University of Kentucky Doctoral Dissertations   Graduate School

2006

STUDIES OF GENETIC VARIATION AT THE KIT LOCUS AND WHITE SPOTTING PATTERNS IN THE HORSE

Samantha Ann Brooks
University of Kentucky, Samantha.Brooks@uky.edu

Recommended Citation

ABSRACT OF DISSERTATION

Samantha Ann Brooks

The Graduate School

University of Kentucky

2006

STUDIES OF GENETIC VARIATION AT THE *KIT* LOCUS AND WHITE SPOTTING PATTERNS IN THE HORSE

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the

requirements for the degree of Doctor of Philosophy in the

College of Agriculture

At the University of Kentucky

By

Samantha Ann Brooks

Lexington, Kentucky

Director: Dr. Ernest Bailey, Department of Veterinary Science

Lexington, Kentucky

2006

ABSTRACT OF DISSERTATION

STUDIES OF GENETIC VARIATION AT THE *KIT* LOCUS AND WHITE
SPOTTING PATTERNS IN THE HORSE

There are numerous different white spotting patterns in the horse, including two of particular interest tobiano and sabino. In the mouse, genetic variation in the gene *KIT* causes many white spotting patterns. Due to the phenotypic similarity among white spotting patterns in horses and mice, *KIT* was investigated as the cause of the tobiano and sabino spotting patterns in horses.

Initially, the *KIT* cDNA sequences from horses with several spotting patterns were compared. Three single nucleotide polymorphisms (SNPs) were identified, though none were associated with a spotting pattern. Three novel splicing variants were also observed: exon 17 skipping, exon 18 skipping and alternative splicing of exon 3.

Families segregating for a sabino spotting pattern (designated *Sabino 1*) and exon 17 skipping were discovered. Sequencing revealed a SNP (KI16+1037) within intron 16 that was completely associated with skipping of exon 17. Using a PCR-RFLP for KI16+1037, linkage was discovered for sabino spotting (LOD=9.02 for Θ=0) and presence of the *Sabino 1* allele detected in seven breeds. While all horses with this SNP exhibited the *Sabino 1* phenotype, some horses with a sabino phenotype did not possess the SNP. This is most likely due to genetic heterogeneity of the phenotype.

Fluorescent *in situ* hybridization (FISH) was used to investigate the possibility of chromosome inversion in the region of KIT. A chromosomal inversion was discovered spanning ECA3q13 to 3q21 using BAC clones containing *KIT* and other genes in the same region. The ECA3q inversion was completely associated with *Tobiano* in the eight horses tested by FISH. This inversion may disrupt regulatory sequences of the *KIT* gene and thereby cause tobiano spotting.

Spotting patterns are important to horse breeders for aesthetic as well as economic reasons. Spotting patterns in the horse may also be an interesting scientific model. The two genetic variants discovered in this work are good examples of genetic diversity due to mechanisms other than SNPs. Study of these variants may be valuable for examining the effects of the *KIT* gene on health traits. In particular, the *KIT* gene directs many functions of the mast cell, a cell that is involved in the etiology of inflammation.

Samantha Ann Brooks

May 3, 2006

STUDIES OF GENETIC VARIATION AT THE *KIT* LOCUS AND WHITE SPOTTING PATTERNS IN THE HORSE

By

Samantha Ann Brooks

Ernest Bailey

Director of Dissertation

Barry P. Fitzgerald

Director of Graduate Studies

May 3, 2006

RULES FOR THE USE OF DISSERTATIONS

DISSERTATION

Samantha Ann Brooks

The Graduate School

University of Kentucky

2006

STUDIES OF GENETIC VARIATION AT THE *KIT* LOCUS AND
WHITE SPOTTING PATTERNS IN THE HORSE

DISSERTATION

A dissertation submitted in partial fulfillment of the

requirements for the degree of Doctor of Philosophy in the

College of Agriculture

at the University of Kentucky

By

Samantha Ann Brooks

Lexington, Kentucky

Director: Dr. Ernest Bailey, Department of Veterinary Science

Lexington, Kentucky

2006

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: Background and Review

Introduction

Most mammalian species exhibit variants in coat color and spotting pattern. From spotted domestic cats and dogs, to humans with a white hair tuft in their widows peak, one only needs to look around daily life to appreciate how common spotting patterns are. The horse is no different. In fact for many breeders spotting patterns have become big business. Horse owners value distinctive coloration, and in this case beauty is firmly lodged in the wallet of the beholder. Buyers will pay a hefty premium for an attractively marked horse. Spotting patterns, from the mild to the bizarre, have become a popular way to stand out in the "herd".

The forerunner of genetics, Gregor Mendel, chose color variants in the pea as one of his primary traits of study in the mid-nineteenth century. Thomas Hunt Morgan used fruit flies with mutations for eye and body color to demonstrate the theory of associative inheritance. Subsequently, when scientists wished to learn if similar laws governed inheritance in mammals they investigated color in the mouse. Since then, the mouse has established itself as the mammal of choice for study due to is small size, inexpensive upkeep, short generation interval, and the ease with which it can be genetically manipulated. Many spotting and coat color mutations have been found in the mouse, some with pleiotropic effects (traits acting on a different physiological system but with the same causative allele) important for health. Indeed, many disease models in mice were first noticed by the "white flag" marking them in the litter.

In contrast, the horse is not an ideal organism in which to experimentally study inheritance due to its low reproductive rate and expensive costs of upkeep. However there is ample genetic variation in the horse to study. At the same time, any novel mutation must not affect the fundamental ability of the horse to perform his job. Sickly or unsuitable individuals are simply too expensive to maintain. This set of circumstances selects against spotting patterns associated with serious disease traits while selecting strongly for spotting patterns without adverse health consequences.

In connection with this dissertation the tobiano and sabino spotting patterns were investigated in horses. *KIT* was a strong candidate for these spotting patterns based on

comparison to known *KIT* variants in the mouse. The overarching hypothesis for this work is that the *KIT* gene is responsible for these spotting patterns. The *KIT* gene encodes the mast cell growth factor receptor and is involved in directing the development and maintenance of melanocytes, gametes and haemopoietic cells. Work on *KIT* can yield valuable information on the function of this gene in maintaining the overall health and fitness of the horse. With that in mind the specific aims undertaken in the succeeding chapters are as follows:

- Chapter 2: Sequence the coding sequence of the *KIT* gene in order to identify any genetic variation associated with several spotting patterns.
- Chapter 3: Investigate a SNP, connected to skipping of *KIT* exon 17, for linkage to *Sabino* in Tennessee Walking Horse families and association in the population.
- Chapter 4: Characterize a chromosome inversion associated with *Tobiano*.

This chapter is intended to provide the necessary background material to gain a basic understanding of the topics in the following chapters. Though by no means completely comprehensive, the more relevant concepts and previous research has been carefully highlighted. The first section in this chapter describes the most common spotting patterns in the horse, their identifying characteristics and previous study. The second section describes principles of genetics and important genomic research critical to this work. The final section reviews the function and variation of the *KIT* gene and its product, the mast cell growth factor receptor.

Section I: White Spotting Patterns in the Horse

Tobiano

The name "tobiano" originated in 19th century South America (Sponenberg, 2003). The rescue of Buenos Aries from the Spanish by a revolutionary cavalry division mounted mostly on spotted horses with the tobiano pattern and commanded by General "Tobías" forever linked the name of "Tobías" with the spotting pattern (Sponenberg, 2003). The Tobiano pattern has a distribution among a variety breeds world wide including the American Paint Horse Association (APHA) and Pinto Horse registries, as well as a variety of gaited horses, draft and pony breeds (Brooks *et al*., 2002). Nationally, the APHA (registering a total of nearly a million horses in the US) recorded about 60% of its spotted horses as Tobiano or "Tovero" (tovero is a combination of tobiano and another spotting pattern usually overo or sabino) (S. Imhof, APHA, personal communication, August 2005.)

The tobiano pattern is characterized by depigmented patches of skin (and the associated hair,) usually crossing the dorsal midline and covering the legs (Figure 1.1). The head of a tobiano patterned horse often retains the base coat color. Expression of *Tobiano* (*TO),* is a dominant trait and the phenotype can vary from only four obvious white stockings to the extreme of 90% white across the body (American Paint Horse Association, 2005). Homozygous *Tobiano* horses often have dime to quarter sizes spots interspersed among the larger white patches. Termed "ink spots" or "paw prints," they are not a reliable indicator of zygosity (Sponenberg, 2003).   Figure 1.1 shows a typical homozygous *Tobiano* horse; note the "paw print" spots in the white area of the neck.

Figure 1.1: A young mare homozygous for *Tobiano.*

Linkage of *Tobiano* to *Albumin (ALB)* was first reported in a pony family (Trommershausen-Smith, 1978). Further investigation revealed linkage disequilibrium between the *Tobiano*, *Albumin* and *Vitamin D Binding Factor (Gc)* loci. Specifically, the *TO* allele was strongly associated with the presence of the *AL-B* and *Gc-S* alleles (Bowling, 1987). The association was termed "phase conservation" and the frequency of the *TO:ALB-B* linkage phase subsequently measured at 0.92 among individuals homozygous for the *TO* allele (Duffield and Goldie, 1998). Bowling theorized that a chromosome inversion may have caused the phase conservation of *TO:ALB-B:GC-S.* Furthermore she stated that, as there are many horses with the *ALB-B* and *GC-S* allele who do not have *Tobiano,* it is not possible that a biochemical interaction between the two causes *Tobiano* (Bowling, 1987). Subsequently, Raudsepp and coworkers investigated the presence of an inversion using G-banding and florescent *in situ* hybridization (FISH) with four gene markers from ECA3: *KIT, PDGFRA, ALB,* and *MC1R* (Raudsepp *et al*., 1999). However no rearrangement in the order of markers was detected.

The phase conservation *TO:ALB-B:GC-S* has, for many years, been used to predict the presence of *Tobiano*. Essentially, horses homozygous for *Tobiano* are usually homozygous for *ALB-B:GC-S* as well. However, due to the fact that *TO* occasionally

exists in phase with the *ALB-A* and/or *GC-F* alleles, these typing methods are not precise indicators of *Tobiano* zygosity. We recently found a SNP within the *mast/stem cell growth factor receptor* (*KIT*) gene that was also in phase with *TO*. This marker, KM1, has the advantage over blood typing markers in that the SNP has been found in all horses with *Tobiano*. The marker was infrequently associated with the *non-tobiano* allele as well (frequency of 0.03 among Thoroughbreds), consequently it is clearly not the cause of tobiano (Brooks *et al*., 2002).

Frame Overo/ Overo Lethal White Foal Syndrome

The pattern described as "frame overo" (*Frame* allele) is phenotypically similar to tobiano and sabino spotting patterns. *Frame* produces a white spotting pattern characterized by white patches on the midsection that usually does not cross the dorsal midline. This leaves the legs, hooves, back, belly, chest and hindquarters of the horse pigmented according to the base coat color, producing a "framed" effect. The head and eyes are often depigmented. The expression of the frame overo pattern can vary from white patches covering more than 90% of the body surface to virtually no white spotting.

However, a recent report described the *Frame* allele among 18% of non-spotted horses registered with the American Paint Horse Association (Santschi *et al*., 2001). Frame overo patterning can also appear spontaneously in breeds that restrict their registries to non-spotted horses (Sponenberg*,* 2003). This is probably due to the presence of minimally marked horses that carry the gene but do not express the typical phenotype, a situation described as incomplete penetrance. Some horses with *Frame* who have a minimal phenotype can appear similar to the *Sabino* phenotype and it can be difficult to distinguish between a minimally marked frame overo and a sabino-type pattern. In other words, the range of gene expression overlaps between *Sabino* and *Frame*. To further confuse the issue it is customary in the United States to use the term "overo" to describe all spotting patterns except *Tobiano*. Nevertheless, each of these patterns has an independent set of characteristics, mode of inheritance, and at least among those patterns studied so far, genetic origin.

Figure 1.2: The frame overo spotting pattern.

Frame overo has long been associated with Overo Lethal White Syndrome (OLWS). While a single copy of the *Frame* allele results in a spotting pattern, homozygosity for *Frame* causes OLWS (McCabe *et al*., 1990). Foals with OLWS are distinguished by a complete lack of body pigmentation and intestinal aganglionosis (Santschi *et al*., 2001). This lack of intestinal innervation results in impaction of the bowls and death before the foal reaches several days in age (Santschi *et al*., 2001). This syndrome is similar to Hirschsprung's disease in humans, which is caused by mutations in the *endothelin receptor type B* gene (*EDNRB)* (Santschi *et al*., 1998). The phenotypic similarity between Hirschsprung's disease and OLWS led to the investigation of this gene as a candidate for OLWS. An *EDNRB* mutation causing frame overo/OLWS was reported in 1998 by three individual laboratories as a dinucleotide substitution in *EDNRB* (Metallinos *et al*., 1998; Santschi *et al*., 1998; Yang *et al*., 1998). However, the location within the gene for the OLWS mutation is different than those associated with Hirschsprung's disease. A DNA test based on the substitution is commercially available and is used extensively to identify carriers and prevent production of foals carrying OLWS. A summary of those labs offering genetic testing services for the horse and their

contact information is listed in Appendix A. So far only one mutation has been associated with OLWS/*Frame*. This suggests that *Frame* originates from a single mutation event in a founding individual.

Sabino-type patterns

"Sabino" is a term of Spanish origin. Literally translated it means "roan" or "pale red", which would seem to describe the roaning effect sabino has on a chestnut base color. Sponenberg described Sabino as a pattern typified by extensive white on the legs and face (Sponenberg, 2003). Body white generally begins at the belly, and has an "upwards" progression. Expression can vary from a horse with an inconspicuous blaze (white strip down the face) and socks (white markings on the lower legs) to completely white.

Figure 1.3: A Belgian with the Sabino-type pattern common in draft horse breeds.

Sabino inheritance has been described in two small studies summarized by Sponenberg in the appendix of his book; one indicating sabino is polygenic, while the other implying sabino is due a single dominant gene. In the first study, photographs of 5 APHA registered sabino stallions and their offspring from non-spotted dams were used to demonstrate that for 4 out of 5 of the stallions the production of spotted foals was statistically unlikely to be due to a single gene trait (Appendix 4, page 175 Sponenberg,

2003). In the second study which examined breeding records from two stallions, each carrying both tobiano and a sabino-type pattern, both demonstrated that the sabino-type pattern was linked to *TO* and due to only one gene (Appendix 4, page 176 Sponenberg, 2003). The conflicting conclusions on inheritance of sabino drawn from these two studies, as well as the broad expression of sabino-type patterns, supports the idea that the term "sabino" encompasses multiple patterns with varied genetic origins.

The work in chapter three of this volume describes one such sabino-type pattern, *Sabino 1,* controlled by a single gene and nearby to markers linked to *TO*. Heterozygotes for this allele are marked with the traditional sabino-type spotting, and expression varies from 10%-75% of the body depigmented. Homozygotes for *Sabino 1* are completely or nearly completely white. Family studies demonstrated linkage of this pattern to an intronic SNP in *KIT* and associated with exon skipping within the *KIT* gene (Brooks and Bailey, 2005).

Figure 1.4: The dam in this photo shows the typical moderate expression of *Sabino 1* in the heterozygote. The foal is a *Sabino 1* homozygote. (Photo courtesy of the owner.)

Splashed White

Splashed white is another pattern with characteristics similar to sabino patterns. Typically, horses with splashed white have extensive white areas on the head, white horizontally distributed across the lower body, and often have one or more blue eyes

(Sponenberg, 2003). The margins of the white areas tend to smoother and better defined than in sabino-type patterns. Some anecdotal evidence suggests an association between the splashed white pattern and deafness (Sponenberg, 2003). The gene for splashed white is not known, nor has it been mapped, though limited studies have suggested it is controlled by a dominant gene and unlinked to the *Tobiano* locus (Appendix 4, page 176 Sponenberg, 2003).

Figure 1.5: A splashed white pony. (Photo courtesy of the owner.)

Dominant White

Though not a typical spotting pattern, "white" in the horse deserves mention as it is easily confused with other phenotypes. Despite the fact that white was originally described as a solid white body with both skin and hair lacking pigment, some studies include incidents of "incomplete penetrance" where some pigmented areas remain across the top line (Mau *et al*., 2004; Pulos and Hutt, 1969). It is also possible to see a similar completely white phenotype through the action of other genes, most notably *Sabino 1*, as well as by means of an additive effect of several different spotting patterns in the same individual. *White* can also be easily confused with similar phenotypes caused by the *Cream* and *Grey* loci.

9

Figure 1.6: A white thoroughbred.

The occurrence of completely white horses was reported in 1912 (Sturtevant, 1912). Since then there has been little agreement as to the genetic heritability of the trait. Early work suggested the possibility that *W* was recessive in some cases and dominant in others (Guerts, 1977). A family study concluded that white (*W*) in horses, like many *W* alleles in the mouse, is dominant lethal based on the occurrence of non-mendelian phenotypic ratios in the offspring of white parents (Pulos and Hutt, 1969). The authors suggested the altered ratios were due to the loss of offspring homozygous for white at an early embryonic stage. Operating under the premise that white is in fact dominant lethal, Mau *et al.* found a weak linkage ($Z=3.04$; $\Theta=0$) between *White* and the region near *KIT* in the Franches-Montagne breed (Mau *et al.*, 2004). However, neither study seems to have adequately controlled for the presence of other spotting patterns within the families, potentially complicating the phenotypic classification of offspring.

Section II: The Equine Genome

Comparative Genomics

The equine genome consists of 31 pairs of autosomal chromosomes and two sex chromosomes. The complete genome of the horse may span about 3740 cM, or approximately 3740 million bases (Penedo *et al*., 2005). This is just slightly larger that the human genome which spans 3076 millions bases (Karolchik *et al*., 2003b). Though the equine genome is not yet fully sequenced, both physical [FISH and radiation hybrids (RH)] and genetic (linkage) maps have been constructed (Chowdhary *et al*., 2003; Milenkovic *et al*., 2002; Penedo *et al*., 2005).
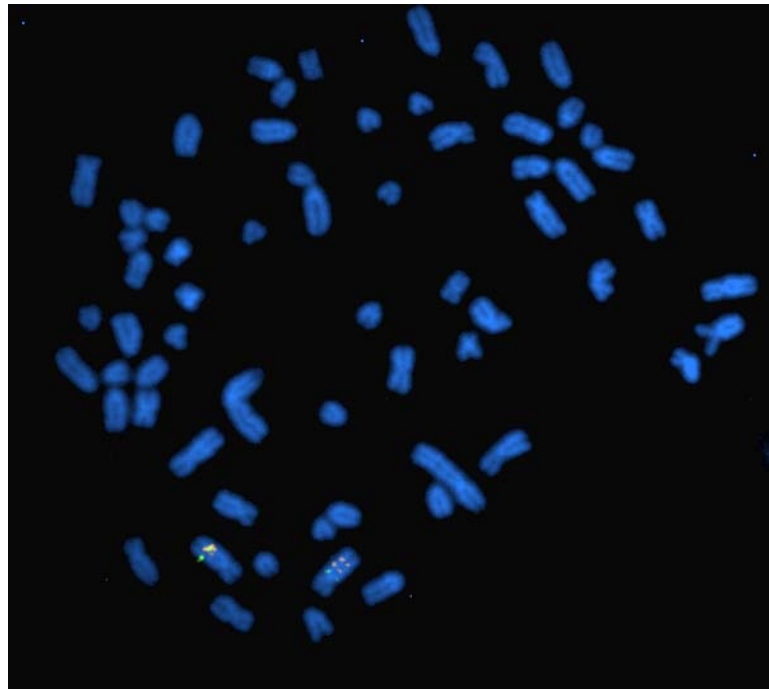
Figure 1.7: Equine metaphase chromosomes labeled by FISH.

Comparative mapping (identifying differences in homology between the maps of two species) illustrates the genomic rearrangements that occur as species diverge over time (Nadeau and Sankoff, 1998). Comparative mapping can also be a useful tool to those studying species with less well characterized genomes, as is the case in the horse (Brown, 1999). The equine genome maps have provided a wealth of information on the relationships between the horse and other species. Comparison of the horse and human

11

genomes by FISH mapping showed that more than a hundred conserved chromosome segments exist between the horse and human genome (Milenkovic *et al*., 2002). The radiation hybrid map found a similar number of conserved regions between horse and human (Chowdhary *et al*., 2003).

One application of comparative genomics is the use of the gene sequence and genomic organization of one species to predict that of another species (Brown, 1999). The use of this strategy in gene hunting has already greatly benefited horse research. Hyperkalaemic periodic paralysis (HYPP), a disease causing muscle weakness, tremors and a temporary (periodic) paralysis, was the first genetic disease in horses traced to a causative mutation. An identical disease exists in humans for which the mutation is known, leading investigators to sequence a gene important to the function of sodium channels in the horse. This led to the discovery of the causative mutation, enabling breeders to identify affected individuals, manage them to avoid episodes, and limit the transmission of the gene to the next generation (Rudolph *et al*., 1992). Severe combined immunodeficiency (SCID) is a homozygous lethal condition in which affected foals lack the ability to produce antibodies and die from infection. Based on a similar condition in the mouse, researchers identified a deletion in the gene for an enzyme essential to the assembly of the genes required for antibodies (Shin *et al*., 1997). As previously mentioned, the mutation for OLWS was also found based on comparative genomics to a similar human disease (Metallinos *et al*., 1998; Santschi *et al*., 1998; Yang *et al*., 1998). Finally, the mutation for a recessive lethal glycogen storage disease (GSD IV) in the horse was found in the glycogen branching enzyme (*GBE1*) gene after similar cellular and biochemical characteristics were compared with the human form of the disease (Ward *et al*., 2004). A comparative approach has been successfully used to identify the genetic origins of several traits in the horse.

Linkage, mapping and haplotypes

The term "linkage" was coined by Thomas Hunt Morgan in 1910 who was later awarded the Nobel Prize for his work on inheritance (Rieger, 1968). Linkage refers to the physical relationship of two loci when they reside on the same chromosome. The smaller the physical distance between genes on a chromosome, the lower the probability that recombination will occur between them during meiosis and therefore the stronger the

linkage. Consequently, alleles at closely linked loci are more likely to be inherited together, in coupling, than randomly assorted. Linkage, measured by the frequency of recombination between loci, can be used to linearly order loci. In this way a linkage map can be constructed describing the order of loci along the chromosome.

Linkage results in the formation of haplotype blocks, which are groups of alleles between which recombination occurs more rarely and thus tend to be inherited as a collection. Beyond the association of haplotype blocks is a state of "phase conservation". Phase conservation exists when a haplotype block of alleles is maintained in "linkage disequilibrium". Linkage disequilibrium describes a situation in which a strong association is present that can not be explained by linkage alone. Phase conservation can be caused by either biochemical or physical relationships that restrict viable recombination. Physical relationships can include chromosome rearrangements such as deletions, insertions, translocations and inversions.

Section III: Biology of the KIT gene and melanocytes

In the early years of genetic study of the mouse, white spotting patterns were a staple trait for studies of heredity. The ease with which white spotting patterns could be identified, even at an early age, provided an ideal research tool for following the inheritance of these mutations. In his book on coat color variations in mice, Willys Silvers describes *White (W, KIT)* as the coat color locus that has received the most attention over the years (Silvers, 1979). He ascribes this recognition to the fact that the gene causes not only spotting but also various health defects. First described in 1908 (Silvers, 1979), the locus has been known for a long time.

KIT was first identified as the transforming oncogene (then called *v-kit)* of a feline sarcoma virus (Besmer *et al*., 1986). Studies of the genomic sequence and organization of the *KIT* gene have since revealed that it is a member of a highly conserved family of tyrosine kinase receptors (Andre *et al*., 1992). These receptors are structurally characterized by three functional segments; an extracellular region comprised of five immunoglobulin-like binding domains, a transmembrane domain, and an intracellular catalytic kinase region (Gokkel *et al*., 1992). Comprised of 21 exons, the genomic sequence spans 82kb in the human (Karolchik *et al*., 2003a).

The extracellular portion of the receptor functions in binding of its ligand, mast cell growth factor *(MGF,* also know as stem cell factor, *SCF).* MGF exists in two isoforms distinguished by alternative splicing of exon six; a membrane bound form and a soluble form (Philo *et al.*, 1996). The extracellular portion of the KIT receptor can also be cleaved to produce a soluble form capable of binding MGF (Dahlen *et al.*, 2001). The soluble form of the receptor functions as a competitive inhibitor of signaling through the MGF/KIT pathway (Dahlen *et al.*, 2001). Signal transduction occurs via the assembly of a complex containing a MGF dimer and two KIT receptors (Philo *et al.*, 1996). Ligand binding results in activation of the kinase domains, autophosphorylation, and initiation of signal transduction through several intracellular pathways. The receptor is then ubiquitinated, targeting it for transport to the lysosome and destruction (Blume-Jensen *et al.*, 1991). This prerequisite dimerization of the receptor adds another dimension to the variability of *KIT* phenotypes. The presence of abnormal receptors can affect the function of any remaining normal receptors they dimerize with.

Signaling through the KIT receptor is essential for normal pigmentation. Melanocytes, the pigment producing cells, originate in the embryo from precursor cells found in the trunk region of the neural crest. These precursors (melanoblasts) migrate out to the periphery as the neural tube closes. Melanoblasts then proliferate to cover the growing body and take up their ultimate position in the skin and hair follicle (reviewed by Erickson, 1993). Signaling by KIT has been shown by many experiments to be obligatory to melanoblast development. Induction of apoptosis has been observed in melanoblasts treated with anti-KIT antibody (Ito *et al.*, 1999). Mature melanocytes fail to proliferate when incubated with *KIT* antisense oligonucleotides (Spritz *et al.*, 1994). Work in W^v/W^v mice has shown that this mutation impairing the function of KIT results in the disappearance of developing melanocytes by the time embryos reach the age of eleven days (Mackenzie *et al.*, 1997). KIT expression has also been shown to accompany differentiation and migration of melanoblasts from small stem cell reserves in the developing hair follicle (Peters *et al.*, 2002). These same stem cells survive independent of KIT signaling until they begin to differentiate into melanocytes (Nishimura *et al.*, 2002).

Many genetic variants of *KIT* have been identified in a variety of species, most notably mice, pigs and humans.   In the mouse, at least 87 different phenotypic alleles have been described at the *W* locus (Blake *et al*., 2003; The Jackson Laboratory). Those phenotypes of known genetic origin can be broadly categorized by the nature of the mutation. Mutations in the *KIT* coding sequence tend to produce patterns that restrict pigmentation in a directional fashion by reducing the ability of melanocytes to successfully complete migration from the neural tube to the appendages (reviewed in Nishimura *et al*., 2002). These patterns are distinguished by depigmentation of the extremities and ventral abdomen (Geissler *et al*., 1981). A second category of alleles encompasses those caused by mutations affecting the regulation of *KIT*.   The patterns produced by this type of mutation have patches of color with smooth boundaries and depigmentation across the body, primarily involving the back as well as the extremities (Berrozpe *et al*., 1999).

Intercellular signaling through the KIT receptor is important for a variety of processes aside from melanogenesis, including haematopoiesis and gametogenesis. Interestingly, the system affected varies with the location of the *KIT* mutation. While one mutation might produce sterility, mild anemia, and a complete lack of pigment, another might result in fertile mice with severe anemia and a spotted coat. For example, W^{41}, a very mild phenotype, is characterized by mottled coat and mild anemia in the homozygote. In contrast, W^{42} is lethal in the homozygote while heterozygotes have very severe macrocytic anemia (Geissler *et al*., 1981). A unique *KIT* varistion is responsible for the familiar pink-skinned phenotype of domesticated pigs. In this case the entire gene has been duplicated, or even triplicated, likely resulting in abnormal gene regulation. Though they entirely lack pigment, these pigs are fertile and have only very mild anemia (Marklund *et al*., 1998). Many *KIT* mutations are found in humans, though fortunately most are mild and result in piebald spotting across the forehead and extremities (Spritz, 1994). Mutations that eliminate production of the complete protein (frameshift mutations for example) are often more mild in heterozygotes than those that produce a defective protein. Pairing of a normal receptor and a defective one will result in an inactive dimer (Philo *et al*., 1996). The defective receptor essentially ties up an otherwise useful

receptor, resulting in only 25% of the functional receptor dimers compared to normal, rather than the 50% available when no defective receptor is present.

The *KIT* receptor also plays an important role in several specialized cell types. Development of interstitial cells of Cajal, a type of pacemaker cell in the intestine is also directed by *KIT* (Torihashi *et al.*, 1995). Spermatogenesis and oogenesis are both disrupted without proper *KIT* signaling (Berrozpe *et al.*, 1999). As part of its function in haematopoiesis *KIT* directs expansion of mast cells, an important element of the innate immune system (Tsujimura *et al.*, 1993). Deregulation of the *KIT* receptor is inherent to several cancers including mastocytoma, gastrointestinal stromal tumors, melanoma and some types of leukemia.

Several specific mouse *KIT* alleles show striking similarity to the tobiano and sabino spotting patterns in the horse. Due to the phenotypic similarity of these murine spotting patterns and the spotting patterns of the horse, *KIT* is an ideal candidate for equine spotting genes.

16

CHAPTER TWO: Sequencing the *KIT* Gene from Horses with White Spotting

Summary

Due to similarities between alleles of the *KIT* gene in the mouse and white spotting phenotypes in the horse, the equine *KIT* coding sequence was investigated for variations that might be associated with several horse spotting patterns. Previous work has linked *KIT* to four spotting patterns in the horse (roan, tobiano, white and possibly sabino) though causal mutations for these patterns have yet to be discovered. The work in this chapter describes sequencing studies on the *KIT* gene, discovery of three novel single nucleotide polymorphisms (SNPs), and the investigation of their relationship to white spotting patterns in the horse.

Introduction

Several coat spotting patterns exist in the in the horse, each with its own set of characteristics. Many of these patterns resemble the phenotypes produced by the *White (W)* locus in the mouse. *W* has been placed at the *KIT* gene, which encodes the mast cell growth factor receptor (Geissler *et al*., 1988). This receptor is pivotal to the differentiation, migration, proliferation and survival of melanocytes, the pigment producing cells in the skin (reviewed in Ashman, 1999). While some *KIT* mutations in the mouse lead to negative health effects like anemia and sterility, these effects are specifically related to the location and nature of the mutation (Geissler *et al*., 1981). Spotting patterns in the horse have not yet been associated with the types of health defects seen in some of the more severe *KIT* alleles in the mouse.

Previous work in the horse has shown that the gene for tobiano spotting (*TO*) is linked to an intronic SNP in *KIT* as well as other genes near that region of ECA3 (Brooks *et al*., 2002; Bowling, 1987). So far all tobiano horses typed for the intronic SNP possess the variant (*KM1*) associated with *TO*. However, this SNP is not responsible for *TO* due to the observation of non-tobiano horses also possessing the *KM1* allele (Brooks *et al*., 2002). Limited family studies have also shown that at least one type of sabino spotting is linked to *Tobiano* and therefore near *KIT* (Appendix 4, Sponenberg, 2003). Linkage studies associated *White* with the *KIT* gene as well (Mau *et al*., 2004). The *KIT* gene has been previously cloned and sequenced in connection with a study of the roan coat color in horses (Marklund *et al*., 1999). Roan is characterized by the interspersion of white hairs among the colored ones. The face and extremities tend to have less white hair than the midsection. Five SNPs and several splice variants were identified but none were thought to be responsible for all roan phenotypes. Nevertheless, linkage of *KIT* and *Roan* and co-segregation within families was demonstrated using the discovered SNPS.

Materials and Methods

Horses

Fifteen horses from ten different breeds were chosen for sequence comparison of the *KIT* gene. Spotting patterns for each of the horses were determined by comparison

with the descriptions published by Sponenberg (Sponenberg, 2003). The horses comprised five groups based on spotting pattern; non-spotted, Sabino-type, Tobiano, Splash, and White. These spotting patterns were selected based on similarities to phenotypes in the mouse known to be caused by variations in the *KIT* gene.

RNA Extraction

RNA was isolated from whole blood using two methods. The first used the Purescript RNA purification kit (Gentra Systems Inc., Minneapolis, MN) following the manufacturer's directions for 10mL whole blood. For samples received by mail or which could not be processed within 6 hours, we used the PAX gene blood RNA kit (Qiagen Inc., Valencia, CA) for RNA isolation. The manufacturer's protocol was modified by repeating the proteinase K digestion step to optimize the protocol for the greater protein content of equine blood.

Reverse Transcriptase PCR (RT-PCR)

To eliminate DNA contamination, RNA samples were DNase I (Invitrogen Corp., Carlsbad, CA) treated immediately before RT-PCR with the Titan One Tube RT-PCR kit (Roche Diagnostics Corp., Indianapolis, IN). Approximately 1 μg total RNA was treated with 1 unit DNase I in a cocktail containing only the water and PCR buffer reagents for RT-PCR. The samples were then incubated at 37°C for 10 minutes followed by 5 minutes at 65°C. A cocktail containing the remaining PCR components was then added and RT-PCR performed according to the manufacturer's directions. The primers KIT22F and KIT22R were used (Marklund *et al*., 1999) with the following thermocycler program: 45°C for 45m, 94°C for 5m, 50 cycles of 94°C for 30s, 60°C for 30s, 68°C for 2m 30s (plus 5s for each cycle after an initial 10 cycles), and a final extension of 68°C for 10m. This PCR product was diluted 1:20 with PCR grade water and used as a template for nested PCR with one of several sets of internal primers (Table 2.1). All primers were designed using the horse cDNA sequence as published in GenBank (Marklund *et al*., 1999) accession #AJ224645. All PCR, including RT-PCR was carried out on a PTC-200 thermocycler (MJ Research Inc., Boston, MA.) All nested PCR used FastStart taq DNA polymerase (Roche) according to the manufacturer's directions. Products were separated

by electrophoresis on 1% agarose gels and visualized after staining with ethidium
bromide under UV illumination.

Table 2.1: Primers used for nested RT-PCR of *KIT* cDNA

Pair	Name	Sequence 5'-3'	Exons	Product length in bp	Annealing Temp °C
1	"KIT22F"	GCTCGGATCCCATCGCAGCTACCG	5'UTR-3	565	60
1	KIT 544R	GCCATGAGGGAGTAATTGG			
2	KIT 381F	GGGCAGTTACACATGCAC	2-5	518	50
2	KIT 899R	CCCTGATGTGAGCTACTACTCT			
3	KIT 723F	AGCCATCAGAGCTGTGC	4-6	449	56
3	KIT 1172R	GTGGAGGTCCTGTTCATATATA			
4	KIT 1021F	GAAGTATTAGATAAAGGGTGCA	5-9	445	55
4	KIT 1466R	ACAGGAATGGAACATCTCTG			
5	KIT 1322F	ATGTTTACGTGAACACAAAACC	7-10	342	50
5	KIT 1664R	AGGCGTGAACAGAGTGTGGGGATGG			
6	KIT 1563F	CGGCACGGTCGAGTGTAGGGCTTAC	9-21	1399	63
6	KIT 2962R	CCGCACAGAATGGTCCACCGC			

Sequencing

PCR products for sequencing were amplified using the BigDye Terminator v1.1
Cycle sequencing kit (Applied Biosystems, Foster City, CA), cleaned using Centri-Sep
columns (Princeton Separations Inc., Adelphia, NJ), and run on a ABI 310 genetic
analyzer (Applied Biosystems) as directed by the manufacturers. The results were
analyzed and compared using the Vector NTI Suite 9 software package (InforMax Inc.,
Frederick, MD).

Determination of Exon Boundaries

The exon boundaries for the equine *KIT* cDNA sequence were predicted using the
UCSC Blat search tool to align the equine cDNA with the human genome (Kent, 2002).
The viewer highlights potential splice sites based on the presence of large stretches of
sequence found in the genomic DNA but not in the cDNA query.

Results

Single Nucleotide polymorphisms (SNPs)

The cDNA sequence, spanning from nucleotide (nt) 29 after the start of translation to 78nt before the stop codon for a total of 2821 bases was generated for five individual horses with the following phenotypes: a tobiano, a sabino, two whites and one non-spotted. Complete sequences were not obtained for all 15 horses due to changes in the research priorities as the project progressed. The majority of mouse and human *KIT* mutations have been found in the intracellular domain of the gene (Geissler *et al*., 1981; Fleischman *et al*., 1996). For that reason, sequencing efforts on the remaining 10 horses focused on the intracellular half of the gene (exons 10-21). Three novel SNPs were identified in *KIT* in exon 14, and one SNP previously described by Marklund and co workers in exon 20 found to be polymorphic in the horses tested (Marklund *et al*., 1999) (Table 2.2). Each of the novel SNPs was designated according to the nucleotide number from the start of translation. SNP 2150 is predicted to change the resulting amino acid from arginine to histidine. None of the four polymorphic SNPs could be specifically associated with a phenotype (present in all individuals with that spotting pattern and not in the others).

Table 2.2: SNPs among *KIT* cDNAs. Breeds used were abbreviated as follows: Thoroughbred-TB, Quarter Horse-QH, English Shire-ES, Clydsdale-Clydes, Tennessee Walking Horse-TWH, Gypsy Vanner-Vanner, and American Paint Horse-APHA. SNPs previously described are preceeded with "M-" (Marklund *et al.*, 1999). Horse ID beginning with "A" are accession numbers for sequences obtained from Genebank (type for only one strand available for these sequences.)

Horse ID	Breed	Spotting Pattern	M-SNP 632	SNP 2150	SNP 2175	M-SNP 2181	SNP 2205	M- SNP 2613	M-SNP 2739	M-SNP 2878
K	TB	Non-spotted						T/T	C/T	
GO	QH	Non-spotted	T/T	G/G	C/C	C/C	G/A	T/T	C/C	
S	Arab	Non-spotted						T/T	C/C	
A224645	NST	Non-spotted	C	G	C	C	G	C	T	G
A224642	Belgian	Roan	T	G	C	T	G	T	C	G
A224643	Belgian	Roan	T	G	C	T	G	T	C	G
A224644	Belgian	Roan	T	G	C	C	G	C	C	A
HT	TB	Roan?	T/T	G/G	C/C	G/G	G/G	T/T	C/C	
BN	ES	Sabino						T/T	C/C	
BAS	ES	Sabino						T/T	C/C	
GMR	TWH	Sabino	T/T	G/G	C/C	G/G	G/G	T/T	T/T	
TH	Clyds	Sabino						T/T	C/C	
GL	Pony	Splash		G/G	C/C	C/C	G/A	T/T	C/C	
V1A	Vanner	Tobiano						T/T	C/C	
PT	APHA	Tobiano	T/T	G/G	A/A	C/C	A/A	T/T	T/T	
7	APHA	Tobiano		A/A	C/C	C/C	G/G	T/T	T/T	
10	APHA	Tobiano (TO/TO)		G/A				T/T	C/C	
561	TB	White		G/G	C/C	C/C		T/T	T/T	
564	TB	White	T/T	G/G	C/C	C/C	A/A	T/T	C/C	

Splice Variants

The majority of RT-PCR products in one horse exhibiting a maximum sabino spotting pattern lacked exon 17. Some exon 17 was also detected in a related horse with moderate sabino spotting. Further investigations of the exon 17 splice variant are described in chapter three of this work. Complete skipping of exon 18 was observed in

one individual, an English Shire with a sabino spotting pattern typical of many draft breeds. However, skipping of exon 18 was not apparent in the sequences from a second individual from the English Shire breed with a sabino-type pattern. A minority of RT-PCR products (band intensity less than 10% the intensity of the major band) with the first 41 bases in exon 3 alternatively spliced out were detected in two individuals, one white and one splash.

cDNA Exon Boundaries and Sequence Conservation
The predicted exon boundaries agreed well with those published for the human mRNA (Andre *et al.*, 1992) (Figure 2.1). The alignment generated to predict exon boundaries 91% sequence identity between the horse cDNA and corresponding human sequence. Sequence identity by the same analysis was 89% with the mouse.

Figure 2.1: (Following pages) Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

```
                                                          1
        Ala Met Arg Gly Ala Arg Gly Ala Trp Asp Phe Leu Cys Val Leu Leu Leu
  +1
   1    GCAATGAGAG GCGCTCGCGG CGCCTGGGAT TTTCTCTGCG TCCTGCTTCT
        CGTTACTCTC CGCGAGCGCC GCGGACCCTA AAAGAGACGC AGGACGAAGA

                      1                              2
        Leu Leu Phe Arg Val Gln Thr Gly Ser Ser Gln Pro Ser Val Ser Pro Gly Glu
  +1
  51    CCTGTTCCGC GTCCAGACAG GCTCTTCTCA ACCATCTGTG AGTCCAGGGG
        GGACAAGGCG CAGGTCTGTC CGAGAAGAGT TGGTAGACAC TCAGGTCCCC

                                          2
        Glu Leu Ser Pro Pro Ser Ile His Pro Ala Lys Ser Glu Leu Ile Val Ser
  +1
 101    AACTGTCTCC ACCATCCATC CATCCAGCAA AATCAGAGTT AATCGTCAGT
        TTGACAGAGG TGGTAGGTAG GTAGGTCGTT TTAGTCTCAA TTAGCAGTCA

                                          2
        Val Gly Asp Glu Ile Arg Leu Leu Cys Ala Asp Pro Gly Phe Val Lys Trp
  +1
 151    GTTGGCGACG AGATTAGGCT GTTATGCGCT GATCCAGGCT TTGTCAAATG
        CAACCGCTGC TCTAATCCGA CAATACGCGA CTAGGTCCGA AACAGTTTAC

                                          2
        Trp Thr Phe Glu Thr Trp Gly Gln Leu Ser Glu Asn Thr His Lys Glu Trp Val
  +1
 201    GACTTTTGAG ACCTGGGGTC AGTTGAGTGA GAACACACAC AAAGAATGGG
        CTGAAAACTC TGGACCCCAG TCAACTCACT CTTGTGTGTG TTTCTTACCC
                                                 381F
                                          2
        Val Thr Glu Lys Ala Glu Ala Thr Asn Thr Gly Ser Tyr Thr Cys Thr Asn
  +1
 251    TCACGGAGAA AGCAGAGGCC ACAAACACGG GCAGTTACAC ATGCACCAAT
        AGTGCCTCTT TCGTCTCCGG TGTTTGTGCC CGTCAATGTG TACGTGGTTA
                                                           3
        Glu Gly Gly Leu Ser Ser Ser Ile Tyr Val Phe Val Arg Asp Pro Ala Lys
  +1
 301    GAAGGCGGCC TAAGCAGTTC CATTTATGTG TTTGTTAGAG ATCCTGCAAA
        CTTCCGCCGG ATTCGTCAAG GTAAATACAC AAACAATCTC TAGGACGTTT

                                          3
        Lys Leu Phe Leu Phe Asp Pro Ser Leu Tyr Gly Lys Glu Gly Ser Asp Thr Leu
  +1
 351    GCTTTTCCTC TTTGATCCTT CCTTGTATGG GAAAGAAGGC AGTGATACGC
        CGAAAAGGAG AAACTAGGAA GGAACATACC CTTTCTTCCG TCACTATGCG
                                           554R
                                          3
        Leu Val Arg Cys Pro Leu Thr Asp Pro Glu Val Thr Asn Tyr Ser Leu Met
  +1
 401    TGGTCCGCTG TCCTCTGACA GACCCAGAGG TGACCAATTA CTCCCTCATG
        ACCAGGCGAC AGGAGACTGT CTGGGTCTCC ACTGGTTAAT GAGGGAGTAC
        554R
                                          3
        Ala Cys Glu Gly Lys Ser Leu Pro Lys Asp Leu Thr Phe Val Ala Asp Pro
  +1
 451    GCGTGCGAGG GGAAATCTCT TCCCAAGGAC TTGACGTTCG TCGCTGATCC
        CGCACGCTCC CCTTTAGAGA AGGGTTCCTG AACTGCAAGC AGCGACTAGG
```

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

```
                                                        3
      +1  Pro Lys Ala Gly Ile Thr Ile Arg Asn Val Lys Arg Glu Tyr His Arg Leu Cys
      501 CAAGGCCGGC ATCACGATCA GAAACGTGAA GCGCGAGTAT CATCGACTGT
          GTTCCGGCCG TAGTGCTAGT CTTTGCACTT CGCGCTCATA GTAGCTGACA

                                                        3
      +1  Cys Leu Arg Cys Ser Ala Asp Lys Asp Gly Lys Ser Val Leu Ser Asn Lys
      551 GCTTGCGCTG CTCTGCAGAC AAGGATGGCA AGTCAGTGCT GTCGAATAAG
          CGAACGCGAC GAGACGTCTG TTCCTACCGT TCAGTCACGA CAGCTTATTC

                       3                              723F
                                                         4
      +1   Phe Thr Leu Lys Val Arg Ala Ala Ile Arg Ala Val Pro Val Val Ser Val
      601 TTCACCCTGA AAGTGAGGGC AGCCATCAGA GCTGTGCCAG TTGTGTCTGT
          AAGTGGGACT TTCACTCCCG TCGGTAGTCT CGACACGGTC AACACAGACA

                                                        4
      +1  Val Ser Lys Ala Ser Tyr Leu Leu Arg Glu Gly Glu Glu Phe Ser Val Thr Cys
      651 CTCCAAAGCA AGCTATCTTC TTAGGGAGGG GGAAGAATTT TCAGTGACGT
          GAGGTTTCGT TCGATAGAAG AATCCCTCCC CCTTCTTAAA AGTCACTGCA

                                                        4
      +1  Cys Leu Ile Lys Asp Ser Ser Ser Val Asp Ser Met Trp Ile Arg Glu
      701 GCTTGATAAA AGATGTGTCT AGTTCCGTGG ACTCAATGTG GATAAGGGAA
          CGAACTATTT TCTACACAGA TCAAGGCACC TGAGTTACAC CTATTCCCTT

                       4                              5
      +1  Asn Ser Arg Thr Lys Glu Gln Val Lys Ser Ser Ser Ser His Gln Gly Asp
      751 AACAGCCGGA CTAAAGAACA GGTAAAGAGT AGTAGCTCAC ATCAGGGTGA
          TTGTCGGCCT GATTTCTTGT CCATTTCTCA TCATCGAGTG TAGTCCCACT
                                          899R
                                                        5
      +1  Asp Phe Asn Phe Val Arg Gln Glu Arg Leu Thr Ile Ser Pro Ala Arg Val Asn
      801 CTTCAATTTT GTACGTCAGG AAAGACTGAC TATCAGCCCA GCAAGAGTTA
          GAAGTTAAAA CATGCAGTCC TTTCTGACTG ATAGTCGGGT CGTTCTCAAT

                                                        5
      +1  Asn Asp Ser Gly Val Phe Met Cys Tyr Ala Asn Asn Thr Phe Gly Ser Ala
      851 ACGATTCTGG AGTGTTCATG TGTTACGCCA ATAATACTTT TGGATCAGCA
          TGCTAAGACC TCACAAGTAC ACAATGCGGT TATTATGAAA ACCTAGTCGT

                       5                              1021F
                                                         6
      +1  Asn Val Thr Thr Thr Leu Glu Val Val Asp Lys Gly Phe Ile Asn Val Phe
      901 AATGTCACAA CAACCTTGGA AGTAGTAGAT AAAGGGTTCA TTAATGTCTT
          TTACAGTGTT GTTGGAACCT TCATCATCTA TTTCCCAAGT AATTACAGAA

                                                        6
      +1  Phe Pro Met Met Asn Thr Thr Val Phe Val Asn Asp Gly Glu Asn Val Asp Leu
      951 CCCTATGATG AATACTACCG TATTTGTGAA CGATGGAGAG AATGTAGATC
          GGGATACTAC TTATGATGGC ATAAACACTT GCTACCTCTC TTACATCTAG
```

```
                                                          6
      +1  Leu  Ile  Val  Glu  Tyr  Glu  Ser  Tyr  Pro  Lys  Pro  Glu  His  Gln  Gln  Trp  Ile
1001     TGATTGTTGA ATATGAATCC TACCCCAAAC CTGAACACCA GCAGTGGATA
         ACTAACAACT TATACTTAGG ATGGGGTTTG GACTTGTGGT CGTCACCTAT
                                                                   1172R

                                                          6
      +1  Tyr  Met  Asn  Arg  Thr  Ser  Thr  Asp  Lys  Trp  Glu  Asp  Tyr  Pro  Lys  Ser  Glu
1051     TATATGAACA GGACCTCCAC TGATAAGTGG GAAGATTATC CCAAGTCTGA
         ATATACTTGT CCTGGAGGTG ACTATTCACC CTTCTAATAG GGTTCAGACT
                       1172R

                             6
      +1  Glu  Asn  Glu  Ser  Asn  Ile  Arg  Tyr  Val  Ser  Glu  Leu  His  Leu  Thr  Arg  Leu  Lys
1101     GAACGAAAGT AATATCAGAT ATGTGAGCGA ACTTCATCTG ACTAGATTAA
         CTTGCTTTCA TTATAGTCTA TACACTCGCT TGAAGTAGAC TGATCTAATT

                                                          7
      +1  Lys  Gly  Thr  Glu  Gly  Gly  Thr  Tyr  Thr  Phe  Leu  Val  Ser  Asn  Ser  Asp  Val
1151     AAGGCACCGA AGGAGGCACT TACACGTTTC TAGTGTCCAA TTCTGATGTC
         TTCCGTGGCT TCCTCCGTGA ATGTGCAAAG ATCACAGGTT AAGACTACAG
                                                    1322F

                                                          8
      +1  Asp  Ser  Ser  Val  Thr  Phe  Asn  Val  Tyr  Val  Asn  Thr  Lys  Pro  Glu  Ile  Leu
1201     GATTCTTCCG TGACATTTAA TGTTTACGTG AACACAAAC CAGAAATCCT
         CTAAGAAGGC ACTGTAAATT ACAAATGCAC TTGTGTTTTG GTCTTTAGGA

                                                          8
      +1  Leu  Thr  Arg  Asp  Arg  Leu  Met  Asn  Gly  Met  Leu  Gln  Cys  Val  Ala  Ala  Gly  Phe
1251     GACTCGTGAC AGGCTCATGA ATGGCATGCT CCAGTGTGTG GCAGCAGGAT
         CTGAGCACTG TCCGAGTACT TACCGTACGA GGTCACACAC CGTCGTCCTA

                                                          8          9
      +1  Phe  Pro  Glu  Pro  Thr  Ile  Asp  Trp  Tyr  Phe  Cys  Pro  Gly  Thr  Glu  Gln  Arg
1301     TCCCAGAGCC CACAATAGAT TGGTATTTTT GTCCAGGAAC TGAGCAGAGA
         AGGGTCTCGG GTGTTATCTA ACCATAAAAA CAGGTCCTTG ACTCGTCTCT
                                                           1466R

                                                          9
      +1  Cys  Ser  Ile  Pro  Val  Gly  Pro  Val  Asp  Val  Lys  Ile  Gln  Asn  Ser  Ser  Val
1351     TGTTCCATTC CTGTGGGGCC AGTGGATGTG AAGATACAAA ACTCATCCGT
         ACAAGGTAAG GACACCCCGG TCACCTACAC TTCTATGTTT TGAGTAGGCA
                       1466R

                                                          9
      +1  Val  Ser  Pro  Phe  Gly  Lys  Leu  Val  Val  Gln  Ser  Ser  Ile  Asp  Tyr  Ser  Ala  Phe
1401     GTCACCGTTT GGAAAACTCG TGGTTCAAAG CTCCATCGAT TACAGTGCAT
         CAGTGGCAAA CCTTTTGAGC ACCAAGTTTC GAGGTAGCTA ATGTCACGTA
```

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

26

9

| +1 | Phe | Lys | His | Asn | Gly | Thr | Val | Glu | Cys | Arg | Ala | Tyr | Asn | Asp | Val | Gly | Lys |

```
1451   TCAAGCACAA CGGCACGGTC GAGTGTAGGG CTTACAATGA TGTGGGCAAG
       AGTTCGTGTT GCCGTGCCAG CTCACATCCC GAATGTTACT ACACCCGTTC
```

9   1635F

10

| +1 | Ser | Ser | Ala | Phe | Phe | Asn | Phe | Ala | Phe | Lys | Glu | Gln | Ile | His | Pro | His | Thr |

```
1501   AGTTCGGCCT TTTTTAACTT TGCATTTAAA GAGCAAATCC ATCCCCACAC
       TCAAGCCGGA AAAAATTGAA ACGTAAATTT CTCGTTTAGG TAGGGGTGTG
```
1664R

1635F

10

| +1 | Thr | Leu | Phe | Thr | Pro | Leu | Leu | Ile | Gly | Phe | Val | Ala | Ala | Ala | Gly | Met | Met | Cys |

```
1551   TCTGTTCACG CCTTTGCTGA TTGGGTTCGT GGCCGCCGCT GGCATGATGT
       AGACAAGTGC GGAAACGACT AACCCAAGCA CCGGCGGCGA CCGTACTACA
```
1664R

11

| +1 | Cys | Val | Ile | Val | Met | Val | Leu | Thr | Tyr | Lys | Tyr | Leu | Gln | Lys | Pro | Met | Tyr |

```
1601   GCGTCATCGT GATGGTTCTT ACCTACAAAT ATTTACAGAA ACCCATGTAT
       CGCAGTAGCA CTACCAAGAA TGGATGTTTA TAAATGTCTT TGGGTACATA
```

11

| +1 | Glu | Val | Gln | Trp | Lys | Val | Val | Glu | Glu | Ile | Asn | Gly | Asn | Asn | Tyr | Val | Tyr |

```
1651   GAAGTACAGT GGAAGGTTGT TGAGGAGATA AATGGAAACA ATTATGTGTA
       CTTCATGTCA CCTTCCAACA ACTCCTCTAT TTACCTTTGT TAATACACAT
```

11

| +1 | Tyr | Ile | Asp | Pro | Thr | Gln | Leu | Pro | Tyr | Asp | His | Lys | Trp | Glu | Phe | Pro | Arg | Asn |

```
1701   CATAGACCCA ACGCAACTTC CTTATGATCA CAAATGGGAG TTTCCCAGAA
       GTATCTGGGT TGCGTTGAAG GAATACTAGT GTTTACCCTC AAAGGGTCTT
```
1849R

12   1873F

| +1 | Asn | Arg | Leu | Ser | Phe | Gly | Lys | Thr | Leu | Gly | Ala | Gly | Ala | Phe | Gly | Lys | Val |

11

```
1751   ACAGACTGAG TTTTGGGAAA ACCTTGGGTG CTGGCGCCTT CGGAAAGGTT
       TGTCTGACTC AAAACCCTTT TGGAACCCAC GACCGCGGAA GCCTTTCCAA
```
1893R

12

| +1 | Val | Glu | Ala | Thr | Ala | Tyr | Gly | Leu | Ile | Lys | Ser | Asp | Ala | Ala | Met | Thr | Val |

```
1801   GTTGAGGCCA CTGCCTACGG CTTAATTAAG TCGGATGCGG CCATGACTGT
       CAACTCCGGT GACGGATGCC GAATTAATTC AGCCTACGCC GGTACTGACA
```
1963R

1959R

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

```
                                                                     13
                                                                   1973F
                              12
     +1  Val Ala Val Lys Met Leu Lys Pro Ser Ala His Leu Thr Glu Arg Glu Ala Leu
   1851  TGCTGTGAAG ATGCTCAAGC CAAGCGCTCA TTTAACGGAA CGAGAAGCCC
         ACGACACTTC TACGAGTTCG GTTCGCGAGT AAATTGCCTT GCTCTTCGGG
                          1963R
                      1959R
     1973F

                                              13
     +1  Leu Met Ser Glu Leu Lys Val Leu Ser Tyr Leu Gly Asn His Met Asn Ile
   1901  TAATGTCTGA ACTCAAAGTC TTGAGTTACC TTGGTAATCA CATGAATATT
         ATTACAGACT TGAGTTTCAG AACTCAATGG AACCATTAGT GTACTTATAA
                                                      2076R
                                                       14
                                                      2084F

                                      13
     +1  Val Asn Leu Leu Gly Ala Cys Thr Val Gly Gly Pro Thr Leu Val Ile Thr
   1951  GTGAATCTTC TCGGAGCATG CACCGTTGGA GGGCCCACCC TGGTCATTAC
         CACTTAGAAG AGCCTCGTAC GTGGCAACCT CCCGGGTGGG ACCAGTAATG
                        2076R                                  SNP 2150
           2084F
                                              14
     +1  Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu Arg Arg Lys Arg Asp
   2001  GGAATATTGT TGCTATGGTG ATCTTCTGAA TTTTTTGAGA AGAAAACGTG
         CCTTATAACA ACGATACCAC TAGAAGACTT AAAAAACTCT TCTTTTGCAC
                                      SNP 2175
                                              14
     +1  Asp Ser Phe Ile Cys Ser Lys Gln Glu Asp His Ala Glu Ala Ala Leu Tyr
   2051  ATTCATTTAT TTGCTCAAAG CAGGAAGATC ACGCAGAAGC AGCACTTTAT
         TAAGTAAATA AACGAGTTTC GTCCTTCTAG TGCGTCTTCG TCGTGAAATA
                               2196R
                                      14                     2245F
                                                              15
              SNP 2205                                       2235F
     +1  Lys Asn Leu Leu His Ser Lys Glu Ser Ser Cys Asn Asp Ser Thr Asn Glu
   2101  AAGAACCTTC TGCATTCAAA GGAGTCCTCC TGCAATGATA GTACTAATGA
         TTCTTGGAAG ACGTAAGTTT CCTCAGGAGG ACGTTACTAT CATGATTACT
                  2245F
                                              15
              2235F
     +1  Glu Tyr Met Asp Met Lys Pro Gly Val Ser Tyr Val Val Pro Thr Lys Thr Asp
   2151  GTACATGGAC ATGAAACCTG GAGTTTCTTA TGTCGTACCC ACCAAGACAG
         CATGTACCTG TACTTTGGAC CTCAAAGAAT ACAGCATGGG TGGTTCTGTC
                                                              2326R
```
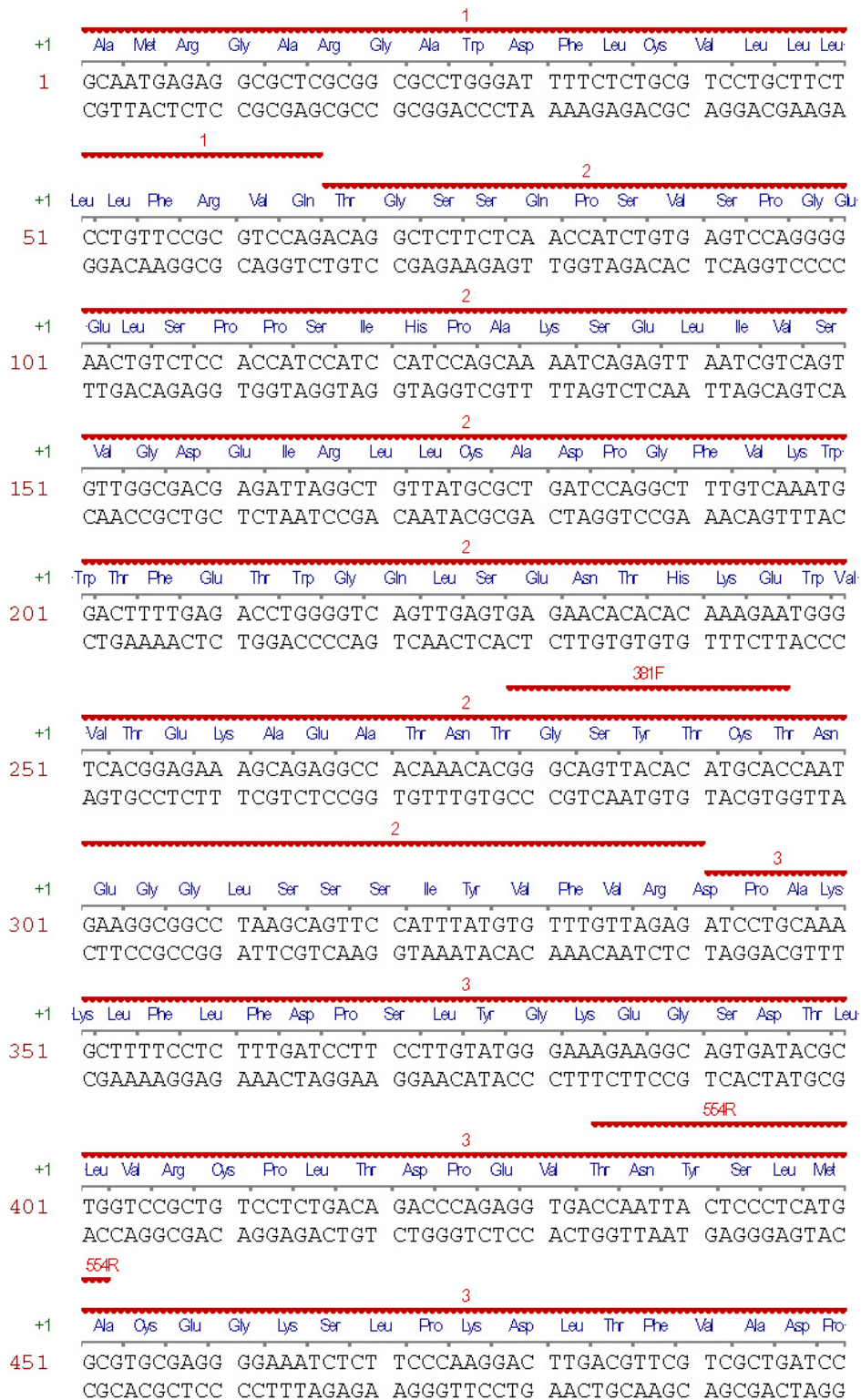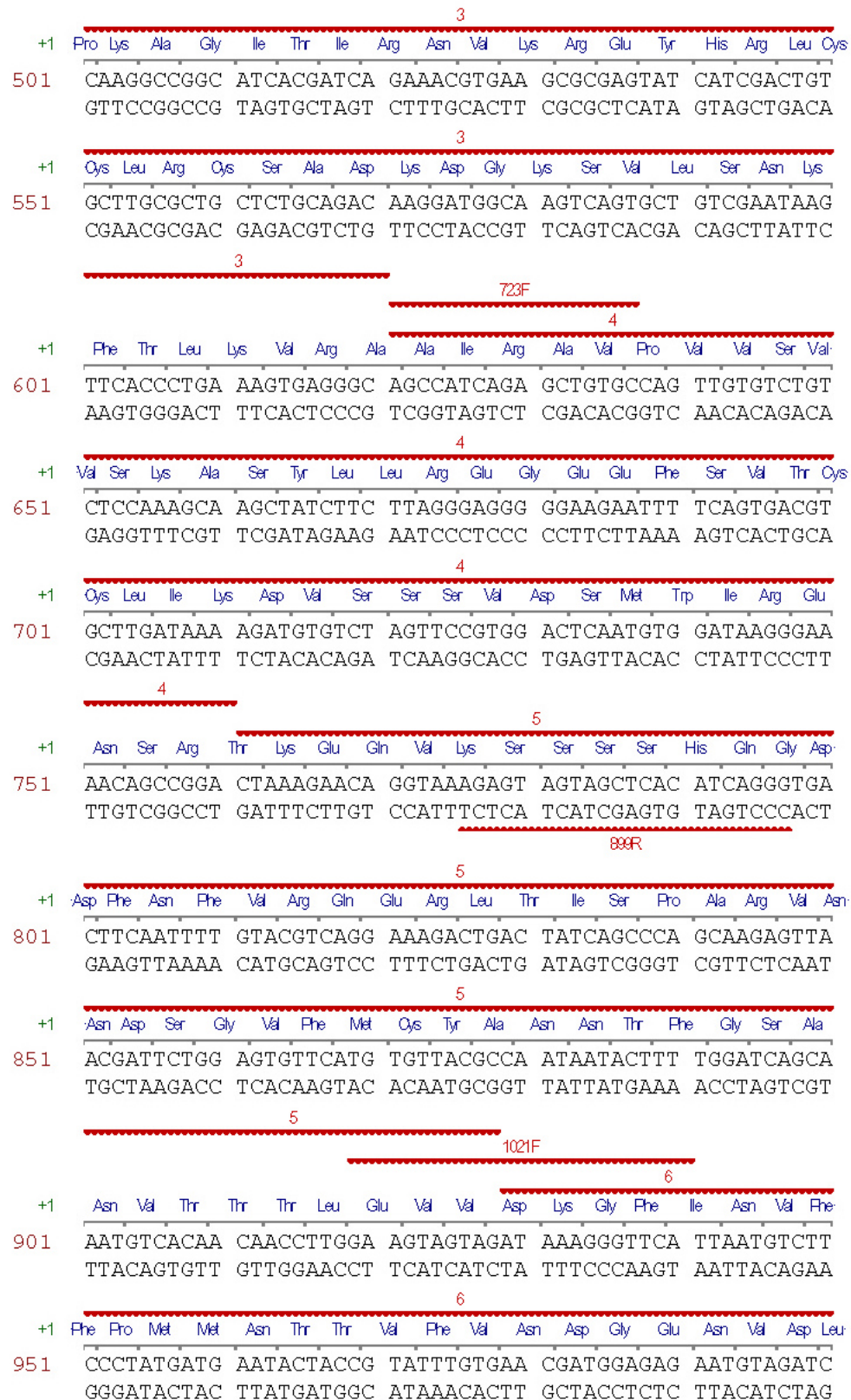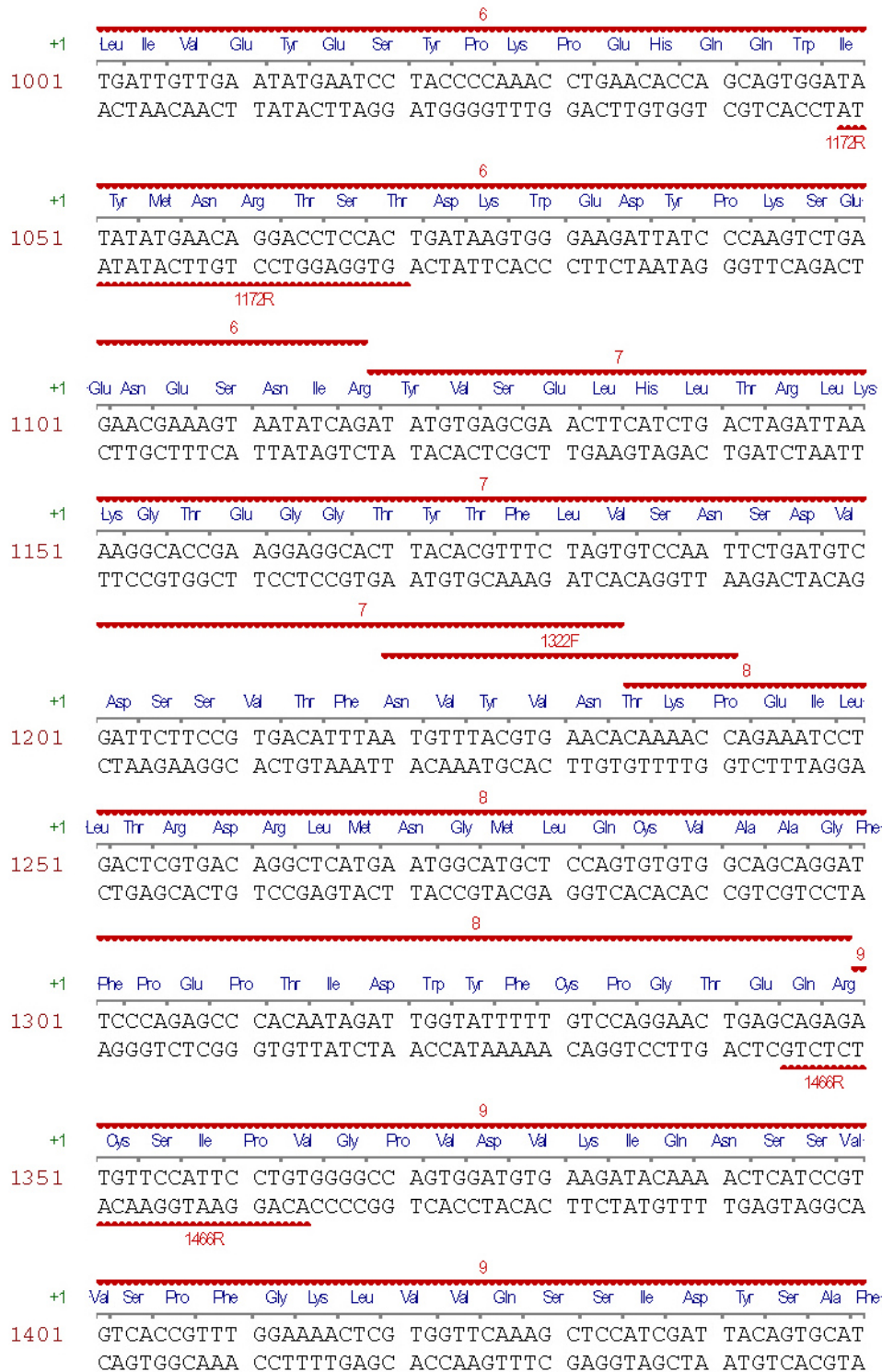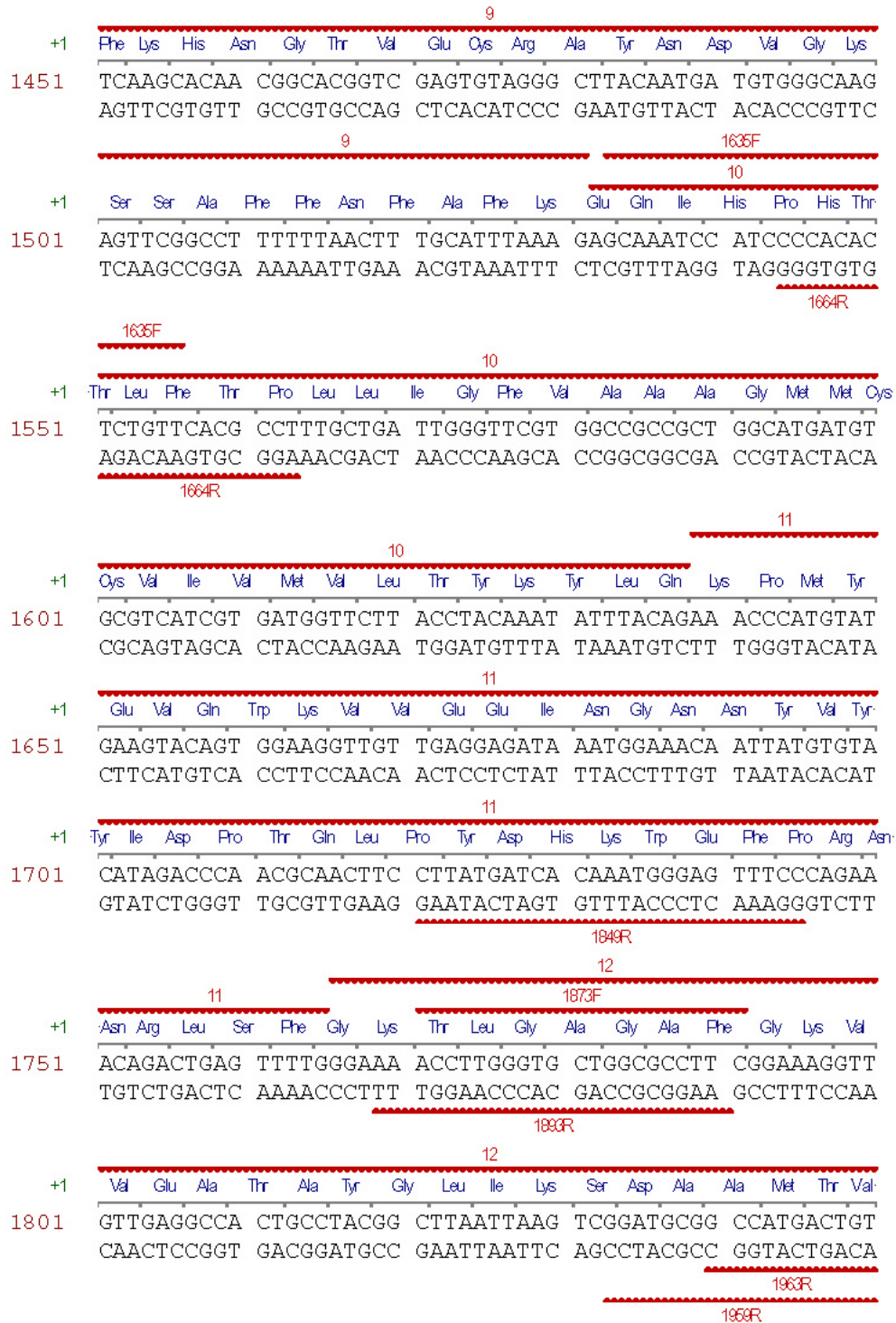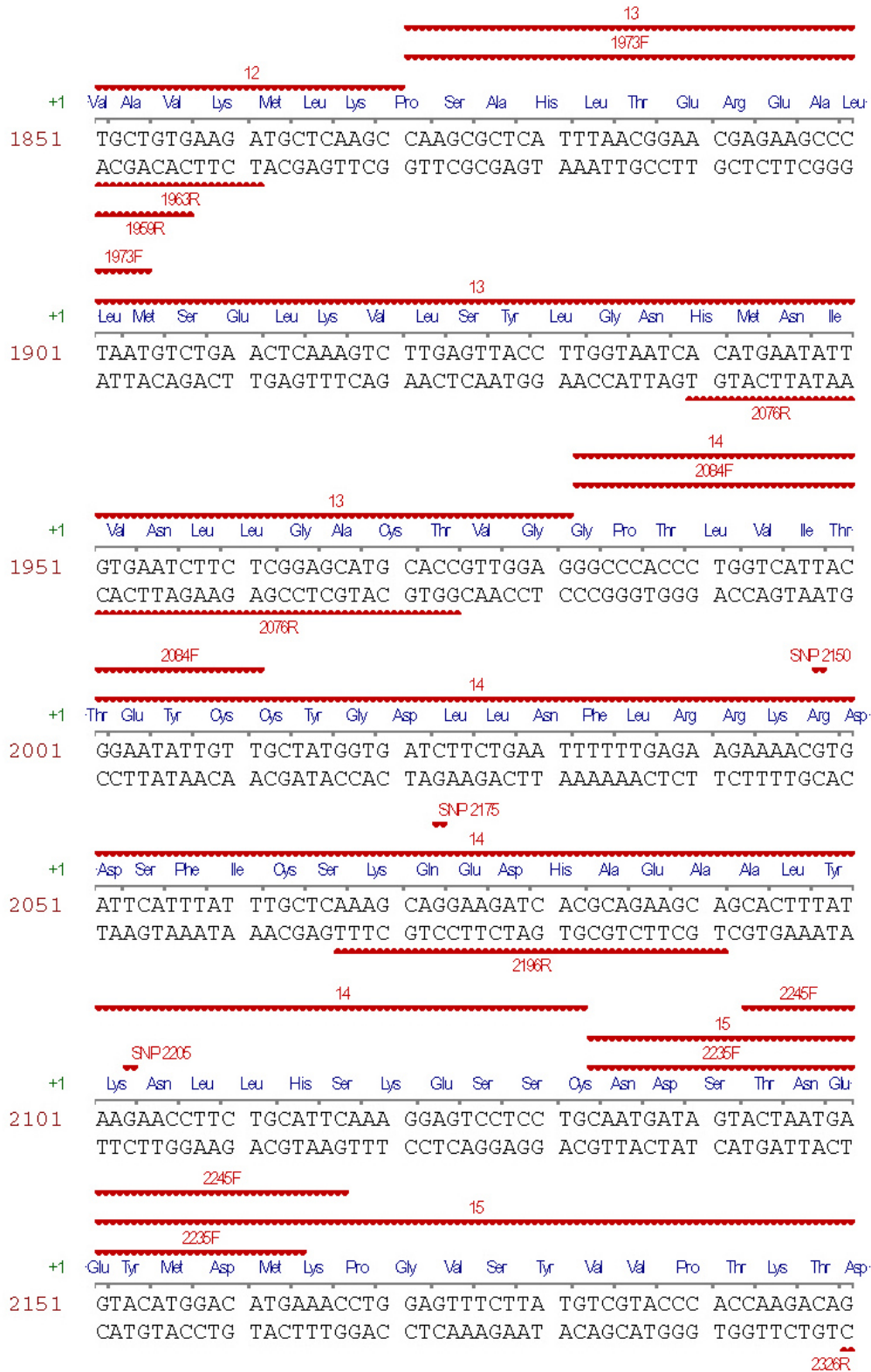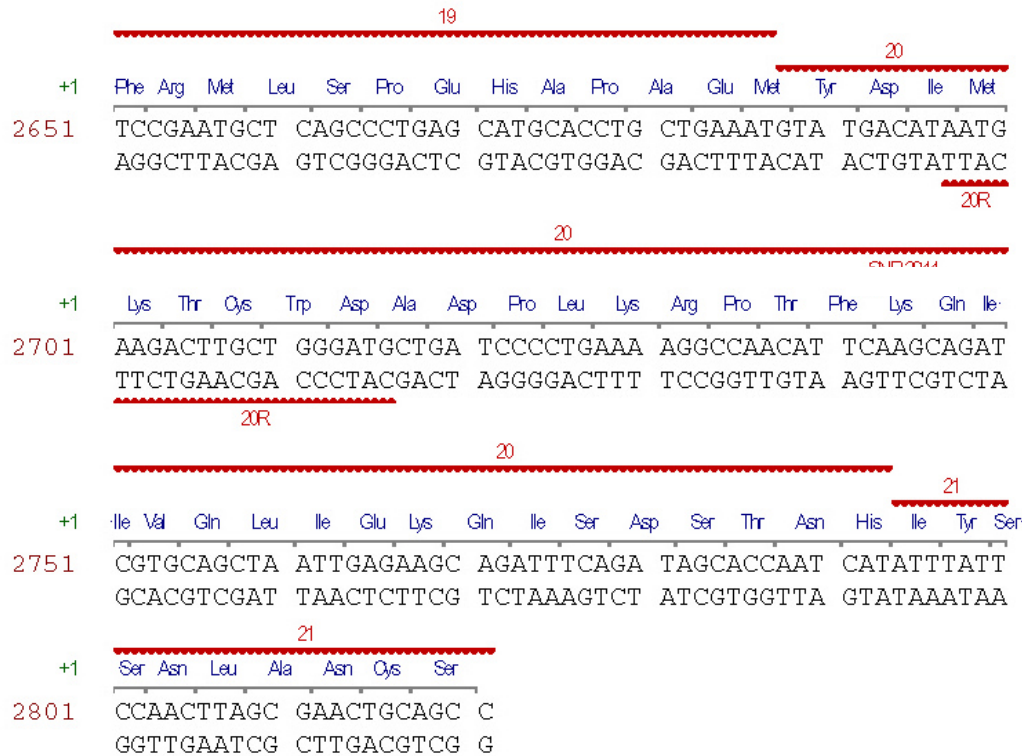
28

2327F

2327Fb

15

+1 Asp Lys Arg Arg Ala Ala Arg Ile Gly Ser Tyr Ile Glu Arg Asp Val Ala

2201 ACAAGAGGAG AGCTGCGAGA ATAGGCTCGT ACATAGAAAG AGACGTGGCT
     TGTTCTCCTC TCGACGCTCT TATCCGAGCA TGTATCTTTC TCTGCACCGA

2326R

16

+1 Pro Ser Ile Met Glu Asp Asp Glu Leu Ala Leu Asp Leu Glu Asp Leu Leu

2251 CCGTCCATCA TGGAAGATGA CGAGCTGGCC CTAGACCTGG AAGACCTGCT
     GGCAGGTAGT ACCTTCTACT GCTCGACCGG GATCTGGACC TTCTGGACGA

16

+1 Leu Ser Phe Ser Tyr Gln Val Ala Lys Gly Met Ala Phe Leu Ala Ser Lys Asn

2301 CAGCTTTTCT TACCAGGTGG CAAAGGGCAT GGCGTTCCTT GCCTCAAAGA
     GTCGAAAAGA ATGGTCCACC GTTTCCCGTA CCGCAAGGAA CGGAGTTTCT

16

e17-17F

17

+1 Asn Cys Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Thr His Gly

2351 ACTGTATTCA TAGGGACTTG GCAGCCAGAA ATATCCTCCT TACTCATGGT
     TGACATAAGT ATCCCTGAAC CGTCGGTCTT TATAGGAGGA ATGAGTACCA

2571R (e17-117R)

17

+1 Arg Ile Thr Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Lys Asn Asp

2401 CGAATCACAA AGATTTGTGA TTTTGGTCTA GCCAGAGACA TCAAGAATGA
     GCTTAGTGTT TCTAAACACT AAAACCAGAT CGGTCTCTGT AGTTCTTACT

17

2571R (e17-117R)

18

+1 Asp Ser Asn Tyr Val Val Lys Gly Asn Ala Arg Leu Pro Val Lys Trp Met Ala

2451 TTCTAATTAT GTGGTCAAAG GAAATGCTCG GCTACCTGTG AAATGGATGG
     AAGATTAATA CACCAGTTTC CTTTACGAGC CGATGGACAC TTTACCTACC

18

+1 Ala Pro Glu Ser Ile Phe Asn Cys Val Tyr Thr Phe Glu Ser Asp Val Trp

2501 CACCTGAAAG CATTTTCAAC TGTGTGTACA CATTTGAAAG TGATGTCTGG
     GTGGACTTTC GTAAAAGTTG ACACACATGT GTAAACTTTC ACTACAGACC

18

+1 Ser Tyr Gly Ile Phe Leu Trp Glu Leu Phe Ser Leu Gly Ser Ser Pro Tyr

2551 TCCTATGGGA TTTTTCTGTG GGAGTTGTTC TCTTTAGGAA GCAGCCCCTA
     AGGATACCCT AAAAAGACAC CCTCAACAAG AGAAATCCTT CGTCGGGGAT

e18-104R

19

+1 Tyr Pro Gly Met Pro Val Asp Ser Lys Phe Tyr Lys Met Ile Lys Glu Gly Phe

2601 CCCTGGAATG CCAGTCGATT CTAAGTTCTA CAAGATGATC AAGGAAGGTT
     GGGACCTTAC GGTCAGCTAA GATTCAAGAT GTTCTACTAG TTCCTTCCAA

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

```
                                                                           20
        +1  Phe Arg Met Leu Ser Pro Glu His Ala Pro Ala Glu Met Tyr Asp Ile Met
     2651   TCCGAATGCT CAGCCCTGAG CATGCACCTG CTGAAATGTA TGACATAATG
            AGGCTTACGA GTCGGGACTC GTACGTGGAC GACTTTACAT ACTGTATTAC
                                                                       20R

                                                    20
        +1  Lys Thr Cys Trp Asp Ala Asp Pro Leu Lys Arg Pro Thr Phe Lys Gln Ile·
     2701   AAGACTTGCT GGGATGCTGA TCCCCTGAAA AGGCCAACAT TCAAGCAGAT
            TTCTGAACGA CCCTACGACT AGGGGACTTT TCCGGTTGTA AGTTCGTCTA
                          20R

                                           20                          21
        +1  ·Ile Val Gln Leu Ile Glu Lys Gln Ile Ser Asp Ser Thr Asn His Ile Tyr Ser
     2751   CGTGCAGCTA ATTGAGAAGC AGATTTCAGA TAGCACCAAT CATATTTATT
            GCACGTCGAT TAACTCTTCG TCTAAAGTCT ATCGTGGTTA GTATAAATAA
                          21
        +1  Ser Asn Leu Ala Asn Cys Ser
     2801   CCAACTTAGC GAACTGCAGC C
            GGTTGAATCG CTTGACGTCG G
```

Figure 2.1 con't: Consensus *KIT* cDNA sequence annotated with exon boundary, primer and SNP locations.

Discussion

None of the SNPs found could be strictly associated with splash, tobiano or white. Unless the causative mutation for these phenotypes is located in the 180 unsequenced bases on either end of the cDNA, the cause for these phenotypes must reside elsewhere. The origin of these phenotypes may be heterogeneous among the horse population, as reflected in the genetic marker associations for these traits. Thus a single causative mutation may not exist and would not be detected by looking for variations unique to the phenotype as a whole. Studies of the inheritance of each spotting pattern within a family, insuring identity by decent, may be more successful in identifying associated variations for these phenotypes.

In the case of tobiano, an unusual conservation of the *TO:ALB-B:GC-S* haplotype has been previously observed by A.T. Bowling (Bowling, 1987). Closer examination of

the prevalence of the *TO:ALB-B* haplotype among tobiano American Paint horses revealed a strong linkage disequilibrium as well as prevalence of this haplotype in 92% of all homozygous horses tested (Duffield and Goldie, 1998). Bowling theorized that a chromosomal inversion would explain the conservation of phase observed in tobiano horses. Should a chromosomal inversion in the horse disrupt the regulatory region of the *KIT* gene similar to those found in rumpwhite, banded and sash mice it could potentially explain both the tobiano phenotype and the phase conservation. The heterogeneity observed in the *KIT* gene (Table 2.2) may reflect mutations that occurred since the creation of the *Tobiano* allele. The use of the SNPs found in this study to generate additional haplotypes in phase with tobiano can be used in the future to further characterize the linkage disequilibrium associated with *Tobiano.*

Chromosome rearrangements near *KIT* have also been attributed to white phenotypes in pigs (Marklund *et al*., 1998). Two unique duplications and an additional two triplications have been identified as the cause of dominant white in pigs (Pielberg *et al*., 2002). It is not likely that a recent duplication event would be detected by sequencing; therefore other quantitative methods could be employed to investigate the possibility of a *KIT* duplication in the white horse.

The discovery of the splice variant missing exon 17 identified here led to additional work described in chapter three of this volume. The observation that *KIT* exon 18 was deleted in one of the draft horses was intriguing, however no SNP was detected that could explain this gene expression variant and no family or population studies were conducted to determine whether sabino in draft breeds was associated with deletion of *KIT* exon 18. An experimentally created mouse strain containing a mutation that leads to skipping of exon 18 exhibits a belly spot and depigmentation of the feet similar to the sabino phenotype in draft horses (Greber *et al*., 2005). Further investigation of this splice variant and its association with sabino spotting in the horse with more sensitive methods of detection are warranted.

The alternative splicing of exon 3 was not associated with a particular phenotype. The deletion of the 41 base pair portion of exon 3 would result in truncated protein. As evidenced by the presence of the exon 3 splice variant in horses of varying phenotype, truncation of a small percentage of mRNAs does not likely affect spotting in the horse.

Furthermore, direct sequencing of the RT-PCR product is not an efficient method to detect or quantify relatively rare splice variants. Due to the exponential nature of amplification by PCR the most common template is better represented in the product than rare templates. However, this preference for more numerous templates limits the number of variants detected due to random erroneous splice events. By cloning the RT-PCR product in to a vector and then sequencing several clones additional rare splice variants could be detected. However, further experiments would need to be conducted in order to determine the relative frequency of these rare variants in the transcriptome and their potential relevance to spotting phenotype.

Four out of five of the *KIT* variants described by Marklund *et al.* were not found to be polymorphic among the horses for which sequence was generated in those areas. This could be due to the divergence between the European breeds used by Marklund and the American breeds chosen for this study. Alternatively, the SNPs in that study may represent haplotypes not found among horses with the phenotypes in this study. Owing to the proximity of the *Roan* and *Tobiano* loci (both linked to *KIT*), it is possible that little recombination has occurred between the two. Selection of exclusive sample groups based on these two phenotypes could therefore simultaneously select for unique sets of associated SNPs.

CHAPTER THREE: Exon Skipping in the *KIT* Gene and Sabino Spotting

Summary

Sabino (*SB*) is a white spotting pattern in the horse characterized by white patches on the face, lower legs, or belly, and interspersed white hairs on the midsection. Based on comparable phenotypes in humans, mice and pigs, the *KIT* gene was investigated as the origin of the sabino phenotype. Here is reported the dominant genetic basis of one type of sabino spotting pattern in horses called *Sabino 1*, with the alleles represented by the symbols *SB1* and *sb1*. Genomic DNA sequencing of *KIT* revealed a single nucleotide polymorphism (SNP) caused by a base substitution for T with A in intron 16, 1037 bases following exon 16. The SNP associated with the *Sabino 1* phenotype was designated KI16+1037A. This substitution eliminated a *Mnl* I restriction site and allowed the use of PCR-RFLP to characterize individuals for this base change. Complete linkage was observed between this SNP and *Sabino 1* in the Tennessee Walking Horse families (LOD=9.02 for Θ=0). Individual horses from other breeds were also tested. All 5 horses homozygous for this SNP were white, and all 68 horses with one copy of this SNP either exhibited the *Sabino 1* phenotype or were multi-patterned. Some multi-patterned individuals appeared white due to the additive effect of white spotting patterns. However, 13 horses with other sabino-type patterns did not have this SNP. Transcripts of *KIT* were characterized by RT-PCR and sequencing cDNA from horses with the genotypes *SB1/SB1, SB1/sb1* and *sb1/sb1*. Horses with the *Sabino1* trait produced a splice variant of *KIT* that did not possess exon 17. Based on these results we propose the following: 1) this SNP, found within intron 16, is responsible for skipping of exon 17 and the *SB1* phenotype, 2) the White and Sabino phenotypes are heterogeneous and this mechanism is not the only way to produce the pattern described as "sabino" or "white" and 3) homozygosity for *SB1* results in a complete or nearly completely white (sabino-white) phenotype.

Introduction

White spotting patterns, and the flashy spotted coats they produce, are valued in the horse for their aesthetic quality. One such spotting pattern, sabino, is characterized by irregularly bordered white patches that begin at the extremities and face (Figure 1). These white patches often extend to include the belly and midsection, either as distinct areas of white hair, or as a diffuse scattering of white hairs resembling roan (Guerts, 1977). White areas lack pigment both in the hair and skin. In the United States the sabino spotting pattern is found in many light horse breeds, most notably the Tennessee Walking Horse and Missouri Foxtrotter (Sponenberg, 2003).

Figure 3.1: A *Sabino 1* mare and her homozygous *Sabino 1* (sabino-white) foal.

Spotting patterns similar to sabino have been found in mice, pigs and humans. Among the many documented spotting patterns of the mouse those most similar to sabino are caused by the *KIT* gene (*W* locus). Most closely resembling sabino, *Viable dominant spotting* (W^v) in the mouse is a semi-dominant trait characterized in heterozygotes by white markings along the mid-ventral line often extending to the extremities, white

headspots, and some dilution of the remaining body color. Homozygotes have black eyes but are otherwise completely white. W^v and other similar patterns at the W locus are caused by mutations affecting the kinase domain of the KIT receptor (Nocka $et\ al$., 1990). These mutations in mice often have detrimental effects that include anemia, mast cell deficiency, sterility and deafness (Geissler and Russell, 1983). The location of the particular mutation can alter the severity of these negative effects (Spritz, 1994). In pigs, a dominant white phenotype has been attributed to both a KIT gene duplication and a splice mutation leading to skipping of KIT exon 17 (Marklund $et\ al$., 1998). In humans, several different mutations of the KIT gene have been shown to cause piebaldism, many of which show striking similarity in depigmentation pattern to sabino horses. In humans the depigmentation also begins at the extremities, has jagged margins, can extend in to the midsection, and very often includes a white forelock (Syrris $et\ al$., 2002).   Due to the similarities in phenotype between sabino spotting in horses, alleles at the *dominant white spotting* (*W)* locus in mice, and piebaldism in humans we hypothesized that the KIT gene was responsible for sabino spotting. KIT has previously been linked to two genes for spotting patterns in the horse, namely *Roan* (Marklund $et\ al$., 1999) and *Tobiano* (Brooks $et\ al$., 2002).

The KIT gene encodes the mast/stem cell growth factor receptor. A member of the tyrosine kinase receptor family, it has an immunoglobulin-like extracellular binding domain, a transmembrane region and two intracellular kinase domains (Andre $et\ al$., 1992). Melanoblasts, or melanocyte precursor cells, are derived from both sides of the neural crest. These cells migrate away from the neural crest distally, towards the extremities, before entering and colonizing the dermis, and finally entering their final location in the epidermis (Mayer, 1973). Though the differentiation of melanoblasts to mature melanocytes is under the control of several factors, KIT signaling is essential if proper migration, proliferation and survival is to occur (Peters $et\ al$., 2002; Mackenzie $et\ al$., 1997; Ito $et\ al$., 1999).

The partial depigmentation seen in W spotting patterns of mice is caused by two different mechanisms that disrupt the normal function of the KIT receptor. Regulatory mutations leading to inappropriate expression of the receptor can change the amount of available ligand, leading to inefficient signal transmission and altered timing of

melanocyte proliferation, as occurs in the *W-sash* mouse (Duttlinger *et al*., 1993). In contrast, mutations affecting the function of the receptor itself have more drastic effects, reducing or eliminating the tyrosine kinase activity, and thus the perception of the signal by the melanocyte. This leads to reduced migration, proliferation or survival of melanocytes in the embryo, and reduced pigmentation in the adult (Peters *et al*., 2002). Here evidence is described for exon skipping in the *KIT* gene as a cause of a white spotting pattern we designate *Sabino 1*.

Materials and Methods

Horses

Linkage studies were conducted using three families of Tennessee Walking Horses. These families consisted of: 1) a sabino stallion and his 12 half-sibling offspring out of solid mares, 2) a sabino stallion, 5 mares, and their 11 offspring and 3) a solid stallion mated to 2 sabino mares with 2 full-siblings each (4 total). LOD scores were calculated for the three families with $\Theta=0$.

The *KIT* alleles described in this report were designated KI16+1037A and KI16+1037T. The designation was abbreviated from the description; "*KIT,* intron 16, base pair 1037, nucleotide A or T." The distributions of the KI16+1037 alleles were based on direct gene counting from their RFLP phenotypes as described below. DNA samples from 52 Thoroughbred horses and 41 American Standardbred horses were randomly selected without regard for color from those previously collected and archived at the Genomics and Immunogenetics Laboratory, Gluck Equine Research Center.

Samples from the Tennessee Walking Horse breed as well as others tested in connection with this study were submitted by breeders with pedigree information and photographs showing coat color patterns. These samples came from the following breeds: American Miniature Horse (28), American Paint Horse (27), Arabian (11), Azteca (3), Missouri Fox Trotter (20), Shetland Pony (9), Spanish Mustang (12), Tennessee Walking Horses (79 in addition to the families), and Thoroughbred (24). This manner of collecting samples precludes reporting allele frequencies, but does provide an indication of the presence of the KI16+1037A SNP in a variety of breeds. Additional

samples were collected from the Belgian (1), English Shire (3), and Clydesdale (2) breeds.

Color Determination

Horses characterized as having a *Sabino 1* white spotting pattern possessed 3 out of 4 of the following characteristics: 1) two or more white feet or legs, 2) blaze (white patch extending the length of the face), 3) jagged margins around white areas, and 4) spots or roaning in the midsection. These characteristics agreed with the description of sabino as published by Guerts (Guerts, 1977).

In the course of this investigation, family studies suggested that homozygosity for the *Sabino 1* gene may result in a phenotype exhibiting extensive white coat color, hereafter referred to as sabino-white in order to distinguish it from *White (W)*. Candidates for sabino-white appeared at least 90% depigmented from birth (thus excluding G*rey,* a pattern of progressive depigmentation that in aged horses can result in a nearly white phenotype (Sponenberg, 2003) and had two parents with a sabino spotting pattern. All white-colored horses were also tested using PCR-RFLP for the presence of other coat color variants that might confound white color determination. Namely, *Frame* causing the frame overo spotting pattern, *Cream* (*CR*) causing a near white Cremello phenotype in the homozygotes for *CR*, and *Tobiano* (*TO*) which can produce white spotting patterns that could mimic or obscure the effects of a sabino gene (Sponenberg*,* 2003).

RNA Extraction

RNA was isolated from whole blood using two methods. The first used the Purescript RNA purification kit (Gentra Systems Inc., Minneapolis, MN) following the manufacturer's directions for 10mL whole blood. For samples received by mail or which could not be processed within 6 hours, we used the PAX gene blood RNA kit (Qiagen Inc., Valencia, CA) for RNA isolation. The manufacturer's protocol was modified by repeating the proteinase K digestion step to optimize the protocol for the greater protein content of equine blood.

Reverse Transcriptase PCR (RT-PCR)

To eliminate DNA contamination RNA samples were DNase I (Invitrogen Corp., Carlsbad, CA) treated immediately before RT-PCR with the Titan One Tube RT-PCR kit (Roche Diagnostics Corp., Indianapolis, IN). Approximately 1 μg total RNA was treated with 1 unit DNase I in a cocktail containing only the water and PCR buffer reagents for RT-PCR. The samples were then incubated at 37°C for 10 minutes followed by 5 minutes at 65°C. A cocktail containing the remaining PCR components was then added and RT-PCR performed according to the manufacturer's directions. The primers KIT22F and KIT22R were used (Marklund *et al*., 1999) with the following thermocycler program: 45°C for 45m, 94°C for 5m, 50 cycles of 94°C for 30s, 60°C for 30s, 68°C for 2m 30s (plus 5s for each cycle after an initial 10 cycles), and a final extension of 68°C for 10m. This PCR product was diluted 1:100 with PCR grade water and used as a template for nested PCR with one of several sets of internal primers (Table 3.1).   All primers were designed using the horse cDNA sequence as published in GenBank (Marklund *et al*., 1999) accession #AJ224645. To detect exon skipping the primers KITe16-51F and KITe18-104R were used for exon 17 and KITe17-17F and KIT20R (Venta *et al*., 1996) for exon 18. All PCR, including RT-PCR was carried out on a PTC-200 thermocycler (MJ Research Inc., Boston, MA.) All nested PCR used FastStart taq DNA polymerase (Roche) according to the manufacturer's directions. Products were separated by electrophoresis on 1% agarose gels and visualized after staining with ethidium bromide under UV illumination.

Table 3.1. Primers used for nested PCR.

Pair	Name	Sequence 5'-3'	Exon	Product length in bp	Annealing Temp. °C
1	"KIT22F"	GCTCGGATCCCATCGCAGCTACCG	5'UTR	565	60
1	KIT 544R	GCCATGAGGGAGTAATTGG	3		
2	KIT 381F	GGGCAGTTACACATGCAC	2	518	50
2	KIT 899R	CCCTGATGTGAGCTACTACTCT	5		
3	KIT 723F	AGCCATCAGAGCTGTGC	4	449	56
3	KIT 1172R	GTGGAGGTCCTGTTCATATATA	6		
4	KIT 1021F	GAAGTATTAGATAAAGGGTGCA	5	445	55
4	KIT 1466R	ACAGGAATGGAACATCTCTG	9		
5	KIT 1322F	ATGTTTACGTGAACACAAAACC	7	342	50
5	KIT 1664R	AGGCGTGAACAGAGTGTGGGGATGG	10		
6	KIT 1563F	CGGCACGGTCGAGTGTAGGGCTTAC	9	1399	63
6	KIT 2962R	CCGCACAGAATGGTCCACCGC	21		

Sequencing

PCR products for sequencing were amplified using the BigDye Terminator v1.1 Cycle sequencing kit (Applied Biosystems, Foster City, CA), cleaned using Centri-Sep columns (Princeton Separations Inc., Adelphia, NJ), and run on a ABI 310 genetic analyzer (Applied Biosystems) as directed by the manufacturers. For products longer than 500bp internal primers were designed for sequence walking (Table 3.2). The results were analyzed and compared using the Vector NTI Suite 9 software package (InforMax Inc., Frederick, MD).

Table 3.2: Additional primers used for sequencing.

Name	Sequence 5'-3'	Exon
KIT 1873F	ACCTTGGGTGCTGGCGCCTTC	12
KIT 1935R	CACAGCAACAGTCATGGCCGCATCC	12
KIT 2245F	ACTAATGAGTACATGGACATGAAAC	15
KIT 2717R	GCATCCCAGCAAGTCTTCAT	20

DNA Extraction

DNA samples obtained from blood were extracted using the Puregene DNA extraction kit (Gentra Sys.) according their protocol. Hair samples submitted by breeders were processed for PCR using 5-7 hair bulbs as previously published by Locke *et al*. (2002).   Briefly, the hair bulbs were incubated at 60°C for 45m followed by 95°C for 45 minutes in 100μL of a buffer consisting of 1x FastStart Taq Polymerase PCR buffer (Roche), 2.5 mM MgCl$_2$ (Roche), 0.5% TWEEN 20 (JT Baker, Phillipsburg NJ) and 0.01mg Proteinase K (Sigma-Aldrich, St Louis MO).

PCR-RFLP for KI16+1037

To screen for the polymorphism of the *KIT* gene in intron 16 (KI16) a portion of the 3' end of intron 16 and beginning of exon 17 was amplified using PCR with FastStart taq DNA polymerase (PE manufacturer) and the following primers: KITi16aF (5'-GGT CCT GAC GAT GAG AAA CAC AAG T-3') and KITe17-117R (5'-TTT GAC CAC ATA ATT AGA ATC ATT C-3') at an annealing temperature of 57°C. A "T" to "A" substitution, this polymorphism eliminates a *Mnl I* site when the *Sabino 1* allele is present. The allele with the SNP KI16+1037T was designated *sb1* and the allele with the

SNP KI16+1037A (and associated with the sabino pattern) was designated *SB1*. This PCR product was also designed to include an additional *Mnl I* site, unrelated to the polymorphism that could be used as a positive control for cutting by the restriction enzyme. The PCR product was restriction digested using approximately 5mU *Mnl I* enzyme (New England Biolabs Inc., Beverly, MA) per 1ng PCR product at 37°C for 1h. Fragments were separated by electrophoresis on 2% agarose gels and visualized after staining with ethidium bromide under UV illumination.

PCR-RFLP for CR and Frame

To aid in the color and spotting pattern determination of horses submitted to this study we used PCR-RFLP tests for a SNP linked to *TO*, *CR* and *Frame*. Testing for *Tobiano* was conducted as described in Brooks *et al.*, 2002.

The genetic variant responsible for Cream (*CR)* was reported to be a base substitution in the gene called *underwhite (UW)* in mice and *Membrane Associated Transporter Protein* (*MATP)* gene in humans (Mariat *et al.*, 2003). To test for the mutation in *MATP* we devised the following PCR-RFLP test. The primers MATP-F (5'-CTT TGA TTG CTG ACC GAA GGA AGA A-3') and MATP-R (5'-TAA TAC CGA TGC TAC ATT GCT GTC T-3') were used for PCR with an annealing temperature of 56°C followed by restriction digestion of the 525bp PCR product with an excess of the enzyme *Mse I*. Following electrophoresis on a 2% gel and staining with ethidium bromide the wild type produced 2 bands 478bp and 26bp (actually comprised of a 22 and a 26bp band) in size, heterozygotes 4 bands of 478, 406, 72, and 26bp and homozygous *CR* 3 bands 406, 72, and 26bp in size. The 26 bp fragment was a positive control for success of the restriction digest and present regardless of which alleles were present.

A base substitution in the gene *Endothelin Receptor B (EDNRB)* is responsible for OLWS. We encountered some problems reliably detecting the *EDNRB* SNP using allele specific PCR as described by Metallinos *et al.* (1998). Therefore, an alternative method was devised to detect the same SNP. A multiplex PCR was designed to use the published E1.2F and E1.R primers (Metallinos *et al.*, 1998) to amplify a 175bp product containing the mutation, as well as a second set (OC-F 5'-CAA ATT GAA GTT GCA AGA C-3' and OC-R 5'-ATA TTT CTG GCT GCC AAG TC-3') to amplify a 430bp control fragment. The multiplex PCR was performed at an annealing temperature of 54°C using

a two to one ratio of the OC primer set to the E1 primer set. The product was then restriction digested using an excess of the enzyme *HPY 188 I*. Following electrophoresis on a 2% gel and staining with ethidium bromide the non-spotted allele produced 4 bands 229, 201, 97 and 78bp in size, while the presence of a single copy of the *Frame* allele produced only 3 bands, 229, 201 and 175bp. No homozygotes for the *Frame* allele were observed due to the lethal effects of Overo Lethal White Foal Syndrome among homozygotes.

Results

RT-PCR and cDNA Sequencing

RNA was isolated from lymphocytes of Tennessee Walking horses categorized into three phenotypes: 1)those without the sabino spotting pattern, 2)those with the *Sabino 1* pattern, and 3) those with mostly white body color that were known to be offspring of two sabino parents. cDNA was produced and the product sequenced by primer walking using the primer pairs described in Tables 3.1 and 3.2. Sequence comparisons showed that horses possessing the gene for sabino, both white and spotted, produced an abnormal *KIT* transcript lacking exon 17 in addition to the normal transcript. This deletion was apparent when PCR products for *KIT* exons 16-18 were compared on agarose gels for horses with and without the sabino pattern (Figure 2). As can be seen in lanes 1 and 4 of Figure 2, the horses without the sabino pattern produced a single band of approximately 356 base pairs (bp). Lane 2 shows the results from a horse with traditional sabino markings that exhibited a major band at 356 bp and a minor band at 233 bps. The RT-PCR fragments from a sample from a white (sabino-white) offspring of two sabino parents is shown in lane 3; the image shows two bands, specifically a fainter band 356 bp long and a brighter band 233 bp long.
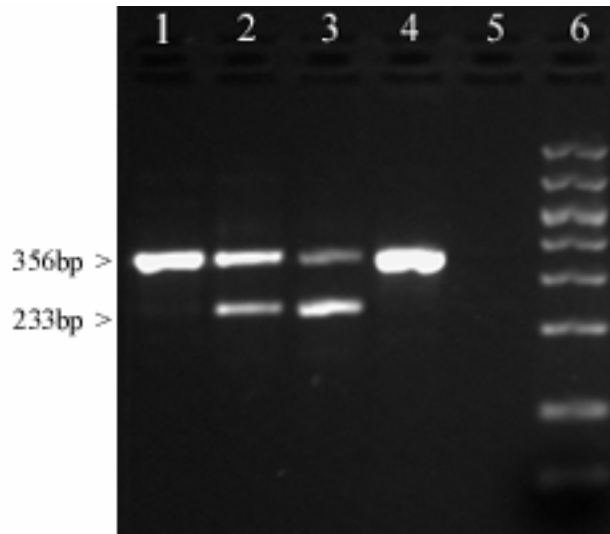
Figure 3.2: Agarose gel visualized with ethidium bromide, illuminated by ultraviolet light, showing the RT-PCR products of *KIT* exon 16-18. The cDNA band in lane 1 was from a horse with no sabino-type markings, lane 2 from a horse with *Sabino1* markings, lane 3 from a sabino-white colored horse (homozygous for *Sabino 1*), lane 4 another horse with no sabino-type markings, lane 5 a negative control, and lane 6 a size standard ladder. In lanes 1 and 4 only the normal-length 356-bp product is present. In lanes 2 and 3 the 233bp product lacking exon 17 can be seen in addition to the normal-length product.

Direct sequencing of the RT-PCR products confirmed that the difference between the larger and smaller fragments was the absence of exon 17 in the smaller fragment. Spot densiometry measurements of the intensity of the bands in the gel pictured (Figure 2) revealed that the ratio of normal products to those lacking exon 17 was close to 2:1 (1.95:1.00) for lane 2 (a horse with the sabino pattern). In lane 3, with DNA from the sabino-white individual, the ratio of normal to exon 17 skipped products was close to 1:2 (1.00:1.96). Densitometry readings were not taken for all samples tested and it is likely that the ratio observed in these tests would depend on the PCR conditions. However, the general principal was observed that samples from horses with *Sabino 1* patterns produced some product without exon 17 and that horses that appeared homozygous (sabino-white) produced mostly, but not entirely, the product without exon 17. No other polymorphisms were found among the *KIT* cDNA (bases 103-2909) associated with sabino. The cDNA

sequence was deposited in GenBank under accession numbers AY874543 and AY910688.

Sequencing of Genomic DNA

To determine the cause of the exon 17 skipping, the entire intron 16 and the first 290 bases of an estimated 5000 total in intron 17 were amplified and sequenced for 6 sabinos (including two sabino-whites) and 8 non-sabino horses. Within the 1049 bases of intron 16, seven SNPs were found, only one of which was present in all four horses known to have *KIT* transcripts lacking exon 17 (two sabino and two sabino-white) as well as the two additional sabino horses who had not been previously tested by RT-PCR for exon 17 skipping. This site was designated KI16+1037 denoting that it occurred at the *KIT* gene (K) intron (I) number 16 (16) at position 1037 (+1037). The SNP associated with sabino-white (KI16+1037), a T to A substitution, was 13 base pairs upstream of the beginning of exon 17 and resulted in the elimination of a *Mnl I* restriction site. This allowed the use of PCR-RFLP to quickly screen individuals for KI16+1037A or T. The sequence of intron 16 was deposited in GenBank under accession number AY874542. Two other SNPs were identified in intron 17; neither showed associations with exon 17 skipping.

PCR-RFLP

Digestion of DNA from horses without a sabino pattern (*sb1/sb1*) produced bands of 207, 74 and 47 bp in size while sabino-white individuals (*SB1/SB1*) only produced bands of 252 and 74 bp. DNA fragments from sabino patterned (*Sb1/sb1*) horses had bands of 252, 207, 74, and 47 bp (Figure 3). The presence of the 252 bp fragment corresponded to the presence of SNP KI16+1037A and the dominant *SB1* allele while the 207 bp band corresponded to the SNP KI16+1037T and the recessive allele *sb1*. The presence of the 74 bp band demonstrated that the restriction digest was successful regardless of which alleles were present.
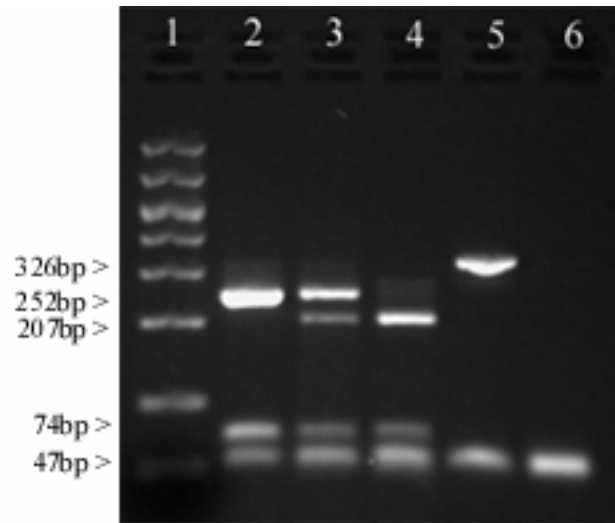
Figure 3.3: Agarose gel visualized with ethidium bromide showing the PCR-RFLP results for KI16+1037. Lane 1 contains a size standard ladder, lane 2 contains DNA from a sabino-white homozygous for *SB1* with only KI16+1037-A, lane 3 contains DNA from a horse with sabino markings and heterozygous for *SB1* with KI16+1037-A/T, lane 4 contains DNA from a horse without sabino-type markings (*sb1/sb1*) and with only KI16+1037-T, lane 5 contains undigested DNA from the PCR amplification, and lane 6 was a negative control for the PCR reaction.

Family Data and LOD scores

Three Tennessee Walking Horse families, segregating for the *Sabino 1* spotting pattern, were tested for the KI16+1037 RFLP. In all three families sabino was the only spotting pattern present. Tobiano, Frame Overo and Cream phenotypes were not observed and their associated PCR-RFLP alleles were not detected. LOD score calculations for the linkage of the gene including the KI16+1037A SNP and a sabino phenotype from the three families were combined for a total LOD score of 9.02 for $\Theta=0$ (Table 3.3).

Table 3.3a and 3.3b: LOD score results for Θ=0

3a: LOD Score from Sire Segregation

Sire			Dams			Offspring- Sire Contribution						
ID	Geno[a]	Pheno[b]	N	Geno[a]	Pheno[b]	A,SB	T,SB	A,sb	T,sb	N[c]	Uninf.[d]	LOD
MBF	A/T	SB/sb	8	T/T	sb/sb	7	0	0	5	12	0	3.61
597	A/T	SB/sb	3	T/T	sb/sb	3	0	0	4	7	0	2.11
597	A/T	SB/sb	2	A/T	SB/sb	2	0	0	2	7	3	1.20
									LOD Score for Sires =			6.92

3b: LOD Score from Dam Segregation

Sire			Dams			Offspring- Dam Contribution						
ID	Geno[a]	Pheno[b]	ID	Geno[a]	Pheno[b]	A,SB	T,SB	A,sb	T,sb	N[c]	Uninf[d]	LOD
597	A/T	SB/sb	601	A/T	SB/sb	0	0	0	1	2	1	0
597	A/T	SB/sb	ASB	A/T	SB/sb	2	0	0	1	5	2	0.90
600	T/T	sb/sb	606	A/T	SB/sb	2	0	0	0	2	0	0.60
600	T/T	sb/sb	610	A/T	SB/sb	1	0	0	1	2	0	0.60
									LOD Score for Dams=			2.10

a. KI16+1037 genotype
b. sabino spotting phenotype
c. Number of offspring
d. Number of uninformative meioses

Total LOD Score = 9.02

Breed distribution of KITI16+1037A (SB1)

A total of 320 horses from 13 breeds were tested for the KITI16+1037A SNP (Table 3.4). The SNP was found among American Miniature Horses, American Paint Horses, Azteca, Missouri Fox Trotters, Shetland Ponies, and Spanish Mustangs. All 5 horses homozygous for KI16+1037A were sabino-white. Among those horses possessing a single KITI16+1037A allele, all were either *Sabino 1*, multi-patterned, or sabino-white. Horses that were phenotypically white but heterozygous for the SNP also tested positive for other spotting patterns, namely *Tobiano* and *FRAME*. Therefore it appears that there is an additive effect between *SB1* and other white spotting loci. The KI16+1037A SNP was not found among Arabian horses (N=11), Thoroughbred horses (N=52) or American Standardbred horses (N=41).

Table 3.4: Results for PCR-RFLP typing of KI16+1037

Breeds	A/A	T/A	T/T	Total
American Miniature Horse	0	11	18	29
American Paint Horse	0	4	23	27
Arab	0	0	11	11
Azteca	0	1	2	3
Belgian	0	0	1	1
Cyldesdale	0	0	2	2
English Shire	0	0	3	3
Missouri Fox Trotter	1	6	13	20
Shetland Pony	0	3	6	9
Spanish Mustang	0	4	8	12
Standardbred	0	0	41	41
Tennessee Walking Horse	4	39	67	110
Thoroughbred	0	0	52	52
Total	**5**	**68**	**248**	**320**

RT-PCR and sequencing in other breeds

Two American Paint Horses, two Arabian, one Belgian, two Clydesdales, three English Shires and three Thoroughbreds with sabino-type phenotypes were tested by RT-PCR for exon 17 skipping. No transcripts lacking exon 17 were visible on an agarose gel after electrophoresis and staining with ethidium bromide.   One English Shire horse with sabino type spotting patterns lacked only exon 18 in a majority of *KIT* transcripts.

Analysis of splice site

The splice score for both the mutant and wild type 3' sequence of intron 16 was calculated using an online applet (Lauridsen *et al*., 2000) based on the formula developed by Shapiro and Senapathy (Shapiro and Senapathy, 1987).   The wild type splice site was stronger than the KI16+1037A containing site with scores of 79.32 and 77.14, respectively.

Discussion

When the KI16+1037 SNP was present, transcripts of the *KIT* gene were expressed without exon 17 in peripheral blood lymphocytes.   Although the SNP was 13bp upstream from the 3' splice site, KI16+1037A appears to reduce the overall strength of the 3' splicing consensus sequence without directly affecting the acceptor or branch

sites (Lauridsen *et al.*, 2000). Evidence for this notion is supported by the base conservation at this position reported for other genes and species. In a study of introns among mammals, the position 13 bases before splice junctions possessed T 41% of the time while A was least common, occurring only 5% of the time (Shapiro and Senapathy, 1987). Consequently, it appears that this base change could hinder the splicing machinery.

However, the loss of exon 17 was not complete. Even horses that were homozygous for the intron 16 SNP produced some product with exon 17 in their peripheral blood cells (Figure 2). This suggests that the mechanism for splicing exon 17 in the presence of this mutation is not completely abolished; even with this mutation, some *KIT* gene products including exon 17 can be produced. This differs from exon 17 skipping in white pigs. The mutation in that case abolishes the 5' donor site of intron 17, resulting in complete failure to recognize exon 17. In conjunction with a *KIT* gene duplication in the pig, exon 17 skipping was found to be associated with a reduced white blood cell count (Marklund *et al.*, 1998). Viable adult pigs carrying as high as 60% of *KIT* transcripts lacking exon 17 have been observed (Pielberg *et al.*, 2002), suggesting that the minimum percentage of normal *KIT* transcripts required for viability is less than 40%. It was theorized that the duplication provides a protective effect and that presence of the splice mutation without the duplication is lethal in pigs (Pielberg *et al.*, 2002). These findings have also demonstrated that homozygosity for this splice mutation is, on its own, sufficient to produce a white phenotype, independent of the duplication seen in pigs. Further work is needed to precisely quantify the expression of the *KIT* transcript lacking exon 17 in *Sabino 1* horses. Better quantification could be obtained using real time RT-PCR.

The presence of KI16+1037A was completely associated with the presence of one sabino-type phenotype we designate sabino 1, and encoded by *Sabino 1*, (*SB1*). Furthermore, in each case where a horse was homozygous for the SNP (and *SB1*), they had a complete or nearly complete white phenotype (sabino-white). This SNP is therefore potentially useful to identify *SB1* carriers. This is especially important in identifying phenotypically white horses that in fact carry multiple white spotting patterns rather than dominant white. The sabino pattern has also been frequently misidentified. It

has been suggested that many sabino Tennessee Walking horses are registered as "roan" (Sponenberg, 2003). The data in Table 4 demonstrates that the KI16+1037A allele is present in a wide variety of breeds from the United States. However, due to the limitations of ascribing phenotype based on a photograph, especially in multi-patterned individuals, our ability to draw conclusions relating the genotype of individuals to their apparent phenotype is limited.

Loss of exon 17, though it does not change the reading frame, may result in a partially or completely nonfunctional protein. Exon 17 encodes a portion of the second tyrosine kinase domain (TK2) including one of only three tyrosine residues in the TK2 that are conserved across the tyrosine kinase receptors (Rousset *et al*., 1995). Loss of this tyrosine could affect receptor function. While we observed this altered gene expression in blood, the effect on spotting pattern is probably a consequence of gene action in other tissues. Most likely, this alteration of gene expression affects migration or expression of melanocytes during early development.

Interestingly, sabino and sabino-white horses appear to have no health defects. Owners have not seen symptoms of anemia or deafness, and many sabino spotted horses are successful, highly valued breeding animals. Exon 17 skipping has been observed in piebald humans, with no hearing loss (Syrris *et al*., 2002).   Additional data is needed in order to draw any conclusions on the health effects of this altered transcript in the horse.

This study also demonstrated that the sabino phenotype has heterogeneous genetic origins as this SNP did not explain all sabino phenotypes. Presumably variation at other genetic sites within *KIT,* or another gene, is responsible for those sabino phenotypes. For example, Clydesdale horses are well known for their sabino phenotypes, but the mutation was not found among two horses of that breed exhibiting a sabino-type pattern. Genetic heterogeneity of sabino-type patterns is also supported by the observation that *Sabino 1* results in a sabino-white phenotype in the homozygous condition while white Clydesdale horses are uncommon despite the high frequency of a gene for a sabino-type pattern among horses of that breed (Guerts, 1977). Consequently, we anticipate future discoveries of other genes encoding different sabino phenotypes (possibly designated *SB2, SB3*, etc).

Dominant White was identified as a homozygous lethal among a horse herd in Nebraska (Pulos and Hutt, 1969). Recently, Mau and co-workers (2004) studied dominant white in Swiss Montagnes Horses and identified linkage with the *KIT* gene. Their definition of white horses included those with incomplete penetrance of the white gene expression.   In some respects their description was quite similar to that advanced by Gower, who expressed the opinion that most horses characterized as dominant white were really homozygous for a *Sabino* gene(Gower, 1999).   However, there are subtle differences in the description of the inheritance of the white phenotypes described by Pulos, Gower, and Mau (Pulos and Hutt, 1969; Gower, 1999; Mau *et al.*, 2004). In light of their present data, it would be instructive to test the white herd studied by Pulos and Hutt, as well as the families used by Mau and co-workers for the KI16+1037 SNP. However, this SNP was not observed in 5 thoroughbred horses described as possessing a dominant white phenotype.

In summary, the results of this study underline the importance of combining coat color studies with molecular research. In the face of heterogeneous genetic origins of similar color phenotypes, breeders will benefit greatly from the precision provided by development of molecular tests for the different coat color genes. The variant of *KIT* described in this study can explain some but not all the phenotypes described as "sabino". However, the KI16+1037A associated sabino pattern (*SB1*) was found in horses from a variety of diverse breeds.

CHAPTER FOUR: Discovery by FISH of a Paracentric Inversion of ECA3 Associated with Tobiano Spotting

Summary

The tobiano white spotting pattern is the product of a single dominant allele (*TO)* which was previously mapped to chromosome ECA3. *Tobiano* is characterized by large patches of white across the body and limbs. The genetic basis for *Tobiano* pattern has not yet been identified. Phase conservation of the haplotype *TO:ALB-B:GC-S* has been previously reported and is used to determine zygosity for *TO*. Based on this phase conservation it was hypothesized that Tobiano was due to a chromosomal inversion. We used florescent probes for thirteen genes and one intergenic sequence to discover a paracentric inversion on ECA3q. This inversion spans a region beginning in ECA3q13 and ending in ECA3q21. The breakpoints of this inversion occur between the *ADH1C* and *WDFY3* genes near the centromere and between the *KIT* and *KDR* genes at the end nearest the telomere. The homologous region in the human gene map (*ADH1C-KIT*) spans ~45million bases. Similar white spotting patterns in the mouse associated with genomic rearrangements near *KIT* have been shown to cause altered *KIT* gene expression leading to the spotted phenotype. It is hypothesized that this end of the inversion disrupts a regulatory sequence for *KIT* gene expression, thus causing the *Tobiano* pattern.

Introduction

A dominant trait, the tobiano spotting pattern in horses is characterized by distinct depigmented patches across the body and legs. Heterozygotes for *Tobiano* (*TO/to*) and homozygotes (*TO/TO*) are phenotypically indistinguishable. Tobiano is present in many diverse breeds of horse from around the world. Tobiano patterns, along with other coat spotting patterns, are valued in the horse for their aesthetic appeal. Breeders can increase their profits by increasing the percentage of their foals that are spotted. To better accomplish this horse breeders need a reliable method to genotype potential breeding stock and select for individuals homozygous for *TO*. *Tobiano* has been previously associated with a conserved haplotype of two blood protein markers on ECA 3: *Albumin* (*ALB*)-*B* and *Vitamin D binding factor* (*GC*)-*S* (Bowling, 1987). Although these markers can often be used to infer the *Tobiano* genotype of a horse, conservation of the opposite phase (*TO:ALB-A*) was observed with a frequency of 8% and can confound typing efforts (Duffield and Goldie, 1998). In 2002 we discovered a SNP marker within the *mast/stem cell growth factor receptor* (*KIT*) gene that is also in phase conservation with *Tobiano* (Brooks *et al*., 2002). This marker (KM1) was present in all 129 tobiano horses tested, including at least 2 in which *TO* was out of phase with *ALB-B* and/or *GC-S*. At the same time, the KM1 marker was excluded as the cause of *Tobiano* because it was also found at a very low frequency (0.03) in non-tobiano thoroughbreds. However, the KM1 marker can be used to detect zygosity with a higher precision than the *ALB-B* and *GC-S* markers.

To explain the unusually high level of linkage disequilibrium between the *Tobiano, ALB* and *GC* loci it has been proposed that an inversion on ECA 3 could be preventing recombination in this region (Bowling, 1987). A recombination event within the inverted region leads to the production of one acentric and one dicentric chromosome. This consequentially causes interference with meiosis resulting in gametes with abnormal chromosome content (Madan, 1995). Inversions have been documented in many other species, including humans. A survey of 446 paracentric inversions in humans found that less than half were associated with traits detrimental to health (Pettenati *et al*., 1995). The majority of paracentric inversions were discovered incidentally.

Several similar spotting patterns at the *W* locus in the mouse have been shown to be due to chromosomal rearrangements near the *KIT* gene. Two inversions (rump white (Stephenson *et al*., 1994) and sash (Nagle *et al*., 1995)) and four deletions, patch, 19h (Nagle *et al*., 1994), 57, and banded (Kluppel *et al*., 1997)) occur within 200 kb upstream of the *KIT* gene. All have been shown to disrupt the tissue-specificity or temporal expression of *KIT* during embryogenesis (Nagle *et al*., 1994; Kluppel *et al*., 1997; Hough *et al*., 1998; Berrozpe *et al*., 1999) rather than the coding sequence of the *KIT* gene. The similarities between tobiano and this particular group of mouse spotting patterns, previously demonstrated linkage of *Tobiano* and *KIT* (Brooks *et al*., 2002) and the lack of a difference in the *KIT* cDNA of tobiano horses (Brooks, this work, chapter two) all suggest that a nearby chromosomal rearrangement could be the cause of tobiano.

Materials and Methods

Horses

Eight horses were selected for cytogenetic analysis based on *Tobiano* phenotype and predicted *Tobiano* genotype using the associated SNP marker in the *KIT* gene (Brooks *et al*., 2002). Two were homozygous non-*Tobiano* (KM1 -/-), four heterozygous (KM1 +/-) and two homozygous *Tobiano* (KM1 +/+). One *TO/to* heterozygote possessed only the *AL-A* and *GC-F* alleles rarely found in phase with *Tobiano* (blood typing conducted by UC Davis and provided by owner). All *Tobiano* horses used were APHA registered with the exception of one *TO/to* heterozygote who was an unregistered miniature horse. Both non-*tobiano* horses were Thoroughbreds.

Cell culture and Slide Preparation

Approximately 0.5mL buffy coat from whole blood was cultured with phytohaemagglutinin and/or pokeweed mitogen and phorbol in RPMI-1640 with 10% fetal bovine serum. Cells were harvested and FISH experiments conducted according to standard protocols as previously described (Lear *et al*., 2001).

Probes

         Probes were chosen based on homologous position in the human sequence using the UCSC Genome Browser (http://genome.ucsc.edu/index.html)(Karolchik *et al*., 2003a). Previous work demonstrated that the majority of the q arm is contiguous between these two species (Chowdhary *et al*., 2003). Primers were designed based on published equine gene sequences and equine ESTs from NCBI Genbank when available, or consensus sequence from human, cow, dog and mouse from their draft genome sequences on the UCSC genome browser. The Vector NTI Suite 9 software package (InforMax Inc., Frederick, MD) was used for primer design. Primers for *CCNI*, LOC51170 (*retinal short-chain dehydro-genase/reductase, retSDR2*)(Chowdhary *et al*., 2003), and *Clock* (Murphy *et al.,* submitted) markers were from previously published mapping reports. Primers used to probe the CHORI-241 equine BAC library using PCR are listed in Table 4.1. Positive clones were grown under standard culturing conditions and their DNA extracted using the Perfectprep kit (Eppendorf, Westbury NY.) The Nick Translation Kit, SpectrumGreen-dUTP, SpectrumRed-dUTP, and SpectrumOrange-dUTP from Vysis Inc. (Downers Grove, IL.) were used to label the extracted BAC DNA following the manufacturers' recommendations. Abbreviations for all gene markers follow the human standards published on the UCSC Genome Browser. The two nonstandard abbreviations used are as follows: "558", conserved intergenic sequence corresponding to HSA 4q55.558Mb, and KDR5'Ue1, *kinase insert domain receptor* 5' untranslated sequence and first exon.

Table 4.1: Primers used to probe the BAC library
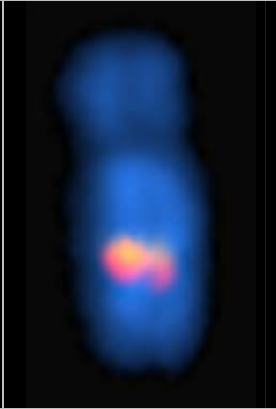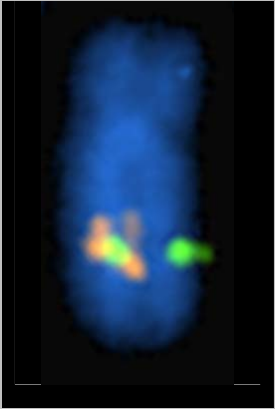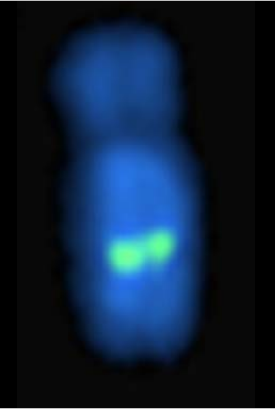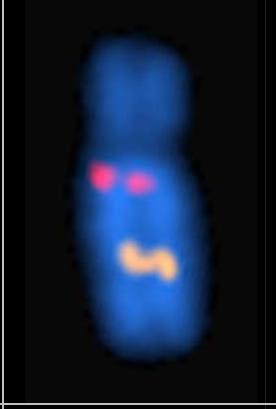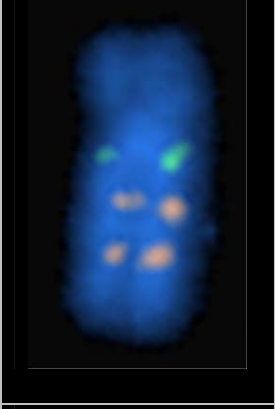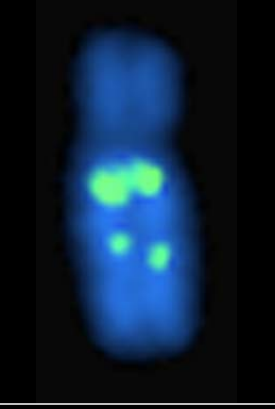
Marker	Source sequence	F Primer	F Seq 5'-3'	R Primer	R Seq 5'-3'
ADH1C	Genbank #M64864	ADH-F	GGTGATAAAGTCATCCCACTCTT	ADH-R	TCGATCTTGGCCACTGAGAT
ALB	Genbank #AY008769	ALB-F	ACCTGAGAGAAATGAATGCTTCC	ALB-R	CAAAAGCTTCCAAGTCTTAATGTTG
CCNI	Chowdhary *et al*. 2003	CCNI-F	GGTGCTCACCAAAGAAACATGTAAC	CCNI-R	CATCAGTACATGGTTCTGACCTGC
Clock	B. Murphy	ClockF1	TTTGATGGGTTGGTGGAAGAAGATGA	ClockR1	TGGGTCTATTGTTCCTCGCAGCAT
GABRB1	UCSC Consensus	GABRB1-3F	GGCTTCTCTCTTTCCCTGTGATGATTACCA	GABRB1-3F	CCGGCCGCAAGCGAATGTCATAT
Intergenic Seq. "558"	UCSC Consensus	558-1F	CACTTGGGTTTGTGTTTGGA	558-1R	TCTCCCTCTGTTACCATTAGAACC
KDR	UCSC Consensus	KDR-3F	CTCCATTTATTGCTTCTGTTAGTGACCA	KDR-3R	AACTTTCATCATTAATTTTTGCTTCACAGA
KDR5'Ue1	UCSC Consensus	KDR5u-1F	AGCTCAGTGTGGTCCCAGAG	KDRe1-1R	GAAGGCTCAAACCAGACGAG
KIT e21	Genbank #AY874543	KITe21-24F	CAGCCCCCGCCAGGAGAACT	KITstp-R	TCAGACATCTTCGTGGACAAGCAGAGG
LOC51170	Chowdhary *et al*. 2003	LOC51170-F	GCTCTGGCAGTCCTCGTTTCT	LOC51170-R	GCATTGAAATGTGAGGTATGTCTCAA
MASA	Genbank #CX603024	MASA-F	TGAAAATGGTGCATTGATGC	MASA-R	GGGCTCTTCTAGGTCTTATGGAGT
PDGFRA	UCSC Consensus	conPDGFRA-1188F	ARAGGCAAAGGCATCACMATG	conPDGFRA-1484R	CAGTGAGATTTTCAATCAGAGTCAGRTTG
PDLIM5	UCSC Consensus	PDLIM5-1F	GCCCACCAAGAAAACACATT	PDLIM5-1R	AGTGATCTGGGCGAGGATTC
TEC	UCSC Consensus	TEC-3F	GCCCAGTACAAAGTCGCAATCAAAGCTAT	TEC-3R	TCGGAGGAAATTCAGAAGGCAGCC
WDFY3	Genbank #CX599384	WDFY3-F	GCTGCCTGTAGCAGCCCATCCATCT	WDFY3-R	GCCTTTTACATAACCACCAAAGAACAGATATTAG

55

Results

        FISH analysis of metaphase chromosomes from *Tobiano* and non-*Tobiano* horses revealed that chromosomes from *Tobiano* horses carried a paracentric inversion on the q arm of chromosome 3. Figure 4.1 shows a sampling of the images collected. The left most set of chromosomes shows a clear difference in the distance between the red (*KDR)* and orange *(KIT)* markers on the normal and *Tobiano* chromosome. The center set of images illustrates three markers with different relative order on each chromosome. The marker found to cross the distal breakpoint ("558") hybridizes to a single point in the normal ECA3 (Figure 4.1, top right) and to two distinct locations in the *Tobiano* (bottom right). All four heterozygous *Tobiano* horses tested carried one normal and one inverted chromosome. Two homozygous *Tobiano* horses possessed only the chromosome containing the inversion. No evidence of an inversion was found in two non-tobiano horses. The markers used and their approximate chromosomal position is shown in Table 4.2.

        The telomeric end of the inversion is located between the *KIT* and *KDR* genes in band 3q21. The region between *KIT* and *KDR* in the human spans about 340 kilobases and contains no other annotated genes. The centromeric end of the inversion lies between the *ADH1C* and *UNC5C* genes in band 3q13. In the human, the region between these two genes is rather sparse. A contiguous section spanning nearly half of the 3.8 megabases in this region is entirely without known genes.

Figure 4.1: ECA3 FISH images from normal (top row) and *Tobiano* (middle row) chromosomes. The markers used are listed in the bottom row below the chromosomes (the text is in the same color with which the marker was labeled). The markers used are described in Table 4.1.

**Table 4.2:** Results for markers used in this study

Marker	CHORI-241 BAC Clones	HSA	ECA	Relation to Inversion	Sequence Source
GABRB1	49:M13, 70:P3	4q046.8mb	3q21	Telomeric	UCSC Consensus
TEC	38:G1	4q047.9mb	3q21	Telomeric	UCSC Consensus
PDGFRA	23:F11, 58:D10	4q054.9	3q21	Telomeric	UCSC Consensus
KIT e21	39:N17, 90: F8	4q055.447mb	3q21	Telomeric	Genbank #AY874543
Intergenic Seq. "558"	102:M1	4q055.558mb	3q21	At Breakpoint	UCSC Consensus
KDR5'Ue1	129:D4, 127:D23	4q055.787mb	3q21	Within	UCSC Consensus
KDR	127:D23	4q055.7mb	3q21	Within	UCSC Consensus
Clock	11:A9, 52:E21	4q056.2mb	3q21	Within	B. Murphy
ALB	21:K5, 18:J5	4q074.6mb	3q14.3	Within	Genbank #AY008769
CCNI	31:L21, M19 99:B9	4q078.4mb	3q14.2	Within	Chowdhary *et al.* 2003
MASA	69:C10, 130:C15	4q083.7mb	3q13	Within	Genbank #CX603024
WDFY3	44:L24	4q085.9mb	3q13	Within	Genbank #CX599384
LOC51170	105:B19	4q088.7mb	3q13	Within	Chowdhary *et al.* 2003
PDLIM5	19:G11, 23:E13	4q095.7mb	3q13	Within	UCSC Consensus
ADH1C	189:L20, 141:M23	4q100.6mb	3q13	Centromeric	Genbank #M64864

Discussion

Thirteen total genes and one intergenic sequence were mapped within and surrounding the inversion. The locations of 7 previously mapped (either RH or FISH) genes in the non-tobiano horses agreed with their previously reported map positions (Table 4.2). The normal gene order was conserved between horse and human as expected. Both ends of the inversion were located in a homologous human region that contains few, if any genes. Therefore, it seems unlikely that either breakpoint disrupts the coding sequence of a gene in the horse.

Previous work by Raudepp and coworkers did not identify this inversion, probably due to the choice of markers and absence of the necessary resolution (Raudsepp *et al.*, 1999). As can be seen in Figure 4.2, the inversion would not be apparent on G-banded chromosomes because the pattern of the inverted segment is symmetrical. The

ECA3 markers available at the time were also insufficient as three of the four were outside of the inversion and the remaining marker, *ALB* happens to be near the center. Though the orientation of *ALB* is in fact reversed, the change in relative distance from markers outside of the inversion is too small to be detectable with metaphase FISH.
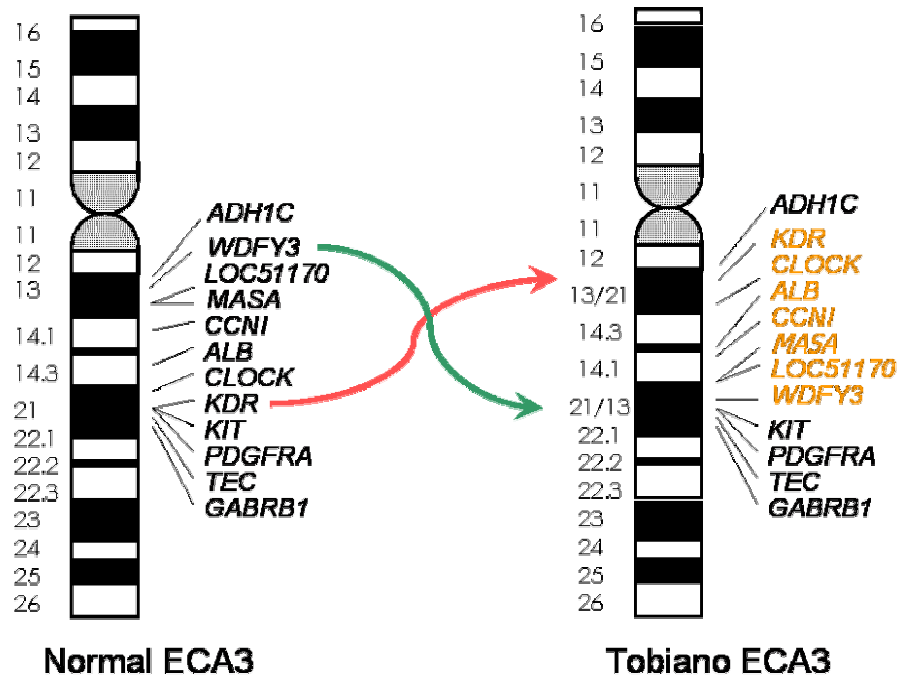
Figure 4.2: Diagram summarizing the markers used to characterized ECA3q and their relative positions.

The biochemical basis for the *Tobiano* spotting pattern requires further study. However, the position of the inversion could disrupt regulatory regions of the *KIT* gene similar to the rump-white (Stephenson *et al*., 1994) and sash (Nagle *et al*., 1995) inversions in the mouse. Both of these inversions disrupt regulatory elements near the *KIT* gene and lead to inappropriate expression in the embryo (Duttlinger *et al*., 1993; Nagle *et al*., 1995). The *KIT* gene has been linked to several other white patterns in the horse including roan *(*Marklund *et al., 1999)*, dominant-white (Mau *et al*., 2004), and one type of sabino (Brooks and Bailey, 2005). It seems likely that the *KIT* gene is indeed responsible for *Tobiano* despite the lack of a change in coding sequence (Brooks, this work, chapter two.)

In light of these findings, a cytogenetic test for this rearrangement would seem to be a more accurate diagnostic test for genotyping the trait than the biochemical or DNA detection systems currently in use. However, there are two important disadvantages to this approach. First, cytogenetic analysis is expensive and time consuming. Second, cytogenetic testing requires fresh tissues, rather than hair follicles or preserved blood commonly used for PCR. Breeders prefer the ability to collect hair rather than blood or other tissues for DNA analysis because it is faster, cheaper, and less stressful for their animals. Therefore, it will be useful to design a PCR test based on one of the two breakpoints of the inversion.

This inversion is the first chromosomal inversion described in a domestic animal species to be associated with a trait under positive selection. Selection for the *Tobiano* spotting pattern, and the associated inversion, by man has overcome any evolutionary disadvantages associated with meiosis in individuals heterozygous for the inversion. Recent work on a polymorphic inversion in man has attributed a selective advantage to its carriers as well (Stefansson *et al*., 2005). Women who carry that inversion actually have more children, rather than fewer as would be expected in individuals heterozygous for an inversion.

Horses with *Tobiano* are normal and healthy suggesting that this chromosome rearrangement is not associated with obvious deleterious health effects. Presence of this inversion in an individual from a pony breed indicates that this inversion and *Tobiano* predate divergence of the horse and pony breeds. Additionally, presence of the inversion in a horse with linkage to the less common phase containing *ALB-A:GC-F* alleles supports the idea that the *Tobiano* trait was acquired first, followed by the *ALB:GC* alleles, possibly by double recombination. While a single recombination event within the inversion in a heterozygote could result in a non-viable gamete, if recombination occurs twice simultaneously the normal chromosome structure would be maintained. In this way a *Tobiano* horse could have acquired the *ALB-A or GC-F* alleles despite the reduction in viable recombination between the three loci caused by the inversion. Individuals heterozygous for chromosomal inversions can have reduced fertility due to the formation of non-viable gametes (Madan, 1995). This may or may not be the case in the tobiano horse. However, it is unlikely that if the inversion does reduce fertility it

would be detected. First, the presence of the tobiano pattern is such a strong selective advantage that this likely overcomes any reduced reproductive efficiency in the population. Second, homozygotes are often chosen as breeding animals rather than heterzygotes due to the increased ability to produce spotted foals. Finally, the reproductive rate of the mare is relatively low compared to experimental species; a reduction due to chromosome abnormality would have to be sizeable in order to be detected by the breeder.

Inversions have also been shown to play an integral role in the derivation of species.   Rieseberg and Livingstone recently reviewed this topic (Rieseberg and Livingstone, 2003). They describe two different mechanisms by which chromosomal rearrangements have precipitated speciation between humans and chimpanzees, rather than resulted from it. Initially it was proposed that heterozygosity for inversions reduces fitness, thus reducing the ability of two populations differing by an inversion to interbreed. However, a reduction in fitness means rearrangements can not be established at significant frequencies to exert much of an effect, except for in highly inbred populations. Instead, they conclude that rearrangements drive speciation not by restricting fitness, but by restricting gene flow, thus allowing parallel evolution of two distinct chromosome homologues.

A polymorphic chromosome inversion at a high population frequency such as this one is in the horse could potentially be an interesting model for the effect of inversions on recombination rates and chromosome evolution. Moreover, this inversion demonstrates the importance of considering alternative genome organizations as contributors to genetic heterogeneity in particular breeds or populations.

CHAPTER FIVE: Implications and Potential Future Work

Summary

Coat color provides a clearly recognizable phenotype for genetic studies in horses. Molecular genetic studies on the sabino and tobiano color patterns demonstrated probable genetic basis for the traits based on changes in the *KIT* gene. Alternative splicing is suggested as the cause of *Sabino 1*. In the case of tobiano, the genetic cause could be related to disruption of *KIT* regulation by a large chromosome inversion. These two variants produce spotting patterns that are very attractive to some horse breeders and owners. However, the consequences of these *KIT* alterations bear investigation as well. One well established function of the KIT receptor is in modulating the development of mast cells through the action of its ligand, mast cell growth factor. Mast cells mediate important pathways in the innate immune system and the process of inflammation. While there are no known adverse health consequences for horses with these two *KIT* variants, an exon 17 splice polymorphism and the ECA3q paracentric inversion, they might provide an opportunity for investigation of alteration in the KIT receptor and its effect on mast cells.

Why Study Spotting Patterns in the Horse?

Gregor Mendel chose to study traits that could be easily phenotyped in order to grasp the fundamentals of heredity. Though the color of the bloom on the pea plant may not have affected the taste of the pea crop, it was none the less instrumental in describing the selective breeding that has resulted in the thousands of pea varieties we have today. Four out of the six traits mapped to the first chromosome map were for body and eye color (Morgan, 1915). Would Thomas Hunt Morgan have been able to construct his theories on chromosomes if his fruit flies didn't possess mutations in body and eye color? Probably, but he would have had a tougher time doing it. Advances in genetics have often been made with pigmentation variants. Coat color variants in horses provide an opportunity to link a clear genetic phenotype with molecular markers and genomic information.

Aside from the scientific benefits of studying coat color in the horse there is a more immediate advantage. People like the look of unusual coat colors, and they are willing to pay for a way to improve their ability to get more of them. The value in a flashy colored horse isn't as easily measured as that in a fast race horse, but it does translate in to real dollars at the point of sale. For example, the average price of an APHA registered yearling (young horses with undetermined athletic ability) without a spotting pattern on Equine.com was $1540. A yearling APHA registered horse with the tobiano pattern averaged $2803 (Appendix B). This price difference creates an enormous interest among breeders to increase their production of spotted foals. Interestingly, unlike studies performed on disease traits, breeders are more than willing to submit their horses for inclusion in research on positive traits like spotting patterns.

Overall, the horse industry is a large and important part of the U.S. agricultural community. Today there are over 9 million horses in the United States with a little over 3 million found in the 6 states (in order) Texas, California, Florida, Oklahoma, Kentucky and Ohio. In a study sponsored by the American Horse Council and published in 2005, the horse industry was reported to contribute $39.2 billion in direct economic impact in the U.S., have a $102 billion impact considering indirect factors, provide more than 453,000 full time jobs, generate 1.4 million more jobs indirectly when considering induced impacts, pays $1.9 billion in taxes annually, and stimulates an additional $4.1

billion in taxes (American Horse Council, 2005.) Research that benefits the economy of the horse industry, benefits the US economy as a whole.

Splice Variants as a Source of Genetic Variability

The initial estimate for the number of genes in the human genome was around a 100,000 based on EST data. The completed draft human genome sequence revealed that in fact there were only around 20,000 genes. How are only 20,000 genes able to do the job we thought was being done by 100,000? The answer probably lies within the regulatory sequences. Alternative exon splicing in particular is established as a means for one gene to encode multiple proteins carrying out multiple functions. It has been recently estimated that as many as 50% of protein coding genes undergo some form of alternative splicing (Stetefeld and Ruegg, 2005). Splicing polymorphisms also contribute significantly to diversity in the population. They have been implicated in diseases as variable as panic disorders (Battaglia and Ogliari, 2005), neurofibromatosis (Raponi *et al*., 2006), cardiac arrhythmia (Shang *et al*., 2005) and hemophilia (Chen *et al*., 1991). Splicing polymorphisms can have broad implications for the field of pharmacogenetics. For example, three splicing variants have been identified that lead to resistance to glucocorticoid drugs in humans (Bray and Cotton, 2003). In this study, a SNP was found in intron 16 of the equine *KIT* gene that was associated with skipping of exon 17. Other splice variants of *KIT* have been reported for horses, although the causes of these splice variants have not been identified nor any phenotype associated (Marklund *et al*., 1999). The *Sabino 1* variant is of special interest because it is polymorphic within the horse population. Identification of the *Sabino 1* splicing polymorphism in horses underlines the importance of considering more than just variation within the coding sequence when searching for causative mutations.

Structural Variation

The discovery of polymorphic inversions in humans through use of the human whole genome sequence and the HapMap project (www.genome.gov/10001688) has enlightened the scientific view of chromosome rearrangements. Previously, only those rearrangements associated with disease were detected in the population. Chromosome rearrangements as a whole were therefore viewed as a negative feature. It is now

64

apparent that far more rearrangements exist than previously thought, and that they may in some cases exert a positive force on selection. One study of a polymorphic inversion in Europeans has shown that women carrying the inversion have more children on average than those who do not, indicating that there may be a selective advantage associated with the inversion (Stefansson *et al.*, 2005).

In light of this new outlook on rearrangements, a new term is coming in to use: structural variation. To gain an idea of how common structural variations are in humans, skim sequencing projects will compare normal, healthy individuals to the reference complete human genome sequence. In one study already completed, sample genome sequencing for a single woman identified 297 novel structural variations, with 56 of these being inversions (Tuzun *et al.*, 2005). More comprehensive work is already underway to further characterize structural variation among people of more diverse origins. In the mouse, the majority of known structural variations have been discovered due to an association with a disease trait. Extensive genome wide studies to identify benign, polymorphic, structural variations in the mouse have not yet been conducted. The ECA3q paracentric inversion in the horse therefore represents, at the present time, the only known large, benign, inversion at a high population frequency in a model species. Inversions, like the one in *Tobiano* horses, may have been integral players in the processes of evolution and speciation as well. It is theorized that inversions drive divergence by restricting gene flow in the formation of large haplotype blocks (Rieseberg and Livingstone, 2003). A recent report comparing the human and available chimpanzee sequence revealed 23 confirmed inversions, 3 of which are still polymorphic in a sample human population today (Feuk *et al.*, 2005). Clearly, chromosome inversions may be a common polymorphic feature within populations, and it is incorrect to assume that there is only one genomic organization for a species.

Spotting Genes and Pleotropism

Coat color is more than skin deep. The genes that cause variation in coat color often also affect other functions. Examples of pleotropism between coat color traits and health are common. As was reviewed in chapter one, *KIT* mutant mice have been used to study a number of health defects from anemia to sterility. In the horse, the association of Overo Lethal White Foal syndrome and the frame overo pattern expedited the

identification of heterozygotes and the eventual discovery of the causative mutation. Even in humans, recent work has shown an association between an *MC1R* polymorphism (red hair color) and reduced response to morphine analgesia (Mogil *et al.*, 2005). Coat color variants, including spotting patterns can indicate changes in many underlying physiologic systems.

Theory on Expression of the KIT gene, Mast Cells, and Laminitis

These studies were focused on finding the cause of white coat color pattern variation in horses with the tobiano and sabino patterns. The *KIT* variation found among horses with the *Sabino 1* (*SB1)* allele is similar to that found in mast cell deficient mice (Tsujimura *et al.*, 1993). Horses carrying the *Sabino 1* variation of the *KIT* gene appear outwardly to be quite healthy; they are at least without serious anemia or sterility (Brooks and Bailey, 2005). The genetic and gene expression variants of *KIT*, found with and without white spotting pattern associations, may be one determinant for mast cell function and number in horses. Horses with *KIT* gene variations, specifically mast cell deficiency, could therefore be a valuable tool in studying the role of mast cells in disease.

The KIT receptor and its ligand (*MGF)* are the primary regulators for the proliferation, differentiation and migration of mast cells. In the mouse, research has shown a significant correlation between genetic variation in the expression of the *KIT* gene and mast cell deficiency (Tsujimura *et al.*, 1993). Mast cell deficiency on its own has few serious drawbacks, most notably a reduced ability to fight off bacterial and parasitic infection (Gruber and Kaplan, 2005).

Mast cells are a specialized subset of immune cells of bone marrow origin. Specifically, connective tissue mast cells (CTMCs) act as "gatekeepers of the microvasculature", are located in vascularized connective tissues, and are often a first defense against attack (Gruber and Kaplan, 2005). Mast cells respond to a variety of stimuli, just a few of which include toxins, bacteria, and mechanical injury (Gruber and Kaplan, 2005). Once activated, they quickly degranulate, releasing powerful preformed mediators including histamine, heparin, superoxide, tryptase, chymase, platelet activating factor, tumor necrosis factor alpha (TNF-α), and many inflammatory cytokines (Gruber and Kaplan, 2005). The mediators produced by CTMCs lead to rapid inflammation in

their resident tissues followed by recruitment of leukocytes. Though equipped to protect the body from a variety of assaults, mast cell activation, particularly in a confined area, can cause more damage than good (Lazarus *et al*., 2000). CTMCs are essential players in many different diseases including chronic allergy, asthma, ischemia (esp. heart attack and stroke), arthritis, cancer and autoimmune conditions including rheumatoid arthritis and multiple sclerosis.

Laminitis is a devastating disease of the horse. It is characterized by a breakdown of the laminae, the structure that suspends the distal phalanx within the hoof capsule. This results in destabilization of the foot leading to rotation and/or displacement of the distal phalanx. Accompanied by bruising and edema, it is extremely painful and debilitating. Vascular changes are associated with the disease, both in the form of increased regional blood flow and hypoperfusion of the capillary bed, which can lead to vascular damage and poor perfusion of the tissues (Moore *et al*., 2004). The triggering mechanisms for laminitis are varied and poorly understood. They can include carbohydrate overload (over-consumption of grain or lush pasture), black walnut toxicity, serious infections, hormonal imbalance, metabolic stress, mechanical injury and glucocorticoid administration (Animal Health Foundation, 2005).

Previously, two primary theories for the pathology of laminitis have been proposed (Hood, 1999; Moore *et al*., 2004). Ischemia and reperfusion (IR) injury form the first potential pathway. A toxin or bacterial product, such as endotoxins, acting as trigger factors is the second. The unifying theme underlying both pathways is vascular change leading to damage of the lamellar tissues, characterized by enzymatic activity and cytokine release. Regardless of the triggering event, be it physical trauma, obesity/metabolic stress, endocrine imbalance, or bacterial assault, the end result is the same. Mast cells play a central role in inciting tissue damage during inflammatory responses like those seen in laminitis.

A role for mast cells in vascular disease precipitated by obesity or metabolic dysregulation in humans has recently come to light. Angina and myocardial infarction as a result of systemic mast cell degranulation following allergic reaction has been well documented (Kounis, 2005). In fact, tracking of mast cell activation by measuring serum tryptase has recently shown promise for diagnostic use in coronary artery disease

(Deliargyris *et al*., 2005). Locally increased mast cell numbers are commonly seen in the pathogenesis of atherosclerosis (Kaartinen *et al*., 1994). In horses, chronic laminitis is a frequent complication of pituitary dysfunction and Cushings syndrome (equine metabolic syndrome) in the horse (Donaldson *et al*., 2004; McGowan and Neiger, 2003). Recent work has shown significantly higher numbers of mast cells in adipose tissue from human subjects with similar metabolic syndromes compared to healthy controls (Chaldakov *et al*., 2001). Hyperglycemia, a significant characteristic of equine metabolic syndrome, has been shown experimentally to lead directly to degranulation of perivascular mast cells (Johnson, 2003; Kalichman *et al*., 1995). One study identified pituitary dysfunction, characterized by abnormally high levels of adrenocorticotropic hormone (ACTH), in 70% of cases of chronic laminitis (Donaldson *et al*., 2004). Several different hormones secreted by the pituitary, including ACTH and endogenous glucocorticoids, activate mast cells directly (Paus *et al*., 2005). Hormonal activation of mast cells is linked to several stress induced inflammatory conditions in humans including psoriasis (Paus *et al*., 2005) and migraines (Theoharides *et al*., 2005). The innate sensitivity of mast cells to glucocorticoids may explain the association of laminitis with the therapeutic administration of exogenous glucocorticoids for unrelated conditions (Johnson *et al*., 2002).

Current research in laminitis has provided additional evidence that CTMCs could be a pivotal part of the disease process. Recently, Black and colleagues showed that the emigration of leukocytes is a significant early step in development of laminitis (Black *et al*., 2005). Mast cells release several chemotactic factors that direct the infiltration of leukocytes in to the site of inflammation (Gruber and Kaplan, 2005). Pharmacologic stabilization of CTMCs reduces the expression of adhesion molecules on circulating leukocytes (Zhao *et al*., 2005). Additionally, neutrophil infiltration in to wounds is significantly reduced in mast cell deficient mice, though wound healing was otherwise normal (Egozi *et al*., 2003). Mast cell chymases are crucial in the activation of matrix metalloproteinase (MMP)-2 and MMP-9, two central enzymes in the destruction of lamellar attachments (Tchougounova *et al*., 2005; Mungall *et al*., 2001). Most interestingly, chronically laminitic horses were recently shown to have significantly higher responses to intradermal allergen tests compared to unaffected horses (Wagner *et*

al., 2003). CTMCs are vital to the immediate response to allergens, and their degranulation is responsible for the wheal formation measured in allergen tests (Gruber and Kaplan, 2005).

Observations of histological (Hood *et al*., 1993) and hoof temperature changes (van Eps and Pollitt, 2004) support the role of reduced digital perfusion followed by reperfusion injury (IR injury) in the pathogenesis of laminitis. Though Van Eps and Pollitt concluded that the success of cyrotherapy in treating the early stages of laminitis (despite the associated cold induced hypoperfusion) suggests that ischemia is not involved in the development of laminitis (van Eps and Pollitt, 2004) this observation fits well in the mast cell mediated IR injury model.   In fact the protective effect of hypothermia against IR injury is well documented in a number of situations (Mowlavi *et al*., 2003; Takata *et al*., 2005; Stefanutti *et al*., 2005).

A bacterial mechanism for the initiation of laminitis has also been proposed. In cases of experimentally induced endotoxemia in the horse changes in vascular reactivity and reduced digital perfusion have been observed (Ingle-Fehr and Baxter, 1998; Zerpa *et al*., 2005). Vascular changes associated with endotoxemia could lead to IR injury in the delicate vasculature of the hoof capsule. An important element in innate immunity, CTMCs can also be directly activated by many different bacterial products including LPS (Gruber and Kaplan, 2005).

If mast cell numbers are locally, as well as systemically, elevated in individuals suffering from chronic laminitis, then measurements of mast cell numbers could prove to be a valuable diagnostic tool. Either blood protein or gene expression measurements could easily be adapted for clinical use. Measuring the level of mast cell elevation could give the clinician insight into the success of their treatment strategy in stabilization and reduction of inflammation in the foot. Elevated mast cell numbers could also identify those horses most likely to experience a relapse of inflammation. Potentially, elevated mast cells numbers could be used to identify individuals in the pre-clinical stages of laminitis. The ability to determine which horses are at risk for developing laminitis, and intervene by changing management practices or preemptive anti-inflammatory treatment, would be a huge development in the battle against laminitis. This could be the first tool available to clinicians aimed at prevention of laminitis.

Current treatments for laminitis focus on reducing pain and damage and can do little to prevent reoccurrence. Though removing known trigger factors can help, there is a desperate need for a way to gain specific control of the inflammatory processes within the hoof. Increased mast cell numbers could not only indicate, but also contribute to the inflammation associated with laminitis. Hypersensitivity to inflammation has already been demonstrated in the chronically laminitic horse (Wagner *et al*., 2003). This hypersensitivity likely contributes to the chronicity of the disease. Many drugs have been developed to control both the number and sensitivity of mast cells in order to combat the myriad of diseases to which they contribute. Several different plant extracts and even ethanol have been shown to have mast cell stabilizing activity (Harish *et al*., 2001; Kim *et al*., 2003; Toivari *et al*., 2000). Genetic or pharmacological induction of mast cell deficiency does have a few disadvantages, but it is most remarkable because of its protective effect against inflammatory and allergic disease. Recent work in the mouse has shown that genetic mast cell deficiency can reduce the tissue damage caused by IR injury by as much as 85% (Lazarus *et al*., 2000). One mast cell stabilizing drug has been shown to control inflammation related increases in blood flow (Ruh *et al*., 2000), and reduce damage and improve survival following intestinal IR injury (Kalia *et al*., 2005). The failure of laminae to heal properly from the damage caused by laminitis leads to the formation of fibrous scar tissue in the hoof and may affect the future stability of the foot (Mobasheri *et al*., 2004). Experimental pharmacologic stabilization of mast cells in rats reduces myocardial fibrosis following myocarditis and vascular fibrosis in experimentally induced diabetes (Palaniyandi Selvaraj *et al*., 2005; Jones *et al*., 2004). The same drug has been proposed as a novel treatment of insulin resistance in man (Namazi and Soma, 2005). Pharmacologic treatment targeting mast cells likewise reduced development of fibrosis during the healing of burns (Jones *et al*., 2004). Adaptation of one of these available drugs would provide help potentially decades earlier rather than developing a novel compound against one of the molecular events within the hoof.

Laminitis in horses is one of the most common causes of death. The precise etiology of the disease is unknown although it can occur following a wide variety of inflammatory stimuli. Finding a method to prevent the condition or to effectively treat it in the early stages of disease will alleviate considerable pain for the animal, as well as

recover some use of animals that would otherwise be entirely lost. This disease is common among riding and performance horses and has been responsible for the deaths of such well-known horses as Thoroughbred racing stallion, Secretariat, after he retired to stud. Though the vascular component of the pathological course of laminitis is well documented (Deliargyris *et al.*, 2005), and preliminary work has suggested systemic mast cell involvement (Wagner *et al.*, 2003), the role of the mast cell in laminitis, and the use of the *KIT* gene as a marker for the mast cell in the horse has yet to be investigated.

APPENDIX A: Some Labs Offering Genetic Testing Services in the US

Company	Contact Info	Web Site	Tests Offered
Animal Genetics, Inc.	1336 Timberlane Road Tallahassee, FL 32312-1766 USA 1-866-922-6436	http://www.animalgenetics.us	OLWS Tobiano HYPP Cream Agouti Red Factor (MC1R)
Pet DNA Services of AZ	PO Box 7809 Chandler, AZ 85246-7809 602-380-8552	http://www.petdnaservicesaz.com/Equine.html	Red Factor (MC1R) Agouti Cream OLWS Tobiano HYPP
Shelterwood Labs	626 Bear Drive Timpson, TX 75975 936-254-2228	http://www.dnadiagnostics.com	OLWS Chesnut (MC1R) Sabino 1 Tobiano Agouti Cream
The University of Tampa, Dr. Rebecca Bellone	Box U 401 W. Kennedy Blvd. Tampa, FL. 33606-1490 813-253-3333 X3551	http://www.ut.edu/directory/bio_detail.cfm?userid=77949l	Red Factor (MC1R) Agouti OLWS Tobiano HYPP Sabino 1
University of California, Davis-Veterinary Genetics Laboratory	Veterinary Genetics Laboratory One Shields Avenue Davis, CA 95616-8744 (530) 752-2211	http://www.vgl.ucdavis.edu	Red Factor (MC1R) Agouti Cream OLWS Sabino 1 Tobiano HYPP Junctional Epidermolysis Bullosa (JEB) Glycogen Branching Enzyme Deficiency (GBED)
University of Kentucky Equine Parentage Testing and Research Laboratory	101 Dimock Animal Pathology Building University of Kentucky Lexington, KY 40546-0076 USA 859-257-1165	http://www.ca.uky.edu/gluck/EPTRL.asp	Tobiano E locus (MC1R) Agouti Cream Sabino 1

APPENDIX B: Brief Survey of Prices for APHA Registered Yearlings

Table A.1a and 1b: Average sale price for Solid APHA yearlings (Table 1a) vs. Tobiano
APHA yearlings (1b) as advertised on Equine.com

1a. Solid Horses			1b. Tobiano Horses	
Big Sonny Mac	$3,000.00		Registered coming…	$2,000.00
Peptoboonsmal grandson…	$1,500.00		Sheza Attraction	$7,200.00
Flashy…	$500.00		Sheza Spicy Bandit	$4,500.00
Ima Sonny Investment	$1,100.00		Gorgeous registered…	$3,500.00
This Kids at the Bar	$750.00		Yearling APHA…	$1,800.00
Longe line…	$1,000.00		Bullseye	$3,500.00
VBR Sky Lilly	$1,800.00		For sale…	$1,500.00
Daughter of delta…	$1,750.00		APHA-breeders	$3,500.00
Gorgeous athletic…	$1,500.00		Geldings for sale…	$700.00
Mac Dunit Cajun Style	$2,500.00		BlueEyed Cherokee	$2,750.00
Average Price	**$1,540.00**		Two Lady	$1,500.00
			Naomi	$1,000.00
			Bay Tobiano Gelding…	$3,000.00
			Average Price	**$2,803.85**

In order to quickly get an idea of what value a spotting pattern adds to a horse, a
sample group was taken from ads listed on Equine.com, a nationwide horse sale site
(Primedia Company). All horses were unsold, registered with the APHA, and between 1
and 2 years of age. Ads were those listed during the week of March 10th through the
17th, 2006. Individuals were recorded by either their registered name or the first few
words of the ad. After removing duplicate ads for the same horse the prices were
averaged.

BIBLIOGRAPHY

American Paint Horse Association: *Guide to Coat Color Genetics*.
http://www.apha.com/forms/PDFFiles/guidebooks/05ColorGen.pdf, 2005.

American Horse Council: *Horsepower*. American Horse Council, Washington D.C.,
2005.

Andre, C., Martin, E., Cornu, F., Hu, W.X., Wang, X.P. and Galibert, F.: Genomic
organization of the human c-kit gene: evolution of the receptor tyrosine kinase
subclass III. *Oncogene* 7 (1992) 685-91.

Animal Health Foundation: Laminitis-Founder Syndrome.
http://www.animalhealthfoundation.com/pages/laminitis.html, 2005.

Ashman, L.K.: The biology of stem cell factor and its receptor C-kit. *Int J Biochem Cell
Biol* 31 (1999) 1037-51.

Battaglia, M. and Ogliari, A.: Anxiety and panic: from human studies to animal research
and back. *Neurosci Biobehav Rev* 29 (2005) 169-79.

Berrozpe, G., Timokhina, I., Yukl, S., Tajima, Y., Ono, M., Zelenetz, A.D. and Besmer,
P.: The W(sh), W(57), and Ph Kit expression mutations define tissue-specific
control elements located between -23 and -154 kb upstream of Kit. *Blood* 94
(1999) 2658-66.

Besmer, P., Murphy, J.E., George, P.C., Qiu, F.H., Bergold, P.J., Lederman, L., Snyder,
H.W. Jr, Brodeur, D., Zuckerman, E.E. and Hardy, W.D.: A new acute
transforming feline retrovirus and relationship of its oncogene v-kit with the
protein kinase gene family. *Nature* 320 (1986) 415-21.

Black, S.J., Lunn, D.P., Yin, C., Hwang, M., Lenz, S.D. and Belknap, J.K.: Leukocyte
emigration in the early stages of laminitis. *Vet Immunol Immunopathol* (2005)

Blake, J.A., Richardson, J.E., Bult, C.J., Kadin, J.A. and Eppig, J.T.: MGD: the Mouse
Genome Database. *Nucleic Acids Res* 31 (2003) 193-5.

Blume-Jensen, P., Claesson-Welsh, L., Siegbahn, A., Zsebo, K.M., Westermark, B. and
Heldin, C.H.: Activation of the human c-kit product by ligand-induced
dimerization mediates circular actin reorganization and chemotaxis. *EMBO J* 10
(1991) 4121-8.

Bowling, A.T.: Equine linkage group II: phase conservation of To with AlB and GcS. *J
Hered* 78 (1987) 248-50.

Bray, P.J. and Cotton, R.G.: Variations of the human glucocorticoid receptor gene (NR3C1): pathological and in vitro mutations and polymorphisms. *Hum Mutat* 21 (2003) 557-68.

Brooks, S.A. and Bailey, E.: Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. *Mamm Genome* 16 (2005) 893-902.

Brooks, S.A., Terry, R.B. and Bailey, E.: A PCR-RFLP for KIT associated with tobiano spotting pattern in horses. *Anim Genet* 33 (2002) 301-3.

Brown, T.A.: *Genomes*. Wiley-Liss, New York, 1999.

Chaldakov, G.N., Fiore, M., Stankulov, I.S., Hristova, M., Antonelli, A., Manni, L., Ghenev, P.I., Angelucci, F. and Aloe, L.: NGF, BDNF, leptin, and mast cells in human coronary atherosclerosis and metabolic syndrome. *Arch Physiol Biochem* 109 (2001) 357-60.

Chen, S.H., Zhang, M., Thompson, A.R., Bray, G.L. and Scott, C.R.: Splice junction mutations in factor IX gene resulting in severe hemophilia B. *Nucleic Acids Res* 19 (1991) 1172

Chowdhary, B.P., Raudsepp, T., Kata, S.R., Goh, G., Millon, L.V., Allan, V., Piumi, F., Guerin, G., Swinburne, J., Binns, M., Lear, T.L., Mickelson, J., Murray, J., Antczak, D.F., Womack, J.E. and Skow, L.C.: The first-generation whole-genome radiation hybrid map in the horse identifies conserved segments in human and mouse genomes. *Genome Res* 13 (2003) 742-51.

Dahlen, D.D., Lin, N.L., Liu, Y., and Broudy, V.C.: Soluble c-kit receptor blocks stem cell factor bioactivity in vitro. *Leukemia Research* 25 (2001) 413-21.

Deliargyris, E.N., Upadhya, B., Sane, D.C., Dehmer, G.J., Pye, J., Smith, S.C. Jr, Boucher, W.S. and Theoharides, T.C.: Mast cell tryptase: a new biomarker in patients with stable coronary artery disease. *Atherosclerosis* 178 (2005) 381-6.

Donaldson, M.T., Jorgensen, A.J. and Beech, J.: Evaluation of suspected pituitary pars intermedia dysfunction in horses with laminitis. *J Am Vet Med Assoc* 224 (2004) 1123-7.

Duffield, D.A. and Goldie, P.L.: Tobiano spotting pattern in horses: linkage of To with AlA and linkage disequilibrium. *J Hered* 89 (1998) 104-6.

Duttlinger, R., Manova, K., Chu, T.Y., Gyssler, C., Zelenetz, A.D., Bachvarova, R.F. and Besmer, P.: W-sash affects positive and negative elements controlling c-kit expression: ectopic c-kit expression at sites of kit-ligand expression affects melanogenesis. *Development* 118 (1993) 705-17.

Egozi, E.I., Ferreira, A.M., Burns, A.L., Gamelli, R.L. and Dipietro, L.A.: Mast cells modulate the inflammatory but not the proliferative response in healing wounds. *Wound Repair Regen* 11 (2003) 46-54.

Erickson, C.A.: From the crest to the periphery: control of pigment cell migration and lineage segregation. *Pigment Cell Res* 6 (1993) 336-47.

Feuk, L., MacDonald, J.R., Tang, T., Carson, A.R., Li, M., Rao, G., Khaja, R. and Scherer, S.W.: Discovery of human inversion polymorphisms by comparative analysis of human and chimpanzee DNA sequence assemblies. *PLoS Genet* 1 (2005) e56

Fleischman, R.A., Gallardo, T. and Mi, X.: Mutations in the ligand-binding domain of the kit receptor: an uncommon site in human piebaldism. *J Invest Dermatol* 107 (1996) 703-6.

Geissler, E.N., McFarland, E.C. and Russell, E.S.: Analysis of pleiotropism at the dominant white-spotting (W) locus of the house mouse: a description of ten new W alleles. *Genetics* 97 (1981) 337-61.

Geissler, E.N. and Russell, E.S.: Analysis of the hematopoietic effects of new dominant spotting (W) mutations of the mouse. II. Effects on mast cell development. *Exp Hematol* 11 (1983) 461-6.

Geissler, E.N., Ryan, M.A. and Housman, D.E.: The dominant-white spotting (W) locus of the mouse encodes the c-kit proto-oncogene. *Cell* 55 (1988) 185-92.

Geurts, R.: *Hair Colour in the Horse*. Sporting Book Center, Inc., Canaan N.Y., 1977.

Gokkel, E., Grossman, Z., Ramot, B., Yarden, Y., Rechavi, G. and Givol, D.: Structural organization of the murine c-kit proto-oncogene. *Oncogene* 7 (1992) 1423-9.

Gower, J.: *Horse Color Explained*. Trafalgar Square Publishing, North Pomfret, Vermont, 1999.

Greber, B., Lehrach, H. and Himmelbauer, H.: Mouse splice mutant generation from ENU-treated ES cells--a gene-driven approach. *Genomics* 85 (2005) 557-62.

Gruber, B.L. and Kaplan, A.P.: Mast Cells and Rheumatic Diseases. In Koopman, W.J. and Moreland, L.W. (Eds.), *Arthritis and Allied Conditions: a textbook of rheumatology*. Lippincott Williams & Wilkins, Philadelphia, 2005, pp.393-423.

Harish, M.S., Nagur, M. and Badami, S.: Antihistaminic and mast cell stabilizing activity of Striga orobanchioides. *J Ethnopharmacol* 76 (2001) 197-200.

Hood, D.M.: The pathophysiology of developmental and acute laminitis. *Vet Clin North Am Equine Pract* 15 (1999) 321-43.

Hood, D.M., Grosenbaugh, D.A., Mostafa, M.B., Morgan, S.J. and Thomas, B.C.: The role of vascular mechanisms in the development of acute equine laminitis. *J Vet Intern Med* 7 (1993) 228-34.

Hough, R.B., Lengeling, A., Bedian, V., Lo, C. and Bucan, M.: Rump white inversion in the mouse disrupts dipeptidyl aminopeptidase-like protein 6 and causes dysregulation of Kit expression. *Proc Natl Acad Sci* U S A 95 (1998) 13800-5.

Ingle-Fehr, J.E. and Baxter, G.M.: Evaluation of digital and laminar blood flow in horses given a low dose of endotoxin. *Am J Vet Res* 59 (1998) 192-6.

Ito, M., Kawa, Y., Ono, H., Okura, M., Baba, T., Kubota, Y., Nishikawa, S.I. and Mizoguchi, M.: Removal of stem cell factor or addition of monoclonal anti-c-KIT antibody induces apoptosis in murine melanocyte precursors. *J Invest Dermatol* 112 (1999) 796-801.

Johnson, P.J., Slight, S.H., Ganjam, V.K. and Kreeger, J.M.: Glucocorticoids and laminitis in the horse. *Vet Clin North Am Equine Pract* 18 (2002) 219-36.

Johnson, P.J.: Peripheral Cushingoid Syndrome ('Equine Metabolic Syndrome"). In Robinson, E.N. (Ed.), *Current Therapy in Equine Medicine*. Saunders, New York, 2003, pp.812-6.

Jones, S.E., Gilbert, R.E. and Kelly, D.J.: Tranilast reduces mesenteric vascular collagen deposition and chymase-positive mast cells in experimental diabetes. *J Diabetes Complications* 18 (2004) 309-15.

Kaartinen, M., Penttila, A. and Kovanen, P.T.: Accumulation of activated mast cells in the shoulder region of human coronary atheroma, the predilection site of atheromatous rupture. *Circulation* 90 (1994) 1669-78.

Kalia, N., Brown, N.J., Wood, R.F. and Pockley, A.G.: Ketotifen abrogates local and systemic consequences of rat intestinal ischemia-reperfusion injury. *J Gastroenterol Hepatol* 20 (2005) 1032-8.

Kalichman, M.W., Powell, H.C., Calcutt, N.A. and Mizisin, A.P.: Mast cell degranulation and blood-nerve barrier permeability in rat sciatic nerve after 7 days of hyperglycemia. *Am J Physiol* 268 (1995) H740-8.

Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., Weber, R.J., Haussler, D. and Kent, W.J.: The UCSC Genome Browser Database. *Nucleic Acids Res* 31 (2003) 51-4.

Kent, W. J. BLAT - The BLAST-Like Alignment Tool. http://genome.ucsc.edu/cgi-bin/hgBlat?command=start, 2002.

Kim, M.S., Na, H.J., Han, S.W., Jin, J.S., Song, U.Y., Lee, E.J., Song, B.K., Hong, S.H. and Kim, H.M.: Forsythia fructus inhibits the mast-cell-mediated allergic inflammatory reactions. *Inflammation* 27 (2003) 129-35.

Kluppel, M., Nagle, D.L., Bucan, M. and Bernstein, A.: Long-range genomic rearrangements upstream of Kit dysregulate the developmental pattern of Kit expression in W57 and Wbanded mice and interfere with distinct steps in melanocyte development. *Development* 124 (1997) 65-77.

Kounis, N.G.: Kounis syndrome (allergic angina and allergic myocardial infarction): A natural paradigm? *Int J Cardiol* 21 (2005).

Lauridsen, C., D'Alberti, T., and Robinson, P. Shapiro and Senapathy Splice Score: An Applet. **http://home.snafu.de/probins/Splice/splice.html,** 2000.

Lazarus, B., Messina, A., Barker, J.E., Hurley, J.V., Romeo, R., Morrison, W.A. and Knight, K.R.: The role of mast cells in ischaemia-reperfusion injury in murine skeletal muscle. *J Pathol* 191 (2000) 443-8.

Lear, T.L., Brandon, R., Piumi, F., Terry, R.R., Guerin, G., Thomas, S. and Bailey, E.: Mapping of 31 horse genes in BACs by FISH. *Chromosome Res* 9 (2001) 261-2.

Mackenzie, M.A., Jordan, S.A., Budd, P.S. and Jackson, I.J.: Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev Biol* 192 (1997) 99-107.

Madan, K.: Paracentric inversions: a review. *Hum Genet* 96 (1995) 503-15.

Mariat, D., Taourit, S. and Guerin, G.: A mutation in the MATP gene causes the cream coat colour in the horse. *Genet Sel Evol* 35 (2003) 119-33.

Marklund, S., Kijas, J., Rodriguez-Martinez, H., Ronnstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I. and Andersson, L.: Molecular basis for the dominant white phenotype in the domestic pig. *Genome Res* 8 (1998) 826-33.

Marklund, S., Moller, M., Sandberg, K. and Andersson, L.: Close association between sequence polymorphism in the KIT gene and the roan coat color in horses. *Mamm Genome* 10 (1999) 283-8.

Mau, C., Poncet, A., Bucher, B., Stranzinger, G. and Rieder, S.: Genetic mapping of dominant white (*W*), a homozygous lethal condition in the horse (*Equus caballus*). *J Anim. Breed. Genet.* 121 (2004) 374-383.

Mayer, T.C.: The migratory pathway of neural crest cells into the skin of mouse embryos. *Dev Biol* 34 (1973) 39-46.

McCabe, L., Griffin, L.D., Kinzer, A., Chandler, M., Beckwith, J.B. and McCabe, E.R.: Overo lethal white foal syndrome: equine model of aganglionic megacolon (Hirschsprung disease). *Am J Med Genet* 36 (1990) 336-40.

McGowan, C.M. and Neiger, R.: Efficacy of trilostane for the treatment of equine Cushing's syndrome. *Equine Vet J* 35 (2003) 414-8.

Metallinos, D.L., Bowling, A.T. and Rine, J.: A missense mutation in the endothelin-B receptor gene is associated with Lethal White Foal Syndrome: an equine version of Hirschsprung disease. *Mamm Genome* 9 (1998) 426-31.

Milenkovic, D., Oustry-Vaiman, A., Lear, T.L., Billault, A., Mariat, D., Piumi, F., Schibler, L., Cribiu, E. and Guerin, G.: Cytogenetic localization of 136 genes in the horse: comparative mapping with the human genome. *Mamm Genome* 13 (2002) 524-34.

Mobasheri, A., Critchlow, K., Clegg, P.D., Carter, S.D. and Canessa, C.M.: Chronic equine laminitis is characterised by loss of GLUT1, GLUT4 and ENaC positive laminar keratinocytes. *Equine Vet J* 36 (2004) 248-54.

Mogil, J.S., Ritchie, J., Smith, S.B., Strasburg, K., Kaplan, L., Wallace, M.R., Romberg, R.R., Bijl, H., Sarton, E.Y., Fillingim, R.B. and Dahan, A.: Melanocortin-1 receptor gene variants affect pain and mu-opioid analgesia in mice and humans. *J Med Genet* 42 (2005) 583-7.

Moore, R.M., Eades, S.C. and Stokes, A.M.: Evidence for vascular and enzymatic events in the pathophysiology of acute laminitis: which pathway is responsible for initiation of this process in horses? *Equine Vet J* 36 (2004) 204-9.

Morgan, T.H., Sturtevant, A.H., Muller, H.J. and Bridges, C.B.: *The Mechanism of Mendelian Heredity*. Henry Holt and Company, New York, 1915.

Mowlavi, A., Neumeister, M.W., Wilhelmi, B.J., Song, Y.H., Suchy, H. and Russell, R.C.: Local hypothermia during early reperfusion protects skeletal muscle from ischemia-reperfusion injury. *Plast Reconstr Surg* 111 (2003) 242-50.

Mungall, B.A., Kyaw-Tanner, M. and Pollitt, C.C.: In vitro evidence for a bacterial pathogenesis of equine laminitis. *Vet Microbiol* 79 (2001) 209-23.

Nadeau, J.H. and Sankoff, D.: Counting on comparative maps. *Trends Genet* 14 (1998) 495-501.

Nagle, D.L., Kozak, C.A., Mano, H., Chapman, V.M. and Bucan, M.: Physical mapping of the Tec and Gabrb1 loci reveals that the Wsh mutation on mouse chromosome 5 is associated with an inversion. *Hum Mol Genet* 4 (1995) 2073-9.

Nagle, D.L., Martin-DeLeon, P., Hough, R.B. and Bucan, M.: Structural analysis of chromosomal rearrangements associated with the developmental mutations Ph, W19H, and Rw on mouse chromosome 5. *Proc Natl Acad Sci* U S A 91 (1994) 7237-41.

Namazi, M.R. and Soma, J.: Tranilast: a novel weapon against insulin resistance. *Med Hypotheses* 64 (2005) 1135-7.

Nishimura, E.K., Jordan, S.A., Oshima, H., Yoshida, H., Osawa, M., Moriyama, M., Jackson, I.J., Barrandon, Y., Miyachi, Y. and Nishikawa, S.: Dominant role of the niche in melanocyte stem-cell fate determination. *Nature* 416 (2002) 854-60.

Nocka, K., Tan, J.C., Chiu, E., Chu, T.Y., Ray, P., Traktman, P. and Besmer, P.: Molecular bases of dominant negative and loss of function mutations at the murine c-kit/white spotting locus: W37, Wv, W41 and W. *EMBO J* 9 (1990) 1805-13.

Palaniyandi Selvaraj, S., Watanabe, K., Ma, M., Tachikawa, H., Kodama, M. and Aizawa, Y.: Involvement of mast cells in the development of fibrosis in rats with postmyocarditis dilated cardiomyopathy. *Biol Pharm Bull* 28 (2005) 2128-32.

Paus, R., Theoharides, T.C. and Arck, P.C.: Neuroimmunoendocrine circuitry of the 'brain-skin connection'. *Trends Immunol* (2005)

Penedo, M.C., Millon, L.V., Bernoco, D., Bailey, E., Binns, M., Cholewinski, G., Ellis, N., Flynn, J., Gralak, B., Guthrie, A., Hasegawa, T., Lindgren, G., Lyons, L.A., Roed, K.H., Swinburne, J.E. and Tozaki, T.: International Equine Gene Mapping Workshop Report: a comprehensive linkage map constructed with data from new markers and by merging four mapping resources. *Cytogenet Genome Res* 111 (2005) 5-15.

Peters, E.M., Tobin, D.J., Botchkareva, N., Maurer, M. and Paus, R.: Migration of melanoblasts into the developing murine hair follicle is accompanied by transient c-Kit expression. *J Histochem Cytochem* 50 (2002) 751-66.

Pettenati, M.J., Rao, P.N., Phelan, M.C., Grass, F., Rao, K.W., Cosper, P., Carroll, A.J., Elder, F., Smith, J.L., Higgins, M.D. and et, a.l.: Paracentric inversions in humans: a review of 446 paracentric inversions with presentation of 120 new cases. *Am J Med Genet* 55 (1995) 171-87.

Philo, J.S., Wen, J., Wypych, J., Schwartz, M.G., Mendiaz, E.A. and Langley, K.E.: Human stem cell factor dimer forms a complex with two molecules of the extracellular domain of its receptor, Kit. *J Biol Chem* 271 (1996) 6895-902.

Pielberg, G., Olsson, C., Syvanen, A.C. and Andersson, L.: Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics* 160 (2002) 305-11.

Primedia Company: Equine.com. http://www.equine.com, 2006.

Pulos, W.L. and Hutt, F.B.: Lethal dominant white in horses. *J Hered* 60 (1969) 59-63.

Raponi, M., Upadhyaya, M. and Baralle, D.: Functional splicing assay shows a pathogenic intronic mutation in neurofibromatosis type 1 (NF1) due to intronic sequence exonization. *Hum Mutat* 27 (2006) 294-5.

Raudsepp, T., Kijas, J., Godard, S., Guerin, G., Andersson, L. and Chowdhary, B.P.: Comparison of horse chromosome 3 with donkey and human chromosomes by cross-species painting and heterologous FISH mapping. *Mamm Genome* 10 (1999) 277-82.

Rieger, R.: *A Glossary of Genetics and Cytogenetics*. Springer-Verlag, New York, 1968.

Rieseberg, L.H. and Livingstone, K.: Evolution. Chromosomal speciation in primates. *Science* 300 (2003) 267-8.

Rousset, D., Agnes, F., Lachaume, P., Andre, C. and Galibert, F.: Molecular evolution of the genes encoding receptor tyrosine kinase with immunoglobulinlike domains. *J Mol Evol* 41 (1995) 421-9.

Rudolph, J.A., Spier, S.J., Byrns, G., Rojas, C.V., Bernoco, D. and Hoffman, E.P.: Periodic paralysis in quarter horses: a sodium channel mutation disseminated by selective breeding. *Nat Genet* 2 (1992) 144-7.

Ruh, J., Vogel, F. and Schmidt, E.: In vivo assessment of villous microcirculation in the rat small intestine in indomethacin-induced inflammation: role of mast-cell stabilizer Ketotifen. *Acta Physiol Scand* 170 (2000) 137-43.

Santschi, E.M., Purdy, A.K., Valberg, S.J., Vrotsos, P.D., Kaese, H. and Mickelson, J.R.: Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. *Mamm Genome* 9 (1998) 306-9.

Santschi, E.M., Vrotsos, P.D., Purdy, A.K. and Mickelson, J.R.: Incidence of the endothelin receptor B mutation that causes lethal white foal syndrome in white-patterned horses. *Am J Vet Res* 62 (2001) 97-103.

Shang, Y.P., Xie, X.D., Wang, X.X., Chen, J.Z., Zhu, J.H., Tao, Q.M. and Zheng, L.R.: A novel splice mutation of HERG in a Chinese family with long QT syndrome. *J Zhejiang Univ Sci* B 6 (2005) 626-30.

Shapiro, M.B. and Senapathy, P.: RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15 (1987) 7155-74.

Shin, E.K., Perryman, L.E. and Meek, K.: A kinase-negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. *J Immunol* 158 (1997) 3565-9.

Silvers, W.K.: *The Coat Colors of Mice*. Springer, Verlag, 1979.

Sponenberg, D.P.: *Equine Color Genetics*, Second Edition. Iowa State University Press, Ames, Iowa, 2003.

Spritz, R.A.: Molecular basis of human piebaldism. *J Invest Dermatol* 103 (1994) 137S-140S.

Spritz, R.A., Ho, L. and Strunk, K.M.: Inhibition of proliferation of human melanocytes by a KIT antisense oligodeoxynucleotide: implications for human piebaldism and mouse dominant white spotting (W). *J Invest Dermatol* 103 (1994) 148-50.

Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A., Gudnadottir, V.G., Desnica, N., Hicks, A., Gylfason, A., Gudbjartsson, D.F., Jonsdottir, G.M., Sainz, J., Agnarsson, K., Birgisdottir, B., Ghosh, S., Olafsdottir, A., Cazier, J.B., Kristjansson, K., Frigge, M.L., Thorgeirsson, T.E., Gulcher, J.R., Kong, A. and Stefansson, K.: A common inversion under selection in Europeans. *Nat Genet* 37 (2005) 129-37.

Stefanutti, G., Pierro, A., Vinardi, S., Spitz, L. and Eaton, S.: Moderate hypothermia protects against systemic oxidative stress in a rat model of intestinal ischemia and reperfusion injury. *Shock* 24 (2005) 159-64.

Stephenson, D.A., Lee, K.H., Nagle, D.L., Yen, C.H., Morrow, A., Miller, D., Chapman, V.M. and Bucan, M.: Mouse rump-white mutation associated with an inversion of chromosome 5. *Mamm Genome* 5 (1994) 342-8.

Stetefeld, J. and Ruegg, M.A.: Structural and functional diversity generated by alternative mRNA splicing. *Trends Biochem Sci* 30 (2005) 515-21.

Sturtevant, A.: A critical examination of recent studies on colour inheritance in horses. *Journal of Genetics* (1912) 41-51.

Syrris, P., Heathcote, K., Carrozzo, R., Devriendt, K., Elcioglu, N., Garrett, C., McEntagart, M. and Carter, N.D.: Human piebaldism: six novel mutations of the proto-oncogene KIT. *Hum Mutat* 20 (2002) 234

Takata, K., Takeda, Y., Sato, T., Nakatsuka, H., Yokoyama, M. and Morita, K.: Effects of hypothermia for a short period on histologic outcome and extracellular glutamate concentration during and after cardiac arrest in rats. *Crit Care Med* 33 (2005) 1340-5.

Tchougounova, E., Lundequist, A., Fajardo, I., Winberg, J.O., Abrink, M. and Pejler, G.: A key role for mast cell chymase in the activation of pro-matrix metalloprotease-9 and pro-matrix metalloprotease-2. *J Biol Chem* 280 (2005) 9291-6.

The Jackson Laboratory, Bar Harbor Maine. Mouse Genome Database (MGD), Mouse Genome Informatics Web Site. http://www.informatics.jax.org, 2006.

Theoharides, T.C., Donelan, J., Kandere-Grzybowska, K. and Konstantinidou, A.: The role of mast cells in migraine pathophysiology. *Brain Res Brain Res Rev* 49 (2005) 65-76.

Toivari, M., Maki, T., Suutarla, S. and Eklund, K.K.: Ethanol inhibits IgE-induced degranulation and cytokine production in cultured mouse and human mast cells. *Life Sci* 67 (2000) 2795-806.

Torihashi, S., Ward, S.M., Nishikawa, S., Nishi, K., Kobayashi, S. and Sanders, K.M.: c-kit-dependent development of interstitial cells and electrical activity in the murine gastrointestinal tract. *Cell Tissue Res* 280 (1995) 97-111.

Trommershausen-Smith, A.: Linkage of tobiano coat spotting and albumin markers in a pony family. *J Hered* 69 (1978) 214-6.

Tsujimura, T., Koshimizu, U., Katoh, H., Isozaki, K., Kanakura, Y., Tono, T., Adachi, S., Kasugai, T., Tei, H., Nishimune, Y. and et, a.l.: Mast cell number in the skin of heterozygotes reflects the molecular nature of c-kit mutation. *Blood* 81 (1993) 2530-8.

Tuzun, E., Sharp, A.J., Bailey, J.A., Kaul, R., Morrison, V.A., Pertz, L.M., Haugen, E., Hayden, H., Albertson, D., Pinkel, D., Olson, M.V. and Eichler, E.E.: Fine-scale structural variation of the human genome. *Nat Genet* 37 (2005) 727-32.

van Eps, A.W. and Pollitt, C.C.: Equine laminitis: cryotherapy reduces the severity of the acute lesion. *Equine Vet J* 36 (2004) 255-60.

Venta, P.J., Brouillette, J.A., Yuzbasiyan-Gurkan, V. and Brewer, G.J.: Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. *Biochem Genet* 34 (1996) 321-41.

Wagner, I.P., Rees, C.A., Dunstan, R.W., Credille, K.M. and Hood, D.M.: Evaluation of systemic immunologic hyperreactivity after intradermal testing in horses with chronic laminitis. *Am J Vet Res* 64 (2003) 279-83.

Ward, T.L., Valberg, S.J., Adelson, D.L., Abbey, C.A., Binns, M.M. and Mickelson, J.R.: Glycogen branching enzyme (GBE1) mutation causing equine glycogen storage disease IV. *Mamm Genome* 15 (2004) 570-7.

Yang, G.C., Croaker, D., Zhang, A.L., Manglick, P., Cartmill, T. and Cass, D.: A dinucleotide mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome (LWFS); a horse variant of Hirschsprung disease. *Hum Mol Genet* 7 (1998) 1047-52.

Zerpa, H., Vega, F., Vasquez, J., Ascanio, E., Campos, G., Sogbe, E., Romero, E., Ascanio, M. and Garcia, H.: Effect of acute sublethal endotoxaemia on in vitro digital vascular reactivity in horses*. J Vet Med A Physiol Pathol Clin* Med 52 (2005) 67-73.

Zhao, X., Dib, M., Wang, X., Widegren, B. and Andersson, R.: Influence of mast cells on the expression of adhesion molecules on circulating and migrating leukocytes in acute pancreatitis-associated lung injury. *Lung* 183 (2005) 253-64.

VITA: Samantha Ann Brooks

• Born: July 7th 1979, Lexington, KY•

EDUCATION

June 2001-May 2006   Lexington, Kentucky

University of Kentucky, Graduate School
- Otis A. Singletary Fellowship
- Academic Excellence In-State Tuition Scholarship
- Kentucky Opportunity Fellowship
- Geoffery C. Hughes Fellowship

August 1997-June 2001   Lexington, Kentucky

University of Kentucky, B.S. Agricultural Biotechnology, May 2001
- GPA 3.3/4.0
- Fayette County Farm Bureau Scholarship
- William A. Seay Scholarship

ACADEMIC HONORS

Graduate School
- Otis A. Singletary Fellowship
- Academic Excellence In-State Tuition Scholarship
- Kentucky Opportunity Fellowship
- Geoffery C. Hughes Fellowship

Undergraduate
- GPA 3.3/4.0
- Fayette County Farm Bureau Scholarship
- William A. Seay Scholarship

PROFESSIONAL POSTITIONS

June 2001- Present   Vet. Science Dept. Univ. of KY
    Graduate Research Assistant

January 2000- June 2001   Vet. Science Dept. Univ. of KY
    Lab Technician/ Undergraduate Researcher

November 1998- May 1999   Agronomy Dept. Univ. of KY
    General Lab Technician

SCIENTIFIC PUBLICATIONS

A PCR-RFLP for KIT associated with the tobiano spotting pattern in horses.
  Brooks SA, Terry RB, and Bailey E.
  Animal Genetics 2002, August; 33 (4):301-3

Assignment of the appaloosa coat color gene (*LP*) to equine chromosome 1.
  Terry RB, Archer S, Brooks S, Bernoco D, and Bailey E.
  Animal Genetics 2004, April; 35 (2):134-137.

Exon skipping in the *KIT* gene causes a Sabino spotting pattern in horses.
  Brooks S and Bailey E.
  Mammalian Genome 2005, 16 (11):893-902.

SCIENTIFIC PRESENTATIONS

January, 2004 International Plant and Animal Genome Conference- Poster:
  "CHEF Studies of the Equine Chromosome Region Encoding the Tobiano Coat Color Pattern"

January, 2005 International Plant and Animal Genome Conference- Poster:
  "Exon skipping in the *KIT* gene causes the Sabino-White spotting pattern in Tennessee Walking horses"

July, 2005 International Equine Gene Mapping Workshop- Invited Speaker:
  "Discovery of a Chromosome Inversion on ECA3 Associated with the Tobiano White Spotting Pattern"

January, 2006 International Plant and Animal Genome Conference- Poster:
  "Characterization of a Paracentric Inversion on ECA3 Associated with the Tobiano Spotting Pattern in Horses"

Samantha Ann Brooks

May 3, 2006