final read for sequencing.

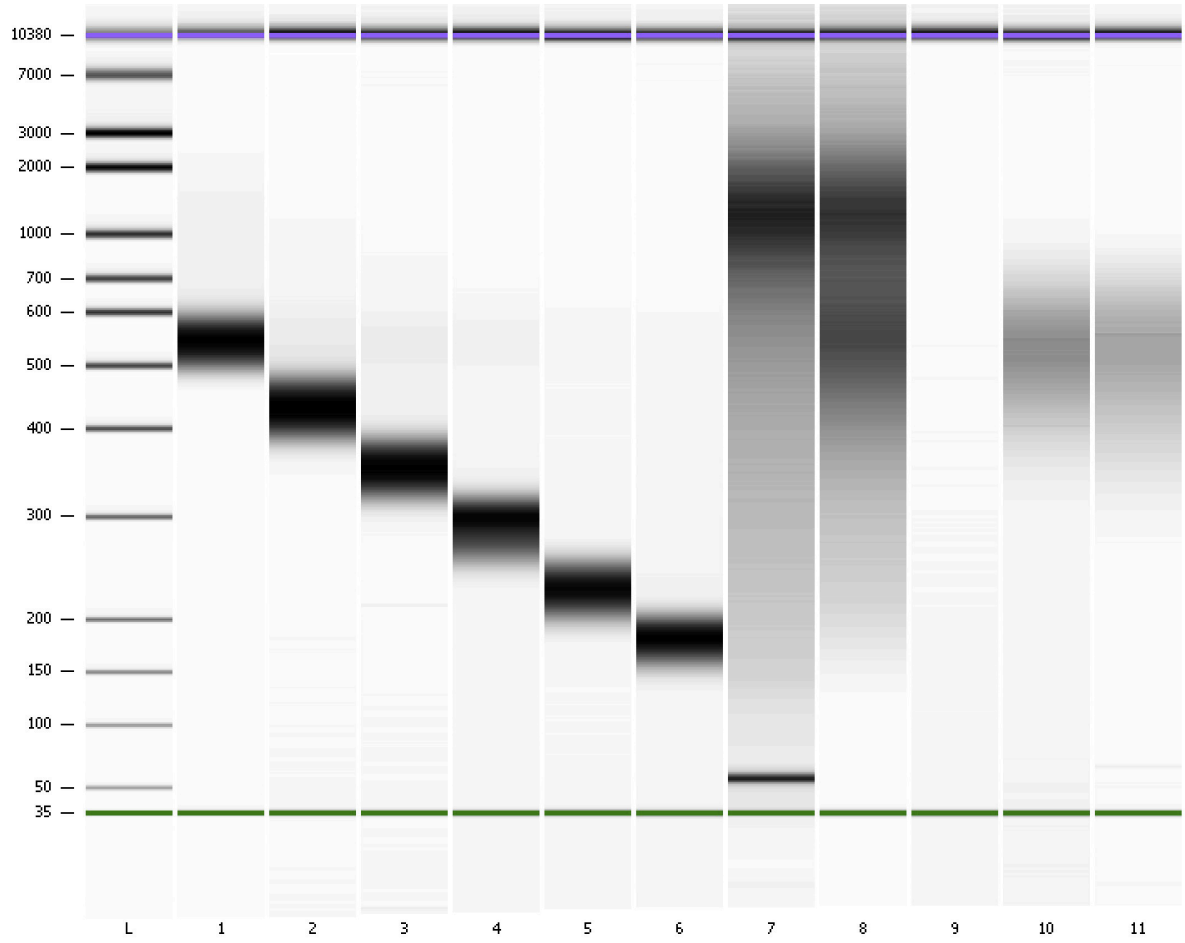
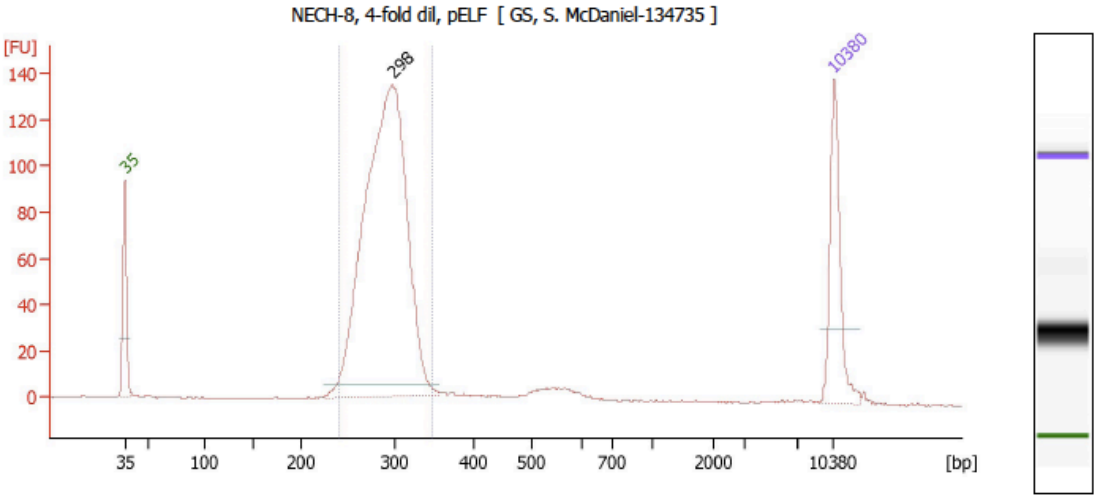


Fig 2. Example of a Bioanalyzer output from fractions of a Pippin ELF size selection. Note the stair step nature of each fraction and the minimal overlap between fractions.

Assay Class: High Sensitivity DNA Assay
Data Path: C:\...gh Sensitivity DNA Assay_DE13701186_2015-04-07_09-54-11.xad

Created: 4/7/2015 9:54:10 AM
Modified: 4/7/2015 11:18:42 AM

Electropherogram Summary Continued ...



Overall Results for sample 4 : NECH-8, 4-fold dil, pELF

Number of peaks found:	1	Corr. Area 1:	1,033.5
Noise:	0.1		

Peak table for sample 4 : NECH-8, 4-fold dil, pELF

Peak	Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	35	125.00	5,411.3	Lower Marker
2	298	796.99	4,049.4	
3	10,380	75.00	10.9	Upper Marker

Region table for sample 4 : NECH-8, 4-fold dil, pELF

From [bp]	To [bp]	Corr. Area	% of Total	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/µl]	Molarity [pmol/l]	Color
241	346	1,033.5	89	291	7.2	822.23	4,310.1	

Fig 3. Representative example of a Bioanalyzer profile of a Pippin ELF size selected fraction.

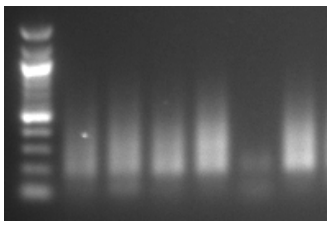


Fig 4. Agarose gel of 6 libraries after PCR1. The 5th sample illustrates what a failed sample looks like, all other samples are successful library constructions. 1.4% agarose gel 100bp ladder.