**Expanding the Definition of Model Animal Organisms**

Joseph D. Romano1 and Indra Neil Sarkar1,2,3

1Department of Microbiology and Molecular Genetics;

2Center for Clinical and Translational Science; and

3Department of Computer Science

University of Vermont

Burlington, VT 05405

To Whom Correspondence Should be Addressed:

Indra Neil Sarkar, PhD, MLIS

Center for Clinical and Translational Science

University of Vermont

89 Beaumont Avenue

Given Courtyard N309

Burlington, VT 05405 USA

Email: [neil.sarkar@uvm.edu](mailto:neil.sarkar@uvm.edu)

Phone: +1-802-656-8283

Facsimile: +1-802-656-4589

**Description of Authors**

**Joseph D. Romano** is an undergraduate honors student majoring in Molecular Genetics at the University of Vermont.

**Indra Neil Sarkar** is the Director of Biomedical Informatics in the Center for Clinical and Translational Science, in addition to being an Assistant Professor in Microbiology and Molecular Genetics as well as Computer Science at the University of Vermont. His research involves the development of biomedical informatics methods across the entire spectrum of life, from molecules to populations.

**Competing Interests**

The authors declare that they have no competing interests.

ABSTRACT

Since the earliest days of of biomedical research, model organisms have been essential for providing standardized platforms to study physiologic and genetic processes associated with human health. The array of currently used model organisms have changed little over the past half century. In light of the recent growth in biological data due in part to advances in biotechnology, it may be appropriate to consider additional animal models that may yield significant insights to human health. This review explores the history of animal model organisms and proposes that there may be an opportunity to identify potentially new model animal organisms. In particular, this review considers the potential of agricultural species as additional animal model organisms, given their careful breeding for agricultural purposes. The review then concludes with a discussion of the types of criteria that may need to be considered to characterize the ‘model-ness’ of a candidate species.

INTRODUCTION

Continued advances in biotechnology have resulted in a range of new comparative biology tools, including many that can be used in the study of human health. Some of these platforms include single-molecule real-time sequencing [1] and publicly accessible annotated genomes of a rapidly growing number of species [2], including humans [3]. The increasing volume of available biological data challenges the biomedical enterprise to develop approaches to transform these data into meaningful information that can lead to novel biological insights.

The life sciences have relied on the use of “model” organisms to observe and study conserved functions across species. Two significant reasons to use model organisms in place of human subjects include: (1) the obstacles presented by the ethics of performing research on human subjects; and (2) the slowness and technical obstacles associated with acquiring results from human subjects in a controlled setting. The second of these is partially an artifact of the lifespan of humans, the low yield in offspring, and the difficulty to adequately inbreed human populations to allow for controlled genetic experiments. Model organisms have enabled researchers to circumvent these kinds of challenges while providing an insightful estimate of human physiology and genetics.

It is current practice to select a well-established model organism for research in comparative genomics. However, it might be worthwhile to consider whether or not a chosen model is best suited for modeling the particular human function of interest. Furthermore, it may be necessary to identify a new model species for research of a particular trait. The selection of such a new species must be subject to scrutiny by the biomedical community, and should be chosen using a carefully selected set of criteria.

Using the history of model organisms in biomedical research as a basis for perspective, this review investigates the efficacy of specific species in fulfilling their intended roles. The discussion then shifts to explore the potential to use other species as model organisms, with a particular emphasis on agricultural species, which may bridge ‘gaps’ left by traditional biomedical model organisms. Finally, this review explores the potential of evaluating new model organisms from a comparative genomics perspective.

MODEL ANIMAL ORGANISMS IN BIOMEDICAL RESEARCH

The first work evaluating the potential utility of nonhuman species as model organisms was in 1900 by C. W. Woodworth, who believed *Drosophila melanogaster* (the fruit fly) could be an ideal animal for studying patterns of genetic inheritance. Consequently, Morgan quantified the inheritance of traits in *D. melanogaster* to define and explain many of the basic concepts of Mendelian genetics [4]. As these processes were researched, the mouse (*Mus musculus*) was also adopted to see if inheritance followed the same laws in mammals as it did in simpler organisms. For decades, *D. melanogaster* and *M. musculus* remained the primary organisms of interest for such research. However, as detailed molecular and chromosomal studies of inheritance became a topic of interest (from the 1930’s through the 1960’s), additional species were identified as model organisms. Simpler life forms – such as phages, *Caenorhabditis elegans* (roundworm), and *Saccharomyces cerevisiae* (baker’s yeast)– were used in the hopes that they would enable the elucidation of basic molecular genetic processes, such as the behavior and mechanics of DNA and the protein products of individual genes. By the early 1980’s, a canonical set of model organisms had been established by practice and by unofficial convention. These organisms – along with the aforementioned *D. melanogaster,* *S. cerevisiae* [5]*, C. elegans*[6]*,* and *M. musculus* [7,8] –included the bacterium *Escherichia coli* [9] and the plant *Arabidopsis thaliana* [10]. A graphical summary of when these models were introduced is shown in Figure 1.

While these species together constitute a diverse representation of the tree of life, many comparative genomic studies in the context of human health are focused on the use of animal model organisms. This is because the genes (and their related products) of animals are accepted to be more similar to their human counterparts than those found in plants, fungi, and bacteria. As such, this review focuses its discussion on animal model organisms, and, more specifically, mammalian models.

Historically, the fundamental criterion desired in the selection of a model organism has been to identify those species whose gene expression is particularly easy to observe. During the time period where analysis of genetic traits primarily relied on direct observation of phenotypic variations in large populations of individuals, the fruit fly was chosen as a model organism because of its ease of observation and high reproductive rate. Mice and rats (*Rattus norvegicus*) were subsequently chosen as mammalian models because they also have a high reproductive rate and are small enough to study in a laboratory setting. It is also important to note that each of these species was well researched prior to their adoption by the scientific community as model organisms. Status as a “model organism” was only achieved after sufficient evidence supported the notion that a given species could be used to study patterns of genetic inheritance and expression, and thus could be used to infer knowledge about human genetics.

An important consideration in comparative genomics for utility in human health studies is the feasibility of using model organisms to model human disease. Both *M. musculus* and *R. norvegicus* have proven to be useful tools in modeling human diseases with a high degree of relevance [11,12]. This can be illustrated through the study of gene homology, where more closely related species share more gene sequence features that may be involved in similar functions. Therefore, species that are evolutionarily close ancestors often share many gene homologues with one another. It is for this reason that mammalian model species generally share more homologies with humans than species elsewhere in the tree of life, and are thus often more useful in comparative genomics studies. For example, many monogenic diseases have clear gene homologues shared between mice and humans, including: achondroplasia [13], polydactyly [14], Li-Fraumeni syndrome [15], cystic fibrosis [16], and sickle cell anemia [17]. The majority of diseases that have been successfully modeled and studied to date are those that are due to monogenic traits. This stands in opposition to diseases influenced by complex networks of genes, known as multigenic traits, where the interaction of numerous polymorphic genes may result in a disease phenotype. Studying this type of genetic influence on disease is significantly more complex [18].

CURRENT STATE OF MODEL ORGANISMS IN BIOMEDICAL RESEARCH

The currently accepted model organisms have been of great use in modeling human genetics and disease expression; however, this set of species may not necessarily reflect the full complement of available knowledge. Amidst some discussion for alternative model organisms (e.g., dogs [19]), the list of model organisms used in biomedical research has generally remained the same for decades. However, the selection of these model organisms predates the biotechnology advances that have enabled the generation of “-omics”-based information that can be leveraged about a wider range of organisms. Thus, the unprecedented growth in –omics data (e.g., as shown in Figure 1, the growth in completed genomes as grown exponentially in recent years) places the scientific community in a position to explore the potential of additional model organisms.

Thanks to advances in biotechnology, the amount of biological sequence data being generated continues to grow at an exponential rate [20]. The National Center for Biotechnology Information’s (NCBI) GenBank database [21], as of its release 191 on August 15th, 2012, contains 156,424,033 distinct nucleic acid sequences [22]. In complement to the data available from NCBI, the genome annotation database Ensembl (operated jointly between EMBL-EBI and WTSI) enumerates, and makes publicly available, detailed reference and sequence information for the complete genomes of 352 species [23]. This wealth of data, while still primarily in the process of being curated and annotated, provides an unprecedented opportunity to use *in silico* approaches to identify and study evolutionary phenomena that span the Tree of Life.

*M. musculus* has a mammalian genome that is comparable to humans [24], however there are several notable differences that pose potential limitations to research. For example, size and life span differ greatly between mice and humans. Additionally, *M. musculus* metabolism and human metabolism differ significantly [25]. *M. musculus* has also been evolved to reproduce at a high rate over a short lifespan, while humans reproduce less frequently over a longer lifespan. Finally, the difference between the *M. musculus* and human immune system necessitated the need to develop transgenic *M. musculus* strains to allow for limited generation of useful immunologic data for clinical application to human health [26,27].

The majority of model species used in biomedical research have been chosen using criteria that were suitable before the genomic data deluge. Thus, while research methods and understanding of biological processes have evolved dramatically over the past few decades based on available model species, there may be an opportunity to advance model organism-based studies that better leverage the available data. To this end, the continued growth in genomic data across a wide spectrum of species provides an opportunity to consider alternative model organisms that can complement knowledge based on long-established model organisms in the study of processes and phenomena associated with human health.

CONSIDERING ALTERNATIVE SPECIES AS GENETIC MODELS

A model organism for biomedical research must possess multiple traits to be useful to researchers. The availability of data prior to adopting the model is the first criterion, required to mitigate the time of adoption by the research community. Second, the proposed model organism needs to be one that is easy to work with, and therefore easy to observe for target traits and genotypes/phenotypes. The third major consideration is whether the proposed model organism’s genetics and molecular physiology are similar enough to those of a human to allow for generation of useful and relatable data. The promise of the future applications of biomedical research increasingly suggest that clinical diagnoses and procedures will enable the leverage of genomic knowledge in the context of medicine [28]. In support of this, there will be an increasing need to identify model systems that can provide additional views of biological processes in light of biomedicine. Such perspectives will only be possible with the expansion of model organisms beyond the current choices (predominantly *M. musculus* and *R. norvegicus*).

Although the molecular mechanics of genetics compose a relatively new field of knowledge, humans have been manipulating the genomes of domesticated animals for tens of thousands of years [29]. Through successive generations of selection for desirable traits (e.g., dairy output), the genomes of these species have been effectively standardized. Inbreeding of cattle, swine, and sheep have thus resulted in genomic standardization necessary for successful agriculture. Animal species of agricultural significance may thus be suitable candidates to consider as potential models.

A small community of researchers has specifically explored the potential of agricultural animals as model organisms [30,31]. While these recommendations are significant in their implications, the focus has been primarily towards the economic prospects of using these species, with particular emphasis on the increased ability to acquire grant funding from agricultural research institutes [32]. This is certainly a useful motivating factor for considering the use of these species, but does not explicitly consider the utility of generating significant data that may promote deeper understanding of phenomena associated with human health.

Nonetheless, there may be biomedical utility to consider. The consideration of bovines (*Bos* Taurus) as a potential model organism can be used to illustrate this claim. In support of its potential as a model organism, the full bovine genome has been sequenced [33,34], and is publicly available via multiple online database browsers. The amount of available data on *B. Taurus*‘s genome is significant; the NCBI provides a wide array of annotated sequences, including two complete genome assemblies [35]. The Map Viewer interface linked to by the NCBI genome pages allows for browsing data on each chromosome with links to other NCBI pages corresponding to individual genes. The other primary resource for browsing annotated cattle genome data is hosted by Georgetown University, and operates under the name of The Bovine Genome Database [36]. While this plays host to the same genome assemblies available on NCBI (Assembly Btau\_4.0 and UMD3.1 Assembly), it offers greater functionality, including more ways to search for specific regions in the genome, and more diverse information regarding genes. The gene pages are organized with respect to gene function and genomic location, with links to related NCBI and Ensembl pages. In addition to the ability to download highly specified sequences, the site links to a dedicated BLAST server with 30 bovine sequence databases. Other features include the ability to download full genome assemblies, various annotation utilities, and composite maps that visualize human and bovine genomes concurrently for ease of comparison.

Using peer-reviewed and indexed literature as a proxy for scientific knowledge, one can compare the amount of relative knowledge in biomedicine versus agriculture using respective literature catalogues: MEDLINE (maintained by the National Library of Medicine) for biomedicine and AGRICOLA (maintained by the National Agriculture Library) for agriculture. A search for the respective subject heading (MeSH for MEDLINE; NALT for AGRICOLA) “Cattle” on journal articles published between 2008 and 2011 returns 30,708 results on MEDLINE and 6,765 results on AGRICOLA. This represents 3,379 unique journals on MEDLINE, and 385 unique journals on AGRICOLA. Although these sets are not mutually disjoint, only 181 journals (5% of the 3,601 total journals) are present in both sets. This suggests that, perhaps contrary to intuition, a significant majority of the scientific data published on *B. taurus* is already in the biomedical realm. To researchers looking for an alternative model organism to *M. musculus*, this may be seen as an advantage – as mentioned earlier, one of the primary considerations for designating a model organism is that it should have a vast array of existing data. Regardless, the nature of comparative genomics enables one to develop computational algorithms for predicting degrees of similarity without necessitating the viewing of individual supporting rationale *in vivo* over a long period of time – these advanced computational methods are reaching the point of refinement where predictions can be made with some degree of certainty [37-39]. Therefore, the adoption of a new genomic model species could theoretically be expedited considerably, especially with availability of multiple annotated genomes along with an array of gene transcripts and protein sequences.

More specific analysis of the *B. taurus* genome is also required to assess the potential of bovines to model human disorders, such as those that can be done by enumerating homologies, specifically identifying commonalities in disease genes and disease gene networks associated with multigenic diseases [40]. Some of this type of work has been explored by previous research [41], but needs to be compiled and re-interpreted from the viewpoint of human diseases or disorders. One example that is of a high degree of interest to both agricultural and clinical researchers is the phenotypic effect of the bovine BRCA1 gene. In humans, women with single nucleotide polymorphisms (SNPs) in either BRCA1/2 genes are 60% likely to develop breast cancer during their lives [42,43]. In *B. taurus*, BRCA1 has been linked to mastitis [44], but has not yet been linked to cancer of the mammary glands (analogous to breast cancer in humans). While the link to mastitis suggests some degree of homology to BRCA1 in humans, it would be of clinical significance to investigate linkages specifically to cancer in *B. taurus*. The protein product of the bovine BRCA1 is notably similar to the protein in humans, and likewise shows a similar expression pattern [45]. Yet mammary gland cancer incidence in *B. taurus*, regardless of carrier status for BRCA1, is lower than breast cancer in humans [46]. Since each genome contains homologous BRCA1 sequences, a genomic approach should be considered to interpret which other factors result in such a low occurrence of this cancer in *B. taurus*. This translational approach to gaining insight into the development of a genetically complex heritable cancer holds interest to the clinical realm, and could be relevant to future cancer research. While this is merely one gene homology of potential significance, further research in bovine genomic physiology may yield more implications regarding complex diseases. Aside from the aforementioned prospect of lucrative grant funding potential, this is one major incentive for specifying new mammalian models [47]

In addition to gene homologies with human orthologues, the metabolism and anatomy of *B. taurus* (or of other mammalian agricultural species, such as *Ovis aries*, *Sus scrofa*, or *Equus ferus*) may be reasons to consider it as a suitable model organism. While model organisms should represent estimations of human systems, a significant factor for selection has historically been ease-of-use in the laboratory. However, while short lifespans and easily observable phenotypes may be essential in early studies of disease processes, there may be merit in advancing subsequent studies to include those that have more physiological characteristics similar to humans. To this end, humans and many domesticated or agricultural species are closer in physical size, reproduce at controlled rates (indicative of “stabilizing” reproduction), and genetically are evolved to live in stable environments (e.g., in well-managed farm facilities) [48]. By contrast, many current model organisms (e.g., mice) are significantly different in size than humans, have a short lifespan, designed to reproduce in very large numbers quickly (“opportunistic” reproductive strategy), and genetically are evolved to live in highly dynamic environments [49].

Logistically, implementing a given agricultural species as an additional model organism should be simpler than other types of non-model animals (e.g., wild animals). Agricultural animals are already domesticated and confined within farms and agricultural facilities, so researchers could gather blood and tissue samples on an as-needed basis, and make observations of the specimens as dictated by the specific research endeavor. Tangentially, the involvement of farmers in enrolling in such a program could yield economic benefits, such as a monetary incentive for providing researchers with samples. The bioethical impact of this process could be designed to be minimal – the life of animal specimens would not need to be compromised, and their living conditions should not need to be altered. Nonetheless, it is still important to underscore that mice and other rodents are essential as initial mammalian model organisms for many reasons, including the ability to control laboratory environments. Mice and other rodents also have certain genetic traits that have been highly selected for through many generations of inbreeding and documentation of phenotypic expression[50]. Agricultural species have undergone a similar process of selection, with the aim of isolating traits with agriculturally beneficial traits, and are similarly well-observed [51]. Agricultural species may prove to supplement model organism based studies before studies involving human subjects. The addition of agricultural animals to the list of model organisms could thus serve to bridge the gap between knowledge inferred by studies on mice and better enable the translation of that knowledge to the clinical realm [52,53]. This is aptly demonstrated by the aforementioned fact that transgenic *M. musculus* strains are required for studies in human immunological processes. While species such as *B. taurus* have not yet demonstrated to have a “more human” immune system, an intermediate species could alleviate current efforts required to make these inferences between mice and humans [54].

IDENTIFYING ‘MODEL-NESS’ OF MAMMALIAN SPECIES

Were the academic community to adopt additional model organisms, one of the first considerations of a given model organism candidate would need to include quantitative measures of suitability. Such measures could then be used to rank the suitability of a potential model organism given the range of available candidates (including those that are established) and researcher preferences. Bioinformatics-based methods have been developed and discussed for the analysis of constellations of parameters that could be adapted for this purpose [55,56]. Conceptually, the development of a ranking scheme would require the creation of a model of analysis that quantifies features that are characteristic of ‘suitable’ model organisms (e.g., genome similarity relative to a disorder of interest, physiological parameters, fecundity values, or costs). The model could then be primed with established model organisms and then expanded to include additional candidate organisms. A system could then be devised that reports a ranking or statistical significance score (e.g., p-value) of ‘model-ness’ relative to parameters that align with a researcher’s preferences. One or more thresholds may be specified to exclude certain candidate species except those that meet predetermined criteria that are deemed essential in new model organisms (e.g., cost, physiological, or fecundity parameters). Such a framework could also potentially accommodate for objective evaluation of the efficacy of currently established model organisms in research. As this review suggests, agricultural species may emerge as early candidates of interest, based upon the reasons that have been discussed above.

Many human diseases of strong clinical interest are influenced by highly complex genetic factors [57-59]. Therefore, there is limited benefit to identifying additional patterns of simple or Mendelian inheritance. Many diseases have underlying ‘gene networks’ that may be organism-specific, which poses potential barriers to the current practice of using model organisms with significant genetic divergence from humans. Therefore agricultural species, such as *B. taurus,* may be more suitable candidates compared to the dominantly used rodent species, due to similarities in disease etiology and presentation. An initial approach to test this theory could be to assemble disease gene networks that are known or hypothesized to exist in species of interest, and then attempt to align them according to relative degrees of similarity to human disease gene networks. Such an approach may also be used to indicate the suitability of these species as test subjects for putative disease therapies. Breast cancer, as mentioned earlier, is one potential target disease; others may include Alzheimer’s disease [60], amyotrophic lateral sclerosis [61], and autism [62]; all of which are examples of diseases known to be influenced by a multitude of genetic polymorphisms.

CONCLUSION

A review of the use of model organisms in genetic research shows that their selection was based primarily upon ease of observation and data analysis, yet they have remained virtually unchanged for more than three decades. Moving into the era of comparative genomics, it has become essential to reevaluate standard model organisms and possibly identify new species that may provide more meaningful insights to human biology. Domestic animal species would satisfy many of the characteristics desired in ideal model organisms, due to selective inbreeding and careful documentation of traits over numerous generations. Still more specifically, agricultural animals are of particular interest to the field, and preliminary data suggest they may be of utility to solving the mechanics of complex diseases of high interest to the clinical realm.

KEY POINTS

- Biomedical research depends on the use of model organisms to enable detailed study of physiologic and genomic phenomena associated with human disease.

- Historically, model organisms represent those that are both well studied and easy to observe for traits.

- The established set of model organisms has remained virtually static for decades.

- Agricultural species, such as *Bos taurus*, may merit consideration as new model animal organisms, as they share many of the characteristics desired in an ideal animal model.

- There is a need for developing empirical “model-ness” tests to quantify the suitability of potential model organisms.

**FIGURE LEGEND**

**Figure 1: Temporal history of canonical model organisms**. The introduction of commonly used model organisms is shown along a timescale along with other significant biology milestones. Also shown is the number of completed genomes (according to the Genomes OnLine Database [GOLD]) from the mid-1990’s to present.

1. Schadt, EE, Turner, S, Kasarskis, A. A window into third-generation sequencing. Hum. Mol. Genet. 2010; 19:R227–40

2. Genome List - NCBI.

3. Lander, ES, Linton, LM, Birren, B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409:860–921

4. Thomas H. Morgan - Biography.

5. Botstein, D, Fink, GR. Yeast: an experimental organism for modern biology. Science 1988; 240:1439–1443

6. Brenner, S. In the beginning was the worm ... Genetics 2009; 182:413–415

7. Paigen, K. One hundred years of mouse genetics: an intellectual history. I. The classical period (1902-1980). Genetics 2003; 163:1–7

8. Paigen, K. One hundred years of mouse genetics: an intellectual history. II. The molecular revolution (1981-2002). Genetics 2003; 163:1227–1235

9. Müller, B, Grossniklaus, U. Model organisms--A historical perspective. J Proteomics 2010; 73:2054–2063

10. Steinitz-Sears, LM. Chromosome Studies in Arabidopsis Thaliana. Genetics 1963; 48:483–490

11. Bedell, MA, Jenkins, NA, Copeland, NG. Mouse models of human disease. Part I: techniques and resources for genetic analysis in mice. Genes Dev. 1997; 11:1–10

12. Bedell, MA, Largaespada, DA, Jenkins, NA, et al. Mouse models of human disease. Part II: recent progress and future directions. Genes Dev. 1997; 11:11–43

13. Velinov, M, Slaugenhaupt, SA, Stoilov, I, et al. The gene for achondroplasia maps to the telomeric region of chromosome 4p. Nat. Genet. 1994; 6:314–317

14. Radhakrishna, U, Blouin, JL, Mehenni, H, et al. Mapping one form of autosomal dominant postaxial polydactyly type A to chromosome 7p15-q11.23 by linkage analysis. Am. J. Hum. Genet. 1997; 60:597–604

15. Malkin, D, Li, FP, Strong, LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990; 250:1233–1238

16. Kerem, B, Rommens, JM, Buchanan, JA, et al. Identification of the cystic fibrosis gene: genetic analysis. Science 1989; 245:1073–1080

17. Rees, DC, Williams, TN, Gladwin, MT. Sickle-cell disease. Lancet 2010; 376:2018–2031

18. Glazier, AM, Nadeau, JH, Aitman, TJ. Finding genes that underlie complex traits. Science 2002; 298:2345–2349

19. Ostrander, EA. Franklin H. Epstein Lecture. Both ends of the leash--the human links to good dogs with bad genes. N. Engl. J. Med. 2012; 367:636–646

20. GenBank Statistics (ca. 2008).

21. Benson, DA, Karsch-Mizrachi, I, Clark, K, et al. GenBank. Nucleic Acids Res. 2012; 40:D48–53

22. NCBI-GenBank Flat File Release 189.0.

23. Ensembl Genomes.

24. Waterston, RH, Lindblad-Toh, K, Birney, E. Initial sequencing and comparative analysis of the mouse genome. 2002; 1–43

25. Demetrius, L. Of mice and men. When it comes to studying ageing and the means to slow it down, mice are not just small humans. EMBO Rep. 2005; 6 Spec No:S39–44

26. Taneja, V, David, CS. HLA transgenic mice as humanized mouse models of disease and immunity. J. Clin. Invest. 1998; 101:921–926

27. Kotturi, MF, Assarsson, E, Peters, B, et al. Of mice and humans: how good are HLA transgenic mice as a model of human immune responses? Immunome Res 2009; 5:3

28. Khoury, MJ, Gwinn, M, Yoon, PW, et al. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? Genet. Med. 2007; 9:665–674

29. Zeder, MA. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. Proc. Natl. Acad. Sci. U.S.A. 2008; 105:11597–11604

30. Roberts, RM. The Place of Farm Animal Species in the New Genomics World of Reproductive Biology. Biology of Reproduction 2001; 409–417

31. Ireland, JJ, Roberts, RM, Palmer, GH, et al. A commentary on domestic animals as dual-purpose models that benefit agricultural and biomedical research. Journal of Animal Science 2008; 86:2797–2805

32. Reynolds, LP, Ireland, JJ, Caton, JS, et al. Commentary on Domestic Animals in Agricultural and Biomedical Research: An Endangered Enterprise. Journal of Nutrition 2009; 139:427–428

33. Gibbs, R, Weinstock, G, Kappes, S, et al. Bovine Genomic Sequencing Initiative. 2002; 1–12

34. Bovine Genome Sequencing and Analysis Consortium, Elsik, CG, Tellam, RL, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 2009; 324:522–528

35. Bos Taurus (ID 82) - Genome - NCBI.

36. Welcome to the Bovine Genome Database | BovineGenome.org.

37. Karp, PD, Paley, SM, Krummenacker, M, et al. Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. Briefings in Bioinformatics 2010; 11:40–79

38. Lee, TJ, Paulsen, I, Karp, P. Annotation-based inference of transporter function. Bioinformatics 2008; 24:i259–67

39. Romero, PR, Karp, PD. Using functional and organizational information to improve genome-wide computational prediction of transcription units on pathway-genome databases. Bioinformatics 2004; 20:709–717

40. Sarkar, IN. A vector space model approach to identify genetically related diseases. Journal of the American Medical Informatics Association 2012; 19:249–254

41. Lopez-Corrales, N, Sonstegard, T, Smith, T. Comparative gene mapping: cytogenetic localization of PROC, EN1, ALPI, TNP1, and IL1B in cattle and sheep reveals a conserved rearrangement relative to the human genome. Cytogenet Cell Genet 1998; 83:35–38

42. SEER Cancer Statistics Review, 1975-2005. 2007;

43. Chen, S, Parmigiani, G. Meta-analysis of BRCA1 and BRCA2 penetrance. J. Clin. Oncol. 2007; 25:1329–1333

44. Yuan, Z, Chu, G, Dan, Y, et al. BRCA1: a new candidate gene for bovine mastitis and its association analysis between single nucleotide polymorphisms and milk somatic cell score. Mol. Biol. Rep. 2012; 39:6625–6631

45. Krum, SA, Womack, JE, Lane, TF. Bovine BRCA1 shows classic responses to genotoxic stress but low in vitro transcriptional activation activity. Oncogene 2003; 22:6032–6044

46. Dallenbach, FD. [Etiology and epidemiology of breast carcinoma]. Fortschr. Med. 1975; 93:1323–1325

47. O'Brien, SJ. The Promise of Comparative Genomics in Mammals. Science 1999; 286:458–481

48. Fadiel, A, Anidi, I, Eichenbaum, KD. Farm animal genomics and informatics: an update. Nucleic Acids Res. 2005; 33:6308–6318

49. Pianka, E. On r-and K-selection. The American Naturalist 1970;

50. JAX Mice Database - Inbred Strain.

51. G2911, Inbreeding: Its Meaning, Uses and Effects on Farm Animals.

52. Attfield, KE, Dendrou, CA, Fugger, L. Bridging the gap from genetic association to functional understanding: the next generation of mouse models of multiple sclerosis. Immunol. Rev. 2012; 248:10–22

53. Schmouth, J-F, Bonaguro, RJ, Corso-Diaz, X, et al. Modelling human regulatory variation in mouse: finding the function in genome-wide association studies and whole-genome sequencing. PLoS Genet. 2012; 8:e1002544

54. Watanabe, Y, Takahashi, T, Okajima, A, et al. The analysis of the functions of human B and T cells in humanized NOD/shi-scid/gammac(null) (NOG) mice (hu-HSC NOG mice). Int. Immunol. 2009; 21:843–858

55. Sreenivasaiah, PK, Kim, DH. Current trends and new challenges of databases and web applications for systems driven biological research. Front Physiol 2010; 1:147

56. Mazza, T. Editorial: Accelerating systems biology. Briefings in Bioinformatics 2010; 11:267–269

57. Davies, JL, Kawaguchi, Y, Bennett, ST, et al. A genome-wide search for human type 1 diabetes susceptibility genes. Nature 1994; 371:130–136

58. Berrettini, WH. Molecular linkage studies of bipolar disorder. Dialogues Clin Neurosci 1999; 1:12–21

59. Burbach, JPH, van der Zwaag, B. Contact in the genetics of autism and schizophrenia. Trends Neurosci. 2009; 32:69–72

60. OMIM Entry - #104300 - ALZHEIMER DISEASE; AD.

61. OMIM Entry - #105400 - AMYOTROPHIC LATERAL SCLEROSIS 1; ALS1.

62. OMIM Entry - %209850 - AUTISM.