Mapping Out the Dogs Genetic Future

Building a road map of our dog's makeup through the Canine Genome Project eventually will lead to the eradication of genetic disease

By Susan Thorpe-Vargas, Ph.D. and John Cargill, M.A., M.B.A., M.S.

In 1990, the greatest intellectual task ever attempted by humans began. Even more of a challenge than walking on the moon, the Human Genome Map Project staggers the imagination in terms of concept and complexity.

The ultimate intent of the project is to ascertain the definitive sequence of the more than 3 billion base-pairs comprising the human genome (a genome is all the genetic material in the chromosomes of a particular organism; its size generally is given as its total number of base pairs). This enormous effort has "spilled over" to other species, and our dogs' will reap the benefits.

Results of the Canine Genome Project, which hasn't the social, ethical and legal implications that muddy the waters of the human genome work, may be used to enhance the quality of our dogs' lives and help us back out of the genetic cul-de-sac in which we now find ourselves. Pet owners spend billions of dollars every year diagnosing and treating genetic diseases afflicting their pets. We now have in our hands the elementary tools to prevent or ameliorate our dogs' physical suffering. And in a truly international endeavor in January of this year, the dog community took another small step forward: the publishing of the first linkage map of the dog far earlier then ever expected.

The First of Many Hurdles

One of the biggest hurdles to overcome when mapping a genome, human or canine, is to assign a gene or genetic marker to a particular chromosome. Chromosomes are discrete self-replicating units of DNA composed of four different nucleotide bases: cytosine (C), which always pairs with guanine (G); and adenine (A), which always pairs with thymine (T). The sequence of these DNA sub-units determines the gene products that give all life its physical expression. The human has 46 chromosomes; the dog has 78. Unfortunately assigning a genetic marker has been much more difficult because most of the canine chromosomes are the same shape and many are quite similar in size.

In mammals, the genetic material is sequestered in a special area of the cell called the nucleus. As the cell grows and eventually divides, it duplicates this genetic material so each new cell contains the same amount. This process is called mitosis, and any organism in which this occurs is called an eukaryote.

An additional cell division (meiosis) occurs when gametes (the mature male or female reproductive cell; i.e., sperm or ovum) are formed; however, the resultant four cells contain only half the number of chromosomes (called haploid). This makes sense when you consider that when the sperm and egg join, they each are contributing half of the genetic information from the mother and half from the father.

It is possible to see chromosomes only in certain phases of the cell cycle. The best time to see them is during metaphase. Normally, chromosomes exist in a dispersed state that cannot be seen

with an ordinary light microscope. Just before the cell divides, it tightly gathers up its chromosomes. While in this state of metaphase, it is possible to take a picture of all the chromosomes in the cell. We call this picture the karyotype, and in this picture we can see the number of chromosomes, their size and their physical appearance.

Standardization of the canine karyotype was necessary before researchers could relate genes or genetic markers to their chromosomal origins. Development of chromosome-specific markers will ensure that all of the canine chromosomes will be represented within the map and that the linkage groups are correctly orientated physically.

This difficult barrier has been overcome by some brilliant work in England. Knowing the dog has 39 haploid chromosomes, researchers used a male dog in their experiment because it is easy to see the Y chromosome. That left 38 chromosomes to identify.

Researchers first separated the chromosomes by their DNA content and the use of two fluorescent dyes that distinguish the base-pair ratio by preferentially staining either A-G- or C-T rich regions. Using a technique called dual-laser flow cytometry, they were able to resolve their sample into 32 different components. Twenty-two of the portions contained single chromosomes, and the remaining eight had two each. Thus all of the chromosomes were accounted for.

To identify the chromosome type, they then used these fractions to "paint" a normal metaphase chromosome spread (a cell that has been "fixed" chemically so it no longer cycles) and highlight the chromosomes using another technique called FISH (fluorescent in-situ hybridization).

Hybridization is a very important concept to understand because it is the basis for many of the methods used to study DNA. Recalling that the two possible DNA base-pairs are C-G and A-T, if you have a DNA strand that reads AATGGCTAT, its complimentary strand would have a base-pair sequence of TTACCGATA. In FISH, complementary strands of DNA or RNA preferentially bind to each other. If one of the strands is tagged with a fluorescent dye, it can be used to locate its equivalent complement on another DNA strand. Other types of probes use radiographic or immunological labels. Hybrid probes will be addressed in greater detail when we discuss mapping strategies.

Mapping Our Way

Just like maps we use every day, genetic maps establish spatial organization and symbolize a wide variety of information. They also are similar in that there are different types of genetic maps, each with a corresponding range and level of precision.

The karyotype (also known as a cytogenic map) is the lowest resolution of what is known as the physical map. The highest resolution would be to know any post transcription modifications once we know the entire base-pair sequence. Another type of genome map is a linkage map. The final genetic map will be a synthesis of physical and linkage maps. This new map will let us know which chromosome a gene is on, how many base-pairs separate each genetic marker, their positions relative to each other and ultimately the complete base-pair sequence. Once the entire sequence has been resolved, we will need to find all the genes and use this information to determine their function.
The medical applications of this map alone are overwhelming. These data, used as diagnostic

tools to identify deleterious mutations, combined with future gene therapy technologies, could lead to the eventual eradication of genetic disease. We also could learn how certain behavioral traits are transmitted, which is of interest to dog breeders.

Making The Linkage

In 1865, the father of modern genetics, a young monk named Gregor Mendel, published a paper in which he described the inheritance of certain traits he had observed while growing peas. In choosing which attributes to follow, Mendel was very lucky that he chose the characteristics he did, as they all turned out to be on different linkage groups. As a rule, we equate linkage groups with individual chromosomes, and the number of linkage groups corresponds to the haploid number of chromosomes. Thus, the dog has 39 linkage groups.

Mendel's observations led him to postulate two "laws" The first law says "particular factors", (genes) come in different forms (alleles). When gametes are formed, these alternative alleles are inherited independently from each other.

Mendel's second law predicts that different genes (i.e., different traits that are not on the same chromosome) will assort themselves into two different types of progeny in statistically equal amounts. These two types are the parental type and the recombinant type. However, when the genes that code for those traits are on the same chromosome, the percentage of recombinant types would be less than the anticipated 50 percent.

American geneticist Thomas Hunt Morgan suggested this lowered recombination rate simply was a function of how far apart the genes were from each other. The closer together they were, the more likely they would stay 'linked' We can use this information to predict the relative distance between the loci of two genes.

Today we measure the distance that separates genetic markers in centimorgans (cM). Two loci are said to be 1 cM apart if they are separated by a recombination event 1 percent of the time. This roughly corresponds to a physical distance of 1 million base-pairs.

The next step in the mapping process is to determine the linear order of the genetic markers. For instance, let's say Gene A is 5 cM from Gene B and Gene B is 7 cM from Gene C. If we then find out Gene A is 12 cM from Gene C, we can assume their relative positions are Gene A.....Gene B.......Gene C.

It would be nice if it were this easy, but unfortunately it is not. Coding regions, called exons, are just too far apart to be linked conveniently, and so we need to use other types of genetic markers. Another problem is that compared with bacteria or even fruit flies, the dog has too few progeny to generate the statistical recombination data needed. As a rule, humans have even fewer offspring.

The discovery and use of microsatellites (a genetic marker appearing in the noncoding region, or intron, of the gene) has overcome this barrier. So far, about 285 canine microsatellites have been characterized of the 1,000 microsatellites researchers think they need to saturate the canine genome, but more are expected to be mapped. Interspersed between genes are long stretches of introns. Within these sections are DNA sequences that can be followed because their unique patterns are inherited. In order to be useful, these genetic markers must be similar within species, breed and family groups, yet be different enough (polymorphic) to detect the

differences among individuals.

There exist di-, tri- and tetra repeat patterns, but because of founder effect and the tight linebreeding inherent in the purebred dog, the most useful microsatellites for elucidating the canine linkage map have been tetra repeats.(Founder effect is simply a form of acute genetic drift, which is the variation in gene frequency from one generation to another due to chance) These singular patterns of four nucleotides usually appear in blocks that vary from 10 to 30 units long. For example, a puppy might receive an (ATTG) 14 pattern from its dam and a (GACA) 22 element from its sire. His littermate may have inherited the exact same tetra repeat from the dam, but a mutation within this section of DNA could have resulted in a (GACA) 21 pattern instead of the one expected from the sire. Although these areas are not genes, differences in the number of copies of the basic repeat unit also are called alleles (length polymorphism). Because of technological advances, it is fairly easy to ascertain the difference between the two genotypes, and these procedures are the basis for some parentage tests now available. The more alleles a microsatellite has, the more likely it is to be useful.

Mutations that occur in these noncoding regions do not cause changes in the dog's appearance; however, linking these genetic markers to disease alleles or genes that characterize a specific trait will lead to diagnostic tests to identify carriers or affected individuals. Several of these tests already are available. These microsatellites especially are useful for identifying the carrier status of genetic disorders that arise from mutations at different sites within the same gene. This is why breed-specific tests often are required for the same disease.

Recombinant DNA technology promises to make higher resolution linkage maps possible. A lab in France has used radiation to fragment human chromosomes and has fused these fragments with cells from other species. These hybrid cells can be manipulated so that only specific human chromosomal components are retained. Determining the frequency of genetic markers that stay together after being irradiated places their order and the distance between them at a finer resolution. These techniques also have overlapped into the canine mapping effort, and work is progressing rapidly on a radiation hybrid panel specific to the dog.

The Physical Mapping Realm

In addition to the linkage maps, there are several types of physical maps:

Chromosomal maps

The lowest-resolution physical map is called a cytogenetic map (karyotype). During the metaphase and the interphase stage of the cell cycle, it is possible to stain the chromosomes with various dyes that result in distinctive banding patterns. Using radioactive or fluorescent labels it is possible to assign genes or other identifiable DNA fragments to their respective chromosomes and to estimate the distance between them, measured in base-pairs. Improved FISH methodology now allows identification of genetic markers from as close as 2Mb to 5 Mb apart (one Megabase, or Mb, equals approximately 1 cM).

With FISH, we can observe chromosomal mutations and abnormalities associated with disease states. German researchers have discovered a translocation on the first canine chromosome (a type of mutation) that is linked to mammary tumors in dogs. Cytogenetic analysis may prove useful for comprehending the underlying genetic mechanisms for other types of cancers for dogs and humans.

Complementary DNA (cDNA) maps

Although two genetic markers may have a recombination rate higher than 50 percent, this does not preclude them from being on the same chromosome. This further complicates the mapping issue. The trick is finding out which of the 39 unique dog chromosomes to assign a particular gene to. One of the methods used depends on knowing the protein the gene is responsible for making and working backward to figure out the approximate DNA sequence. Using a tagged complementary hybrid probe made from a synthetic DNA sequence, it is possible to see where the gene is located on the chromosome.

Another way to map a gene to a chromosome is to know the base-pair sequence of the gene that codes for the same trait in another but genetically related species. For example, all mammals have some genes in common (we even share conserved sequences, base sequences in a DNA molecule that have remained essentially unchanged through evolution, with the lower orders of animals). Although entire chromosomes are not conserved among species, parts of chromosomes, called syntenic groups, are observed.

Homologous genes and genetic markers from the human mapping project have been beneficial to the canine map effort, and the canine map can be expected to be useful to the Human Genome Project. Knowing the function or position of a certain gene in one species makes it a possible candidate gene for the same ailment or trait in another species. One such ailment, Severe Combined Immunodeficiency (SCID), is caused in humans and canines by a mutation in one of the proteins that form the receptor site for interleukin-2. This defect causes a profound inability to mount both a cell-mediated and humoral (antibody) immune response and often is called the "Boy in the Bubble" disease because of the movie about a boy with this affliction who lived in an isolation bubble.

Several different laboratory techniques are used to localize differences to a smaller region of the genome. Once such an area has been identified, it then is possible to use automated sequencing methods to distinguish any base-pair mutations. If these mutations result in an amino acid substitution within the coded protein product, it becomes a likely suspect. The candidate gene approach can save a lot of time, not only by providing a model for the progression and course of a disease, but by suggesting treatment strategies also.

Nature's Scissors

More than 30 years ago, a handy tool for the geneticist was discovered. Several proteins were isolated from various strains of bacteria and were named restriction enzymes because they cut DNA. The normal function of these enzymes was to protect the bacteria from attack by phage (virus that infect bacteria) or other foreign DNA. Each restriction enzyme recognizes a particular double-stranded DNA sequence, and this specificity has been extremely useful for mapping the genome. Hundreds of restriction enzymes have been isolated. Depending upon the source, these enzymes 'see' restriction sites that vary from four to eight base-pair recognition sites. Some rare-cutter enzymes cut DNA very infrequently, which results in a small number of very big pieces.

One approach to high-resolution physical mapping is to cut a single chromosome into large sections, place these fragments in their proper order then cut those pieces into smaller and smaller sections until their sequence is determined. Although this approach produces more continuity and fewer breaks, it is not as effective for finding a specific gene.

A better technique for obtaining finer mapping details is the contig map. The contig map is produced by cutting a chromosome into very small pieces, cloning these pieces and constructing an overlapping clone 'library'. Cloning is a recombinant technique that involves inserting a DNA segment into another host cell, called a cloning vector, and using that cell's own replication apparatus to generate multiple copies of foreign DNA. This provides large amounts of experimental material.

Cloning vectors often are bacteria such as E.coli, but recent technological advances have made it possible to clone larger segments of DNA by using an artificial cloning vector packed into a lamda phage. This virus normally infects bacteria and inserts its own DNA into that cell's genome, where it is replicated along with the normal cellular DNA. So using nature's own tricks often has worked well for us in this endeavor.

Once contig mapping of a particular section of the genome is accomplished in one laboratory, the resulting genetic library can be published so other researchers can use the same information. A common reference system called sequence-tagged sites makes this sharing of information possible. STS are short DNA fragments (200 to 500 base-pairs long) whose unique base sequence and location make them useful landmarks.

A variation of this method is to sequence cDNA partially instead of random sections of DNA. Since they are "tagging" a transcription product of an expressed gene, they are called expressed sequence tags, or EST. These are especially useful for finding candidate genes.

We Have Only Just Begun

Current mapping strategies, as brilliant and innovative as these techniques are, still have left gaping holes that need to be filled. Genetic mapping is technology-driven, but technology costs money and time. Faster, more precise mapping methods are needed. Funding for the canine project generally takes a back seat to funding for the human mapping effort and for species that are more agriculturally and economically important than the dog.

Where is the money to come from? We feel that if genetic testing is embraced by the dog fancy, market pressures will result in development of new tests, some of the profits from which will be used to fund further research. If breeders do not test for genetic disease, where will these profits come from? There are few viable sources. Most of us will look to our parent clubs and the American Kennel Club's Canine Health Foundation. Perhaps the AKC could do more. Several companies such Ralston Purina Co. already are funding genetic research and should be commended not only for the money they have donated, but for the use of their breeding records and kennels. These questions and more will be discussed in the next article of the series.
We encourage breeders to use currently available genetic tests in their selection for breeding process. Pedigrees and blood samples from known carriers of genetic disease, their parents, siblings and offspring need to be available in open registries. Debates may rage over who has the authority to pull this international effort together, but here in the United States, the AKC, the United Kennel Club and various registries such as the Orthopedic Foundation for Animals and the Canine Good Citizen program seem to be taking the lead. Individual parent clubs in some breeds actively are working the genetic disease issue.

For breeds at great risk, certain breeding strategies such as introgression, a very complex method that involves going back to the original stock and selecting for or against a particular

gene trait, should be considered. We need to contemplate opening up the studbooks. At present, studbooks are closed, and the only way to get back some of our lost genetic diversity is to breed those dogs that made up the breed originally. The AKC should have a set procedure for going back to the original foundation stock, even when there is no actual breed registry in the original country. As it now stands, each breed club is being asked to "reinvent the wheel" in this endeavor. The Samoyed, Saluki and the Basenji are perfect examples of this predicament because is seldom have nomadic tribes, from whence these dogs came, maintained written records. Some hard questions need to be asked about the validity of our pedigrees and what must be done to protect those records.

The fancy must address these problems if our beloved animals are to have a viable future. We are the custodians of our various breeds, thus the responsibility for finding those answers to these genetic problems is ours.