Příloha

Appendix 1. **Výstup z aplikace RADtag counter from GenePool**, pomocí které byla vybrána nejvhodnější restriktáza (SbfI) pro štěpení ježčího genomu.

RADtag counter from GenePool, Edinburgh

To use this counter:

1	Enter the GC content of your target genome here:	**0,42**	proportion GC
2	Enter the size in megabases of your genome here:	**2708**	megabases genome
3	Enter the fold coverage of RADtags you require here:	**30**	fold coverage
4	Enter the per-pool plexity you plan to use here:	**25**	plexity
5	Enter number of million reads per lane		
	(please contact the GenePool for throughput currently	**100**	million reads per lane
	achieved on the GAIIx and HiSeq platforms)		

Overhang	**TGCA**			**GGCC**			**AATT**			**GATC**				**TCA**
Enzyme	**SbfI**	**PstI**	**NsiI**	**NotI**	**EaeI**	**EagI**	**EcoRI**	**ApoI**	**MfeI**	**BamHI**	**BclI**	**BglII**	**BstYI**	**BbvCI**
Site	CCTGCA*GG	CTGCA*G	ATGCA*T	GC*GGCCGC	Y*GGCCR	C*GGCCG	G*AATTC	R*AATTY	C*AATTG	G*GATCC	T*GATCA	A*GATCT	R*GATCY	CC*TCAGC
Site frequency	7,21E-06	0,000164	0,000312	3,78E-06	0,000486	8,58E-05	0,000312	0,0017682	0,000312	0,000164	0,000312	0,000312	0,0009272	3,43E-05
Sites/Mb	7	164	312	4	486	86	312	1768	312	164	312	312	927	34
Number of sites in genome	**19533**	**442916**	**844655**	**10242**	**1316636**	**232255**	**844655**	**4788292**	**844655**	**442916**	**844655**	**844655**	**2510864**	**93012**
Number of tags	39065	885833	1689310	20485	2633273	464509	1689310	9576585	1689310	885833	1689310	1689310	5021729	93012
Num sequences for coverage	1171957	26574988	50679286	614546	78998182	13935279	50679286	287297542	50679286	26574988	50679286	50679286	150651862	2790374
Million sequences per pool	29,3	664,4	1267,0	15,4	1975,0	348,4	1267,0	7182,4	1267,0	664,4	1267,0	1267,0	3766,3	69,8
does your pool fit in one lane?	**YES**	**NO**	**NO**	**YES**	**NO**	**NO**	**NO**	**NO**	**NO**	**NO**	**NO**	**NO**	**NO**	**YES**

Appendix 2. **Sequenced RAD Markers for Rapid SNP Discovery and Genetic Mapping**

Paul D. Etter (modified by L. Choleva for *SbfI* rextriction enzyme)

2. Materials

2.1. DNA extraction and RNase A treatment

1. DNeasy Blood & Tissue Kit (Qiagen).
2. RNaseA (Qiagen).

2.2. Restriction endonuclease digestion

1. Restriction enzyme (NEB): ***Sbf*I-HF**.
2. Clean, intact high-quality genomic DNA: 25 ng/µl.

2.3. P1 Adapter ligation

1. NEB Buffer 2.
2. rATP (Promega): 100 m*M*.
3. P1 Adapter: 100 n*M.* A modified Solexa© adapter (2006 Illumina, Inc., all rights reserved).

Prepare 100nM stocks of P1 adapters in 1X Annealing Buffer (AB, *see* **Note 4**).

Below, example barcoded *Sbf*I P1 adapter sequences. "P" denotes a phosphate group and "x" refers to barcode nucleotides.

P1 top:

5´-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTxxxxTGC*A-3´

P1 bottom:

5´-Phos-xxxxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T-3´

The oligos are given in the excel file "Oligo Table RAD_P1_adapters_Lukas Choleva.xls"

4. Concentrated T4 DNA Ligase (NEB): 2,000,000 U/ml.

2.4. Purification steps

1. QIAquick or MinElute PCR Purification Kit (Qiagen).

2.5. DNA shearing

1. Bioruptor, nebulizer or Branson sonicator 450.

2.6. Size selection/agarose gel extraction

1. Agarose (Sigma)
2. 5X TBE: 0.45 *M* Tris-Borate, 0.01 *M* EDTA, pH 8.3.
3. 6X Orange Loading Dye Solution (Fermentas).
4. GeneRuler 100 bp DNA Ladder Plus (Fermentas).
5. Razor blades.
6. MinElute Gel Purification Kit (Qiagen).

2.7. Perform end repair

1. Quick Blunting Kit (NEB).

2.8. 3´-dA overhang addition

1. NEB Buffer 2.
2. dATP (Fermentas): 10 m*M.*
3. Klenow Fragment (3´ to 5´ exo$^-$, NEB): 5,000 U/ml.

2.9. P2 Adapter ligation

1. NEB Buffer 2.
2. rATP: 100 m*M.*
3. P2 Adapter: 10 µ*M.* A modified Solexa© adapter (2006 Illumina, Inc., all rights reserved).

Prepare 10 µ*M* double-stranded adapter in 1X AB (*see* **Note 4**). Asterisk denotes a phosphorothioate bond introduced to confer nuclease resistance to the double-stranded oligo ***(14)***.

Paired End

P2 top:

5´- /5Phos/GATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA-3´

P2 bottom:

5´CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T -3´

The oligos are also given in the excel file "Oligo Table RAD_P2_adapters_Lukas Choleva.xls"

4. Concentrated T4 DNA Ligase.

2.10. RAD tag Amplification/Enrichment

1. Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB).

2. Modified Solexa© Amplification primer mix (2006 Illumina, Inc., all rights reserved): 10 µ*M.*

   P1-forward primer: 5´- AATGATACGGCGACCACCG*A -3´

   P2-reverse primer: 5´- CAAGCAGAAGACGGCATACG*A -3´

The oligos are also given in the excel file "Oligo Table RAD_P2_adapters_Lukas Choleva.xls"

3. Methods (GENERAL)

The protocol described below, outlined in **Fig. 1,** prepares RAD tag libraries for high-throughput Illumina sequencing (*see* **Note 1**). In short, genomic DNA is digested with a restriction enzyme and an adapter (P1) is ligated to the fragment's compatible ends (**Fig. 1A**). This adapter contains forward amplification and Illumina sequencing primer sites, as well as a nucleotide barcode 4 or 5 bp long for sample identification. To reduce erroneous sample assignment due to sequencing error, all barcodes differ by at least two nucleotides. The adapter-ligated fragments are subsequently pooled, randomly sheared, and size-selected (**Fig. 1B**). DNA is then ligated to a second adapter (P2), a Y adapter ***(13)*** that has divergent ends (**Fig. 1C**). The reverse amplification primer is unable to bind to P2 unless the complementary sequence is filled in during the first round of forward elongation originating from the P1 amplification primer. The structure of this adapter ensures that only P1 adapter-ligated RAD tags will be amplified during the final PCR amplification step (**Fig. 1D**). The protocol for mapping of the lateral plate locus in stickleback using *EcoR*I RAD markers used in Baird et al., 2008 ***(12)*** is described here in detail as an example of the multiplexing approach. For bulk-segregant analysis pooled samples are combined prior to digestion and treated as a single library with one barcode.

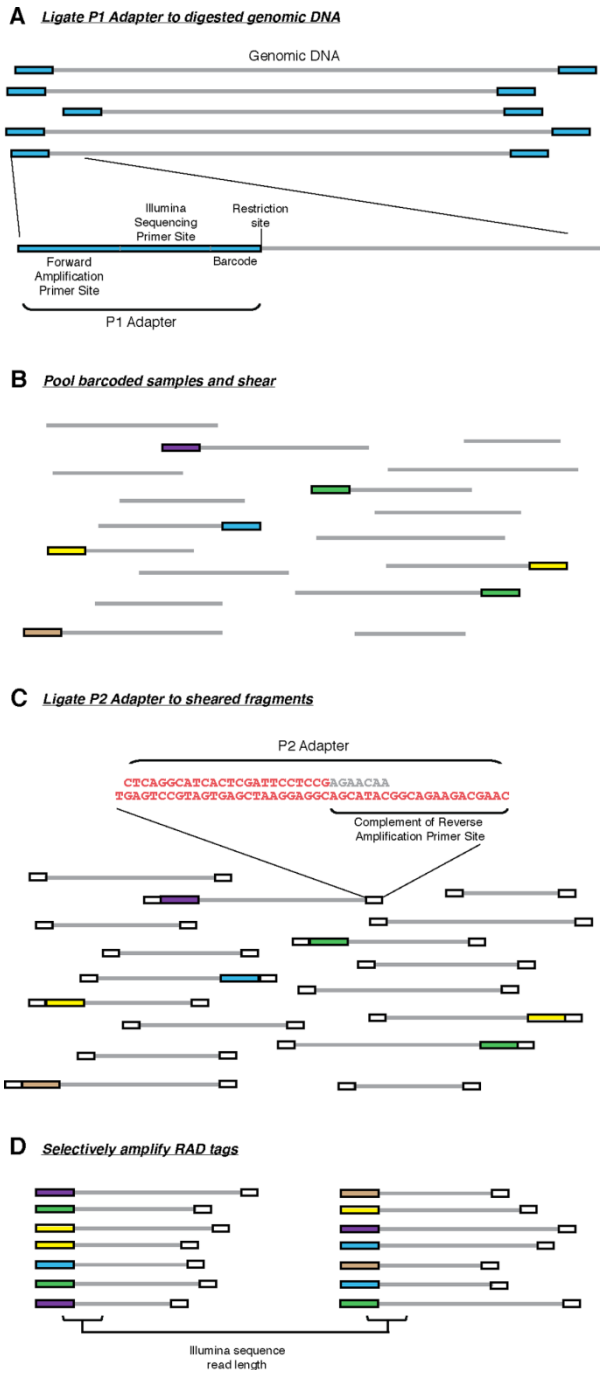

Fig. 1. RAD tag library generation. (**A**) Genomic DNA is digested with a restriction enzyme and a barcoded P1 adapter is ligated to the fragments. The P1 adapter contains a forward amplification primer site, an Illumina sequencing primer site, and a barcode (colored boxes represent P1 adapters with different barcodes). (**B**) Adapter-ligated fragments are combined (if multiplexing), sheared and (**C**) ligated to a second adapter (P2, white boxes). The P2 adapter is a divergent "Y" adapter, containing the reverse complement of the P2 reverse amplification primer site, preventing amplification of genomic fragments lacking a P1 adapter. (**D**) RAD tags, which have a P1 adapter, are selectively and robustly enriched by PCR amplification. (Reproduced from **ref.** ***12*** with permission from PLoS).

3.1. DNA extraction and RNase A treatment

1. **IGNORE** We recommend extracting genomic DNA samples using the DNeasy Blood & Tissue Kit (Qiagen) or a similar product that produces very pure, high molecular weight, RNA-free DNA. High-quality DNA is required for optimal restriction endonuclease digestion and is of utmost importance for the overall success of the protocol. Follow the manufacturers instructions for extraction from your tissue type. Be sure to treat samples with RNase A following manufacturer's instructions to remove residual RNA. Quantify the DNA using a fluorimeter to get the most accurate concentration readings (*see* **Note 3**). The optimal concentration after elution is 25 ng/µl or greater.

3.1.5 Anneal adapters

Single-stranded oligos need to be annealed with their appropriate partner before ligation. We provide sequences for 48 uniquely barcoded adapter P1 oligo pairs (oligos P1-FOR and P1-REV) and common adapter P2 (oligos P2-FOR and P2-REV), see **OLIGO TABLE RAD_P1_adapters.**

1. To create Adapter P1, combine each oligo P1-FOR with its complementary oligo P1-REV and in a 1:1 ratio in working strength annealing buffer (final buffer concentration 1x) for a total annealed adapter concentration of 40uM (for example, if purchased oligos are resuspended to an initial concentration of 200uM, use 20ul oligo P1-FOR, 20ul oligo P1-REV, 10ul 10x annealing buffer and 50ul nuclease-free water). Do the same for oligos P2-FOR and P2-REV to create the common adapter P2.

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

3. Prepare final working strength concentrations of annealed adapters from this annealed stock (the appropriate working stock dilution for your experiment can be determined from our ligation molarity calculator). For convenience, it is possible to store the adapters at 4°C while in active use.

3.2. Restriction endonuclease digestion

1. Digest 0.1-1µg genomic DNA for each individual sample (**for 60 min at 37°C in a 50 µl reaction volume containing 5.0 µl 10× Buffer 4 and 10 units (U) *Sbf*I-HF (New England Biolabs [NEB])**, following the manufacturers instructions.

(alternatively, to ensure complete digestion, run digests for 3 hours at 37°C, holding at 4°C. Do not heat kill the enzymes, as this may skew base composition in the resulting fragment library. Before

proceeding with step “Clean the double digest with AMPure XP beads”, cool the reaction to room temperature. Alternatively, reactions can be stored at 4°C overnight).

2. Heat-inactivate the restriction enzyme following manufacturer’s instructions. **(Heat-inactivate for 20 min at 65°C).** Allow reaction to cool slowly to ambient temperature (30-60 min). If the enzyme cannot be heat-inactivated, purify with a QIAquick column following manufacturer’s instructions prior to ligation.

3.3. P1 Adapter ligation

1. This step in the protocol ligates barcoded, restriction-site overhang specific P1 adapters onto complementary compatible ends on the genomic DNA created in the previous step (*see* **Note 5**). 48 different barcoded P1 adapters are used to make ***Sbf*I-HF** RAD tag libraries for 48 individuals (*ignore* Note 7).

2. To each inactivated digest, add:

3.0 µl (2.5 µl – used by Baird et al. 2008) Barcoded SbfI-P1 Adapter (100 nM), a modified Illumina© adapter (2006 Illumina, Inc., all rights reserved; top oligo: 5’-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTxxxxxTGC*A-3’ [xxxxx = barcodes, * = phosphorothioate bond]; bottom oligo: 5’-Phos-xxxxxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCAT*T-3’), added to each sample along with 0.6 µl rATP (100 m*M*, Promega), 1.0 µl 10× SbfI Buffer 4, 0.5 µl (1000 U; 2,000,000 U/ml) T4 DNA Ligase (high concentration, NEB), 4.9 µl H2O; 60.0 µl total volume; and incubate reaction at room temperature (RT) for 30 min.

(note: some authors add 750 pmol P1 adaptor per 1 ug)

3. NEB Buffer 4 is used in the ligation reactions in this protocol instead of ligase buffer because the salt it contains (50 m*M* NaCl) ensures the double-stranded adapters remain annealed during the reactions (*see* **Note 4**). T4 DNA Ligase is active in all 4 NEB Buffers if supplemented with 1m*M* rATP. If the restriction buffer used for digestion does not contain at least 50 m*M* potassium or sodium ions, or if the endonuclease cannot be heat-inactivated and the reaction must be purified in a column prior to P1 ligation, add 6.0 µl NEB Buffer 4.

4. **IGNORE:** Reduce the amount of P1 used in the ligation reaction if starting with less than 1µg genomic DNA or if cutting with an enzyme that cuts less frequently than *EcoR*I (**yes, this is the case for SbfI, as this enzyme cuts more frequently than *EcoR*I . Therefore, 4.0 µl Barcoded SbfI-P1 Adapter is used instead of 5.0 µl, as typical when *EcoR*I enzyme is used**). It is critical to optimize the

amount of P1 adapter added when a given restriction enzyme is used for the first time in an organism (*see* **Note 6**).

5. Heat-inactivate T4 DNA Ligase for 20 min at 65° C. Allow reaction to cool slowly to ambient temperature (30 min).

3.4. Sample multiplexing

1. This step allows multiple individually barcoded samples to be combined and processed as one to cut down on cost, work time, and differences in amplification efficiency that may arise between different library preparations when processing many at once.

2. Combine barcoded samples at desired ratio. Use a 100-300 µl aliquot containing 1-2 µg DNA total to complete the protocol and freeze the rest at -20° C.

(**IGNORE:** In Baird et al., 2008 ***(12)*** F_0 parent samples, as well as the F_2 pools used for bulk-segregant analysis, were combined at equal volumes to create one library (*see* **Fig. 2**, lanes 2, 3 & 5). *EcoR*I libraries containing barcoded samples from F_2 individuals sharing a given lateral plate phenotype were pooled and processed as separate libraries after P1 ligation (*see* **Note 7**)).

3.5. DNA shearing

1. Shear DNA samples to an average size of 500 bp to create a library of P1/restriction-site-ligated molecules with random variable ends for amplification. This step requires some optimization for different DNA concentrations and each time a different restriction endonuclease is used. The following protocol has been optimized to shear Stickleback DNA digested with (either *EcoR*I) or ***Sbf*I** using the Bioruptor and is a good starting point for any study. The goal is to create sheared product that is predominantly smaller than 1 kb in size (*see* **Fig. 2**).

2. Dilute ligation reaction to 100 µl in water (or take 100-300 µl aliquot from multiplexed samples) and shear in Bioruptor 10 times for 30 sec on high following manufacturer's instructions.

(note: some authors shear using five 30 s on-and-off cycles)

3. Clean up sheared DNA sample(s) using a MinElute column following manufacturer's instructions. This purification is performed in order to remove the ligase and restriction enzymes, and to concentrate the DNA so that the entire sample can be loaded in a single lane on an agarose gel. Elute in 20 µl EB.

3.6. Size selection/agarose gel extraction

1. This step in the protocol removes free un-ligated or concatomerized P1 adapters and restricts the size range of tags to that which can be sequenced efficiently on an Illumina Genome Analyzer flow cell. Run the entire sheared sample in 1X Orange Loading Dye on a 1.25% agarose, 0.5X TBE gel for 45 min at 100 V, next to 2.0 µl GeneRuler 100 bp DNA Ladder Plus for size reference (*see* **Fig. 2**, **Note 8**).

2. Being careful to exclude any free P1 adapters and P1 dimers running at ~130 bp and below, use a fresh razor blade to cut a slice of the gel spanning 300-500 (-700) bp. Extract DNA using MinElute Gel Purification Kit following manufacturer's instructions with the following modification: to improve representation of A + T-rich sequences, melt agarose gel slices in the supplied buffer at room temperature (18-22° C) with agitation for 30 min ***(14)***. **Elute in 19 µl EB into eppendorf tube containing 2.5 µl 10X Blunting Buffer from Quick Blunting Kit used in the following step.**

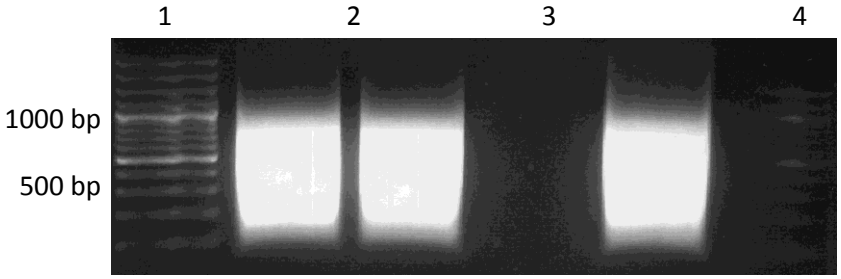

Fig. 2. Three barcoded and multiplexed RAD tag libraries. **2**, **3** & **5** each contain two DNA samples that were restriction digested, ligated to barcoded P1 adapters, combined, sheared, purified and then loaded on an agarose gel. **2** - F_0 parental DNA samples cut with *Sbf*I. **3** - F_2 pools cut with *Sbf*I. **4** - blank. **5** - F_0 parental DNA samples cut with *EcoR*I. Libraries contain 2 µg total combined genomic DNA each. **1** & **6** - 2.0 µl GeneRuler 100 bp DNA Ladder Plus.

3.7. Perform end repair (The Quick Blunting Kit (NEB) was used to polish the ends of the DNA.)

1. The Quick Blunting Kit protocol converts 5´ or 3´ overhangs, created by shearing, into phosphorylated blunt ends using T4 DNA Polymerase and T4 Polynucleotide Kinase.

2. To the eluate from the previous step, add: 2.5 µl dNTP mix (1m*M*), 1.0 µl Blunt Enzyme Mix. Incubate at RT for 30 min.

3. Purify with QIAquick column. Elute in 43 µl EB into eppendorf tube containing 5.0 µl 10X NEB Buffer 2.

3.8. 3´-dA overhang addition

1. This step in the protocol adds an 'A' base to the 3´ ends of the blunt phosphorylated DNA fragments, using the polymerase activity of Klenow Fragment (3´ to 5´ exo$^-$). This prepares the DNA fragments for ligation to the P2 adapter, which possesses a single 'T' base overhang at the 3´ end of its bottom strand.

2. To the eluate from the previous step, add: 1.0 µl dATP (10m*M*), 3.0 µl Klenow (exo$^-$). Incubate at 37º C for 30 min. Allow reaction to cool slowly to ambient temperature (15 min).

3. Purify with QIAquick column. Elute in 45 µl EB into eppendorf tube containing 5.0 µl 10X NEB Buffer 2.

3.9. P2 Adapter ligation

1. This step in the protocol ligates the Paired_End-P2 adapter, a "Y" adapter with divergent ends that contains a 3´ dT overhang, onto the ends of blunt DNA fragments with 3´ dA overhangs.

2. To the eluate from previous step, add: 1.0 µl Paired_End-P2 Adapter (10 µ*M*), 0.5 µl rATP (100 m*M*), 0.5 µl concentrated T4 DNA Ligase. Incubate reaction at room temperature for 30 min.

3. Purify with QIAquick column. Elute in 52(50) µl EB.

OPTIONAL: 25 ml of the eluate was digested again with SbfI for 30 min to remove rare genomic DNA concatemers formed from re-ligation of short fragments with two SbfI restriction sites within 500 bp. The sample was purified, eluted in 50 ml and quantified

using the Quant-iT TM dsDNA HS Assay Kit and Qubit TM fluorometer (Invitrogen).

3.10. RAD tag Amplification/Enrichment

1. In this step you will perform high-fidelity PCR amplification on P1 and P2 adapter-ligated DNA fragments, enriching for RAD tags that contain both a P1 and P2 and preparing them to be hybridized to an Illumina Genome Analyzer flow cell (*see* **Fig. 1**).

2. **OPTIONAL?:** Perform a test amplification to determine library quality. In thin-walled PCR tube, combine: 10.5 µl H_2O, 12.5 µl Phusion High-Fidelity Master Mix, 1.0 µl Solexa primer mix (10 µ*M*), 1.0 µl RAD library template (eluate from last step). Perform 18 (14) cycles of amplification in thermal cycler: 30 sec 98° C, 18X [10 sec 98° C, 30 sec 65° C, 30 sec 72° C], 5 min 72° C, hold 4° C. Run 5.0 µl PCR product in 1X Orange Loading Dye out on 1.0% agarose gel next to 1.0 µl RAD library template and 2.0 µl GeneRuler 100 bp DNA Ladder Plus (**Fig. 3**).

3. If the amplified product is at least twice as bright as the template, perform a larger volume amplification (typically 50-100 µl) to create enough to retrieve a large amount of the RAD tag library from the final gel extraction in the protocol.

(NOTE AND SUGGESTION: (circa 40 ng is used as template in a 100 µl PCR): some authors carry out the final amplification in two separate 50 µL PCRs per library each (25 µl Phusion High-Fidelity Master Mix, 2.5 µl Solexa primer mix (10 µ*M*), 2.5 µl RAD library template (eluate from last step), 20 µl H_2O. The two aliquots (2 x 50 µL) are combined before the final size selection).

If amplification looks poor, use more library template in a second test PCR reaction (*see* **Note 9**). **Fig. 3** shows three libraries that amplified well, which is apparent when comparing the amplified product to the amount of template loaded in the lane to the right of each sample. Template should be dim, yet visible on the gel. Purify large volume reaction with a MinElute coulumn. Elute in 20 µl EB.

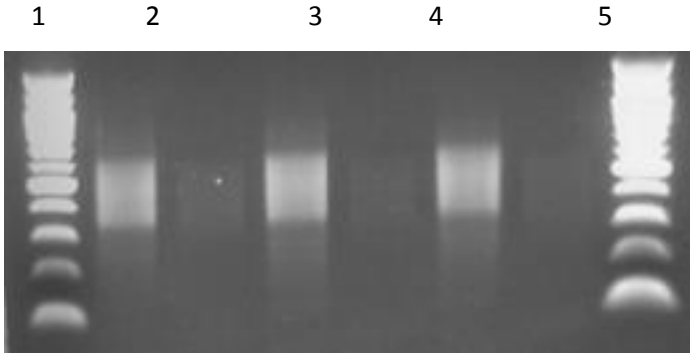

Fig. 3. Test amplification PCR product from the three libraries shown in **Fig. 2**. **2**, **4** & **6** contain 5.0 µl amplified PCR product. **2** - F_0 *Sbf*I library. **4** - F_2 *Sbf*I library. **6** - F_0 *EcoR*I library. **3**, **5** & **7** contain 1.0 µl template used for amplification in the lane to the left. Template was loaded at 5X the amount used in the equivalent volume loaded for amplified reactions. **1** & **8** - 2.0 µl GeneRuler 100 bp DNA Ladder Plus. Libraries are 3-600 bp in size.

4. This purification step is performed to eliminate any contaminant bands that may appear due to an improper ratio of P1 adapter to restriction-site compatible ends (*see* **Note 6**). Load entire sample in 1X Orange Loading Dye on a 1.25% agarose, 0.5X TBE gel and run for 45 min at 100 V, next to 2.0 µl GeneRuler 100 bp DNA Ladder Plus for size reference (**Fig. 4**). Being careful to exclude any free adapters or P1 dimer contaminants running at ~130 bp and below, use a fresh razor blade to cut a slice of the gel spanning 300-500(850) bp in size in an inverted triangle shape. PCR amplification of a wide-range of fragment sizes often results in biased representation of amplified products with an increased number of short fragments. We found this to be true in our current protocol, but reduced

the effects by selecting a triangular slice during gel extraction to reduce the level of short fragment lengths from the PCR reaction. Extract DNA using MinElute Gel Purification Kit following manufacturer's instructions. Melt agarose gel slices in the supplied buffer at room temperature. Elute in 20 µl EB. (optional: diluted to 10 nM).

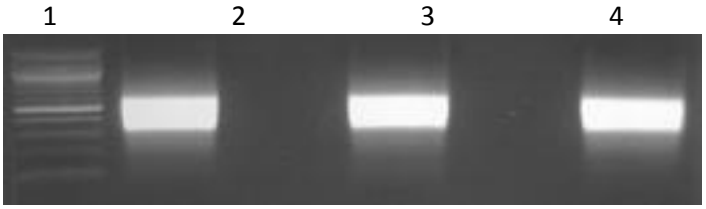

Fig. 4. PCR product from the three libraries shown in **Figs. 2 & 3** after the final large volume amplification and purification. **2**, **4** & **6** each contain 20 µl purified PCR product from 100 µl amplifications. **2** - F_0 *Sbf*I library. **4** - F_2 *Sbf*I library. **6** - F_0 *EcoR*I library. **1** - 2.0 µl GeneRuler 100 bp DNA Ladder Plus. **3**, **5** & **7** are blank. Libraries are 3-600 bp in size.

5. Quantify the DNA using a fluorimeter to get the most accurate concentration readings. Concentrations will range from 1-20 ng/µl. Determine the molar concentration of the library by examining the gel image and estimating the median size of the library smear, which should be around 400 bp. Multiply this size by 650 (the molecular mass of a base-pair) to get the molecular weight of the library. Use this number to calculate the molar concentration of the library (*see* **Note 10**).

Sequenced on the Paired-end module of the Genome Analyzer following Illumina protocols for 2x100 bp reads.

6. **IGNORE**: Validate library by cloning 1.0 µl of the gel purified library into a blunt-end compatible sequencing vector. Sequence individual clones by conventional Sanger sequencing. Verify that the insert sequences are from the genomic source DNA.

7. Sequence libraries on Illumina Genome Analyzer following manufacturer's instructions.

4. Notes

1. This protocol has been modified from that used in Baird et al., 2008 ***(12)*** and now incorporates critical improvements we have made since publication, including ones adopted from Quail et al., 2008 ***(14)***. Although we recommend following the described protocol exactly as stated, using the reagents we suggest, competing companies may offer cheaper versions or reagents that come at lower enzyme concentrations that will work just as well. Use of these reagents may require additional optimization, including increased incubation time or larger reaction volumes, for optimal

RAD tag library preparation. For instance, QIAquick columns may be substituted for MinElute columns in many places; however, reaction volumes in the subsequent step will have to be increased because of the increased elution volumes required for maximum recovery from the QIAquick columns.

2. Unless stated otherwise, all solutions should be prepared in water that has a resistivity of 18.2 MΩ-cm and total organic content of less than five parts per billion. This standard is referred to as "water" or "H_2O" in this text.

3. We recommend using a fluorescence-based method for DNA quantification in order to get the most accurate concentration readings. Since they bind specifically to double-stranded DNA, the dyes used in fluorimetric assays are not as affected by RNA, free nucleotides or other contaminants commonly found in DNA preparations (which can lead to inaccurate concentration predictions when using absorbance). If using another form of DNA quantification, such as UV spectrometer 260/280 absorbance readings, be sure to confirm the concentration by running a sample on an agarose gel and comparing to a known quantity of DNA or ladder. We recommend checking the integrity of at least a subset of samples on a gel prior to embarking on this protocol regardless of the quantification method, especially when working with many samples. Genomic DNA should consist of a fairly tight high molecular weight band without any visible degradation products or smears. When working with degraded DNA samples is the only option we have found that parameters of the protocol can be optimized (such as using more input DNA to start with and shearing less) to create usable libraries. These libraries often don't amplify as well as ones made with intact, high-quality genomic DNA.

4. Prepare 100 µ*M* stocks for each single stranded oligonucleotide in 1X Elution Buffer (EB: 10m*M* Tris-Cl, pH 8.5). Combine complementary adapter oligos at 10 µ*M* in 1X AB (10X AB: 500 m*M* NaCl, 100 m*M* Tris-Cl, pH 7.5-8.0). Place in beaker of water just off the boil, cool slowly to room temperature to anneal. Dilute to desired concentration in 1X AB. The presence of some salt is necessary for the double-stranded adapters used in this protocol to hybridize and to remain stable at ambient temperatures and above. At a 1m*M* salt concentration the P1 adapter, which has 62 bases of complementary double-stranded sequence (assuming a 4 base pair barcode), has a T_m of approximately 40º C (depending on the barcode composition). P2, which has only 24 complementary bases, has a T_m of only 27º C at the same salt concentration. At 50 m*M* salt the T_ms jump up to ~69º and 56º, respectively.

5. In general, making master mixes, using multi-channel pipettes and dealing with samples in 96- or 384-well plates will speed up the restriction digest and P1 ligation steps when multiplexing multiple barcoded individuals.

6. *EcoR*I has been shown to work robustly in multiple organisms in our lab. **Restriction enzymes that cut less frequently create fewer RAD tags, and thus require more input DNA and less P1 adapter to keep the molar ratio approximately equal.** Less frequent cutters are more difficult to amplify in general and protocol parameters may take some optimization for favorable results. It is critical to optimize the amount of P1 adapter used when a given restriction enzyme is used for the first time in an organism, unless the actual number of sites is known. Otherwise, some optimization may be required to ensure enough P1 is used to get robust RAD library amplification without using too much. If the ratio of P1 adapter overhangs to available genomic compatible ends is too low, you will get insufficient amplification and/or biased representation of some RAD tags. However, if the ratio of P1 to genomic overhangs is too high, **a contaminant band that runs around 130 bp will appear after the final PCR reaction**. If this contaminant overwhelms the amplification reaction it can lead to significant adapter sequence reads in the final sequencing output (even after gel extraction following the final PCR). This phenomenon is completely dependent upon the number of actual cut sites present in that genome. Our *Sbf*I study in stickleback used 2.5 µl P1 per microgram starting material and performed very well for library construction (*see* **Figs. 3** & **4**; lanes 2 & 4); however, this is likely to due to the fact that there are actually more *Sbf*I sites than expected by chance. Therefore, starting with less P1 may be preferable for genomes with closer to the expected number of sites.

ligation molarity calculator - guide for calculating appropriate adapter concentration for ligations.

7. DNA samples from 96 recombinant F_2 individuals were uniquely barcoded, which allowed us to track RAD markers and associate them with differing lateral plate or pelvic structure phenotypes. F_2 individuals used in the mapping analysis included 60 fish possessing the complete lateral plate phenotype, 31 low lateral plate individuals. The barcoded samples from fish possessing the same lateral plate phenotype were combined and treated as one library after P1 ligation. In order to genotype all F_2 individuals with a low pelvic structure phenotype, the DNA from 5 individuals that had a low pelvic score, but that had a partial lateral plate phenotype, were barcoded and processed with the low plate group. The two multiplexed libraries included 67 individuals demonstrating the high pelvic structure phenotype and 29 with a low pelvic score that were resorted *in silico* to map this second trait. For the bulk-segregant analysis using *Sbf*I, one library was prepared using two pooled DNA samples from recombinant F_2 individuals, combined according to lateral plate phenotype prior to restriction digestion. The digested pools were labeled with different barcodes, combined and treated as one library after P1 ligation.

8. We have found it is unwise to run more than one library sample on the same agarose gel, as is shown in the figures, unless they will be combined and sequenced in the same lane on the flow

cell, because it can lead to contamination between samples. This is especially important when dealing with samples following PCR amplification. We recommend using aerosol-resistant filter tips for all amplification and downstream steps in the protocol to avoid library contamination.

9. Libraries that amplify robustly, such as those shown in **Fig. 3**, can be amplified with only 16 or fewer cycles of PCR to avoid skewing the representation of the library ***(14)***. Robust libraries can often times be cleaned up without the final gel extraction step if there are no visible contaminant bands running below 300 bp on the gel after the test amplification. We have retrieved good sequence read numbers from libraries that amplified well with only 4 µl of template in a 100 µl reaction as well as ones that amplified very poorly and required up to 30 µl template in the same volume. The first scenario is preferable, as you can be more confident of the true concentration of RAD tag molecules in your sample, which have both P1 and P2 sequences, and are therefore able to bind to adapter oligonucleotides present on the Illumina flow cell. Poorly amplified libraries will contain a greater number of background sheared genomic DNA fragments with only P2 adapters attached, which cannot bind to the flow cell.

10. For long-term storage of DNA samples, Illumina recommends a concentration of 10 n*M* and adding Tween 20 to the sample to a final concentration of 0.1% Tween. This helps to prevent adsorption of the template to plastic tubes upon repeated freeze-thaw cycles, which would decrease the cluster numbers from a sample over time.

References

1. Berger J, Suzuki T, Senti KA, Stubbs J, Schaffner G et al. (2001) Genetic mapping with SNP markers in Drosophila. Nat Genet 29: 475-481.

2. Stickney HL, Schmutz J, Woods IG, Holtzer CC, Dickson MC et al. (2002) Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. Genome Res 12: 1929-1934.

3. Wicks SR, Yeh RT, Gish WR, Waterston RH, Plasterk RH (2001) Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat Genet 28: 160-164.

4. Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E et al. (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci U S A 101: 9915-9920.

5. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331.

6. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-4414.

7. Chen D, Ahlford A, Schnorrer F, Kalchhauser I, Fellner M et al. (2008) High-resolution, high-throughput SNP mapping in Drosophila melanogaster. Nat Methods 5: 323-329.

8. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res 17: 240-248.

9. van Orsouw NJ, Hogers RC, Janssen A, Yalcin F, Snoeijers S et al. (2007) Complexity reduction of polymorphic sequences (CRoPS): a novel approach for large-scale polymorphism discovery in complex genomes. PLoS ONE 2: e1172.

10. Miller MR, Atwood TS, Eames BF, Eberhart JK, Yan YL et al. (2007) RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biol 8: R105.

11. Lewis ZA, Shiver AL, Stiffler N, Miller MR, Johnson EA et al. (2007) High-density detection of restriction-site-associated DNA markers for rapid mapping of mutated loci in Neurospora. Genetics 177: 1163-1171.

12. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE. 3(10):e3376. Epub Oct 13.

13. Coyne KJ, Burkholder JM, Feldman RA, Hutchins DA, Cary SC (2004) Modified serial analysis of gene expression method for construction of gene expression profiles of microbial eukaryotic species. Appl Environ Microbiol 70: 5298-5304.

14. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ (2008). A large genome center's improvements to the Illumina sequencing system. Nat Methods. Dec;5(12):1005-10.