[

What you get, is NOT necessarily what you see!

by
Susan Thorpe-Vargas, Caroline Coile, John Cargill

Have you ever wondered at the extraordinary diversity in the appearance of the various dog breeds? How is it that a Yorkshire terrier can be the same species as a Bull Mastiff, or Pug be related to a Saluki? What are the factors that can lead to this incredible range and variety in appearance, not to mention behavior and temperament? In this second breeding genetics article, in a series of six, the authors will attempt an explanation. In the process, we will clarify some basic genetic principles at a higher level of complexity and in perhaps more detail then ever before attempted in a dog magazine. It is not simply a question of phenotype versus genotype, or dominant vs. recessive. Some of this material may be heavy going, but we hope you will persevere. The modern dog breeder, who wants to lay claim to that title, has no excuse for not using the very latest information available or taking advantage of the most recent technological advances. But we do not ask you to except these tools uncritically and it is our goal to help inform you.

.

Why so many dog breeds?

The concept of a pure breed is a relatively recent one; most early dog breeds consisted of local populations of relatively similar dogs and were bred for s specific purpose. Although there were some exceptions, the dog breeders of the time did not hesitate to breed their dog of one type to the newly arrived dog from another area that happened to be of another type. Those dog populations that did breed true, were those that lived in physical isolation. Thus, into the 19th century the various dog breeds were more often than not strains of closely related and similar looking or working dogs that as a population had a great deal of genetic diversity.

In searching for cultures without dogs since pre-historic times we came up empty-handed. Some of the earliest breeds come from the Middle East, Africa and Asia. The Middle Eastern coursing hounds had become well established by 2,000 or better BC The plains hunting dog of Africa, the Basenji, may even predate the dogs of the Pharaohs. In

the Far East, isolated Tibet and Mongolia, a number of breeds are of ancient origin. Malta was occupied as early as 3,500 B.C, and the dog brought to Malta may have had earlier Egyptian origins. The point here, is, that since relatively early times in human recorded history, there has been a tremendous diversity in dogs once they became associated with humans. Contrast the Roman Mollosus (or what we think it looked like) with the Maltese or the Tibetan Terrier and with the Lhaso Apso, and it immediately becomes obvious that there may be no "standard" dog. The tiny kingdom of Tibet, produced many different breeds, some now probably extinct, but which include the following breeds and their ancestors: Kuvasz (before Hungary); Lhaso Apso, Tibetan Terrier, Tibetan Spaniel, Tibetan Mastiff, just to mention a few familiar to Western dog fanciers.

Even dogs we think of as "English" as Mastiffs have had ancient origins. Mastiffs may be recognized on Egyptian monuments circa 3,000 BC They were in China circa 1100 BC and eventually went to England with invading Roman forces in the first century AD We can safely say that certain dog breeds have been breeding true for a very long time.

In the mid 1800s a new pastime caught the attention of the European upper class: the exhibition of dogs. Prior to this time it had been common to arrange contests between dogs based on coursing, fighting, or hauling ability, but the idea that any dog's worth could be estimated by its physical appearance was a relatively new one. It had its roots in the breeding of identifiable strains of livestock, in which visual uniformity served as an identifying trademark of an "improved" strain, and thus a source of pride. As competition among breeds grew, the search for "new and exotic" breeds from afar began. Travelers and dog buyers would spot prospective breed candidates, acquire a few, and exhibit them. The dogs thus selected represented a very small minority of the existing population of that breed. Using one of our own breeds as an example, the Samoyed was brought to England by way of the polar exploration teams. However, severe selection for an all white dog, limited the English Samoyed foundation stock to less then twenty animals. Subsequent importation would have added to these founding dogs, but at some point the studbooks were closed to additional foundation stock and the opportunity for genetic diversity was lost.

It's All in the Timing

There is a saying among dog breeders that "All puppies look alike" and it is true that newborn puppies, except for size and of course color, look remarkably alike. How is it that these puppies grow up to look so different? The vast array of physical and behavioral differences in dogs is most likely not due to selection for each individual trait, but more likely to selection for groups of traits that are all similarly affected by the same hereditary mechanisms. One such mechanism is the regulation and timing of developmental processes. Selection for one trait affected by developmental timing could inadvertently select for other traits also affected by developmental timing. The retention of immature characteristics in adults is known as **neoteny**, and it is possible that this process has played a vital role in the initial domestication and later diversification of dogs.[i] As animals mature, they pass through different stages, each uniquely adapted to its particular circumstance. In wolves, the neonates and juveniles are dependent upon parents to care for them, and they are extremely successful at eliciting that care. In comparison to adults, they are relatively tame and subservient. Wolves (or the wild ancestors of wolves and dogs) that tended to retain these immature qualities of tameness and subservience into adulthood would have been favored by early humans and would have formed the core of primitive domesticated dogs. By choosing the individuals that reproduced while still immature behaviorally, concurrent selection for other juvenile traits may very likely have occurred, laying the basis for the diversity seen in dogs. The rounded head and shortened muzzle of some breeds is reminiscent of the neonatal wolf; floppy ears, too, are a neonatal wolf trait. The dog's smaller head, brain, and teeth, in comparison to the wolf's, are comparable to those of the immature wolf. Many of the herding, hunting, and playing behaviors humans have found so useful and entertaining can be found in immature forms of hunting behavior in wolves. Barking is not seen in adult wolves, but is a trait of juveniles---as well as adult dogs.

Further crossing of dogs showing differing degrees and influences of neoteny could produce novel combinations of adult and immature characteristics, so that domestic dogs may be regarded as a blend of immature and adult characteristics. Sometimes this creates

problems for the dog breeder; as in the case of toy breeds with disproportionately large eyes. The eye seems to be relatively immune to neoteny, and is difficult to reduce in size through selection, in contrast to body and skull size, both of which have been induced to retain immature dimensions. How, then, does neoteny occur? Breeders are used to thinking of genes as being either present or absent, but in fact, a major feature of genes has to do with timing, or regulation.

The number of genes in the entire dog genome is estimated at about 30,000. Every gene **codes** for a different protein or polypeptide, but these gene products are not being made all the time nor at the same time. Regulation of expression involves turning genes on and off at various intervals and in a particular temporal sequence. Comparatively, very little is known about this process, in fact it is considered one of the hottest topics in Molecular Biology. What we do know is rather complex, but is well worth a quick survey as this information will clarify some of the concepts being discussed later in this article and future articles in the genetic series.

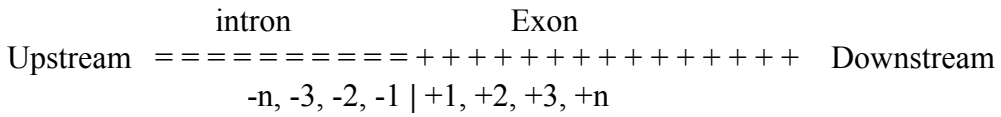
The first step in gene expression begins with **transcription.** This is the process of copying a DNA sequence called the **template**, into a single strand of RNA known as the primary transcript. This operation is initiated by an enzyme called **RNA polymerase**. [ii] Genes come in two flavors, **structural** and **regulatory.** RNA polymerase is a protein that is coded for by a regulatory gene. Transcription starts when this enzyme binds to a special region at the start of the gene called the **promoter** and continues until it reaches a **terminator** sequence. The first point of control in this process is therefore the binding of the enzyme to this specific site. Your first questions should be **HOW** does RNA polymerase recognize the promoter site and **WHAT** is it that tells it to duplicate that particular gene? Before we can answer these questions we should first talk about something called cell **differentiation**. Have you ever asked yourself why are the cells in my fingernails only producing fingernail proteins and not, lets say, eye proteins? The simple answer is that all the other genes in the cell, except those coding for fingernail proteins are somehow turned off.

In the process of maturation, a cell progressively and irreversibly becomes more committed to a certain line of development. One of the ways they think a cell can 'remember' what it has decided to be seems to depend on the chromosomes. Control of gene expression is the result of regulating transcription initiation and chromosomes play rather a unique role in this process. It is possible to visualize cellular DNA only during certain phases. Most of the time it exists in a relatively uncondensed form and it is only during this dispersed phase that transcription can occur. However, even during this stage, some parts of the chromosomes stay tightly wound up and condensed. The part that is unwound is called **euchromatin** and it is transcriptionally active. The part that stays condensed cannot be transcribed because the transcriptional factors cannot physically get to the DNA. But there are two types of inactive chromosome. One is called **constitutive heterochromatin** and it always transcriptionally inert. The other is referred to as **facultive heterochromatin** and it varies in a tissue-specific manner. So depending on which cell type it is, large blocks of chromosomes are physically prevented from being transcribed. This constitutes regulation at rather a gross level, a finer aspect of control exists in the specific sequence of the DNA it's self.

Generally, initiation of transcription is mediated by cell-specific elements. Modulation of gene expression involves these transcription factors recognizing certain particular base-pair patterns, both before and within the gene coding region. This recognition process can be compared to a plug and a socket because the DNA promoter and enhancer sequences must 'fit' the transcription elements. The better the fit, the more often transcription occurs. So transcription factors help RNA polymerase to recognize the gene to be transcribed and modulate that gene's transcription frequency. Some function by directing the RNA polymerase to the correct initiation site. Other transcription factors orient the polymerase properly so that it travels along the DNA sequence in the correct direction. Whatever their various roles, they are all essential to binding of RNA polymerase and initiating transcription. These elements are thought to also be involved in repressing gene expression. What is really amazing is that these transcription factors cooperate with each other and often act thousands of base pairs away from the initiation site. Another way gene regulation occurs can result in an entirely new protein being made or even in some

cases no gene product at all. This can happen through the selection of alternative transcription initiation sites or optional splice sites. An additional control mechanisim has been suggested by the processing of messenger RNA. It is *m*RNA that is actually **translated** into the final gene product. Whether or not messenger RNA makes it out of the nucleus so that it can be made into a protein, or how long it lasts in the cytosol before it is degraded, would definitely affect the final gene product. However, research has barely begun on these topics, so we will leave it for now to discuss another pathway to phenotypic differences.

****
Figure 1.



Reading from left to right, the DNA sequences before the start of a gene are situated **upstream** and those base pairs that lie within the gene or to the right of the transcription unit are said to be **downstream.** Those base-pairs at the start of the transcription site are numbered from left to right, +1, +2, +3....etc. The sequence before the transcription unit is numbered from right to left, -1, -2, -3....etc..

****

All Alleles Are Not Created Equal

Control of gene expression also depends on how genes interact and their alternative alleles. Because chromosomes are present in pairs, it stands to reason that the genes on them are also present in pairs. Genes in corresponding locations on homologous chromosomes are called homologous genes, and when these homologous genes can code for different proteins, they are called alleles. Sometimes we are aware of only two possible alleles for a particular gene, but often several possible alleles exist. Only two at a time can be present in one individual, however. The possibility of having either identical or nonidentical allele members of a pair creates an array of different ways these alleles can interact. Briefly, one allele can complete mask the presence of the other (complete or simple dominance), both alleles can be expressed equally (codominance), or

the end result may be intermediate between the products of the two alleles (incomplete dominance).

The gene can be considered a small business with two partners. Sometimes both partners share the same desires, just as in some case both alleles code for the same products: this is the situation with homozygous alleles. Sometimes partners, and alleles, don't agree, such as with heterozygous alleles. These cases can have several outcomes. As in any "partnership", decision making can take several forms. In some cases one partner (the dominant allele) calls all the shots, regardless of the wishes of the other (recessive allele). In genetics this is known as simple dominance. In other partnerships, compromise is the order of the day, and when the two partners are not in agreement, they settle on an intermediate solution (incomplete dominance). In yet other partnerships, both members go ahead and do what they want to do regardless of what the other does. In genetics such a solution is termed co-dominance.

Simple dominance: Dog breeders sometimes fall into the trap of assuming a trait is due to a dominant allele because" even after being hidden for generations it just popped back out…can't seem to get rid of it". In fact, they have put their finger on the signature of the recessive allele. Consider the case of black versus liver hair color. A single dominant allele B codes for black pigmentation; dogs that are either BB or Bb will be black and indistinguishable from one another. Only if two recessive alleles, bb, are present will liver coloration result. If two liver (bb) dogs were bred together, they could only produce liver offspring. If two black dogs were bred, the possibility exists that both of those dogs could be heterozygous (Bb) and produce a bb offspring that would be liver---not because the liver was dominant, but because it was recessive and thus hidden in the parents. A trait caused by a dominant allele can be traced directly from one ancestor to the next through a pedigree, although, as we will see later, other genes can also act on the dog's color to possibly modify or obscure it.

Not all traits are inherited in this manner, however. In fact, most traits do not show simple dominance.

Incomplete dominance: In contrast to simple dominance, in which two alleles produce three possible genotypes but only two possible phenotypes, incompletely dominant allele pairs produce three possible genotypes and phenotypes. The merle coat color pattern (found in breeds such as the Australian shepherd, dachshund, and collie) is an example of an intermediate phenotype created by two non-identical (M and *m*) alleles. Dogs that are *mm* have "normal" non-merle coat colors determined by genes at other locations. Dogs that are Mm display the classic merle color, in which areas of the coat have loss of normal pigmentation, resulting in the appearance of flecks of normally colored hair interspersed among lighter hair. Dogs that are MM have greater pigment loss and may be nearly white, and very often have visual and auditory problems that are pigment related. Breeders thus usually discourage merle to merle breeding, since ¼ of the progeny of a Mm x Mm breeding would be MM. Instead, taking advantage of incomplete dominance, merles (Mm) are best obtained by breeding non-merles (mm) to merles (Mm), resulting in litters consisting on average of 50% Mm merles and 50% mm non-merles. Two simple tests can determine if a trait is incompletely dominant. For one, crosses between two different parental types should always result in the intermediate type. For another, crosses between two intermediate types should result in both intermediate as well as parental types.

Co-dominance: In yet others the alleles code for products that can both be distinguished in the individual. The most common examples of this codominance are usually found in certain blood proteins expressed in both people and dogs. Perhaps the simplest and most familiar are human blood groups. In humans, three possible alleles exist: A, B and O. A and B are dominant over O, but are codominant with each other, thus resulting in AB blood type.

Penetrance and Expressivity: Just when early researchers though they had dominant and recessive inheritance clearly defined, they kept coming across cases where an allele that should have been expressed wasn't. The most obvious were in identical twins that weren't quite identical. One would exhibit a trait known to run in that family while the

other would not, yet is identical in all other respects. This is known as variable penetrance Related to this is the concept of variable expressivity, where both twins would share the same trait, but one would have a more pronounced version of it than the other. Two dogs that both carry the same alleles for spotting may have very different spotting patterns. For some reason some alleles will not always be expressed, or will be expressed to varying extents, in an individual that should normally express them. For the breeder, these two phenomena can make tracking the hereditary pattern of a trait more complicated.

Pleiotropism: Some genes affect widely disparate traits. Chinese Cresteds come in a hairless and powderpuff varieties, with the hairless caused by a single allele H. In fact, this is a homozygous lethal allele, because dogs with HH die before birth so hairless dogs are all Hh. The H allele not only results in hairlessness, but also in tooth abnormalities, which is why allowances are made for hairless Cresteds with bad bites. Because these two traits are pleiotropic effects of one allele, they cannot be separated and one must always go with another.

In addition to the interactions that occur between alleles at the same locus, interactions can also occur between alleles at different loci. Examples of traits involving different loci include the concepts of phenocopies, epistasis, and perhaps most important, polygenic effects.

Phenocopies: Sometimes two dogs will seem to share the same trait but in fact the trait is the result of totally different genes. White dogs can result from the alleles for extreme white spotting (basically a spotted dog without any spots showing) or from a dog with several alleles at different loci for factors that make the coat pale (basically a cream dog that is so pale it appears white). In many breeds progressive Retinal Atrophy (PRA) exists, sometimes appearing clinically identical. Although within each breed PRA is recessively inherited, crossbreeding affected dogs of different breeds may yield normal offspring because different genes in the two breeds cause the disease. (If an affected dog

of breed A is pp RR, and an affected dog of breed B is PP rr, then their offspring would all be Pp Rr, and appear normal).

Epistasis: Not only can alleles interact with other alleles at the same locus, but in some cases, with alleles at other loci. While dominance can be considered intralocus interaction, epistasis can be considered interlocus interaction. The simplest case of epistasis occurs when the presence of one trait effectively masks the presence of another trait. Such an example occurs with Labrador Retriever coat colors. At the B locus, the dominant B allele codes for black fur (BB or BB) and the recessive b allele for chocolate fur (bb). But at a totally different locus, E, the presence of the dominant E allows either black or brown fur (according to what is determined at the B locus), but ee restricts the formation of any dark pigment, thus resulting in a yellow dog no matter what is coded for at the B locus.

Polygenes: The problem dog breeders have with using ideas of dominant and recessive genes in breeding dogs is that most traits don't appear in discreet intervals, but instead are continuously distributed over a range of values. For instance, dogs don't come in just short, medium, and tall, they come in all sizes. Even within a breed, height is normally distributed in a bell curve. This is because many important traits are the result of many pairs of genes acting together. In these cases, the extent of a trait is determined by gene dosage, which is the number of particular alleles present in a genotype. Imagine that height is controlled by incompletely dominant alleles at three different loci, A, B, and C, with A+, B+, and C+ all coding for an additional half inch of height. A dog with the genotype A+A+, B+B+, C+C+ would be three inches taller than one with the genotype A-A-,B-B-,C-C-. In fact, 27 different genotypic combinations are possible in this example, resulting in seven different heights. The more loci involved the greater the number of possible genotypes and phenotypes, until the phenotypes become so numerous that they appear to be continuously distributed. This blending is further influenced by environmental factors. Hip dysplasia is thought to be polygenic.

Linkage and Linkage Disequalibrium: In a highly inbred population genetic defects can become fixed rather rapidly if they happen to be on the same chromosome as a gene that codes for a desirable trait. The closer they are physically on the chromosome the tighter they are 'linked'. These genes and their respective alleles will be inherited together unless they become 'unlinked' in a procedure called **crossing over** or **recombinatio**n. This is a process that occurs during the formation of gametes, **whereby** homologous chromosome pairs exchange segments of their DNA structure. Such closely linked genes are said to be in a state of linkage disequilibrium. When a breeder selects for or against a specific gene trait, he or she is also choosing those traits or not, which are located on the same chromosome. One should remember this when making a breeding decision. Severe selection pressure against an unwanted trait, could result in throwing the baby out with the bathwater and the permanent loss of a necessary or desirable attribute.

Sex Linkage: A special case of linkage exists when genes are located on the sex chromosomes. Unlike the other 38 pairs of chromosomes, the sex chromosomes are not always paired in a homologous fashion. This is simply because sex is determined by whether an individual has two X-chromosomes (XX= female) or an X and a Y chromosome (XY=male). The Y chromosome is a very small chromosome and until recently there were doubts that any significant information was contained on it. The X chromosome is larger and is known to carry on it genes that code for several important traits. Genes on the X chromosome are not matched by genes on the Y chromosome, negating the possibility of allelic pairs. In the male, whatever alleles are on his single X-chromosome will be expressed (a condition known as hemizygous). In the female, the situation is still a little different from what is seen in the other autosomal (non-sex) chromosomes.

For many years it was assumed that these X-linked alleles acted just the same as autosomal alleles. But they don't. Instead of acting in a standard dorminant-recessive way, these alleles act more like codominant alleles. In placental mammals one of the two X-chromosomes is randomly inactivated in each cell of the body. The remnants of these inactivated chromosomes can be seen as dark spots (Barr bodies) in almost every cell of a

normal (XX) female, but not in normal (XY) males. Very early in embryonic development both X-chromosomes are apparently active, but then one of the two is rendered dysfunctional by staying tightly condensed in the heterochromitin state. It is entirely a matter of chance whether it is the paternal or maternally derived X chromosome, but once inactivated, all subsequent cells derived from that cell will continue to have the same inactivated X chromosome. In individuals with visible sex-linked traits the results can be clearly seen as patches of paternally and maternally derived traits. Thus all female mammals are mosaics. The best known example is the calico cat, which is almost always females (the few males are abnormal XXY individuals) and which displays a patchwork of black and orange colors, each patch representing a clone of an original cell that randomly inactivated either the X chromosome with an allele for orange fur or a the X chromosome with the allele for black fur. In dogs, we have to look a little more carefully for such evidence. Examples include X-linked muscular dystrophy in Golden retrievers and X-linked hereditary nephritis. Because these female carriers are mosaics for the abnormalities seen in these diseases, they may exhibit attenuated signs of the disorder, with the severity depending upon the proportion of the mosaic derived from the X-chromosome carried the abnormal allele.

Sex-linked traits will be passed from dams to sons via one of her X chromosome. Because sires can only pass their X-chromosomes to their daughters, in order for the trait to be fully expressed in a female she must have an affected sire and carrier (or affected) dam. The degree of mosaicism that the dam expresses is random and does not affect the chances of her offspring being affected or the severity of that trait if affected.

Misunderstandings about sex-linked inheritance have given rose to many breeding myths, the most widespread of which place greater emphasis on the "sire line" (sire to grandsire) in the belief that "what you see is what you get" due to the single X-chromosome, as well as the belief that important breed attributes are carried on the Y chromosome; or which contend that whether an ancestor is on the dam versus the sire's side of the pedigree is of prime importance. These theories neglect the fact that the Y-chromosome contains little,

if any, identified genes apart from those involved with male reproduction, and that the sex chromosomes are but one of 39 pairs of chromosomes. These ideas served the 19th century breeder well, but they have no place in the 21st century breeder's arsenal.

Conclusion

So the variety of appearance between dogs of different breeds is controlled at several different levels. Some types of expression seems to depend upon turning control/regulatory genes on and off so that a specific developmental cascade is expressed. Other phenotypic differences must rely upon the interaction of genes, their various alleles and where these hereditary units are located on the chromosomes. Hopefully, the modern breeder will be able to use this knowledge to make more informed choices when planning a breeding or to understand why certain breeding decisions went awry. In the next article in this series we will discuss the techniques and concepts used in the physical and lingage mapping of the canine genome. We will cover such subjects as conserved sequences, synteny and homology - how the mouse and human genome projects will help us in our efforts to find the genetic basis of the diseases that afflict our dogs and why the canine genome project is so very important to the future of our canine companions.

---

[i] R Coppinger & R Schneider: Evolution of working dogs. In: JSerpell (Ed): The Domestic Dog: Its evolution, behaviour, and interactions with people. Cambridge University Press. Cambridge. 1995.

[ii] There are three classes of genes and each class is transcribed by a different RNA polymerase.