

# ESTADO DA ARTE DO SEQUENCIAMENTO GENÔMICO NA PECUÁRIA

## *STATUS OF THE ART SEQUENCING GENOME ON LIVESTOCK*

G. L. PEREIRA<sup>1\*</sup>, K. O. ROSA<sup>1</sup>, R. A. CURI<sup>2</sup>, L. C. A. REGITANO<sup>3</sup>, M. D. S. MOTA<sup>2</sup>

### RESUMO

Desde que a possibilidade de determinar a sequência nucleotídica de genomas surgiu em 1975, muitos foram os avanços da genômica. Na década de noventa teve início o sequenciamento do genoma humano, viabilizado em grande parte pelos avanços nos métodos computacionais. Nos últimos dez anos, o advento de tecnologias de sequenciamento de nova geração permitiu o sequenciamento de milhões de bases a baixo custo e em curto espaço de tempo quando comparadas às técnicas de sequenciamento anteriores. Após a conclusão do projeto genoma humano, muitas iniciativas foram tomadas para a realização do sequenciamento de diversas espécies domésticas, gerando grande volume de dados, e redirecionando os estudos genéticos. Esta revisão teve como objetivo descrever o estado atual do sequenciamento das principais espécies de animais domésticos de interesse econômico, bem como de expor as ferramentas mais utilizadas no acesso às informações genômicas.

**PALAVRAS-CHAVE:** BAC. Genoma. Next Generation Sequencing. SNP. WGS.

### SUMMARY

There has been a lot of advance in genomics since 1975 when the possibility to determine the nucleotide sequence of a genome was described. In the 90's the human genome sequencing was started and it was greatly favored by advances in computer technologies. In the last ten years the development of next generation sequencing technologies allowed the sequencing of millions of bases at low cost and in a shorter time compared to the previous technologies. After the conclusion of the human genome project, several initiatives to sequence the genome of domestic animal species were taken, resulting in a large amount of data that is redirecting the goals of genetic studies in domestic animals. The aim of this review was to describe the present situation of the sequencing initiatives on the main domestic animal species of economical interest as well as to list the most important tools available to access the genomic information.

**KEY-WORDS:** BAC. Genome. Next Generation Sequencing. SNP. WGS.

A história do sequenciamento genômico teve início em meados da década de 70, primeiramente com a elaboração da metodologia descrita por Frederick Sanger e colaboradores em 1975, a qual ficou conhecida por método de Sanger, didesoxi ou terminação de cadeia. De forma independente, Maxam e Gilbert desenvolveram, em 1976, a técnica de sequenciamento químico, a qual foi preterida devido à maior dificuldade técnica para automatização (MAXAM & GILBERT, 1977; SANGER & COULSON, 1975). A metodologia de Sanger é fundamentada no princípio da terminação controlada da polimerização pela utilização de didesoxinucleotídeos, possibilitando assim determinar a ordem sequencial dos nucleotídeos que compõem fragmentos da molécula do DNA. Em sua versão original não era ferramenta apropriada para o sequenciamento de genomas complexos (FURLAN *et al.*, 2007; SANGER & COULSON, 1975). Porém, no final da década de 80, após o método de Sanger ter sido modificado, o sequenciamento passou a ser realizado por meio automatizado e integrado a sistemas de leitura computadorizados (VENTER *et al.*, 2001).

Em meados da última década, plataformas baseadas em tecnologias de sequenciamento de próxima geração, ou de nova geração (*Next-Generation Sequence* – NGS), começaram a ser comercializadas e têm sido amplamente empregadas em substituição ao método de Sanger. Ao sequenciar genomas extensos, como o humano, a técnica de Sanger se mostra pouco eficiente frente aos métodos baseados em NGS, com grande demanda de tempo e altos custos.

Muitos animais domésticos de interesse econômico já têm os genomas em sua maior parte sequenciados. Alguns projetos que tiveram início antes da era NGS, adotaram tais tecnologias para finalizar ou melhorar a qualidade dos dados gerados. Outras espécies, por exemplo a caprina (*Capra hircus*), uma das últimas a ter o genoma decifrado, foram sequenciadas inteiramente por tecnologia NGS (WADE *et al.*, 2009; DOAN *et al.*, 2012; DONG *et al.*, 2012). Essas tecnologias vêm sendo amplamente utilizadas para sequenciamento de animais de diferentes raças de espécies já sequenciadas, visando a descoberta de variações no número de cópias (*copy number variants* – CNVs), inserções e deleções (*insertion/deletion* – InDels) e novos polimorfismos de nucleotídeo único (*single nucleotide polymorphism* – SNPs), entre outros fins. Esses estudos têm contribuído muito para melhorar a quantidade e a qualidade de dados genômicos para estas espécies (DOAN *et al.*, 2012; BARRIS *et al.*, 2012).

Diante do exposto, o objetivo desta revisão foi relatar o atual estado do sequenciamento genômico das principais espécies de interesse zootécnico, bem como descrever ferramentas computacionais de acesso aos dados genômicos e tecnologias desenvolvidas a partir destas informações.

### **Estratégias de sequenciamento e sequenciamento de próxima geração**

Antes do sequenciamento propriamente dito, seja pelo método de Sanger ou NGS, é necessário adotar estratégias para obtenção dos fragmentos de DNA do genoma a ser sequenciado, já que não é possível o sequenciamento de grandes segmentos de DNA.

Para tal, são utilizadas duas estratégias principais o sequenciamento *shotgun* de genoma inteiro (*whole genome shotgun* – WGS) e outra chamada de sequenciamento de clones ordenados. Ambas as estratégias visam gerar bibliotecas genômicas de clones, obtidas na fragmentação do genoma em milhares de pequenos segmentos aleatórios, seguido da leitura de cada sequência e sobreposição destas, por meio computacional (GRIFFITHS, 2008).

A diferença básica entre estratégias está no conhecimento prévio da localização dos fragmentos sequenciados no genoma. Em estratégias que se utilizam clones de sequências conhecidas, como sequenciamento de clones ordenados, as sequências podem ser obtidas por meio de cromossomos artificiais de bactérias (BAC) que são isolados e mapeados para o genoma, depois individualmente sequenciados. O WGS sequencia fragmentos aleatórios do genoma, obtidos por clones não mapeados, para posteriormente, após sobreposição de sequências lidas e formação de *contigs* serem mapeados para o genoma (VENTER *et al.*, 2001; GRIFFITHS, 2008).

Após 12 anos da publicação do sequenciamento do genoma humano em 15 de fevereiro de 2001 na revista *Nature* (VENTER *et al.*, 2001), muitos avanços ocorreram em relação às técnicas de sequenciamento e aos estudos genômicos. Os métodos de sequenciamento de próxima geração tornaram-se realidade, com rápidos avanços tecnológicos na última década. Com o auxílio de ferramentas de bioinformática, o NGS permite o sequenciamento e análise de milhões de pares de bases (pb) em rodada única de leitura. Além da velocidade de sequenciamento, deverá haver redução considerável dos custos por genoma sequenciado, atingindo, provavelmente, nos próximos cinco ou 10 anos, queda de até 10.000 vezes em comparação aos primeiros sequenciamentos completos de genomas (PAREEK *et al.*, 2011).

A primeira geração de NGS teve início com o desenvolvimento do equipamento *454 life science* em 2000 por Jonathan Rothberg ([www.454.com](http://www.454.com)). Essa nova plataforma de sequenciamento massivo de DNA foi validada por Margulies *et al.* (2005), que logo no início do seu artigo publicado na revista *Nature* descreve um "sistema de sequenciamento altamente paralelo e escalável com automatização significativamente maior do que os instrumentos padrão de eletroforese capilar", a qual permite produzir 25 milhões de pares de bases em uma única leitura com mais de 99% de precisão. Nesta ocasião foi realizado o resequenciamento de 96% do genoma da bactéria *Mycoplasma genitalium* com altíssima acurácia (99.96%) em uma única leitura. Esse procedimento, diferentemente da técnica que adota capilares (necessidade de eletroforese), tem abordagem de

sequenciamento em tempo real, por síntese, e se baseia na detecção de pirofosfatos luminométricos inorgânicos (PPi) liberados com a incorporação dos desoxiribonucleotídeos trifosfatados (RONAGHI *et al.*, 1998).

A segunda geração de NGS é representada principalmente por três plataformas: 454 FLX (Roche, EUA); Solexa (Illumina Inc., EUA) e SOLiD (Applied Biosystems, EUA), as quais possuem elevada taxa de sequenciamento, gerando de 150 a 200 Gb por semana e tamanho médio de fragmentos de 100 pb (CARVALHO & SILVA, 2010). A terceira geração de NGS, diferentemente da segunda, não necessita de amplificação do DNA, cujo objetivo é fortalecer o sinal luminoso para captação baseado em câmeras CCD, o que pode gerar distorções pela abundância de fragmentos. Na terceira geração, a miniaturização em nano escala e utilização mínima de reagentes, tornaram possível o sequenciamento de uma única molécula de DNA (SCHADT *et al.*, 2010). O primeiro sequenciador dessa geração foi o Heliscope™ (Pacific Bioscience, EUA), seguidos de outros como SMRT™ (Pacific Bioscience, EUA) (PAREEK *et al.*, 2010). As plataformas de terceira geração produzem sequências entre 30 e 200 vezes mais longas do que as de segunda. Essa característica pode ser de grande vantagem na montagem de genomas por bioinformática (ROBERTS *et al.*, 2013).

### Sequenciamento genômico na pecuária

Nas duas últimas décadas ocorreu, e vem ocorrendo, verdadeira explosão de projetos de sequenciamento de genomas. Dessa forma, todas as espécies que de alguma forma representam interesse para o homem, incluindo o próprio genoma humano, tiveram ou estão tendo o seu código genético decifrado. Entre estas espécies estão aquelas de exploração na pecuária. O estado do sequenciamento genômico das espécies de interesse econômico na pecuária encontra-se resumido na Tabela 1.

### Genoma das aves domésticas

Anterior ao sequenciamento completo do genoma, em 2002 foram sequenciados 339.314 *Expressed sequence tags* (ESTs), a partir de 64 bibliotecas de DNA complementar (cDNA), originárias de diferentes tecidos de galinha doméstica, os quais foram agrupados e montados em 85.486 sequências contíguas (*contigs*). Observou-se ao menos 38% de *contigs* ortólogos em outros organismos e 13.000 novos genes foram documentados (BOARDMAN, 2002).

Financiado pela *National Human Genome Research Institute* (NHGRI, EUA), o projeto de sequenciamento do genoma da ave doméstica foi iniciado em março de 2003, onde foram gastos cerca de US\$ 13 milhões (NHGRI, 2012). Com 38 pares de cromossomos autossomos e um par de sexuais, o genoma da galinha doméstica (*Gallus gallus*) foi, em 2004, o primeiro dentre os animais de produção a ter genoma sequenciado.

Composto por cerca de um bilhão de pares de bases, cerca de 40% do tamanho do genoma humano,

tem em sua estrutura aproximadamente 23.000 genes. Esta sequência rascunho do genoma do frango foi obtida a partir do DNA proveniente de uma fêmea endogâmica de galinha vermelha-da-selva (*Red Junglefowl*). Isto, para minimizar a heterozigosidade e prover sequências de ambos os cromossomos sexuais, Z e W. As sequências foram geradas com cobertura de 6,6x por meio de WGS e bibliotecas de plasmídeos e BACs (HILLIER *et al.*, 2004).

Em Maio de 2006 foi produzido pelo Centro de Sequenciamento do Genoma da *Washington University School of Medicine* e submetido pela *International Chicken Genome Sequencing Consortium* o conjunto de dados *Gallus\_Gallus2.1*. Aproximadamente 95% das sequências foram ancoradas a cromossomos autossomos 1-24, 26-28, e 32, e sexuais W e Z. Já em setembro de 2011 o mesmo Centro de Sequenciamento, visando corrigir lacunas e duplicações da versão 2.1, apresentou um novo conjunto de dados a partir do indivíduo RJF # 256, já usado em montagens anteriores. Com cobertura de 12x, obtida a partir da plataforma Roche™ 454®, esta nova montagem representou melhoria significativa nas informações do conjunto de dados *Gallus\_Gallus 2.1*, originando assim a versão 4.0 (UCSC, 2012).

Dois mil setecentos e trinta e seis locos de características quantitativas (*Quantitative Trait Loci* – QTL), representando 257 características diferentes foram descritos em 132 publicações, as quais encontram-se incluídas na base de dados *chickenQTLdb* (NAGRP, 2012).

Mais de 2,8 milhões de SNPs foram identificados por meio de comparação da sequência do genoma do ancestral (*Gallus gallus*) com sequências obtidas de três linhagens domesticadas: um macho de corte (*White Cornish*); uma fêmea de postura (*White Leghorn*); e uma fêmea de espécie ornamental (*Silkie chinesa*) (NCBI, 2012). A identificação de SNPs a partir de *contigs* derivados de NGS resultou em aumento da cobertura do genoma da galinha, permitindo o desenho do *Illumina SNP BeadChip* (Illumina Inc., EUA) de moderada densidade com aproximadamente 64 mil SNPs (60k) (GROENEN *et al.*, 2011).

### Genoma bovino

Em dezembro de 2003, o *International Bovine Genome Sequencing Project* foi lançado em decorrência de um “*workshop*” realizado em junho de 2003 na cidade de Montreal no Canadá. Já em dezembro de 2004, a primeira sequência rascunho do genoma bovino havia sido depositada em bancos de dados públicos e gratuitos. Entre as muitas entidades envolvidas, a *Genome British Columbia*, juntamente com o *National Human Genome Research* (NHGRI) e *United States Department of Agriculture* (USDA) foram alguns dos mais importantes parceiros, no que diz respeito ao financiamento e desenvolvimento do projeto (Genome Canada, 2013).

O sequenciamento do genoma bovino empregou abordagem WGS combinado à BAC. Muitos dos BACs do projeto foram sequenciados utilizando-se a estratégia de sequenciamento de clones agrupados,

**Tabela 1** – Situação atual do sequenciamento Genômico na Pecuária.

Animal	Chrs (n)	Extensão sequenciada	Estratégia de sequenciamento	Plataforma	Cobertura	Versões mais recentes disponíveis	Painéis comerciais de SNPs	Publicação da primeira sequência de alta qualidade	Coordenação dos projetos de sequenciamento
Frango	33	1.2 Gb	Cromossomo artificial bacteriano (BAC), fosmídeos, e plasmídeos baseado em WGS	-	6,6x	Gallus gallus 4.0	60k	<i>Nature</i> 09/12/2004	<i>Washington University Genome Sequencing Center</i>
Suíno	20	2.8 Gb	Baseado em clones	Sanger, Illumina / Solexa	-	Sscrofa 10.2	60k	<i>BMC genomics</i> 19/07/2010	<i>Swine Genome Sequencing Consortium</i>
Bovino	30	2,67Gb	Combinação de sequenciamento de clones ordenados e WGS	-	9,5x	UMD 3.1	3 a 778k	<i>Genome Biology</i> 24/3/2009	<i>Baylor College of Medicine</i> em Houston, Texas O genoma foi remontado por Salzberg in Baltimore, Maryland
	31	2,98Gb	Montagem combinada de WGS 7.15x e sequências BAC	-	7.15x	BTAU 4.6.1		<i>Science</i> 24/4/2009	<i>Bovine Genome Sequencing Project</i> conduzido por <i>Baylor College of Medicine's Human Genome Sequencing Center</i> em Houston, Texas
Ovino	27/27	2.6Gb/2,86Gb	WGS	454-GS-FLX, Illumina GAI	-	OAR 3.1/Ovis Aries 1.0	50k	<i>Animal Genetics</i> 30/8/2010	<i>International Sheep Genomics Consortium</i>
Equino	32	2.47Gb	WGS	Sanger, Illumina GAI	6.79x	EquCab 2.0	70k	<i>Science</i> 06/11/2009	<i>Broad Institute e Horse Genome Project</i>
Caprino	30	2.66Gb	Fosmídeos, WGS	Illumina GA-II/ HiSeq 2000	> 5x	CHIR 1.0	-	<i>Nature Biotechnolgy</i> 23/12/2012	<i>International Goat Genome Consortium</i>

Fontes: Bai *et al.*, 2012; JGI, 2012; Navegador Genoma UCSC, 2012; Animal Genome, 2012; NCBI, 2013.

“Chrs” = cromossomos sequenciados, integral ou parcialmente, incluindo cromossomos autossomos, sexuais e mitocondriais; “-” = não disponibilidade do produto ou informação.

denominada *Clone-Array Pooled Shotgun Strategy* (CAPSS), em vez de individualmente, como medida de redução de custos (LIU *et al.*, 2009).

Em 2004, mapa com cobertura de 3x foi gerado pela estratégia WGS com pequenos insertos. Em março de 2005, um segundo mapa com cobertura de 6,2x foi gerado com a utilização de WGS de pequenos clones e BACs com extremidades sequenciadas. Esses foram os mapas preliminares BTAU 1.0 e 2.0, respectivamente. Em 2006 o BTAU 3.1 com cobertura de 7.1x foi obtido utilizando WGS com pequenas inserções e BACs sequenciados individualmente ou por CAPSS. Em 2007, o BTAU 4.0 com cobertura de 7.1x foi gerado com as mesmas técnicas utilizadas para criar o mapa BTAU 3.1. Atualmente o mapa BTAU se encontra na versão 4.6.1, lançada em 2012, após as versões 4.2 e 4.5 publicados no ano de 2009. Em abril de 2009, a sequência de alta qualidade do genoma do gado taurino (*Bos primigenius taurus*) foi publicada na revista *Nature*, com cobertura de 7.1x, pelo *Bovine Genome Sequencing and Analysis Consortium* (BCM, 2012).

Os mapas de sequência do genoma bovino UMD 2.0 e UMD 3.1 foram construídos em Baltimore, Maryland, EUA, pela equipe do laboratório de Salzberg, utilizando informações do NCBI reforçadas com modificações realizadas por poderosas ferramentas de montagem e mapeamento. Com 24 milhões de sequências WGS e 11 mil clones BAC, o UMD 2.0 foi construído. Após sete meses da publicação do UMD 2.0, o UMD 3.1 foi lançado, em dezembro de 2009, com cobertura de 9,5x e 2,85Gb sequenciadas, das quais mais de 2,64Gb foram alocadas em algum cromossomo, 30Mb a mais que o UMD 2.0. O UMD 3.1 corresponde ao mesmo conjunto UMD 3.0, lançado em agosto de 2009, com pequena modificação somente para ser incorporado ao GeneBank, sem que houvesse alteração em seu conteúdo (ZIMIN *et al.*, 2009; BAI *et al.*, 2012).

Em dezembro de 2011, o sequenciamento do genoma de bovinos zebuínos (*Bos primigenius indicus*), realizado com a utilização da plataforma de sequenciamento SOLiD e com cobertura de 52x, foi apresentado. Para o sequenciamento foi utilizado DNA de um touro Nelore registrado na Associação Brasileira de Criadores de Zebu (ABCZ). Noventa e sete por cento dos genes codificantes de proteínas dos taurinos foram cobertos (CANAVEZ *et al.*, 2011), e, como esperado, os genomas taurino e zebuíno apresentaram elevada similaridade de nucleotídeos nos cromossomos autossomos e no X, o que não ocorreu em relação ao Y, que na montagem taurina ainda está incompleto.

Em 2012 foram sequenciados os genomas de bovinos pertencentes a três raças também adaptadas ao clima tropical: Brahman; Africander e Tuli Bull. Nessa iniciativa, realizada com a plataforma de sequenciamento Illumina GA-II (Illumina Inc., EUA), cada genoma foi coberto mais que 6x, entre 86,7% e 88,8% das bases de cada genoma foram cobertas por uma ou mais leituras, e o total de 3,56 milhões de SNPs foram identificados.

Também em 2012 foram sequenciados, por meio de tecnologia NGS, sete bovinos provenientes de quatro raças: Angus; Holandesa; Hereford e Nelore. Porém, o objetivo principal do estudo foi descoberta de CNVs (BICKHART *et al.*, 2012).

O Illumina *BovineSNP50 BeadChip v2*, ainda comercialmente disponível, apresenta 54.609 SNPs altamente informativos e uniformemente distribuídos ao longo do genoma bovino. Além deste, a Illumina disponibiliza o chip de alta densidade *BovineHD* (777 mil SNPs) e o de baixa densidade *BovineLD v1.1* (6.912 SNPs), sendo este último ferramenta para imputação e genotipagem de todo o rebanho (ILLUMINA, 2013). A empresa Neogen oferece ainda chips de baixa (10 mil SNPs) e média (80 mil SNPs) densidades específicos para zebuínos e taurinos (NEOGEN, 2013).

### Genoma suíno

Com o objetivo de coordenar o sequenciamento do genoma suíno, o *Swine Genome Sequencing Consortium* (SGSC) foi formado em setembro de 2003 por acadêmicos, representantes de governos e de indústrias. Em 2005 foi publicada a estratégia a ser adotada. A justificativa para sequenciar o genoma dessa espécie reside em sua importância destacada na produção de proteína animal e por ser utilizado como organismo modelo, já que possuem fisiologia muito próxima a dos humanos (SGSC, 2013).

Em novembro de 2009 foi montada a sequência rascunho de alta qualidade do genoma suíno (*Sus scrofa domestica*) com tamanho aproximado de 2,7Gb. (ARCHIBALD *et al.*, 2010; BAI *et al.*, 2012). Clones BAC da biblioteca *Chori-242*, preparados a partir do DNA de uma única fêmea da raça Duroc (Duroc 2-14), foram preferencialmente escolhidos para o sequenciamento. Entretanto, alguns BACs com fragmentos do genoma de outras raças também foram utilizados. O objetivo inicial era determinar a sequência do genoma suíno com cobertura de 3x. Entretanto, as duas extremidades de 768 subclones de cada BAC, com comprimento médio de 707 pb, foram sequenciadas para fornecer cobertura aproximada de 4x. Os dados do WGS foram gerados utilizando-se tanto o método de sequenciamento por terminação de cadeia (ou de Sanger) em sistema capilar no *Korea Livestock Research Institute*, como o Solexa/Illumina no *Beijing Genomics Institute* e no *Wellcome Trust Sanger Institute* (SCHOOK *et al.*, 2005; ARCHIBALD *et al.*, 2010).

Ainda em 2009, Ramos e colaboradores (RAMOS *et al.*, 2009) sequenciaram por meio do *Illumina's Genome Analyzer* o DNA proveniente de quatro grupos de suínos: as raças Duroc, Large White, Pietrain e o javali. A partir desses sequenciamentos, cerca de 272.000 SNPs foram encontrados.

Em setembro de 2011, o Consórcio de Sequenciamento do Genoma Suíno divulgou uma montagem mista do genoma suíno baseado em BACs e WGS, o *Sscrofa10.2*. Este mapa incluiu o conjunto de cromossomos 1 a 18, X e Y. O NCBI *Sus scrofa build 4.1* incluiu a montagem *Sscrofa10.2* e o genoma

mitocondrial completo derivado de um porco da raça Landrace (WELLCOME TRUST SANGER INSTITUTE, 2012).

O banco de dados do genoma suíno continua a receber atualizações consideráveis e novos QTL continuam sendo adicionados ao *PigQTLdatabase*. Atualmente esse banco disponibiliza para suínos o total de 8.315 QTL, citados em 355 publicações, para 622 características de interesse diferentes (NAGRP, 2013). Com o uso desta ferramenta é possível localizar nos cromossomos suínos o local mais provável para os genes responsáveis por características quantitativas importantes para a produção e reprodução.

Mais de 549 mil SNPs foram usados para criar o *PorcineSNP60* (Illumina Inc., EUA), chip comercial com 64.621 SNPs.

### Genoma equino

O projeto para sequenciar o genoma equino teve início em meados da década de 90, mais especificamente em outubro de 1995, quando em Lexington, Kentucky, EUA, nasceu o *Horse Genome Project*. Em janeiro de 1997 a comunidade internacional participante do projeto reuniu-se como parte de uma pesquisa nacional patrocinada pelo *United States Department of Agriculture* (USDA) e em 2005 foi apresentado requerimento junto ao *National Human Genome Research Institute* (NHGRI), mostrando os benefícios da pesquisa do genoma dos equídeos na compreensão de mecanismos do genoma humano (HORSE GENOME PROJECT, 2011).

O sequenciamento e a organização do genoma equino foram realizados pelo *Broad Institute* do Instituto de Tecnologia de Massachussets (MIT) e pela Universidade de Harvard. Foi produzindo um rascunho de alta qualidade com cobertura de 6,8x a partir do DNA genômico de uma égua Puro-sangue Inglês, chamada *Twilight*, de propriedade da Universidade de Cornell em Ithaca, Nova Iorque, EUA (NHI, 2007).

Aproximadamente 300 mil BACs com extremidades sequenciadas foram gerados na Alemanha pela Universidade de Medicina Veterinária em Hanover, e pelo centro Helmholtz para a Pesquisa de Infecções em Braunschweig. Entre janeiro de 2006 e janeiro de 2007 os BACs foram sequenciadas e ordenadas gerando um mapa inicial, o *EquCab1.0*. Em setembro de 2007, com a sequência rascunho de alta qualidade finalizada, foi concluído o mapa virtual *EquCab2.0*. Perto de 2,7 Gb foram sequenciados, ao custo aproximado de 15 bilhões de dólares, mostrando que o genoma equino é menor que o humano e pouco maior que o do cão doméstico (WADE *et al.*, 2009).

Está disponível em [www.broad.mit.edu/mammals/horse/snp](http://www.broad.mit.edu/mammals/horse/snp) o mapa contendo localização de pouco mais de 940 mil SNPs distribuído por todo o genoma equino, exceto no cromossomo Y. Este conteúdo foi compilado pelo *Broad Institute's Equine Genome Sequencing Project*, para o qual foram utilizadas informações dos genomas das raças Árabe, Andaluz, Akhal-teke, Islandesa, Standardbred, Puro Sangue Inglês e Quarto de Milha.

Em 2012 foi publicada a sequência do genoma de uma égua da raça Quarto de Milha. Por meio da

plataforma de sequenciamento Illumina foram gerados 59,6 Gb de DNA sequenciado, com cobertura de 24,7x do genoma. As sequências foram mapeadas em 97% do genoma referência do cavalo Puro-Sangue Inglês. Além disso, foram identificados 3,1 milhões de SNPs, 193 mil InDels e 282 CNVs (DOAN *et al.*, 2012). Utilizando informações do conjunto *EquCab2.0* a Illumina construiu inicialmente o *Equine SNP50 BeadChip* com 54.602 SNPs. Este chip foi descontinuado no início de 2011 em razão da inserção comercial do *Equine SNP70 BeadChip*, este mais denso, contendo mais de 74 mil SNPs uniformemente distribuídos ao longo do genoma (NEOGEN, 2013)

### Genoma ovino

Ao contrario do projeto do genoma bovino que foi financiado pelo USDA, NIH e outros (*Bovine Genome Sequencing and Analysis Consortium*, 2009), o sequenciamento do genoma ovino dispôs de recursos limitados.

Com a possibilidade de obter grandes quantidades de sequências a custo reduzido, as tecnologias NGS pareceram ser a solução. Porém, em razão da incapacidade dessas abordagens de gerarem leituras de grandes segmentos, não foram imediatamente aprovadas pelos membros do consórcio. Contudo, no ano de 2009, o genoma do urso panda foi sequenciado e mapeado unicamente a partir de tecnologia NGS com sequenciamento das extremidades de fragmentos. Desse modo, o consórcio aceitou iniciar o sequenciamento de um mapa referência em 2009, durante *workshop* do *Internacional Sheep Genome Consortium* (ISGC) (ISGC, 2010).

A fase de geração de dados referência do projeto genoma ovino foi iniciada em dois centros de sequenciamento. O *Kunming Institute of Zoology*, Yunnan e o *Benjin Genomics Institute* (BGI), Shenzhen, ambos localizados na China, encomendaram o sequenciamento WGS de uma ovelha da raça Texel, empregando a plataforma da empresa Illumina. Bibliotecas com clones variando de 170pb a 40kb foram construídas e usadas para gerar aproximadamente 220Gb de sequências com extremidades pareadas e cobertura de 75x. Simultaneamente o *ARK-Genomics Center for Comparative and Functional Genomics* do Instituto Roslin (Midlothian, Escócia) utilizando um macho da raça Texel produziu 140GB (45x de cobertura) de sequências com extremidades pareadas a partir de bibliotecas de clones variando de 200 a 500pb. Adicionalmente, sequências foram geradas a partir de biblioteca de clones (3-8kb) com ambas as extremidades sequenciadas, providas do mesmo macho Texel mencionado acima. Uma primeira montagem *de novo* usando sequências derivadas da ovelha Texel for realizada por meio de leituras de pequenas sequências gerando *contigs* e *scaffolds*. Informações do macho Texel foram utilizadas para completar espaços, aumentar o número de SNPs identificados e facilitar a análise do cromossomo Y (ARCHIBALD *et al.*, 2010).

Comparações com informações genômicas de cães, bovinos e humanos foram feitas para refinar o

mapa de radiação híbrida (RH) e construir um mapa físico integrado. Estas comparações permitiram avaliar a qualidade de *contigs* e *scaffolds* (DARYMPLE *et al.*, 2007; WU *et al.*, 2009). Para auxiliar na caracterização dos componentes de transcrição do genoma (genes), o BGI (Shenzhen, China) gerou aproximadamente 15Gb de sequências derivadas de sete diferentes tecidos corporais. Estas informações juntamente com conjuntos de dados provenientes de NGS de outros tecidos foram usadas para predir o número total de genes, possibilitando anotar suas estruturas e posições no genoma ovino (ARCHIBALD *et al.*, 2010).

Informações do sequenciamento de 60 animais de 15 raças diferentes permitiram a descoberta de 18.000 novos SNPs, os quais juntamente com outros 37.000 SNPs obtidos da sequência rascunho (OARv1.0) foram utilizados para a construção do *BeadChip Ovine SNP50* comercializado pela empresa Illumina (ILLUMINA, 2010). O mapa virtual mais recentemente publicado foi o OARv3, gerado após a versão OARv2.1 com a intenção de preencher lacunas, atribuir corretamente maior número de sequências aos cromossomos e fornecer dados de 5.000.000 SNPs identificados. Com os mesmos objetivos do OARv3.1, está em curso a construção o mapa virtual subsequente OARv4.0, previsto para ser publicado no segundo semestre de 2013 (LIVE STOCK GENOMICS, 2012).

### Genoma caprino

O *international goat genome consortium* (IGGC) foi criado em março de 2010 na cidade de Shenzhen, China, e é coordenado por pesquisadores chineses e franceses. A atividade inicial do consórcio concentrou-se em três áreas, a saber: sequenciamento do genoma caprino e a montagem *de novo* e; mapeamento e desenvolvimento de um painel híbrido de radiação; e produção de chip com alta densidade de SNPs. A montagem do genoma da cabra foi realizada no *Benjin Genomics Institute* (BGI) (Ruminant Genome Biology Consortium, 2010).

No final de 2012 foi publicada a sequência rascunho do genoma caprino com tamanho de 2,66 Gb, ou seja, 91% dos 2,92 Gb estimados. O DNA de uma fêmea *Yunnan Black Goat* de três anos de idade foi utilizado para o sequenciamento, realizado por meio da plataforma Illumina® GAI (Illumina Inc., EUA). Foram gerados 191,5Gb de leituras de alta qualidade (cobertura de 65x) vindas de fragmentos correspondentes a duas bibliotecas: uma com sequências curtas (180, 350 e 800pb); e outra com sequências longas (2, 5, 10 e 20kb). Também foram criadas bibliotecas de fosmídeos, com fragmentos de tamanhos entre 400 e 700pb, sequenciados como as bibliotecas de sequências curtas descritas acima. Além disso, foi realizado *RNAseq* a partir de dez tecidos orgânicos diferentes da *Yunnan Black Goat* e do tecido folicular de uma cabra Cashmere do interior da Mongólia, por meio da plataforma Illumina HiSeq 2000 (Illumina Inc., EUA). Com o alinhamento de *contigs* foram construídos *Scaffolds*, e *Scaffolds* aumentados foram gerados a partir de extremidades de fosmídeos sequenciadas. Por não haver mapa físico para caprinos, uma nova tecnologia de mapeamento

óptico denominada *whole-genome mapping* (WGM), desenvolvida pelo grupo *OpGen* (OpGen, EUA), foi utilizada para facilitar a reunião de *Super Scaffolds*. Estes *Super Scaffolds* foram ancorados nos cromossomos com base em informações de regiões sintênicas entre caprinos e bovinos. Por fim, 22.175 genes codificadores foram anotados, sendo que a maior parte foi recuperada a partir de dados do transcriptoma dos dez tecidos utilizados para o sequenciamento do RNA (DONG *et al.*, 2012).

## CONSIDERAÇÕES FINAIS

Nos últimos anos, diferentes métodos de sequenciamento de DNA têm sido desenvolvidos e utilizados amplamente, revolucionando estudos genéticos e genômicos e gerando grandes quantidades de informações para as mais diversas espécies de interesse para o homem. Estas técnicas têm sido aplicadas aos animais domésticos de produção, proporcionando ferramentas importantes para estudos filogenéticos e de melhoramento genético animal, bem como para a descoberta e compreensão de diversos processos biológicos.

O baixo custo e menor tempo despendido pelos novos métodos de sequenciamento proporcionaram agilidade na descoberta de polimorfismos de DNA. Nesse sentido, plataformas para detecção de milhares a centenas de milhares de variações no DNA (chips de SNPs de alta densidade) foram desenvolvidas e estão sendo rotineiramente utilizadas junto a metodologias estatísticas e ferramentas computacionais cada vez mais adequadas, mudando drasticamente as estratégias usualmente empregadas por programas de melhoramento genético animal. Como exemplo pode-se citar o uso destas ferramentas em rebanhos de raças de vacas leiteiras de países europeus e norte-americanos, que utilizam a seleção genômica para gerar precocemente valores genéticos genômicos (EGBV) de alta acurácia pela soma dos efeitos de marcadores espalhados por todo o genoma do animal genotipado. Dessa forma, aumenta-se o ganho genético pela diminuição do intervalo de gerações, o que reduz drasticamente os custos destes programas.

Os estudos de associação ampla do genoma (GWAS), com a utilização dos SNP chips, têm possibilitado a descoberta de genes com grande efeito em algumas características de interesse econômico, permitindo além da utilização da informação destes marcadores na seleção, a compreensão dos mecanismos genéticos que governam tais características. Ainda na busca da compreensão destes mecanismos genéticos, muito se tem feito na área da genômica funcional, principalmente pelo sequenciamento de RNA ou ensaios com *Microarrays*. Estes estudos têm permitido encontrar alterações nos níveis de expressão gênica relacionados à diferentes fenótipos ou tratamentos, além de possibilitar a identificação de variações genéticas que podem ser diretamente responsáveis por parte da variação de uma característica.

Certamente ainda há muito a se avançar, principalmente em relação a grandes transformações advindas de tecnologias de sequenciamento. Com a diminuição ainda maior dos custos e o aumento na velocidade de geração de dados, dentro de poucos anos novas ferramentas tecnológicas estarão disponíveis, trazendo possibilidades cada vez mais promissoras, e impensáveis há alguns anos, para o estudo genético dos animais domésticos.

## REFERÊNCIAS

- ARCHIBALD, A. L.; BOLUND, L.; CHURCHER, C.; FREDHOLM, M.; GROENEN, M. A. M.; HARLIZIUS, B.; LEE, K. T.; MILAN, D.; ROGERS, J.; ROTHSCHILD, M. F.; UENISHI, H.; WANG, J.; SCHOOK, L. B. Pig genome sequence - analysis and publication strategy. **BMC Genomics**, v.11, n.438, p.1-5, 2010.
- BAI, Y.; SARTOR, M.; CAVALCOLI, J. Current status and future perspectives for sequencing livestock genomes. **Journal of Animal Science and Biotechnology**, v.3, n.1, p.3-8, 2012.
- BARRIS, W. HARRISON, B.E. MCWILLIAM, S. BUNCH, R. J. GODDARD M. E. BARENDSE, W. Next generation sequencing of African and Indicine cattle to identify single nucleotide polymorphisms. **Animal Production Science**, v.52, n.3, p.133-142, 2012.
- BAYLOR COLLEGE OF MEDICINE (BCM). **Genome Project Bovine**. Disponível em: <<http://www.hgsc.bcm.tmc.edu/project-species-mBovine.hgsc?pageLocation=Bovine>> acesso em: 21/06/2012.
- BICKHART, D. M.; HOU, Y.; SCHROEDER, S. G. *et al.* Copy number variation of individual cattle genomes using next-generation sequencing. **Genome Research**, v.22, n.4, p.778-790, 2012.
- BOARDMAN, P. E.; SANZ-EZQUERRO, J.; OVERTON, I. M.; BURT, D. W.; BOSCH, E.; FONG, W. T.; TICKLE, C.; BROWN, W. R.; WILSON, S. A.; HUBBARD, S. J. A comprehensive collection of chicken cDNAs. **Current Biology**, v.12, n.22, p.1965-1969, 2002.
- BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM *et al.* The genome sequence of taurine cattle: a window to ruminant biology and evolution. **Science**, v.324, n.5926, p.522-528, 2009.
- CANAVEZ, F. C.; LUCHE, D. D.; STOTHARD, P.; LEITE R. M. K.; SOUSA-CANAVEZ, J. M.; PLASTOW, G.; MEIDANIS, J.; SOUZA, M. A.; FEIJÃO, P.; MOORE, S. S.; CAMARA-LOPES, L. H. Genome Sequence and Assembly of *Bos indicus*. **Journal of Heredity**, doi: 10.1093/jhered/esr153, 2012.
- Disponível em: <<http://jhered.oxfordjournals.org/content/early/2012/02/06/jhered.esr153.full>> acesso em 13/01/2013.
- CARVALHO, M. C. C. G.; SILVA, D. C. G. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciencia Rural**, v.40, n.3, p.735-744, 2010.
- DALRYMPLE, B. P.; KIRKNESS, E. F.; NEFEDOV, M.; MCWILLIAM, S.; RATNAKUMAR, A.; BARRIS, W.; ZHAO, S.; SHETTY, J.; MADDOX, J. F.; O'GRADY, M.; NICHOLAS, F.; CROWFORD, A. M.; SMITH, T.; JONG, P. J.; McEVAN, J.; ODDY, V. H.; COCKETT, N. E. and INTERNATIONAL SHEEP GENOMICS CONSORTIUM. Using comparative genomics to reorder the human genome sequence into a virtual sheep genome. **Genome Biology**, v.8, n.7, p.152.1-152.20, 2007.
- DOAN, R.; COHEN, N. D.; SAWYER, J.; GHAFFARI, N.; JOHNSON, C. D.; DINDOT, S. V. Whole-Genome Sequencing and Genetic Variant Analysis of a Quarter Horse Mare. **BMC Genomics**, v.13, n.78, p.1471-2164, 2012.
- DONG, Y.; XIE, M.; JIANG, Y.; XIAO, N.; DU, X.; ZHANG, W.; TOSSER-KLOPP, G.; WANG, J. YANG, S.; LIANG, J.; CHEN, W.; CHEN, J.; ZENG, P.; HOU, Y.; BIAN, C.; PAN, S.; LI, Y.; LIU, X.; WANG, W.; SERVIN, B.; SAYRE, B.; ZHU, B.; SWEENEY, D.; MOORE, R.; NIE, W.; SHEN, Y.; ZHAO, R.; ZHANG, G.; LI, J.; FARAUT, T.; WOMACK, J.; ZHANG, Y.; KIJAS, J.; COCKETT, N.; XU, X.; ZHAO, S.; WANG, J.; WANG, W. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). **Nature Biotechnology**, v.31, p.135-143, 2013.
- FURLAN, L. R.; FERRAZ, A. L.; BORTOLOSSI, J. C. A genômica funcional no âmbito da produção animal: estado e perspectivas. **Revista Brasileira de Zootecnia**, v.36, suppl.0 2007.
- GENOME CANADA. Bovine Genomics. Disponível em: <<http://www.genomecanada.ca/en/portfolio/research/bovine.aspx>> acesso em: 07/04/2013.
- GROENEN, M. A. M.; MEGENS, H. J.; ZARE, Y.; WARREN, W. C.; HILLIER, L. W.; CROOIJMANS, R. P. M. A.; VEREIJKEN, A.; OKIMOTO, R.; MUIR, W. M.; CHENG, H.H. The development and characterization of a 60K SNP chip for chicken. **BMC genomics**, v.12, p.1-9, 2011.
- HILLIER, L. W. MILLER, BIRNEY W. E. *et al.* Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. **Nature**, v.432, n.7018, p.695-716, 2004.
- HORSE GENOME PROJECT (HGP). 2011. Disponível em: <

- <http://www.uky.edu/Ag/Horsemap/abthgp.html> > acesso em: 21/02/2013
- ILLUMINA. BovineHD Genotyping BeadChip, 2013 Disponível em: <[http://www.illumina.com/documents/products/datasheets/datasheet\\_bovineHD.pdf](http://www.illumina.com/documents/products/datasheets/datasheet_bovineHD.pdf)> acesso em 13/03/2013.
- ILLUMINA. BovineLD v1.1 Genotyping BeadChip, 2013. Disponível em: <[http://res.illumina.com/documents/products/datasheet/datasheet\\_bovineld.pdf](http://res.illumina.com/documents/products/datasheet/datasheet_bovineld.pdf)> acesso em 29/08/2013.
- ILLUMINA. Ovine SNP50 Genotyping BeadChip, 2010. Disponível em: <[http://res.illumina.com/documents/products/datasheet/datasheet\\_ovinesnp50.pdf](http://res.illumina.com/documents/products/datasheet/datasheet_ovinesnp50.pdf)> acesso em: 10/03/2013.
- INTERNATIONAL SHEEP GENOMICS CONSORTIUM; ARCHIBALD, A. L.; COCKETT, N. E.; DALRYMPLE, B. P.; FARAUT, T.; KIJAS, J. W.; MADDOX, J. F.; MCEWAN, J. C.; HUTTON ODDY, V.; RAADSMA, H. W.; WADE, C.; WANG, J.; WANG, W.; XUN, X. The sheep genome reference sequence: a work in progress. **Animal Genetics**, v.41, n.5, p.449–453, 2010.
- INTERNATIONAL SHEEP GENOME CONSORTIUM (ISGC). 2010. Disponível em: <<http://www.sheepmap.org/>> acesso em: 10/12/2012.
- LIU G. E. Applications and case studies of the next-generation sequencing technologies in food, nutrition and agriculture. **Recent Patents on Food, Nutrition & Agriculture**, v.1, n.1, p.75–79, 2009.
- LIVE STOCK GENOMICS (LSG). **OAR v2.0**. 2010. Disponível em: <<http://www.livestockgenomics.csiro.au/sheep/oar2.0.php>> acesso em: 24/01/2013.
- LIVE STOCK GENOMICS (LSG). **OAR v3.1**. 2012. Disponível em: <<http://www.livestockgenomics.csiro.au/sheep/oar3.1.php>> acesso em: 23/11/2012.
- MARGULIES, M.; EGHOLM, M.; ALTMAN, W. E.; ATTIYA, S.; BADER, J. S.; BEMBEN, L. A.; BERKA, J.; BRAVERMAN, M. S.; CHEN, Y. J.; CHEN, Z.; DEWELL, S. B.; DU, L.; FIERRO, J. M.; GOMES, X. V.; GODWIN, B. C.; HE, W.; HELGESEN, S.; HO, C. H.; IRZYK, G. P.; JANDO, S. C.; ALENQUER, M. L. I.; JARVIE, T. P.; JIRAGE, K. B.; KIM, J. B.; KNIGHT, J. R.; LANZA, J. R.; LEAMON, J. H.; LEFKOWITZ, S. M.; LEI, M.; LI, J.; LOHMAN, K. L.; LU, H.; MAKHIJANI, V. B.; MCDADE, K. E.; MCKENNA, M. P.; MYERS, E. W.; NICKERSON, E.; NOBILE, J. R.; PLANT, R.; PUC, B. P.; RONAN, M. T.; ROTH, G. T.; SARKIS, G. J.; SIMONS, J. F.; SIMPSON, J. W.; SRINIVASAN, M.; TARTARO, K. R.; TOMASZ, A.; VOGT, K. A.; VOLKMER, G. A.; WANG, S. H.; WANG, Y.; WEINER, M. P.; YU, P.; BEGLEY, R. F.; ROTHBERG, J. M. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v.437, p.376-380, 2005.
- MAXAM, A. M.; GILBERT, W. A new method for sequencing DNA. **Proceedings of the National Academy of Science of the United States of America**, v.74, n.2, p.560-564, 1977.
- MCEWAN, J. C.; PAYNE, G. M.; BRAUNING, R.; MCCULLOCH, A. F.; STANTON, JO-A.; ODDY, V. H.; NICHOLAS, F. W.; DALRYMPLE, B. P.; INTERNATIONAL SHEEP GENOMICS CONSORTIUM. Using Roche 454 Flx Sequencing For SNP Discovery In Sheep. In: Plant & Animal Genomes XVI Conference. 12 a 16 de Janeiro de 2008. Town & Country Convention Center. San Diego, CA. Disponível em: <[http://www.intl-pag.org/16/abstracts/PAG16\\_P051\\_543.html](http://www.intl-pag.org/16/abstracts/PAG16_P051_543.html)> acesso em: 12/02/2013.
- NATIONAL ANIMAL GENOMES RESEARCH PROGRAM (NAGRP). **QTLdb chicken**. Disponível em: <<http://www.animalgenome.org/cgi-bin/QTLdb/GG/index>> acesso em: 12/01/2013.
- NATIONAL ANIMAL GENOMES RESEARCH PROGRAM (NAGRP). **QTLdb Sus Scrofa**. Disponível em: <<http://www.genome.iastate.edu/cgi-bin/QTLdb/SS/index>> acesso em: 13/02/2013.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION (NCBI). **SNPdb**. Rockville Pike, Bethesda MD, USA. [2012]. Disponível em: <<http://www.ncbi.nlm.nih.gov/projects/SNP/>> acesso em 25/06/2012.
- NATIONAL INSTITUTE OF HEALTHS (NHI). **NHI News**. 2007. Disponível em: <<http://www.nih.gov/news/pr/feb2007/nhgri-07.htm>> acesso em: 23/03/2013.
- NATIONAL HUMAN GENOME RESEARCH INSTITUTE. **NHI news advisory: chicken genome assembled**. 2012. Disponível em: <<http://www.genome.gov/11510730>> acesso em 23/11/2012.
- NEOGEN. Agrigenomic Solutions for Dairy Cattle, 2013. Disponível em: <<http://www.neogen.com/Agrigenomics/Dairy.html>> acesso em: 15/08/2013.
- NEOGEN. Agrigenomic Solutions for Beef Cattle, 2013. Disponível em: <<http://www.neogen.com/Agrigenomics/Beef.html>> acesso em 15/08/2013.
- NEOGEN, Equine SNP70 BeadChip Whole Genome SNP profiling, 2013. Disponível em: <<http://www.neogen.com/Agrigenomics/pdf/Slicks/EquineGenotypingFlyer.pdf>> acesso em 10/07/2013.

- PAREEK, C. S.; SMOCZYNSKI, R.; TRETYN, A. Sequencing technologies and genome sequencing. *Journal of Applied Genetics*, v.52, n.4, p.413–435, 2011.
- RAMOS, A. M.; CROOIJMANS, R. P. M. A.; AFFARA, N. A.; AMARAL, A. J.; ARCHIBALD, A. L.; BEEVER, J. E.; BENDIXEN, C.; CHURCHER, C.; CLARK, R.; DEHAIS, P.; HANSEN, M. S.; HEDEGAARD, HU, J. Z.; KERSTENS, H. H.; LAW, A. S.; MEGENS, H. J.; MILAN, D.; NONNEMAN, D. J.; ROHRER, G. A.; ROTHSCHILD, M. F.; SMITH, T. P. L.; SCHNABEL, R. D.; VAN TASSELL, C. P.; TAYLOR, J. F.; WIEDMANN, R. T.; SCHOOK, L. B.; GROENEN, M. A. M. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. *Plos one*, v.4, n.8, p.1-13, 2009.
- ROBERTS, J. R.; CARNEIRO, M. O.; SCHATZ, M. C. The advantages of SMRT sequencing. *Genome Biology*, v.14, n.8, p1-4, 2013.
- RONAGHI, M.; UHLEN, M.; NYREN, P. A sequencing method based on real-time pyrophosphate. *Science*, v.281, n.5375, p.363–365, 1998.
- RUMINANT GENOME BIOLOGY CONSORTIUM. **The International Goat Genome Consortium**. 2010. Disponível em: <[http://www.ruminants.org/index.php?option=com\\_content&view=article&id=4:the-goat-champion-is-&catid=5&Itemid=2](http://www.ruminants.org/index.php?option=com_content&view=article&id=4:the-goat-champion-is-&catid=5&Itemid=2)> acesso em 17/04/2013.
- SANGER, F.; COULSON, A. R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, v.94, n.3, p.441-448, 1975.
- SCHADT, E. E.; TURNER, S.; KASARSKIS, A. A window into third-generation sequencing. *Humane Molecular Genetics*, v.19, n.2, p.227–240, 2010.
- SCHOOK, L. B.; BEEVER, J. E.; ROGERS, J.; HUMPHRAY, S.; ARCHIBALD, A.; CHARDON, P.; MILAN, D.; ROHRER, G.; EVERSOLE K. Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome. *Comparative and Functional Genomics*, v.6, n.4, p.251-255, 2005.
- SWINE GENOME SEQUENCING CONSORTIUM (SGSC). SGSC mission and goals. Disponível em: <<http://piggenome.org/>> acesso em: 23/11/2012.
- UNIVERSITY OF CALIFORNIA SANTA CRUZ (UCSC). ChickenSNPdb. Santa Cruz - CA 2012. Disponível em: <<http://genome.ucsc.edu/cgi-in/hgGateway?org=Chicken&db=galGal3&hgside=250839081>> acesso em: 03/10/2012.
- VENTER, J. C.; ADAMS, M. D.; MYERS, E.; LI, P. W.; MURAL, R. J.; SUTTON, G. G.; SMITH, H. O.; et al. The sequence of the human genome. *Science*, v.291, n.5507, p.1304–1351, 2001.
- WADE, C. M.; GIULOTTO, E.; SIGURDSSON, S.; ZOLI, M.; GNERRE, S.; IMSLAND, F.; LEAR, T. L.; ADELSON, D. L.; BAILEY, E.; BELLONE, R. R.; BLÖCKER, H.; DIST, O.; EDGAR, R. C.; GARBER, M.; LEEB, T.; MAUCELI, E.; MACLEOD, J. N.; PENEDO, M. C. T.; RAISON, J. M.; SHARPE, T.; VOGEL, L.; ANDERSSON, D. F.; ANTCZAK, T.; BIAGI, M. M.; BINNS, B. P.; CHOWDHARY, S. J.; COLEMAN, J.; DELLA VALLE, G.; FRYC, S.; GUÉRIN, G.; HASEGAWA, T.; HILL, E. W.; JURKA, J.; KIIALAINEN, A.; LINDGREN, G.; LIU, J.; MAGNANI, E.; MICKELSON, J. R.; MURRAY, J.; NERGADZE, S. G.; ONOFRIO, R.; PEDRONI, S.; PIRAS, M. F.; RAUDSEPP, T.; ROCCHI, M.; RØED, K. H.; RYDER, O. A.; SEARLE, S.; SKOW, L.; SWINBURNE, J. E.; SYVÄNEN, A. C.; TOZAKI, T.; VALBERG, S. J.; VAUDIN, M.; WHITE, J. R.; ZODY, M. C.; BROAD INSTITUTE GENOME SEQUENCING PLATFORM; BROAD INSTITUTE WHOLE GENOME ASSEMBLY TEAM; LANDER, E. S.; LINDBLAD-TOH, K. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, v.326, n.5954, p.865-867, 2009.
- WELLCOME TRUST SANGER INSTITUTE (WTSI). Swine Genome Sequencing Project. Disponível em: <<http://www.sanger.ac.uk/resources/downloads/othervertebrates/pig.html>> acesso em 12/12/2012.
- WU, C. H.; JIN, W.; NOMURA, K.; GOLDAMMER, T.; HADFIELD, T.; DALRYMPLE, B. P.; MCWILLIAM, S.; MADDOX, J. F.; COCKETT, N. E. A radiation hybrid comparative map of ovine chromosome 1 aligned to the virtual sheep genome. *Animal Genetics*, v.40, p435–8, 2009.
- ZIMIN, A. V.; DELCHER, A. L.; FLOREA, L.; KELLEY, D. R.; SCHATZ, M. C.; PUIU, D.; HANRAHAN, F.; PERTEA, G.; VAN TASSELL, C. P.; SONSTEGARD, T. S.; MARCAIS, G.; ROBERTS, M. SUBRAMANIAN, P.; YORKE, J. A.; SALZBERG, S. L. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biology*, v.10, R42, 2009.