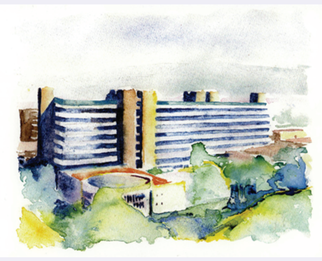
Approaching the Three - Dimensional Organization and Dynamics of the Human Genome

Tobias A. Knoch [1, 2]

in cooperation with C. Cremer [2], T. Cremer [3], K. Fejes-Toth [1, 2], A. Friedl [4], M. Göcker [5], A. Kellerer [4], R. Lohner [6], A. Monier [7], J. Langowski [1], P. Lichter [1], P. Quicken [1], J. Rauch [2], K. Rippe [1, 2], K. Sullivan [7], W. Waldeck [1], M. Wachsmuth [2], T. Weidemann [1]

[1] Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280 and [2] Kirchhoff Institut, Ruprecht-Karls Universität, Im Neuenheimer Feld 227
D - 69120 Heidelberg, Federal Republic of Germany
http://www.taknoch.org or TA.Knoch@taknoch.org

[3] Institut für Anthropologie und Human Genetik, Munich, FRG
[4] Strahlenbiologisches Institut and GSF, Munich, FRG
[5] Spezielle Botanik, Universität Tübingen, FRG
[6] Rechenzentrum Universität Karlsruhe, Karlsruhe, FRG
[7] The Scripps Research Institut, La Jolla, CA, USAnich, FRG
[8] Symantec Education Services, Munich, FRG



FISH

Fluorescence in situ hybridization (FISH) is used for the specific marking of chromosome arms (Fig. 1A) and pairs of small chromosomal DNA regions (Fig. 1B). The labeling is visualized with confocal laser scanning microscopy followed by image reconstruction. Chromosome arms show only small overlap and globular substructures, as predicted by the MLS-model (Fig. 1A & 9A). A comparison between simulated and measured spatial distances between genomic regions as function of their genomic distances results in a good agreement with the MLS-model having loop sizes of arround 126 kbp and linker sizes between 63 kbp and 126 kbp (Fig. 2).

Fig. 1A & 1B: FISH-images of a territory painting of chromosome 15 (left, 1A) and genomic markers YAC-48 and YAC60 (right 1B) with a genomic separation of 1.0 Mbp in interphase of fibroblast cells.

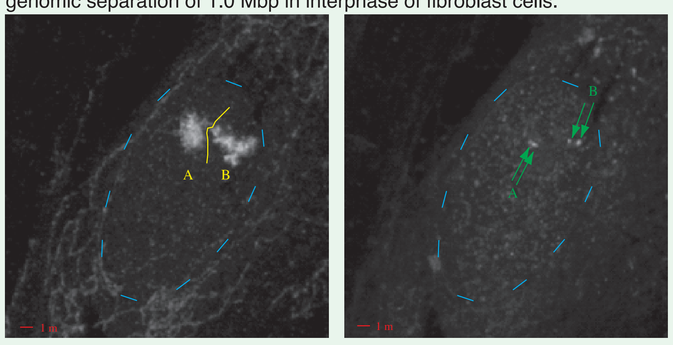
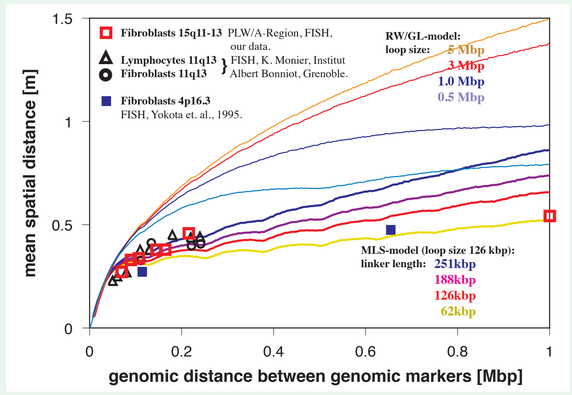


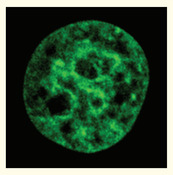
Fig. 2: Comparison of the RW/GL- and the MLS-model with experimentally determined interphase distances.



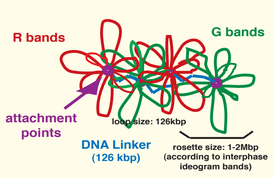
INTRODUCTION

Despite the successful linear sequencing of the genome its three dimensional structure is widely unknown although its importance for gene regulation and replication. By integration of experiments and simulations ranging from the DNA sequence to the nuclear morphology we show here an interdisciplinary approach leading to the determination of the three-dimensional organization and dynamics of the human genome.

Nuclear chromatin morphology by histone H2A-YFP *in vivo* labelling



Multi-Loop-Subcompartment (MLS) model of chromatin folding (Münkel *et al.*, 1997)



CONCLUSION

Simulations of chromsomes and the whole cell nucleus show that only the MLS-model leads to the formation of non-overlapping chromosome territories and distinct functional and dynamic sub-compartments. Spatial distances between FISH labeled pairs of genomic markers as function of their genomic distance agrees with an MLS-model with loop sizes of 120 kbp and linker sizes of 63 to 126 kbp. The in vivo chromatin distribution visualized by histone-GFP fusion proteins is similar to those found in the simulation of whole cell nuclei. Fractal analysis of the simulations reveal the multifractality of chromosomes. It is possible to quantify the in vivo chromatin distribution with fractal analysis and to relate the result to differences in morphology. The simulated diffusion of particles in the nucleus is only moderately obstructed by the chromatin fiber topology in agreement with FCS experiments. Completely sequenced genomes show fine-structured multi-scaling long-range correlations favouring again an MLS like model. Beyond, all these aspects of genome organization are hollistically connected.

SIMULATION

For the prediction of experiments we simulated various models of human interphase chromosome 15 with Monte Carlo and Brownian Dynamics methods. The chromatin fiber was modelled as a flexible polymer fiber. Only stretching, bending and excluded volume interactions are considered. Chromosomes are further confined by a spherical potential representing the surrounding chromosomes or the nuclear membrane. Only the MLS model leads to clearly distinct functional and dynamic subcompartments in agreement with experiments (Fig. 8B & 1A) in contrast to the RW/GL models where big loops are intermingling freely and featureless (Fig. 8C & 8D).

Fig. 8A: Starting configuration with the form and size of a metaphase chromosome.

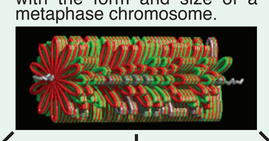


Fig. 8B: MLS model with 126 kbp loops and linkers.

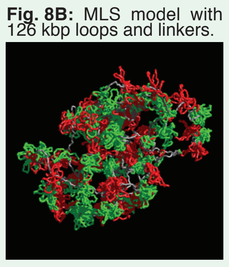


Fig. 8C: RW/GL model with 126 kbp loops.

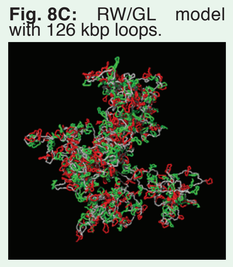


Fig. 8D: RW/GL model with 5 Mbp loops.

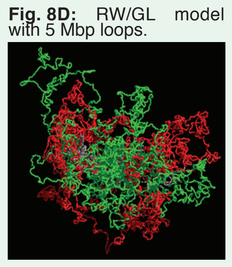
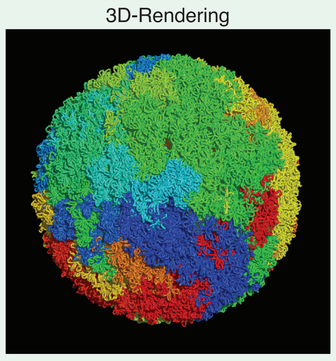
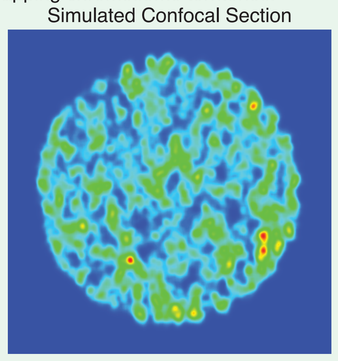


Fig. 9A & 9B: Simulation of a human interphase nucleus containing all 46 chromosomes with 1,200,000 polymer segments. The MLS-model leads to the formation of distinct and non-overlapping chromosome territories.

3D-Rendering

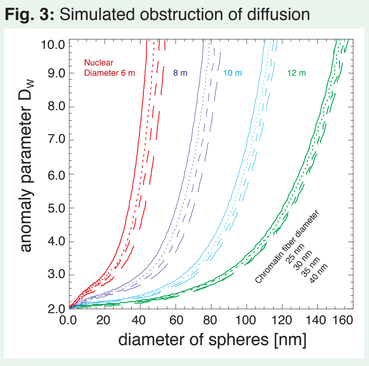


Simulated Confocal Section



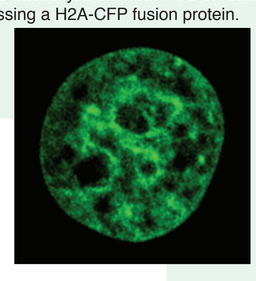
PARTICLE DIFFUSION

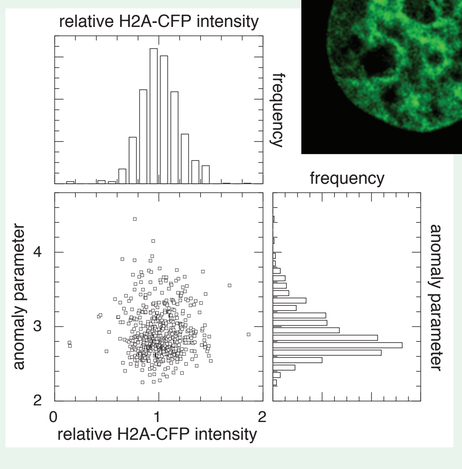
Fig. 3: Simulated obstruction of diffusion



The diffusion of spherical particles within a nucleus was simulated using Brownian Dynamics methods. The mean square displacement of the particles depends on the diameter, the radius of the nucleus, i.e. the obstacle concentration, and also critically on the interaction between particles and structure. For typical biological particles <10 nm the degree of obstruction D_W is moderate (Fig. 3). Thus such particles reach most nuclear locations in less than 10 - 20 ms. This agrees with the volume occupancy and mean chromatin fiber spacing. The diffusion of particles in living interphase nuclei depends on the local structure. In vivo chromatin markers allow to investigate this relation using fluorescence correlation spectroscopy (FCS). The correlation between diffusion obstruction and structure vanishes for small particles (Fig. 4) and increases with increasing particle size.

Fig. 4: The degree of diffusion obstruction plotted against the chromatin density, represented by the H2A-CFP fluorescence intensity. Data from FCS of Alexa568 dye in LCLS103H cell nuclei stably expressing a H2A-CFP fusion protein.





DNA Sequence Correlations

Correlation analysis of completely sequenced genomes reveals fine-structured multi-scaling long-range correlations which are linked to the three-dimensional genome organization (Fig. 5). The general multi-scaling behaviour is due to a block organization and the fine-structure is attributable to the codon usage and to nucleosomal binding. Computer generated random sequences agree with these results. Mutation by sequence reshuffling destroyed all correlations. Trees constructed from the species specific correlation behaviour were as expected for Eukarya (Fig. 6) and led to a new classification system for Archaea and Bacteria (Fig. 7).

Fig. 5: Comparison of the average correlation behaviour of Eukarya, Archaea and Bacteria classes.

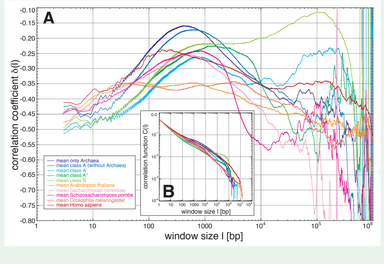


Fig. 6: Tree of Eukarya.

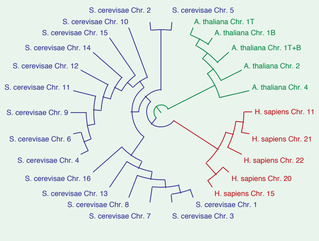
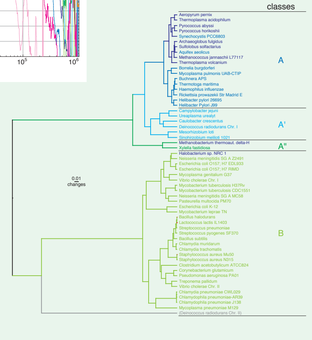


Fig. 7: Tree of Archaea & Bacteria.



FRACTAL ANALYSIS

Fractal analysis is especially suited to quantify the unordered and non-euclidean chromatin distribution of the nucleus. The dynamic behaviour of the chromatin structure and the diffusion of particles in the nucleus are also closely connected to the fractal dimension. The fractal analysis of the simulation of chromosome 15 lead to multifractal behaviour in agreement with porous network research (Fig. 10). Therefore chromosome territories show a higher degree of determinism than previously thought. First tests of fractal analysis of chromatin distributions by histone fusions to fluorescent proteins in vivo result in significant differences for different morphologies (Fig. 11) and might favour an MLS-model like chromatin distribution.

Fig. 10: Comparison of RW/GL- and MLS- model with fractal dimension of the chromatin fiber from simulations.

Fig. 11: Fractal Dimension as function of the intensity threshold.

Approaching the Three-Dimensional Organization and Dynamics of the Human Genome

Knoch, T. A.

Chromatin - Assembly and Inheritance of Functional States, Transregio 5 Symposium, Munich, Germany, 9th - 11th October, 2003.

Abstract

Genomes are one of the major foundations of life due to their role in information storage, process regulation and evolution. However, the sequential and three-dimensional structure of the human genome in the cell nucleus as well as its interplay with and embedding into the cell and organism only arise scarcely. To achieve a deeper understanding of the human genome the three-dimensional organization of the human cell nucleus, the structural-, scaling- and dynamic properties of interphase chromosomes and cell nuclei were simulated and combined with the analysis of long-range correlations in completely sequenced genomes as well as the chromatin distribution *in vivo*. Using Monte Carlo and Brownian Dynamics methods, the 30 nm chromatin fibre was simulated according to the Multi-Loop-Subcompartment (MLS) model, in which ~100 kbp loops form rosettes, connected by a linker, and the Random-Walk/Giant-Loop (RW/GL) topology, in which 1-5 Mbp loops are attached to a flexible backbone. Both the MLS and the RW/GL model form chromosome territories but only the MLS rosettes result in distinct subcompartments visible with light microscopy and low overlap of chromosomes, -arms and subcompartments. The MLS morphology, the size of subcompartments and chromatin density distribution of simulated confocal (CLSM) images agree with the expression of fusionproteins from the histones H1, H2A, H2B, H3, H4 and mH2A1.2 with the auto-fluorescent proteins CFP, GFP, YFP, DsRed-1 and DsRed-2 which also revealed different interphase morphologies for different cell lines. Even small changes of the model parameters induced significant rearrangements of the chromatin morphology. Thus, pathological diagnoses, are closely related to structural changes on the chromatin level. The position of interphase chromosomes depends on their metaphase location, and suggests a possible origin of current experimental findings. The scaling behaviour of the chromatin fibre topology and morphology of CLSM stacks revealed fine-structured multi-scaling behaviour in agreement with the model prediction and correlations in the DNA sequence. Review and comparison of experimental to simulated spatial distance measurements between genomic markers as function of their genomic separation also favour an MLS model with loop and linker sizes of 63 to 126 kbp. Simulated and experimental DNA fragment distribution after ion-irradiation revealed also best agreement with such an MLS. Correlation analyses of completely sequenced Archaea, Bacteria and Eukarya chromosomes revealed fine-structured positive long-range correlation due to codon, nucleosomal or block organization of the genomes, allowing classification as well as tree construction. This shows a complex sequential organization of genomes closely connected to their three-dimensional organization. Visual inspection of the morphology reveals also big spaces between the chromatin fibre allowing high accessibility to nearly every spatial location, due to the chromatin occupancy <30% and a mean mesh spacing of 29 to 82 nm for nuclei of 6 to 12 µm diameter. This agrees with a simulated displacement of 10 nm sized particles of ~1 to 2 µm takes place within 10 ms, i. e. a moderately obstructed diffusion of biological molecules in agreement with experiments. Thus, the local, global and dynamic characteristics of cell nuclei are not only tightly inter-connected, but also are integrated holisticly to fulfill the overall function of the genome.

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.

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, 2rd ed.), ISBN 3-00-035857-9 and ISBN 978-3-00-0358857-0 (hard cover, 2rd ed.), ISBN 3-00-035858-7, and ISBN 978-3-00-035858-6 (DVD, 2rd ed.), 1998.

Knoch, T. A., Münkel, 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., Münkel, 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., Münkel, 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 Perspectiven 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.