TONIA'S FLOWCHART (THE MOLECULAR END) FOR A PARENTAGE PROJECT USING MICROSATELLITES.

Information on microsatellites: (Selkoe and Toonen 2006; Zhang and Hewitt 2003; Chambers and Macavoy 2000; Goldstein and Schotterer 1999)

Protocols are available for each of these steps.

DNA EXTRACTION (1 – 3 MONTHS DEPENDING ON PROTOCOL AND NUMBER OF SAMPLES)

1) **Find a Protocol** that will likely work for your samples. The best way to do this is to contact someone who has extracted DNA from similar tissue/species. *Megan – I have a salt-based DNA extraction protocol that we will use.*
2) **Test the Protocol** on a few samples. Look at those samples on a 0.8% agarose gel to make sure the DNA looks good. ALSO try to amplify your DNA using some universal type primers. Nothing is worse than extracting DNA from 1000 samples that look beautiful on a gel but won't amplify because the extraction protocol failed to remove some type of inhibitor.
3) **Once you have a good protocol, extract All Your Samples at the same time – but you may want to use only part of your sample.** It will make things easier in the long run if you try to keep all the samples at the same concentration (i.e. start with the same amount of tissue, resuspend the DNA in the same volume, etc). It is a good idea to only use part of your sample in case something goes wrong you have a back up sample for each individual.
4) **Make the DNA Dilutions.** Generally (depending on the DNA extraction protocol) you end up with heaps more DNA than you need, thus you will need to make a dilution of your DNA in to another tube before using it in a PCR reaction. Even if your DNA is at the right concentration you always want to make a working aliquot of your DNA so you don't accidentally contaminate your only tube of DNA for that individual. I find it easiest in the long run to make these dilutions (or aliquots) in 96-well plates. You want your working stock of DNA at a concentration between 5 ng/ul and 30 ng/ul.

FIND MICROSATELLITE LOCI (1-9 MONTHS – DEPENDING ON THE METHOD AND THE SPECIES)

You can either do this yourself or have it outsourced commercially. Methods for doing this your self could include microsatellite enrichment, making a cloned library, and/or new methods using massively paralleled sequencing (454 GS-FLX) (e.g. Abdelkrim et al 2009 Biotechniques 46(3):185; Meglecz *et al.*, 2010. *Bioinformatics* 26**:** 403-404).
Alternatively, you can attempt to use microsatellite markers developed in a closely related species.

OPTIMIZING THE PRIMERS

(2-9 MONTHS DEPENDING ON THE SPECIES/POPULATION AND THE NUMBER OF LOCI NEEDED)

Identify the Microsatellite Markers to be Used. Microsatellite markers are found through development of a microsatellite library for the species of interest.

1) **Test the Markers**. Two things you want to test the markers for: a) that they amplify in a large sample of your animals, and b) that they are variable (i.e have more than one allele). If the microsatellite development was outsourced, this will be done for you (usually in 10 animals).
2) **Design PCR and Genotyping Multiplexes.** Every PCR reaction costs around ~$0.40, and each genotyping run on the capillary sequencer cost ~$1.5. Also, each of these steps takes (nearly) the same amount of time whether you are amplifying/genotyping 1 locus or 10 multiplexed loci. So the more loci you can get to amplify consistently in a PCR reaction, and that can be run in the same Genotyping reaction the better! This is also the step where you decide which colors to label you primers. There are 4 colors to choose from, the main point is that you don't want overlapping loci to have the same colors.

One option is to use a fluorescently labeled M13 primer attached to your primer, see these references (Heyden *et al.* 2008; Boutin-Ganache *et al.* 2001).

3) **Test the Multiplexes and Redesign as Needed.** This step can either be quick and easy, or long and arduous, but almost never do you end up with the same multiplex design as in the beginning.

COLLECTION OF DATA (1-6 MONTHS DEPENDING ON THE NUMBER OF SAMPLES)

1) **Organize Your Samples!** By this point you should have an excel sheet of all your samples, each with a unique identifier (name: less then 10 characters) that you will use for the length of your project. You will be working in a 96-well format (i.e. a 96 well plate that has 12 columns (1-12) and 8 rows (A-H)), so you will need to have your samples are in this format as well so you know exactly were you can find each sample. Note: I have an excel worksheets that can transfer back and forth between columns and 96-well format. Also, I always put a positive control (DNA from a reliable sample) in well A1 and negative controls (just water) in wells A2 and H12. This not only gives your plate directionality but also checks for contamination and PCR reaction screw ups.
2) **Genotyping.**
    a. **Multiplex PCRing.** PCR reactions will be done in 96-well plates using 7 or 10 ul reactions, thus you will do 1 positive control, 2 negative controls, and 93 of your test samples at one time. This process is the same whether you put primers from one locus into your PCR master mix or primers for 10 loci into your PCR master mix.
    b. **Electrophoresis on ABI3130xl.** Mix PCR samples with formamide and size standard and run on sequencer.
3) **Set up Panels in Genemapper for Scoring.** Using the results files from the genotyping on the sequencer to set up panels with bins for your allele at each locus. This is a way to standardize and to some extent automate the scoring of your samples.
4) **Check Mom-Baby genotypes.** Make sure your alleles are being inherited in a Mendelian manner and that your method of allele calling is consistent from mom to baby.
5) **Scoring Your Genotypes.** Use the panels you set up in GeneMapper to score all your data. After the automated scoring, you need to go through and check all the scoring by eye – you CAN NOT simply trust the GeneMapper program.
6) **Export you Genotypes into Excel** and identify the individuals that didn't work satisfactory.
7) **Re-genotype the Individuals that didn't Work the First Time** (steps 1 or 2 through 5).
8) **Have Someone Double-score Your Data.**
9) **Dataset.** Get all your data into one final file ready for analyses. Double check any alleles that don't quite look right or are very rare. Finalize your dataset.

PARENTAGE ANALYSES (1-3 MONTHS)

Check your data for Hardy-Weinberg and Linkage-Disequilibrium. This can be done using a variety of pop gen programs: ARLEQUIN, GENEPOP, GENALEX, CERVUS, just to name a few. Particularly you want to check for null alleles as this can very negatively affect the reliability of your parentage analyses.

1) **Adjust your Dataset According to the Results from Above.** You may need to eliminate some loci or delete your homozygote genotypes to compensate for null alleles.
2) **Begin Parentage Analyses.** My program of choice (at this point in time) is CERVUS (Kalinowski *et al.* 2007; Slate *et al.* 2000; Marshall *et al.* 1998) but there are a lot of other great programs out there that may be better for your project, such as ML-RELATE, PAPA, NEWPAT, PARENTE, FAMOS, GERUND, COLONEY….
3) **Check your assignments.** Of course, regardless of what program you use to assist in the assignment of paternity, you always need to compare the data yourself. Make sure the father assigned makes sense with the genotyping data, and the field data or experiments.

REFERENCES

Heyden, Matthew J, Thao M Nguyen, Amanda Waterman and Kenneth J Chalmer. 2008. Multiplex-Ready PCR: A new method for multiplexed SSR and SNP genotyping ***Bmc Genomics*** 9**:** 1-12.

Kalinowski, Steven T., Mark L. Taper and Tristan C. Marshall. 2007. Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment doi:10.1111/j.1365-294X.2007.03089.x. ***Molecular Ecology*** 16**:** 1099-1106.

Selkoe, Kimberly A., and Robert J. Toonen. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. ***Ecology Letters*** 9**:** 615-629.

Zhang, De-Xing, and Godfrey M. Hewitt. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects doi:10.1046/j.1365-294X.2003.01773.x. ***Molecular Ecology*** 12**:** 563-584.

Boutin-Ganache, Isabelle, Marco Raposo, Martine Raymond and Christian F Deschepper. 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. ***BioTechniques*** 31**:**

Slate, Jon, Tristan Marshall and Josephine Pemberton. 2000. A retrospective assessment of the accuracy of the paternity inference program cervus doi:10.1046/j.1365-294x.2000.00930.x. ***Molecular Ecology*** 9**:** 801-808.

Chambers, Geoffrey K, and Elizabeth S Macavoy. 2000. Microsatellites: consensus and controversy. ***Comparative Biochemistry and Physiology Part B*** 126**:** 455-476.

Goldstein, David B., and Christian Schotterer, 1999 *Microsatellites: evolution and applciations*. Oxford University Press.

Marshall, T. C., J. Slate, E. B. Kruuk and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. ***Molecular Ecology*** 7**:** 639-655.