**Science Education Collection**

# SNP Genotyping

URL: https://www.jove.com/science-education/5544

## Abstract

Single nucleotide polymorphisms, or SNPs, are the most common form of genetic variation in humans. These differences at individual bases in the DNA often do not directly affect gene expression, but in many cases can still be useful for locating disease-associated genes or for diagnosing patients. Numerous methodologies have been established to identify, or "genotype", SNPs.

JoVE's introduction to SNP Genotyping begins by discussing what SNPs are and how they can be used to identify disease-associated genes. Several common SNP genotyping methods are then examined, including direct hybridization, PCR-based methods, fragment analysis, and sequencing. Finally, we present several examples of how these techniques are applied to genetic research today.

## Transcript

Genetic variation occurs both within and between different populations, and it contributes to the variation in traits, including disease susceptibility, among individuals. The most common variations are single nucleotide polymorphisms, or SNPs. SNP genotyping involves determining the specific sets of variants, in this case SNPs, present in an individual.

This video will discuss some of the principles behind SNP genotyping, give an introduction to several common SNP identification methods, and finally, some applications of these techniques.

First, let's discuss what a SNP is and how it can be used in genotyping.

"Polymorphisms" are alleles, or different versions of a genetic locus, found at an appreciable frequency within a population, and include base substitutions as well as insertions or deletions. Single nucleotide polymorphisms are variations at a single DNA base. SNPs can occur throughout the entire genome, and most do not alter protein structure or gene expression. However, they may still be associated with diseases due to their physical proximity, or "linkage," to a disease-causing genetic variant. On the other hand, there are several prominent SNPs that are known to directly cause diseases such as cystic fibrosis, sickle cell anemia, and β-thalassemia.

One practical application for SNP genotyping is personalized medicine, where identification of SNPs that are associated with diseases, or with differential responses to a drug or therapy, may help doctors to better diagnose and treat patients.

SNPs are also useful for genome-wide association studies, which examine hundreds of thousands of SNPs across the entire genome to identify ones that are associated with traits or diseases. GWAS takes advantage of groups of SNPs known as "haplotypes." SNPs within a haplotype are in "linkage disequilibrium," which means that the particular combination of variants are more likely to occur together than expected by chance, often the result of their physical proximity on a chromosome. This means that haplotypes can be differentiated by genotyping only a small number of individual SNPs, making genetic association studies less cost- or labor-intensive.

After looking at what SNPs are and how they can be used, let's go through several of the numerous methodologies that have been developed to genotype SNPs. These include direct hybridization, PCR-based methods, fragment analysis, and sequencing.

Direct hybridization techniques like SNP arrays use short oligonucleotide probes to identify individual SNPs directly. Hundreds of thousands of SNPs can be tested on a single biochip. The probes are chemically fixed to the chip matrix, typically a glass slide. To detect SNPs in a sample, isolated genomic DNA is amplified by PCR and fragmented to improve probe binding-efficiency. The DNA is then labeled, typically with a fluorescent marker, followed by hybridization to the chip. After the sample DNA is allowed to bind to the probes, all unbound or non-specifically bound DNA is washed away. By measuring the signal at each probe spot, the presence of the target allele in the sample can be determined.

Another SNP-typing method is allele-specific PCR, where primers are designed against SNP-containing sequences so that they only hybridize to the perfectly complementary allele. The PCR products can be detected by methods such as fluorescent tags. Alternatively, PCR primers flanking the SNP site of interest can be designed in conjunction with a SNP-specific probe, as in the TaqMan PCR assay. Here, the probe contains a fluorescent marker, as well as a quencher molecule that suppresses the fluorescence from the nearby marker. During PCR, only the specifically bound probe sequence will be degraded by the 5′ exonuclease activity of the polymerase, separating the fluorescent tag from the quencher and resulting in a fluorescent signal that indicates the presence of the particular variant.

A third category of methods, called fragment analysis, involves the creation of DNA fragments of various sizes or labels, based on the presence or absence of a particular SNP. Restriction fragment length polymorphism analysis takes advantage of the strict specificity of restriction endonucleases for their target sequences, where one allele is cleaved and the other is not. Ligation assay uses two probes situated directly next to each other, one of which terminates at the target SNP basepair. If the SNP-containing probe binds perfectly, ligation can occur. Primer extension assay involves

primers that bind one nucleotide short of the target SNP. This primer is then variably extended based on the individual allele present. The fragments generated by these methods can then be differentiated using gel or capillary electrophoresis, or mass spectrometry.

Finally, sequencing can detect SNPs very specifically, as well as identify novel SNPs with unknown sequences. However, extra care and additional experimental replicates must be performed to distinguish actual SNPs from sequencing read errors. As sequencing technologies become cheaper and more accessible, researchers are increasingly utilizing sequencing for genotyping.

Having seen how SNPs can be genotyped, let's look at some specific applications for these techniques.

Genotyping can be used to distinguish between varieties of a species with nearly identical appearance. This is accomplished by first isolating genomic DNA, followed by PCR using primers designed to bind sequences with known SNPs or other genetic variants. These primers will only amplify DNA of the variety containing the specific polymorphism, and thereby can be used to distinguish between these look-alikes.

Researchers also use genotyping to identify pathogenic bacteria harboring specific drug-resistance mutations. A streamlined SNP array protocol for low-resource settings is established by adding genomic DNA and PCR master-mix directly to a chip. Sample amplification and hybridization occur in one step. The arrays are then washed, imaged, and results are analyzed to determine the presence or absence of tuberculosis bacteria, as well as any drug-resistant mutations.

Finally, SNPs can be used to assess the function of genetic variants found to increase disease risk by association studies. Here, scientists experimented with a cell line that is heterozygous for a neutral "marker" SNP located in the transcribed portion of a gene of interest, as well as an untranscribed test SNP that is disease-associated. A negative control cell line is chosen that is heterozygous for the marker SNP, but homozygous for the non-disease associated variant of the test SNP. Allele-specific primer extension was performed on genomic DNA, as well as cDNA reverse transcribed from expressed mRNA. The primer extension products were then quantitated using MALDI-TOF mass spectrometry. By comparing the relative abundance of the fragments associated with the two marker SNP alleles, researchers can determine if the disease-associated SNP results in reduced gene expression.

You've just watched JoVE's video on SNP genotyping. This video introduced the concept and uses of SNPs, several basic methods used to identify and characterize SNPs, and highlighted three applications of SNP genotyping. As always, thanks for watching!