THE NUCLEAR GENETIC DIVERSITY OF TURKISH NATIVE HORSE BREEDS AND ITS CONSERVATION IMPLICATIONS

by

Melis Denizci

BS. in BIO., Middle East Technical University, 2006

Submitted to the Institute of Environmental Sciences in partial fulfillment of

the requirements for the degree of

Master of Science

in

Environmental Sciences

Boğaziçi University

2010

THE NUCLEAR GENETIC DIVERSITY OF TURKISH NATIVE HORSE BREEDS AND ITS CONSERVATION IMPLICATIONS

APPROVED BY:

Assist. Prof. Raşit Bilgin (Thesis Supervisor)	.....
Dr. Berrin Erdağ (Thesis Co-supervisor)	.....
Doç. Dr. Andrzej Furman	.....
Prof. Dr. Bahar İnce	.....
Dr. Evren Koban	.....

DATE OF APPROVAL 09/06/2010

ACKNOWLEDGEMENTS

In the first place, I would like to thank my advisor, Assist. Prof. Raşit Bilgin. With his enthusiasm and his great efforts to explain things clearly and simply, he was always accessible and willing to help. Throughout my thesis writing period, he provided encouragement, advice, fun teaching and good ideas. Special thanks go to my co-advisor, Dr. Berin Erdağ, who is most responsible for helping me during my research. She provided supervision, advice, and guidance from the very early stages of this research as well as sharing her valuable experience with me throughout the work. Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Dr. Bahar İnce, Doç. Dr. Andrzej Furman and Dr. Evren Koban, for their encouragement, insightful comments, and questions. Furthermore, I would like to thank Sezen Arat for having me as a part of the TURKHAYGEN-I project and letting me write my thesis on this subject. I gratefully acknowledge Dr. Evren Koban for her advice, supervision, and crucial contribution throughout my research and writing part, which made her a backbone of this research and so to this thesis. Her involvement with her originality has triggered and nourished my field and research experience. I would like to say "thank you" to Koray Balcıoğlu and Aylin Ö. Bahadır for helping me at any time, and solving any unsolvable problems concerning lab work or private life. I am grateful to Serkan Uzyol, Diğdem Aktopraklıgil, Soner Aksu, Gazi Turgut, Aydın Bahar for supporting my study in several ways. My special thanks go to Özgür Aslan for being such a good friend. He made my research and life better through our interactions during the long hours in the lab or office. My deepest gratitude goes to my family for their unflagging love and support throughout my life; this thesis would have simply been impossible without them. I am indebted to my father, İsmail Denizci, for his care and love. He had never complained in spite of all the hardships in his life. I cannot ask for more from my mother, Ferah Denizci, as she is always a friend or a parent whenever, I needed one. I have no suitable word that can fully describe her everlasting love to me. I feel proud of my sister, Merve Denizci, for her talents. She has always been one of my best counselors. In particular, I am forever grateful to my husband, Levent Öncü, for supporting and encouraging me to pursue this degree. I would like to thank his friendship and help in the past nine years.

ABSTRACT

Anatolia has been home to many ancient civilizations. More importantly, Southeast Anatolia was part of the place, called the Fertile Crescent, where hunter gatherers established permanent settlements and domesticated some plant and animal species. The period of initial domestication and shift from hunting-gathering to farming and agriculture has not been well defined, yet. However, it is well known that Anatolia was associated with the domestication of some of the livestock animals; goat, sheep and pig. Hence, Anatolian populations of many livestock species might contain genetic evidence to fill the gaps and unravel their domestication stories. One of these species is the horse (*Equus caballus*). Due to rapid mechanization in Turkish agriculture, the native Turkish horses have been in a decline, lately. However, the use of horse power in field cultivation and transportation is still a necessity. The genetic characterization of the present native breeds, and assessing the prioritization of the breeds are the first steps to preserve the present genetic diversity, to meet the demands of future breeding programs and to develop conservation strategies. In this study, a total of 425 horse samples were genotyped at 21 microsatellite loci. These samples included five Anatolian domestic breeds of known phenotypes (Hinis, Canik, Malakan, Cukurova horses and Ayvacık Pony) and two Anatolian horse samples without defined phenotypic characteristics (Erzurum and East Anatolian horses), and two foreign breeds (English and Arabian horses), for comparison. The present genetic variation measured by heterozygosity and number of alleles revealed high diversity in Anatolian horse breeds. Factorial Correspondence Analysis detected some level of population differentiation between breeds, though was not significant except for Ayvacık Pony. The Principal Component Analysis supported the FCA results; whereas, Structure analysis and Neighbour Joining tree could not make a clear differentiation within the Anatolian breeds. The analyses also suggested a unidirectional gene flow from the Arabian horses into the Anatolian breeds. The results indicated that stronger measures should be undertaken to conserve the genetic identity of the Anatolian breeds, which have already been compromised by the lack of clear breeding strategies.

ÖZET

Anadolu birçok antik uygarlığa ev sahipliği yapmıştır. Daha önemlisi, Güneydoğu Anadolu avcı-toplayıcıların kalıcı yerleşim kurduğu ve bazı bitki ve hayvanları evcilleştirmeye başladıkları, Bereketli Hilal olarak isimlendirilen bölgenin bir parçasıdır. Evcilleştirme süreci ve avcı-toplayıcılıktan, çiftçilik ve tarıma geçiş süreci henüz tamamen aydınlatılamamıştır. Ancak, Anadolu'nun keçi, koyun ve domuz gibi çiftlik hayvanlarının evcilleştirilmesinde rol aldığı iyi bilinmektedir. Bu nedenle, evcilleştirme hikayelerindeki boşlukları dolduracak genetik bulgulara sahip olabileceklerinden, birçok evcil türün Anadolu popülasyonu önemlidir. Bu türlerden bir tanesi de attır (*Equus caballus*). Son zamanlarda, Türk tarımının hızla makineleşmesi nedeniyle Türk yerli atlarında bir düşüş yaşanmaktadır. Ancak tarla yetiştiriciliğinde ve ulaşımda beygir gücünün kullanımı bu hayvanlara olan ihtiyacı hala devam ettirmektedir. Mevcut yerli at ırklarının genetik karakterizasyonun yapılması ve öncelikli ırkların belirlenmesi, mevcut genetik çeşitliliğin korunmasında ve gelecekteki ıslah programlarının gereksiniminin karşılanmasında ve koruma stratejileri geliştirilmesinde atılması gereken ilk adımlardır. Bu çalışmada, toplamda 425 adet at örneği 21 mikrosatelit lokus açısından genotiplendirilmiştir. Bu örnekler içinde fenotipleri bilinen beş yerli Anadolu ırkı (Hınıs, Canik, Malakan, Çukurova atları ve Ayvacık Midillisi), ve karşılaştırma amaçlı olarak belirgin fenotipik özellikleri olmayan iki ayrı Anadolu örneği (Erzurum ve Doğu Anadolu atları) ile iki yabancı ırk (Arap ve İngiliz atları) yer almaktadır. Alel sayısı ve heterozigotluk değerleri ile ölçülen mevcut genetik çeşitlilik, Anadolu at ırklarında yüksek varyasyon olduğunu göstermiştir. Faktöryel Birleşim Analizi (FCA), Ayvacık Midillisi dışında ırklar arasında anlamlı olmayan düşük seviyede bir farklılaşma tespit etmiştir. Temel Bileşenler Analizi (PCA)'nin sonuçları FCA sonuçlarını desteklerken, komşu birleştirme metodu ile çizilen filogenetik ağaç ve Structure analizi ırklar arasında belirgin bir ayrım yapamamıştır. Ayrıca, analizler Arap atlarından Anadolu ırklarına doğru tek taraflı bir gen akımına işaret etmektedir. Sonuçlar, yetersiz ıslah stratejilerinin tehlikeye attığı Anadolu at ırklarının genetik kimliğini korumak için daha güçlü önlemler alınması gerektiğini göstermektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

ABSTRACT iv

ÖZET v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES ix

LIST OF ABBREVIATIONS xi

1. INTRODUCTION 1

1.1. Significance of Molecular Markers 3

1.2. Microsatellite Markers 4

1.3. Turkish Horse Breeds 5

2. LITERATURE REVIEW 8

3. MATERIALS AND METHODS 13

3.1. Samples 13

3.2. Laboratory Methods 15

3.2.1. DNA Isolation 15

3.2.2. Microsatellites 16

3.2.3. Polymerase Chain Reaction (PCR) Conditions 18

3.2.4. Data Analysis 19

4. RESULTS AND DISCUSSIONS 23

5. CONCLUSION AND RECOMMENDATIONS 48

6. REFERENCES 53

APPENDIX A: OBSERVED VS EXPECTED GENOTYPE COUNTS OF ALL LOCI FOR ANATOLIAN POPULATIONS 65

APPENDIX B: ALLELE FREQUENCIES WITH GRAPHS BY POPULATION AND LOCUS 92

APPENDIX C: EXPECTED VS OBSERVED HETEROZYGOSITY VALUES FOR EACH LOCUS 98

APPENDIX D: EXPECTED VS OBSERVED HETEROZYGOSITY VALUES FOR EACH POPULATION 105

LIST OF FIGURES

Figure 1.1. Pictures of five native breeds of Turkey 6
Figure 3.1. The distribution of the five native horse breeds across Turkey 15
Figure 4.1.1. An example from nine loci multiplex PCR analyses 23
Figure 4.1.2. An example from five loci multiplex PCR analyses 24
Figure 4.1.3. An example from seven loci multiplex PCR analyses 24
Figure 4.2. Factorial Correspondence Analysis of native Anatolian breeds 40
Figure 4.3. Factorial Correspondence Analysis of all individuals 41
Figure 4.4. PCAgen results for the Anatolian horse breeds 42
Figure 4.5. PCAgen results for all the breeds included in the study 43
Figure 4.6. The bar chart by population ID 44
Figure 4.7. Plot of estimates in multiple lines representing the individual genotypes analysed in the study 45
Figure 4.8. NJ tree of samples based on DA genetic distances 46

LIST OF TABLES

Table 3.1. List of sampling location, breed name, abbreviation and number of samples collected. 14

Table 3.2. List of microsatellite markers in the study, their chromosome number in Equus caballus (*ECA*), primers, size range. 17

Table 4.1. Total number of alleles observed for each locus in each breed and sample the mean number of alleles (MNA) observed for each breed/ sample (MNA/pop) and for each locus (MNA/locus). 26

Table 4.2. Allelic richness per locus and population. 28

Table 4.3. The distribution and frequencies of private alleles. 29

Table 4.4. The observed heterozygosity (H_o) values of each breed for each locus. 30

Table 4.5. The estimated expected heterozygosity (H_e) values of each breed for each locus. 32

Table 4.6. The average expected (H_e) and observed heterozygosity (H_o) values for the breeds analyzed. 33

Table 4.7. List of loci that have significant differences between observed and expected heterozygosities (P<0.000265). 33

Table 4.8. Breed specific F_{IS} indices (1023 permutations) and their significance test results. 34

Table 4.9. All loci evaluated for information content for breed identification. 35

Table 4.10. Loci pair that have significant linkage disequilibrium (P< 0.000265). 36

Table 4.11. The result of AMOVA. 36

Table 4.12. The result of AMOVA. 37

Table 4.13. Pairwise F_{ST} values of nine breeds with 21 loci (*P*-values obtained after 720 permutations, Indicative adjusted nominal level (5%) for multiple comparisons is: 0.001389). 38

Table 4.14. Pairwise F_{ST} values of nine breeds with 14 loci (excluding null alleles). 39

Table 4.15. Population assignment outcomes to 'Self' or 'Other' breed. 39

LIST OF ABBREVIATIONS

Abbreviation	Explanation
AP	Ayvacık pony
CKR	Cukurova horse
CNK	Canik horse
EAH	East Anatolian Horse
ERZ	Erzurum horse
HH	Hınıs horse
MLK	Malakan horse
ARP	Arabian Hose
ING	English Horse
mtDNA	Mitochondrial deoxyribonucleic acid
SSR	Simple sequence repeats
STR	Short tandem repeats
VNTR	Variable number of tandem repeats
scnDNA	Single copy nuclear DNA
PCR	Polymerase chain reaction
D-loop	Displacement loop
DNA	Deoxyribonucleic Acid
DNTP	Deoxyribonucleotidetriphosphate
PCA	Principal Component Analysis
FCA	Factorial Correspondence Analysis
AMOVA	Analysis of Molecular Variance
ECA	*Equus caballus*
MNA	Mean Number of Allele
DF	Degrees of Freedom

1. INTRODUCTION

Biological diversity has an intrinsic value for the richness of nature and life. Nowadays, investigation and conservation of biodiversity is becoming more and more important. The scientists are now aware of the importance of avoiding biodiversity loss (Civanova et al., 2006). As the alterations in our planets' biodiversity can result in ethical and aesthetic concerns along with a powerful effect to change the ecosystems and the goods and services it contributes to us, scientific community is now focusing on the conservational issues (Hooper et al., 2005).

The results of human activities are threatening our world by reducing the number of species that exist. The biological diversity is decreasing, many species are getting extinct and many of them have only few remaining individuals, which are all endangering the survival of the remaining populations. It is clear that the rate of extinction is beyond the rate of establishment of new species (Qiu-Hong et al., 2004). As the present extinction levels are close to the past five mass extinctions in the history, the current situation is referred to as the "sixth mass extinction" (Frankham et al., 2002).

In small populations, decreased genetic diversity is known to increase the risk of extinction because it reduces the populations' ability to cope with environmental changes. So, as environmental change is experienced by all species, it is important to investigate the levels of genetic diversity to manage successful conservation strategies for different species (Frankham and Kingslover, 2004).

Being situated on the Silk Road and between some of the most important civilizations of the world, Anatolia has been a geopolitically important area in the history. More importantly, Anatolia was part of Fertile Crescent, which is described as one of the most favorable environment for agricultural societies. The earliest and best-preserved

Neolithic settlements were discovered in Anatolia (Esin, 1999). This is important because the Neolithic Age spanned the time when humans passed from hunting and gathering to the domestication of plants and animals (Hongo et al., 2004). It is clear that Anatolia has played a major role in animal domestication. According to the archeological and genetic studies, Anatolia has been one of the main areas for sheep, goat and pig domestication (Naderi et al., 2008, Zeder, 2008, Perkins, 1969, Buitenhuis, 1997, Peters et al., 1999, 2004). Accordingly, new evidence from Turkey has critical importance. Horses, one of the domesticated animals, played an extensive role in the movements of human populations in Anatolia and the surrounding regions, as they served in transportation and warfare. They also served as food items. The course of civilization has been drastically affected by the domestication of horse and vice versa, as they supplied meat and milk besides their role in transportation (Anthony, 1986; Diamond, 1991).

To initiate a quick action plan for conservation of the native breeds, a management plan is necessary. However, as our financial sources are limited, it is important to determine the priority breeds which are carrying the maximum diversity or to identify the breeds with unique genetic characters. Thus, the use of molecular markers can be a powerful tool for designing sustainable breeding strategies.

It is now possible to study the extent of genetic diversity of many domestic species by using a wide range of molecular markers which provide valuable information in investigating the center of origin for various domestic species, their migration routes and defining the areas where admixture occurred. Within and between diversity parameters can be taken into account for determining priority of the breeds. By using this kind of information, the most appropriate breeds/individuals and geographic areas can be selected for successful management and conservation purposes (Hanotte et al., 2005). While prioritizing breeds, it is important to choose the maximum genetic diversity for potential future use (Gibson et al., in press).

1.1. Significance of Molecular Markers

To develop a sustainable strategy for conservation and management of the breeds, genetic characterization is an essential and powerful tool**.** Choosing the most appropriate marker is one of the most critical steps in the conservation genetics analysis. The available markers that can be used for genotyping include sequence analysis of mitochondrial DNA (mtDNA)'s D-loop and cytochrome-b regions, single nucleotide polymorphisms (SNPs), Y chromosome markers and microsatellites (paternal or autosomal) (Avise, 1994). Y chromosome polymorphism studies determine history of paternal lineages. Y chromosome has been found to be less variable than other genomic sequences leading to a few identified Y chromosome specific polymorphic markers for livestock species, mostly because low number of males have contributed to the gene pool of the most of the species (Petit et al., 2002). On the other hand, there is a wide range of autosomal microsatellite markers available, which are recommended by Food and Agriculture Orfanisation (FAO) and International Society for Animal Genetics (ISAG) for genetic diversity assessment studies in livestock species. In the literature, microsatellites can also be called as simple sequence repeats (SSR), short tandem repeats (STR), or variable number tandem repeats (VNTR).

Single-copy nuclear DNA (scnDNA) and mtDNA carry signatures of past genetic mutations and they are appropriate for analyzing taxonomic relationships. scnDNA is diploid and bi-parentally inherited in most of the vertebrates, whereas mtDNA is a single locus with maternal inheritance (Birky et al., 1983). mtDNA sequences can be used to analyze the number of maternal lineages and geographic origins. mtDNA and VNTRs are more suitable for studying recent genetic variation when compared to scnDNA. In this respect VNTRs are commonly used and can be more useful in understanding contemporary genetic patterns.

Since microsatellite analyses can be used for gathering information from different genomic regions, microsatellites are suitable to model the divergence of populations. Moreover, individual specific fingerprints can be obtained, as well as distant ones, due to

the rapid rate of microsatellite evolution. For these reasons, microsatellite analyses are the one of the most promising approaches to measure genetic variation over relatively recent period of times. In the literature, microsatellites are typically used for individual genetic identification, paternity testing in breeding populations, measuring genetic diversity, and determining population differentiation, genetic bottlenecks, genetic relationships and admixture (Qiu-Hong et al., 2004).

1.2. Microsatellite Markers

For the estimation of genetic diversity within and among populations, microsatellites are commonly used (Aberle et al., 2004; Glowatzki-Mullis et al., 2005). They are highly abundant and virtually uniformly distributed throughout the genomes of species (Orti et al., 1997). These markers are short base pair repeats that are interspersed throughout the entire eukaryotic genomes. They have a repeat size of two to six base pairs, with the most common form being (CA)n or (TG)n. (Litt et al, 1989; Weber et al., 1989) They have a diverse number of repeats presented in different alleles, and are co-dominantly inherited (MacHugh et al., 1994; Li., 1997). They are highly polymorphic and have a high mutation rate. They also have a simple and stable inheritance mechanism while being passed onto the next generation.

Microsatellites are short sequences that can be easily amplified by PCR and it is possible to assign the specific bands to certain loci and allele frequencies. Hence, recently microsatellites have attracted attention for their use in construction of genetic maps of a great variety of organisms (Knapik et al., 1998; Cregan et al., 1999), and understanding the relationship between human genetic diseases and the instability of repeat numbers (Mahadevan et al., 1992; Stallings 1994; O'Donnel and Warren, 2002). They are also efficient genetic tools in paternity testing, forensic studies, individual assignments of unknown samples to species, breeds or populations and assessing genetic variation, (Bjornstad et al., 2000; Tozaki et al., 2003; Achmann et al., 2004; Glowatzki-Mullis et al., 2006), genetic characterization of breeds, construction of pedigrees, linkage analysis and

determining evolutionary relationships (Tozaki et al., 2003; Aberle et al., 2004; Achmann et al., 2004; Glowatzki-Mullis et al., 2005; Vega-Pla et al., 2006 ; Lee & Cho 2006; Moodley et al., 2006).

Although microsatellites were initially designed for human research, they also turned out to be efficient tools for animal studies (Schlötterer et al., 1991). Heywood and Iriondo (2003) proposed microsatellites as to provide information for the identification of conservation units and investigation of the genetic histories including gene flow and genetic drift.

Even though microsatellites have been extensively used in several studies, a single mutation model to determine the allelic variation has not been determined, yet. There are many different mutation models that can be used for microsatellites. In the infinite allele model (IAM, Kimura and Crow, 1964), each mutation randomly creates a new allele; hence the mutations alter the number of repeats. In the stepwise mutation model (SMM, Kimura and Ohta, 1978), a microsatellite gains or loses a repeat if the locus mutates implying that two alleles differing by only one motif are more related than the ones differing in several alleles. In the two-phase model (TPM, DiRienzo et al., 1994) most mutational events result in an increase or decrease of one repeat unit. The SMM is the most preferred model when exploring the association between individuals and population structure (Moxon et al., 1999).

1.3. Turkish Horse Breeds

Many horse breeds like the Arabian, Persian, Thracian, Russian, Mongolian and Caucasian horses from geographically close or historically related civilizations have contributed to the formation of the Anatolian horse breeds (Batu, 1938). In Turkey, Anatolian, Cukurova and Midilli breeds are classified as the major breeds. Çamardı Kulası, Malakan Horse, Canik Horse, Hınıs'ın Kolukısası Horse are the other breeds which are

placed in Anatolia region (Batu, 1938). After the establishment of the Republic of Turkey, Nonius, Ardene and Halflinger breeds were imported to contribute to the Anatolian horse breeds. In addition to these, Thoroughbred and Arabian horses were also brought to Turkey for the improvement of the native breeds, creating half-breed populations (Batu, 1938). After the development of industry and transportation, the need for horse-power decreased significantly. Consequently, Anatolian horse livestocks face the threat of extinction. General features of the five native breeds of Turkey included in the present study are described below, with a representative picture for each breed provided in Figure 1.1.

Canik Horse	Malakan Horse	Hınıs Horse	Ayvacık Pony	Cukurova Horse

Figure 1.1. Pictures of five native breeds of Turkey

Canik Horse: Canik breed is known to comprise alert, aggressive, swift, speedy and strong horses. Although every common color can be seen, the most common color is bay. They are generally used as riding horses. Their heads are medium-sized and muscular, with wide foreheads, they have gently sloping croup, powerful legs and pasterns, and hard nails. The height of Canik horses at withers is between 140-145 cm.

Malakan Horse: Malakan Horse is the only heavy breed of Turkey. These horses are very strong, friendly natured and a good draft and farm work horse. This breed is characterized by a large and powerful body and bone structure. They have a short, thick and muscular necks and a rough head, longer and are thicker than the extinct Anatolian Horse. They have an average-sized waist, short legs and thick bone structure and a large chest. They are very resistant to cold weather and are very strong. They have shorter waists

and more strong nails compared to Haflingers. Although until recently they were predominantly black, currently they frequently exhibit red, bay and ice colors. The height of Malakan horses at withers is about 140 cm.

Hınıs Horse: Hınıs horse is preferred to be used as a riding horse. The most significant characteristic of Hınıs breeds, which the basis of their name as "short aimed of Hinis", is their front legs being shorter than back legs. They are very powerful and swift. These horses have a drastically deep and wide chest that allows great capacity for heart and lungs. They have short wrists, thick joints, strong bones and a short neck. They also have a wide chest and hard nails. These horses have all the common horse colors. The height of Hınıs horses at withers is between 135-138 cm.

Ayvacık Pony: These horses have a unique property of being able to walk in a leisurely fashion. Their short height is an advantage for them, while walking under olive trees, which are found frequently in the Ayvacık area. Children can ride these horses as they are calm and friendly. They have short head and neck, strong and short legs and a large chest. Their colors are usually ice bay gray. The height of these horses at withers is between 116 - 120 cm.

Cukurova Horse: These horses were hybridized by the crosses of Turkish-Arab horse and Uzunyayla horse and then crossed to the native Anatolian horse. However, Cukurova breeds remain uncrossed in the Cukurova area. These horses are larger than the extinct Anatolian horse. They have long neck and shoulders, and a powerful breast and bone structure. They are resistant to hot and humid climates and they can be used for heavy duties. They are mostly reddish, dune-colored or gray like. Their withers' height is about 130-140 cm.

2. LITERATURE REVIEW

Horses have been an integral part of the human life for millennia. They have been used as a food source (Zeder and Hesse, 2000; Ouatram et al, 2009), for transportation, labor, and even recreation (Mills and McDonnell, 2005). According to the archeological studies, it is likely that first domestication of the horse has occurred about 6000 years ago on the Eurasian steppes between Ukraine and Turkestan and spread to the rest of the world from this area (Anthony, 1996; Bennett and Hoffmann, 1999; Clutton-Brock, 1999).

mtDNA haplotypes of ancient domestic horses are almost similar to haplotypes of existing domestic horses. Also, the frequency of common haplotypes of existing horses is similar to those haplotypes of horses of early domestication (Jansen et al., 2002). However, this does not implicate that other haplotypes were not used in domestication, rather, some of the maternal individuals with specific haplotypes died out because of random genetic drift and selection (Avise, 1994). The observed mtDNA diversity in ancient domestic horses is also higher than the existing ones (Kavar and Dove, 2008).

Two models of horse domestication have been suggested (Clutton-Brock, 1999). These can be outlined as follows:

According to Model I; as mentioned above, the wild stock horses that inhabited the steps between Ukraine to Turkestan were the first to be domesticated. Subsequently, the domestic horses were spread from that area to the different parts of the world, and the different types and breeds that survived until today were developed as a result of artificial selection and also affected by natural selection in order to adapt to their local environments.

According to Model II; a multiple origin hypothesis involves a large number of founders. In this model, horses may have been independently captured from diverse wild populations and bred in captivity. This model assumes that there was a geographical cline in the wild horses. The horses on the northern regions were smaller and stronger than the ones in the southern regions. This model is also supported by the earliest findings of domestic horses, which were in different size and proportions in different parts of the world (Kavar et al., 2008).

According to model I, domestic horses were derived from wild horses that had high genetic diversity but which were not geographically structured. In contrast, model II suggests that the wild horses from which domesticated horses originated were initially separated following a geographical cline or presence of subspecies. Migration is probably the main factor for the lack of phylogeographic structure (Kavar et al., 2008). As horses are very mobile animals, high migration rates are expected also in the history of domestication. However, the absence of migrations particularly in Western Europe has been suggested by morphometric studies, which implicate a regional fragmentation (Bignon et al., 2005).

The Neolithic remains of wild horses are found mainly in Switzerland, Sweden, Denmark, the Netherlands, Spain, Italy, France, Germany, Hungary, Serbia, Ukraine, Kazakhstan and Russia. These results implicated that earliest domestication must have occurred in Eastern Europe or Central Asia (Levine, 2005). However, the horses from the Holocene seem to have a wider distributional area, especially in western and central Europe (Levine, 2005). Again, according to Levine (2006), horse domestication might have occurred in a long time frame, where it has been affected by genetic events that occurred in captive areas. In addition, as has been implicated by historical remains, wild horses had a wider distribution in the post iron age, which made it possible that their genes might have been introduced into domestic breeds after domestication occurred.

There is recent evidence which supports the second model. According to Vila (2001), wild horse populations were phylogeographically structured, particularly on

different continents. Their inference is due to the diversity found in existing domestic horses not explaining origination from only two discrete groups. Hence, they argue that this scenario, consistent with the archeological and genetic data, was a more likely explanation for the horse domestication. This scenario posits that as wild horse populations diminished due to environmental changes, breeding in captivity increased, which resulted in the integration of multiple matrilines into the gene pool of domestic horses. The phylogeographic structure in Eurasian wild horse population was explained by Jansen *et al.* (2002). According to this study, several of the breeds were clustered with their corresponding breeds and geographic areas such as the A2 cluster; specific to Przewalski horses, C1 specific to northern European ponies, D1 specific to Iberian and northwest African breeds. For example, as mentioned above, one of the clusters (C1) was geographically restricted to central Europe, British Isles and Scandinavia. Most of the documented horses from C1 were Northern European ponies and several undocumented horses in C1 were observed to be ponies. C1 demonstrates the explicit association between cluster and breed. These results suggest that, the geographic separation of mtDNA clusters from a wide range of samples were in accord with the scenario of contribution of wild mares for domestication from geographically different areas.

It is possible that the most appropriate explanation for domestication of horse is the intermediate model of domestication. According to this intermediate model, the domestic horses were originated from the wild horses distributed over a moderately wide geographical cline and had a large enough pre-existing haplotype diversity (Lister et al., 1998).

In addition to the studies on mtDNA lineages, the patrilineal diversity in horses was explored using Y chromosome markers. Wallner et al. (2003) identified only a single haplotype suggesting that a limited number of patrilines were involved in horse domestication. These data indicating high matrilineal and low patrilineal diversity can be interpreted so that horse domestication began with the “appropriate male” or by selection. The appropriate male may have had an extraordinary characteristic or there might not have been so specific differences between haplotypes (Kavar et al., 2008)

In recent years, microsatellite markers have been frequently used for evaluating genetic diversity, structure, distances and characterizing local breeds (Solis et al., 2005, Aberle et al., 2004). A summary of these studies is presented below.

Leroy et al. (2009) analyzed the genetic diversity and structure of French horses by using 11 microsatellite loci. The populations were classified into more or less differentiated four clusters; warm-blooded horses, draught horses, Nordic horses and pony breeds. The clusters were meaningful when compared with the morphologic characters, geographic origins and the use of the breeds. In this study, it has been suggested that five local breeds, namely the Boulonnais, Landais, Merens, Poitevin and Pottok breeds should have been conserved in priority. Particularly, it has been asserted that genetic variability should have been managed by using breeds that contribute original genes, such as Boulonnais, Merens and Poitevin.

Tozaki et al. (2003) investigated the genetic relationships of seven Japanese, four mainland-Asian and two European populations by using 20 microsatellite loci. Three different groups clustered separately in the phylogenetic analysis; European cluster, Hokkaido-Kiso cluster and Mongolian cluster. The clustering was consistent with the geographic separations. According to the phylogenetic tree and distribution of genetic variation, it has been suggested that Japanese horses originated from Mongolian horses migrating through the Korean Peninsula. Furthermore, their results indicated that many populations in Japan had low genetic variability.

In a study by Azor et al. (2007), genetic characterization of Spanish Trotter horse was carried out by using 16 microsatellite loci. Strong heterozygote deficiency was found in all breeds, which was indicated by high and positive F_{IS} values. Spanish Trotter was a candidate for being under high selection pressure for racing performance (Gomez et al., 2005) and Andalusian breed was reported to be under pressure for morphological traits

(Valera et al., 2005). The highest genetic distance values were obtained for the Andalusian outgroup. Spanish Trotters were also seen to be more distant from Mallorquina than Menorquina breeds, which was in contrast with the geographic location of these two breeds. It has been asserted that although Spanish Trotter horses were originally affected by the Balearic native horses, new generations have not been influenced by Balearic horses.

In Aberle et al. (2004), genetic diversity and distance among six German draughts to wild, primitive and riding horse breeds were analyzed using 30 microsatellite markers. According to the differentiation test results, highly significant genetic differences were observed among all draught horse breeds, except the Mecklenburg and Saxon Thuringa Coldbloods. The Schleswig Draught Horse was found to be the most distinct draught horse breed.

In this perspective, the main objectives of this study are; (i) to make genetic characterization of five native Anatolian horse breeds and (ii) to use these results to contribute to the development of conservation strategies for domestic Turkish horse breeds. Microsatellite markers were used to determine genetic diversity within and among horse breeds, to explore the variability of breeds compared to the data available in the literature and to assess and compare the genetic distinctiveness of Anatolian horse breeds with two imported breeds (Arabian and English). Furthermore, the assignment test of the individuals to pre-defined morphological groups was performed and differentiation between the breeds was evaluated.

3. MATERIALS AND METHODS

3.1. Samples

In this project, blood samples were collected using 10 ml K_3EDTA containing vacuum tubes from selected individuals of the five local breeds included in the study: Canik horse, Malakan horse, Hınıs horse, Ayvacık pony and Cukurova horse. A total of 425 individuals were analyzed (Table 3.1). The horse breeds were chosen according to their phenotypes described in old text books (Batu, 1938), and through personal communications with academicians and veterinarians. The individuals that represent the best phenotypic characteristics were predominantly chosen. However, some individuals that didn't represent the described phenotypic characteristics (grade horses) were also taken for comparison. Furthermore, two breeds, not native to Anatolia, were chosen for comparison and verification of the results. The first one was the Arabian horse, samples of which were collected from the General Directorate of Agricultural Enterprises (TIGEM), Karacabey and the other one was the English Horse, whose samples were collected from the Jockey Club of Turkey, Izmit (Table 3.1). In Figure 3.1., the distribution of the five native horse breeds across Turkey is given.

Table 3.1. List of sampling location, breed name, abbreviation and number of samples collected.

No	Name of the Breed	Abbreviation of the Breeds	Sampling Location	Number of the Sample
1	Hınıs Horse	HH	Erzurum	60
2	Canik Horse	CNK	Samsun	64
3	Malakan Horse	MLK	Kars Ardahan Igdır	64
4	Ayvacık Pony	AP	Canakkale	49
5	Erzurum Horse	ERZ	Erzurum	17
6	East Anatolian Horse	EAH	Kars Ardahan Kayseri Iğdır Ağrı Van	71
7	Cukurova Horse	CKR	Adana Osmaniye	60
8	Arabian Horse	ARP	Bursa	20
9	English Horse	ING	İzmit	20

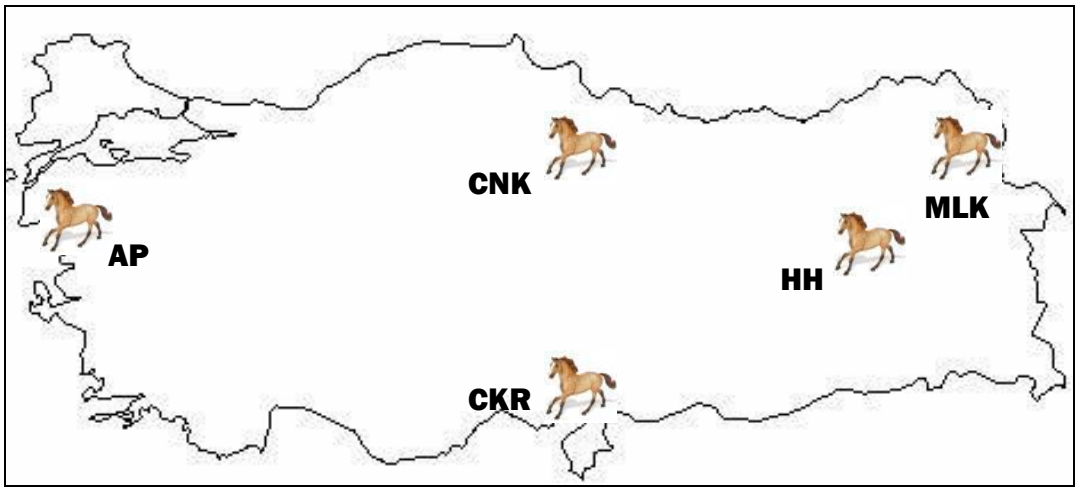


Figure 3.1. The distribution of the five native horse breeds across Turkey. AP and HH, MLK, CNK, CKR are the breeds of known phenotypes, EAH and ERZ are the grade horses that did not have defined phenotypic characters.

3.2. Laboratory Methods

3.2.1. DNA Isolation

A standard phenol:chloroform:isoamyl alcohol protocol (Sambrook et al., 1989) was performed for isolating total DNA from the collected blood samples. All of the DNAs were isolated at TUBITAK GMBE laboratories. First, 10 ml blood sample was put into a tube containing 0.5 ml EDTA (0.5 M, pH 8.0), and 40 ml of 2X lysis buffer (10X Lysis Buffer: 770 mM NH_4Cl, 46 mM $KHCO_3$, 10 mM EDTA) was added. The tubes were mixed by inverting for 10 minutes and were subsequently centrifuged at 3000 rpm at 4°C for 10 minutes. Then, the supernatant was discarded, 3 ml salt/EDTA (75 mMNaCl, 25 mM EDTA) was added in to the tubes and mixed by vortexing. After the addition of 0.3 ml of 10% SDS and 150 µl of proteinase K (10 mg/ml) solution, samples were incubated at 55°C for 1-3 hours. After the incubation, 3 ml phenol (pH 8.0) was added to each tube and tubes were shaken vigorously for 20 seconds and then by gentle inversions for 5 minutes.

Tubes were centrifuged at 3000 rpm at 4°C for 10 minutes. Supernatant phase was transferred to the new tubes labelled properly and 3 ml phenol: chloroform: isoamyl alcohol (25:24:1) was added into the each tube. Then, the tubes were shaken vigorously for 20 seconds, mixed well for 5 minutes, and centrifuged at 3000 rpm at 4°C for 10 minutes. Moreover, the supernatant phase was transferred into new sterile and labelled 15 ml falcon tubes, ice cold 96% EtOH (kept at -20°C) was added at a volume of twice the supernatant, and the tubes were shaken well to help the extracted DNA precipitate. Finally, the precipitated DNA was taken by a micropipette, transferred into 1.5 ml eppendorf tubes, washed once with 70% alcohol, air dried and dissolved in 0.3-0.5 ml 10 mMTris-HCl (pH 8.00) solution. The extracted DNA samples were either stored at +4°C for short-term use or at -20°C for long term storage.

The extracted DNA samples were 1:10 diluted with Tris-HCl or Tris-EDTA and run in a 0.8 % agarose gel. In addition, the quality and quantity of the DNA samples were measured with Nanodrop device (Thermo) and the results were compared with the agarose gel photographs.

3.2.2. Microsatellites

21 microsatellite loci were used in this study (Table 3.2). The microsatellites chosen were all polymorphic and had dinucleotide repeats. They were chosen from the most commonly used loci in the literature, from ISAG/FAO recommendation list, and they included the loci used for the ISAG horse parentage tests.

Table 3.2. List of microsatellite markers in the study, their chromosome number in *Equus caballus* (*ECA*), primers, expected size ranges and citations.

Loci	Chr (*ECA*)	Primer 5'-3'	Size Range	Reference
Aht4	24	AACCGCCTGAGCAAGGAAGT GCTCCCAGAGAGTTTACCCT	142-164	Binns et al., 1995
I18	16	CAACAAAGATGTTGCAAGGG GTGTGCCTCTTGTCTCTTAGG	81-109	Marti et al., 1998
Cor2	14	CTTGAGCACCCAGTAACACC CCAGGAATCTTCTCTACCGA	235-243	Hopman et al., 1999
Lex33	4	TTTAATCAAAGGATTCAGTTG GGGACACTTTCTTTACTTTC	192-222	Coogle et al., 1996
Asb2	15	CACTAAGTGTCGTTTCAGAAGG GCACAACTGAGTTCTCTGATAGG	218-256	Breen et al., 1994
Hms3	9	CCAACTCTTTGTCACATAACAAGA GCCATCCTCACTTTTTCACTTTGTT	152-180	Guerin et al., 1994
Hms5	5	TAGTGTATCCGTCAGAGTTCAAGG GCAAGGAAGTCAGACTCCTGGA	100-110	Guerin et al., 1994
Hms6	4	CTCCATCTTGTGAAGTGTAACTCA GAAGCTGCCAGTATTCAACCATTG	155-169	Guerin et al., 1994
Htg6	15	CCTGCTTGGAGGCTGTGATAAGAT GTTCACTGAATGTCAAATTCTGCT	79-105	Ellegren et al., 1992
Aht33	31	CTGAGGGCGTAAGTCGAGTC GTTAATAGGAGCGGTTGTTTGG	151-167	Swinburne et al., 2000b
Asb43	29	TCACTTAGTAGGGGCATGC GTGTTTGTCCTTGACTCTCC	77-103	Irvin et al., 1998
Nev79	17	ATTGCCTGTGCTGAGATGG GCAAATTGCCTCTGTATCACAC	175-197	Bjornstad et al., 2000
CA425	28	CTCATGTCCGCTTGTCTC AGCTGCCTCGTTAATTCA	220-242	Eggleston_Stott et al., 1997
Hms2	10	CTTGCAGTCGAATGTGTATTAAATG ACGGTGGCAACTGCCAAGGAAG	216-244	Guerin et al., 1994
Asb17	2	ACCAGTCAGGATCTCCACCG GAGGGCGGTACCTTTGTACC	81-125	Breen et al., 1997
Asb23	3	ACATCCTGGTCAAATCACAGTCC GAGGGCAGCAGGTTGGGAAGG	181-209	Breen et al., 1997

Tky301	23	AATGGTGGCTAATCAATGGG GTGTATGATGCCCTCATCTC	144-166	Tozaki et al., 2001
Vhl20	30	CAAGTCCTCTTACTTGAAGACTAG AACTCAGGGAGAATCTTCCTCA	85-109	Van Haeringen et al., 1994
Htg4	9	CTATCTCAGTCTTGATTGCAGGAC GCTCCCTCCCTCCCTCTGTTCTC	126-142	Ellegren et al., 1992
Hms7	1	CAGGAAACTCTCATGTTGATACCATC GTGTTGTTGAAACATACCTTGACTGT	171-189	Guerin et al., 1994
Cor58	12	CACCAGGCTAAGTAGCCAAG GGGAAGGACGATGAGTGAC	203-229	Ruth et al., 1999

3.2.3. Polymerase Chain Reaction (PCR) Conditions

All DNA samples were amplified using the 21 primer pairs described above. Biometra 3000 Thermal Cycler machine was used to perform the PCR reactions. Three genotyping panel sets were optimized to analyze the samples at 21 microsatellite loci. The first panel consisted of the microsatellite loci I18, AHT4, LEX33, COR02, ASB2, HTG6, HMS3, HMS5, HMS6. The second panel consisted of the microsatellite loci ASB43, AHT33, HMS2, NEVHEQ79, UCDEQ425. Amplification of microsatellites for the first and second multiplex PCR reactions were performed in a 25 µl total reaction volume containing 30-50 ng of genomic DNA, 0.2 to 0.8 pmol of the primer pairs, 1X PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs and 1.5U Taq Polymerase. The third panel consisted of the microsatellite loci ASB17, TKY301, ASB23, VHL20, HTG4, HMS7 and COR58. Amplification of microsatellites for the third multiplex PCR reaction was performed in a 15 µl total reaction volume containing 30-50 ng of genomic DNA, 0.1 to 1 pmol of the primer pairs, 0,2 µl of BSA (10 mg/ml), 1X PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs and 1U Taq Polymerase.

The PCR conditions for the first and second panel were as follows: One cycle of denaturation at 95°C for 2 min, 30 cycles of amplification process where the samples were incubated at 94°C for 30 s, at annealing temperature (T_A) for 45 s (T_A=55°C for the first

panel, T_A=60°Cfor the second panel), elongation at 72°C for 1 min, and one cycle of final elongation at 72°C for 20 min. The PCR conditions for third panel was as follows: One cycle of denaturation at 95°C for 3 min, 35 cycles of amplification process where the samples were incubated at 94°C for 20 s, annealing at 60°C for 1 min, elongation at 72°C for 50 s, and one cycle of final elongation at 72°C for 20 min.

The PCR products were visualized on a 2% agarose gel electrophoresis, followed by the detection through capillary electrophoresis. Beckman Coulter CEQ8800 Genetic Analysis System was used for the detection and allele sizing. One µl of each PCR product was mixed with 0.2 µl of Beckman Coulter's DNA Size Standard Kit - 400 and 30 µl of sample loading buffer. Samples were loaded into the plates and data was collected using the Beckman's fragment analysis tool.

3.2.4. Data Analysis:

Genotyping errors may occur due to null alleles. Null alleles are non-amplified alleles that result in an apparent homozygote when segregating with another allele (Oosterhout et al., 2005*).* MICRO-CHECKER 2.2.3 (Oosterhout et al., 2004) indicates the presence of null alleles if there is an overall significant excess of homozygotes in microsatellite data of diploid populations and was used to test for genotyping errors.

Population differentiation among and within populations was estimated based on F-statistics according to Weir and Cockerham (1984), using the FSTAT 2.9.3 computer program (Goudet, 2001). The observed number of alleles, allele frequencies *per* population, overall frequencies and allelic richness *per* population were calculated using the same program. F_{IS} is defined as the relationship between homologous alleles within individuals relative to samples, which explores the inbreeding within subpopulations. Thus, F_{IS} measures if a deficit of heterozygotes within subpopulations exists or not. F_{IT} measures how much population structure has affected the average heterozygosity of

individuals within the population. F_{IT} is defined as the correlation of corresponding alleles within individuals relative to total sample. F_{ST} measures how differentiated the subpopulations are. It basically deals with the proportion of total genetic variation that is distributed among subpopulations, rather than within subpopulations.

Genetic Analysis in Excel (GenAlEx 6) (Peakal et al., 2006) is a package for population genetic analysis which runs within Microsoft Excel, enabling population genetic analyses of co-dominant, haploid and binary data. Using this software, descriptive statistics such as expected and observed heterozygosity values were calculated and population assignments were made.

In addition, the linkage disequilibrium between pairs of loci was explored by using GENEPOP 4.0 genetics software (Raymond and Rouset, 1995). In the software, the null hypothesis of "genotypes at one locus are independent from genotypes at the other locus" is assumed. A Contingency Table for all pairs of loci was created and probability tests were performed for each value using a Markov Chain Method.

Another program (WHICHLOCI) (Banks et al., 2003) was used to discriminate the relative power among genetic loci. WHICHLOCI program analyses the efficiency of loci for correct population assignment by ranking empiric data from the real data.

The conversion of the data input file to be analyzed by Fstat and Arlequin softwares was performed by using the GENETIX 4.0.5 program (Belkhir et al., 2004). Possible admixture with and without prior population information was estimated with the Factorial Correspondence Analysis (FCA) method in the same program. Hence, by using Factorial Correspondence Analysis, it was possible to visualize individuals in a multidimensional space and to analyze how they were related to each other. In the analysis, the samples were examined on 3D graphics with different triple combinations of the 10 factors (each represented by an axis) estimated by the software.

Principle Component Analysis (PCA) is a method, which reduces the dimensionality of the data by applying a covariance analysis between factors. It performs this reduction by identifying directions, principal components, in which the data variation is kept at maximum. In this analysis, individuals of a population can be represented as a single value, which makes simple visualization and exploration of the data possible. PCA for the microsatellite data was implemented with the program PCAGEN.

Genetic differentiation within and among populations, and within and among groups based on geographic location data was estimated using the Analysis of Molecular Variance (AMOVA) test, implemented with the statistics program ARLEQUIN 3.0 (Excoffier et al., 2005).

For detecting the signatures of the recent genetic bottlenecks, BOTTLENECK program (Cornuet and Luikart, 1996) was used. The analysis is based on the hypothesis that allelic diversity is reduced faster than the heterozygosity if the population experienced a recent reduction of their effective population size. When a population experiences bottleneck, rare alleles are lost and allelic diversity is decreased. However, heterozygosity is not reduced proportionally as heterozygosity is not affected by rare alleles. In this study, a frequently preferred method in literature, Wilcoxon rank test was performed and TPM approach was used to test the significance of the results by permutating the data 1000 times (Luis et al, 2007, Giacomoni et al., 2008, Shahsavarani et al., 2010, Gupta et al., 2005).

For the detection of population expansion, an excel macro program called KGTESTS was used. The excel macro used in this study implements the k and g tests to analyze the population expansion (Bilgin, 2007). The significance of the g test can be checked by the Table 1 (p. 455) from Reich et al. (1999). The analyses were done based on a stepwise model. The within locus k test is based on the point that different modes of allele-length distribution at a locus is expected for a constant sized population, whereas, a

single and more peaked distribution is expected for an expanding population. The interlocus g test is based on the notion that constant-sized populations' high interlocus variability increases the variance across loci, however, in expanded populations, mutations have lower effect on the oldest lineages (Reich and Goldstein et al., 1998; Reich et al.; 1999).

A Bayesian Markov chain Monte Carlo (MCMC) method approach was used to examine the distinctiveness of the populations and clustering of individual genotypes using the STRUCTURE 2.2 program (Pritchard et al., 2000; Evanno et al., 2005), which assesses the genetic clustering within a whole data set and enables a comparison of phenotypic and genotypic groupings.

Finally, a phylogenetic tree based on D_A genetic distances estimated from the allele frequency data was constructed by using Neighbour joining algorithm (POPTREE2 software; Takezaki *et al*., 2009). Bootstrap test (Saitoi and Nei, 1987, Nei et al., 1983), 10,000 replicates, was employed to statistically test the confidence of the resulting tree topology.

4. RESULTS AND DISCUSSION

Fluorescently labeled PCR products produced by fluorescently labeled primer pairs were separated by using a capillary electrophoresis instrument (Beckman Coulter CEQ8800 Genetic Analysis System). Allele sizes were determined by processing raw data with the Beckman fragment analysis tool in comparison to the internal size standard (Beckman Coulter's DNA Size Standard Kit - 400) used within every PCR product sample loaded on the machine. Each horse sample's DNA was analyzed by three multiplex panels. One representative resultant graphic from each of the multiplex panels is given below (Figures 4.1.1-3).

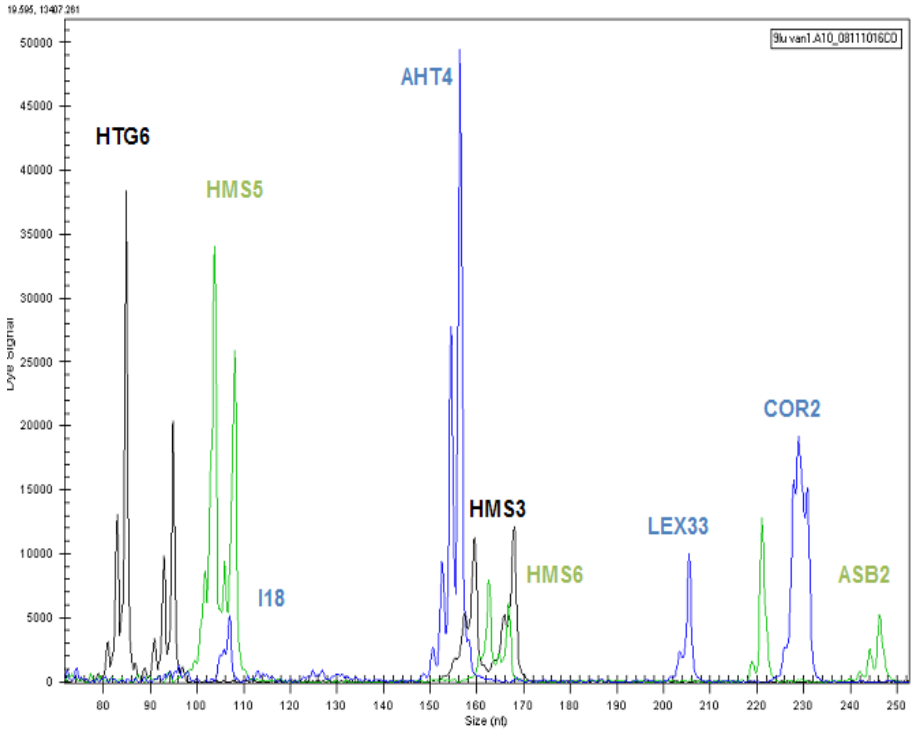


Figure 4.1.1. An example from nine loci multiplex PCR analyses.

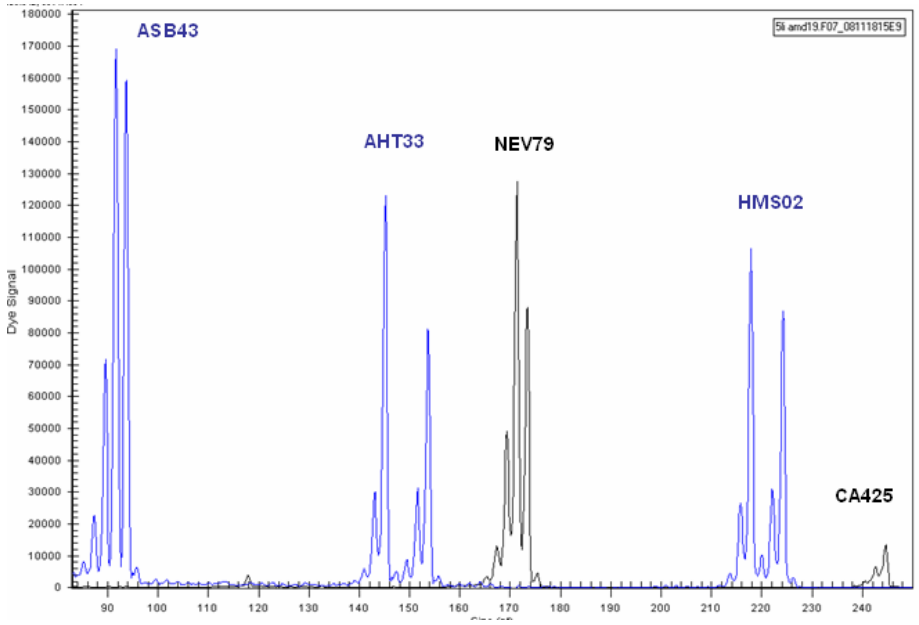


Figure 4.1.2. An example from five loci multiplex PCR analyses.

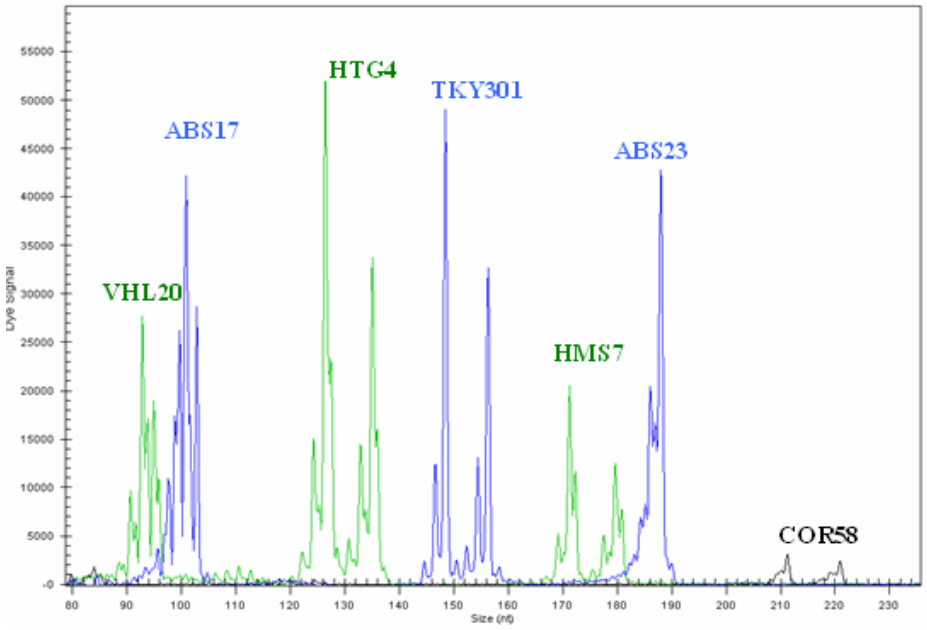


Figure 4.1.3. An example from seven loci multiplex PCR analyses.

Potential incorrect assignment of microsatellite genotyping was explored using the Micro-checker program. Analyses indicated the likely presence of null alleles at seven loci including AHT33, ASB23, ASB43, CA425, HMS6, TKY301, and VHL20. Some of

the analyses, see below, were subsequently made with the entire data set and after excluding the null alleles from the analyses. Analyses of these two data sets gave different results in certain cases.

The total number of alleles for all loci used in the study, the number of alleles observed for each locus in each breed, and the average numbers alleles *per* locus are given in Table 4.1. The highest and lowest numbers of alleles are 24 for ASB17 locus and six for COR2 locus, respectively. A total of 185 alleles for HH, 202 alleles for CNK, 197 alleles for MLK, 189 alleles for AP, 190 alleles for CKR were detected, which resulted in means of 8.8, 9.6, 9.4, 9.0, and 9.0 alleles *per* locus, respectively.

The highest and lowest mean numbers of alleles detected *per* locus were 13.1 for ASB17 locus and 4.3 for HMS5 locus. In addition, the highest and lowest mean numbers of alleles detected *per* population were 9.6 for Canik and 6.8 for Erzurum Horse breeds. The mean number of alleles for the Arabian horses was 5.3, and 5.0 for the English horses. Among the Anatolian Horse breeds, the one that showed the highest allelic variability was Canik Horse, with 202 observed alleles.

Table 4.1. Total number of alleles observed at each locus in each breed and sample, the mean number of alleles (MNA) observed for each breed/sample (MNA/pop) and for each locus (MNA/locus).

Na	HH	CNK	MLK	AP	ERZ	EAH	CKR	ARP	ING	MNA/ Locus	All
HTG6	6	9	8	7	5	6	11	3	4	6.6	13
HMS3	9	8	8	9	7	9	9	5	5	7.7	11
HMS5	5	6	5	5	4	5	3	3	3	4.3	8
HMS6	6	7	5	8	5	6	7	5	5	6.0	8
ASB2	10	9	12	10	9	11	11	7	10	9.9	14
I18	9	11	10	10	7	10	9	5	4	8.3	12
AHT4	10	10	10	9	5	11	10	5	4	8.2	11
LEX33	10	11	10	10	7	10	10	4	5	8.6	13
COR2	5	5	6	5	5	6	6	4	3	5.0	6
ASB43	6	6	6	7	5	8	7	4	4	5.9	11
AHT33	11	13	13	11	11	15	11	9	6	11.1	16
NEV79	10	10	11	12	8	11	9	4	2	8.6	19
HMS2	8	8	10	10	7	9	11	7	6	8.4	14
CA425	9	13	9	8	6	10	11	5	4	8.3	13
ASB17	16	18	16	14	9	17	15	7	6	13.1	24
TKY301	9	8	8	7	7	10	7	5	7	7.6	10
ASB23	10	10	11	10	6	11	9	6	5	8.7	15
VHL20	11	12	10	9	8	12	8	6	5	9.0	13
HTG4	6	7	8	7	5	7	7	5	3	6.1	8
HMS7	8	9	9	9	7	10	8	4	6	7.8	10
COR58	11	12	12	12	10	12	11	9	7	10.7	13
MNA/pop	8.8	9.6	9.4	9.0	6.8	9.8	9.0	5.3	5.0	8.1	

Ranging between 8.8 and 9.6, Anatolian Horse breeds have a greater allelic diversity compared to the domestic horse breeds in the literature. The mean number of alleles *per* locus reported for other breeds are: 7.06 for Lipizzaner Horses (Achmann et al., 2004), 6.0 for Spanish Trotter Horses (Azor et al., 2007), 7.25 for Lithuanian Heavy Draught Horses (Juras and Cothran, 2004), 8.08 for Polish Heavy Horses (Iwanczyk et al., 2006), 6.83 for Canadian Horses and 4.67 for Sable Island Horses (Plante et al., 2007),

4.82-8.09 for 19 breeds of French origin or raised in France (Leroy et al., 2009), 2.1-5.1 for seven breeds of Japanese origin (Tozaki et al., 2003), and 5.2-6.7 for seven breeds of German Horses (Aberle et al., 2004). In another study, Cunningham et al. (2001) reported 7.5 as the average number of alleles *per* locus for 43 Turkish Horses from non-registered populations and 4.7 for Thoroughbred Horses (English origin). Although these values are lower than the results obtained in this study, they reflect a similar pattern in which Turkish Horses have greater allelic diversity. These results show that Anatolian horse breeds have higher allelic diversity and suggest a wider gene pool compared to the Arabian and English horse breeds. The potential reason why Arabian and English horse breeds have lower number of alleles compared to Anatolian breeds is that Arabian and English horse breeds have been originated by more intensive selective breeding.

As the sample sizes in various breeds were different from each other, variations in the sample sizes were also taken into account when assessing allelic richness. These corrected allelic richness values are given in the Table 4.2., below. The mean number of allelic richness results *per* locus differed between 3.711 and 10.011. The mean number of allelic richness results *per* population changed between 6.810 and 7.568. The expected allelic richness values were generally lower than the observed allelic richness values. The genotypic counts of expected *vs.* observed alleles are given in the Appendix A, for the visualization of the results.

Table 4.2. Allelic richness *per* locus and population

Allelic Richness	HH	CNK	MLK	AP	ERZ	EAH	CKR	ARP	ING	Mean/ Locus
HTG6	4.718	5.716	5.562	5.462	5.000	5.315	6.988	3.000	3.850	5.068
HMS3	7.247	7.070	7.053	7.728	7.000	6.981	7.437	4.998	4.850	6.707
HMS5	3.771	4.190	3.728	4.400	4.000	4.313	3.000	3.000	3.000	3.711
HMS6	5.227	5.577	4.605	6.745	5.000	5.363	6.284	4.846	4.831	5.386
ASB2	7.665	8.074	8.893	8.151	9.000	8.668	9.572	6.810	9.542	8.486
I18	7.842	7.857	7.988	8.048	7.000	8.300	7.245	4.848	3.998	7.014
AHT4	8.369	8.306	8.591	8.484	5.000	8.279	7.735	4.981	4.000	7.083
LEX33	7.219	8.243	7.873	6.991	7.000	7.259	7.567	4.000	4.850	6.778
COR2	4.868	4.690	5.144	4.715	5.000	4.897	5.465	3.998	3.000	4.642
ASB43	5.941	5.225	5.852	5.874	5.000	5.822	5.436	3.981	3.850	5.220
AHT33	9.347	9.530	9.716	10.323	11.000	10.390	9.642	8.661	5.998	9.401
NEV79	7.557	7.401	7.565	8.489	8.000	7.785	7.152	4.000	1.998	6.661
HMS2	7.135	7.089	7.689	8.990	7.000	8.069	7.974	6.531	5.531	7.334
CA425	6.709	8.051	6.214	7.557	6.000	7.256	7.499	4.998	4.000	6.476
ASB17	11.168	11.801	11.905	10.467	9.000	11.681	11.546	6.829	5.700	10.011
TKY301	7.698	6.984	6.798	6.552	7.000	7.492	6.454	4.981	6.662	6.736
ASB23	7.519	7.636	8.245	8.526	6.000	7.843	7.038	5.829	5.000	7.071
VHL20	8.799	8.993	8.756	8.103	8.000	8.749	7.809	5.848	4.981	7.782
HTG4	5.527	5.566	6.799	6.223	5.000	6.300	6.029	5.000	2.850	5.477
HMS7	7.164	7.523	7.531	7.167	7.000	7.540	6.659	4.000	5.850	6.715
COR58	10.229	9.778	9.941	9.936	10.000	9.858	9.776	8.844	6.850	9.468
Mean/pop	7.225	7.395	7.450	7.568	6.810	7.531	7.348	5.237	4.819	6.820

An examination of the frequencies of the observed alleles for each breed, given in Appendix B, showed that some of the alleles were present only in one of the breeds, namely breed specific or private alleles. The distributions and frequencies of private alleles are given in Table 4.3., below. In total, all loci except for COR2 (which have the lowest number of alleles) have private alleles. The highest number of private alleles was found in NEV79 locus. The distribution of private alleles *per* population are; 11 for Canik, four for Hinis, five for Malakan, five for Cukurova, eight for Ayvacık Pony, 13 for East Anatolian

horse, and two for English horses. However, as the frequencies are low, they cannot be used for the identification of the breeds. These low frequency private alleles must be checked for size-calling errors. Moreover, they might be present in other breeds and populations, which were not included in this present study. In such a case, they could not be called private alleles.

Table 4.3. The distribution and frequencies of private alleles.

Locus	Allele	Frequency	Breed		Locus	Allele	Frequency	Breed
HTG6	75	0.008	CNK		NEV79	157	0.008	HH
	81	0.008	CKR			163	0.008	CKR
	89	0.008	CKR			177	0.007	EAH
	99	0.016	CNK			179	0.008	CNK
HMS3	143	0.02	AP			191	0.007	EAH
	173	0.007	EAH			193	0.021	EAH
HMS5	93	0.008	MLK			199	0.01	AP
	95	0.021	EAH		HMS2	208	0.008	MLK
	111	0.016	CNK			214	0.01	AP
HMS6	172	0.02	AP			234	0.008	CKR
ASB2	235	0.061	AP			244	0.025	ING
I18	93	0.01	AP		CA425	216	0.008	CNK
	111	0.016	CNK			224	0.008	CNK
AHT4	138	0.007	EAH		ASB17	77	0.007	EAH
LEX33	193	0.016	CNK			117	0.008	HH
	223	0.007	EAH			121	0.02	AP
ASB43	79	0.025	ING			123	0.008	CNK
	83	0.007	EAH			125	0.014	EAH
	103	0.01	AP			127	0.008	CNK
	115	0.008	CKR		TKY301	160	0.007	EAH
AHT33	175	0.008	MLK		ASB23	194	0.017	HH
VHL20	111	0.007	EAH			198	0.008	HH
HTG4	140	0.016	MLK			210	0.008	MLK
HMS7	191	0.007	EAH		COR58	207	0.016	CNK

Observed heterozygosity values (H_o) for all breeds are given in the Table 4.4., below. Similar to the expected heterozygosity values given in the Table 4.5., most of the observed heterozygosity values are higher in Anatolia than in the Arabian and English horse breeds. Canik Horses have the highest average of H_o, with 0.797. The lowest average H_o value within the Anatolian herds is in the Malakan breed, with 0.753. Arabian and English horses have low average H_o values, 0.695 and 0.655, respectively, when compared to the Anatolian Horses. The average H_o values among Anatolian breeds are quite similar to each other, differing at most around a marginal value of 0.04.

Table 4.4. The observed heterozygosity (H_o) values of each breed for each locus.

H_o	**HH**	**CNK**	**MLK**	**AP**	**ERZ**	**EAH**	**CKR**	**ARP**	**ING**
HTG6	0.767	0.672	0.672	0.531	0.647	0.592	0.683	0.7	0.6
HMS3	0.8	0.797	0.75	0.755	0.765	0.803	0.783	0.75	0.7
HMS5	0.65	0.656	0.672	0.796	0.529	0.662	0.65	0.75	0.7
HMS6	0.8	0.844	0.625	0.694	0.706	0.761	0.7	0.8	0.65
ASB2	0.817	0.844	0.781	0.898	0.941	0.803	0.8	0.9	0.95
I18	0.85	0.734	0.781	0.796	0.765	0.789	0.7	0.4	0.65
AHT4	0.833	0.859	0.828	0.939	0.765	0.901	0.833	0.75	0.7
LEX33	0.867	0.859	0.781	0.714	0.765	0.803	0.767	0.9	0.85
COR2	0.633	0.688	0.656	0.714	0.588	0.676	0.617	0.6	0.6
ASB43	0.783	0.734	0.734	0.714	0.824	0.817	0.767	0.35	0.6
AHT33	0.8	0.891	0.781	0.857	0.882	0.831	0.883	0.75	0.8
NEV79	0.733	0.844	0.734	0.735	0.824	0.718	0.733	0.75	0.05
HMS2	0.833	0.766	0.75	0.816	0.882	0.775	0.85	0.65	0.45
CA425	0.733	0.828	0.672	0.837	0.647	0.69	0.6	0.8	0.45
ASB17	0.767	0.813	0.859	0.837	0.824	0.803	0.85	0.85	0.75
TKY301	0.717	0.828	0.766	0.796	0.882	0.817	0.767	0.7	0.7
ASB23	0.917	0.734	0.844	0.918	0.824	0.845	0.85	0.6	0.9
VHL20	0.833	0.859	0.813	0.837	0.941	0.746	0.867	0.55	0.75
HTG4	0.733	0.75	0.703	0.694	0.706	0.761	0.733	0.6	0.4
HMS7	0.783	0.844	0.75	0.776	0.882	0.845	0.733	0.55	0.7
COR58	0.85	0.891	0.859	0.939	0.824	0.873	0.883	0.9	0.8
Mean/pop	0.786	0.797	0.753	0.79	0.782	0.777	0.764	0.695	0.655

Considering the published reports of microsatellites in other horse breeds, the data obtained in this study indicates a high level of heterozygosity for the Anatolian horse breeds (0.75-0.79). The observed heterozygosities in other equine breeds are 0.66 for Lipizzan Horse (Achmann et al., 2004), 0.63 for Pantaneiro Horse (Giacomoni et al., 2008), 0.65 for Spannish Trotters (Azor et al., 2007), 0.61 for Zanskari Horse (Behl et al., 2006). This again suggests a relative lack of selection in the Anatolian horse breeds.

The expected heterozygosity values for all breeds are given in the Table 4.5., below. The expected heterozygosity values (H_e), which also can indicate gene diversity, are generally higher in Anatolian Horses than in the Arabian and English horse breeds sampled from Turkey. Different expected heterozygosities are reported in the literature ranging between 0.69-0.72 for Thouroughbred Horse (English) (Jakabova et al., 2002, Iwancyzik et al., 2006), and 0.64-0.72 for Arabian Horses (Plante et al., 2007, Iwancyzik et al., 2006). Considering the values given in the literature, Arabian and English Horses still have lower expected heterozygosity values when compared to the Anatolian Horses. East Anatolian Horse, Hinis and Canik Horses have the highest average of H_e, with values of 0.795, 0.792 and 0.792, respectively. The lowest average H_e value between Anatolian herds is in the Erzurum Horse, Ayvacık Pony and Cukurova Horse, with 0.762, 0.783 and 0.783. The average H_e values among Anatolian breeds are very similar to each other, differing at most with a marginal value of 0.03. Observed *vs.* expected heterozygosity values of each locus and of each population for visualization of the results are given as bar charts in the Appendix C and Appendix D.

Table 4.5. The estimated expected heterozygosity (H_e) values of each breed for each locus.

H_e	HH	CNK	MLK	AP	ERZ	EAH	CKR	ARP	ING
HTG6	0.694	0.691	0.645	0.595	0.675	0.674	0.738	0.635	0.554
HMS3	0.811	0.794	0.794	0.833	0.787	0.818	0.783	0.734	0.673
HMS5	0.67	0.679	0.663	0.656	0.652	0.666	0.616	0.64	0.614
HMS6	0.74	0.776	0.741	0.732	0.747	0.759	0.766	0.64	0.639
ASB2	0.822	0.839	0.86	0.845	0.829	0.856	0.871	0.764	0.824
I18	0.798	0.778	0.807	0.832	0.76	0.798	0.752	0.484	0.664
AHT4	0.838	0.809	0.853	0.872	0.735	0.839	0.821	0.746	0.741
LEX33	0.777	0.81	0.804	0.734	0.754	0.791	0.785	0.723	0.726
COR2	0.715	0.676	0.67	0.681	0.682	0.698	0.716	0.638	0.599
ASB43	0.775	0.779	0.773	0.763	0.753	0.779	0.766	0.595	0.618
AHT33	0.872	0.867	0.866	0.891	0.881	0.872	0.882	0.828	0.786
NEV79	0.763	0.813	0.751	0.711	0.725	0.732	0.741	0.643	0.139
HMS2	0.825	0.828	0.807	0.855	0.817	0.842	0.816	0.595	0.385
CA425	0.767	0.791	0.733	0.779	0.685	0.737	0.743	0.734	0.509
ASB17	0.829	0.844	0.883	0.828	0.844	0.852	0.859	0.733	0.695
TKY301	0.825	0.816	0.809	0.806	0.758	0.834	0.769	0.741	0.745
ASB23	0.837	0.834	0.852	0.845	0.794	0.837	0.834	0.725	0.785
VHL20	0.845	0.85	0.86	0.851	0.83	0.853	0.847	0.708	0.756
HTG4	0.713	0.69	0.66	0.705	0.697	0.75	0.685	0.754	0.471
HMS7	0.823	0.808	0.769	0.766	0.792	0.828	0.789	0.709	0.794
COR58	0.89	0.857	0.877	0.86	0.811	0.879	0.872	0.866	0.809
Mean/pop	0.792	0.792	0.785	0.783	0.762	0.795	0.783	0.697	0.644

When average expected and observed heterozygosity values are compared (Table 4.6), different results are obtained for different breeds. The mean H_o values are lower than the mean H_e results for Hinis Horse, Malakan Horse, East Anatolian Horse and Cukurova Horse. The mean H_o values are greater for the Ayvacık Pony, Erzurum Horse and English Horse. The H_o and H_e values are almost identical for Canik and Arabian horse breeds. In addition, the list of loci that have significant differences between observed and expected heterozygosities are given in Table 4.7. The most significant differences are obtained for the Malakan and Cukurova Horses, which is consistent with the F_{IS} results (Table 4.8.). It should be noted that the *P* values for this test, and the HWE and linkage disequilibrium

tests (see below) were adjusted by the Bonferroni correction (Rice, 1989). As multiple tests were made with nine breeds and 21 loci, the corrected cut-off *P* value was recalculated as 0.05/(21x9) = 0.000265.

Table 4.6. The average expected (H_e) and observed heterozygosity (H_o) values for the breeds analyzed.

	H_o	**H_e**
HH	0.786	0.792
CNK	0.797	0.792
MLK	**0.753**	**0.785**
AP	0.79	0.783
ERZ	0.782	0.762
EAH	0.777	0.795
CKR	**0.764**	**0.783**
ARP	0.695	0.697
ING	0.655	0.644

Table 4.7. List of loci that have significant differences between observed and expected heterozygosities (*P*<0.000265).

Breed	**Locus**	**DF**	**ChiSq**
HH	HMS3	36	137.498
CNK	HTG6	36	74.341
	HMS5	15	69.952
	I18	55	113.086
	VHL20	66	187.168
MLK	I18	45	89.103
	HMS2	45	94.424
	LEX33	45	118.171
EAH	ASB17	136	266.086
CKR	HTG6	55	154.105
	ASB23	36	80.981

Again, to test for HWE, the within breed variation was explored using F_{IS} index and permutation tests were performed to test the significance of the results, which are given in the Table 4.8., below. The lack of any significant deviation from Hardy-Weinberg

equilibrium, after Bonferroni corrections, suggests the lack of inbreeding in the populations. This result is expected for the populations where there is no artificial selection for breed improvement. However, the lowest F_{IS} value (0.00113) being that for the phenotypically distinct Ayvacık pony should still be noted. Considering that Ayvacık Pony breed is found only in a very restricted area (Canakkale region), these horses are very well adapted to their region. According to these results, the local breed management for the Ayvacık Pony horses could be the reason for this observation.

Table 4.8. Breed specific F_{IS} indices (1023 permutations) and their significance test results.

#	**Breed**	**F_{IS}**	***P* (Rand F_{IS}>=Obs F_{IS})**
1	HH	0.01628	0.163245
2	CNK	0.00163	0.471163
3	MLK	0.04813	0.000978
6	CKR	0.03272	0.012708
7	AP	0.00113	0.513196
8	ARP	0.02778	0.211144
9	ING	0.00892	0.415445

Whichloci program was used for comparing the information content of the microsatellite loci, in terms of their contribution in detecting differentiation of breeds. The differentiation power increases through ranking loci according to their discriminatory ability for the breeds under study. Although more polymorphic and heterozygous loci perform better, specific loci with unique frequencies and ranking higher for particular breeds may affect the discrimination power. This method allows maximizing the discrimination power and minimizing the costs. As a result, VHL20, ASB17, HMS2, AHT33 and COR58 were the first five loci that ranked the highest, in that order, among the 21 loci in terms of their discriminatory power. HMS5 locus was placed at the end of the ranking.

Table 4.9. All loci evaluated for information content for breed identification.

Rank	Locus	Score	Score %
1	VHL20	0.3122	5.7979
2	ASB17	0.3073	5.7071
3	HMS2	0.3013	5.5957
4	AHT33	0.2949	5.4760
5	COR58	0.2784	5.1706
6	AHT4	0.2731	5.0716
7	HMS3	0.2716	5.0427
8	HTG4	0.2669	4.9561
9	NEV79	0.2660	4.9395
10	LEX33	0.2564	4.7621
11	HMS6	0.2533	4.7043
12	ASB2	0.2509	4.6589
13	I18	0.2478	4.6012
14	ASB43	0.2478	4.6012
15	HMS7	0.2436	4.5228
16	TKY301	0.2384	4.4278
17	HTG6	0.2336	4.3371
18	ASB23	0.2242	4.1637
19	CA425	0.2202	4.0895
20	COR2	0.2084	3.8708
21	HMS5	0.1887	3.5035

In the Table 4.10., below, the pairs of loci and *P* values for the breeds in which significant linkage disequilibrium was observed are given. 11 significant linkage disequilibrium values were observed, after the Bonferroni corrections. Four loci in linkage disequilibrium were observed in the Ayvacık Pony horses. No significant linkage disequilibrium values were observed for Erzurum Horses, Cukurova Horses and English Horses.

Table 4.10. Loci pair that have significant linkage disequilibrium ($P< 0.000265$).

Breed	Locus # Locus		Breed	Locus # Locus
HH	LEX33 - NEV79		AP	HMS3 - HMS6
	ASB43 - ASB17			HMS3 -NEV79
CNK	COR2 - ASB17			ASB2 - CA425
MLK	AHT4 - AHT33			HMS2-TKY301
	ASB2 - HMS2		EAH	ASB2 - CA425
			ARP	TKY301-ASB23

Analysis of Molecular Variance was performed to explore the genetic variation within and between breeds. In the first AMOVA analysis, each native breed was assumed to be one group and (EAH) East Anatolian Horse and (ERZ) Erzurum Horse were treated as one group together. Based on the results, about 96% of the total variation was attributed to within individuals, 2.07% of the variation was among groups, and 2.14% of the variation was attributed to among individuals within breeds (Table 4.11.).

Table 4.11. The result of AMOVA.

Source of Variation	Sum of Squares	Variance Components	Percentage Variation
Among Groups	170.433	0.17376 Va	2.07
Within Breeds	3503.310	0.17953 Vc	2.14
Within Individuals	3426.500	8.0635 Vd	96.05
Total	7107.485	8.39414	

In the second AMOVA analysis, with groupings based on the FCA results (Figure 4.3); the Ayvacık Pony (AP), the Arabian (ARP) and the English (ING) breeds were treated as separate groups and the rest of the breeds were assigned to one single group. In the results, about 94% of the total variation was attributed to within individuals, 3.56% of the variation was among groups, 2.09% of the variation was attributed to among

individuals within breeds, and only 0.28% of the variation was attributed to among breeds within groups (Table 4.12.).

Table 4.12. The result of AMOVA.

Source of Variation	Sum of Squares	Variance Components	Percentage Variation
Among Groups	122.362	0.30479 Va	3.56
Among Breeds Within Groups	55.313	0.02407 Vb	0.28
Within Breeds	3503.310	0.17953 Vc	2.09
Within Individuals	3426.500	8.06235 Vd	94.07
Total	7107.485	8.5705	

The AMOVA results did not indiacte a significant amount of variation between breeds. The highest differentiation was attributed to within individuals in both of the analyses. When each native breed was separated as a group in the data, there was no significant difference among the groups.

The pairwise F_{ST} values between all of the breeds were estimated and the results are given in the Tables 4.13 and 4.14 below. This analysis was performed to see if there were significant genetic differences among Anatolian breeds, as well as between the Anatolian breeds and the Arabian breed, and finally between the Anatolian breeds and the English Thoroughbred horses. Different results were observed in terms of population differentiation when the analysis was done with all of the loci (Table 4.13) and loci without null alleles (Table 4.14). In both analyses, Arabian and English Horse populations were significantly differentiated from the other breeds. There were also significant differences between Canik and Hinis, Malakan and Hinis, Malakan and Canik breeds, and Cukurova and Canik breeds. Ayvacık Pony was significantly differentiated from all other breeds except for the Erzurum breed. On the other hand, when loci with null alleles were excluded, Canik and Malakan breeds could not be significantly differentiated from Hinis breeds. Furthermore, the F_{ST} values between Ayvacık Pony and Malakan breeds were not significant when loci with null alleles were excluded from the analyses. The differentiation

of Ayvacık pony from most of the other breeds in Anatolia is in accordance with both the geographic location of the breed, which is quite distant from the other breeds in the study, and its diagnostic phenotypic characteristic of low body height.

Studies in the literature revealed the mean F_{ST} estimates as 0.065 for Indian breeds (Behl et al., 2007), 0.09 for Spanish breeds (Azor et al., 2007), 0.099 for French breeds (Leroy et al., 2009), 0.02-0.08 for Lipizzan horses from diverse European countries (Achmann et al., 2004), 0.08-0.25 for Norwegian horse breeds (Bjornstad et al., 2000), and 0.05-0.14 for Danish breeds (Thirstup et al., 2008). Population differentiation estimates of European breeds seem to be higher than that of Anatolian breeds.

Table 4.13. Pairwise F_{ST} values of nine breeds with 21 loci (*P*-values for: 720 permutations, Indicative adjusted nominal level (5%) :0.001389).

	HH	**C N K**	**M L K**	**A P**	**E R Z**	**E A H**	**C K R**	**A R P**	**I N G**
H H		*	*	*	NS	NS	NS	*	*
C N K	0.0030		*	*	NS	NS	*	*	*
M L K	0.0041	0.0068		*	NS	NS	NS	*	*
A P	0.0078	0.0089	0.0042		NS	*	*	*	*
E R Z	-0.0025	0.0015	0.0043	0.0091		NS	NS	*	*
E A H	0.0005	0.0028	0.0023	0.0048	-0.0029		*	*	*
C K R	0.0019	0.0049	0.0027	0.0053	-0.0006	0.0023		*	*
A R P	0.0425	0.0461	0.0536	0.0573	0.0437	0.0464	0.0408		*
ING	0.0849	0.0887	0.0980	0.1056	0.0937	0.0909	0.0876	0.1308	

(*; P<0.05, NS: not significant)

Table 4.14. Pairwise F_{ST} values of nine breeds with 14 loci (excluding null alleles).

	HH	C N K	M L K	A P	E R Z	E A H	C K R	A R P	I N G
H H		NS	NS	*	NS	NS	NS	*	*
C N K	0.0015		*	*	NS	*	*	*	*
M L K	0.0052	0.0077		NS	NS	NS	NS	*	*
A P	0.0077	0.0084	0.0014		NS	*	*	*	*
E R Z	-0.0046	0.0017	0.0048	0.0112		NS	NS	*	*
E A H	-0.0005	0.0032	0.0029	0.0034	-0.0014		NS	*	*
C K R	0.0009	0.0028	0.0027	0.0059	-0.0008	0.0015		*	*
A R P	0.0343	0.0364	0.0512	0.0593	0.0386	0.0411	0.0386		*
ING	0.1057	0.1087	0.1252	0.1363	0.1195	0.1160	0.1142	0.1577	

(*; P<0.05, NS: not significant)

In the assignment test analysis with using leave one out option, two different data sets were used. In the first analysis, all the breeds were included. In the second analysis, only native Anatolian breeds were included. According to the results, only individuals from the Arabian and English horse breeds were assigned to their own populations, 90% and 100% of the time, respectively. However, all the other samples were assigned to some other breeds in the study without any significant pattern. Considering the native Anatolian breeds, Ayvacık Pony had the comparatively largest self-breed assignment with a value of 47%. The order of the other breeds in terms of having most of its samples assigned correctly to itself were, Cukurova, Canik, Malakan and Hinis horse.

Table 4.15. Population assignment outcomes to 'Self' or 'Other' breed.

Breed Assignment with All Breeds			
All breeds	**Self Breed**	**Other Breed**	**Percent**
HH	11	49	18%
CNK	24	40	38%
MLK	15	49	23%
AP	20	29	41%
ERZ	2	15	12%
EAH	11	60	15%
CKR	17	43	28%
ARP	18	2	90%
ING	20	0	100%
Total	138	287	
Percent	32%	68%	

Breed Assignment with Native Anatolian Breeds			
Native Breeds	**Self Breed**	**Other Breed**	**Percent**
HH	20	40	33%
CNK	27	37	42%
MLK	25	39	39%
AP	23	26	47%
CKR	26	34	43%
Total	121	176	
Percent	41%	59%	

In addition, Factorial Correspondence Analysis (FCA) of samples was performed for only Anatolian pure breeds based on population data with prior phenotypic breed information. FCA analysis with prior phenotypic data showed some differentiation and grouping between different breeds from Anatolia (Figure 4.2). In this analysis, the Ayvacık pony, Malakan and Canik horses seemed to separate pretty distinctively. Hinis and Cukurova horses seemed also to be somewhat differentiated, although they showed the greatest overlap among all.

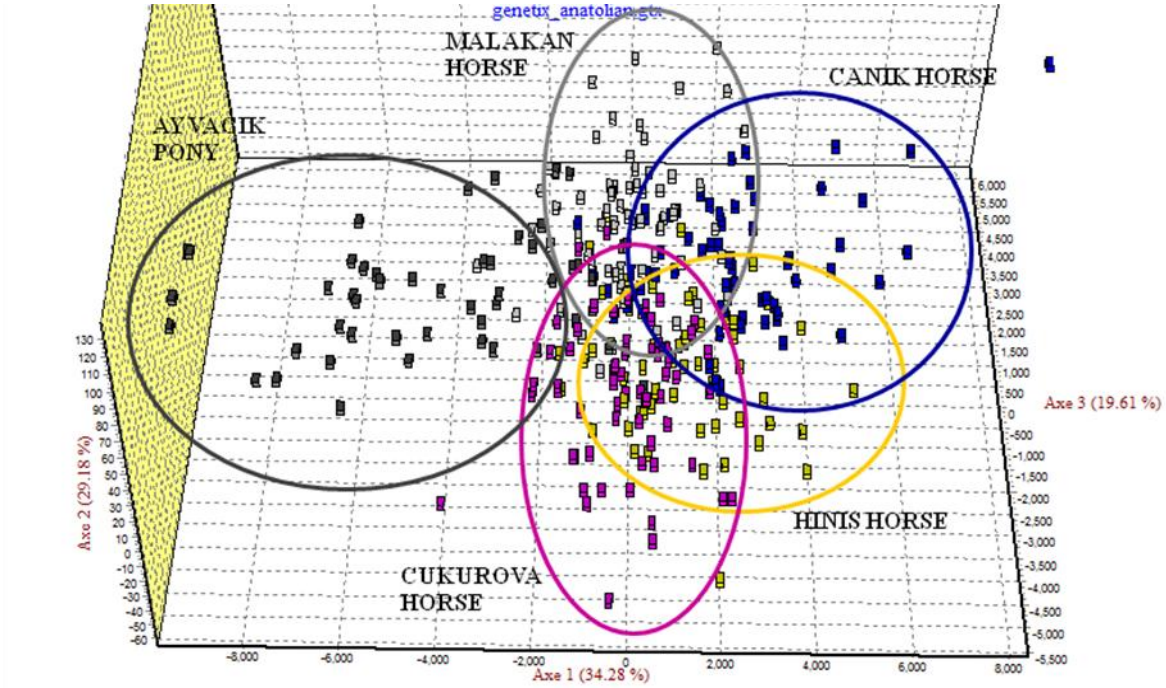


Figure 4.2. Factorial Correspondence Analysis of native Anatolian breeds

FCA was also performed by using all the breeds analyzed in the study. Among all the breeds analyzed, Arabian and English horses separated as two distinctive groups. Also, Ayvacık Pony fell apart from the rest of the Turkish horses, as well as the Arabian and English horse samples. However, in this analysis all of the rest of the Anatolian horse breeds clustered together (Figure 4.3). This probably is due to the samples from Arabian and English breeds crunching the differences between the various Anatolian breeds, which

are relatively closer to each other than these non-Anatolian breeds. It should be noted that even under this analysis, the Ayvacık pony somewhat separated from the other Anatolian horses, showing that this breed is more differentiated than the others within Anatolia.

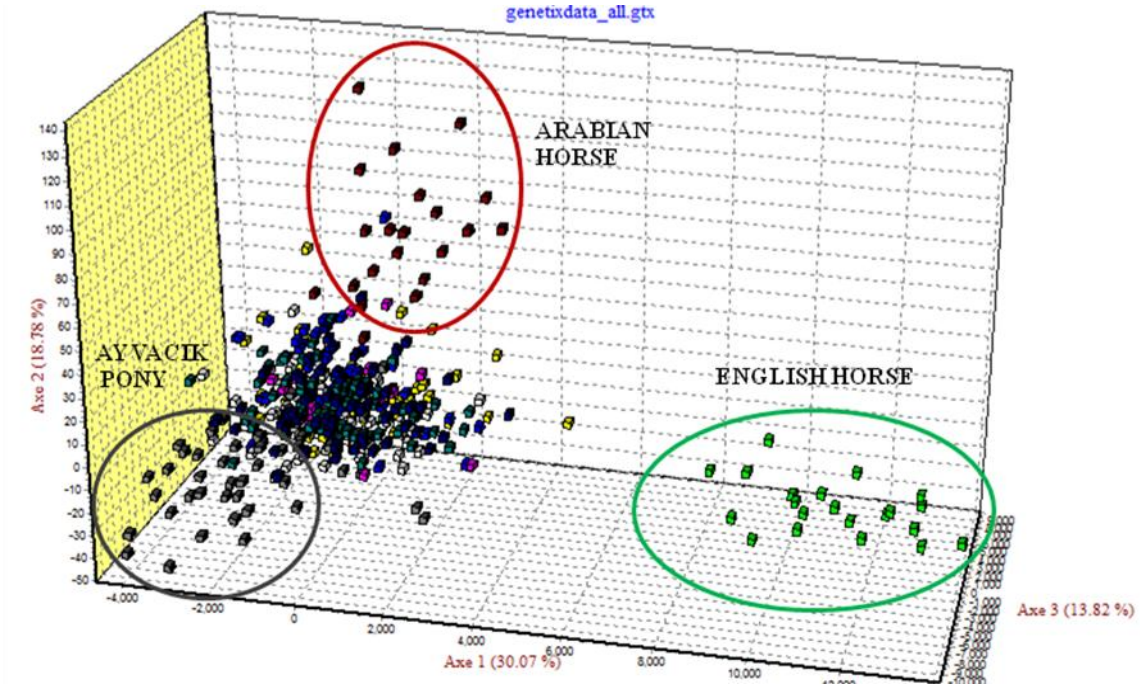


Figure 4.3. Factorial Correspondence Analysis of all individuals

Considering the results of FCA analysis including all of the samples, all breeds can be classified into four clusters; Arabian horses, English horses, Ayvacık Pony and rest of the Anatolian horses. English horses formed a quite homogenous group and were located more distantly from other breeds, as they were originated from a relatively small and inbred population. Arabian horses were also differentiated from the Anatolian breeds; however, they did not fall much far apart from the Anatolian native breeds. These results suggest that Arabian horses were probably used for crossbreeding with Anatolian horses. Ayvacık pony breed formed a separate group as can be expected from its phenotypic character. In addition, the differentiation between native Anatolian horses could be observed by excluding the Arabian and English horses, and providing prior phenotypic

data. Hence FCA was able to detect some differentiation within the Anatolian horse breeds.

Principal Component Analysis was also used for visualization and interpretation of the data by making two-dimensional plots. In a similar manner to the FCA analysis, most of the breeds could be separated from each other by the PCA analysis (Figure 4.4). Ayvacık Pony and Malakan were separated from the Hinis and Canik Horse samples with respect to the X axis, with the Cukurova Horses being in between and especially Ayvacık Pony being the farthest breed. The grouping of Canik and Hinis breeds is also concordant with that seen in the FCA analysis. Again in the PCA, Malakan Horse samples were separated from the other breeds with respect to the Y axis. Finally, when Arabian and English Horse breeds were included in the analysis (Figure 4.5.), they were separated from the rest of the Anatolian horse breeds with respect to the X axis.

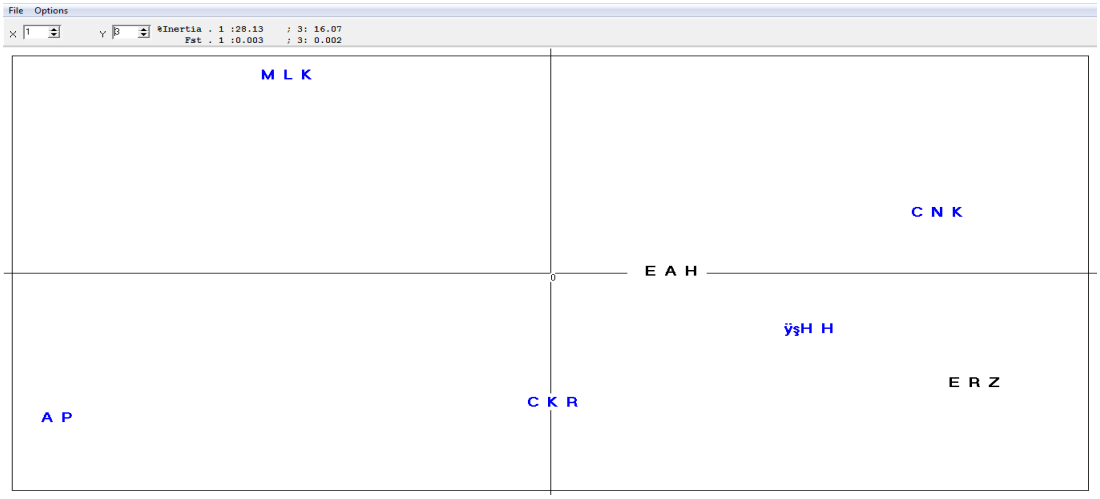


Figure 4.4. PCAgen results for the Anatolian horse breeds

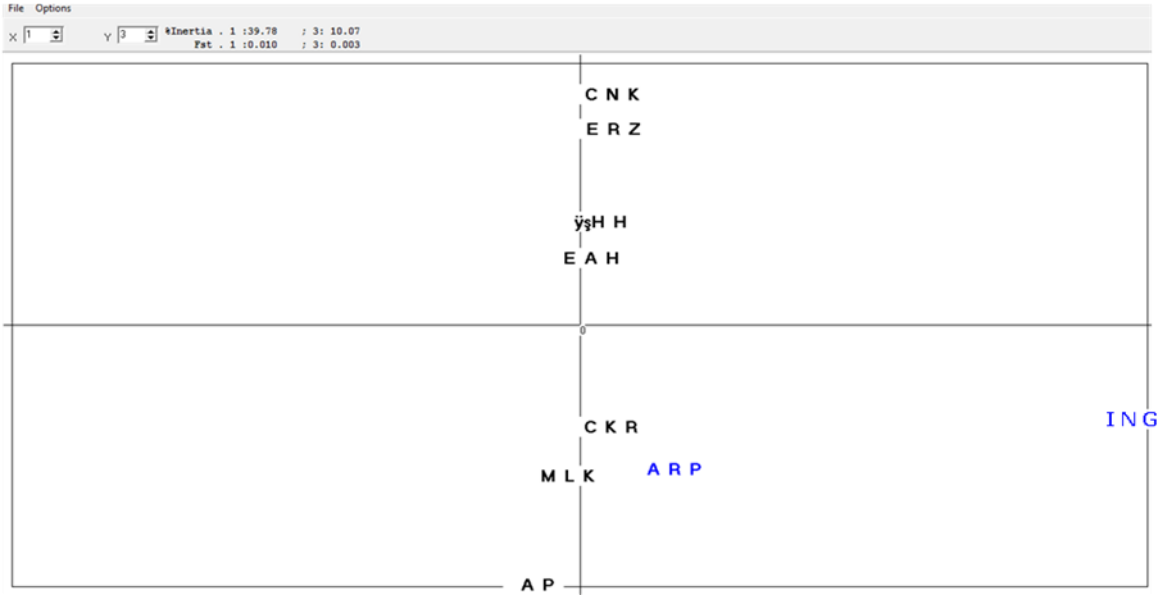


Figure 4.5. PCAgen results for all the breeds included in the study

The Bottleneck analysis results based on all of the loci, rejected the null hypothesis that "the population was at mutation-drift equilibrium" as the probabilities estimated were lower than 0.05. These results indicate that all breeds had been through a recent bottleneck. In the second bottleneck analysis, carried out by excluding the null alleles, some different results were obtained. The null hypothesis "the population was at mutation-drift equilibrium" was rejected for most of the breeds including Hinis, Erzurum, East Anatolian, Cukurova, and English breeds, as the probabilities estimated were lower than 0.05. For the Ayvacık Pony breed, the probability value for the heterozygosity excess was 0.27081, which was not statistically significant. These results are consistent with our knowledge that native breeds decreased rapidly in the past decades. Those bottleneck events may have resulted in an increase in the rate of inbreeding and in the loss of genetic variation. Some alleles may have been lost that were unique to Anatolian native breeds.

The k and g tests, implemented as a test of population expansion, indicated 14 loci with negative k values; however, the probability value of the k test was 0.0733. So, the population expansion was not significant according to the k test. The g test value was larger than the cut off value that was estimated in the Table 1 of Reich et al., 1999 (smaller

than 0.27) for the interlocus test. This also indicates that there was no significant expansion according to the interlocus test.

For studying the population structure of Anatolian breeds, the data was also analyzed using the Structure program. Using the whole data set, a Bayesian MCMC approach (10^5 iterations and burn-in of 10^5) was used to infer the number of likely clusters (*K*; 1-9) under an admixture model with correlated allele frequencies. The most likely cluster in the study was found to be K=4. The observation of the structure results of Anatolian breeds showed that breeds are highly variable and difficult to assign to a unique cluster (Figures 4.6. and 4.7.). In the multiple line plots of estimates (Figure 4.7), it is clear that the individuals of the Anatolian breeds were assigned to two or more clusters which indicated that they were admixed. On the other hand, the Arabian and English horse samples formed separate clusters from the rest of the Anatolian breeds.

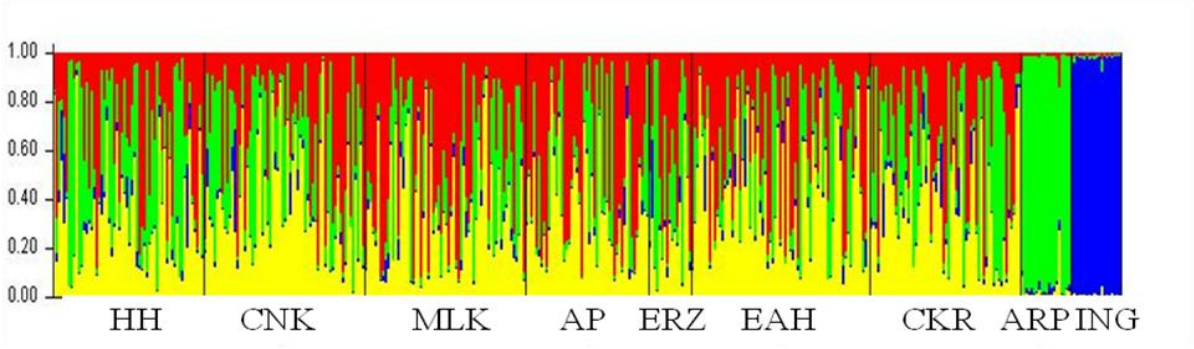


Figure 4.6. The bar chart by breed ID

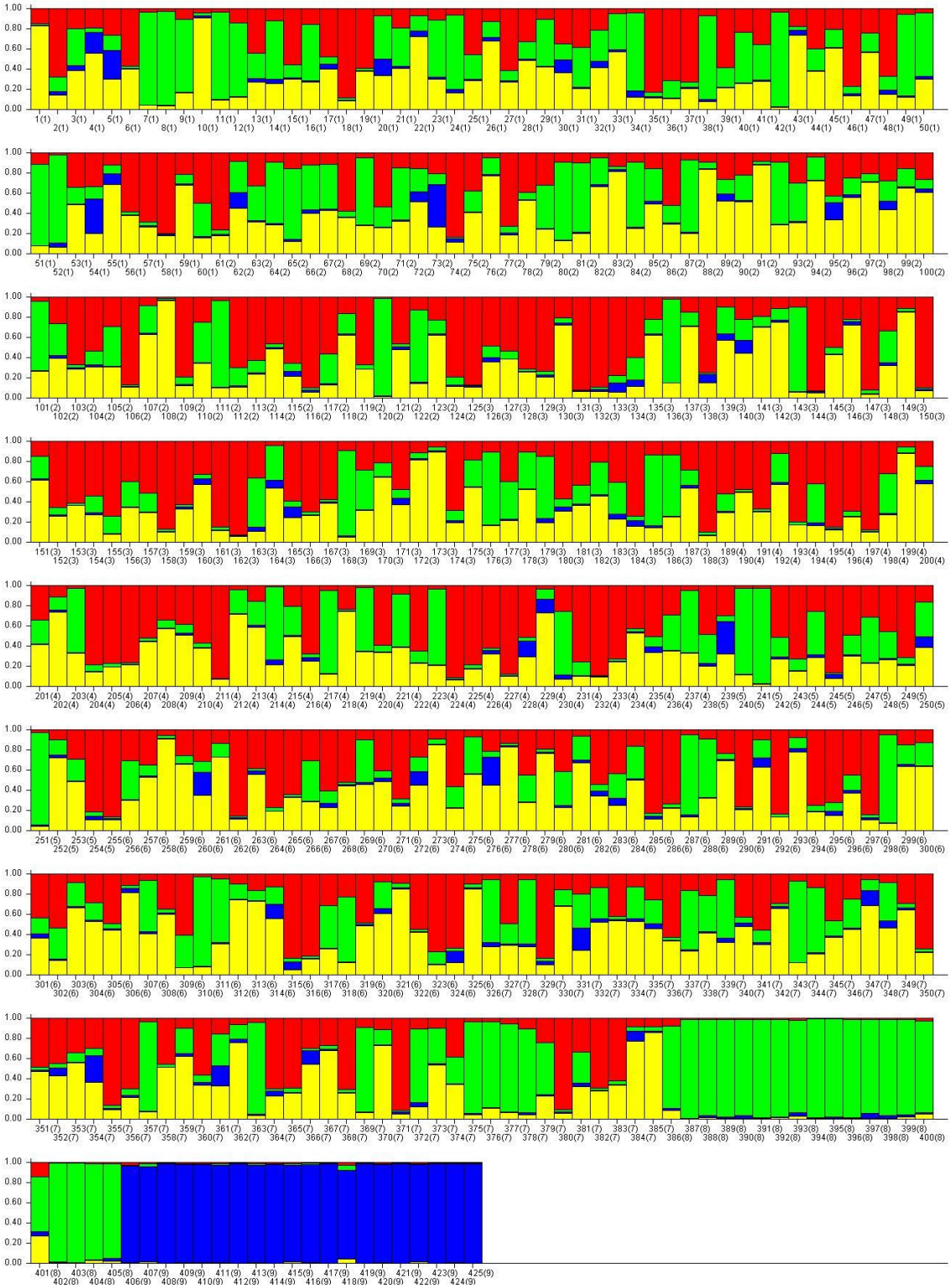


Figure 4.7. Plot of estimates in multiple lines representing the individual genotypes analysed in the study

The structure results indicated that there is no differentiation between Anatolian breeds. Only, English and Arabian horses were differentiated as separate clusters. These results reflect the close genetic relationship and gene flow between these five Anatolian breeds. The results suggest that none of the Anatolian breeds included in the study comprise a unique group of individuals, on the contrary, they all share common genetic diversity. Arabian horse seems to have contributed to Anatolian breeds to some extent, as there was evidence of assignment of the individuals of the Anatolian breeds (shown as green in Figure 4.6) to the Arabian breed cluster. These findings reflect the lack of proper breeding strategies for the Anatolian horses. Considering the structure results, Anatolian horses are very motile, admixed and genetically not distinguishable from each other.

Considering the comparable results from other countries, French breeds could be classified into four differentiated clusters (Leroy et al., 2009), Indian breeds formed separate clusters for their pony breeds (Behl et al., 2007), Feral Island population in Canada represented a unique group of individuals (Plante et al., 2007), Danish horses formed three distinct breeds (Thistrup et al., 2008) and in this study Arabian and Thoroughbred breeds formed separate clusters, but this was not the case for the Anatolian breeds. Hence, comparatively speaking, there seems to be lower levels of differentiation between the Anatolian breeds.

The neighbor-joining (NJ) tree shows the deepest branch connecting to the English Horse breed, which is consistent with the results from the other analyses. Arabian Horses were also separated from the Anatolian breeds with a bootstrap value of 84%. Considering the native breeds, they were separated from each other with a bootstrap range of 43-51 which is indicative of low levels of differentiation among breeds. Erzurum horses separated from the rest of the Anatolian breeds with a high bootstrap value, which might be due to the relatively low number of samples, and sampling of the breeds from grade horses without defined phenotypic characters.

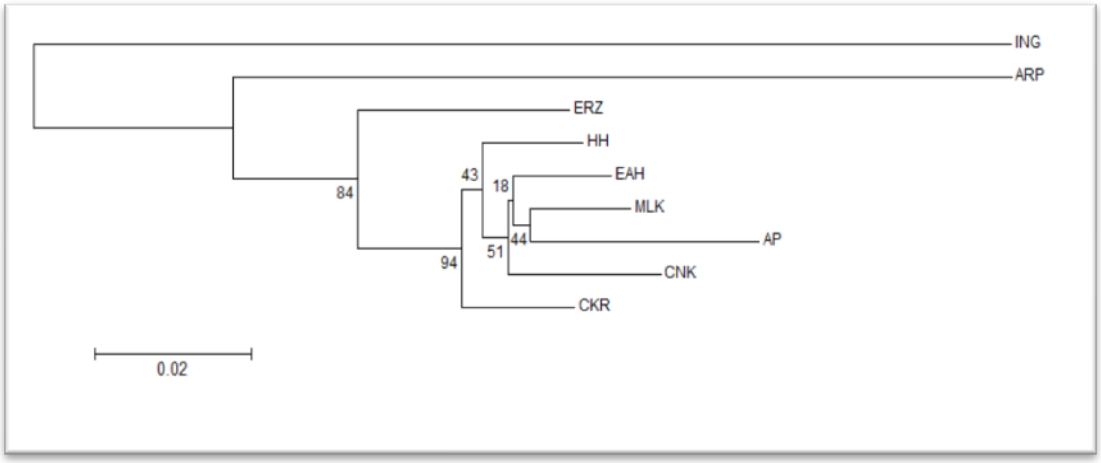


Figure 4.8. NJ tree of the samples based on pairwise D_A genetic distances

5. CONCLUSION AND RECOMMENDATIONS

The detection of high degree of differentiation between Anatolian breeds and Arabian and English breeds proves the reliability of the markers used and the genotype data obtained in this study. The present high allelic variability reflects the large genetic variation within Anatolian livestock horses when compared to the other breeds in literature; Lipizzaner Horses (Achmann et al., 2004), Spanish Trotter Horses (Azor et al., 2007), Lithuanian Heavy Draught Horses (Juras and Cothran, 2004), Polish Heavy Horses (Iwanczyk et al., 2006), Canadian Horses and Sable Island Horses (Plante et al., 2007), French Horses (Leroy et al., 2009), Japanese Horses (Tozaki et al., 2003), German Horses (Aberle et al., 2004), Haflinger Horses, Fjord Horses, Iceland Pony, Arab Horses (Leroy et al., 2009). Even the breed with the highest allelic diversity reported in these studies, the Polish Heavy Horses (Iwanczyk et al., 2006), have an average allelic diversity value (8.08) lower than that of the Anatolian Horse breeds (8.8-9.6). Under the selection process, as certain traits are being picked, and others are not, certain alleles can be expected to be lost from the gene pool. The higher allelic diversity in the Anatolian breeds probably reflects the relative lack of selection in Anatolia when compared to the other regions of the world.

In addition to the high levels of observed allelic diversity, higher levels of observed and expected heterozygosity estimates, the fact that the Anatolian horse breeds are not well differentiated in accordance with their origin and phenotypic characteristics are likely the results of the lack of proper breeding strategies in Anatolia. Although they seem to be carrying the morphological characteristics of the selected breeds, structure and assignment test results show that they are not genetically exhibiting homogenous populations. For instance, in Europe where artificial selection is undertaken more rigorously, various horse breeds are generally well differentiated from each other; Thoroughbred horses, Spanish horses (Canon et al., 2000). In the present study, Ayvacık Pony had some clustering and differentiation from the other breeds of Anatolian livestock. Various types of analyses indicated this differentiation. These include F_{IS} estimates, pairwise F_{ST} estimates, assignment test results, and FCA and PCA analyses results. This is

probably due to the selection process undertaken for this particular breed, in conjunction with their use in the collection of olives in Ayvacık region. This differentiation reflects the phenotypic selection process, which resulted in the very unique morphology, especially the low body stature of this breed. In the structure and NJ analyses, a strong population differentiation in Anatolian breeds could not be observed. Furthermore, the results revealed that there is no deviation from the Hardy-Weinberg equilibrium among the Anatolian breeds analyzed. These results indicate that there is high admixture of Anatolian breeds.

The results also indicate that Arabian horses are genetically closer to Anatolian horses than to the English horses. This can be expected due to geographic proximity of the mentioned regions, resembling an artificial case of 'isolation-by-distance". There has always been a connection between Arabians and Turks through trading, diplomacy and war in the past and present history, and horses have been an important subject of these interferences. Arabian Horses were obtained to ensure a supply for cavalry horses in the times of Ottoman Empire as well as to make use of the Arabian horse characteristics including speed and endurance in horse races. For these reasons, numerous Arabian Horses were imported and cross bred to Anatolian Horses. This relationship is clear in the FCA analysis, which clustered Anatolian and Arabian horses in two close groups on the 3D space, which were quite away from the British horses. In addition, the structure analysis resulted in the assignment of Anatolian breeds to the Arabian breed, but not *vice versa*. This finding supports the archive records in the history that horses from Arabia were actually imported into Anatolia to improve the quality of the Anatolian breeds. However, the direction of the gene flow was only in one direction, and not from Anatolia to Arabia, as almost no Arabian horses were seen to be assigned to the Anatolian breeds.

Results of the present study are in partial accordance with an earlier study which analyzed the same sample DNAs based on sequence diversity at their mtDNA control region (Aslan, 2009). mtDNA results revealed that individuals of the same breed had different haplotypes grouped in different clades. Also, individuals from geographically distant breeds were contained within the same clade. Consequently, none of the breeds had

individuals coming from only one mtDNA clade. Furthermore, all seven globally recognized haplogroups were identified in the Anatolian samples. These results could be attributed to the Anatolia's land-bridge position between Asia and Europe that played an important role in the historical trade routes when horses were used as a means of transportation. In the mtDNA results, some of the haplogroups show clines from Europe to Asia through Anatolia and *vice versa*. In this regard, the mtDNA findings also unveiled that Anatolia has been a passageway for horse breeds between Asia and Europe (Koban et al., submitted, 2010).

However, mtDNA analyses could not differentiate between domestic Anatolian breeds in the level of mtDNA clades. In other words, it was not possible to distinguish between the breeds by analyzing their mtDNA sequence data. Similar results were obtained by using the microsatellite markers. Yet, FCA analysis, performed by using prior population information parameter, could capture some differentiation between domestic Anatolian breeds, especially for Ayvacık Pony. Furthermore, Canik and Malakan Horses also showed some differentiation, whereas, Cukurova and Hinis samples had the greatest overlap.

In combination with the mtDNA results investigating maternal lineages and evolutionary history, the present study based on microsatellites revealed the high genetic diversity and population structure present among the Anatolian horse breeds. However, the polymorphic loci used in the present study provided little information for intraspecific variation and assignment tests. For further and deeper analysis, a higher number of polymorphic loci and more samples may provide a better understanding at the intraspecific level. As a complementary study, Y-chromosome genotyping analysis to investigate the relationships among the breeds' paternal lineages could be carried out.

Global warming and associated climate changes increase the concern related to conservation genetics of locally adapted species that represent the adaptive potential for specific habitats, which acquire particular phenotypic traits and continue exhibiting the

historical patterns of gene flow. Accordingly, the high genetic diversity revealed both by mtDNA and microsatellite analysis suggests that Anatolian native breeds, which are known as not having been exposed to intense artificial selection, may possess unique genetic traits which can be useful to study and explore genes related with the heritable diseases as well as the genes related with the resistance to harsh environmental conditions.

From a conservation point of view, the results suggest that an immediate action plan is needed for Malakan breed as the heterozygosity levels were found to be lower than expected. This result may indicate both population substructuring (Wahlund effect) as the samples were collected from different areas or it may indicate non-random mating resulting in the loss of neutral or locally adapted alleles. Considering, the fact that Malakan horses have experienced a sharp population decline in the last two decades and low heterozygosity levels, it will be a better strategy to revise the management of that breed. Furthermore, Canik breed could be given priority as it shows the highest levels of genetic variation (Weitzman, 1993; Reist-Marti et al., 2003). The high genetic diversity may provide material for developing successful breeding programs in the future and may provide genetic resources for the evolution of unique adaptations. Among all, Ayvacık Pony is important for the Ayvacık region as a working horse for olive collection. This breed is dwindling in number as the olive trees are cut down each year and the mechanization in the field is increasing. For this reason, it is essential to protect Ayvacık Pony breed, which has a unique morphology and genotypic distinctness.

Rapid changes in production systems are decreasing the number of domestic horse breeds and resulting in the undocumented loss of locally adapted alleles. Besides, active trading habits, lack of proper breeding strategies and replacement and crossbreeding of our native horses with foreign breeds for better race characters are leading to the loss of genetic differentiation and clear separation among the breeds. It is important to ensure that all the remaining domestic horse breeds are effectively protected as they carry alleles selected during the domestication events, as well as locally adapted alleles. It is equally important to preserve the breeds that have large genetic diversity and genetically differentiated breeds. Anatolian breeds present in this study have high allelic diversity,

whereas; they lack genetic differentiation among each other. However, from a global point of view, preserving the Anatolian breeds would be sound as genetically distinct breeds from different domestication centers may represent the migration routes and centers of origin. Here, the primary objective should be the conservation of the present genetic diversity for potential future uses. Hence, the best method for conservation can be *in-situ* preservation of these animals in their native environments so that they can continue adapting to their natural habitats. Ideally, providing an environment as close as possible to the ancestral conditions including selective pressures is the best solution. However, as many of the native habitats have been altered and local populations changed their way of life, ex-situ conservation by cryopreserving the DNA, embryo or sperm samples is also recommended.

6. REFERENCES

Aberle, K.S., Hamann, H., Drögemüller, C., Distl, O., 2004. Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. Animal Genetics, 35, 270-277.

Achmann, R., Curik, I., Dovc, P., Kavar, T., Bodo, I., Habe, F., Marti, E., Solkner, J., Brem, G., 2004. Microsatellite diversity, population subdivision and gene flowin the Lipizzan horse.Animal Genetics, 35, 285–292.

Anthony, D.W., 1996. In: Olsen, S.L. (Ed.), Horses Through Time. Roberts Rinehart for Carnegie Museum of Natural History, Boulder, CO, 57-82.

Aslan, O., 2009. The conservation of genetic diversity in Turkish native horse breeds. M.S. Thesis, Boğaziçi University.

Avise, J.C., 1994. Molecular markers, natural history and evolution.Chapman and Hall, New York.

Azor,P.,Valera, M., Gómez, M., Goyache, F., Antonio, M., 2007. Genetic characterization of the Spanish Trotter horse breed using microsatellite markers. Genetics and Molecular Biology, 30, 37-42.

Banks, M.A., Eichert, W., Olson, J.B., 2003. Which genetic loci have greater population assignment power? Bioinformatics, 19, 11, 1436-8.

Batu, S., 1938. Turk horses and horse breeding knowledge. Ankara Yüksek Ziraat Enstitüsü, Lecture Book, Number: 3, Ankara.

Behl, R., Behl, J., Gupta, N., Gupta, S.C., Ahlawat, S.P.S., Ragnekar, M., Ahmed, Z., 2006. Genetic characterization of Zanskari breed of horse. Journal of Genetics, 85, 3.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 1996-2004. GENETIX 4.05, logicielsous Windows TM pour la genetique des populations. Montpellier (France): Laboratoire Genome, Populations, Interactions, CNRS UMR 5000, Universite de Montpellier II.

Bennett, D., Hoffmann, R.S., 1999. Equuscaballus. Mammalian Species, 628, 1-14.

Bignon, O., Baylac, M., Vigne, J.-D., Eisenmann, V., 2005. Geometric morphometrics and the population diversity of Late Glacial horses inWestern Europe (Equuscaballusarcelini): phylogeographic and anthropological implications. Journal of Archaeological Science, 32, 375-391.

Bilgin, R., 2007. Kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites. Molecular Ecology Notes, 7, 416-417.

Binns, M.M., Holmes, N.G., Holliman, A., Scott, A.M., 1995. The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing.British Veterinary Journals, 151, 9-15.

Birky, C.W., Maruyama, T., Fuerst, P., 1983. An approach to population genetic and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics, 103, 513–527.

Bjornstad, G., Midthjell, L., Roed, K.H., 2000. Characterization of 10 equine dinucleotide microsatellite loci: *NVHEQ21*, *NVHEQ54*, *NVHEQ67*, *NVHEQ70*, *NVHEQ75*, *NVHEQ77*, *NVHEQ79*, *NVHEQ81*, *NVHEQ82* and *NVHEQ83*. Animal Genetics, 31**,** 78-79.

Bjornstad, G., Gunby, E., Roed, K.H., 2000. Genetic structure of Norwegian horse breeds. Journal of Animal Breeding and Genetics, 117, 307-317.

Breen, H., Downs, P., Irwin, Z., Bell, K., 1994. Intrageneric amplification of horse microsatellite markers with emphasis on the Przewalski_s horse (E. przewalskii). Animal Genetics, 25, 401-5.

Breen, M., Lindgren, G., Binns, M.M., Norman, J., Irvin, Z., Bell, K., Sandberg, K., Ellegren H., 1997. Genetical and physical assignments of equine microsatellites first integration of anchored markers in horse genome mappings. Mammalian Genome, 8, 267-73.

Buitenhuis, H., 1997. Aşıklı Höyük: A'Prododomestication' Site. In: M. Kokabi and J. Wahl (eds.). Proceedings of the 7th ICAZ Conference Actes du 7e Colloque International d'Archeozoologie. Societe de Recherche Interdisciplinaire. Anthropozoologica 25-26:655-662.

Civanova, K., Putnova, L., Dvorak, J., 2006. Analysis of microsatellite set for biodiversity studies in horses. Acta Technica et Zootechnica, 39.

Clutton-Brock, J., 1999. A Natural History of Domesticated Mammals, (Second Ed.) Cambridge University Press, Cambridge.

Cock, O., Bill, H., Derek, W., Peter, S., 2005. Microsatellite Data Checking Software , Micro-Checker, University of Hull, Department of Biological Sciences and the Department of Computer Science.

Coogle, L., Bailey, E., Reid, R., Russ, M., 1996. Equine dinucleotide repeat loci from LEX025 to LEX033. Animal Genetics, 27, 289-290.

Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144, 2001-2014.

Cunningham, E.P., Dooley, J.J., Splan, R.K., Bradley, D.G., 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Animal Genetics, 32, 360–364.

Cregan, B., Mudge, J., Fickus, E.W., Marek, L.F., Danesh, D., Denny, R., Shoemaker, R.C., Matthews, B.F., Jarvik, T., Young, N.D., 1999. Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. Theoretical and Applied Genetics, 98, 919-928.

Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences of the USA, 91, 3166- 3170.

Eggleston-Stott, M.L., DelValle, A., Bautista, M., Dileanis, S., Wictum, E., Bowling, A.T., 1997. Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. Animal Genetics, 28, 370-371.

Ellegren, H., Johannsson, M., Sandberg, K., Andersson, L., 1992. Cloning of highly polymorphic microsatellites in horses. Animal Genetics, 23, 133-42.

Esin, U., 1999, Introduction, The Neolithic in Turkey: A General Review. In: M. Özdoğan, N.Başgelen (eds.). Neolithic in Turkey. The Cradle of Civilization. Arkeoloji ve Sanat Yayınları. İstanbul:13-23.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1, 47-50.

Frankham, R., Ballou, J.D., Briscoe, D.A., 2002. Introduction to Conservation Genetics, Cambridge University Press, New York.

Gibson, J.P., Ayalew, W., Hanotte, O., (in press). Measures of diversity as inputs for decisions in conservation of livestock genetic resources. CABI.

Gomez, M.D., Moll, P., Roca, B., Azor, P.J., Valera, M., 2005. Genetic evaluation systems of trotting horses in Europe. Medicina Military, 61, 228-230.

Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2.9.3. http://www.unil.ch/izea/softtwares/fstat.html. (Accessed January 2007).

Guerin, G., Bertaud, M., Amigues, Y., 1994. Characterization of seven new horse microsatellites: HMS1, HMS2, HMS3, HMS5, HMS6, HMS7 and HMS8. Animal Genetics, 25, 62.

Gregoire, L., Lucille, C., Etienne, V., Jean-Claude, M., Anne, R., Coralie, D.-B., Xavier R., 2009. Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism, Genetics Selection Evolution, 41, 5.

Hanotte, O., Jianlin, H., 2005. Genetic characterization of livestock populations and Its use in conservation decision making, The Role of Biotechnology, Villa Gualino, Turin, Italy, 5-7.

Heywood, V.H., Iriondo, J.M., 2003. Plant conservation: Old problems, new perspectives. Biological Conservation, 113, 321-335.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setala, H., Symstad, A.J., Vandermeer, J., Wardle, D.A., 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge, Ecological Monographs, 75, 1, 3-35.

Hopman, T.J., Han, E.B., Story, M.R., Schug, M.D., Aquadro, C.F., Bowling, A.T., Murray, J.D., Caetano, A.R., Antczak, D.F., 1999. Equine dinucleotide repeat loci COR001–COR020. Animal Genetics, 30, 225-6.

Hongo, H., Meadow, R.H., Öksüz, B., İlgezdi, G., 2004. Animal Exploitation at Çayönü Tepesi, Southeastern Anatolia. TÜBA-AR, VII, 107-119.

Irvin, Z., Giffard, J., Brandon, R., Breen, M., Bell, K., 1998. Equine dinucleotide repeat polymorphisms at loci ASB21, 23, 25 and 37–43. Animal Genetics, 29, 67.

Juras, R., Cholewinski, G., Cothran, E.G., 2006. Genetic structure and phylogenetic relationships of the Polish Heavy horse**.** Journal of Applied Genetics, 47, 4, 353-9.

Jansen, T., Forster, P., Levine, M.A., Oelke, H., Hurles, M., Renfrew, C., Weber, J., Olek, K., 2002. Mitochondrial DNA and the origins of the domestic horse.Proceedings of the National Academy of Sciences of the United States of America, 99, 10905-10910.

Juras, R., Cothran, E.G., 2004. Microsatellite in Lithuanian native horse breeds: usefulness for parentage testing. Biologija, 4, 6-9.

Kavar, T., Dovc, P., 2008. Domestication of the horse: Genetic relationships between domestic and wild horses. Livestock Science, 116, 1-14.

Kimura, M., Crow, J.F., 1964. The number of alleles that can be maintained in a finite population. Genetics, 49, 725-738.

Kimura, M., Ohta, T., 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proceedings of the National Academy of Sciences of the USA, 75, 2868-2872.

Knapik, E.W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M.C., Jacob, H.J., 1998. A microsatellite genetic linkage map for zebrafish (*Daniorerio*). Nature Genetics, 18, 338-343.

Koban, E., Denizci, M., Aslan, O., Aktopraklıgil, D., Aksu, S., Bilgin, R., Bower, M., Erdag, B., Balcıoglu, K., Bahadır, A., Bagıs, H., Arat, S. 2010 (submitted). High microsatellite and mitochondrial diversity in Anatolian native horse breeds shows Anatolia as a genetic conduit between Europe and Asia. Animal Genetics.

Levine, M.A., 2005. Domestication and early history of the horse. In: Mills, D.M., McDonnell, S.M. (Eds.), The Domestic Horse: The Origins, Development, and Management of Its Behaviour. Cambridge University Press, Cambridge, 5-22.

Levine, M.A., 2006. mtDNA and horse domestication: the archaeologist's cut. In: Mashkour, M. (Ed.), Equids in Time and Space. Proc. 9th Int. Conf. of Archaeozoology.Oxbow Books, Oxford, 192-201.

Li, W-H.,1997. Molecular evolution. Sinauer Associates, Inc.

Lindgren, G., Backstrom, N., Swinburne, J., Hellborg, L., Einarsson, A., Sandberg, K., Cothran, G., Vila, C., Binns, M., Ellegren, H., 2004. Limited number of patrilines in horse domestication. Nature Genetics, 36, 335-336.

Lister, A.M., Kadwell, M., Kagan, L.M., Jordan, W. C., Richards, M.B., Stanley, H.F., 1998.Ancient and modern DNA in a study of horse domestication. Ancient Biomolecules, 2, 267-280.

Litt, M.,Luty, J.A., 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics, 44, 397-401.

Machugh, D.E., Loftus, R.T., Bradley, D.G., Sharp, P.M., Cunningham, P., 1994. Microsatellite DNA variation within and among European cattle breeds Proceedings of the Royal Society B, 256, 25-31.

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., Leblond, S., Earle Macdonald, J., De Jong, J., Wieringa, B.,1992. Myotonic dystrophy mutation: An unsTable CTG repeat in the 38 untranslated region of the gene. Science, 255, 1253-1258.

Marti, E., Breen, M., Fischer, P., Swinburne, J., Binns, M.M., 1998. Six new cosmid derived and physically mapped equine dinucleotide repeat microsatellites. Animal Genetics, 29, 236-8.

Mills, D.S., McDonnell, S.M., 2005. The Domestic Horse: The Origins, Development and Management of Its Behaviour. Cambridge University Press.

Moxon, Richard, E.,Christopher, W., 1999. DNA Microsatellites:Agents of Evolution? Scientific American, 94-99.

Naderi, S., Rezaei, H.R., Pompanon, F., Blum, M.G., Negrini, R., Naghash, H.R., Balkiz, O., Mashkour, M., Gaggiotti, O.E., Ajmone-Marsan, P., Kence, A., Vigne, J-D., Taberlet, P., 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. Proceedings of the National Academy of Sciences USA, 105(46), 17659-17664.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89, 583-590.

O'Donnell, W.T., Warren, S.T., 2002. A decade of molecular studies of fragile X syndrome. Neurosciences, 25, 315-338.

Ortí, G., Pearse, D. E., Avise, J. C., 1997. Phylogenetic assessment of length variation at a microsatellite locus. Proceedings of the National Academy of Science, 94, 10745-10749.

Outram, A.K., Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., Evershed, R.P.**,** 2009. The Earliest Horse Harnessing and Milking. Science, 323, 1332 – 1335.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288-295.

Pedro-Javier, A., Mercedes, V., Maria-Dolores, G., Felix, G., Antonio, M., 2007. Genetic characterization of the Spanish Trotter horse breed using microsatellite markers. Genetics and Molecular Biology, 30, 1.

Perkins, D., J.R., 1969. Fauna of catal huyuk: evidence for early cattle domestication in anatolia. Science, 164, 177-179.

Peters J., Helmer, D., von den Driesch, A., Segui, M.S., 1999, Early Animal Husbandry in the Northern Levant. Paleorient, 25/2, 27-47.

Peters, J., Schmidt, K., 2004. Animals in the Symbolic World of Pre-Pottery Neolithic Göbekli Tepe, South-Eastern Turkey: a Preliminary Assessment. Anthropozoologica *39/1*, 179-218.

Petit, E., Balloux, F., Excoffier, L., 2002. Mammalian population genetics. Why not Y. Trends in Ecology and Evolution, 17, 28-33.

Plante, Y., Vega-Pla, J.L., Lucas, Z., Colling, D., de March, B., Buchanan, F., 2007. Genetic diversity in a feral horse population from sable island, Canada. Journal of Heredity, 98, 6, 594-602.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Qiu-Hong, W., Hua, W., Tsutomu, F., Sheng-Guo, F., 2004. Which genetic marker for which conservation genetics issue? Electrophoresis, 25, 2165-2176.

Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Heredity, 86, 248-249.

Reich, D.E., Goldstein, D.B., 1998. Genetic evidence for a Paleolithic human population expansion in Africa. Proceedings of the National Academy of Sciences, USA, 95, 8119-8123

Reich, D.E., Feldman, M.W., Goldstein, D.B., 1999. Statistical properties of two tests that use multilocus data sets to detect population expansions. Molecular Biology and Evolution, 16**,** 453-466.

Reist-Marti, S.B., Gibson, J., Rege, J.E.O., Simianer, H., Hanotte, O., 2003. Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. Conservation Biology, 17, 1299–1311.

Ruth, L.S., Hopman, T.J., Schug, M.D., Aquadro, C.F., Bowling, A.T., Murray, J.D., Caetano, A.R., Antczak, D.F., 1999. Equine dinucleotide repeat loci COR041-COR060. Animal Genetics, 30, 320-321.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: A laboratory manual, Second Ed., 3 vol., Cold-Spring Harbor, New York.

Schlötterer, C., Amos, B., Tautz, D., 1991. Conservation of polymorphic simple sequence loci in cetacean species. Nature, 354, 63-65.

Solis, A., Jugo, B.M., Mériaux, J.C., Iriondo, M., Mazon, L.I., Aguirre, A.I., Vicario, A., Estomba, A., 2005. Genetic diversity within and among four south European native horse breeds based on microsatellite DNA analysis: implications for conservation. Journal of Heredity, 96, 6, 670-678.

Stallings, R.L., 1994. Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequence: Implication for human genetic diseases. Genomics, 21, 116-121.

Swinburne, J.E., Lockhart, L., Aldridge, V., Marti, E., Breen, M., Binns, M.M., 2000b. Characterization of 25 new physically mapped horse microsatellite loci: AHT24–48. Animal Genetics, 31, 23-38.

Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., Miura, N., Choi-Miura, N.H., Tomita, M., 2001a. Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. Journal of Veterinary Medical Science, 63, 1191-7.

Tozaki, T., Takezaki, N., Hasegawa, T., Ishida, N., Kurusawa, M., Saitou, N., Mukoyama, H., 2003. Microsatellite variation in Japanese and Asian horses and their phylogenetic relationship using a European horse out group. Journal of Heredity, 94, 374-380.

Valera, M., Molina, A., Gutierrez, J.P., Gomez, J., Goyache, F., 2005. Pedigree analysis in the Andalusian horse: Population structure, genetic variability and influence of the Carthusian strain. Livestock Production Science, 95, 57-66.

Van Haeringen, H., Bowling, A.T., Scott, M.L., Lenstra, J.A., Zwaagstra, K.A., 1994. A highly polymorphic horse microsatellite locus: VHL20. Animal Genetics, 25, 207.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P.F., 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538.

Vila, C., Leonard, J.A., Gotherstrom, A., Marklund, S., Sandberg, K., Liden, K., Wayne, R.K., Ellegren, H., 2001.Widespread origins of domestic horse lineages.Science, 291, 474-477.

Wallner, B., Piumi, F., Brem, G., Muller, M., Achmann, R., 2004. Isolation of Y chromosome-specific microsatellites in the horse and cross-species amplification in the genus Equus. Journal of Heredity, 95, 158-164.

Weber, J.L., May, P.E., 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics, 44, 388–396.

Weir, B.S., Cockerham, C.C.,1984. Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358-1370.

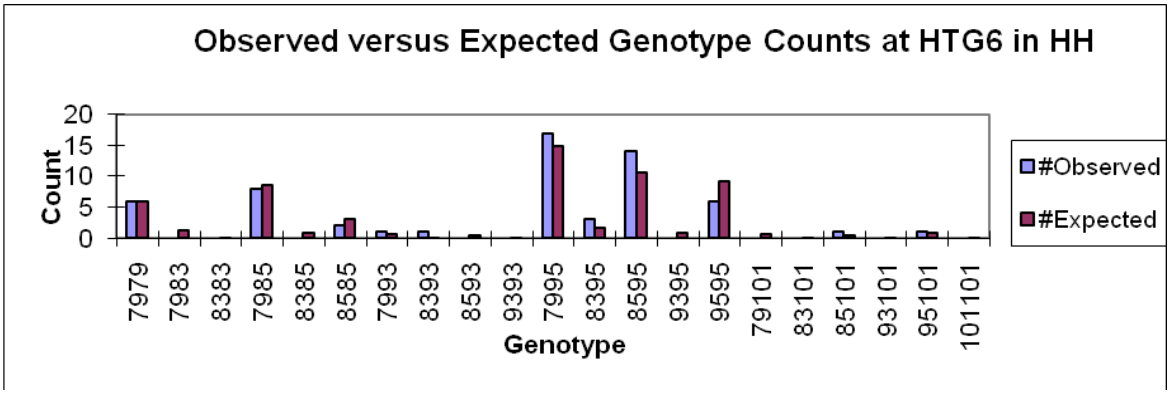
Weitzman, M., 1993. What to preserve? An application of diversity theory to crane conservation. Quarterly Journal of Economics, 108, 157–183.

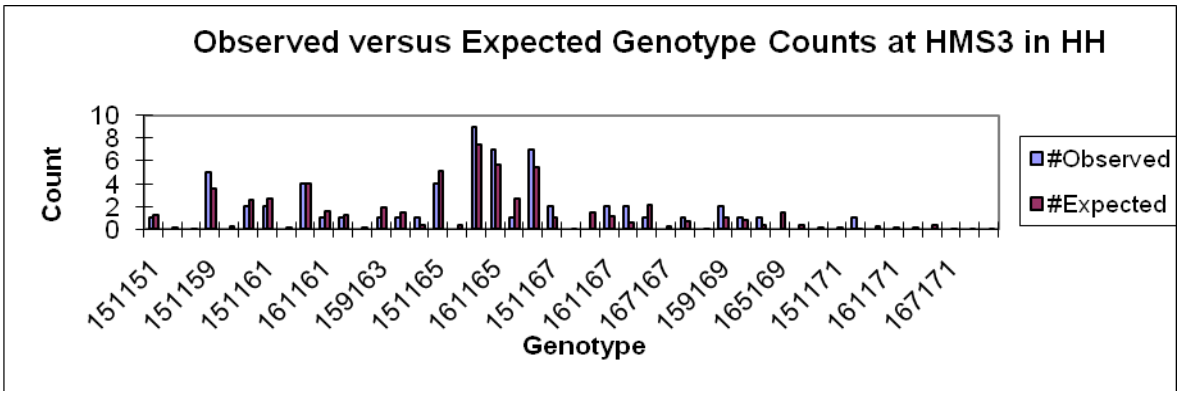
Zeder, M.A., Hesse, B., 2000. The initial domestication of goats (*Capra hircus*) in the Zagros Mountain 10,000 years ago. Science, 287, 2254–2257.

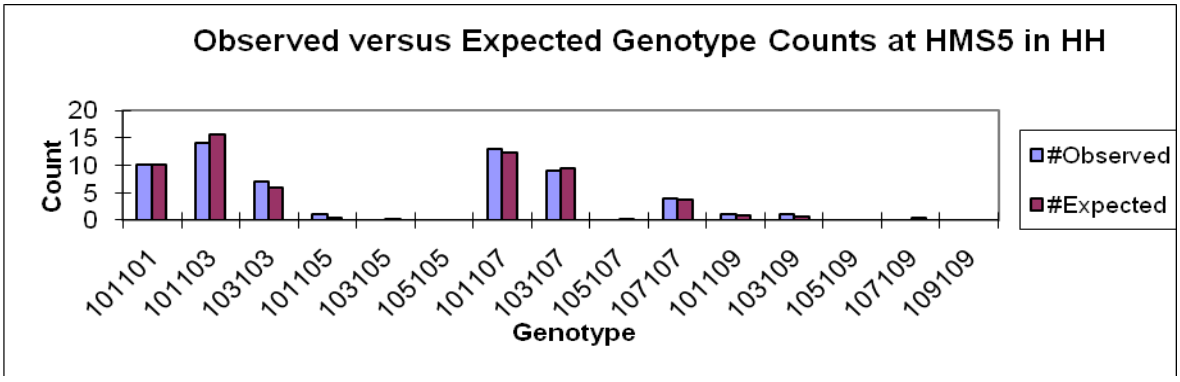
Zeder, M.A., 2008. Domestication and early agriculture in the Mediterranean basin: Origins, diffusion and impact. Proceedings of the National Academy of Sciences USA, 105, 11597-11604.

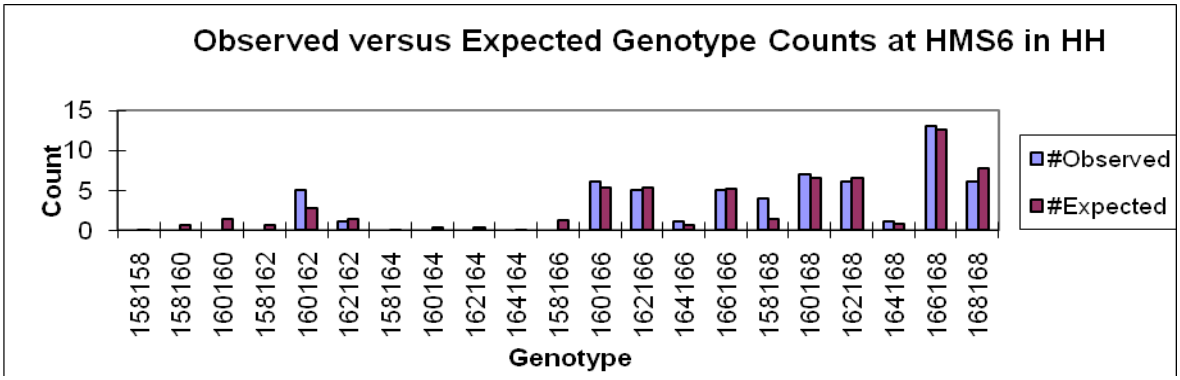
APPPENDIX A: OBSERVED VS EXPECTED GENOTYPE COUNTS OF ALL LOCI FOR ANATOLIAN POPULATIONS

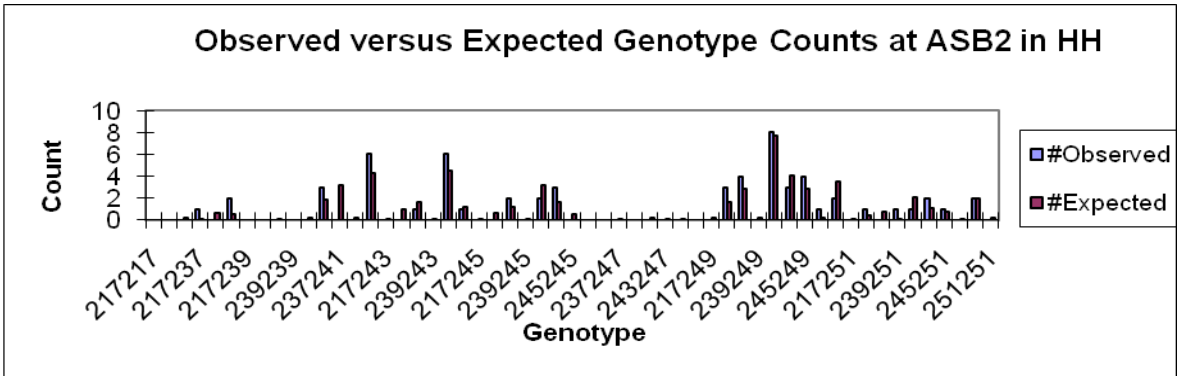
HINIS HORSE

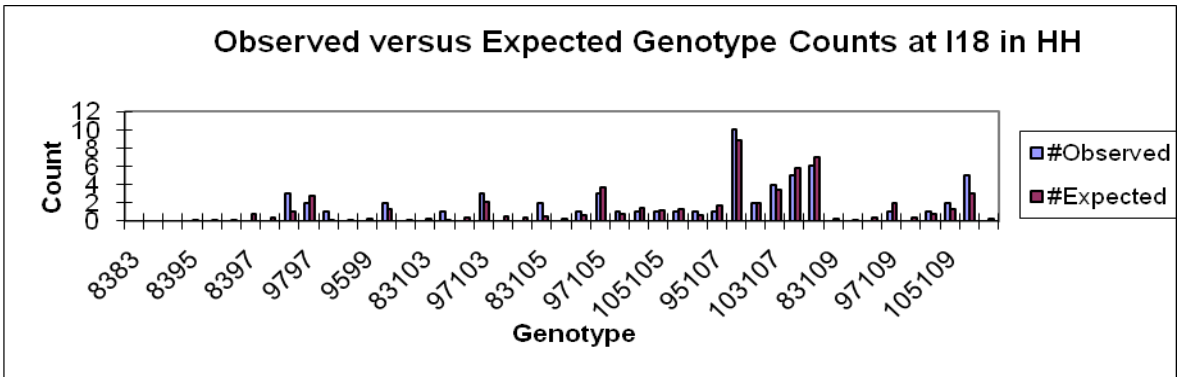


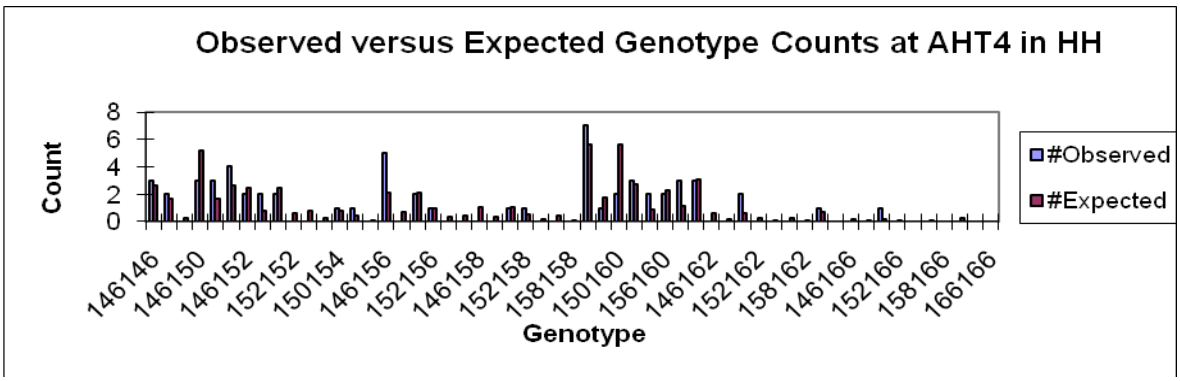


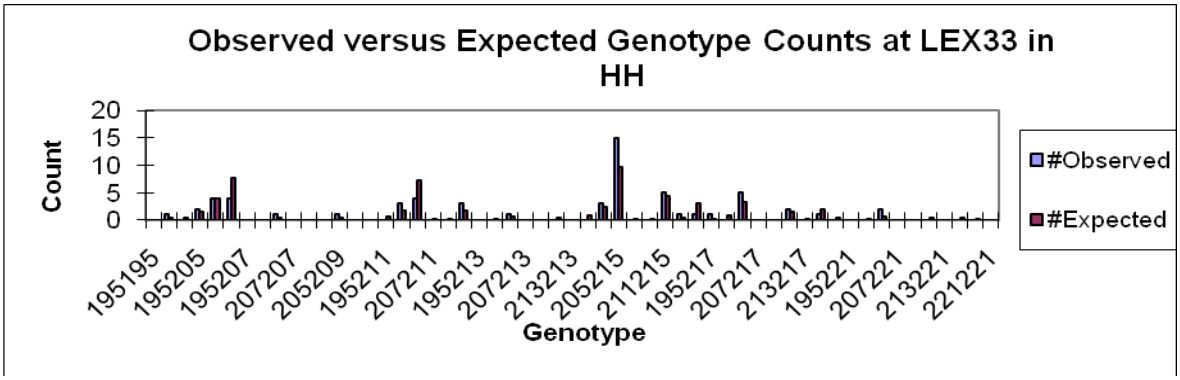


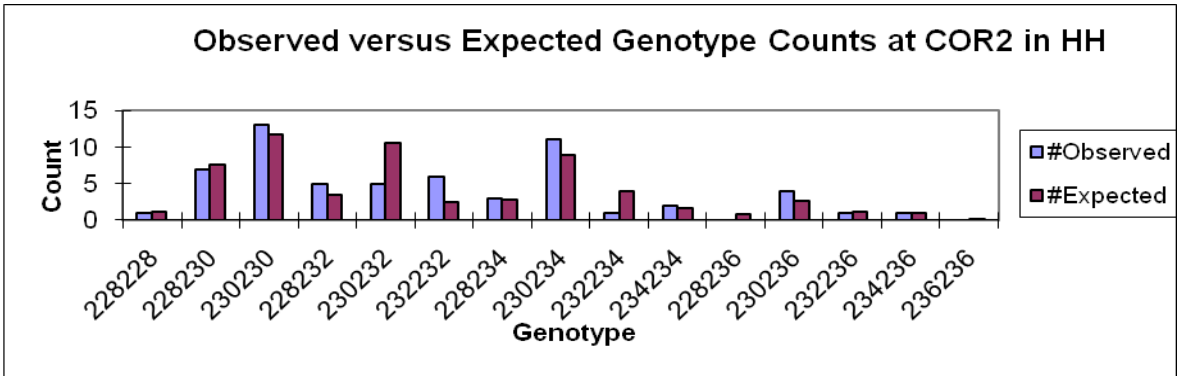

Observed versus Expected Genotype Counts at HMS6 in HH
Count
Genotype
#Observed
#Expected
158158 158160 160160 158162 160162 162162 158164 160164 162164 164164 158166 160166 162166 164166 166166 158168 160168 162168 164168 166168

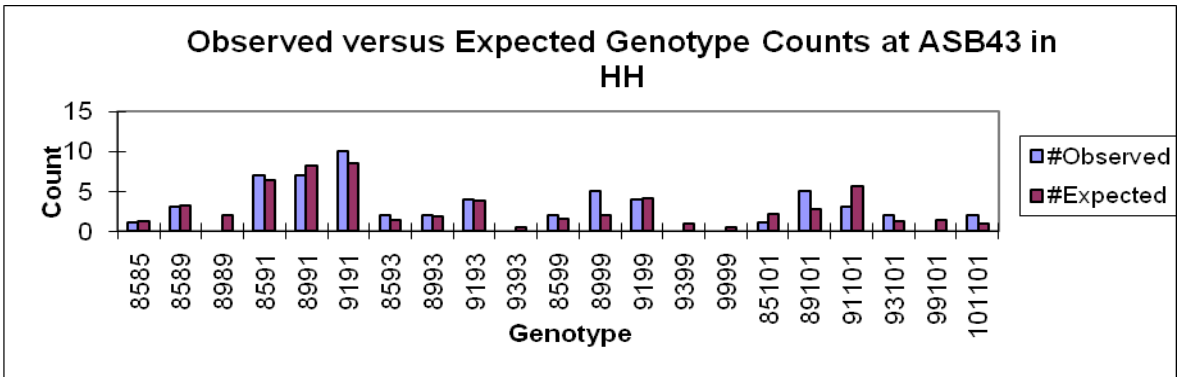

Observed versus Expected Genotype Counts at ASB2 in HH
Count
Genotype
#Observed
#Expected
217217 217237 217239 239239 237241 217243 239243 217245 239245 245245 237247 243247 217249 239249 245249 217251 239251 245251 251251

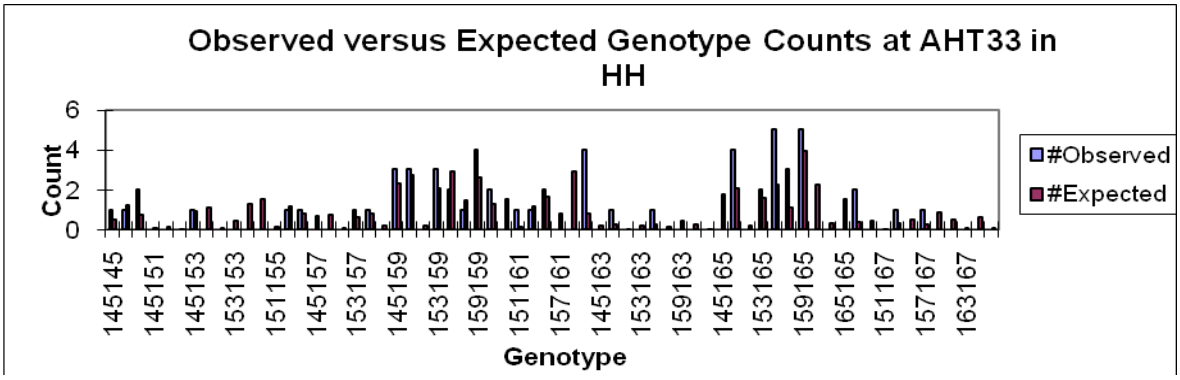

Observed versus Expected Genotype Counts at I18 in HH
Count
Genotype
#Observed
#Expected
8383 8395 8397 9797 9599 83103 97103 83105 97105 105105 95107 103107 83109 97109 105109

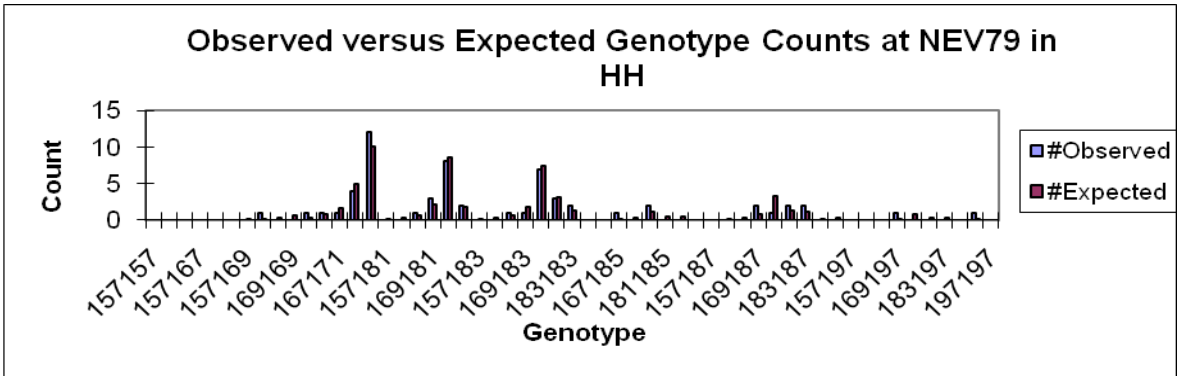

Observed versus Expected Genotype Counts at AHT4 in HH
Count
Genotype
#Observed
#Expected
146146 146150 146152 152152 150154 146156 152156 146158 152158 158158 150160 156160 146162 152162 158162 146166 152166 158166 166166

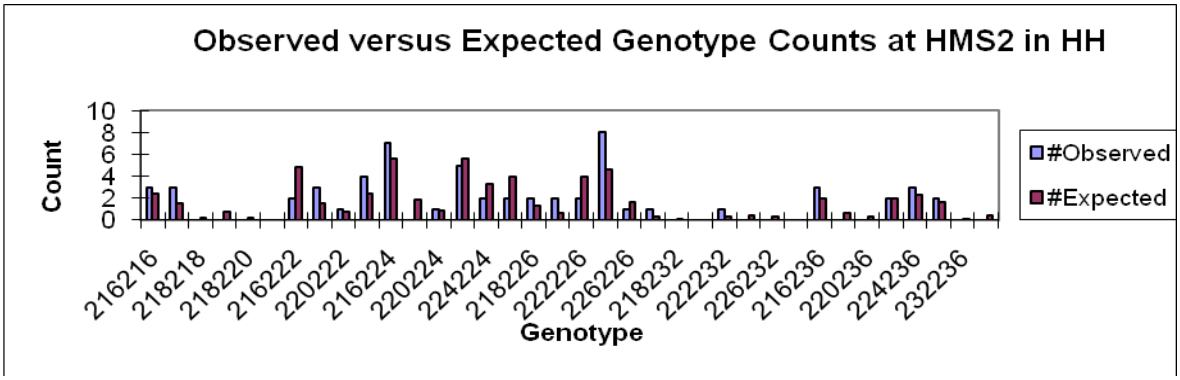

Observed versus Expected Genotype Counts at LEX33 in HH
Count
Genotype
#Observed
#Expected
195195
195205
195207
207207
205209
195211
207211
195213
207213
213213
205215
211215
195217
207217
213217
195221
207221
213221
221221

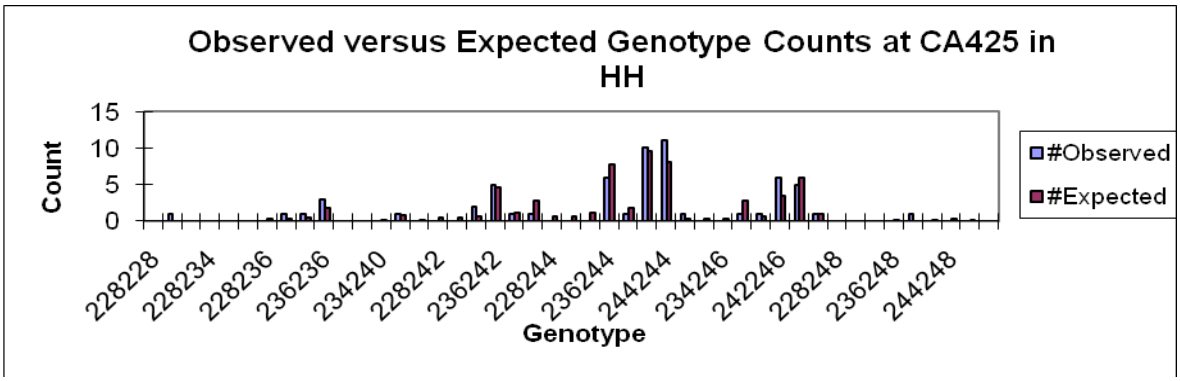

Observed versus Expected Genotype Counts at COR2 in HH
Count
Genotype
#Observed
#Expected
228228
228230
230230
228232
230232
232232
228234
230234
232234
234234
228236
230236
232236
234236
236236

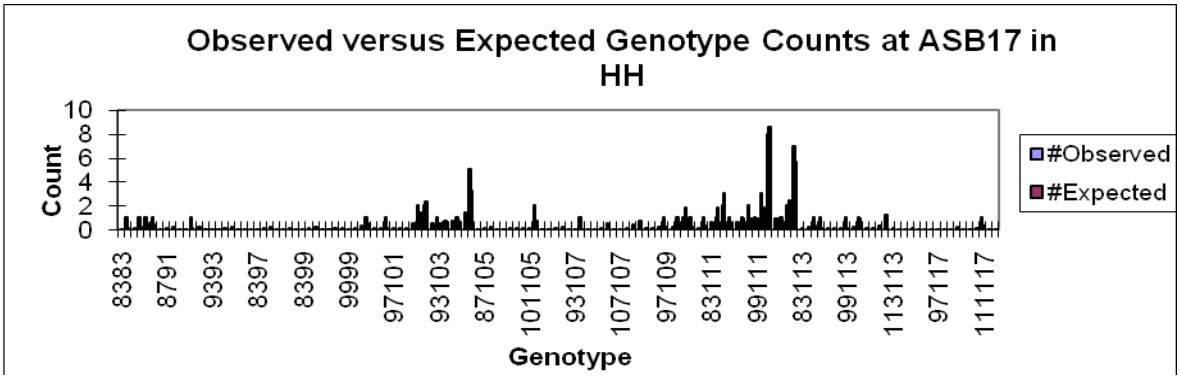

Observed versus Expected Genotype Counts at ASB43 in HH
Count
Genotype
#Observed
#Expected
8585
8589
8989
8591
8991
9191
8593
8993
9193
9393
8599
8999
9199
9399
9999
85101
89101
91101
93101
99101
101101

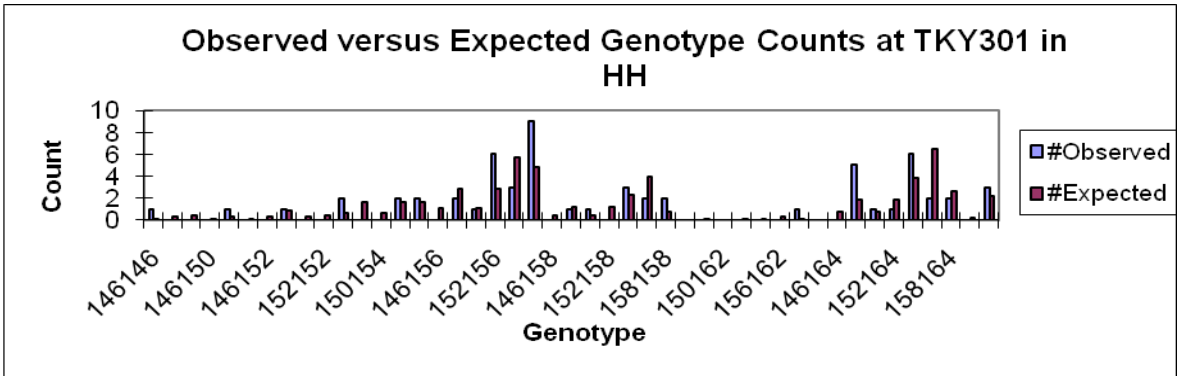

Observed versus Expected Genotype Counts at AHT33 in HH
Count
Genotype
#Observed
#Expected
145145
145151
145153
153153
151155
145157
153157
145159
153159
159159
151161
157161
145163
153163
159163
145165
153165
159165
165165
151167
157167
163167

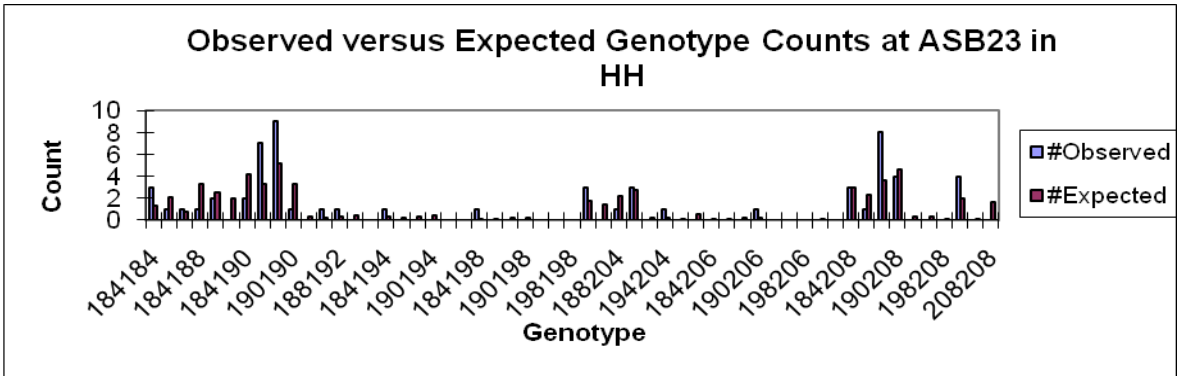

Observed versus Expected Genotype Counts at NEV79 in HH
Count
15
10
5
0
157157
157167
157169
169169
167171
157181
169181
157183
169183
183183
167185
181185
157187
169187
183187
157197
169197
183197
197197
Genotype
#Observed
#Expected

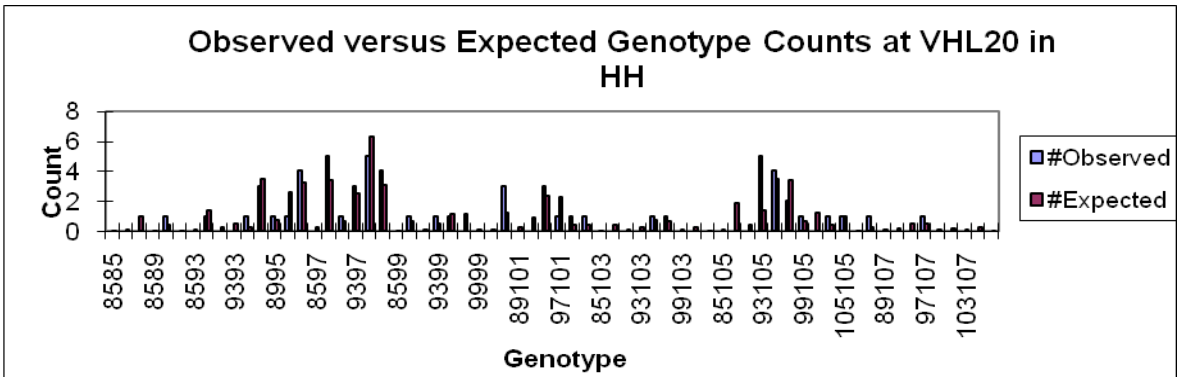

Observed versus Expected Genotype Counts at HMS2 in HH
Count
10
8
6
4
2
0
216216
218218
218220
216222
220222
216224
220224
224224
218226
222226
226226
218232
222232
226232
216236
220236
224236
232236
Genotype
#Observed
#Expected

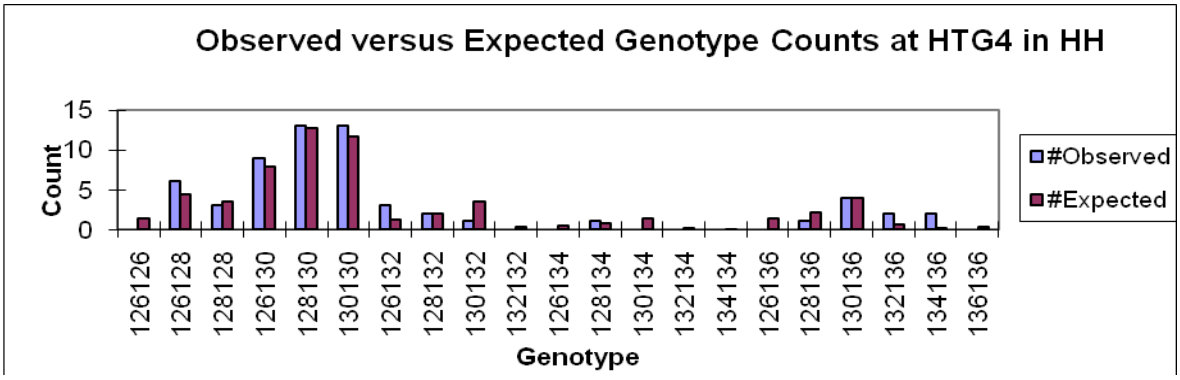

Observed versus Expected Genotype Counts at CA425 in HH
Count
15
10
5
0
228228
228234
228236
236236
234240
228242
236242
228244
236244
244244
234246
242246
228248
236248
244248
Genotype
#Observed
#Expected

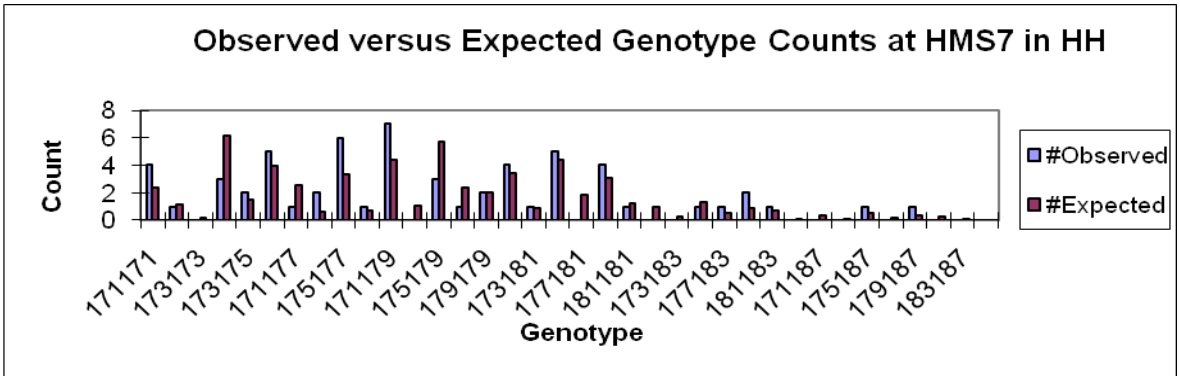

Observed versus Expected Genotype Counts at ASB17 in HH
Count
10
8
6
4
2
0
8383
8791
9393
8397
8399
9999
97101
93103
87105
101105
93107
107107
97109
83111
99111
83113
99113
113113
97117
111117
Genotype
#Observed
#Expected

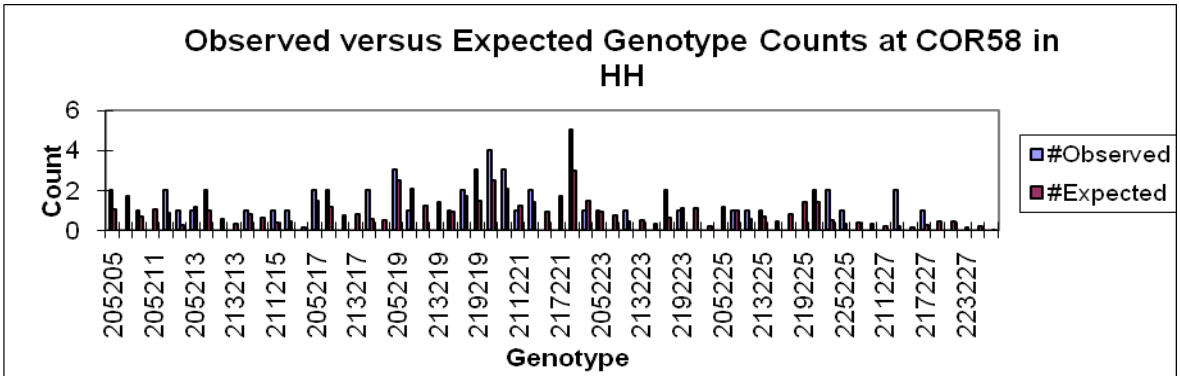

Observed versus Expected Genotype Counts at TKY301 in HH
Count
Genotype
#Observed
#Expected


Observed versus Expected Genotype Counts at ASB23 in HH
Count
Genotype
#Observed
#Expected

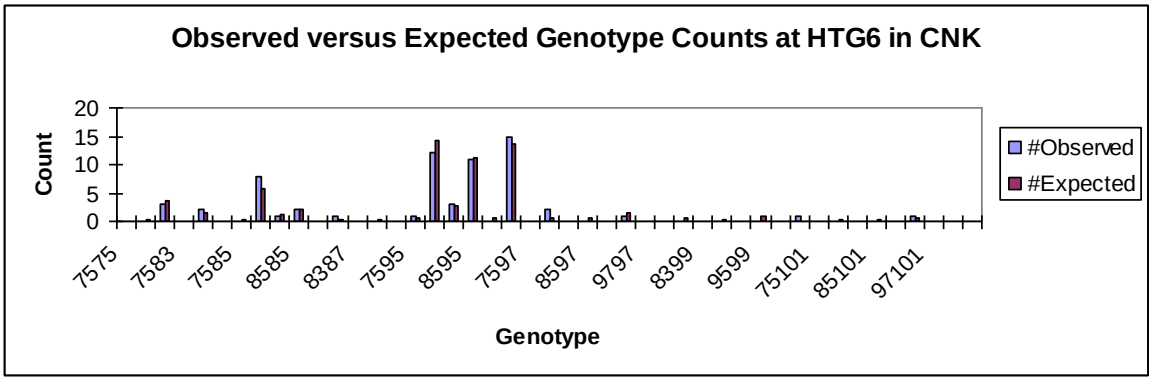

Observed versus Expected Genotype Counts at VHL20 in HH
Count
Genotype
#Observed
#Expected

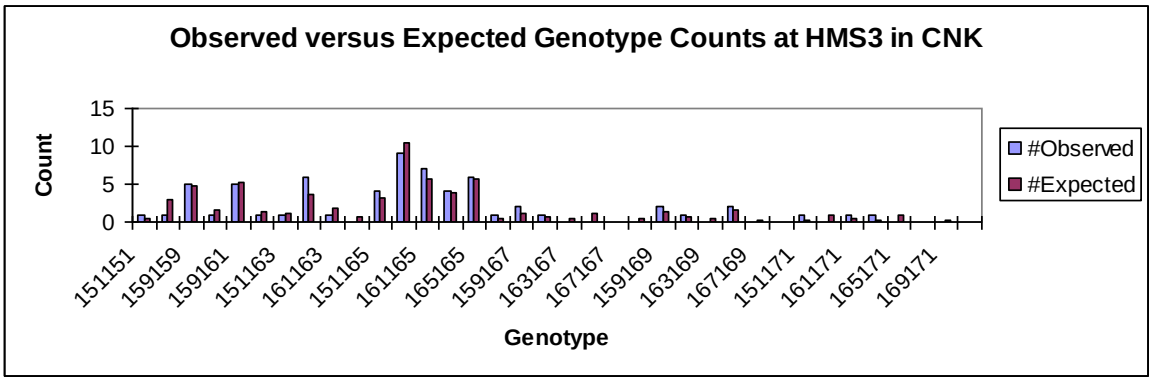

Observed versus Expected Genotype Counts at HTG4 in HH
Count
Genotype
#Observed
#Expected

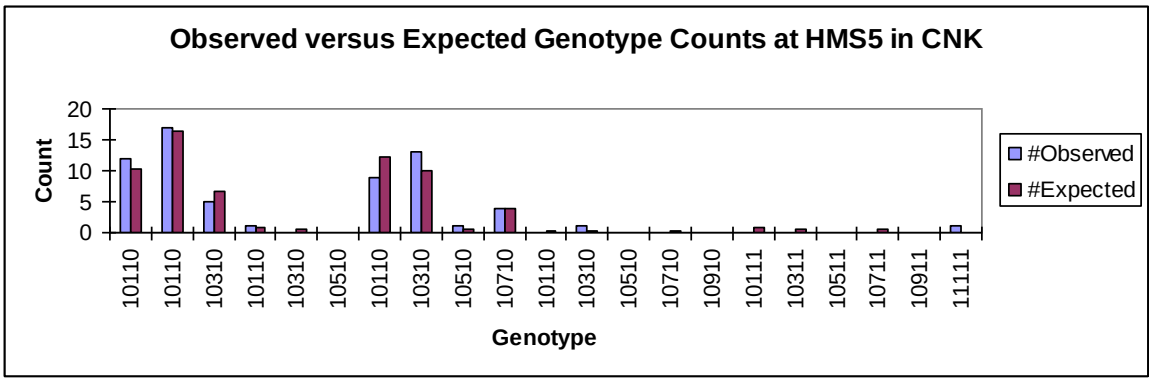

Observed versus Expected Genotype Counts at HMS7 in HH
Count
Genotype
#Observed
#Expected

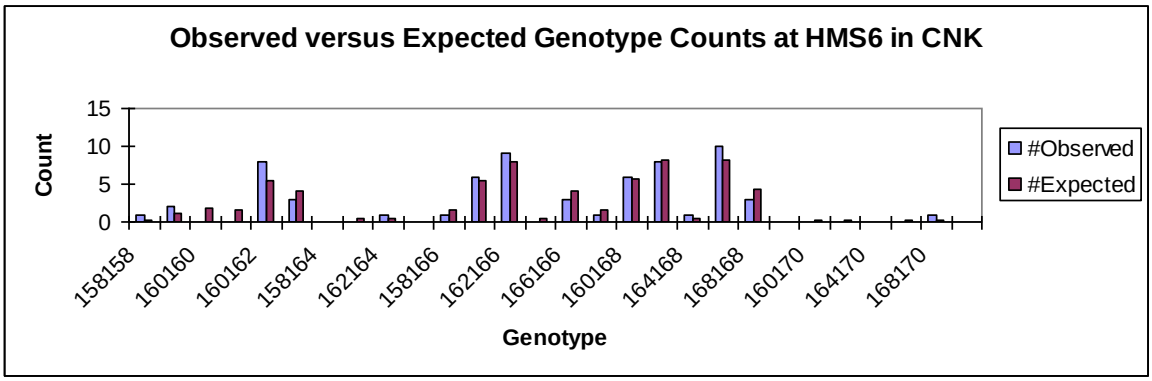

Observed versus Expected Genotype Counts at COR58 in HH
Count
Genotype
#Observed
#Expected

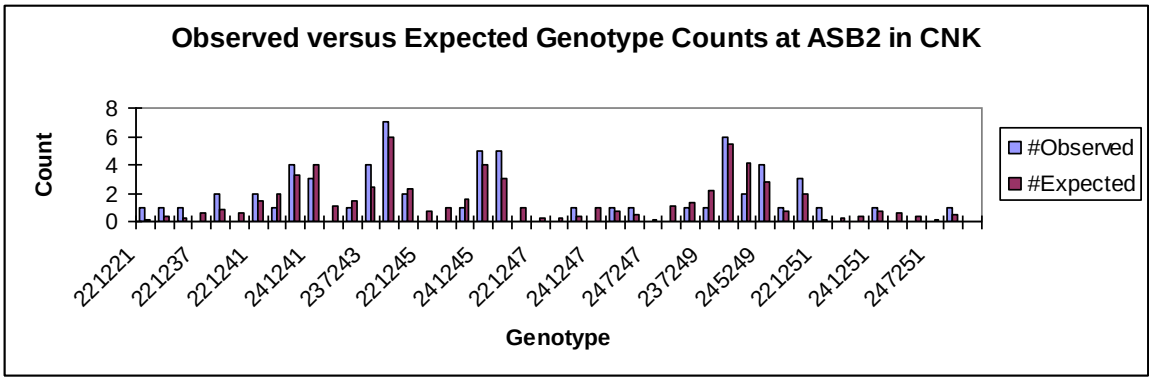
CANİK HORSE

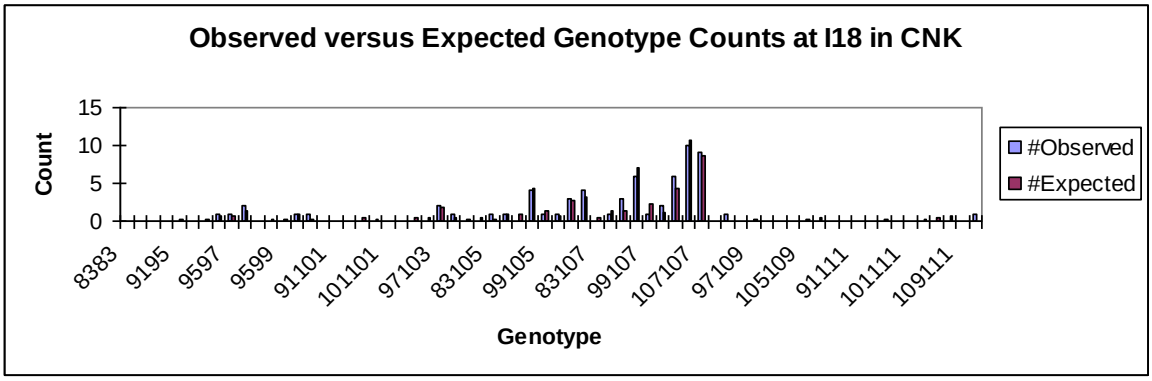

Observed versus Expected Genotype Counts at HTG6 in CNK
Count
Genotype
#Observed
#Expected

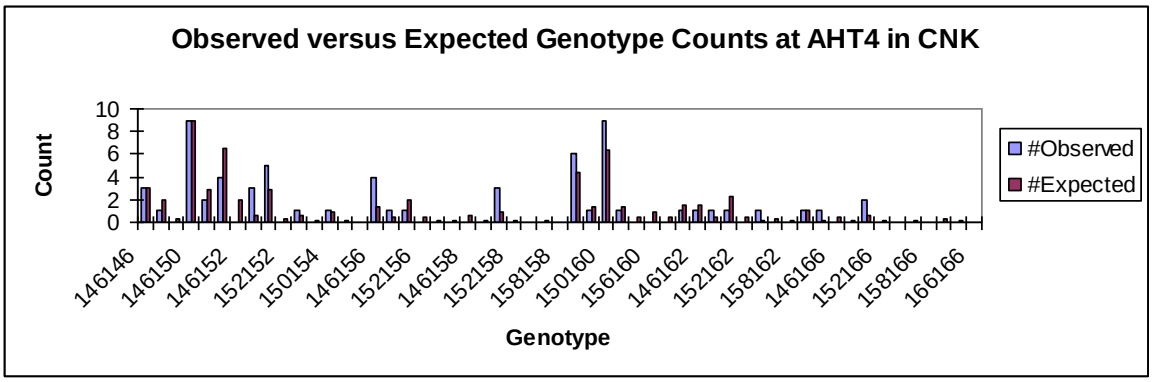

Observed versus Expected Genotype Counts at HMS3 in CNK
Count
Genotype
#Observed
#Expected

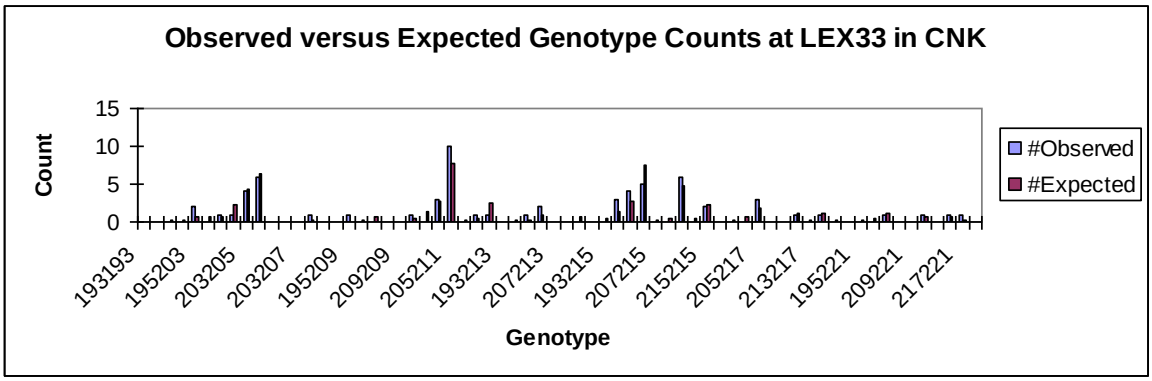

Observed versus Expected Genotype Counts at HMS5 in CNK
Count
20
15
10
5
0
10110
10110
10310
10110
10310
10510
10110
10310
10510
10710
10110
10310
10510
10710
10910
10111
10311
10511
10711
10911
11111
Genotype
#Observed
#Expected

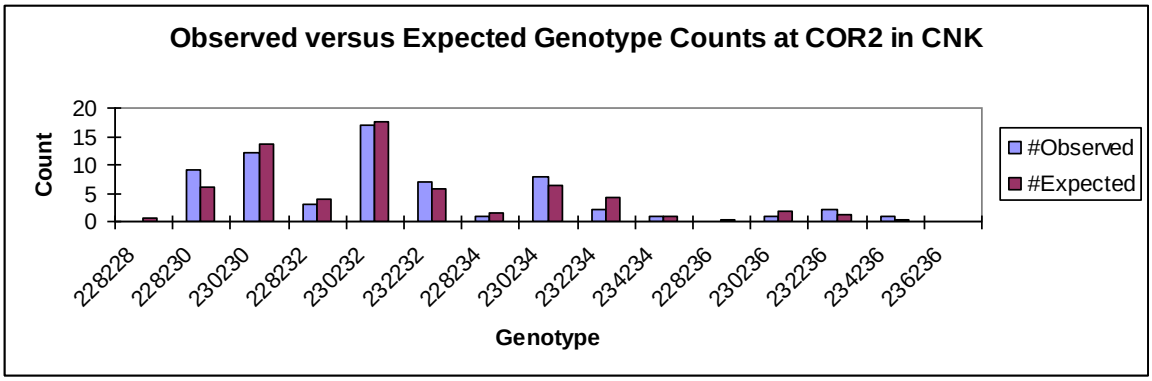

Observed versus Expected Genotype Counts at HMS6 in CNK
Count
15
10
5
0
158158
160160
160162
158164
162164
158166
162166
166166
160168
164168
168168
160170
164170
168170
Genotype
#Observed
#Expected

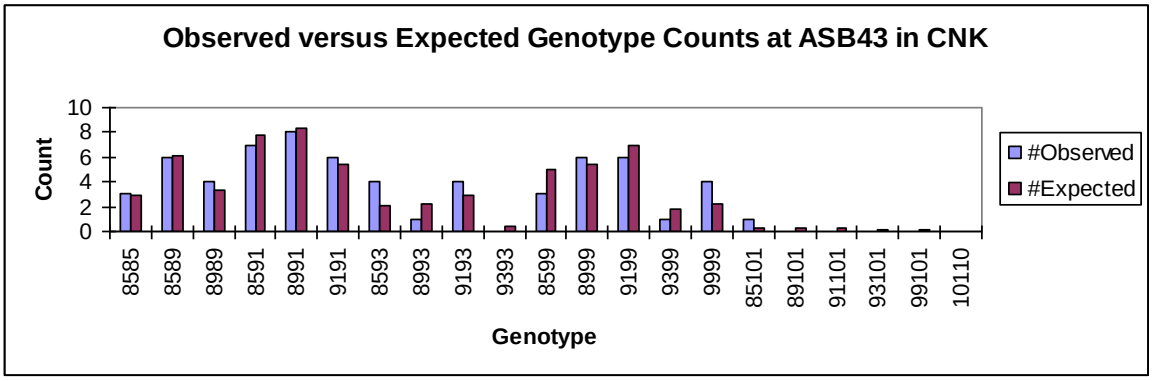

Observed versus Expected Genotype Counts at ASB2 in CNK
Count
8
6
4
2
0
221221
221237
221241
241241
237243
221245
241245
221247
241247
247247
237249
245249
221251
241251
247251
Genotype
#Observed
#Expected

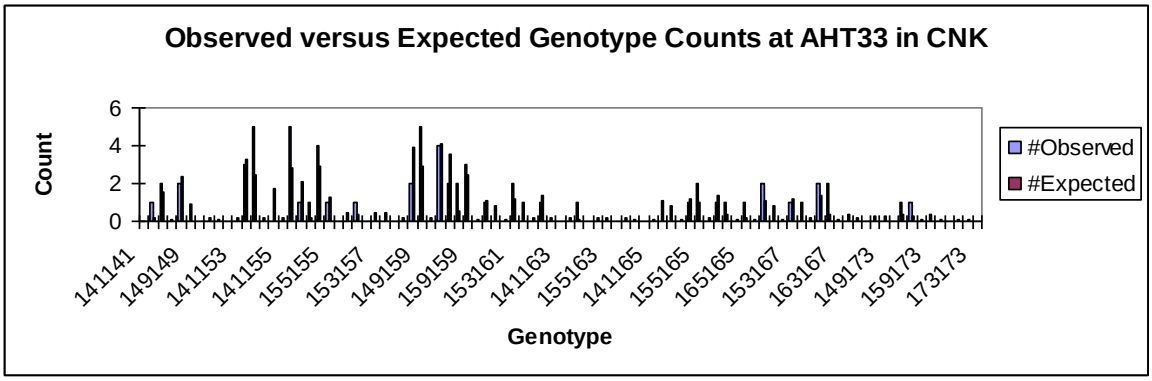

Observed versus Expected Genotype Counts at I18 in CNK
Count
15
10
5
0
8383
9195
9597
9599
91101
101101
97103
83105
99105
83107
99107
107107
97109
105109
91111
101111
109111
Genotype
#Observed
#Expected

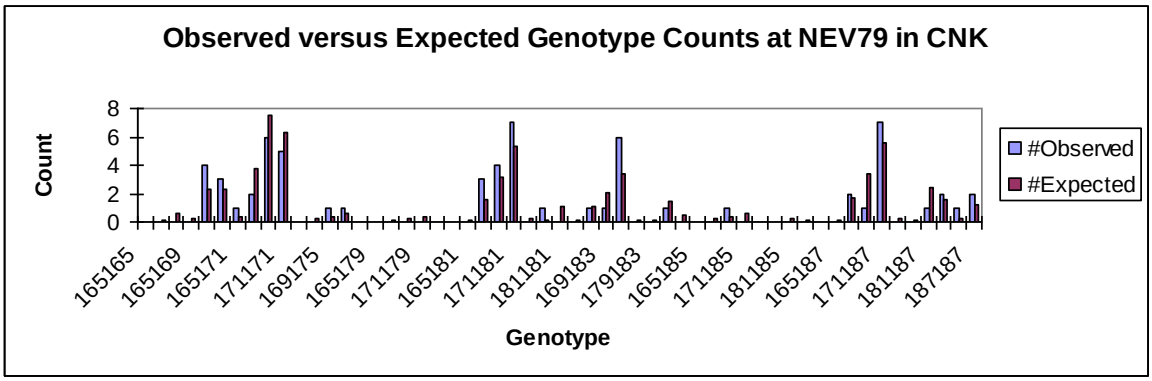

Observed versus Expected Genotype Counts at AHT4 in CNK
Count
Genotype
#Observed
#Expected
146146
146150
146152
152152
150154
146156
152156
146158
152158
158158
150160
156160
146162
152162
158162
146166
152166
158166
166166

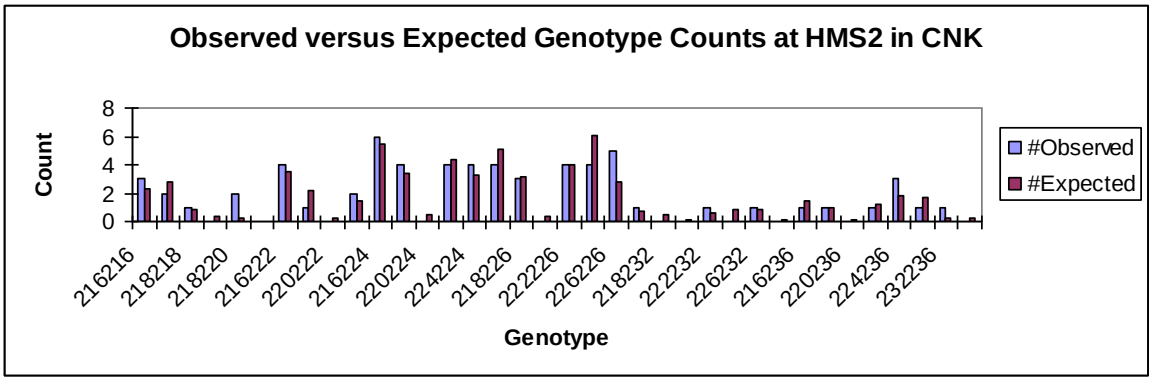

Observed versus Expected Genotype Counts at LEX33 in CNK
Count
Genotype
#Observed
#Expected
193193
195203
203205
203207
195209
209209
205211
193213
207213
193215
207215
215215
205217
213217
195221
209221
217221

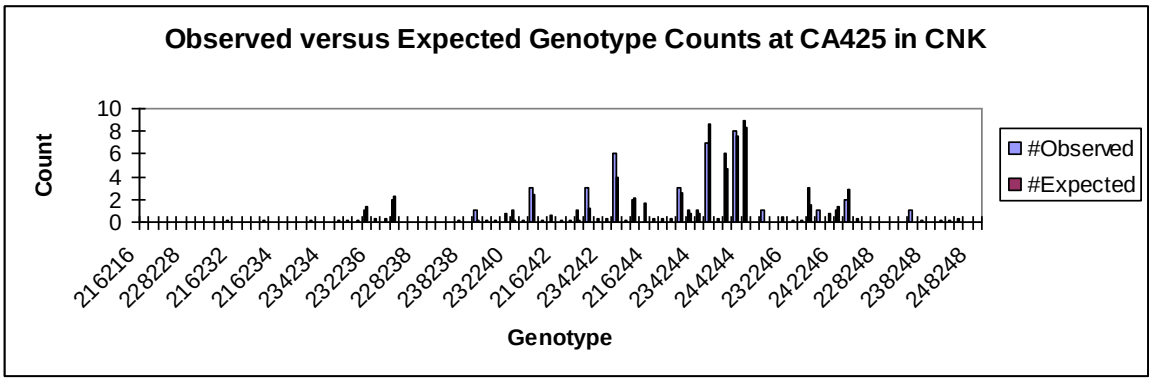

Observed versus Expected Genotype Counts at COR2 in CNK
Count
Genotype
#Observed
#Expected
228228
228230
230230
228232
230232
232232
228234
230234
232234
234234
228236
230236
232236
234236
236236

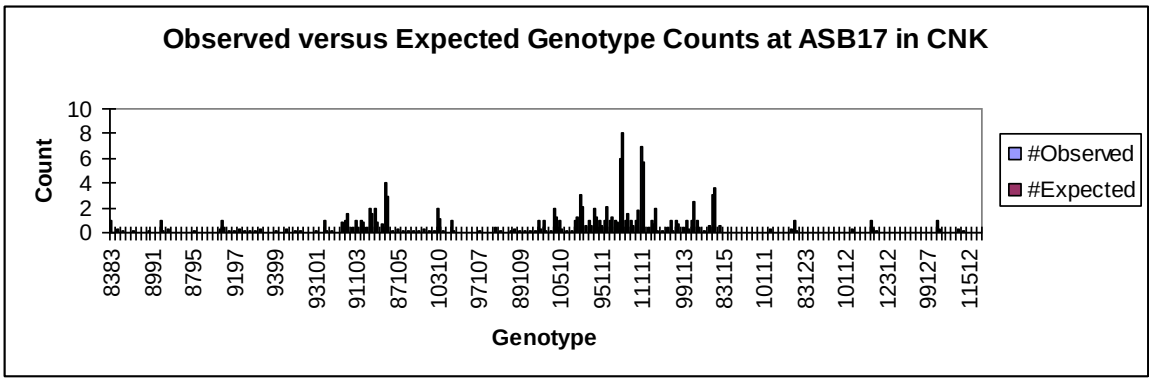

Observed versus Expected Genotype Counts at ASB43 in CNK
Count
Genotype
#Observed
#Expected
8585
8589
8989
8591
8991
9191
8593
8993
9193
9393
8599
8999
9199
9399
9999
85101
89101
91101
93101
99101
10110

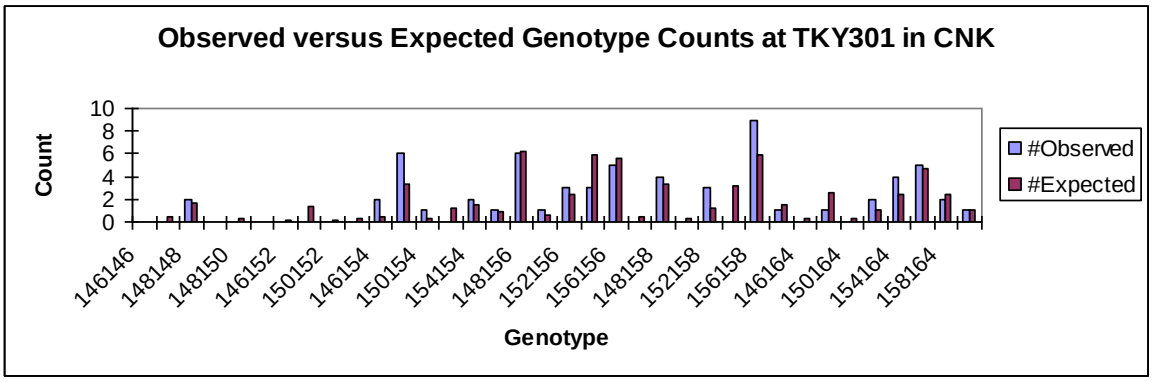

Observed versus Expected Genotype Counts at AHT33 in CNK
Count
Genotype
#Observed
#Expected
141141
149149
141153
141155
155155
153157
149159
159159
153161
141163
155163
141165
155165
165165
153167
163167
149173
159173
173173

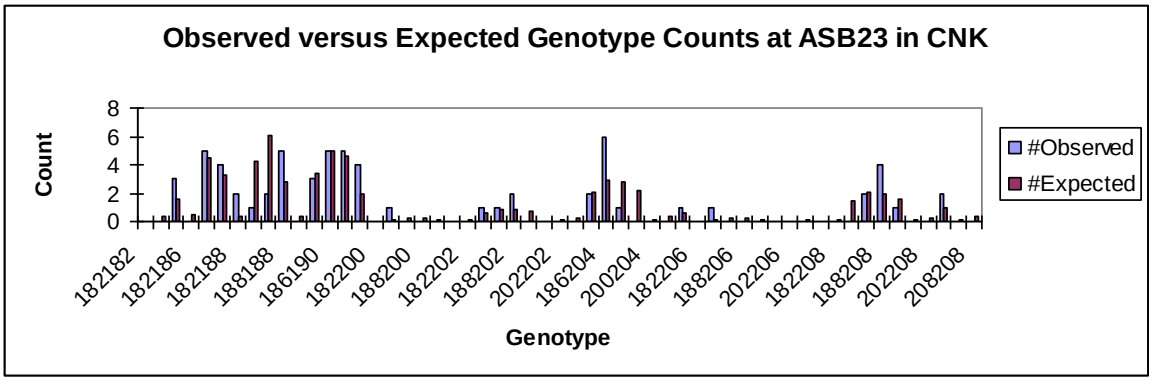

Observed versus Expected Genotype Counts at NEV79 in CNK
Count
Genotype
#Observed
#Expected
165165
165169
165171
171171
169175
165179
171179
165181
171181
181181
169183
179183
165185
171185
181185
165187
171187
181187
187187

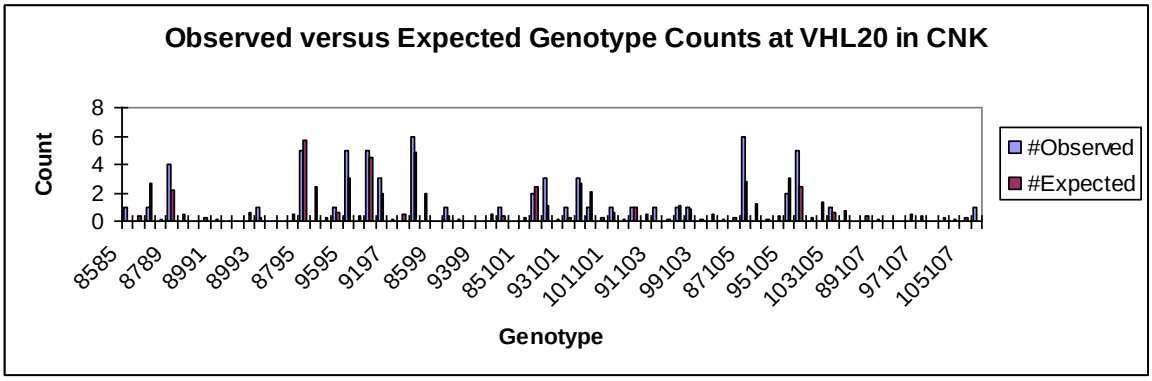

Observed versus Expected Genotype Counts at HMS2 in CNK
Count
Genotype
#Observed
#Expected
216216
218218
218220
216222
220222
216224
220224
224224
218226
222226
226226
218232
222232
226232
216236
220236
224236
232236

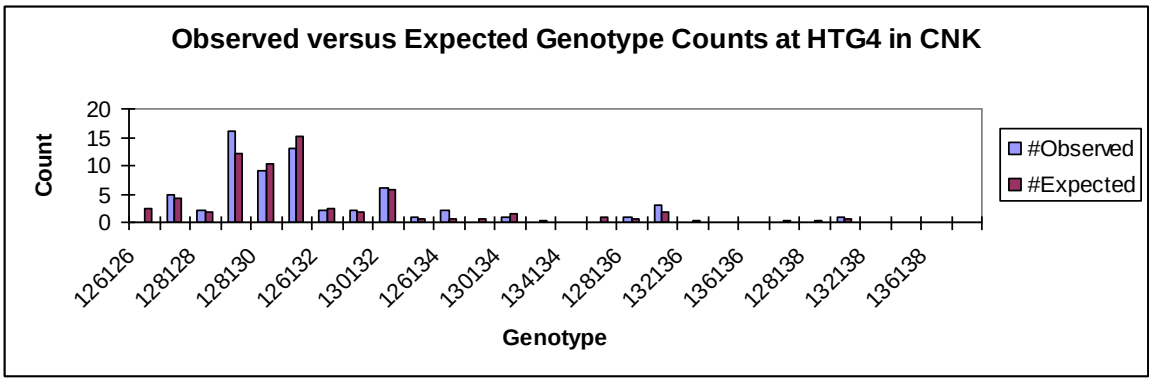

Observed versus Expected Genotype Counts at CA425 in CNK
Count
Genotype
#Observed
#Expected
216216
228228
216232
216234
234234
232236
228238
238238
232240
216242
234242
216244
234244
244244
232246
242246
228248
238248
248248

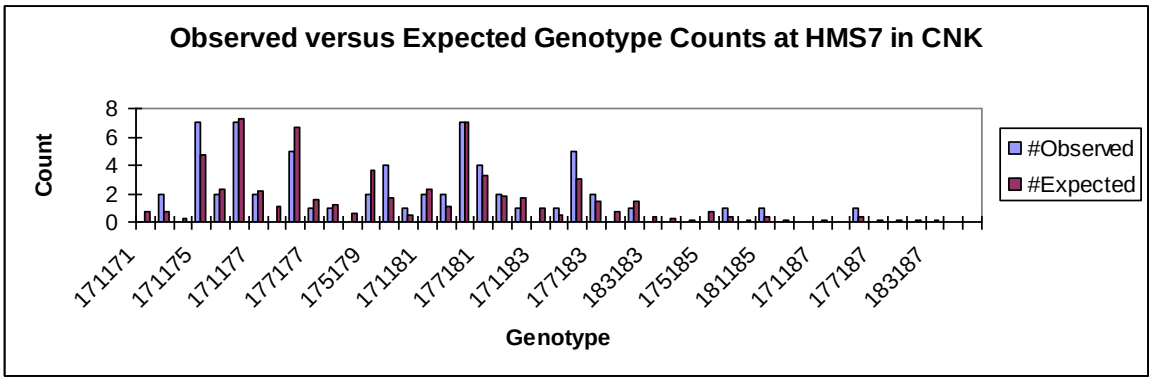

Observed versus Expected Genotype Counts at ASB17 in CNK
Count
Genotype
#Observed
#Expected

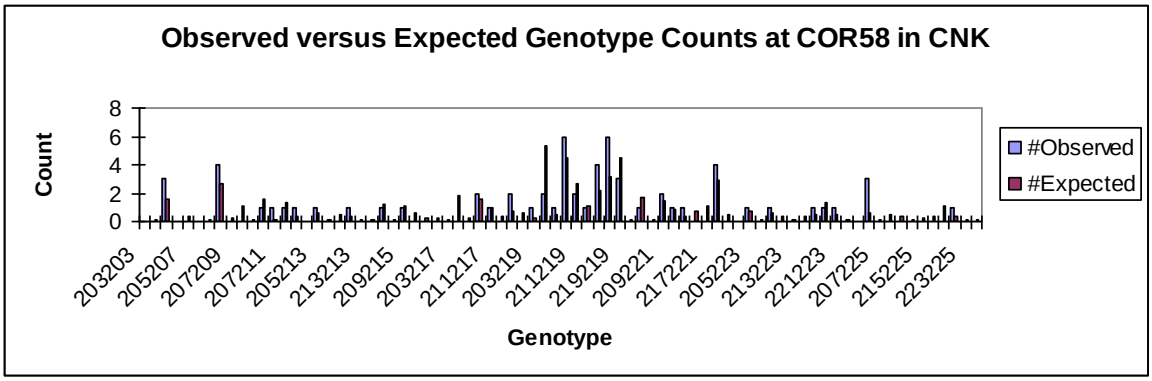

Observed versus Expected Genotype Counts at TKY301 in CNK
Count
Genotype
#Observed
#Expected


Observed versus Expected Genotype Counts at ASB23 in CNK
Count
Genotype
#Observed
#Expected

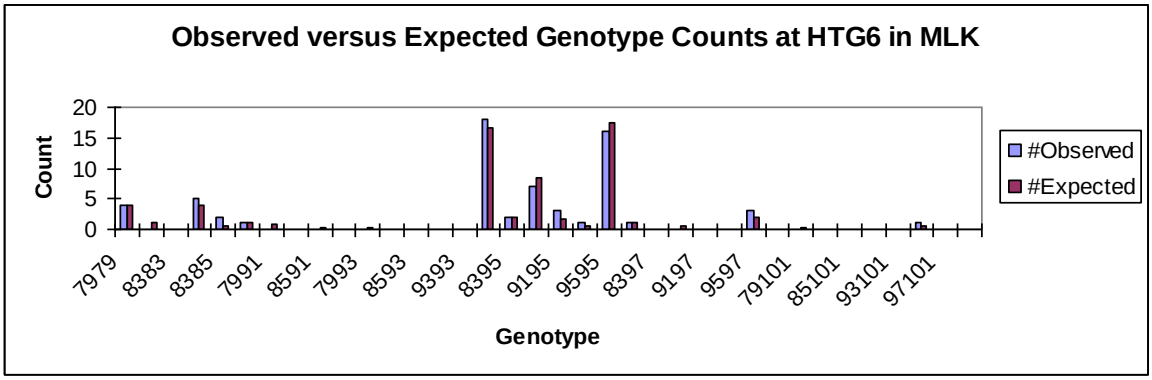

Observed versus Expected Genotype Counts at VHL20 in CNK
Count
Genotype
#Observed
#Expected

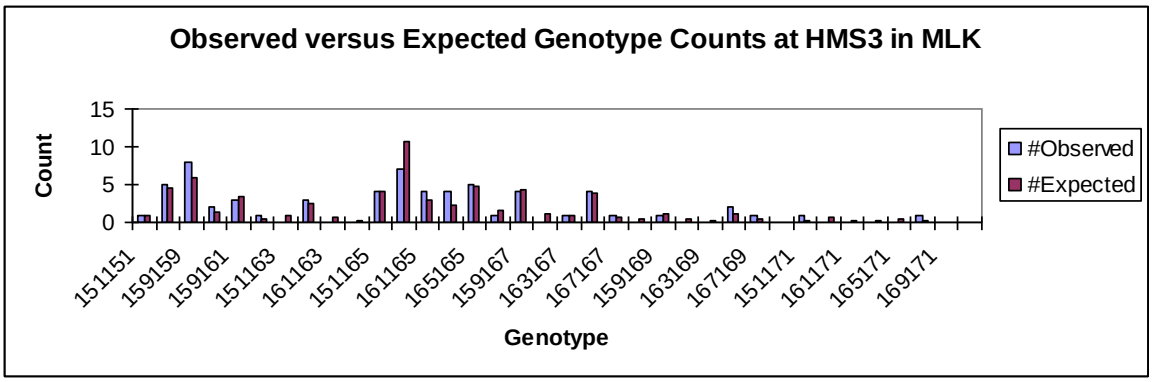

Observed versus Expected Genotype Counts at HTG4 in CNK
Count
20
15
10
5
0
126126
128128
128130
126132
130132
126134
130134
134134
128136
132136
136136
128138
132138
136138
Genotype
#Observed
#Expected

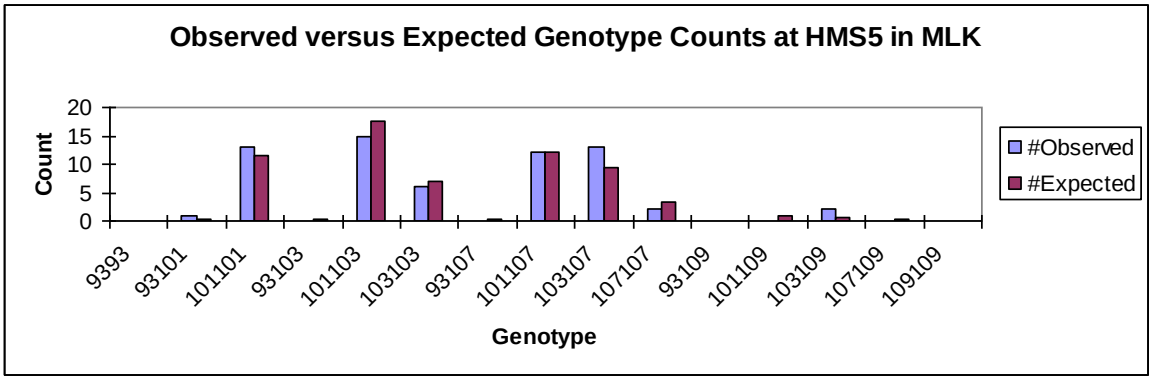

Observed versus Expected Genotype Counts at HMS7 in CNK
Count
8
6
4
2
0
171171
171175
171177
177177
175179
171181
177181
171183
177183
183183
175185
181185
171187
177187
183187
Genotype
#Observed
#Expected

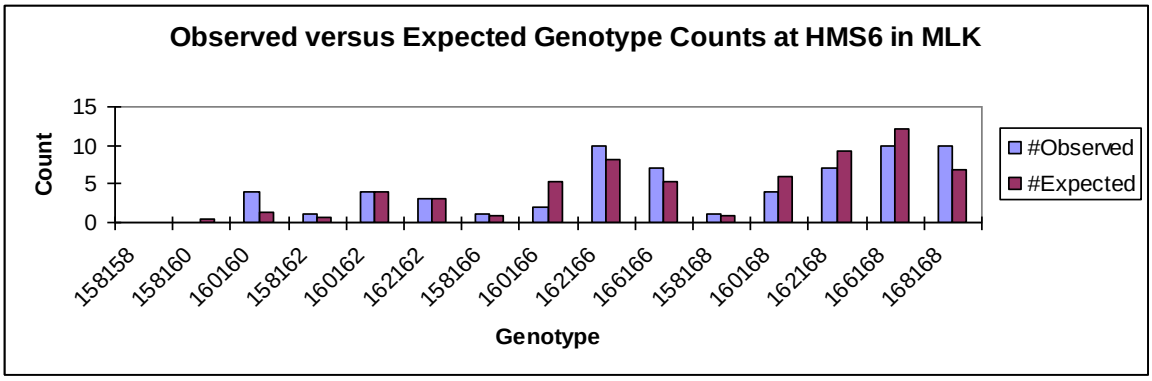

Observed versus Expected Genotype Counts at COR58 in CNK
Count
8
6
4
2
0
203203
205207
207209
207211
205213
213213
209215
203217
211217
203219
211219
219219
209221
217221
205223
213223
221223
207225
215225
223225
Genotype
#Observed
#Expected

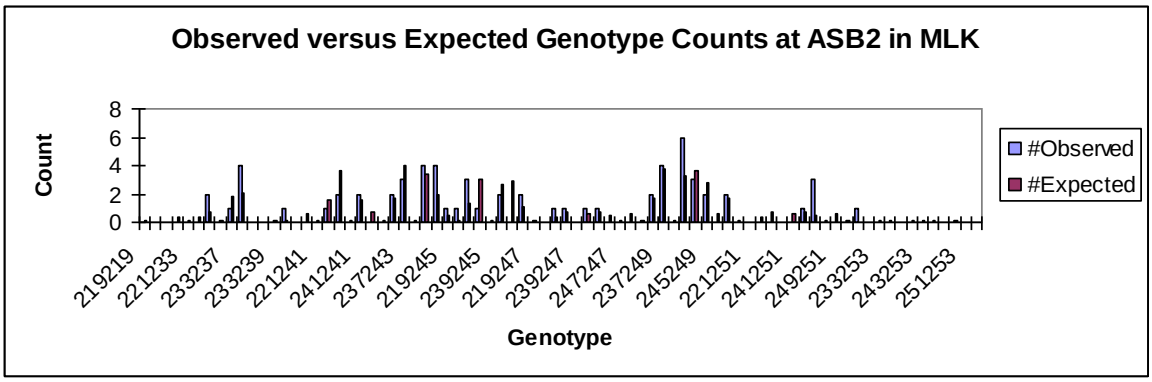
MALAKAN HORSE

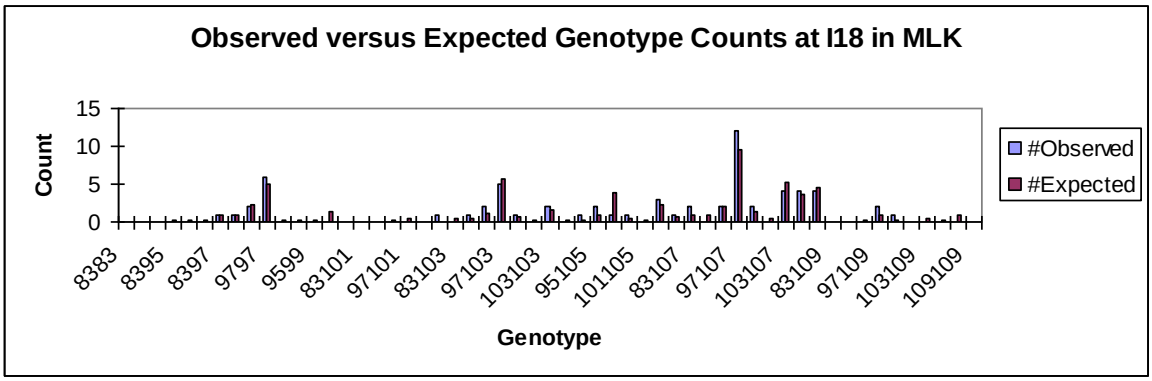


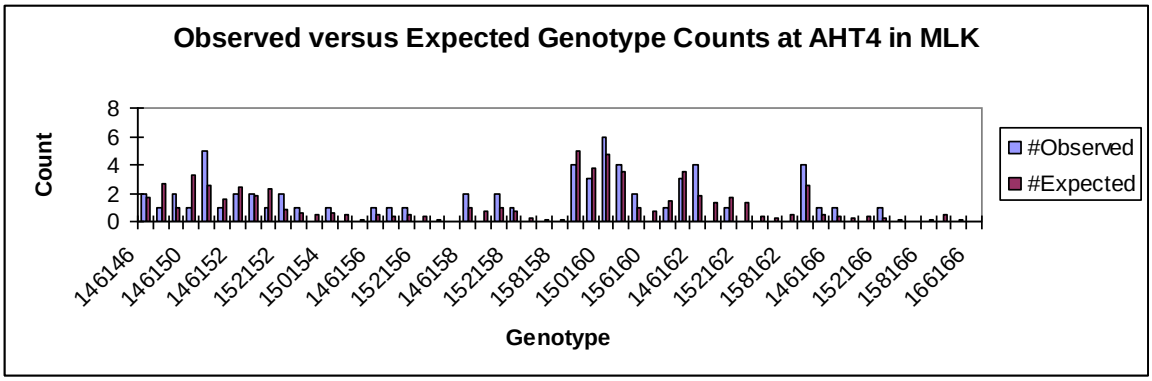


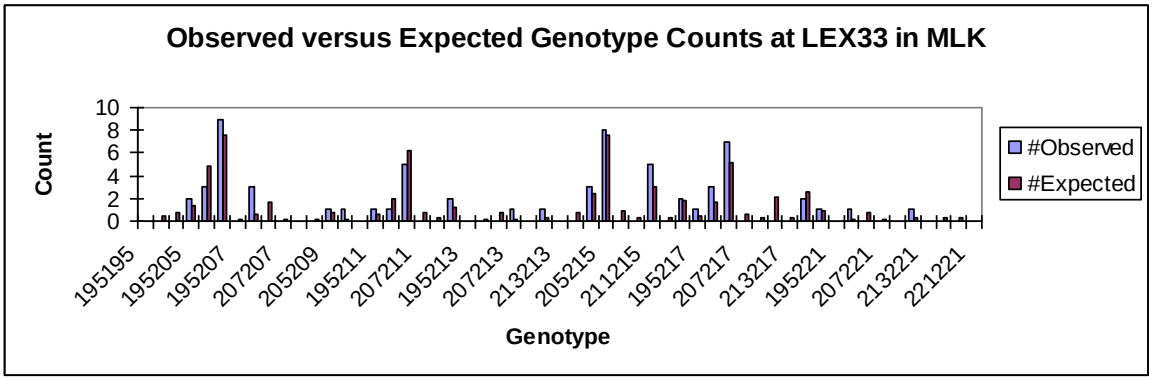


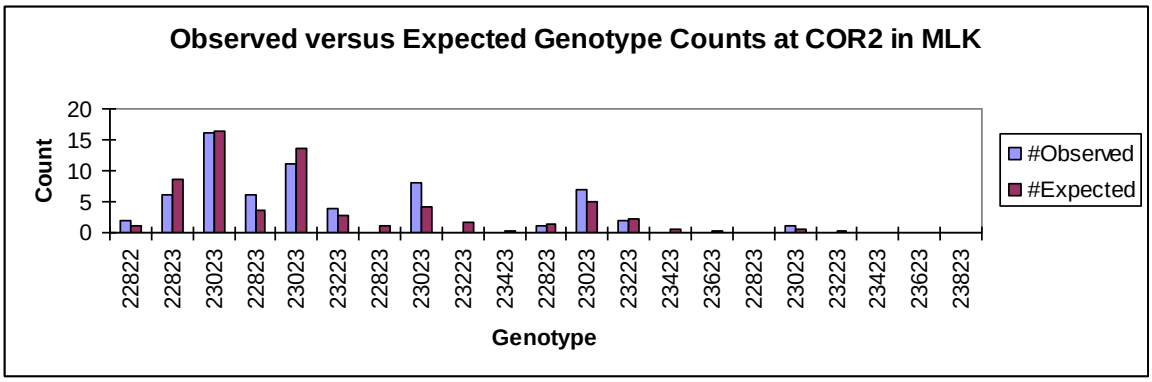


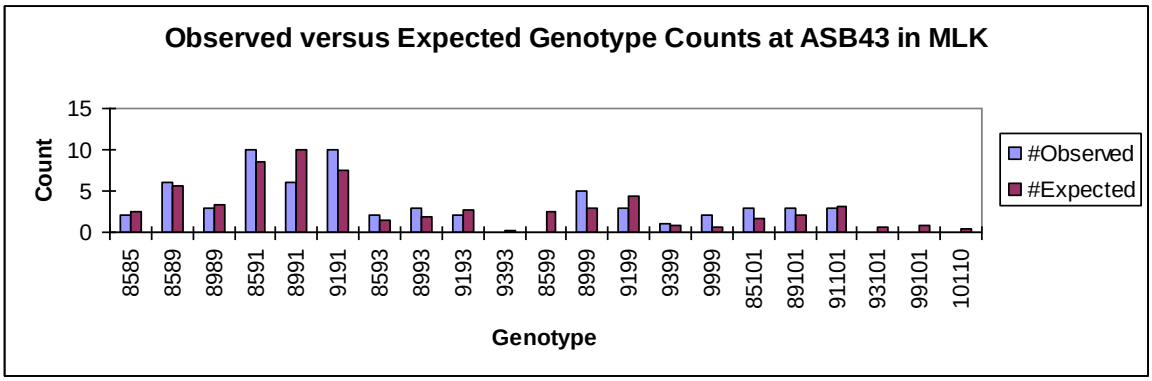

Observed versus Expected Genotype Counts at ASB2 in MLK
Count
Genotype
#Observed
#Expected
0
2
4
6
8
219219
221233
233237
233239
221241
241241
237243
219245
239245
219247
239247
247247
237249
245249
221251
241251
249251
233253
243253
251253

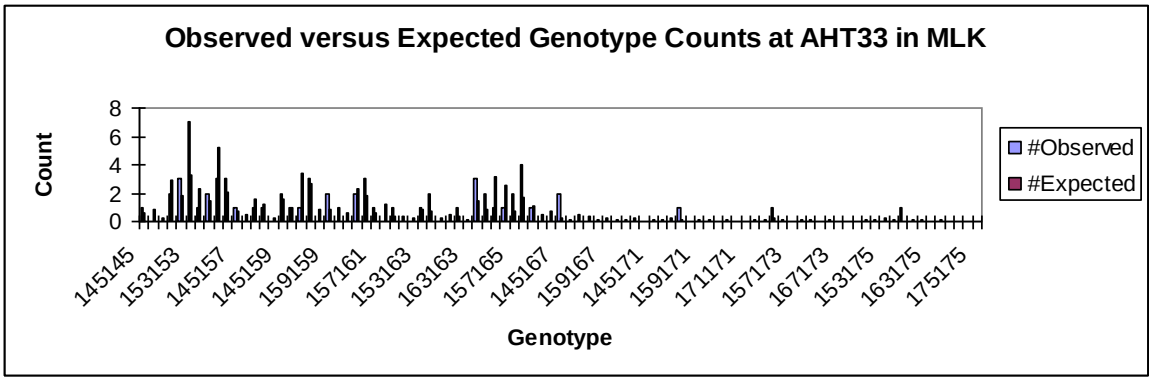

Observed versus Expected Genotype Counts at I18 in MLK
Count
Genotype
#Observed
#Expected
0
5
10
15
8383
8395
8397
9797
9599
83101
97101
83103
97103
103103
95105
101105
83107
97107
103107
83109
97109
103109
109109

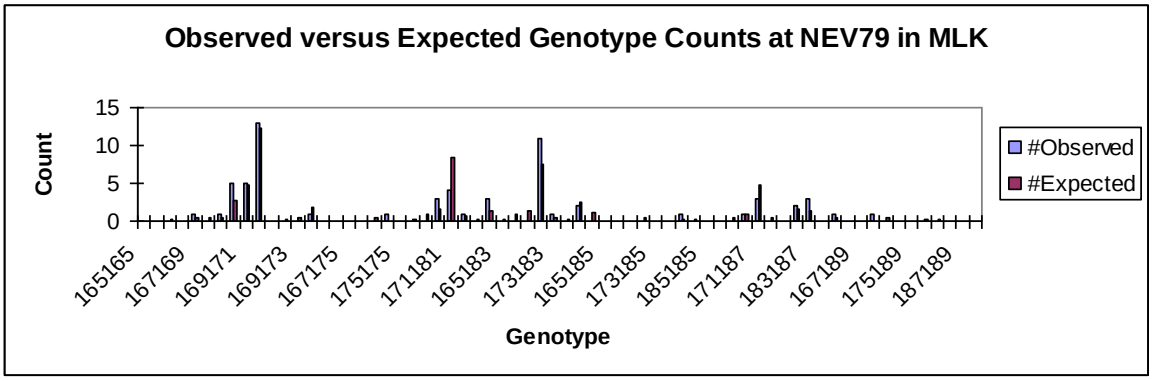

Observed versus Expected Genotype Counts at AHT4 in MLK
Count
Genotype
#Observed
#Expected
0
2
4
6
8
146146
146150
146152
152152
150154
146156
152156
146158
152158
158158
150160
156160
146162
152162
158162
146166
152166
158166
166166

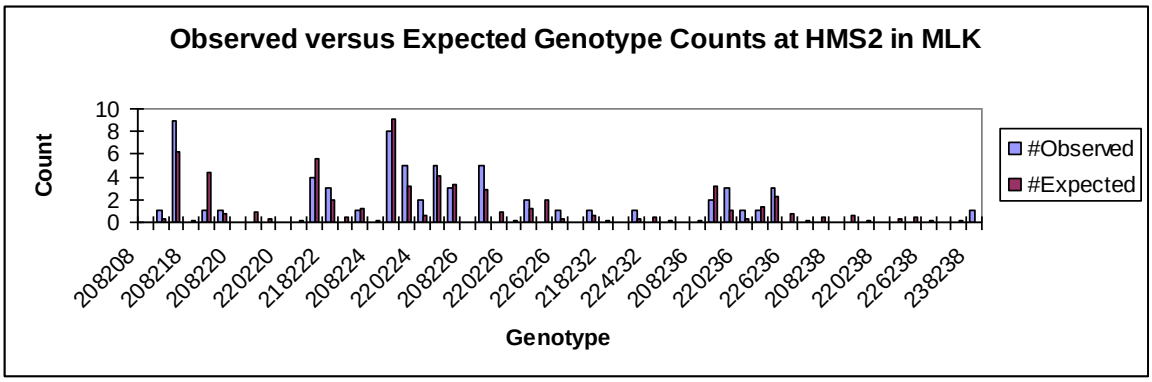

Observed versus Expected Genotype Counts at LEX33 in MLK
Count
Genotype
#Observed
#Expected
0
2
4
6
8
10
195195
195205
195207
207207
205209
195211
207211
195213
207213
213213
205215
211215
195217
207217
213217
195221
207221
213221
221221

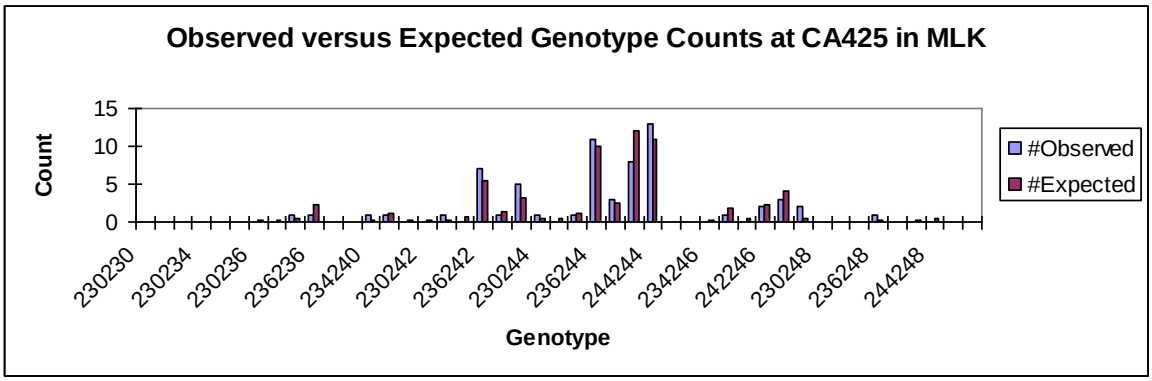

Observed versus Expected Genotype Counts at COR2 in MLK
Count
20
15
10
5
0
22822
22823
23023
22823
23023
23223
22823
23023
23223
23423
22823
23023
23223
23423
23623
22823
23023
23223
23423
23623
23823
Genotype
#Observed
#Expected

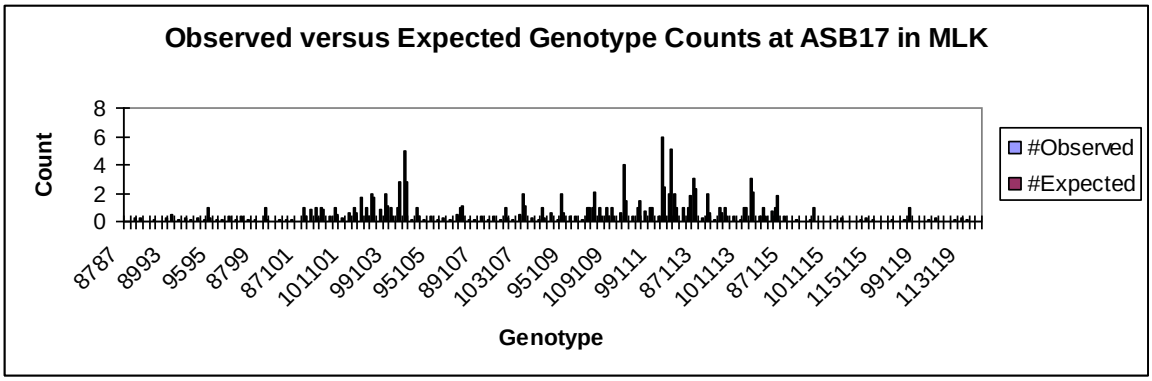

Observed versus Expected Genotype Counts at ASB43 in MLK
Count
15
10
5
0
8585
8589
8989
8591
8991
9191
8593
8993
9193
9393
8599
8999
9199
9399
9999
85101
89101
91101
93101
99101
10110
Genotype
#Observed
#Expected

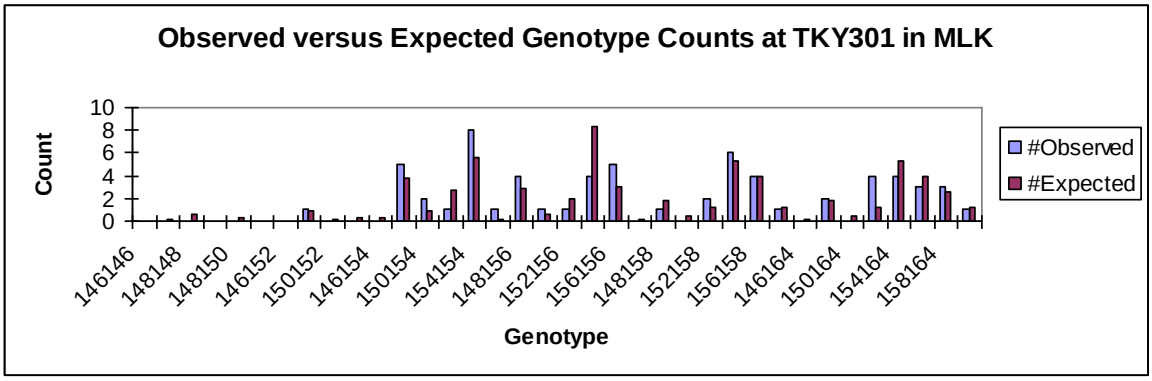

Observed versus Expected Genotype Counts at AHT33 in MLK
Count
8
6
4
2
0
145145
153153
145157
145159
159159
157161
153163
163163
157165
145167
159167
145171
159171
171171
157173
167173
153175
163175
175175
Genotype
#Observed
#Expected

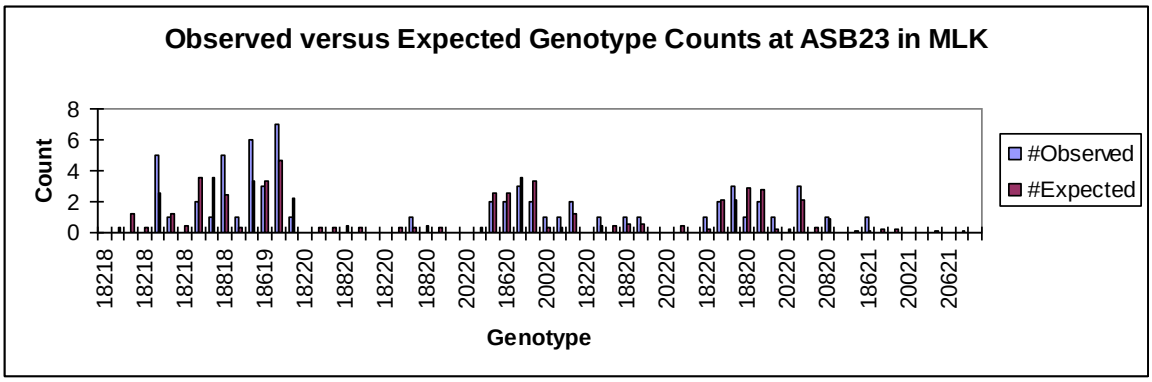

Observed versus Expected Genotype Counts at NEV79 in MLK
Count
15
10
5
0
165165
167169
169171
169173
167175
175175
171181
165183
173183
165185
173185
185185
171187
183187
167189
175189
187189
Genotype
#Observed
#Expected

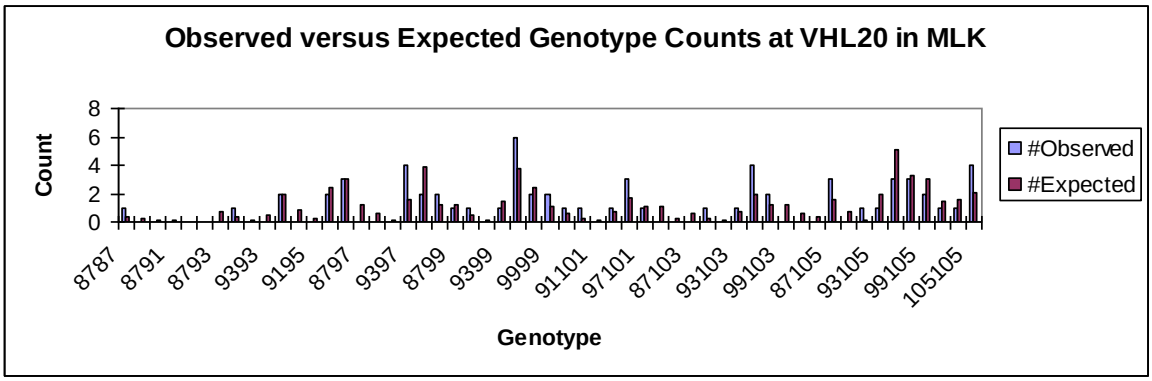

Observed versus Expected Genotype Counts at HMS2 in MLK
Count
Genotype
#Observed
#Expected
208208
208218
208220
220220
218222
208224
220224
208226
220226
226226
218232
224232
208236
220236
226236
208238
220238
226238
238238

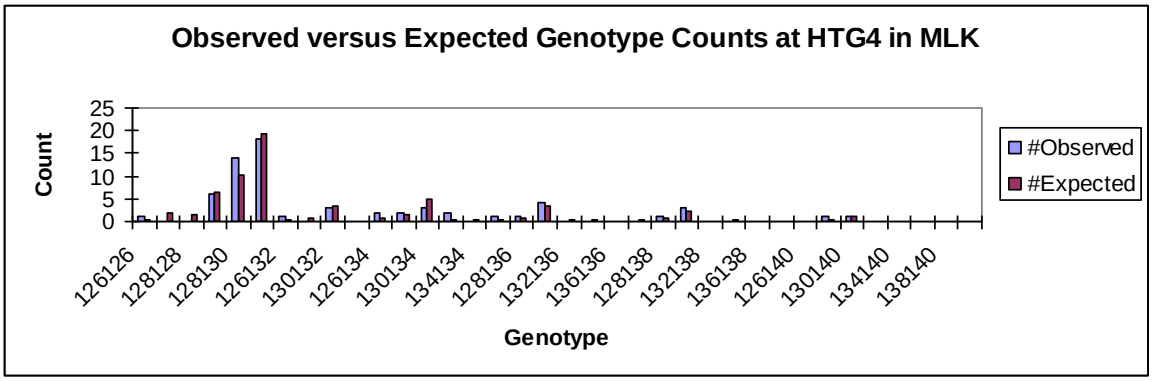

Observed versus Expected Genotype Counts at CA425 in MLK
Count
Genotype
#Observed
#Expected
230230
230234
230236
236236
234240
230242
236242
230244
236244
244244
234246
242246
230248
236248
244248

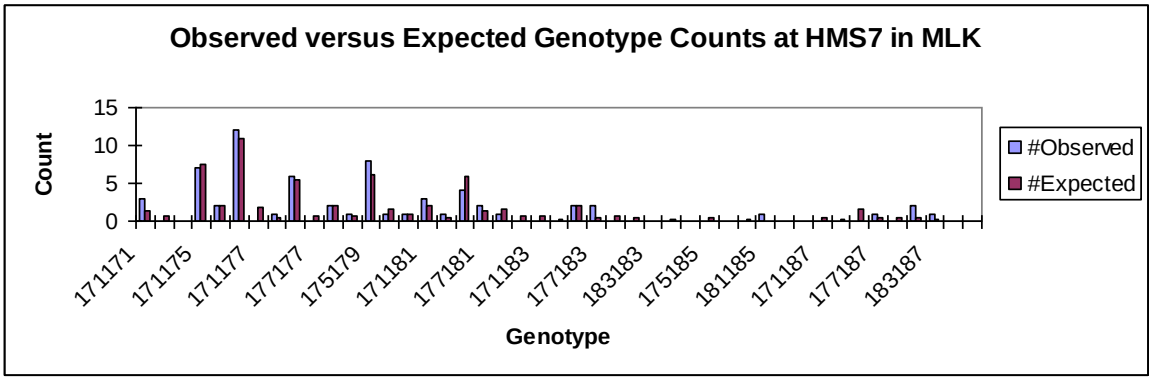

Observed versus Expected Genotype Counts at ASB17 in MLK
Count
Genotype
#Observed
#Expected
8787
8993
9595
8799
87101
101101
99103
95105
89107
103107
95109
109109
99111
87113
101113
87115
101115
115115
99119
113119

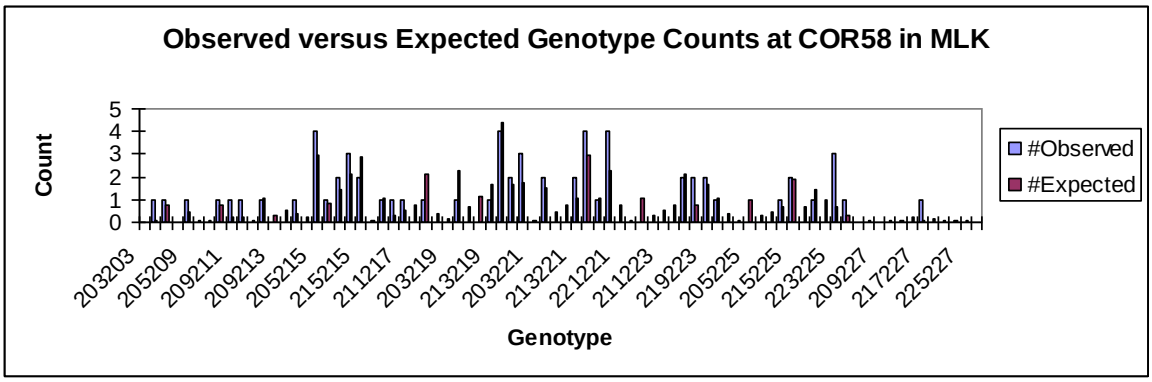

Observed versus Expected Genotype Counts at TKY301 in MLK
Count
Genotype
#Observed
#Expected
146146
148148
148150
146152
150152
146154
150154
154154
148156
152156
156156
148158
152158
156158
146164
150164
154164
158164


Observed versus Expected Genotype Counts at ASB23 in MLK
Count
Genotype
#Observed
#Expected
18218
18218
18218
18818
18619
18220
18820
18220
18820
20220
18620
20020
18220
18820
20220
18220
18820
20220
20820
18621
20021
20621

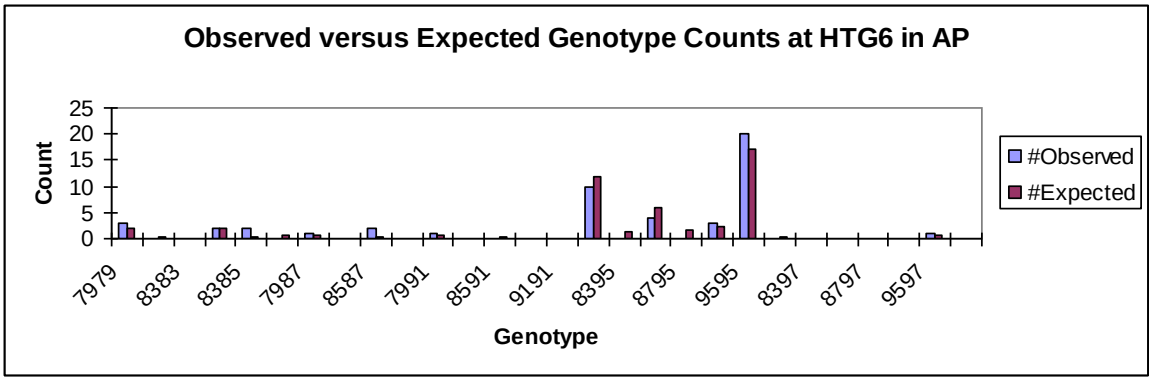

Observed versus Expected Genotype Counts at VHL20 in MLK
Count
Genotype
#Observed
#Expected
8787
8791
8793
9393
9195
8797
9397
8799
9399
9999
91101
97101
87103
93103
99103
87105
93105
99105
105105

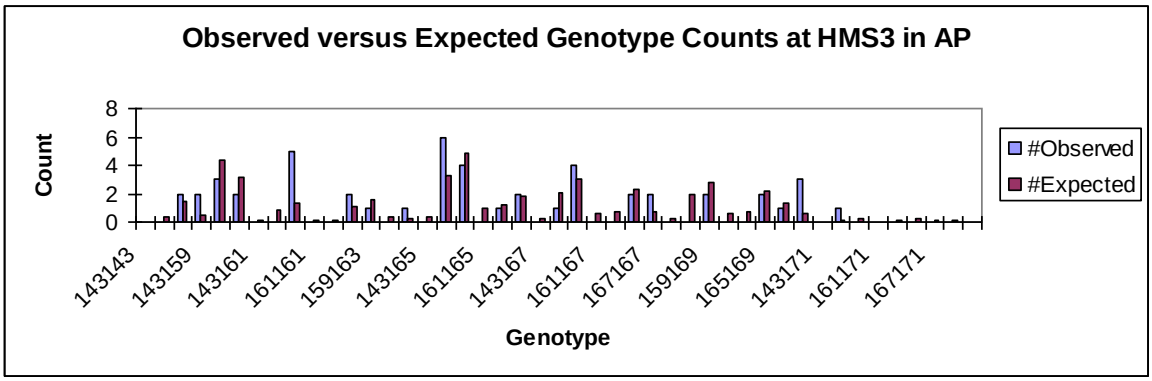

Observed versus Expected Genotype Counts at HTG4 in MLK
Count
Genotype
#Observed
#Expected
126126
128128
128130
126132
130132
126134
130134
134134
128136
132136
136136
128138
132138
136138
126140
130140
134140
138140

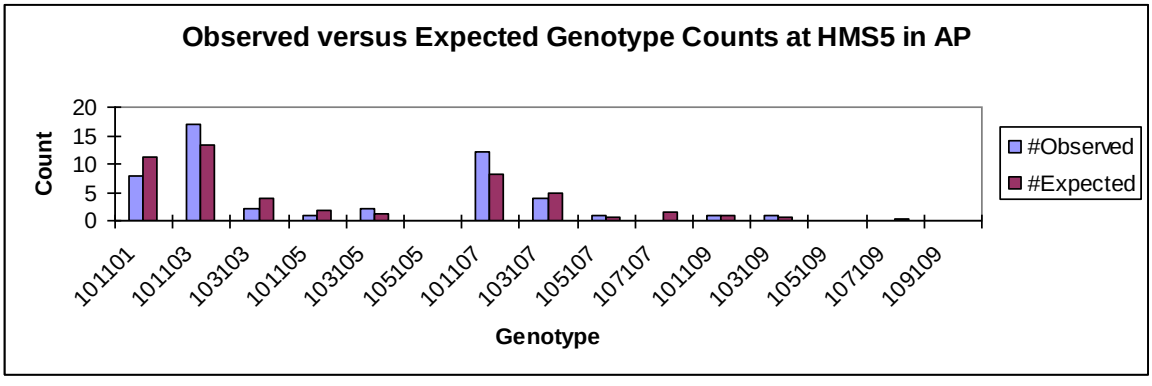

Observed versus Expected Genotype Counts at HMS7 in MLK
Count
Genotype
#Observed
#Expected
171171
171175
171177
177177
175179
171181
177181
171183
177183
183183
175185
181185
171187
177187
183187

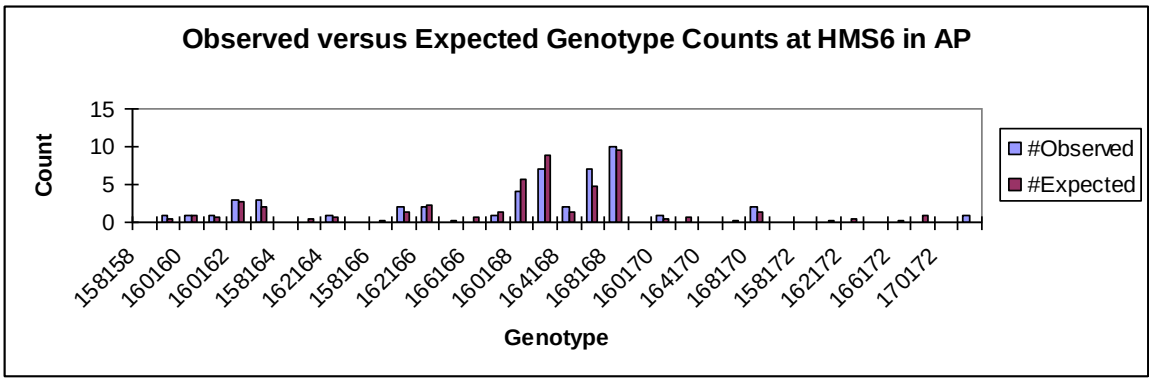


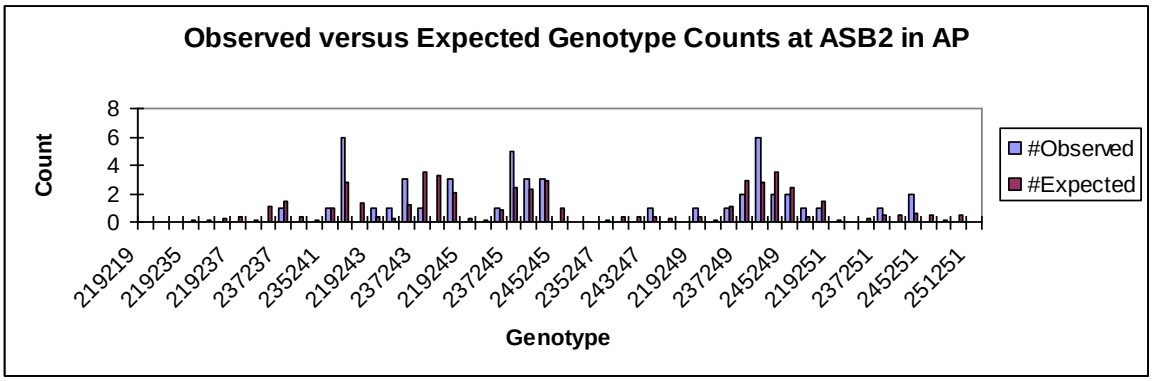
AYVACIK PONY

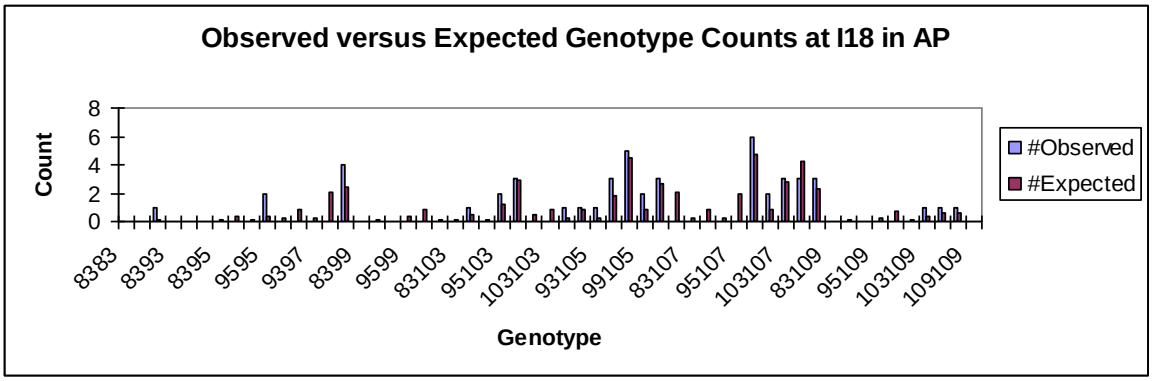


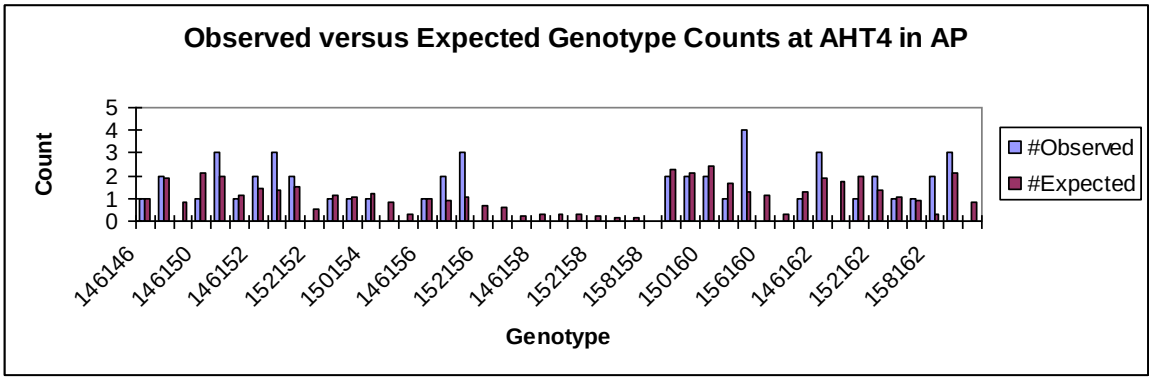


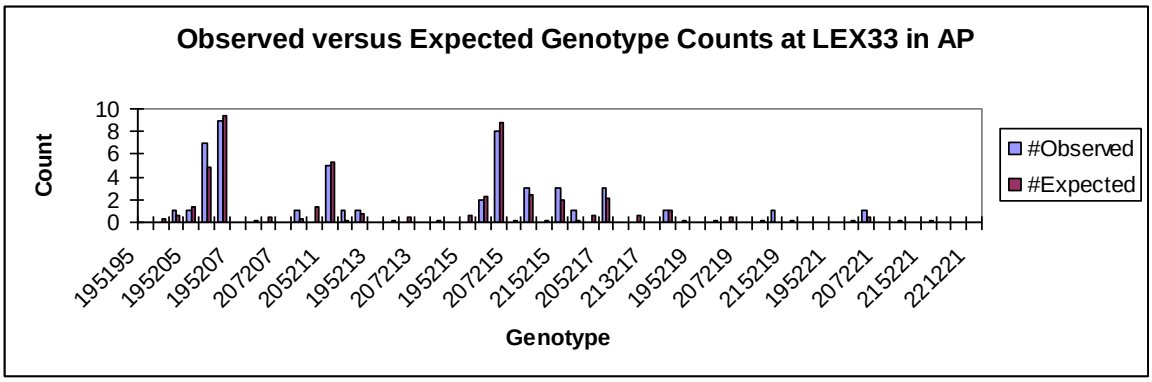


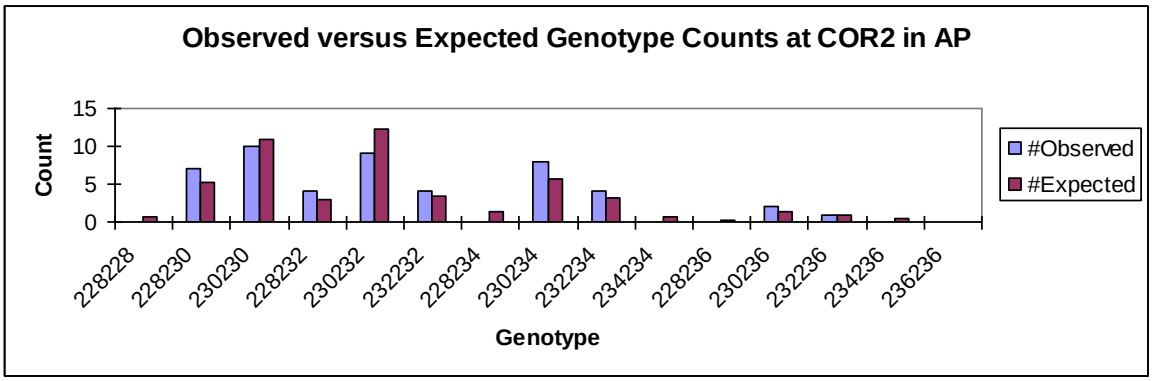

Observed versus Expected Genotype Counts at HMS6 in AP
Count
15
10
5
0
158158
160160
160162
158164
162164
158166
162166
166166
160168
164168
168168
160170
164170
168170
158172
162172
166172
170172
Genotype
#Observed
#Expected

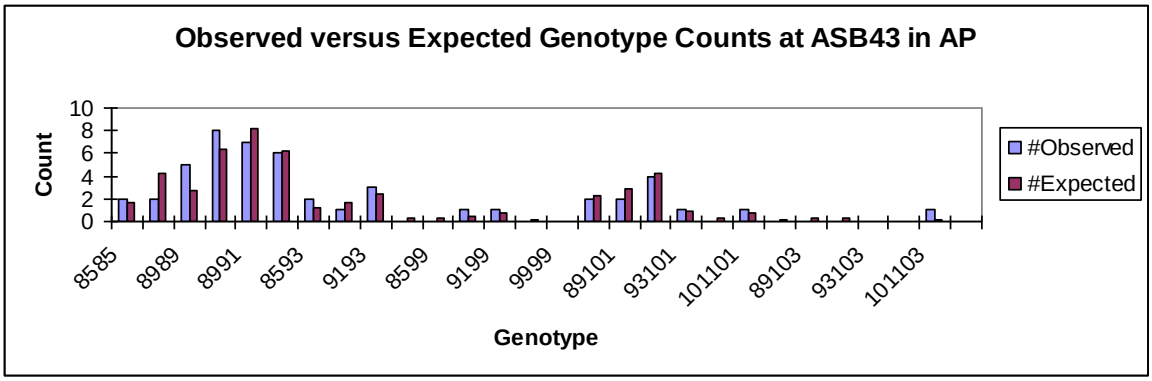

Observed versus Expected Genotype Counts at ASB2 in AP
Count
8
6
4
2
0
219219
219235
219237
237237
235241
219243
237243
219245
237245
245245
235247
243247
219249
237249
245249
219251
237251
245251
251251
Genotype
#Observed
#Expected

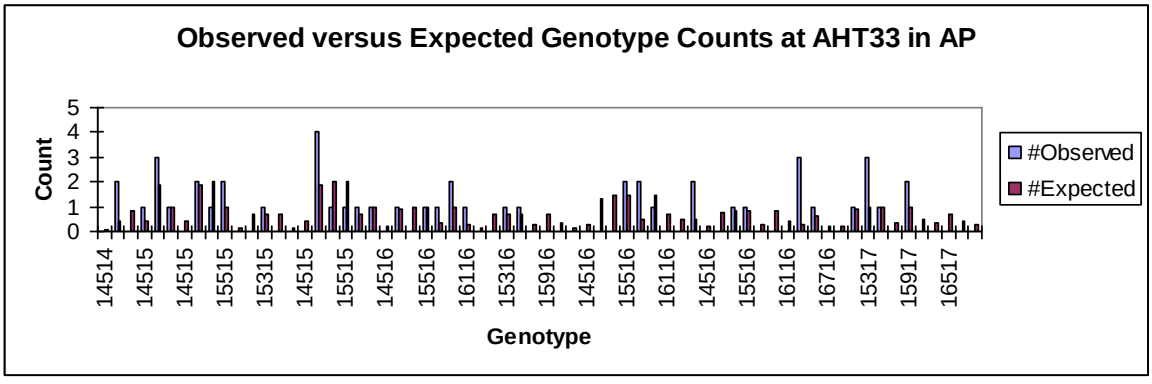

Observed versus Expected Genotype Counts at I18 in AP
Count
8
6
4
2
0
8383
8393
8395
9595
9397
8399
9599
83103
95103
103103
93105
99105
83107
95107
103107
83109
95109
103109
109109
Genotype
#Observed
#Expected

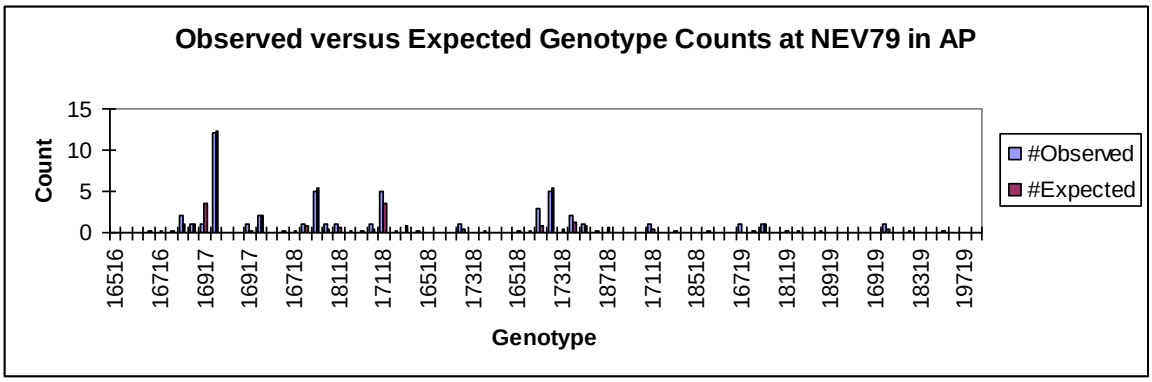

Observed versus Expected Genotype Counts at AHT4 in AP
Count
5
4
3
2
1
0
146146
146150
146152
152152
150154
146156
152156
146158
152158
158158
150160
156160
146162
152162
158162
Genotype
#Observed
#Expected

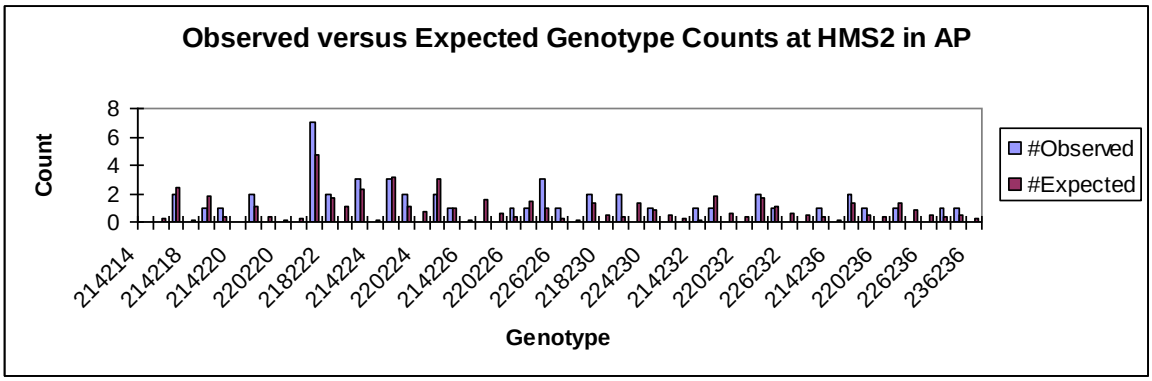

Observed versus Expected Genotype Counts at LEX33 in AP
Count
10
8
6
4
2
0
195195
195205
195207
207207
205211
195213
207213
195215
207215
215215
205217
213217
195219
207219
215219
195221
207221
215221
221221
Genotype
#Observed
#Expected

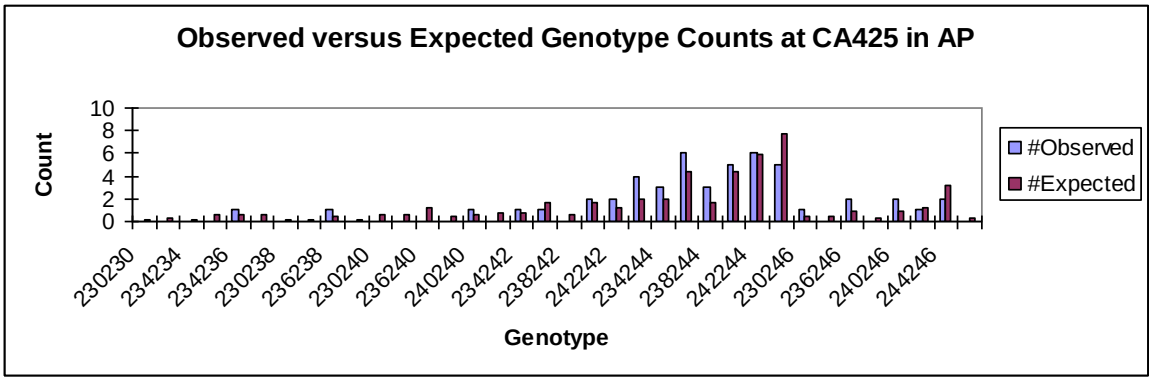

Observed versus Expected Genotype Counts at COR2 in AP
Count
15
10
5
0
228228
228230
230230
228232
230232
232232
228234
230234
232234
234234
228236
230236
232236
234236
236236
Genotype
#Observed
#Expected

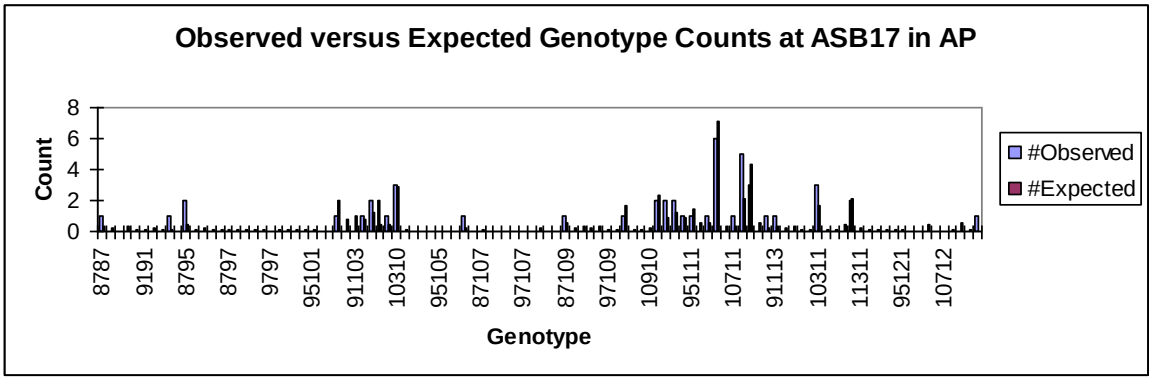

Observed versus Expected Genotype Counts at ASB43 in AP
Count
10
8
6
4
2
0
8585
8989
8991
8593
9193
8599
9199
9999
89101
93101
101101
89103
93103
101103
Genotype
#Observed
#Expected

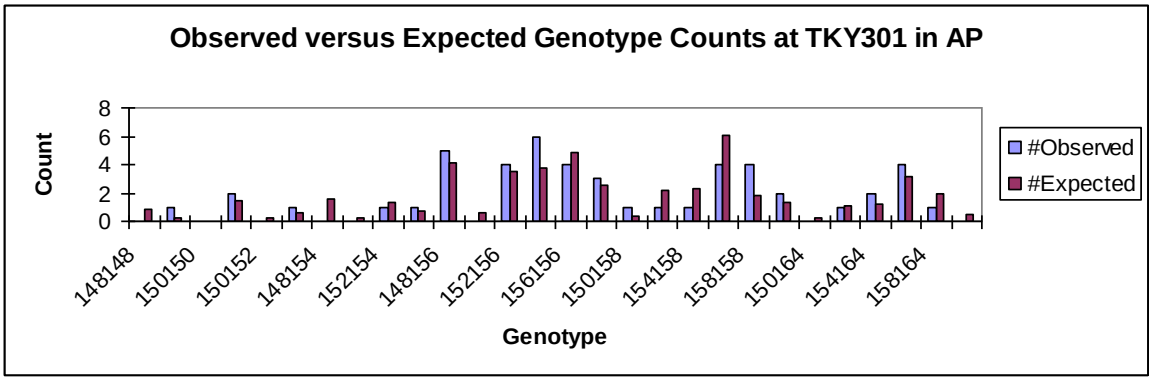

Observed versus Expected Genotype Counts at AHT33 in AP
Count
5
4
3
2
1
0
14514
14515
14515
15515
15315
14515
15515
14516
15516
16116
15316
15916
14516
15516
16116
14516
15516
16116
16716
15317
15917
16517
Genotype
#Observed
#Expected

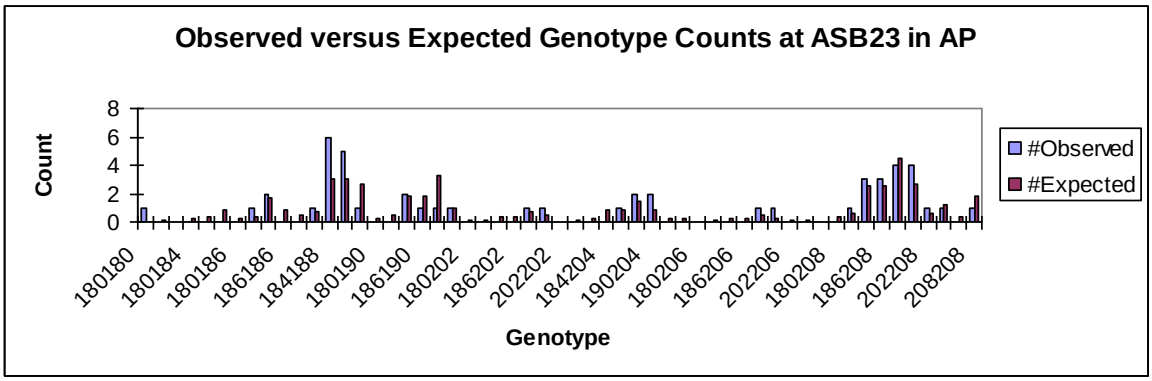

Observed versus Expected Genotype Counts at NEV79 in AP
Count
15
10
5
0
16516
16716
16917
16917
16718
18118
17118
16518
17318
16518
17318
18718
17118
18518
16719
18119
18919
16919
18319
19719
Genotype
#Observed
#Expected

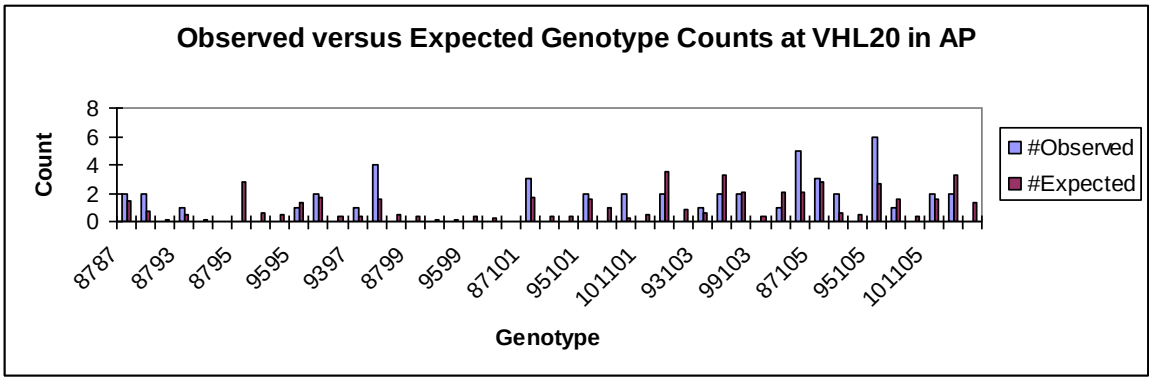

Observed versus Expected Genotype Counts at HMS2 in AP
Count
8
6
4
2
0
214214
214218
214220
220220
218222
214224
220224
214226
220226
226226
218230
224230
214232
220232
226232
214236
220236
226236
236236
Genotype
#Observed
#Expected

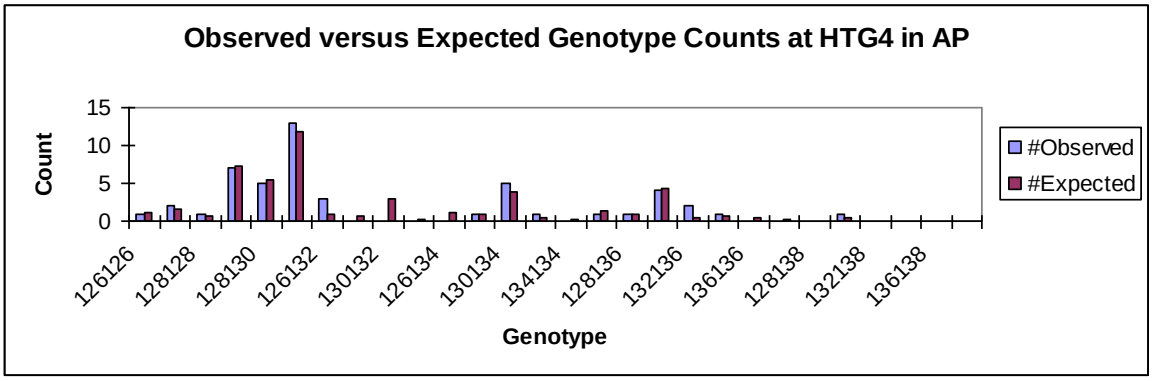

Observed versus Expected Genotype Counts at CA425 in AP
Count
10
8
6
4
2
0
230230
234234
234236
230238
236238
230240
236240
240240
234242
238242
242242
234244
238244
242244
230246
236246
240246
244246
Genotype
#Observed
#Expected

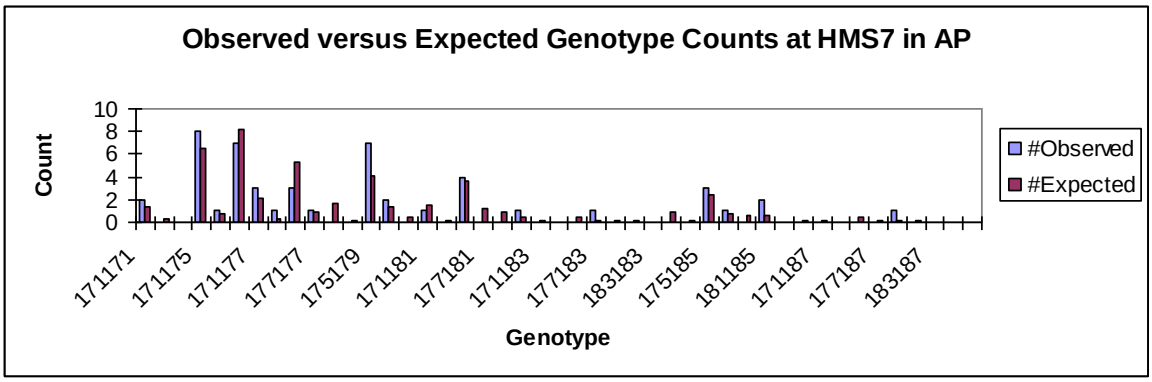

Observed versus Expected Genotype Counts at ASB17 in AP
Count
8
6
4
2
0
8787
9191
8795
8797
9797
95101
91103
10310
95105
87107
97107
87109
97109
10910
95111
10711
91113
10311
11311
95121
10712
Genotype
#Observed
#Expected

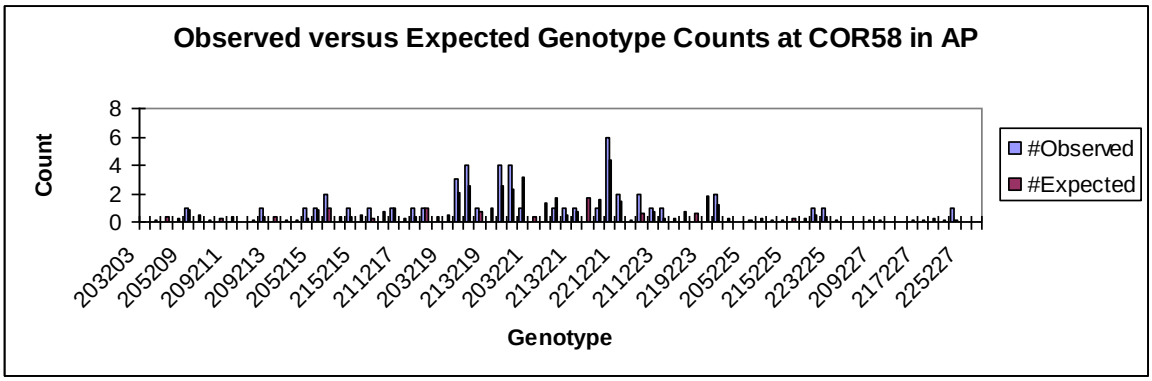

Observed versus Expected Genotype Counts at TKY301 in AP
Count
Genotype
#Observed
#Expected
148148
150150
150152
148154
152154
148156
152156
156156
150158
154158
158158
150164
154164
158164


Observed versus Expected Genotype Counts at ASB23 in AP
Count
Genotype
#Observed
#Expected
180180
180184
180186
186186
184188
180190
186190
180202
186202
202202
184204
190204
180206
186206
202206
180208
186208
202208
208208

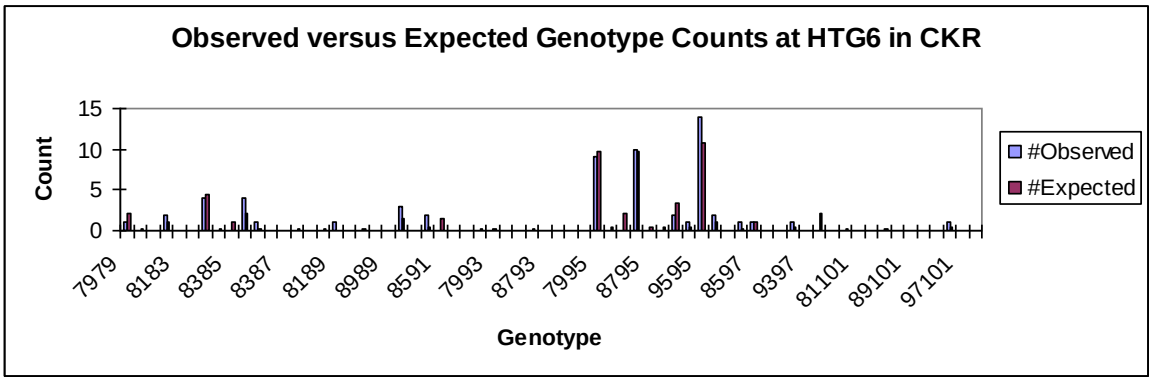

Observed versus Expected Genotype Counts at VHL20 in AP
Count
Genotype
#Observed
#Expected
8787
8793
8795
9595
9397
8799
9599
87101
95101
101101
93103
99103
87105
95105
101105

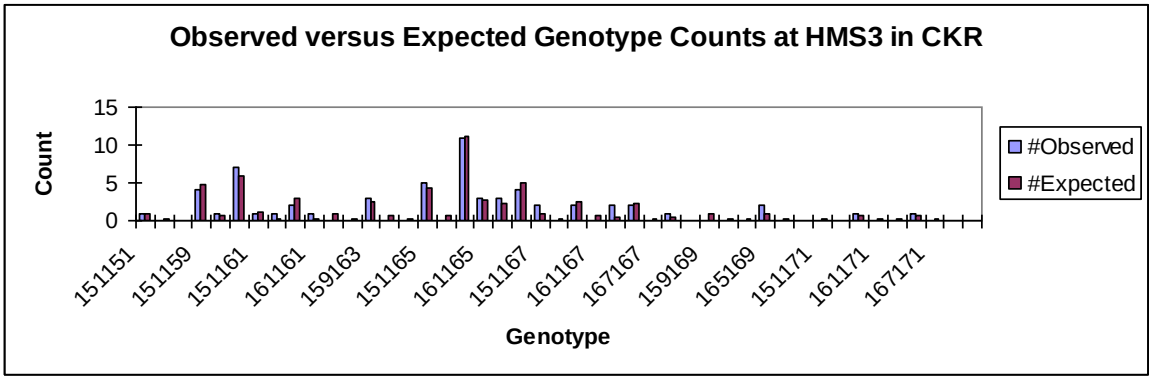

Observed versus Expected Genotype Counts at HTG4 in AP
Count
Genotype
#Observed
#Expected
126126
128128
128130
126132
130132
126134
130134
134134
128136
132136
136136
128138
132138
136138

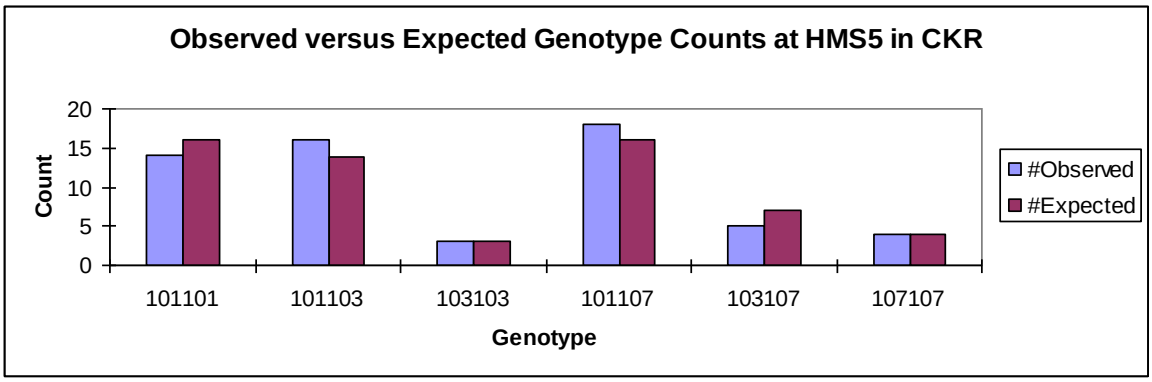


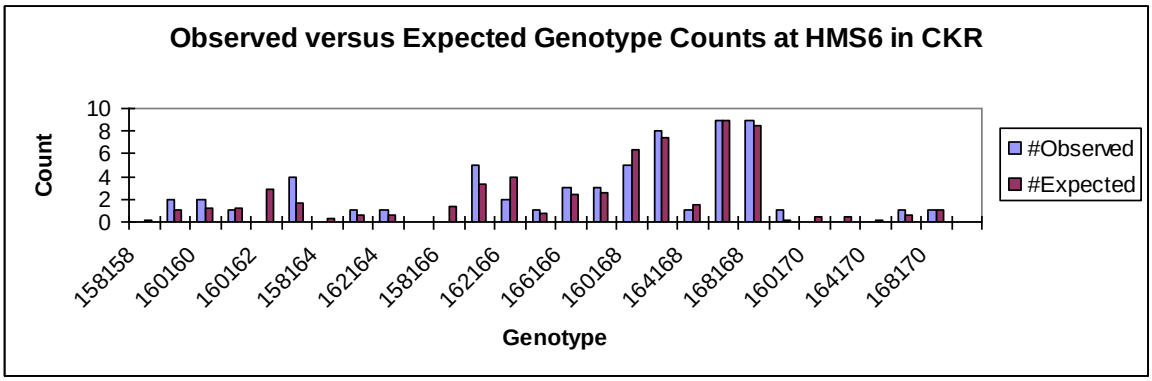


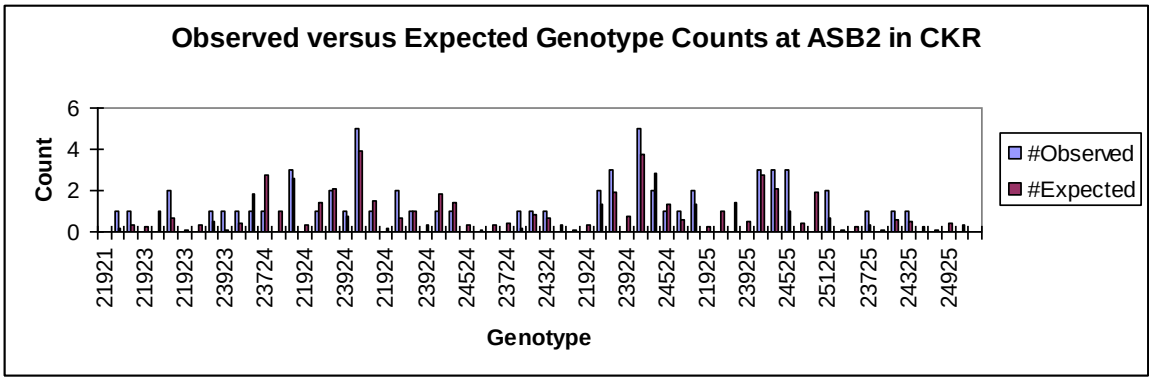
CUKUROVA HORSE

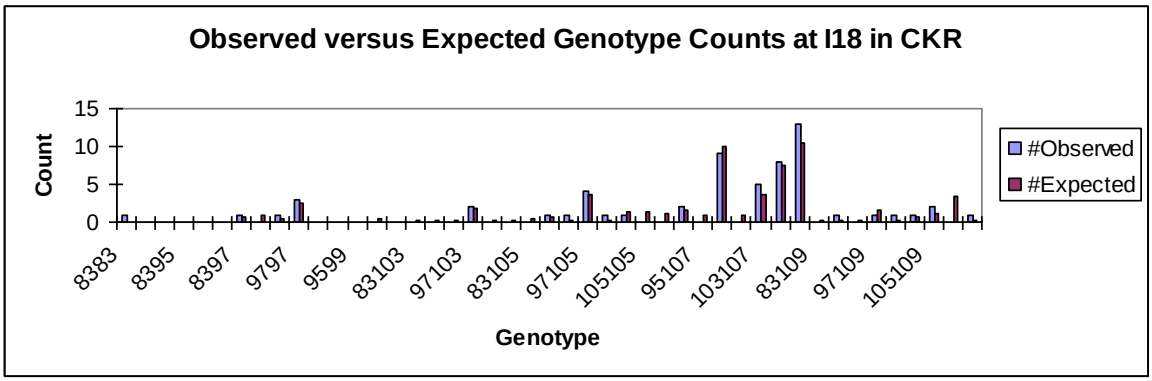


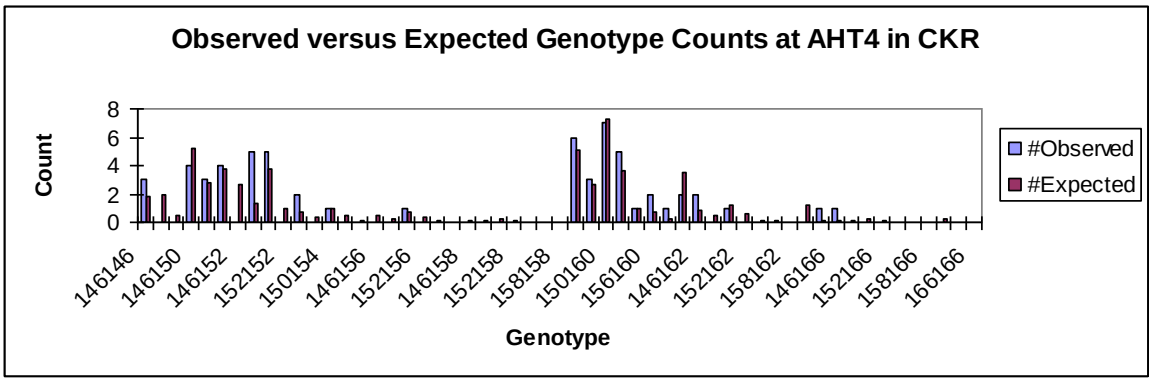


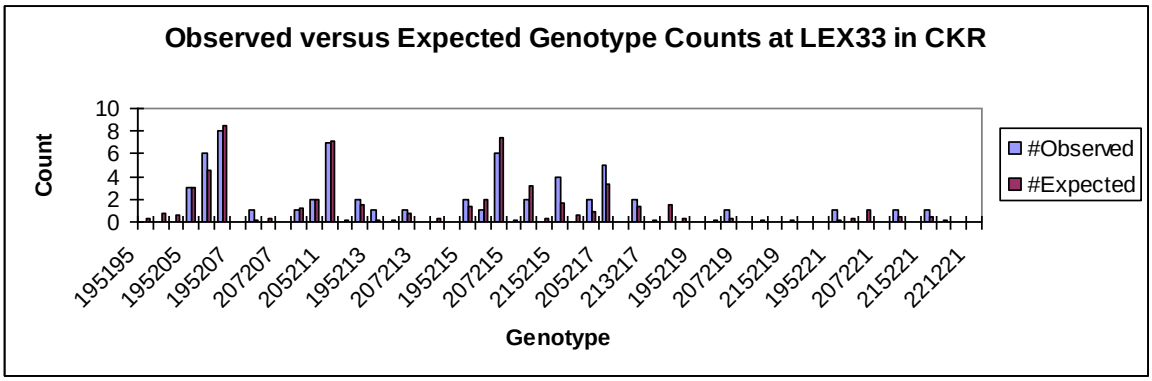

Observed versus Expected Genotype Counts at HMS5 in CKR
Count
20
15
10
5
0
101101
101103
103103
101107
103107
107107
Genotype
#Observed
#Expected

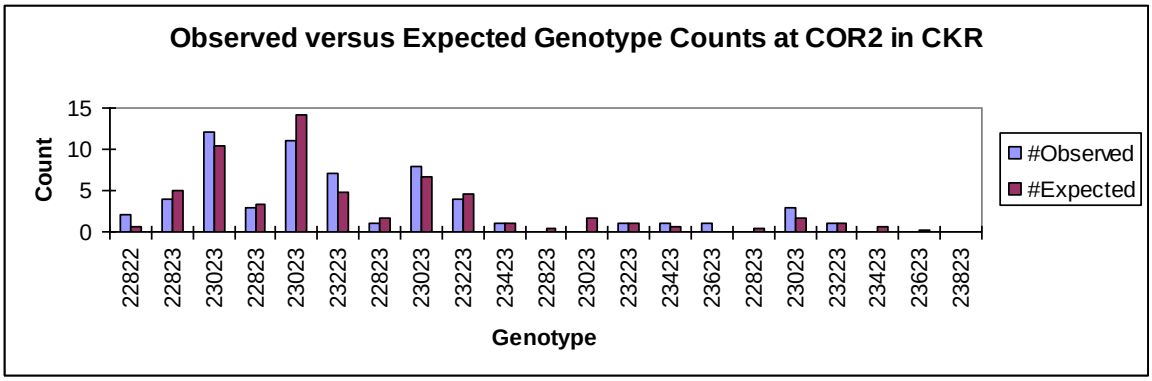

Observed versus Expected Genotype Counts at HMS6 in CKR
Count
10
8
6
4
2
0
158158
160160
160162
158164
162164
158166
162166
166166
160168
164168
168168
160170
164170
168170
Genotype
#Observed
#Expected

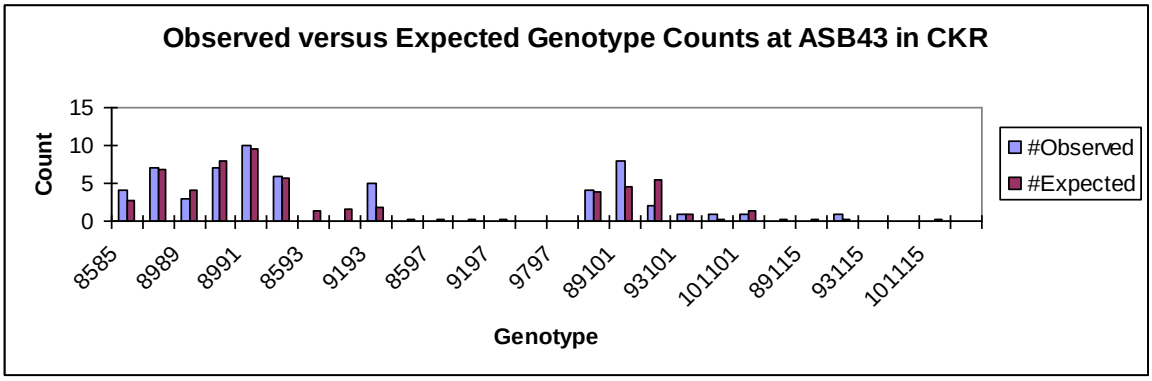

Observed versus Expected Genotype Counts at ASB2 in CKR
Count
6
4
2
0
21921
21923
21923
23923
23724
21924
23924
21924
23924
24524
23724
24324
21924
23924
24524
21925
23925
24525
25125
23725
24325
24925
Genotype
#Observed
#Expected

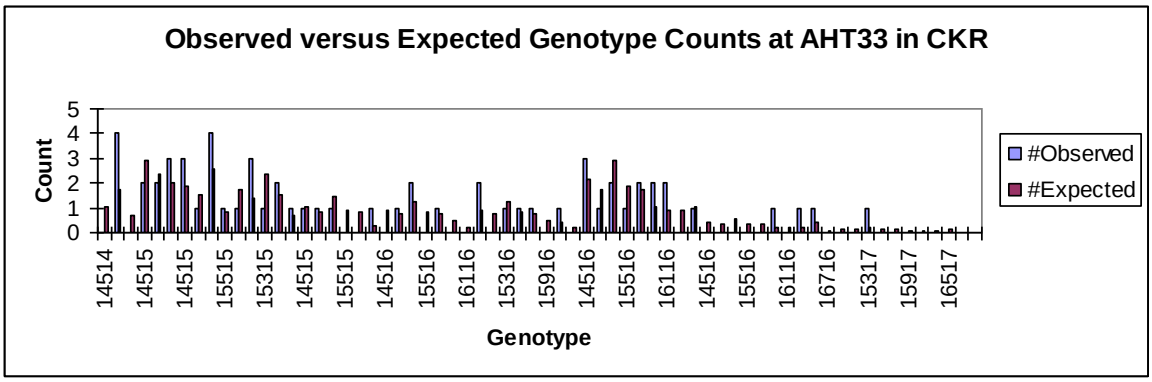

Observed versus Expected Genotype Counts at I18 in CKR
Count
15
10
5
0
8383
8395
8397
9797
9599
83103
97103
83105
97105
105105
95107
103107
83109
97109
105109
Genotype
#Observed
#Expected

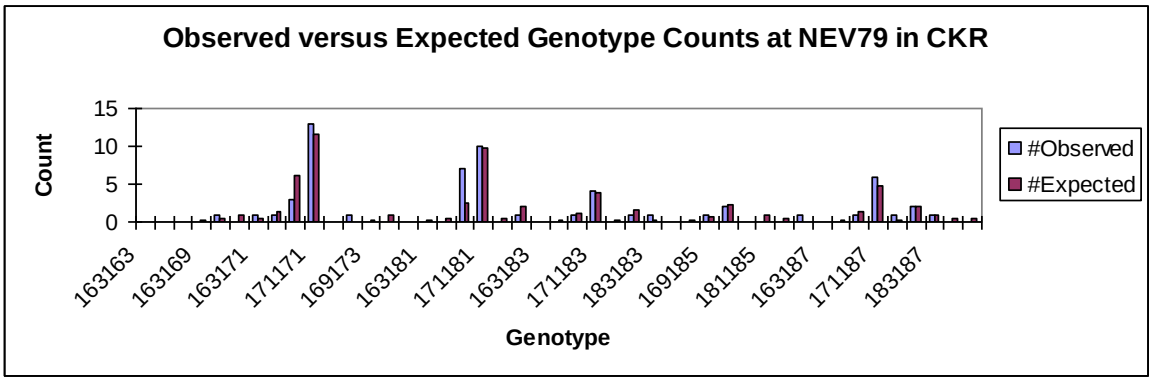

Observed versus Expected Genotype Counts at AHT4 in CKR
Count
Genotype
#Observed
#Expected
146146
146150
146152
152152
150154
146156
152156
146158
152158
158158
150160
156160
146162
152162
158162
146166
152166
158166
166166

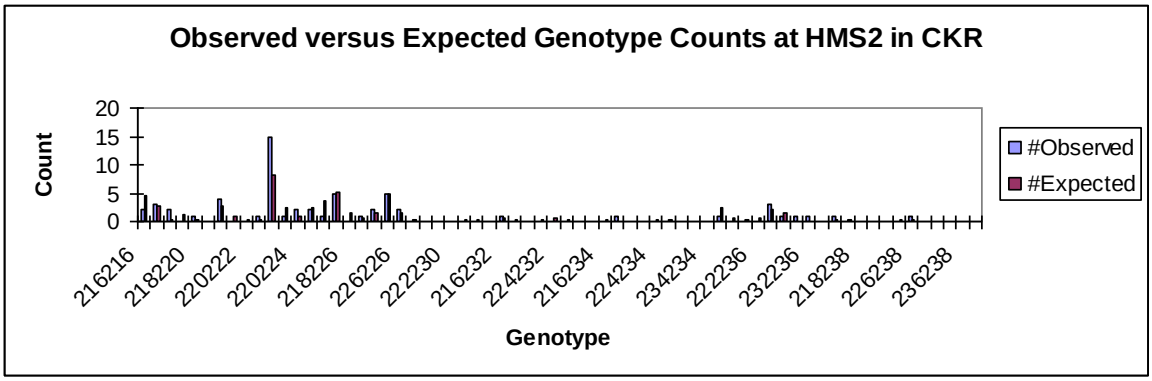

Observed versus Expected Genotype Counts at LEX33 in CKR
Count
Genotype
#Observed
#Expected
195195
195205
195207
207207
205211
195213
207213
195215
207215
215215
205217
213217
195219
207219
215219
195221
207221
215221
221221

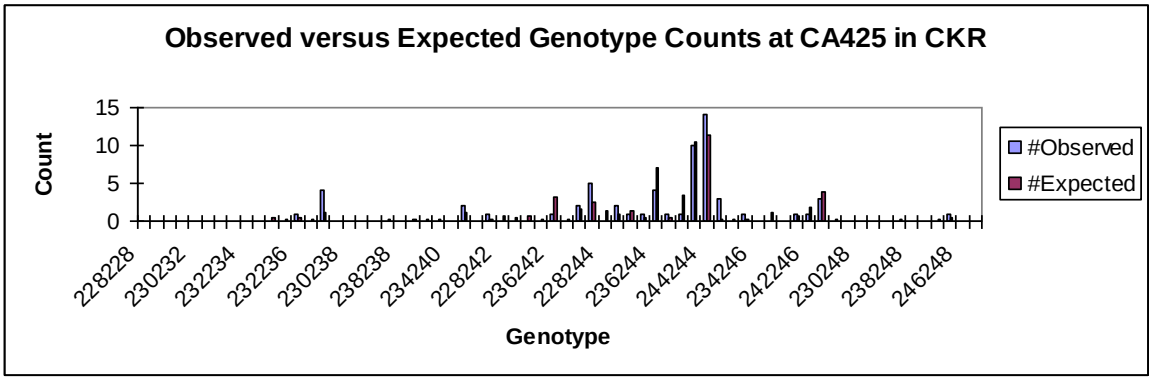

Observed versus Expected Genotype Counts at COR2 in CKR
Count
Genotype
#Observed
#Expected
22822
22823
23023
22823
23023
23223
22823
23023
23223
23423
22823
23023
23223
23423
23623
22823
23023
23223
23423
23623
23823

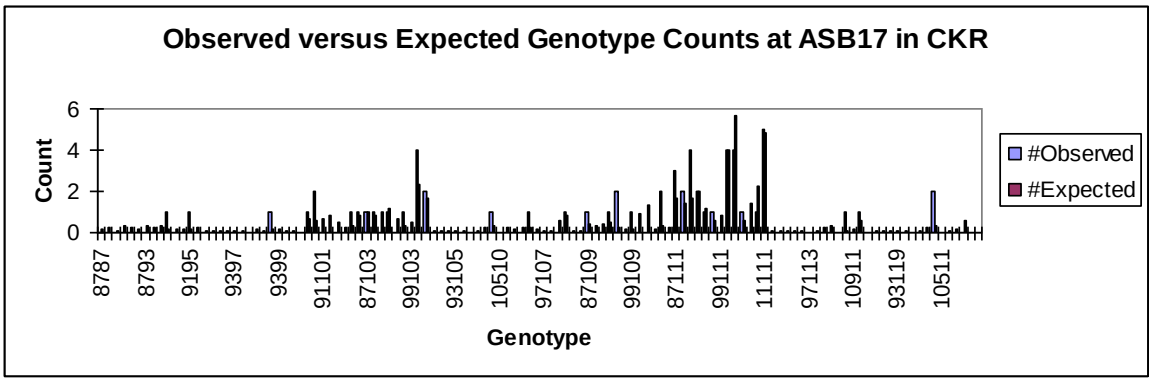

Observed versus Expected Genotype Counts at ASB43 in CKR
Count
Genotype
#Observed
#Expected
8585
8989
8991
8593
9193
8597
9197
9797
89101
93101
101101
89115
93115
101115

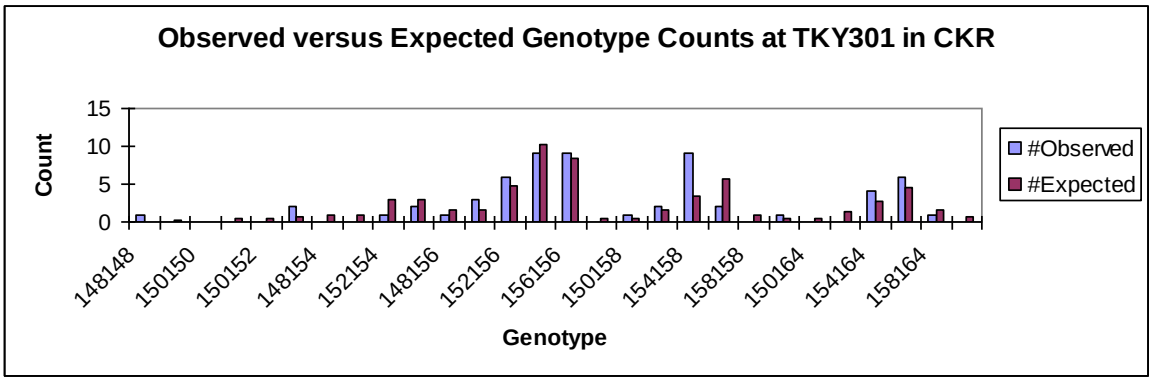

Observed versus Expected Genotype Counts at AHT33 in CKR
Count
0
1
2
3
4
5
14514
14515
14515
15515
15315
14515
15515
14516
15516
16116
15316
15916
14516
15516
16116
14516
15516
16116
16716
15317
15917
16517
Genotype
#Observed
#Expected

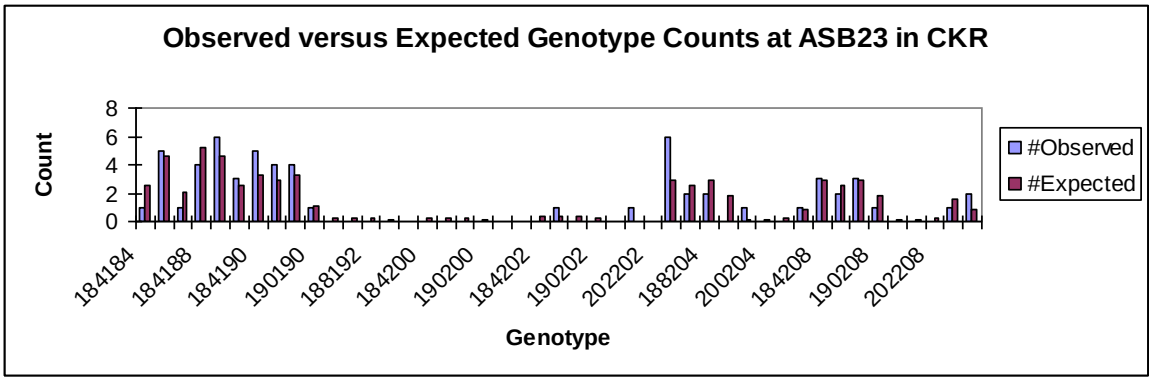

Observed versus Expected Genotype Counts at NEV79 in CKR
Count
0
5
10
15
163163
163169
163171
171171
169173
163181
171181
163183
171183
183183
169185
181185
163187
171187
183187
Genotype
#Observed
#Expected

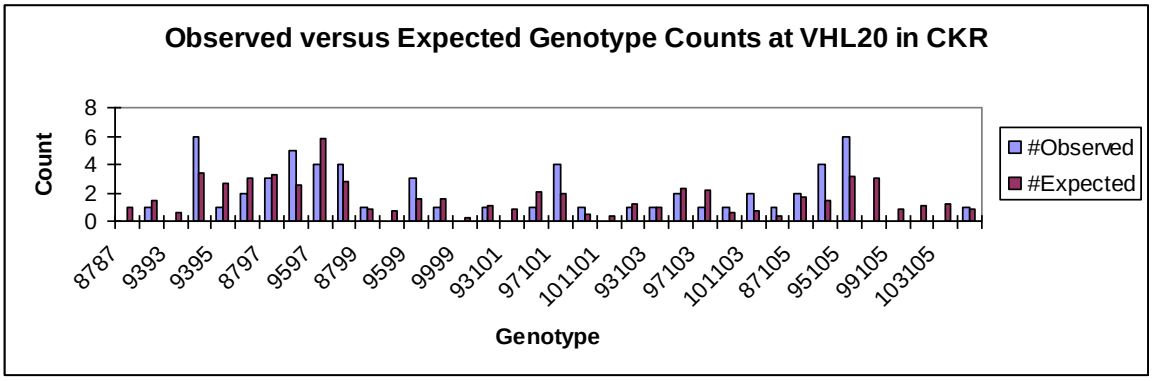

Observed versus Expected Genotype Counts at HMS2 in CKR
Count
0
5
10
15
20
216216
218220
220222
220224
218226
226226
222230
216232
224232
216234
224234
234234
222236
232236
218238
226238
236238
Genotype
#Observed
#Expected

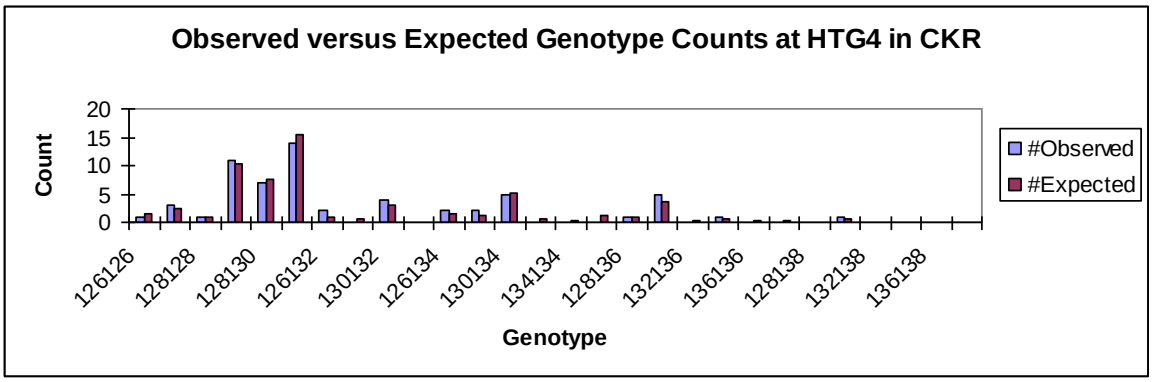

Observed versus Expected Genotype Counts at CA425 in CKR
Count
0
5
10
15
228228
230232
232234
232236
230238
238238
234240
228242
236242
228244
236244
244244
234246
242246
230248
238248
246248
Genotype
#Observed
#Expected

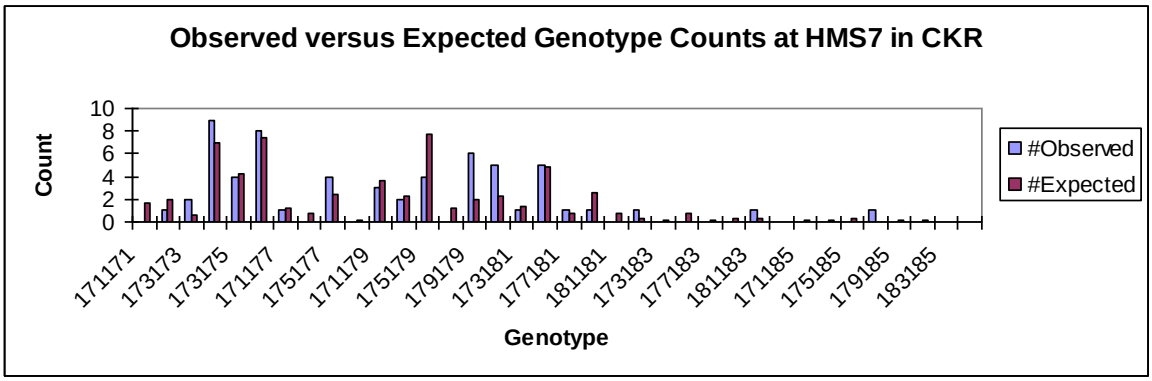

Observed versus Expected Genotype Counts at ASB17 in CKR
Count
Genotype
#Observed
#Expected
8787
8793
9195
9397
9399
91101
87103
99103
93105
10510
97107
87109
99109
87111
99111
11111
97113
10911
93119
10511

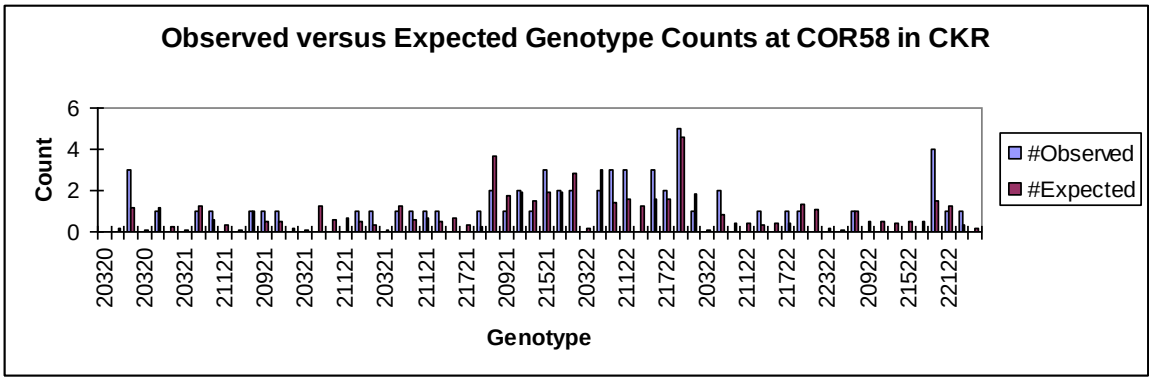

Observed versus Expected Genotype Counts at TKY301 in CKR
Count
Genotype
#Observed
#Expected
148148
150150
150152
148154
152154
148156
152156
156156
150158
154158
158158
150164
154164
158164


Observed versus Expected Genotype Counts at ASB23 in CKR
Count
Genotype
#Observed
#Expected
184184
184188
184190
190190
188192
184200
190200
184202
190202
202202
188204
200204
184208
190208
202208

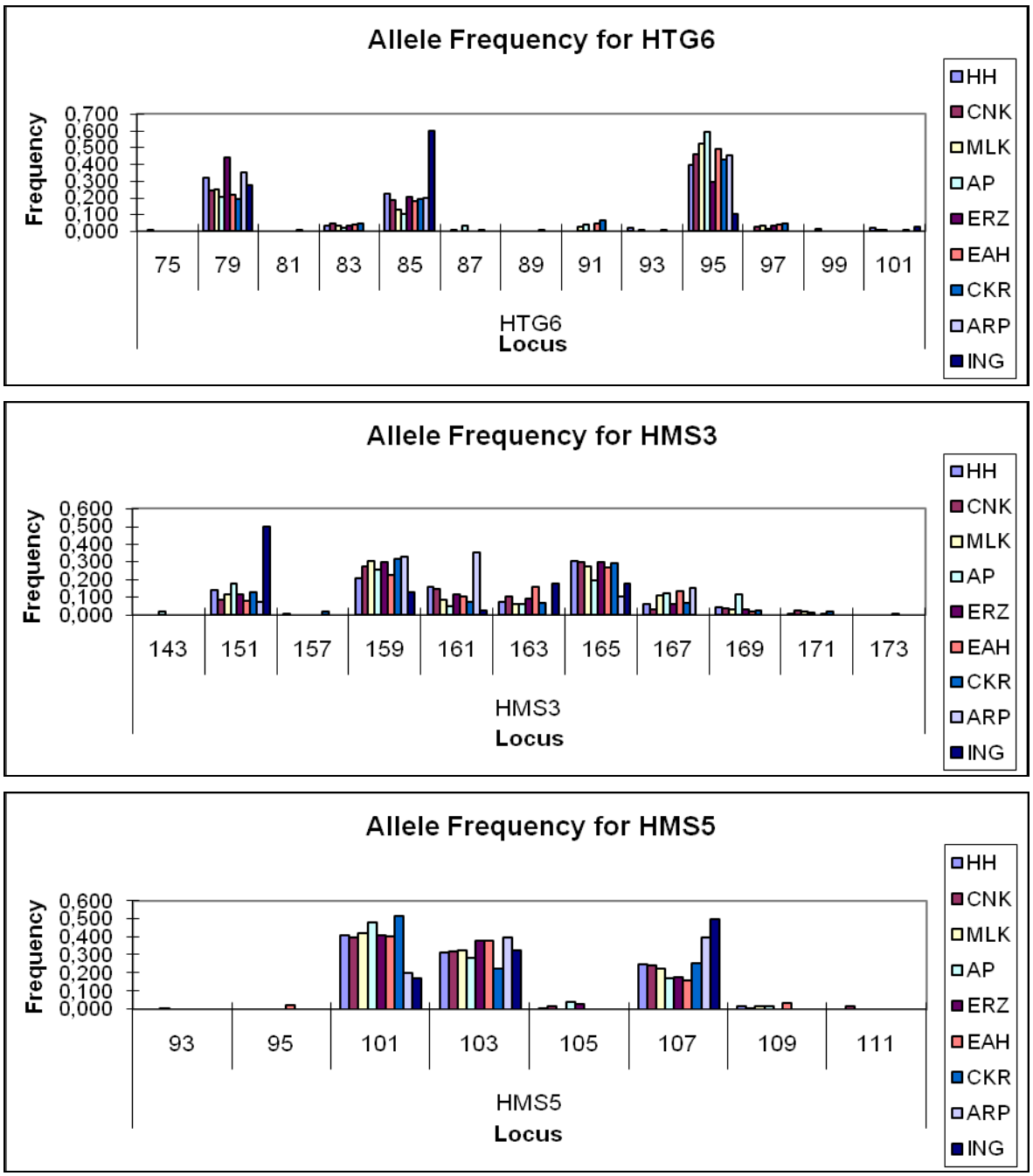

Observed versus Expected Genotype Counts at VHL20 in CKR
Count
Genotype
#Observed
#Expected
8787
9393
9395
8797
9597
8799
9599
9999
93101
97101
101101
93103
97103
101103
87105
95105
99105
103105

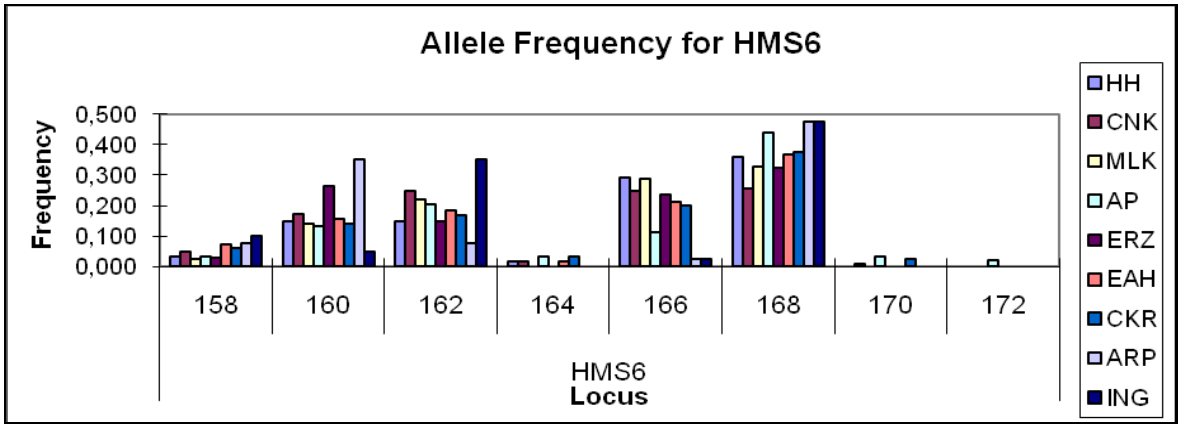

Observed versus Expected Genotype Counts at HTG4 in CKR
Count
20
15
10
5
0
126126
128128
128130
126132
130132
126134
130134
134134
128136
132136
136136
128138
132138
136138
Genotype
#Observed
#Expected

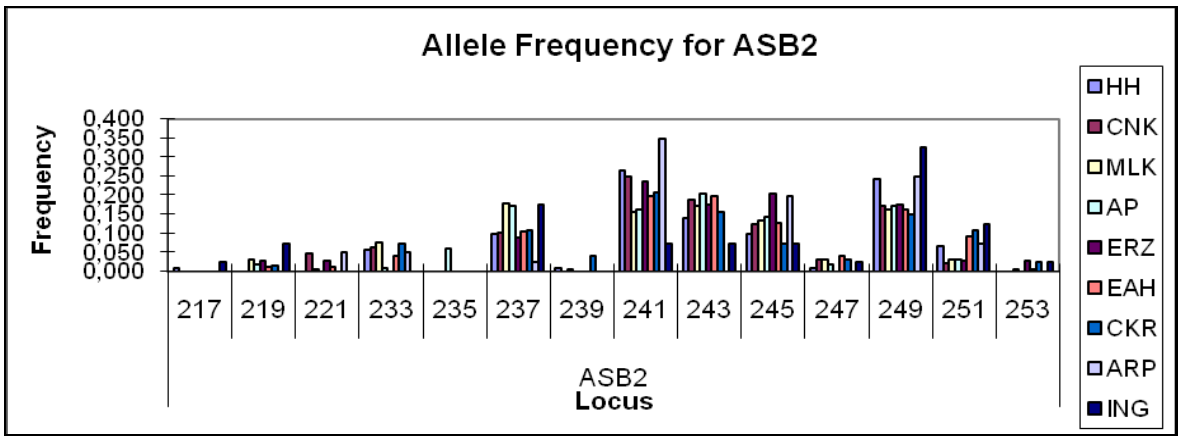

Observed versus Expected Genotype Counts at HMS7 in CKR
Count
10
8
6
4
2
0
171171
173173
173175
171177
175177
171179
175179
179179
173181
177181
181181
173183
177183
181183
171185
175185
179185
183185
Genotype
#Observed
#Expected

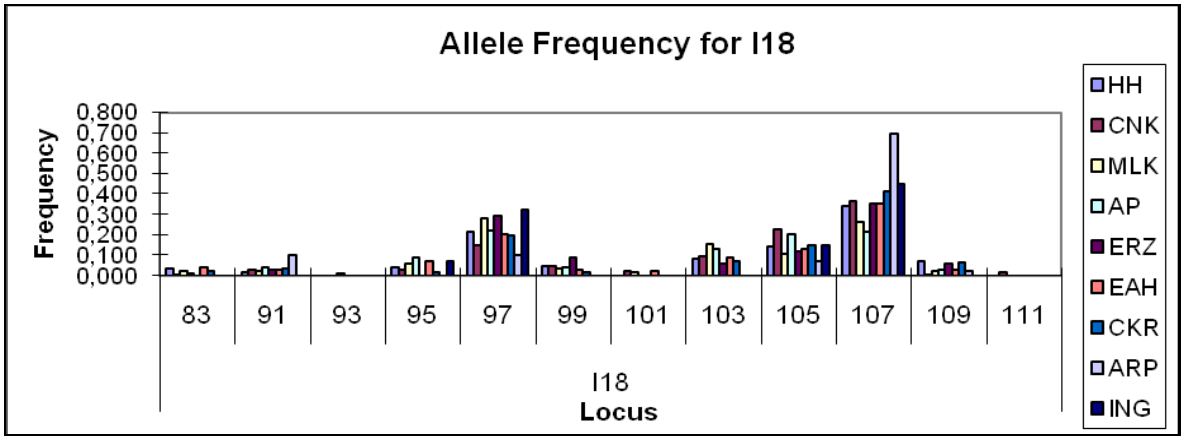

Observed versus Expected Genotype Counts at COR58 in CKR
Count
6
4
2
0
20320
20320
20321
21121
20921
20321
21121
20321
21121
21721
20921
21521
20322
21122
21722
20322
21122
21722
22322
20922
21522
22122
Genotype
#Observed
#Expected

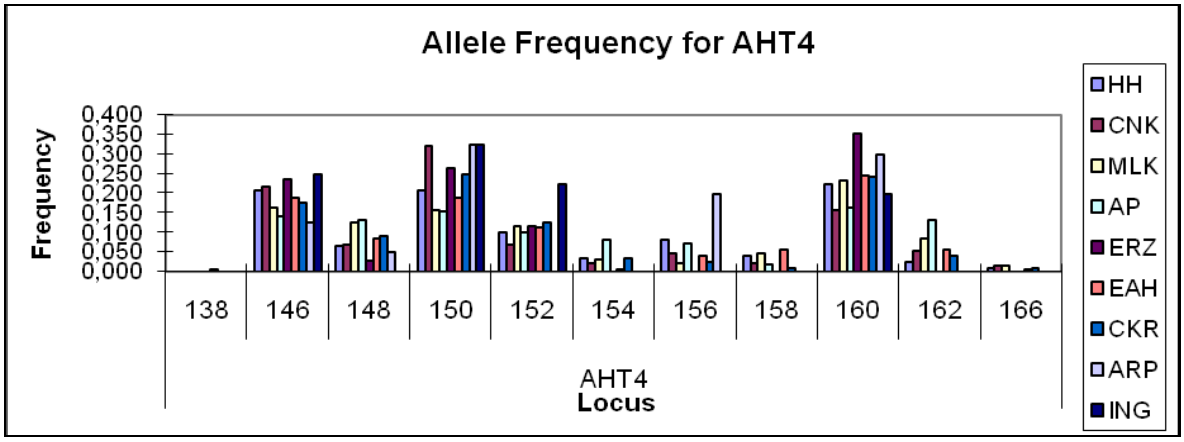
APPENDIX B: ALLELE FREQUENCIES WITH GRAPHS BY POPULATION AND LOCUS

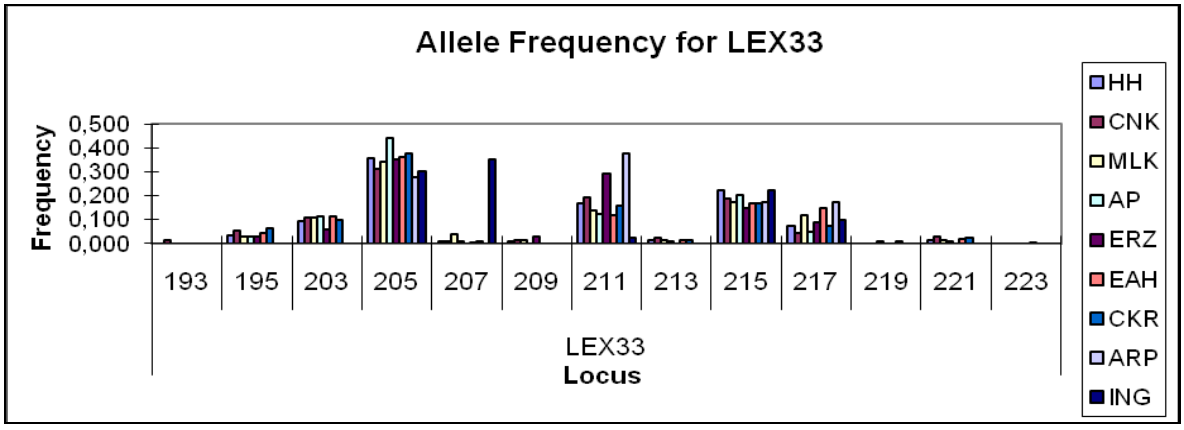


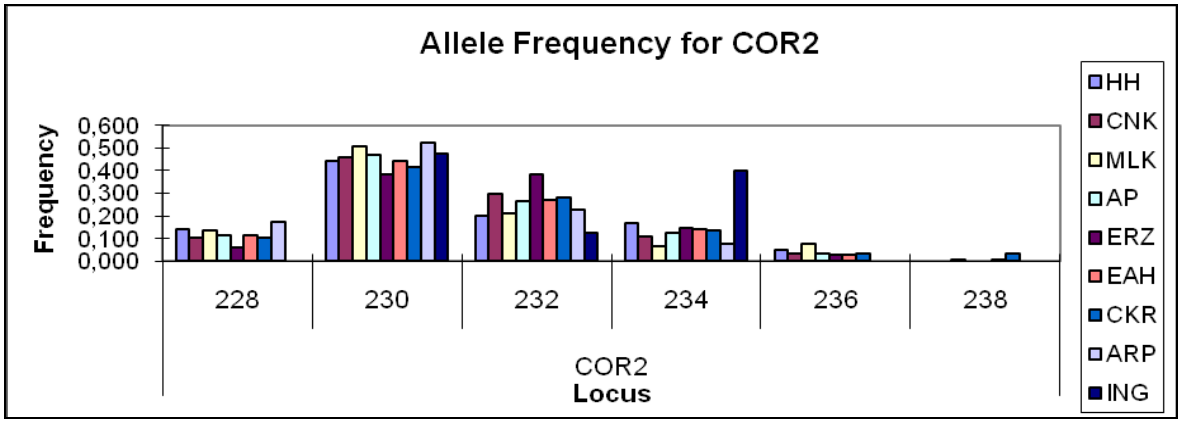

Allele Frequency for HMS6
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
158
160
162
164
166
168
170
172
HMS6
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

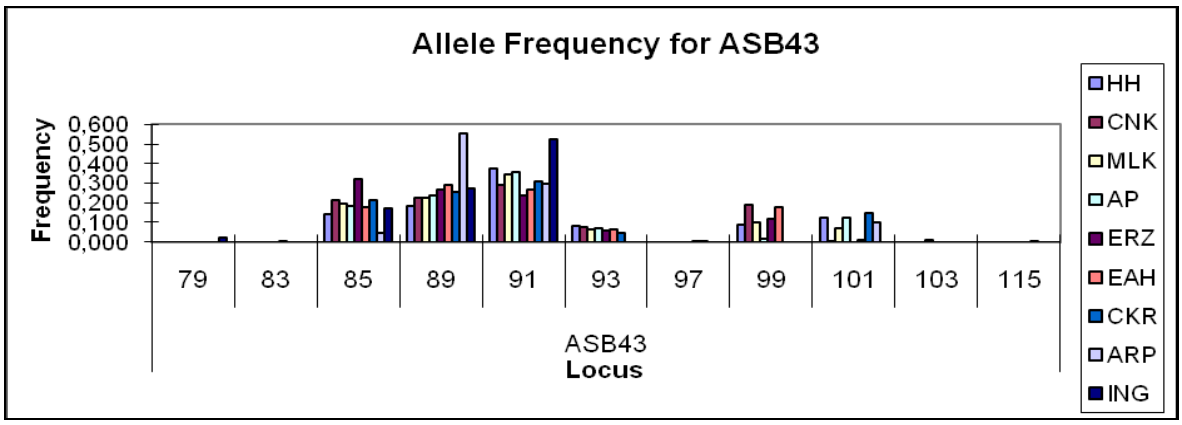

Allele Frequency for ASB2
Frequency
0,400
0,350
0,300
0,250
0,200
0,150
0,100
0,050
0,000
217
219
221
233
235
237
239
241
243
245
247
249
251
253
ASB2
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

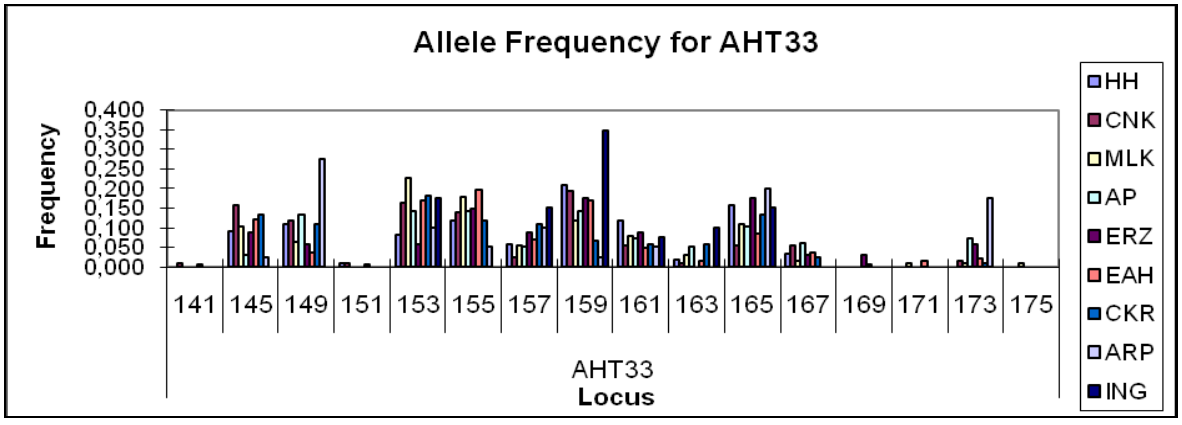

Allele Frequency for I18
Frequency
0,800
0,700
0,600
0,500
0,400
0,300
0,200
0,100
0,000
83
91
93
95
97
99
101
103
105
107
109
111
I18
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

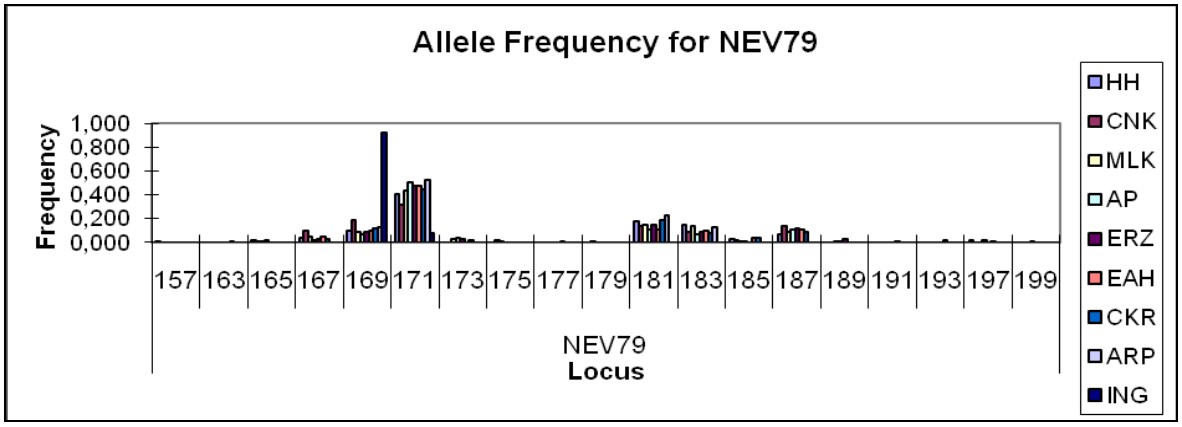

Allele Frequency for AHT4
Frequency
0,400
0,350
0,300
0,250
0,200
0,150
0,100
0,050
0,000
138
146
148
150
152
154
156
158
160
162
166
AHT4
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

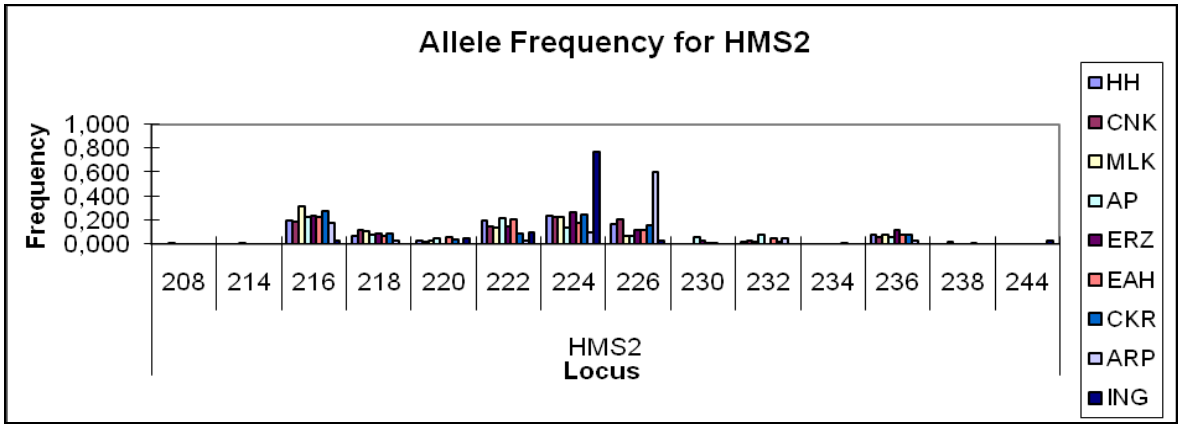

Allele Frequency for LEX33
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
193 195 203 205 207 209 211 213 215 217 219 221 223
LEX33
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

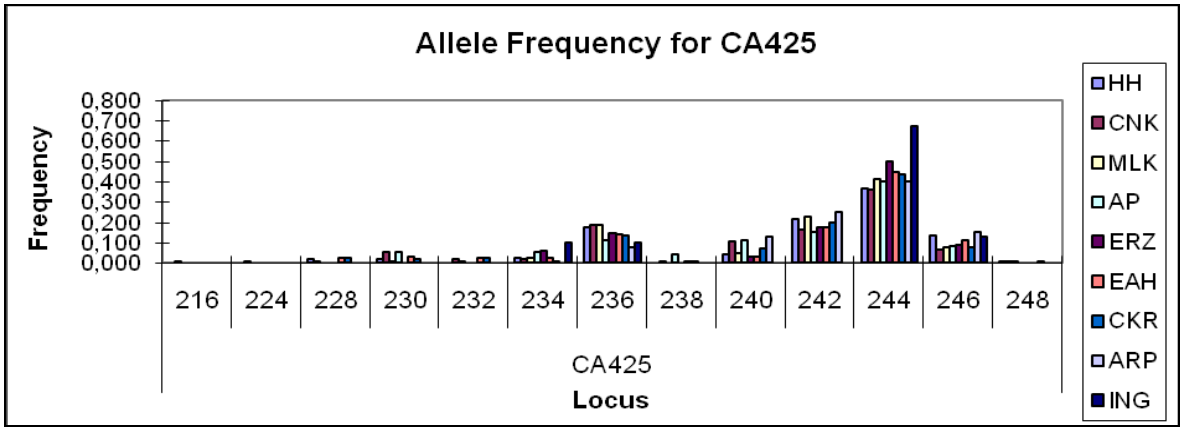

Allele Frequency for COR2
Frequency
0,600
0,500
0,400
0,300
0,200
0,100
0,000
228 230 232 234 236 238
COR2
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

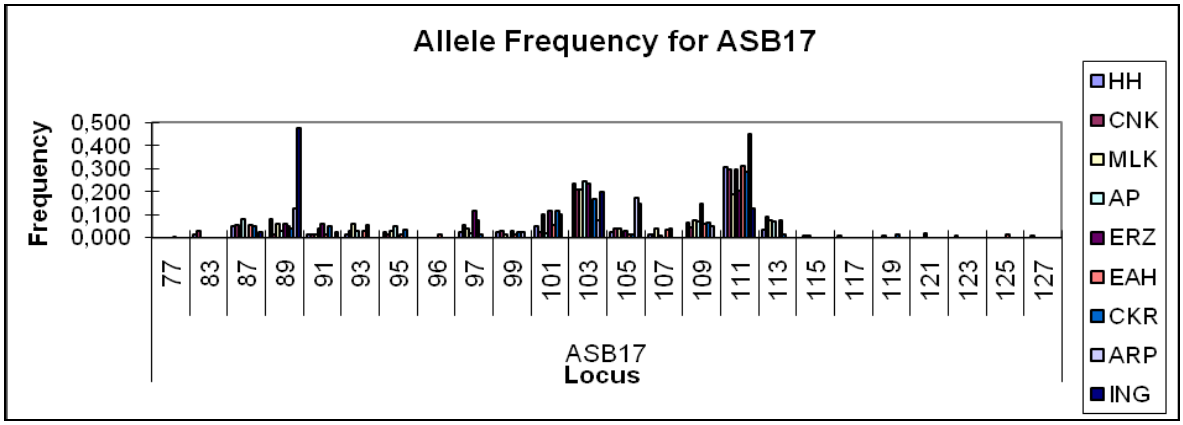

Allele Frequency for ASB43
Frequency
0,600
0,500
0,400
0,300
0,200
0,100
0,000
79 83 85 89 91 93 97 99 101 103 115
ASB43
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

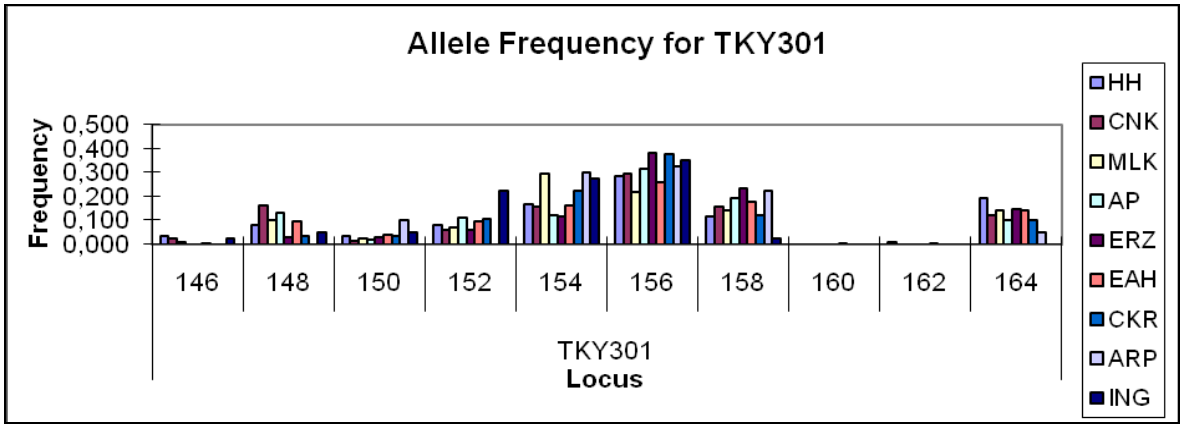

Allele Frequency for AHT33
Frequency
0,400
0,350
0,300
0,250
0,200
0,150
0,100
0,050
0,000
141 145 149 151 153 155 157 159 161 163 165 167 169 171 173 175
AHT33
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

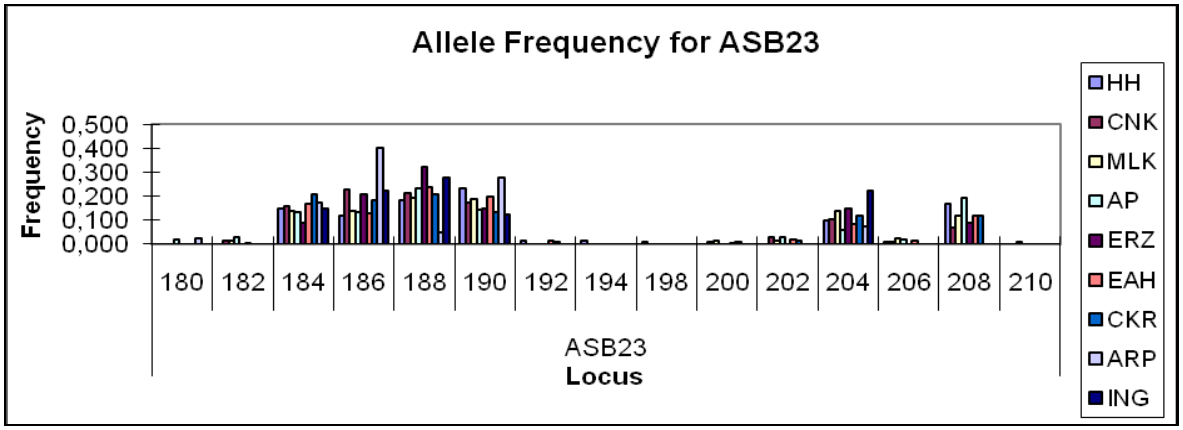

Allele Frequency for NEV79
Frequency
1,000
0,800
0,600
0,400
0,200
0,000
157 163 165 167 169 171 173 175 177 179 181 183 185 187 189 191 193 197 199
NEV79
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

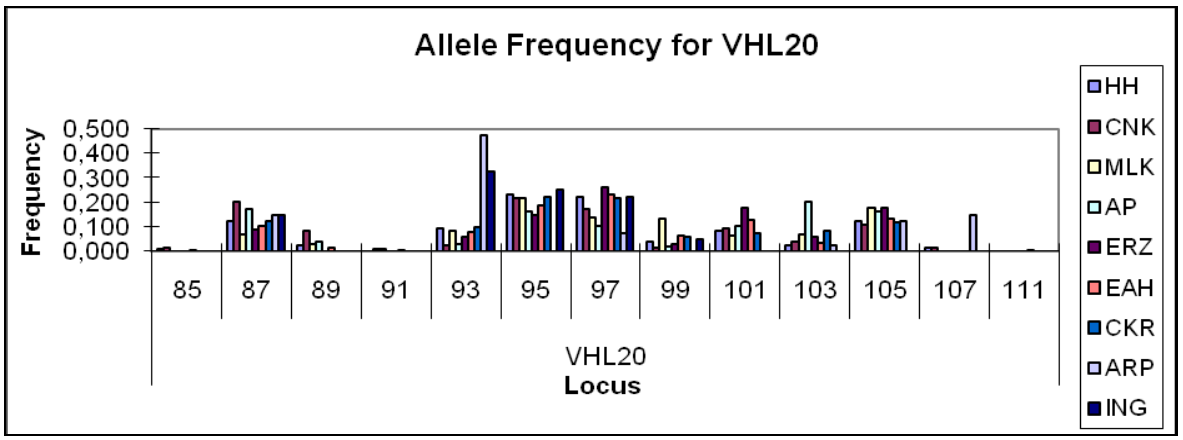

Allele Frequency for HMS2
Frequency
1,000
0,800
0,600
0,400
0,200
0,000
208 214 216 218 220 222 224 226 230 232 234 236 238 244
HMS2
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

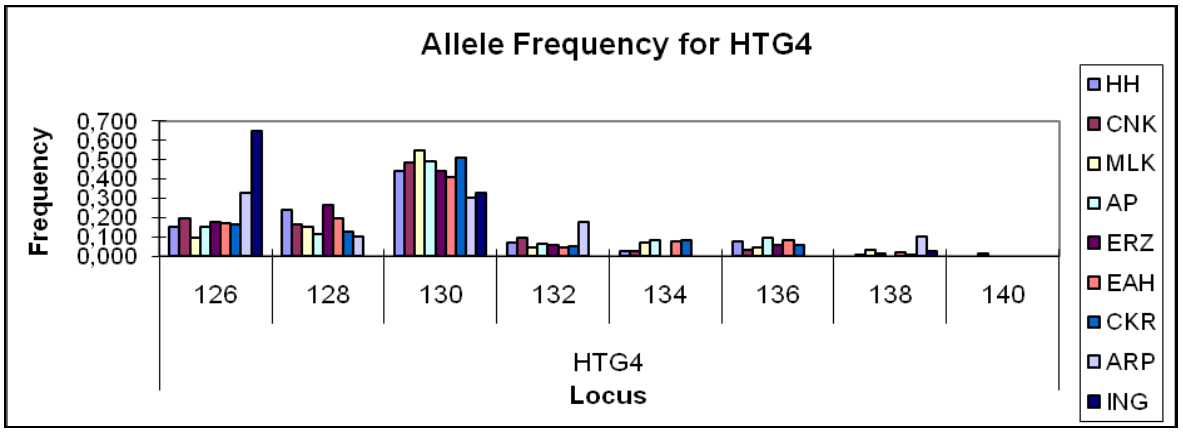

Allele Frequency for CA425
Frequency
0,800
0,700
0,600
0,500
0,400
0,300
0,200
0,100
0,000
216 224 228 230 232 234 236 238 240 242 244 246 248
CA425
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

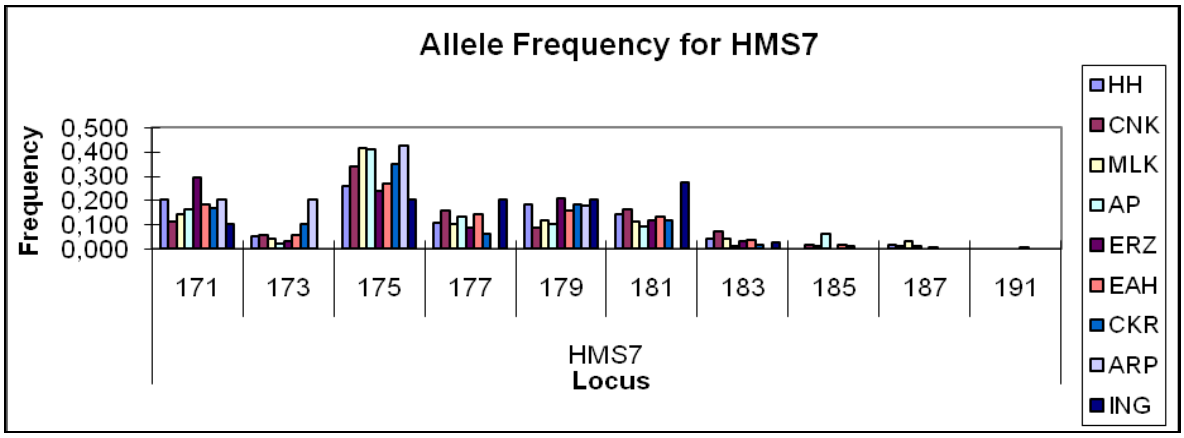

Allele Frequency for ASB17
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
77 83 87 89 91 93 95 96 97 99 101 103 105 107 109 111 113 115 117 119 121 123 125 127
ASB17
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

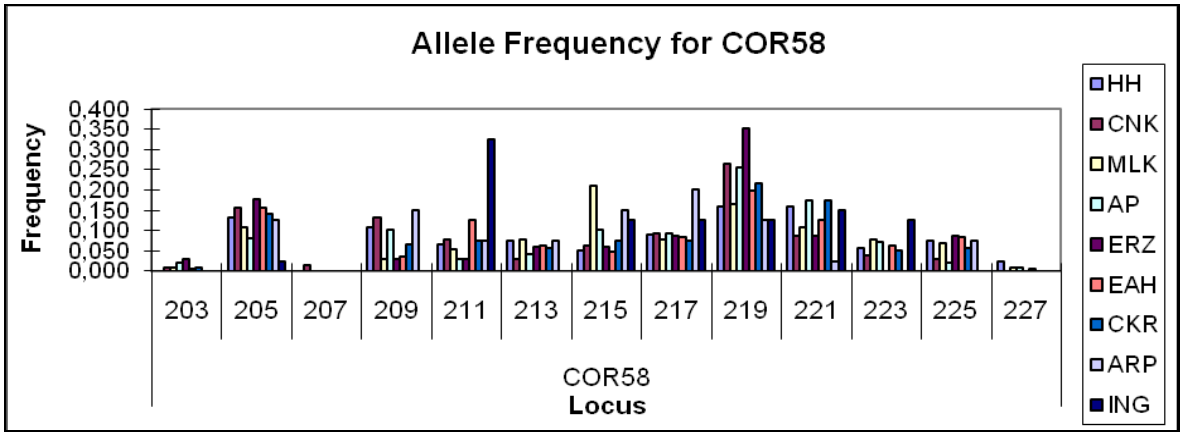

Allele Frequency for TKY301
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
146
148
150
152
154
156
158
160
162
164
TKY301
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING


Allele Frequency for ASB23
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
180
182
184
186
188
190
192
194
198
200
202
204
206
208
210
ASB23
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

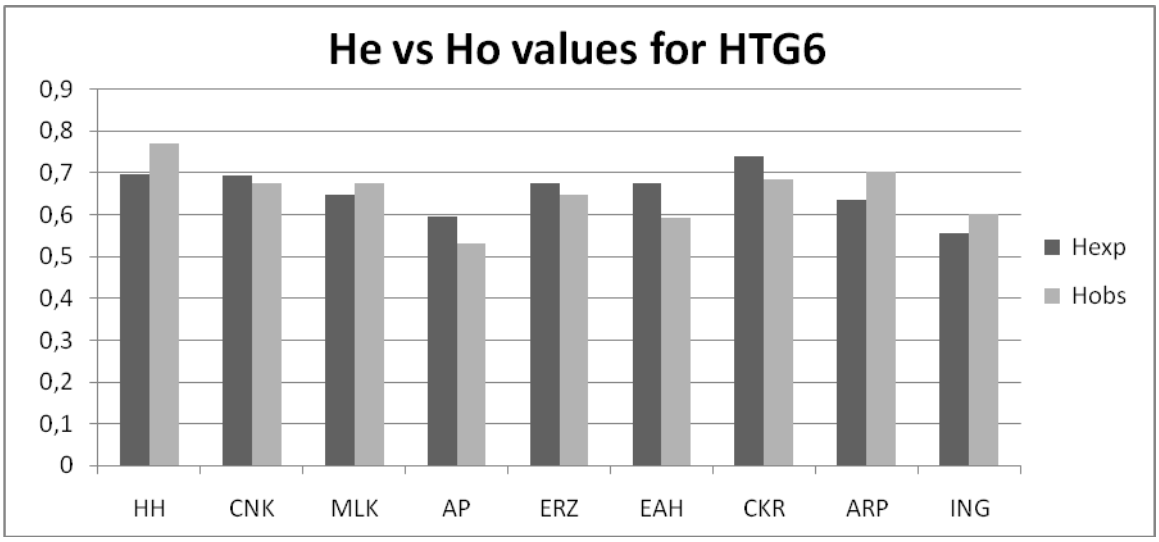

Allele Frequency for VHL20
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
85
87
89
91
93
95
97
99
101
103
105
107
111
VHL20
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

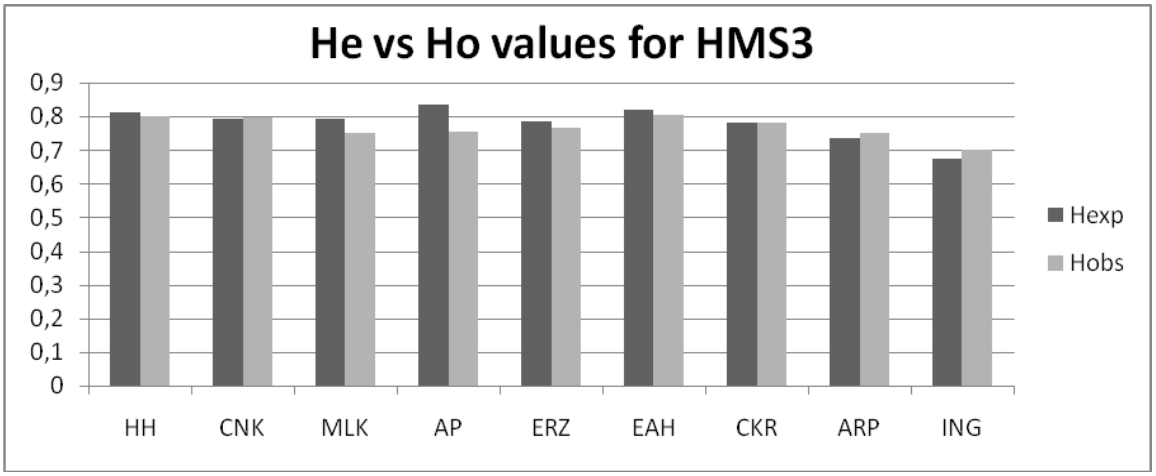

Allele Frequency for HTG4
Frequency
0,700
0,600
0,500
0,400
0,300
0,200
0,100
0,000
126
128
130
132
134
136
138
140
HTG4
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

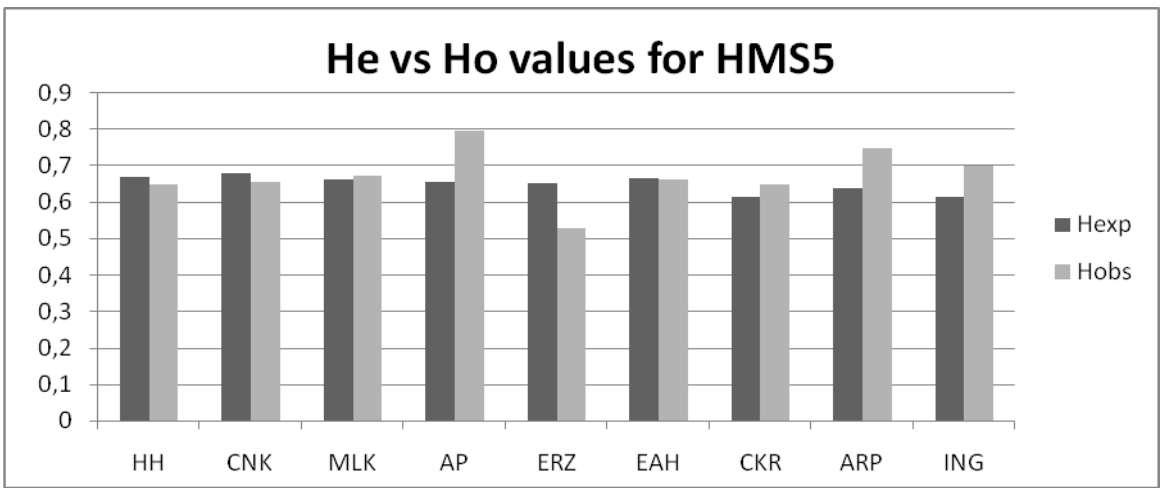

Allele Frequency for HMS7
Frequency
0,500
0,400
0,300
0,200
0,100
0,000
171
173
175
177
179
181
183
185
187
191
HMS7
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

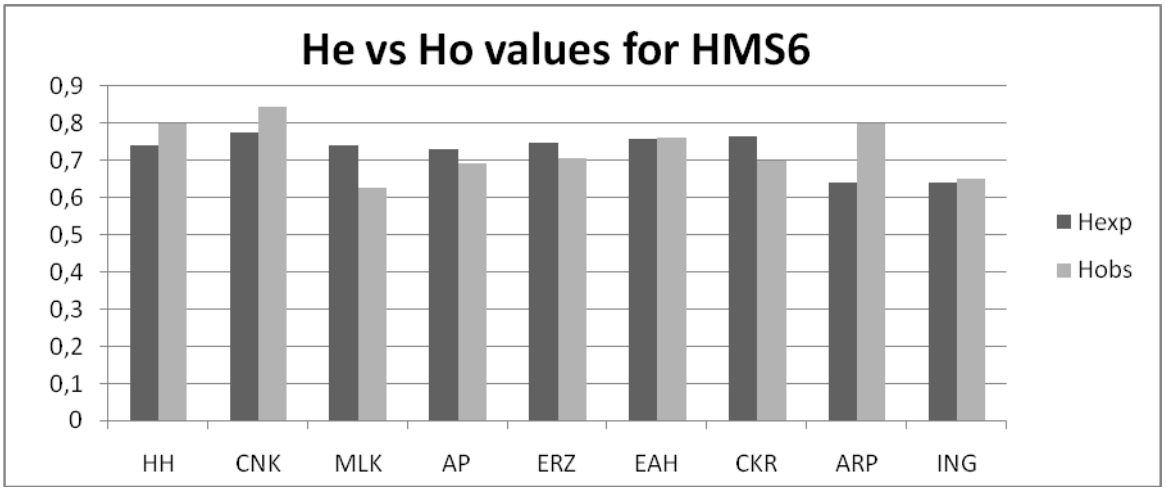

Allele Frequency for COR58
Frequency
0,400
0,350
0,300
0,250
0,200
0,150
0,100
0,050
0,000
203
205
207
209
211
213
215
217
219
221
223
225
227
COR58
Locus
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING

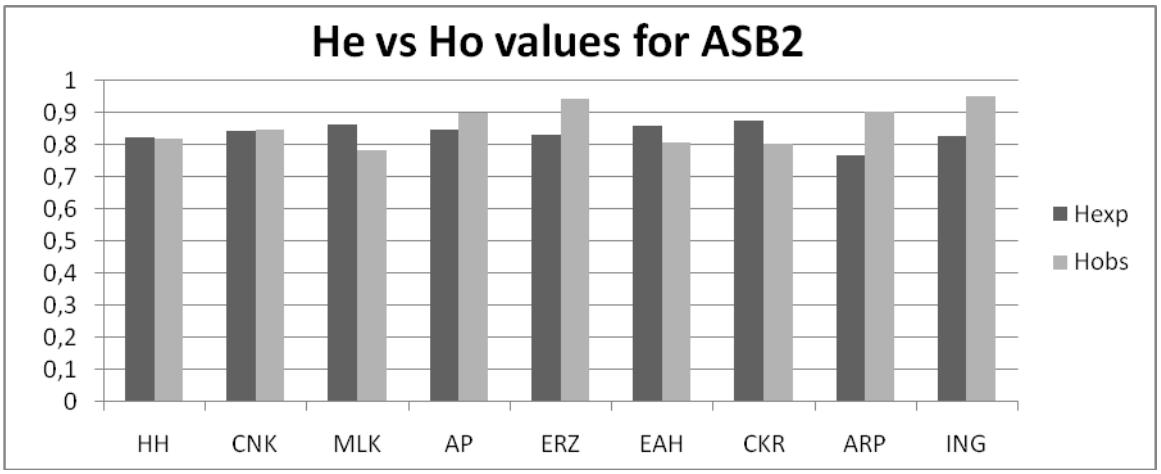
APPENDIX C: EXPECTED VS OBSERVED HETEROZYGOSITY VALUES FOR EACH LOCUS

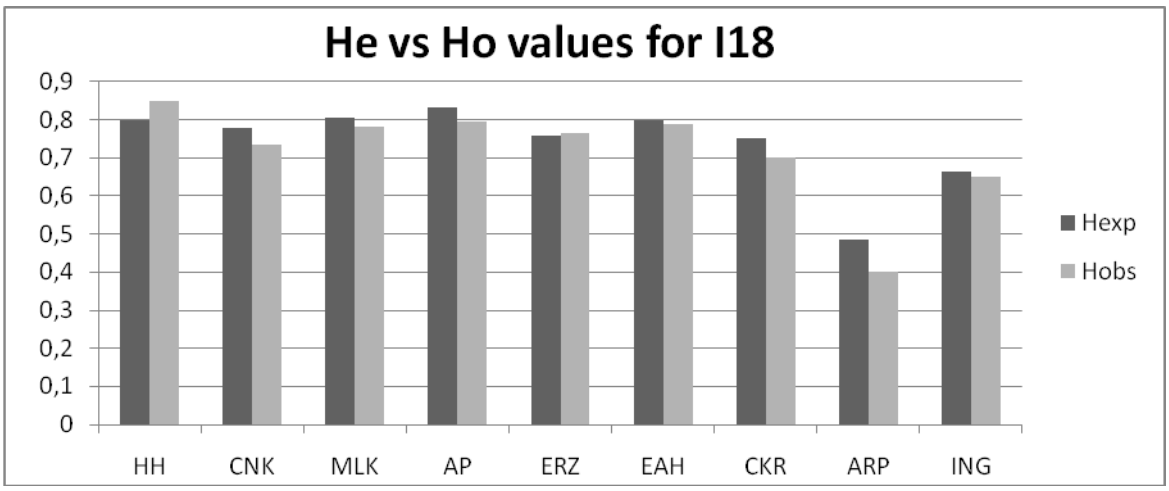


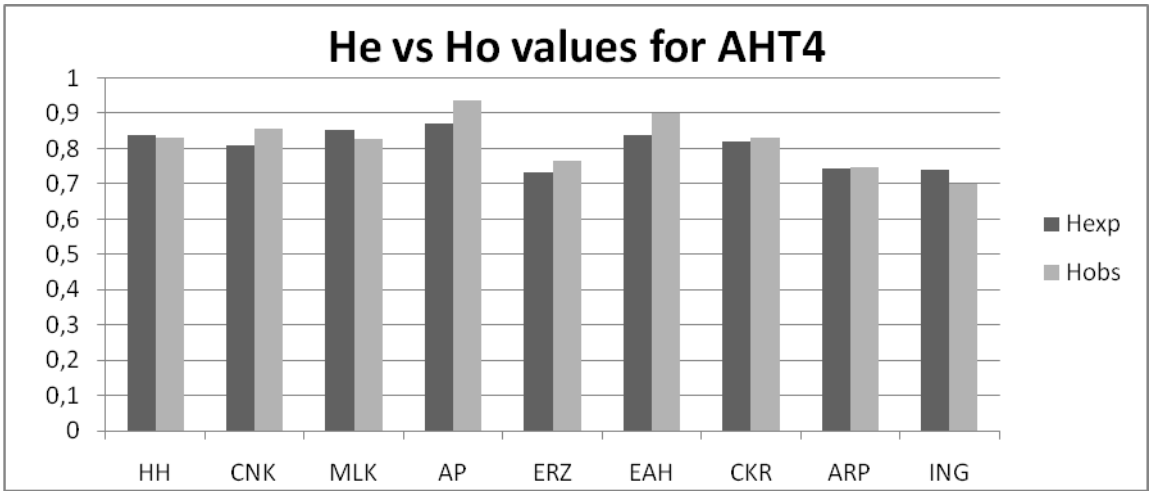


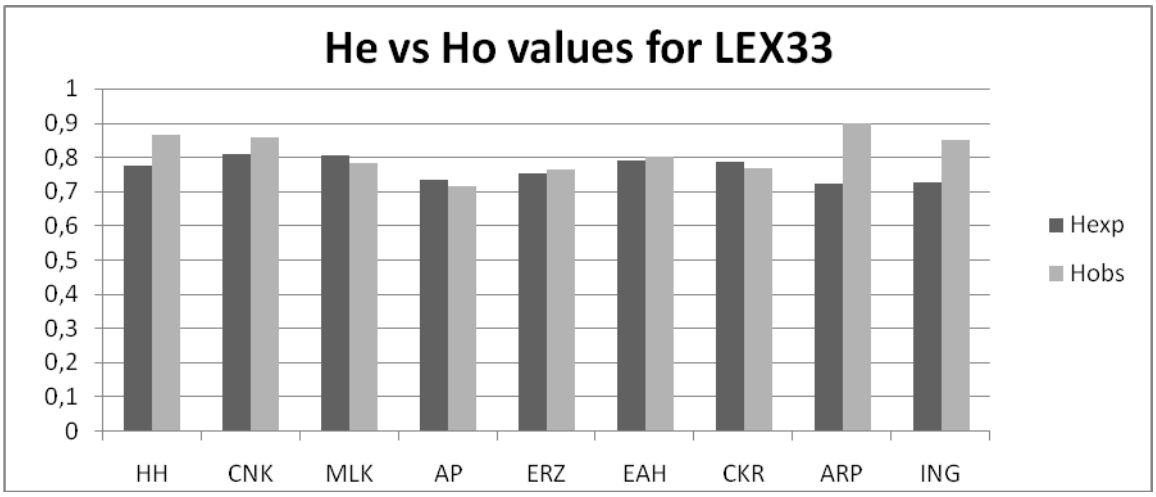


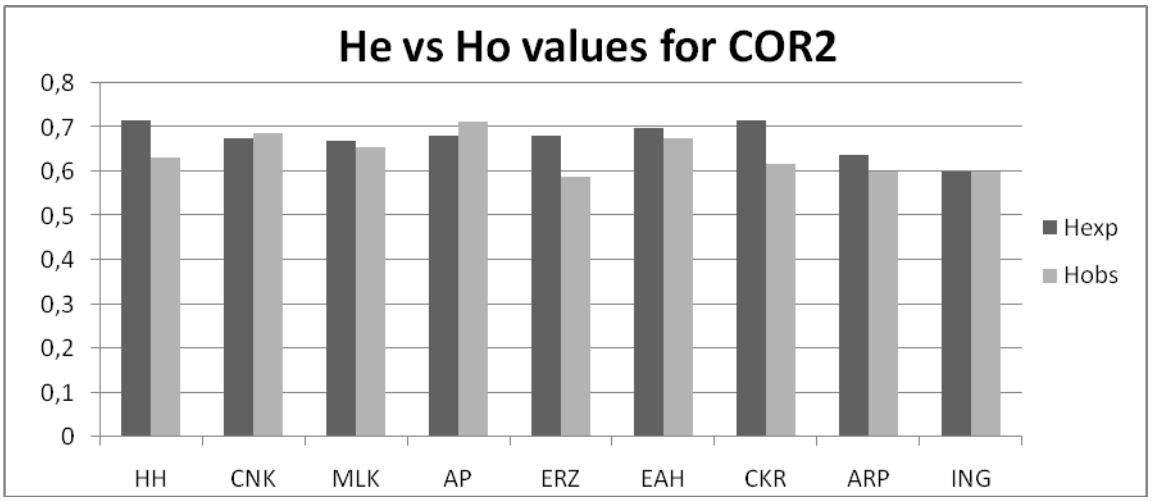

He vs Ho values for HMS6
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

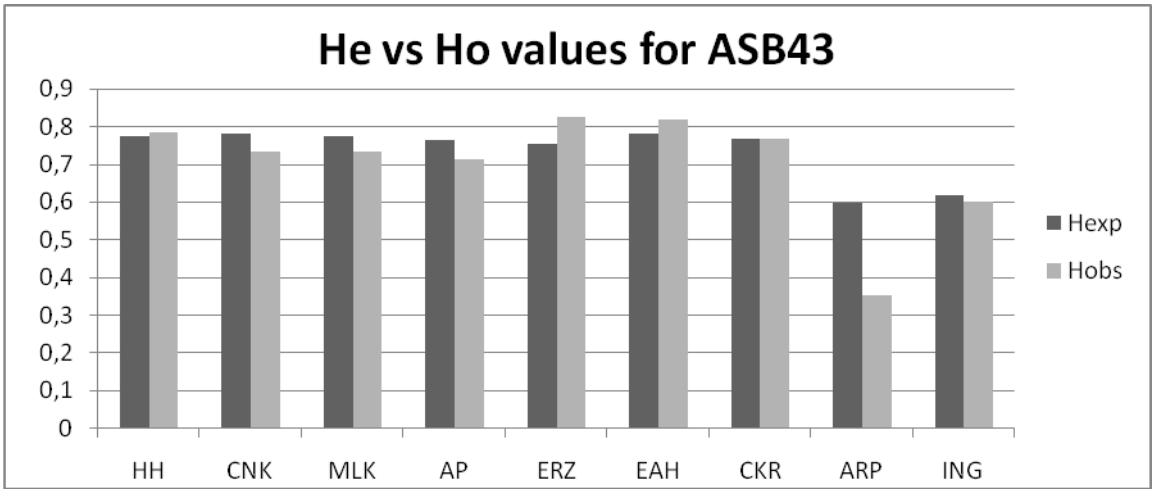

He vs Ho values for ASB2
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

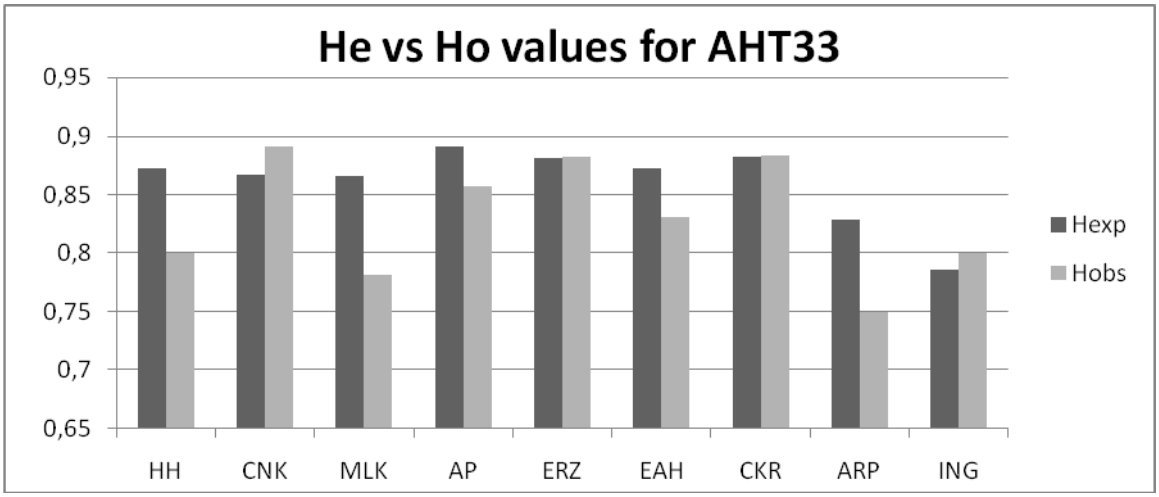

He vs Ho values for I18
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

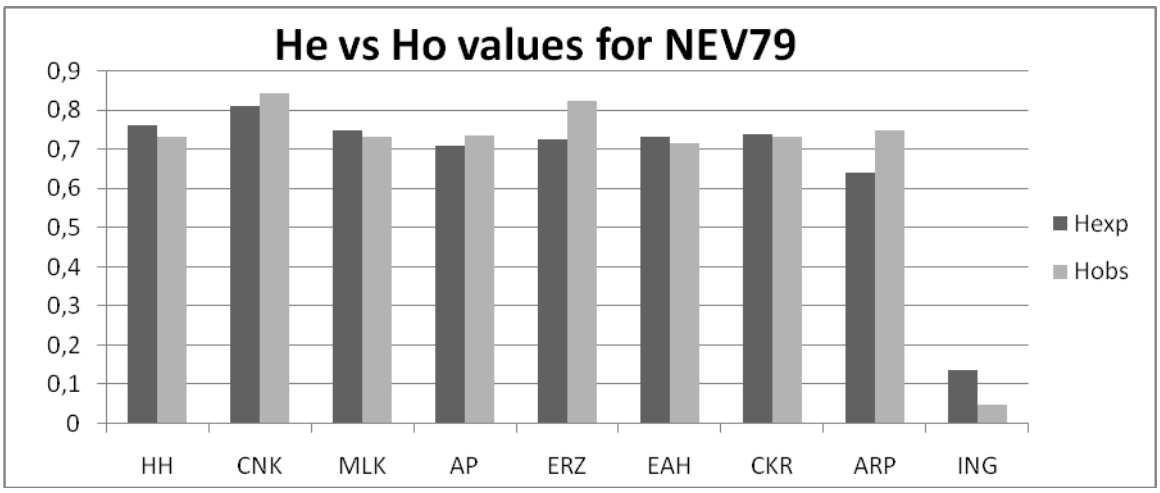

He vs Ho values for AHT4
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

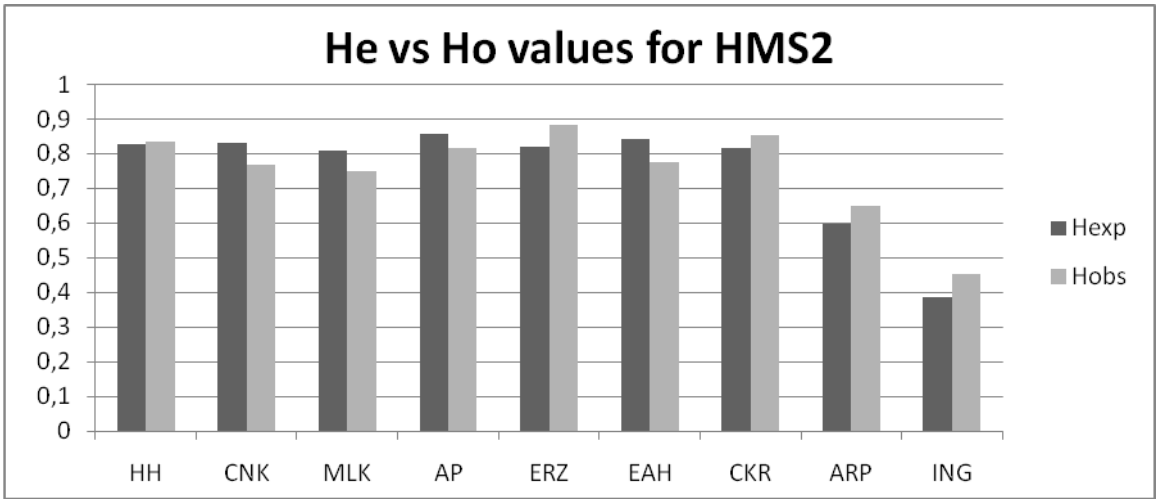

He vs Ho values for LEX33
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

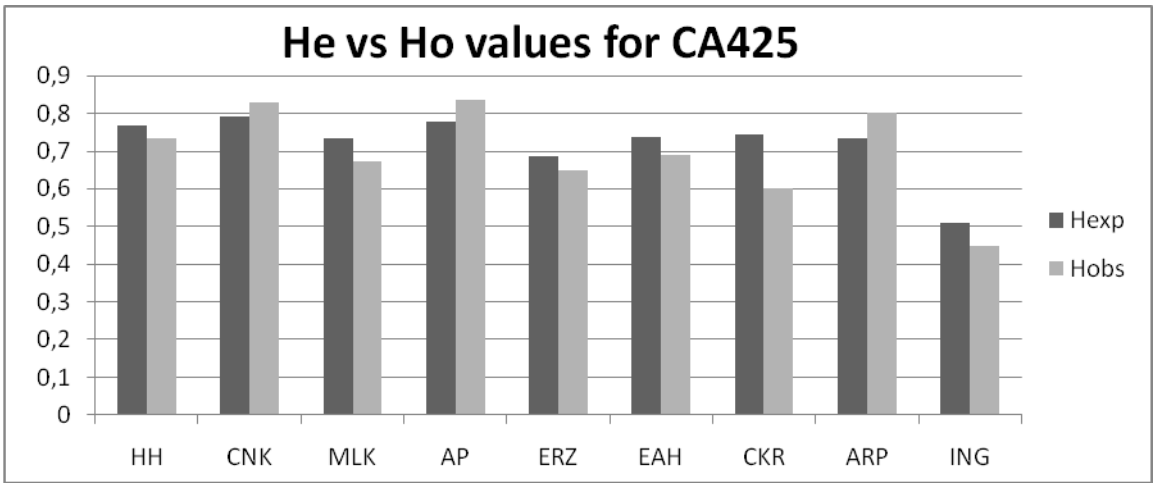

He vs Ho values for COR2
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

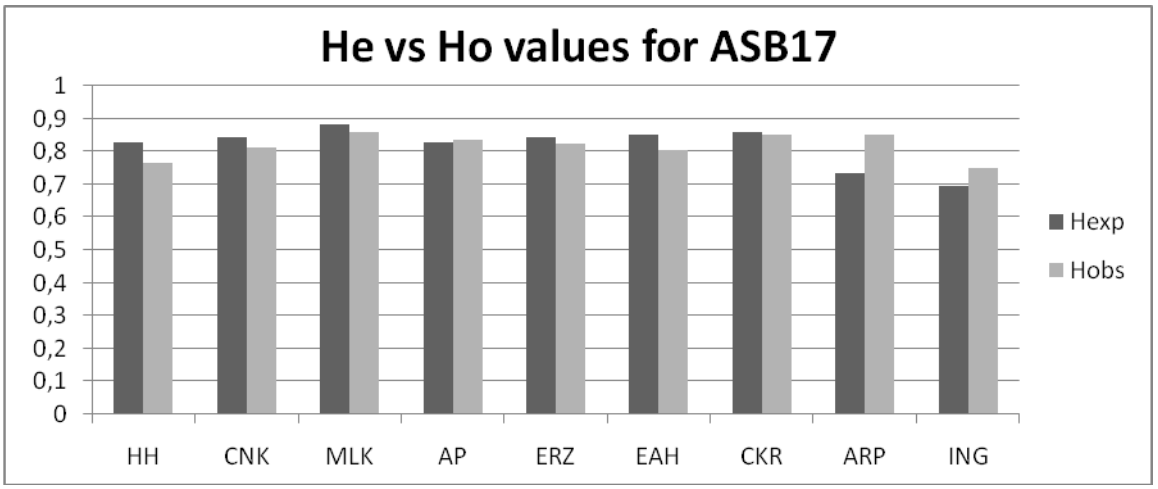

He vs Ho values for ASB43
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

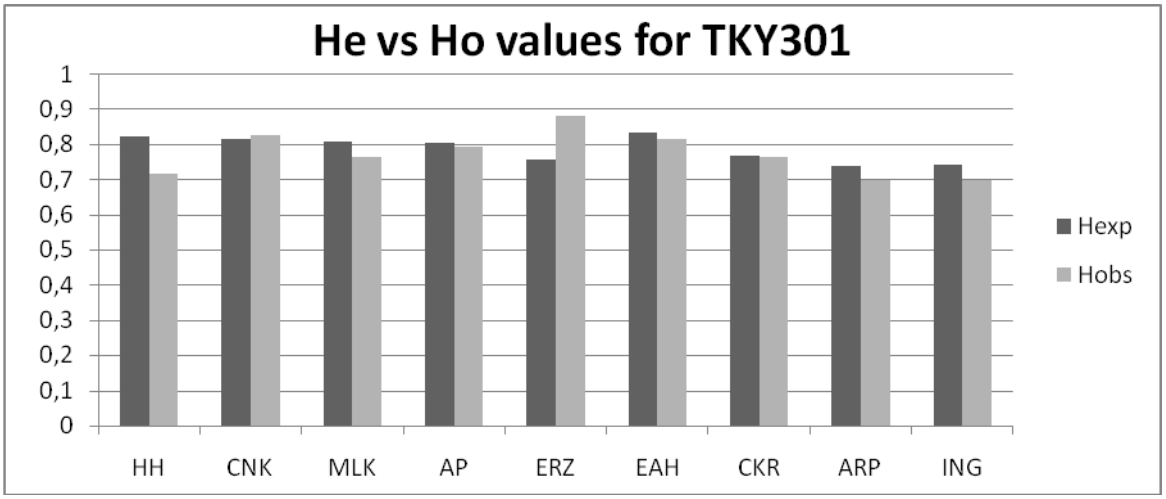

He vs Ho values for AHT33
0,95
0,9
0,85
0,8
0,75
0,7
0,65
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

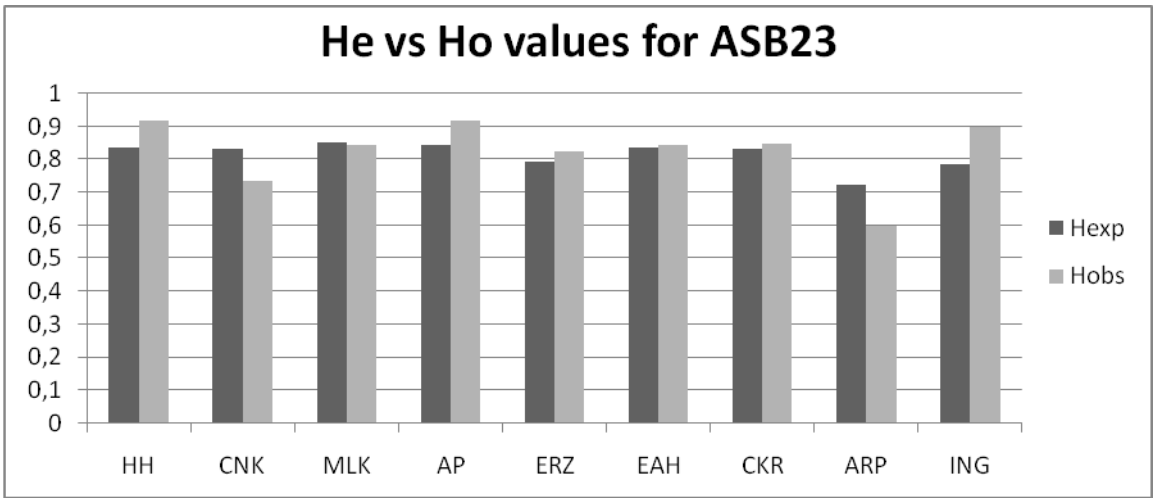

He vs Ho values for NEV79
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

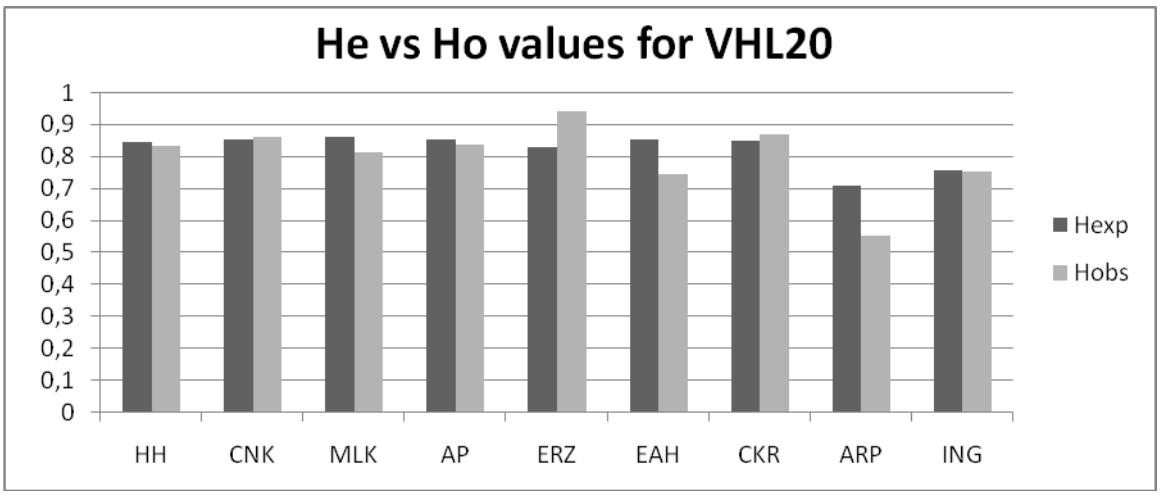

He vs Ho values for HMS2
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

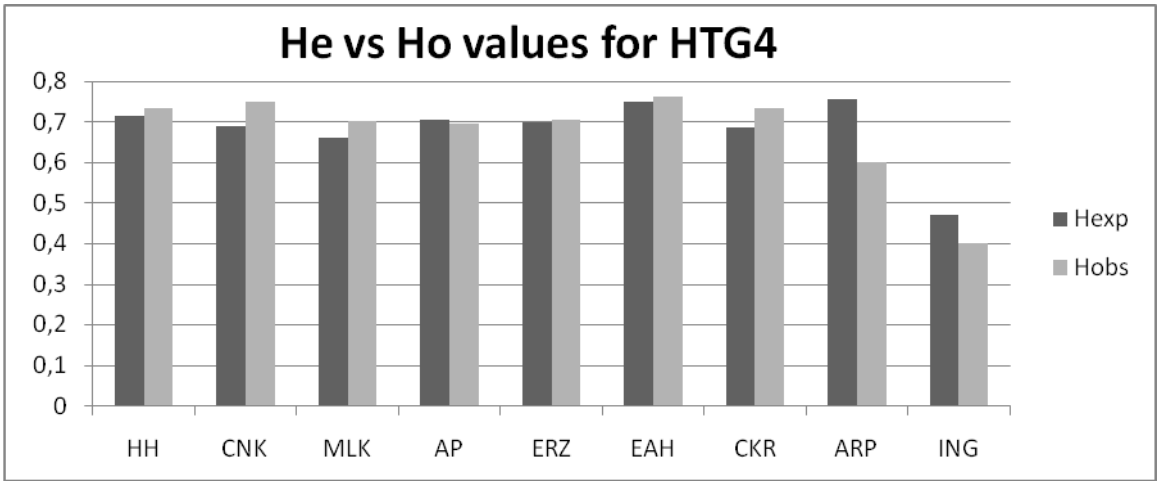

He vs Ho values for CA425
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

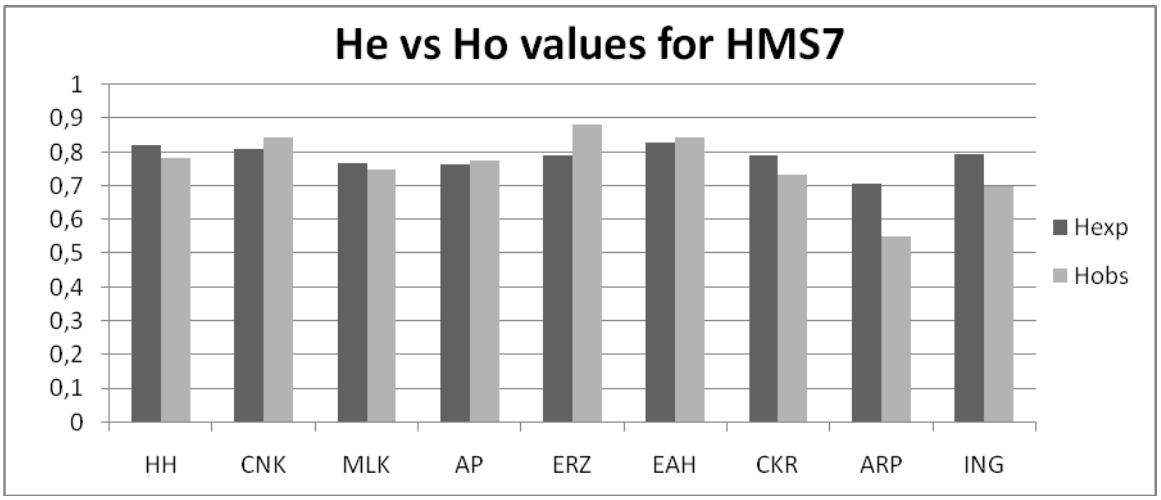

He vs Ho values for ASB17
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

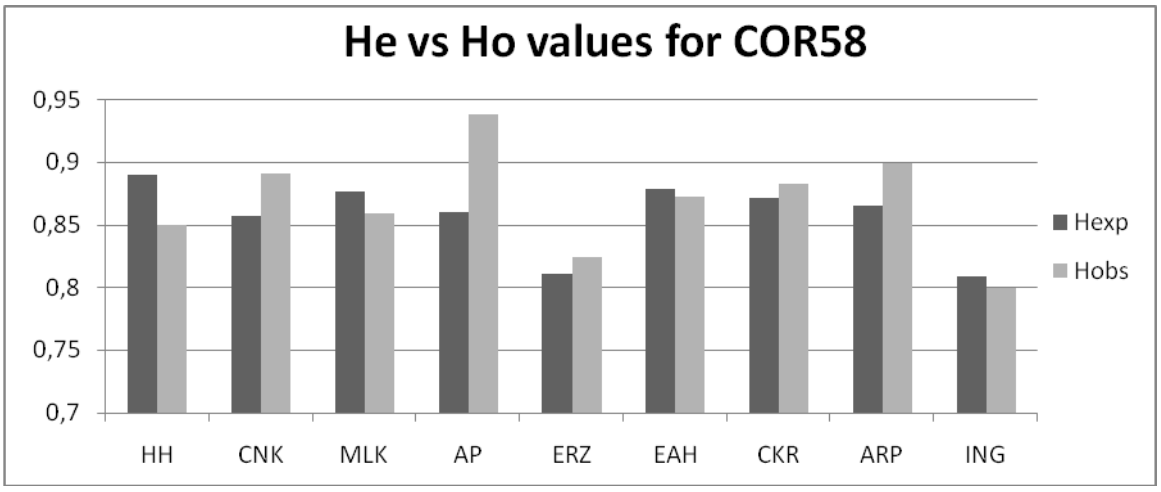

He vs Ho values for TKY301
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs


He vs Ho values for ASB23
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

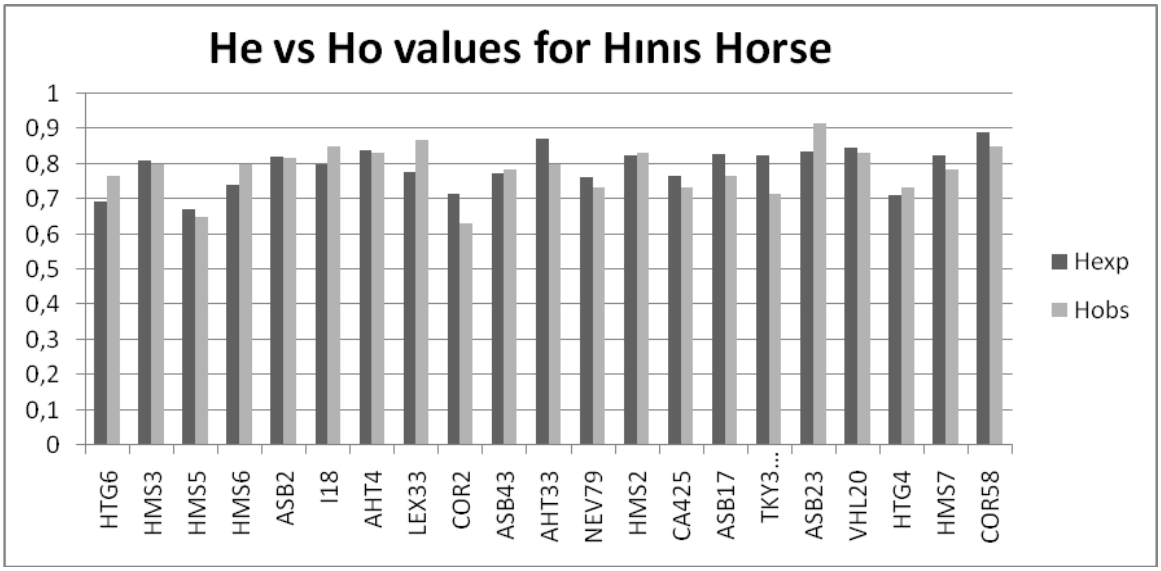

He vs Ho values for VHL20
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

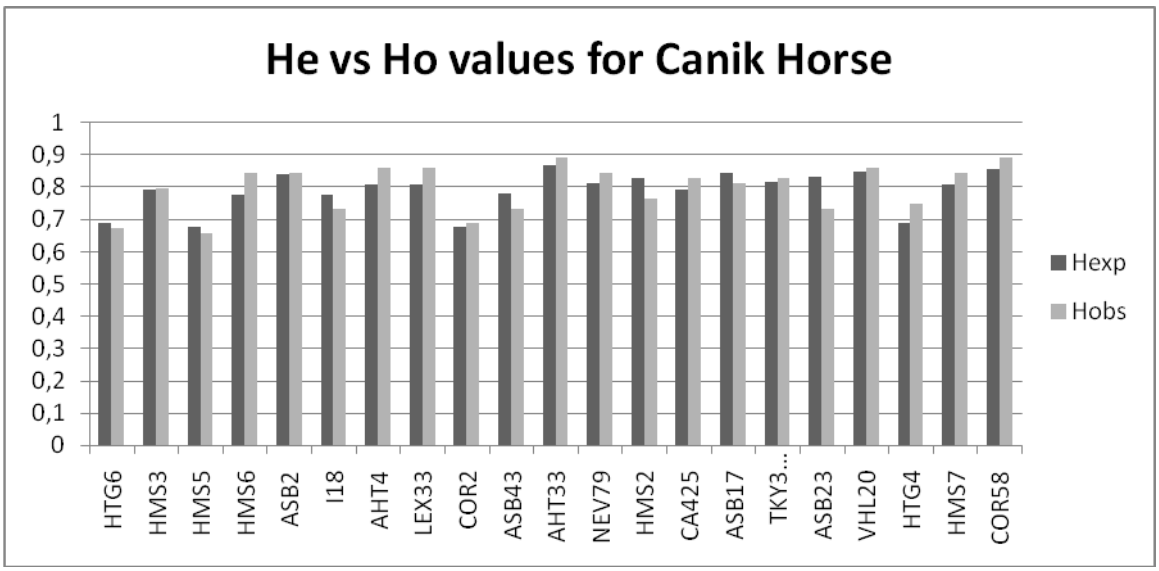

He vs Ho values for HTG4
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

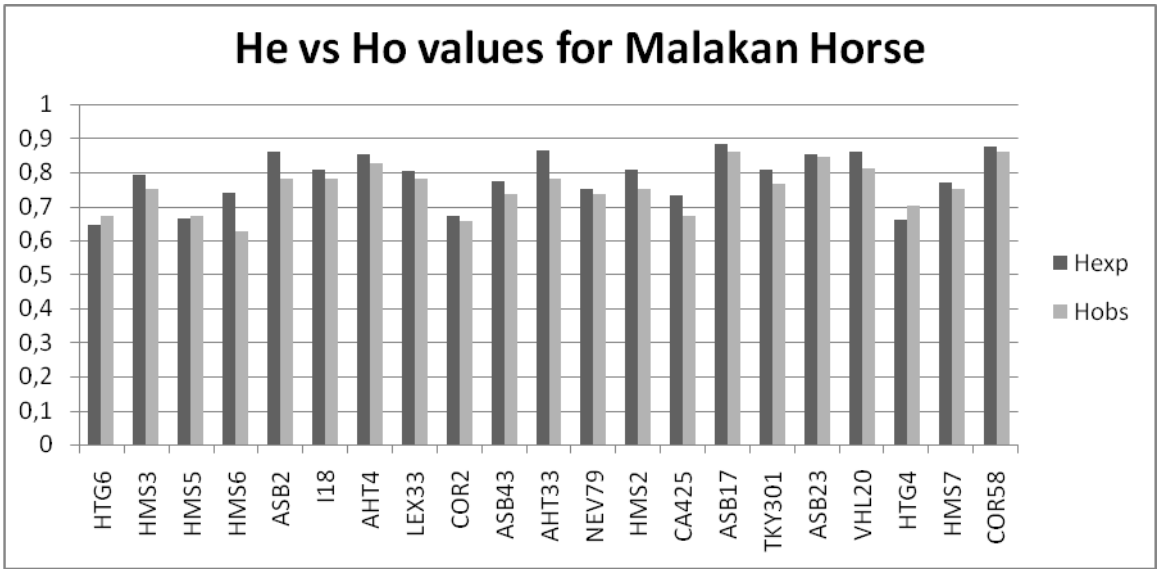

He vs Ho values for HMS7
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

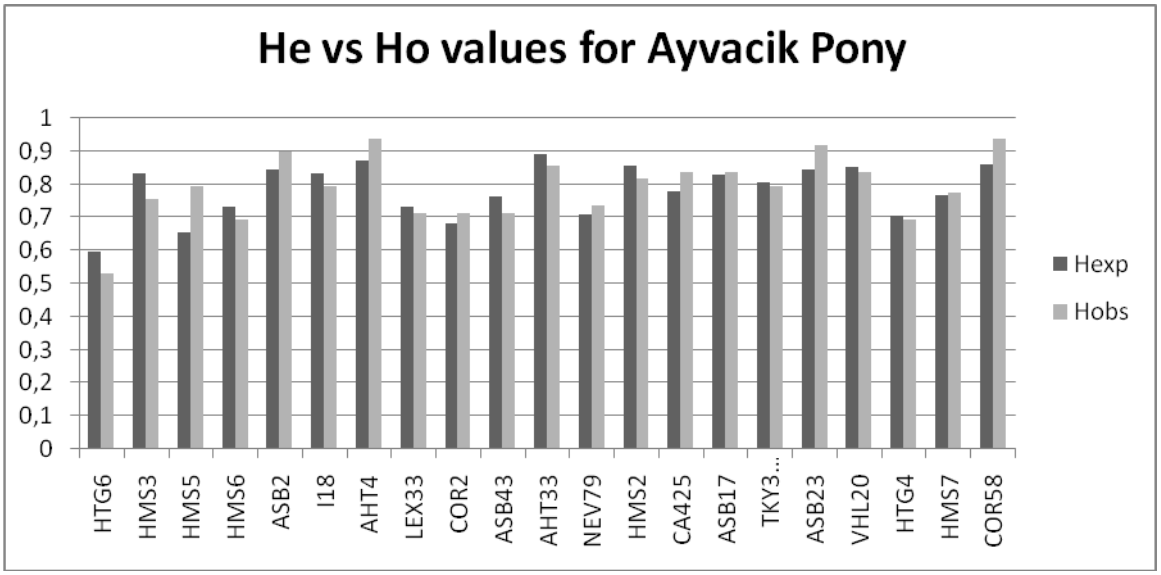

He vs Ho values for COR58
0,95
0,9
0,85
0,8
0,75
0,7
HH
CNK
MLK
AP
ERZ
EAH
CKR
ARP
ING
Hexp
Hobs

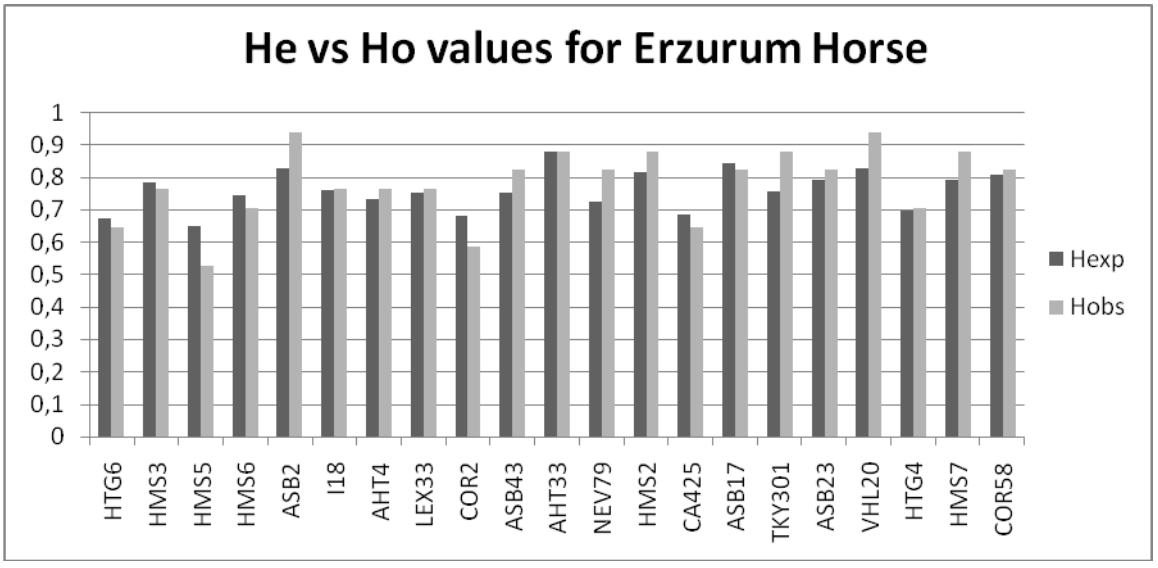
APPENDIX D: EXPECTED VS OBSERVED HETEROZYGOSITY VALUES FOR EACH POPULATION

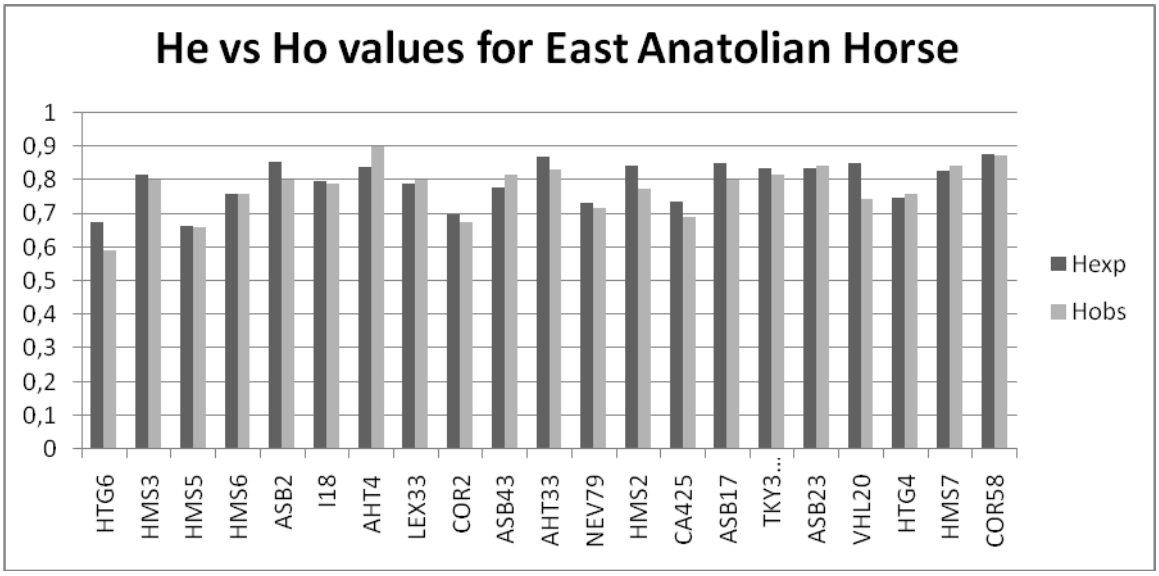


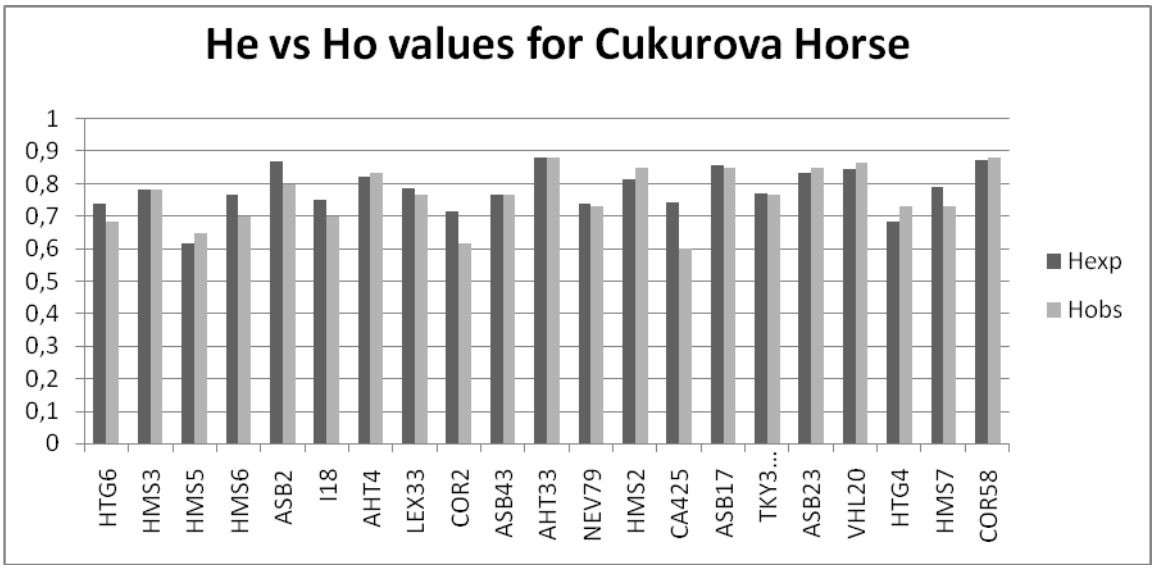


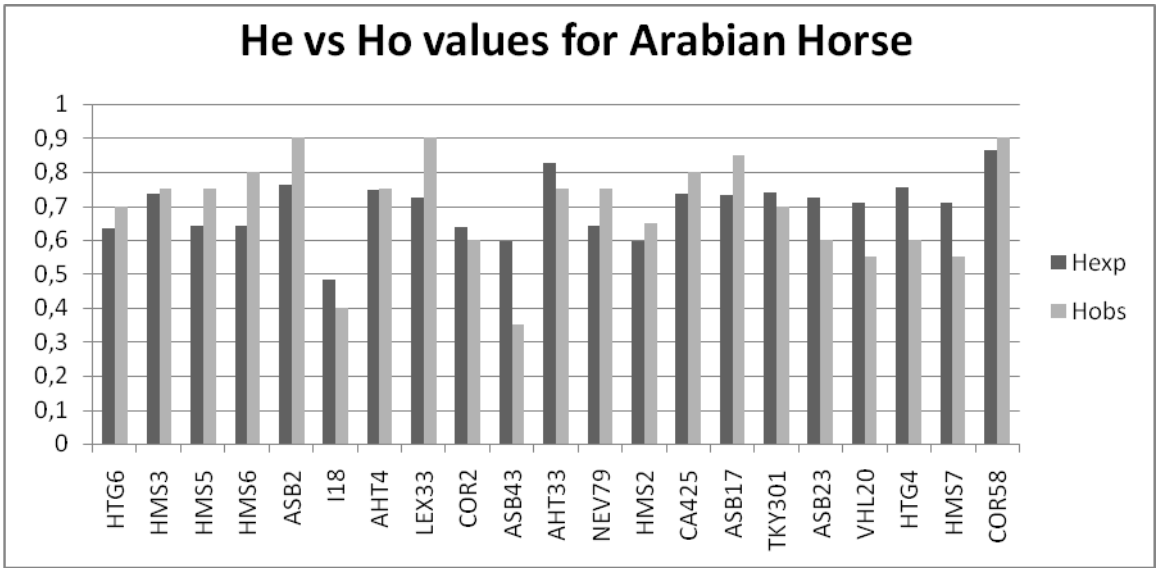

He vs Ho values for Malakan Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY301
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs

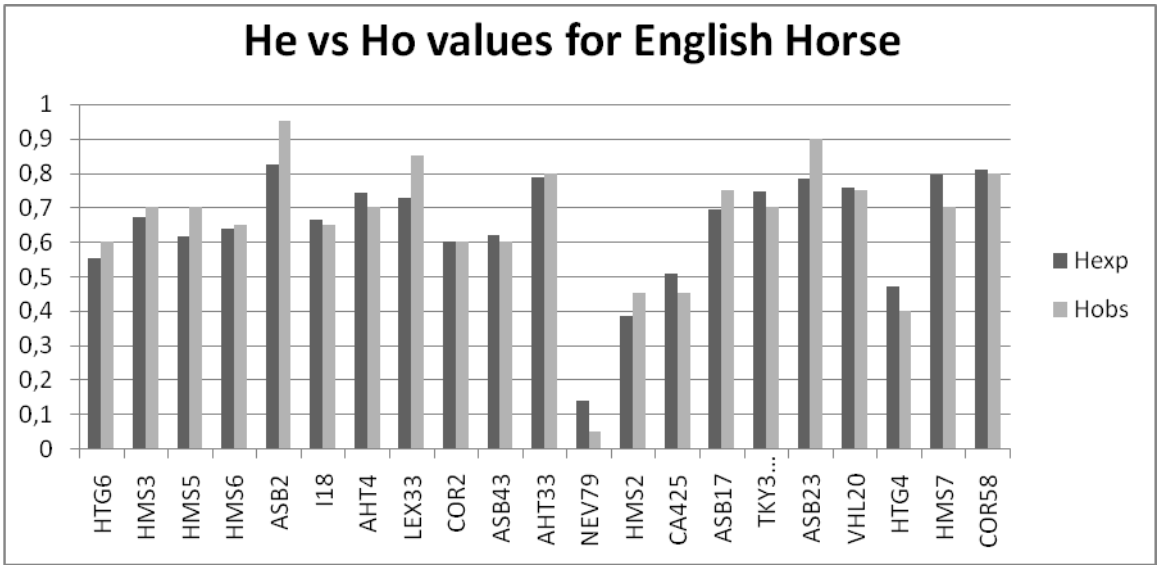

He vs Ho values for Ayvacik Pony
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY3...
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs


He vs Ho values for Erzurum Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY301
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs


He vs Ho values for East Anatolian Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY3...
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs


He vs Ho values for Cukurova Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY3...
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs


He vs Ho values for Arabian Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY301
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs


He vs Ho values for English Horse
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
HTG6
HMS3
HMS5
HMS6
ASB2
I18
AHT4
LEX33
COR2
ASB43
AHT33
NEV79
HMS2
CA425
ASB17
TKY3...
ASB23
VHL20
HTG4
HMS7
COR58
Hexp
Hobs