Microsatellite-Enhanced Genomic Library Construction
Streptavidin-Biotin Bench Version
SNX Linker (version 1.1 December, 2002)

Purpose To construct and clone genomic libraries with increased proportions of inserts that contain tandem repeat arrays. Thus, a greater number of microsatellite repeat regions can be detected, sequenced and subsequently used to design species-specific flanking primers for microsatellite amplification. This protocol greatly improves enrichment by increasing the efficiency of linker ligation to genomic DNA. This prevents cloning of repeat sequences with only one flanking region and increases the yield of repeat sequences. Based on Armour et al., 1994 and Fleischer and Loew, 1995.

These methods are described in: Hamilton M. B., E. L. Pincus , A. DiFiore , and R. C. Fleischer. 1999. A universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. BioTechniques 27(3): 500-507.

We would also be grateful for any suggestions or improvements. We thank A. Colwell, T. Burg and K. Helenurm for insightful questions and comments and their willingness to try something new.

Disclaimer: Persons using this protocol do so at their own risk. Every effort has been made to verify the procedures described here and the technique has been used successfully on multiple occasions, in multiple labs. However, the authors cannot be held liable for failure of the protocol to produce microsatellites. Users must exercise proper professional judgment when carrying out these procedures and are encouraged to make modifications that are appropriate for their starting materials and end goals. We have not exhaustively described every aspect of these molecular genetic procedures and required safety precautions. This protocol is not intended for those without background in molecular genetic laboratory procedures.

Detailed Protocol

First, a word about the source of restriction and modifying enzymes for this protocol. After much frustration when using enzymes from a variety of suppliers, we obtain all of our enzymes for these procedures from New England BioLabs (NEB). We do this because NEB tests enzymes for activity in all of their buffer systems, provides a catalog with pertinent technical information, and gives excellent technical support over the phone. Specifically, NEB ligase is active in three of their four restriction enzyme buffers, allowing simultaneous digestion and ligation required by the linker design below. Contrary to speculation, MBH does not own shares of NEB stock.

We also use GelStar (FMC) to stain agarose gels for these procedures. This is because GelStar is more sensitive than regular ethidium bromide and much less DNA is used up

in running a gel. This helps to save products such as ligations for subsequent procedures.

I. SNX linkers

I A. Have complementary linkers synthesized (make sure the 5′ end of the SNX reverse linker is purchased 5′ phosphorylated) so that double-stranded they have a blunt end and a 3′ overhang.

1. SNX: 5'- CTA AGG CCT TGC TAG CAG AAG C -3'
2. SNX rev: 5'- pGCT TCT GCT AGC AAG GCC TTA GAA AA -3'

```
                          Stu I      Nhe I       Xmn I
e.g. 5′-       CT AAG GCC TTG CTA GCA GAA GC  -3'
     3′- AAA AGA TTC CGG AAC GAT CGT CTT CGp -5'
```

Notice that these linkers have several restriction enzyme sites. The *Xmn* I site is used to prevent the linker from ligating to itself to form blunt-ended dimers during Step IV A. This maintains more linker available to ligate to genomic DNA. *Stu* I and *Nhe* I do not cut pBS SK and therefore can be used to prevent linker-ligated genomic DNA fragments from ligating together during ligation into vector. The *Stu* I site is complementary to the *EcoR* V site in pBS SK (or any blunt site in another vector) while the *Nhe* I site is complementary to the *Xba* I site in pBS SK.

The A tail polarizes the linker so that only one end can serve as a blunt end for ligation. This tail must be left in place or both ends will be blunt. Imagine that two SNX linkers are ligated together (the "-" indicates the ligation location):

```
5′-      CTAAGGCCTTGCTAGCAGAAGC-GCTTCTGCTAGCAAGGCCTTAGAAAA - 3'
3′- AAAAGATTCCGGAACGATCGTCTTCG-CGAAGACGATCGTTCCGGAATC     - 5'
```

In this way the SNX linkers can form dimers, but not larger polymers because of the A tails. Notice that when two linkers ligate together they form an *Xmn* I site (bold and underlined above). The use of *Xmn* I in the ligation reaction will cut these linker dimers apart. Thus the use of *Xmn* I will make sure that the population of SNX linkers in a ligation reaction contains a large proportion of monomers. This feature of the linkers makes blunt-end ligations many times more efficient.

Now imagine that the linker ligates to a blunt-end genomic DNA fragment (indicated by N's):

```
5′-      CTAAGGCCTTGCTAGCAGAAGC-NNNNNNNNNNNNN
3′- AAAAGATTCCGGAACGATCGTCTTCG-NNNNNNNNNNNNN
```

The *Xmn* I will not cut this ligation product as long as the genomic DNA fragment is NOT produced by *Xmn* I.

Now imagine that we have enriched this population of linker-ligated genomic DNA fragments. They will look like:

```
5'-     CTAAGGCCTTGCTAGCAGAAGC-NNNNNNNNNN-GCTTCTGCTAGCAAGGCCTTAGAAAA-3'
3'-AAAAGATTCCGGAACGATCGTCTTCG-NNNNNNNNNN-CGAAGACGATCGTTCCGGAATC     -5'
```

If we use *Nhe* I (the site is underlined above) to cut the linkers this will give a four-base overhang for ligation into our vector. The inserts will look like this after digestion with *Nhe* I:

```
5'- CTAGCAGAAGC-NNNNNNNNNN-GCTTCTG      - 3'
3'-     GTCTTCG-NNNNNNNNNN-CGAAGACGATC - 5'
```

It is important to remember that the genomic DNA MUST be digested with *Nhe* I initially or else some of the genomic inserts may contain *Nhe* I sites. (Any genomic inserts containing *Nhe* I sites would be cut and therefore would ligate into vector as inserts with only one SNX linker section. This is not an enormous problem but might tend to reduce the flanking regions available for primer design. It is best to avoid the unpredictability of this situation by using *Nhe* I initially on genomic DNA in addition to other enzymes and then chewing any overhangs back with mung bean exonuclease.).

This linker design assumes that *Xmn* I will cut very infrequently in the genomic DNA of the target organism. This is because any genomic DNA end that is the result of a cut by *Xmn* I will not be ligatable to the linker (ligation of the linker to the genomic end will re-create the cut site, which will then be re-cleaved by the enzyme). Test digests using genomic DNA and *Xmn* I should be conducted to confirm this assumption. The SNX linker and *Xmn* I can still be used if *Xmn* I cuts frequently in a given genome. Notice that EcoR I cuts within the *Xmn* I site. Just add this enzyme to your genomic digestion reactions and then chew back the overhangs with mung bean nuclease (see below). This will destroy all *Xmn* I sites within your genomic DNA and *Xmn* I will not cut genomic DNA during linker ligation.

A Note on the Dephosphorylation Step Before ligating SNX: Several people have contacted me with questions about using dephosphorylated insert DNA. If the insert DNA is dephosphorylated, then when the SNA linker is ligated to the insert DNA one strand will carry a nick. The construct will look like:

```
5'-     CTAAGGCCTTGCTAGCAGAAGC  NNNNNNNNNN-pGCTTCTGCTAGCAAGGCCTTAGAAAA-
3'
3'-AAAAGATTCCGGAACGATCGTCTTCGp-NNNNNNNNNN  CGAAGACGATCGTTCCGGAATC     -
5'
```

where the space indicates base pairs that cannot ligate due to lack of 5' phosphates on the insert DNA. This nick will cause the shorter SNX strand to "fall off" if the ligation products are heated or purified (e.g. with QiaQuick columns). In practice this possibility was not observed in our earlier lab work. This can be explained by at least

two possibilities: 1) the SNX linker attaches to both ends of genomic DNA fragments only to the extent that the CIP is not 100% efficient, or 2) the ligase we use is regularly contaminated kinase or has some type of kinase activity.

In any event, I tried simply omitting the dephosphorylation of insert DNA before ligating the SNX linker. We did not observe a high rate of genomic DNA polymerization when trying this strategy on one species (as of 12/2002) but did get some nice microsatellite loci. I hypothesize that chimerization is relatively uncommon even in the absence of CIP'ed inserts because the linker is in large excess in the ligation reaction.

A second possibility is to repair the nick using a PCR reaction (which does not contain primers) before proceeding on to subtractive hybridization. This latter method was proposed by Travis Glenn (see http://129.252.89.20/Msats/Microsatellites.html). I have not yet used this method but I hope to try it soon. See review by Zane et al. 2002.

I thank Travis Glenn and several others for asking questions about this point. It forced me to think the process through even though the protocol had worked fine. This may help explain why some people have had considerable difficulty with the SNX protocol when using different suppliers of CIP and ligase.

II. Processing insert DNA

II A. Digest to completion at least 5 µg of clean, high molecular weight genomic DNA (from individual of heterogametic sex, if possible) with a 10-fold excess of restriction enzyme. With the linkers given above, you can use restriction enzymes that are blunt cutters, or ones that leave an overhang, or a combination of both blunt and overhang cutters.

One can and should employ an enzyme (or set of enzymes) that cuts genomic DNA so that a large proportion of fragments are in the 1000 to 200 base pair range. The best way to accomplish this is to conduct a series of test restriction digests with single and multiple enzymes to determine which enzyme or combination of enzymes will give the desired size distribution. For example, the combination of *Hae* III, *Rsa* I, *Alu* I and *Nhe* I is used to digest genomic DNA from a tropical tree, giving the majority of fragments within the desired size range. NEB provides a buffer compatibility chart in their catalog that will help you select a buffer when combining enzymes.

A commonly employed strategy is to size-select genomic DNA by cutting the 1000-200 bp range from an agarose gel. We do not recommend this procedure for at least two reasons. First, it is difficult to reclaim the genomic DNA from the agarose and much DNA is lost in the process. Any residual ammonium salts (such as NH_4OAC) used to precipitate the genomic DNA from digested agarose will inhibit the action of ligase.

Second, using only a portion of the genome for microsatellite cloning is a sampling step. We have shown (Hamilton and Fleischer, 1999) that cloning from size-selected digests of *Dpn* II and α*Taq* I gives significantly different numbers of hybridization-positives for the same repeat sequence after enriched DNA is cloned and transformed into *E. coli*. Since most people are interested in random markers for population genetic inferences, size-selection is detrimental because this sampling will bias the resulting marker pool.

If using the *Stu* I site in the SNX linker to ligate to a blunt site in a vector, the restriction enzyme *Hae* III should be used in all genomic digests. The *Hae* III site (GGCC) is present inside the *Stu* I site, but *Hae* III leaves a blunt end. Any combination of blunt cutters can be used with *Hae* III to digest DNA to desired size range (see NEB catalog for efficiency of particular restriction enzymes in alternate buffers). Alternatively, a combination of *Hae* III and overhanging restriction enzymes can be used to digest genomic DNA. If this strategy is followed, mung bean exonuclease should be used to chew off sticky ends prior to linker ligation. The same logic applies if planning to use the *Nhe* I site to ligate to the *Xba* I site in a vector. The genomic restriction digest must contain *Nhe* I and the resulting overhangs must be chewed to blunt ends with mung bean exonuclease. We highly recommend using *Nhe* I for ligation into vector since the four base overhang ligates very efficiently. IMPORTANT - See II B for specific mung bean exonuclease protocol.

Digest a large amount of genomic DNA after determining which combination of enzymes yields most fragments in the 1000-200 bp size range. Note that the total volume of restriction enzyme added to a digest should not exceed 1/10 of the total volume of the digest or the glycerol storage buffer will inhibit digestion. Use the manufacturer's conditions and buffer and digest overnight at 37˚C in a total volume of 20-100 µl. Use DNA that has been RNAse A digested during extraction. An example blunt-end digest is:

1. genomic DNA, 5-25 µg	81 µl
2. 10x buffer (NEB #2)	10 µl
3. BSA (NEB 100x)	1 µl
4. *Hae* III	2 µl
5. *Rsa* I	2 µl
6. *Alu* I	2 µl
7. *Nhe* I	2 µl

Run a small aliquot of this digest on 2% agarose to check completeness of digestion. If digestion is not complete, allow more time.

II B. If you have used restriction enzymes that generate overhangs, chew off these single-stranded overhangs with mung bean nuclease after digestion. This step can be carried out in NEB restriction buffers 1, 2 and 4. NEB recommends using mung bean nuclease at 30˚C for 30 minutes. At higher temperatures the DNA will "breathe" (meaning some regions of double stranded DNA become single stranded), so it should not be added directly to restriction digests which are carried out at 37˚C. The sample

should then be QiaQuicked (a spin column from QiaGen) to remove the mung bean nuclease before dephosphorylating. Elute DNA with 50 µl of the EB Buffer provided with the columns.

II C. One could now **dephosphorylate** the 5' ends of the digested genomic DNA. This step greatly decreases the likelihood of insert fragments ligating to themselves, thereby creating chimeric sequences (sequences from two locations in the genome that are made contiguous) that will be useless when designing microsatellite primers (see Sambrook et al. 1989).

A Note on the Dephosphorylation Step You have two options here: i) This step can be skipped, ii) if dephosphorylation is carried out, an additional (but easy) PCR step will have to be done to repair the nicked strand after linker ligation.

To the 50 µl of purified DNA from above, add

restriction enzyme buffer 2, 3 or 4	6 µl	
sterile water		3 µl
CIP (short for "alkaline phosphatase, calf intestinal")		1 µl

Vortex gently and incubate at 37˚C for at least 2 hours.

II D. Purify the CIP'ed genomic DNA of all modifying enzymes using a QiaQuick PCR purification column following the instructions. This time elute the DNA from the column in 30 µl of the EB Buffer provided with the columns.

III. Make linker oligo stocks

III A. Make a 100 µM stock of each linker oligo in fresh, sterile T.E. (see appendix for recipe). Newly synthesized oligos are provided with yield information (usually the number of pmoles synthesized). Add T.E to make a 100 µM stock solution in the tube the dried oligos were shipped in. Then make 200 µl of 10 µM stock solutions of each linker oligo by diluting the 100 µM stuff 1 to 10 in T.E. Store all oligo stocks at -20˚C.

III B. Combine equimolar amounts of SNX linker and SNX linker reverse to construct a double-stranded linker which will be ligated to the digested, dephosphorylated insert DNA from step II above. For example, combine 50 µl of each 10 µM linker in one tube to make a 5 µM double-stranded linker stock. Mix and leave at room temperature for a few minutes so strands will anneal, then store at -20˚C.

IV. Ligation of insert to linkers

IV A. Cohesive termini or "sticky end" ligation proceeds rapidly at linker to genomic fragment molar end ratios near 1:1. For blunt end ligations, however, linker to genomic fragment molar end ratios need to be in the range of 40:1 to 100:1 and a larger amount of ligase is needed. See appendix 1 for methods to calculate molar end concentrations and discussion in Sambrook et al. 1-68. Below is an **example** of a 30 µl blunt end linker-ligation reaction.

1. double stranded linker (5 µM)	11.7 µl
2. NEB Buffer #2	3.0 µl
3. 100X BSA	0.3 µl
4. digested, CIP'ed, QiaQuick cleaned genomic DNA	10.0 µl
5. rATP (10 mM)	3.0 µl
6. *Xmn* I	1.0 µl
7. ligase (high concentration, 1000 U)	1.0 µl

Incubate overnight in thermocycler programmed for cycles of 30 minutes at 16˚C and 10 minutes of 37˚C. The ligation can also be left at room temperature. NEB states that sticky end ligation reactions may be enhanced by using 50 µM rATP instead of 1 mM (however, we found 1 mM to work best). See ligase description in NEB catalog. At the end of the ligation heat kill *Xmn* I and ligase by heating at 65˚C for 20 minutes.

We use ATP from Sigma: catalog A-6559 (100mM, pH 7). You can also make it yourself from powder but the pH'ing step is difficult for small volumes. The time savings justifies the high cost. By the way, we call it "r"ATP just to signify that it must have a full ribose to be active as a phosphate donor in a ligation. The dATP that you would use in a PCR reaction will not work.

PCR nick repair: If you carried out a CIP reaction to remove 5' phosphates from the insert DNA, you should now repair the nick with a PCR reaction. The reaction would just be a PCR mix like that that given below BUT WITHOUT ANY PRIMERS. The "template" would be the entire 30 uL ligation reaction. Note that the repair is carried out by nick translation. You must use *E. coli* DNA polymerase I or Taq to accomplish this. Vent will not work! (see polymerases - Properties of DNA Polymerases in the NEB technical literature section http://www.neb.com/neb/frame_tech.html)

IV B. Run 5 µl of ligation on a 2.0% agarose gel with size marker. Include digested genomic DNA and 5 µM double-stranded SNX linker on gel as controls. Successful ligation is confirmed by linker-ligated insert DNA being slightly larger than the

unligated genomic DNA. Linker monomers and dimers (i.e., bands at about 22 and 44 bp) should be visible in the ligation lanes. We recommend using GelStar (FMC) to stain this gel (see introductory note). This gel is optional. The PCR control below is a better test for efficiency of ligation. We usually omit this step but do include digested genomic DNA and double-stranded SNX linker lanes on the PCR gel of step IV D.

IV C. The linker-ligated inserts can be amplified directly from the ligation in a PCR reaction with the SNX linker oligo as a primer. **Conducting this control is IMPERATIVE since a poor ligation will render enriched DNA impossible to recover and clone.** Here is the PCR recipe for a 50µl PCR reaction

Place in PCR tube

1. Sterile water	33.7 µl
2. ligation reaction template	2.0 µl

Make a mix of the following (multiplied by the number of tubes)

3. 10X Thermopol buffer (20 mM $MgSO_4$)	5.0 µl
4. dNTP mix (2 mM each base =8 mM)	5.0 µl
5. SNX linker (10 µM)	4.0 µl
6. Vent exo- Polymerase	0.3 µl

Thermal profile is: 96˚C for 5 minutes followed by 40 cycles of 96˚C for 45 seconds - 62˚C for 1 minute - 72˚C for 1 minute. Be sure to include double-stranded SNX linker, digested genomic DNA and water as PCR controls.

IV D. Run a 10 µl aliquot of the amplified insert PCR on a 2.0% agarose gel to check amplification. Also load digested genomic DNA on gel as a control. Successful amplification indicates successful ligation.

V. Repeat-enrichment with biotin-labeled oligos and streptavidin beads

The use of biotin-labeled oligos and a streptavidin matrix is described in: Kandpal et al., 1994. Our work also benefited from communication with Shou-Hsien Li of the State University of New York at Albany.

At this point we need to increase the proportion of genomic DNA fragments containing repeat motifs. This is accomplished by hybridizing genomic DNA (with attached linkers) to synthesized oligonucleotide repeats. Genomic DNA that is complementary to the oligos will be captured and much of the genomic DNA lacking repeats will be removed. Biotin labeled oligos, which bind strongly to streptavidin, are used in the hybridization. Streptavidin-coated iron beads are then added to the hybridization mixture. The biotinylated oligos, along with bound strands of genomic DNA, bind to the beads. A magnet can then be used to separate the beads (and attached DNAs) from

the wash solutions. After washing away genomic DNA not bound to the oligos, the repeat-enriched genomic DNA is made double stranded and multiplied by PCR.

First, you'll need to order repeat oligos **with biotin labeling** for the sequence motifs you wish to enrich for. Enrichment oligos can be **3' modified** so that any residual oligo that carries over from the enrichment cannot act as a PCR primer (see Koblizkova et al., 1998). This is not absolutely critical, especially if you already own 5' biotinylated oligos. The 3' modification was suggested to prevent potential amplification from one linker primer to the microsatellite repeat (primed by residual oligo from the enrichment). This would produce clones without flanking region on one side of the microsatellite repeat (e.g. Gardner et al., 1999)

We find that the increased efficiency of linker ligation prevents any PCR artifacts due to the minute amount of enrichment oligo that remains after subtractive hybridization. If linkers are attached to both ends of a genomic DNA fragment, amplification from the two linker primers will easily out-compete amplification from one linker primer to the microsatellite repeat. This is because the linker primer is available in great excess compared to any residual enrichment oligo. We feel that the lack of flank is actually due to poor ligation of the linker, not contamination by enrichment oligo.

It is also possible to use repeat oligos that you already have even if they are not biotin labeled. In this case, use terminal transferase (TdT) to attach a biotin labeled dATP or ddATP (available from Sigma and Gibco-BRL) to the 3' end of the oligos. Notice that these oligos will be biotin labeled and 3' modified. Using a dideoxy nucleotide is the best protection against PCR artifacts and having multiple ATPs attached to the repeat oligos. The unincorporated biotin-ddATP should be removed with Sephadex columns after TdT labeling.

We used oligos around 30 base pairs, regardless of repeat size. These oligos could be HPLC purified, although such purification is expensive. It separates oligos that are not full length (as much as 25% of the total yield according to Operon) or lack a biotin end-label from those that are full-length and properly biotinylated. We do not use HPLC purified oligos. If the biotinylated oligos are not HPLC purified it is necessary to add more oligo for hybridization. You may also want to buy one or two **complements** of the repeat oligos that are **not biotinylated** (this is optional but can be used for hybridization controls when screening bacterial colonies for repeat-containing inserts). You will also need streptavidin beads and a fancy magnet, which are available from Dynal or Promega. The magnet is a critical item, even if it is ridiculously expensive. Remember to ask Dynal for protocol information, even though it won't help you much.

V. A. Hybridize linker-ligated genomic DNA and biotinylated repeat oligos.
First, make up hybridization buffer (12X SSC, 0.1% SDS). The hybridization buffer may need to be warmed (e.g. in a 37˚C water bath) before use if the SDS has precipitated. The amount of genomic DNA in the hybridization depends on how much is available.

More is generally better, but make sure you have enough for all repeat hybridizations. Although all repeat oligos can be added to a single hybridization, we suggest an independent hybridization for each repeat. This is a bit more work but allows tighter control over hybridization and wash temperatures and will give more product for repeats with very different melting temperatures or repeats that are rare. We used enough biotinylated oligo for a 1:1 fragment ratio between genomic DNA and the oligo (this includes a generous fudge-factor for short oligos that will not be biotinylated). This is obviously more oligo than necessary, which is fine as long as enough beads are added to bind all the oligo.

1. linker-ligated genomic DNA (about 100 ng)	6 µl
2. biotinylated repeat oligo (1 µM)	2 µl
3. hybridization buffer?	50 µl
4. sterile, distilled water	42 µl
Total Volume:	**100 µl**

This hybridization reaction has a NaCl concentration of 1 M. Overlay these hybridizations with a few drops of mineral oil to prevent evaporation. Heat hybridizations to 95˚C for 15 mins. to denature genomic DNA. Place at hybridization temperature (depending on repeat composition and sodium concentration) for several hours to overnight. Hybridization temperatures should be 5 to 10 degrees below calculated T_m (use the rule-of-thumb that an A/T pair contributes 2˚ and a G/C pair contributes 4˚).

V. B. Prepare beads.
Aliquot into a single 1.5 ml microfuge tube enough beads for all hybridization reactions (we use 300 µg per 2 pmoles of biotinylated oligo, which assumes a binding capacity around 10 pmoles of hybrids per mg of beads). Rinse beads with "binding & washing buffer" (B&W buffer; 10 mM Tris [pH 7.5], 1 mM EDTA, 1 M NaCl). Rinse 3 or 4 times using magnet to hold beads when removing wash solutions. Vortex beads during each wash (beads will settle out of suspension).

V. C. Transfer hybridization reactions to new tubes (**1.5 ml tube to fit in magnet**) to get rid of mineral oil. Add beads to hybridization reactions. Agitate tubes several hours at 43˚ C (mainly to keep SDS from precipitating) so biotin and streptavidin can bind. Try to keep beads suspended during binding.

V. D. Make wash solutions in advance (2X SSC, 0.1% SDS and 1X SSC, 0.1% SDS) and aliquot into 1.5 ml tubes that can be kept at temperature in a heat block. Samples should be briefly vortexed during each wash. Place tubes in heat block during washes to maintain temperature. All washes are for 5 minutes.

a. Place tube in magnet (remember to wait about 1 min. for beads to stick to side of tube) and pipet off hybridization solution.

b. Wash twice for 5 minutes with 200 µl room temp 2X SSC, 0.1% SDS. Place tube in magnet to remove wash solution each time.

c. Wash twice for 5 minutes at 45˚ C with 200 µl 1X SSC, 0.1% SDS. Place tube in magnet to remove wash solution each time.

d. Wash twice for 5 minutes at 60˚ C with 1X SSC, 0.1% SDS. Place tube in magnet to remove wash solution each time.

Wash temperatures can be adjusted to increase (hotter) or decrease (cooler) hybridization stringency. We use 60˚ C for repeats such as CAG, GGA and CT or 45˚ C for AAT and AAG.

e. Elute genomic DNA bound to oligo by adding 60 µl of preheated T.E (that is tee-dot-eee). Heat to 95˚ C for 10 minutes. Working quickly, place sample in magnet and remove T.E to a new tube. A second elution may recover additional genomic DNA. Save beads in 60 µl of T.E at -20˚ C.

V. E. Amplify repeat-enriched DNA. This makes the single-stranded elution double-stranded and will increase the amount. We recommend using 2 mM $MgSO_4$.

Place in PCR tube

1. Sterile water	25.7 µl
2. post-hyb eluted DNA	10.0 µl

Make a mix of the following (multiplied by the number of tubes)

3. 10X Thermopol buffer (20 mM $MgSO_4$)	5.0 µl
4. dNTP mix (2 mM each base =8 mM)	5.0 µl
6. SNX linker (10 µM)	4.0 µl
7. Vent exo- Polymerase	0.3 µl

Thermal profile is: 96˚C for 5 minutes followed by 40 cycles of 96˚C for 45 seconds - 62˚C for 1 minute - 72˚C for 2 minutes. Be sure to include double-stranded SNX linker, digested genomic DNA and water as PCR controls.

V. F. Run a 5 µl aliquot of the amplified insert on a 2.0% agarose gel to check amplification.

Note: enrichment steps can be repeated using PCR products as starting material to increase repeat enrichment further. Just repeat V. A. through V. F. This should not be necessary.

VI. Digestion of linkers and ligation of genomic DNA into plasmid

VI. A. Prepare the SNX linkers from the post-enrichment amplified inserts to form ends for cloning. Multiple PCR reactions can be pooled to increase DNA amounts (scale digest as necessary). *Stu* I or *Nhe* I are about 50% active in Thermopol PCR buffer so it is possible to digest the PCR products directly in the PCR reaction mix.

1. Digest the PCR reaction with *Stu* I or *Nhe* I to give a cloning end.
    1. post-hyb insert DNA PCR 45.0 µl
    2. *Stu* I or *Nhe* I 1.0 µl
    3. 100x BSA for *Nhe* I 0.46 µl

    Incubate overnight at 37˚C.

    2. Run an aliquot (5 µl) of the the digest on a 2.0% agarose gel to verify digestion. Load an aliquot of double-stranded SNX linker digested with *Stu* I or *Nhe* I as a size control. Successful digestion of the post-enrichment amplified inserts is indicated by a band the same length as the smaller band in the control linker digest. This gel is optional. The digested end of the linker is very small (14 bp) and is difficult to see on an agarose gel (even with GelStar) without loading a large amount of the digest.
    3. If digestion of linkers is successful, remove linker fragments from insert DNA using a Qiagen Qiaquick PCR purification column according to manufacturer's protocol. Elute DNA with 30 µl of the EB Buffer provided with the columns.

    If this digest does not work well in Thermopol the PCR reactions should be cleaned up with a QiaQuick column and the digest then performed in NEB buffer #2 with *Stu* I or *Nhe* I. Then clean out digested off linker ends with a QiaQuick column.

VI. B. Digestion and dephosphorylation of vector DNA:

1. Digest 3 µg of pBluescript with *EcoR* VI or *Xba* I (remember that *EcoR* V will ligate to the *Stu* I site and *Xba* I to the *Nhe* I site). Allow to digest for several hours to overnight at 37˚C. Remove 5′ phosphates from vector by adding 1 µl of NEB CIP directly to digest and incubating again at 37˚C for several hours. Purify the vector DNA of all modifying enzymes using a QiaQuick PCR purification column following the instructions. Elute the DNA from the column in 30 µl of the EB Buffer provided with the columns. The resulting digested vector should then be at a concentration of about 100 ng/µl.

Check for complete digestion by running digest on a 1.2 % agarose gel adjacent to uncut pBluescript. Recall that uncut plasmid has multiple bands due to secondary

structure and the digest should be a single band because it is all linear double-stranded.

VI. C. Ligation of vector and insert:

For complementary overhangs ligation proceeds well at 1:1 or 2:1 molar ratio of insert to vector DNA. About 100 ng of vector with about 100 ng of insert DNA was used successfully but conditions will vary and DNA amounts should be estimated from concentration information and molar end ratios. We highly recommend using the *Nhe* I/*Xba* I sites since the resulting four base overhang ligates very well.

1. Construct a ligation reaction using the *Xba* I cloning site:

1. linker digested, post-enriched insert PCR DNA	12.8 µl
2. NEB Buffer #2	2.0 µl
3. rATP (10 mM)	2.0 µl
4. 100x BSA	0.2 µl
5. *Xba* I digested, pBS SK (100 ng/µl)	1.0 µl
6. *Nhe* I	1.0 µl
7. ligase (regular concentration, 400 U)	1.0 µl

Note that rATP stands for ribose ATP. dATP (deoxyribose ATP) used for PCR will not work as a phosphate donor!

Incubate overnight in thermocycler programmed for cycles of 30 minutes at 16˚C and 10 minutes of 37˚C. The ligation can also be left at room temperature. At the end of the ligation heat kill restriction enzyme and ligase by heating at 65˚C for 20 minutes in the thermocycler. Ligation controls with no insert and with ϕX 174 DNA are suggested.

If you are using the *Stu* I/*EcoR* V cloning sites, the ligation reaction should be set up like the original ligation to attach the SNX linkers. Use the reaction above but substitute high concentration ligase and *Stu* I (as in step IV A).

VII. Transformation of recombinant plasmids into competent *E. coli*

You must prepare a large number of LB plates with 75 µg/mL ampicillin before carrying out transformations. Make sure ampicillin is fresh!!! We spread X-gal and IPTG directly on the plates (see Sambrook et al.).

A. We used Stratagene Epicurian Coli® XL1-Blue MRF' supercompetent cells. Don't be cheap - buy supercompetent cells! Inoculate three plates per library (i.e., try 25, 50 and 75 µl of transformation mixture). Save the left-over cells at 4° C. We grow out the cells overnight from the test plates to determine the appropriate amount of transformation to plate. The next day we plate the remaining cells to get the desired density of colonies.

We recommend that plates be at low to moderate densities to make it easier to pick positive colonies. We used 50 µl of cells and 0.85 µl of ß-mercaptoethanol very successfully to save money. The following is from the Strategene instructions. Note that the Stratagene instructions have a typo in step 9. The SOC should be heated to 37° C **not** 42° C.

1. Thaw supercompetent cells on ice.
2. Gently mix cells by hand. Aliquot 100 µl of cells into a **prechilled** 15-ml Falcon 2059 polypropylene tube. (Note: using 50 ul cells will work but may require more transformation mixture to be plated).
3. Add 1.7 µl of ß-mercaptoethanol (supplied with cells) to 100 ul of bacteria to reach a final concentration of 25 mM.
4. Swirl the contents of the tube gently. Incubate the cells on ice for 10 minutes, swirling gently every 2 minutes.
5. Add 01-50 ng ligation mix to the cells and swirl gently.
6. Incubate on ice for 30 minutes.
7. Heat pulse the tubes in a 42 °C waterbath for **45 seconds.** The length of time of the heat pulse is critical for highest transformation efficiency.
8. Incubate the tubes on ice for 2 minutes.
9. Add 0.9 ml of preheated (37° C) SOC medium (**not** LB) and incubate tubes at 37 °C for 1 hour with shaking at 225-250 rpm.
10. Use a sterile spreader to plate transformation mix on LB-agar/ampicillin plates in presence of IPTG/XGAL.
11. Place plates at 37 °C for 16 to 18 hours.

VIII. Selection of colonies containing microsatellite inserts

A. Screening of library plates for positive clones containing microsatellites (see Sambrook et al.).

1. Use nylon hybridization membrane (MSI MagnaLift) circles for colony lifts off library plates. This brand of membrane is compatible with the NEB Phototope detection kit we used. See Sambrook et al. for colony lift details. Use optional SDS soak to reduce background. Do not bake filters (this causes the bacterial debris to adhere to the filter and will cause increased background). Rather air dry filters and then UV crosslink (we used two "autocrosslink"s in a Stratagene crosslinker). Filters can be stored at room temp between pieces of filter paper.

2. Set hybridization oven to 55° C. Colony lift filters are first incubated with proteinase K to remove bacterial debris that causes high background with non-radioactive detection methods. Make up protK buffer: add 2.5 mL 1 M tris pH 7.6, 2.5 mL 0.5 M EDTA; 12.5 mL 10% SDS in a total volume of 250 mL. See Sambrook et al. B.16 for recipe. Prewet filters in dH_2O and then place filters into a hybridization tube. Add 5.0 mL of protK buffer and 25 µl protinase K (10 mg/mL) to each tube.

Incubate filters for 1 hour (longer is fine) in hybridization oven. This step assures very clean results with the NEB detection kit.

3. Pour out the protinase K solution and add 25 mL per hyb tube of hybridization solution (0.25 M Na_2HPO_4, pH 7.2; 7% SDS; 1 mM Na-EDTA, pH 8.0; 1% BSA, fraction 5; from Westneat et al. 1988). Prehybridize (at hybridization temp) for 4 hours to overnight.

7. Combine 2 µl of a 1 µM solution of each biotinylated repeat oligo (the same ones used for the bead enrichment) into a clean microfuge tube. We used all enrichment oligos for hybridization for all filters. Pour out the prehybridization solution and add 10 ml fresh hybridization solution. Pipet the combined oligos into the hybridization tube. Hybridize probe to filter at 45-65˚C overnight. Hyb temp is generally 5 to 10˚C below T_m, but usually no higher than 65˚C in practice.

8. Pour hybridiztion solution into a 15 mL disposable tube. This solution can be reused with additional filters.

Wash filters for 15 minutes each:

a. once in 2X SSC, 0.1% SDS @ room temp
b. once in 2X SSC, 0.1% SDS @ 45˚C
c. twice in 1X SSC, 0.1% SDS @ 65˚C (A/T rich enrichments get 1X wash only)

Heat wash solutions in advance. Washes are best carried out in tupperware containers with covers placed in a heated shaker so that wash temperature is constant. 2X SSC, 0.1% SDS can be used as a less stringent final wash or the last stringent wash can be skipped.

9. Carry out detection steps exactly as specified in the NEB kit. We used a maximum dilution for the detection reagent and the wash conditions specified for colony lifts.

B. Amplification of insert DNA.

Inserts from positive clones are PCR amplified for direct sequencing (recommended) or cells are minipreped for plasmid isolation (not recommended because it is too much work). Miniprep protocols for plasmid isolation can be found in Sambrook et al. 1989 and Applied Biosystems "Bulletin 18".

Cells are disrupted by boiling to release plasmids and the plasmid DNA in the cell-boil solution is used as a template in a PCR reaction with the insert-flanking plasmid primers T7 and T3.

1. Label 600 µl microfuge tubes and fill them with 200 µl T.E.

2. Using a sterile pipet tip, carefully pick cells from the center of a positive colony. Wash off the cells into the T.E by flicking the pipet tip in the solution.

3. Place the tubes with cells and T.E in a 100˚C heat block for 10 minutes.

4. Vortex tubes while hot to further disrupt cells.

5. Spin tubes for 1 minute at max speed to pellet cell debris (optional).

6. Use 1 to 4 µl of this solution as template in a 50 µl PCR reaction (remember to include a blue colony as a control for polylinker amplification).

7. Run PCR products on a 2.0% gel to check amplification.

8. Clean up PCR reactions with QiaQuick PCR purification preps and sequence using T7 and/or T3 as primers.

IX. Sequencing of positive clones and primer design.

A. We typically sequence plasmids or insert amplification products using a standard double-stranded protocol. We use automated (*Taq* Cycle Sequencing, Applied Biosystems Incorporated) methods. We refer the reader to the vendor's detailed protocols.

B. Primers are designed from sequences flanking the microsatellite arrays identified within the inserts. Occasionally a microsatellite array begins on the edge of the plasmid, or is incomplete or too large for continued development. In addition, some clones contain the same insert. In general only a fraction (>30%?) will yield useful microsatellite primers.

C. Design rules are available from a number of sources (e.g. Hoelzel and Green 1992). Free software such as Amplify for the Mac is useful to design primers (http://www.wisc.edu/genetics/CATG/amplify/#download).

D. Primers must be tested and optimized in order to guarantee faithful and consistent amplification. In addition, microsatellites should be assessed with pedigreed families to ensure Mendelian inheritance and independent assortment.

Large repeats that lack sufficient flanking sequence to design primers:
Sequencing often reveals some repeats that lack enough flanking region to design a primer. In this case it is possible to extend the known sequence and get enough flank for a reliable primer. See for example Ochman et al., 1990. This should be unnecessary. If enrichment is successful you will have many clones to choose from.

Not So Standard Solutions:
T.E: 10mM Tris-HCl, pH 8.0; 0.1mM EDTA, pH 8.0 (reduced conc. of EDTA will not inhibit PCR reactions)

Appendix 1

The molecular weight of double stranded DNA,

$$MW_{ds} = \#\text{base pairs})(650 \text{ daltons/bp})$$

The moles of **ends** of double stranded DNA = 2(grams DNA)/(MW in daltons)

For DNA that is of variable length such as a genomic digest you must estimate the average base pair length of the entire sample. For example, a digest that is size selected for 200 to 800 dp fragments has an average length of about 500 bp.

a 500 bp fragment is 500(650) = 325,000 daltons

1 μg of this genomic DNA has

$2(1 \times 10^{-6} \text{ grams})/325{,}000 \text{ daltons} = 6.154 \times 10^{-12}$ moles of ends

For DNA of fixed length such as oligonucleotides calculate the MW in daltons:

a 20 bp oligo is 20 bp(650 daltons/bp) = 13,000 daltons

A 1:1 molar end ratio for genomic DNA ends to oligo DNA ends will require 6.154×10^{-12} moles of ends of oligo in solution. So for a 20 bp oligo

Y grams oligo = $(6.154 \times 10^{-12}$ moles of ends)(13,000 daltons)/2 =

4.0001×10^{-8} grams or 40 ng

In general

$$Y \text{ grams} = (\text{moles of ends})(\text{ratio factor})(\text{MW in daltons})/2$$

For a cohesive termini ligation the ratio factor is 1 or 2. In the case of a blunt end ligation the ratio factor for the oligo would be between 40 and 100.

NEB recommends a DNA concentration of 0.1 to 1 uM of 5′ termini. USB says that "for each 10 ul of reaction volume, a combination of 20-40 ng vector and a three to ten-fold excess of foreign DNA will produce an adequate yield of recombinant genomes for most cloning purposes."

Literature cited

Armour JAL, Neumann R, Gobert S, and Jeffreys AJ. 1994. Isolation of human simple repeat loci by hybridization selection. Human Molecular Genetics 3: 599-605.

Fleischer RC, and Loew S. 1995. Construction and screening of microsatellite-enriched genomic libraries. Pages 459-468 in J. Ferraris and S. Palumbi (eds.), Molecular Zoology: Advances, Strategies and Protocols. Wiley-Liss, New York.

Gardner, MG, Cooper, SJB, Bull, CM, and Grant, WN. 1999. Isolation of microsatellite loci from a social lizard, Egernia stokesii, using a modified enrichment procedure. Journal of Heredity 90:301-304.

Hamilton MB, and Fleischer, R.C. 1999. Cloned microsatellite repeats differ between 4-base restriction endonucleases. Journal of Heredity 90:561-563.

Kandpal RP, Kandpal G, Weissman SM. 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. Proc. Natl. Acad. Sci. USA 91:88-92.

Koblizkova A, Dolezel J, and Macas J. 1998. Subtraction with 3' modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. BioTechniques 25:32-38

Ochman H, Medhora MM, Garza D, and Hartl DL. 1990. Amplification of flanking sequences by inverse PCR. Pp. 219-227 in M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White (eds.). PCR Protocols. A Guide to Methods and Applications. Academic Press, San Diego.

Sambrook J, Fritsch EF, and Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Westneat DF, Noon WA, Reeve HK, and Aquadro CF. 1988. Improved hybridization conditions for DNA 'fingerprints' probed with M13. Nucleic Acids Res 16:4161.

Zane, L, L Bargelloni, and T Patarnello. 2002. Strategies for microsatellite isolation: a review. Molecular Ecology 11:1-16.