ddRADseq Protocol (Rheindt Lab 2014, National University of Singapore)

-modified for pangolin DNA by Helen Nash & Wirdateti

1. Adapter preparation

Materials:

- 1x Duplex buffer
- Sterile PCR tubes
- Oligos (from Peterson et al. 2012):
  - P1.1E, P1.2E (to make P1E adapters) => 200uM stock
  - P2.1M, P2.2M (to make P2M adapters) => 200uM stock

Desired products:

5uM P1 adapter working stock, 5uM (with EcoRI cut site)

5uM P2 adapters working stock, 50uM (with MspI cut site)

* *Note that the working stock final concentrations are DIFFERENT.*

To get these, we will need to anneal P1.1E + P1.2E adapters and P2.1M + P2.2M adapters together respectively.

Making the 200uM oligo stock

Resuspend lyophilized oligos in 1x Duplex buffer to make the 200uM stock Equation for volume of duplex buffer to add:
amount of duplex buffer (ul) = [amount of oligo(nmol) x 10]/2

STEP-BY-STEP GUIDE

Step 1. Combine the following reagents, in two separate PCR tubes:

For P1E adapter working stock:	For P2M adapter working stock:
• 200uM stock P1.1E adapter, 5ul • 200uM stock P1.2E adapter, 5ul • 1x Duplex buffer, 190ul	• 200uM stock P2.1M, 25ul • 200uM stock P2.2M, 25ul • Duplex buffer, 50ul
Total 200ul of 5uM working stock	Total 100ul of 50uM working stock

Step 2. To anneal the adapters, heat the oligo mixes to 97.5°C for 2.5 min, then reduce temperature by 2°C per minute until it reaches 15°C. Hold at 4°C. We used model C1000 Bio-Rad Thermocycler.

Step 3. Store annealed primers at -20°C.

2. Sera-Mag beads for size selection

Buffer ingredient list:
18% PEG-8000
1M NaCl
10mM Tris-HCL, pH 8.0
1mM EDTA, pH 8.0

**To prepare 20ml stock of 0.1M EDTA stock first: Dissolve 0.9306g EDTA salt in 25ml water.*

Easy-read table for making bead buffer

	Stock conc.	Dilution factor	For 50ml buffer	For 500ml buffer	Final working conc
PEG-8000	Powder	-	9g	90g	18% w/v
NaCl	Powder	-	2.922g	29.22g	1M
Tris-HCl, pH 8	1M	100x	500ul	5ml	10mM
EDTA-disodium salt	0.1M	100x	500ul	5ml	1mM
Sterile water	-	-	Top up to 50ml	Top up to 500ml	

As the original buffer contains sodium azide, we need to prepare the beads for use.
Wash the Sera-Mag beads twice with equal volume of sterile water.

STEP-BY-STEP GUIDE

Step 1. Wait until the bead solution has reached room temperature, then mix the bead solution very thoroughly. Ensure that there are no bead pellets and that the solution is uniform. This may take some time.

Step 2. Pipette out the desired volume of beads. This volume has to be large enough for the ratio testing as well as the actual NGS library preparation.

Step 3. After the bead pellet has formed, pipette out as much of the clear liquid as possible without disturbing the bead pellet.

Step 4. Add in the sterile water. Remove the tube from the magnetic stand, and mix thoroughly. Replace on the magnetic stand.

Step 5. Repeat the wash twice, then resuspend the beads in the prepared buffer instead of water.
For our pangolin runs we prepared 1ml of beads plus 50ml of buffer so total 51ml.

Step 6. Combine the cleaned 5% bead solution into the same Falcon tube.

3. Bead ratio selection

The purpose of this step is to determine the optimal bead to DNA solution ratios that should be used in order to obtain the targeted bead fragment sizes.
The target DNA fragment size for Illumina RAD-Seq runs is between 250-650bp.

The steps taken are the same as for the bead selection size during the actual Library prep, just that DNA ladder is used instead of sample DNA, and different pairs of bead ratio (eg. 0.68x to 0.85x, 0.70 to 0.90x, etc) are tested to discover which ratio is the most optimal.

Materials list

- Molecular grade water
- 1 strip tube of washed Sera-Mag beads, suspended in PEG buffer
- 8-well strip tubes/PCR plate columns with lids (8 samples = 3 columns, 16 = 6 col, 24 = 9 col, 32 = 12 col, 40 = 15 col, 48 = 18 col)
- 1 strip tube/boat of 85% ethanol
- Magnetic plate
- DNA ladder (made by mixing 2ul of DNA ladder with 48ul of molecular grade water); 50ul for each bead ratio pair tested.

**Note: ensure the bead mixture is at room temperature before mixing it with the DNA.*

**Note: ensure that the bead mixture is <u>very well mixed/completely uniform</u> before use.*

Guide to calculating volumes of bead solution/DNA ladder to use

*Complete all these calculations before actually starting on the bead cleanup and selection step.

Example: <u>0.75x - 0.92x</u>.

Total volume of DNA solution for each sample = 50ul

Volume of beads to add to the DNA ladder (1st round): = 0.75 x 50 = <u>37.5ul</u>.

Total volume per bead+ DNA ladder = vol. of beads added + DNA ladder volume = 37.5 + 50 = <u>87.5ul</u>.

Volume of 1st round supernatant to transfer after first round of incubation = Total volume of bead and DNA ladder -5 = 87.5 - 5 = <u>82.5ul</u>

Volume of beads to add to 1st round supernatant = (higher ratio - lower ratio) x 1st round supernatant volume = (0.92-0.75) x 82.5 = <u>14.025ul</u>

STEP-BY-STEP GUIDE

Step 1. Mix beads thoroughly, ensuring that there are no bead pellets and the solution is homogeneous.

Step 2. Carefully mix the DNA ladder. Pipette 2ul of ladder and 48ul of molecular grade water into each of the test wells.

Step 3. Add the requisite volume of beads for the 1st round of size selection into each well. Mix carefully.

Step 4. Incubate at r.t.p. for 45 minutes. DO NOT place the wells/PCR plate on the magnetic plate during incubation.

Step 5. Place the wells/PCR plate on the magnetic plate and wait for the solution to clear. The beads will form a pellet on one side of the wells. The fragments we are interested in are in the supernatant.

Step 6. Carefully remove the calculated volumes of supernatant from the 1st round wells, and transfer them into the 2nd round wells.

Step 7. Add the required volume of beads to the 2nd round wells. Mix carefully.

Step 8. Incubate at r.t.p. for 45 minutes. DO NOT place the wells/PCR plate on the magnetic plate during incubation.

Step 9. Place the wells/PCR plate on the magnetic plate and wait for the solution to clear. The beads will form a pellet on one side of the wells. The fragments we are interested in are now on the beads.

Step 10. Remove as much of the supernatant as possible, while being careful not to disturb the bead pellets. Discard the supernatant. Alternatively, you can keep the supernatant for visualization on gels.

Step 11. Wash the beads carefully in 85% ethanol. Place the wells/PCR plate on the magnetic plate, and remove the ethanol after washing, while being careful not to disturb the bead pellets. Do this washing two times. *We used 150ul 85% ethanol per wash.*

Step 12. Allow the beads to air-dry at room temperature. This usually takes between 3 to 30 minutes, depending on how much of the ethanol you managed to remove. Be very careful at this stage, as the wells have to be uncovered in order for the beads to dry. Do not allow contamination to occur.

Step 13. Re-suspend the beads in 10ul of molecular grade water. The size-selected DNA from the ladder is now in the water.

Step 14. Make 2% agar gels, and run the bead selection results. This will allow you to determine which bead ratio pair works best for you.

For our pangolin runs, we tested numerous ratios. The best ratio we discovered which gave the clearest gel band of 250 to 650bp fragments is given in yellow below. Therefore, we used 39ul and 11.76ul later in the protocol for bead/buffer volumes for the real pangolin DNA (see later Part 2: Sera-Mag Bead clean-up and size selection.

Low ratio (1st round)	*High ratio (2nd round)*	*DNA Ladder (ul)*	*Low volume of bead/buffer (ul) First Round wells*	*Supernatant (ul)*	*High volume bead/buffer (ul) Second round wells*
0.75	*0.87*	*50*	*37.5*	*82.5*	*9.9*
0.75	*0.9*	*50*	*37.5*	*82.5*	*12.375*
0.75	*0.95*	*50*	*37.5*	*82.5*	*16.5*
0.75	*0.98*	*50*	*37.5*	*82.5*	*18.975*
0.7	*0.92*	*50*	*35*	*80*	*17.6*
0.75	*0.92*	*50*	*37.5*	*82.5*	*14.025*
0.78	*0.92*	*50*	*39*	*84*	*11.76*
0.82	*0.92*	*50*	*41*	*86*	*8.6*

Part 1) RELIG reactions

PREPARATION

Materials list:

- Excel printout with sample list (DNA concentrations + volumes, molecular grade water volumes)
- 8-well strip tubes/PCR plate columns with lids (8 samples = 4 columns, 16 = 8 col, 24 = 12 col, 32 = 16 col, 40 = 20 col, 48 = 24 col)
- Eppendorf tubes
- Primers to be used (P2M, and P1E = 48 different ones) *NB number 16 in the oligos list of Peterson et al. 2012 didn't work so we only used 47.*
- T4 DNA Ligase and buffer
- EcoRI-HF
- MspI
- NaCl (0.5M)
- BSA (10mg/ml)
- DNA samples

Always label both lids and wells of the PCR plates/strip tubes. Number the columns, and indicate what the well will contain (Replicate 1/2, -ve control, gDNA), as well as the date.

Arrangement of wells: Replicate 1 columns, Replicate 2 columns, -ve controls, gDNA

Reagent amounts (PER SAMPLE REPLICATE or CONTROL)

Reagent	**Volume**	**Master mix**
Molecular grade water	1.235 ul *x3 = 3.705ul*	**MM1**: Prepare for 3n++ worth of reactions. Separate out 2n+ worth to make MM2.
0.5M NaCl	1.3 ul *x3 = 3.9ul*	
10 mg/ml BSA	0.065 ul *x3 = 0.195ul*	
10x T4 DNA ligase buffer	1.4 ul *x3 = 4.2ul*	
50uM P2M adapter	1.0ul *x3 = 3ul* *0.5 =1.5ul	
EcoRI-HF	0.25 ul *x2 = 0.5ul*	**MM2**: Add 2n+ worth of these to the 2n+ aliquot of MM1.
MspI	0.05 ul *x2 =0.1ul*	
T4 DNA Ligase	0.2 ul *x2 = 0.4ul*	
5uM P1E adapter	1.0 ul x3.5 =3.5ul*0.5 to avoid primer dimers we halved the adapters =1.75ul	**D**: Make 3.5x reactions worth of these.
gDNA	50 ng (≤6.5 ul) Due to the varied concentrations of each DNA sample, the ul of DNA required per sample to reach 50ng is unique to each sample. Calculate in an Excel spreadsheet.	
Molecular grade water	Use to top up DNA to 6.5 ul. Use an Excel spreadsheet.	
Total	13 ul	

Guide to calculating MM volumes

*Multiply number of reactions required by the volumes per reagent for each of the reagents in the table above to get the total amounts you need. Extra is factored in to account for pipetting errors.

No. of reactions' worth of **MM1** to make:

8 samples = [8x3]+5 extra = 29x

16 samples = [16x3]+10 extra = 58x

24 samples = [24x3]+15 extra = 87x

32 samples = [32x3]+20 extra = 116x etc....

No. of reactions worth of **MM1** to pipette out for **MM2**:

8 samples = [8x2]+ 2 extra = 18x

16 samples = [16x2]+5 extra = 37x

24 samples = [24x2]+10 extra = 58x

32 samples = [32x2]+ 25 extra = 79x etc....

MM1 goes into the wells for Replicate 1, Replicate 2, and the negative controls. It contains NaCl, BSA, P2M adapter, T4 DNA Ligase buffer, and water.

MM2 goes into the wells for Replicate 1 and Replicate 2. It contains the restriction enzymes necessary for the RELIG reaction.

D: Make 3.5x reactions' worth of D (2x replicates, 1x negative control, 0.5x for pipette error).
D is made first in the gDNA wells, then transferred into the Rep. 1, Rep. 2, and negative control wells. It contains the gDNA, P1E adapters, and water. Each sample will have different volumes of gDNA and water, depending on the starting concentration of your DNA samples. The protocol requires a minimum of 40 ng of DNA; this translates into a minimum gDNA concentration of 6.15 ng/ml for this step. The total amount of DNA should be kept the same across all the samples.

**Note: calculate all volumes before starting the actual labwork, and have it displayed clearly.*

**Note: always add the enzymes in last if possible, and take them out of the freezer for as short a time as possible. When out of the freezer, always keep them in ice*.

**Note: for pangolin samples we used 50ng of DNA per sample.*

STEP-BY-STEP GUIDE

Step 1. Make **D** according to the volumes of sample DNA and molecular grade water in your Excel printout, and then add the P1E adapters. Each sample/lane needs to receive its own P1E adapter. There can be no P1E adapter repeats within each lane.

Step 2. Make **MM1** according to the volumes calculated during preparation. Add reagents in the following order: water, NaCl, BSA, T4 DNA Ligase buffer, P2M adapter. Mix well.

Step 3. Pipette out 5.0ul of **MM1** into the wells for the negative controls ONLY.

Step 4. Pipette out the requisite volume of **MM1** into a separate sterile Eppendorf tube. This will be used to make **MM2**.

Step 5. Add the requisite volumes of enzymes (EcoRI and MspI) to the **MM2** tube. Mix gently.

Step 6. Pipette out 5.5ul of **MM2** into wells for Replicates 1 and 2.

Step 7. Mix solution in wells containing **D** gently. Pipette 7.5ul into wells for Replicates 1, 2, and negative control.

Step 8. Incubate the reaction mixture at 37°C for 3.5 hours. We used a Kryatec Supercycler for the incubation.

Step 9. Store at 4 deg C.

Part 2: Sera-Mag Bead clean-up and size selection

PREPARATION

From bead ratio selection (done prior to RELIG), you already know the optimal bead/buffer:DNA ratios you need to use to select the desired DNA fragment length of 250-650bp (including adapter lengths of 76+ bp).

For pangolin samples we used:

Low ratio (1st round)	*High ratio (2nd round)*	*DNA Ladder (ul)*	*Low volume of bead/buffer (ul) First Round wells*	*Supernatant (ul)*	*High volume bead/buffer (ul) Second round wells*
0.78	*0.92*	*50*	*39*	*84*	*11.76*

Although later we discovered that some of the pangolin samples worked better with this alternative ratio:

Low ratio (1st round)	*High ratio (2nd round)*	*DNA Ladder (ul)*	*Low volume of bead/buffer (ul) First Round wells*	*Supernatant (ul)*	*High volume bead/buffer (ul) Second round wells*
0.75	*0.88*	*50*	*37.5*	*82.5*	*10.725*

Materials list

- 1 strip tube of molecular grade water
- 1 strip tube of Sera-Mag beads, suspended in PEG-8000
- 8-well strip tubes/PCR plate columns with lids (8 samples = 3 columns, 16 = 6 col, 24 = 9 col, 32 = 12 col, 40 = 15 col, 48 = 18 col)
- 1 strip tube/boat of 85% ethanol
- Magnetic plate
- Adapter-ligated sample DNA (topped up to 50ul volume with molecular grade water) - FROM PART (1)

Note: ensure the bead mixture is at room temperature before mixing it with the DNA.

Note: ensure that the bead mixture is very well mixed/completely uniform before use.

Label all columns/wells.
Arrangement: "Ist round of bead selection" columns, "Discarded supernatant" columns, "2nd round of bead selection" columns

STEP-BY-STEP GUIDE

Step 1. Mix beads thoroughly, ensuring that there are no bead pellets and the solution is homogeneous.

Step 2. Pipette 32 ul of water into the "1st round" wells, then pipette in the required volume of bead mixture (following the example above, this would be 37.5ul of bead mixture). Pipette the required volume of bead mixture into the "2nd round" wells (following the example above, this would be 10.725ul of bead mixture). This order of adding reagents minimizes contamination, while allowing us to save expensive filter tips.

Step 3. Pipette 9ul of adapter-ligated DNA samples from each of the 2 replicates (from the RELIG step)into the "1st round" wells. Carefully mix the samples.

Step 4. Incubate at r.t.p. for 45 minutes. DO NOT place the wells/PCR plate on the magnetic plate during incubation. *For pangolins we modified this first incubation to 40mins.*

Step 5. Place the wells/PCR plate on the magnetic plate and wait for the solution to clear. The beads will form a pellet on one side of the wells. DNA fragments which are larger than 650bp are bound to the beads, while DNA fragments between 0 to 650 bp in length are in the supernatant. The fragments we are interested in are in the supernatant. *For pangolins we modified this first wait time on the magnetic stand to 10mins.*

Step 6. Carefully remove the calculated volumes of supernatant from the "1st round" wells (e.g. 82.5ul), and transfer them into the "2nd round" wells. The "2nd round" wells should already contain the required volumes of bead mixture.

Step 7. Incubate at r.t.p. for 45 minutes. DO NOT place the wells/PCR plate on the magnetic plate during incubation. *For pangolins we modified this second incubation to 25mins.*

Step 8. Place the wells/PCR plate on the magnetic plate and wait for the solution to clear. The beads will form a pellet on one side of the wells. DNA fragments which are between 250bp to 650 bp in length are bound to the beads, while DNA fragments smaller than 250bp are in the supernatant. The fragments we are interested in are now on the beads. *For pangolins we modified this second wait time on the magnetic stand to 7mins.*

Step 9. Remove as much of the supernatant as possible, while being careful not to disturb the bead pellets. Discharge the supernatant into the "Discarded supernatant" columns.

Step 10. Wash the beads carefully in 85% ethanol. Place the wells/PCR plate on the magnetic plate, and remove the ethanol after washing, while being careful not to disturb the bead pellets. Do this washing two times. *We used 150ul 85% ethanol per wash.*

Step 11. Allow the beads to air-dry at room temperature. This usually takes between 3 to 30 minutes, depending on how much of the ethanol you managed to remove. Be very careful at this stage, as the wells have to be uncovered in order for the beads to dry. Do not allow contamination to occur.

Step 12. Re-suspend the beads in 10ul of molecular grade water. The size-selected, adapter-ligated DNA template is now in the water. You can now either remove this supernatant into fresh tubes, or remove it only when doing the PCR step.

Part 3: PCR enrichment of fragments

Three PCR reactions and one -ve control per sample are performed to reduce the chances of PCR bias.

Q5 HF DNA Polymerase (hot start) is used in this protocol.

PREPARATION

Materials list

- 8-well strip tubes/PCR plate columns with lids (8 samples = 4 columns, 16 = 8 col, 24 = 12 col, 32 = 16 col, 40 = 20 col, 48 = 24 col)
- Eppendorf tubes for making the Master Mix (MM)
- 5x Q5 reaction buffer
- dNTPs (10mM total; 200uM each dNTP)
- PCR1 primer (made 200ul of stock) *Then took 10ul of the stock plus 390ul of molecular grade water to make 400ul of 5uM.*
- PCR2 primers (2.1, 2.5, 2.7; made 200ul of stock of each index 1, 5 and 7) *Only needed index 1 and 5 for our pangolin samples. Then took 10ul of the stock plus 390ul of molecular grade water to make 400ul of 5uM, for each index.*
- Q5 HF DNA Polymerase (2U/ul)
- Adapter-ligated DNA template - FROM PART (2)
- Molecular grade water

Reagent amounts (PER SAMPLE REPLICATE or CONTROL)

Reagent	**Volume**	**Master mix**
5x Q5 reaction buffer	2.0 ul x4 =8ul	**MM**: Prepare for 4n++ reactions. Make MM in a Eppendorf tube.
10mM total dNTPs	0.8ul x4 =3.2ul	
5uM PCR 1 Primer	0.4ul x4 =1.6ul	
5uM PCR 2 Primer	0.4ul x4 =1.6ul	
Q5 HF DNA Polymerase	0.1ul x4 =0.4ul	
Molecular grade water	3.8 ul x4 =15.2ul	
Adapter-ligated DNA template	2.5ul	FROM PART (2)
Total	**10ul**	

Label all columns/wells.
Arrangement: Replicate 1 columns, Replicate 2 columns, Replicate 3 columns, -ve control columns

STEP-BY-STEP GUIDE

Step 1. Calculate the reagent volumes required for making the MM. Make sure the calculations are clearly written and easily accessed.

Step 2. Carefully mix your adapter-ligated DNA template, and pipette 2.5ul of it into the wells for Replicates 1, 2 and 3. As always, ensure that the wells/PCR plates are sterile, and make sure contamination does not occur (leave wells uncovered for as short a time as possible etc).

Step 3. Add 2.5ul of molecular grade water into each of the -ve control wells.

Step 4. Make the MM in a clean, sterile Eppendorf tube. Before adding each reagent, mix the reagent carefully but thoroughly. Add the Q5 HF DNA Polymerase last. As it is an enzyme, keep it in the freezer/on ice at all times, and only remove it from the cold for as short of a time as possible.

Step 5. Carefully but thoroughly mix the MM, then pipette 7.5ul of MM into each well (Replicates 1, 2, 3, and the -ve control).

Step 6. Place all the wells into the PCR machine, and run the following protocol. We used Bio-Rad Thermocylcer C1000.

PCR protocol for Q5 HF DNA Polymerase:

Hotlid 105, 30

Volume 10

Temp 98 °C for 30s

Temp 98 °C for 10s
Temp 64 °C for 30s } 12 cycles
Temp 72 °C for 30s

Go to 2, 11

Temp 72 °C for 300s

Temp 12 °C, 0

End

Step 7. Store the PCR-enriched samples at 4°C.

Part 4: Further size purification of PCR products

Repeat **Part 2: Sera-Mag Bead cleanup and size selection** on the PCR product with the following differences:

- Pool 6ul x 3 (from each of the PCR triplicates) into the "1st round" wells
- during the final DNA elution step, elute in 16ul of sterile water instead of 10ul.

This round of bead selection is to ensure that no fragments outside of the range we are interested in, which may have been amplified during PCR, remain within the samples.

Part 5: Pooling of samples

The samples will have to be run through:

i) the AATI Fragment Analyzer machine, to ascertain the DNA fragment sizes
ii) Qubit BR dsDNA kit for quantification.

Each of these steps requires the use of 2ul of each sample.

Create an Excel sheet containing the DNA concentrations of each sample and the average DNA fragment size of each sample.

Part 6: Summary

We modified a protocol for ddRADseq (Peterson et al. 2012) following Tay et al. (2016) and used the restriction enzymes EcoRI-HF and MspI because they worked well for other taxa (Garg et al. 2016; Ng et al. 2017). During optimisation of the RADseq protocol, we selected a Sera-Mag® bead ratio that produced DNA fragments within a range of 250–650 base pairs (bp). Following optimisation, we took 50ng of DNA from each pangolin sample, digested each one separately with the two restriction enzymes, and ligated all DNA fragments with two oligonucleotides, including a unique barcode per sample, plus a batch index code. Next, we size-selected fragments using the optimal bead ratio, and amplified them via a polymerase chain reaction (PCR). After PCR, we used a second round of size selection with the Sera-Mag beads® to again remove any unwanted fragments outside the target range of 250–650 bp. Finally, we used the AATI Fragment Analyzer™ to determine peak average length of our DNA fragments in each sample, and Qubit 2.0™ to record the double-stranded DNA concentrations of each sample. We used additional control samples with molecular grade water instead of DNA throughout all procedures to confirm there had been no contamination.

We divided the 97 processed samples into two pools, or libraries, with samples of lower concentration in one pool, and samples of higher concentration in the other. We combined all DNA from each sample instead of using equimolar amounts. Singapore Centre on Environmental Life Sciences Engineering checked the quality of each DNA library, and then sequenced each library in two lanes of one flowcell of an Illumina HiSeq 2500 Rapid Sequencing Run to produce 2 x 150 bp

paired end reads. We spiked both lanes with 5% PhiX to increase the quantity of data obtained, and downloaded the paired end reads for bioinformatic analysis.

FASTQC Reports v0.11.5 (Feb 2016)

pool1_helen_index1_ATCACG_L001_R1_001_fastqc.html

pool1_helen_index1_ATCACG_L001_R2_001_fastqc.html

pool1_helen_index5_ACAGTG_L001_R1_001_fastqc.html

pool1_helen_index5_ACAGTG_L001_R2_001_fastqc.html

pool2_helen_index1_ATCACG_L002_R1_001_fastqc.html

pool2_helen_index1_ATCACG_L002_R2_001_fastqc.html

pool2_helen_index5_ACAGTG_L002_R1_001_fastqc.html

pool2_helen_index5_ACAGTG_L002_R2_001_fastqc.html

We used a Phred Score of 20 as our quality threshold. Therefore, according to the above FASTQC reports we needed to truncate our 150 bp reads to 135 bp for further analysis.

Scripts used for process_radtags in STACKS v1.35

cd /Volumes/GABE/Helen/

process_radtags -P -p /Users/RheindtLab/Desktop/HelenRaw/FastQfiles16Mar2016/Lib2_index1 -o /Volumes/GABE/Helen/output/Lib2_index1 --renz_1 ecoRI --renz_2 mspI -b ./Lib2_index1 -i gzfastq -t 135 -c -q –r

t = truncation of read length

We repeated this script for each of our four folders of FASTQ files of 150bp paired-end reads. There was one folder per index per pool:

Sample_pool1_helen_index1

Sample_pool1_helen_index5

Sample_pool2_helen_index1

Sample_pool2_helen_index5

Scripts used for bwa_memscript (Li et al. 2013)

#We indexed the Sunda pangolin nuclear genome:

cd ./Desktop/HelenRaw/m_javanica

#index ref genome

bwa index ./sga_scaff_min1k.fa

#Then we aligned our renamed 135 bp reads to the genome using Burrows-Wheeler Aligner #http://bio-bwa.sourceforge.net/ Specifically the BWA-MEM algorithm.

cd /Volumes/GABE/bwa-0.7.12

bwa mem -t 20 -M /Users/RheindtLab/Desktop/HelenRaw/m_javanica/sga_scaff_min1k.fa \

/Volumes/GABE/Helen/output/Renamed1/Hong_Kong_2.1.fq.gz \

/Volumes/GABE/Helen/output/Renamed1/Hong_Kong_2.2.fq.gz >
/Volumes/GABE/Helen/output/Aligned/Hong_Kong_2.sam

bwa mem -t 20 -M /Users/RheindtLab/Desktop/HelenRaw/m_javanica/sga_scaff_min1k.fa \

/Volumes/GABE/Helen/output/Renamed1/MZBR_0271.1.fq.gz \

/Volumes/GABE/Helen/output/Renamed1/MZBR_0271.2.fq.gz >
/Volumes/GABE/Helen/output/Aligned/MZBR_0271.sam

bwa mem -t 20 -M /Users/RheindtLab/Desktop/HelenRaw/m_javanica/sga_scaff_min1k.fa \

/Volumes/GABE/Helen/output/Renamed1/MZBR_0275.1.fq.gz \

/Volumes/GABE/Helen/output/Renamed1/MZBR_0275.2.fq.gz >
/Volumes/GABE/Helen/output/Aligned/MZBR_0275.sam

We repeated this script for each sample

Scripts for Samtools v1.4 (Li et al. 2009)

1) To convert SAM to BAM files

2) To sort BAM files

#convert SAM to BAM files

cd /Volumes/GABE/

samtools view -bShq 20 /Volumes/GABE/Helen/output/Aligned/Hong_Kong_2.sam >
/Volumes/GABE/Helen/output/Aligned/Hong_Kong_2.bam

samtools view -bShq 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0271.sam >
/Volumes/GABE/Helen/output/Aligned/MZBR_0271.bam

samtools view -bShq 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0275.sam >
/Volumes/GABE/Helen/output/Aligned/MZBR_0275.bam

samtools view -bShq 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0276.sam >
/Volumes/GABE/Helen/output/Aligned/MZBR_0276.bam

samtools view -bShq 20 /Volumes/GABE/Helen/output/Aligned/MZBR_1069.sam >
/Volumes/GABE/Helen/output/Aligned/MZBR_1069.bam

#Repeated for each sample

#sort BAM files

cd /Volumes/GABE/

samtools sort -m 4G -@ 20 /Volumes/GABE/Helen/output/Aligned/Hong_Kong_2.bam
/Volumes/GABE/Helen/output/Sorted/Hong_Kong_2

samtools sort -m 4G -@ 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0271.bam
/Volumes/GABE/Helen/output/Sorted/MZBR_0271

```
samtools sort -m 4G -@ 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0275.bam
/Volumes/GABE/Helen/output/Sorted/MZBR_0275


samtools sort -m 4G -@ 20 /Volumes/GABE/Helen/output/Aligned/MZBR_0276.bam
/Volumes/GABE/Helen/output/Sorted/MZBR_0276


samtools sort -m 4G -@ 20 /Volumes/GABE/Helen/output/Aligned/MZBR_1069.bam
/Volumes/GABE/Helen/output/Sorted/MZBR_1069

#Repeated for each sample
```

Scripts for ref_map.pl pipeline in STACKS v1.35

#Our ref_map.pl pipeline in STACKS, included pstacks, cstacks and sstacks (Catchen et al. 2013)

cd /Volumes/GABE/

-T: specify number of threads to execute

-b: batch ID representing this dataset in the database

-m: minimum depth coverage to report a stack in pstacks

-D: batch description

-S: disable recording SQL data in database

-s: path to a SAM/BAM file containing an individual sample from a population

-o: path to write pipeline output files

-O: specifying populations for each individual

```
ref_map.pl -T 20 -b 1 -m 5 -D "helenrefmap" -S \

-s /Volumes/GABE/Helen/output/Sorted/Hong_Kong_2.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_0271.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_0275.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_0276.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1069.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1078.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1082.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1085.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1162.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1163.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1166.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1167.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1179.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1183.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1184.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1187.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_0272.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1034.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1035.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1036.bam \
```

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1040.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1041.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1043.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1044.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1045.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1047.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1055.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1056.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1060.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1062.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1065.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1066.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1067.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1068.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1087.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1180.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1181.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_0274.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1032.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1033.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1079.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1080.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1081.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1083.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1157.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1159.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1160.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1161.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1164.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1182.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1186.bam -s /Volumes/GABE/Helen/output/Sorted/MZBR_1188.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1216.bam -s /Volumes/GABE/Helen/output/Sorted/rescue_1.bam \

-s /Volumes/GABE/Helen/output/Sorted/rescue_2.bam -s /Volumes/GABE/Helen/output/Sorted/rescue_3.bam \

-s /Volumes/GABE/Helen/output/Sorted/rescue_5a.bam -s /Volumes/GABE/Helen/output/Sorted/rescue_6b.bam \

-s /Volumes/GABE/Helen/output/Sorted/rescue_7b.bam -s /Volumes/GABE/Helen/output/Sorted/rescue_8b.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_0270.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_0273.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1030.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1031.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1037.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1038.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1042.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1048.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1049.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1051.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1052.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1053.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1054.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1061.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1063.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1070.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1071.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1072.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1073.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1074.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1075.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1076.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1088.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1177.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1178.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1185.bam \

-s /Volumes/GABE/Helen/output/Sorted/MZBR_1189.bam -s
/Volumes/GABE/Helen/output/Sorted/MZBR_1190.bam \

-o /Volumes/GABE/Helen/output/refmap_m5 -O /Volumes/GABE/Helen/pangolin_populations.txt

#our first pangolin populations text file differentiated the Chinese pangolin from the Sunda pangolins, plus Singaporean versus Indonesian/Bornean samples

Hong_Kong_2 1

MZBR_0271 3

MZBR_0275 3

MZBR_0276 3

MZBR_1069 3

MZBR_1078 3

MZBR_1082	3
MZBR_1085	3
MZBR_1158	3
MZBR_1162	3
MZBR_1163	3
MZBR_1166	3
MZBR_1167	3
MZBR_1179	3
MZBR_1183	3
MZBR_1184	3
MZBR_1187	3
MZBR_0272	3
MZBR_1034	3
MZBR_1035	3
MZBR_1036	3
MZBR_1040	3
MZBR_1041	3
MZBR_1043	3
MZBR_1044	3
MZBR_1045	3
MZBR_1046	3
MZBR_1047	3
MZBR_1050	3
MZBR_1055	3
MZBR_1056	3
MZBR_1060	3
MZBR_1062	3
MZBR_1064	3
MZBR_1065	3
MZBR_1066	3
MZBR_1067	3

MZBR_1068	3
MZBR_1087	3
MZBR_1165	3
MZBR_1180	3
MZBR_1181	3
MZBR_0274	3
MZBR_1032	3
MZBR_1033	3
MZBR_1079	3
MZBR_1080	3
MZBR_1081	3
MZBR_1083	3
MZBR_1084	3
MZBR_1086	3
MZBR_1157	3
MZBR_1159	3
MZBR_1160	3
MZBR_1161	3
MZBR_1164	3
MZBR_1182	3
MZBR_1186	3
MZBR_1188	3
MZBR_1216	3
rescue_1	2
rescue_2	2
rescue_3	2
rescue_5a	2
rescue_6b	2
rescue_7b	2
rescue_8b	2
MZBR_0270	3

MZBR_0273	3
MZBR_1030	3
MZBR_1031	3
MZBR_1037	3
MZBR_1038	3
MZBR_1042	3
MZBR_1048	3
MZBR_1049	3
MZBR_1051	3
MZBR_1052	3
MZBR_1053	3
MZBR_1054	3
MZBR_1061	3
MZBR_1063	3
MZBR_1070	3
MZBR_1071	3
MZBR_1072	3
MZBR_1073	3
MZBR_1074	3
MZBR_1075	3
MZBR_1076	3
MZBR_1077	3
MZBR_1088	3
MZBR_1177	3
MZBR_1178	3
MZBR_1185	3
MZBR_1189	3
MZBR_1190	3