

G3DV: A New 3D Genome Browser and Experimental Data Viewer

Michael J. Moorhouse¹, Hubertus J. F. M. M. Eussen², Tobias A. Knoch³, Michael Lesnussa³
Albert D.M.E. Osterhaus¹, Frank G. Grosveld³, Annelies J. E. M. M. de Klein²

¹Department of Virology; ²Department of Clinical Genetics; ³Department of Cell Biology Erasmus MC, PostBus 2040, 3000 CA ROTTERDAM, The Netherlands.

Abstract

With the increasing complexity of biological studies and sophistication of molecular diagnostic tests, new visualization methods are needed to display experimental results in an appropriate genomic context so that they can be properly understood.

We have developed a new type of genome browser, the GLOBE 3D Viewer (G3DV) that presents genomics experimental data alongside public annotation in an interactive 3D virtual space to give a more 'holistic' representation of the genome.

Here, we demonstrate its utility for the selection of BACs in a repetitive element-rich region of the human genome that is suitable for use in medical diagnostic tests using classical FISH techniques.

Introduction

In modern biological and medical investigations there is an increasing need to combine data from different sources in a genomics context to aid the interpretation of such investigations.

While other genome browsers exist such as the popular Ensembl and UCSC Genome Browsers (see: <http://genome.ucsc.edu/> & <http://www.ensembl.org>), they focus on the presentation on historical data and the output of their associated annotation pipelines. Other specific analysis tools often concentrate on the display of results of their specific analysis.

Here we describe a system that uses the metaphor of a '3D Virtual World' for genomics data and results of genomics experiments.

On standard commodity computer systems it runs fast enough to display 250,000+ genomics 'features' from historical annotation or experimental results.

As it uses the classical "feature defined by base-pair: start & end" paradigm it is compatible with many existing annotation types.

Results

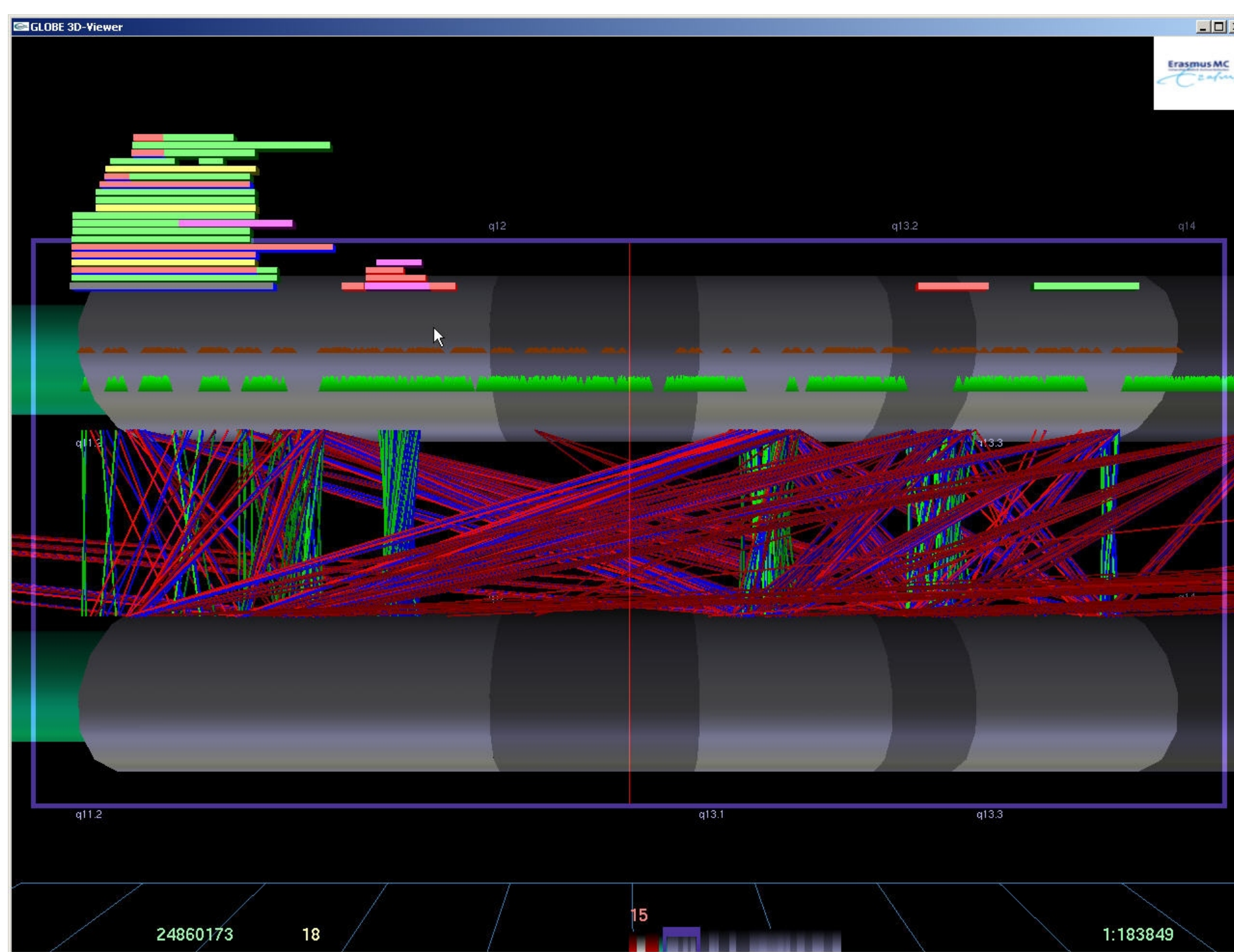


Figure 1: Comparison of the Affymetrix 500K GeneChip (tm) Mapping array (green triangles) and the Agilent 105K Human Oligo Array (brown triangles) with HapMap Data (Colored Bars, see Redon et al, Nature 2006, PMID: 171228750) in the context of repetitive sequence regions (red, green and blue lines).

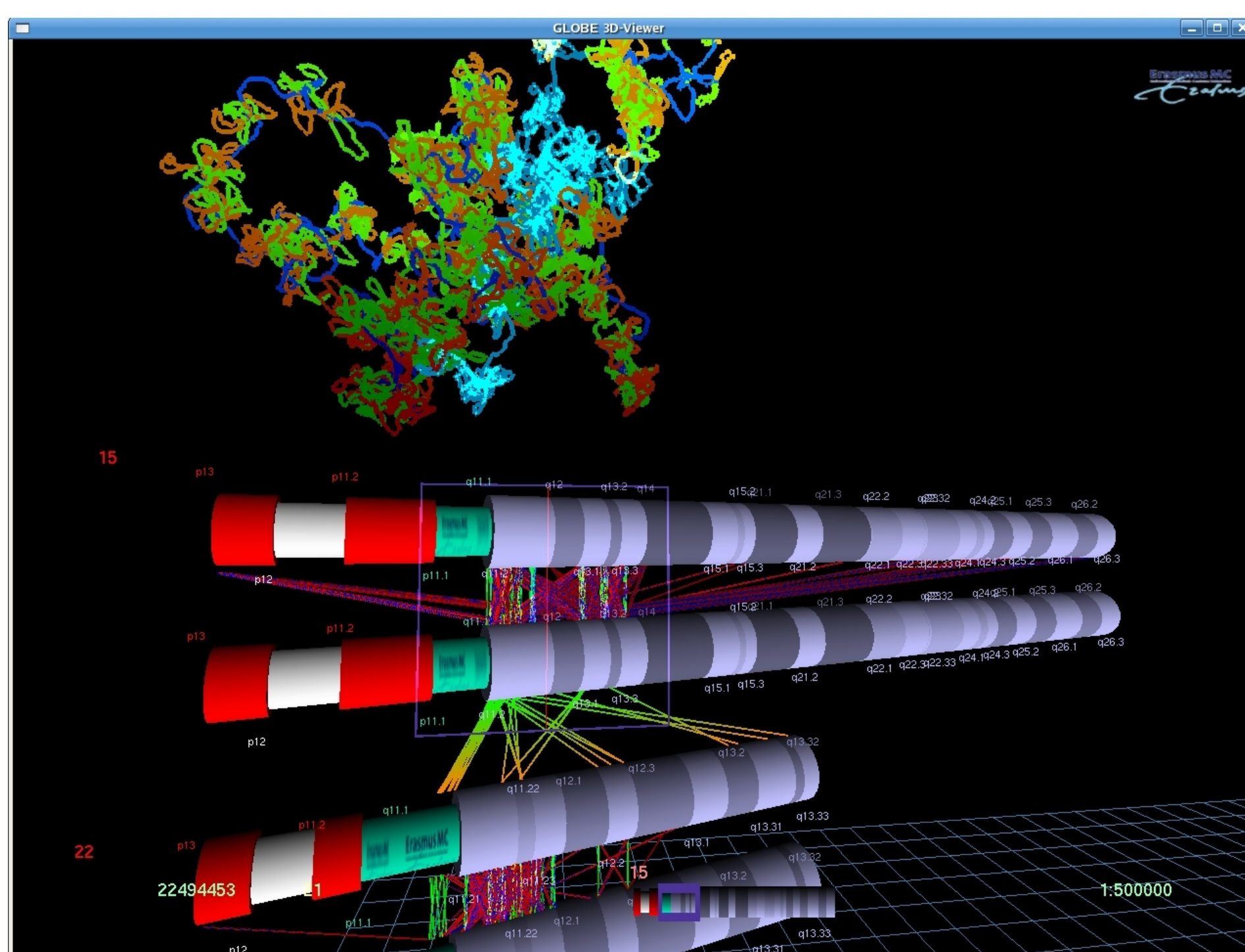


Figure 2: Demonstration of the sequence similarities between Human Chromosomes 15 & 22 (yellow/green lines); other colors as Figure 1. The selected region (purple box) is highlighted in a simulated 3D Chromosome structure (multiple colors, highlighted region shaded in cyan).

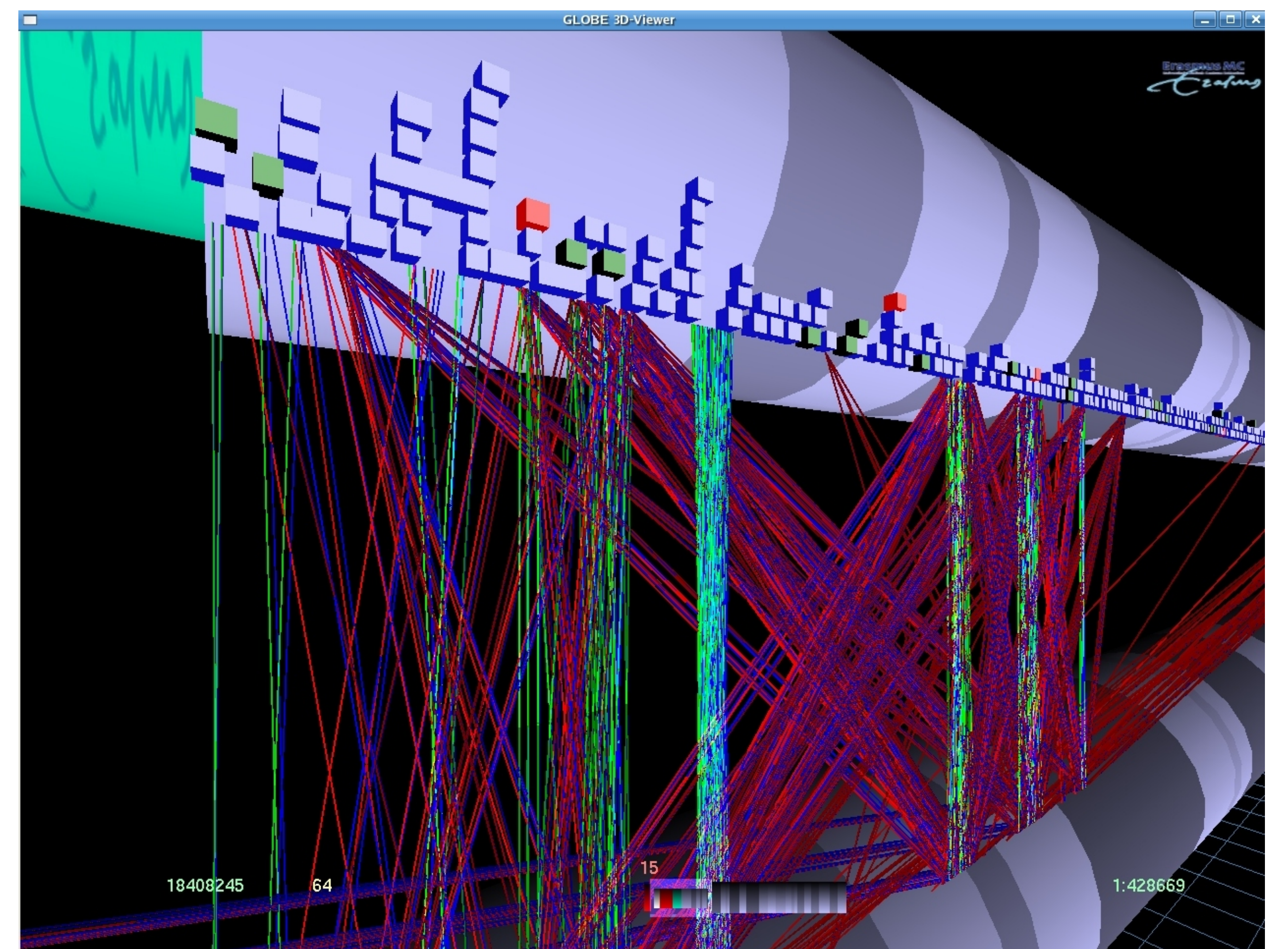


Figure 3: Location of BAC Clones / FISH Probes colored by validation status (green = confirmed, unique signal; red = aberrant signals observed; grey/blue = untested at this time); other colors as Figure 1. Aberrant Probes often - typically those giving multiple FISH signals in negative tests - are often associated with regions of high similarity, i.e. Duplicons. Probes can be selected and their identifiers exported for use.

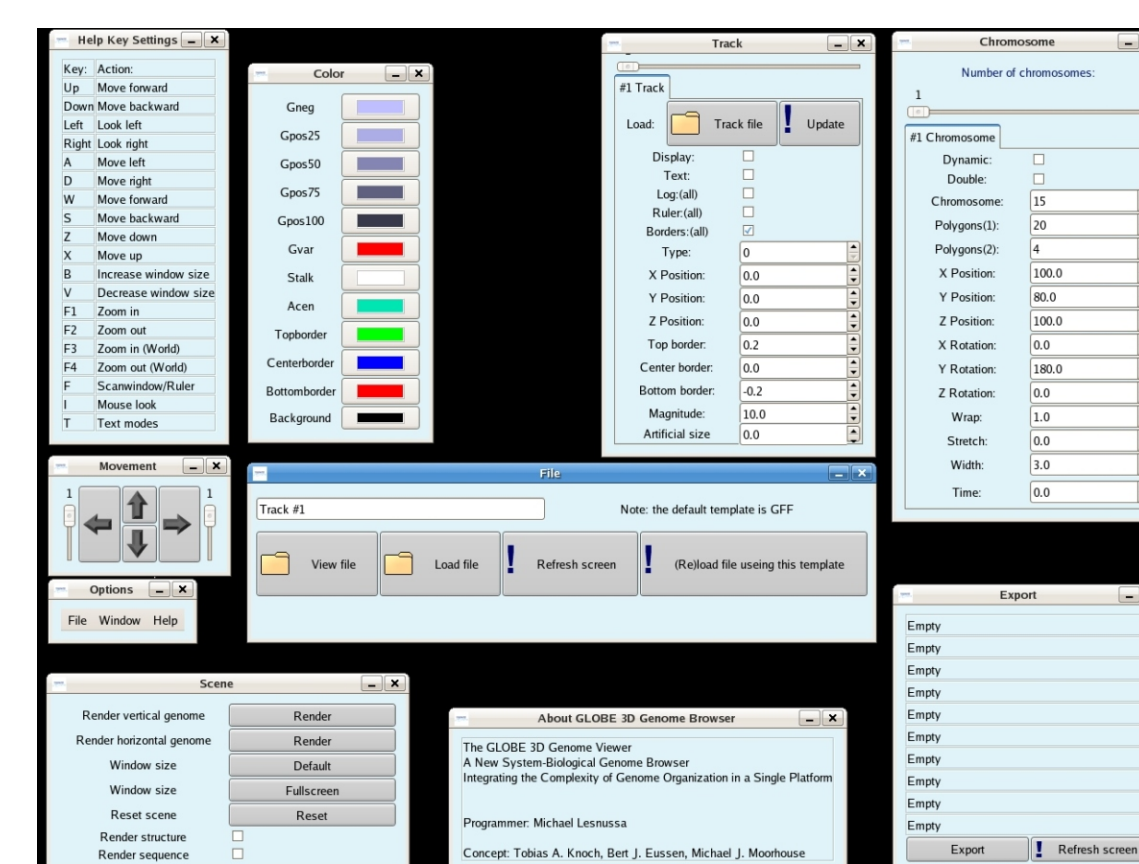


Figure 4: The G3DV Graphical User Interface showing the range of options implemented. The detached panels are ideal for "multi-headed" PCs but are also convenient on single-headed displays as they can be opened / closed individually to reduce screen clutter.

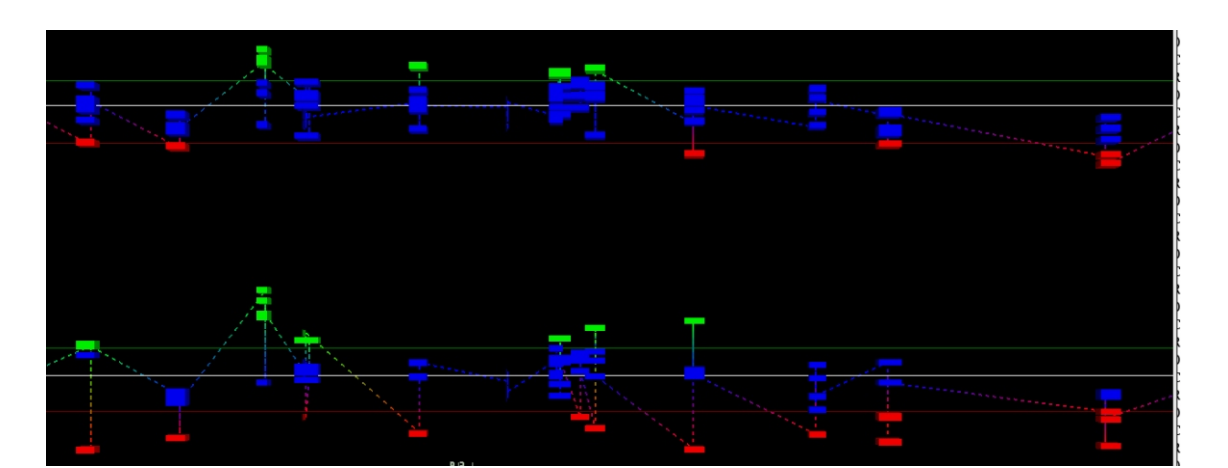


Figure 5: Example of ratio data results measured using multiple arrays. Markers can be colored by thresholds as demonstrated here.

Key Features

- ★ Uses the metaphor of '3D Virtual World' that is intuitive to navigate around.
- ★ "Mirroring" of chromosomes allows clear representation of even complex association between features.
- ★ Can present many different types of existing annotation: genes, proteins, SNPs, BACs, deletion & medically interesting regions.
- ★ Results of 'ratio' experiments (transcriptomics, or copy-number changes from SNP or aCGH) can be presented alongside existing annotation in different forms: ratio plots, color thresholds.
- ★ Runs on standard PC Hardware (Microsoft Windows(tm) or GNU/Linux) sufficient to display 250,000+ genomics features.

G3DV: A New 3D Genome Browser and Experimental Data Viewer

Moorhouse, M. J., Eussen, H. F. M. M., Lesnussa, M., Osterhaus, A. D. M. E.,
Grosveld, F. G., de Klein, A. J. E. M. M. & **Knoch, T. A.**

5th European Conference on Computational Biology, Eilat, Israel, 21st - 24th January, 2006.

Abstract

Genomes are tremendous co-evolutionary holistic systems for molecular storage, processing and fabrication of information. Their system-biological complexity remains, however, still largely mysterious, despite immense sequencing achievements and huge advances in the understanding of the general sequential, three-dimensional and regulatory organization. Here, we present the GLOBE 3D Genome Platform a completely novel grid based virtual “paper” tool and in fact the first system-biological genome browser integrating the holistic complexity of genomes in a single easy comprehensible platform: Based on a detailed study of biophysical and IT requirements, every architectural level from sequence to morphology of one or several genomes can be approached in a real and in a symbolic representation simultaneously and navigated by continuous scale-free zooming within a unique three-dimensional OpenGL and grid driven environment. In principle an unlimited number of multi-dimensional data sets can be visualized, customized in terms of arrangement, shape, colour, and texture etc. as well as accessed and annotated individually or in groups using internal or external data bases/facilities. Any information can be searched and correlated by importing or calculating simple relations in real-time using grid resources. A general correlation and application platform for more complex correlative analysis and a front-end for system-biological simulations both using again the huge capabilities of grid infrastructures is currently under development. Hence, the GLOBE 3D Genome Platform is an example of a grid based approach towards a virtual desktop for genomic work combining the three fundamental distributed resources: i) visual data representation, ii) data access and management, and iii) data analysis and creation. Thus, the GLOBE 3D Genome Platform is the novel system-biology oriented information system urgently needed to access, present, annotate, and to simulate the holistic genome complexity in a unique gateway towards a real understanding, educative presentation and curative manipulation planning of this tremendous evolutionary information grail – genomes.

Corresponding author email contact: TA.Knoch@taknoch.org

Keywords:

Genome, genomics, genome organization, genome architecture, structural sequencing, architectural sequencing, systems genomics, coevolution, holistic genetics, genome mechanics, genome function, genetics, gene regulation, replication, transcription, repair, homologous recombination, simultaneous co-transfection, cell division, mitosis, metaphase, interphase, cell nucleus, nuclear structure, nuclear organization, chromatin density distribution, nuclear morphology, chromosome territories, subchromosomal domains, chromatin loop aggregates, chromatin rosettes, chromatin loops, chromatin fibre, chromatin density, persistence length, spatial distance measurement, histones, H1.0, H2A, H2B, H3, H4, mH2A1.2, DNA sequence, complete sequenced genomes, molecular transport, obstructed diffusion, anomalous diffusion, percolation, long-range correlations, fractal analysis, scaling analysis, exact yard-stick dimension, box-counting dimension, lacunarity dimension, local nuclear dimension, nuclear diffuseness, parallel super computing, grid computing, volunteer computing, Brownian Dynamics, Monte Carlo, fluorescence in situ hybridization, confocal laser scanning microscopy, fluorescence correlation spectroscopy, super resolution microscopy, spatial precision distance microscopy, auto-fluorescent proteins, CFP, GFP, YFP, DsRed, fusion protein, in vivo labelling, information browser, visual data base access, holistic viewing system, integrative data management, extreme visualization, three-dimensional virtual environment, virtual paper tool.

Literature References

- Knoch, T. A.** Dreidimensionale Organisation von Chromosomen-Domänen in Simulation und Experiment. (Three-dimensional organization of chromosome domains in simulation and experiment.) *Diploma Thesis*, Faculty for Physics and Astronomy, Ruperto-Carola University, Heidelberg, Germany, 1998, and TAK Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-010685-5 and ISBN 978-3-00-010685-9 (soft cover, 2nd ed.), ISBN 3-00-035857-9 and ISBN 978-3-00-0358857-0 (hard cover, 2nd ed.), ISBN 3-00-035858-7, and ISBN 978-3-00-035858-6 (DVD, 2nd ed.), 1998.
- Knoch, T. A., Münkkel, C. & Langowski, J.** Three-dimensional organization of chromosome territories and the human cell nucleus - about the structure of a self replicating nano fabrication site. *Foresight Institute - Article Archive*, Foresight Institute, Palo Alto, CA, USA, <http://www.foresight.org>, 1- 6, 1998.
- Knoch, T. A., Münkkel, C. & Langowski, J.** Three-Dimensional Organization of Chromosome Territories and the Human Interphase Nucleus. *High Performance Scientific Supercomputing*, editor Wilfried Juling, Scientific Supercomputing Center (SSC) Karlsruhe, University of Karlsruhe (TH), 27- 29, 1999.
- Knoch, T. A., Münkkel, C. & Langowski, J.** Three-dimensional organization of chromosome territories in the human interphase nucleus. *High Performance Computing in Science and Engineering 1999*, editors Krause, E. & Jäger, W., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3-540-66504-8, 229-238, 2000.
- Bestvater, F., **Knoch, T. A.**, Langowski, J. & Spiess, E. GFP-Walking: Artificial construct conversions caused by simultaneous cotransfection. *BioTechniques* 32(4), 844-854, 2002.
- Knoch, T. A. (editor)**, Backes, M., Baumgärtner, V., Eysel, G., Fehrenbach, H., Göker, M., Hampl, J., Hampl, U., Hartmann, D., Hitzelberger, H., Nambena, J., Rehberg, U., Schmidt, S., Weber, A., & Weidemann, T. Humanökologische Perspektiven Wechsel - Festschrift zu Ehren des 70. Geburtstags von Prof. Dr. Kurt Egger. Human Ecology Working Group, Ruperto-Carola University of Heidelberg, Heidelberg, Germany, 2002.
- Knoch, T. A.** Approaching the three-dimensional organization of the human genome: structural-, scaling- and dynamic properties in the simulation of interphase chromosomes and cell nuclei, long- range correlations in complete genomes, *in vivo* quantification of the chromatin distribution, construct conversions in simultaneous co-transfections. *Dissertation*, Ruperto-Carola University, Heidelberg, Germany, and TAK†Press, Tobias A. Knoch, Mannheim, Germany, ISBN 3-00-009959-X and ISBN 978-3-00-009959-5 (soft cover, 3rd ed.), ISBN 3-00-009960-3 and ISBN 978-3-00-009960-1 (hard cover, 3rd ed.), ISBN 3-00-035856-9 and ISBN 978-3-00-010685-9 (DVD, 3rd ed.) 2002.
- Knoch, T. A.** Towards a holistic understanding of the human genome by determination and integration of its sequential and three-dimensional organization. *High Performance Computing in Science and Engineering 2003*, editors Krause, E., Jäger, W. & Resch, M., High-Performance Computing Center (HLRS) Stuttgart, University of Stuttgart, Springer Berlin-Heidelberg-New York, ISBN 3- 540-40850-9, 421-440, 2003.
- Wachsmuth, M., Weidemann, T., Müller, G., Urs W. Hoffmann-Rohrer, **Knoch, T. A.**, Waldeck, W. & Langowski, J. Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophys. J.* 84(5), 3353-3363, 2003.
- Weidemann, T., Wachsmuth, M., **Knoch, T. A.**, Müller, G., Waldeck, W. & Langowski, J. Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J. Mol. Biol.* 334(2), 229-240, 2003.
- Fejes Tóth, K., **Knoch, T. A.**, Wachsmuth, M., Frank-Stöhr, M., Stöhr, M., Bacher, C. P., Müller, G. & Rippe, K. Trichostatin A induced histone acetylation causes decondensation of interphase chromatin. *J. Cell Science* 117, 4277-4287, 2004.
- Ermler, S., Kronic, D., **Knoch, T. A.**, Moshir, S., Mai, S., Greulich-Bode, K. M. & Boukamp, P. Cell cycle-dependent 3D distribution of telomeres and telomere repeat-binding factor 2 (TRF2) in HaCaT and HaCaT-myc cells. *Europ. J. Cell Biol.* 83(11-12), 681-690, 2004.

- Kost, C., Gama de Oliveira, E., **Knoch, T. A.** & Wirth, R. Spatio-temporal permanence and plasticity of foraging trails in young and mature leaf-cutting ant colonies (*Atta spp.*). *J. Trop. Ecol.* 21(6), 677- 688, 2005.
- Winnefeld, M., Grewenig, A., Schnölzer, M., Spring, H., **Knoch, T. A.**, Gan, E. C., Rommelaere, J. & Czipluch, C. Human SGT interacts with BAG-6/Bat-3/Scythe and cells with reduced levels of either protein display persistence of few misaligned chromosomes and mitotic arrest. *Exp. Cell Res.* 312, 2500-2514, 2006.