



**UNIVERSITY**  
*of*  
**GLASGOW**

Hunt, E. and Hanlon, N. (2004) SyntenyVista. In, *Third Nordic Conference on Human-Computer Interaction, October 23-27 2004*, pages pp. 455-456, Tampere, Finland.

<http://eprints.gla.ac.uk/3509/>

# SyntenyVista

Ela Hunt

Department of Computing Science  
University of Glasgow  
Glasgow G12 8QQ, UK  
ela@dcs.gla.ac.uk

Neil Hanlon

BHF Glasgow Cardiovascular Research Centre  
University of Glasgow  
Glasgow G12 8QQ, UK  
hanlonn@dcs.gla.ac.uk

## ABSTRACT

Visualization of genome comparisons is an important research tool in biology, medicine and agricultural research. We present a new visualization system, SyntenyVista, which allows for interactive exploration of genome comparisons. It incorporates a novel feature called *cartoon scaling* designed to improve the user's ability to understand gene relationships. SyntenyVista is being deployed in the context of e-Science where it will provide an entry point to integrated data on genomes, animal models of disease, and disease genes.

## Author Keywords

Genomics, visualization, cartoon scaling, synteny.

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

SyntenyVista software supports the biologist wanting to use the results of a comparative analysis of two or more genomes [2,5,8]. Such analyses produce large quantities of data that are currently not easy to understand and could benefit from improved visualization support. SyntenyVista shows relationships between genes, groups of genes, chromosomes, and mapped genetic features across the human, mouse and rat genomes. It presents accurate up to date information in a novel, visually attractive metaphor.

## UNDERLYING DATA

The software automatically downloads genomic information from Ensembl [2,4]. We focus on three species: the human, the mouse and the rat. We show an overview of the genomes by allowing the user to see all chromosomes in the three species, see Figure 1. Since the Ensembl database focuses on genes, and for each of the

genes holds positional information, we visualize the genes along each of the chromosomes in the detailed view of each chromosome. This facility is provided by Ensembl, and also by other databases which hold similar data [5,8]. The novelty of our approach is that we go beyond the static portrayal of chromosomes and genes, and focus on the similarities between groups of similar genes in three species. In this context *conserved syntenly* is defined as a group of genes on one chromosome corresponding to a group of similar genes on another chromosome from a different species.

## SYNTENY LAYOUT

We show chromosomes in a vertical orientation which allows us to display all gene names positioned beside the genes, with additional features positioned along the chromosome axis. Other layout considerations included the combined view of overview and detail [3], showing gene relationships as lines, and the use of smooth zooming [1]. We developed a new distortion technique, cartoon scaling, see Figure 2, to enable the user to see more genes and gene comparisons than one would normally see in the physical representation of the genome which preserves gene lengths and distances. Cartoon scaling is a data distortion technique which condenses the data space along chromosome length [4], and at the same time makes sure that labels are easily associated with genes. Genes are represented as being equidistant and of equal length, and zooming is applied to change feature resolution.

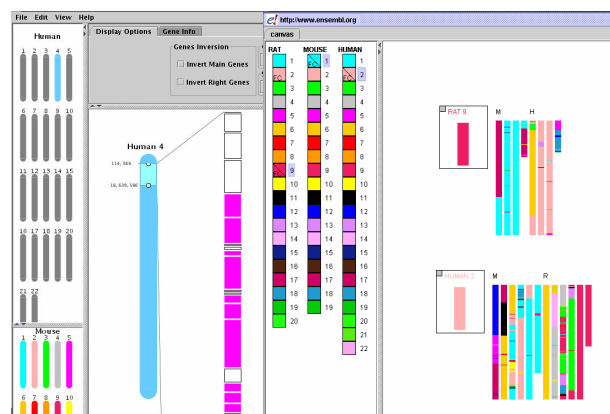
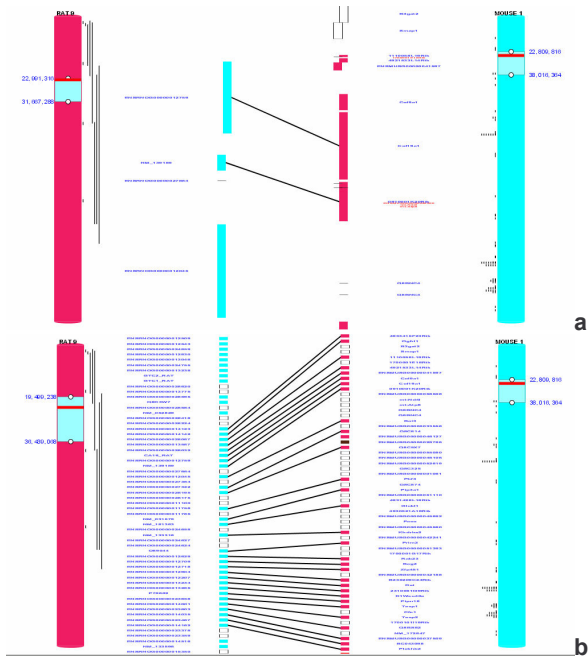


Figure 1. Alternative representations of genome overviews in SyntenyVista.



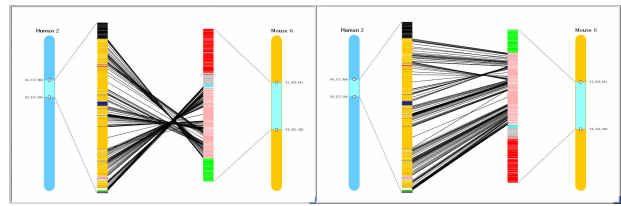
**Figure 2. Genes represented according to physical scale (a) and using cartoon scaling (b) which shows them as being of equal size and equidistant.**

#### USER INTERACTION

The user can move in the space of species, chromosomes and genes, by clicking on the objects, zooming, and panning. On mouse-over additional information can be displayed in a text panel. We support both mouse-driven and keyboard-driven interaction and data selection. Panning and zooming improve significantly on the static presentations which use server-generated graphics [2,4], and the presentation of synteny extends the visualization techniques described in [6,8]. Chromosome inversion can be used to disambiguate an area where too many crossing lines representing gene relationships obstruct the view, see Figure 3.

#### CONCLUSION

SyntenyVista is a novel representation of gene relationships between species. It offers a high degree of interactivity, and combines the overview and detail on the same screen area, which ensures a frame of reference for the user. It improves significantly on the existing representations of synteny [2,5,8] and introduces cartoon scaling as a new distortion technique which can be switched on and off by the user, depending on the viewing context. The chromosome inversion facility allows the user to compare alternative views and perspectives available for the three genomes. In fact, the visualization technique can be used to compare any pair of similar genomes, provided data are available on how genes in those genomes are related. We are about to start



**Figure 3. Chromosome inversion improves the representation of gene relationships. The chromosome shown on the right is shown before and after inversion.**

using SyntenyVista to show both data acquired from Ensembl, and data produced in the lab, alongside novel information derived using computational methods.

#### ACKNOWLEDGMENTS

This research was funded by: Medical Research Council, British Heart Foundation, Wellcome Trust, Department of Trade and Industry, and Engineering and Physical Sciences Research Council, UK. We thank M. Ferrier of the Bridges e-Science project for the connection to Ensembl. A video of our prototype is available at [7].

#### REFERENCES

1. Bederson, B.B., Hollan, J.D., Perlin, K., Meyer, J., Bacon, D. and Furnas, G. Pad++: A Zoomable Graphical Sketchpad For Exploring Alternate Interface Physics. *J. Vis. Lang. Comput.*, 7 (1996), 3-32.
2. Birney, E., Andrews, T.D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., et al. An overview of Ensembl. *Genome Res.* 14(5) (2004), 925-8.
3. Card, S.K., Mackinley, J.D. and Shneiderman, B., *Readings in information visualization: using vision to think.* Morgan Kaufmann, 1999.
4. Hunt, E., Hanlon, N., Leader, D. and Dominiczak, A.F., The visual language of synteny. *OMICS* 8(4) (2004), to appear.
5. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., et al. The Human Genome Browser at UCSC. *Genome Res.* 12(6) (2002), 996-1006.
6. Pook, S., Vaysseix, G. and Barillot, E. Zomit: biological data visualization and browsing. *Bioinformatics*, 14(9) (1998), 807-14.
7. SyntenyVista. <http://www.dcs.gla.ac.uk/~ela/Synteny>.
8. Twigger, S.N., Nie, J., Routti, V., Yu, J., Chen, D., Li, D., et al. Integrative genomics: in silico coupling of rat physiology and complex traits with mouse and human data. *Genome Res.* 14(4) (2004), 651-60.