**LAB 4: SNP Genotyping**

Single nucleotide polymorphisms, or SNPs (pronounced “snips”), are the most common type of genetic variation among all organisms including humans. Each SNP represents a difference in a single nucleotide – e.g. the replacement of a cytosine (C) with a thymine (T) in a certain stretch of DNA. SNPs are found throughout the germline DNA. In humans, they occur once every 300 nucleotides on average, and there are roughly 10 million SNPs in the human genome. Most commonly, these variations are found in the non-coding DNA or in ‘introns’ within structural genes – although they can be found in coding sequences as well. A SNP is distinguished form a ‘mutation’ by its frequency in the population; by convention, a frequency of >1% is considered a SNP while <1% is considered a mutation.

While most SNPs are silent and have no effect on phenotype, many *do* affect phenotype – and it is largely the differences in our SNPs that distinguish us as individuals in a population. SNPs that are close to each other in the genome tend to be inherited together in blocks called “haplotypes,” and SNPs and haplotypes are widely used to follow inheritance in genetic studies.

Some SNPs are closely associated with genetic diseases (e.g. sickle cell anemia) or other health-related phenotypes such as response to certain drugs, susceptibility to environmental factors such as toxins, and risk of developing particular diseases. These days, SNPs are mostly analyzed by DNA sequencing, but there are other methods as well, and in this lab we will be using a technque called ‘Restriction Fragment Length Polymorphism’, or RFLP analysis. We will be analyzing a particular SNP called rs4988235 (yes, there are databases for SNPs too – the ‘rs’ designation refers to the Ref-SNP number that is maintained for every known SNP in the human genome in a database cleverly called, dbSNP). This particular SNP is also known as *LCT* -13910C/T – which refers to the position of the SNP relative to the structural gene that it regulates, and the two possible nucleotides that can be found at this position. *LCT* -13910C/T is associated with the phenotype of lactose tolerance or intolerance, and in this lab, each student will be genotyping their own DNA at this locus.

**Lactose tolerance/intolerance**

Lactose, sometimes called ‘milk sugar,’ is a disaccharide consisting of glucose and galactose. In order to be used as fuel for the body, it must first be broken down into its monosaccharide components by the enzyme lactase, encoded by a gene called *LCT*, which is expressed in the brush border cells of the small intestine. In virtually all mammals, lactase expression in these cells is limited to infancy, and transcription of the *LCT* gene declines after weaning. Hence, most mammals are lactose intolerant as adults – unable to break down lactose. This is also the ancestral condition for humans, and drinking milk for many adult humans results in indigestion, nausea, bloating, gas and general intestinal distress.

However, some eight to thirteen thousand years ago, coincident with the rise of animal husbandry and the domestication of milk-producing animals such as cows, goats, camels, etc., a mutation developed in certain human populations that resulted in persistence of lactose expression throughout life. This mutation of course allowed its bearers to continue to drink the nourishing milk of their domesticated animals into adulthood, and was strongly selected for. The frequency of the mutation thus increased in these populations over time – now far exceeding the 1% threshold required to call it a SNP.

In fact, such mutations developed and spread independently in at least three separate populations of humans – one in the Middle East, one in West Africa, and one – the one we are testing for – in the Balkans region of North Central Europe. Each of these mutations had the same phenotypic effect – persistence of lactase expression into adulthood, but they are all different SNPs found in slightly different spots in the genome – a remarkable example of convergent evolution. All of them occur upstream of the *LCT* gene in a non-coding region that contains a regulator element (an “enhancer’) that regulates lactase gene expression. In fact, these SNPS are all located in the introns of an entirely unrelated gene adjacent to *LCT* called *MCM6* on chromosome 2. For the *LCT* -13910C/T SNP, the presence of a C at this position is associated with lactose intolerance, while the presence of a T is associated with lactase persistence. The lactase persistence gene is dominant and heterozygotes are lactose tolerant.

**Lab Overview**

In this lab you will first isolate genomic DNA from your cheek using swabs and a kit provided. This DNA will then be amplified by PCR using a particular set of primers with the sequence:

Forward: 5’-GTTGAATGCTCATACGACCATG

Reverse: 5’-TGCTTTGGTTGAAGCGAAGATG

The PCR product will then be purified using a spin column, and subjected to restriction enzyme digestion with the enzyme BsmF1 which will cut the DNA of only one of the two possible SNP sequences. The digested DNA will then be run on an agarose gel to visualize the size of the bands which will tell you whether you are homozygous for one or the other sequence, or are heterozygous.

For the lab write-up, you will be expected to utilize the tools and techniques you have already been taught to identify the sequence of the PCR product amplified by these primers, and the site of BsmF1 digestion in the context of each genotype (and the expected fragment sizes for each). Finally, you will analyze the pattern of the gel with *your* DNA and decode your own genotype. Does it match what you know (or think you know) about your own reaction to milk or milk products? If not, speculate as to why.

**Lab Protocol:**

**I. DNA Purification from a Buccal Brush**

Before starting: Preheat water bath to 65°C for use in steps 3 and 12 of the procedure.

1. To collect buccal cells, scrape the inside of the mouth 10 times with a Buccal Collection Brush. For best results, wait at least 1 h after eating or drinking to collect buccal cells. DNA may be purified immediately or samples may be stored on the collection brush for up to 1 month at room temperature (15–25°C).
2. Dispense 300 µl Cell Lysis Solution into a 1.5 ml microcentrifuge tube. Remove the collection brush from its handle using scissors or a razor blade, and place the detached head in the tube.
3. Complete cell lysis by incubation at 65°C for at least 15 min (up to 60 min for maximum yield).
4. Remove the collection brush head from the Cell Lysis Solution, scraping it on the sides of the tube to recover as much liquid as possible.
5. Add 100 µl Protein Precipitation Solution, and vortex vigorously for 20 s at high speed. Incubate for 5 min on ice.
6. Centrifuge for 3 min at 13,000–16,000 x g. The precipitated proteins should form a tight pellet. If the protein pellet is not tight, incubate on ice for 5 min and repeat the centrifugation.
7. Pipet 300 µl isopropanol and 0.5 µl Glycogen Solution (cat. no. 158930) into a clean1.5ml microcentrifuge tube, and add the supernatant from the previous step by pouring carefully. Be sure the protein pellet is not dislodged during pouring.
8. Mix by inverting gently 50 times. Centrifuge for 5 min at 13,000–16,000 x g.
9. Carefully discard the supernatant, and drain the tube by inverting on a clean piece of absorbent paper, taking care that the pellet remains in the tube.
10. Add 300 µl of 70% ethanol and invert several times to wash the DNA pellet. Centrifuge for 1 min at 13,000–16,000 x g.
11. Carefully discard the supernatant. Drain the tube on a clean piece of absorbent paper, taking care that the pellet remains in the tube. Allow to air dry for 5 min. The pellet might be loose and easily dislodged. Avoid over-drying the DNA pellet, as the DNA will be difficult to dissolve.
12. Add 100 µl DNA Hydration Solution and vortex for 5 s at medium speed to mix. Incubate at 65°C for 1 h to dissolve the DNA.
13. Incubate at room temperature overnight with gentle shaking. Ensure tube cap is tightly closed to avoid leakage. Samples can then be centrifuged briefly and transferred to a storage tube.
14. Quantitate 1 l of DNA using the micro-drop procedure. Determine the quantity of DNA remaining in the tube and calculate volume required to provide 50 ng.

**II. PCR reaction**

1. In a 25 l PCR tube mix:

10X PCR buffer 2.5l

10 mM dNTPs 0.5 l

10 M Forward primer 0.5 l

10 M Reverse primer 0.5l

Taq Polymerase 1 unit

DNA 50 ng

Water to 25 l

1. Incubate in thermal cycler 2 minutes at 94° to completely denature template
2. Perform 35 cycles of PCR amplification:

Denature: 94° for 30s

Anneal: 55° for 30s

Extend: 68° for 30s

**III. PCR Product clean-up**

1. Add 5 volumes of Buffer PB (i.e. 125 l) to 1 volume of the PCR sample and mix.
2. Check that the color of the mixture is yellow. If the color of the mixture is orange or violet, add 10 µl of 3 M sodium acetate, pH 5.0, and mix. The color of the mixture will turn to yellow.
3. Place a QIAquick spin column in a provided 2 ml collection tube. To bind DNA, apply the sample to the QIAquick column and centrifuge for 30–60 s.
4. Discard flow-through. Place the QIAquick column back into the same tube. To wash, add 0.75 ml Buffer PE to the QIAquick column and centrifuge for 30–60 s.
5. Discard flow-through and place the QIAquick column back in the same tube. Centrifuge the column for an additional 1 min. IMPORTANT: Residual ethanol from Buffer PE will not be completely removed unless the flow-through is discarded before this additional centrifugation.
6. Place QIAquick column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 30 µl Buffer EB (10 mM Tris·Cl, pH 8.5) to the center of the QIAquick membrane, let the column stand for 1 min, then centrifuge the column for 1 min IMPORTANT: Ensure that the elution buffer is dispensed directly onto the QIAquick membrane for complete elution of bound DNA. The average eluate volume is 28 µl from 30 µl elution buffer.
7. Quantitate DNA concentration in 1 µl using the micro-dot method.

**IV. Restriction digest, ethanol precipitation, electrophoresis**

1. In a microfuge tube mix:

10X buffer 5.0l

DNA volume containing 1g or 15 l maximum volume.

Water to 49 l

BsmF1 1 l (10 units)

Incubate at 65° for 1hour.

Note: Be sure to retain the rest of the DNA sample for uncut control.

1. Add 5 l 3M Sodium acetate and 125 l ice-cold ethanol. Place tube on dry ice for 10 minutes, then spin at maximum speed.
2. Wash pellet by adding 200 l cold ethanol, re-spin for 5 minutes.
3. Decant ethanol allow tube to dry inverted for 5 minutes.
4. Resuspend pellet in 10 l of electrophoresis loading buffer.
5. Load uncut and cut DNA into adjacent lanes of a 1.5% agarose gel with Sybr-Safe.Ensure that 1kb ladder is present on gel.