Projeto Genoma Equino

Priscila M. M. de Leon

Dra., Médica Veterinária - Biotecnologia/UFPel

Projeto Genoma Equino

- 1980:

  ✓ Surgiu o plano de sequenciamento do genoma equino;

  Objetivo inicial:

  ✓ mapa genético do cavalo (32 cromossomos);

  ✓ elucidar doenças, criar vacinas e tratamentos;

  ✓ características genéticas de cor da pelagem;

  ✓ desenvolvimento de testes diagnósticos de doenças;

Projeto Genoma Equino
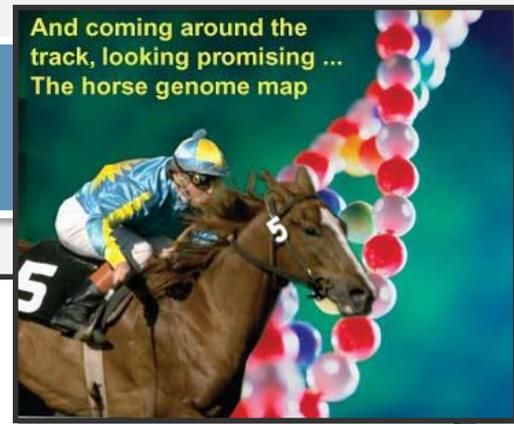
- **Em 1990:**

  - ✓ Aprovados US$ 3 bilhões;

  - ✓ Maior parte da verba era destinada a modernos equipamentos de sequenciamento, programas e computadores de bioinformática;

  - ✓ **Projeto do Genoma Equino baseado no Projeto Genoma Humano**;

- **Em 1995:**

  - ✓ 70 cientistas de 20 países se reuniram em Lexington/Kentucky para montar o plano de mapeamento do genoma do cavalo;

  - ✓ mapeados 7 genes de 1 cromossomo;

  - ✓ 15 cavalos foram coletados e distribuídos em 22 laboratórios de 12 países;

- ✓ Elaboração do Projeto Genoma Equino
- ✓ Foto da reunião de outubro de 1995 em Lexington/Kentucky;
- ✓ Havia 70 cientistas de 20 países
- ✓ foram 7 reunião de 1995 a 2007

- 1995 in Lexington/Kentucky
- 1997 in San Diego/ California
- 1999 in Uppsala/Sweden
- 2001 in Brisbane/Australia
- 2003 in Kruger Park/South Africa
- 2005 in Newbridge/Ireland
- 2007 at Lake Tahoe/California

Sequenciamento do Genoma Equino

- Em 2002:

  ✓ **900 genes mapeados**

  ✓ US$ 7 milhões e US$ 10 milhões para completar o mapa do genoma

  * Financiamento:
  Equine Research Coordination Group - AAEP, AQHA, Grayson-Jockey Club Research Foundation e Morris Animal Foundation

  * Principais Universidades envolvidas:
  University of California-Davis, Cornell University, University of Kentucky, University of Minnesota, Texas A&M University, Tufts University

  * Laboratórios Privados:
  Australia, New Zealand, France, Sweden, Germany, South Africa, Japan, Czech Republic, Poland, Denmark, Norway, and the United Kingdom.

Sequenciamento do Genoma Equino

* se estimava 100.000 a 300.000 genes
   Mas... na verdade 20.000 genes (representando 2% dos cromossomos)

- Em 2005:
   o Instituto Nacional de Pesquisa do Genoma Humano alertou para os benefícios do sequenciamento do genoma equino na Medicina.

- Em 2006:
   - ✓ o sequenciamento começou em fevereiro;
   - ✓ o genoma foi cortado em 30 milhões de pedaços;
   - ✓ sequenciamento de cada pedaço estava completo em julho;
   - ✓ montagem da sequência concluída em janeiro de 2007;

   UCSC Genome Browser Home

genome.gov
National Human Genome Research Institute
National Institutes of Health

Research Funding | Research at NHGRI | Health | Education | Issues in Genetics | Newsroo

Highlights RSS

The heart of the matter: genomics and cardiovascular disease

Please join NHGRI at a 'genomics and cardiovascular disease' lecture by Les Biesecker, M.D., Genetic Disease Research Branch, NHGRI. The April 13 lecture at Suburban Hospital in Bethesda, Md., focuses on the genetically distinct components of cardiovascular disease and how genomic diagnostic approaches can provide the basis for individualized patient treatment. **Read more**

- benefícios do sequenciamento do genoma do cavalo para a pesquisa médica e veterinária;
- sequenciamento de 24 genomas de mamíferos para identificar semelhanças e diferenças com o genoma humano;

Institute Diversity Action program, has taken big steps towards his dream

http://genome.ucsc.edu/index.html

UCSC Genome Bioinformatics

Genomes - Blat - Tables - Gene Sorter - PCR - VisiGene - Session - FAQ - Help

Horse Genome Browser – equCab2 assembly (sequences)

The Sep. 2007 *Equus caballus* draft assembly EquCab2 (UCSC version equCab2) was produced by The Broad Institute. For more information about this assembly, see EquCab2.0 in the NCBI Assembly database.

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic region, an mRNA or EST, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the Horse genome. See the User's Guide for more information.

"Twilight" - *Equus caballus*
Photo courtesy of NHGRI (press photos)

Request:	Genome Browser Response:
chr13	Displays all of chromosome 13
chr13:1-1000000	Displays first million bases of chromosome 13
chr13:1000000+20000	Displays a region of chr 13 that spans 20,000 bases, starting with position 1,000,000
AF506972	Displays region of mRNA with GenBank accession number AF506972
BM734875	Displays region of EST with GenBank accession BM734875 on chr chr11
pseudogene mRNA	Lists transcribed pseudogenes, but not cDNAs
zinc finger	Lists many zinc finger mRNAs
kruppel zinc finger	Lists only kruppel-like zinc fingers
smith	Lists mRNAs deposited by scientist named Smith
Collins,M.E.	Lists mRNAs deposited by co-author M.E. Collins

Use this last format for author queries. Although GenBank requires the search format *Collins ME*, internally it uses the format *Collins,M.E.*.

Assembly details

Broad Institute:

- Instituto de pesquisa biomédica
- Centro colaborador no sequenciamento para NHGRI

BROAD INSTITUTE

Horse Genome Project

The domestic horse, *Equus caballus,* is a member of the mammalian order Perrisodactyla. The horse genome is being sequenced for two reasons. It is expected to help identify functional genome features common to all mammals, and it will serve as a tool for researchers to better understand the diseases that affect equines. The recent creation of modern breeds and the presence of specific diseases within certain breeds together suggest that trait mapping may be relatively easy within horse breeds. Since horses and humans share a number of medical conditions (such as allergies and arthritis), mapping disease genes using horse populations may in turn benefit human health.

The equine genome sequencing project, performed mostly at the Broad Institute, has produced a high-quality draft sequence of a female thoroughbred horse. BAC end reads for the project were generated by the University of Veterinary Medicine, in Hanover, and the Helmholtz Centre for Infection Research in Braunschweig, Germany. In addition, the project generated a large collection of single-nucleotide polymorphisms (SNPs), to allow the mapping of genetic traits in horse populations. SNPs will be identified from a variety of modern and ancestral breeds, including the Akal-teke, Andalusian, Arabian, Icelandic, Quarterhorse, Standardbred, and Thoroughbred.

Current Status

Projeto Genoma Equino

http://www.uky.edu/Ag/Horsemap/welcome.html

Horse Genome Project

home | about the hgp | the people ▾ | the horses | genomics 101 (terms and definitions) | applications of genome study ▾ | faq ▾ | links | contact us

The Horse Genome Project is a cooperative international effort by over 100 scientists in 20 countries to define the genome, the DNA sequence, of the domestic horse. With this knowledge, scientists can begin to understand the genetic aspects of equine physiology and disease. Genetic tools have the potential to help researchers find new therapies and treatments for diseases such as laminitis, respiratory diseases, etc.

Some horsemen may worry that once the horse genome is sequenced, all the mystery and magic will be gone from horse breeding and ownership -- colorful characters around a racetrack replaced by colorless scientists with computer printouts and test tubes. Fear not. The study of the horse genome is more like studying the weather than inventing a sports car. Consider the following. All our research on the weather has shown we can experience a sunny day in Florida knowing that a hurricane will arrive in two days, but we cannot change the hurricane. Still the knowledge allows us to make choices; when the hurricane approaches we can abandon town or we can shore up our foundations. Successful use of genomic tools will help the horse breeders, veterinarians and horse owners to do better what they already do... anticipate problems, predict outcomes and enjoy the unique interaction between horses and people.

RECENT NEWS

International Equine Genome Mapping Workshop

The 8th Dorothy Russell Havemeyer Foundation International Equine Genome Mapping Workshop took place near Newmarket, UK from July 22 to 25, 2009...[more]

MEDIA ECLIPSE AWARD

Writers at the Louisville Courier Journal recently won the 2008 Media Eclipse Award for journalism based

http://www.uky.edu/Ag/Horsemap/hgphorses.html

http://www.uky.edu/Ag/Horsemap/hgppeople.html

Twilight

Égua Puro-Sangue, da Cornell University Ithaca/ NY

Amostra de sangue

Foi escolhida devido ter menor diversidade de marcadores genéticos entre 10 cavalos testados

Kerstin Lindblad-Toh

Eli and Edythe L. Broad Institute, Massachusetts Institute of Technology and Harvard University / Cambridge

Senquenciamento do genoma de mamíferos: camundongo, cão, cavalo, chipanzé, gambá

Claire Wade

Computational Biology and Animal Genetics, The University of Sydney

Projeto Genoma Equino

✓ gastos US$ 15 milhões;

✓ Tamanho do genoma: 2,7 bilhões de pb de DNA;

✓ Mapa de variação genética entre raças (Andaluz, Puro-Sangue, Árabe, Quarto de Milha, Akel Teke, Icelandic, Standardbred)

✓ 1 milhão de SNPs - http://www.broadinstitute.org/mammals/horse

✓ 90 variantes genéticas similares aos distúrbios observados em seres humanos (musculoesqueléticas, doenças neuromusculares, cardiovasculares, respiratórias e de infertilidade)

→ **Pesquisas Concluídas/Andamento**:

- ✓ Imunodeficiência combinada severa do cavalo Árabe
- ✓ Paralisia periódica hipercalêmica do Quarto de Milha
- ✓ Síndrome branca letal no Paint Horse
- ✓ Laminite
- ✓ Deformidades ósseas
- ✓ Cólica
- ✓ Hipoplasia cerebelar
- ✓ Megacólon
- ✓ Doenças musculares
- ✓ Doenças ósseas
- ✓ Cor da pelagem
- ✓ Performance Atlética
- ✓ Fertilidade
- ✓ Comportamento

* Publicado dia 6 de novembro de 2009.

Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse

C. M. Wade,[1,2,3]* E. Giulotto,[4] S. Sigurdsson,[1,5] M. Zoli,[6] S. Gnerre,[1] F. Imsland,[5] T. L. Lear,[7] D. L. Adelson,[8] E. Bailey,[7] R. R. Bellone,[9] H. Blöcker,[10] O. Distl,[11] R. C. Edgar,[12] M. Garber,[1] T. Leeb,[11,13] E. Mauceli,[1] J. N. MacLeod,[7] M. C. T. Penedo,[14] J. M. Raison,[8] T. Sharpe,[1] J. Vogel,[15] L. Andersson,[5] D. F. Antczak,[16] T. Biagi,[1] M. M. Binns,[17] B. P. Chowdhary,[8] S. J. Coleman,[7] G. Della Valle,[6] S. Fryc,[1] G. Guérin,[19] T. Hasegawa,[20] E. W. Hill,[21] J. Jurka,[22] A. Kiialainen,[23] G. Lindgren,[24] J. Liu,[25] E. Magnani,[4] J. R. Mickelson,[26] J. Murray,[27] S. G. Nergadze,[4] R. Onofrio,[1] S. Pedroni,[14] M. F. Piras,[4] T. Raudsepp,[8] M. Rocchi,[28] K. H. Røed,[9] O. A. Ryder,[30] S. Searle,[15] L. Skow,[18] J. E. Swinburne,[31] A. C. Syvänen,[23] T. Tozaki,[32] S. J. Valberg,[26] M. Vaudin,[31] J. R. White,[1] M. C. Zody,[1,5] Broad Institute Genome Sequencing Platform,[1] Broad Institute Whole Genome Assembly Team,[1] E. S. Lander,[1,33] K. Lindblad-Toh[1,5]*

Supporting Online Material for

Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse

C. M. Wade, * E. Giulotto, S. Sigurdsson, M. Zoli, S. Gnerre, F. Imsland, T. L. Lear, D. L. Adelson, E. Bailey, R. R. Bellone, H. Blöcker, O. Distl, R. C. Edgar, M. Garber, T. Leeb, E. Mauceli, J. N. MacLeod, M. C. T. Penedo, J. M. Raison, T. Sharpe, J. Vogel, L. Andersson, D. F. Antczak, T. Biagi, M. M. Binns, B. P. Chowdhary, S. J. Coleman, G. Della Valle, S. Fryc, G. Guérin, T. Hasegawa, E. W. Hill, J. Jurka, A. Kiialainen, G. Lindgren, J. Liu, E. Magnani, J. R. Mickelson, J. Murray, S. G. Nergadze, R. Onofrio, S. Pedroni, M. F. Piras, T. Raudsepp, M. Rocchi, K. H. Røed, O. A. Ryder, S. Searle, L. Skow, J. E. Swinburne, A. C. Syvänen, T. Tozaki, S. J. Valberg, M. Vaudin, J. R. White, M. C. Zody Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, E. S. Lander, K. Lindblad-Toh*

*To whom correspodence should be addressed. E-mail: c.wade@usyd.edu.au (C.M.W.); kersli@broadinstitute.org (K.L.-T.)

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 to S11
Tables S1 to S14
References

Sequenciamento

- O DNA de 1 égua (amostra de sangue)

- Biblioteca de BACs (314,972 BAC)

- Sequenciador ABI 3730 (Applied Biosystems)

- Arachne 2.0

- EquCab 2.0

- Genomas Humano, Bovino, Canino e Camundongo

- **Resultados:**

  ✓ O tamanho do genoma é de 2.7 Gb

  ✓ maior que o cão (2.5 Gb) e menor que o bovino e humano (2,9Gb)

  ✓ 2% do DNA codifica proteína (98% "DNA Lixo")

  ✓ 46% do genoma são sequências repetitivas

Animais utilizados nos estudos genômicos

Table S1. Sources of Samples

Strain	Number	Use †	Contributor
Twilight - Thoroughbred	1	GSDRL	Cornell University (USA)
Andalusion Horse	1	SDRL	University of Kentucky (USA)/Kentucky Horse Park
Icelandic Horse	1	SDRL	University of Kentucky (USA)/Kentucky Horse Park
Akhal teke	1	SDRL	University of Kentucky (USA)/Kentucky Horse Park
Thoroughbred Horse	1	SDRL	University of Kentucky (USA)
Standardbred-Pacer	1	SDRL	University of Kentucky (USA)
Quarter Horse	1	SDRL	University of California -Davis (USA)
Arabian Horse	1	SDRL	University of Kentucky (USA)/Kentucky Horse Park
Przewalskii Horse (Equid)	1	RL	San Diego Zoo's Institute for Conservation Research (USA)
Donkey (Equid)	1	RL	San Diego Zoo's Institute for Conservation Research (USA)
Paso Fino	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Exmoor Pony	1	RL	University College of Dublin (Ireland)
Trakehner	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Friesian	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Belgian draft	1	RL	University of Minnesota (USA)
Norwegian Fjord	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Shire	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Suffolk Punch	1	RL	University of Kentucky (USA)/Kentucky Horse Park
Lusitano	1	RL	University of Kentucky (USA)/Kentucky Horse Park
American Saddlebred	1	RL	University of Kentucky (USA)
French Trotter	1	RL	INRA (France)
Selle Francais	1	RL	INRA (France)
Thoroughbred	24	L	University of Kentucky (USA)
Arabian	24	L	University of Kentucky (USA)
Icelandic Horse	24	L	Uppsala University (Sweden)
Hokkaido	24	L	Laboratory of Racing Chemistry (Japan)
Belgian draft	24	L	University of Minnesota (USA)
Quarter horse	24	L	University of California -Davis (USA)
French Trotter	24	L	INRA (France)
Standardbred-Pacer	24	L	University of Kentucky (USA)
Hanoverian	24	L	Institute of Animal Breeding and Genetics (Germany)
Norwegian Fjord	24	L	Norwegian School of Veterinary Science (Norway)
Andalusian	24	L	University of Kentucky (USA)
Przewalskii Horse	8	R	San Diego Zoo's Institute for Conservation Research (USA)

† Genome sequencing (G), SNP discovery (S), Re-sequencing (R), L=Linkage Disequilibrium and Haplotype analysis (L)

Equino x Humano

- revela sintenia conservada entre as espécies

- 17 cromossomos (53%) do cavalo compõem o material de 1 cromossomo humano (cão é 29%)

- 15.027 genes ortólogos com humanos

- 90 hereditariedades podem servir para modelo humano:
  - ✓ Infertilidade
  - ✓ Doenças inflamatórias
  - ✓ Desordens musculares

Sintenia: *propriedade de dois ou mais genes estarem localizados no mesmo cromossomo em espécies diferentes.*
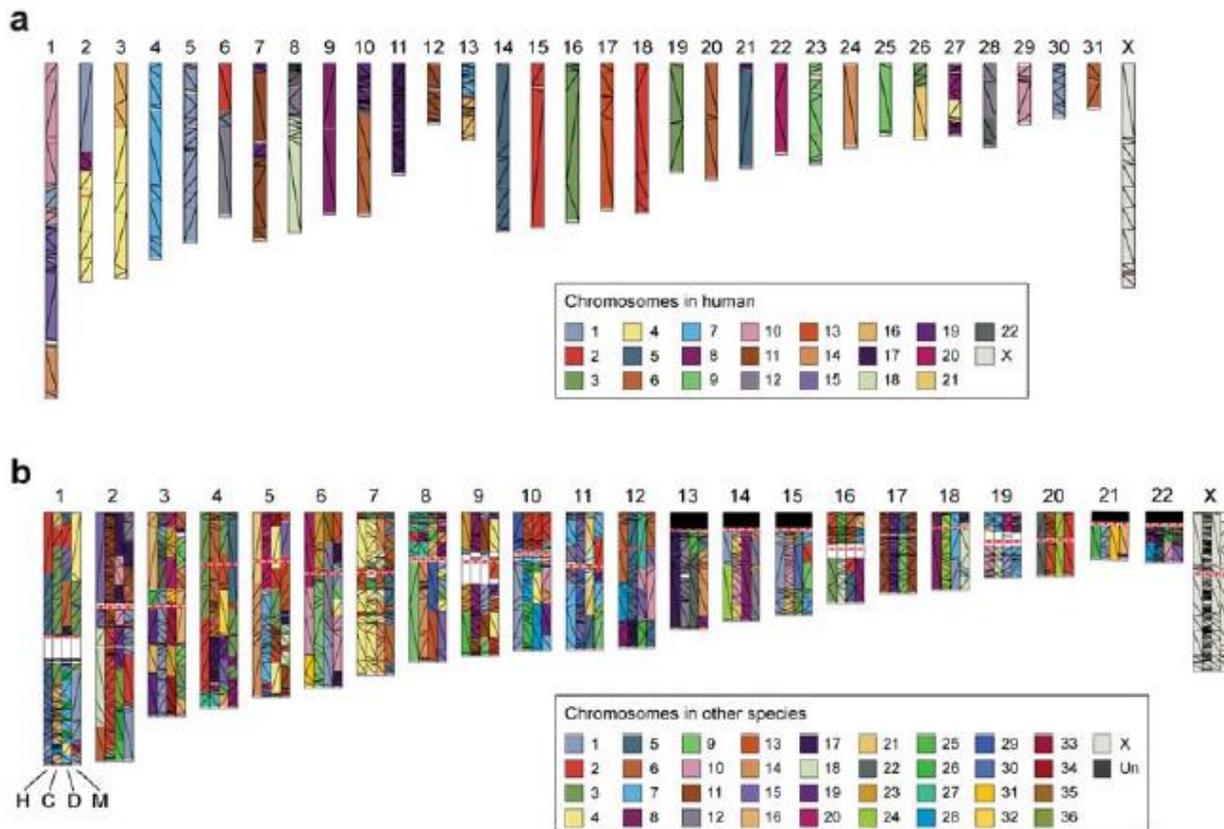
Sintenia Conservada

A - Cromossomos do Cavalo é colorido com o correspondente Humano

B - Cromossomos Humanos são coloridos com Cavalo (H), Bovino (C), Cão (D) e Camundongo(M).

Característica Inesperada

➤ Novo Centrômero Evolucionário (**ENC**) no cromossomo 11 (**ECA11**)

➤ mudança de posição centromérica sem rearranjo do cromossomo

  ✓ ECA11 região conservada em mamíferos

  ✓ formação evolutiva muito recente em equinos

  ✓ é funcional e estável em todos os cavalos

* Região Centromérica de mamíferos é tipicamente complexa, com repetições satélite em tandem
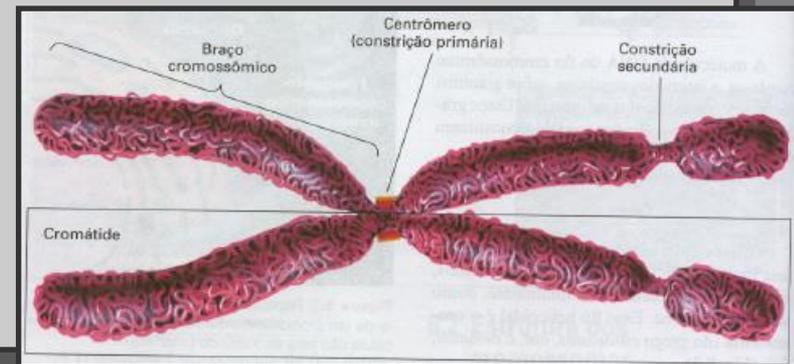
a
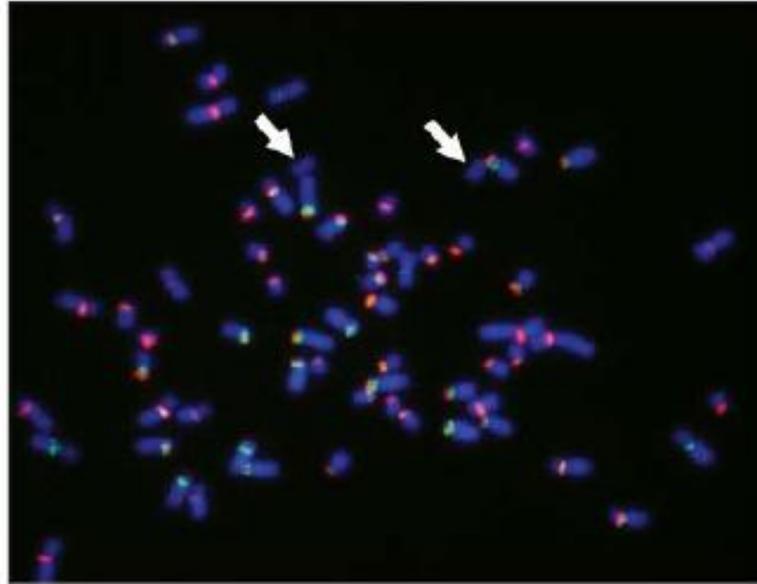
Fig. S2. Localization and sequence analysis of the ECA11 centromere. (a) Hybridization of the two major horse satellite sequences on horse chromosomes; the 23 bp repeat (2p1) is labeled in red and the 221 bp repeat (37cen) is labeled in green. All centromeres are labeled with one or both satellite probes except chromosome 11 (arrows). **(b)** Schematic representation of the cytogenetic localization of the primary
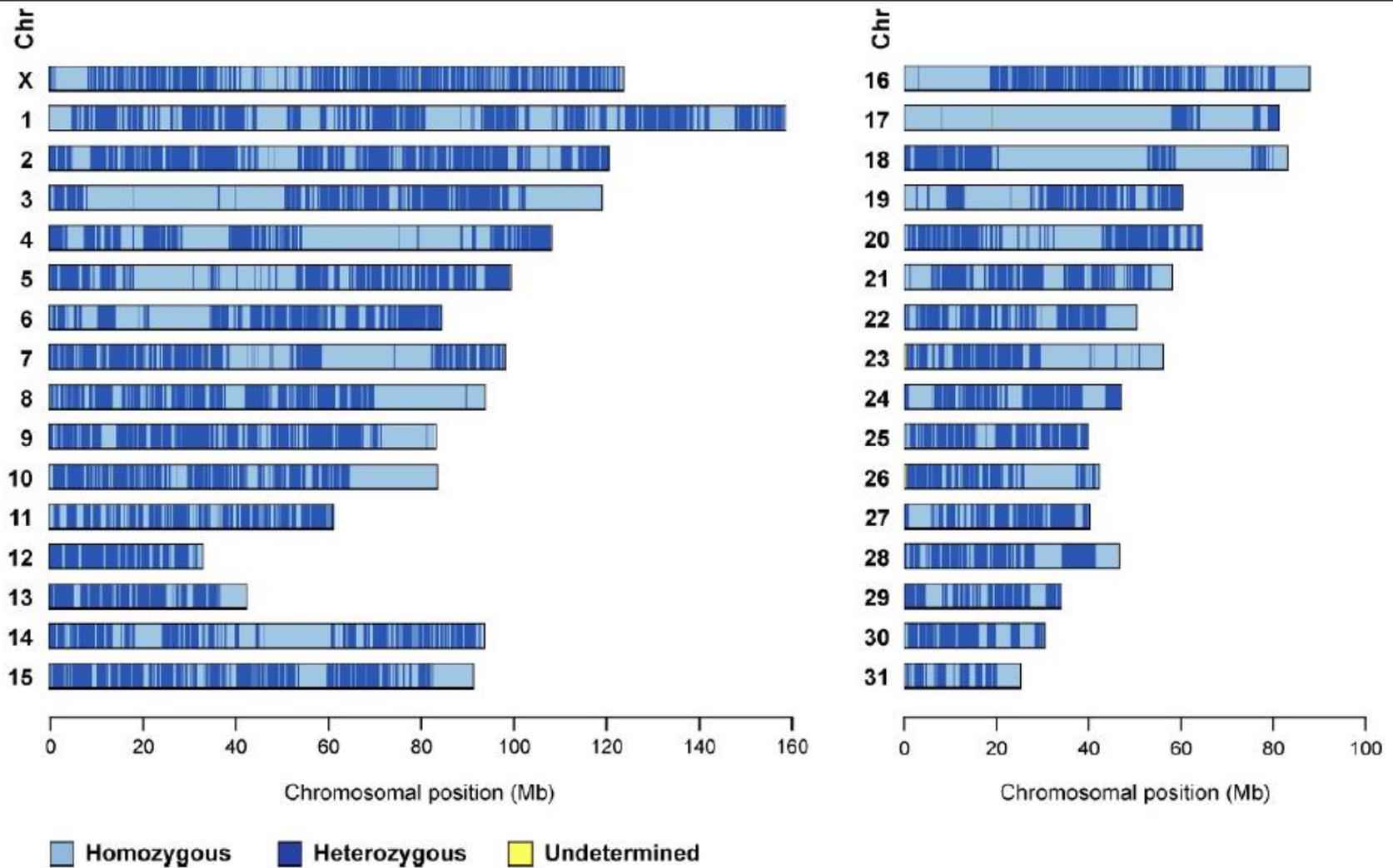
Homozigose no genoma do cavalo

✓ 46% homozigoto

✓ maior % observado entre os genomas seqüenciados

LD (Linkage Desequilibrium) por raça.

✓ Maioria das raças apresentam LD semelhante

✓ Puro-Sangue LD mais distante

✓ 100.000 SNPs para mapeamento de raças

Desequilíbrio de Ligação: associação não-aleatória de polimorfismos em dois ou + locus, não necessariamente no mesmo cromossoma

a Filogenia

b eca4

c eca5

d eca11

e eca14
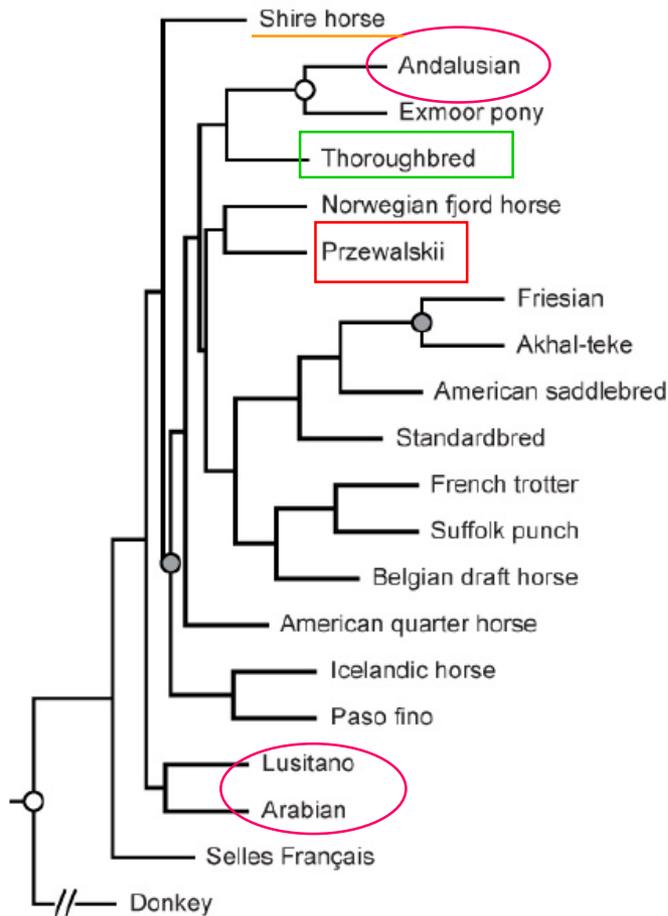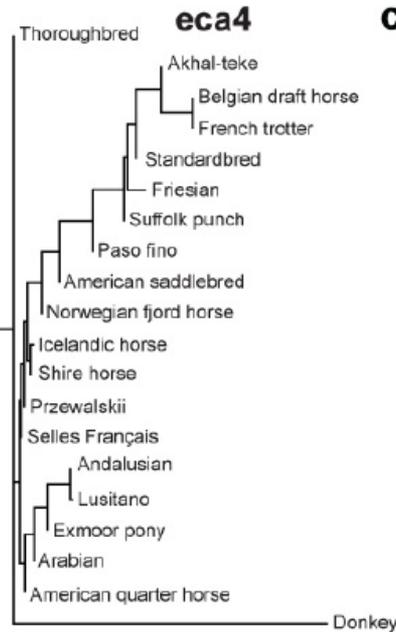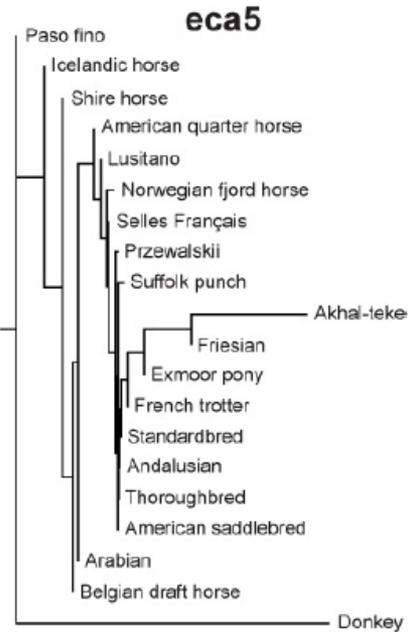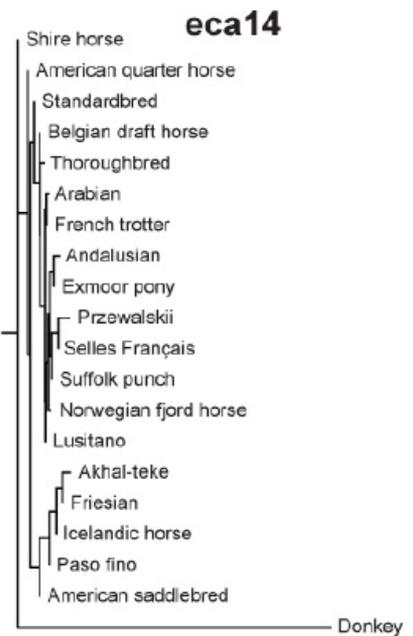
Shire

Przewalski

PSI

Lusitano

Andaluz
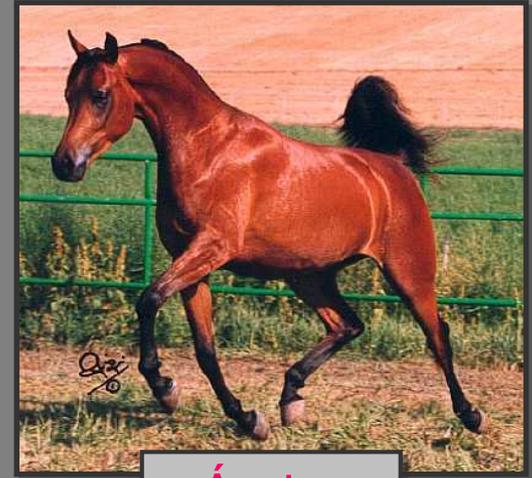
Árabe

SNP

- 98% dos SNPs são polimórficos (verdadeiros);

- sugerindo a expansão da população e estrangulamento da seleção genética;

- Raça Akhal Tekel apresenta menor nível de polimorfismo (61%) e SNPs específicos;

- Tx de SNP no cavalo é 1 / 1.200 pb ;

- Tx SNP similar ao homem e o cão

- 100.000 SNP para o mapeamento genético;

- Mapeamento de raças antigas;

Akhal Tekel

SNP

Standardbred

Akhal Tekel

- 74,3% de polimorfismos
- maior nível de polimorfismo

- 61% de polimorfismos
- menor nível de polimorfismo

Pelagem Leopardo do Apaloosa
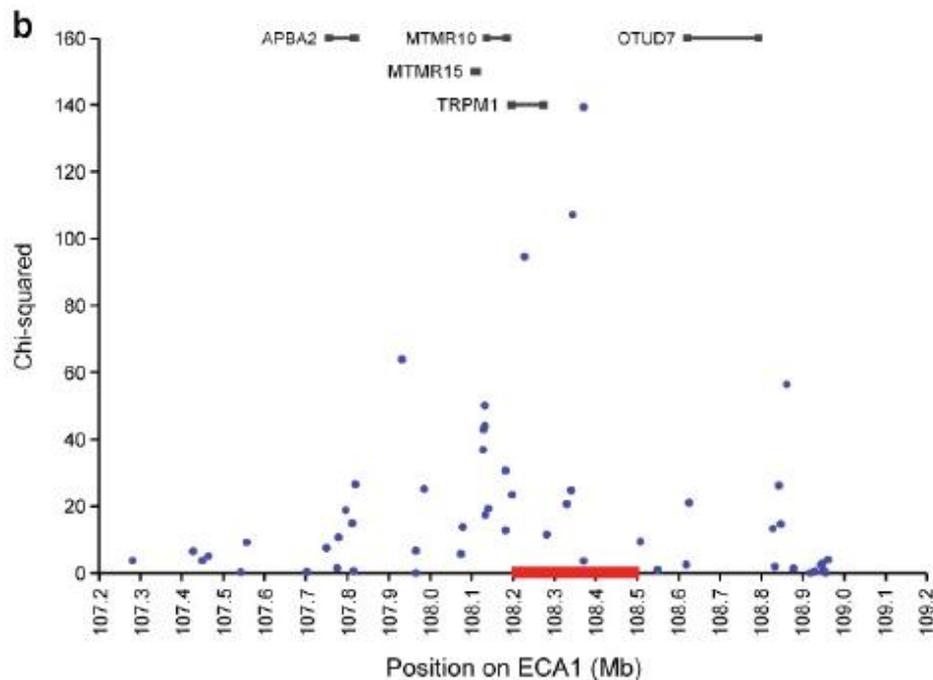
✓ Heterozigoto – muitas manchas brancas

✓ Homozigoto – poucas manchas brancas (cegueira)

✓ 85 indivíduos Apaloosa – 70 SNP identificados, sendo
2 em regiões conservadas

ANIMAL
REPRODUCTION
SCIENCE

Potential applications of equine genomics in dissecting diseases and fertility[☆]

Bhanu P. Chowdhary *, Nandina Paria, Terje Raudsepp

*Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences,
Texas A&M University, College Station, TX 77843-4458, USA*

Introdução

- **<u>Importância do mapeamento genético</u>**

  *isolamento de genes e marcadores associados com características importantes economicamente, como:

  - ✓ Resistência ou suscetibilidade a doenças
  - ✓ Crescimento
  - ✓ Reprodução
  - ✓ Desempenho Atlético
  - ✓ SNP (identificação rápida de doenças específicas dos cromossomos)
  - ✓ Desenvolvimento de SNP-chip

Table 1a
Equine genetic disorders with available genetic tests

Trait/condition	Locus	Chromosome
Glycogen storage disease IV, GBED	GBE1	ECA26q
Hereditary equine regional dermal asthenia, HERDA	PPIB	ECA1
Herlitz junctional epidermolysis bullosa, H-JEB	LAMC2	ECA5p
Hyperkalemic periodic paralysis, HYPP	SCN4A	ECA11p
Overo lethal white foal syndrome, OLWFS	EDNRB	ECA17q
Polysaccharide storage myopathy, PSSM	GYS1	ECA10
Severe combined immunodeficiency, SCID	PRKDC	ECA9p

GBED: Quarto de Milha e Paint Horse, letal e recessiva

HERDA: Quarto de Milha, recessivo, potros de 1,5 anos.

H-JEB: semelhança Homem

HYPP: Quarto de Milha, semelhança Homem.

OLWFS: gene Overo

PSSM: 10% de Quarto de Milha e 36% de Belgas

SCID: 8% dos cavalos Árabes

Equine genetic disorders with known linked markers		
Trait/condition	Locus	Chromosome
Anterior segment dysgenesis, ASD	ms	ECA6
Cerebellar abiotrophy	DMAP1, PRNPIP	ECA2p
Degenerative suspensory ligament desmitis, DSLD	ms	ECA14qter
Endurance performance	ACE	ECA11
Epitheliogenesis imperfecta	LAMA3	ECA8
Fertility (sperm-egg fusion)	CRISP1	ECA20q
Insect bite hypersensitivity, IBH	HMS01	ECA15
Laminitis	KIT downregulation	ECA3
Osteochondrosis and navicular disease	8 candidate genes, SNPs	Five chromosomes
Recurrent airway obstruction	ms	ECA13
Recurrent exertional rhabdomyolysis, RER	ms, gene exclusion	ECA4/ECA12?
Sex reversal	Y deletion including SRY	ECAY

✓ Projeto em andamento no Texas para determinar SNPs e expressão gênica ligados a regulação da fertilidade e determinação do sexo. Analisando em condições de anormalidades.

✓ 10 – 15% da infertilidades do Homem são causadas em mutações no Y

Banco de Dados

Banco de Dados

EquineSNP50 Genotyping BeadChip

The EquineSNP50 Genotyping BeadChip features more than 54,000 evenly spaced and validated SNPs derived from the EquCab2 assembly. This 12-sample Infinium® BeadChip presents a cost-effective and high-quality genotyping solution for equine research.

EquinesSNP50 BeadChip Highlights

- **Uniform Coverage:**
  High-density, evenly distributed coverage across the entire genome

- **High Throughput:**
  12 samples per BeadChip, 54,602 SNP assays per sample

- **Streamlined Workflow:**
  PCR-free protocol

- **Low Sample Input:**
  200 ng per sample

- **Integrated Data Analysis:**
  BeadStudio Software

Introduction

The EquineSNP50 Genotyping BeadChip was developed in collaboration with the International Equine Genome Mapping Workshop[1] and the Morris Animal Foundation's Equine Genome Consortium[2]. This

Figure 1: EquineSNP50 BeadChip