PstI–MspI GBS Library Preparation in the Rieseberg Lab
Begun by Brook Moyers, 11 June 2014
Modified by Kate Ostevik, Marco Todesco

This protocol is largely based on Poland et al. 2012. The steps to GBS are: digestion (DIG) of DNA by restriction enzymes, ligation (LIG) of barcoded and common adapters to the ends of the digested fragments, amplification (PCR) of the adapter-ligated fragments, size selection (SS) of fragments to a specific range, and (possibly) depletion (DSN) of high-copy fragments from the library (which is followed by additional amplification). The order of these steps and the stage at which samples are pooled has varied among libraries in our lab. In particular, we have pooled samples after ligation, after amplification, or by row after ligation and then rows pooled after amplification. Pooling early saves time and effort, but pooling later allows individual variation in amplification, etc. efficiency to be better controlled. Size selection has been done after ligation, after amplification, and/or after depletion. Size selecting early theoretically makes downstream steps more efficient, but must be done carefully with thought to changes that will occur in fragments size (e.g. lengthening during amplification). Choose your own adventure!

Marco's approach:
DNA --> DIG --> LIG --> Clean & Concentrate --> PCR --> Quantification, Pool all --> Gel SS --> DSN --> Bead Clean --> QC

I. DNA

Use 10 mM Tris-HCl as your elution/storage buffer, which buffers but does not inhibit downstream applications. Note that if you use water as an elution buffer, or anywhere throughout this protocol, you must ensure that has a pH of ~7. Distilled water tends to be acidic (pH 5–6), which, along with compounds like EDTA, can inhibit downstream applications.We use 200 ng starting DNA material in the lab, in volume of 11.7 µl or 10 µl. Normalize your DNA samples to whichever of these concentrations is most sensible using whichever buffer your DNA is already stored in (e.g. Qiagen AE, water, 10 mM Tris-HCl, TE). To save yourself later effort, it can be useful to create enough volume of normalized DNA for several library preps in a single 96 well plate (e.g. 50 µl of 10 ng/µl of each sample).

II. Digestion

Aliquot an appropriate volume (e.g. 11.7 µl of 8.5 ng/µl) of normalized DNA into a 96 well plate using Qubit. Each well will receive a different barcode, so it is important to **keep track of where samples are located**, and to make sure to always check that the orientation of the new plate matches the old when transferring samples.

Create a double digestion master mix (your **total reaction volume will be 20 µ**l). This protocol calls for 8 units of each enzyme, which is likely more than sufficient. We have found that this particular combination of enzymes and buffer works well, and allows for the same buffer to be

used during ligation, but it is possible that other sources of enzymes or buffers may work as well.

DOUBLE DIGESTION MASTER MIX	**Catalogue Number**	**1 sample (µl) for 11.7 / 10 µl DNA**	**1 plate / 100 samples (µl)**
H_2O, pH 7 OR 10mM Tris-HCl, pH 8		5.5 / 7.2	550 / 720
NEB 10X Cutsmart buffer	B7204	2.0	200
NEB Pst I–HF	R3140	8 units: 0.4	40
NEB Msp I	R0106	8 units: 0.4	40
Total		**8.3 / 10.0**	**830 / 1000**

Add the appropriate volume of double digestion master mix to each sample, bringing the total volume in each to 20 µl. Digest for 5 hrs at 37 ˚C (5 hrs seems to work well), followed by 20 min at 65 ˚C to (partially) heat inactivate the enzymes, then a hold at 4 ˚C. You may store the digested product at 4 ˚C, but it is probably a good idea to proceed to ligation as soon as possible.

III. Ligation

The common adapter used here is the same Y-shaped adapter as the Poland et al. 2012 protocol:
Adapt2(PE) for MspI (C|CGG)
5’ nnnnnnnnC CGAGATCGGAAGAGCGGGGACTTTAAGC
3’ nnnnnnnnGGC TCTAGCCTTCTCGCCAAGTCGTCCTTACGGCTCTGGCTAG

Oligos to order:
>Adpt2PE_top_cg
cgAGATCGGAAGAGCGGGGACTTTAAGC

>Adpt2PE_bot
GATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

To anneal, (rehydrate the oligos to 100 uM and) add 10 uL of each oligo to 80 uL of Tris-HCl (pH 8.0). In the thermocycler, hold the reaction at 95C for two minutes, then cool to 30C at 1C per

minute, then hold at 4C. This gives you 100 uL of 10 uM double-stranded adapters, or about 94 ng/uL. This is your working concentration.

The barcoded adapters were already developed for the single-enzyme system in the lab. We have up to 192 barcoded adapter options, in two sets of 96. Our lab is responsible for the stock for barcodes 97–192, and the Irwin lab for barcodes 1–96.

BEFORE USING BARCODES: Please double-check that you are using either a single-use plate with 4.5 µl per well, or that you are confident about the labeling/concentration of whatever stock you are plating from.

Either aliquot 4.5 µl of barcoded adapters (either set 1–96 or 97–192 at 0.4 ng/µl) into the double digestion plate OR transfer the entire volume of digested samples (18.5–20 µl depending on evaporation) into a pre-aliquoted plate of barcoded adapters. The key here is that each sample well contains the same concentration of barcodes, so that ligation occurs evenly across samples. Make sure to keep the plate orientations the same before transfer.

Make a master ligation mix:

LIGATION MASTER MIX	**Catalogue Number**	**1 sample (µl)**	**1 plate / 100 samples (µl)**
H_2O, pH 7 **OR** 10mM Tris-HCl, pH 8		8.0	800
NEB 10X Cutsmart buffer	B7204	2.0	200
10 mM ATP	P0756	4.0	400
Common adaptor (10 uM)		1.0	100
NEB T4 DNA Ligase	M0202	~200 U: 0.5	50
Total		**15.5**	**1550**

Aliquot 15.5 µl to each well. This, along with the barcoded adapters will add 20 µl to your original digestion volume (18.5–20), for a total volume of 38.5–40 µl. Ligate for 2–3 hrs at 22 ˚C (3 hrs seems to work well), followed by 20 min at 65 ˚C to heat inactivate the ligase, then a hold at 4 ˚C.

At this point, the next step varies. Kate pools rows and performs a size selection (and a clean/concentrate) using beads, Brook pools all samples, then cleans and concentrates using beads, and Marco cleans individually and proceeds to PCR (his approach is outlined in sections IV & VI). Marco's approach is the most likely to get equal sequencing effort per sample, while Brook's is the most efficient use of time.

IV. Cleaning and Concentration

There are several methods (e.g. column-based, etc.) to clean and concentrate DNA or DNA libraries. As we are able to make our own SPRI beads, drastically lowering the cost, we tend to prefer bead clean/concentrating over other methods. Cleaning needs to be done after ligation and after PCR; concentration may need to happen at any point. Both of these can be done in concert with a bead size selection (see section V, Kate's protocol, for an example). When using the beads, make sure you allow them time to come to room temperature (~30 min.) and MIX VERY WELL EACH TIME before using.

In general, a ~1.5–1.8X volume bead clean with two 80% EtOH washes and one 100% EtOH wash will recover almost all of your DNA except the very small stuff (e.g. adapters, primers). We all generally use 1.6X when cleaning and/or concentrating.

<u>Brook's protocol for pooling, cleaning, and concentrating one plate of digested, ligated DNA samples, modified from Kristin's:</u>

1. Remove homemade AMPure XP (a.k.a. SPRI) beads from fridge and let stand for 30 minutes to come to room temperature.
2. Make 20 ml of fresh 80% EtOH.
3. Pool the contents of one digested/ligated plate into a 15 mL falcon tube (~38µl from each of 96 wells, or ~3650 µl total).
4. Mix beads VERY well and aliquot 5900 µL (~1.6X) into the 15 ml tube. Mix well and leave on bench for 15 minutes.
5. Mix gently, and aliquot contents into eight 1.5 mL tubes (~1.2 ml/tube).
6. Put tubes on magnet and allow pellet to form (~2–4 min).
7. Remove supernatant.
8. Add 1 mL of 80% ethanol to each tube.
9. Remove supernatant.
10. Add 1 mL of 80% ethanol to each tube.
11. Remove supernatant.
12. Add 1 mL 100% ethanol to each tube.
13. Remove as much supernatant as you can, e.g. with a smaller volume pipette.
14. Allow tubes to dry for ~7 minutes or until the pellet/tube no longer smells like EtOH. Cover with a kimwipe to prevent accidental contamination.
15. Close caps to prevent over-drying.
16. Add 200 µL 10 mM Tris-HCl to the first tube. Mix by pipetting and running the pipetted volume over the pellet. Vortex and allow tube to sit off the magnet for 2 minutes.
17. Put tube back on magnet and allow supernatant to clear (~2 min). Remove supernatant (you may not be able to remove the whole volume) and add it to the second tube.

18. Repeat steps 16 and 17 for all the tubes, reusing the original supernatant. This supernatant will slowly become more and more concentrated as it is moved through each tube, although you will gradually lose some volume.
18. Finally, carefully move supernatant from the eighth tube into a new, clean tube, doing your best to avoid transferring any beads. The final volume is usually ~125–140 µl.

Marco's approach for cleaning a plate of digested, ligated DNA samples WITHOUT pooling:

After ligation, purify and concentrate the individual reaction in a 96-well plate, using SPRI beads. To each well add 64 µl of beads using a multichannel pipette (aliquot the beads beforehand in a 8-strip of PCR tubes, and then dispense them from there), and then proceed using the 96-well magnet (since they are slightly different, here are steps of the protocol written for WGS). All steps use multichannel pipettes:

- Add 64 µl of SPRI beads, mix well, leave at room temperature for 5 minutes.
- Move the tubes to the magnet, leave them there for 5 minutes or until the solution is clear, then remove the supernatant.
- Keeping the tubes on the magnet, add 200 µl of 75% EtOH to each tube. Wait 15-30 seconds, then remove the ethanol.
- Add another 200 µl of 75% EtOH to each tube. Wait 15-30 seconds, then remove the ethanol.
- Dry the tubes for 5-10 minutes at 37°C in a thermal cycler, with the lid open.
- Resuspend well in 12 µl of 10 mM Tris-HCl pH 8.0. You might need to pipette up and down the solution on the sides of the wells to get all the beads. Leave at room temperature for 10 minutes.
- Put back on the magnet, wait for a couple of minutes or until the solution is clear. Move 10 µl of the supernatant to a clean plate.

Marco then proceeds to PCR (see section VI for his protocol).

V. Size Selection

This step can be done using a gel extraction kit or our homemade SPRI beads, and at several stages during the protocol (see intro). Gel extractions have lower yield (typically 10%), but are more precise, while bead extractions have very good yield by result in "shoulders" of fragments outside of the selected range. The size selection protocol with beads can vary depending on your aim (e.g. fragment range), but generally follows the protocol in GBS 2.0 (0.8 or 0.85X, wash, elute and transfer, then 0.5X and transfer supernatant, then 1X, wash, elute, and keep eluant—see original GBS protocol for more details). Below are two protocols used by the lab that may be helpful:

Kate's post-ligation clean and size select:

- Start with a plate of digested/ligated DNA and homemade SPRI beads at room temperature.

- Pool 36 µl from each well across each row into a 1.5 mL tube, for a total of 432 µl per row. You should have eight tubes for Rows A–H.
- Mix beads well. Add 345.6 µl (0.8X) of beads to each pooled tube.
- Let stand 5 mins.
- Place on magnet and allow pellet to form (2 mins).
- Discard supernatant (this contains small fragments).
- Wash 3X with 1 mL of 80% EtOH, then 1 ml 80% EtOH, and then 1 ml 100% EtOH. Remove supernatant after each wash. After 100% EtOH wash, carefully remove as much supernatant as possible.
- Let dry for 15 mins with caps open.
- Remove tubes from magnet. Add 50 µl of H_2O or 10 mM Tris-HCl to each tube and pipette 10X to mix. Let stand 5 minutes off magnet.
- Place on magnet and allow pellet to form (~2 min).
- Transfer supernatant (~50 µl, contains remaining larger fragments) of each tube to a new, labeled PCR tube.
- Mix beads well. Add 32.5 µl (0.65X) beads to each new tube and pipette to mix.
- Let stand 5 mins.
- Place on plate magnet and allow pellet to form (~2 mins).
- Transfer 72 µl of supernatant (contains smaller fragments that did not bind to the beads at 0.5X) from each tube to a fresh tube (off the magnet).
- Mix beads well. Add 72 µl (1X) beads to each tube. Mix and let stand 5 mins.
- Place on magnet and allow pellet to form (~2 min).
- Remove and discard supernatant.
- Wash 3X with 1 mL of 80% EtOH, then 1 ml 80% EtOH, and then 1 ml 100% EtOH. Remove supernatant after each wash. After 100% EtOH wash, carefully remove as much supernatant as possible.
- Let dry for 15 mins with caps open.
- Remove tubes from magnet. Add 32 µl of H_2O or 10 mM Tris-HCl to each tube and pipette 10X to mix. Let stand 5 minutes off magnet.
- Place on magnet and allow pellet to form (~2 min).
- Carefully transfer supernatant to fresh tubes.
- Quantify with Qubit and then you're ready for PCR!

<u>Qiagen Gel Extraction Kit Protocol</u>

- Pour a 1.5% gel using 2.25 g high MW agarose and 150 ml SB, along with 1–1.5 µl ethidium bromide, into one of the small (orange) rigs. Use the comb with 12 teeth, 1 mm (e.g. thinner) width at the top of the gel.
- Note: Do not add more than 10 µg DNA per lane or the library will not separate well. Aim for ~20 µg DNA across 3 wells, or ~14 µg across two.
- Load the gel. Leave at least 1 lane between your library and other occupied wells (e.g. ladder, other libraries). It is best to run only one library per gel, or two if you will multiplex them later. Add 25–30 µL of something to each well, including the empty ones. For empty wells, load e.g. 20–25 µl water mixed with 5 µl dye. For the ladder row(s), load 2–3 µl 1kb ladder mixed

with 17–23 µl water and 5 µl dye. For your library wells, add a volume of 20–25 of your sample mixed with 5 µl dye.f

- Add 5 µl EtBr to the bottom reservoir of the gel rig. Because EtBr is attracted to the negative pole (opposite to DNA), this helps ensure that the lower portion of the gel is well-stained.
- Run for ~10 minutes at a lower voltage, e.g. 100 V.
- Increase to a higher voltage (125–150 V) and run until the desired size range seems well enough separated (~30-60 mins).
- To excise bands:

○ Label and pre-weigh a number of 2ml tubes (3–8 per library, depending on the width of your selected band and the number of wells).○ Clean blue light viewer before placing gel there. You will need to wear the orange glasses OR place the (cleaned) orange plate over the gel to visualize the stained DNA, and will likely need to turn out the light.

○ Use new, sterile blades—a new one for each library. Cut carefully and as precisely as possible—you want to have the lowest possible gel volume that contains all of your desired fragments. Do your best to avoid contaminating the tubes with EtBr by using the blade to carefully transfer your excised gel fragments into the pre-weighed tubes (no more than ~400 g per tube).
○ Weigh the tubes containing the gel fragments to get the mass of the excised fragment. Rearrange slices as needed.
○ Proceed to follow the Qiagen kit protocol. Note that the 3X (step 2) and 1X (step 4) volumes both refer to the original mass of gel in each tube. Pre-warm Buffer QG to 50 ˚C. Wait 2–5 minutes after applying the wash buffer (step 7), as suggested by the note. For maximum yield, pre-warm your elution buffer (10 mM Tris-HCl recommended) to 50 ˚C and allow to sit on column for a few minutes before centrifugation.

- You will likely need to concentrate your extracted library before using. Use a 1.6X bead concentration, wash, and elute into your desired volume.

VI. Amplification (PCR)

There are several approaches that have been successfully used by the lab. One thing to keep in mind is that increasing the number of PCR cycles increases the likelihood of PCR bias significantly affecting your sequencing data. Especially if you plan to do a depletion treatment (section VII), which involved further PCR, try to keep your number of PCR cycles low (< 15).

All PCR approaches use the same primers (from Poland et al. 2012):

>IlluminaFor_PE (Tm=70) 58bp (Tm=57)
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

>IlluminaRev_PE (Tm=70) 46bp (Tm=62)
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAA

Brook's approach with NEB Phusion Taq Master Mix, following Poland et al. 2012

This has worked well with between 9 and 120 ng of template DNA, with similar yields. Running several (eight, here) PCR reactions per library both increases the final yield and should decrease the amount of random variation in amplification efficiency among samples.

AMPLIFICATION MIX	**Catalogue Number**	**1 rxn (µl)**	**1 library (8 rxns, µl)**
H_2O, pH 7 **OR** 10mM Tris-HCl, pH 8		7.5*	X8.5 = 63.75
10 µM Illumina Forward Primer		1	8.5
10 µM Illumina Forward Primer		1	8.5
2X NEB MM Phusion	M0531	12.5	106.25
Total		**22**	**187**

*You can easily vary the mix according to the volume of template you wish to add by changing the volume of water/Tris added to the master mix

Add 3 µl of DNA template to each of eight PCR tubes (strip tubes work best). Add 22 µl of amplification to mix to each tube for a total volume of 25 µl. Mix carefully and spin down quickly before placing on thermocycler.

PCR protocol:
95 C (30 sec)
{ 95 C (30 sec), 62 C (20 sec), 68 C (30 sec)* } 12–16 cycles (Brook uses 14)
72 C (5 min)
4 C (forever)
*The extension time can be altered to change the distribution of fragments

After PCR, pool and clean with a 1.6 bead clean (23 µl X 8 rxns = 184 µl plus 294.4 µl beads). Wash and elute with 30–50 µl 10 mM Tris-HCl.

Kate's PCR protocol using Kapa HIFI Master Mix

AMPLIFICATION MIX	**1 rxn**	**2 libraries (16 rxns)**

Kapa HIFI HotStart 2X	12.5	X17 = 212.5
Primer F	1.0	17.0
Primer R	1.0	17.0
Total	14.5	246.5
DNA template	10.5	in each well
H_2O / 10MmTris		
Total	**25.0**	

Amplification cycles	^0C	time
	98	00.00.30
	98	00.00.30
	62	00.00.20
	72	00.00.30
repeat steps 2-4 12–16X		
	72	00.05.00
	4	forever

<u>Marco's PCR protocol with an unpooled plate of DIG/LIG/Cleaned samples:</u>

After cleaning (see section IV) I do individual PCRs in a 10 µl volume (so the total volume of PCR reaction is about 5 times more than in Brook's and Kate's approach). For one plate, the master mix is:

AMPLIFICATION MIX	**1 rxn (µl)**	**1 library (100 rxns)**
Kapa HIFI HotStart 2X	5	X100 = 500

Primer F	0.4	40
Primer R	0.4	40
H_2O	0.2	20
Total	6	600
DNA template (purified ligation)	4	in each well
Total	10.0	

I find the re-usable silicon lids work well for PCR plates. The program I use is the one for Kate's above for the Kapa enzyme, and I found 13-14 cycles to give a reasonable yield. I use the Kapa enzyme because in a few tests Kate and I made gave quite a higher yield than Phusion and the NEBNext Master Mix (like the NEBNext Master Mix, it supposedly also amplifies well over a much broader range of GC contents than Phusion).

I then quantified the individual PCRs and pooled them accordingly (you'll have to normalize the quantities to the least abundant sample), and proceed to gel size selection as in Brook's protocol (step V). With 13-14 cycles I get between 25 and 45 µg of library before size selection, which is plenty. If you are not planning to do a depletion step (which requires high library concentrations to work effectively) 5-10 µg would be actually more than enough (after gel size-selection you'll end up with about 10% of the amount of library you loaded in a gel), so you could remove one or two cycles. If you are not planning to do a depletion step at this point, the library is done.

While for some plates all samples are relatively similar (up to 2-fold differences), other plates had up to 4-5-fold difference between the most and least concentrated samples. I guess things could get worst if you have DNA from different sources or of very different quality. This variant of the protocol is more time consuming and more expensive (I would think an additional 120 dollars per library), but should give more even representation of different samples. If coverage is a concern (I am planning to run 192 samples in a single Illumina lane), I think this makes sense. Otherwise, Brook's and Kate's versions are less work-intensive. Of course, all of this is based on wild speculations, since none of us had results from the latest iterations of the GBS protocol :)

VII. Depletion of high copy fragments (DSN)

Coming soon. The basic gist is that we do it like the WGS method, but only incubate for 5 hours rather than 44 or 22 (because the library is less complex). From limited sequence data so far, this seems to reduce the high copy fragment percentage of the total number of reads by about half (from 10% → ~5%).

VIII. Quality Control (QC)

The first step is to quantify your library using the Qubit. You need a minimum of ~2 ng/µl and at least 17 µl total for Biodiversity Centre sequencing with Ana, or >75 ng in 25–75 µl for Genome Quebec. If you have that, proceed to the Bioanalyzer. Check the distribution of your fragments—are they in the right size distribution for your project? If you depleted, does the distribution look more smooth than jagged? If yes, you may choose to do qPCR, or submit the library directly for sequencing.