

Commentary

Open Access

## Unlocking the bovine genome

Ross L Tellam\*<sup>1</sup>, Danielle G Lemay<sup>2</sup>, Curtis P Van Tassell<sup>3</sup>, Harris A Lewin<sup>4</sup>, Kim C Worley<sup>5</sup> and Christine G Elsik<sup>6</sup>

Address: <sup>1</sup>CSIRO Livestock Industries, 306 Carmody Rd, St Lucia 4067, QLD, Australia, <sup>2</sup>Department of Food Science and Technology, University of California, One Shields Ave, Davis, CA 95616, USA, <sup>3</sup>Bovine Functional Genomics laboratory, USDA/ARS, BARC-East, Beltsville, MD, 20705, USA, <sup>4</sup>Institute for Genomic Biology and Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, <sup>5</sup>Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, MS BCM226, Houston, TX 77030, USA and <sup>6</sup>Department of Biology, 406 Reiss, 37th & O Streets NW, Georgetown University, Washington, DC 20057, USA

Email: Ross L Tellam\* - ross.tellam@csiro.au; Danielle G Lemay - dglemay@ucdavis.edu; Curtis P Van Tassell - Curt.VanTassell@ars.usda.gov; Harris A Lewin - h-lewin@uiuc.edu; Kim C Worley - kworley@bcm.edu; Christine G Elsik - ce75@georgetown.edu

\* Corresponding author

Published: 24 April 2009

Received: 20 April 2009

BMC Genomics 2009, 10:193 doi:10.1186/1471-2164-10-193

Accepted: 24 April 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/193>

© 2009 Tellam et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

The draft genome sequence of cattle (*Bos taurus*) has now been analyzed by the Bovine Genome Sequencing and Analysis Consortium and the Bovine HapMap Consortium, which together represent an extensive collaboration involving more than 300 scientists from 25 different countries.

### Commentary

The cattle genome [1,2] was originally selected for sequencing because of the unique biology of ruminants, the importance of ruminants as a major source of protein nutrition for humans, and the evolutionary position of cattle relative to other mammals [3]. Wide-ranging and detailed analyses of the genome sequence and population based variation within the sequence are reported in 22 companion publications, the majority of which appear in the *BioMed Central* group of journals published in April 2009, appropriately, the Year of the Ox. A full list of companion papers and other materials can be found at [4]. These reports provide remarkable insight into mammalian evolution, ruminant biology, and the origins of cattle breeds and impacts of domestication on the species.

At present, there are approximately 3.4 billion domesticated cattle, buffalo, sheep and goats in the world which provide a major source of protein nutrition for 6.6 billion humans [5]. About three quarters of the world's agricultural land produces forage suitable for grazing by these

ruminants, which have an astonishing capacity to efficiently convert low quality plant fibre that is often unsuitable for human use into energy dense fat, muscle and milk. This biological process has long been exploited by humans. Ruminants are also valued for their hides, hair and work capacity. Indeed, the social and cultural evolution of humans, particularly as they adopted an agrarian lifestyle, is intimately intertwined with ruminant domestication, which began in the Near East some 8,000–10,000 years ago [6,7]. The bull and cow have been symbols of human prosperity and political power and, in some cultures, have attained the status of religious icons.

The growth of domesticated ruminant populations, especially cattle, has mirrored the rapid expansion of the human population. More than 800 cattle breeds have been selected by humans since domestication for different economic, social and religious reasons and these represent an important world heritage and a unique scientific resource [8]. While the hand of humans has certainly sculpted the genetic architectures of modern cattle breeds,

it is interesting to note that the human genome has also undergone selective adaptation for tolerance to lactose in response to domestication of ruminants for milk production [9].

The cattle genome sequence fills an evolutionary void among mammalian genome sequences as it is the first complete high coverage genome sequence from the Cetartiodactyl order of eutherian mammals which first appeared ~60 million years ago when grasslands began to cover large parts of the earth's surface. The order includes ~220 species of even-toed ungulate mammals (suborders Ruminantia, Suina and Tylopoda), hippopotamus and 88 cetacean species (whales, dolphins and porpoises). The Cetartiodactyla have adapted to an extraordinary diversity of terrestrial and aquatic environments. Cattle are representative of the Ruminantia, a phylogenetically distant clade to humans and rodents, the two most intensively investigated mammalian species. The evolutionary relationships between these three mammalian groups are well positioned to inform genome similarities and differences in each group and comment on mammalian biology. The evolution of genome organization, particularly the expansion and contraction of paralogous gene families and the molecular bases for economically important adaptive traits of ruminants such as lactation, metabolism, reproduction and disease resistance, are all important scientific issues that can explicitly be addressed by comparative analysis of the cattle genome with other mammalian genomes.

The unique biology of ruminants was a predominant factor influencing the decision to sequence the cattle genome [3]. Ruminants are characterized by their ability to convert ingested complex plant carbohydrates into volatile fatty acids such as acetate, propionate, butyrate, lactate and valerate, which are then used as the major energy source [10]. This remarkable fermentative transformation is performed by a complex community of symbiotic microorganisms present in the rumen. Microbial fermentation of plant material is also the main source of dietary protein and all B vitamins [10]. The immune system of cattle appears to be exquisitely adapted to the extensive microbial communities that populate its gut and exposed epithelial surfaces. For example, up to 40% of peripheral blood lymphocytes of cattle have  $\gamma/\delta$  T-cell receptors as compared to less than 1% in humans and rodents [11]. Thus, cattle are an excellent model species to study the evolution and function of these enigmatic  $\gamma/\delta$  T-cells, which appear to play a major role in immune regulation and suppression. Cattle are also used as models for human reproductive biology and infectious diseases. The latter application is also critical to maintaining a safe, secure and healthy food supply. Assisted reproductive technologies used by humans such as artificial insemina-

tion, estrous synchronization, embryo transfer and in vitro fertilization, were first developed for use in cattle. Furthermore, cattle and other ruminants can be efficiently cloned from adult somatic cells and therefore they represent excellent models to decipher the roles of epigenetic modifications on development and differentiation of stem cells [12].

The assembled 7.1× cattle genome sequence primarily represents DNA from a single inbred female Hereford animal, L1 Dominette [1,13,14]. In addition, limited sequence information was obtained from individuals representing six breeds, which allowed identification of putative SNP, a major resource for the HapMap Project which used 34,470 of these in its analyses. The sequencing, assembly, placement of sequence scaffolds on bovine chromosomes and assessment of assembly quality were assisted by a range of livestock research community resources including radiation hybrid maps, chromosome specific and whole genome linkage maps, physical maps based on BAC end sequence information, extensive EST collections and approximately 10,000 full length cDNA sequences [1,13]. An independent high density linkage map of the bovine genome is also reported [15]. The Bovine Genome Sequencing and Analysis Consortium and the Bovine HapMap Consortium used a variety of genome wide computational approaches, comparative genomics analyses and focused gene specific investigations to highlight the evolutionary history, unique biology and population diversity associated with modern cattle.

What has the initial analysis of the bovine genome sequence revealed? Evolutionary breakpoint regions in mammalian chromosomes, which define the boundaries of conserved syntenic blocks of sequence, are associated with considerable evolutionary plasticity in mammals. Large segmental duplications are over-represented and some types of repetitive sequence elements are selectively enriched or excluded in these regions [1,16]. Moreover, the segmental duplications are enriched for genes whose protein products often directly interface with the external environment e.g. immune proteins and sensory receptors. This observation suggests that these genes represent evolutionary adaptations to the immediate environment of cattle. Comparative analyses also revealed that sequences of bovine proteins are generally more similar to human orthologs than are rodent orthologs. There are also cattle specific changes in the organization of genes involved in digestion, immunity, reproduction and lactation [1].

Milk is exquisitely formulated to suit the differing developmental needs of newborn mammals. Analysis of the genes encoding proteins present in bovine milk and the genes expressed in mammary tissue during the lactation

cycle provide the first comprehensive overview of lactation genomics [17]. The integration of this information with population genetic resources generated in the dairy industry provides an enabling resource for future discovery of genes underpinning complex dairy production traits. These analyses clearly indicate that milk is not just a convenient source of high quality nutrients but it also contains a variety of bioactive proteins which probably play key roles in the antimicrobial defense system of the newborn suckling calf and possibly also in gut maturation.

Several of the companion papers focused on changes in the organization of specific gene families in the bovine lineage and their potential impact on ruminant biology. Compared with humans and rodents, there are significant changes in the organization of several immune gene families [1] exemplified by the serum amyloid A3 [18] and type 1 interferon [19] families, as well as the major histocompatibility complex [20] and T cell receptor B locus [21]. The extensive genomic rearrangements of these and other immune gene families suggest that during evolution of the bovine lineage there has been adaptation to changing infectious disease challenges [1]. There are also alterations in the organization of gene families involved in reproduction compared with the human and rodent genomes [22]. These probably reflect the distinctive reproductive program in cattle, particularly differences in placental structure. In addition, a number of genes are implicated in various aspects of bovine reproductive function and early development [23,24].

Metabolic reconstruction analysis revealed that most genes encoding metabolic enzymes are conserved in mammals however in cattle a few gene losses compared with other mammalian species were identified [25]. These evolutionary changes may be important for understanding metabolic adaptations to the ruminant lifestyle. Examination of defined regions of chromosome 6 revealed a number of novel transcripts [26] suggesting that the full complement of transcriptional activity of the genome is yet to be defined. This is a reminder of a major future challenge – the mapping of the complete repertoire of transcriptional activity across the genome.

The Bovine HapMap Consortium analysed the frequency of over 37,000 SNP in 497 cattle from 19 geographically and biologically diverse cattle breeds [2]. The resulting hapmap is literally a map of the genetic diversity among different cattle populations. Analysis of the hapmap revealed a picture of cattle history in which the ancient population size was relatively large but genetic diversity was dramatically reduced by the processes of domestication and selective breeding [2]. Thus, while the overall number of domesticated cattle has increased in parallel

with the expanding human population, the genetic diversity of those animals has decreased. Nevertheless, genetic diversity among cattle remains greater than among dog breeds and is similar to the diversity among humans [2]. In a companion report, the haplotype structures for several cattle breeds were determined at high resolution [27]. Other reports describe the use of population genetics to investigate the origins and evolution of the different breeds [28,29], effective population sizes [30] and signatures of positive selection associated with domestication [31]. Another report highlights the potential for direct industry application of population genetic information, namely the association of variation of the Integrin Beta 5 gene with bull fertility [32]. A cattle QTL meta-analysis was also performed to identify QTL confirmed in multiple studies using different populations [33]. This analysis provides the scientific community with additional confidence to pursue the identification of causal genetic polymorphisms underlying many complex production traits. Another report describes the development and performance of a higher-density SNP assay that leveraged the information gleaned from the HapMap Consortium efforts [34].

Although these collective studies generated a multitude of discoveries, this is only the beginning of a much larger body of future research that will be catalyzed by the availability of the bovine genome sequence and haplotype map. The genome sequence will be refined, reanalyzed and integrated with biological data in ways we can not yet imagine to examine fundamental questions relating to the linkage between genes and cattle phenotypes, the reasons for the evolutionary success of ruminants, the history of cattle domestication, and the mechanisms of mammalian evolution, to name but a few. The availability of very large well managed cattle populations and sire based breeding systems will accelerate the discovery of genes contributing to complex traits of direct relevance to humans such as energy partitioning, muscle formation, milk production and disease resistance. Undoubtedly, the bovine genome sequence and haplotype map will revolutionize the beef and dairy industries by providing genetic and genomic tools to help address the pivotal livestock issues of the 21<sup>st</sup> century – efficient and sustainable production systems with smaller environmental footprints.

## Acknowledgements

We thank the many people who have contributed to the Bovine Genome Sequencing and Analysis Project and the Bovine HapMap Project.

## References

1. Bovine Genome Sequencing and Analysis Consortium: **The genome sequence of taurine cattle: A window to ruminant biology and evolution.** *Science* 2009, **324**:522-528.
2. Bovine HAPMAP Consortium: **The genetic history of cattle.** *Science* 2009, **324**:529-532.

3. **Bovine genome project** [<http://www.hgsc.bcm.tmc.edu/project-species-m-Bovine.hgsc?pageLocation=Bovine>]
4. **The Bovine Genome Database** [[http://BovineGenome.org/bovine\\_genome\\_consortium/](http://BovineGenome.org/bovine_genome_consortium/)]
5. **FAO: The state of the world's animal genetics resources for food and agriculture.** [<http://www.fao.org/docrep/010/a1250e/a1250e00.htm>].
6. Willham RL: **From husbandry to science: A highly significant facet of our livestock heritage.** *J Anim Sci* 1986, **62**:1742-1758.
7. Diamond J: *Guns, Germs and Steel: the Fates of Human Societies* New York: WW Norton & Company; 1997.
8. Mason IL: *World Dictionary of Livestock Breeds* Wallingford, UK: CAB International; 1998.
9. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P: **Convergent adaptation of human lactase persistence in Africa and Europe.** *Nature Genetics* 2007, **39**:31-40.
10. Delority BA: *Rumen Microbiology* Nottingham, UK: Nottingham University Press, UK; 2003.
11. Hein WR, Mackay CR: **Prominence of gamma delta T cells in the ruminant immune system.** *Immunol Today* 1991, **12**:30-34.
12. Renard JP, Maruotti J, Jounneau A, Vignon X: **Nuclear reprogramming and pluripotency of embryonic cells: Application to the isolation of embryonic stem cells in farm animals.** *Theriogenology* 2007, **68**:S196-205.
13. Liu Y, Qin X, Song X, Jiang H, Shen Y, Durbin KJ, Lien S, Kent MP, Sodeland M, Ren Y, Zhang L, Sodergren W, Havlak P, Worley KC, Weinstock GM, Gibbs RA: **Bos taurus genome assembly.** *BMC Genomics* 2009, **10**:180.
14. MacNeil MD: **Research contributions from 75 years of breeding Line 1 Hereford cattle at Miles City, Montana.** *J Anim Sci* in press.
15. Arias JA, Keehan M, Fisher P, Coppieters W, Spelman R: **A high density linkage map of the bovine genome.** *BMC Genetics* 2009, **10**:18.
16. Kemkemer C, Kohn M, Cooper DN, Froenicke L, Högel J, Hameister H, Kehrer-Sawatzki H: **Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution.** *BMC Evolutionary Biology* 2009, **9**:84.
17. Lemay DG, Lynn DJ, Martin WF, Neville MC, Casey TM, Rincon G, Kriventseva EV, Barris WC, Hinrichs AS, Molenaar AJ, Pollard KS, Maqbool NJ, Singh K, Murney R, Zdobnov EM, Tellam RL, Medrano JF, German JB, Rijnkels M: **The bovine lactation genome: insights into the evolution of mammalian milk.** *Genome Biology* 2009, **10**:R43.
18. Molenaar AJ, Harris DP, Rajan GH, Pearson ML, Callaghan MR, Sommer L, Farr VC, Oden KE, Miles MC, Petrova RS, Good LL, Singh K, McLaren RD, Prosser CG, Kim KS, Wieliczko RJ, Dines MH, Johannessen KM, Grigor MR, Davis SR, Stelwagen K: **The acute-phase protein serum amyloid A3 is expressed in the bovine mammary gland and plays a role in host defence.** *Biomarkers* 2009, **14**:26-37.
19. Walker AM, Roberts RM: **Characterization of the Bovine Type I IFN Locus: Rearrangements, Expansions, and Novel Sub-families.** *BMC Genomics* 2009, **10**:187.
20. Brinkmeyer-Langford C, Childers CP, Fritz K, Gustafson-Seabury A, Cothran M, Raudsepp T, Womack JE, Skow LC: **A high resolution RH map of the bovine major histocompatibility complex.** *BMC Genomics* 2009, **10**:182.
21. Connelley T, Aerts J, Law A, Morrison WI: **Genomic analysis reveals extensive gene duplication within the bovine TRB locus.** *BMC Genomics* 2009, **10**:192.
22. Telugu BPVL, Walker AM, Green JA: **Characterization of the Pregnancy-Associated Glycoprotein gene family in the bovine genome.** *BMC Genomics* 2009, **10**:185.
23. Rodriguez-Osorio N, Wang Z, Kasinathan P, Page GP, Robl JM, Memili E: **Transcriptional reprogramming of gene expression in bovine somatic cell chromatin transfer embryos.** *BMC Genomics* 2009, **10**:190.
24. Ulzun N, Rodriguez-Osorio A, Kaya H, Wang JJ, Parrish VA, Ilyin E, Memili E: **Functional Genomics of HMG3a and SMARCA1 in Early Mammalian Embryogenesis.** *BMC Genomics* 2009, **10**:183.
25. Seo S, Lewin HA: **Reconstruction of Metabolic Pathways for the Cattle Genome.** *BMC Systems Biology* 2009, **3**:33.
26. Weikard R, Goldammer T, Eberlein A, Kühn C: **Novel transcripts discovered by mining genomic DNA from defined regions of bovine chromosome 6.** *BMC Genomics* 2009, **10**:186.
27. Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP, Grefenstette JJ: **High-Resolution Haplotype Block Structure in the Cattle Genome.** *BMC Genetics* 2009, **10**:19.
28. MacEachern S, McEwan J, McCulloch A, Mather A, Savin K, Goddard M: **Molecular evolution of the Bovini tribe (Bovidae, Bovinae): Is there evidence of rapid evolution or reduced selective constraint in Domestic cattle?** *BMC Genomics* 2009, **10**:179.
29. MacEachern S, McEwan J, Goddard M: **Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae).** *BMC Genomics* 2009, **10**:177.
30. MacEachern S, Hayes B, McEwan J, Goddard M: **An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (Bos taurus) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle.** *BMC Genomics* 2009, **10**:181.
31. Barendse W, Harrison BE, Bunch RJ, Thomas BM, Turner LB: **Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits.** *BMC Genomics* 2009, **10**:178.
32. Feugang JM, Kaya A, Page GP, Chen L, Mehta T, Hirani K, Nazareth L, Topper E, Gibbs RA, Memili E: **Two stage genome wide association study identifies Integrin Beta 5 as having a potential role in bull fertility.** *BMC Genomics* 2009, **10**:176.
33. Salih H, Adelson DL: **QTL Global Meta-Analysis.** *BMC Genomics* 2009, **10**:184.
34. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Sonstegard TS, Smith TPL, Moore SS, Van Tassell CP: **Development and characterization of a high density SNP genotyping assay for cattle.** *PLoS ONE* in press.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:

[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

