Genotyping by sequencing protocol

July 27, 2015

1 DNA extraction

1.1 First session

You need a DNeasy Blood & Tissue Kit (Qiagen), and RNase A. Have some liquid nitrogen ready in a container, clean plastic pestles and an electric screw driver. Buffer ATL may form precipitate upon storage. If necessary, warm to 56°C until precipitates fully dissolve.

1. Put one mosquito in a 1.5 ml tube and immediately, on ice. Add a small amount of autoclaved sea sand. Keep the tube open and deep the bottom of it in the liquid nitrogen to freeze the sample, with care not to freeze your fingertips. Use the screw driver with a pestle fitted on it to grind the mosquito. Press hard the pestle against the tube. Alternate between grinding and cooling the tube down in the liquid nitrogen.

2. When the mosquito is powdered, let the tube recover a temperateu >0°C. Add 180 μl of buffer ATL and remove the pestle with care of leaving the liquid in the tube. Then, add 20 μl of proteinase K, vortex, and incubate at 56°C overnight on a rocking platform. Make sure to seal the tubes with parafilm to reduce evaporation.

1.2 Second session

1. Just like the ATL buffer, buffer AL may also form precipitates. Warm it up if necessary to dissolve them. Make sure that buffers AW1 and AW2 have the ethanol added.

2. After the incubation, bring samples to room temperature. Centrifuge for 2 min at 8000 $\times g$ (10000 rpm?), at 20°C. Transfer supernatant to a new tube. This removes the sand.

3. Add 4 μl RNase A (100 mg/ml), mix by vortexing, and incubate for 2 min at room temperature.

4. Vortex for 15 s. Add 200 μl buffer AL to the sample, and mix thoroughly by vortexing. Then add 200 μl ethanol (96-100%), and mix again thoroughly by vortexing. It is essential that the sample, buffer AL, and ethanol are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. Buffer AL and ethanol can be premixed and added together in one step to save time when processing multiple samples.

5. Pipet the mixture from the previous step (including any precipitate) into the DNeasy Mini spin column placed in a 2 ml collection tube. Centrifuge at $\geq 6000 \times g$ (8000 rpm) for 1 min. Discard flow-through and collection tube.

6. Place the DNeasy Mini spin column in a new 2 ml collection tube, add 500 μlk buffer AW1, and centrifuge for 1 min at $\geq 6000 \times g$ (8000 rpm). Discard flow-through and collection tube.

7. Place the DNeasy Mini spin column in a new 2 ml collection tube, add 500 μl buffer AW2, and centrifuge for 3 min at $20000 \geq \times g$ (14000 rpm) to dry the DNeasy membrane. Discard flow-through and collection tube. It is important to dry the membrane, since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol will be carried over during the following elution. Following the centrifugation step, remove the DNeasy Mini spin column carefully so that the column does not come into contact with the flow-through, since this will result in carryover ethanol. If carryover of ethanol occurs, empty the collection tube, then reuse it in another centrifugation for 1 min at $20000 \times g$ (14000 rpm).

8. Place the DNeasy Mini spin column in a clean 1.5 ml or 2 ml microcentrifuge tube (not provided in the DNeasy kit), and pipet 200 μl buffer AE directly onto the DNeasy membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at $\geq 6000 \times g$ (8000 rpm) to elute. Elution with 100 μl (instead of 200 μl) increases the final DNA concentration in the eluate, but also decreases the overall DNA yield. For maximum DNA yield, repeat elution once more. Do not elute more than 200 μl into a 1.5 ml microcentrifuge tube because the DNeasy Mini spin column will come into contact with the eluate.

Table 1: Typical restriction digest.

Restriction enzyme	10 units is sufficient, generaly 1μl is used.
DNA	$x\mu$l ($\sim$ 1 μg)
10$\times$ NEBuffer	$0.4x$ μl.
Total reaction volume	$4x$ μl.
H_2O	$2.6x - 1\mu$l
Incubation time	1 hour (I don't think it hurts if more).
Incubation temperature	37°C for most enzymes.

2 Digestion

I digest the DNA with only one enzymee, NlaIII (5'-CATG↓-3'; NEB). Table 1 shows the composition of a typical digestion reaction. The DNA volume varies among samples. When deciding the final volume, it is important to keep in mind the following advice:

1. If DNA is in a buffer with some salt (such as TE), then the DNA volume should **not** be more than 25% of the reaction volume. Otherwise, the salt in the buffer can inhibit the enzyme.

2. If DNA is very diluted and the DNA volume is higher than 100 μl, it may be a good idea to concentrate the DNA first. However, a highly diluted DNA is apparently the last of the concerns in restriction reactions (mental note: I wonder if it can be compensated to some extent with longer incubation time).

3. As a rule of thumb: use 1 μl of enzyme (10000 units/ml) per 1 μg of DNA.

4. Enzymes are in 50% glycerol (to avoid freezing). Always keep the total concentration of glycerol in the reaction below 5%.

5. DNA concentration in the reaction should be at 20-100 ng/μl.

Do not stop the reaction by heat-inactivation of the restriction enzyme. Instead, use a Spin Column or magnetic beads (see below) to clean the reaction product and concentrate the DNA. This way, you prevent bias against shorter and AT-rich fragments, which would de-naturalize during heating and then loose the chance to ligate the adapters.

2.1 Clean up with magnetic beads

1. Prepare twice as much 70% ethanol as total DNA volume you have among all the tubes, or a bit more. Prepare at least 400 μl of 70% ethanol per sample.
2. Vortex High Prep PCR beads thoroughly.
3. Add beads in a 1.5:1 ratio[1] to the restriction (or PCR) reaction and vortex well (30 s; or pipette mixing).
4. Incubate at room temperature for 5 min (or more; this is an important step) without shaking.
5. Place tube on magnetic rack for 2 min or until the solution is clear (the time depends on the volume of the solution).
6. Remove supernatant carefully without disturbing the beads. Make sure you have NO beads in the pipet.
7. Add the same amount of freshly prepared ethanol as supernatant you removed (at least 200 μl), so that the alcohol covers all the beads.
8. Let stand for a minute.
9. Remove supernatant.
10. Add the same amount of 70% ethanol again.
11. Let stand for another minute.
12. Remove the supernatant.
13. Air dry pellet (10-20 min) or incubate at 37°C for ∼10 min. You may see a crack in the pellet. If you over-dry the beads, you will see many cracks. If you under-dry the beads, the DNA recovery rate will be lower, due to the remaining ethanol.
14. Add 1× TE buffer.
15. Vortex for 30 s or pipette mix 10 times. The liquid level should be high enough to contact the magnetic beads. If not, extra vortexing is required, and may not be sufficient to fully elute all the DNA.

[1] 1.5 ratio is generous and allows binding of short fragments as well. If removing short fragments, use an 0.8:1 ratio instead.

16. Incubate 2 min at room temperature.

17. Place tube on magnetic rack for 2 min, or until the solution is clear.

18. Transfer supernatant to a fresh tube.

About the final volume to elute the DNA, I have been successful with very low ammounts (10-20 μl), if the pellet was not too dry. It's easier with more. After the digestion reaction, I recover the samples with the amount of TE that would give me the same concentration in all samples, in the order of 10 ng/μl, assuming that all the original amount of DNA is recovered.

2.2 Estimate distribution of fragment sizes with Agilent BioAnalyzer

This is a tricky machine. Before using it, it is convenient to clean the electrodes: just put 450 μl H_2O in a cleaning chip, place it in the machine (with lid closed), and let it sit for a little while. After removing the water chip, leave the lid of the machine open for a couple of minutes to dry de electrodes. The machine does not have to be turned on for this. You can rinse the chip with destilled water and re-use it later.

Before the assay, you should also set up the computer: turn it on, open the software, choose the machine, select the type of assay, and name the samples.

The assay chip has 11 spots available for samples. When loading it, it is recommended not to press the pippette to the second stop, to avoid the introduction of air bubbles. Loading should be accurate and not too slow. Once the chip is ready, it should be used immediately. Every assay costs on the order of €50.00.

Follow the instructions in the original protocol.

3 Ligation

3.1 Prepare annealed adapters stock

This is a combination of three protocols [1, 2, 3]. Oligos are delivered dry, and need to be suspended in any desired volume or concentration. The delivery documentation suggests 100 pmol/μl (that is, 100 μM).

1. Prepare the annealing buffer stock (10X): 100 mM Tris HCl (pH 8), 500 mM NaCl, 1 mM EDTA (see table 2).

Table 2: Annealing buffer (100 μl, 10X).

Reagent	Amount to add	Final concentration
Tris-HCl, pH 8, 1 M	10 μl	100 mM
EDTA, 0.5 M	2 μl	10 mM
NaCl (58.44 g/mol)	2.92 mg	500 mM
H_2O	88 μl	

Table 3: Volumes of reactants in the ligation reaction. All in μl. T4 ligase is at 2000000 U/ml.

Sample	DNA	10× Buffer	Adapter	1.5 M NaCl	T4	Total

2. Suspend the oligos in the volumes required to have them in 100 μM (100 pmol/μl). Use TE buffer.

3. To produce 10 μl of annealed stock from each adapter, at 15 μM (15 pmol/μl), combine: 1.5 μl of the top oligo (100 μM), 1.5 μl of the bottom oligo (100 μM), 1 μl 10× annealing buffer (table 2), and 6 μl of H_2O.

4. In a thermocycler, incubate at 97.5 °C for 2.5 minutes, and then cool at a rate not greater than 3 °C per minute, until the solution reaches a temperature of 21 °C. Hold at 4 °C.

5. The final working strength concentration of annealed adapters is calculated for each sample to have 7 times more adapters than fragment ends in the sample, in a volume of 1 or 2 μl. The required working concentration must be lower than the concentration of the annealed adapter stock (it can be lowered by increasing the volume added per reaction).

3.2 Adapter ligation

Following [2], I was planning to use NEB Buffer 2, supplemented with rATP, to run the ligation. I find that I don't have NEB buffer 2, but only T4 DNA

ligase buffer. The reason argued to suggest NEB buffer 2 is that it contains 50 mM NaCl, which is about the maximum that T4 DNA ligase can stand without being inhibited. The presence of some salt helps keep the DNA fragments annealed. I just supplement the reaction with 50 mM NaCl.

1. Prepare a 1.5 M solution of NaCl in a 10 ml Falcon tube, using 10 ml H_2O and 0.87 g NaCl (molecular weight 58.44 g/mol).

2. Add to the DNA samples the corresponding amounts of 10× T4 ligase reaction buffer, working stock of adapters, 1.5 M NaCl, and T4 DNA ligase (2000000 U/ml), according to table 3. Be sure to add the adapters to the reaction before the ligase, to prevent religation of the genomic DNA.

3. Incubate the reaction at 16 °C overnight (or at least 2 hours).

4. Heat inactivate T4 DNA ligase for 10 min at 65 °C. Allow reactions to cool slowly (2 °C per 90 s, or 0.02°C/s) to ambient temperature.

4 Clean up of ligation reactions

Before pooling the samples together, it is important to check that the ligation worked. Unfortunately, this involves another clean-up step, with magnetic beads (see above). If we did not clean the reaction, the excess of adapter dimers would dominate the fragment size distributions and the Bioanalyzer would not get precise measures in the size range of interest. This time, the ratio of beads to sample must be lower, 0.8:1, in order to remove as many adapter dimers as possible. Based on the results of the last Bioanalyzer run, the 'elution' volumes must be calculated with the aim of leveling up (and optimizing) the concentrations of fragments *in the desired size range* among samples.

5 Check size distributions with Agilent Bioanalyzer

This second run of the Bioanalyzer helps determine if the ligation worked. If it did, only a small shift of the distribution of fragment sizes would be observed, with respect to the previous assay. If the distribution of fragment sizes changes shapes significantly, it is indicative of re-ligation of genomic fragments. Another reason to run all samples through the Bioanalyzer is to

get accurate estimates of the molarity of the fragments in the desired size range, which are needed for pooling.

6 Choose a size range

The adapters ligated have divergent ends. They are composed of a bottom strand of 39-42 bases and a top strand of 42-45 bases. They are annealed along a stretch of 16-19 base pairs. At this point the same adapter is ligated at both ends. Thus each genomic DNA fragment has been extended by 81-87 'base pairs'. Sequencing will start at base 30 of the top strand of the adapters, in both directions, thus sequencing 12-15 bases of the adapter (codeword and restriction site) and into the genomic fragment for as long as 301 bases. The overlap between the two reads coming from opposite ends will be 662 - X, where X is the fragment size. In principle, an overlap longer than just needed to merge the reads serves only the purpose of improving the quality of the bases in the middle of the fragment. I consider that aiming at overlaps between 15 and 60 (fragment range 602 - 647) would be ideal. Let's round it to 600-650.

The molarity of the fragments observed in the previous step should not be a concern now. In order to increase the coverage obtained per genomic fragment, we need to aim at relatively short ranges with necessarily low concentrations. The final PCR should take care of this.

7 Pooling

To determine the amount of DNA from each sample to pool, it is desirable to collect equivalent amounts (at desired fragment size) across samples. However, it is even more important to balance the amounts of fragments with adapters of different sizes. There are 3 sets of adapters, with balanced base composition within each set. Thus, even if not three sets are equally represented, the 4 different sizes within each set must be equally represented.

Apparently, 0.002 pmol of fragments of the desired size range per sample should be enough, and maybe even too much.

8 Size selection

This is done with Pippin Prep, at the IZW.

Table 4: Amplification primers

	Primer 1	Primer 2
Length (bp)	61	55
Molecular weight (g/mol)	18973.2	16782.0
Basic T_M First cycle (°C)	65.7	64.4
52 mM Na^+ T_M 1^{st} cycle (°C)	75.3	73.6
Basic T_M Later cycles (°C)	74.7	74.3
52 mM Na^+ T_M Later (°C)	86.7	86.0

9 Another Bioanalyzer run

It would be convenient to make sure that size selection worked. At this point a Qubit quantification may give more accurate estimates of concentration, which may be lower than 0.5 ng/μl.

10 Amplification PCR

The protocol for Phusion High-Fidelity PCR Master Mix with HF Buffer recommends amplification primers (table 4) to be at 0.5 μM in PCR reactions of either 20 or 50 μl, which means that there should be either 10 or 25 pmols of each primer per reaction. About template DNA, the protocol only says that it should be less than 250 ng. However, the double digest protocol suggests 2.0 μM primers in 20 μl of final volume; that is, 40 pmols of primers. And only 20 ng of DNA, which would be about 0.06 and 0.1 pmols of fragments, depending on their average size (500 or 200). I think the latter is closer to what I need. I run 5 PCR reactions like the one shown on table 5.

I run the product of the first PCR reaction through the Bioanalyzer, and apparently the result is good. The profile shows higher concentration of DNA in the expected size range than the left over concentration of primers. I had to manually add the peak of the upper marker. I do not know if that affects the estimates of concentrations: 15107.76 pg/μl and 51954.8 pmol/l. Since the volume was 50 μl, that is 2.6 pmols of DNA yield per reaction, out of 0.03 (87-fold increase).

I run the other four PCR reactions in the same way, pooled the products, and cleaned them up with magnetic beads. I eluted in 50 μl. Either the last 4 PCR did not work as well, or the clean up was not very efficient. In all, and according to Qubit, I got a 20.4 ng/μl ($\times$ 50 μl = 1020 ng, which is

Table 5: PCR reactions in 50 μl. 10 μl of 10 μM Primer Mix may be prepared with 1 μl 100 μM Primer 1 + 1 μl 100 μM Primer 2 + 8 μl nuclease-free water.

Component	Volume (μl)	Final conc. (pmols/l)
10 μM Primer Mix	1.0	200000.00
Template DNA	15.2	580.24
2× Phusion Master Mix	25.0	–
Nuclease free water	8.8	–
Total	50.0	

25.7 times the original DNA mass. Assuming an average size of 454 bp (from Bioanalyzer; equivalent to a molecular mass of 275917.5), the 20.4 ng/μl are 73.94 nmol/l, and 3.697 pmols (again, around 25.5 times the original number of molecules).

11 Clean up with magnetic beads

12 Quantification with Qubit or Bioanalyzer

13 Sequencing

References

[1] P. Andolfatto, D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast, T. Sunayama-Morita, and D. L. Stern. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.*, 21:610–617, 2011.

[2] P. D. Etter, S. Bassham, P. A. Hohenlohe, E. A. Johnson, and W. A. Cresko. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol. Biol.*, 772:157–178, 2011.

[3] B. K. Peterson, J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5):e37135, May 2012.