BY THE NUMBERS

by Duc Lu, Angus Genetics Inc.

How are haplotypes that affect fertility detected?

Mendelian Law of Segregation can be used to discover lethal recessive or semi-lethal recessive haplotypes.

In a previous "By the Numbers" article published in the February *Angus Journal*, three common terms (SNP, allele and haplotype) often used in this genomic era were introduced.

This article will focus on a tool commonly used in detecting haplotypes affecting fertility in cattle — the Mendelian Law of Segregation.

Gregor Mendel, the father of Mendelian Law, stated a diploid individual possesses a pair of alleles for any particular trait, and each parent passes one of those alleles randomly to each gamete (sperm or egg). A sire's gamete will unite with a dam's gamete, resulting in a diploid offspring. The principle is illustrated in Figure 1.

Assume *w* is the haplotype of interest at a segment in the bovine genome. This haplotype does not affect an animal unless he or she carries two copies of *w* (homozygous *ww*). There are other haplotypes at this segment too, but they are not the target one and thus are grouped all together under W. Animals that carry one copy of w are called a carrier, WW animals are non-carrier, and the homozygous *ww* are affected individuals.

According to the Mendelian Law of Segregation, both carrier sires and dams produce gametes containing either W or *w* randomly, which upon union result in a diploid offspring either WW (25% chance), or W*w* (50% chance), or *ww* (25% chance). Assuming there were 1,000 W*w* x W*w* matings in a herd and all progeny are genotyped after birth, then one would expect to see 25% of their progeny being WW, 50% being W*w*, and 25% being *ww*.

Assuming scenario 2 in Figure 2, where zero homozygous *ww* was observed in the herd, it suggests *w* is potentially a lethal recessive. All *ww* might have died prior to birth, such as early embryonic death. This is an example of a complete missing of homozygotes *ww* in the progeny. If this mechanism was embryonic death, then we would also expect a lower pregnancy rate in these carrier by carrier matings, especially on a per service basis.

Are such discovered haplotypes different from genetic conditions Angus breeders are familiar with?

Yes, they are. Most of the genetic variants that cause currently known genetic conditions are on commercial single nucleotide polymorphism (SNP) chips or tests (HD 50K or Angus GS), so if a breeder would like to test

Continued on page 42

Fig. 1: ***Mating between a sire and a dam carrying a haplotype of interest — Scenario 1***

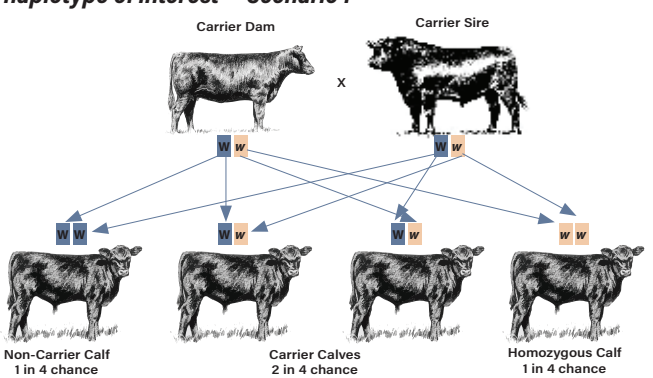
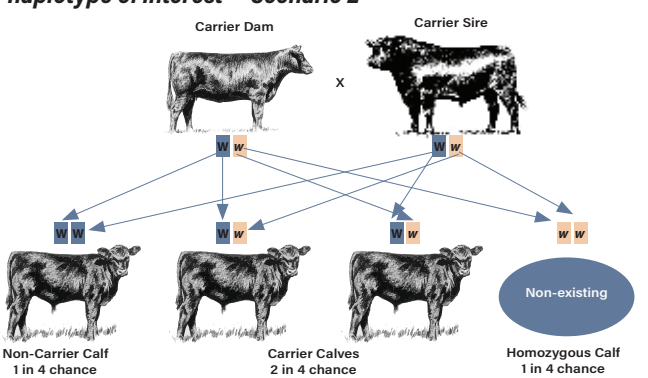


Fig. 2: ***Mating between a sire and a dam carrying a haplotype of interest — Scenario 2***



 continued from page 41

their animals for a certain genetic condition, he or she can request a specific test.

Haplotypes, on the other hand, are used to track underlying variants, but the causal variants are unknown. Determining the underlying variants that affect embryo survival, for example, will require whole genome sequencing and other scientific efforts to pin down the exact mutation.

The assumption behind a haplotype is alleles within the haplotype (or chunk of DNA) must be tightly linked together, and are, therefore, consistently inherited as a segment in its entirety from parent to offspring. This inherited segment includes the causal variant so the haplotype is not broken up due to recombination events that may ocurring during meiosis.

Other factors contributing to the quality of a haplotype include:

- Length of the haplotype,
- Number of SNP within the haplotype, and
- Genotyping errors.

Haplotype length and SNP density

Given a haplotype length, higher-density SNP chips often mean more SNP in the haplotype, potentially resulting in tighter linkage between SNP and the causal variant in the haplotype. The tighter the linkage, the more accurately an animal's status can be predicted. We have firsthand experience with our database when calling a haplotype using two different SNP panels — the standard set of SNP we use in our weekly evaluation vs. the new Angus GS[SM] v.2. AGI called a 20 SNP haplotype using 50K genotypes with a 4.77% frequency in the Angus population (approximately 9.5% of the animals carry that haplotype).

With that information in hand, Angus GS v.2 was designed to include 199 SNP in that same haplotype region. As a result of the additional SNP, the haplotype of interest now has a reduced frequency of 0.64% (approximately 1.2% of the animals carry that haplotype). The frequency of carriers went from 9.5% down to 1.2%, meaning many original carrier statuses were false positives and are really non-carriers, once additional SNP data was included from Angus GS v.2. This is an exercise performed to illustrate the importance of haplotype length and SNP density in haplotype discovery. It also implies haplotypes discovered today with current chips could change status if SNP density increases, whereas SNP density does not affect the resulting accuracy of known genetic conditions.

Genotyping errors

Genotyping errors or imputation errors can be a contributing factor as well. Angus GS v.2 was launched in December 2019, which means more animals are genotyped with other SNP chips than animals genotyped with Angus GS v.2. Designed to tackle fertility, Angus GS v.2 has approximately 72,000 SNP available to be used in these analyses. Therefore, animals with other genomic profiles since the beginning of Angus genotyping were imputed to Angus GS v.2.

While the accuracy of genotype imputation has never been 100% for any genomic profile available, the accuracy of imputing other genotypes to Angus GS v.2 is between 98-99%, which is considered highly accurate. Nevertheless, 1-2% of SNP are being mis-imputed in the process. The error does not affect the accuracy of a haplotype if it does not occur within the genomic region that contains the haplotype of interest. For this reason, AGI is taking an extra step to re-genotype a subset of animals that were previously tested with other genomic profile products with Angus GS v.2 to confirm that our regions of interest were imputed correctly.

Opportunity in the making

By looking at the frequency of haplotype status, AGI research efforts are finding haplotypes with missing or greatly reduced homozygotes. As genotyping technology increases in density, the research surrounding these haplotypes becomes more accurate. Compared to known genetic conditions in Angus, breeders will need to be aware haplotypes are not 100% accurate. This means haplotype status could change as more information becomes available, such as a denser genotyping array or an updated reference genome, which happens periodically.

The value of identifying haplotypes is to set up safeguards for Angus for the future. While this research is identifying haplotypes to be relatively low in frequency to date, left unchecked, the frequency of these haplotypes could grow to concerning levels. Luckily for Angus members, more than 800,000 animals have been genotyped in the Angus breed. This vast genomic database provides the opportunity to harness genomic technology to strategize around these nuances — an opportunity not available with smaller genomic databases. With this knowledge, Angus breeders will be in the driver's seat for the foreseeable future. AJ

dlu@angus.org

Editor's note: If you have questions, please contact the AGI team at 816-383-5100.