Published in final edited form as:

J Am Vet Med Assoc. 2019 November 15; 255(10): 1126–1136. doi:10.2460/javma.255.10.1126.

Genome-wide association studies and genetic testing: understanding the science, success, and future of a rapidly developing field

Lauren Baker, DVM, MS, Peter Muir, BVSc, MedVetClinStud, PhD, Susannah J. Sample, DVM, PhD

Department of Surgical Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, Madison, WI 53706.

Abstract

Dog owners are increasingly interested in using commercially available testing panels to learn about the genetics of their pets, both to identify breed ancestry and to screen for specific genetic diseases. Helping owners interpret and understand results from genetic screening panels is becoming an important issue facing veterinarians. The objective of this review article is to introduce basic concepts behind genetic studies and current genetic screening tests while highlighting their value in veterinary medicine. The potential uses and limitations of commercially available genetic testing panels as screening tests are discussed, including appropriate cautions regarding the interpretation of results. Future directions, particularly with regard to the study of common complex genetic diseases, are also described.

The fields of canine genetics and canine genomics are rapidly growing as scientific advances have resulted in accessible, affordable, and increasingly powerful genotyping and DNA sequencing methods. This growth can be seen in the scientific literature, as the number of publications in these fields increases, and in the consumer market, as direct-to-consumer genetic testing gains popularity among pet owners. How should veterinary practitioners use the information provided by current genetic screening panels? In this review, we focus on the basics of Mendelian diseases (also known as simple genetic diseases) to provide practitioners practical information about genetic testing that may be requested by owners and demystify genomic studies while highlighting their value in veterinary medicine. We will focus on genetic screening tests and GWAS for Mendelian diseases and discuss genetic conditions that have been successfully studied in dogs. Also discussed are future directions in the area of canine genomics, particularly with regard to common complex genetic diseases.

Mendelian Disease, Genetic Testing, and GWAS

Dog owners are increasingly interested in using commercially available genetic screening panels to learn about the genetics of their pets, both to identify breed ancestry of their pets

Address correspondence to Dr. Sample (susannah.sample@wisc.edu).

The authors have no conflicts of interest to disclose.

and to screen for specific genetic diseases. When discussing commercially available genetic screening panels, it is important for practitioners to appreciate that these panels are not regulated and that no quality guidelines currently exist for how test results must be reported.

¹ Because public demand for genetic testing of dogs continues to increase, creation of a specialty specifically devoted to veterinary genetics, genomics, and genetic counseling will likely prove necessary. However, until regulations and quality guidelines are universally adopted by genetic testing laboratories, including regulations and guidelines mandating that laboratories transparently report what the results of specific tests truly mean, owners will likely turn to general practitioners to interpret and explain the results they receive. Gaining a basic understanding of the genetics and science underlying the tests that make up these panels will help practitioners appreciate both the value of the individual tests and the context in which the results of those tests should be presented.

Veterinarians use genetics in everyday practice. Prioritizing differential diagnoses on the basis of signalment and knowledge of disease prevalence in various breeds may seem intuitive but reflects the connection between genetics and disease. Among health-care professionals, veterinarians are uniquely trained in this type of phenotypic insight.

In addition, genetic testing is increasingly being used in general practice as a tool for disease screening and disease diagnosis. For example, it is now common to undertake genetic testing for von Willebrand disease in Doberman Pinschers² or for the MDR1 mutation that results in ivermectin sensitivity in Collies and related breeds.³ Various universities and companies provide genetic screening tests for a number of other diseases in dogs, and a few websites have been developed as resources for breed-specific diseases.⁴

Still, despite the availability of these tools, the impact of genetic testing in general practice is currently somewhat limited. This lack of translation between science and practice is a continual challenge in veterinary medicine. With a decreasing number of veterinary clinician-scientists to help bridge the gap between science and practice,⁵ there is a growing need for scientists to reach out to practitioners and for practitioners to voice their needs through advocacy for scientific advancements.

To understand the use, value, limitations, and needs in the area of companion animal genetic testing, it is important for practitioners to have a basic understanding of the tools that are commonly used to identify genetic diseases. One of the most commonly used approaches in this regard is a GWAS. Through the use of GWAS, genomic regions responsible for or associated with genetic diseases or traits in companion animals, particularly recessive conditions, have been identified. To date, many canine GWAS have focused on breed-specific phenotypes that are uncommon or not associated with a particular disease.⁶ Although this type of research may not seem particularly impactful for general practice, it provides important information regarding specific breeds that could potentially increase our understanding of diseases that correspond to human conditions.^{7–9} These studies also yield important insight regarding the evolution of dogs and humans and reveal key information about the canine genome.^{10,11}

Most of the diseases in dogs studied using GWAS have been rare diseases that appear within a single breed or small subset of breeds as a result of intense artificial selection.¹² Although these diseases are important for the individual breeds in which they occur, GWAS can also be used to understand diseases that are common across dog breeds, most of which are complex genetic diseases.

Basics of genetics and genomics

To understand the basics of GWAS and the basis of genetic testing, there are a few definitions worth knowing. Most importantly, the terms genetics and genomics, although often used interchangeably, refer to different things. Genetics is the study of heredity or how traits are passed from one generation to the next. In contrast, genomics is the study of the genome. The genome is the DNA of a given organism, which is comprised of a combination of DNA from that organism's parents. Excluding some genes on sex chromosomes, an individual has 2 copies of each gene, 1 inherited from each parent. These copies are termed alleles; an allele is most simply defined as a variant copy of a gene. A combination of a given set of alleles is referred to as a genotype, and a genotype results in a particular phenotype. A good example of this can be seen with human blood types. An individual with the AO blood type inherited an A allele from one parent and an O allele from the other, which resulted in an AO genotype and a phenotype with A erythrocyte antigens.

Modes of inheritance

To date, individual genetic tests and screening panels for genetic diseases in veterinary medicine have been principally focused on Mendelian disease conditions. Mendelian diseases are conditions that are regulated by a single mutation in a single gene. These conditions are generally inherited in a recessive or dominant fashion, although sex-linked and more complicated forms of Mendelian inheritance also exist. A recessive Mendelian disease is one that appears when an individual has 2 mutant alleles, which typically results from having inherited 1 mutant allele from each parent. In contrast, a dominant Mendelian disease is one that appears when an individual has only 1 mutant allele. Generally speaking, the mutations associated with recessive Mendelian diseases most commonly inactivate the affected gene (also known as loss-of-function mutations), whereas the mutations associated with dominant Mendelian diseases most commonly result in increased gene expression (also known as gain-of-function mutations).¹³ Notably, any type of mutation can result in alterations to gene products that can result in new or interfering functions, but this is most commonly associated with dominant, gain-of-function mutations.

As an example, consider progressive rod-cone degeneration, a type of progressive retinal atrophy that is the result of an autosomal recessive mutation in the *PRCD* gene. In this case, the previously uncharacterized gene was named for the disease that led to its discovery.¹⁴ Affected animals typically have to inherit the mutated gene from both parents, whereas animals with only 1 copy of the mutated gene do not develop disease. The mutation has been identified in > 29 breeds.¹⁵ Interestingly, the identical mutation has been found to cause retinal atrophy in human beings, highlighting the value of genetic studies in companion animals.¹⁵

Genetic screening panels

With the ongoing rise in popularity of companies such as 23andMe and AncestryDNA, it is not surprising that owners are showing greater interest in understanding the genetic background of their dogs. Just as many people enjoy learning about their family history, dog owners are often interested in learning the genetic makeup of their mixed-breed dogs. Currently available genetic panels can provide general information regarding breed contribution, which is interesting to owners and could also potentially provide information to practitioners in the context of breed-associated diseases.

The more pressing issue for veterinarians, however, is helping owners interpret results of disease-screening tests offered by these companies. To do this, veterinarians need to understand how these panels screen for various diseases through the use of SNP (pronounced “snip”) markers. Single nucleotide polymorphism markers are the basis for GWAS and have become one of the most influential tools used to study genetic disease. Thus, understanding SNPs is important for understanding both how genetic screening panels work and how GWAS is used to identify disease-associated genetic variants.

History and basics of GWAS

A GWAS is an experimental research approach in which the entire genome is screened to find associations between a genomic region and a given phenotypic trait. An advantage of this type of approach is that it does not require any previous knowledge of the underlying pathophysiology of the trait of interest. This is in contrast with candidate gene studies, in which knowledge of gene function is used to exclusively investigate potential candidate genes that may contribute to the trait. Therefore, GWAS can help discover phenotype-associated genes that were not previously suspected to contribute to a disease mechanism and provide disease-associated markers that can be used as a method of genetic screening.

Use of GWAS is relatively new, with the idea behind the method first proposed in 1996 by Risch and Merikangas.¹⁶ The first human GWAS was published in 2002.¹⁷ In 2005, the first assembly of the canine genome was published, with the first GWAS in dogs following shortly afterward. In January 2019, searching PubMed with the term “GWAS dog” resulted in 289 papers, reflecting the rapid gain in popularity and enthusiasm for GWAS as a means to study phenotypes in dogs.

Single nucleotide polymorphisms are the basis for GWAS and are also used in commercial screening panels. A SNP is simply a location on the genome where a single DNA base pair varies in at least 1% of a population of study subjects (Figure 1). There are millions of SNPs spread across the genome. Some SNPs are mutations that result in changes to protein structure or gene expression and thus can be used to test for Mendelian genetic diseases. However, the vast majority of SNPs are benign variants and are not biologically important. Nevertheless, these SNPs can be used as biological markers to identify a region of the genome that associates with a phenotypic trait. It is this concept that forms the basis of GWAS.

The ability to use SNPs as markers for a large genomic region is possible because of LD. Linkage disequilibrium is typically defined as a nonrandom association of alleles at 2 or

more genomic loci in a population. Simplified, LD results in segments of DNA that are inherited together. These segments of DNA are called haplotype blocks. When a SNP is identified within a haplotype block, it serves as a tag for the entire haplotype block. When a SNP resides in a haplotype block that contains genes or elements of the genome that regulate gene expression, the SNP can be used as a tag for the phenotypic trait (Figure 2).

The size of haplotype blocks varies across the genome and is population dependent, meaning that LD differs among different groups of animals. In addition, LD is broken down as a result of generations of random mating.¹⁸ For GWAS in veterinary patients, it is expected that purebred animals are not randomly mated. Therefore, within a given breed, LD is extensive, resulting in large haplotype blocks that may contain several genes.¹⁹

General levels of LD across populations have been studied. The typical haplotype block size in human beings is in the order of tens of kilobases.^{20,21} However, for a population of purebred dogs, haplotype block sizes can range from around 400 kilobases to > 3 megabases.²² For GWAS in dogs, the specific average haplotype block size range is variable and is dependent on the combination of breeds being studied.²² Populations of mixed-breed dogs have smaller haplotype blocks than purebred dogs.

When using GWAS to scan the genome for an association between a haplotype block and a phenotypic trait, in theory only a single SNP is needed for each haplotype block. Because purebred dog populations have larger haplotype blocks (ie, bigger blocks of DNA that are inherited together), fewer SNP markers are needed to identify a genomic region of interest than for a population of mixed-breed dogs. It is for this reason that using purebred dogs for GWAS translates into a substantial scientific advantage.²² Furthermore, compared with studies involving mixed-breed dogs or multiple breeds, the cost of a GWAS using a single purebred dog population is much less, and the number of animals needed for a well-powered study is decreased.^{22,23}

When discussing genetic testing, it is important to understand the differences between genetic tests based on linkage to a particular phenotype and direct genetic tests for specific mutations. A SNP can be the basis of either type of test. When a GWAS identifies a SNP that associates with a disease, the SNP isn't typically a part of the actual disease-causing genetic mutation but instead resides within the same haplotype block as the genetic mutation of interest. In this case, the SNP associates with the disease through linkage and should only be used for genetic screening. In contrast, if a SNP is part of a genetic mutation that causes a disease, it can be used to directly test for that disease.

Unfortunately, when interpreting the results of commercially available genetic screening panels, whether the test used for each disease in the panel is a linkage test or a direct test is unclear, especially with tests that do not have associated peer-reviewed literature. Because commercial testing is not currently regulated, there is no legal need for companies offering genetic testing to disclose this information. To further complicate matters, even genetic tests for disease-causing mutations can have limitations, particularly when testing for diseases that have incomplete penetrance. For these reasons, the tests in commercially available genetic screening panels should only be considered screening tests,²⁴ and it is critical that

any positive results for these tests be confirmed through additional clinical evaluation, as determined by the disease in question. Clinical signs, signalment, and history are also important considerations when interpreting the results of these screening tests.²⁴

GWAS design

Most GWAS use a single population of purebred dogs because this increases study power and decreases the number of animals needed to evaluate a disease condition.²² This is true for both Mendelian and complex genetic diseases. Using a single breed of dog is not required for GWAS. What is critical, however, is accurate phenotyping. A GWAS can be undertaken for a quantitative trait (eg, height or weight) or as a case-control study. In either type of study, if phenotyping is not accurate, any genetic association that exists will be weakened or not detected. For studies involving quantitative traits (eg, height or weight), this means accurately measuring the trait of interest. For case-control studies, this means accurately determining which dogs actually are (cases) and are not (controls) affected. If phenotyping is not accurate, any genetic association that exists will be harder to detect, and spurious associations may be identified.

In critically evaluating reports of genetic studies, readers must be aware of these potential weaknesses and take time to evaluate these important aspects of study design. A good example of the importance of scrutinizing case and control selection would be for late-onset conditions such as laryngeal paralysis. Laryngeal paralysis can occur at any age in dogs but is most common in older dogs. For such a study, therefore, using a control group consisting of young dogs without any evidence of laryngeal paralysis would not be appropriate because some of these dogs could go on to develop the condition at a later age. Understanding the age of onset and risk of disease development over a breed's lifetime is an important part of selecting appropriate controls.

Case selection is also important. Again, using an example of laryngeal paralysis, including a dog that developed laryngeal paralysis after undergoing surgery for removal of an invasive thyroid tumor would not be appropriate because the cause of laryngeal paralysis in that dog could potentially have been environmental (ie, a result of surgery) rather than genetic.

Understanding GWAS results

Results of a GWAS are most commonly presented graphically as a Manhattan plot, so named because these graphs resemble the Manhattan skyline (Figure 3). The x-axis represents locations in the genome, placed in chromosomal order. The y-axis is typically the negative logarithm of the associated *P* value for each SNP. The plot itself consists of a series of dots, with each dot representing an individual SNP. In veterinary GWAS, the data are most often analyzed by means of LMM analysis, with each SNP evaluated in cases versus controls and a *P* value then assigned. Those SNPs with the lowest *P* values are the most highly associated with the phenotype in question. Typically, a genomic region that associates with a disease condition will consist of a peak of SNPs that have gradually decreasing and then increasing *P* values. Notably, *P* values considered significant in GWAS are much lower than the $P < 0.05$ cutoff commonly considered a marker of statistical significance. This is because each SNP is being tested individually and hundreds of thousands to millions of

SNPs may be included in the study. Thus, the cutoff for significance must be substantially lower to account for the multiple comparisons being undertaken.

The appropriate significance cutoff for studies involving populations that have undergone artificial selection, such as purebred dogs, is not well established. A common method is to use a Bonferroni correction based on either the number of SNPs analyzed or the number of haplotype blocks in the genome of the population included in the study. The methods for GWAS data filtering and the algorithms used to account for familial relatedness and ancestral population structure are also important considerations but are beyond the scope of this review.

Although LMM analyses have been the most common approach for analyzing GWAS data from dogs, there are many other algorithms and methods that can be used. These alternative methods must also be considered, particularly when the genetic contribution toward complex trait disease is being analyzed.

What GWAS provides

As the use of GWAS becomes more popular, it is important for general practitioners to understand the strengths and limitations of these studies. A GWAS allows identification of a SNP marker that segregates (to at least some degree of accuracy) cases and controls for a specific disease condition in the population being studied. However, SNP markers identified through GWAS are rarely the actual disease mutations, but rather highlight a genomic region that is in LD with the mutation of interest (ie, the SNP marker and the genetic mutation are in the same haplotype block; Figure 2). Because of genomic crossover events that can happen during meiosis, genetic tests based solely on GWAS-identified SNP markers cannot be expected to have 100% accuracy, and such tests must be interpreted in light of their sensitivity and specificity. In other words, a GWAS-identified SNP marker may tag a mutation for a large portion of a purebred population, but it is not a perfect test, and false-positive and false-negative results may occur. Most diagnostic tests that are currently available for dogs are based on an actual genetic mutation, but this is not always the case. For example, the tests for primary hyper-parathyroidism in Keeshonden and cerebellar ataxia in Spinone Italianos are both based on linkage. When considering the value of a genetic test, it is worthwhile to understand whether the test is based on a SNP marker that is in linkage with a causal mutation or a test based on an actual causal mutation.

Additionally, when GWAS is undertaken in a particular breed and a disease-associated SNP marker is identified, it is not appropriate to use that SNP marker to test for the disease in dogs of another breed. This is because of changes in LD across populations; a SNP associated with disease in one population may be tagging a very different region of the genome in another population.

In contrast, if a disease-associated genetic mutation is found, then the test for that specific mutation can be used across breeds, assuming the genetic basis of that disease is shared across breeds. Unfortunately, this may not always be the case. Clinical validity refers to how strongly a genetic variant is related to the presence, absence, or risk of disease. A disease-associated variant may be identified in a specific population of dogs, but its clinical validity

must be established before it can be used as a general test for that disease. Unfortunately, genetic testing companies are not currently required to establish the clinical validity of the tests included in commercially available screening panels, again highlighting a need for regulation.

With increasing access to next-generation sequencing methods, such as whole genome sequencing, GWAS is often considered a first step toward identification of a causal genetic mutation when a suspected Mendelian disease is being studied. Given that the canine genome contains about 2.4 billion base pairs, use of GWAS to narrow the disease-associated region of interest to closer to 1 million base pairs is highly advantageous. Once a candidate locus is found, further analysis can be used in a targeted manner to thoroughly analyze the genomic region of interest to identify a disease-associated genetic mutation.

GWAS for Mendelian disease

To date, most GWAS involving dogs have focused on autosomal recessive Mendelian diseases. There are 3 primary reasons for this. First, because of the intense selection of purebred dogs, recessive diseases are common in particular breeds, relative to their prevalence in outbred populations, and are therefore relatively important for these specific breeds. Second, in the absence of a genetic test for a recessive disease, animals that are carriers cannot be identified before they are bred with another carrier, which means that the disease is easily maintained in the population. Third, the number of subjects required to identify a genetic locus associated with a recessive disease is fairly small, generally ranging around 20 cases and 20 phenotype-negative controls in a purebred population, making identification of recessive diseases potentially easier to undertake than identification of diseases with other modes of inheritance.²⁵ Examples of autosomal recessive diseases for which mutations have been found with an approach that incorporated GWAS include cerebellar hypoplasia in Eurasier dogs²⁶; canine hereditary ataxia in Old English Sheepdogs,²⁷ Gordon Setters,²⁷ and Parson and Jack Russell Terriers²⁸; and ichthyosis in Golden Retrievers.²⁹

In contrast to recessive diseases, dominant diseases are less common in dogs. This is particularly true for dominant diseases that have an early onset, because animals with these disease conditions are unlikely to be bred. The exception is late-onset dominant diseases, which require a substantial and long-term research approach that is often challenging to execute. Also, GWAS of dominant diseases requires more cases and controls, with 40 cases and 40 controls being a minimum estimate if using a purebred population.²⁵ Examples of dominant diseases in dogs for which mutations have been identified with an approach that incorporated GWAS include polycystic kidney disease in Bull Terriers, early-onset cataracts in Australian Shepherds,³⁰ and craniomandibular osteopathy in West Highland White Terriers, Scottish Terriers, and Cairn Terriers.³¹

Importantly, many of the recessive and dominant diseases identified in dogs can also be seen in humans. Thus, many of these studies have been of translational value by helping to advance genetic discovery in human patients.

Genetic screening panels and Mendelian diseases

Most commercially available genetic screening panels that are currently available consist of tests based on SNPs associated with Mendelian diseases. These panels are advertised as a means of screening dogs for a large number of genetic diseases. Because large numbers of dogs have been tested, the resulting DNA profile collections have substantial scientific potential. Unfortunately, they have not yet been made readily available to the general scientific community.

A recent study³² from individuals affiliated with canine DNA testing companies highlights both the value and practical limitations of genetic screening panels. This study evaluated 83,000 mixed-breed dogs and 18,102 purebred dogs for the prevalence of 152 Mendelian diseases. The study primarily focused on carriers of recessive diseases (ie, dogs that had 1 disease-associated allele), meaning its emphasis was on dogs that did not necessarily have clinical evidence of disease but that would, if bred to another carrier, theoretically have an approximately 25% chance of passing the disease on to its offspring, assuming that the recessive mutation causes the disease in that individual dog's breed. Interestingly, alleles for only 127 of the 152 diseases included in the screening panel were found in the > 100,000 dogs evaluated, and just 30 of the 152 diseases made up 96% of the disease-associated alleles that were found. The authors concluded that about 30% of the mixed-breed dogs and 18% of the purebred dogs were carriers for one of the 127 diseases. The authors also specifically looked at the 9 diseases in their panel that most commonly affected both mixed-breed and purebred dogs and found that 1.4% of the mixed-breed dogs and 3.9% of the purebred dogs were homozygous for disease-associated alleles for at least 1 of these 9 genetic diseases and therefore at risk of developing clinical disease. It is also important to note that the mutations associated with many of the diseases being tested for in this panel have not been evaluated across many breeds. Whether a mutation is relevant to breeds other than the breed in which it was discovered is often not known.

To our knowledge, this was the first study to evaluate the frequencies of common recessive diseases in a large population of dogs. The findings helped put into context the distribution of recessive disease carriers in the pet population and also identified purebred dog populations that may have a risk of recessive diseases even though they were historically not considered to be at risk. However, the study also highlighted some of the limitations of these genetic screening panels. First, many of the diseases screened for with current panels are exceedingly rare or specific to a particular breed. Thus, the value of routinely screening dogs for these recessive diseases is unclear. Second, for any dog that is neutered or not intended for breeding, its status as heterozygous for a recessive disease is of little practical value.

Another major concern with commercially available genetic screening panels is the way that they are marketed to the public. Advertisements for these panels often focus on the number of diseases for which they test. The reality is that many of these tests are not relevant to the vast majority of dogs. Additionally, disease penetrance is not well understood for many of the diseases included in these panels. Penetrance refers to the proportion of individuals testing positive for a genetic disease variant that actually go on to develop disease. For many diseases, penetrance is not 100%, and the disease is referred to as having incomplete penetrance. The reason for incomplete penetrance is often not well understood, but knowing

disease penetrance is vitally important when interpreting the results of genetic tests. It should be the responsibility of the genetic testing company to provide clear interpretations of test results for pet owners and veterinarians. Unfortunately, this is not the current reality.

Pet owners must understand that most of the tests included in commercially available genetic screening panels are only screening tests and that, for these tests, a negative result does not necessarily mean their pet is free from genetic disease and a positive result must be confirmed through other diagnostic testing before any particular action is considered.²⁴ Genetic screening tests provide an assessment of an individual's genotype, but are only reflective of disease risk. This is in contrast to diagnostic genetic tests, which are used to identify a specific genetic disease on the basis of an established disease-causing genotype and can be undertaken after a disease is suspected on the basis of clinical signs. Treatment decisions, breeding decisions, and most certainly any decisions associated with life or death should not be made solely on the basis of results of commercially available genetic screening panels. Even for those diseases for which direct genetic tests for the known disease-causing mutation are available, testing should generally only be considered when clinical signs are present. Even with direct genetic tests, the results must be interpreted in the context of incomplete penetrance and with an understanding of the current gaps in knowledge regarding how genetic variants associate with diseases across breeds.

As veterinarians, we must ensure that this message is made clear to owners. Furthermore, although these panels can provide helpful information, they are not testing for the most common diseases in pet dogs because, unfortunately, the most common diseases in pet dogs generally fall into the category of complex genetic diseases.

Approaches to Investigate Common Complex Diseases

GWAS for complex disease

Complex diseases are common and inherently more interesting than Mendelian diseases but are also more challenging to understand both conceptually and scientifically.^{33,34} Most conditions seen in general practice that have some genetic basis fall into the category of complex diseases, including most cancers,^{35,36} many orthopedic conditions,^{37,38} and diseases such as atopic dermatitis,^{39,40} epilepsy,³⁶ and inflammatory bowel disease,⁴¹ among many others. It is important to understand that the genetic contributions to these complex diseases are not single gene mutations, but rather the cumulative effects of tens to thousands of genetic variants spread across the genome. The contribution of many regions of the genome to the risk of disease is called polygenicity, and the combination of these variants is often referred to as the genetic architecture of the trait.³⁴ Complex genetic diseases are also multifactorial. That is, they are not solely dependent on genetics, and environmental modifiers play important roles in shaping the final phenotype. A good example of a complex trait in human beings is height.⁴² A child with tall parents is also likely to be tall because height is influenced by the genes that child inherits. However, environmental factors such as nutrition, overall health, and degree of activity will also influence the child's growth and eventual height.

Use of GWAS to dissect complex diseases is a current standard genomic approach in the fields of human genetics and animal science. The ability to predict the risk that a dog will develop clinical manifestations of various complex genetic diseases is arguably of much greater value to general practitioners on a day-to-day basis than the ability to test for certain Mendelian diseases. Advancing our knowledge of the genetic contribution to common complex diseases in companion animals is important for several reasons:

- Complex traits are common. Americans spend billions of dollars each year on treatment of dogs with these conditions.⁴³ Owners and veterinarians need to have a greater understanding of why these diseases occur, how they can be optimally managed, and how disease-modifying treatments can help their pets and patients.
- Developing a complete picture of all genomic regions associated with a complex disease provides opportunities for pharmacological interventions, as has been the case for certain complex diseases in humans. Importantly, the relative effect that a candidate locus may have on a complex disease is not always correlated with how successful pharmacological intervention will be. Genes or biological pathways that may initially seem less important to a disease process can still prove to be effective drug targets.⁴⁴
- Undertaking a robustly powered GWAS can enable successful development of models for predicting disease risk. Once established, such models can be used to accurately assess an individual animal's genetic disease risk, providing an opportunity for selective breeding and preemptive care of at-risk dogs.
- One-health medicine is an important and popular concept that deserves more action. Understanding genetic contributions to complex diseases in purebred companion animals can help us understand similar conditions in humans. The cost of such studies in purebred companion animals is orders of magnitude less than the cost for studies in human populations because of the substantially smaller sample sizes required, the relatively easy access to affected animals, and the decreased computational resources required for data analysis.

Unfortunately, use of GWAS to investigate complex genetic diseases in companion animals is lagging substantially behind its use in humans and food animals. Approaches incorporating GWAS results to predict complex traits are commonly used for genetic selection in production animals and are currently being applied to human traits.^{45,46} In these fields, development of statistical and predictive algorithms remains an active area of research. In contrast, although studies^{47,48} creating prediction models for hip dysplasia in dogs have been published, these models have either not been validated or not successfully translated into clinical veterinary practice.

A limited number of studies have been undertaken to investigate genetic contributions to complex diseases in dogs. Many of these studies focus on orthopedic diseases,^{37,47,49} although other diseases or phenotypic traits have been investigated, such as genetic control of the response to *Leishmania* infection,⁵⁰ obsessive compulsive disorder,⁵¹ and osteosarcoma.³⁵ A large study³⁶ has also been performed evaluating genetic contributions to a number of conditions, including idiopathic epilepsy, lymphoma, and mast cell tumors,

among others. However, an optimal approach for studying genetic contributions to complex diseases in dogs has not been established. A particular challenge in a population of purebred dogs is the unique feature of subject relatedness, adding to the magnitude of challenges associated with undertaking GWAS for complex disease. To understand the basis of this challenge, an understanding of the concept of heritability is necessary.

Heritability is defined as the amount of phenotypic variance that can be explained by genetic variance. More simply, it is a measure of how much a given phenotype can be explained by genetics alone. It is important to consider heritability in the context of complex genetic diseases, as it determines the extent to which genomic predictions may be useful and whether selective breeding could be applied to reduce disease prevalence.⁵² The heritability of complex diseases ranges from 0 to 1. A heritability of 0 reflects a nonheritable, purely environmental disease (eg, infectious disease); a heritability of 1 reflects a purely genetic disease. Intuitively, the more heritable a disease is, the more successful genetic studies of that disease will be. Methods for calculating genetic heritability from pedigrees have existed for over a century, but with the greater focus on GWAS to investigate complex diseases, heritability is now often calculated by analyzing SNP data, with results referred to as SNP heritability. Generally, SNP heritability is lower than heritability estimated from pedigrees, likely because pedigrees may capture a certain amount of shared environmental risk (ie, related individuals may be raised in similar circumstances) and also because commercial SNP arrays do not completely capture other potentially important forms of genetic variation, like rare variants or structural variation.

It is important to understand that heritability is not correlated with polygenicity, which is a limiting factor when discussing predictive genetic testing for complex traits.⁵³ Polygenicity refers to the idea that many SNPs are likely to be associated with a disease. The term SNP effect is used to refer to the relative importance of a particular SNP, as determined by a GWAS. A SNP that has a large effect is one that strongly associates with the complex disease being studied, whereas a SNP that has a small effect is one that has a weaker association with the disease. Polygenicity is challenging because traditional statistical methods for analyzing GWAS results make it difficult to tell whether a nonsignificant SNP result is truly not significant or represents a small, but real, effect. In general, if a trait is very polygenic, most of the SNPs that affect the trait will have very small effects and, thus, will be difficult to identify.⁵⁴ Furthermore, although SNPs are good at tagging common genetic variants in a population, they are not good at identifying rare variants or other forms of genetic variation, such as copy number variants, that likely also influence phenotype.⁵⁵ Indeed, all of these forms of genetic variation likely interact with each other to increase susceptibility to complex diseases in ways we do not yet understand or have the technology to identify, adding to the challenge of genetic dissection of complex diseases.

Missing heritability has become a popular term when talking about the genetics of complex traits. It is a common phenomenon that significant SNPs identified with GWAS only explain a small proportion of heritability, regardless of whether heritability is calculated from pedigrees or through analysis of SNPs. There is growing evidence that missing heritability is actually an artifact of the approach taken to study the genetic basis of disease. Put simply, “missing” heritability is more likely hiding within genetic study results. Additional

heritability can be explained through improved study design aimed at overcoming the challenges posed by polygenicity, such as larger sample sizes, better phenotyping, consideration of nongenetic risk factors, and use of alternative statistical approaches, for example.⁵⁵ Nevertheless, even if study designs improved to the point that all heritability could be accounted for, the fact that heritability of complex disease is, by definition, < 1 limits the accuracy we can achieve with genetic tests. Ultimately, the best prediction performance that can be expected from a purely genetic test is the square root of the heritability of the disease or trait.⁵⁶ For example, if a disease has an estimated heritability of 0.4 (ie, 40% of the disease risk can be explained by genetic variance), the best that can be expected from a genetic test for that disease is a predictive accuracy of about 63%. A test with this degree of accuracy can still be useful from the perspective of identifying dogs that are at higher risk or making breeding decisions, as is the case for production animals. With greater understanding of the contributions of environmental variables to disease risk, additional environmental variables can be added to an algorithm to improve prediction ability. Still, these tests will not be as accurate as direct genetic tests for the mutations associated with Mendelian diseases. Rather, they will be estimates of the risk of developing a disease. However, predictive test results will still be useful in deciding whether any individual animal should be monitored over time. They can also potentially be useful in deciding whether particular dogs should be used for breeding, although ultimately the choice of whether to breed a dog with high genetic risk should be made with consideration of the whole animal and on an individual basis.

Study design and statistical approaches for complex diseases

When evaluating published reports of genetic studies of complex diseases, the general practitioner should consider 2 major issues in study design. The first is sample size. In GWAS of complex traits in humans, the number of trait-associated genetic loci may reach into the thousands.^{57,58} Thus, these types of studies now often use a consortium approach to analyze SNPs from hundreds of thousands of subjects, with some investigators suggesting that such numbers are still too low.³³ Arguably, for a robust multibreed study of a complex trait in dogs, given the genetic diversity across breeds and correspondingly small haplotype block size, many thousands of dogs would be necessary to reach adequate statistical power. For this reason, use of purebred dogs from a single breed is advantageous for dissecting the genetic contributions of complex traits. The number of dogs needed for a robust single-breed study is not yet defined, but perhaps 1,000 dogs might be a reasonable start. However, the realities of data collection and the cost of sample processing can make it difficult to recruit, phenotype, and genotype a sufficiently large number of study subjects. It is likely that collaboration between multiple universities and specialty centers will be needed to amass veterinary medical data sets of this size.

The second issue practitioners must consider when reading published reports of genetic studies of complex diseases is the statistical approach. To date, most veterinary genetic studies of complex diseases and traits have used LMMs for statistical analysis. The LMM method is popular because it quite easily corrects for population structure (eg, relatedness) within a study sample by incorporating a genetic relationship matrix or principal components analysis into the statistical model. Correction for population structure is very

important; however, the LMM method also tests SNPs individually for associations with the phenotype of interest. Correction for this multiple testing means that *P* values must meet a stringent threshold to be considered significant. The reality is, though, that many of these SNPs may have small effect sizes, are correlated with each other, and may be interacting in ways we don't fully understand and are not captured by the LMM method. Alternative methods, such as gene-based testing, by which SNP effects are aggregated onto genes, or Bayesian statistical approaches, which model the combined effect of all SNPs tested, may be more appropriate. Delving into the details of alternative statistical approaches for complex trait GWAS is beyond the scope of the present review, but practitioners should be aware that this is a hotly debated and active area of research. Further work investigating statistical approaches to investigate the genetic contributions of common complex diseases in dogs is needed.

Future of GWAS for complex diseases

A common criticism of GWAS of complex diseases in companion animals is that the work undertaken thus far has not yet successfully provided an accessible predictive genetic test. Ultimately, further work needs to be undertaken. Use of a consortium approach through collaboration between research groups will likely be needed to develop sufficiently powered data sets that will yield models capable of making reasonable predictions. Such approaches are common in human medicine but are yet to be implemented in the veterinary community. Although the utility of genetic testing is limited by trait heritability, this does not mean that information gained from genetic testing is not useful. By providing information on the risk of disease development, testing can help identify dogs with a high genetic risk for which additional screening and discussion with the owners regarding mitigating environmental and other nongenetic risk factors might be warranted. Additionally, large-effect SNPs that are identified through GWAS highlight biological pathways that are important in understanding disease pathophysiology. Such pathways could be targeted for pharmaceutical development.

Conclusions

Genome-wide association study is a popular and powerful method to facilitate understanding of genetic traits across species but is particularly useful in purebred animals. Although most GWAS undertaken in veterinary medicine to date have focused on less common breed-specific recessive diseases, important discoveries continue to be made that have substantial benefit to both the veterinary and the human medical community.

As more genetic and genomic prediction tests become available and as direct-to-consumer genetic screening panels continue to gain popularity, veterinarians must be prepared to critically evaluate this information as it applies to medical practice. Most of the conditions evaluated with these panels are recessive, and identifying that a pet is a carrier for such a condition is likely of limited value to most pet owners. Although a select number of conditions tested for with direct-to-consumer genetic panels can be seen in mixed-breed dogs, most are rare and breed specific.³² If an animal is found to possess a risk variant, veterinarians must be able to counsel owners about what these results mean and, perhaps more importantly, convey to the public that although useful, these panels do not screen for

the most common genetic diseases in animals, which are complex in nature. Overall, genetic screening tests are a new tool that holds enormous promise. However, the veterinary scientific community must continue to critically evaluate the most pressing needs of veterinary practitioners and overcome research challenges in the field of GWAS, particularly when it comes to studying inheritance and prediction of common complex diseases. This work is a substantial endeavor and carries some risk but has the potential for lasting impacts on our patients' and owners' lives. For a general practitioner, understanding how common complex diseases can be managed, prevented, or treated through the development of disease-modifying treatments or use of precision medicine to optimize management of individual patients will be transformative.

Acknowledgments

Dr. Baker received support from a National Library of Medicine training grant to the Computation and Informatics in Biology and Medicine Training Program (NLMT15LM007359). Dr. Sample received support from the National Institutes of Health (K01OD019743-01A1).

ABBREVIATIONS

GWAS	Genome-wide association study
LD	Linkage disequilibrium
LMM	Linear mixed model
SNP	Single nucleotide polymorphism

References

1. Shaffer LG. Standards and guidelines for canine clinical genetic testing laboratories. *Hum Genet* 2019;138:493–499. [PubMed: 30426199]
2. Gentilini F, Turba ME. Two novel real-time PCR methods for genotyping the von Willebrand disease type I mutation in Doberman Pinscher dogs. *Vet J* 2013;197:457–460. [PubMed: 23911791]
3. Mealey KL, Bentjen SA, Gay JM, et al. Ivermectin sensitivity in collies is associated with a deletion mutation of the *mdr1* gene. *Pharmacogenetics* 2001;11:727–733. [PubMed: 11692082]
4. Burns K. Unlocking the genetic secrets of your dog. Available at: www.avma.org/News/JAVMANews/Pages/170315a.aspx. Accessed Aug 27, 2018.
5. Physician-Scientist Workforce (PSW) Report 2014 Available at: report.nih.gov/workforce/psw/index.aspx. Accessed Aug 27, 2018.
6. Schoenebeck JJ, Ostrander EA. Insights into morphology and disease from the Dog Genome Project. *Annu Rev Cell Dev Biol* 2014;30:535–560. [PubMed: 25062362]
7. Plassais J, Lagoutte L, Correard S, et al. A point mutation in a lincRNA upstream of *GDNF* is associated to a canine insensitivity to pain: a spontaneous model for human sensory neuropathies. *PLoS Genet* 2016;12:e1006482. [PubMed: 28033318]
8. Wolf ZT, Brand HA, Shaffer JR, et al. Genome-wide association studies in dogs and humans identify *ADAMTS20* as a risk variant for cleft lip and palate. *PLoS Genet* 2015;11:e1005059. [PubMed: 25798845]
9. Frischknecht M, Niehof-Oellers H, Jagannathan V, et al. A *COL11A2* mutation in Labrador Retrievers with mild disproportionate dwarfism. *PLoS One* 2013;8:e60149. [PubMed: 23527306]
10. Talenti A, Dreger DL, Frattini S, et al. Studies of modern Italian dog populations reveal multiple patterns for domestic breed evolution. *Ecol Evol* 2018;8:2911–2925. [PubMed: 29531705]

11. Parker HG, Dreger DL, Rimbault M, et al. Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Reports* 2017;19:697–708. [PubMed: 28445722]
12. Marsden CD, Ortega-Del Vecchyo D, O'Brien DP, et al. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proc Natl Acad Sci U S A* 2016;113:152–157. [PubMed: 26699508]
13. Hijikata A, Tsuji T, Shionyu M, et al. Decoding disease-causing mechanisms of missense mutations from supramolecular structures. *Sci Rep* 2017;7:8541. [PubMed: 28819267]
14. Zangerl B, Goldstein O, Philp AR, et al. Identical mutation in a novel retinal gene causes progressive rod-cone degeneration in dogs and retinitis pigmentosa in humans. *Genomics* 2006;88:551–563. [PubMed: 16938425]
15. Downs LM, Hitti R, Pignolato S, et al. Genetic screening for PRA-associated mutations in multiple dog breeds shows that PRA is heterogeneous within and between breeds. *Vet Ophthalmol* 2014;17:126–130. [PubMed: 24255994]
16. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517. [PubMed: 8801636]
17. Ozaki K, Ohnishi Y, Iida A, et al. Functional SNPs in the lymphotoxin- α gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;32:650–654. [PubMed: 12426569]
18. Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 2008;9:477–485. [PubMed: 18427557]
19. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 2005;438:803–819. [PubMed: 16341006]
20. Reich DE, Cargill M, Bolk S, et al. Linkage disequilibrium in the human genome. *Nature* 2001;411:199–204. [PubMed: 11346797]
21. Weiss KM, Clark AG. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 2002;18:19–24. [PubMed: 11750696]
22. Sutter NB, Eberle MA, Parker HG, et al. Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res* 2004;14:2388–2396. [PubMed: 15545498]
23. Parker HG, Shearin AL, Ostrander EA. Man's best friend becomes biology's best in show: genome analyses in the domestic dog. *Annu Rev Genet* 2010;44:309–336. [PubMed: 21047261]
24. Shaffer LG, Geretschlaeger A, Ramirez CJ, et al. Quality assurance checklist and additional considerations for canine clinical genetic testing laboratories: a follow-up to the published standards and guidelines. *Hum Genet* 2019;138:501–508. [PubMed: 30982136]
25. Karlsson EK, Baranowska I, Wade CM, et al. Efficient mapping of Mendelian traits in dogs through genome-wide association. *Nat Genet* 2007;39:1321–1328. [PubMed: 17906626]
26. Gerber M, Fischer A, Jagannathan V, et al. A deletion in the VLDLR gene in Eurasier dogs with cerebellar hypoplasia resembling a Dandy-Walker-like malformation (DWLM). *PLoS One* 2015;10:e0108917. [PubMed: 25668033]
27. Agler C, Nielsen DM, Urkasemsin G, et al. Canine hereditary ataxia in Old English Sheepdogs and Gordon Setters is associated with a defect in the autophagy gene encoding RAB24. *PLoS Genet* 2014;10:e1003991. [PubMed: 24516392]
28. Gast AC, Metzger J, Tipold A, et al. Genome-wide association study for hereditary ataxia in the Parson Russell Terrier and DNA-testing for ataxia-associated mutations in the Parson and Jack Russell Terrier. *BMC Vet Res* 2016;12:225. [PubMed: 27724896]
29. Grall A, Guaguère E, Planchais S, et al. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in Golden Retriever dogs and humans. *Nat Genet* 2012;44:140–147. [PubMed: 22246504]
30. Ricketts SL, Pettitt L, McLaughlin B, et al. A novel locus on canine chromosome 13 is associated with cataract in the Australian Shepherd breed of domestic dog. *Mamm Genome* 2015;26:257–263. [PubMed: 25894238]
31. Hytönen MK, Arumilli M, Lappalainen AK, et al. Molecular characterization of three canine models of human rare bone diseases: Caffey, van den Ende-Gupta, and Raine syndromes. *PLoS Genet* 2016;12:e1006037. [PubMed: 27187611]

32. Donner J, Anderson H, Davison S, et al. Frequency and distribution of 152 genetic disease variants in over 100,000 mixed breed and purebred dogs. *PLoS Genet* 2018;14:e1007361–e20. [PubMed: 29708978]
33. Wray NR, Wijmenga C, Sullivan PF, et al. Common disease is more complex than implied by the core gene omnigenic model. *Cell* 2018;173:1573–1580. [PubMed: 29906445]
34. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169:1177–1186. [PubMed: 28622505]
35. Karlsson EK, Sigurdsson S, Ivansson E, et al. Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome Biol* 2013;14:R132. [PubMed: 24330828]
36. Hayward JJ, Castelano MG, Oliveira KC, et al. Complex disease and phenotype mapping in the domestic dog. *Nat Commun* 2016;7:10460. [PubMed: 26795439]
37. Baker LA, Kirkpatrick B, Rosa GJM, et al. Genome-wide association analysis in dogs implicates 99 loci as risk variants for anterior cruciate ligament rupture. *PLoS One* 2017;12:e0173810. [PubMed: 28379989]
38. Huang M, Hayward JJ, Corey E, et al. A novel iterative mixed model to remap three complex orthopedic traits in dogs. *PLoS One* 2017;12:e0176932. [PubMed: 28614352]
39. Wood SH, Ke X, Nuttall T, et al. Genome-wide association analysis of canine atopic dermatitis and identification of disease-related SNPs. *Immunogenetics* 2009;61:765–772. [PubMed: 19838693]
40. Tengvall K, Kierczak M, Bergvall K, et al. Genome-wide analysis in German Shepherd Dogs reveals association of a locus on CFA 27 with atopic dermatitis. *PLoS Genet* 2013;9:e1003475. [PubMed: 23671420]
41. Peiravan A, Bertolini F, Rothschild MF, et al. Genome-wide association studies of inflammatory bowel disease in German Shepherd Dogs. *PLoS One* 2018;13:e0200685. [PubMed: 30028859]
42. Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014;46:1173–1186. [PubMed: 25282103]
43. Wilke VL, Robinson DA, Evans RB, et al. Estimate of the annual economic impact of treatment of cranial cruciate ligament injury in dogs in the United States. *J Am Vet Med Assoc* 2005;227:1604–1607. [PubMed: 16313037]
44. Cao C, Moul J. GWAS and drug targets. *BMC Genomics* 2014;15:S5.
45. Vazquez AI, de los Campos G, Klimentidis YC, et al. A comprehensive genetic approach for improving prediction of skin cancer risk in humans. *Genetics* 2012;192:1493–1502. [PubMed: 23051645]
46. de los Campos G, Hickey JM, Pong-Wong R, et al. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 2013;193:327–345. [PubMed: 22745228]
47. Bartolomé N, Segarra S, Artieda M, et al. A genetic predictive model for canine hip dysplasia: integration of genome-wide association study (GWAS) and candidate gene approaches. *PLoS One* 2015;10:e0122558. [PubMed: 25874693]
48. Guo G, Zhou Z, Wang Y, et al. Canine hip dysplasia is predictable by genotyping. *Osteoarthritis Cartilage* 2011;19:420–429. [PubMed: 21215318]
49. Baker LA, Rosa GJM, Hao Z, et al. Multivariate genome-wide association analysis identifies novel and relevant variants associated with anterior cruciate ligament rupture risk in the dog model. *BMC Genet* 2018;19:39. [PubMed: 29940858]
50. Quilez J, Martínez V, Woolliams JA, et al. Genetic control of canine leishmaniasis: genome-wide association study and genomic selection analysis. *PLoS One* 2012;7:e35349. [PubMed: 22558142]
51. Tang R, Noh HJ, Wang D, et al. Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biol* 2014;15:R25. [PubMed: 24995881]
52. Evans LM, Tahmasbi R, Vrieze SI, et al. Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet* 2018;50:737–745. [PubMed: 29700474]
53. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 2008;9:255–266. [PubMed: 18319743]

54. Moser G, Lee SH, Hayes BJ, et al. Simultaneous discovery, estimation and prediction analysis of complex traits using a Bayesian mixture model. *PLoS Genet* 2015;11:e1004969. [PubMed: 25849665]
55. Robinson MR, Wray NR, Visscher PM. Explaining additional genetic variation in complex traits. *Trends Genet* 2014;30:124–132. [PubMed: 24629526]
56. Wray NR, Yang J, Hayes BJ, et al. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet* 2013;14:507–515. [PubMed: 23774735]
57. International Schizophrenia Consortium Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748–752. [PubMed: 19571811]
58. Zeng J, Vlaming R, Wu Y, et al. Signatures of negative selection in the genetic architecture of human complex traits. *Nat Genet* 2018;50:746–753. [PubMed: 29662166]

Figure 1—.

A schematic representation of a SNP. A SNP is a variation of a single nucleotide that occurs in at least 1% of a population of study subjects; SNPs are the basis of GWAS.

Figure 2—.

A schematic representation of LD. Linkage disequilibrium is a nonrandom association between alleles on the same chromosome. Regions of the genome that are in strong LD are referred to as haplotype blocks, and SNPs located within a haplotype block that also contains a genetic variation associated with a phenotypic trait can be used as a means to tag that haplotype block. Populations with intense artificial selection, such as purebred companion animals, have substantially larger haplotype blocks than do outbred populations.

Figure 3—.

An example of a Manhattan plot, the standard graphical means to display results of a GWAS. Each dot represents a SNP. The x-axis represents locations in the genome, placed in chromosomal order. The y-axis represents scaled P values. Genomic regions that are associated with a given phenotype can be identified through the formation of a peak of SNPs, suggesting the locus contains a candidate genetic variant.