INTEGRATED HIGH-RESOLUTION PHYSICAL AND COMPARATIVE GENE MAPS IN HORSES

A Dissertation

by

CANDICE LEA BRINKMEYER LANGFORD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2006

Major Subject: Genetics

INTEGRATED HIGH-RESOLUTION PHYSICAL AND COMPARATIVE GENE MAPS IN HORSES

A Dissertation

by

CANDICE LEA BRINKMEYER LANGFORD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Bhanu P. Chowdhary
Committee Members,	James E. Womack
	Loren C. Skow
	Terje Raudsepp
	William J. Murphy
Chair of Genetics Faculty,	James Wild

December 2006

Major Subject: Genetics

ABSTRACT

Integrated High-Resolution Physical and Comparative

Gene Maps in Horses. (December 2006)

Candice Lea Brinkmeyer Langford, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Bhanu Chowdhary

High-resolution physically ordered gene maps for the horse (*Equus caballus*, ECA) are essential to the identification of genes associated with hereditary diseases and traits of interest like fertility, coat color, and disease resistance or susceptibility. Such maps also serve as foundations for genome comparisons across species and form the basis to study chromosome evolution. In this study seven equine chromosomes (ECA6, 7, 10, 15, 18, 21 and X) corresponding to human chromosomes (HSA) 2, 19 and X were selected for high-resolution mapping on the basis of their potential involvement in diseases and conditions of importance to horses. To accomplish this, gene- and sequence-specific markers were generated and genotyped on the TAMU 5000_{rad} horse x hamster RH panel. Additionally, screening of a BAC library by overgoes and subsequent STS content mapping and fingerprinting approaches were used to assemble and verify a BAC contig along a ~5 Mb span on ECA21.

Dense gene maps were generated for each of the seven equine chromosomes by adding 408 new markers (285 type I and 123 type II) to the current maps of these chromosomes, thereby greatly improving overall map resolution to one mapped marker every 960kb on average (range: 700 kb – 1.3 Mb). Moreover, the contig on ECA21

contained 47 markers (42 genes and 5 microsatellites) as well as 106 STS markers distributed along 207 BAC clones. Comparisons of these maps with other species revealed a remarkably high level of horse-human X chromosome conservation, as well as two evolutionary breakpoints unique to Perissodactyls or Equids for the equine homologues of HSA19 and HSA2, one of which has been more precisely localized by the ECA21 contig. Thus, high resolution maps developed for these chromosomes i) provide a basis to map traits of interest rapidly to specific chromosomal regions, ii) facilitate searches for candidate genes for these traits by fine comparisons of the equine regions with corresponding segments in other species, and iii) enable understanding the evolution of the chromosomes. Expansion of this work to the entire equine genome will be important for developing novel strategies for diagnosis, prevention, and treatment of equine diseases.

DEDICATION

To the glory of God, through whom all things are possible.

I asked God for strength that I might achieve; I was made weak that I might learn humbly to obey.
I asked God for health that I might do greater things; I was given infirmity that I might do better things.
I asked for riches that I might be happy; I was given poverty that I might be wise.
I asked for power that I might have the praise of men; I was given weakness that I might feel the need of God.
I asked for all things that I might enjoy life; I was given life that I might enjoy all things.
I got nothing that I asked for – but everything I had hoped for…
Almost despite myself my unspoken prayers were answered.
I am among all men most richly blessed.
-- attributed to an unknown Confederate soldier

ACKNOWLEDGEMENTS

I am overwhelmingly fortunate and grateful for a number of people in my life who have both led and encouraged me throughout my graduate education. Firstly, the members of my committee have provided priceless guidance and support. My advisor, Dr. Bhanu Chowdhary, has been both mentor and friend. His patience and confidence in me have been unparalleled; his heartfelt kindness, without bounds. His contributions to my education and life constitute far more than a mere "drop in the bucket" (as he says). Next, Dr. Loren Skow has been very generous in sharing his lab space, cabin, and wisdom. Though a bit timid of him at first, I have come to know him as a kindhearted teacher and buddy. Dr. Terje Raudsepp has also been immensely benevolent in her support, guidance, and helpfulness. I feel very lucky to have had the opportunity to work with Dr. James Womack and Dr. William Murphy, as each of them is an exceptional and internationally known researcher, and they have provided valuable advice and assistance.

Of great importance to me are the contributions of those who have been co-authors or otherwise played a role in the works reported in this dissertation. Drs. Mi-Kyung Lee and Hongbin Zhang provided aid in fingerprinting for the ECA21 contig; Dr. Ashley Gustafson-Seabury taught me all about the magic of overgoes; Dr. Chris Seabury and Keith Durkin supplied help for sequencing; Drs. Michele Wagner and Jim Mickelson provided microsatellite markers for every study and graciously let me contribute to the HSA2 paper; Dr. Eun-Joon Lee taught me the art of primer design and genotyping; Drs. Alejandro Schäffer and Richa Agarwala kindly demonstrated their processes for me and have done great work in generating the radiation hybrid maps, as did Glenda Goh and

Srinivas Kata. Also, those members of my committee who have provided assistance or been co-authors on these papers are gratefully acknowledged.

My fellow lab-mates in both the Chowdhary and Skow labs have been wonderful coaches and friends. They have been great for discussions, commiserating, support, mischief, and laughter. Thank you and big hugs go to Krista Fritz (my "little sister"), Chris Childers, Gayle Linger, Marian Cothran, Dr. Avni Santani, Glenda Goh, Nandina Paria, Yujane Wu, Dr. Eun-Joon Lee (and Hyun-Ju), Dr. Deeann Wallis, and Jan Elliot. For all that you have shared with me, I am forever indebted; the memories will last a lifetime.

I am quite grateful to the College of Veterinary Medicine for the fellowship awarded for my first two years of graduate study, the Association of Former Students for the Academic Excellence Award scholarship, and the Genetics Graduate Program. The research reported in this dissertation was funded by grants from the Texas Higher Education Board (ARP 010366-0162-2001 BPC; ATP 000517-0306-2003 BPC, JEW), NRICGP/USDA Grant 2003-03687 (BPC), NRICGP_USDA Grant 2000-03510 (to L.C.S.), the Morris Animal Foundation D02EQ-21 (JRM), Texas Equine Research Foundation (BPC, LCS), Link Endowment (BPC, LCS), American Quarter Horse Association, the Dorothy Russell Havemeyer Foundation, and available from the USDA-NRSP-8 Coordinators' Fund.

My family has given me tremendous strength and encouragement, for which I am eternally thankful. My husband Mikey is nothing short of a saint; his tolerance and support are far beyond that of any mere mortal and he has reassured me countless times when I felt afraid or weak. My profound gratitude goes to him for doing much of the

cooking and always accompanying me to my countless appointments. I cannot express in words how much his love and friendship have made an impact on my life. My parents, Mike and Bonnie Brinkmeyer, brought me up believing that I am completely capable of doing anything and were unwavering in their confidence in me. Both have been excellent role models for me: my dad in his creativity, unique genius, and heroism in the MS150, and my mom with her own determination in pursuing her goals – her mantra "There are no problems, only challenges" has come to mind numerous times during my graduate career. All four of my grandparents (Granny and Pop, and Nana and Grandpa Tom) have shown bravery and endurance, as well as wisdom and insightfulness, which have motivated me my entire life. My brother Wesley and sister-in-law Aubrey have given me many reasons to smile and give me at least one good, long laughing fit every time I see them. All of my aunts, uncles, and cousins have so sweetly given me encouragement in their own ways, with cards, notes, and their beautiful smiles. I love you all lots and lots!!

The members of the MS150 bike team are my heroes. Their physical exertion and perseverance on behalf of me and others with multiple sclerosis have inspired and sustained me in more ways than one. Special thanks go to my aunt and uncle, Lori and Josh Benn, for organizing the team.

My heartfelt appreciation also extends to Dr. Victor Rivera and the rest of the group at the Maxine Mesinger MS Clinic in Houston, for tirelessly answering most of my endless questions. A big thank-you goes to Dr. Rivera for laughing whenever I asked him whether I should or shouldn't do something, knowing it wasn't going to make a difference to me anyway.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

I. INTRODUCTION

The horse-human relationship

The domestication of the horse

The paths of man and horse converged some 50,000 years ago when Cro-Magnon man learned to hunt and kill horses, sometimes by herding them into inescapable situations (Levine 1999). Ancient cave paintings from that time suggest that humans had a certain degree of reverence for the horse and its beauty, even in these early days. Archeological evidence suggests that by around 4000-3000 BC man living on the steppes of present-day Europe and Asia discovered how to tame or domesticate horses (Levine 1999; Bowling and Ruvinsky 2000), though exactly how and when this took place is unclear.

While other domesticated livestock such as cattle, sheep, and pigs can trace their roots to a few founding populations, suggesting that their domestication occurred in only a handful of locations (Vilà et al. 2001), mitochondrial DNA (mtDNA) from a wide variety of horse breeds tells a story that is somewhat different. Vilà et al. (2001) examined and compared the maternally inherited genomes of horses from a wide variety of modern breeds with mtDNA extracted from the preserved remains of ancient horses from an Alaskan location dated 12,000 to 28,000 years ago. Unlike the other domesticated livestock species, horses show an abundance of diversity in their mtDNA, indicating that their maternal lineages are widespread and that a number of separate

---

This dissertation follows the style and format of Mammalian Genome.

domestication events must have occurred.

In contrast to the relatively calm and good-natured horses bred in modern times, the wild Przewalski's horses and free range mustangs found today are known to be particularly difficult to tame or domesticate. This leads to the postulation that wild horses in ancient times were similarly obdurate, and individuals were selected for taming and breeding on the basis of their gentleness and trainability (Jansen et al. 2002) – some of the earliest examples of humans inadvertently exploiting the genetics of horses through selective breeding.

Usefulness of the horse for humans

Once the horse was domesticated, it was subsequently adapted for a variety of uses. The utility of horses as beasts of burden was probably among the first to be realized, with horses put to work pulling and carrying heavy loads for – and including – humans (Levine 1999). At some point, humans discovered that these strong animals could be ridden for herding, hunting, or transportation. Also, agricultural functions were developed for the horse, many of which are employed to this day in some parts of the world. Each of the services performed by the horse for man required certain characteristics, such as strength and manageability – features for which horses were initially selected and bred. As a result, human emphasis shifted to selective mating to obtain horses with desired characteristics. This can be considered as the beginning of the application of basic 'genetic' approaches to breed horses. Eventually man observed that certain conditions and diseases could also be passed down through the generations, and

came to use this to his advantage as well, carefully breeding horses for selected qualities and health.

Breeding resulting from increased understanding of the potential inherited causes behind certain traits and diseases is becoming all the more important as time has progressed. In today's industrialized nations, where electrical and mechanical devices such as automobiles have usurped the role of beast of burden, horses are primarily used in different contexts such as racing, showing, and recreation – a role that has increased exponentially during the past couple of decades. Because of these developments, the need for improved breeding programs and increased knowledge of diseases and conditions of interest to the horse community has continued to grow along with these changes.

Economic and social impacts of the horse

Statistics published by the American Horse Council in 2005 highlight the relevance of the horse to modern society in the United States. Currently the horse industry contributes over $101 billion to the economy annually and involves around 4.6 million people nationwide (http://www.horsecouncil.org/statistics.htm). These numbers clearly reflect a prominent role for the horse in the lives of many Americans. In addition to its economic influence, the horse has an invaluable effect on the mental and physical well-being of countless numbers of people. Many people find horses to be excellent companion animals and the therapeutic benefits of riding or simply sharing the company of a horse are recognized (Henricksen 1971; Potter et al. 1994; North American Riding for the Handicapped Association, Inc., http://www.nahra.org). Additionally, horses

contribute to the physical safety of people through their services in police forces (http://www.mountedpolice.com/) and even in guiding the blind (http://www.guidehorse.org/).

Throughout history, horses and humans have maintained a relationship that has proven beneficial through economic progress and enhanced well-being for both. As a result, mankind is obligated to strive for the advancement of horse welfare by continuing to seek ways to improve equine medicine, along with advancement in management, nutrition, breeding and selection methods. Additionally, an in-depth knowledge of horse biology as a whole is essential. All this can be facilitated with a detailed understanding of the genetic constitution of the horse, and the roles, functions, and interactions of individual genetic factors (genes) in regulating various body systems, development, and disease mechanisms.

Gene maps and their significance

Mapping genes and genomes – an overview

To better understand the hereditary components of various traits in the horse and utilize the information in a way that is advantageous to the species and the people with whom they are involved, it is essential to have comprehensive knowledge of the horse genome including precise locations of genes and other elements critical for genome function. For all species where genomes have been analyzed, gene maps are the foundational tool that fulfills this primary requirement by depicting the locations of the approximately 25,000 genes in typical mammalian genomes. When sequence

information is available for several species, comparative gene maps provide additional details on the structural constitution of individual genes and the regulatory elements that govern their expression. Using the comparative information, genetic data may be extrapolated from the maps of one species and used to rapidly develop or expand gene maps in other species by identifying regions where common sets of genes are present in genomic segments of different species (conserved synteny). Comparisons made at the sequence level help to identify additional features such as evolutionarily conserved functional or regulatory elements and breakage or fusion points that are an integral part of chromosome evolution. Jointly, such information is critical to searches for genes governing traits of interest in a given species along with understanding chromosome evolution. Importantly, identification of regions that are highly conserved across mammals as well as regions that are unique to a certain species (or even individual populations of a species) is relevant to the study of mammalian evolution in general and species divergence and evolution in particular.

Progress in equine medicine, breeding programs, and advances in equine management and health care can substantially benefit from improved knowledge about the genome of the horse. Identification of causative genes and gene variants is critical, as understanding the molecular basis of a condition or disease can result in improved accuracy of diagnostic tests. This will prove useful in breeding programs, and also in determining the appropriate treatment for an affected horse. Moreover, information and additional benefits (such as improved treatment methods) provided by gene maps from other species like humans, cattle, dogs, etc. may be used in the horse for clinically

identical diseases. Detailed comparative maps can prove highly valuable in such circumstances, as details about homology and evolution are provided which can help to determine, for example, whether a species (and therefore its genetic composition) is closely related to the horse. It is therefore apparent that gene maps are of utmost importance to the future health and well-being of horses.

Gene mapping in other species

The power and utility of gene maps have already been demonstrated in a number of mammals, particularly humans, mice, rats, and various livestock and companion animal species where sequence information has become available in recent years (e.g. cattle and dogs). Genome maps and/or whole genome sequence data are publicly available for over 1000 organisms (http://www.ncbi.nlm.nih.gov); consequently, it is possible for genomic research to proceed at a rapid pace with an increasing interface with medicine. For example, causal mutations are routinely being identified for a number of diseases in humans and the frequency of such reports in domestic animals is also on the rise (described below). In recent years, comparative gene mapping data has facilitated the transfer of information from maps of one species to that of another. An excellent illustration of this is Wilson disease in humans and its analysis in dogs where it is known as copper toxicosis (Su et al. 1982; van de Sluis et al. 2002; Muller et al. 2003; Stuehler et al. 2004). Moreover, for species where sequence data is available, programs have been or are being developed to further exploit the genomic information by determining the locations and roles of functional elements which regulate gene activity, the significance of variations in splicing, non-coding RNAs, etc. Overall, gene maps

confer a number of benefits – both realized and potential – to those species where they are available.

The Human Genome Project was officially established in 1990 with the goals of identifying and studying every part of the genome and the intent to make the discoveries publicly available for novel biomedical and functional genomics research (Collins et al. 1998; 2003a). The resulting genome sequence sparked a new era in biomedical research (Collins et al. 2003b). The ever-increasing volume of information collected and disseminated is resulting in the discovery and exploration of genes associated with thousands of diseases and conditions, with nearly 17,000 diseases listed in the database Online Mendelian Inheritance in Man (OMIM; http://www.ncbi.nlm.nih.gov/Omim.html). Public availability of the genome sequence has encouraged subsequent global research projects such as the International HapMap Project (The International HapMap Consortium, 2003) and Encyclopedia of DNA Elements (ENCODE; The ENCODE Project Consortium, 2004), which have been launched with goals of identification and characterization of elements of the human genome which contribute to diseases or traits. Ultimately such by-products of the Human Genome Project will enable us to understand the regulation of causative genes. The latter are presently prime targets for the development of novel, more effective drugs that render promise for refinement of treatment methods.

The success of the Human Genome Project formed a foundation upon which genome projects were established for a number of species, including important agricultural and companion species such as cattle, chicken, pig, and dog (discussed in

ensuing paragraphs). Additionally, medium- to high-resolution gene maps are presently available for a wide range of livestock species such as sheep, goat, turkey, and buffalo (Womack 2005). These maps have led to the identification of genes and mutations responsible for several traits of interest in some of the species studied in greater detail. The increasing amount of genome information available for these species may be accessed in the form of databases such as the Ensembl (http://www.ensembl.org), University of California Santa Cruz (http://genome.ucsc.edu/), and National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) genome browsers.

A plethora of resources has become available in other species as well. Linkage maps marked the beginning of the development of dense gene maps in the chicken and this eventually led to integration of different genomic resources (Groenen et al 2000; Groenen and Crooijmans 2003; Aerts et al. 2003) that provided the foundation for the complete sequence of the chicken genome. Other tools such as a 6000_{rad} radiation hybrid panel have been used for integrating linkage and cytogenetic information and to align the maps to the sequence information (Morisson et al. 2002; Leroux et al. 2005; Morisson et al. 2005). Several BAC libraries including the TAM31, TAM32, and TAM33 (Lee et al. 2003b) and CHORI-261 (http://bacpac.chori.org/chicken261.htm) libraries have been generated. Of these, the three TAM libraries were used to construct the first BAC-based map of the complete chicken genome (Ren et al. 2003). The next year, a clone-based physical map of the genome providing a 20X coverage enabled anchoring of draft sequence information obtained for the entire chicken genome (Wallis et al. 2004; International Chicken Genome Sequencing Consortium 2004). Along with the

sequencing of the genome, a map showing nearly 2.8 million SNPs has been published for the chicken genome (Wong et al. 2004; International Chicken Polymorphism Map Consortium), which is anticipated to prove useful for investigating genetic traits such as QTLs (Gao et al. 2006; Rowe et al. 2006). Gene expression analysis has also become possible for the chicken in recent years with the availability of an extensive collection of chicken ESTs and cDNA clones (Boardman et al. 2002; http://www.chick.umist.ac.uk/) used to generate cDNA microarrays (Burnside et al. 2005; Bourneuf et al. 2006; Smith et al. 2006).

Although linkage (e.g., Barendse et al. 1994) and syntenic maps (see Womack 2005) have been key to the expansion of genomic information in cattle, radiation hybrid maps have expanded rapidly in the recent past. Most recently, a third-generation radiation hybrid map of the entire cattle genome was published with 3,484 markers (including ~3,000 BAC end sequences) mapped using the Illinois-Texas $5{,}000_{rad}$ RH panel (Everts-van der Wind et al. 2004, 2005). Another RH map was constructed that same year using a $7{,}000_{rad}$ RH panel ($SUNbRH_{7000\text{-}rad}$) to map 5,593 markers (Itoh et al. 2005). A $12{,}000_{rad}$ RH panel has also been established for cattle (Rexroad et al. 2000). In addition, the three BAC libraries available for cattle (RPCI-42, http://bacpac.chori.org /mbovine42.htm; CHORI-241, http://bacpac.chori.org/bovine240.htm; and TAMBT, Cai et al. 1995) have been used to generate a 294,651 whole clone HindIII fingerprint map of the entire genome (http://www.bcgsc.ca/lab/mapping/bovine). Altogether, these resources have aided in the assembly of the cattle genome sequence

(http://www.hgsc.bcm.tmc.edu/projects/bovine/), which presently gives an ~7.1X coverage of the bovine genome (Btau_3.1; ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20060815-freeze). Additional sequence information collected from a variety of cattle breeds has enabled detection of SNPs (see ttp://www.ncbi.nlm.nih.gov/SNP/snp_batchSearch.cgi?org=9913&type=SNP; http://www.animalgenome.org/bioinfo/resources/util/q_bovsnp.html). Moreover, cattle EST projects have proven to be a powerful resource for bovine research (http://racerx00.tamu.edu/bovine/cattle_est_db.html), leading to the construction of cDNA microarrays such as the 18,263-element array generated by the National Bovine Functional Genomics Consortium (NBFGC; Suchyta et al. 2003), and two others by the University of Illinois with 3800 and 7872 elements, respectively (Band et al. 2002; Everts et al. 2005).

For the cat, a linkage map containing 253 markers on 34 linkage groups was published by Menotti-Raymond et al. (1999) and provided a platform for genetically mapping diseases and conditions. Since then, three generations of radiation hybrid maps have been developed for the cat genome using the $5{,}000_{rad}$ whole genome radiation hybrid panel developed by Murphy et al. (1999b). The first whole-genome radiation hybrid map for the cat included 424 type I and 176 type II markers (Murphy et al. 2000); the second generation map added 585 type I and 279 type II markers to this (Menotti-Raymond et al. 2003b). A resolution of 1 marker mapped every 1.5Mb was achieved by the latest map (Murphy et al. 2006). These maps provided a framework for sequence information acquired for the cat as a result of the feline genome project initiated in 2002 (O'Brien et al. 2002). From this, a 2X draft coverage sequence has become available

and SNP discovery is underway (see http://pre.ensembl.org/Felis_catus/index.html); a 7X coverage genome sequence is anticipated sometime in 2007. In addition, the feline CHORI BAC library (RPCI-86; http://www.bacpac.chori.org) has been exploited to generate a high-resolution contig map of the feline major histocompatibility complex (Beck et al. 2005).

Gene mapping and genome-related research in the dog have evolved rapidly as well in recent years – once again with linkage and cytogenetic mapping taking the initial lead (Mellersh et al. 1997; Neff et al. 1999; Breen et al. 1999; Mellersh et al. 2000). High-resolution radiation hybrid maps with a 1 marker per Mb coverage of the genome have been developed for the dog using the RHDF5000-2 whole genome RH panel (Guyon et al. 2003). This panel was next used for an integrated map of the genome including 4,249 markers (Breen et al. 2004), complemented by the 1.5X whole-genome sequence from a standard poodle dog which was available as early as 2003 (Kirkness et al. 2003). Moreover, Hitte et al. (2005) constructed a 10,000-marker RH map using this survey sequence by deriving gene markers from the available sequence and mapping them using a $9{,}000_{rad}$ RH panel. In this way, the somewhat limited poodle sequence data was organized to yield more information. The 7.5X draft sequence of the boxer genome was made public in July 2004 along with a catalogue of SNPs established using information from multiple breeds (Lindblad-Toh et al. 2005). Of particular interest in the dog community are extensive regions of linkage disequilibrium (LD), resulting from the homogeneity due to an imposed lack of diversity and/or inbreeding within individual dog breeds (Ostrander and Wayne 2005); investigation of these regions provide clues to

canine evolution (Lindblad-Toh et al. 2005). Additionally, haplotypes that are shared across multiple breeds were uncovered which may reveal heritable elements such as risk factors that are common to the breeds (Lindblad-Toh et al. 2005).

The gene maps and genomes of mice have proven useful for making comparisons with human and for modeling a number of diseases (for review, see Guénet 2005). A framework provided by a clone-based physical map (Gregory et al. 2002) was used to organize sequence information obtained for the mouse in the first draft of the mouse genome sequence (Mouse Genome Sequencing Consortium 2002). The present assembly of the mouse genome (NCBI build 36) is approximately 2.66 Gb in length and includes nearly 22,000 genes. Availability of sequence information has facilitated investigations into features of the mouse genome, such as those that could give clues about events in murine and mammalian evolution. These include segmental duplications (Cheung et al. 2003), gene deletions and evolutionary breakpoints (Fitzgerald and Bateman 2004), and differences in genetic variation between strains of laboratory (e.g., Wade et al. 2002; Yalcin et al. 2004) and "wild-derived" (Ideraabdullah et al. 2004) mice, among others. Mouse sequence data has also made possible closer examinations into the transcriptome, including the identification of non-coding RNAs (ncRNAs; e.g., The Fantom Consortium 2005).

For pigs, integration of cytogenetic and linkage maps (Archibald et al. 1995; Rohrer et al. 1996; Goureau et al. 1996) was first achieved using the INRA-University of Minnesota 7000_{rad} whole-genome radiation hybrid panel (IMpRH; Yerle et al. 1998; Hawken et al. 1999). This panel was also utilized in the generation of EST RH

comparative maps of the porcine genome (Rink et al. 2002, 2006). The development of a $12{,}000_{rad}$ RH panel (IMpRH2; Yerle et al. 2002) has allowed finer mapping of regions of interest (e.g., Demars et al. 2006). At present parts of the pig genome are being sequenced, with a 0.66X coverage map already published (Wernersson et al. 2005); it is expected that a 6X coverage of sequence data for the pig genome will probably be available by next year. Finally, the generation of porcine microarrays has provided valuable resources for carrying out functional analyses of the pig genome (Niewold et al. 2005; Dvorak et al. 2005).

Impacts in humans and other mammalian species

The domestic species mentioned above have profited from higher-resolution maps which have facilitated improvements in medicine and understanding molecular bases behind traits of interest. For example, genes responsible for morphological differences between dog breeds (Chase et al. 2002) have been mapped and examined. Gene expression data has clarified the genetic basis behind behavior patterns in different canids (e.g., Saetre et al. 2004, 2006). Also, dense gene maps greatly enhance insight into the evolution of these species and mammals as a whole (Liu et al. 2006; Kijas et al. 2006), which can ultimately aid in the discovery and exploitation of functional elements. High resolution sequence-level gene maps increase knowledge of the evolution of certain sequence features, such as repetitive elements (Schibler et al. 2006) and gene families (e.g. natural killer receptor gene complex, Hao et al. 2006). Analysis of gene sequences in dogs has led to an improved elucidation of the evolution of canids as well as mammals (Parker et al. 2004). Importantly, causative genes and mutations

responsible for traits of interest have been identified and investigated in these species. Examples of such conditions include double muscling in cattle (Grobet et al. 1997), resistance to edema disease (Frydendahl et al. 2003) and membranoproliferative glomerulonephritis type II (Hegasy et al. 2002) in pigs, spider lamb syndrome in sheep (chondroplasia; Beever et al. 2006), beta-mannosidosis in goats (Leipprandt et al. 1996), coat color (Kerns et al. 2003; Berryere et al. 2005) and leukocyte adhesion deficiency in dogs (Kijas et al. 2000) – among others (for a detailed list, refer to "Online Mendelian Inheritance in Animals [OMIA]," http://omia.angis.org.au/). Causes behind resistance to retroviral infection have been decoded in cattle (Si et al. 2006).

During recent years, a range of novel powerful research tools and resources such as microarrays and single nucleotide polymorphism (SNP) chips are expanding our knowledge of more complex, multi-gene conditions. For example, a bovine innate immune microarray has been developed to help ascertain molecular aspects of varying disease responses in cattle (Donaldson et al. 2005). Gene expression profiles in cattle embryos have been analyzed for improving nuclear transfer efficiency using microarray technology (Smith et al. 2005). Genomic information generated for the pig has been used in various applications, including gene expression analyses in phenotypically different types of muscles (Bai et al. 2003), ovary function (Caetano et al. 2004), and microarray development (e.g., Zhao et al. 2005). Canine sequence data have recently been used in a number of applications, such as SNP detection (Chevreux et al. 2004), identifying and determining the function of microRNAs (Wang and Wang 2006), and gene expression profiles of tumors (Thomas et al. 2003, 2005; Thomson et al. 2005).

Altogether, these examples represent cases of the practical application of genomic information in these species.

Several factors serve as prerequisites for the development of dense gene maps. First is a rich pool of markers that can be rapidly utilized for mapping to individual chromosomes. This includes polymorphic microsatellite markers, sequence-tagged site (STS) markers and gene specific markers. While a large number of polymorphic markers are being generated and verified for their degree of variability (e.g., Ellegren et al. 1992; Marklund et al. 1994), gene specific markers can either be generated from ESTs obtained from cDNA libraries developed for some of the horse tissues (e.g., Lieto and Cothran 2001; Pascual et al. 2002; Stratagene; others also available at http://www.ncbi.nlm.nih.gov/) or through multiple alignment of sequence data from other species to obtain equine orthologs of human genes. BAC-end sequences with comparative information can also be used as markers. These markers may be genotyped on an RH panel and a dense linear map can be obtained for individual horse chromosomes. Assuming the need for 1 marker/Mb of the genome, a 2,800 marker physical map will provide the desired resolution. However, to have this resolution, the actual number of mapped markers needs to be higher (~4,000) so that a high-confidence ordered framework can be obtained. Further, selected markers will then have to be mapped via cytogenetic methods to anchor RH groups and determine their orientation. Only then can these maps be of the desired resolution where they can be used for mapping and identifying genes governing traits of interest and making detailed

comparisons with other species to elucidate features associated with chromosome evolution.

In short, other species have profited significantly from the level of detail supplied by their gene maps. In order to see similar progress for the horse, it is clearly necessary for gene map resolution to improve. Increased marker density and nucleotide-level information will result in an enhanced utility of the gene maps for practical applications in horses, such as those observed in other species.

Gene mapping approaches

Several approaches are employed for developing gene maps; these may be broadly classified as genetic and physical. While most genetic and physical maps provide the relative order of loci in a given region of a chromosome, each measures distance between markers differently and has different requirements for map generation.

Genetic maps

Linkage maps fall under the category of genetic (or meiotic) maps. Thomas Hunt Morgan's observation of the non-random segregation of genetic "characters" led him to hypothesize that such characters are associated – or linked – with one another (Morgan 1911). His student, Alfred H. Sturtevant, subsequently developed the first genetic linkage map for *Drosophila* (Sturtevant 1913; refer also to Sturtevant 1917), ordering markers on the basis of crossover frequency between them. This mapping required characteristics (like physical traits) that could be inherited in different forms which today we call alleles.

Linkage mapping requires a set of highly polymorphic markers, which are used to screen a reference family where the segregation of specific alleles is traced to identify loci (or actually alleles of loci) that are segregating together. However, because only polymorphic markers – primarily microsatellites – may be used, linkage maps are not very useful for making comparisons of maps across different species (which may not share the same polymorphic markers) and subsequently inferring evolutionary events. Also, because recombination occurrence is not uniform throughout the genome, distance estimations – measured as frequency of recombination between markers and expressed as centiMorgans, cM – are not always reliable. As a result, map sizes may not be accurately judged in relation to one another.

Linkage maps provide an assortment of polymorphic markers (referred to as a genome scan panel) that are distributed uniformly over the genome and are routinely used to map traits of interest including diseases to specific chromosomes or chromosomal regions. Researchers use this data to identify candidate genes implicated in the phenotype. As a result, the method is particularly useful for finding genes and mutations/variations potentially associated with diseases or traits.

Physical maps

Synteny, cytogenetic, radiation hybrid, contig, and sequence-level maps are all examples of physical maps. These maps can incorporate any marker, regardless of polymorphic status. The term "syntenic" describes those loci which reside on the same chromosome; accordingly, synteny maps group genes on the basis of their common chromosome address. Synteny mapping involves the use of a somatic cell hybrid panel,

constructed by fusing donor (e.g., horse) cells with recipient (primarily rodent) cells to form hybrid cells (see Ephrussi and Weiss 1969). The chromosomes from the donor genome are lost at random as individual hybrid cells multiply, until they at last become more or less stable resulting in each hybrid clone having an established and partly variable complement of donor chromosomes. Clones can be screened by PCR, enzyme electrophoresis, or Southern blotting for the presence or absence of genes or gene products. A gene may be assigned to a chromosome when screening reveals that the chromosome is consistently positive or negative for that gene or gene product in every cell line where the chromosome is present.

Cytogenetic mapping involves labeling a DNA probe and then hybridizing it to its corresponding location on a chromosome. Fluorescence *in situ* hybridization, or FISH, involves fluorescently labeling DNAs for specific markers, which then may be hybridized individually or in groups using different labeling and reporter molecules (Pinkel et al. 1986). FISH can vary in resolution from 3-5Mb when done using metaphase spreads (Trask et al. 1989; Yokota et al. 1995), 25-750kb when carried out on interphase chromatin (Lawrence et al. 1988), or 1-500kb when using stretched DNA fibers (Laan et al. 1995). Irrespective of the approach used, FISH helps to establish physical locations of the markers on the chromosome and has the ability to depict their relative order. When performed using DNA material from an entire chromosome, the technique is known as chromosome painting (Pinkel et al. 1986; Lichter et al. 1988). When chromosome painting is done across evolutionarily closely or distantly related

species, it is referred to as ZooFISH (Scherthan et al. 1994). The latter facilitates detection of homology between chromosomes of different species.

Radiation hybrid panels are developed by fusing irradiated donor cells (e.g., horse cells) with non-irradiated recipient cells (e.g. hamster cells), forming hybrid cells (Goss and Harris 1975; Cox et al. 1990). Clones of hybrid cells contain both whole hamster chromosomes and donor chromosomes fragmented by radiation. The donor chromosomes tend to fuse with recipient chromosomes. Because not all donor chromosomes can be stably duplicated and segregated during mitosis, usually due to lack of centromeres, the unincorporated donor chromatin is lost at random until the individual hybrid cell lines (collectively termed "panel") are stabilized. The radiation dose given to the donor cells may be adjusted to control the frequency of breaks and therefore the achievable map resolution. Typically, radiation hybrid panels of 5000_{rad} have been made to be able to deduce linear order of ~5000 or less markers in a range of species (e.g., Womack et al. 1997; Murphy et al. 1999b). In some instances RH panels have been generated with a radiation dosage of 12,000 or even $20{,}000_{rad}$ (e.g., Rexroad et al. 2000; Yerle et al. 2002). A higher radiation dose will result in more chromosomal breakages, thus increasing the chances of a break occurring between two markers and enhancing map resolution. Such panels can resolve the order of ~10,000 markers (Hitte et al. 2005). Typically, the amount of rads used in generating the panel indicates the number of markers than can be mapped (e.g., a $5{,}000_{rad}$ panel will allow up to 5,000 markers to be mapped).

The panel of hybrid cell lines generated as described above can be genotyped by PCR, and the results analyzed using statistical software (e.g., RHMAPPER, Slonim et al. 1997; RHMAP 3.0, Boehnke 1992; Lunetta et al. 1995; rh_tsp_map, ftp://ftp.ncbi.nih.gov/pub/agarwala/rhmapping/rh_tsp_map.tar; Agarwala et al. 2000; CONCORDE, http://www.isye.gatech.edu/~wcook/rh/; Applegate et al. 1998; Qsopt, http://www.isye.gatech.edu/~wcook/qsopt). This results in physically ordered gene maps with distances between markers measured in centiRays (cR) – which is a statistical probability indicating the likelihood of a radiation-induced break occurring between the two markers (Cox et al. 1990). Because primers for any marker that can be screened by PCR may be used for genotyping the RH panel, regardless of polymorphic status, radiation hybrid maps facilitate the integration of markers from various sources (i.e. genetic linkage, synteny, cytogenetic) into a single consensus map. Importantly, such integrated maps contain a substantial proportion of gene specific markers that are very useful for making detailed comparisons with maps of different species.

A third type of physical map is the contig map, constructed from contiguous, large overlapping clones of DNA fragments, such as the BACs. Contigs provide a framework for the in-depth analysis of a region of interest (for example, the equine major histocompatibility complex [MHC]; Gustafson et al. 2003) or for the efficient development of nucleotide-level maps, as they facilitate the sequencing of a large section of DNA by partitioning it into more manageable, sequencing-ready pieces (e.g. the human genome; see below). Most BAC clones may be put in order based on shared characteristics such as known marker content or sequence. This can be done via

digesting the BAC DNA with a restriction enzyme and then comparing resultant patterns in a process known as fingerprinting (Marra et al. 1997; Soderlund et al. 2000), or by screening clones by PCR for the presence of certain unique sequences or markers. Alternatively, end sequences generated from the BACs are used to generate STS markers, which in turn provide the basis to substantiate the overlaps determined by fingerprinting (Olson et al. 1986; Venter et al. 1996; Zhao et al. 2000). Establishing clone order by default provides the order of the markers they contain though order of markers within a clone may not be deduced in some instances. The development of contigs results in highly detailed, linearly ordered physical maps which may include any type of marker.

The ultimate physical map is the sequence-level map because it provides the highest level of resolution (at the nucleotide level). The development of sequence maps is considerably facilitated by a clone library. The clones may be organized into contigs prior to sequencing to anchor the sequence data onto previously-established physical maps, as done by the public endeavor to sequence the human genome (International Human Genome Sequencing Consortium 2001). Alternatively, whole genome DNA may be digested and cloned, and the clones may be randomly sequenced to eventually assemble the data (shotgun approach; Venter et al. 1998). This method was employed by Celera to sequence the human genome in 2001 (Venter et al. 2001). Such nucleotide-resolution maps enable the identification and localization of all genes within the genome of a species, and make possible the description of their structures including regulatory elements (e.g., Margulies et al. 2005; Kimura et al. 2006). These maps also provide

complete insight into non-coding regions of the genome and facilitate research to understand their significance in genome structure and function (e.g., Johnson et al. 2005a; Kapranov et al. 2005). Sequence information also elucidates features such as single nucleotide polymorphisms (SNPs; The International SNP Map Working Group 2001), haplotypes (The International HapMap Project 2003, 2005), and gene regulatory elements (The ENCODE Project Consortium 2004). In addition, sequence information provides valuable evolutionary insights by allowing analysis of breakpoint regions that are signatures of evolutionary rearrangements from an ancestral genome.

Overview of gene mapping in the horse

Efforts to analyze the horse genome and develop gene maps started much later compared to some of the other domestic species such as cattle, pig, and chicken. Gene mapping for the horse began with the mapping of the glucose-6-phosphate dehydrogenase (G6PD) gene to the equine X chromosome by two groups (Trujillo et al. 1965; Mathai et al. 1966). The next decade heralded the first mapping of autosomal genes (Sandberg, 1974). Unfortunately, an inundation of maps did not follow, and equine gene mapping made slow progress over the next several years. Some of the significant contributions made during this period included developing genetic linkage maps (of three coat color genes; Andersson and Sandberg 1982) and physical assignment of the equine major histocompatibility complex to chromosome 20 by *in situ* hybridization (Ansari et al. 1988; Mäkinen et al. 1989). Eventually the pace of equine gene mapping was accelerated after the first International Equine Gene Mapping

Workshop in 1995, where researchers from around the world formed a foundation for the organized and collaborative study of the equine genome at an international level. One of the findings reported at this meeting was the whole genome horse-human comparative map generated by Zoo-FISH (published later as Raudsepp et al. 1996) that formed the key basis for the development of gene maps in horses, primarily by permitting the use of the human gene maps as the template for targeted expansion.

Synteny map

Somatic cell hybrid panels were among the first organized tools developed for generating gene maps of horse chromosomes. Panels of hybrid cells were generated by several groups (Lear et al. 1992; Williams et al. 1993; Bailey et al. 1995) by fusing horse and mouse cells to form hybrids which contained a complete complement of mouse chromosomes and a subset of equine chromosomes. These panels were used to obtain early synteny maps for some of the equine chromosomes. However, the most used panel developed for the horse was that reported by Shiue et al. (1999), which contributed significantly in the construction of synteny maps for the all equine autosomes and the sex chromosomes (Caetano et al. 1999a, b; Shiue et al. 1999). These maps provided gene mapping evidence to reaffirm the homology between horse and human genomes that was previously reported by Raudsepp et al. (1996). As synteny mapping does not allow markers to be placed in a linear order, this particular method became obsolete by the emergence of the radiation hybrid mapping technique.

Cytogenetic map

The technique of in situ hybridization was first utilized by Ansari et al. (1988) and Mäkinen et al. (1989) to localize the equine major histocompatibility complex. During the next 4-5 years only ~15 additional genes were mapped using this approach, some of which included calcium release channel, glucosephosphate isomerase, etc. (Harbitz et al. 1990). Oakenfull et al. (1993) broke new ground by mapping the alpha globin gene complex by fluorescence *in situ* hybridization to horse chromosome 13. This was the first gene to be mapped by this method (FISH) in the horse. Subsequently, the development of the first horse-human Zoo-FISH map (Raudsepp et al. 1996) represented an advancement of the FISH approach to identify evolutionarily conserved chromosomal segments between the two species. As mentioned earlier, this provided the opportunity to utilize the well-developed human gene map for identifying homologous regions in the horse and facilitated provisional assignments of linkage and synteny to specific chromosomes. The findings also contributed to the elucidation of a possible ancestral karyotype by comparing Zoo-FISH and gene mapping data from a range of other species (human, pig, cattle, Indian muntjac, cat, American mink, harbor seal, and horse; Chowdhary et al. 1998). During the past 5-6 years, a number of studies have added to the expansion of cytogenetic maps for horse chromosomes, noteworthy among which are Godard et al. (2000), Lear et al. (2001), Mariat et al. (2001), Milenkovic et al. (2002), Chowdhary et al. (2003), Lee et al. (2003a), Gustafson-Seabury et al. (2005) and Perrocheau et al. (2005, 2006). The primary contribution of these studies is the anchoring of genetic linkage, synteny, and RH groups to specific

chromosomes or chromosomal locations identified by other maps, and identification of additional regions of homology. These studies have also provided valuable estimations and/or verifications of marker orders and distances.

Linkage map

The first autosomal linkage map constructed for the entire horse genome employed the Uppsala half-sib reference family (Lindgren et al. 1998). Almost at the same time, the International Horse Reference Family Panel was analyzed and facilitated the establishment of two more linkage maps for the horse genome (Guérin et al. 1999, 2003) and the Animal Health Trust three-generation full-sib reference family was used for an additional whole-genome linkage map (Swinburne et al. 2000). Recently, Penedo et al. (2005) combined data from all three half-sib reference families into a single comprehensive map containing 766 markers spanning 3,740cM and distributed on thirty-one equine autosomes . Furthermore, the AHT three-generation full-sib reference family was utilized for the generation of a significantly improved resolution 2,772cM linkage map comprising a total of 742 markers on the autosomes and the X chromosome (Swinburne et al. 2006). Altogether these genetic maps have provided two sets of genome scan panels: the International Horse Reference Family Panel (IHRFP; Guérin et al. 1999, 2003) and the panel resulting from the map generated using the AHT family (Swinburne et al. 2000). Together these provide a total of 742 markers (Swinburne et al. 2006) for scanning the genome with polymorphic markers uniformly distributed over the genome, to map or find genes or markers potentially associated with traits of interest.

Importantly, the map published by Swinburne et al. (2006) had only one linkage group per chromosome.

Radiation hybrid map

Radiation hybrid maps have the advantage of integrating different kinds of maps, irrespective of the type of markers. The first RH panel developed for the horse was a 3000_{rad} panel used to construct preliminary RH maps for horse chromosomes 1 and 10 (ECA1 and 10; Kiguwa et al. 2000). A short time later, a 5000_{rad} panel was constructed by Chowdhary et al. (2002) and used to generate an RH map for ECA11 before being used to build a comprehensive radiation hybrid map for the entire horse genome (Chowdhary et al. 2003). This represented the first high-throughput gene map for the horse genome that integrated type I and type II markers, thereby amalgamating synteny, cytogenetic, and linkage maps into a single consensus map. Subsequently, this same panel was used to create high-resolution maps (≥1 marker per Mb) for several equine chromosomes: ECAX (Raudsepp et al. 2002; Raudsepp et al. 2004a), ECA17 (Lee et al. 2003a), ECAY (Raudsepp et al. 2004b), ECA22 (Gustafson-Seabury et al. 2005), and the equine homologs of HSA5 (Goh et al. 2006), and has been integral to the work reported in this dissertation. Maps of this resolution are essential to rapidly and efficiently identify genes responsible for conditions of significance to the horse industry, such as disease resistance or susceptibility, fertility, and performance-related traits. Also, such maps offer a comparative platform in relation to the gene maps of other species that provides insight into chromosome evolution.

Sequence map

A landmark event in the history of the horse came in July 2006 with the announcement that the entire equine genome is being sequenced by scientists at the Broad Institute of MIT and Harvard, Boston, USA. Using DNA from an inbred Thoroughbred mare named Twilight, the horse genome is currently in the process of being deciphered, with a 7X coverage anticipated for the ~2.7 Gb genome by the end of the year. Simultaneously, a catalogue of SNPs is being generated using material from various modern and ancestral horse breeds. As a member of Perissodactyla, a mammalian order which presently lacks any sequenced genome, the horse is expected to supply valuable information that will contribute to ongoing efforts to understand evolution of mammalian genomes. Comparing whole genome sequence data from diverse genomes enables the identification of not only the coding sequences but also a variety of regulatory elements that are evolutionarily conserved and potentially serve key roles in regulating gene function (The ENCODE Project Consortium 2004; Margulies et al. 2005). Such elements, identified across mammals and vertebrates, are expected to serve as targets for drug delivery in future biomedical research and treatment approaches.

Relevance

The over-arching goal of horse gene mapping is the discovery and exploration of genes associated with equine diseases and traits of interest (such as disease resistance and susceptibility, performance, and fertility) along with the elements that regulate the manifestation of these conditions. As early as 1992, the causative mutation responsible

for a disorder inherited by simple Mendelian means (hyperkalemic periodic paralysis) was identified and described (Rudolph et al. 1992). Subsequently, gene maps and comparative genomics have led to identification of mutations implicated in severe combined immunodeficiency disease (Bailey et al. 1997; Shin et al. 1997), overo lethal white foal disease (the equine version of Hirschsprung disease in humans; Yang et al. 1998; Santschi et al. 1998; Metallinos et al. 1998), junctional epidermolysis bullosa (Spirito et al. 2002; Milenkovic et al. 2003), glycogen storage disease (Ward et al. 2004), and malignant hyperthermia (Aleman et al. 2004). Additionally, genes and their variants or polymorphic markers associated with various coat colors have been described: agouti and brown (Rieder et al. 2001), cremello (Locke et al. 2001; Mariat et al. 2003), tobiano (Brooks et al. 2002; Mau et al. 2004; Brooks and Bailey 2005), chestnut (Marklund et al. 1996), grey (Henner et al. 2002; Locke et al. 2002; Swinburne et al. 2002), silver (Mikko et al. 2006), and appaloosa (Terry et al. 2002, 2004). These developments testify the significance and use of gene maps in the horse. However, they also underline the need of higher resolution maps to facilitate the rapid detection of genes directly associated with the conditions. High-resolution maps presently being generated by us and reported in this dissertation will provide essential details for deciphering the causes of these conditions at the molecular level. High-resolution maps also serve as the basis for refined comparisons of the equine genome with genomes of different species, which in turn also means an improved understanding of chromosome evolution.

The availability of whole genome sequence information for the horse in the near future will enable identification of all genes present in the horse genome. This will pave

the way for the construction and use of expression microarrays - most likely based on long oligos of ~70-mer length, and will lead to functional analysis of the equine genome, which will be key to the investigation of changing levels of gene expression under normal developmental as well as pathological states. Further, the genomewide SNP chips comprising 30-50,000 SNPs distributed over the equine genome will serve as an important platform for association studies for Mendelain as well as complex traits. Jointly, the knowledge resulting from all these developments are of major relevance to a better comprehension of the mechanisms associated with manifestation and progression of diseases and their subsequent prevention, diagnoses, and treatments.

Improving gene maps for the horse – need and possible strategies

Despite consistent advances in analyzing the equine genome, current maps in the horse lack the adequate density essential for candidate gene discovery and association studies. While the first generation radiation-hybrid map for the horse genome represented a major advancement by providing gene maps of all autosomes and the X chromosome, the resolution of one marker mapped every 4Mb or greater provided by the map (Chowdhary et al. 2003) was insufficient for future practical applications. The most noteworthy drawback in the map is the number and distribution of both gene specific and polymorphic microsatellite markers and lack of alignment between maps developed by different groups using different approaches.

Rationale for the present study

In this dissertation, the generation of detailed gene maps for seven equine chromosomes (ECA6, 7, 10, 15, 18, 21 and X) was undertaken. These chromosomes were selected for high-resolution mapping on the basis of their potential involvement in harboring genes governing diseases and conditions important to equine welfare. Firstly, horse chromosomes 7, 10, and 21 share homology with human chromosome 19 (HSA19), which is the most gene-dense chromosome in humans (Dehal et al. 2001; Grimwood et al. 2004). This human chromosome has been found to contain genes implicated in conditions which are also of interest to horses, including insulin-like 3 (INSL3) - implicated in cryptorchidism and reduced fertility in humans and mice (Nef and Parada, 1999; Zimmermann et al. 1999; Ivell et al. 2005; Ferlin et al. 2006); insulin receptor (INSR), involved in diabetes mellitis (Taira et al. 1989; described for horses in Johnson et al. 2005b); and ryanodine receptor 1 (RYR1), which has a role in malignant hyperthermia (MacLennan et al. 1989; described for horses in Aleman et al. 2004).

Expanding on the study of the horse homologs of HSA19, a section of ECA21 was selected for contig development. This particular genomic segment warrants special interest due to a Perissodactyl-specific breakpoint at the boundary in homology between HSA19 and HSA5 on ECA21. Also, this ~5Mb segment contains genes responsible for various diseases and conditions that are of interest to horse owners and breeders. These include the aforementioned INSL3 gene; cartilage oligomeric matrix protein (COMP), connected with pseudoachondroplasia (Briggs et al. 1995; Posey et al. 2004); and exostoses (multiple) 3 (EXT3), which is implicated in a bone growth disorder called

multiple exostoses (Le Merrer et al. 1994). Hence it was decided to develop a BAC contig on this region to decipher the gene content, obtain complete sequence and make the information available for future equine research. An additional goal of this study was to precisely demarcate homology of this region in relation to HSA19 and characterize the evolutionary break/fusionpoint connecting to the segment corresponding to HSA5.

The other chromosome chosen for development of a dense gene map was the X chromosome. This chromosome is the most conserved chromosome in mammals (Ohno 1967; Charlesworth 1991), with an organization, constitution, and gene content that have remained very much the same throughout evolution in most mammals. The chromosome is known to harbor a preponderance of sex- and reproduction-related genes. It also contains genes that are associated with sex-linked diseases, development, and/or performance (Ross et al. 2005). Moreover, dosage compensation is one of the key phenomena associated with this chromosome (Lyon 2002). All these factors motivated us to develop a high resolution map of the chromosome in the horse such that an improved understanding of its organization and content can pave a way to study the genes present on this chromosome.

Finally, the equine homologs of human chromosome 2 were selected for generation of dense gene maps. This human chromosome carries genes that are implicated in thyroid- and reproduction-related disorders. Some of these include hyperparathyroidism 3 (HRPT3), involved in familial isolated hyperparathyroidism (Warner et al. 2006); CELIAC3 and CTLA4, involved in susceptibility to celiac disease (Djilali-Saiah et al. 1998; Naluai et al. 2000; van Belzen et al. 2004);

preeclampsia/eclampsia 2 (PEE2), implicated in the reproductive disorder preeclampsia (Arngrimsson et al. 1999); CDHS2, involved in coronary heart disease (Falchi et al. 2004; BHF Family Heart Study Research Group 2005); and susceptibility to essential hypertension 3 (HYT3; Angius et al. 2002). In humans, this chromosome harbors a unique ancestral chromosome fusion event (Fan et al. 2002a, b; Hillier et al. 2005) and in horses one of the boundaries of homology with HSA2 appears to be unique to Equids.

The detailed mapping of each of these seven equine chromosomes (including parts or all of ECA6, 7, 10, 15, 18, 21, and X) is described in this dissertation, along with methods employed, inferences made, and potential benefits.

II. OBJECTIVES

The overall goals of this dissertation were to generate high resolution physical and comparative maps that could eventually be used for improved understanding of equine diseases and conditions, disease resistance, fertility, reproduction, coat color, and other traits that are of economic significance to the equine industry. Generation of such maps will not only lead to the development of novel biomedical tools for equine health and welfare, but will also provide insight into the evolution of equine chromosomes. These overall goals were accomplished using the following specific objectives, listed along with the significance of each.

The first objective was to develop high resolution integrated physical gene maps for seven horse chromosomes: ECA6p, ECA7, ECA10p, ECA15, ECA18, ECA21 and ECAX. Dense gene maps with a targeted resolution of one marker per Mb or greater are needed for every horse chromosome. These can be efficiently generated via analysis of genotyping data for both previously-mapped as well as new gene-specific markers, STSs, and microsatellites on the 5000_{rad} horse x hamster RH panel. The linear map thus generated can be anchored to specific chromosomes and appropriately oriented by FISH localization of selected markers. This will ultimately lead to high resolution consensus maps which will integrate markers from a variety of sources (including those on different linkage and synteny maps), thereby maximizing their usefulness.

The second objective was to generate comprehensive comparative maps for these equine chromosomes and finely align them to corresponding chromosomes in humans (viz., HSA2, HSA19 and HSAX) and other sequenced mammals. Gene mapping data

and sequence positions of comparative markers (where available) in a range of mammals and vertebrates provide information about the organization of the corresponding regions of the genomes of these species. This in turn helps to finely align the equine genome with that of sequenced species or species with dense gene maps. Such comparative maps are useful for candidate gene searches, as genes implicated in a condition in one species may also be responsible for that or a similar condition in another.

The third objective was to postulate the comparative evolution of the seven equine chromosomes from a common ancestor in relation to other mammalian species and orders. Comparative gene maps give useful information for evaluating regions of homology across diverse species. This elucidates their likely ancestral state and provides an account of the evolution of the chromosome(s) leading to configurations in individual species. It also enables the identification of evolutionary breakage and/or fusion points and rearrangements. Extrapolating information on the evolution of horse chromosomes together with other species can lead to an improved understanding of mammalian and vertebrate evolution.

The fourth objective was to generate a sequence ready Bacterial Artificial Chromosome (BAC) contig over a 5 Mb region of the ECA21 that corresponds to the proximal half of HSA19p, and characterize it for gene content and order. Contigs provide a template upon which to anchor sequence data. In addition, they are essential for fine mapping of regions of interest, such as those known or suspected to contain candidate genes (e.g., INSL3 on ECA21, implicated in cryptorchidism). Also, contigs facilitate more precise localization, sequence analysis and characterization of

evolutionary breakage or fusion points, such as the boundary in homology between HSA19 and HSA5 on ECA21. They also highlight minute rearrangements in gene order that can be easily missed by lower-resolution maps.

III. SUMMARY OF METHODS USED

The high-resolution physical, comparative, and BAC contig maps for selected equine chromosomes were prepared using a range of approaches. Details of these approaches are provided in sections IV-VII containing three published articles and a manuscript describing a high resolution map for an individual group of chromosomes. In the sections below, a summary of these approaches is provided. The materials and methods are classified into three broad categories:

- Radiation hybrid (RH) mapping
- BAC isolation, contig development, and fingerprinting
- Fluorescence in situ hybridization (FISH)

Radiation hybrid (RH) mapping

This was the primary approach for generating the detailed physical gene maps in this dissertation. The method involves the use of the 5000_{rad} horse x hamster radiation hybrid panel developed at Texas A&M University (described in Chowdhary et al. 2002). The panel comprises 92 hybrid cell lines and has been characterized and used extensively to generate whole genome radiation hybrid maps for the horse (e.g., Chowdhary et al. 2003).

Marker generation

A variety of markers was genotyped on the panel for the generation of linear maps of equine chromosomes. These include gene-specific markers, microsatellites, and sequence-tagged sites (STSs).

Gene-specific markers

Two sources were used to obtain this type of markers. The first included expressed sequence tags (ESTs) generated from the equine skeletal muscle and the testis cDNA libraries. Sequence data from ~25,000 ESTs was analyzed using bioinformatic approaches (BLAST hits and <1e-10 E-value when compared with available mammalian sequences) to identify loci that would potentially map to the targeted chromosomes or chromosomal regions based on previously described homology between horse and human chromosomes (Raudsepp et al. 1996). For each of the chromosomes, genes were selected along the corresponding human chromosome at regular intervals (~1-2 Mb) so as to ensure the desired resolution. In addition to ESTs, gene specific markers were also developed by obtaining equine orthologs for human genes, following alignment of multiple sequences (http://www.ddbj.nig.ac.jp/search/clustalw-e.html) of the desired gene from a range of mammalian species (www.ncbi.nlm.nih.gov/BLAST; www.ensembl.org; genome.ucsc.edu). Subsequent primer design was performed in such a way as to obtain equine specific amplification. In such cases, the PCR product was verified by sequencing.

Microsatellite markers

All published polymorphic microsatellite markers on the chosen chromosomes were used for genotyping and integration in the map. For this, published data from Guérin et al. 2003, Penedo et al. 2005 and Swinburne et al. 2006 were used. Additionally, in some cases (e.g., Wagner et al. 2006) new microsatellite markers were generated by us.

STS markers

All published sequenced tagged sites (STSs) expected to map on the selected chromosomes were genotyped in this study. Additionally, a number of STS markers were generated by end-sequencing of individual BACs. However, the majority of these STSs were not genotyped in the RH panel primarily due to their close physical proximity which could lead to problems in resolving their physical order. These STSs were instead used for chromosome walking and identification of new BAC clones

Primer design and optimization

PCR primers for all types of markers mentioned above were designed using the PRIMER3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). All primers were optimized for equine specific DNA amplification in a hamster DNA background. This was essential to avoid amplification of hamster DNA (a part of hybrid cell lines) that could lead to false positive scoring of the genotyping results. This was primarily accomplished by selecting for primers within conserved sequences, with 1-3 mismatches in rodent sequence. PCR products of heterologous primers were sequenced to verify the identity of the expected gene for genotyping on the RH panel.

Genotyping, data analysis, and map development

Following optimization, primers for individual markers were genotyped by PCR on the RH panel in a 10 µl volume for each reaction. All PCRs were carried out in duplicate. The amplified products were visualized on a 2% agarose gel and all results were scored manually. The resultant data was analyzed using one of two different software packages. The ECAX RH map was generated using RHMAPPER (Slonim et al.

1997) and RHMAP 3.0 software (Boehnke 1992; Lunetta et al. 1995), as described in Chowdhary et al. (2003). Compared to this, the maps of the equine homologues of HSA19 and HSA2 were constructed using the software packages rh_tsp_map (ftp://ftp.ncbi.nih.gov/pub/agarwala/rhmapping/rh_tsp_map.tar; Agarwala et al. 2000), CONCORDE (http://www.isye.gatech.edu/_wcook/rh/; Applegate et al. 1998), and Qsopt (http://www.isye.gatech.edu/_wcook/qsopt). The maps were drawn using the program MapChart, a software for the graphical presentation of maps developed by Voorrips et al. (2002; see also http://www.biometris.wur.nl/UK/Software/MapChart/). Comparative maps were developed using either available gene mapping or annotated sequence information, with preference to the latter. The comparative data was used to identify conserved linkages – regions of chromosomes across species with similar gene order; or conserved syntenies – regions of chromosomes carrying the same group of genes. Finally, this data was used together with available Zoo-FISH information to identify the ancestral chromosomal configuration and the putative evolution (fission and/or fusion events) that led to the present arrangement in compared species.

BAC libraries

Overgo primer development

For generating overgo primers to be used in contig construction on ECA21, genes were selected at ~150kb intervals along the 14-20Mb sequence segment of human chromosome 19 and the 65-70 Mb segment of HSA5 (NCBI build 35). Sequence information from a variety of mammalian and vertebrate species (identified using

ENSEMBL; http://www.ensembl.org/) was screened for repetitive elements using RepeatMasker (http://repeatmasker.genome.washington.edu/cgi-bin/Repeat-Masker) and aligned with the human gene sequence using CLUSTALW (http://www.ddbj.nig.ac.jp /search/clustalw-e.html). Conserved regions of the sequence with four or fewer mismatches across all species were selected for design of o*ver*lapping oli*go*nucleotide primers, or overgoes. This was accomplished using the Overgo Maker program (http://www.genome.wustl.edu/tools/?overgo.html), and the gene identities of the resultant overgoes were confirmed by BLASTn analysis (http://www.ncbi.nlm.nih.gov /BLAST/).

BAC library screening (by PCR and filter hybridization)

In order to obtain DNA for FISH and contig construction, two equine BAC libraries – the Texas A&M University horse BAC library (http://hbz7.tamu.edu/ homelinks/bac_est/bac.htm) and the CHORI-241 horse BAC library (http://bacpac. chori.org/equine241.htm) – were screened using PCR or overgo primers for selected markers.

For the three radiation hybrid mapping studies, oligonucleotide primers for selected markers were used to identify BACs by screening superpools and plate pools of the BACs from each library by PCR. Subsequent PCR screens of row and column pools ultimately provided the BAC clones containing the specific markers. BACs included in the contig along ECA21 were identified by screening CHORI-241 BAC library filters using labeled overgo primers (probes) for markers in the region. The eleven filters

together contained the entire collection of BAC clones included in the library, spotted onto the filters in a grid pattern.

DNA extraction

Selected BAC clones were cultured in 2YT media containing 30µl/ml chloramphenicol and subsequently plated on LB agar plates which also contained 30µl/ml chloramphenicol. The identities of colonies were confirmed by PCR. Positively verified single colonies for BACs were used to inoculate 100ml of 2YT + 30µl/ml chloramphenicol. DNA was obtained for these BACs via alkaline lysis extraction (e.g., Birnboim and Doly 1979; Birnboim et al. 1983) using a Qiagen midi-prep kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions.

BAC end sequencing and STS generation

BAC end sequencing was accomplished using the T7.29 (5'-GCCGCTAATACGACTCACTATAGGGAGAG) and SP6.26 (5'-CCGTCGACATTTAGGTGACACTATAG) primers for BACs from the CHORI-241 library and T7.19 (5'-TAATACGACTCACTATAGGG) and M13 reverse (5'-CAGGAAACAGCTATGACC) primers for the Texas A&M University equine BAC library. A Gene Amp (Applied Biosystems) PCR system 9700 was employed for dye terminator sequencing reactions (total volume 10 µl) as previously described (Gustafson et al. 2003). Purification of reaction products was carried out by passing the products through spin columns (Spin-50, BioMax, Odenton, MD). The purified products were then loaded on an ABI 3730 automated capillary sequencer (PE Applied Biosystems).

Contig development

STS content mapping

STSs were developed from end sequences for all BACs included in the ECA21 contig. The sequences were screened for repeats using RepeatMasker and analyzed by BLASTn to ensure that they were either horse-specific or located on a homolog of ECA21 in other species. Primers were designed from BAC end sequences using PRIMER3 software and optimized on horse genomic DNA as well as the BAC of sequence origin. These STS primers were then used to screen other BACs in the contig, in particular those adjacent to the parent BACs, for STS content mapping. The results were used to verify positions and orientations of the BACs.

Chromosome walking

Chromosome walking was performed for the ECA21 contig using STS primers for BACs flanking the observed gaps. The primers were used to screen pooled BAC libraries by PCR to identify new BACs extending into the gap. This "walking" continued with the end-sequencing of new BACs, STS primer design, and screening by PCR until the gaps were filled.

Fingerprinting

To further verify the order of BACs in the ECA21 contig, an automated capillary electrophoresis-based fingerprinting method (Xu et al. 2004, 2005) was used. Fingerprinting experiments were carried out using a three-restriction enzyme and one-labeling color kit (Xu et al. 2004). The BAC DNA was digested with HindIII, BamHI

and HaeIII; ddATP-NED or ddATP-HEX were then used to label ends resulting from digestion by BamHI and HindIII. Next, the labeled fragments were fractionated on an ABI 3100 capillary sequencer and the FPC program was used to edit and assemble the BAC fingerprints (Soderlund et al. 2000; Xu et al. 2004, 2005). Questionable clones (Qs) were removed from the automatic contigs initially assembled and subsequent examination was performed to disassemble the chimeric contigs. Finally, the FPC program was used to add singletons and merge overlapping neighbor contigs into larger contigs.

Fluorescence *in situ* hybridization (FISH)

Approximately 1 μg of DNA from BACs isolated for the presence of specific markers were individually labeled either by biotin (Bio-Nick Mix – Roche Molecular Biochemicals) or digoxigenin (Dig-Nick Mix – Roche Molecular Biochemicals) and hybridized in the presence of unlabeled competitive DNA onto horse metaphase chromosome spreads or interphase chromatin, either singly or in pairs. *In situ* hybridization, signal detection, microscopy, and image analysis were performed as per Chowdhary et al. (2003). The single color hybridizations on metaphase chromosomes were carried out to anchor and orient the RH groups, while the double color hybridizations with two or three probes on metaphase or interphase chromatin were used for finding the relative order of closely located markers.

IV. A HIGH-RESOLUTION PHYSICAL MAP OF EQUINE HOMOLOGUES OF HSA19 SHOWS DIVERGENT EVOLUTION COMPARED TO OTHER MAMMALS*

Synopsis

A high-resolution (1 marker/700 kb) physically ordered RH and comparative map of 122 loci on equine homologues of human chromosome 19 (HSA19) shows a variant evolution of these segments in equids/Perissodactyls as compared to other mammals. The segments include: parts of both the long and the short arm of horse chromosome 7 (ECA7), proximal part of ECA21, and the entire short arm of ECA10. The map includes 93 new markers, of which 89 (64 gene-specific and 25 microsatellite) were genotyped on a 5000rad horse x hamster radiation hybrid (RH) panel, and 4 were mapped exclusively by FISH. The orientation and alignment of the maps was strengthened by 21 new FISH localizations, of which 15 represent genes. The ~seven-fold improved map resolution attained in this study will prove extremely useful for candidate gene discovery in the targeted equine chromosomal regions. The highlight of the comparative map is the fine definition of homology between the four equine chromosomal segments and corresponding HSA19 regions specified by physical coordinates (bp) in the human genome sequence. Of particular interest are the regions on

---

*This section is reprinted with kind permission of Springer Science and Business Media from "A high resolution physical map of equine homologs of HSA19 shows divergent evolution compared with other mammals" by Candice Brinkmeyer-Langford, Terje Raudsepp, Eun-Joon Lee, Glenda Goh, Alejandro A. Schäffer, Richa Agarwala, Michelle L. Wagner, Teruaki Tozaki , Loren C. Skow, James E. Womack, James R. Mickelson and Bhanu P. Chowdhary, 2005, *Mammalian Genome*, volume 16, pages 631-49.

ECA7 and ECA21 that correspond to the short arm of HSA19 – a genomic rearrangement discovered to date only in equids/Perissodactyls as evidenced through comparative Zoo-FISH analysis of the evolution of ancestral HSA19 segments in 8 mammalian orders involving ~50 species.

Introduction

The purpose of genome analysis in the horse (*Equus caballus* - ECA) is to identify and analyze genetic factors governing disease resistance, fertility, performance and other traits significant to the equine industry, for which highly informative gene maps are essential. These maps comprise an orderly collection of gene specific and polymorphic markers representative of the entire genome and are instrumental in facilitating discovery of genes of interest. The currently available low- to medium-resolution synteny (Caetano et al. 1999a, b; Shiue et al. 1999), linkage (Lindgren et al. 1998; Swinburne et al. 2000; Guérin et al. 1999, 2003), and cytogenetic maps (Raudsepp et al 1996; Godard et al. 2000; Lear et al. 2001; Mariat et al. 2001; Milenkovic et al. 2002) for each horse autosome and the sex chromosomes have limitations in terms of uneven distribution of markers and low density in several regions. Consequently, the possibility to use the candidate gene approach is limited. This limitation is evident even in the first-generation radiation hybrid (RH) and comparative maps (Chowdhary et al. 2003) and the newly published consensus linkage map (Penedo et al. 2005). The recent construction of ~1-Mb resolution physically ordered maps for some of the horse chromosomes e.g., ECA17 (Lee et al. 2003a), ECAX (Raudsepp et al. 2002, 2004a),

ECAY (Raudsepp et al. 2004b), and ECA22 (Gustafson-Seabury et al. 2005) is providing an excellent platform to initiate detailed analysis of genomic regions suspected to harbor genes of significance to horse breeders and owners. Maps of this resolution are indeed essential for all equine chromosomes.

HSA19 corresponds to small parts of both the short (p) and long (q) arms of horse chromosome 7 (ECA7), proximal one-quarter of ECA21, and the entire short arm of ECA10 (Raudsepp et al. 1996; Chowdhary et al. 2003; Yang et al. 2004) and is one of the most gene dense chromosomes in the human genome (size: 64 Mb; contains 1750 genes) (Dehal et al. 2001; Grimwood et al. 2004). The most recently published map of the HSA19-homologous regions in the horse includes only 5 genes assigned to ECA7, 7 to ECA10, and none to the proximal part of ECA21 (Chowdhary et al 2003). This low resolution of gene specific markers is insufficient for candidate gene based searches in the human and mouse genomes for various equine conditions. The number and distribution of polymorphic markers in these regions is also inadequate. To resolve these shortcomings, the present study aims to generate a comprehensive high- resolution map for equine homologues of HSA19 (ECA7, ECA21, and ECA10) using an established 5000rad horse x hamster RH panel (Chowdhary et al. 2002). The findings will facilitate the identification of genes of significance in these regions of the equine genome and contribute to improved understanding of the evolution of HSA19 homologues in horses, equids and other mammals.

Materials and methods

Primer design and marker generation

Genes located on human chromosome 19 were selected at approximately 1 Mb intervals from the telomere of the short arm to the telomere of the long arm along the human draft sequence (http://genome.ucsc.edu, version May 2004). A representative sequence for each gene was selected from ENSEMBL (http://www.ensembl.org/) and used to obtain homologous sequences in mouse, rat, cattle, pig, and other mammalian species using BLASTn (http://www.ncbi.nlm.nih.gov/BLAST/). Following multiple alignment of these sequences using ClustalW (http://www.ddbj.nig.ac.jp/search/clustalw-e.html), conserved regions were identified for primer design. Either a single exon or two adjacent exons were selected and primers were designed in highly conserved regions between different mammalian species with two to three mismatches with mouse or rat at potential primer sites. Primers were designed using the Primer3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and were optimized using hamster and horse genomic DNAs to ensure horse-specific amplification. All equine PCR products were verified by sequencing and analyzed using BLASTn (Altschul et al. 1997) to confirm their homologies with the corresponding human gene. Additionally, primer pairs for 25 microsatellite markers were obtained from various sources (see Table A4-1 for details; note: tables with the designation 'A' are found in the Appendix) using approaches previously described (Tozaki et al. 2004; Wagner et al. 2004a, b, c).

Genotyping of markers on a radiation hybrid panel

PCR typing of the markers obtained above was performed in duplicate on the 5000rad horse x hamster RH panel (Chowdhary et al. 2002) as described earlier. Each marker was typed using 50ng DNA as template, 1X buffer (Sigma Aldrich), 0.3 pmols of each primer, 0.2mM dNTPs, 1.5, 2.0, or 3.0mM $MgCl_2$, and 0.25 units JumpStart REDTaq DNA polymerase (Sigma Aldrich) in a 10µl PCR reaction. PCR amplification was performed as follows: an initial 30 s denaturation at 94 °C; 1 cycle of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s; followed by 30 cycles of 94 °C for 30 s, annealing temperature (ranging from 50-60 °C) for 30 s, and 72 °C for 30 s; ending with a final extension for 5 min at 72 °C. For those primers whose annealing temperatures are listed as TD60, a touchdown program was used wherein the annealing temperature decreased by 1°C increments each cycle during the initial 11-cycle segment from 60 °C to 50 °C, followed by 30-cycle segment with the lowest (50 °C) annealing temperature. The amplification products were resolved on 2.5% agarose gels (containing 0.25µg/mL ethidium bromide) and manually scored.

Computation of radiation hybrid (RH) maps

RH maps were computed using the software packages rh_tsp_map (ftp://ftp.ncbi.nih.gov/pub/agarwala/rhmapping/rh_tsp_map.tar; Agarwala et al. 2000), CONCORDE (http://www.isye.gatech.edu/~wcook/rh/; Applegate et al. 1998), and Qsopt (http://www.isye.gatech.edu/~wcook/qsopt). The maps were constructed to optimize the maximum likelihood (MLE) similarly to the cat maps in Menotti-Raymond et al. (2003b). However, we also took advantage of the new Qsopt package and recent

enhancements to rh_tsp_map. The rh_tsp_map package performed well in an independent comparison (Hitte et al. 2003). Pairwise LOD scores were computed and inter-marker distances were estimated. Linkage groups were identified with pairwise LOD score thresholds of 6.0, 7.0 and 7.0 for ECA7, ECA10 and ECA21, respectively. These thresholds yielded large enough groups, with the exception of one size 3 group on ECA7, and left few singleton markers. We dropped 8 singleton markers on ECA7, none on ECA10 and 2 on ECA21.

Ben-Dor and Chor (1997) showed that one is unlikely to correctly order dozens of markers with a panel size of 92, so we selected a subset of markers in each linkage group for creating a robust map. The program frame_markers eliminated from initial consideration markers that have too many 2 entries – representing uncertainty regarding presence (1) or absence (0) of the marker in a specific RH line - and one marker out of each pair of markers that are too close to each other. These markers are candidates for placing relative to the robust map at a later stage. For this study, the intergenic markers were preferentially retained over microsatellite markers to obtain more useful comparative maps. Robust maps were created by finding a globally optimum order for an automatic translation of the RH mapping problem to the well-studied traveling salesman problem (TSP), and then using the CONCORDE package (Applegate et al. 1998) linked together with Qsopt to solve the TSP instances to guaranteed optimality, as described earlier (Agarwala et al. 2000). We call these robust maps for a subset of markers "MLE-consensus maps" (instead of the more generally used term “framework map”) because each map is an optimal order for all three definitions of MLE that differ

in how the 2 entries are treated. The robustness of the maps was tested using the map_eval (LOD ≥0.25) and flips programs, comparing the best order to the second best with a threshold of LOD ≥0.25 and a window size of 8.

Next, the placement program was used to find for each marker dropped by frame_markers its best placement relative to the MLE-consensus map. Placed markers with vectors identical to MLE-consensus markers were set aside at this stage. Markers with non-identical vectors that could not be placed with a LOD > 0.1 were dropped. Following this, maps for multiple linkage groups on the same horse chromosome were concatenated using intermarker distances estimated earlier. The order and orientation of linkage groups on a chromosome was determined by FISH data and available genetic maps (Swinburne et al. 2000; Guérin et al. 2003; Penedo et al. 2005).

Finally, to generate cR positions, we constructed and solved a restricted TSP instance in which 1) the MLE-consensus markers were required to stay in the same order and 2) the placed markers were required to lie in the preferred interval between MLE-consensus markers or between the extreme MLE-consensus marker and the telomere. The only flexibility in these reduced TSP instances occurs when multiple placed markers are assigned to the same interval by the placement program. Although we placed as many markers as possible with cR positions, we used only the markers on the MLE-consensus map for comparison of marker/gene order with other mammalian genomes.

BAC library screening and FISH mapping

Primer pairs for selected markers mapped to ECA7, ECA21, and ECA10 were used to screen the CHORI-241 (http://bacpac.chori.org/equine241.htm) equine BAC

libraries. About 1µg DNA from BAC clones thus obtained was individually labeled with either biotin or digoxigenin using the Bio- or DIG-Nick Translation Mix (Roche Molecular Biochemicals). Labeled probes were hybridized either separately or in pairs to horse metaphase chromosome spreads in order to determine their physical location on the chromosome. DNA labeling, in situ hybridization, signal detection, microscopy, and image analysis were performed as previously described (Chowdhary et al. 2003).

Results

Development of gene-specific markers

A total of 117 primer pairs corresponding to human chromosome 19 genes were generated, of which 74 gave horse-specific amplification against a hamster background. Products from the latter were sequenced, and their identities were confirmed in relation to human genes using BLAST. These markers were genotyped with respect to the 5000rad horse x hamster RH panel. Information about all markers used is summarized in Table A4-1.

RH mapping

Of the 74 HSA19 gene specific markers developed above, 56 were successfully genotyped on the RH panel. Genotyping data from 20 additional genes (5, 1 and 14 genes from HSA5, 11 and 19, respectively) and 25 microsatellite markers were added to these results (see Table A4-1 for details). Following initial analysis of the 101 markers for the development of linkage groups, 12 gene specific markers were dropped (located too close to another marker), leaving 89 new markers mapped on equine chromosomal

regions corresponding to HSA19 (for details see Table A4-2). The average retention frequency of the markers mapped to the specific regions of the three chromosomes is 18.2% (ECA7), 19.6% (ECA10), and 15.1% (ECA21). Physically ordered RH maps were developed for the 3 chromosomes (Fig. 4-1; Table A4-3). A summary for each chromosome is provided below.

ECA7 map

Analysis of 71 markers led to the identification of two RH groups containing 54 markers spanning the entire length of the chromosome. The proximal group contained 16 markers distributed over 419.47 centiRay (cR) length, while the distal group contained 38 markers spread over 818.23 cR, of which only 20 loci are shown in the truncated version in Fig. 4-1. Thus two small regions corresponding to parts of the short arm of HSA19 were identified on ECA7 (Fig. 4-1). The remaining part of the chromosome corresponds to HSA11 (Raudsepp et al. 1996). A large proportion of the markers (74%) was present in the MLE-consensus map and confidently ordered.

ECA10 map

Analysis of 122 markers on this chromosome led to their distribution into two RH groups. Figure 4-1 shows only the group present on the ECA10p because this arm corresponds to HSA19q (ECA10q corresponds to part of HSA6; Raudsepp et al. 1996). The group comprises 57 markers, and spans 756.99 cR. Twenty-three of these markers with distinct RH vectors were in the MLE-consensus map and confidently ordered.

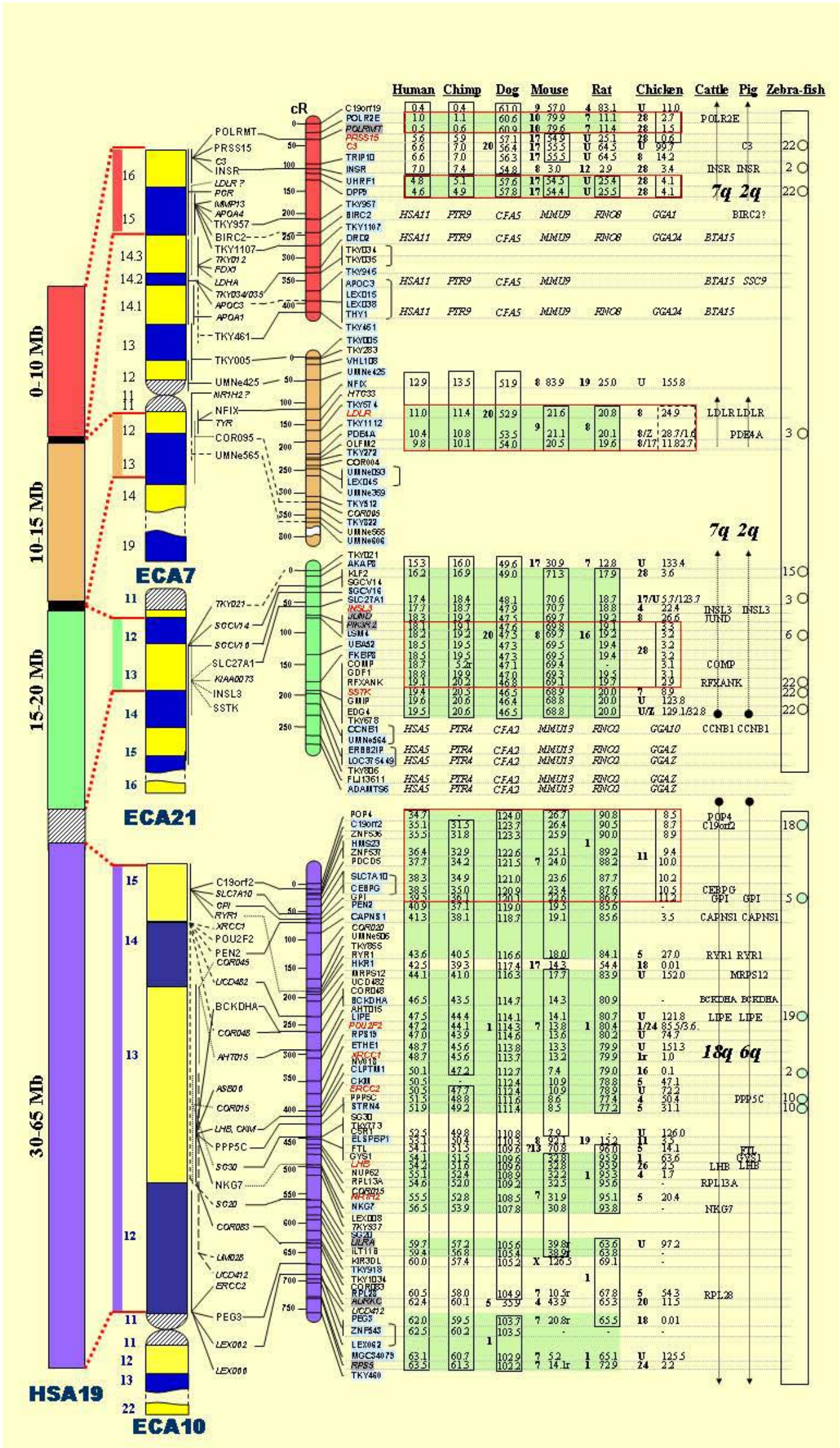


Figure 4-1. High-resolution radiation hybrid (RH) and comparative maps of equine chromosomal regions corresponding to human chromosome 19 (HSA19). To the far left is a schematic representation of HSA19 showing regions (demarcated in megabasepairs-Mbp on the human genomic sequence) that correspond to parts of horse chromosomes 7, 10 and 21 (ECA7, ECA10, ECA21). The G-banded ideograms of the equine chromosomes depict these regions by a color bar (*left*), and show all FISH-mapped loci (*right*). The 21 newly FISH-mapped loci are shown in bold. Next to the FISH-mapped markers are RH maps spanning only those regions of the equine chromosomes that correspond to HSA19. A total of four RH groups with color-codes corresponding to homologous regions on HSA19 are shown. The distal group on ECA7 is truncated at ~365 centirays (cR) to restrict to HSA19 homologous region. The RH groups are demarcated at 50 cR intervals. The MLE-consensus markers (framework) are shaded in light blue, with those having a lod score ≥3.0 shown in bold font. The remaining markers (unshaded) were placed on this framework. Markers shown in italics were placed at a lod score ≤0.5, and can therefore have alternative adjacent locations. Markers with same physical/cR position have a vertical bar to the left and are connected by a common line to the map. Markers with identical vectors have a bracket to the right. Gene specific markers in red font represent FISH mapped loci that were tentatively placed in the map on the basis of RH-map location of adjacent FISH mapped markers. "?" indicates disagreement between FISH and RH assignments. Next to the RH map are the megabases locations for human (HSA), chimp (PTR), dog (CFA), mouse (MMU), rat (RNO) and chicken (GGA) orthologs of the mapped horse genes (http://genome.ucsc.edu/cgi-bin/hgGateway). Vertical rectangles on Mb position of genes within each species essentially show conserved gene order in relation to the derived order in the horse. The light green shaded horizontal rectangles show conserved linkage across sequenced mammals. The red horizontal rectangles depict putative conserved linkages across vertebrates. A status of the comparative mapped loci in cattle (BTA) and pig (SSC) is provided for orthologs with available mapping information. Arrows indicate putative order of loci in these species. Mapping data for HSA19 orthologs in zebrafish (DRE) are depicted to the far right, showing the comparative genome organization of this evolutionarily distantly related species in relation to mammals.

ECA21 map

Analysis of 103 markers assigned to this chromosome provided 2 RH groups of which the proximal group comprising 22 markers (10 with distinct RH vectors in the MLE-consensus map) corresponds mainly to HSA19 (Raudsepp et al. 1996), and spans 278.07 cR. The distal RH group corresponds to HSA5 and is therefore not presented.

BAC library screening and FISH mapping

To accurately align and orient the RH groups to specific chromosomal regions, FISH was carried out with BAC clones isolated for 20 genes and 9 microsatellites (Fig. 4-2). In general, the FISH locations of markers coincide well with the order of loci in the RH groups. Of the 15 genes successfully FISH mapped in this study, 14 represent equine orthologs of HSA19 genes.

Comparative map

A comparative map of equine genomic segments corresponding to HSA19 in a range of mammalian/vertebrate species was developed as described earlier (Chowdhary et al. 2003; Gustafson-Seabury et al. 2005). For this, the physical orders of 9, 16 and 35 putative equine gene loci from ECA7, ECA21 and ECA10p RH maps, respectively, were used. Additionally, 3 loci (*PRSS15,C3*, *LDLR*) from ECA7, 3 (*KIAA0073, INSL3, SSTK*) from ECA21 and 5 (*XRCC1, POU2F2, LHB, ERCC2* and *NR1H2*) from ECA10p were 'placed' in the physical order, based on their location in the FISH map in relation to adjacent markers (Fig. 4-1 markers in red). Megabase positions of orthologs for the 70 putative equine genes were obtained for different species (see Fig. 4-1 legend for details). As described earlier (Chowdhary et al. 2003) maximally contiguous

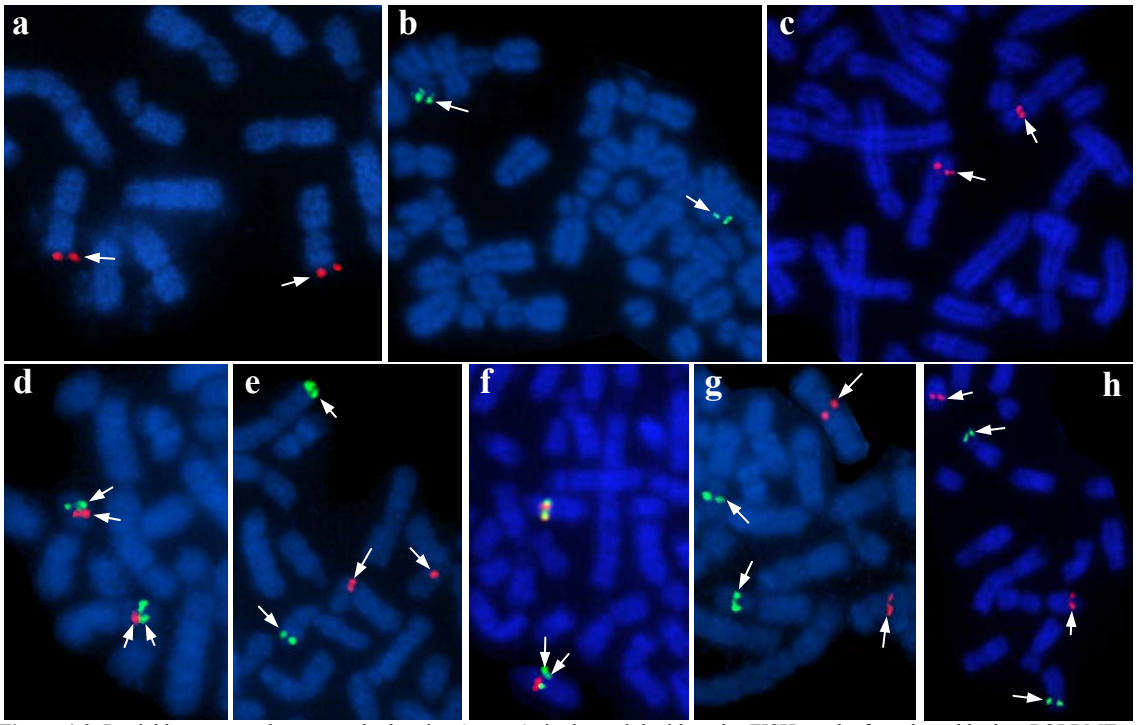


Figure 4-2. Partial horse metaphase spreads showing (arrows) single- and double-color FISH results for selected loci. a: POLRMT - ECA7p16, b: C19orf2 - ECA10p15, c: TKY957 - ECA7p15-14.3, d: BCKDHA (green) and NKG7 (red) - ECA10p13 and 10p13-p12, respectively e: PRSS15 (green) - ECA7p16 and PEG3 (red) - ECA10p12, f: SSTK (red) and SLC27A1 (green) - ECA21q13, g: PPP5C (green) - ECA10p13, and NFIX (red) - ECA7q11-q12, h: POU2F2 (green) - ECA10p14-p13 and INSL3 (red) - ECA21q13-q14.

chromosomal regions with identical gene content and order – referred to as conserved linkages – were identified by clustering the contiguous Mb locations of individual loci in human/chimpanzee, mouse, rat, dog and chicken (see Fig. 4-1). In other species, only comparative data for mapped equine genes was used.

Briefly, the two ECA7 segments and the ECA21 segment corresponding to the short arm of HSA19 are present as a single syntenic group in human, chimp, dog, cattle and pig. The segment corresponding to ECA21, however, is present as a separate genomic segment in mouse, rat and, largely also in chicken (a non mammalian

vertebrate). Compared to this, the segment corresponding to ECA10p shares conserved synteny in human, chimp and dog, and to some extent also in mouse and rat. In chicken, only a small region corresponding to the terminal part of ECA10p (HSA19q proximal segment) is conserved as a syntenic block on GGA11. The remaining ECA10p (HSA19q), however, is scattered on several chicken chromosomes.

Discussion

High-resolution map of equine chromosomal segments corresponding to HSA19

The generation of high-resolution RH and comparative maps for the two small segments on ECA7, complete short arm of ECA10, and a small sub-centromeric region on ECA21 provides a comprehensive, physically-ordered map for those equine genomic regions corresponding to HSA19. The results confirm and refine our earlier observations (Raudsepp et al. 1996; Chaudhary et al. 1998; Chowdhary et al. 1998; Chowdhary et al. 2003) regarding synteny conservation between HSA19 and corresponding equine chromosomal segments. The results also provide the first gene mapping evidence for the recently detected Zoo-FISH correspondence between parts of ECA21 and the human chromosome (Yang et al. 2004).

Mapping of 89 new loci by RH and 4 exclusively by FISH in this study considerably increases the density of mapped markers on the four equine chromosomal segments corresponding to HSA19. Compared to the previously published map (Chowdhary et al. 2003), the number of mapped markers *(RH and/or FISH)* in these segments increased by 4-fold, and the gene-specific markers by 7-fold. Based on the

current size estimate of HSA19, the four equine chromosomal segments are together anticipated to be ~64 Mbps. Our map for these segments collectively provides a resolution of one marker every 561 kb, which represents a ~7-fold improvement over the previous map that had on average 1 marker/4 Mbps of the equine genome (Chowdhary et al. 2003).

The dense mapping of HSA19 markers in the horse proves extremely useful for the fine alignment of the four corresponding equine chromosomal segments with specific regions on the human chromosome. Our findings show that the terminal part of the short arm of HSA19 (from ~0 to 7 Mb) is represented on the telomeric end of the short arm of ECA7, while the ~9 to 13 Mb region of the human chromosome is represented on the sub-centromeric part of the long arm of the same horse chromosome. The remaining part of HSA19p (~15-20 Mb) is represented on the proximal two bands of ECA21. Compared to this, the long arm of HSA19 (from 34-64 Mb) is represented as a conserved syntenic block on ECA10p (Fig. 4-1).

Over two-thirds of the primer pairs generated for HSA19 genes in this study gave horse-specific amplification in the presence of a hamster DNA background. This reflects distinct success in generation of equine orthologs for HSA19 genes such that the loci uniformly represent the human chromosome. The only exception, however, is the 19-26 Mb segment of HSA19 (band p12). The region primarily comprises gene families that have highly conserved sequences across evolutionarily distantly related species (Grimwood et al. 2004). Hence, for these genes it was difficult to design horse-specific primers that could amplify a distinct band in the presence of hamster DNA. Incidentally,

the region, together with the 27-33 Mb segment across the centromere, also shows a relatively lower density of genes compared to the rest of the human chromosome (GeneLoc: http://genecards.weizmann.ac.il/cgi-bin/geneloc/gene_densities.pl?chr=19).

The FISH localization of 21 new markers in the HSA19-homologous regions increases the tally of cytogenetically mapped markers to 60. Overall, the location of FISH mapped markers is consistent with the physical order obtained by RH analysis. The localizations proved useful to accurately position different RH groups in relation to individual chromosomes. Of particular interest were the FISH assignments on ECA7 where two small but distinct regions of homology with HSA19 were observed. Next, FISH mapping of *BIRC2, TKY1107*, and *TKY957* helped to position the markers in the RH group on ECA7p. Lastly, redesigning of primers for *POU2F2* and *INSL3* (previously FISH mapped by Milenkovic et al. 2002), sequence confirmation of the amplification products, isolation of new BACs and subsequent FISH mapping show that *POU2F2* unambiguously mapped to the distal half of ECA10p while *INSL3* mapped to the proximal part of ECA21. The localizations rectify previous erroneous mapping data and are in better conformity with adjacent genes mapped from HSA19 (see Fig. 4-1 for human Mb positions).

Comparative map

Adding a total of 61 new HSA19 orthologs in the horse gene map increases the number of gene-specific/Type I markers by >5-fold over the previously reported map (Chowdhary et al. 2003). On average, the targeted regions have one Type I marker every Mbp, which is a resolution comparable to that reported by us for ECA17 (Lee et al.

2003a), ECAX (Raudsepp et al. 2004a) and ECA22 (Gustafson-Seabury et al. 2005). This improved resolution enables refined demarcation of correspondence (in Mbps) between the four equine segments and HSA19.

Comparison of gene order in the four equine chromosomal segments with corresponding segments on HSA19 showed conserved linkage between ECA10p and HSA19q. Further, the high density of markers on the small proximal segment of ECA21 is too dense for resolving gene order with current RH panel and metaphase FISH. Nevertheless, the framework markers suggest conserved linkage. Next, the gene order for the two small ECA7 segments corresponding to HSA19 (7 markers on ECA7p and 3 on ECA7q) also could not be resolved optimally by RH analysis. However, tight linkage of the same group of loci in horse and humans suggests conservation. Lastly, the arrangement of the equine HSA19 syntenic segments and their gene order in chimpanzee (PTR20; Richard et al. 2000; http://www.ensembl.org/Pan_troglodytes/syntenyview?otherspecies=Homo_sapiens&chr=20) was predictably the same as in humans. Comparison of gene order in the four equine chromosomal segments with corresponding regions in the dog showed that ECA10p/HSA19q shares conserved linkage with the terminal part of CFA1 (102-124 Mb). The *cen→tel* arrangement of loci on both chromosomes is also the same. However, the three segments on ECA7 and ECA21 (corresponding to HSA19p) are present as a single distal block on CFA20. Interestingly, the telomeric group of loci on ECA7p, CFA20 and HSA19p is the same. Comparison of equine homologues of HSA19 with corresponding mouse and rat chromosomes (Dehal

et al. 2001; Grimwood et al. 2004) showed that ECA10p/HSA19q shares conserved synteny with parts of MMU7 and RNO1 (Fig. 4-1 and 4-3). Though the two rodents show a rearranged gene order in relation to horse/dog/human, their loci are present in two contiguous conserved linkage blocks (Fig. 4-1). Next, the ECA21/HSA19p segment shows distinct conserved linkage both in mouse (MMU8) and rat (RNO16). Lastly, the ECA7 loci are distributed on at least 4 different chromosomes in both mouse and rat, due to which no clear inference on synteny/linkage conservation can be drawn.

Expanding the comparisons to chicken illustrates that ECA7 and ECA21 (HSA19p) genes are predominantly present on a single chicken chromosome - GGA28 (Smith et al. 2002). Though it is difficult to infer comparative order for the ECA7 loci, the ECA21 loci do share similar order with orthologs on GGA28 indicating conserved linkage despite divergence of mammals and birds ~300-350 MYrs ago (Kumar and Hedges 1998; van Tuinen and Hadly 2004). With regards to the ECA10p (HSA19q) loci, the most prominent conserved syntenic and linkage block comprising eight genes was detected on GGA11. The remaining genes seem to be scattered across at least 5 chicken chromosomes.

Despite availability of sequence data in cattle, zebrafish etc., contigs have not yet been put together for individual chromosomes to draw meaningful conclusions. Nevertheless, based on RH map order available in cattle (Goldammer et al. 2002; Gautier et al. 2003; Everts-van der Wind et al. 2004), it is anticipated that ECA10p/HSA19q genes largely (though not completely) share a conserved linkage with BTA18 (Goldammer et al. 2002). Next, ECA7+ECA21/HSA19p genes located on

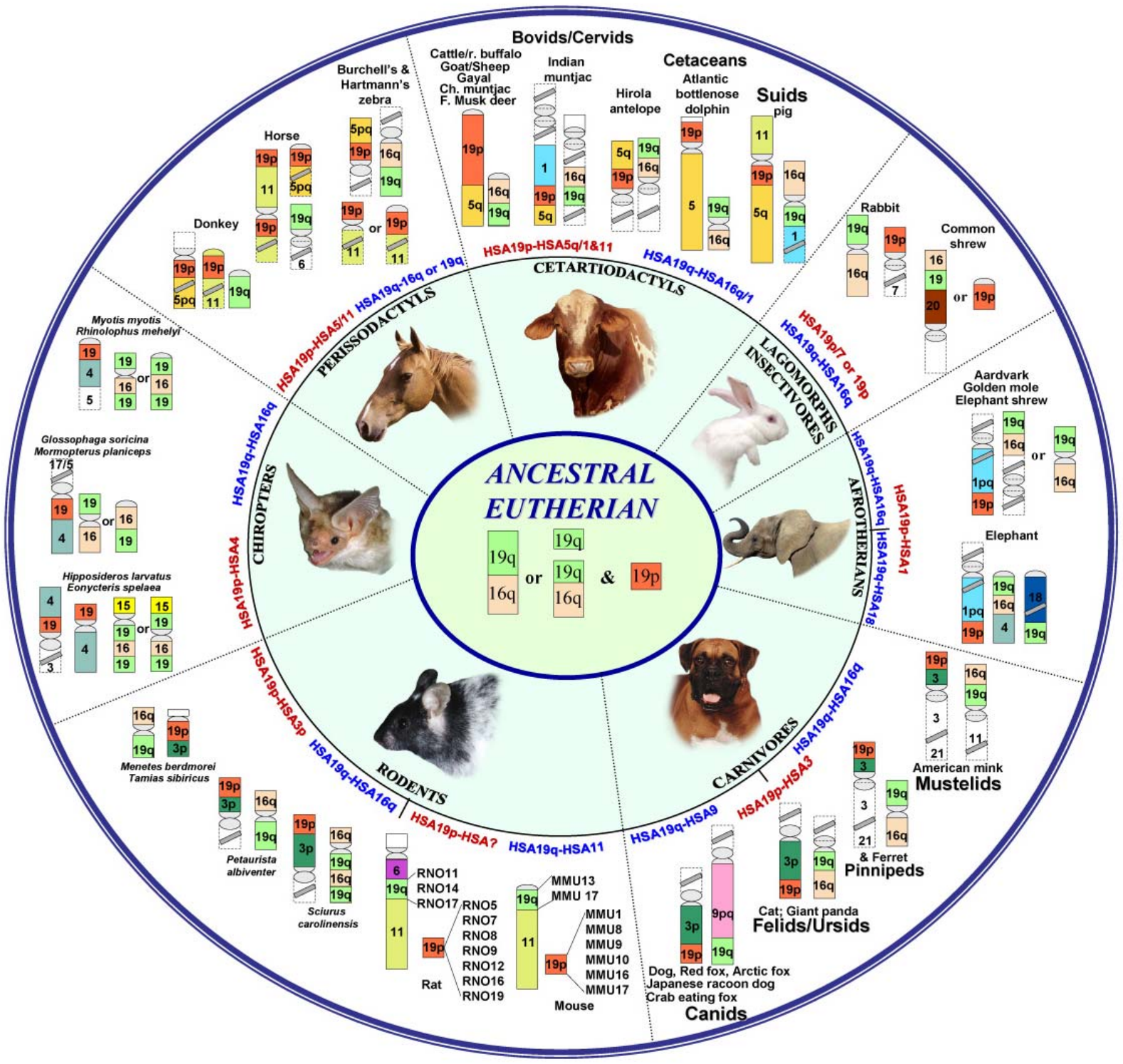


Figure 4-3. An overview of the evolution of putative ancestral HSA19 chromosomal segments (central circle) in >50 mammalian species (hitherto analyzed by Zoo-FISH) belonging to 8 evolutionarily diverse mammalian orders. Perissodactyla is presently the only mammalian order with species that have four distinct chromosomal segments corresponding to HSA19. On the periphery of the second circle are combinations of human chromosomal segments that invariably form tandem/neighboring segments to genomic regions corresponding to HSA19p (red font) and HSA19q (blue font) in the karyotypes of species within individual mammalian orders. This depicts diversity in the evolution of these segments from the putative ancestral configuration. Color codes for correspondence with individual human chromosomes is shown (right bottom). A summary of references from which data were used to develop this figure is provided at the end of the reference list.

References for Fig. 4-3: Perissodactyls: horse and donkey - Raudsepp et al. 1996, Yang et al. 2004; Hartmann's and Burchell's zebras - Yang et al. 2003b; cetartiodactyls: dolphin - Bielec et al. 1998, pig - Rettenberger et al. 1995, Frönicke et al. 1996, Goureau et al. 1996, Schmitz et al. 1998, Indian muntjac - Yang et al. 1997, Burkin et al. 1997, Chinese muntjac and gayal - Chi et al. 2005, Hirola antelope - Chaves et al. 2004, cattle - Chowdhary et al. 1996, river buffalo - Iannuzzi et al. 1998, goat - Schibler et al. 1998, sheep - Iannuzzi et al. 1999; lagomorphs: rabbits/hares - Korstanje et al. 1999; insectivores: common shrew - Dixkens et al. 1998; afrotherians: aardvark and elephant - Frönicke et al. 2003, Yang et al. 2003a, golden mole and elephant shrew - Robinson et al. 2004; carnivores: American mink - Graphodatsky et al. 2000, ferret - Hameister et al. 1997, cat and giant panda - Nash et al. 1998, Tian et al. 2004, arctic fox - Graphodatsky et al. 2000, 2001, crab-eating fox and Japanese raccoon dog - Nash et al. 2001, red fox and dog - Yang et al. 1999, Breen et al. 1999; rodents: Eastern gray squirrel (*Sciurus carolinensis*) - Stanyon et al. 2003, Li et al. 2004, Asian squirrel (*Menetes bermorei*) - Richard et al. 2003, flying squirrels (*Petaurista albiventer*) and chipmunks (*Tamias sibiricus*) - Li et al. 2004, rat and mouse - Nilsson et al. 2001, Helou et al. 2001; and chiropters: bats - Volleth et al. 1999, 2002.

BTA7 show a rearranged order relative to human (Gautier et al. 2003) indicating the likelihood of rearrangement compared also to the horse. In cats, minor rearrangements compared to ECA10q/HSA19q gene order have been detected on chromosome E2; however, HSA19p orthologs on chromosome A2 maintain an evolutionarily conserved linkage (Menotti-Raymond et al. 2003a). Lastly in pigs (SSC), despite high-resolution RH maps of SSC6q1.2 (Martins-Wess et al. 2003) and SSC6q13.3-q13.4 (Bosak et al. 2005), only synteny conservation can presently be detected between ECA10p/HSA19q and part of SSC6q; the remaining three equine segments correspond to SSC2q (Rattink et al. 2001).

Comparative organization of HSA19 homologues

Number of segments in different species

The results in the present study show that HSA19 corresponds to 4 distinct segments distributed on three equine chromosomes. Three of the segments (two on ECA7 and one on ECA21) share synteny with the short arm, and the remaining segment (on ECA10p) shares synteny with the long arm of the human chromosome. The only other known species where HSA19p corresponds to more than one segment are donkey and zebra (Yang et al. 2003b, 2004), though indirect evidence suggests the presence of two segments also in cattle and sheep (Frönicke and Wienberg 2001; Fig. 4-3). In these species, presently a total of 3 segments correspond to HSA19. Other species with 3 segments include elephants (Frönicke et al. 2003, Yang et al. 2003a), Eastern grey squirrel (Stanyon et al. 2003) and some bats (Volleth et al. 1999, 2002). However, contrary to equids (and probably the two bovids), HSA19*q* (rather than HSA19*p*) is

present as two segments in these species. In all other mammalian species/orders studied thus far (except mouse and rat), two distinct syntenic segments of HSA19 are observed (Fig. 4-3), each of which corresponds to one of the arms of the human chromosome.

Evolutionarily conserved neighboring segment combinations

Genomic segments corresponding to six human chromosomes emerge as the *primary* contiguous neighbors (tandemly located or as sole homologue of the other arm) of genomic segments corresponding to HSA19 in different mammalian species analyzed up till now (Fig. 4-3). These include HSA1, HSA3, HSA4, HSA5, HSA11 and HSA16. In Perissodactyls and Cetartiodactyls, segments equivalent to HSA5 and HSA11 form the primary contiguous neighbors to parts sharing homology with HSA19p. In carnivores (including canids, mustelids, pinnipeds, felids and ursids) and non mouse/rat rodents (squirrels, chipmunks), segments corresponding to HSA3 are partnered to the HSA19p-homologous segments. In Chiropters (bats), the neighboring segments are homologous to HSA4, while in Afrotherians (elephants, aardvark etc.) the neighboring segments correspond to HSA1.

Compared to HSA19p, chromosomal segments sharing homology with HSA19q were invariably found to be contiguous with segments corresponding to HSA16q as depicted earlier by us in a cross species analysis (Chowdhary et al. 1998). The two segments are either tandemly placed or have a centromere in between. The only exceptions lacking this combination are horse, donkey, mouse/rat and canids. It is noteworthy that donkey is the only species known wherein the HSA19q equivalent segment exists as an independent chromosome (Raudsepp and Chowdhary 2001, Yang

et al. 2004). In summary, the contiguous neighboring segments provide brief insight into the likely fusion/fission events and rearrangements potentially involved in the evolution of the HSA19 homologues in different species.

Putative ancestral chromosome(s) of HSA19 and their evolution in horses, equids and other Perissodactyl karyotypes

The data summarized in Fig. 4-3 enables the drawing of some basic conclusions about the ancestral chromosomal configuration of HSA19 homologues:

i) The HSA19q-HSA16q combination is unambiguously an ancestral configuration that is preserved in a range of evolutionarily distantly related mammalian species: Hirola antelope (Chaves et al 2004), Indian muntjac (Yang et al. 1997), common shrew (Dixkens et al. 1998), giant panda, spectacled bear (Nash et al. 1998), harbor seal (Frönicke et al. 1997), bats (Volleth et al. 2002), and elephants (Frönicke et al. 2003). It is also present in chicken (GGA11; Smith et al. 2002). As far as known, domestic and Przewalski horse (Myka et al. 2003), donkey (Raudsepp et al. 1996, Yang et al. 2004), canids (arctic fox - Graphodatsky et al. 2000, 2001; crab-eating fox and Japanese raccoon dog - Nash et al. 2001; red fox and dog - Yang et al. 1999; Breen et al. 1999) and mouse/rat (Nilsson et al. 2001; Helou et al. 2001) are the only species where this configuration is broken (see Fig. 4-3). In elephants, squirrels and some of the bats, two small syntenic chromosomal segments jointly correspond to HSA19q (Stanyon et al. 2003; Richard et al. 2003; Li et al. 2004; Volleth et al. 1999, 2002). This is most likely attributed to a fission event, followed by a fusion with two different chromosomes in the

elephant (Frönicke et al. 2003; Yang et al. 2003a) and an inversion in squirrels and bats (Fig. 4-3).

ii) The segment corresponding to HSA19p is evidently an independent ancestral chromosome that is present as a single block in the majority of the species analyzed. Equids, where this segment is present on two different chromosomes, are exceptions to this rule (Yang et al. 2003b, 2004). The Burchell's zebra – rhinoceros Zoo-FISH indirectly indicates a similar trend also in other Perissodactyls (Trifonov et al. 2003).

Starting with these ancestral configurations, the HSA19 homologues in equids/Perissodactyls seem to have evolved as follows:

i) A *fission* event in the ancestral 19p equivalent chromosome led to two segments that can be characterized as 0-15 Mb and 15-25 Mb regions in the current HSA19p. In donkeys and zebras, the 0-15 Mb segment underwent *fusion* with the HSA11 equivalent chromosome, leading to EAS20, EZH14 and EBU14, while the 15-25 Mb segment fused with HSA5pter-5q13 equivalent chromosome, leading to EAS10, EZH5p and EBU9p (Fig. 4-3). In horses, the latter fusion event was the same, leading to ECA21. However, the 0-15 Mb segment underwent an additional fission event, followed by a pericentric inversion event, leading to the current ECA7.

ii) The ancestral 19q-16q equivalent configuration remained intact in the two zebra species hitherto analyzed. However, in the horse and donkey a fission event seems to have separated the two segments. In horses, the 19q related segment fused with parts corresponding to HSA6, leading to ECA10, while the HSA16q segment fused with a segment corresponding to part of HSA4, resulting in ECA3. In donkeys, the fission led

to two independent chromosomes – EAS26 equivalent to HSA19q, and EAS28 equivalent to HSA16q (Raudsepp and Chowdhary 2001; Yang et al. 2004).

Conclusions

The findings of this study provide high-resolution physically ordered RH and comparative maps for equine homologues of HSA19. The maps will be useful for precise physical assignment of traits of interest within these chromosomal regions and for the discovery of candidate loci affecting the traits. Discovery of four genomic segments corresponding to HSA19 (first in any of the mammalian species studied thus far), and their fairly precise Mbp demarcation in relation to the human chromosome, will augment such efforts. The demarcations also provide valuable insights into the pattern of evolution of HSA19 homologous segments in equids/Perissodactyls. A comparative overview of the evolution of these chromosomal segments in different mammals adds to the current knowledge of likely fusion/fission events associated with these segments in different species.

V. BAC CONTIG OVER A 5 MB REGION OF THE HORSE GENOME FOR DISCOVERY OF FUNCTIONAL ELEMENTS AND EVOLUTIONARY BREAKPOINTS

Introduction

Dense gene maps are essential for the discovery, characterization and analyses of genes important for equine health, fertility, performance and other traits that are of vital economic significance to the equine industry. Although significant strides have been made in the recent past to develop low- to medium-resolution genetic maps for all equine autosomes and the sex chromosomes (Chowdhary et al. 2002; Chowdhary et al. 2003; Penedo et al. 2005; Perrocheau et al. 2006; Swinburne et al. 2006), the level of detail available from these maps is not sufficient for most practical applications. Consequently, high resolution gene maps (1 marker/Mb) have recently been reported for some of the horse chromosomes, including ECA17 (Lee et al. 2003a), ECAX (Raudsepp et al. 2004a), ECA22 (Gustafson-Seabury et al. 2005), ECAY (Raudsepp et al. 2004b), and the equine homologues of human chromosomes 19, 2, and 5 (HSA19, HSA2, and HSA5; Brinkmeyer-Langford et al. 2005; Wagner et al. 2006; Goh et al. 2006). Once available for all equine chromosomes, these maps will be useful in a candidate gene approach to help identify loci implicated in traits important to the equine community. However, one of the major drawbacks of these maps is their inability to provide the same level of intricate detail regarding genome structure and organization that can be

found with whole genome sequence information. Such details are essential to study the structures and functions of genes associated with important traits.

Sequence information is presently available for a variety of mammalian species including humans and mice and some of the domesticated animals like cattle, dog, and chicken (http://www.ensembl.org). While sequence information in domesticated animals is proving to be extremely valuable in generating novel tools for the rapid discovery of genes associated with disease and other traits of economic significance, recent studies carried out in humans and mice show that it is also important for proper understanding of gene function and regulation (Bejerano et al. 2004; Margulies et al. 2005; Xie et al. 2005; Kimura et al. 2006) as well as evolution (Medina 2005; Murphy et al. 2005; McEwen et al. 2006). Large scale efforts presently being invested in humans to identify all genomic elements that regulate gene function (ENCODE Project Consortium 2004) further emphasize the significance of this work, and underscore the need for sequence data. Presently, sequence data is being generated for the horse (http://www.broad.mit.edu/). It is anticipated that the complete assembled and annotated sequence along with gene predictions will be available by 2007. This should considerably improve the ability of equine researchers to use genomic information for developing a research platform for horses that could facilitate the development of novel diagnostic, preventive, and therapeutic approaches – the way it is being used in humans and other species. However, support data will be required from a variety of independent sources using complementary mapping approaches to validate and annotate the assembled sequence.

We recently reported a high-resolution (1 marker/700kb) gene map for equine homologues of human chromosome 19 (HSA19) – namely, small segments of ECA7, the entire short arm of ECA10, and the proximal one-quarter of ECA21 (Brinkmeyer-Langford et al. 2005). HSA19 is the most gene-dense chromosome in humans (~1750 genes along its 64Mb length; Dehal et al. 2001; Grimwood et al. 2004). It harbors several genes for which mutations or variations are implicated in diseases of various body systems in humans. Some of these conditions such as dwarfism (Naviaux 1999; Frankeny 2003; Marcella 2005), multiple exostoses (Morgan 1968; Gardner et al. 1975), diabetes mellitis (Johnson et al. 2005), and hypothyroidism (Frank et al. 2002) are documented also in horses. However, due to lack of adequate genome information, very little has been carried out to find their underlying genetic causes.

As a first step towards addressing such issues, we embarked upon generating a BAC-contig based map for the ~5Mb segment of ECA21 that corresponds to HSA19, which contains a number of genes known or suspected to be associated with conditions in humans which may also be of economic significance to the equine industry. These genes include exostoses (multiple) 3 (EXT3) - proposed to be associated with a bone growth disorder known as multiple exostoses (Le Merrer et al. 1994), cartilage oligomeric matrix protein (COMP) - associated with pseudoachondroplasia (Briggs et al. 1995; Posey et al. 2004), and insulin-like 3 (INSL3) - implicated in cryptorchidism and reduced fertility in humans and mice (Nef and Parada 1999; Zimmermann et al. 1999; Ivell et al. 2005; Ferlin et al. 2006). The development of the proposed contig will facilitate searches for such genes and/or closely linked markers associated with various

genetic conditions. It will also serve as a base to assemble and verify the draft sequence of the region and fill in gaps to eventually obtain a finished sequence that could provide complete information on the gene content and various components associated with their regulation (functional elements). Further, the availability of the contig will enable detailed comparisons of the equine segment with corresponding genomic segments in other species and help to detect differences and/or rearrangements it has gathered since evolving independently from a common mammalian ancestor. The contig will also form the basis to characterize two flanking evolutionary breakage/fusion points in relation to human and other mammalian genomes.

Materials and methods

Overgo primer design

Genes at approximately 150kb intervals along the 15-20Mb segment of the draft sequence of human chromosome 19 (NCBI 35 assembly) were selected for overgo design. Representative sequences for each gene in human, mouse, rat, cattle, dog, chicken, and other available mammalian and vertebrate species were selected from ENSEMBL (http://www.ensembl.org/). The sequences were analyzed by RepeatMasker for the presence of repetitive elements (http://repeatmasker.genome.washington.edu/cgi-bin/Repeat-Masker), and aligned using ClustalW (http://www.ddbj.nig.ac.jp/search/clustalw-e.html). Regions conserved across species were identified and overgo primers were designed with four or fewer mismatches using the Overgo Maker program (http://www.genome.wustl.edu/tools/?overgo.html). Individual primer sequences were

screened using BLASTn (Altschul et al. 1997) to confirm their gene specificity. Detailed information on all overgo probes is presented in Table 5-1.

Overgo probe labeling and filter hybridization to screen BAC library

The overgo primers were radioactively labeled as previously described (Gustafson et al. 2003). Briefly, overgo primers were radioactively labeled in a 10 µl labeling reaction containing 1 µM forward primer, 1 µM reverse primer, 150 Ci/mmol each of ^{32}P dATP and ^{32}P dCTP (Amersham Biosciences, Piscataway, NJ), 2 U Klenow fragment DNA polymerase (Roche, Indianapolis, IN), and 1× DNA Polymerase Buffer (Promega, Madison, WI). For fill-in labeling, 1 µl of a 250-µM dATP and dCTP mixture was added to each reaction (Han et al. 2000). Unincorporated nucleotides were removed using Sephadex G-10 gravity flow columns. Labeled probes were then individually checked with a scintillation counter. Approximately the same amounts of each probe were pooled and used for screening high-density filters from the USDA CHORI-241 equine BAC library (http://bacpac.chori.org/equine241.htm). The labeled overgo probes were added to the hybridization solution (20× SSPE, 10% SDS, 5% milk, and 100× Denhardt's Solution) containing 50% formamide, denatured by boiling for 10 min, chilled, and then hybridized onto CHORI-241 BAC library filters at 42° C for 16 h. After hybridization, filters were washed three times at 55° C for 15 min in 2× SSPE. The filters were exposed to film over intensifying screens for 2-7 days at –80° C and the autoradiograms were developed subsequently. Positive BAC clones thus identified were

Table 5-1. Overgo primers used to identify BAC clones containing selected genes. Listings in parentheses are alternate names used for the gene.

GENE NAME	GENE ID (ENSEMBL OR NCBI)	FORWARD OVERGO PRIMER	REVERSE OVERGO PRIMER
ADAMTS6	ENSG00000049192	GTGTGTAGAGAGCTCTGGTGTCTC	AGCGGTTGCTTTTGCTGAGACACC
AKAP8	ENSG00000105127	TCTGACTCTCTCATTGCCAAGATC	TGTCCAAACGCTGGTTGATCTTGG
ASF1B	ENSG00000105011	CGAGCCCTTTCCACAGCCCCTTCC	AAGCTGATCTCGAACCGGAAGGGG
BPY2IP1	ENSG00000130479	GTCAATGGCTTCACTGTGCTGGTC	GGTTTGAGCCACCGTTGACCAGCA
CDK7	ENSG00000134058	AGTCTTGTGCTGACACCATCACAC	ACATGTAGGCTTTGATGTGTGATG
CHERP	ENSG00000085872	ATGACAGCAAGCCTCCCATCCAGA	TCTGAAGAGCCAGGCATCTGGATG
COMP	ENSG00000105664	TTCTCCCAGGAGAACATCATCTGG	GGTAACGCAGGTTGGCCCAGATGA
COPE	ENSG00000105669	TGAGAAGCTGCAGGATGCCTACTA	CATCTCCTGGAAGATGTAGTAGGC
CRLF1	ENSG00000006016	CTCTTTACGCCCTATGAGATCTGG	GGTTGGTGGCCTCCACCCAGATCT
CYP4F11	ENSG00000171903	CACTGCCCATCCTTGCCCCTTTCC	AAGCTCCAAGGACAGTGGAAAGGG
DDX39	ENSG00000123136	CATGGAGGTGTTTGTAGACGACGA	CAGTGTGAGCTTGGTCTCGTCGTC
ELL	ENSG00000105656	GAAGGTTCAGTTTCGGAAACCAGC	GTCTGTTGCACCTGGGGCTGGTTT
EMR3	ENSG00000131355	GACAAGGGATTCATCTGGAGTTTC	CACAGACAGGGCCCAGGAAACTCC
ERBB2IP	ENSG00000112851	TGGCACAGTGATAGAGAATTGCTG	GCTCTACAACTCTCCACAGCAATT
F2RL3	ENST00000248076	ACTCTATGGTCACATGTATGGCTC	GGCCAGCAGCAGCACTGAGCCATA
FKBP8	ENSG00000105701	GCCCTGTACCGGAAAATGCTGGGC	GCAGCCGGCTGGGGTTGCCCAGCA
FLJ39501	OTTHUMG00000070837	CGTCCTGCCACTGTTGGTTCTGGT	GATATAGTCAGGGTGCACCAGAAC
FLJ40365	OTTHUMG00000070840	CCCTGGCGCTTTCGTGTGGAGATG	TGCTGCCCCCTTTGAGCATCTCCA
GMIP	ENSG00000089639	TGAGGTTATCCGCTCGCTGAAGAC	CAGCTGTACCAAGAGGGTCTTCAG
GTPBP3	ENSG00000130299	TCTGACTTGGCCTCTCCGTACAGC	TGTCCAGGAAGTTGCAGCTGTACG
HAPLN4	ENSG00000187664	AACGCCGAGGAACGCTACGACGCC	TGGACGTGAAGCAGAAGGCGTCGT
INSL3	ENSG00000105639	ACAGCGGGCCAGACCAGCAGAGGG	ATCTCCCGCTGAAAGTCCCTCTGC
JUND	ENSG00000130522	AGAAAGTCCTCAGCCACGTCAACA	AGCAGCTGGCAGCCGCTGTTGACG
KCNN1	ENSG00000105642	GGGTTCGGAAACACCAGCGTAAGT	TGGATGGCTTGGAGGAACTTACGC
KIAA0303 (MAST4)	ENSG00000069020	TGGAACAGTTTGCTGAGACAGAAG	GGGGAATAAATTCTGCCTTCTGTC
KIAA0892	ENSG00000129933	TCTTTCACTGGCTGCCCAAGGAAC	ACAAGCACACACATGTGTTCCTTG
KLF2	ENST00000248071	ACGACCTCAACAGCGTGCTGGACT	CCCATGGACAGGATGAAGTCCAGC
LOC345667 (Q6ZNM4)	ENST00000314351	GCATACTTCAGACAGATGTGTTGT	GTCCTTGGCAGGTCTTACAACACA
LOC388515	XM_373795	ATGTGGATTGCAGTGGAGACGCTG	CCGCATTCTCTAACGCCAGCGTCT
LOC388524	NM_001005472	TCTCTTCTGTTCCGCGTCCTCCAC	ACCAGGGGCTACCGAGGTGGAGGA
LOC390898	XM_497612	ACAGGGCATGGAGTTCACGGCAAG	GAAGGTTGATGGGCTGCTTGCCGT
LOC390913	XM_497619	TGGCCAAGCTTCTAACCAATGCTC	CGGTGCAATAAGCTATGAGCATTG
LOC440515	XM_496298	ATGTACATGTTCGTGTTCATGGCC	TAAATTGCAATGGTCCGGCCATGA

Table 5-1. Continued.

GENE NAME	GENE ID (ENSEMBL OR NCBI)	FORWARD OVERGO PRIMER	REVERSE OVERGO PRIMER
LOC91120	OTTHUMG00000071382	GCAATGTGGCAAAGTCTTTAAATC	AGAAAGGCCTGAGTGAGATTTAAA
LSM4	ENST00000252816	AGTGCTACATCCGCGGCAGCACCA	ATGCGCAGGTACTTGATGGTGCTG
MDS032	ENSG00000053501	AACCAGTTCCTGGCCCCTGGCCGT	TGGCTGTGGTTGGCACGCGGCCAG
MECT1	ENSG00000105662	GAGCAGTTCAACATGATGGAGAAC	TGCTGGAGCTGATGGCGTTCTCCA
MEF2B	ENSG00000064489	AAGTTCGGGCTGATGAAGAAAGCT	GCACGCTCAACTCGTAAGCTTTCT
MYO9B	ENSG00000099331	TTCCGGGAGAAGAACATGGACTAC	CGATGTCTGGCCGCATGTAGTCCA
NLN	ENSG00000123213	ATGGACATGCTCCACAATTTCTTG	GGTTTGGCTCACGTTTCAAGAAAT
NOTCH3	ENSG00000074181	CCCTTTGCAACGTGGAGATCAATG	GGGCTGGACGCACACTCATTGATC
OCLN	ENSG00000197822	TCCACCTATCACTTCAGATCAACA	CTTGTACAGTTGTCTTTGTTGATC
PGLS	ENSG00000130313	TGGTCCAGCCCCACACTGGGAAAC	TCCAGGAACCAGCAGAGTTTCCCA
PIK3R1	ENSG00000145675	GACTGTGAATAAAGGGTCCTTAGT	ACTGAATCCAAGAGCTACTAAGGA
PIK3R2	ENSG00000105647	GCAGTACCAGGACAAGAGCCGCGA	ATAAAGCTGGTCATACTCGCGGCT
PTGER1	ENSG00000160951	TGTACATCCTGCTGCGCCAGGCGG	AGCAGTTGGCGCAGCACCGCCTGG
SDCCAG10	ENSG00000153015	TTGCACTGCTGAACCAGTTTAAAT	GCTTGAGTGAGTTTAGATTTAAAC
SFRS12	ENSG00000153914	CTCAGGCTGCAGCTAAGGAGTTAG	CGCTTCATTACTTCTTCTAACTCC
SFRS14	ENSG00000064607	TCGGAAGAGGATCAGCAGCAAGTC	CATGCCAACCTTCAATGACTTGCT
SLC30A5	ENSG00000145740	GATTAATATCATACCGAGACCCTC	GAATGACGCCAAAAATGAGGGTCT
SSTK	ENSG00000178093	CTCTACGTCATGGTCACCGGGTGC	AGTCATCGAAGGGCATGCACCCGG
TAF9	ENSG00000085231	AACCACGCTAGGCAAAGAACTTGC	CAGTCCTGATCTTGATGCAAGTTC
TMEM38A (MGC3169)	OTTHUMG00000071467	ACATGTCTTTCCCCACCAAGGCCA	ATGGCTCCATACAGGCTGGCCTTG
TPM4	ENSG00000167460	GTGGGCTTACATCAGACACTGGAT	GTTCGTTTAGTGTCTGATCCAGTG
TRIM23	ENSG00000113595	TTGGGATGTAGGTGGAAAACACAA	CCACAATGGTCTTAATTTGTGTTT
UNC13A	ENSG00000130477	TACTACGCACACACCACCGCCTCC	AGGCAGACACGTTGGTGGAGGCGG
ZNF14	ENST00000344099	TTCCTGGCCCATGGGAAGATTGCG	TCAGGAGCAGGTGAAACGCAATCT
ZNF539	ENSG00000183850	CTGTCGCCGGAGTCCCAGGTCTGT	ACACAGAGCAGTGAAGACAGACCT

used to create dot-blot filters, which were screened in the same way with individual (non-pooled) overgo probes to differentiate BAC clones based on gene content.

BAC DNA isolation and end sequencing

DNA was isolated for putative minimum-tiling-path BACs using a Qiagen midi-prep kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. End sequencing of all BACs (including those obtained from chromosome walking; see below) was carried out using the T7.29 (5'-GCCGCTAATACGACTCACTATAG GGAGAG) and SP6.26 (5'-CCGTCGACATTTAGGTGACACTATAG) primers for BACs from the CHORI-241 library and T7.19 (5'-TAATACGACTCACTATAGGG) and M13 reverse (5'-CAGGAAACAGCTATGACC) primers for those from the Texas A&M University equine BAC library (http://hbz7.tamu.edu/homelinks/bac_est/bac.htm). Dye terminator sequencing reactions (total volume 10 μl) were set up and performed in a Gene Amp (Applied Biosystems) PCR system 9700 using previously described protocols (Gustafson et al. 2003). Reaction products were purified by passing through spin columns (Spin-50, BioMax, Odenton, MD) and loaded on an ABI 3730 automated capillary sequencer (PE Applied Biosystems) for obtaining the sequences.

STS content mapping and chromosome walking

Primer pairs were designed from end sequences following RepeatMasker analysis and BLAST comparisons. All primers were designed using Primer3 software (http://www-genome.wi.mit.edu/cgibin/primer/primer3_www.cgi), and optimized on horse genomic DNA and DNA from the BAC of sequence origin. The primers were used to screen adjacent BACs in the contig to confirm the positions and orientations of individual BACs. Primer pairs from the BAC end sequences bordering gaps between contigs were used to screen the CHORI-241 and Texas A&M equine BAC libraries to identify new BACs. The latter were subsequently end-sequenced, and primers were designed to screen neighboring BACs (to verify location) and/or for additional library screens for further walking. In some instances, internal sequencing was carried out on BACs adjacent to the gaps to develop additional STS primers for chromosome walking. Detailed information on all STSs and primers generated in this study, along with their PCR conditions, etc., is summarized in Table 5-2.

Table 5-2. PCR primers designed from STSs for STS content mapping and chromosome walking.

STS	F PRIMER	R PRIMER	EXPECTED SIZE
001P14SP6	AATATAACTGCCTGAACTTTGTCA	CATGAAGACTTTTCAATTTGTGC	230
001P14T7	CAGGTAGGACCTTGGGTTCA	CCTCCCTTGCTCACTCTCTG	194
009F01 int 1	TGGTGTATCGTGTGGCAGAT	AACGTACTTGCACGCAACAG	225
009F01 int 2	AGACGGCTGCTTCTGTGAAT	CTCCATGTCAGGAACACACG	177
009F01SP6	AGGGCATTCTTTGTCGCTTA	GCTCAGAACACGCAGCTCTT	208
009I01SP6	CCTTGTTCCAGGCTTCTGTT	CTGCCAACAATTTTGGAGGT	230
011O15T7	CATCTTGTCTTGTGAGTTTGTGG	GGGTTACAAAGCCTTGAACAAA	179
015A14/181B11 int 1	TGCTCCCCACCTCTAGTCAA	TGAGCCCTGCCCAACTGGA	250
015A14/181B11 int 2	AGCTGGTCCAGTGTCCTCAC	AATTCCCCTTTTAAATTTCCAC	195
015A14/181B11 int 3	CACATGTGAGAGGAGACAGCA	TGAAAACTTCGCATCTGGTG	152
015A14/181B11 int 4	GTTTGAGCTTTTGGGGTTTT	TGAAAACTTCGCATCTGGTG	217
015A14T7	GGCAGGAAGATGAGCGACTA	TGGATGACCTCTGTGGTGAA	173
017A17SP6	AAGTGTGGGGTTCCTCTGGT	GGCCTAACCAATTCCTACCC	150
017N12SP6	TGGCTGATGTCAGCAATGAT	TTTGACTTGCTGCTCTTGGA	157
017N12T7	TTCCAGGTCAAAGCTGAAGG	GGCCTCTGAGTATGGGTTGA	157
021G18SP6	CCGAGACATGAATCACCAAA	CAACGGGGAAAGAGTCTGTG	151
021G18T7	ATAGGTGGGAAATGGTGCAG	TCAGGGTTGCTTCCTCAGAC	199
34.1C10M13	TGTTTTGAAGTTTTGATGTTTGG	CATGATGTAATCTACACGAAATTTGA	247
034.1C10T7	GGGGAAAGAAGCTTGCATTAT	ACCCTACCTGGTCACTTGATT	212
038L24T7	CGGCCTAACACTCATTCCTC	CCAAGTTCCTTGAATCTTAACAA	291
046M21SP6	CAGGCTTGACTCTTGGGAAG	AGAAATGCCAGGGAGTTGTG	225
046M21T7	TTAATCACCTGGAGGCCAAG	GAAGTGGGCTGTTCTGCTTC	207
047B23/186F05 int 1	AGTGACTGAGGGGCAGCTAA	ACTCAAAAGTGCCCATGTCC	158
047B23/186F05 int 2	AGCTGCCAAGACCAAACTGT	CTTAGCTGCCCCTCAGTCAC	178
047B23SP6	GGCACCCATAAGGAGTCTCA	GCTCCATCGGAGTGCTGTAT	212
047B23T7	TCATTGCCTTACAGACTGTTGAA	GAAGAGGGATTTGGACTTGC	208
049J07SP6	CACTTTTTCTGAGCGCCTTC	CGCTGTCTGTCTCTCTCCACT	161
049J07T7	TGCCCCTCACACACAATTTA	TGTGCTCCCTCATTTCCTTC	168
056C12SP6	ATAGGGAATTAAGGGCATGG	AGCAGGCGGATTCATGTTT	150
056C12T7	ATGGTCTGGACTCTGGATGC	TGTTTGCAAAGCGTTTCACT	159
060I06SP6	AATGCAGACCCAAGCTCATC	AGTCCATTTCACTGCCCAAC	229
060I06T7	TCCTAACACACAGTGGAGCAG	ACGGTCATTCAGACGAGGAG	52
064I21SP6	TCAATGGGCTTTTCTCTAGCA	CCAAGACCCCAGTTGACTCT	174
064I21T7	CAAGTGTCAATGAGCTCAGGAAT	TTTTCCTTCATAATGTCATCTTTGTT	171
081G24SP6	TTCCAGGCCAATTTAAGTTCA	GCCAAGTCAGGTGGAGAGTG	159
081G24T7	AGAGGGGCGAGCTAGGTAAC	GTCGTCTGGATTTTGCCAAT	158
085N19SP6	ATCCCAGGAGAGGATCTTGG	TGTGTGAGCCCACAACGAC	172
085N19T7	GACTCCACATCTGGCTTTGC	CAGGGGGAGCAGTCTTCAT	103
087G16SP6	AACAACACTCCCTGGCTCAC	AGGAATGGGGTCACTCACAG	158
087G16T7	AGCTTGAGCCCTTAGGCATT	TTTTCCACCCAGGCATTTAG	207
097.1G07T7	GCCAGTCCCACTGGATCTTA	AAGTCAACAGCTCCCCTTCC	160
101K14SP6	ATCAGCCTGCATTCTCAGGT	CTGGAATTGGAGTCCCAAGA	189
101K14T7	GATCTTGGGGGACCTTGAG	CTTTTCTCTGGCCTCCACTG	202
107D18SP6	TTCTAGGGCATTAGGGAGCA	CTGAAGGGAGACTGGCAAAG	188

Table 5-2. Continued.

STS	F PRIMER	R PRIMER	EXPECTED SIZE
109A22SP6	CCTTGCGCTGAGAAACCTAC	ATCGTGTTTGTGGGGAATGT	180
109A22T7	GATCTTGGGACTTGCTGCTC	TGCAGATGGATCAAACCAAA	215
114F19SP6	TCCGGGTGAAGAGAAATGAC	CAGAATGGAGAGGGATTGGA	225
115E14SP6	TGGCACCAGGAATATGAACA	TTGCTTTTGCAGCTTCTTGA	250
115E14T7	TCCACAATCCGAAATTAGCC	AAGGAAGAGTCGTGGCGTTA	153
125O11T7	GGACCCTCTCTGTGGAATGT	CGGTCGGTGTTAGACGTCAG	245
127O05SP6	AGGGTATGCATGAGGTGTCC	AACCCTGTCCACAAGACCTG	226
139F16T7	CGCCTGGGTGATAGAGAGAC	CAGGCTGGTTGGATCCTTTA	151
152N03SP6	GCAGGGACAGTGCTTAGAGG	CTTCCTCGAGCCCTCAGAC	168
152N03T7	TGCCTTCCAGGTCTCTTTTC	CATCTGGGGCTGAAAGGTTA	167
153C05SP6	GACTGCAATTCCCAGTGACA	GTGGGCCAGCAGGTAGAAAC	154
162B03SP6	CTCCCTGGGAATCCTAGCTT	TCACAGGTAGCTGTCGAAATG	152
167K09SP6	TTCAAAGGGAACACATTCCTG	CACACCAGCCTGTCACTCTT	56
168H06SP6	CTTCTACCCAAACCCCACCT	GGGTTACAGGTCAGCAGGAA	153
168H06T7	GCTCCTTCCTTCATCTGCTG	GGGAAGGATCCCTTTGATCT	184
172F16SP6	TGTGCAGGCATCATAAGACC	AGGAACGATTCCATCACAGC	194
177I18SP6	TTCCTGTGTCTTCGTGTTGG	CCAAGTGGTGGTAGGAGGAA	150
177J08SP6	TTTGGGAAGCCTCAGAAGAA	CGTGAGGTGGGTGAGAAAAT	209
177J08T7	CAGACACATGGACCAGCACT	TTTGAAGGCAAAATAACAGGTG	291
181B11SP6	TGATGAATTGATCCTCTGCTG	CTCAGGTGTTGCTTGGGACT	152
182E20SP6	GGACACCCTGGGATAGATCA	CAGAGCAGACATGGGGATG	100
182E20T7	TTGGACCAGAAGTCGAGGTT	AACGTGTCCACTGCCTCTTC	91
186F05 int 1	GCACAGTGGGTGAGACTCCT	CCCCTCTTTAACAACTAAGTAGGC	110
186F05 int 2	CCTTCTGGTCTCCATGGTTT	TCCCAAGTTGAGAGAAGGACA	223
186F05T7	GACCTGCCACCACTGAGAAT	AAAGTGTCCCTCCTGCACAC	197
195I22T7	TCTCCTCTCTGGCTCTTTCG	ATTTGAATGGACTCGGGAAA	100
219J06SP6	AGAGGTGACCCGGGTTTAGT	ACATCACCTCCCTCCCTTCT	117
219J06T7	CCTTCGGTTCATCTCTCGTC	ACGCTGAGGCACCATCTTAT	206
229L17T7	GCCTCACCACCTTGTCCTAA	AGAGGGTCCAAGCTCTCCAG	186
230C10SP6	ATCGGAACTCACAACCCATC	CCGCCTCTTTCTTTTCCTTC	132
230C10T7	GTCCAGCTTCCCTGTGACC	CCTTGACAACGTCCAGTCCT	116
230H17SP6	TCAGAGTGGACCACAGCATC	GTCGGACTTCGTCCTTGTGT	171
230H17T7	CGACCTGGTCATTCAGGAAC	CATGAGAGGGCCCAGAGTT	150
232K12SP6	CAGGGCCACTCATGTGTCTA	ACTCATCAGCCTCCACACCT	103
238A18SP6	GAAAAAGATGGGGCATCTGA	GAACGCTCCTTGTGAATGGT	151
238A18T7	ACAGGGATGCAGAAGTGTCC	GGACTGGGCAGGGAAGATAC	157
238I22SP6	CCTGCATTCGGGTACTGTTT	ACATCCAGCTGACCTCGTCT	141
238I22T7	AGGAAGGAGCCTGACCCTAC	CTGTTTCAGGAAACCCAAGG	189
242J14T7	GCAGAGGCAAGAGAAATTGG	TCCCATCAGGAATCCAGAAA	152
265C11SP6	TTTCACAGTTTTGGCAATATGAA	TGCTTGGGGGTTAATAGCAC	102
265C11T7	GGGGGCACTATGCTTTATCA	GGAAATGACCGTCCTTAGCA	182
276A04SP6	CCACTGGAACTGTGGGAGAT	AGCAGCAGGAAATCTCACTTG	194
311I15SP6	CGACCTGGTCATTCAGGAAC	CATGAGAGGGCCCAGAGTT	150
311I15T7	CAGCTCACCTCGGAGTACAG	CTCAGAGGTGGTGGCAGAC	153

Table 5-2. Continued.

STS	F PRIMER	R PRIMER	EXPECTED SIZE
325D24SP6	AAACTGAGGCCAGGAGAGGT	ATTCCCAGGATGAAGGTTCC	123
325D24T7	CTTAGTCAAACCCGGCCATA	AAACAGGGAGACACCCAAGA	203
326D04SP6	ACACCAGAAAATGCCACCAT	GTGCTTGGGACTGAAGGAAG	216
326D04T7	GGAGAGTGACGTCAGCAACA	TCCGATGTATAGCAGCATGG	153
338D15SP6	CTAGGGGGTGAGCACAAAAG	TGCCTATGCCCTCATCTTTC	163
338D15T7	TGCACATCCCTTGTAGGTGA	CCCTACATCATCCTCGCAAT	176
394J21SP6	GTAGGTGGGCCCCTAAACTC	GAAGGACCATCCTCAGACCA	109
394J21T7	CTGAACTCCAGCACCATCCT	TACCGTGCACTTGGTGATGT	153
408C14SP6	TCAATGGGCTTTTCTCTAGCA	CCTAACCATCCAATGCCAAG	221
408M24SP6	GGAAGAAATTCCCGGCTTAG	GGCCCTTTTTCTCTTTGCTT	242
408M24T7	AGGCCACTCTCTGCACTACG	CTCAAGCCAAGCAGGAAAAC	178
412B24SP6	TAGGAGTTGCCTGGTCTGCT	ACTTGTGGAGTGCCCAAAAC	156
431M22SP6	TATCGTCAGGGAAAGGCAAC	TGCTTCCATCTCATCATTGC	132
431M22T7	CATAAATCCAGCCCCTGGTA	TCCTGTCGGTCAGCCTTAAT	156
493B19SP6	CAAACCAATGCCATTTTGAA	CACTGGCCATGTCTCAGCTA	158
501I16T7	CGATCCTTTTCTGGGATTCA	TTCGGCTGGAAAGAGAGAAA	202
506J16SP6	ATATCGTAACGCACCAAGCA	AGTCCTCTCGGCAATTCAAA	151
506J16T7	ACAAGTGAGTCCCCATGTCC	CTGGCCACCTTCATCTAGGA	165

Primers for further sequencing of BACs:

Originating STS	Sequence	BAC(s) used for sequencing
009I06SP6	TTTTGGGAGAGTGGTAGTCACATA	047B23; 186F05
009I06SP6	CCAACAATTTTGGAGGTTAGATTT	047B23; 186F05
047B23SP6	TATTGGTCCTGTAGCCTGTAGTGA	186F05
047B23SP6	TCCTTGAGAAGGCTGGACTTATAC	186F05
087G16SP6	TCCTCACTCTGTCACGAAACTAAG	015A14; 181B11
087G16SP6	AGGAATGGGGTCACTCACAG	015A14; 181B11
238I22T7	AGCTGCCACTCTCTTCCTCTATTA	009F01
431M22SP6	TTGGTCCGTGGTGTCTTTTATTA	015A14; 181B11
431M22SP6	AAACAGAAATCCTTAATTTCAATGC	015A14; 181B11

Fluorescence* in situ *hybridization (FISH)

About 1µg of DNA from selected BACs was individually labeled with either biotin or digoxigenin using the Bio- or DIG-Nick Translation Mix (Roche Molecular Biochemicals). Labeled probes were hybridized either separately or in pairs to horse metaphase chromosome spreads or interphase cells to verify their locations and to estimate approximate physical distances between BACs flanking a gap. *In situ* hybridization, signal detection, microscopy, and image analysis were performed as previously described (Chowdhary et al. 2003).

Designing gene specific PCR primers and verifying the identity of amplicons by sequencing

Primer pairs for 33 genes expected to be present in the region were developed to establish their likely physical order on the contig. Of these, 21 primer pairs were designed from horse ESTs. The remaining 12 pairs were designed by aligning sequences for orthologs from various mammalian species as described above. Following PCR, the amplified products were sequenced as described above to confirm the identities of individual genes. Detailed information on markers, primers, PCR conditions, etc. is summarized in Table A5-3.

DNA fingerprinting and contig assembly

A recently reported automated, capillary electrophoresis-based fingerprinting method (Xu et al. 2004, 2005) was used for additional independent verification of the BAC contig map. A three-restriction enzyme and one-labeling color kit (Xu et al. 2004) was employed for fingerprinting experiments. BAC DNA was digested with *Hin*dIII,

*Bam*HI, and *Hae*III. The fragment ends of *Bam*HI and *Hin*dIII were labeled with ddATP-NED or ddATP-HEX and the labeled fragments were fractionated on an ABI 3100 capillary sequencer. The BAC fingerprints were edited and assembled into contigs using the FPC program (Soderlund et al. 2000; Xu et al. 2004, 2005). The automatic contigs initially assembled were examined to remove the questionable clones (Qs) and to disassemble the chimeric contigs. Singletons were added to the contigs and overlapping neighbor contigs merged into larger contigs with the FPC program.

Results

Screening of BAC library with overgos and verification of BAC identity

Overgo primers were developed for 58 genes uniformly distributed on the 14.0 to 20.0 Mb sequence map of HSA19 and 64.1 to 68.8 Mb region of HSA5. These human chromosomal regions share homology with the proximal one-third of ECA21. Screening the CHORI-241 equine BAC library with pools of 20 overgo probes per hybridization gave a total of 337 positive BAC clones for the loci. Next, screening of secondary filters from these BACs with individual overgo probes helped to associate 243 BACs with 47 of the 58 genes. On the basis of their common gene content, it was possible to group a number of BACs into clusters that would potentially overlap. The remaining 94 BACs seemed to be false positives and were therefore discarded from future analysis. No BACs could be obtained for 11 genes viz., CRLF1, CYP4F11, KIAA0892, LOC388515, LOC388524, LOC390898, LOC390913, LOC440515, LOC91120, ZNF14, and ZNF539.

The presence of the above specified 47 genes in the remaining BACs was further verified by PCR, using equine specific and heterologous primers for individual genes, and sequence analysis of the heterologous amplicons. Seven of the genes did not show consistent amplification with BAC DNA and were therefore discarded from next phase of analysis. Next, to ensure that the equine genomic region corresponding to HSA19 is accurately defined on ECA21, BACs for genes lying towards the end of the homologous segment were mapped by FISH to horse metaphase chromosomes. This led to the removal of 7 more loci because they mapped either to ECA7 (HSA19 homolog) or to ECA14 (HSA5 homolog). An additional 5 loci from the ECA21 region corresponding to HSA5 were removed because they were located >1 Mb away from the HSA19-HSA5 evolutionary breakpoint on ECA21 investigated in this study. Last, two more genes were deleted because BLAST analysis did not provide a reliable match of their amplicon sequences with the expected human ortholog. Finally a total of 25 genes contained in 190 BACs were available for further analysis. The coverage and distribution of these genes over the corresponding sequence template of HSA19 and HSA5 were examined. Following close scrutiny, 17 additional genes were chosen to provide a uniform representation of the human orthologs at regular intervals along the ~5 Mb segment on ECA21. Equine ESTs were identified for the 17 human genes and PCR primers were designed, eventually leading to their assignments to specific BACs. Further, 5 microsatellite markers previously mapped to this region of ECA21 (Brinkmeyer-Langford et al. 2005) were similarly assigned to individual BACs in the pool. Thus, the total number of markers contained in the 190 BACs increased to 47. Complete

information on these markers, their primers and PCR conditions, etc. is summarized in Table A5-3.

STS development, STS content mapping and chromosome walking

The 190 BACs obtained above were clustered into groups based on common sets of marker content. Using comparative markers as alignment points, individual clusters were arranged on the HSA19/HSA5 sequence template. This provided a provisionally ordered set of contigs on the proximal one-third of ECA21. Representative overlapping BACs from each contig were identified, resulting in a collection of 47 BAC clones that were selected for expansion of individual contigs and for filling gaps between them. The BACs covered the majority of the HSA19 homologous region on ECA21, the HSA19/HSA5 breakpoint, and part of the adjacent HSA5 homologous region. DNA was isolated for the 47 BAC clones. End sequences from these clones were analyzed, leading to the development of 74 sequence-tagged sites (STSs; Table 5-2). No STSs could be developed from the remaining 20 end sequences because they largely contained repetitive elements. Primer pairs for individual STSs were used for content mapping by PCR on DNA from specific BACs. The findings permitted the arrangement of the 47 BACs into 4 contigs spanning the targeted region on ECA21.

The gaps between the contigs were filled by chromosome walking as described earlier (Raudsepp et al. 2004b). Briefly, STSs from end sequences of the most proximal BACs of the contigs flanking a gap were used to screen the CHORI-241 and/or TAMU equine genomic BAC libraries by PCR. This provided new BACs that were end sequenced to obtain new STSs for repeating the process until the gap was covered with

overlapping BACs. In total, 17 BACs were obtained via chromosome walking (15 from the CHORI-241 library and 2 from the TAMU library). End sequencing of these BACs added 28 STSs to the contig. These STSs validated the locations of existing and new BACs, and helped to orient the new clones in relation to others. All gaps (2) except the one at the evolutionary break/fusion point between HSA19 and HSA5 were filled with overlapping BACs. Finally, two contigs were obtained over the targeted ECA21 region: the larger over the region corresponding to HSA19 ("contig A"), and the smaller over the adjacent part from where homology with HSA5 begins ("contig B"; Fig. 5-1). The gap signifies the evolutionary breakage/fusion point and its putative size is discussed below. Overall, the two contigs have a total of 207 BACs (196 in contig A and 11 in contig B), of which 22 and 4 formed the minimum tiling paths for contig A and contig B, respectively. A total of 106 STS markers were generated during the course of contig development, of which 87 are present in contig A and 19 in contig B.

Fluorescence* in situ *hybridization (FISH)

FISH on equine metaphase chromosomes was carried out to verify the chromosomal locations of some of the loci present at the boundary of homology between HSA19/HSA5 and corresponding segment on ECA21. As indicated in previous sections for 7 genes previously presumed to be located on ECA21, FISH mapping of local BACs showed that 5 of the loci are actually present on ECA7 (ASF1B, DDX39, PTGER1, EMR3, and FLJ40365) and 2 on ECA14 (OCLN and TAF9). Thus, BACs for

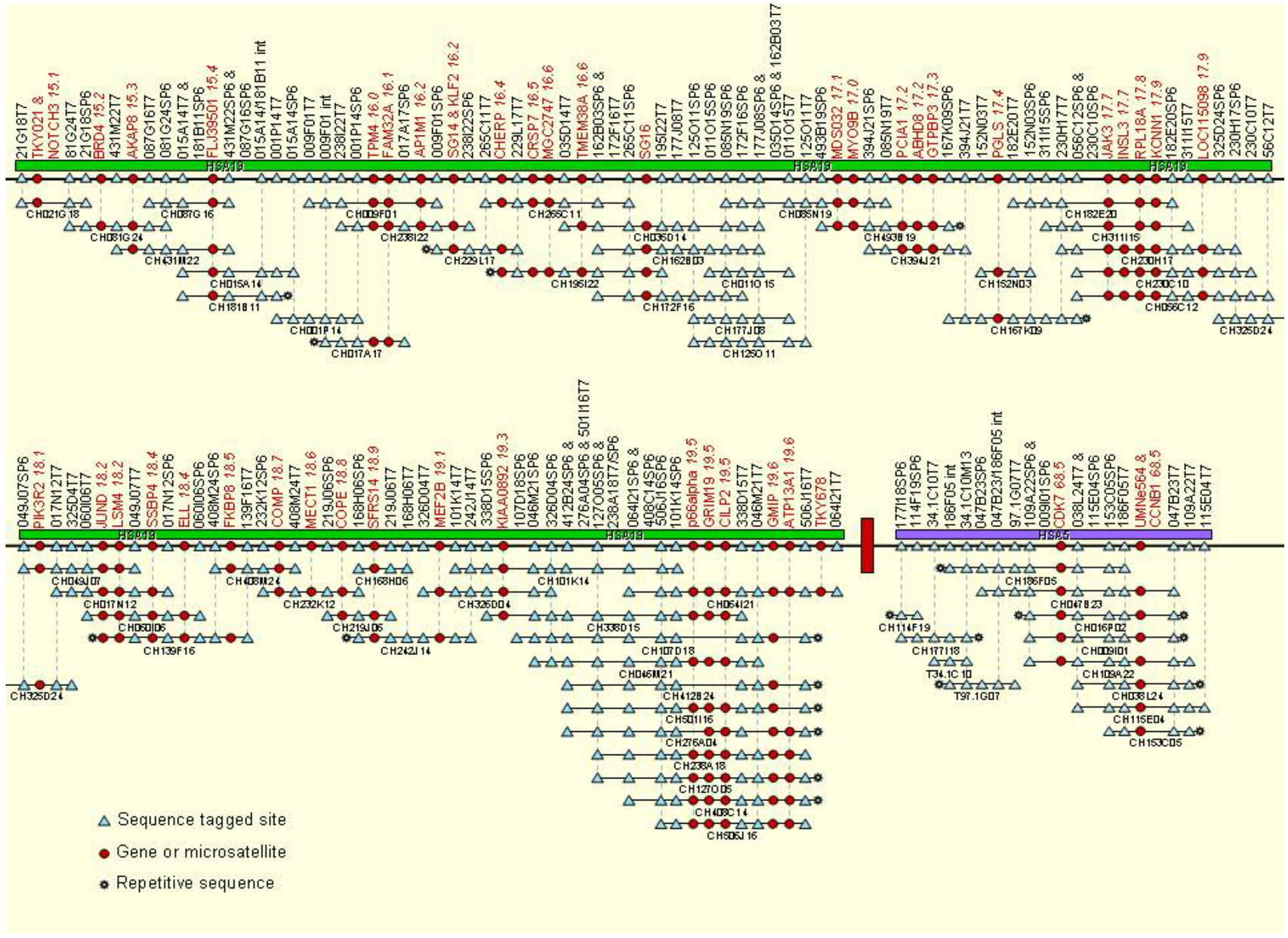


Figure 5-1. BAC contig along the section of ECA21 which corresponds to parts of HSA19 (green bar; referred to in text as contig A) and HSA5 (purple bar; referred to in text as contig B). Blue triangles denote STSs; red circles denote genes and microsatellites; and starbursts indicate repetitive end sequences

these loci were excluded from development/expansion of the contig. Further, metaphase FISH using BACs containing NOTCH3, CDK7, and CCNB1 showed that these genes are indeed present within the region for which the contig is being developed.

FISH was also carried out to estimate physical sizes of two of the gaps during the course of contig development. Dual-color metaphase and interphase FISH approaches indicated that gaps between markers FLJ39501 and TPM4 and between ATP13A1 and CDK7 did not exceed more than ~200 kb, as denoted by overlapping signals on metaphase chromosomes that were further corroborated by closely located signals on interphase chromatin (Fig. 5-2 d, e, and f). STSs from BACs flanking the first gap (between FLJ39501 and TPM4) eventually led to the identification of BAC clones that ultimately filled the gap. New BACs identified through chromosome walking for these gaps were also labeled and hybridized to confirm their locations within the gaps. Finally, the only remaining gap between contig A and contig B is rather small (not exceeding ~100-150 kb). This is evident from dual color interphase FISH results using BACs 506J16 and 114F19 (Fig. 5-2 e and f). These BACs face each other across the gap (Fig. 5-1).

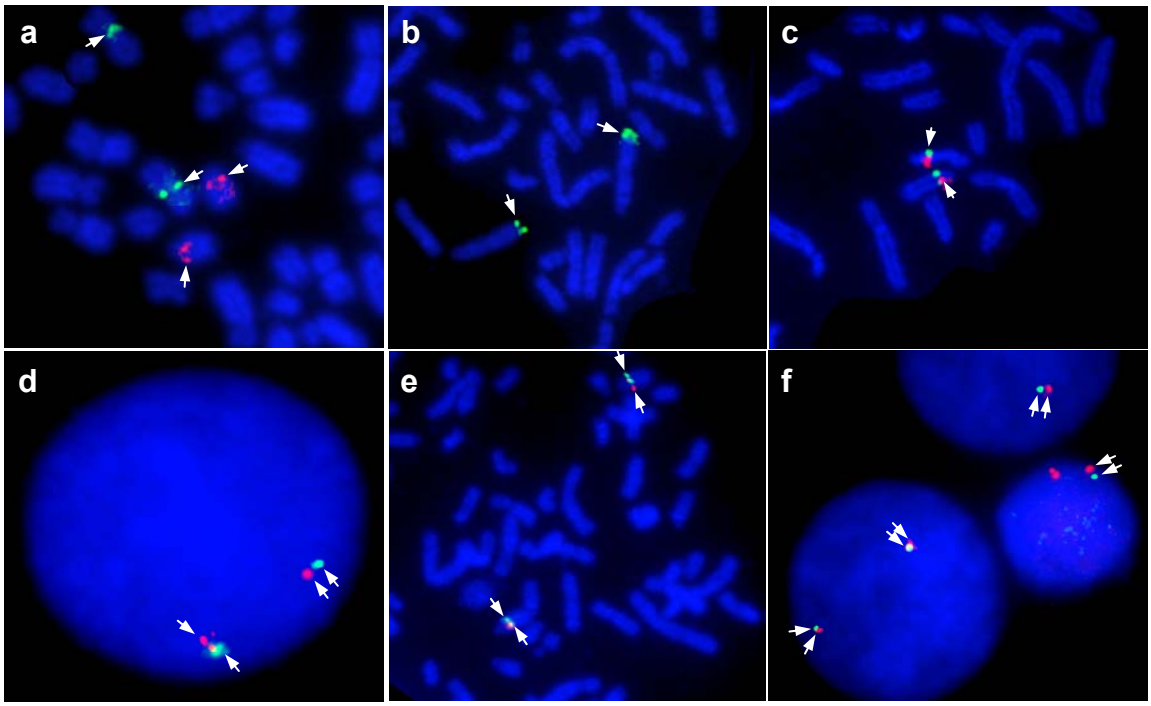


Figure 5-2. Partial horse metaphase spreads or interphase chromatin showing (arrows) single- and double-color FISH results for selected loci. a: CH329M17 (green) and 009F01 (red), b: CH035N16 on ECA14ter, c: NOTCH3 (red) and GMPI green), d: CH015A14 (green) and 009F01 (red), e: CH506J16 (red) and CH144F19 (green) showing size of remaining gap in contig, f: CH506J16 (red) and CH144F19 (green) showing size of remaining gap in contig.

Fingerprint analysis

A total of 26 BACs comprise the minimum tiling path on contigs A and B. Fingerprint analysis pooling the BACs from the two contigs clearly showed that the minimum tiling path obtained via STS content mapping did indeed accurately represent the region, resulting in the assembly of the BACs into 2 contigs. The fingerprint map affirmed the contigs obtained by STS content mapping and provided added confidence and robustness to the map.

Discussion

Contig overview

The development of this contig began with 47 overgoes designed for genes in the region of interest, which returned positive BACs after screening CHORI-241 BAC library filters. Provisional contigs were established from the positive BACs and chromosome walking provided additional BACs to close the gaps. Metaphase and interphase FISH were then used to determine the boundaries of the region based on the locations of BACs at the far ends of the contigs. Sequence-tagged sites were used to screen the BACs and validate the contigs. Additional genes, microsatellites, and STS markers were added to the contigs via PCR to make the map more informative. Interphase FISH using BACs extending into the one remaining gap in the region, which encompasses the breakpoint in homology between HSA19 and HSA5, indicate that it is around 30-40kb wide. Finally, BAC fingerprinting was performed for contig verification. The final result is a contig containing 104 new STSs, 42 genes (of which

33 are newly added to the map), and 5 microsatellites, which provides a thorough coverage of the region.

The contig described in this study provides a detailed map of the proximal segment of ECA21 that corresponds to the 15.1 to 19.7Mb sequence positions of the proximal part of the short arm of HSA19. The 26 BACs in the minimum tiling path give the contig a span of ~4.3 Mb. Presently, it is the largest contig available in horses and is over twice the size compared to the only other region in the horse for which a contig has been established, the equine major histocompatibility complex (Gustafson et al. 2003). BAC contigs have also been constructed in at least four other domesticated species for which whole-genome sequence (beyond draft-level) is not currently available. These include several regions in the pig, such as the region on porcine chromosome 7 which contains quantitative trait loci related to fat (Barbosa et al. 2004; Demars et al. 2006; Tanaka et al. 2006; Sato et al. 2006), and the major histocompatibility complexes of sheep and cats, with three contigs spanning approximately 1900, 400, and 300 kb covering the region in sheep (Liu et al. 2006) and two contigs of ~2.85 Mb and ~0.50 Mb built for the same region in the cat (Beck et al. 2005). In addition, six contigs have been developed for the Polled Intersex Syndrome region of the goat; these are approximately 400, 1000, 1500, 1000, 400, and 300 kb in length (Schibler et al. 2000). Contigs such as these are essential for anchoring sequence information to physical maps and provide nucleotide-level resolution for fine mapping regions of interest.

The development of a contig for the segment of ECA21 which correlates to HSA19 has significantly increased the level of map detail in the region, providing

additional markers to improve the overall utility of the map. The previously-established high-resolution radiation hybrid (RH) map for this segment had an average of 1 marker per 380kb; however, the high density of markers in the region prevented their accurate ordering with the 5000rad horse x hamster radiation hybrid panel and metaphase FISH (Brinkmeyer-Langford et al. 2005). Despite this, overall marker order of the RH and FISH maps appear to be in agreement with that found in this study. The map reported herein shows a ~3.8-fold improvement in resolution over the RH map, with 1 marker mapped per ~100 kb. Even at this density, not every gene in this very gene-dense region was selected for mapping. Rather, one gene was selected every ~100kb from the corresponding human region; more closely packed genes could not have been adequately resolved with respect to each other. Additionally, five previously described microsatellite markers were added to the contig, giving them more precise physical locations with respect to the mapped coding genes. The order and positions of three of these microsatellites (TKY021, SG14, and SG16) are in agreement with the most recent linkage map developed for the horse (Swinburne et al. 2006; TKY678 and UMNe564 are presently not on any linkage map).

Comparative

The equine genome is estimated to be around 10% smaller than that of humans. This has been found to hold true for the section of the genome on ECA21 which shares homology with HSA19. The region covered by the contig on ECA21 is approximately 4.3 kb in size, 10% smaller than the corresponding region in humans, which is around 4.7kb (~4.6kb HSA19 and ~0.1kb HSA5). Summing the lengths of synteny blocks for

several species gave estimates of the approximate sizes of the regions homologous to the segment of ECA21 corresponding to HSA19 (however, it should be noted that this method of estimation does not take into account individual sequence variations). This region also appears to be smaller in size than in humans for the other mammalian species compared: chimpanzee, ~4.0Mb (PanTro 1.0); cattle, ~3.6Mb (Btau 2.0); rat, ~3.0Mb (RGSC 3.4); dog, ~2.6Mb (CanFam 1.0); and mouse, ~2.3Mb (NCBI m35). The size differences between species can be attributed to any of several factors. Firstly, the region is enriched for various repetitive elements, particularly short interspersed nuclear elements (SINES) in humans, so there may be differences between humans and other species (including horses) in the number of repetitive elements within and between genes (as seen between human and mouse; Dehal et al. 2001). Next, much of the region (over 25% in humans) is comprised of gene family members, many of which have arisen through tandem *in situ* duplications of ancestral copies (Dehal et al. 2001; Grimwood et al. 2004) resulting from their propensity for nonhomologous pairing. Lineage-specific changes within these families may also result in differences in lengths of the corresponding regions.

One such gene family, a cluster of olfactory receptor genes, is located at the 14.7-14.9Mb position in humans. This region probably encompasses the division in HSA19 homology between ECA7 and ECA21. Two more families are located at the human positions 15.6-15.8Mb and 15.5-15.9Mb; in humans these contain 5 members of the OR10H olfactory receptor family and 9 members of the CYP4F gene family, respectively (Grimwood et al. 2004). Due to its lineage-specific nature, human

homologues in this region could not be mapped in the horse. In addition to these three gene family-rich regions, humans and chimpanzees possess a 4.6Mb cluster of ZNF genes, located in the pericentromeric region of HSA19p from ~19.5-24.1Mb. This region is generally accepted to be primate-specific (Bellefroid et al. 1995; Eichler et al. 1998; Dehal et al. 2001; Grimwood et al. 2004) and is therefore not found on the corresponding contigs of the horse or any other non-primate species.

Overall HSA19 gene order (based on NCBI build 36) appears to be conserved in horses and chimps, where the region of interest is present on a single chromosome in each. Gene orders of the homologous regions in dog, cattle, mouse, rat, chicken, and zebrafish are not as well conserved. In general for the dog and chicken, the entire region is intact though inverted. There are several breakpoints in the region involving inversion or transposition events in cattle, mouse, and rat.

Interestingly, the region of interest is situated differently in each species with regards to homology with humans. As illustrated by the contig described in this study, in horses, the end homologous to human position 15.1 Mb is adjacent to the centromere; the other end of HSA19 homology is next to the 68.5Mb region of HSA5. As expected, chimp chromosome 19 and HSA19 share an almost one-to-one homology along the entire lengths of the chromosomes. In dogs, the entire p arm of HSA19 is located on dog chromosome 20, with the end of the segment homologous to position 19.7 Mb on HSA19 associated with a region homologous to the 44.7 Mb sequence location next to the centromere on HSA3q. Predictably, HSA19p homology is scattered in both mouse and rat, with segments frequently bounded by lineage- or rodent-specific regions.

The breakpoint between HSA19 and HSA5 homology on ECA21 is of particular interest because of its Perissodactyl-specific nature. While HSA19p homology is present on a single chromosome in most mammalian species, it has been found on two separate chromosomes in Perissodactyls (Yang et al. 2003b, 2004; Brinkmeyer-Langford et al. 2005). In cattle, HSA19p is located entirely on chromosome 7, though the 2.2-19.5 Mb sequence segment of HSA19 is separate from the 0-1.9Mb segment (Everts-van der Wind et al. 2005). Both ends of the 2.2-19.5Mb block are bounded by regions homologous to HSA5; however, neither is the same as in horses. The 179.7Mb sequence position is next to the HSA19 19.5Mb position at the distal end in cattle, while the proximal end of the HSA19 block is bounded by the 132.0Mb position of HSA5. In each of the Perissodactyl species studied (horses, donkeys, Burchell's zebra, and Hartmann's zebra), the ~15.1-19.6Mb HSA19p-homologous segment is adjacent to a region corresponding to HSA5pter-5p13 (68.5Mb sequence position). While HSA19p and HSA5 share adjacent homology in artiodactyls, the corresponding HSA5 region is different; in no other mammalian species has this particular connection been observed.

Lineage-specific gene families are frequently observed at breakpoints in each of the species examined. The presence of olfactory receptor, CYP4F, and ZNF gene families and L1 repeat or LTR sequences at the margins of syntenic rearrangements is not uncommon for HSA19 homologous segments (Grimwood et al. 2004), suggesting that non-homologous recombination may have been a driving force in the evolution of the region (Dehal et al. 2001). The olfactory receptor gene family cluster at human position 14.7-14.9Mb is probably located at the place where HSA19p homology is

divided between ECA7 and ECA21 (genes within this span could not be mapped in the horse). A breakpoint at this location is not observed in dogs or cattle, though it is present in the mouse and rat. Additionally, the OR10H olfactory receptor and CYP4F gene families located within the 15.5-16.0Mb sequence segment of HSA19 are located at synteny breakpoints in cattle, mouse, and rat. In cattle, this breakpoint is intrachromosomal; in mouse and rat, it is interchromosomal. It is interesting to note that an intrachromosomal boundary is also present within the section homologous to the 16.6-17.0Mb segment in humans in each of these three species, despite the lack of gene family presence there; there are relatively few genes in this region (none of which were mapped in this contig). Presumably there is some sequence feature that has predisposed this section to breakage. Other synteny breakpoints not associated with gene families include three breakpoints in cattle, localized between the human positions of 17.30-17.48Mb, 18.12-18.28Mb, and 19.43-19.48Mb, and one other breakpoint in mouse located between the human positions 17.78-17.83Mb. Finally, the ZNF family located on HSA19 from ~19.5-24.1Mb demarcates another popular site for synteny breakpoints: each of the mammalian species studied here contains a boundary within this region. This includes the horse, where this particular breakpoint defines the boundary in homology between HSA19 and HSA5. The other side of this boundary is homologous to the 68.5Mb sequence position on HSA5. Breakpoints in HSA5 homology are found in similar locations in the pig (71.65Mb sequence position in humans; Lahbib-Mansais et al. 2006) and dog (74.2Mb sequence position in humans).

The development of this contig on ECA21 represents a significant step towards achieving nucleotide-level detailed resolution for a region in the horse by providing a template on which sequence information can be assembled. In addition to learning more about the evolution of ECA21 and other homologues of HSA19 in the horse, the level of detail provided by the contig facilitates the identification of species- and lineage-specific features. Another benefit will be improved accuracy in discovering functional elements, which have been conserved throughout evolution. The ENCODE project seeks to identify and describe all functional elements in the human genome sequence (ENCODE Project Consortium 2004), an endeavor which is facilitated by the comparison of sequences of homologous regions in many, distantly-related species (Margulies et al. 2005). Although this particular region is not currently selected for analysis by the ENCODE project—presumably due to its extraordinarily high gene density—the presence of medically-important genes, at least one cancer breakpoint involving the gene ELL, and its Perissodactyl-specific nature may make it of special interest in future searches for regulatory elements. The addition of data from the horse to the ever-expanding phylogenetic tree of mammalian species with sequence information will further increase the utility of this contig for identifying and characterizing functional elements in the region which will ultimately form targets for disease research and drug development.

VI. EXCEPTIONAL CONSERVATION OF HORSE-HUMAN GENE ORDER ON X CHROMOSOME REVEALED BY HIGH-RESOLUTION RADIATION HYBRID MAPPING*

Synopsis

Development of a dense map of the horse genome is key to efforts aimed at identifying genes controlling health, reproduction, and performance. We herein report a high-resolution gene map of the horse (*Equus caballus*) X chromosome (ECAX) generated by developing and typing 116 gene-specific and 12 short tandem repeat markers on the 5,000-rad horse x hamster whole-genome radiation hybrid panel and mapping 29 gene loci by fluorescence *in situ* hybridization. The human X chromosome sequence was used as a template to select genes at 1-Mb intervals to develop equine orthologs. Coupled with our previous data, the new map comprises a total of 175 markers (139 genes and 36 short tandem repeats, of which 53 are fluorescence *in situ* hybridization mapped) distributed on average at ~880-kb intervals along the chromosome. This is the densest and most uniformly distributed chromosomal map presently available in any mammalian species other than humans and rodents. Comparison of the horse and human X chromosome maps shows remarkable conservation of gene order along the entire span of the chromosomes, including the location of the centromere. An overview of the status of the horse map in relation to

---

*This section is reproduced with permission from Proceedings of the National Academy of Sciences, USA, Exceptional conservation of horse-human gene order on X chromosome revealed by high resolution radiation hybrid mapping, Terje Raudsepp, Eun-Joon Lee, Srinivas R. Kata, Candice Brinkmeyer, James R. Mickelson, James E. Womack, Loren C. Skow and Bhanu P. Chowdhary, volume 101, 2386-2391, © 2004 by the National Academy of Sciences, U.S.A.

mouse, livestock, and companion animal species is also provided. The map will be instrumental for analysis of X linked health and fertility traits in horses by facilitating identification of targeted chromosomal regions for isolation of polymorphic markers, building bacterial artificial chromosome contigs, or sequencing.

Introduction

Equine genome analysis has proceeded at an unprecedented pace during recent years. From the initial horse (*Equus caballus*, ECA) gene map 6-7 years ago, organized international efforts have led to a ten fold expansion in the map. This is evident from the recently published meiotic (Guérin et al. 1999, 2003; Swinburne et al. 2000), cytogenetic (Milenkovic et al. 2002) and radiation hybrid maps (Chowdhary et al. 2003) that provide an array of polymorphic and gene specific markers distributed over all equine autosomes and the X chromosome. Comparative information available through these maps is proving critical for accurate alignment of the horse genome with the sequenced genomes of human and mouse (Chowdhary et al. 2003; Milenkovic et al. 2002) and with the gene maps of other livestock species. The maps, most of which are of low to medium-density, have served as a starting point to initiate research aimed at identifying genes responsible for valuable traits associated with equine biology, health and performance including genes responsible for base coat color, overo lethal white, hyperkalemic periodic paralysis, and severe combined immunodeficiency (Marklund et al. 1996; Rudolph et al. 1992; Santschi et al. 1998; Wiler et al. 1995). Further, marker based studies are in progress to dissect molecular causes of various coat colors [e.g., grey, (Henner et al. 2002; Locke et al. 2002; Swinburne et al. 2002); appaloosa, (Terry et al. 2001, 2002)], genetic diseases

[e.g., exertional rhabdomyolysis, (MacLeay et al. 1999a, b); polysaccharide storage myopathy (Valberg et al. 1996)] and other traits of interest. The major impediment associated with the latter group of studies is the lack of adequate resolution maps for individual equine chromosomes. Such maps could facilitate rapid targeted hunts for candidate genes associated with the traits, once they are mapped by genetic linkage analyses with highly polymorphic markers.

The X chromosome is the most conserved mammalian chromosome (Charlesworth 1991; Ohno 1967). Extensive analyses/comparisons of structure, organization and gene content of this chromosome in evolutionarily diverse mammals have revealed a remarkable degree of conservation (Graves et al. 2002; Murphy et al. 1999a; Raudsepp et al. 2002). Until now, the chromosome has been best studied in humans and mice, where the focus of research has been the intriguing patterns of X-inactivation and the involvement of various X-specific genes in genetic diseases (Boyd et al. 2000; Lyon 2002), female and male fertility (Vaiman 2002; Vialard et al. 2002; Wang et al. 2001) and embryonic development (Burgoyne et al. 2002). Relatively very little applied research has yet been conducted in livestock and companion/pet species. This is primarily attributed to the limited information available through the medium to low resolution X chromosome maps that are presently available in pigs (McCoard et al. 2002; Rink et al. 2002), cattle (Amaral et al. 2002; Iannuzzi et al. 2000), cats (Menotti-Raymond et al. 2003b; Murphy et al. 1999a), dogs (Breen et al. 2001; Everts et al. 2002; Guyon et al. 2003; Kirkness et al. 2003; Spriggs et al. 2003) and sheep/goats/buffalo (Iannuzzi et al. 2000; Piumi et al. 1998; Schibler et al. 1998), which are insufficient to permit comprehensive studies aimed at dissecting traits or phenomenon of interest.

The current horse X chromosome (ECAX) map is comparable to those of other livestock and companion species. Beginning with the physical assignment of three gene specific loci viz., *G6PD, HGPRT* and *PGK* during early 1970s (Deys 1972), the ECAX map developed mainly through synteny (Caetano et al. 1999b; Shiue et al. 2000) and genetic linkage mapping approaches (Swinburne et al. 2000). During recent years, the map has benefited considerably through the availability of the RH panel (Chowdhary et al. 2002). The current ECAX map comprises 26 gene and 17 short tandem repeat loci (Chowdhary et al. 2003; Raudsepp et al. 2002), offering a basic RH map for comparative analysis of ECAX. However, the resolution of this map is inadequate to facilitate identification of economically important genes over the ~153 Mbp span of the chromosome (Chowdhary et al. 2003).

Building on our recent success in developing high resolution gene maps for ECA17 (Lee et al. 2003a), ECA22 (Gustafson-Seabury et al. 2005) and a targeted region of ECA26 (Ward et al. 2003), we undertook development of a high-resolution gene map of ECAX by stepwise selection of gene specific markers from the human and mouse X-chromosome sequence templates. This effort has resulted in a dense and comprehensive map that is second only to the highly detailed maps in human and mice, and will lay the foundations to identify and analyze X-linked genes involved in equine reproduction and genetic disorders.

Materials and methods

Marker selection and primer design

The human genome sequence data available from NCBI build 34 of the human reference sequence (http://genome.ucsc.edu/cgi-bin/hgGateway) and the Ensembl l version (also from July 2003; http://www.ensembl.org/Homo_sapiens/) were jointly interrogated for known genes from the human X chromosome. The genes were selected at approximately 1 Mb intervals, beginning at 0 Mbps (distal tip of the short arm) and ending at 153 Mbps (distal end of the long arm). Whenever possible, human exonic sequences from selected genes were compared with orthologous sequences from mouse, rat, cattle, pig and other available mammals using BLASTn ***nr*** and ***est_others*** search engines (http://www.ncbi.nlm.nih.gov/BLAST/). This was followed by multiple alignment of the sequences in CLUSTALW (http://bioweb.pasteur.fr/seqanal/interfaces/clustalw-simple.html). The alignments were used to design heterologous primers for PCR amplification of horse DNA in a hamster DNA background, as described previously (Jiang et al. 1998, 2002; Lee et al. 2003a). Briefly, all primers were derived either from a single exon or from two adjacent exons leaving an ~ 500-700 bp intron in between, and chosen for 100% sequence identity among human, cattle, pig etc. orthologues, but with 1 to 3 mismatches with the rodent (mouse, rat) sequences. Primers were designed using the Primer3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). A total of 88 gene specific markers were developed using this approach. Additionally, species specific primers for all microsatellites (12) and some genes (28) were designed based on the available equine genomic or EST sequences. Detailed information on all new markers (128) is presented in Table A6-1.

Primer optimization and sequencing

Horse and hamster genomic DNAs were used to optimize the PCR conditions for individual primer pairs, such that only horse specific DNA amplification was obtained and all equine PCR amplification products were verified by sequencing. The identities of the sequences were confirmed through BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) searches as described earlier (Lee et al. 2003a) and all comparisons were revalidated against the latest built of sequence data at UCSC and Ensembl. For *RPS6KA3* (see Table A6-1), new primers were designed using sequence data obtained from the first primer pair. This provided robust horse-specific amplification compared to the weak amplification observed using the initial primer pair. The identity of this PCR product was revalidated by sequencing, as described above.

RH typing analysis, bacterial artificial chromosome (BAC) library screening, and fluorescence* in situ *hybridization (FISH) mapping

PCR typing on the 5000_{rad} horse x hamster RH panel and data analysis was performed as described (Chowdhary et al. 2003). The TAMU and CHORI-241 equine genomic BAC libraries were screened by PCR to obtain clones specific for 29 ECAX genes. The selection of these markers ensured physical anchors for the RH map at regular intervals along the entire length of the chromosome. Therefore, loci were chosen in regions having no or very few FISH mapped markers. Individual BAC clones were labeled with biotin and/or digoxigenin using BIO- and DIG-Nick Translation Mixes (Roche Molecular Biochemicals) and separately hybridized to horse metaphase chromosomes to confirm probe origin and determine precise physical locations.

Additionally, 9 closely positioned or overlapping markers in the RH map were co-hybridized in differently labeled pairs or triplets on metaphase and interphase chromatin to refine their relative physical order. *In situ* hybridization, signal detection, microscopy and image analysis were carried out as described (Chowdhary et al. 2003).

Results

Generation of gene-specific markers on ECAX

A total of 167 equine gene-specific primer pairs were designed by using the human X chromosome sequence map (Ensembl, www.ensembl.org; Human Genome Browser, http://genome.ucsc.edu/index. html?org=Human&db=hg16&hgsid=27764832). After the first round of optimization, 40 primer sets were excluded due to weak amplification of horse DNA, multiple PCR products, or amplification products of equal sizes for horse and hamster. Of the remaining 127 primer pairs, 28 were identified from equine EST or gene sequences, and the remaining 99 originated from multiple alignments of mammalian sequences (Table A6-1). Sequencing the PCR amplification products of individual primer pairs to verify the identity of the markers resulted in 11 primer pairs being discarded, because the sequences did not correspond to the expected genes. This yielded 116 equine orthologs for human X chromosome genes and represented a ~70% success rate in developing horse-specific markers for use on the 5,000- rad RH panel.

Generation of a composite RH map

A total of 128 new markers (116 genes and 12 microsatellites) were typed on the 5000rad horse x hamster whole genome RH panel (Chowdhary et al. 2002). When

integrated with 42 loci from the two previous RH maps (Chowdhary et al. 2003; Raudsepp et al. 2002), the final map for the equine X chromosome comprised 169 markers (135 Type I and 34 Type II) uniformly distributed along the length of the chromosome (Fig. 6-1). With the estimated size of ECAX at ~153 Mbp (Chowdhary et al. 2003) the average marker density is 1 every 900 kbps and the average gene marker density is 1 every ~1 Mbps. The retention frequency (RF) of the 169 markers in the RH panel ranged from 5.4% (*BRIA3*, *BIRC4* and *ODZ1*) to 31.5% (*Adlican*), with an average of ~13% (Fig. 6-1). This retention frequency is satisfactory considering the RH panel was made from a male horse. A relatively high retention of markers was observed towards the proximal and distal end of the short arm, while retention was relatively lower than average in the distal part of the long arm (Fig. 6-1).

At logarithm of odds score 7 (2PT-RHMAP), the 169 markers clustered in six RH-linkage groups arranged tandemly from *pter* to *qter* on the chromosome (LGI–LGVI; Fig. 6-1, which is published as supporting information on the PNAS web site), each comprising 48, 11, 15, 30, 31, and 34 markers, respectively (Fig. 6-1). The proximal three linkage groups comprising 74 markers were located on the short arm, whereas the distal three linkage groups comprising 95 markers were located on the long arm. The order of markers within each RH linkage group was deduced by using the equal retention probability and stepwise locus ordering models (RHMAXLIK). Twenty-nine loci served as frameworks (odds of 1,000:1; shaded orange in Fig. 6-1). The comprehensive map thus generated spans of 1,344 centiRay $(cR)_{5000}$, uniformly covering the entire length of ECAX (see Fig. 6-1 legend for details on map construction). The only two human X (HSAX) regions (each spanning ~5–6 Mbp) not represented on ECAX

were the centromeric region and the HSAXp21.1 region (spanning between 54–62 and 31–37 Mb positions, respectively, on the human sequence map). Otherwise, the representation/alignment of HSAX on ECAX was quite uniform, averaging one gene-specific marker every ~1-Mbp interval.

Seven genes from the upper terminal part of the RH group 1 (a region corresponding to ECAXpter) produced PCR patterns characteristic to both the X and Y chromosomes. These genes gave the same-size PCR amplification products in males and females and also amplified in RH cell lines known to be Y specific (results not shown). Redesign of primers for most of these genes did not change the PCR typing pattern. However, the second set of primers generated by using sequence data from the first amplification products of *STS* produced an X specific pattern, as indicated with two adjacent map positions for this gene: *STSX* and *STSXY* (Fig. 6-1).

FISH map

Equine BAC clones containing 29 selected genes FISH mapped to the expected chromosomal location (Table A6-1) based on theRHmap and the previously FISH mapped loci (Raudsepp et al. 2002). These localizations bring the total number of cytogenetically mapped markers on ECAX to 53 (46 gene-specific and 7 short tandem repeats; Fig. 6-1). The FISH markers are fairly uniformly distributed along the chromosome, from Xp*ter* to Xq*ter* (Fig. 6-1 and Fig. 6-2), except in band Xp21. Two-color FISH on metaphase chromosomes provided the following physical order for overlapping loci:. *Adlican-TMSB4* (*pter*→*cen*; Fig. 6-2*h*), *CHMDIAPH2* (*cen*→*qter*; Fig. 6-2*f*) and *FMR1-MTM1* (*cen*→*qter*; Fig. 6-2*h*). Further, interphase FISH with combinations of differently labeled probes helped refine relative order of three loci

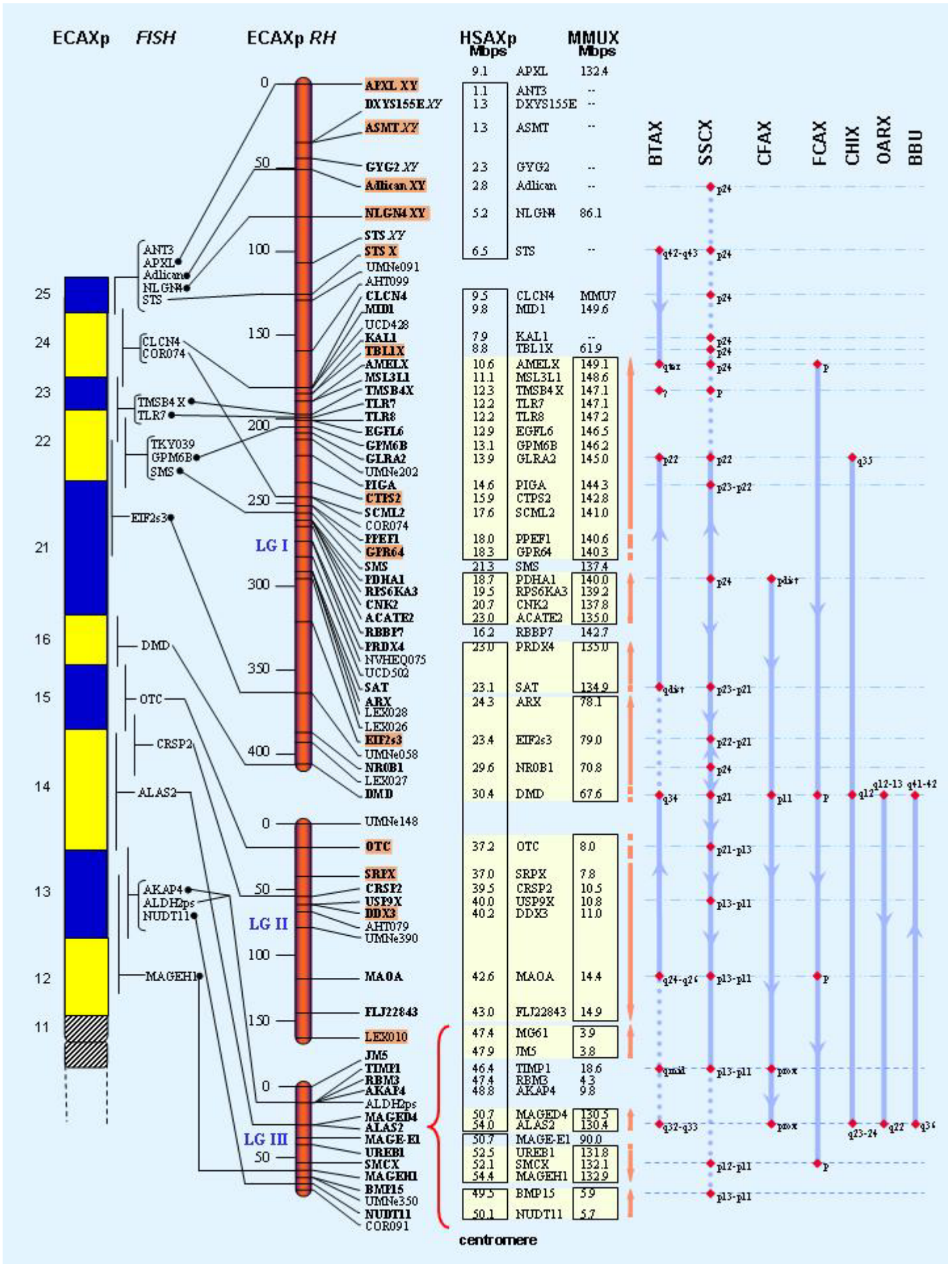


Figure 6-1. A high-resolution radiation hybrid (RH) and comparative map of the *Equus caballus* (ECA) X chromosome (ECAX). Map of the short arm (Xp) is presented on the first sheet and of the long arm (Xq) on the second sheet. To the extreme left of the RH map is the diagrammatic representation (ideogram) of the G banded X chromosome. Next to the ideogram are all of the fluorescence *in situ* hybridization (FISH) localizations (53 loci), of which 29 were conducted in this study (denoted by •). Most of the FISH loci are connected by diagonal lines to corresponding loci in the RH map. The six RH groups (RHMAP2PT; logarithm of odds, 7.0) are marked as LG I–LG VI and are depicted by vertical color bars calibrated on the left at 50-centiRay (cR) intervals. To the right of the bar are mapped equine loci; genes are shown in bold and microsatellites in normal font. Framework markers (odds, 1:1,000) are shaded. Next to the RH map are the sequence location in megabase pair for all human (HSA) and mouse (MMU) orthologs of the mapped horse genes (human, http://genome.ucsc.edu/cgi-bin/hgGateway; mouse, Genome gateway at same site). Orthologs showing conserved order compared to the derived order of equine genes are grouped in boxes (conserved linkages) demonstrating the degree of gene order conservation in human and mouse compared to that seen in the horse. The overall *pter⇒qter* order of the horse loci corresponds closely to that observed in humans; however, the reshuffle in mouse can be inferred from the arrows next to the conserved linkage blocks. The direction of the arrows shows the *pter⇒qter* arrangement of loci within each block. Yellow-shaded horizontal regions represent loci that have preserved gene order in horse, human, and mouse and signify core ancestral segments. To get an overview of the status of ECAX gene map in relation to other livestock/companion species, cattle, pig, dog, cat, and sheep/goat/buffalo orthologs of the mapped horse genes are presented next to the mouse comparative map. Solid vertical bars with arrows show known *pter⇒qter* arrangement of loci, whereas dotted lines show an uncertain order. Map information for cattle (BTA), pig (SSC), dog (CFA), cat (FCA), sheep (OAR), goat (CHI), and buffalo (BBU) genes was obtained from the respective databases.

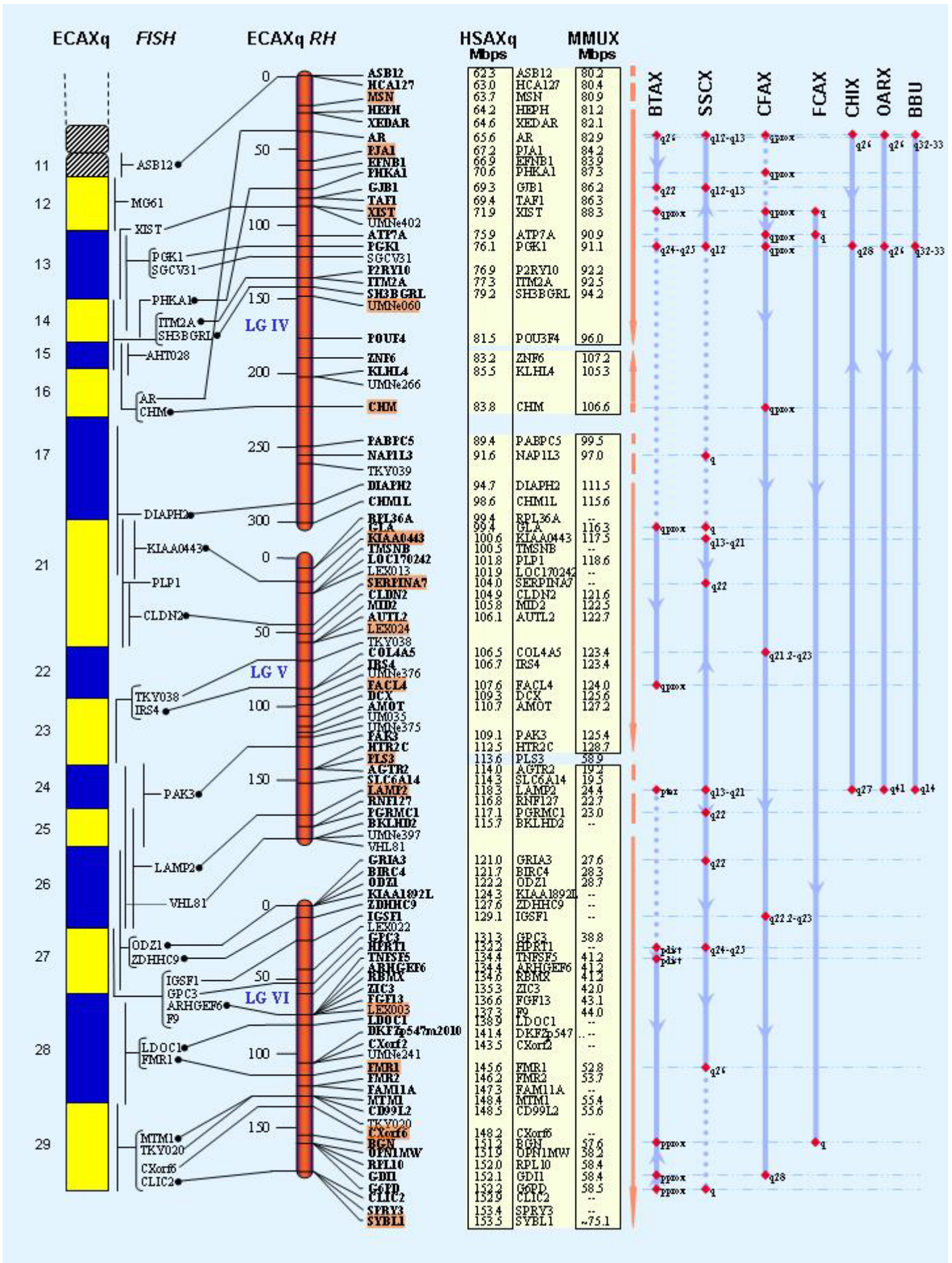


Figure 6-1. Continued.

(*Adlican, NLGN4*, and *TMSB4*) located in the terminal region of the short arm (Xp*ter*). The interphase-FISH order for the three loci was *pter*→*Adlican*→*NLGN4*→*TMSB4*→*cen* (Fig. 6-2*i*). All FISH results were in agreement with the physical order of loci deduced by theRHmap. However, there were minor exceptions: although double-color FISH showed that *NUDT11* is distal to *MAGEH1* on ECAXp13 (Fig. 6-2*e*), RH analysis gave a reversed order. Reexamination of typing results of these and adjacent loci did not show any genotyping_scoring errors. Last, our FISH assignment of *LAMP2* is in agreement with our RH map but does not concur with earlier localization (Milenkovic et al. 2002).

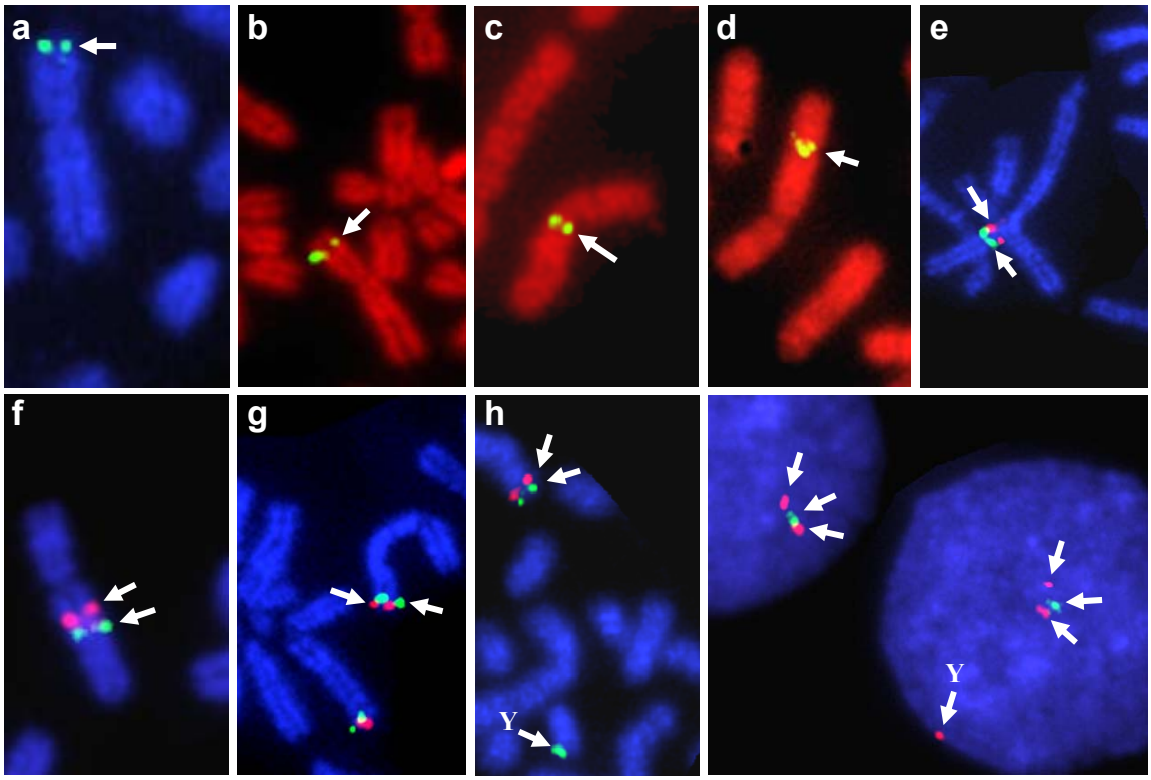


Figure 6-2. Single-color fluorescence *in situ* hybridization (FISH) show (arrows) location of (*a*) *NLGN4* on Xp25, (*b*) *TLR7* on Xp23-p22, (*c*) *PHKA1* on Xq14, and (*d*) *IRS4* on Xq23. Double-color FISH on metaphase chromosomes show relative order of (*e*) *MAGEH1* (red) on Xp13-p12 and *NUDT11* on Xp13, (*f*) *CHM* (red) on Xq15-q16 and *DIAPH2* (green) on Xq17-q21prox, (*g*) *FMR1* (green) on Xq28 and *MTM1* (red) on Xq29, and (*h*) *Adlican* (green) on Xp25/Yqter and *TMSB4* (red) on Xp23-p22. (*i*) Relative order of *Adlican* (red)-*NLGN4* (green)-*TMSB4* (red) in interphase nuclei.

Comparative mapping

The mapping of 137 equine orthologs of HSAX genes demonstrated complete synteny conservation between the X chromosomes of the two species (Fig. 6-1). Of these, 114 loci have also been mapped in mouse: 113 to MMUX and one (*CLCN4*) to MMU7 (49). The results reiterate conservation of the X chromosome across three evolutionarily distantly related mammals, with a remarkably higher degree of conservation between horse and human than between horse and mouse or human and mouse. As yet no mouse orthologs have been found for 23 of the horse/human X specific genes. Of these, nine genes are from the human pseudoautosomal regions (PAR), PAR1 and -2 (see Fig. 6-1).

To obtain a refined comparative map among horse, human, and mouse X chromosomes, we determined the precise sequence location of human and mouse orthologs for all 133 physically ordered equine genes from the available genomic draft sequences (see *Materials and Methods*). Four FISH-mapped genes (*ANT3*, *MG61*, *PLP1*, and *F9*) not mapped by RH analysis were also included in the comparison by placing them on the comparative map based on adjacent FISH markers, increasing the number of comparative loci to 137. As described earlier (Chowdhary et al. 2003), if sequence locations of a group of human or mouse orthologs indicated conservation of gene order in relation to the derived order of equine genes, the data were clustered in boxes (see Fig. 6-1). A total of seven clusters were observed in human and 13 in mouse. These clusters, referred to as conserved linkages [maximally contiguous chromosomal region with identical gene content and order (Nadeau and Sankoff 1998)], showed groups of genes with similar physical order in horse–human and horse–mouse. The clustering also

highlighted smaller evolutionarily conserved segments among the three species that could be observed within the larger conserved linkages (Fig. 6-1, cream-shaded regions over human/mouse orthologs).

Discussion

We herein report a physical map of the horse X chromosome comprising 175 markers (139 genes and 36 microsatellites). The 128 markers mapped in this study by RH analysis, 90% of which are functional genes, expand the existing ECAX map (Raudsepp et al. 2002; Chowdhary et al. 2003) by >4-fold. Further, a total of 53 FISH localizations substantially improve the number of physical anchor points between the chromosome and the RH map and concurrently provide excellent corroboration to the RH data. The map is the first high resolution integrated gene map for a horse chromosome. With markers distributed on average at 880-kb intervals along the chromosome and gene-specific loci dispersed on average at ~1-Mb intervals, this is the densest and most uniformly distributed map of the X chromosome presently available in any of the mammalian species other than humans and mice. The currently available X chromosome maps of various livestock and pet/companion animals show a range of 12 (sheep/goat/buffalo) to 92 (pig) genes mapped using different approaches [cattle, ~37 genes (Amaral et al. 2002; Band et al. 2000; Iannuzzi et al. 2000; Piumi et al. 1998); pig, ~92 genes (Cirera et al. 2003; Davoli et al. 2002; Hawken et al. 1999; Hu et al. 1997; Lahbib-Mansais et al. 2003; McCoard et al. 2002; Quilter et al. 2002; Rink et al. 2002); dog, 22 genes (Breen et al. 2001; Everts et al. 2002; Spriggs et al. 2003); cat, 30 genes (Menotti-Raymond et al. 2003; Murphy et al. 1999a); sheep/goat/buffalo, 12 genes (Iannuzzi et al. 2000; Piumi et

al. 1998; Schibler et al. 1998)]. In most of these species except cat (Menotti-Raymond et al. 2003) and dog (Guyon et al. 2003), either the different maps are not integrated and physically ordered into a single map or the distribution of the loci is not uniform. In contrast, the 137 gene human–horse comparative map presented in this study is 2- to 4-fold more informative for gene content and is conspicuously integrated into a single physically ordered map. This is evident also from the comparative status of the X chromosome map for all ECAX genes hitherto mapped in different species (see Fig. 6-1). Last, despite containing only ~10% of the expected X homologs from human (1,695 genes http://genome.ucsc.edu/cgi-bin/hgGateway) or mouse (1,116 genes; draft sequence, www.ensembl.org/Mus_musculus), the map provides excellent alignment of ECAX with the X chromosome map of the two species.

Despite a total of 169 markers, the ECAX RH map was divided into six RH linkage groups, due to the threshold chosen (logarithm of odds, 7 RHMAP2PT) to develop the linkage groups. Assuming the linkage groups to be separated by at least 50 cR5000, the five gaps add ~250 cR to the current map of 1,344 cR, thus increasing the effective map size to ~1,600 cR. This implies that with the size of ECAX as 153 Mbp, the average span of the map for this chromosome in our 5,000-rad panel is ~10 cR_Mbp. Further, of the 153 Mbp, ~2–3 Mbp spans the centromeric region, whereas 5–6 Mbp spans the intercalary heterochromatic region on the long arm of the chromosome (Xq21). It is presumed that these regions are gene poor. These assumptions stem from the fact that gene density in the predominantly heterochromatic regions of HSA1, -9, and -16 is zero for stretches spanning 20, 16, and 9 Mbp of heterochromatin, respectively (http://bioinformatics.weizmann.ac.il/cards). A similar condition was also seen for

satellite-bearing chromosome HSA22, where the proximal ~14 Mbp are barren, and HSAY, where the distal ~20 Mbp do not contain any genes (except for three genes in the PAR2 region).

XY amplification and the pseudoautosomal region

Of significant interest in the ECAX map is the distal_terminal region of the short arm. Seven markers from this region (*APXL*, *DXYS155E*, *ASMT*, *GYG2*, *Adlican*, *NLGN4*, and *STS*) demonstrate both X and Y specific amplification with RH cell line DNA. In humans, all seven genes are known to have both X and Y homologs with sequence identities ranging from 60% to 96% (Skaletsky et al. 2003; Gene Cards, http://bioinformatics.weizmann.ac.il/cards). Presently 31 such genes are located on HSAY. These genes, referred to as X degenerates, are considered as surviving relics of the ancestral autosome from which X and Y chromosomes evolved (Skaletsky et al. 2003). Four more X degenerate genes (*TMSB4X*, *USP9X*, *KAL1*, and *SMCX*; Fig. 6-1) were also mapped in this study, but primers for these genes show only X specific amplification. We suggest that, because the latter four loci are further away from the XY recombining region, sequences of these genes might have had more evolutionary time for X-Y divergence (Lahn and Page 1999); however, this needs further investigation. Nevertheless, the identification of equine X-Y pattern genes in horse is an important starting point to analyze shared sequences between equine X and Y chromosomes, thus shedding light on their evolution.

Synaptonemal complex analysis in horses shows that the ECAXp*ter* region is pseudoautosomal [PAR (Power et al. 1992; Safronova and Pimenova 1988)]. The three main characteristics of the PAR genes are that they (*i*) represent functional

homologs of genes present on both X and Y chromosomes; (*ii*) escape X inactivation; and (*iii*) recombine during male meiosis (Blaschke and Rappold 1997; Charchar et al. 2003; Ellison et al. 1996; Gianfrancesco et al. 2001). The human PAR1 spans ~2.6 Mb and contains 13 genes (Charchar et al. 2003; Ellison et al. 1996; Gianfrancesco et al. 2001; Rappold 1993). Of these, two, *DXYS155E* and *ASMT*, are mapped to ECAX in this study. The human PAR2, located terminally on HSAXq, spans 320 kb and contains four genes [*HSPRY3*, *SYBL1*, *IL9R*, and *CXYorf1* (Charchar et al. 2003)]. Of these, two have been mapped in this study to the terminal end of ECAXq (Fig. 6-1). Analysis hitherto carried out in a range of mammalian species show significant species-specific differences in the PAR genes. A gene that is pseudoautosomal in one species may not be pseudoautosomal in the other. Of the 13 human PAR genes, eight do not have a known mouse homolog (Gianfrancesco et al. 2001). Further, *STS* is a PAR gene in cattle (Moore et al. 2001), pig (Quilter et al. 2002), dog, sheep and mouse, but not in human and primates (Blaschke and Rappold 1997; Gianfrancesco et al. 2001); *KAL1* is pseudoautosomal in pigs (Quilter et al. 2002) and cattle (Moore et al. 2001) but not in human and primate. Surprisingly, this locus is not even found in mouse (Gianfrancesco et al. 2001). The horse PAR has not yet been defined for gene content. The genes mapped in this study provide a foundation to begin investigating the Xp*ter* region in horse for identification of prospective PAR genes. Subsequent expression/methylation studies of these genes can help validate their pseudoautosomal status and contribute to precise characterization and physical demarcation the equine PAR.

Comparative map

The physically ordered ECAX map, comprising 139 equine genes, is in close agreement with our previous maps (Raudsepp et al. 2002; Chowdhary et al. 2003) and provides a robust comparative map in relation to the observed order of these genes in humans and mice. The most striking feature of the horse–human comparison is that, barring minor exceptions involving four to five interruptions on the short arm, the relative order of loci in the two species is exceptionally conserved from Xp*ter* to Xq*ter*. This degree of gene order conservation has not yet been observed for any other chromosome between humans and other nonprimate mammals. The horse–mouse comparison, however, shows noticeably less conservation in gene order. The rearrangements are evident from the 13 conserved linkage blocks (boxes containing ordered loci) originating from different regions of MMUX, some with reversed centromere–telomere orientation in relation to horse and human (Fig. 6-1).

The comparative map of the equine X chromosome presented here helps to identify 13 blocks/clusters of loci demonstrating conserved gene order across horse, human, and mouse (Fig. 6-1, yellow shaded regions). These clusters potentially represent the most conserved X chromosome regions of the ancestor common to horse, human, and mouse. Additionally, the blocks provide a quick comparative overview of smaller conserved linkages and their relative organization in the three species. Although an exceptional degree of gene order conservation is observed between ECAX and HSAX, there are also minor rearrangements. The most noticeable of these is a small segment corresponding to ECAXp12–13 and the human sequence positions 46–54 Mb (Fig. 6-1, red bracket). Interestingly, the horse and mouse genomes are also relatively more

rearranged in the same region. Based on studies focused on the evolution of mammalian X chromosome, it is proposed that the human X chromosome is composed of different evolutionary strata. These are roughly divided into the X conserved region XCR, corresponding to the proximal one-quarter of short arm and complete long arm of HSAX, and the recently added region XRA, corresponding to distal three-quarters of the short arm of HSAX (Wilcox et al. 1996; Lahn and Page 1999). Rearrangements observed by us in the horse, human, and mouse genomes are incidentally located close to the suggested ancestral fusion point of the conserved and the ancestral regions on HSAXp11.23 (Wilcox et al. 1996). A detailed study of this region aimed at identifying evolutionary breakpoints may also lead to discovery of signatures of these rearrangements in horse/mouse, which in turn will be useful for understanding the origin of gene order differences in this region across the three species.

Future uses

The mammalian X chromosome contains a disproportionately high number of genes influencing development, female/male fertility, reproduction and diseases (OMIM, http://www.ncbi.nlm.nih.gov/Sitemap/index.html#OMIM; OMIA, http://morgan.angis.su.oz.au/Databases/BIRX/omia/; Graves and Delbridge 2001; Liao et al. 2003; Vaiman 2002; Wang et al. 2001) that are also of significance in horses. The human and pig X chromosomes carry an unexpectedly high number of genes specifically expressed in the skeletal muscle (Bortoluzzi et al. 1998; Davoli et al. 2002; Pallavicini et al. 1997). Analysis of these genes could have implications for performance of horses. Structural and numerical aberrations of the X chromosome are the most common documented chromosome abnormalities in the horse that invariably lead to reproductive failures,

intersexuality, hermaphroditism or sex reversal (Mäkinen et al. 1999, 2001). Additionally, a number of X-linked conditions/diseases have been described in the horse, e.g., *G6PD* deficiency (Stockham et al. 1994), X-linked severe immunodeficiency (Felsburg et al. 1992), fragile X (Ronne 1992), agammaglobulinemia (Perryman et al. 1983), hydrocephalus (Ojala and Ala-Huikku 1992). Presently very little information is available concerning the underlying molecular causes of these conditions. The high resolution RH and comparative map of ECAX presented in this study will, in conjunction with detailed map/sequence information on human and mouse X chromosomes (Boyd et al. 2000; Liao et al. 2003; Vaiman 2002; Vialard et al. 2002), serve as a basis to rapidly converge on X-specific genes significant in the horse. Our current efforts to isolate BAC clones for all mapped markers on this chromosome will considerably facilitate localized isolation of polymorphic markers, building BAC contigs in areas of interest and sequencing targeted chromosomal regions.

VII. A 1.3-Mb INTERVAL MAP OF THE EQUINE HOMOLOGS OF HSA2*

Synopsis

A comparative approach that utilizes information from more densely mapped or sequenced genomes is a proven and efficient means to increase our knowledge of the structure of the horse genome. Human chromosome 2 (HSA2), the second largest human chromosome, comprising 243 Mb, and containing 1246 known genes, corresponds to all or parts of three equine chromosomes. This report describes the assignment of 140 new markers (78 genes and 62 microsatellites) to the equine radiation hybrid (RH) map, and the anchoring of 24 of these markers to horse chromosomes by FISH. The updated equine RH maps for ECA6p, ECA15, and ECA18 resulting from this work have one, two, and three RH linkage groups, respectively, per chromosome/chromosome-arm. These maps have a three-fold increase in the number of mapped markers compared to previous maps, and an increase in the average marker density to one marker per 1.3 Mb. Comparative maps of ECA6p, ECA15, and ECA18 with human, chimpanzee, dog, mouse, rat, and chicken genomes reveal blocks of conserved synteny across mammals and vertebrates.

---

*This section reproduced with kind permission of Springer Science and Business Media from Cytogenetics and Genome Research, A 1.3-Mb interval map of equine homologs of HSA2, Michelle L. Wagner, Terje Raudsepp, Glenda Goh, Richa Agarwala, Alejandro A. Schäffer, Patricia K. Dranchak, Candice Brinkmeyer-Langford, Loren C. Skow, Bhanu P. Chowdhary, and James R. Mickelson, volume 112, 227-234, © 2006 by S. Karger AG, Basel.

Introduction

A major goal of the horse (*Equus caballus*) genome mapping effort is to identify genes impacting equine well-being. Genes responsible for deleterious traits have been and continue to be propagated coincidentally with the line-breeding of desirable traits in various breeds of horses. Thus far, the bases for fatal Mendelian disorders including Overo Lethal White Foal Syndrome in Paint Horses (Yang et al. 1998; Santschi et al. 2001), Severe Combined Immunodeficiency in Arabians (Wiler et al. 1995), and Glycogen Branching Enzyme Deficiency in Quarter Horses (Ward et al. 2004), as well as the rarely fatal but debilitating Hyperkalemic Periodic Paralysis in Quarter Horses (Rudolph et al. 1992), have been defined. Although a candidate gene approach was utilized to discover the genes mutated in these disorders, this method is applicable only when similar disorders caused by homologous genes are present in other species. Since many inherited equine disorders are likely to be unique to *E. caballus*, as well as non-Mendelian, a positional cloning approach will be required to map genes contributing to many equine diseases. A comprehensive equine genome map will be useful in identifying positional candidate genes responsible for traits of interest when the gene has been mapped to a chromosomal segment by genetic linkage or cytogenetic analysis.

The first generation whole genome equine radiation hybrid (RH) map includes 472 microsatellite and 258 gene/EST markers (Chowdhary et al. 2003). This map has an average density of one marker every 4 Mb, but many large gaps between markers exist. Subsequently, the marker density of ECA17 was increased to an average interval of every 1.4 Mb by the mapping of 75 markers, 56 of which were genes, making this the

first equine autosome to be densely mapped (Lee et al. 2003a). The power of RH mapping to make a high-resolution map of selected equine chromosomes has been further demonstrated with the recent publication of a 175 marker (139 gene) map of the equine X chromosome, an 83 marker (52 gene) map of ECA22, and the preparation of maps of four equine chromosome segments corresponding to HSA19 (Raudsepp et al. 2004a; Gustafson-Seabury et al. 2005; Brinkmeyer-Langford et al. 2005). The ECAX and ECA22 maps now contain markers at 880 kb and 770 kb mean intervals, respectively, making them two of the densest farm animal chromosomal maps available at the present time (Raudsepp et al. 2004a; Gustafson-Seabury et al. 2005). Additionally, a recent report of the physical mapping of 77 markers on ECAY has made this one of the most densely mapped animal Y chromosomes (Raudsepp et al. 2004b).

In this study, likely equine orthologs of HSA2 genes and known microsatellites were typed on a horse radiation hybrid panel, accompanied by additional FISH mapping, to enable direct comparison of the maps of HSA2 with its equine counterparts ECA6p, ECA15, and ECA18 (Milenkovic et al. 2002; Chowdhary et al. 2003; Yang et al. 2004). These maps reveal rearrangements between human and horse chromosomes, and regions of conserved synteny between mammals, which have not yet been detected due to the low resolution of previous maps. In so doing, we bring the horse genome mapping project closer to its long-term goal of dense maps of all chromosomes.

Materials and methods

Primer design and PCR conditions

In one approach, primers were designed to amplify conserved adjacent exons of genes at regularly spaced (1-10 Mb) intervals along HSA2. Previous results suggested that 84% of such primers would work in different mammalian species, including the horse (Venta et al. 1996; Shubitowski et al. 2001). PCR products from horse genomic DNA were separated on 1% agarose gels, extracted, purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA), and sequenced. Sequences that had BLAST results (Altschul et al. 1997) supporting the correct equine orthologous gene were then used to design horse specific RH mapping primers within introns using Prime (GCG, Wisconsin Package Version 10.2, Madison, WI).

In related approaches, gene specific markers were generated from HSA2 genes selected from NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?taxid = 9606&chr = 2), or the human reference sequence (http://genome.ucsc.edu, version May 2004). HSA2 gene cDNA sequences were BLASTed against the publicly available horse EST and cDNA sequences, and strong equine matches were then BLASTed against the non-redundant (nr) database to find non-conserved regions with rodents (typically 3' untranslated regions) for horse-specific primer design with Prime. HSA2 gene exon sequences were also compared to sequences for mouse, rat, cattle, and pig using BLASTn (http://www.ncbi.nlm.nih.gov/BLAST/) and conserved regions among mammals were identified by multiple alignments in CLUSTALW (Thompson et al. 1994; http://bioweb.Pasteur.fr/seqanal/interfaces/clustalw-simple.html). PCR primers

were then designed from regions, in single exons or adjacent exons with 500–700 bp introns, with 100% identity in coding sequence for humans, cattle, and pigs, but with 1–3 mismatches in rodents. Primer3 software was used for these primer designs (http://wwwgenome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Lastly, primer pairs for 62 microsatellite markers were obtained from multiple published sources (Mickelson et al. 2003; Tozaki et al. 2003; Wagner et al. 2004a, b, c).

All primers were typed on the Texas A & M University 5000 rad horse x hamster RH panel comprising 92 cloned hybrid cell lines (Chowdhary et al. 2002, 2003). PCR primers were optimized to specifically amplify horse DNA, with hamster DNA and no DNA controls included in each experiment. Basic PCR conditions were per 15 μl reaction: 50 ng each hybrid cell DNA, 1.5 mM $MgCl_2$, 1 nmol dNTPs, 0.25 U Jump Start Sigma Red Taq Polymerase (Sigma, St. Louis, MO) or 0.25 U HotStar Taq Polymerase (Qiagen, Valencia, CA). All markers were typed in duplicate, and products were resolved on 2% agarose gels, and scored manually using Gel Score Software (Wes Barris). Ambiguities were resolved with PCR of individual hybrid cell DNA in duplicate, along with horse DNA and hamster DNA controls.

BAC library screening and FISH mapping

Primer pairs for Type I and Type II markers spaced along ECA15 and ECA18 were used to screen the CHORI-241 (http://bacpac.chori.org/equine241.htm) equine BAC library. Thirty-seven BAC clones were isolated, with the goal of obtaining two clones per marker to confirm FISH results. Approximately 1 microgram DNA from BAC clones was individually labeled with biotin and/or digoxigenin (BIO and DIG Nick

Translation Mix, Roche Biochemicals). Labeled probes were hybridized separately or in pairs to horse metaphase chromosome spreads to visualize their physical chromosomal locations. DNA labeling, in situ hybridization, signal detection, microscopy, and image analysis were performed as previously described (Chowdhary et al. 2003).

Computation of radiation hybrid maps

Radiation hybrid maps were computed using the software packages rh_tsp_map (Agarwala et al. 2000; ftp://ftp.ncbi.nih.gov/pub/agarwala/rhmapping/rh_tsp_map.tar), CONCORDE (Applegate et al. 1998;http://www.isye.gatech.edu/~wcook/rh/), and Qsopt (http://www.isye.gatech.edu/~wcook/qsopt), as well as some auxiliary programs. The methods are similar to those used in (Menotti-Raymond et al. 2003) and identical to those used in (Brinkmeyer-Langford et al. 2005).

Radiation hybrid maps were computed with the goal of optimizing the maximum likelihood (MLE) criterion. When making linkage groups by single linkage clustering based on pairwise LOD scores, we used thresholds of 7.0, 6.0, and 7.5 for ECA6p, ECA18, ECA15, respectively. There were 1, 3, and 2 linkage groups respectively. Three singleton markers for ECA15 were dropped, but there were no singleton markers for ECA6p and ECA18. We verified that there was a space of at least 1 LOD unit between the smallest intra-group score and the largest inter-chromosomal score.

We used the program frame_markers to eliminate from initial consideration markers that have too many 2 entries (indicating uncertainty in the genotyping) and one out of each pair of markers that are too close to each other, preferentially retaining an intergenic marker over a closely linked microsatellite. We tested this set of markers to

see if it could be robustly ordered and dropped markers until that was the case. By "robustly ordered" we mean that: 1) the optimal ordering is the same for all three mathematical formulations of the MLE criterion (defined in Agarwala et al. 2000); 2), for each marker its best placement relative to all other markers is at least 0.25 LOD units better than the second best placement; and 3), the optimal order is at least 0.25 LOD units better than any other order obtained by permuting up to 8 consecutive markers. A map of markers that satisfies these three requirements is called an "MLE-consensus" map because of the first requirement.

The remaining steps included: 1) placing markers not on the MLE-consensus maps relative to those maps; 2), ordering and orienting linkage groups; 3), concatenating linkage groups on the same chromosome; and, 4) producing a final map by solving restricted instances of the traveling salesman problem (TSP). These four steps were done exactly as in Brinkmeyer-Langford et al. 2005, except that step 2 was unnecessary for chromosome 6p because it had only 1 linkage group. The order and orientation of linkage groups on a chromosome was determined by considering FISH data and consulting a genetic linkage map (Penedo et al. 2005).

Results

Development of gene-specific markers

Seventy-four of the 87 primer pairs from highly conserved coding regions in HSA2 genes gave successful equine amplifications. Sequencing of the products showed that 93% (69) gave BLASTn matches to the intended human gene (or orthologs in other

species) with E-values < 1e-10. Horse specific PCR primers were designed from these sequences. Thirty-six primer pairs were designed from the 3' untranslated regions of putative orthologs of human genes obtained from equine EST sequences, or from BAC end sequences. Lastly, 14 gene-specific markers were generated by multiple alignments of genomic DNA from non-equine species in single exons or two adjacent exons where there was 100% identity between human, cow, and pig but one to three mismatches with the rodent sequences.

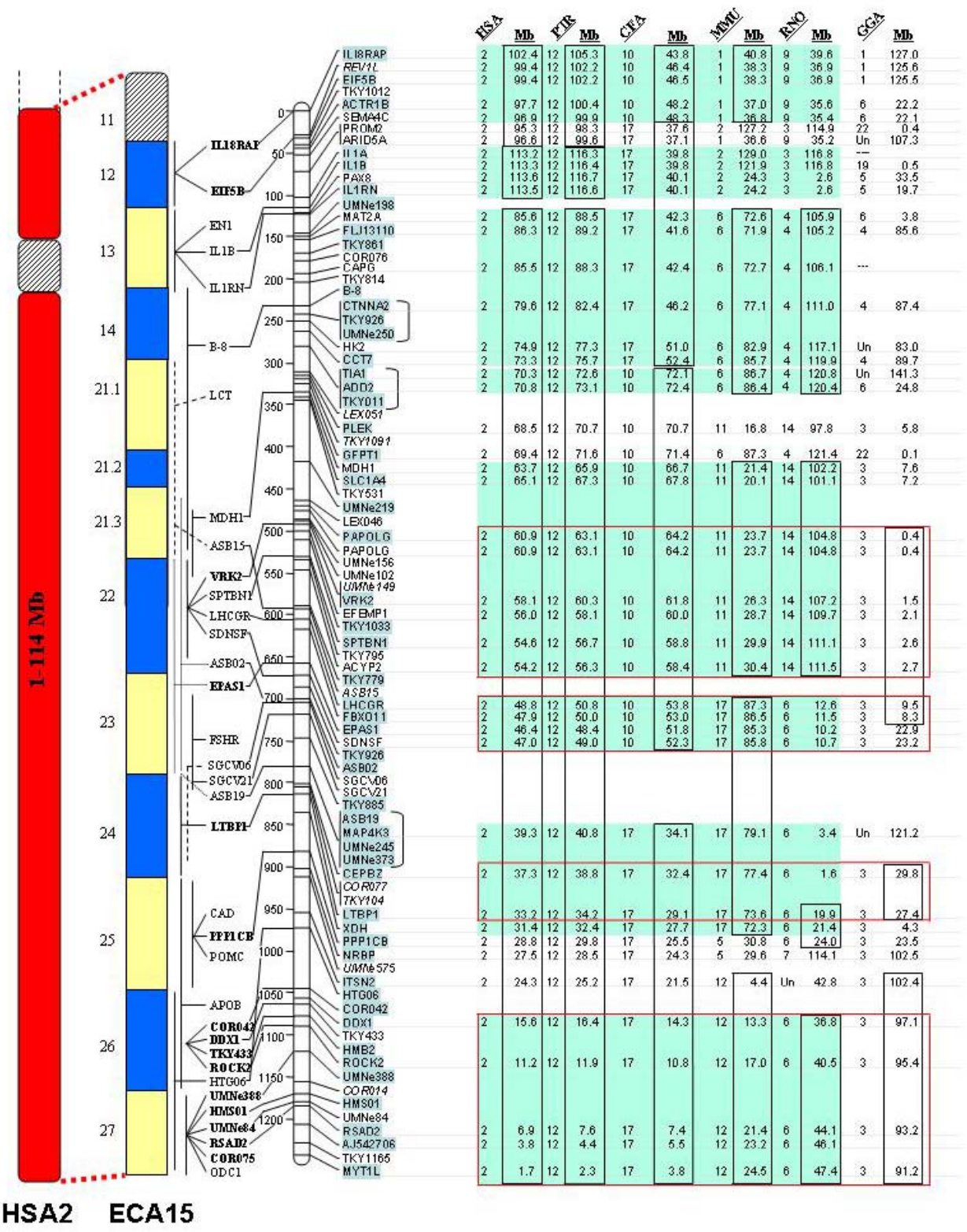
RH typing

Sixty-one of the 69 primer pairs derived from the cross-intron approach produced readily scorable horse-specific products in a hamster DNA background. Of these, 14 were eliminated in the statistical analysis of vector data, leaving 47 markers. Of the 36 primer pairs designed for horse EST and BAC sequences, 17 markers were added to the maps. Further, all 14 gene-specific markers generated by multiple alignments of genomic DNA from non-equine origins could be genotyped on the RH panel. Thus, the total number of newly mapped Type I markers on the RH maps of ECA6p, 15 and 18 was 78 (Table 7-1). Additionally, 62 Type II markers were also added to the RH maps of these equine chromosomes. Average retention frequency for new HSA2-based gene markers on the panel was 18%, with an average of 21% for ECA6p, 16% for ECA18, and 19 % for ECA15.

ECA6p map

The gene markers mapped to ECA6p are derived from the 210-235 Mb region of HSA2q (Figure 7-1). The updated RH map of this chromosome comprises 29 markers

Figure 7-1. Radiation hybrid (RH) and comparative maps of (a) ECA18 and ECA6p: corresponding to a majority of HSA2q and (b) of ECA15: corresponding to HSA2p and the proximal region of HSA2q. G-banded ideograms of individual chromosomes show FISH mapped markers, with new additions to the cytogenetic map in bold. To the left of the ideograms, regions of correspondence of individual equine chromosome with HSA2 are shown by approximately demarcating Megabase (Mb) positions on the human chromosome. To the right side of the ideogram, the RH map is presented. The MLE-consensus markers are shaded in light blue. Markers with a LOD score >3.0 are depicted in bold. Markers in normal font were placed on this framework, while those that were placed with a LOD of <0.5, and thus less reliable, are in italics. Markers with the same physical/cR position have a vertical bar to the left of them, while those with identical RH vector coordinates are shown with a bracket to the right. To the right of the RH map are Mb positions for orthologs in human (HSA), chimp (PTR), dog (CFA), mouse (MMU), rat (RNO), and chicken (GGA) as found in the Human Genome Browser (http:// genome.ucsc.edu/cgi-bin/hgGateway). Vertical black rectangles on Mb positions depict conservation of gene order within a species. Green shaded rectangles show conserved gene order among sequenced mammals and horizontal red rectangles show conserved gene order among vertebrates relative to the horse. Dotted lines show continuation/break in chromosomal region.



	HSA	Mb	PTR	Mb	CFA	Mb	MMU	Mb	RNO	Mb	GGA	Mb
IL18RAP	2	102.4	12	105.3	10	43.8	1	40.8	9	39.6	1	127.0
REV1L	2	99.4	12	102.2	10	46.4	1	38.3	9	36.9	1	125.6
EIF5B	2	99.4	12	102.2	10	46.5	1	38.3	9	36.9	1	125.5
TKY1012												
ACTR1B	2	97.7	12	100.4	10	48.2	1	37.0	9	35.6	6	22.2
SBMA4C	2	96.9	12	99.9	10	48.3	1	36.8	9	35.4	6	22.1
PROM2	2	95.3	12	98.3	17	37.6	2	127.2	3	114.9	22	0.4
ARID5A	2	96.6	12	99.6	17	37.1	1	36.6	9	35.2	Un	107.3
IL1A	2	113.2	12	116.3	17	39.8	2	129.0	3	116.8	---	
IL1B	2	113.3	12	116.4	17	39.8	2	121.9	3	116.8	19	0.5
PAX8	2	113.6	12	116.7	17	40.1	2	24.3	3	2.6	5	33.5
IL1RN	2	113.5	12	116.6	17	40.1	2	24.2	3	2.6	5	19.7
UMNe198												
MAT2A	2	85.6	12	88.5	17	42.3	6	72.6	4	105.9	6	3.8
FLJ13110	2	86.3	12	89.2	17	41.6	6	71.9	4	105.2	4	85.6
TKY861												
COR076												
CAPG	2	85.5	12	88.3	17	42.4	6	72.7	4	106.1	---	
TKY814												
B-8												
CTNNA2	2	79.6	12	82.4	17	46.2	6	77.1	4	111.0	4	87.4
TKY926												
UMNe250												
HK2	2	74.9	12	77.3	17	51.0	6	82.9	4	117.1	Un	83.0
CCT7	2	73.3	12	75.7	17	52.4	6	85.7	4	119.9	4	89.7
TIA1	2	70.3	12	72.6	10	72.1	6	86.7	4	120.8	Un	141.3
ADD2	2	70.8	12	73.1	10	72.4	6	86.4	4	120.4	6	24.8
TKY011												
LEX051												
PLEK	2	68.5	12	70.7	10	70.7	11	16.8	14	97.8	3	5.8
TKY1091												
GFPT1	2	69.4	12	71.6	10	71.4	6	87.3	4	121.4	22	0.1
MDH1	2	63.7	12	65.9	10	66.7	11	21.4	14	102.2	3	7.6
SLC1A4	2	65.1	12	67.3	10	67.8	11	20.1	14	101.1	3	7.2
TKY531												
UMNe219												
LEX046												
PAPOLG	2	60.9	12	63.1	10	64.2	11	23.7	14	104.8	3	0.4
PAPOLG	2	60.9	12	63.1	10	64.2	11	23.7	14	104.8	3	0.4
UMNe156												
UMNe102												
UMNe149												
VRK2	2	58.1	12	60.3	10	61.8	11	26.3	14	107.2	3	1.5
EFBMP1	2	56.0	12	58.1	10	60.0	11	28.7	14	109.7	3	2.1
TKY1033												
SPTBN1	2	54.6	12	56.7	10	58.8	11	29.9	14	111.1	3	2.6
TKY795												
ACYP2	2	54.2	12	56.3	10	58.4	11	30.4	14	111.5	3	2.7
TKY779												
ASB15												
LHCGR	2	48.8	12	50.8	10	53.8	17	87.3	6	12.6	3	9.5
FBXO11	2	47.9	12	50.0	10	53.0	17	86.5	6	11.5	3	8.3
EPAS1	2	46.4	12	48.4	10	51.8	17	85.3	6	10.2	3	22.9
SDNSF	2	47.0	12	49.0	10	52.3	17	85.8	6	10.7	3	23.2
TKY926												
ASB02												
SGCV06												
SGCV21												
TKY885												
ASB19												
MAP4K3	2	39.3	12	40.8	17	34.1	17	79.1	6	3.4	Un	121.2
UMNe245												
UMNe373												
CEPBZ	2	37.3	12	38.8	17	32.4	17	77.4	6	1.6	3	29.8
COR077												
TKY104												
LTBP1	2	33.2	12	34.2	17	29.1	17	73.6	6	19.9	3	27.4
XDH	2	31.4	12	32.4	17	27.7	17	72.3	6	21.4	3	4.3
PPP1CB	2	28.8	12	29.8	17	25.5	5	30.8	6	24.0	3	23.5
NRBP	2	27.5	12	28.5	17	24.3	5	29.6	7	114.1	3	102.5
UMNe575												
ITSN2	2	24.3	12	25.2	17	21.5	12	4.4	Un	42.8	3	102.4
HTG06												
COR042												
DDX1	2	15.6	12	16.4	17	14.3	12	13.3	6	36.8	3	97.1
TKY433												
HMB2												
ROCK2	2	11.2	12	11.9	17	10.8	12	17.0	6	40.5	3	95.4
UMNe388												
COR014												
HMS01												
UMNe84												
RSAD2	2	6.9	12	7.6	17	7.4	12	21.4	6	44.1	3	93.2
AJ542706	2	3.8	12	4.4	17	5.5	12	23.2	6	46.1		
TKY1165												
MYT1L	2	1.7	12	2.3	17	3.8	12	24.5	6	47.4	3	91.2

Figure 7-1. Continued.

(19 new: 11 Type I and 8 Type II) in a single linkage group spanning 355.55 cR in length, and providing a marker every 12 cR. 55% of the markers are found in the MLE-consensus map.

ECA18 map

Gene markers mapped to ECA18 came from the 118-210 Mb region of HSA2q (Figure 7-1). The new ECA18 RH map covers 1265.70 cR, and has 71 markers (33 Type I and 38 Type II) compared to the 24 markers reported in the previous map (Chowdhary et al. 2003). The new ECA18 map has a marker every 18 cR, with 51% of markers making the MLE-consensus map.

ECA15 map

Genes mapped on ECA15 originate from the 1-114 Mb region of HSA2, corresponding to the p-arm and part of the q-cent region (Figure 7-1). The map comprises a total of 86 markers (45 TypeI and 41 Type II) spread over 1242.86 cR distance and had 68 new loci (42 Type I and 26 Type II). The new ECA15 RH map now has a marker every 14 cR, with 60% of the markers present in the MLE-consensus map.

FISH mapping

Thirty-five BAC clones containing 24 markers were isolated from the CHORI-241 BAC library and used for FISH mapping. Mostly gene-containing BACs were used, with the exception of five markers on ECA15 being microsatellites. A single clone for each of the marker resulted in 9 and 15 new FISH localizations that anchored the RH maps to ECA18 and ECA15, respectively. Representative figures are shown for 6 markers each from ECA15 and ECA 18 (Figure 7-2).

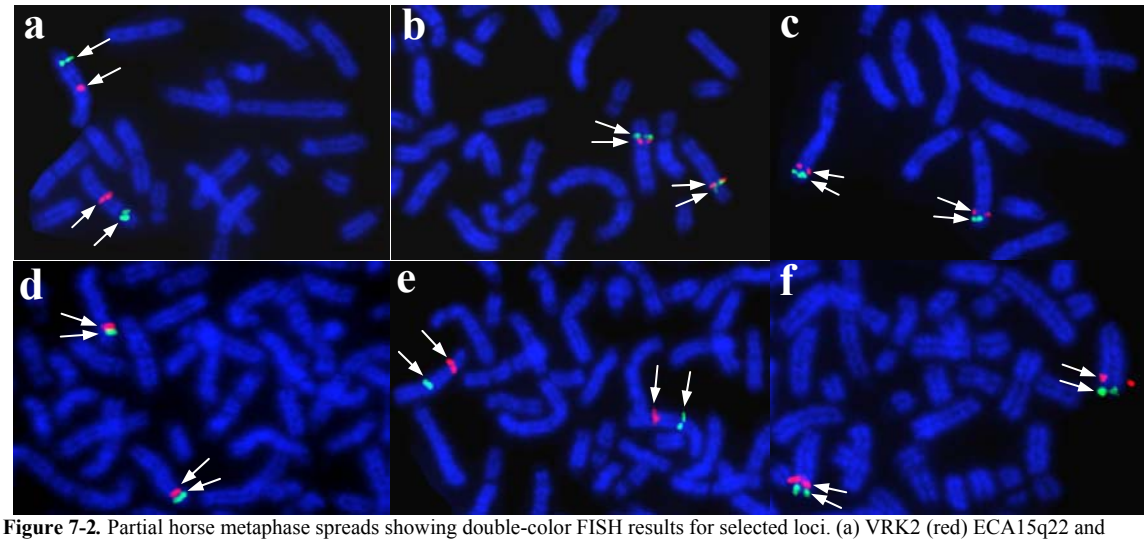


Figure 7-2. Partial horse metaphase spreads showing double-color FISH results for selected loci. (a) VRK2 (red) ECA15q22 and IL18RAP (green) ECA15q12, (b) LTBP1 (red) 15q24 and PPP1CB (green) 15q25, (c) DDX1 (red) 15q26 and RSAD2 (CIG5) (green) 15q27, (d) DDX18 (red) 18q15 and BIN1 (green) 18q12-q13, (e) DDX18 (red) 18q15 and GRB14 (green) 18q23, (f) AJ577071 (red) 18q25 and NDUFS1 (green) 18q26.

Comparative map

Forty-five, 33, and 16 gene markers from the RH maps of ECA15, ECA18, and ECA6p, respectively, were used to identify orthologs in human, chimpanzee, dog, mouse, rat, and chicken, as described previously (Chowdhary et al. 2003; Gustafson-Seabury et al. 2005; Brinkmeyer-Langford et al. 2005). Chromosomal locations and megabase positions of the orthologs in these species were determined using the Human Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway). Conserved linkages (Nadeau and Sankoff 1998) between these species and horse are shown by vertical rectangles around contiguous megabase positions for each species (Figure 7-1).

As a first approximation, our data show that ECA6p is represented by a single nearly conserved linkage of HSA2 gene markers from 210 – 235 Mb; ECA18 is represented by six such conserved linkages, one from 118 – 128 Mb and five between

135 and 210 Mb; and ECA15 is represented by three similar group of markers, one considerably large segment from 1 – 86 Mb, and two smaller segments from 96- 102 Mb and 113Mb regions. Comparison of gene order between horse and chimp showed trends very similar to that of horse and human. The only difference is that HSA2 synteny is broken to PTR12 and PTR13 (Jauch et al. 1992), the former sharing homology with ECA15 and the latter with ECA18/ECA6p.

Of the three equine chromosomes under investigation, ECA6p shows the greatest degree of conservation across compared mammals (Figure 7-1). Genes corresponding to ECA6p are present in a single conserved linkage group in human, chimpanzee, mouse, and rat, with the potential for minor rearrangements, while in dog there are two conserved linkage groups. The chicken is also relatively well conserved with respect to ECA6p.

Although the segment corresponding to ECA18 is contained within single human and chimp chromosomes, it is present on three chromosomes in dog, three in mouse, four in rat and two in chicken (Figure 7-1). The segment nearest to ECA18qcen is inverted in horse, dog, mouse, rat, and chicken compared to human and chimpanzee. Further, an apparent inversion between horse and all other vertebrates is seen near ECA18qter, between 190 and 198 Mb position with respect to HSA2.

ECA15 corresponds to seven identifiable segments conserved across compared mammals (Figure 7-1). There is an apparent inversion in gene order in horse, compared to human, with a noticeable break in the HSA2q segment adjacent to the centromere.

Additionally, there are three regions of conserved linkage with respect to HSA2 across all compared species.

Discussion

RH mapping to generate high-resolution physical maps

The potential to develop comprehensive equine chromosome maps with the 5000 rad horse x hamster radiation hybrid cell panel was demonstrated with the publication of the first generation whole-genome RH map (Chowdhary et al. 2003). More recent efforts have focused on the production of higher-resolution physical maps of specific chromosomes, where the maps of ECA7, 10, 17, 21, 22, and X have an average density of 1 marker per Mb (Raudsepp et al. 2004a; Lee et al. 2003a; Brinkmeyer-Langford et al. 2005; Gustafson-Seabury et al. 2005). Using similar approaches, we herein report a higher-resolution physical and comparative map for equine chromosomes that correspond to HSA2.

RH and FISH map of equine orthologs of HSA2

When combined with the previous map data (Chowdhary et al. 2003), the new RH maps for ECA15, 18 and 6p contain 186 total markers (94 Type I and 92 Type II), of which 78 gene-specific and 62 microsatellites are new. Additionally, we added 24 new FISH mapped markers, which anchor 14 and 7 RH mapped markers, respectively, on ECA15 and ECA18. This nearly doubled the number of FISH mapped markers on both ECA15 (20 vs 35) and ECA18 (12 vs 21) (Milenkovic et al. 2002; Chowdhary et al. 2003).

Despite this large increase in total number of mapped markers, several gaps (e.g., 16-24 Mb region of HSA2 on ECA15 and 127-135 Mb and 177-189 Mb intervals of HSA2, on ECA18) of greater than 5 Mb still exist between Type I markers. This is partly attributed to exclusion of some of these markers from the final maps due to low two-point LOD scores. However, the gaps are well covered with microsatellite markers.

Agreement of the RH maps with previously reported maps

The new RH and FISH assignments are in good agreement with the previously reported RH and genetic linkage maps (Chowdhary et al. 2003; Swinburne et al. 2000, 2006; Penedo et al. 2005). The maps for both ECA15 and ECA6p have only minor rearrangements of a few markers with respect to the previously published RH map (LEX046 and COR042 on ECA15 and CHRNG on ECA6p). Two refinements on ECA18 should be addressed, though. The region from SGCV07 (541.99 cR) to MCM6 (672.30 cR) was previously in the reverse orientation, as was the region from HLM3 to COL3A1. The only FISH/RH discrepancy occurs on ECA15 with ASB15 and SPTBN1. FISH mapping places ASB15 on ECA15q21.1-21.3 and SPTBN1 on ECA15q22, while on the RH map they are inverted. ASB15 was placed onto the RH map with a LOD of <0.5; therefore, its location is not as reliable as that of SPTBN1, which is on the MLE-consensus framework and has a LOD >3.

Conservation of gene order and chromosome structure among vertebrates

The high resolution RH maps presented in this study have helped to refine the locations of two evolutionary break/fusion points between HSA2 and the three equine

chromosomes. Previously, the break between ECA18 and ECA6p was shown to lie in an ~11 Mb segment on HSA2 (Chowdhary et al. 2003). We have now narrowed down this region to approximately 5 Mb. Similarly, the breakpoint region between ECA15 and ECA18 has been narrowed down to within ~5 Mb of HSA2, from the previously reported ~9 Mb segment (Chowdhary et al. 2003). The high degree of homology between ECA6p, HSA2, and MMU1 has been noted previously (Mariat et al. 2001; Chowdhary et al. 2003). Our analysis confi rms this homology and extends it to demonstrate that ECA6p is evolutionarily conserved across four of the five sequenced mammals (Figure 7-1). Dog is an exception where this conservation is split between CFA25 and CFA37. When the comparison is extended to include chicken, there are three regions of conserved linkage across vertebrates, supporting the previously noted conservation of this chromosomal arm/segment across species (Mariat et al. 2001; Chowdhary et al. 2003). The evolutionary breakage/ fusion point between HSA2 segments that separates ECA6p and ECA18 appears unique to horse and other equids (Yang et al. 2003b, 2004) among the vertebrate species examined. ECA18 is relatively well conserved in relation to human (HSA2) and chimp (PTR13) although there are a number of apparent minor rearrangements (Figure 7-1). Two notable exceptions to this conservation are inversions between the subcentromeric (*BIN1* to *DDX18*) and distal (ECA18q24→q25; AJ577071 to *COL3A1*) conserved linkages on ECA18, and corresponding segments on HSA2 and PTR13. Regions sharing conserved synteny with ECA18 are present on three different chromosomes in dog (CFA19, CFA36, and CFA37), two in mouse (MMU1 and MMU2), three in rat (RNO3, RNO9, and RNO13),

and surprisingly with primarily a single chromosome in chicken (GGA7). Overall, six core regions representing conserved linkage are observed across mammals and chicken (Figure 7-1). A high conservation in gene order is seen between ECA15, HSA2 and PTR12 (Figure 7-1). The largest region of conserved linkage of ECA15 with the primate chromosomes spans approximately 85–87 Mb and represents ~80% of the chromosome. A minor rearrangement in gene order with respect to primates involves the *IL1A – L1RN* region. The proximal region of ECA15 has four areas of conserved linkages among all the mammals, one of which spans to chicken. The distal half of ECA15 has four regions of conserved linkages across compared mammals, and notably shows a large segment that corresponds to several segments of GGA3. The distal ECA15 regions sharing conserved synteny with other mammals are split among two different chromosomes in dog (CFA10 and 17), three in mouse (MMU11, 12, and 17) and two in rat (RNO6 and 14). These observations suggest that the distal 2/3 of ECA15 represents an ancestral vertebrate chromosomal segment.

Conclusions

This report details the construction of high-resolution RH and comparative maps for equine homologues of HSA2. We have added an additional 140 markers to the equine RH map and 24 markers to the cytogenetic map in an effort to further refine and enhance the maps of ECA6p, ECA18, and ECA15. The availability of high-resolution comprehensive maps for all equine chromosomes will increase the usefulness of these

maps in identifying the chromosomal locations of genes responsible for traits of interest to breeders and veterinarians.

VIII. SUMMARY OF RESULTS AND SALIENT FINDINGS

First paper: A high-resolution physical map of equine homologs of HSA19 shows divergent evolution compared with other mammals

A dense radiation hybrid and comparative gene map was developed for parts of the horse chromosomes that correspond to human chromosome 19 (HSA19) – small parts of both the short (p) and long (q) arms of ECA7, entire short arm of ECA10, and proximal one-third of ECA21 (Raudsepp et al. 1996; Chowdhary et al. 2003; Yang et al. 2004). The map contained a total of 120 markers, of which 89 markers (64 genes and 25 microsatellites) were generated in this study and genotyped on the 5000_{rad} horse x hamster radiation hybrid panel (Chowdhary et al. 2002). This markedly improved the resolution of the map for these chromosomal regions by four-fold, with one marker mapped on average every 700kb. The orientation and anchoring of the RH groups was accomplished by FISH mapping of 15 genes and 9 microsatellites (including 4 additional markers not included on the RH map).

The density of mapped markers facilitated the fine alignment of horse and human genomes. It is evident from the results that the ~0-7 Mb segment of the HSA19p sequence map is represented on the telomeric end of ECA7p and the ~9-13 Mb sequence segment is represented on the subcentromeric part of ECA7q. The 15-20 Mb sequence region of HSA19p corresponds to the proximal two bands of ECA21, while the entire long arm of HSA19 is preserved in a single syntenic block on ECA10p. The segment of ECA21 which shares homology with HSA19 had previously been identified and

described only via Zoo-FISH (Yang et al. 2004); significantly, the radiation hybrid maps generated in this study provided the first gene mapping evidence for this region.

A comparative map of parts of the equine chromosomes that are similar to HSA19 was also developed in relation to corresponding chromosomal regions in chimpanzee – PTR, dog – CFA, mouse – MMU, rat – RNO, chicken – GGA, cattle – BTA, pigs – SSC, cat – FCA and zebrafish – DRE with the help of either whole-genome sequence data or gene mapping data available for individual species. This led to the identification of 9 conserved linkages across the sequenced mammals and 5 putative conserved linkages in relation to chicken.

The segments corresponding to HSA19p in horse (two on ECA7 and one on ECA21) are present as a single syntenic group in the non-rodent mammalian species studied (HSA, PTR, CFA, BTA, and SSC). However, the section which corresponds to ECA21 is present separately from the ECA7-homologous segments in both rodents studied (MMU and RNO) and chicken. Further, the region which corresponds to the long arm of HSA19 is similar to the short arm of ECA10 in horses. This segment shares conserved synteny with human, chimp, and dog, as well as mouse and rat to some degree. In chicken, only the terminal part of ECA10p is conserved as a syntenic block on GGA11; the rest of ECA10p genes are present on several different chicken chromosomes. The segment corresponding to ECA21 is present entirely on GGA28.

Comparative Zoo-FISH analysis of the evolution of ancestral HSA19 segments in eight mammalian orders involving about 50 species revealed that the segment homologous to the short arm of HSA19 is represented as a single block in most species,

and is therefore an independent ancestral chromosome. Further, regions on ECA7 and ECA21 that correspond to the short arm of HSA19 represent a genomic rearrangement discovered to date only in equids/Perissodactyls.

An interesting and unique characteristic of the homology of equine chromosomes with HSA19 is its presence in four separate segments located on three equine chromosomes: the p arm on two chromosomes, and the q arm on one. HSA19p is known to be present on more than one chromosome in only two other species, donkey and zebra (Yang et al. 2003b, 2004). Moreover, regions adjacent to HSA19p-homologous segments (located in tandem or as the sole homolog on the other arm) correspond to several different human chromosomes across species (see Figure 4-3).

Second paper: Development and characterization of a BAC contig over a 5Mb segment of horse chromosome 21

A total of 190 BACs containing 25 equine orthologs of HSA19p genes were isolated via the overgo approach and mapped to ECA21. For an accurate depiction of homology of the region with HSA19p, 17 additional genes and 5 microsatellite markers were selected and mapped. Genes and corresponding BACs located on ECA7 (HSA19 homolog) or 14 (HSA5 homolog) were excluded from analysis.

End sequences of selected BACs were screened for repetitive elements by RepeatMasker (http://www.repeatmasker.org/). This resulted in 74 unique sequence-tagged sites (STSs; see Table 5-2) from which additional primers were designed. These STS primers were subsequently used to order and orient the BACs.

The STS primers were also used for chromosome walking to fill the gaps separating initial groups of overlapping BACs, as described in Raudsepp et al. 2004b. As a result, 15 new BACs were obtained from the CHORI-241 BAC library and 2 from the TAMU BAC library, which provided 28 more STSs for the contig. Eventually 207 BACs were isolated for the region, encompassing 42 genes, 5 microsatellites, and 106 STSs. Of these, 26 BACs comprise the minimum tiling path of the contig, which spans approximately 4.3Mb. The contig is currently the largest for the horse and nearly twice the size of the only other contig available for horses, described in Gustafson et al. 2003.

With 47 markers mapped in this region, the resolution of the contig is ~1 marker every 100kb, reflecting an ~3.8-fold improvement over the previous map of the corresponding region of ECA21. The marker order in the contig is in agreement with the most recent radiation hybrid and linkage maps of the region (Brinkmeyer-Langford et al. 2005; Swinburne et al. 2006).

In humans, the region corresponding to the contig is ~4.7 Mb in length, but in horses it is estimated to be ~4.3 Mb in size, which is 10% smaller than in humans. A similar smaller size of the region compared to humans is also apparent in several other species compared (PTR, BTA, CFA, RNO, MMU, and GGA).

A large region spanning ~4.6 Mb in the humans comprises a family of zinc finger genes that are found only in primates and are absent in horses. This reflects a duplication or expansion event that is primate specific.

The evolutionary breakage/fusion point between HSA19p and HSA5 segments on ECA21 was more precisely defined in this study. The contig shows that ECA21

shares homology with a stretch of sequence on HSA19p extending from the 15.0-19.6Mb positions, whereafter the adjacent segment corresponds to HSA5 starting at the 68.5Mb sequence position.

Third paper: Exceptional conservation of horse–human gene order on X chromosome revealed by high-resolution radiation hybrid mapping

A high-resolution map comprising 175 markers was developed for the horse X chromosome. For this, a total of 116 gene-specific and 12 short tandem repeat markers were added to the map of the equine X chromosome through genotyping on the 5000_{rad} horse x hamster radiation hybrid panel (Chowdhary et al. 2002).

Twenty-nine gene loci were added to the cytogenetic map to anchor and orient the RH groups. Metaphase and interphase FISH were used to resolve the positions of 9 markers for which the order could not be deduced by RH analysis.

Overall, the map has a resolution of 1 mapped marker every ~880kb, with gene-specific markers evenly distributed at 1Mb intervals. Consequently, this is the densest and most uniformly distributed map currently available for the X chromosome of any non-sequenced species. The work also represents the first high-resolution physically integrated gene map for a horse chromosome.

Several genes from the distal part of ECAX tended to produce X and Y specific PCR patterns on the RH panel. This can be explained i) by their putative PAR status in horses, or ii) by retaining a certain degree of sequence homology due to the common ancestry of the sex chromosomes. Two of these (DXY155E and ASMT) are known

PAR1 genes in humans, while the PAR status in horses of other Xp genes needs further verification. Two genes from human PAR2 were mapped to the terminal end of ECAXq and are not pseudoautosomal in horses.

Information from the X chromosome maps of nine mammalian species (HSA, MMU, BTA, SSC, CFA, FCA, CHI - goat, OAR - sheep, and BBU - river buffalo) was used for comparative analysis. The most prominent observation to come from this comparison is that horses and humans share a remarkably similar order of loci, with the exception of a handful of minor rearrangements. For no other non-primate, non-sequenced mammalian species has such an extraordinary conservation been observed to date.

The few rearrangements in gene order between horses and humans are most evident in the segment corresponding to human sequence positions 46-54Mb, found in the horse at ECAXp12-13 (indicated by a red bracket in Figure 6-3). An increased prevalence of rearrangements between the horse and mouse is also observed in this region. These rearrangements in gene order are situated near the putative fusion point of two different strata of the human X chromosome: the X conserved region (XCR) and the recently added region (XRA). As a result, additional study of this particular region could be profitable for understanding the basis behind these rearrangements.

The maps of the horse and mouse X chromosomes are less similar, with a number of rearrangements and reversals in the orientations of clusters of loci (refer to Figures 6-1 and 6-3). The most dramatic differences are on Xp, where many genes are present in humans and not mice. This includes the PAR: humans and mice share no

PAR genes (5 PAR genes in humans are present in mice on autosomes). However, 13 conserved linkage blocks were identified across horses, humans, and mice, denoting preserved X chromosome regions likely inherited from a common ancestor.

The X chromosomes of other species included in the comparisons show considerable differences in gene order and orientation relative to human and horse. These discrepancies extend to their PARs, as several genes (e.g. STS, KAL1) have been found to be classified PAR in some species but not others.

Fourth paper: A 1.3-Mb interval map of equine homologs of HSA2

This study reported a high-resolution map of parts of ECA6p, ECA15 and ECA18 (homologs of HSA2) by analyzing a total of 186 markers (94 Type I and 92 Type II). PCR primers were designed for 140 new markers (78 genes and 62 microsatellites). These markers were then genotyped on the 5000_{rad} horse x hamster radiation hybrid panel (Chowdhary et al. 2002). Twenty-four of these markers were cytogenetically mapped by FISH for anchoring and orienting the RH groups on ECA15 and 18. The resultant maps of the equine chromosomes had a density of one marker every 1.3Mb which translates to a three-fold increase compared to previous maps of the regions.

The high-resolution RH maps of the three equine chromosomes helped to refine the locations of two evolutionary break/fusion points between HSA2 and the three equine chromosomes. The previous break for HSA2 synteny between ECA18 and ECA6p was detected in an ~11 Mb segment on HSA2 (Chowdhary et al. 2003). This

has now been narrowed down to an approximately 3 Mb region, and appears to be unique to equids as it is not observed in any of the other species compared.. Similarly, the breakpoint region between ECA15 and ECA18 has been narrowed down to within ~4 Mb of HSA2, from the previously reported ~9 Mb segment (Chowdhary et al. 2003).

Comparison of the gene mapping data for the three equine chromosomes with that available in other mammalian and vertebrate species showed that ECA18 is relatively well conserved in relation to human (HSA2), chimp (PTR13), and chicken (GGA7). However, with regards to dog, mouse, and rat, the corresponding segments are present on 2-3 chromosomes. Overall, six core regions representing conserved linkages are observed across mammals and chicken (see Fig. 7-1: green shaded regions and regions demarcated with red rectangles). In contrast, regions corresponding to ECA6p primarily show conserved linkage across four of the five sequenced mammals, with dog as an exception where this conservation is split between CFA25 and CFA37.

ECA15 seems to be well conserved among mammals and notably shows a large segment (distal half of ECA15) that corresponds to part of chicken chromosome 3. The inverted gene order of the proximal part of ECA15 with respect to primates, but similar to that of dog, suggests an evolutionary reorganization between these species. The distal half of ECA15 has four regions of conserved linkages across compared mammals (Fig. 7-1: green shaded rectangles). However, comparison of the gene order with chicken essentially shows conservation implying that the distal 2/3 of ECA15 represents an ancestral vertebrate chromosomal segment.

IX. FUTURE PROSPECTS

A better understanding of the molecular and genetic bases behind equine diseases and disorders is of utmost importance to the equine industry. Horse breeders yearn to find ways to improve breeding strategies such that the health and (in many cases) athletic ability of the foals are maximized for greater gains at shows and race tracks; fertility of the breeding stock is also crucial to a successful and productive industry. Improvements in equine veterinary medicine are necessary as well to diagnose and treat or even prevent a number of maladies in horses, many of which require cures rather than symptomatic treatments. It is true that not every condition or disease in horses has a genetic cause. However, considering that equine biology, like that of any other living being, is governed by genes and their interactions, an in-depth understanding of the genome is expected to equip us with knowledge critical to discerning their roles in causing and mediating these conditions or predisposing horses to specific diseases. Given the substantial size of the equine industry and its impact on the economy, it is apparent that an increase in knowledge of the horse genome – its structure, organization, functions, and interactions – will be of paramount relevance during coming years.

Gene mapping and genome analysis for the horse have undergone rapid and significant advances over the last several years. Physical as well as genetic maps have been developed for the entire horse genome, rivaling those available in other domestic species where the process started much earlier. Despite this, the gene maps currently available for several horse chromosomes lack sufficient marker density or a uniform marker distribution. Hence, high-resolution maps with evenly-spaced markers are

essential for all chromosomes. The second-generation whole-genome radiation hybrid map for the horse as well as other projects focused on improving maps for individual chromosomes will continue to be the major immediate goals for the equine gene mapping community.

Two additional fronts with respect to the analysis of the equine genome will remain the center of focus for the current and next years. These are i) physical assembly of the genome using BAC clones and ii) whole genome sequencing and assembly of the genome which will be followed by annotation. For the former, ~150,000 BAC clones from the CHORI-241 equine BAC library are being end sequenced and will be fingerprinted to eventually obtain a complete physical map of the horse genome (http://www.volkswagenstiftung.de/presse-news/presse05/08122005.pdf). Using comparative mapping information and human gene annotation, the putative gene content of the equine BAC clones is being inferred. Additionally, the equine chromosome assignment of the BAC clones is predicted based on the currently available human-horse comparative maps. With regards to sequencing the equine genome, major progress is presently being achieved with the recently-initiated endeavor at the Broad Institute. While the first part of this initiative is completed and a 7X sequence of the equine genome is available, the next phase will involve assembly of the data and eventual annotation to identify and classify various features of the entire sequence. High resolution maps such as those presented in this thesis will be extremely useful in assembly as well as annotation of the sequence data finally leading to a finished sequence probably by the end of the next year.

The availability of the 7X whole genome sequence of the horse described above will trigger the creation of two major resources: i) long oligonucleotide microarrays and ii) SNP chips. These tools will be extensively used during the coming years respectively for functional analysis of the equine genome and for analyses of Mendelian as well as complex traits. It is expected that the annotated and assembled 7X-coverage sequence of the horse genome and the ~65,000 ESTs publicly available in various databases will be used in the near future to bioinformatically develop 70-mer probes for individual equine genes. These oligos will then be used to generate the first whole genome equine oligoarray, which will be a vital instrument for conducting organized and genome-wide functional genomics research in several branches of equine sciences. The resource will for the first time enable analyses of patterns of gene expression in horses under normal and perturbed physiological and developmental conditions. It will also facilitate comparisons of gene expression patterns between horses that vary with regard to genetically influenced, intrinsic, or environmentally induced health phenotypes.

SNP discovery is currently underway in conjunction with the 7X whole genome shotgun sequencing of the equine genome at the Broad Institute of Harvard and MIT, funded by the National Human Genome Research Institute. This will generate a comprehensive selection of equine SNPs and enable their validation in diverse breeds by re-sequencing predetermined regions of the genome in representative individuals. Eventually this will lead to the development of a highly informative whole genome SNP chip with applications in association studies of health and disease traits in different breeds of horses. This tool will be extensively used in the coming years not only to

understand equine biology, but also to study simple as well as complex diseases for which present testing and diagnostic tools do not provide conclusive results. It is expected that the findings from SNP analysis will allow the equine industry to make informed decisions regarding breeding programs.

An avenue of research that has been under-appreciated until recently is the Y chromosome. While contigs have been developed for part of the euchromatic region of the equine Y (Raudsepp et al. 2004b), complete sequence information for this chromosome will not be available through the whole genome sequencing effort because a female horse was used for this purpose. The existing contigs should serve as a template to ultimately provide a tenable scaffold for organizing future sequence data for the Y chromosome. As this chromosome plays important roles in governing male fertility as well as differences in phenotypic sex, developing a sequence-level map is of prime importance. With such information in hand, the in-depth scrutiny of the chromosome can lead to an improved understanding of these factors in the horse.

As a result of present and future developments in equine genomics, the field of equine biology will be revolutionized. Advances ensuing from horse genomics research will provide valuable resources for clinicians and investigators. These include more accurate and efficient means for diagnosing affected horses, such that proper treatment can commence. Additionally, treatment methods themselves will improve, becoming more focused on the molecular aspects (such as genes) involved in the development and progress of conditions, and tailored to the affected individual. Moreover, more sophisticated preventative measures will be possible, such as avoiding exposure of

susceptible animals to certain risk factors. While genomics will not replace current practices in horse care, it will provide additional beneficial tools which will improve the efficacy of equine medicine. Genomics will facilitate the elucidation of both environmental and genetic components of diseases and traits, providing valuable details particularly about genetic aspects of conditions and how they are affected by external circumstances. Ultimately, progress in horse genomics will usher in a new era in equine medicine and biology.

X. CONCLUSIONS

The work described in this thesis led to the development of high-resolution, integrated gene maps for seven equine chromosomes. The maps permitted detailed comparisons of the horse genome with the genomes of mammalian and vertebrate species. This helped to determine evolutionarily conserved chromosomal segments across species and provided insight into their evolution from an ancestral chromosomal configuration. In the future, these maps will be useful for assigning traits of economic significance to specific chromosomal regions in the horse and subsequently identifying and analyzing genes governing them. Lastly, the maps will serve as templates upon which the newly acquired sequence information for individual chromosomes will be assembled. Overall, the studies described herein advance our current understanding of various aspects of the equine genome. The following observations are key features of this advancement:

- High-resolution integrated physical gene maps were developed for seven equine chromosomes: the three equine homologs of HSA19 (parts of ECA7, 10, and 21), the three equine homologs of HSA2 (parts of ECA6, 15, and 18), and the X chromosome. This was accomplished by generating a total of ~500 markers (99 microsatellites, 106 STSs, and 291 genes) and genotyping them on the TAMU 5000_{rad} horse x hamster RH panel. Selected markers were used to screen the CHORI-241 and TAMU horse BAC libraries to obtain BACs that were localized to specific chromosomal bands by FISH. This in turn facilitated the anchoring

and orientation of the RH groups to the chromosomes. Additionally, a contig comprising 207 overlapping BAC clones was established over the proximal part of ECA21 that corresponds to HSA19p and part of HSA5.

- Comparisons of the resulting maps for the equine chromosomes with those available for corresponding genomic regions in other mammalian and vertebrate species allowed the precise demarcation of correspondences between the equine chromosomal segments and chromosomes of the compared species. It also enabled the identification of regions of conserved synteny or evolutionarily conserved linkages across these species that permitted the detection of putative ancestral chromosomal regions between them.

- Following analysis of Zoo-FISH and comparative gene mapping data from ~50 mammalian species, a putative ancestral mammalian configuration was deduced for HSA19. This enabled postulation of the evolution of HSA19 homologs in the horse and other species from a common ancestor. In addition to facilitating the determination of possible fission and fusion events that led to the present chromosome configurations in individual species, the comparisons made possible the identification of a Perissodactyl-specific fission event. As a result of this event, HSA19p corresponds to two separate chromosomes in the equids, but not in any of the other compared mammalian species. Further, while the adjacency of segments corresponding to HSA19 and HSA5 is observed in artiodactyls as well

as Perissodactyls, different regions are involved in the latter, indicating separate evolutionary events in the two orders.

- The BAC contig developed over the proximal one-third of ECA21 showed that this region corresponds to the ~15-20 Mb positions on the HSA19 sequence map. The BAC contig spans ~4.3Mb in length and is presently the largest assembly of overlapping BAC clones available for the horse. Interesting findings conferred by the contig include the apparent lack of a group of zinc finger genes in horses which is present as a large block in humans and other primates. This was confirmed by screening the CHORI-241 library filters using overgo probes for several zinc finger genes from the human region. In addition, diversity in the size of the contig is apparent particularly in regions of gene families.

- The high-resolution gene map of the equine X chromosome revealed a degree of conservation with the counterpart in humans that has not yet been observed in any of the other non-sequenced mammalian species compared.

- The gene maps for ECA6, ECA15, and ECA18 represent the first high-resolution maps for these equine chromosomes. Comparison of the mapping data with HSA2 led to the identification of two evolutionary breakpoints: one between ECA6 and ECA18, and another between ECA15 and ECA18. The breakpoint

between ECA6 and ECA18 is considered to be equid-specific because it has not been found in any of the other compared mammalian or vertebrate species.

- The radiation hybrid maps described in this dissertation demonstrate the far-reaching utility of the approach for integrating information from a variety of other map types (e.g. synteny, linkage, and cytogenetic) into a single, highly informative, linearly ordered consensus map. As is evident from the results, these maps are tremendously effective for making comparisons with gene maps of other species and extrapolating information, as well as elucidating the details of the evolution of a particular region or chromosome. Most importantly, the resolution attained for the seven chromosomes will considerably facilitate candidate gene searches for a range of equine conditions.

REFERENCES

1. Aerts J, Crooijmans R, Cornelissen S, Hemmatian K, Veenendaal T, et al. (2003) Integration of chicken genomic resources to enable whole-genome sequencing. Cytogenet Genome Res 102, 297-303

2. Agarwala R Applegate DL, Maglott D, Schuler GD, Schäffer AA (2000) A fast and scalable radiation hybrid map construction and integration strategy. Genome Res 10, 350-364

3. Aleman M, Riehl J, Aldridge BM, Lecouteur RA, Stott JL, et al. (2004) Association of a mutation in the ryanodine receptor 1 gene with equine malignant hyperthermia. Muscle Nerve 30, 356-365

4. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST - A new generation of protein database search programs. Nuc Acids Res 25, 3389-3402

5. Amaral ME, Kata SR, Womack JE (2002) A radiation hybrid map of bovine X chromosome (BTAX). Mamm Genome 13, 268-271

6. Andersson L, Sandberg K (1982) A linkage group composed of three coat color genes and three serum protein loci in horses. J Hered 73, 91-94

7. Angius A, Petretto E, Maestrale GB, Forabosco P, Casu G, et al. (2002) A new essential hypertension susceptibility locus on chromosome 2p24-p25, detected by genomewide search. Am J Hum Genet 71, 893-905

8. Ansari HA, Hediger R, Fries R, Stranzinger G (1988) Chromosomal localization of the major histocompatibility complex of the horse (ELA) by in situ hybridization. Immunogenetics 28, 362-364

9. Applegate D, Bixby R, Chvátal V, Cook W (1998) On the solution of traveling salesman problems. Documenta mathematica, extra volume International Congress of Mathematics 1998 III, 645-656

10. Arngrimsson R, Sigurardottir S, Frigge ML, Bjarnadottir RI, Jonsson T, et al. (1999) A genome-wide scan reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13. Hum Mol Genet 8, 1799-1805

11. B. H. F. Family Heart Study Research Group (2005) A genomewide linkage study of 1,933 families affected by premature coronary artery disease: The British Heart

Foundation (BHF) Family Heart Study. Am J Hum Genet 77, 1011-1020

12. Bai Q, McGillivray C, da Costa N, Dornan S, Evans G, et al. (2003) Development of a porcine skeletal muscle cDNA microarray: analysis of differential transcript expression in phenotypically distinct muscles. BMC Genomics 1, 8

13. Bailey E, Graves KT, Cothran EG, Reid R, Lear TL, et al. (1995) Synteny-mapping horse microsatellite markers using a heterohybridoma panel. Anim Genet 26, 177-180

14. Bailey E, Reid RC, Lear TL, Skow LC, Mathiason K, et al. (1997) Linkage of the gene for equine combined immunodeficiency disease to microsatellite markers HTG8 and HTG4; synteny and FISH mapping to ECA9. Anim Genet 28, 268-273

15. Band MR, Larson JH, Rebeiz M, Green CA, Heyen DW, et al. (2000) An ordered comparative map of the cattle and human genomes. Genome Res 10, 1359-1368

16. Band MR, Olmstead C, Everts RE, Liu ZL, Lewin HA (2002) A 3800 gene microarray for cattle functional genomics: comparison of gene expression in spleen, placenta, and brain. Anim Biotechnol 13, 163-172

17. Barbosa A, Demeure O, Urien C, Milan D, Chardon P, et al. (2004) A physical map of large segments of pig Chromosome 7q→q14: comparative analysis with human Chromosome 6p21. Mamm Genome V15, 982-995

18. Barendse W, Armitage SM, Kossarek LM, Shalom A, Kirkpatrick BW, et al. (1994) A genetic linkage map of the bovine genome. Nat Genet 6, 227-235

19. Beck TW, Menninger J, Murphy WJ, Nash WG, O'Brien SJ, et al. (2005) The feline major histocompatibility complex is rearranged by an inversion with a breakpoint in the distal class I region. Immunogenetics V56, 702-709

20. Beever JE, Smit MA, Meyers SN, Hadfield TS, Bottema C, et al. (2006) A single-base change in the tyrosine kinase II domain of ovine FGFR3 causes hereditary chondrodysplasia in sheep. Anim Genet 37, 66-71

21. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, et al. (2004) Ultraconserved elements in the human genome. Science 304, 1321-1325

22. Bellefroid EJ, Marine J, Matera AG, Bourguignon C, Desai T, et al. (1995) Emergence of the ZNF91 Kruppel-associated box-containing zinc finger gene family in the last common ancestor of *Anthropoidea*. Proc Nat Acad Sci USA 92, 10757-10761

23. Ben-Dor A, Chor B (1997) On constructing radiation hybrid maps. J Comput Biol 4, 517-533

24. Berryere TG, Kerns JA, Barsh GS, Schmutz SM (2005) Association of an Agouti allele with fawn or sable coat color in domestic dogs. Mamm Genome 16, 262-272

25. Bielec PE, Gallagher DS, Womack JE, Busbee DL (1998) Homologies between human and dolphin chromosomes detected by heterologous chromosome painting. Cytogenet Genome Res 81, 18-25

26. Birnboim H (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 100, 243-255

27. Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nuc Acids Res 7, 1513-1523

28. Blaschke RJ, Rappold GA (1997) Man to mouse--lessons learned from the distal end of the human X chromosome. Genome Res 7, 1114-1117

29. Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E, et al. (2002) A comprehensive collection of chicken cDNAs. Curr Biol 12, 1965-1969

30. Boehnke M (1992) Multipoint analysis for radiation hybrid mapping. Ann Med 24, 383-386

31. Bortoluzzi S, Rampoldi L, Simionati B, Zimbello R, Barbon A, et al. (1998) A comprehensive, high-resolution genomic transcript map of human skeletal muscle. Genome Res 8, 817-825

32. Bosak N, Yamomoto R, Fujisaki S, Faraut T, Kiuchi S, et al. (2005) A dense comparative gene map between human chromosome 19q13.3→q13.4 and a homologous segment of swine chromosome 6. Cytogenet Genome Res 108, 317-321

33. Bourneuf E, Herault F, Chicault C, Carre W, Assaf S, et al. (2006) Microarray analysis of differential gene expression in the liver of lean and fat chickens. Gene 372, 162-170

34. Bowling AT, Ruvinksy A (2000) Genetic aspects of domestication, breeds and their origins. In The Genetics of the Horse, A.T. Bowling, Ruvinsky, A., ed. (New York, NY: CABI Publishing), pp 25-51

35. Boyd Y, Blair HJ, Cunliffe P, Masson WK, Reed V (2000) A phenotype map of the mouse X chromosome: models for human X-linked disease. Genome Res 10,

277-292

36. Breen M, Thomas R, Binns MM, Carter NP, Langford CF (1999) Reciprocal chromosome painting reveals detailed regions of conserved synteny between the karyotypes of the domestic dog (*Canis familiaris*) and human. Genomics 61, 145-155

37. Breen M, Jouquand S, Renier C, Mellersh CS, Hitte C, et al. (2001) Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. Genome Res 11, 1784-1795

38. Breen M, Hitte C, Lorentzen TD, Thomas R, Cadieu E, et al. (2004) An integrated 4249 marker FISH/RH map of the canine genome. BMC Genomics 5, 65

39. Briggs MD, Hoffman SMG, King LM, Olsen AS, Mohrenweiser H, et al. (1995) Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. Nat Genet 10, 330-336

40. Brinkmeyer-Langford C, Raudsepp T, Lee E-J, Goh G, Schäffer A, et al. (2005) A high-resolution physical map of equine homologs of HSA19 shows divergent evolution compared with other mammals. Mamm Genome 16, 631-649

41. Brooks SA, Terry RB, Bailey E (2002) A PCR-RFLP for KIT associated with tobiano spotting pattern in horses. Anim Genet 33, 301-303

42. Brooks SA, Bailey E (2005) Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. Mamm Genome 16, 893-902

43. Burgoyne PS, Ojarikre OA, Turner JM (2002) Evidence that postnatal growth retardation in XO mice is due to haploinsufficiency for a non-PAR X gene. Cytogenet Genome Res 99, 252-256

44. Burkin DJ, Yang F, Broad TE, Wienberg J, Hill DF, et al. (1997) Use of the Indian muntjac idiogram to align conserved chromosomal segments in sheep and human genomes by chromosome painting. Genomics 46, 143-147

45. Burnside J, Neiman P, Tang J, Basom R, Talbot R, et al. (2005) Development of a cDNA array for chicken gene expression analysis. BMC Genomics 6, 13

46. Caetano AR, Pomp D, Murray JD, Bowling AT (1999a) Comparative mapping of 18 equine type I genes assigned by somatic cell hybrid analysis. Mamm Genome 10, 271-276

47. Caetano AR, Shiue YL, Lyons LA, O'Brien SJ, Laughlin TF, et al. (1999b) A comparative gene map of the horse (*Equus caballus*). Genome Research 9, 1239-1249

48. Caetano AR, Johnson RK, Ford JJ, Pomp D (2004) Microarray profiling for differential gene expression in ovaries and ovarian follicles of pigs selected for increased ovulation rate. Genetics 168, 1529-1537

49. Cai L, Taylor JF, Wing RA, Gallagher DS, Woo SS, et al. (1995) Construction and characterization of a bovine bacterial artificial chromosome library. Genomics 29, 413-425

50. Charchar FJ, Svartman M, El-Mogharbel N, Ventura M, Kirby P, et al. (2003) Complex events in the evolution of the human pseudoautosomal region 2 (PAR2). Genome Res 13, 281-286

51. Charlesworth B (1991) The evolution of sex chromosomes. Science 251, 1030-1033

52. Chase K, Carrier DR, Adler FR, Jarvik T, Ostrander EA, et al. (2002) Genetic basis for systems of skeletal quantitative traits: Principal component analysis of the canid skeleton. Proc Nat Acad Sci USA 99, 9930-9935

53. Chaudhary R, Raudsepp T, Guan XY, Zhang H, Chowdhary BP (1998) Zoo-FISH with microdissected arm specific paints for HSA2, 5, 6, 16, and 19 refines known homology with pig and horse chromosomes. Mamm Genome 9, 44-49

54. Chaves R, Frönicke L, Guedes-Pinto H, Wienberg J (2004) Multidirectional chromosome painting between the Hirola antelope (*Damaliscus hunteri*, Alcelaphini, Bovidae), sheep and human. Chromosome Res 12, 495-503

55. Cheung J, Wilson MD, Zhang J, Khaja R, MacDonald JR, et al. (2003) Recent segmental and gene duplications in the mouse genome. Genome Biol 4, R47.

56. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Muller WEG, et al. (2004) Using the miraEST Assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res 14, 1147-1159

57. Chi J, Fu B, Nie W, Wang J, Graphodatsky AS, et al. (2005) New insights into the karyotypic relationships of Chinese muntjac (*Muntiacus reevesi*) , forest musk deer (*Moschus berezovskii*) and gayal (*Bos frontalis*). Cytogenet Genome Res 108, 310-316

58. Chowdhary BP, Frönicke L, Gustavsson I, Scherthan H (1996) Comparative analysis of the cattle and human genomes: detection of ZOO-FISH and gene mapping-based chromosomal homologies. Mamm Genome 7, 297-302

59. Chowdhary BP, Raudsepp T, Frönicke L, Scherthan H (1998) Emerging patterns of comparative genome organization in some mammalian species as revealed by Zoo-FISH. Genome Res. 8, 577-589

60. Chowdhary BP, Raudsepp T, Honeycutt D, Owens EK, Piumi F, et al. (2002) Construction of a 5000(rad) whole-genome radiation hybrid panel in the horse and generation of a comprehensive and comparative map for ECA11. Mamm Genome 13, 89-94

61. Chowdhary BP, Raudsepp T, Kata SR, Goh G, Millon LV, et al. (2003) The first-generation whole-genome radiation hybrid map in the horse identifies conserved segments in human and mouse genomes. Genome Research 13, 742-751

62. Cirera S, Jorgensen CB, Sawera M, Raudsepp T, Chowdhary BP, et al. (2003) Comparative mapping in the pig: localization of 214 expressed sequence tags. Mamm Genome 14, 405-426

63. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, et al. (1998) New goals for the U.S. human genome project: 1998-2003. Science 282, 682-689

64. Collins FS, Green ED, Guttmacher AE, Guyer MS (2003a) A vision for the future of genomics research. Nature 422, 835-847

65. Collins FS, Morgan M, Patrinos A (2003b) The human genome project: lessons from large-scale biology. Science 300, 286-290

66. Cox DR, Burmeister M, Price ER, Kim S, Myers RM (1990) Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. Science 250, 245-250

67. Davoli R, Fontanesi L, Zambonelli P, Bigi D, Gellin J, et al. (2002) Isolation of porcine expressed sequence tags for the construction of a first genomic transcript map of the skeletal muscle in pig. Anim Genet 33, 3-18

68. Dehal P, Predki P, Olsen AS, Kobayashi A, Folta P, et al. (2001) Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. Science 293, 104-111

69. Demars J, Riquet J, Feve K, Gautier M, Morisson M, et al. (2006) High resolution physical map of porcine chromosome 7 QTL region and comparative mapping of

this region among vertebrate genomes. BMC Genomics 7

70. Deys BF (1972) Demonstration of X-linkage of G6PD, HGPRT, and PGK in the horse by means of mule-mouse cell hybridization. Ph. D. Thesis, University of Leiden

71. Dixkens C, Klett C, Bruch J, Kollak A, Serov OL, et al. (1998) Zoo-FISH analysis in insectivores: "Evolution extols the virtue of the status quo". Cytogenet Cell Genet 80, 61-67

72. Djilali-Saiah I, Schmitz J, Harfouch-Hammoud E, Mougenot JF, Bach JF, et al. (1998) CTLA-4 gene polymorphism is associated with predisposition to coeliac disease. Gut 43, 187-189

73. Donaldson L, Vuocolo T, Gray C, Strandberg Y, Reverter A, et al. (2005) Construction and validation of a bovine innate immune microarray. BMC Genomics 22, 135

74. Dvorak CM, Hyland KA, Machado JG, Zhang Y, Fahrenkrug SC, et al. (2005) Gene discovery and expression profiling in porcine Peyer's patch. Vet Immunol Immunopathol 105, 301-315

75. Eichler EE, Hoffman SM, Adamson AA, Gordon LA, McCready P, et al. (1998) Complex beta-satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12. Genome Res 8, 791-808

76. Ellegren H, Johansson M, Sandberg K, Andersson L (1992) Cloning of highly polymorphic microsatellites in the horse. Anim Genet 23, 133-142

77. Ellison JW, Li X, Francke U, Shapiro LJ (1996) Rapid evolution of human pseudoautosomal genes and their mouse homologs. Mamm Genome 7, 25-30

78. Ephrussi B, Weiss MC (1969) Hybrid somatic cells. Sci Am 220, 26-35

79. Everts RE, van Wolferen ME, Versteeg SA, Zijlstra C, Engelen JJ, et al. (2002) A radiation hybrid map of the X-chromosome of the dog (*Canis familiaris*). Cytogenet Genome Res 98, 86-92

80. Everts RE, Band MR, Liu ZL, Kumar CG, Liu L, et al. (2005) A 7872 cDNA microarray and its use in bovine functional genomics. Vet Immunol Immunopathol 105, 235-245

81. Everts-van der Wind A, Kata SR, Band MR, Rebeiz M, Larkin DM, et al. (2004) A 1463 gene cattle-human comparative map with anchor points defined by human

genome sequence coordinates. Genome Res. 14, 1424-1437

82. Everts-van der Wind A, Larkin DM, Green CA, Elliott JS, Olmstead CA, et al. (2005) A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. Proc Nat Acad Sci USA 102, 18526-18531

83. Falchi M, Forabosco P, Mocci E, Borlino CC, Picciau A, et al. (2004) A genomewide search using an original pairwise sampling approach for large genealogies identifies a new locus for total and low-density lipoprotein cholesterol in two genetically differentiated isolates of Sardinia. Am J Hum Genet 75, 1015-1031

84. Fan Y, Linardopoulou E, Friedman C, Williams E, Trask BJ (2002a) Genomic structure and evolution of the ancestral chromosome fusion site in 2q13-2q14.1 and paralogous regions on other human chromosomes. Genome Res 12, 1651-1662

85. Fan Y, Newman T, Linardopoulou E, Trask BJ (2002b) Gene content and function of the ancestral chromosome fusion site in human chromosome 2q13-2q14.1 and paralogous regions. Genome Res 12, 1663-1672

86. Felsburg PJ, Somberg RL, Perryman LE (1992) Domestic animal models of severe combined immunodeficiency: canine X-linked severe combined immunodeficiency and severe combined immunodeficiency in horses. Immunodefic Rev 3, 277-303

87. Ferlin A, Bogatcheva NV, Gianesello L, Pepe A, Vinanzi C, et al. (2006) Insulin-like factor 3 gene mutations in testicular dysgenesis syndrome: clinical and functional characterization. Mol Hum Reprod 12, 401-406

88. Fitzgerald J, Bateman JF (2004) Why mice have lost genes for COL21A1, STK17A, GPR145 and AHRI: evidence for gene deletion at evolutionary breakpoints in the rodent lineage. Trends in Genetics 20, 408-412

89. Frank N, Sojka J, Messer IVNT (2002) Equine thyroid dysfunction. Veterinary Clinics of North America: Equine Practice 18, 305-319

90. Frankeny RL (2003) Miniature Horses. (North Pomfret: Trafalgar Square Publishing)

91. Frönicke L, Chowdhary BP, Scherthan H, Gustavsson I (1996) A comparative map of the porcine and human genomes demonstrates ZOO-FISH and gene mapping-based chromosomal homologies. Mamm Genome 7, 285-290

92. Frönicke L, Wienberg J, Stone G, Adams L, Stanyon R (2003) Towards the delineation of the ancestral eutherian genome organization: comparative genome maps of human and the African elephant (*Loxodonta africana*) generated by chromosome painting. Proc R Soc Lond B Biol Sci. 270, 1331-1340

93. Frydendahl K, Kåre Jensen T, Strodl Andersen J, Fredholm M, Evans G (2003) Association between the porcine *Escherichia coli* F18 receptor genotype and phenotype and susceptibility to colonisation and postweaning diarrhoea caused by *E. coli* O138:F18. Veterinary Microbiology 93, 39-51

94. Gao Y, Hu XX, Du ZQ, Deng XM, Huang YH, et al. (2006) A genome scan for quantitative trait loci associated with body weight at different developmental stages in chickens. Anim Genet 37, 276-278

95. Gardner EJ, Shupe JL, Leone NC, Olson AE (1975) Hereditary multiple exostosis. A comparative genetic evaluation in man and horses. J Hered 66, 318-326

96. Gautier M, Hayes H, Bønsdorff T, Eggen A (2003) Development of a comprehensive comparative radiation hybrid map of bovine chromosome 7 (BTA7) versus human chromosomes 1 (HSA1), 5 (HSA5) and 19 (HSA19). Cytogenet Genome Res 102, 25-31

97. Gianfrancesco F, Sanges R, Esposito T, Tempesta S, Rao E, et al. (2001) Differential divergence of three human pseudoautosomal genes and their mouse homologs: implications for sex chromosome evolution. Genome Res 11, 2095-2100

98. Godard S, Vaiman D, Oustry A, Nocart M, Bertaud M, et al. (1997) Characterization, genetic and physical mapping analysis of 36 horse plasmid and cosmid-derived microsatellites. Mamm Genome 8, 745-50

99. Godard S, Vaiman A, Schibler L, Mariat D, Vaiman D, et al. (2000) Cytogenetic localization of 44 new coding sequences in the horse. Mamm Genome 11, 1093-1097

100. Goh G, Raudsepp T, Durkin K, Wagner ML, Schäffer AA, et al. (2006) High-resolution gene maps of horse chromosomes 14 and 21: Additional insights into evolution and rearrangements of HSA5 homologs in mammals. Genomics (in press)

101. Goldammer T, Kata SR, Brunner RM, Dorroch U, Sanftleben H, et al. (2002) A comparative radiation hybrid map of bovine chromosome 18 and homologous chromosomes in human and mice. Proc Nat Acad Sci USA 99, 2106-2111

102. Goss SJ, Harris H (1975) New method for mapping genes in human chromosomes. Nature 255, 680-684

103. Goureau A, Yerle M, Schmitz A, Riquet J, Milan D, et al. (1996) Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. Genomics 36, 252-262

104. Graphodatsky AS, Yang F, Serdukova N, Perelman P, Zhdanova NS, et al. (2000) Dog chromosome-specific paints reveal evolutionary inter- and intrachromosomal rearrangements in the American mink and human. Cytogenet Genome Res 90, 275-278

105. Graphodatsky AS, Yang F, O'Brien PCM, Perelman P, Milne BS, et al. (2001) Phylogenetic implications of the 38 putative ancestral chromosome segments for four canid species. Cytogenet Genome Res 92, 243-247

106. Graves JA, Delbridge ML (2001) The X--a sexy chromosome. Bioessays 23, 1091-1094

107. Graves JA, Gecz J, Hameister H (2002) Evolution of the human X--a smart and sexy chromosome that controls speciation and development. Cytogenet Genome Res 99, 141-145

108. Gregory SG, Sekhon M, Schein J, Zhao S, Osoegawa K, et al. (2002) A physical map of the mouse genome. Nature 418, 743-750

109. Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, et al. (2004) The DNA sequence and biology of human chromosome 19. Nature 428, 529-535

110. Grobet L, Royo Martin LJ, Poncelet D, Pirottin D, Brouwers B, et al. (1997) A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. Nat Genet 17, 71-74

111. Groenen MAM, Cheng HH, Bumstead N, Benkel BF, Briles WE, et al. (2000) A Consensus Linkage Map of the Chicken Genome. Genome Res 10, 137-147

112. Groenen MAM, Crooijmans RPMA (2003) Structural genomics: integrating linkage, physical and sequence maps. In Poultry Genetics, Breeding and Technology, W.M. Muir, S.E. Aggrey, eds. (Wallingford: CAB International), pp 497–536

113. Guénet JL (2005) The mouse genome. Genome Res 15, 1729-1740

114. Guérin G, Bailey E, Bernoco D, Anderson I, Antczak DF, et al. (1999) Report of the International Equine Gene Mapping Workshop: male linkage map. Anim Genet 30, 341-354

115. Guérin G, Bailey E, Bernoco D, Anderson I, Antczak DF, et al. (2003) The second generation of the International Equine Gene Mapping Workshop half-sibling linkage map. Anim Genet 34, 161-168

116. Gustafson AL, Tallmadge RL, Ramlachan N, Miller D, Bird H, et al. (2003) An ordered BAC contig map of the equine major histocompatibility complex. Cytogenet Genome Res 102, 189-195

117. Gustafson-Seabury A, Raudsepp T, Goh G, Kata SR, Wagner ML, et al. (2005) High-resolution RH map of horse chromosome 22 reveals a putative ancestral vertebrate chromosome. Genomics 85, 188-200

118. Guyon R, Lorentzen TD, Hitte C, Kim L, Cadieu E, et al. (2003) A 1-Mb resolution radiation hybrid map of the canine genome. Proc Nat Acad Sci USA 100, 5296-5301

119. Hameister H, Klett C, Bruch J, Dixkens C, Vogel W, et al. (1997) Zoo-FISH analysis: the American mink (Mustela vison) closely resembles the cat karyotype. Chromosome Res 5, 5-11

120. Hao L, Klein J, Nei M (2006) Heterogeneous but conserved natural killer receptor gene complexes in four major orders of mammals. Proc Nat Acad Sci USA 103, 3192-3197

121. Harbitz I, Chowdhary BP, Saether H, Hauge JG, Gustavsson I (1990) A porcine genomic glucosephosphate isomerase probe detects a multiallelic restriction fragment length polymorphism assigned to chromosome 10pter in horse. Hereditas 112, 151-156

122. Hawken RJ, Murtaugh J, Flickinger GH, Yerle M, Robic A, et al. (1999) A first-generation porcine whole-genome radiation hybrid map. Mamm Genome 10, 824-830

123. Hegasy GA, Manuelian T, Hogasen K, Jansen JH, Zipfel PF (2002) The molecular basis for hereditary porcine membranoproliferative glomerulonephritis type II: point mutations in the factor H coding sequence block protein secretion. Am J Pathol 161, 2027-2034

124. Helou K, Walentinsson A, Levan GR, Stähl F (2001) Between rat and mouse zoo-FISH reveals 49 chromosomal segments that have been conserved in evolution.

Mamm Genome 12, 765-771

125. Henner J, Poncet PA, Guérin G, Hagger C, Stranzinger G, et al. (2002) Genetic mapping of the (G)-locus, responsible for the coat color phenotype "progressive greying with age" in horses (*Equus caballus*). Mamm Genome 13, 535-537

126. Henricksen JD (1971) Horseback riding for the handicapped. Archives of Physical Medicine and Rehabilitation 52, 282-283

127. Hillier LW, Graves TA, Fulton RS, Fulton LA, Pepin KH, et al. (2005) Generation and annotation of the DNA sequences of human chromosomes 2 and 4. Nature 434, 724-731

128. Hitte C, Lorentzen TD, Guyon R, Kim L, Cadieu E, et al. (2003) Comparison of MultiMap and TSP/CONCORDE for constructing radiation hybrid maps. J Hered 94, 9-13

129. Hitte C, Madeoy J, Kirkness EF, Priat C, Lorentzen TD, et al. (2005) Facilitating genome navigation: survey sequencing and dense radiation-hybrid gene mapping. Nat Rev Genet 6, 643-648

130. Hu Z, Rohrer GA, Murtaugh MP, Stone RT, Beattie CW (1997) Mapping genes to swine X chromosome provides reference loci for comparative mapping. Mamm Genome 8, 608-610

131. Iannuzzi L, Di Meo GP, Perucatti A, Bardaro T (1998) ZOO-FISH and R-banding reveal extensive conservation of human chromosome regions in euchromatic regions of river buffalo chromosomes. Cytogenet Genome Res 82, 210-214

132. Iannuzzi L, Meo GPD, Perucatti A, Incarnato D (1999) Comparison of the human with the sheep genomes by use of human chromosome-specific painting probes. Mamm Genome 10, 719-723

133. Iannuzzi L, Di Meo GP, Perucatti A, Incarnato D, Schibler L, et al. (2000) Comparative FISH mapping of bovid X chromosomes reveals homologies and divergences between the subfamilies bovinae and caprinae. Cytogenet Cell Genet 89, 171-176

134. Ideraabdullah FY, de la Casa-Esperon E, Bell TA, Detwiler DA, Magnuson T, et al. (2004) Genetic and haplotype diversity among wild-derived mouse inbred strains. Genome Res 14, 1880-1887

135. International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on

vertebrate evolution. Nature 432, 695-716

136. International Chicken Polymorphism Map Consortium (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432, 717-722

137. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409, 860-921

138. Itoh T, Watanabe T, Ihara N, Mariani P, Beattie CW, et al. (2005) A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. Genomics 85, 413-424

139. Ivell R, Hartung S, Anand-Ivell R (2005) Insulin-like factor 3: where are we now? , Ann N Y Acad Sci 1041, 486-496

140. Jansen T, Forster P, Levine MA, Oelke H, Hurles M, et al. (2002) Mitochondrial DNA and the origins of the domestic horse. Proc Nat Acad Sci USA 99, 10905-10910

141. Jauch A, Wienberg J, Stanyon R, Arnold N, Tofanelli S, et al. (1992) Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting. Proc Nat Acad Sci USA 89, 8611-8615

142. Jiang Z, Priat C, Galibert F (1998) Traced orthologous amplified sequence tags (TOASTs) and mammalian comparative maps. Mamm Genome 9, 577-587

143. Jiang Z, He H, Hamasima N, Suzuki H, Verrinder G (2002) Comparative mapping of Homo sapiens chromosome 4 (HSA4) and *Sus scrofa* chromosome 8 (SSC8) using orthologous genes representing different cytogenetic bands as landmarks. Genome 45, 147-156

144. Johnson JM, Edwards S, Shoemaker D, Schadt EE (2005a) Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. Trends Genet 21, 93-102

145. Johnson PJ (2002) The equine metabolic syndrome: Peripheral Cushing's syndrome. Veterinary Clinics of North America: Equine Practice 18, 271-293

146. Johnson PJ, Scotty NC, Wiedmeyer C, Messer NT, Kreeger JM (2005b) Diabetes mellitus in a domesticated Spanish Mustang. J Am Vet Med Assoc 226, 584-588

147. Kapranov P, Drenkow J, Cheng J, Long J, Helt G, et al. (2005) Examples of the complex architecture of the human transcriptome revealed by RACE and high-

density tiling arrays. Genome Res 15, 987-997

148. Kerns JA, Olivier M, Lust G, Barsh GS (2003) Exclusion of melanocortin-1 receptor (Mc1r) and agouti as candidates for dominant black in dogs. J Hered 94, 75-79

149. Kiguwa SL, Hextall P, Smith AL, Critcher R, Swinburne J, et al. (2000) A horse whole-genome-radiation hybrid panel: Chromosome 1 and 10 preliminary maps. Mamm Genome 11, 803-805

150. Kijas JMH, Juneja RK, Gafvert S, Andersson L (2000) Detection of the causal mutation for canine leukocyte adhesion deficiency (CLAD) using pyrosequencing. Anim Genet 31, 326-328

151. Kijas JW, Menzies M, Ingham A (2006) Sequence diversity and rates of molecular evolution between sheep and cattle genes. Anim Genet 37, 171-174

152. Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, et al. (2006) Diversification of transcriptional modulation: Large-scale identification and characterization of putative alternative promoters of human genes. Genome Res. 16, 55-65

153. Kirkness EF, Bafna V, Halpern AL, Levy S, Remington K, et al. (2003) The dog genome: survey sequencing and comparative analysis. Science 301, 1898-1903

154. Korstanje R, O'Brien PC, Yang F, Rens W, Bosma AA, et al. (1999) Complete homology maps of the rabbit (Oryctolagus cuniculus) and human by reciprocal chromosome painting. Cytogenet Genome Res 86, 317-322

155. Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392, 917-920

156. Laan M, Kallioniemi OP, Hellsten E, Alitalo K, Peltonen L, et al. (1995) Mechanically stretched chromosomes as targets for high-resolution FISH mapping. Genome Res 5, 13-20

157. Lahbib-Mansais Y, Tosser-Klopp G, Leroux S, Cabau C, Karsenty E, et al. (2003) Contribution to high-resolution mapping in pigs with 101 type I markers and progress in comparative map between humans and pigs. Mamm Genome 14, 275-288

158. Lahbib-Mansais Y, Mompart F, Milan D, Leroux S, Faraut T, et al. (2006) Evolutionary breakpoints through a high-resolution comparative map between porcine chromosomes 2 and 16 and human chromosomes. Genomics (In Press,

Corrected Proof)

159. Lahn BT, Page DC (1999) Four evolutionary strata on the human X chromosome. Science 286, 964-967

160. Lawrence JB, Villnave CA, Singer RH (1988) Sensitive, high-resolution chromatin and chromosome mapping in situ: Presence and orientation of two closely integrated copies of EBV in a lymphoma line. Cell 52, 51-61

161. Le Merrer M, Legeai-Mallet L, Jeannin PM, Horsthemke B, Schlnzel A, et al. (1994) A gene for hereditary multiple exostoses maps to chromosome 19p. Hum. Mol. Genet. 3, 717-722

162. Lear TL, Trembicki KA, Ennis RB (1992) Identification of equine chromosomes in horse x mouse somatic cell hybrids. Cytogenet Cell Genet. 61, 58-60

163. Lear TL, Brandon R, Piumi F, Terry RR, Guérin G, et al. (2001) Mapping of 31 horse genes in BACs by FISH. Chromosome Res 9, 261-262

164. Lee E-J, Raudsepp T, Kata SR, Adelson D, Womack JE, et al. (2003a) A 1.4-Mb interval RH map of horse chromosome 17 provides detailed comparison with human and mouse homologues. Genomics 83, 203-215

165. Lee MK, Ren CW, Yan B, Cox B, Zhang HB, et al. (2003b) Construction and characterization of three BAC libraries for analysis of the chicken genome. Anim Genet 34, 151-152

166. Leipprandt JR, Kraemer SA, Haithcock BE, Chen H, Dyme JL, et al. (1996) Caprine [beta]-mannosidase: sequencing and characterization of the cDNA and identification of the molecular defect of caprine [beta]-mannosidosis. Genomics 37, 51-56

167. Leroux S, Dottax M, Bardes S, Vignoles F, Feve K, et al. (2005) Construction of a radiation hybrid map of chicken chromosome 2 and alignment to the chicken draft sequence. BMC Genomics 6, 12

168. Levine MA (1999) Investigating the origins of horse domestication. Equine Veterinary Journal 28, 6-14

169. Li T, O'Brien PCM, Biltueva L, Fu B, Wang J, et al. (2004) Evolution of genome organizations of squirrels (Sciuridae) revealed by cross-species chromosome painting. Chromosome Res 12, 317-335

170. Liao DJ, Du QQ, Yu BW, Grignon D, Sarkar FH (2003) Novel perspective: focusing on the X chromosome in reproductive cancers. Cancer Invest 21, 641-658

171. Lieto LD, Cothran EG (2001) Characterization of expressed sequence tags generated from skin cDNA clones of *Equus caballus* by single pass sequencing. Anim Biotechnol 12, 87-97

172. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803-819

173. Lindgren G, Sandberg K, Persson H, Marklund S, Breen M, et al. (1998) A primary male autosomal linkage map of the horse genome. Genome Research 8, 951-966

174. Liu GE, Matukumalli LK, Sonstegard TS, Shade LL, Van Tassell CP (2006) Genomic divergences among cattle, dog and human estimated from large-scale alignments of genomic sequences. BMC Genomics 7, 140

175. Locke MM, Ruth LS, Millon LV, Penedo MCT, Murray JD, et al. (2001) The cream dilution gene, responsible for the palomino and buckskin coat colours, maps to horse chromosome 21. Anim Genet 32, 340-343

176. Locke MM, Penedo MC, Bricker SJ, Millon LV, Murray JD (2002) Linkage of the grey coat colour locus to microsatellites on horse chromosome 25. Anim Genet 33, 329-337

177. Lunetta KL, Boehnke M, Lange K, Cox DR (1995) Experimental design and error detection for polyploid radiation hybrid mapping. Genome Res 5, 151-163

178. Lyon MF (2002) X-chromosome inactivation and human genetic disease. Acta Paediatr Suppl 91, 107-112

179. MacLeay JM, Sorum SA, Valberg SJ, Marsh WE, Sorum MD (1999a) Epidemiologic analysis of factors influencing exertional rhabdomyolysis in Thoroughbreds. Am J Vet Res 60, 1562-1566

180. MacLeay JM, Valberg SJ, Pagan JD, de laCorte F, Roberts J, et al. (1999b) Effect of diet on thoroughbred horses with recurrent exertional rhabdomyolysis performing a standardised exercise test. Equine Vet J Suppl 30, 458-462

181. MacLennan DH, Zorzato F, Fujii J, Otsu K, Phillips M, et al. (1989) Cloning and localization of the human calcium release channel (ryanodine receptor) gene to the proximal long arm (cen-q13.2) of human chromosome 19. Am J Hum Genet 45

(suppl), A205

182. Mäkinen A, Chowdhary B, Mahdy E, Andersson L, Gustavsson I (1989) Localization of the equine major histocompatibility complex (ELA) to chromosome 20 by in situ hybridization. Hereditas 110, 93-96

183. Mäkinen A, Hasegawa T, Makila M, Katila T (1999) Infertility in two mares with XY and XXX sex chromosomes. Equine Vet J 31, 346-349

184. Mäkinen A, Suojala L, Niini T, Katila T, Tozaki T, et al. (2001) X chromosome detection in an XO mare using a human X paint probe, and PCR detection of SRY and amelogenin genes in 3 XY mares. Equine Vet J 33, 527-530

185. Marcella KL (2005) Miniature medicine poses big distinctions. In DVM Newsmagazine, Sep 1 2005 issue. <http://www.dvmnewsmagazine.com/dvm /article/articleDetail.jsp?id=181036>

186. Margulies EH, Vinson JP, Program NCS, Miller W, Jaffe DB, et al. (2005) An initial strategy for the systematic identification of functional elements in the human genome by low-redundancy comparative sequencing. Proc Nat Acad Sci USA 102, 4795-4800

187. Mariat D, Oustry-Vaiman A, Cribiu EP, Raudsepp T, Chowdhary BP, et al. (2001) Isolation, characterization and FISH assignments of horse BAC clones containing type I and II markers. Cytogenet Cell Genet 92, 144-148

188. Mariat D, Taourit S, Guérin G (2003) A mutation in the MATP gene causes the cream coat colour in the horse. Genet Sel Evol 35, 119-133

189. Marklund L, Moller MJ, Sandberg K, Andersson L (1996) A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. Mamm Genome 7, 895-899

190. Marklund S, Ellegren H, Eriksson S, Sandberg K, Andersson L (1994) Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. Anim Genet 25, 19-23

191. Marra MA, Kucaba TA, Dietrich NL, Green ED, Brownstein B, et al. (1997) High throughput fingerprint analysis of large-insert clones. Genome Res 7, 1072-1084

192. Martins-Wess F, Rohrer G, Voß-Nemitz R, Drögemüller C, Brenig B, et al. (2003) Generation of a 5.5-Mb BAC/PAC contig of pig chromosome 6q1.2 and its integration with existing RH, genetic and comparative maps. Cytogenet Genome

Res 102, 116-120

193. Mathai CK, Ohno S, Beutler E (1966) Sex-linkage of the glucose-6-phosphate dehydrogenase gene in Equidae. Nature 210, 115-116

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

195. McCoard SA, Fahrenkrug SC, Alexander LJ, Freking BA, Rohrer GA, et al. (2002) An integrated comparative map of the porcine X chromosome. Anim Genet 33, 178-185

196. McEwen GK, Woolfe A, Goode D, Vavouri T, Callaway H, et al. (2006) Ancient duplicated conserved noncoding elements in vertebrates: A genomic and functional analysis. Genome Res 16, 451-465

197. Medina M (2005) Genomes, phylogeny, and evolutionary systems biology. Proc Nat Acad Sci USA 102, 6630-6635

198. Menotti-Raymond M, David VA, Lyons LA, Schäffer AA, Tomlin JF, et al. (1999) A genetic linkage map of microsatellites in the domestic cat (Felis catus). Genomics 57, 9-23

199. Menotti-Raymond M, David VA, Agarwala R, Schäffer AA, Stephens R, et al. (2003a) Radiation hybrid mapping of 304 novel microsatellites in the domestic cat genome. Cytogenet Genome Res 102, 272-276

200. Menotti-Raymond M, David VA, Chen ZQ, Menotti KA, Sun S, et al. (2003b) Second-generation integrated genetic linkage/radiation hybrid maps of the domestic cat (*Felis catus*). J Hered 94, 95-106

201. Metallinos DL, Bowling AT, Rine J (1998) A missense mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome - an equine version of Hirschsprung-disease. Mamm Genome 9, 426-431

202. Mickelson JR, Wu JT, Morrison LY, Swinburne JE, Binns MM, et al. (2003) Eighty-three previously unreported equine microsatellite loci. Anim Genet 34, 71-74

203. Mikko S, Brunberg E, Cothran G, Sandberg K, Watt B, et al. (2006) Linkage mapping of the silver coat color locus in the horse. In Proceedings of the 30th International Conference on Animal Genetics Abstracts

204. Milenkovic D, Oustry-Vaiman A, Lear TL, Billault A, Mariat D, et al. (2002) Cytogenetic localization of 136 genes in the horse: comparative mapping with the human genome. Mamm Genome 13, 524-534

205. Milenkovic D, Chaffaux S, Taourit S, Guérin G (2003) A mutation in the LAMC2 gene causes the Herlitz junctional epidermolysis bullosa (H-JEB) in two French draft horse breeds. Genet Sel Evol 35, 249-256

206. Moore SS, Byrne K, Johnson SE, Kata S, Womack JE (2001) Physical mapping of CSF2RA, ANT3 and STS on the pseudoautosomal region of bovine chromosome X. Anim Genet 32, 102-104

207. Morgan TH (1911) Random segregation versus coupling in Mendelian inheritance. Science 34, 384

208. Morgan JP (1968) Radiographic diagnosis of bone and joint diseases in the horse. Cornell Vet 58 (suppl), 28-47

209. Morisson M, Lemiere A, Bosc S, Galan M, Plisson-Petit F, et al. (2002) ChickRH6: a chicken whole-genome radiation hybrid panel. Genet Sel Evol 34, 521-533

210. Morisson M, Leroux S, Jiguet-Jiglaire C, Assaf S, Pitel F, et al. (2005) A gene-based radiation hybrid map of chicken microchromosome 14: comparison to human and alignment to the assembled chicken sequence. Genet Sel Evol 37, 229-251

211. Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520-562

212. Muller T, van de Sluis B, Zhernakova A, van Binsbergen E, Janecke AR, et al. (2003) The canine copper toxicosis gene MURR1 does not cause non-Wilsonian hepatic copper toxicosis. J Hepatol 38, 164-168

213. Murphy WJ, Sun S, Chen ZQ, Pecon-Slattery J, O'Brien SJ (1999a) Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. Genome Res 9, 1223-1230

214. Murphy WJ, Menotti-Raymond M, Lyons LA, Thompson MA, O'Brien SJ (1999b) Development of a feline whole genome radiation hybrid panel and comparative mapping of human chromosome 12 and 22 loci. Genomics 57, 1-8

215. Murphy WJ, Sun S, Chen Z-Q, Yuhki N, Hirschmann D, et al. (2000) A radiation hybrid map of the cat genome: implications for comparative mapping. Genome Res 5, 691-702

216. Murphy WJ, Larkin DM, Everts-van der Wind A, Bourque G, Tesler G, et al. (2005) Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. Science 309, 613-617

217. Murphy WJ, Davis B, David VA, Agarwala R, Schaffer AA, et al. (2006) A 1.5-Mb-resolution radiation hybrid map of the cat genome and comparative analysis with the canine and human genomes. Genomics (In Press, Corrected Proof)

218. Myka JL, Lear TL, Houck ML, Ryder OA, Bailey E (2003) FISH analysis comparing genome organization in the domestic horse (*Equus caballus*) to that of the Mongolian wild horse (*E. przewalskii*). Cytogenet Genome Res 102, 222-225

219. Nadeau JH, Sankoff D (1998) Counting on comparative maps. Trends Genet 14, 495-501

220. Naluai ÅT, Nilsson S, Samuelsson L, Gudjonsdottir AH, Ascher H, et al. (2000) The CTLA4/CD28 gene region on chromosome 2q33 confers susceptibility to celiac disease in a way possibly distinct from that of type 1 diabetes and other chronic inflammatory disorders. 56, 350-355

221. Nash WG, Wienberg J, Ferguson-Smith MA, Menninger JC, O'Brien SJ (1998) Comparative genomics: tracking chromosome evolution in the family ursidae using reciprocal chromosome painting. Cytogenet Genome Res 83, 182-192

222. Nash WG, Menninger JC, Wienberg J, Padilla-Nash HM, O'Brien SJ (2001) The pattern of phylogenomic evolution of the Canidae. Cytogenet Genome Res 95, 210-224

223. Naviaux B (1999) Miniature Horses: Their Care, Breeding, and Coat Colors. (Fort Bragg, CA: Raintree Publications)

224. Nef S, Parada LF (1999) Cryptorchidism in mice mutant for Insl3. Nat Genet 22, 295-299

225. Niewold TA, Kerstens HHD, van der Meulen J, Smits MA, Hulst MM (2005) Development of a porcine small intestinal cDNA micro-array: characterization and functional analysis of the response to enterotoxigenic *E. coli*. Vet Immunol Immunopathol 105, 317-329

226. Nilsson S, Helou K, Walentinsson A, Szpirer C, Nerman O, et al. (2001) Rat-mouse and rat-human comparative maps based on gene homology and high-resolution Zoo-FISH. Genomics 74, 287-298

227. Oakenfull E, Buckle V, Clegg J (1993) Localization of the horse (*Equus caballus*) alpha-globin gene complex to chromosome 13 by fluorescence *in situ* hybridization. Cytogenet Cell Genet 62, 136-138

228. O'Brien SJ, Menotti-Raymond M, Murphy WJ, Yuhki N (2002) The feline genome project. Annu Rev Genet 36, 657-86

229. Ohno S (1967) Sex Chromosomes and Sex-linked Genes. (Berlin: Springer)

230. Ojala M, Ala-Huikku J (1992) Inheritance of hydrocephalus in horses. Equine Vet J 24, 140-143

231. Olson MV, Dutchik JE, Graham MY, Brodeur GM, Helms C, et al. (1986) Random-clone strategy for genomic restriction mapping in yeast. Proc Nat Acad Sci USA 83, 7826-7830

232. Ostrander EA, Wayne RK (2005) The canine genome. Genome Res 15, 1706-1716

233. Pallavicini A, Zimbello R, Tiso N, Muraro T, Rampoldi L, et al. (1997) The preliminary transcript map of a human skeletal muscle. Hum Mol Genet 6, 1445-1450

234. Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, et al. (2004) Genetic structure of the purebred domestic dog. Science 304, 1160-1164

235. Pascual I, Dhar AK, Fan Y, Paradis MR, Arruga MV, et al. (2002) Isolation of expressed sequence tags from a Thoroughbred horse (*Equus caballus*) 5'-RACE cDNA library. Anim Genet 33, 231-232

236. Penedo MCT, Millon LV, Bernoco D, Bailey E, Binns M, et al. (2005) International equine gene mapping workshop report: a comprehensive linkage map constructed with data from new markers and by merging four mapping resources. Cytogenet Genome Res 111, 5-15

237. Perrocheau M, Boutreux V, Chadi-Taourit S, Di Meo GP, Perucatti A, et al. (2005) Equine FISH mapping of 36 genes known to locate on human chromosome ends. Cytogenet Genome Res 111, 46-50

238. Perrocheau M, Boutreux V, Chadi S, Mata X, Decaunes P, et al. (2006) Construction of a medium-density horse gene map. Anim Genet 37, 145-155

239. Perryman LE, McGuire TC, Banks KL (1983) Animal model of human disease. Infantile X-linked agammaglobulinemia. Agammaglobulinemia in horses. Am J Pathol 111, 125-127

240. Pinkel D, Straume T, Gray JW (1986) Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Nat Acad Sci USA 83, 2934-2938

241. Piumi F, Schibler L, Vaiman D, Oustry A, Cribiu EP (1998) Comparative cytogenetic mapping reveals chromosome rearrangements between the X chromosomes of two closely related mammalian species (cattle and goats). Cytogenet Cell Genet 81, 36-41

242. Posey KL, Hayes E, Haynes R, Hecht JT (2004) Role of TSP-5/COMP in Pseudoachondroplasia. Int J Biochem Cell Biol 36, 1005-1012

243. Potter JT, Evans, JW, Nolt, BH Jr. (1994) Therapeutic horseback riding. J Am Vet Med Assoc 204, 131-133

244. Power MM, Gustavsson I, Switonski M, Ploen L (1992) Synaptonemal complex analysis of an autosomal trisomy in the horse. Cytogenet Cell Genet 61, 202-207

245. Quilter CR, Blott SC, Mileham AJ, Affara NA, Sargent CA, et al. (2002) A mapping and evolutionary study of porcine sex chromosome genes. Mamm Genome 13, 588-594

246. Rappold GA (1993) The pseudoautosomal regions of the human sex chromosomes. Hum Genet 92, 315-324

247. Rattink AP, Faivre M, Jungerius BJ, Groenen MAM, Harlizius B (2001) A high-resolution comparative RH map of porcine Chromosome (SSC) 2. Mamm Genome 12, 366-370

248. Raudsepp T, Frönicke L, Scherthan H, Gustavsson I, Chowdhary BP (1996) Zoo-FISH delineates conserved chromosomal segments in horse and man. Chromosome Res 4, 218-225

249. Raudsepp T, Chowdhary, B.P. (2001) Correspondence of human chromosomes 9, 12, 15, 16, 19 and 20 with donkey chromosomes refines homology between horse and donkey karyotypes. Chromosome Res 9, 623-629

250. Raudsepp T, Kata SR, Piumi F, Swinburne J, Womack JE, et al. (2002) Conservation of gene order between horse and human X chromosomes as evidenced through radiation hybrid mapping. Genomics 79, 451-457

251. Raudsepp T, Lee E-J, Kata SR, Brinkmeyer C, Mickelson JR, et al. (2004a) Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. Proc Nat Acad Sci USA 101, 2386-2391

252. Raudsepp T, Santani A, Wallner B, Kata SR, Ren C, et al. (2004b) A detailed physical map of the horse Y chromosome. Proc Nat Acad Sci USA 101, 9321-9326

253. Ren C, Lee M-K, Yan B, Ding K, Cox B, et al. (2003) A BAC-based physical map of the chicken genome. Genome Res 13, 2754-2758

254. Rettenberger G, Klett C, Zechner U, Kunz J, Vogel W, et al. (1995) Visualization of the conservation of synteny between humans and pigs by heterologous chromosomal painting. Genomics 26, 372-378

255. Rexroad III CE, Owens EK, Johnson JS, Womack JE (2000) A 12,000 rad whole genome radiation hybrid panel for high resolution mapping in cattle: characterization of the centromeric end of chromosome 1. Anim Genet 31, 262-265

256. Richard F, Lombard M, Dutrillaux B (2000) Phylogenetic origin of human chromosomes 7, 16, and 19 and their homologs in placental mammals. Genome Res 10, 644-651

257. Richard F, Messaoudi C, Bonnet-Garnier A, Lombard M, Dutrillaux B (2003) Highly conserved chromosomes in an Asian squirrel (*Menetes berdmorei*, Rodentia: Sciuridae) as demonstrated by Zoo-FISH with human probes. Chromosome Res 11, 597-603

258. Rieder S, Taourit S, Mariat D, Langlois B, Guérin G (2001) Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). Mamm Genome 12, 450-455

259. Rink A, Santschi EM, Eyer KM, Roelofs B, Hess M, et al. (2002) A first-generation EST RH comparative map of the porcine and human genome. Mamm Genome 13, 578-587

260. Rink A, Eyer K, Roelofs B, Priest K, Sharkey-Brockmeier K, et al. (2006) Radiation hybrid map of the porcine genome comprising 2035 EST loci. Mamm

Genome 17, 878-885

261. Robinson TJ, Fu B, Ferguson-Smith MA, Yang F (2004) Cross-species chromosome painting in the golden mole and elephant-shrew: support for the mammalian clades Afrotheria and Afroinsectiphillia but not Afroinsectivora. Proc R Soc Lond B Biol Sci. 271, 1477-1484

262. Ronne M (1992) Putative fragile sites in the horse karyotype. Hereditas 117, 127-136

263. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, et al. (2005) The DNA sequence of the human X chromosome. Nature 434, 325-337

264. Rowe SJ, Windsor D, Haley CS, Burt DW, Hocking PM, et al. (2006) QTL analysis of body weight and conformation score in commercial broiler chickens using variance component and half-sib analyses. Anim Genet 37, 269-272

265. Rudolph JA, Spier SJ, Byrns G, Rojas CV, Bernoco D, et al. (1992) Periodic paralysis in Quarter Horses: a sodium channel mutation disseminated by selective breeding. Nat Genet 2, 144-147

266. Saetre P, Lindberg J, Leonard JA, Olsson K, Pettersson U, et al. (2004) From wild wolf to domestic dog: gene expression changes in the brain. Molecular Brain Research 126, 198-206

267. Saetre P, Strandberg E, Sundgren PE, Pettersson U, Jazin E, et al. (2006) The genetic contribution to canine personality. Genes, Brain and Behavior 5, 240-248

268. Safronova LD, Pimenova TI (1988) [Karyotypes of cattle (*Bos taurus*) and horses (*Equus caballus*) on the basis of synaptonemal complexes]. Genetika 24, 708-714

269. Sandberg K (1974) Linkage between the K blood group locus and the 6-PGD locus in horses. Anim Blood Groups Biochem Genet 5, 137-141

270. Santschi EM, Purdy AK, Valberg SJ, Vrotsos PD, Kaese H, et al. (1998) Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. Mamm Genome 9, 306-309

271. Santschi EM VP, Purdy AK, Mickelson JR (2001) Incidence of the endothelin receptor B mutation that causes lethal white foal syndrome in white-patterned horses. Am J Vet Res 62, 97-103

272. Sato S, Hasebe H, Sato S, Asahi Y, Hayashi T, et al. (2006) High-resolution physical mapping and construction of a porcine contig spanning the intramuscular

fat content QTL. Anim Genet 37, 113-120

273. Scherthan H, Cremer T, Arnason U, Weier H-U, Lima-de-Faria A, et al. (1994) Comparative chromosome painting discloses homologous segments in distantly related mammals. Nat Genet 6, 342-347

274. Schibler L, Vaiman D, Oustry A, Giraud-Delville C, Cribiu EP (1998) Comparative gene mapping: a fine-scale survey of chromosome rearrangements between ruminants and humans. Genome Res 8, 901-915

275. Schibler L, Cribiu EP, Oustry-Vaiman A, Furet J-P, Vaiman D (2000) Fine Mapping suggests that the goat polled intersex syndrome and the human blepharophimosis ptosis epicanthus syndrome map to a 100-kb homologous region. Genome Res 10, 311-318

276. Schibler L, Roig A, Mahe MF, Laurent P, Hayes H, et al. (2006) High-resolution comparative mapping among man, cattle and mouse suggests a role for repeat sequences in mammalian genome evolution. BMC Genomics 7, 194

277. Schmitz A, Oustry A, Vaiman D, Chaput B, Frelat G, et al. (1998) Comparative karyotype of pig and cattle using whole chromosome painting probes. Hereditas 128, 257-263

278. Shin EK, Perryman LE, Meek K (1997) A kinase-negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. J Immunol 158, 3565-3569

279. Shiue Y-L, Bickel LA, Caetano AR, Millon LV, Clark RS, et al. (1999) A synteny map of the horse genome comprised of 240 microsatellite and RAPD markers. Anim Genet 30, 1-9

280. Shiue Y-L, Millon LV, Skow LC, Honeycutt D, Murray JD, et al. (2000) Synteny and regional marker order assignment of 26 type I and microsatellite markers to the horse X-and Y-chromosomes. Chromosome Res 8, 45-55

281. Shubitowski DM, Venta PJ, Douglass CL, Zhou R-X, Ewart SL (2001) Polymorphism identification within 50 equine gene-specific sequence tagged sites. Anim Genet 32, 78-88

282. Si Z, Vandegraaff N, O'Huigin C, Song B, Yuan W, et al. (2006) Evolution of a cytoplasmic tripartite motif (TRIM) protein in cows that restricts retroviral infection. Proc Nat Acad Sci USA 103, 7454-7459

283. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423, 825-837

284. Slonim D, Kruglyak L, Stein L, Lander E (1997) Building human genome maps with radiation hybrids. J Comput Biol 4, 487-504

285. Smith J, Paton IR, Murray F, Crooijmans RPMA, Groenen MAM, et al. (2002) Comparative mapping of human Chromosome 19 with the chicken shows conserved synteny and gives an insight into chromosomal evolution. Mamm Genome 13, 310-315

286. Smith J, Speed D, Hocking PM, Talbot RT, Degen WG, et al. (2006) Development of a chicken 5 K microarray targeted towards immune function. BMC Genomics 7, 49

287. Smith SL, Everts RE, Tian XC, Du F, Sung L-Y, et al. (2005) Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. Proc Nat Acad Sci USA 102, 17582-17587

288. Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC V4.7. Genome Res 10, 1772-1787

289. Spirito F, Charlesworth A, Linder K, Ortonne J-P, Baird J, et al. (2002) Animal models for skin blistering conditions: absence of laminin 5 causes hereditary junctional mechanobullous disease in the Belgian horse. J Invest Dermatol 119, 684-691

290. Spriggs HF, Holmes NG, Breen MG, Deloukas PG, Langford CF, et al. (2003) Construction and integration of radiation-hybrid and cytogenetic maps of dog chromosome X. Mamm Genome 14, 214-221

291. Stanyon R, Stone G, Garcia M, Froenicke L (2003) Reciprocal chromosome painting shows that squirrels, unlike murid rodents, have a highly conserved genome organization. Genomics 82, 245-249

292. Stockham SL, Harvey JW, Kinden DA (1994) Equine glucose-6-phosphate dehydrogenase deficiency. Vet Pathol 31, 518-527

293. Stuehler B, Reichert J, Stremmel W, Schaefer M (2004) Analysis of the human homologue of the canine copper toxicosis gene MURR1 in Wilson disease patients. J Mol Med 82, 629-634

294. Sturtevant AH (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. J. Exp Zool 14, 43-59

295. Sturtevant AH (1917) Genetic factors affecting the strength of linkage in *Drosophila*. Proc Natl Acad Sci USA 3, 555-558

296. Su LC, Ravanshad S, Owen CA, Jr., McCall JT, Zollman PE, et al. (1982) A comparison of copper-loading disease in Bedlington terriers and Wilson's disease in humans. Am J Physiol Gastrointest Liver Physiol 243, G226-230

297. Suchyta SP, Sipkovsky S, Kruska R, Jeffers A, McNulty A, et al. (2003) Development and testing of a high-density cDNA microarray resource for cattle. Physiol Genomics 15, 158-164

298. Swinburne J, Gerstenberg C, Breen M, Aldridge V, Lockhart L, et al. (2000) First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. Genomics 66, 123-134

299. Swinburne JE, Hopkins A, Binns MM (2002) Assignment of the horse grey coat colour gene to ECA25 using whole genome scanning. Anim Genet 33, 338-342

300. Swinburne JE, Boursnell M, Hill G, Pettitt L, Allen T, et al. (2006) Single linkage group per chromosome genetic linkage map for the horse, based on two three-generation, full-sibling, crossbred horse reference families. Genomics 87, 1-29

301. Taira M, Taira M, Hashimoto N, Shimada F, Suzuki Y, et al. (1989) Human diabetes associated with a deletion of the tyrosine kinase domain of the insulin receptor. Science 245, 63-66

302. Tanaka M, Suzuki K, Morozumi T, Kobayashi E, Matsumoto T, et al. (2006) Genomic structure and gene order of swine chromosome 7q1.1→q1.2. Anim Genet 37, 10-16

303. Terry RB, Bailey E, Bernoco D, Cothran EG (2001) Linked markers exclude KIT as the gene responsible for appaloosa coat colour spotting patterns in horses. Anim Genet 32, 98-101

304. Terry RB, Bailey E, Lear T, Cothran EG (2002) Rejection of MITF and MGF as the genes responsible for appaloosa coat colour patterns in horses. Anim Genet 33, 82-84

305. Terry RB, Archer S, Brooks S, Bernoco D, Bailey E (2004) Assignment of the appaloosa coat colour gene (LP) to equine chromosome 1. Anim Genet 35, 134-7

306. The ENCODE Project Consortium (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. Science 306, 636-640

307. The Fantom Consortium et al. (2005) The transcriptional landscape of the mammalian genome. Science 309, 1559-1563

308. The International HapMap Consortium (2003) The International HapMap Project. Nature 426, 789-796

309. The International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437, 1299-1320

310. The International SNP Map Working Group (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409, 928-933

311. Thomas R, Fiegler H, Ostrander EA, Galibert F, Carter NP, et al. (2003) A canine cancer-gene microarray for CGH analysis of tumors. Cytogenet Genome Res 102, 254-260

312. Thomas R, Scott A, Langford CF, Fosmire SP, Jubala CM, et al. (2005) Construction of a 2-Mb resolution BAC microarray for CGH analysis of canine tumors. Genome Res 15, 1831-1837

313. Thompson JD, Higgins DJ, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673-4680

314. Thomson SAM, Kennerly E, Olby N, Mickelson JR, Hoffmann DE, et al. (2005) Microarray Analysis of Differentially Expressed Genes of Primary Tumors in the Canine Central Nervous System. Vet Pathol 42, 550-558

315. Tian Y, Nie W, Wang J, Ferguson-Smith MA, Yang F (2004) Chromosome evolution in bears: reconstructing phylogenetic relationships by cross-species chromosome painting. Chromosome Res 12, 55-63

316. Tozaki T, Takezaki N, Hasegawa T, Ishida N, Kurosawa M, et al. (2003) Microsatellite variation in Japanese and Asian horses and their phylogenetic relationship using a European horse outgroup. J Hered 94, 374-380

317. Tozaki T, Penedo MCT, Oliveira RP, Katz JP, Millon LV, et al. (2004) Isolation, characterization and chromosome assignment of 341 newly isolated equine *TKY*

microsatellite markers. Anim Genet 35, 487-496

318. Trask B, Pinkel D, van den Engh G (1989) The proximity of DNA sequences in interphase cell nuclei is correlated to genomic distance and permits ordering of cosmids spanning 250 kilobase pairs. Genomics 5, 710-717

319. Trifonov V, Yang F, Ferguson-Smith MA, Robinson TJ (2003) Cross-species chromosome painting in the Perissodactyla: delimitation of homologous regions in Burchell's zebra (*Equus burchellii*) and the white (*Ceratotherium simum*) and black rhinoceros (*Diceros bicornis*). Cytogenet Genome Res 103, 104-110

320. Trujillo JM, Walden B, O'Neil P, Anstall HB (1965) Sex-linkage of glucose-6-phosphate dehydrogenase in the horse and donkey. Science 148, 1603-1604

321. Vaiman D (2002) Fertility, sex determination, and the X chromosome. Cytogenet Genome Res 99, 224-228

322. Valberg SJ, Geyer C, Sorum SA, Cardinet GH, 3rd (1996) Familial basis of exertional rhabdomyolysis in quarter horse-related breeds. Am J Vet Res 57, 286-290

323. van Belzen MJ, Mulder CJ, Zhernakova A, Pearson PL, Houwen RH, et al. (2004) CTLA4 +49 A/G and CT60 polymorphisms in Dutch coeliac disease patients. Eur J Hum Genet 12, 782-785

324. van de Sluis B, Rothuizen J, Pearson PL, van Oost BA, Wijmenga C (2002) Identification of a new copper metabolism gene by positional cloning in a purebred dog population. Hum Mol Genet 11, 165-173

325. van Tuinen M, Hadly EA (2004) Calibration and error in placental molecular clocks: a conservative approach using the Cetartiodactyl fossil record. J Hered 95, 200-208

326. Venta PJ BJ, Yuzbasiyan-Gurkan V, Brewer GJ (1996) Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. Biochemical Genet 34, 321-341

327. Venter JC, Smith HO, Hood L (1996) A new strategy for genome sequencing. Nature 381, 364-366

328. Venter JC, Adams MD, Sutton GG, Kerlavage AR, Smith HO, et al. (1998) GENOMICS: Shotgun Sequencing of the Human Genome. Science 280, 1540-1542

329. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The Sequence of the Human Genome. Science 291, 1304-1351

330. Vialard F, Cocquet J, Christin-Maitre S, Veitia R, Fellous M (2002) The X chromosome and ovarian function. Cytogenet Genome Res 99, 218-223

331. Vila C, Leonard JA, Gotherstrom A, Marklund S, Sandberg K, et al. (2001) Widespread Origins of Domestic Horse Lineages. Science 291, 474-477

332. Volleth M, Klett C, Kollak A, Dixkens C, Winter Y, et al. (1999) Zoo-FISH analysis in a species of the order Chiroptera: *Glossophaga soricina* (Phyllostomidae). Chromosome Res 7, 57-64

333. Volleth M, Heller K-G, Pfeiffer RA, Hameister H (2002) A comparative Zoo-FISH analysis in bats elucidates the phylogenetic relationships between Megachiroptera and five microchiropteran families. Chromosome Res 10, 477-497

334. Voorrips RE (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. J Hered 93, 77-78

335. Wade CM, Kulbokas EJ, Kirby AW, Zody MC, Mullikin JC, et al. (2002) The mosaic structure of variation in the laboratory mouse genome. Nature 420, 574-578

336. Wagner ML, Goh G, Wu JT, Morrison LY, Alexander LJ, et al. (2004a) Sixty-seven new equine microsatellite loci assigned to the equine radiation hybrid map. Anim Genet 35, 484-486

337. Wagner ML, Goh G, Wu JT, Raudsepp T, Morrison LY, et al. (2004b) Radiation hybrid mapping of 63 previously unreported equine microsatellite loci. Anim Genet 35, 159-162

338. Wagner ML, Goh G, Wu JT, Raudsepp T, Morrison LY, et al. (2004c) Radiation hybrid mapping of 75 previously unreported equine microsatellite loci. Anim Genet 35, 68-71

339. Wagner ML, Raudsepp T, Goh G, Agarwala R, Schäffer AA, et al. (2006) A 1.3-Mb interval map of equine homologs of HSA2. Cytogenet Genome Res 112, 227-234

340. Wallis JW, Aerts J, Groenen MAM, Crooijmans RPMA, Layman D, et al. (2004) A physical map of the chicken genome. Nature 432, 761-764

341. Wang ZJ, Jeffs B, Ito M, Achermann JC, Yu RN, et al. (2001) Aromatase (Cyp19) expression is up-regulated by targeted disruption of Dax1. Proc Natl Acad Sci

USA 98, 7988-7993

342. Wang X, Wang X (2006) Systematic identification of microRNA functions by combining target prediction and expression profiling. Nuc Acids Res 34, 1646-1652

343. Ward TL, Valberg SJ, Lear TR, Guérin G, Milenkovic D, et al. (2003) Genetic mapping of GBE1 and its association with glycogen storage disease IV in American quarter horses. Cytogenet Genome Res 102, 201-206

344. Ward TL, Valberg SJ, Adelson DL, Abbey CA, Binns MM, et al. (2004) Glycogen branching enzyme (GBE1) mutation causing equine glycogen storage disease IV. Mamm Genome 15, 570-577

345. Warner JV, Nyholt DR, Busfield F, Epstein M, Burgess J, et al. (2006) Familial isolated hyperparathyroidism is linked to a 1.7 Mb region on chromosome 2p13.3-14. J Med Genet 43, e12

346. Wernersson R, Schierup MH, Jørgensen FG, Gorodkin J, Panitz F, et al. (2005) Pigs in sequence space: A 0.66X coverage pig genome survey based on shotgun sequencing. BMC Genomics 6, 70

347. Wilcox SA, Watson JM, Spencer JA, Graves JA (1996) Comparative mapping identifies the fusion point of an ancient mammalian X-autosomal rearrangement. Genomics 35, 66-70

348. Wiler R, Leber R, Moore BB, VanDyk LF, Perryman LE, et al. (1995) Equine severe combined immunodeficiency: a defect in V(D)J recombination and DNA-dependent protein kinase activity. Proc Natl Acad Sci USA 92, 11485-11489

349. Williams H, Richards CM, Konfortov BA, Miller JR, Tucker EM (1993) Synteny mapping in the horse using horse-mouse heterohybridomas. Anim Genet 24, 257-260

350. Womack JE, Johnson JS, Owens EK, Rexroad III CE, Schläpfer J, et al. (1997) A whole-genome radiation hybrid panel for bovine gene mapping. Mamm Genome 8, 854-856

351. Womack JE (2005) Advances in livestock genomics: Opening the barn door. Genome Res 15, 1699-1705

352. Wong GK, Liu B, Wang J, Zhang Y, Yang X, et al. (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432, 717-

722

353. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, et al. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature 434, 338-345

354. Xu Z, Sun S, Covaleda L, Ding K, Zhang A, et al. (2004) Genome physical mapping with large-insert bacterial clones by fingerprint analysis: methodologies, source clone genome coverage, and contig map quality. Genomics 84, 941-951

355. Xu Z, van den Berg MA, Scheuring C, Covaleda L, Lu H, et al. (2005) Genome physical mapping from large-insert clones by fingerprint analysis with capillary electrophoresis: a robust physical map of Penicillium chrysogenum. Nuc Acids Res 33, e50

356. Yalcin B, Fullerton J, Miller S, Keays DA, Brady S, et al. (2004) Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. Proc Nat Acad Sci USA 101, 9734-9739

357. Yang F, Muller S, Just R, Ferguson-Smith MA, Wienberg J (1997) Comparative chromosome painting in mammals: human and the Indian muntjac (*Muntiacus muntjak vaginalis*). Genomics 39, 396-401

358. Yang F, O'Brien PCM, Milne BS, Graphodatsky AS, Solanky N, et al. (1999) A complete comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps. Genomics 62, 189-202

359. Yang F, Alkalaeva EZ, Perelman PL, Pardini AT, Harrison WR, et al. (2003a) Reciprocal chromosome painting among human, aardvark, and elephant (superorder Afrotheria) reveals the likely eutherian ancestral karyotype. Proc Nat Acad Sci USA 100, 1062-1066

360. Yang F, Fu B, O'Brien PCM, Robinson TJ, Ryder OA, et al. (2003b) Karyotypic relationships of horses and zebras: results of cross-species chromosome painting. Cytogenet Genome Res 102, 235-243

361. Yang F, Fu B, O'Brien PCM, Nie W, Ryder O, et al. (2004) Refined genome-wide comparative map of the domestic horse, donkey and human based on cross-species chromosome painting: insight into the occasional fertility of mules. Chromosome Res 12, 65-76

362. Yang GC, Croaker D, Zhang AL, Manglick P, Cartmill T, et al. (1998) A dinucleotide mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome (LWFS); a horse variant of Hirschsprung disease. Hum Mol

Genet 7, 1047-1052

363. Yerle M, Pinton P, Robic A, Alfonso A, Palvadeau Y, et al. (1998) Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. Cytogenet Cell Genet 82, 182-188

364. Yerle M, Pinton P, Delcros C, Arnal N, Milan D, et al. (2002) Generation and characterization of a 12,000-rad radiation hybrid panel for fine mapping in pig. Cytogenet Cell Genet 97, 219-228

365. Yokota H, van den Engh G, Hearst JE, Sachs RK, Trask BJ (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. J Cell Biol 130, 1239-1249

366. Zhao S-H, Recknor J, Lunney JK, Nettleton D, Kuhar D, et al. (2005) Validation of a first-generation long-oligonucleotide microarray for transcriptional profiling in the pig. Genomics 86, 618-625

367. Zhao S, Malek J, Mahairas G, Fu L, Nierman W, et al. (2000) Human BAC ends quality assessment and sequence analyses. Genomics 63, 321-332

368. Zimmermann S, Steding G, Emmen JMA, Brinkmann AO, Nayernia K, et al. (1999) Targeted disruption of the Insl3 gene causes bilateral cryptorchidism. Mol Endocrinol 13, 681-691

APPENDIX

Table A4-1
Detailed information about individual equine HSA19 orthologs and microsatellites typed on the 5000rad equine whole genome RH panel and/or FISH mapped to HSA19-homologous regions of ECA7, ECA10, and ECA21

SYMBOL	NAME	FISH LOCATION	PCR PRIMERS 5' – 3'	REFERENCE	PRODUCT SIZE (BP)	T_A (C°), MgCl (mM)	ACCESSION NO.
AKAP8	A kinase (PRKA) anchor protein 8	-	F: ACCAGCGTTTGGACATGATG R: ATAGTCGTAGCTGTAGCTGG	This study	710	58, 1.5	BI774186
AURKC	aurora kinase C	-	F: TCGATGACTTTGAAATCGGG R: CAGTCCTTCCTTCTCTATCTG	This study	134	58, 1.5	CK463294
BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide	**ECA10p13**	F: GAGAACCAGCCCTTCCTCAT R: TTCTTGCGGGACTGCTTCCT	This study	436	60, 1.5	J03759
C19orf19	chromosome 19 open reading frame 19	-	F: GCCTACAGCCCAGAGAAAGT R: CTGCTGGGCTTGATGAAGAT	This study	164	58, 1.5	BF080172
C19orf2	chromosome 19 open reading frame 2	**ECA10p15**	F: AGAATTTTTGTCGCCCTCCT R: TTTTTCTGCTGCAATCTGGC	This study	153	60, 1.5	CK464292
C5R1	complement component 5 receptor 1 (C5a ligand)	-	F: GGCCACGCGTCGACTAGTAC R: CTGTTTGAGCGGATGAAGGTGAA	Pascual, I. et al. 2002	~130	58, 2.0	BI395175
CAPNS1	calpain, small subunit 1	-	F: TGAGGCCAATGAGAGTGAG R: CCATCAGTCTTCAGATCAGG	This study	408	58, 1.5	BP434626
CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma	-	F: GCCCATGGATCGAAACAGT R: AGCTGATTCACTCTCTGCAG	This study	123	58, 1.5	BI961735
CLPTM1	cleft lip and palate associated transmembrane protein 1	-	F: GGCCACGCGTCGACTAGTAC R: CTGTTTGAGCGGATGAAGGTGAA	Pascual, I. et al. 2002	~200	58, 2.0	BI395231
COMP	cartilage oligomeric matrix protein	-	F: TCGTGCAAACAATGAACAGC R: TCCACATGACCACGTAGAAG	This study	570	60, 1.5	AF325902
DPP9	dipeptidylpeptidase 9	-	F: TCACCACACCCGGCTTCT R: TGGCTGATGAACATGTCGAAGT	This study	741	60, 1.5	BM030481
EDG4	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 4	-	F: CTCTTCCTCATGTTCCACACA R: ACAGCCACCATGAGCAGGAA	This study	365	58, 1.5	CK461960
ELSPBP1	epididymal sperm binding protein 1	-	F: AGTATGGGGGAAATTCTTTC R: TCGGGCACCAGAGCTTGTT	This study	108	55, 1.5	AJ539176
ETHE1	ethylmalonic encephalopathy 1	-	F: ACTGTCTGATCTACCCTGCTCA R: TTGACAAACTCCTCACAGCTGA	This study	823	58, 1.5	BP434586
FKBP8	FK506 binding protein 8, 38kDa	-	F: AACTCCTACGACCTCGCCAT R: GCCTTGATGTTGTCAGGCTG	This study	642	60, 1.5	CB427408
FTL	ferritin, light polypeptide	-	F: CCATGAAAGCCGCCATTGT R: CCTCTGGATGTTGGTCAGATGG	monocyte cDNA library	353	58, 1.5	BI961949

Table A4-1 (continued)

SYMBOL	NAME	FISH LOCATION	PCR PRIMERS 5' – 3'	REFERENCE	PRODUCT SIZE (BP)	T_A (C°), MgCl (mM)	ACCESSION NO.
GDF1	growth differentiation factor 1	-	F: CTCAAGGTCCTGTATGCCA R: TGTACAGGAACCAGTAGAGG	This study	124	58, 1.5	BM286244
GMIP	Gem-interacting protein	-	F: GTTATCCGCTCGCTGAAGA R: ACAATGCCCAGGTTGTTGG	This study	460	58, 1.5	BE682381
GPI†	glucose phosphate isomerase	10p15	F: CTGGGACATCAACAGCTTTG R: GCTCCTGCTTGATGAAGTTG	Chowdhary et al. 2003	261	58, 1.5	X07382
GYS1	glycogen synthase 1 (muscle)	-	F: AAGGCTCGAATCCAGGAGTT R: GGTCTTGTCCAAGTTGAAGT	This study	413	58, 1.5	BX926004
HKR1	GLI-Kruppel family member HKR1	-	F: AATGTGGACGAGGCTTTAC R: TCCTTGCACACATAAGGCTTG	This study	170	60, 1.5	CK834051
ILT11B	immunoglobulin-like transcript 11 protein	-	F: CCTCACGACTGGAATTTGAC R: ATAGCATCTGAACGTCCACC	Takahashi et al. 2004	~200	58, 1.5	-
INSL3*	insulin-like 3 (Leydig cell)	**ECA21q13-q14**	F: GAACTGCTACAGTGGCTGGAAGG R: TCAGTGGGGACAGAGGGTCAG	Milenkovic et al. 2002	~200	58, 1.5	X73636.1
INSR	insulin receptor	**ECA7p16**	F: CTGTGGGACTGGAGCAAACA R: CCTCCATCTTGTGAATTTCC	This study	100	60, 1.5	BM255680
JUND	jun D proto-oncogene	-	F: ATCGACATGGACACTCAGGAG R: GCTGAGGACCTTCTGCTTGA	This study	210	58, 1.5	CA779483
KIR3DL	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail	-	F: TGCAATCAGGACAAAATGTG R: TCACAGGACCCAAGAAGAAG	Takahashi et al. 2004	~150	58, 1.5	-
KLF2	Kruppel-like factor 2 (lung)	-	F: AGAAGCCCTACCACTGCAAC R: CTACATGTGCCGCTTCATG	This study	173	58, 1.5	BI132469
LILRA	leukocyte immunoglobulin-like receptor, subfamily A	-	F: GCCAGTACAGATGCTACGGT R: CCAGATCTCAGAGCGACATT	Takahashi et al. 2004	~450	55, 1.5	-
LIPE	lipase, hormone-sensitive	-	F: TTCCTGCAGACCATCTCCAT R: CTGTATGATCCGCTCAAACTC	This study	266	55, 1.5	AJ000482
LSM4	LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae)	-	F: GGATGCCCGAGTGCTACATT R: GCCCTTCTGCTGCTTCTGCT	monocyte cDNA library	140	50, 1.5	-
MGC34079	hypothetical protein MGC34079	-	F: CCTTCAGGAACATCAAGACC R: CGTTCCAGTTCGTCCTTCT	monocyte cDNA library	196	58, 1.5	BI961498
MRPS12	mitochondrial ribosomal protein S12	-	F: AAGCCCAAGAAGCCCAACT R: ACGACGGTGAGCTTGACG	This study	173	60, 1.5	BI961235
NFIX	nuclear factor I/X (CCAAT-binding transcription factor)	**ECA7q11-q12**	F: ATCAAGCAGAAGTGGGCATC R: TCCCCGTCAGTACTTTCCAG	This study	245	60, 1.5	AW659814
NKG7	natural killer cell group 7 sequence	**ECA10p13dist-p12prox**	F: GTGAGCTTCCTGGTCCTGTC R: GAGAAGAACGTCTGGATCTG	This study	259	60, 1.5	BM734829
NUP62	nucleoporin 62kDa	-	F: TGGAGAAGGTGAAGCTGGAC R: CGAGTTCTGGTCGATCCACT	This study	341	58, 1.5	BM089312
OLFM2	olfactomedin 2	-	F: ACTTCATGGTGGACGAGAGC R: GAGTTGGTGACGTAGAGCAC	This study	185	TD60, 1.5	BF078883

Table A4-1 (continued)

SYMBOL	NAME	FISH LOCATION	PCR PRIMERS 5' – 3'	REFERENCE	PRODUCT SIZE (BP)	T_A (C°), MgCl (mM)	ACCESSION NO.
PDCD5	programmed cell death 5	-	F: CAAGGTTTAATAGAAATCCTCGAAA R: CGTCATCTTCATCAGAGTCCA	This study	426	60, 1.5	BX914933
PDE4A	phosphodiesterase 4A, cAMP-specific (phosphodiesterase E2 dunce homolog, Drosophila)	-	F: CTCATGTACAACGACGAGTC R: ACCCCTGAGCTGGTCACTTT	This study	426	58, 1.5	AV595552
PEG3	paternally expressed 3	**ECA10p12**	F: GCTCACTAAAGGTTGGGCTA R: GATCATGCATACTAGAGAGAACC	This study	245	58, 1.5	BU745796
PEN2	presenilin enhancer 2	**ECA10p14-p13**	F: TTGTCCCGGCATACACGGA R: TACCCAGGGGTATGGTGAA	This study	355	58, 1.5	BX921838
PIK3R2	phosphoinositide-3-kinase, regulatory subunit 2 (p85 beta)	-	F: AGGGAGAGTACACGCTGAC R: TTGGACACTGGGTAGAGGAG	This study	656	58, 1.5	AV663221
POLR2E	polymerase (RNA) II (DNA directed) polypeptide E, 25kDa	-	F: CAGACGCTGGAGGAGTTCAA R: CGGTGGGGTCGTCATTGT	This study	106	55, 1.5	CD471676
POLRMT	polymerase (RNA) mitochondrial (DNA directed)	**ECA7p16**	F: TGATGACCGTGGTGTACGG R: CAGAGAACATCTCCTGGAGG	This study	346	58, 1.5	CF722762
POP4	processing of precursor 4, ribonuclease P/MRP subunit (S. cerevisiae)	-	F: TACAGTCTTTTTCTCCCTCT R: GAAACAATAGCGCCGTGAAG	This study	313	60, 1.5	BP172676
POU2F2*	POU domain, class 2, transcription factor 2	**ECA10p14-p13**	F: TCTGGTGGAACTCTGCCTCT R: GCTTCCTCGCCCTCTTTC	This study	398	58, 1.5	BM734940
PPP5C	protein phosphatase 5, catalytic subunit	**ECA10p13**	F: AGTACACAGCCCAGATGTAC R: TCCGCTCGATCTTCCTGATG	This study	280	58, 1.5	BP161170
PRSS15*	protease, serine, 15	**ECA7p16**	F: ACTGCTTGAGCTGCTGGA R: TTGGCCGTGCAGATGAAC	This study	180	60, 2.0	CK947971
RFXANK	regulatory factor X-associated ankyrin-containing protein	-	F: TGTCCATCCACCAGCTTGCA R: GGCTTGTTGATGAGGTTGT	This study	300	58, 1.5	CK459113
RPL13A	ribosomal protein L13a	-	F: CATGAACACCAACCCATCC R: TGCCATCGAACACCTTGAG	monocyte cDNA library	343	58, 1.5	BI961884
RPL28	ribosomal protein L28	-	F: CTGATCAAGAGGAACAAGCA R: ATCTCCGCTTCATCACCACC	This study	285	60, 1.5	BM780607
RPS19	ribosomal protein S19	-	F: CCCAGCCATTTTAGCAGAGG R: TCCCTGAGGTGTCAGTTTGC	monocyte cDNA library	602	58, 1.5	BI961679
RPS5	ribosomal protein S5	-	F: TACATCGCGGTGAAGGAGAA R: GCTTCTTGCCATTGTTACGG	This study	148	58, 1.5	BI961498
SLC27A1	solute carrier family 27 (fatty acid transporter), member 1	**ECA21q13-q14**	F: TTCAACAGCCGCATCCTGC R: CTGTGGGCGATCTTCTTGCT	This study	305	60, 1.5	CK837240

Table A4-1 (continued)

SYMBOL	NAME	FISH LOCATION	PCR PRIMERS 5' – 3'	REFERENCE	PRODUCT SIZE (BP)	T_A (C°), MgCl (mM)	ACCESSION NO.
SLC7A10†	solute carrier family 7, (neutral amino acid transporter, y+ system) member 10	10p15	F: GGTCTTCAGCTTCATCTCGG R: ACATTGTAGGGAGGCATTCG	Chowdhary et al. 2003	274	58, 2.0	AF425263
SSTK*	serine/threonine protein kinase SSTK	**ECA21q13**	F: TCAAGCTTACCGACTTCGGC R: GACCATGACGTAGAGCACGA	This study	170	60, 1.5	BI344525
STRN4	striatin, calmodulin binding protein 4	-	F: GGCCACGCGTCGACTAGTAC R: CTGTTTGAGCGGATGAAGGTGAA	Pascual, I. et al. 2002	~700	55, 2.0	BI395151
TRIP10	thyroid hormone receptor interactor 10	-	F: TACCCATCATCGCCAAGTGT R: TTCATGGGCTGGCTGAAGT	This study	250	58, 1.5	BP455004
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	-	F: GCAGACATGCAGATCTTTGTG R: CTCCTTGTCTTGGATTTTAGC	This study	108	60, 1.5	CD536556
UHRF1	ubiquitin-like, containing PHD and RING finger domains, 1	-	F: GGGAACTCGTTCACGTACAC R: TGTTGGTGTTGGTGAGTTTC	This study	100	60, 2.0	BP169308
ZNF536 (KIAA0390)	zinc finger protein 536	-	F: CTTGTGTCCGATTGTCCTGT R: CACATTTTGCTTCAATCACC	This study	209	55, 1.5	AV593008
ZNF537 (KIAA1474)	zinc finger protein 537	-	F: CAACGACCAGCCCATAGACT R: TGGAAGGGGTGGAGGACTT	This study	245	58, 1.5	AW417891
ZNF543 (Q8NCX4)	zinc finger protein 543	-	F: TTCACCCACCGCTCCAATTT R: CTTCTCCCCAGTGTGAATCC	This study	390	60, 1.5	BP437102
COR095	microsatellite	**ECA7Q13PROX**	F: TACCTCTGGTGGTGATGCTT R: CCCACACTTACTCCCATCAC	TALLMADGE ET AL. 1999	206-216	58, 2.0	AF154948
HMS23	microsatellite	-	F: GATCCAATATTGTAAACCCCGCC R: CCTTCATAACCCTTATTGCAGCC	Godard et al. 1997	95	60, 1.5	U89810
TKY1034	microsatellite	-	F: GACCTGTTAGTGTAACTTCTG R: ACACCTGATGCAATGCTTTC	Tozaki et al. in preparation	158	58, 1.5	AB104252
TKY1107	microsatellite	**ECA7p15**	F: AGGGGGATGATTGAAAACAG R: TCATGACACTACCTTTCCTG	Tozaki et al. in preparation	197	58, 2.0	AB104325
TKY1112	microsatellite	-	F: GATCACAGTGCGATGACTTG R: AAAATTACCATGTGACCCATC	Tozaki et al. in preparation	149	58, 1.5	AB104330
TKY460	microsatellite	-	F: CAACTAGAAGGACCCACAAG R: AGACTTCTCCATCAGGCAC	Tozaki et al. in preparation	188	58, 1.5	AB103678
TKY461	microsatellite	**ECA7p13-p14**	F: ATGGCCCATCGTAAGAAACA R: GAGGGAGGAAGAAAGGAAGG	Tozaki et al. 2004	162-168	58, 1.5	AB103679
TKY512	microsatellite	-	F: CATTCACAGTTGCTGTGGAG R: GAATACCTACAGTACAGCGT	Tozaki et al. in preparation	158	58, 2.0	AB103730
TKY537	microsatellite	-	F: TCAGGGGTTCCTCTTCAGTG R: TTGCCTGGTGTCTAGGTTCC	Tozaki et al. 2004	147	58, 1.5	AB103755
TKY574	microsatellite	-	F: AGGAACATTGGCAAGAGATG R: TGCAGTCTCTGTATTTGGAT	Tozaki et al. in preparation	97	58, 1.5	AB103792
TKY678	microsatellite	-	F: TAAAAGAAGGGGGTAATGGG R: TGTGGTGCTTGTTCCAGCA	Tozaki et al. in preparation	207	58, 1.5	AB103896

Table A4-1 (continued)

SYMBOL	NAME	FISH LOCATION	PCR PRIMERS 5' – 3'	REFERENCE	PRODUCT SIZE (BP)	T_A (C°), MgCl (mM)	ACCESSION NO.
TKY773	microsatellite	-	F: GAGAACGGCCTTAACCGT R: CCTTGCAACCAAAGCCGTTC	Tozaki et al. in preparation	245	58, 1.5	AB103991
TKY806	microsatellite	-	F: TGGAACTGTGATGATGTTGC R: TCTTTCTTCCCTTCCGAGAG	Tozaki et al. 2004	180	58, 1.5	AB104024
TKY855	microsatellite	-	F: GATCTTGGCAGAAATGCCTG R: TCAAGCAAATAGGAGCTGAG	Tozaki et al. in preparation	142	58, 1.5	AB104073
TKY918	microsatellite	-	F: TGATGCCCAGGGAACAACT R: TGTTGTAGTACTTGTCCAGG	Tozaki et al. in preparation	274	58, 1.5	AB104136
TKY946	microsatellite	-	F: GATCCTAGGGCTTTGTGCG R: CATACCATGGAACACTGCTC	Tozaki et al. in preparation	110	58, 2.0	AB104164
TKY957	microsatellite	**ECA7p15-p14.3**	F: AGAAAGTCAGGGAGCAATC R: CCTTTGAGACCACTCCAAAC	Tozaki et al. in preparation	145	58, 2.0	AB104175
UMNe093	microsatellite	-	F: CGCAGAGTTGGAGACACCTG R: GGTGCCATCCACTGAAACAC	Wagner, M.L. et al. 2004c	167	58, 1.5	AY391287
UMNe359	microsatellite	-	F: GTGTGACGGAGGACGAGG R: TGGTGCCATCCACTGAAAC	Wagner, M.L. et al. 2004c	140	58, 1.5	AY391348
UMNe425	microsatellite	**ECA7p12-p11**	F: GAAAGAGCAAGGAGCCAAAG R: GCTTGCAAATGTTTGGGG	M. Wagner, pers. comm	~200	58; 1.5	-
UMNe506	microsatellite	-	F: AGGGACGTGACCTAACATGG R: CAGCACAGCCTCCTCTCC	Wagner, M.L. et al. 2004a	142	58, 1.5	AY735257
UMNe564	microsatellite	-	F: GAATACAGGGGCTTTTTCTGC R: TTCTGCATCTTGATTGCAGTG	Wagner, M.L. et al. 2004b	223	58, 1.5	AY464530
UMNe565	microsatellite	**ECA7q12-q13**	F: GAGAGCTAACCAGAATTGCCC R: CTATACCGCAACTCTCCTGGG	Wagner, M.L. et al. 2004a	333	58, 1.5	AY735270
UMNe606	microsatellite	-	F: TGCCTTTACCTTACAAACAC R: ATCCAAGAATATCATATTGGC	M. Wagner, pers. comm	~190	58, 1.5	-
VHL108	microsatellite	-	F: TTTCTTTCCCTCAGGCTGGA R: GACCTAAGAGAACCCCTTAAGT	van Haeringen et al. 1998	191	58; 1.5	Y08445

FISH locations shown in bold are from this study. TD- touchdown PCR

* indicates those markers that are on the FISH map only.

† indicates those markers were previously FISH mapped only, now also on the RH map.

Table A4-2. A summary of the number and type of markers mapped to regions of ECA7, ECA10 and ECA21 that were found to share homology with HSA19. Others = 1 HSA11 and 5 HSA5 orthologs.

Horse chromosome segment	RH mapped loci							FISH mapped loci			Comparative loci (RH+FISH)	
	On map		Type I			Type II						
	Total	New	Total	New	HSA19 loci	Total	New	Total (genes)	RH Anchors	New (genes)	Total	New
ECA7pter	*19*	*12*	*11*	*8*	*7*	*8*	*4*	*19 (13)*	*8*	*7 (4)*	*9*	*8*
ECA7qprox	*20*	*13*	*3*	*3*	*3*	*17*	*10*	*7 (3)*	*5*	*4 (1)*	*4*	*3*
ECA7 total	39	25	14	11	10	25	14	26 (16)	13	11 (5)	13	11
ECA21qprox	24	21	18	18	13	6	3	7 (4)	4	3 (all)	15	15
ECA10p	57	43	37	35	37	20	8	27 (14)	17	7 (all)	42	35
TOTAL	**120**	**89**	**69**	**64**	**60**	**51**	**25**	**60 (34)**	**34**	**21 (15)**	**70**	**61**

Table A4-3. Maps of ECA7, ECA10, and ECA21 in tabular form. The computation of the maps is explained in materials and methods. These maps include portions of the horse chromosomes that are not homologous to any part of HSA19. The data homologous to HSA19 (presented in Fig. 1) are separated from the remaining data by a thick horizontal line. The latter are italicized. Map coordinates are in cR. Because gene names have changed over time, the homolog of a gene in the table may appear on another mammalian map with a different name. Therefore, some pertinent gene aliases are shown in the rightmost column. ms = microsatellites; dc = discontinued; BAC-es = BAC end sequence.

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA7	C19orf19	0.00	19	0.4	FLJ40059
ECA7	POLR2E	18.50	19	1.0	
ECA7	POLRMT	28.98	19	0.5	
ECA7	TRIP10	89.09	19	6.6	
ECA7	INSR	99.51	19	7.0	
ECA7	UHRF1	110.22	19	4.8	
ECA7	DPP9	124.15	19	4.6	
ECA7	TKY957	212.21	ms		
ECA7	BIRC2	240.87	11	101.7	
ECA7	TKY1107	271.13	ms		

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA7	DRD2	320.33	11	112.8	
ECA7	TKY034	331.02	ms		
ECA7	TKY035	331.02	ms		
ECA7	TKY946	370.66	ms		
ECA7	APOC3	395.00	11	116.2	
ECA7	LEX015	395.00	ms		
ECA7	LEX038	395.00	ms		
ECA7	THY1	395.00	11	118.7	
ECA7	TKY461	419.47	ms		
ECA7	TKY005	0.00	ms		
ECA7	TKY283	2.80	ms		
ECA7	VHL108	15.09	ms		
ECA7	UMNe425	48.09	ms		
ECA7	NFIX	115.31	19	12.9	
ECA7	HTG33	126.20			
ECA7	TKY574	162.41	ms		
ECA7	TKY1112	184.46	ms		
ECA7	PDE4A	213.74	19	10.4	
ECA7	OLFM2	222.63	19	9.8	
ECA7	TKY272	226.70	ms		
ECA7	COR004	238.88	ms		
ECA7	UMNe093	246.39	ms		
ECA7	LEX045	246.39	ms		
ECA7	UMNe359	260.57	ms		
ECA7	TKY512	306.38	ms		

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA7	COR095	313.85	ms		
ECA7	TKY822	336.95	ms		
ECA7	UMNe565	356.35	ms		
ECA7	UMNe606	364.76	ms		
ECA7	*TKY282*	*437.92*	*ms*		
ECA7	*SGCV28*	*469.33*	*ms*		
ECA7	*UMNe459*	*500.73*	*ms*		
ECA7	*TKY986*	*550.36*	*ms*		
ECA7	*TUB*	*574.70*	*11*	*8.0*	
ECA7	*TKY624*	*598.68*	*ms*		
ECA7	*HBB*	*607.93*	*11*	*5.2*	
ECA7	*UMNe574*	*620.61*	*ms*		
ECA7	*UMNe074*	*623.64*	*ms*		
ECA7	*UMNe100*	*630.03*	*ms*		
ECA7	*XLKD-1*	*644.07*	*11*	*10.5*	*LYVE-1*
ECA7	*TKY1058*	*658.57*			
ECA7	*PTH*	*677.33*	*11*	*13.4*	
ECA7	*TKY793*	*716.80*	*ms*		
ECA7	*TKY839*	*752.26*	*ms*		
ECA7	*AHT019*	*762.51*	*ms*		
ECA7	*TKY1147*	*778.11*	*ms*		
ECA7	*TKY1177*	*782.73*	*ms*		
ECA7	*PAX6*	*818.23*	*11*	*31.7*	
ECA10	POP4	0.00	19	34.7	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA10	C19orf2	6.81	19	35.1	
ECA10	ZNF536	13.84	19	35.5	KIAA0390
ECA10	HMS23	17.39	ms		
ECA10	ZNF537	17.39	19	36.4	KIAA1474
ECA10	PDCD5	25.27	19	37.7	
ECA10	AAT10	38.74	ms		SLC7A10
ECA10	CEBPG	52.06	19	38.5	
ECA10	GPI	60.18	19	39.5	
ECA10	PEN2	68.31	19	40.9	
ECA10	CAPNS1	86.08	19	41.3	
ECA10	COR020	129.79	ms		
ECA10	UMNe506	170.65	ms		
ECA10	TKY855	188.49	ms		
ECA10	RYR1	194.41	19		
ECA10	HKR1	199.93	19	42.5	
ECA10	MRPS12	210.30	19	44.1	
ECA10	UCD482	239.25	ms		
ECA10	COR048	266.81	ms		
ECA10	BCKDHA	266.81	19	46.5	
ECA10	AHT015	300.21	ms		
ECA10	LIPE	316.79	19	47.5	
ECA10	RPS19	356.62	19	47.0	
ECA10	ETHE1	381.53	19	48.7	
ECA10	NV018	402.54	ms		
ECA10	CLPTM1	409.73	19	50.1	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA10	CKM	420.29	19	50.5	
ECA10	STRN4	435.38	19	51.9	ZIN
ECA10	PPP5C	435.38	19	51.5	
ECA10	SGCV30	449.50	ms		
ECA10	TKY773	449.50	ms		
ECA10	C5R1	449.50	19	52.5	
ECA10	ELSPBP1	453.92	19	53.1	
ECA10	GYS1	453.92	19	54.1	
ECA10	FTL	453.92	19	54.1	
ECA10	NUP62	466.65	19	55.1	
ECA10	RPL13A	474.82	19	54.6	
ECA10	COR015	486.75	ms		
ECA10	NKG7	498.61	19	56.5	
ECA10	LEX008	502.51	ms		
ECA10	TKY537	516.29	ms		
ECA10	SGCV20	568.76	ms		
ECA10	LILRA	575.23	19	59.7	LILR14
ECA10	ILT11B	582.09	19	59.4	ILT11-3
ECA10	KIR3DL	592.48	19	60.0	KIRRH
ECA10	TKY918	620.38	ms		
ECA10	TKY1034	636.67	ms		
ECA10	COR083	640.57	ms		
ECA10	RPL28	644.33	19	60.5	
ECA10	AURKC	678.65	19	62.4	
ECA10	UCD412	692.42	ms		

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA10	PEG3	695.84	19	62.0	
ECA10	ZNF543	695.84	19	62.5	Q8NCX4
ECA10	LEX062	712.97	ms		
ECA10	MGC34079	728.73	19	63.1	LOC147678
ECA10	RPS5	731.52	19	63.5	
ECA10	TKY460	756.99	ms		
ECA10	*LEX017*	*0.00*	*ms*		
ECA10	*TKY838*	*57.40*	*ms*		
ECA10	*PHIP*	*64.67*	*6*	*79.7*	
ECA10	*NV007*	*85.20*	*ms*		
ECA10	*ME1*	*93.79*	*6*	*83.9*	
ECA10	*CGA*	*111.69*	*6*	*87.8*	
ECA10	*CNR1*	*115.61*	*6*	*88.9*	
ECA10	*PNRC1*	*123.14*	*6*	*89.8*	
ECA10	*UMNe372*	*134.45*	*ms*		
ECA10	*UMNe209*	*146.19*	*ms*		
ECA10	*MAP3K7*	*158.41*	*6*	*91.2*	
ECA10	*CH241-101G24_T7*	*176.20*	*6*	*BAC-es*	
ECA10	*EPHA7*	*180.44*	*6*	*94.0*	
ECA10	*UMNe426*	*180.44*	*ms*		
ECA10	*CH241-102A14_SP6*	*184.87*	*6*	*BAC-es*	
ECA10	*TKY1101*	*208.39*	*ms*		
ECA10	*PRDM1*	*212.29*	*6*	*106.7*	
ECA10	*ASB09*	*220.42*	*ms*		
ECA10	*CH241-101L6_SP6*	*224.38*	*6*	*BAC-es*	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA10	*HMS02*	*237.42*	*ms*		
ECA10	*CH241-102D7_SP6*	*258.19*	*6*	*BAC-es*	
ECA10	*TKY614*	*258.19*	*ms*		
ECA10	*TKY1050*	*270.47*	*ms*		
ECA10	*CH241-101M4_T7*	*289.26*	*6*	*BAC-es*	
ECA10	*UMNe508*	*378.89*	*ms*		
ECA10	*AMD1*	*423.79*	*6*	*111.3*	
ECA10	*TKY471*	*427.90*	*ms*		
ECA10	*CD164a*	*432.01*	*6*	*109.8*	
ECA10	*SGCV17*	*432.01*	*ms*		
ECA10	*CH241-102D24_SP6*	*436.31*	*6*	*BAC-es*	
ECA10	*TKY999*	*440.60*	*ms*		
ECA10	*NR2E1*	*449.53*	*6*	*108.6*	
ECA10	*LEX009*	*471.93*	*ms*		
ECA10	*HDAC2*	*494.47*	*6*	*114.3*	
ECA10	*CH241-100C6_SP6*	*503.40*	*6*	*BAC-es*	
ECA10	*COL10A1*	*540.22*	*6*	*116.5*	
ECA10	*TKY496*	*557.31*	*ms*		
ECA10	*ROS1*	*571.50*	*6*	*117.7*	
ECA10	*MAN1A1*	*607.34*	*6*	*119.2*	
ECA10	*CH241-100E2_SP6*	*619.14*	*6*	*BAC-es*	
ECA10	*HSF2*	*626.67*	*6*	*122.7*	
ECA10	*CH241-101E16_T7*	*637.59*	*6*	*BAC-es*	
ECA10	*NV067*	*657.21*	*ms*		
ECA10	*HEY2*	*668.52*	*6*	*126.1*	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA10	*TKY503*	*679.09*	*ms*		
ECA10	*ARG1*	*711.23*	*6*	*131.9*	
ECA10	*SGK*	*731.26*	*6*	*134.5*	
ECA10	*TKY632*	*751.87*	*ms*		
ECA10	*PEX7*	*769.00*	*6*	*137.2*	
ECA10	*LOC340567*	*788.63*	*6*	*dc*	
ECA10	*AHT086*	*836.16*	*ms*		
ECA21	TKY021	0.00	ms		
ECA21	AKAP8	9.17	19	15.3	
ECA21	KLF2	24.20	19	16.2	
ECA21	SGCV14	24.20	ms		
ECA21	SGCV16	34.19	ms		
ECA21	SLC27A1	56.66	19	17.4	
ECA21	JUND	72.50	19	18.3	
ECA21	PIK3R2	72.50	19	18.1	
ECA21	LSM4	77.01	19	18.2	
ECA21	UBA52	133.59	19	18.5	
ECA21	FKBP8	178.42	19	18.5	
ECA21	COMP	192.59	19	18.7	
ECA21	GDF1	192.59	19	18.8	
ECA21	EDG4	192.59	19	19.5	
ECA21	GMIP	192.59	19	19.6	
ECA21	RFXANK	192.59	19	19.1	
ECA21	TKY678	192.59	ms		
ECA21	CCNB1	197.31	5	68.5	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA21	UMNe564	197.31	ms		
ECA21	ERBB2IP	213.16	5	65.2	
ECA21	LOC375449	213.16	5	65.9	
ECA21	TKY806	232.69	ms		
ECA21	FLJ13611	263.87	5	64.9	
ECA21	ADAMTS6	278.07	5	64.5	
ECA21	*KIF2*	*0.00*	*5*	*61.7*	
ECA21	*HESTG14*	*17.67*	*ms*		
ECA21	*XTP1*	*28.04*	*5*	*59.9*	
ECA21	*CKN1*	*38.47*	*5*	*60.2*	
ECA21	*UMNe229*	*48.32*	*ms*		
ECA21	*TKY677*	*53.06*	*ms*		
ECA21	*PDE4D*	*76.62*	*5*	*58.3*	
ECA21	*FLJ35954*	*97.76*	*5*	*56.3*	
ECA21	*HTG10*	*131.96*	*ms*		
ECA21	*CTLA3*	*146.20*	*5*	*54.4*	
ECA21	*DDX4*	*150.49*	*5*	*55.1*	
ECA21	*GZMA*	*154.41*	*5*	*54.5*	
ECA21	*FST*	*167.10*	*5*	*52.8*	
ECA21	*PELO*	*189.20*	*5*	*52.1*	
ECA21	*COR073*	*193.26*	*ms*		
ECA21	*LEX060*	*197.51*	*ms*		
ECA21	*FLJ21308*	*206.39*	*5*	*50.1*	
ECA21	*HCN1*	*215.28*	*5*	*45.3*	
ECA21	*OSRF*	*224.18*	*5*	*40.8*	

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA21	*COR068*	*238.20*	*ms*		
ECA21	*HMGCS1*	*247.10*	*ms*		
ECA21	*NNT*	*247.10*	*5*	*43.7*	
ECA21	*GHR*	*251.36*	*5*	*42.5*	
ECA21	*FBXO4*	*251.36*	*5*	*42.0*	
ECA21	*C6*	*255.78*	*5*	*41.2*	
ECA21	*FLJ39155*	*270.47*	*5*	*38.3*	
ECA21	*UMNe139*	*270.47*	*ms*		
ECA21	*UMNe206*	*280.27*	*ms*		
ECA21	*C9*	*284.93*	*5*	*39.3*	
ECA21	*DAB2*	*284.93*	*5*	*39.4*	
ECA21	*GDNF*	*294.74*	*5*	*37.9*	
ECA21	*UMNe327*	*314.30*	*ms*		
ECA21	*SKP2*	*354.15*	*5*	*36.2*	
ECA21	*IDN3*	*365.47*	*5*	*37.0*	
ECA21	*TKY824*	*373.01*	*ms*		
ECA21	*IL7R*	*380.80*	*5*	*35.9*	
ECA21	*PRLR*	*399.44*	*5*	*35.1*	
ECA21	*UMNe464*	*404.34*	*ms*		
ECA21	*BRIX*	*404.34*	*5*	*35.0*	
ECA21	*C1QTNF3*	*413.20*	*5*	*34.1*	
ECA21	*TARS*	*422.05*	*5*	*33.5*	
ECA21	*FLJ11193*	*426.29*	*5*	*31.6*	
ECA21	*UMNe509*	*426.29*	*ms*		
ECA21	*HTG32*	*426.29*	*ms*		

Table A4-3 (continued)

Horse Chromosome	Marker	Map Coordinate	Human Chromosome #	Human (Build 35.1)	Previous Alias
ECA21	*FLJ14054*	*426.29*	*5*	*32.8*	
ECA21	*TKY280*	*430.34*	*ms*		
ECA21	*NPR3*	*438.81*	*5*	*32.7*	
ECA21	*CDH9*	*522.31*	*5*	*26.9*	
ECA21	*LEX031*	*526.99*	*ms*		
ECA21	*CDH12*	*558.00*	*5*	*21.8*	
ECA21	*UMNe603*	*578.71*	*ms*		
ECA21	*MYO10*	*602.25*	*5*	*16.7*	
ECA21	*BASP1*	*607.44*	*5*	*17.3*	
ECA21	*TKY1018*	*607.44*	*ms*		
ECA21	*FBXL7*	*612.91*	*5*	*15.6*	
ECA21	*TRIO*	*622.71*	*5*	*14.2*	
ECA21	*DNAH5*	*627.62*	*5*	*13.7*	
ECA21	*LEX037*	*644.24*	*ms*		
ECA21	*LOC134146*	*655.33*	*5*	*dc*	
ECA21	*SEMA5A*	*671.86*	*5*	*9.1*	
ECA21	*AHT078*	*696.30*	*ms*		
ECA21	*SDHA*	*708.46*	*5*	*2.7*	
ECA21	*ADCY2*	*736.45*	*5*	*7.7*	
ECA21	*POLS*	*760.02*	*5*	*6.8*	

Table A5-3. Detailed information about PCR primers used to place and/or verify gene and microsatellite positions on the ECA21 contig.

ACCESSION NUMBER	ENSEMBL/HORSE EST ID	GENE	PRIMERS 5' → 3'	PRODUCT SIZE	PCR CONDITIONS	REFERENCE
NM_024527	Plate22-H19-M13R.ab1	ABHD8	F: CGCTTACTGGTGGAGAACCT R: GTCTGCCAGCTCCACCTC	105	55, 2.0	
NM_005858	ENSCAFT00000025526	AKAP8	F: ACCAGCGTTTGGACATGATG R: ATAGTCGTAGCTGTAGCTGG	710	58, 2.0	(Brinkmeyer-Langford et al 2005)
NM_032493	Plate32-D10-M13R.ab1	AP1M1	F: ATCACTCAGGAAGGCCACAA R: CCTCGTTCTTCCGGTACTTG	109	60, 1.5	
NM_020410	Plate9-K09-M13.rgy.ab1	ATP13A1	F: AAAGTCACTGGCCCTGGAG R: CCACACATGAGGGTCACGTA	266	60, 1.5	
NM_014299	CT02042A2E08.f1.ab1	BRD4	F: GAGATGCGTTTTGCCAAGAT R: CGCTCCTCCTCAGAGTCATC	170	58, 1.5	
NM_031966		CCNB1	F: CAAAATACCTACTGGGTCGG R: AATTTCTGGAGGGTACATTTCT	704	58, 1.5	(Goh et al 2006)
NM_001799	LeukoS5_2_C03.g1_A027	CDK7	F: CAACTTGTTGCTAGATGAAAATGG R: GGTCACAACCTGATGTGTATAAGC	103	60, 1.5	
NM_006387	ENSCAFT00000024839	CHERP	F: GGAGCTTCGAAATGTCATCG R: CGCCAGCTTGCACTTGTAGT	154	60, 1.5	
NM_153221	CT020018A20B09.ab1	CILP2	F: CTCCTCCGATGGCTTCTCTA R: CATCTCCCTGCGGATATCAT	151	60, 1.5	
NM_000095	ENSCAFT00000023228	COMP	F: GGACAAGGTGGTGGACAAGA R: TTCATGGTCTGCACGATCTC	353	TD60, 2.0	
NM_007263	MONO1_16_E08.g1_A005	COPE	F: TCAACCTCATTGTCCTGTCG R: CTTCAGCTGGGACAGGTAGC	507	62, 1.5	
NM_004831	CT020013B10B07.ab1	CRSP7	F: CCGGCAAAGTGAGTTCAGA R: GTGCTTCGGTTGCTGTGTTA	167	60, 1.5	
NM_006532	ENST00000262809	ELL	F: CTGATTTCATCGACCCCCTA R: CCCAGGTCATTGCTGACAT	232	60, 1.5	
NM_014077	SM0056-2_A06_02.ab1	FAM32A	F: AGGACAAGGACAAGGCGAAG R: CCGTTTCTCCTGCATCTTCT	128	60, 1.5	
NM_012181	ENST00000222308	FKBP8	F: CACGGAGACCGCCTTGTA R: AGCGATGACCACAGAGAGTG	562	60, 1.5	
NM_173483	OTTHUMT00000151074	FLJ39501	F: GCATCGGACAGAGCTTCG R: GCTCCACCTTGAGCCAGA	153	TD60, 2.0	
NM_016573	ENSCAFG00000014170	GMIP	F: TACGACGCCTTCATCTCTCT R: GTTGTTGGCAGACATCTTGTT	570	60, 1.5	
NM_015965	CT020003A10_PLATE_H08_64_075.ab1	GRIM19	F: GGAGAACCTGGAGGAGGAAG R: CCGTAGGTGGCATTGAGAAT	389	60, 1.5	
NM_032620	LeukoS1_6_B12.g1_A023	GTPBP3	F: CGTACAGCTGCAACTTCCTG R: ACTTCAGCCAGCTCCTTCCT	223	60, 1.5	
NM_005543	ENSCAFG00000015159	INSL3	F: CTTCCAGAGGGAGATCCAGA R: CTGGGCAGGTACTCCATCA	562	58, 1.5	(Brinkmeyer-Langford et al 2005)
NM_000215		JAK3	F: CTGCGGTTGGTCATGGAGTA R: GAAGTCGGCGATCTTGACGT	486	58, .15	(Brinkmeyer-Langford et al 2005)
NM_005354		JUND	F: ATCGACATGGACACTCAGGAG R: GCTGAGGACCTTCTGCTTGA	210	58, 1.5	(Brinkmeyer-Langford et al 2005)

Table A5-3 (continued)

ACCESSION NUMBER	ENSEMBL/HORSE EST ID	GENE	PRIMERS 5' → 3'	PRODUCT SIZE	PCR CONDITIONS	REFERENCE
NM_002248	ENSCAFT00000023962	KCNN1	F: AGCAGGAAGTGACCAGCAAC R: ATGCCAGTGAGCAGACACAC	127	60, 1.5	
XM_048457	CT020016B20F06.ab1	KIAA0892	F: TTTCTTTGCTTCTAGGGCATTT R: AGGGTGATCACTAAAAGCACTTTAC	100	55, 2.0	
NM_004831		KLF2	F: AGAAGCCCTACCACTGCAAC R: CTACATGTGCCGCTTCATG	173	58, 1.5	(Brinkmeyer-Langford et al 2005)
NM_138442	CT020014A20D10.ab1	LOC115098	F: CAGGGATGCCCAAGAAGTT R: GCTCCTTCCTCATGACGTGT	156	62, 1.5	
NM_012321	Plate17-F24-M13.rgy.ab1 653	LSM4	F: ACAAGTTCTGGCGGATGC R: CCTTCTGCTGCTTCTGCTG	150	58, 1.5	
NM_018467	ENST00000263897	MDS032	F: GGACCAGAACCTGGAGAAACT R: GCAGACGATAATGAGCATGG	103	60, 1.5	
NM_025021	ENST00000321949	MECT1	F: ATGGAGAACGCCATCAGC R: AGGATGATGTTGGGGATGC	158	60, 1.5	
NM_005919	ENST00000162023	MEF2B	F: CTGAGAGCGTCTACGTCCTG R: CTGCAGGCGTACTCCACAGT	1120	60, 1.5	
NM_024104	CT04003X1G10.f1.ab1	MGC2747	F: CTGGAAAACCGGTGACCTT R: GGAACCCTCCTGAAACTCCT	170	60, 1.5	
NM_004145		MYO9B	F: CTGTTCCTGCAGAGCTGGTT R: CAGGCGGATGATGCTCAGT	213	60, 1.5	(Brinkmeyer-Langford et al 2005)
NM_000435	ENSCAFT00000025565	NOTCH3	F: GTGGGCTCCTTTTCGTGCT R: CAGTGGAAGCCTCCATAACC	161	58, 1.5	
NM_017660	Plate19-A20-M13.rgy.ab1	p66alpha	F: CACACGTTCAGCCAGTCG R: CAGTTGGAGGTGGCGTTG	122	62, 1.5	
NM_024050	CT02041A1F05.f1.ab1	PCIA1	F: CGCCAAGAAGAGAGACCAG R: CGAGGAGCAGACAAGACAGG	150	60, 1.5	
NM_012088	Plate17-K03-M13.rgy.ab1 416	PGLS	F: GGGCCTCATCTCTGTCTTCTC R: CCACCAGTTGGGCTAGGG	62	58, 1.5	
NM_005027		PIK3R2	F: AGGGAGAGTACACGCTGAC R: TTGGACACTGGGTAGAGGAG	656	58, 1.5	(Brinkmeyer-Langford et al 2005)
NM_000980	Plate26-A08-M13R.ab1	RPL18A	F: ACTGCGGGTGAAGAACTTTG R: CTGCGATCTCCTCCACTTTC	879	60, 1.5	
NM_014884		SFRS14	F: CCCAACTTTTCCAGACTCTC R: ACTCTGGGTGTTTTCAAAGC	151	58, 1.5	(Brinkmeyer-Langford et al 2005)
U90593	(microsatellite)	SG14	F: CCCCAGTGGTTCCATTTAGATGT R: GGGGAGAGCATTTTGGTGA	188	58, 3.0	(Godard et al. 1997)
U90594	(microsatellite)	SG16	F: AATTCTCAAATGGTTCAGTGA R: CTCCCTCCCTTCCTTCTA	190	58, 2.0	(Godard et al. 1997)
NM_032627	CT020008A20D09.ab1	SSBP4	F: ACCAACTCCAGCGAGAACAT R: GATCCGTTCACGTGGTGAG	231	60, 1.5	
AB048331	(microsatellite)	TKY021	F: AGGTGAACCCCAGAGAGTCC R: AGTGAGGCCTCGGTTGGGAG	117-132	58, 2.0	(Kakoi et al. 1999; Swinburne et al. 2000)
AB103896	(microsatellite)	TKY678	F: TAAAAGAAGGGGGTAATGGG R: TGTGGTGCTTGTTCCAGCA	207	58, 1.5	(Tozaki et al. in preparation)

Table A5-3 (continued)

ACCESSION NUMBER	ENSEMBL/HORSE EST ID	GENE	PRIMERS 5' → 3'	PRODUCT SIZE	PCR CONDITIONS	REFERENCE
NM_024074	SM0051-2_E05_09.ab1 576	TMEM38A (MGC3169)	F: CTCTCACAGCTCCCCTTTTG R: AGCTCCTCCTTGGACTTGGT	162	58, 1.5	
NM_003290	ENSP00000300933	TPM4	F: TGAAAAGGAGGACAAATATGAAGA R: CCAGTTTTGCAACCGTTCTC	189	TD60, 2.0	
AY464530	(microsatellite)	UMNe564	F: GAATACAGGGGCTTTTTCTGC R: TTCTGCATCTTGATTGCAGTG	223	58, 1.5	(Wagner et al 2004b)

Table A6-1. Composite information about all new genes (ordered according to human sequence map) and microsatellite markers mapped to ECAX. Primer sequences with note "Secondary primers" indicate that the final primers for RH typing were designed after sequencing the initial equine PCR product. Mouse gene sequence positions written with blat indicate that map positions were derived indirectly through BLAT comparison of corresponding human gene sequence in Mouse Genome Browser Gateway, http://genome.ucsc.edu/cgi-bin/hgGateway. Information for all other markers on the RH map but not listed is available from ref. 1. Fluorescence *in situ* hybridization locations shown with bold font are from this study.

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
ASMT *XY*	acetylserotonin O-methyltransferase (homolog on Y)	Xp	Xp22.33	1.3	-	F: GTTCTCTTTGCTGCCTGTGA R: CAGCAGCTTCAGGGACACAC	150	1.5	60	
DXYS155E *XY*	B-lymphocyte antigen precursor (homolog on Y)	Xp	Xp22.33	1.3	-	F: GACATCCCCATGCTGGAC R: CTCGCCCTTGAACATGAG	171	1.5	62	
GYG2 *XY*	glycogenin 2	Xp	Xp22.33	2.3	-	F: GGTTATCCTCTCAAGGGTGT R: TGTCTGCATCCAGGAAGAC	170	1.5	54	
AdlicanX (DKFZp564I1922)	adlican	**Xp25+ Yqter**	Xp22.33	2.8	-	F: TTTGCCAATGGGACCCTG R: CGGACTGCATGAAGGAGTT	274	1.5	62	
NLGN4 *XY*	neuroligin 4 precursor	**Xp25**	Xp22.32	5.2	84.1 BLAT	F: GAGATCATGTCCCTGCAGAT R: TCATGGTGATGGTGTTTGGC	163	1.5	58	
STS *X*	steroid sulfatase	Xp25	Xp22.31	6.5	-	F: CCAATGGATGTTTTTCCAC R: ATGGGTTTCTTTCTCTTGG	111	1.5	58	horse gene AF133204; (Raudsepp et al. 2002)

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
KAL1	Kallmann syndrome 1 sequence	Xp	Xp22.31	7.9	-	F: CAGTCTGGACAGCTGGAGGT R: GTGGCGTCGTCTTCACTAGG	110	1.5	58-60	
TBL1X	transducin (beta)-like 1X-linked	Xp	Xp22.22	8.8	67.2 BLAT	F: CACTCGAAACCAATGGAA R: TTCCATATCCTGGCAGTCGA	500	1.5	58	
APXL *XY*	apical protein-like (Xenopus laevis)	**Xp25**	Xp22.22	9.1	140.9	F: AGGGCATCGTCTTCGACAT R: ACAGGCACTTCAGCTGCT	153	1.5	58	
MID1	midline 1 (Opitz/BBB syndrome)	Xp	Xp22.22	9.8	157.8	F: CCCTATTTGTCTGGAGCTC R: GATGATGTTCTGCAGGGTG	220	1.5	58	
AMELX	amelogenin (X chromosome, amelogenesis imperfecta 1)	Xp	Xp22.22	10.6	157.3 BLAT	F: CACCACCAAATCATTCCCG R: TTGGAGTCATGGAGTGTTGG	151	1.5	60	horse cDNA AB032193
MSL3L1	male-specific lethal 3-like 1 (Drosophila)	Xp	Xp22.22	11.1	156.7	F: CAATGCAGCCTTTTCAGC R: CTTAATCCATCCACCATCTC	~780	1.5	62	
TLR7	toll-like receptor 7	**Xp23-p22**	Xp22.22	12.2	155.3	F: GAGGAAAGGGACTGGTTA R: ACACTGCCAGAAGTACGGGT	300	1.5	62	
TLR8	toll-like receptor 8	Xp	Xp22.22	12.2	155.2	F: TAGCCAAGGTAAAAGGCTAC R: TGCATGAGGTTGTCGATGAT	211	1.5	54	
TMSB4X	thymosin, beta 4	**Xp23-p22**	Xp22.22	12.3	155.2	F: ATTCCACAAGCATTGCCTTC R: CCCACTTCTTCCTTCACCAA	210	1.5	58	
EGFL6	epidermal growth factor-like protein 6	Xp	Xp22.22	12.9	154.5	F: CATTTTTGAAGCAGAACGTG R: TCCACAGATAAAGGGCCATC	102	1.5	58	
GPM6B	glycoprotein M6B isoform 2	**Xp22**	Xp22.22	13.1	154.3	F: ATGCTGCATCAAGTGTCTGG R: CAAGGCATGGTCACTTGTGT	166	1.5	65	horse testis cDNA library BM414631 (Milenkovic et al. 2002)
GLRA2	glycine receptor, alpha 2	Xq17	Xp22.22	13.9	153.1	F: TCCTGGGTTAACTGATGG R: TGAAGTGGTTTGTCTCTAAG	296	1.5	58	
PIGA	phosphatidylinositol glycan, class A	Xp	Xp22.13	14.6	152.4	F: TGCCTGATTGAAAGAGGGCA R: CAGCATTAGGAATGACGGAC	454	1.5	63	
CTPS2	cytidine triphosphate synthase II	Xp	Xp22.13	15.9	150.9	F: GTCTTAAACGATGGTGGAGA R: CCTGGACAGCATCAGTAATG	~1200	2.0	56	
RBBP7	retinoblastoma binding protein 7	Xp	Xp22.13	16.2	150.8	F: TGCATATTTGGGAACAGCAA R: TACTGCTCCATGAACGCTTG	277	1.5	62	horse mesenteric lymph node cDNA library, BM780950

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
SCML2	sex comb on midleg-like 2	Xp	Xp22.13	17.6	149.1	F: GTTGGCATGAAGTTGGAAGC R: TTTTCACATGTCCCAACAGG	176	2.0	55	
PPEF1	protein phosphatase, EF hand calcium-binding domain 1	Xp	Xp22.13	18.0	148.6 BLAT	F: CTCATCTCCATGGAAGAATT R: GCTTCCATCTTTGTTTAAGTCC	122	1.5	58	
GPR64	HE6 receptor long splice variant	Xp	Xp22.13	18.3	148.4	F: TGTTCATTGTGGTCCTGGTT R: GTAAAAATGTAAGGCCAGCG	114	2.0	55	
PDHA1	pyruvate dehydrogenase (lipoamide) alpha 1	Xp	Xp11.12	18.7	148.1	F: AATGTGATCTTCACCGGCTG R: GCTTTTAACTCCATTCGGCG	115	1.5	58	
RPS6KA3	ribosomal protein S6 kinase, 90kDa, polypeptide 3	Xp	Xp22.12	19.5	147.2	F: GGATATTTGGTACCGTGGTG R: AAGAGACGAAAGCAGGAGCA	144	1.5	58	
						SECONDARY PRIMERS				
CNK2	connector enhancer of KSR2	Xp	Xp22.12	20.7	145.8	F: AGAAGGAGTGGCCATTATGA R: AACCAGGCAAGCAGACTCTT	105	1.5	60	
SMS	spermine synthase	**Xp22**	Xp22.12	21.3	145.4	F: GACAGATACTGGCCCACTGC R: GCCAAATCACTTTCCGCCAA	867	1.5	62	
ACATE2	mitochondrial Acyl-CoA Thioesterase	Xp	Xp22.11	23.0	143.2	F: CTTCATAGTTTCCTGGCTAAG R: CCTTACGGTGTTTTGAACAG	138	1.5	50	
PRDX4	peroxiredoxin 4	Xp	Xp22.11	23.0	143.3	F: GACAAGGGAATCCTAAGAC R: TTCTCCGTGTTTGTCAGTGT	111	2.0	55	
SAT HESTX-13	spermidine/spermine N1-acetyltransferase	Xp	Xp22.11	23.1	143.2	F: GGTAGCAGAATGGAATGAACCATC R: AGCAAGTACTCCTTGTCGATCTTG	111	2.0	65	horse monocyte cDNA library BI961507
EIF2s3	eukaryotic translation initiation factor 2, subunit 3 gamma	**Xp22-p21dist**	Xp22.11	23.4	83.9	F: CAGGTGCTTGGTGCAGTTG R: TTTGTCTCCTTCAGTACGTACACC	102	1.5	62	(Hu et al. 1997)
ARX	aristaless related homeobox	Xp	Xp22.11	24.3	83.0	F: GAGCAATCAGTACCAGGAG R: AACAGCCGCATTTTGCAC	128	1.5	60	
NR0B1	nuclear receptor subfamily 0, group B, member 1	Xp	Xp21.2	29.6	75.9	F: GGGCAGCATCCTCTACAACA R: CACCTGGAAGCAGGGTAAGT	201	1.5	62	horse cDNA AF167158
DMD	dystrophin	Xp16	Xp21.1	30.4	74.8	F: TCTGGAGTGAGTCTGTCAAA R: GAACCCAGTACCTGAAAACA	187	1.5	58	(Milenkovic et al. 2002)

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
SRPX	sushi-repeat-containing protein	Xp	Xp11.4	37.0	8.5	F: GGAGACCACAAGATCCAGTA R: GTCACTGGAGCACTTCATGT	838	1.5	58	
USP9X HESTX-7	ubiquitin specific protease 9	Xp	Xp11.4	40.0	11.7	F: GCATGGTAGCTCTCTTCAGCA R: GTTCATCTCCAAGCCATTCC	266	2.0	68	horse monocyte cDNA library, BI961890
DDX3	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3	Xp	Xp11.4	40.2	11.8	F: GAAGCTGATCGGATGTTGGA R: GGTCTCCACAAACACTAAGG	~600	1.5	58	
MAOA	monoamine oxidase A	Xp	Xp11.3	42.6	15.1	F: GGATCTGGTCAAGTAAGCGA R: GGCCTCCTTGTAATACATCA	~600	3.0	58	
FLJ22843	hypothetical protein FLJ22843	Xp	Xp11.3	43.0	15.6 BLAT	F: GAGGATCAGTTTTATACTGTGC R: CATGTAAGGATCTTCTGGAC	165	3.0	58	
TIMP1 HESTX-5	tissue inhibitor of metalloproteinase	Xp	Xp11.3	46.4	19.4	F: CCTCAGAAGTCAACCAGACC R: CTCCGACCTGTGGAAGTATC	345	1.5	63	horse monocyte cDNA library BI961880
JM5 [TUDPEc204]	JM5 protein	Xp	Xp11.23	47.9	6.2 BLAT DXImx38e	F: GGCCACGCGTCGACTAGTAC R: CTGTTTGAGCGGATGAAGGTGAA	107	3.0	55	horse lymphocyte cDNA library BI395200
AKAP4	major sperm fibrous sheath protein	**Xp13**	Xp11.22	48.8	5.6	F: GGTTTCCAACATGCACTGAG R: TCTCCTTGATTTCCTTACGG	346	3.0	58	
BMP15	bone morphogenetic protein 15 precursor	Xp	Xp11.22	49.5	4.9	F: CTCCACCCTTTCCAAGTCAG R: ATGGCATGATTGGGGGAATT	140	3.0	58	
NUDT11	nudix (nucleoside diphosphate linked moiety X)-type motif 11	**Xp13**	Xp11.22	50.1	4.8	F: CTGTTAGTGAGTAGCAGTCG R: CAATGCTAACCGAATCTTCC	241	1.5	58	
MAGED4	MAGED4 protein	Xp	Xp11.22	50.7	138.3 trophinin	F: AGGTACCCTCAAGGCACGTA R: AATGTTCGTCGTATTCTCGG	455	1.5	58	horse testis cDNA library BM414656

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
SMCX	Smcx homolog, X chromosome	Xp	Xp11.22	52.1	139.9	F: GCATCGCTTACCCCTATGAG R: TGTCTTCCTCTGTGGGTTCC	~1200	1.5	63	horse mRNA; U52363
UREB1 [TUDPEc224]	upstream regulatory element binding protein 1	Xp	Xp11.22	52.5	139.6	F: GGCCACGCGTCGACTAGTAC R: CTGTTTGAGCGGATGAAGGTGAA	195	1.5	62	horse cDNA, TUDPEc224, BI395210
MAGEH1	melanoma-associated antigen H1	**Xp13-p12**	Xp11.21	54.4	141.3 BLAT	F: TTATAAGCCGGTGCCCCGT R: GCACCTGAATTGAGGATAGC	198	1.5	58	
ASB12	ankyrin repeat and SOCS box-containing 12	**Xq11**	Xq11.2	62.3	85.1	F: CCTCATGGACATCACCAAGA R: GCAGCTAAGGTGGCCATAA	228	1.5	58	
HCA127	hepatocellular carcinoma-associated antigen 127	Xq	Xq11.2	63.0	85.3 BLAT	F: GAATGACCTAAACAAGCTGC R: ATAGGTCAGGGAGCCTCT	119	1.5	60	
MSN	moesin	Xq	Xq12	63.7	85.8	F: AGGATGTGCGGAAGGAAAG R: TCAGGCGGGCAGTAAATATC	157	1.5	50	
HEPH	hephaestin	Xq	Xq12	64.2	86.1	F: TGAAGATGCTGGGCATGCA R: CTTGAAGCTGTCATCCAGG	179	1.5	60	
XEDAR	X-linked ectodermal dysplasia receptor long isoform	Xq	Xq12	64.6	87.1 BLAT	F: TGGTTATGGAGAGGGTGGAG R: GGGATGCACTCTTGGTCCT	766	1.5	58	
AR	androgen receptor	Xq15-q16	Xq12	65.6	87.9	F: GGAGCCCGGAAGTTAAAGAA R: ACCGTCAGCTTCTGGGTTG	102	1.5	58	horse mRNA, AY032721; (Milenkovic et al. 2002)
PJA1	ubiquitin protein ligase Praja1	Xq	Xq13.1	67.2	89.3	F: GTCAAACCACCAAGAGGAG R: TCCTCCTCACTATCATCTG	202	1.5	58	
GJB1	gap junction protein, beta 1, 32kDa	Xq	Xq13.1	69.3	91.2	F: CTACGACCACTTTTTCCCCA R: CCCTGAGATGTGGACCTTGT	204	2.0	58	horse cDNA, AJ319909
TAF1	TATA box binding protein (TBP)-associated factor	Xq	Xq13.1	69.4	91.4 BLAT	F: TCTTCGAGGCACCTTTGGA R: ATGTGCTTTAGCAAGGGCTG	314	1.5	58	
PHKA1	phosphorylase kinase, alpha 1 (muscle)	**Xq14**	Xq13.1	70.6	92.4	F: GCAGATGATATCATGTTGGC R: ATGGCACAGATGCTGTGG	155	1.5	58	
ATP7A [MNK]	ATPase, Cu++ transporting, alpha polypeptide (Menkes syndrome)	Xq	Xq12-q13	75.9	95.9	F: CACAGCAAAGGAGTCCATCA R: CGGTTTCTTGAGTGAGAGGC	360	2.0	58	Horse cDNA AY011430

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
P2RY10	putative purinergic receptor P2Y10	Xq	Xq21.1	76.9	97.0	F: AGTGCTGCCATATGGGTC R: TATGATAGGGAGTGAAGCAG	328	1.5	58	
ITM2A	integral membrane protein 2A	**Xq14-q15**	Xq21	77.3	97.3	F: AATCATTGATGTGCCTGTCC R: CATCAGATAGCAGTTCCCCA	368	1.5	60	
SH3BGRL HESTX-6	SH3 domain binding glutamic acid-rich protein like	**Xq14-q15**	Xq21.1	79.2	99.0	F: CATTCATTGTCCAAGGCTGGT R: CACTGGAATAGAGGCATAGAGAAG	194	1.5	63	horse monocyte cDNA library, BI961868
POU3F4	POU domain, class 3, transcription factor 4	Xq	Xq21.1	81.5	100.8	F: ATTACTTGCAGGGAGTTCC R: CCAGCCACTTGTTCAGCA	648	2.0	58	
ZNF6	zinc finger protein 6	Xq	Xq21.1	83.2	102.3 BLAT	F: GAGCATATGGGGAACACA R: CATCTTCTCCCGCTGCAT	~1200	1.5	60	
CHM	choroideremia (Rab escort protein 1)	**Xq15-q16**	Xq21.1	83.8	102.8	F: ATTAACCCCCAACCTCCAAT R: ACTGAGGGAGTTCTCCTTG	173	1.5	58	
KLHL4	kelch-like 4 (Drosophila)	Xq	Xq21.31	85.5	104.2 BLAT	F: GGCTACATATTGGCACCATG R: TTGACATGGGAGGCATCAC	165	1.5	58	
PABPC5	poly(A) binding protein, cytoplasmic 5	Xq	Xq21.31	89.4	109.6	F: GGAAAACCATTCCGCCTTAT R: TTGAGCCGCACTCCATTCA	172	1.5	55	
NAP1L3	nucleosome assembly protein 1-like 3	Xq	Xq21.32	91.6	112.0	F: CCAGATCACAATGATCCCTT R: TGATGCATTAGGAACCACTC	169	1.5	TD65	
DIAPH2	diaphanous homolog 2 (Drosophila)	**Xq17-q21prox**	Xq21.33	94.7	118.1	F: GCCTGTGAAGAACTGAAGAA R: CCCAAAGACTGGGCATTTC	107	1.5	60	
CHM1L [TNMD]	tenomodulin	Xq	Xq22.1	98.6	122.3	F: TCACTTCCCTACCAACGACA R: GGCATGATGACACGACAGAT	550	1.5	65	horse cDNA AB059407
GLA	galactosidase, alpha	Xq	Xq22.1	99.4	123.0	F: AATGACCTCCGACACATCAG R: GAAGCAAAACAGTGCCTGTG	~700	1.5	60	
RPL36A HESTX-10	ribosomal protein L36a	Xq	Xq22.1	99.4	-	F: CACAATACAAGAAGGGCAAGG R: CACCATAGCCACTCTGCTTC	269	1.5	65	horse monocyte cDNA library BI961599
TMSNB	thymosin beta	Xq	Xq22.1	100.5	-	F: AGCAGGAGAAAGAGTGTGT R: AGGTGAGAAGACATCAGAAGA	153	1.5	60	
KIAA0443	hypothetical protein KIAA0443	**Xq21 prox**	Xq22.1	100.6	124.3	F: GGTCAGTCAGGGAAATTCG R: AACTCGGGAGGAGGGATAAT	278	1.5	60	

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
LOC170242 (CG7943)	similar to hypothetical protein CG7943	Xq	Xq22.2	101.9	-	F: GGAAAGAAAAGCTGGCAC R: GAAGTATTGAGGGCCTTCGT	168	1.5	55	
SERPINA7	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase	Xq	Xq22.3	104.0	-	F: TCAACCTGTACCGGAGGTTC R: TTCCAGCTCCTTCTTTGGAA	230	2.0	55	
CLDN2	claudin 2	**Xq21dist**	Xq22.3	104.9	128.1	F: ATCACCCAGTGTGACATCTA R: CAACAGGAATGAAGCCCAG	243			
MID2	midline 2	Xq	Xq22.3	105.8	129.1	F: TTGTGCTCACAGCCTCTGC R: CAGAACTGGCAAGCAATTCG	294	1.5	53	
AUTL2	putative autophagy-related cysteine endopeptidase 2	Xq	Xq22.3	106.1	129.3	F: GACCTGGAGGAAGATTTCG R: GTATGCTGTTGGCATATCAG	147	1.5	60	
IRS4	insulin receptor substrate 4	**Xq23**	Xq23	106.7	129.9	F: AAAACCACAAAAGCCCACAC R: AAATGGCACTCCGAACTCAA	196	1.5	60	
PAK3	p21 activated kinase 3	**Xq24-q25**	Xq21.3-q24	109.1	131.7	F: TACTCCCTCGGATTATGTAATTTC R: GGAGGTTCGAATGCAAGAGGA	238	2.0	58	Fetal cDNA library EST#841; G62153; (Chowdhary et al. 2003)
DCX	doublecortex; lissencephaly, X-linked (doublecortin)	Xq	Xq23	109.3	132.1	F: CAAACAGGGCAGCAATGTC R: TTTGAAGGAGGTTCGAATGC	198	1.5	60	
AMOT	angiomotin	Xq	Xq23	110.7	133.7	F: CTGGTAGACCCCCTGATTTA R: CAGTGGACAGGCAGGATACC	153	2.0	58	
HTR2C	5-hydroxytryptamine (serotonin) receptor 2C	Xq	Xq23	112.5	135.4	F: ATGTCTGGCCACTACCTAGA R: GAATTGAAACGGCTATGCTC	148	1.5	58	
PLS3	plastin 3 (T isoform)	Xq	Xq23	113.6	65.5	F: TGTCAGGGCACACCTGAAAT R: CACTGGGTGGAAAGCAAAAT	198	1.5	62	Horse Lambda Zap Express library AW260928

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
AGTR2	angiotensin II receptor, type 2	Xq	Xq23	114.0	20.0	F: TCTTCCTCTATGGGCAACCT R: TGCCAGGGATTTCTTCTTTG	189	1.5	60	
SLC6A14	solute carrier family 6 (neurotransmitter transporter), member 14	Xq	Xq23	114.3	20.2	F: CTGATTGACCACTTCTGTGC R: CCACCATAGCCAGAATATCCA	450	1.5	60	
BKLHD2 [KIAA1309, FLJ10262]	BTB and kelch domain containing 2	Xq	Xq24	115.7	21.8	F: GAGTCATGATGGCCTCTGC R: GCAGCTTCCAGTGTGTCTTG	300	1.5	63	
RNF127 (FLJ22612)	ring finger protein 127 (Hypothetical protein FLJ22612)	Xq	Xq24	116.8	27.6	F: GAATATGGCTGCATCCTAGAG R: TGTGTTGTAACCATCCCGC	126	1.5	60	
PGRMC1	progesterone receptor membrane component 1	Xq	Xq24	117.1	27.9	F: GGGCAAACTGCTGAAGGA R: CTTTCTGGATTTGTGACACAC	293	1.5	60	
LAMP2 HESTX-9	lysosomal-associated membrane protein 2	**Xq25-q26**	Xq24	118.3	29.5	F: CGGTTCTATCTGAAGGAGGTG R: CCTTCACATTGAAAGGCTGA	~600	1.5	58	horse monocyte cDNA library; BI961803
GRIA3	glutamate receptor 3 isoform flip precursor	Xq	Xq25	121.0	32.7	F: CCCTTGGCTTATGAAATCTG R: TTCATTTGGAGGATCCGGAG	159	2.0	58	
BIRC4 [XIAP]	baculoviral IAP repeat-containing 4	Xq	Xq25	121.7	33.4	F: TCAGAACACAGGCGACACTT R: ATCCGTGCTTCGTAATCTGC	155	1.5	63	
ODZ1	odz, odd Oz/ten-m homolog 1(Drosophila)	**Xq26-q27**	Xq25	122.2	33.8	F: CAGCAACGGAGTCCTGAAAA R: GGTTGTCCAGCCATTTTCGT	115	1.5	58	
KIAA1892L (LOC139170)	KIAA1892-like	Xq	Xq25	124.3	36.1 BLAT	F: ACCTGTGGGTGAACTACTTC R: GAGGCCTGAAGGGAGAGG	119	1.5	TD65	
ZDHHC9	zinc finger, DHHC domain containing 9	**Xq26-q27**	Xq26.1	127.6	39.4 BLAT	F: GGACTTTTGATGATGTTTGACC R: TTCTTTCTCACCACCATCAC	150	1.5	58	
ARHGEF6	Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6	**Xq27**	Xq26.3	134.4	46.8	F: AGCATGTTTTTGGGACTTGG R: GGCAAACAGGAACCCATTTA	190	1.5	62	horse monocyte cDNA library BM734718

Table A6-1 (continued)

Symbol GENES	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
TNFSF5	tumor necrosis factor (ligand) superfamily, member 5	Xq	Xq26.2	134.4	46.8	F: CAGCTGGCCGTTAAAAGACA R: CCTCCCAAGTGAATGGATTG	203	1.5	60	
RBMX	RNA binding motif protein, X chromosome	Xq	Xq26.3	134.6	47.0	F: TTCCCTAAAGGCCACTTCCT R: CATAACCCGACCACCCTCTA	153	1.5	53	horse monocyte cDNA library BI961399
ZIC3	Zic family member 3 heterotaxy 1	Xq	Xq26.3	135.3	47.6	F: CAGATGGTTGTTTAACTGCG R: CAGGGATGATTAAATCCAGC	145	1.5	63	
FGF13	fibroblast growth factor 13	Xq	Xq26.3	136.6	48.7	F: CTGTGGAACTGTGATGTTG R: TGTTGTTTAGGGGTAACCAGTC	141	1.5	58	
LDOC1	leucine zipper, down-regulated in cancer 1	**Xq28**	Xq27.2	138.9	random 44.3	F: AGAACAGCCAGCTCATGGAA R: GCCGTCTGCACGATAAACTC	151	1.5	58	
DKFZp547-M2010	hypothetical protein	Xq	Xq27.3	141.4	-	F: GCTAGACTGAACCTGAGGAA R: CCACAGCTTTAATGCCACCA	263	1.5	60	
CXorf2 [KIAA1854]	chromosome X open reading frame 2	Xq	Xq27.3	143.5	-	F: AGGGAGGGGCAGATGCTT R: CCTCCTGGAGAACCCCTCT	151	1.5	TD65	horse sequence containing TKY754 AB103972
FMR1	fragile X mental retardation 1	**Xq28**	Xq27.3	145.6	58.2	F: GTACCTGGGGTCACTGCTA R: AGCCTCAATTCTCACCCTCA	290	1.5	58	
FMR2 [OX19]	fragile X mental retardation 2	Xq	Xq28	146.2	59.2	F: CCCCAGTGTCTCTCAACAAC R: CCCCGTAACACGTTGCTAGT	133	1.5	58	
FAM11A	Family with sequence similarity 11 member A	Xq	Xq28	147.3	60.0 BLAT	F: TGCTGCTGTTCTCTGTGTTG R: ATATTGGGGATTTCGTGCC	185	1.5	58	
MTM1	myotubular myopathy 1	**Xq29**	Xq28	148.4	60.8	F: GGTCTACCAAATAAGACATCTCAAG R: AGCGTTTACAAAAAGCCAGA	100	1.5	58	
CD99L2	CD99 antigen-like 2	Xq	Xq28	148.5	61.0 BLAT Mic2l1	F: GCAGAGACTGGCACCATC R: CTGTATGCTGAAGCAGAACT	111	1.5	58	
OPN1MW	opsin 1 (cone pigments), medium-wave-sensitive	Xq	Xq28	151.9	63.7	F: GACCTGGCAGAGACCATCAT R: CAGCTTGGCATCAAATCTCA	~700	1.5	62	horse mRNA AF132043

Table A6-1 (continued)

Symbol	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
GENES										
RPL10 HESTX-8	ribosomal protein L10	Xq	Xq28	152.0	63.8	F: GTACATGGTGAAAAGCTGTGG R: GGGCTTCAATCACATGCTC	755	1.5	63	horse monocyte cDNA library BI961947
GDI1	GDP dissociation inhibitor 1	Xq	Xq28	152.1	63.9	F: CTACAGTGGGGCGTAAGAGC R: GTTCCCAGCATCTCTGCTTC	259	1.5	60	horse monocyte cDNA library BM735209
CLIC2 [HESTX-11]	chloride intracellular channel 2	**Xq29**	Xq28	152.9	-	F: CAGAGAGTGAGCATATCAGAGAGG R: CTAGGTGTTGAATCCCATGCTAAG	108	2.0	63	horse monocyte cDNA library BI961616
SPRY3	sprouty homolog 3 (Drosophila)	Xq	Xq28	153.4	-	F: AGTGAGCACCTCTTCATC R: CCCTTAACACAGCAGAGACA	158	2.0	55	
SYBL1	synaptobrevin-like 1	Xq	Xq28	153.5	random 75.1	F: CGCGAGCCTTTAATTTTCTG R: TGCAGCCAAGACACTTGAGA	119	3.0	56	
Microsatellites										
UM035	microsatellite	Xq	-	-	-	F: GTGATGGATGACATGAGG R: GCATTTAAAACACTAGAACAC	200	2.0	58	AF195581
UMNe091	microsatellite	Xp	-	-	-	F: GCAACAATAAGATACCCAAAGCAG R: GATCTTGGCCAACAACTCGTC	153	1.5	58	AY391286
UMNe148	microsatellite	Xp	-	-	-	F: GATCAAACACTAGAATGTTCACAC R: CAGCTGTGAGGCAGAGACTG	110	1.5	65	AF536262
UMNe202	microsatellite	Xp	-	-	-	F: ATGATTCCAAATGAGGCCTG R: AGCAATCCTTGCAGGCAG	187	2.0	58	AF536285
UMNe241	microsatellite	Xq	-	-	-	F: TTAGTATGTGCTACCCCCCG R: CTCTTTGCCCTGAATTTTATGG	103	1.5	58	AF536311
UMNe266	microsatellite	Xq	-	-	-	F: TTCCAAAAAACATAACAGGGTG R: CCAATTTGAATGCTTTTCATTG	150	2.0	58	AF536324
UMNe350	microsatellite	Xp	-	-	-	F: GCAAAATAAAAGGGTCACTTGC R: AGTGCTCCAGGTGCTTATATCC	153	2.0	58	AY391343
UMNe375	microsatellite	Xq	-	-	-	F: GGAGCCAGTTCAAAGATTTCC R: ATTCAGTGGGTTGTTTTGCC	130	2.0	58	AY391357
UMNe376	microsatellite	Xq	-	-	-	F: ACGCATATTAAAGTTTGAGAAGTCC R: CAGACCTACTGAATTGGAATACACC	155	2.0	58	AY391358

Table A6-1 (continued)

Symbol Microsatellites	Name	Horse cyto.	Human cyto	Human sequence map	Mouse sequence map	Primers 5' – 3'	PCR product (bp)	$MgCl_2$ (mM)	T_a	Reference and accession No
UMNe390	microsatellite	Xp	-	-	-	F: TATGAGACTCTGACATTTGCTGTG R: ATATTCCCTTCCTGAAACAGTCC	195	2.0	62	AY464476
UMNe397	microsatellite	Xq	-	-	-	F: TGTGGCTCCATCTCTCCAG R: TTTTCATGTCCCTAGGAAATTC	126	2.0	55	AY464478
UMNe402	microsatellite	Xq	-	-	-	F: AAGATGTGGCCTGTTTCAGG R: TTGATTCCTGGAGACTGATGG	244	2.0	58	AY464479

VITA

Candice Lea Brinkmeyer Langford
Texas A&M University
Department of Veterinary Integrative Biosciences
College Station, TX 77843-4458

Education

Texas A&M University, College Station, Texas — Sept 2003 – Dec 2006
Ph.D. in genetics (graduation date: December 2006)
Texas A&M University, College Station, Texas — Jan 2001 – Dec 2002
B.S. in genetics
Blinn College, Brenham, Texas — Aug 1999 – Dec 2000

Honors and Awards

Academic Excellence Award: Assoc. of Former Students Memorial Scholarship 2006-07
Best Graduate Student Platform Presentation, Texas Genetics Society Meeting 2006
Primary author article on cover page of MAMMALIAN GENOME Aug. 2005
CVM Graduate Student Fellowship 2003

Publications

Raudsepp T, Lee EJ, Kata S, **Brinkmeyer C**, Mickelson J, Womack J, Skow LC, Chowdhary BP (2004) Exceptional conservation of horse-human gene order on X chromosome revealed by high resolution radiation hybrid mapping. Proc Natl Acad Sci USA 101, 2386-91

Brinkmeyer-Langford C, Raudsepp T, Lee EJ, Goh G, Schäffer AA, Agarwala R, Wagner M, Tozaki T, Mickelson JR, Womack JE, Skow LC, Chowdhary BP (2005) A high-resolution physical map of equine homologues of HSA19 shows divergent evolution compared to other mammals. Mamm Genome 16, 631-49

Wagner ML, Raudsepp T, Goh G, Schäffer AA, Agarwala R, Dranchak PK, **Brinkmeyer-Langford C**, Venta PJ, Skow LC, Chowdhary BP, Mickelson JR (2006) A 1.3-Mb interval map of equine homologues of HSA2. Cytogenet Genome Res 112, 227-34

Brinkmeyer-Langford C, Raudsepp T, Gustafson-Seabury A, Mickelson JR, Chowdhary BP (2006) BAC contig over a 5 Mb region of the horse genome for discovery of functional elements and evolutionary breakpoints. (In preparation)