Laboratory Exercise: Fetching SNPs: A Dog Genotyping Laboratory for Undergraduate Biology

Keith Hultman*[1,2] and Eve Mellgren[1]

[1]Elmhurst College
190 S Prospect Ave
Elmhurst, Illinois 60126

[2]Washington University in St. Louis
1 Brookings Drive
St. Louis, MO 63130

**Corresponding Author: keithh@gmail.com*

Accepted for publication September 5, 2014

Citation:

Hultman, Keith and Mellgren, Eve (2014). Fetching SNPs: A Dog Genotyping Laboratory for Undergraduate Biology**.** *Genetics Society of America Peer-Reviewed Education Portal (GSA PREP)*: 2014.001; doi:10.1534/gsaprep.2014.001

Synopsis:

This multi-week lab project is designed to increase comprehension of the core concept of genetic variation, and give undergraduate students experience with applying their knowledge of genetic variation using derived Cleaved Amplified Sequence (dCAPS) genotyping on DNA isolated from dogs. Students collect buccal samples of canine DNA, extract genomic DNA, and use PCR followed by restriction digests to determine which SNP alleles are present in each dog. They then analyze the genotype information to answer proposed forensic scenarios and test student-generated hypotheses involving the association of SNP markers with several dog traits. The lab can be used in a sophomore level genetics class where students are just being introduced to the concepts of genetic variation and genotype detection and have very little experience using molecular techniques. It can also be adapted for upper-level laboratories with the inclusion of primer design of additional dCAPS markers.

Introduction:

Undergraduate genetics laboratory courses often introduce students to commonly used model organisms and highlight their importance in advancing our knowledge of genetics. Studies of dogs are a natural fit for initially capturing the attention of undergraduate students while introducing them to the concept of a model organism. With the availability of the canine genome sequence in 2005 (*1*), extensive genotypic and phenotypic variation, and inherited diseases that are shared with humans, dogs have become an important model for human genetic disorders (*2*). This lab makes use of known variation in the dog genome in the form of 7 Dog SNP markers (referred to as DS1-DS7) that have been associated with phenotypic traits (*3*). As with other model systems, students learn techniques that are very similar to isolating and manipulating DNA from humans but avoid the concerns of medical privacy. Additionally, this project does not require animal protocols for the in-class laboratory because students can obtain non-invasive cheek swab samples from their dogs at home. Students are invested in the lab because they have the opportunity to teach and interact with members of their community by involving their dogs in the project.

The genotyping is performed using a PCR-based method called derived Cleaved Amplified Polymorphic Sequences (dCAPS; *4*, *5*). The method involves amplifying genomic sequence with mutagenic primers in order to introduce restriction enzyme recognition sites on one allele of a SNP locus. Like standard CAPS, also called PCR-RFLP (Restriction Fragment Length Polymorphisms), it is an inexpensive and simple method for genotyping that can be performed by undergraduate labs (*6*). We cover these methods in lecture before beginning the lab module. It is useful to compare dCAPs, CAPS, and standard Southern blot RFLP in lecture prior to the lab so students understand there are several available methods for detecting SNPs. We have included the PowerPoint slides we use for covering dCAPS as part of a lecture on advanced PCR methods. The primary advantage of dCAPS is that any SNP can be interrogated using dCAPS even if it does not produce a natural restriction site.

Genotyping dogs for these SNP markers can be used to explore several questions relating to the relationship between genotype and phenotype. We have included four case examples that our labs have used with these markers. The first two cases simulate forensic analysis of DNA, paternity testing and crime scene identification. These close-ended modules have a known answer and are easy for the instructor to evaluate the students understanding of the methods and the idea of a genotype. Furthermore these cases allow students to relate the genetics concepts learned in the classroom to real world examples, an idea promoted by the NSF-AAAS Vision and Change report (*7*). The final two cases, heterozygosity of dog breeds and the association of genotype to familiar traits, are examples of questions that are more open-ended in nature and promote student-generated hypotheses. In both of these cases students may find unexpected results that challenge their ability to evaluate their hypotheses. We suggest incorporating one of the forensic modules for the class and allowing each group or individual student to choose an open-ended question for them to develop a hypothesis.

Many undergraduate biology educators have begun to move away from lab exercises where there are known answers ("cookbook" type labs) towards lab exercises that more closely resemble biological research, incorporating opportunities for students to design the experiment. The benefit of offering students experiences where they are active participants is highlighted in the Vision and Change report (*7*) as well as the AAMC-HHMI generated Scientific Foundations for Future Physicians report (*8*). Unlike many "cookbook" type labs, in this lab students do not know what genotypes will be present in their dogs. This laboratory exercise allows students to apply the knowledge they have acquired in the classroom on genetic variation, detection of variation, and various molecular techniques. By analyzing their results and evaluating their own hypotheses, students take ownership in the project. Finally, students obtain important communication skills by writing a lab report to one of the volunteer dog owners who donated a sample for the class.

Justification:

This dog genotyping lab allows students to gain a deeper understanding of the core concept of genetic variation within a population. Students will learn how to identify the genotype of several markers in diploid organisms and will be able to see variation between different individuals that they sample. Specifically, they will become familiar with a common form of variation, Single Nucleotide Polymorphisms (SNPs), and one method that is used to detect this class of genetic variation (dCAPS). Students will also gain an understanding of the relationship between an organism's genotype and how that genotype is expressed through familiar phenotypic traits. Additionally, the lab will reinforce the concepts of the nature of genetic material, and could include the core concept of how genes are transmitted through inheritance if the paternity testing case study is included. The lab could also include a discussion of how dogs are a particularly useful genetic model organism for studying genetic diseases in humans.

The core competencies covered in this lab include the scientific method (observational strategies, hypothesis testing, and experimental design), evaluation of experimental evidence, and explaining concepts to both expert and non-expert audiences. Since students will be

involved in choosing which dogs to sample based on the questions they pose, and potentially primer design for additional dCAPS loci, they will be involved in shaping the experimental design. They formulate hypotheses involving the relationship between genotype and phenotype and make predictions of which genotypes are expected based on breed and phenotypic data provided by the dog owners. Initial analysis of the results involves interpreting if the dog(s) are homozygous or heterozygous and which alleles are present. Students will then be able to use an included table to see if the alleles they found are what they predicted, and evaluate the likelihood that a particular dog will have that allele. Their lab notebook will include their results and conclusions written for a scientific audience. With the addition of requiring a written report to be given to the dog owner, students will also have an opportunity to explain the questions they have posed and their conclusions to a non-expert audience.

Approach/Method: (Instructor Guidelines)

There are many questions related to forensics and medicine that can be addressed in this lab, depending upon the scope of the course and the interest of the students. The best types of questions are those that can easily be answered using fewer than 10 markers. Here are a few case examples that we have used in our dog genotyping labs.

Who's the Puppy Daddy: A dog paternity test case. This case requires obtaining DNA from a puppy along with samples from the parental dogs and several unrelated dogs. It is best, but not absolutely necessary, if DNA from the Mother is collected so that students can determine the exact haplotype the father contributed. After collecting the DNA samples, the instructor informs the students which samples are from the mother and the puppy, but does not disclose which of the other samples is the father's DNA. By looking at which alleles were contributed from the mother, students determine what alleles must have been contributed from the father. Students can then rule out potential sires if they do not share the paternal alleles. This case can also be used if only DNA from the puppy is available, as students should be able to rule out potential fathers if they have at least one marker that does not share an allele with the puppy. Our class was able to identify the correct father from 10 dogs using the 7 DS markers we have described here, without maternal DNA.

DogPile ID: A who-dropped-it forensic case study. This study is based on an actual service that uses DNA testing to identify the dog owner responsible for leaving dog feces in neighbor's lawns. Although it might be possible to obtain DNA from dog feces in the lab, we recommend the instructor simulate the case by randomly sampling from one of the dog DNA samples collected by the students as the 'evidence'. The instructor records which sample is the perpetrator, but does not disclose this information to the class. The students are then tasked with determining which dog samples are a match. This case study recreates many human forensic cases involving the identification from blood, semen, or other crime scene biological samples. Analysis could also involve calculating the probability of finding a match based on the number of markers tested and the population size sampled. Assuming that genotypic frequencies are equally distributed in the population, with three genotypes possible for each

SNP locus (two homozygous and one heterozygous), there is a one in three chance of carrying any particular genotype at each locus. For n markers, there are 3^n genotypic combinations possible. For the 7 markers we describe, there is a 1 in 2187 chance of finding an exact match by chance. Note that it is not likely that each genotype is equally distributed, and actual allelic frequencies could be used to calculate a more accurate estimate if they were available.

Homozygosity of Purebred Dogs: There are several other questions that can be asked that are open ended. One question we have asked students is whether purebred dogs should be more homozygous at these markers than mixed breed dogs. Since purebred dogs are more likely to be inbred or bred from a smaller gene pool, students hypothesize that the purebred samples will have several SNP markers that are homozygous. In contrast, mutts or mixed breed dogs should show fewer homozygous genotypes. Surprisingly this is not always the case. It is likely that some purebred populations are large enough to retain variation at these loci, and the answer your class finds may depend on which breeds are chosen for sampling.

Association with Traits: Many human genetic studies report finding a gene linked to a disease. Often times this is a QTL locus that is associated with a complex disease like obesity or autism. One of the major obstacles in medicine is how to use these associated markers for diagnosis and treatment, when each locus only contributes a small amount to the phenotype and the identified SNP may not even be involved in the mechanism of the disease. Students can explore the relationship between SNPs that have been associated with particular dog traits and ask whether the dogs they have sampled have the expected genotypes given the appearance reported by the volunteer questionnaire. The associated traits for all dog SNPs are available from the Supplemental Table 4 from (*3*). We ask our students if the phenotypes and genotypes are consistent with the reported results and if not, what may be the reason.

Introduction of Lab:

We designed this lab to use derived cleaved amplified polymorphic sequences, or dCAPS because it is a method of genotyping that incorporates both PCR and restriction digestion, two core molecular techniques covered in molecular biology. It also allows for interrogation of any SNP in the genome even if it is not a natural endonuclease recognition site. It is important that you cover the method of dCAPS genotyping before the lab begins.

Thus far the lab has been used in more upper level molecular courses where the dCAPS method was described in detail in lecture/lab and students were involved in designing dCAPS primers. We have also reused primers that were successful in amplification and had showed variation among tested dogs.

Suggestions and specific instructions for each lab protocol follows:

Prelab - Collection of Canine buccal swabs:

Before the lab exercise, we asked our students if they would like to collect samples from their family dogs. We have also obtained samples from members of the department who volunteered to take samples. We have had no trouble in finding volunteers and have had as many as 10 dogs including a wolf-dog hybrid for a single class. In order to make sample collection simple for the volunteers and to standardize the procedure, we put together small sample kits packaged in zip-top bags to give to volunteers.

Materials included in the sample prep kit:
Instructions on sample collection (see supplemental materials)
A pair of large non-latex gloves
Two Sterile cytology brushes wrapped in foil
Two Sample tubes
Dog biscuit

Although the instructions for the buccal swab include a 2 hour drying time in Step 4, we have obtained quality DNA for PCR with as little as 5 minutes of drying time, and the sample collection can be done during a lab period if desired. If the sample is to be stored for several days, for example over a weekend or spring break, then the 2 hour drying time should be followed.

After collecting the samples from volunteers, determine what types of questions the class will try to answer and organize the samples accordingly. The samples should be assigned a letter code by the instructor so that the students are blind to the identity of the samples until after they have genotyped the samples. Using letters for each dog sample will avoid confusing dog samples with the DNA markers which are numbered DS# in the student manual.

Optional upper level computer lab or homework: Designing dCAPs primers:

For an advanced class, including the primer design for dCAPS analysis will help to reinforce the concept of how almost any SNP can be interrogated with this technique, regardless of whether the SNP is part of a standard restriction enzyme recognition site in the sequence. This exercise can expand the number of markers to include more SNPs from the dog genome, or even to modify this lab to genotype a different organism altogether. We have included student instructions on designing dCAPS primers in the Supplemental materials section. A full lab with access to computers is recommended, as many students will need to be guided in working with DNA sequence. Students will need to design a dCAPS primer using the dCAPS Finder 2.0 website (*9*) and a regular primer using Primer3 (*10*). We also have the students make a positive control dCAPS primer, which is a mutational primer that creates a restriction site regardless of which allele is amplified. This exercise should be completed with enough time before Lab 2 in order to check their work, order, and receive the primers.

We have included primer sequences that have successfully amplified and identified 7 SNPs from the Jones 2008 paper. These SNPs with our nomenclature and the locus ID used in Jones 2008: DS1, gnl|ti|360206886_2; DS2, gnl|ti|354710886_1; DS3, gnl|ti|355951851_2; DS4, BICF229J36361; DS5, BICF232J28587; DS6, BICF232J16557; DS7, seq_6.

Lab 1: Genomic DNA Purification from Buccal Swab using the QIAamp DNA Mini Kit.

We have had much greater success in obtaining quality genomic DNA using a cytology brush than a cotton swab applicator. Although any method for purifying Genomic DNA for PCR should work, we have included the instructions for the QIAamp Mini DNA Kit as it has worked successfully for us several times. The instructions for students has been modified from the buccal swab spin protocol from the Qiagen manual (*11*).

Assign your dog buccal swab samples among the student groups. To avoid sample loss, split up each of the two sample tubes from each dog and assign them to two different student groups.

If possible, have your students check the absorbance of the genomic samples using a nanodrop or spectrophotometer.

Instruct the class to dilute their samples to 10 ng/⌠l prior to Lab 2.

Lab 2: PCR of SNP markers

We have included the PCR protocol for Qiagen's Taq PCR Master Mix Kit. It works well for both inexperienced and upper level undergraduates. However, you may prefer to have students assemble their own reactions from components in order to emphasize the ingredients necessary for PCR.

It is only necessary to run one Dog sample with the positive control primer set. Choose a sample with high yield and good absorbance ratios as the positive control template.

OPTIONAL STEP - Confirm PCR amplification with gel electrophoresis

This step can be done by the instructor in between lab sections, or by the class, depending upon schedule restrictions. We have found that it is best for the instructor, or a teaching assistant, to run a 1% agarose gel with samples from each group in between lab sections. Images of the gels can be given to students at the beginning of the next lab section.

Lab 3: Restriction digest of dCAPS PCR products

Have your students set up digests so that twice as much DNA is digested than you will need for a single resolving gel. This assures they will have enough digested DNA if a mistake is made and they need to run another gel. It may be possible to shorten the restriction digest incubation

time if it helps with time management of the lab. However, it is critical to allow for complete digestion in order to avoid false positives for uncut alleles.

Figure 1

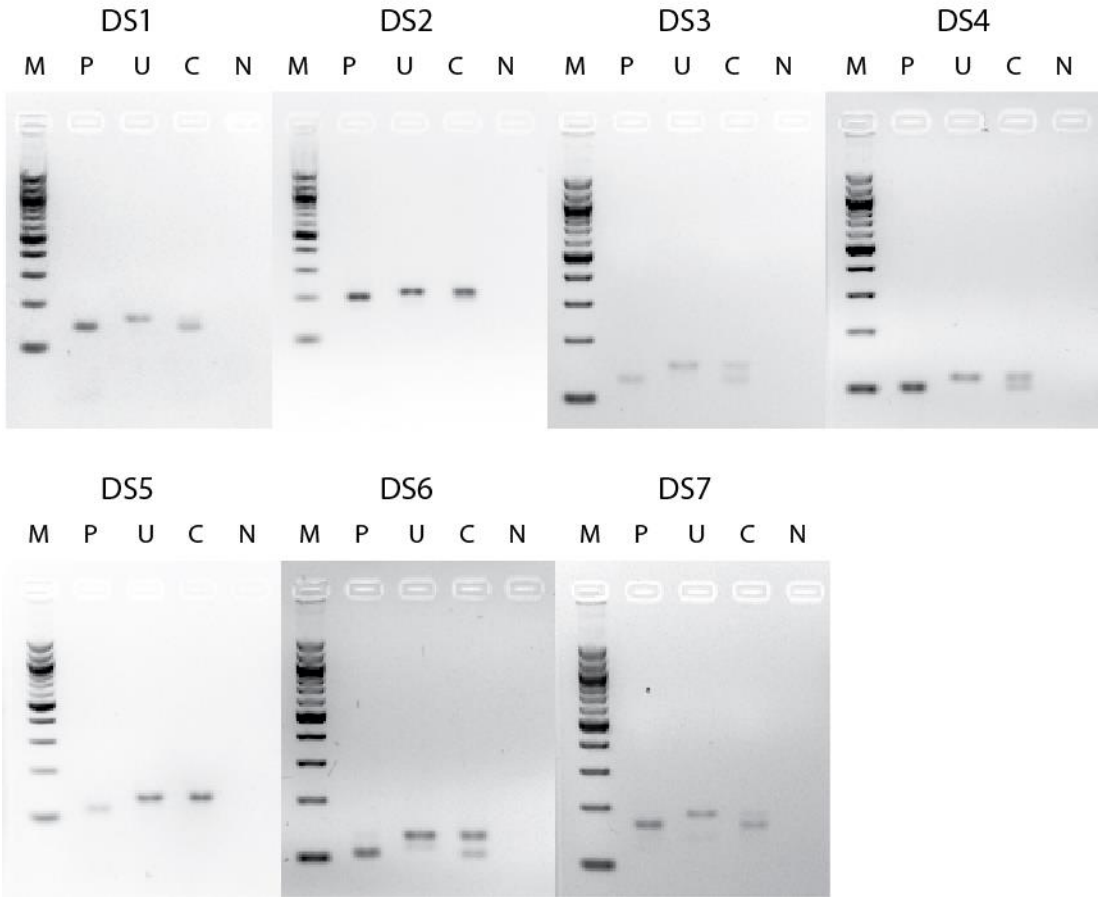

Figure 1: An example of genotyping results for each of the 7 SNP markers discussed in the text for one dog. By comparing the size(s) of products in the digested genotyping lane (C) to the positive control lane (P) and the uncut genotyping lane (U), students can determine the genotype for each marker. This sample was taken from a 70 pound mixed-breed dog whose traits matched Black Labrador (size and shape), Pitt Bull (facial structure), and Chow Chow (black tongue, neck folds, tail curve). M - 100 base pair marker, P- positive control cut, U - uncut genotyping reaction, C - cut genotyping reaction, N - no template negative control.

An example of student results for Table 3 using the above data.

SNP	Positive control		Uncut		DogA	
	Exp	Obs	Exp	Obs	Bands present	Geno -type
DS1	143	✔	163	✔	142/162	C/T
DS2	200	✔	220	✔	220	G/G
DS3	129	✔	152	✔	129/152	C/T
DS4	105	✔	119	✔	105/119	C/T
DS5	114	✔	134	✔	150	A/A
DS6	110	✔	136	136, and faint 130	110/136	A/G
DS7	166	✔	186	✔	166/186	A/G

Example Analysis for Associated Traits Case:

Students can first compare the homozygosity of their samples and test whether mixed breed dogs are more likely to be heterozygous than pure bred dogs. Our example using a mixed breed dog shows homozygosity in just 2 out of 7 SNP markers.

Students can then examine each locus and its associated trait. It is best to concentrate this analysis on homozygous markers. Supplemental Table 4 of the Jones paper lists all associated traits with our DS markers and many others*. According to the Jones paper, the C allele of DS2 is listed as being negatively associated with a short coat. Our sample's short coat and his G/G genotype are therefore consistent with the Jones findings. The second homozygous genotype, A/A at DS5, is positively associated with height at the withers and weight. This is also consistent with our sample's weight of 70 pounds and height of 24 inches. We ask students to look up small and large breeds to get an estimate for size and weight cutoffs. This exercise is often quite difficult for assessing medium sized dogs, and often the traits a dog has may not align with the noted associations. This exercise is useful in highlighting the difficulty of using genetic markers identified in genome wide associations for diagnostic purposes. This is currently a non-trivial problem facing personalized medicine and medical genomics. Students can also analyze the heterozygous markers by hypothesizing which allele is dominant given their sample's phenotype. In our example, the T allele of DS1 might be dominant since the C allele is associated with lower height and weight.

**A note on trait evaluation: we have noticed that several alleles in Supplemental Table 4 of the Jones paper do not match up with the stated SNP nucleotides. A SNP of C/T may*

list the allele associated with height as A(-) for example. Email correspondence with the paper's authors revealed that the complementary strand may have been used during association analysis and in this example the listed allele should be T(-).

Lab Safety:

The lab will require instruction on how to safely work with ethidium bromide and UV light. Students should wear appropriate protective eyewear and lab coat when handling gels using ethidium bromide. We currently use the less toxic Gel Red DNA stain in place of ethidium bromide, however this can affect migration patterns of some ladders so the amount of DNA added may have to be adjusted (*12*). During all other protocols students should wear nitrile gloves to prevent possible contamination of the samples with human DNA.

Student Evaluation:

There are several options for a final written submission from the students. You may want to have them write a formal lab report in scientific paper format. As a way of incorporating the core competency of explaining concepts to different audiences, we have asked students to write an informal report for one of the volunteer dog owners. This exercise involves a typed letter that can include a table or image of their results. We evaluate their ability to explain the methods and their results to a non-expert audience. In addition to this report, students are asked to write their conclusions in proper scientific writing in their lab notebook, which are graded after each lab project. In this way, students communicate their ideas and conclusions with two different audiences.

Acknowledgements:

We would like to thank R. David Hawkins for initially suggesting the use of dogs for our molecular biology courses. Undergraduate students, Steven Van Dorn and Cheryl Chi, provided assistance with trouble-shooting and analysis. Sherpa Hultman contributed DNA for this article and several class experiments. Good boy.

References:

1. K. Lindblad-Toh *et al.*, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).

2. E. K. Karlsson, K. Lindblad-Toh, Leader of the pack: gene mapping in dogs and other model organisms. *Nat Rev Genet* **9**, 713–725 (2008).

3. P. Jones *et al.*, Single-nucleotide-polymorphism-based association mapping of dog stereotypes. *Genetics* **179**, 1033–1044 (2008).

4. M. M. Neff, J. D. Neff, J. Chory, A. E. Pepper, dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. *Plant J* **14**, 387–392 (1998).

5. *Derived Cleaved Amplified Polymorphic Sequences (dCAPS)* (NCBI; http://www.ncbi.nlm.nih.gov/projects/genome/probe/doc/TechDCAPS.shtml).

6. L. H. Hartwell, *Genetics: From Genes to Genomes* (McGraw-Hill Companies, 2004).

7. Vision and Change in Undergraduate Biology Education (2011) (available at http://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf).

8. Scientific Foundations for Future Physicians *aamc.org* (2009) (available at https://www.aamc.org/download/271072/data/scientificfoundationsforfuturephysicians.pdf).

9. *dCAPs Finder 2.0* (; http://helix.wustl.edu/dcaps/dcaps.html).

10. *Primer3* (; http://bioinfo.ut.ee/primer3-0.4.0/primer3/).

11. *QIAamp DNA Mini and Blood Mini Handbook* (Qiagen, ed. 3, 2012), pp. 36–38.

12. Can I use GelRed with the DNA Ladders from NEB? | New England Biolabs*neb.com* (available at https://www.neb.com/faqs/2013/10/22/can-i-use-gelred-with-the-dna-ladders-from-neb).