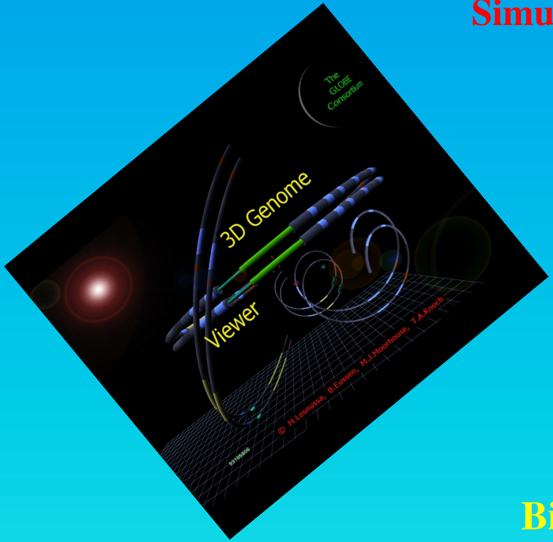
Erasmus Computing Grid

Erasmus MC
University Medical Center Rotterdam

From Sequence to Morphology

-

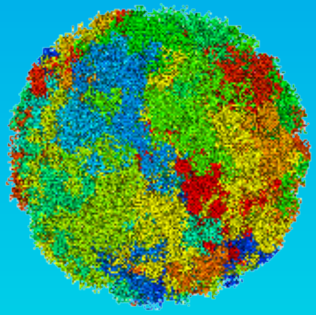
Approaching the Three-Dimensional Organization of the Human Genome



Simulation of Structural-, Scaling and Dynamic-Properties

The Erasmus Computing Grid

The Next Generation Genome Viewer



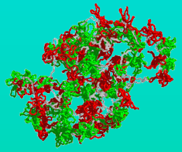
by

Tobias A. Knoch

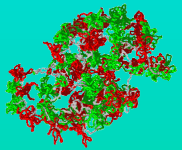
Biophysical Genomics in Cell Biology & Genetics

Bert Eussen, Michael Moorhouse, Annelies de Klein, & Frank Grosveld

Cell Biology & Genetics, Clinical Genetics, & Virology



Erasmus MC, Rotterdam,The Netherlands



Erasmus Computing Grid

Erasmus MC
University Medical Center Rotterdam

HOGESCHOOL ROTTERDAM

Erasmus MC
University Medic l Center Rotterdam

From Sequence to Morphology

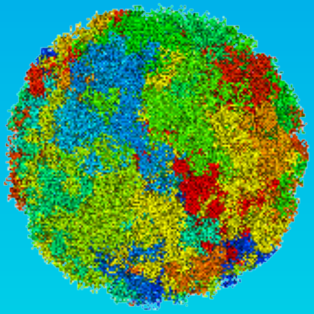
-

Approaching the Three-Dimensional Organizati Genome

Simulation of Structural-, Scalin

The Erasmus

The Next C



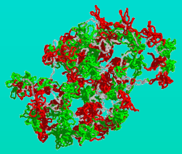
From Sequence to Morphology:
Towards a Holistic Understanding of Genomes!

bias A. Knoch

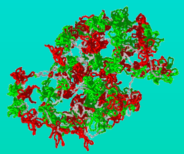
al Genomics in Cell Biology & Genetics

, Michael Moorhouse, Annelies de Klein, & Frank Grosveld

Cell Biology & Genetics, Clinical Genetics, & Virology

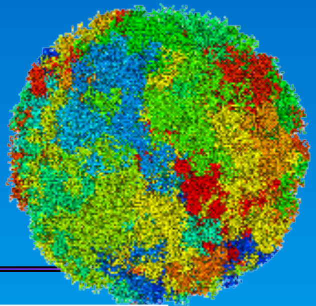


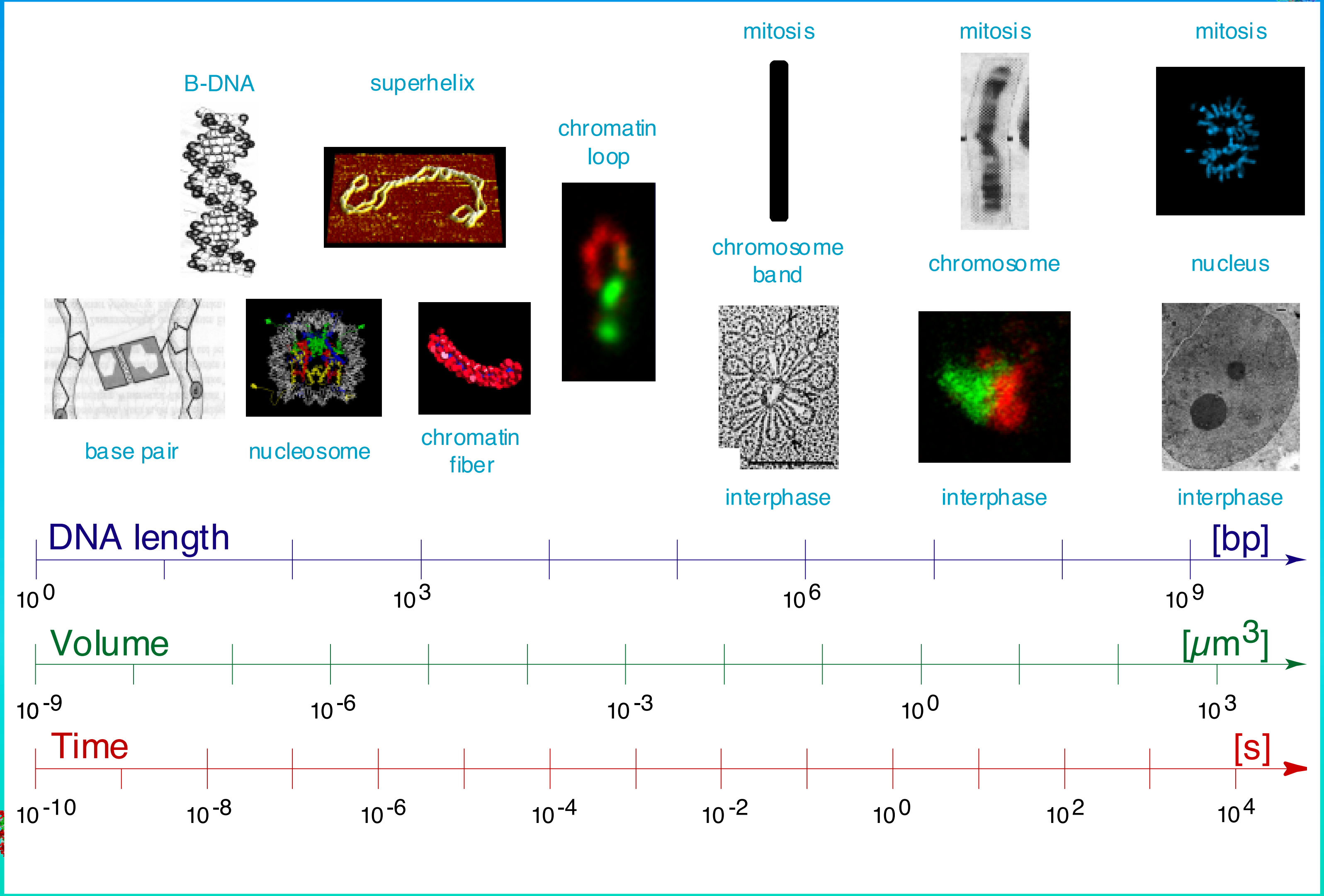
Erasmus MC, Rotterdam,The Netherlands

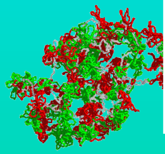


Dynamic and Hierarchical Genome Organization

The different organization levels of genomes bridge several orders of magnitude concerning space and time. How all of these organization levels connect to processes like gene regulation, replication, embryogeneses, or cancer development is still unclear?

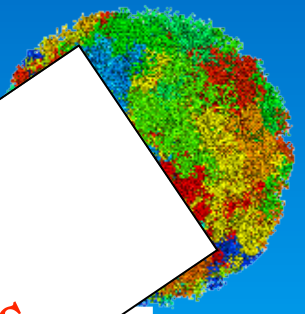


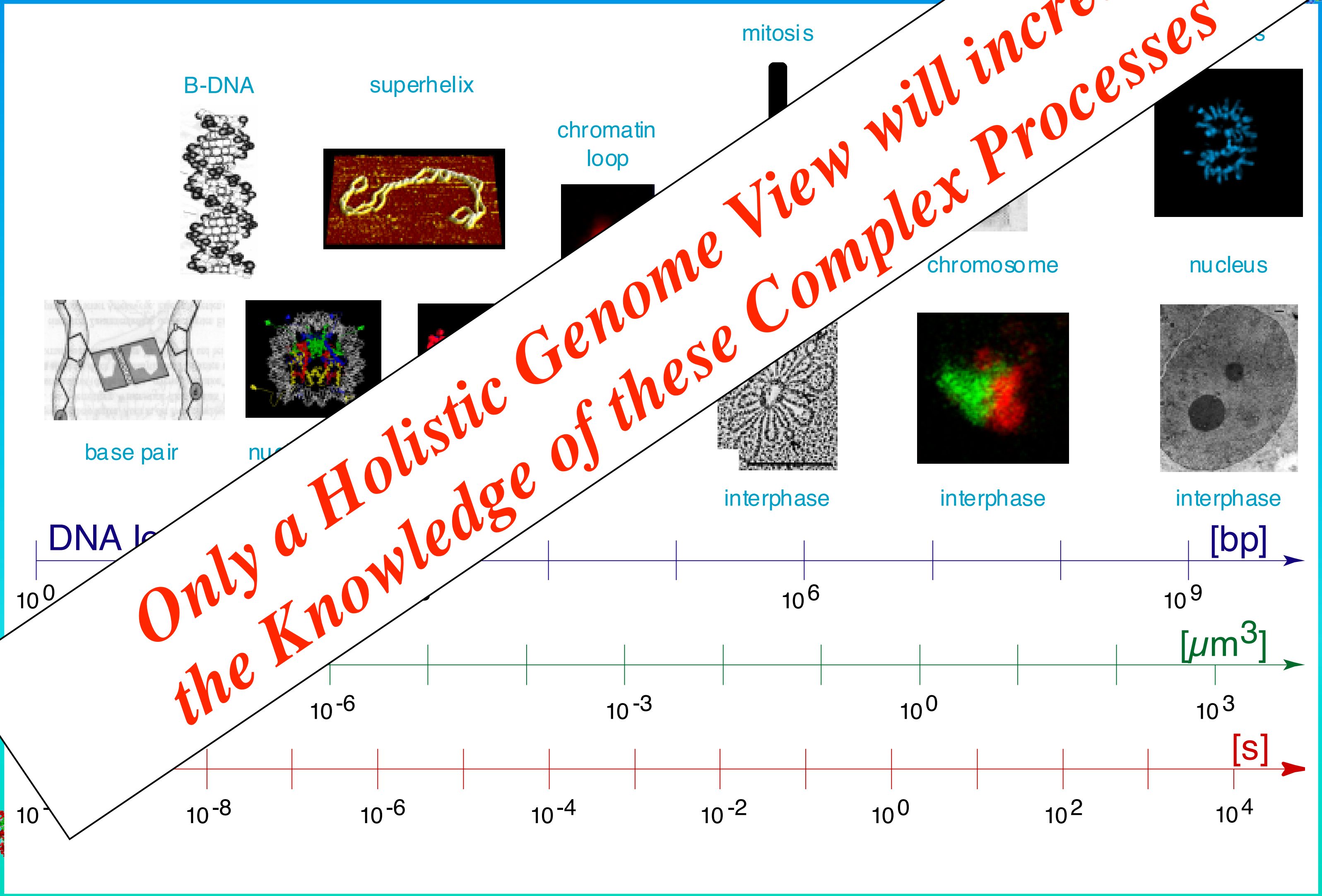




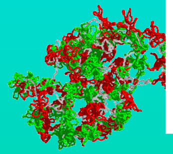
Dynamic and Hierarchical Genome Organization

The different organization levels of genomes bridge several orders of magnitude concerning space and time. How all of these organization levels connect to processes like gene regulation, replication, embryogeneses, or cancer development is still unclear?



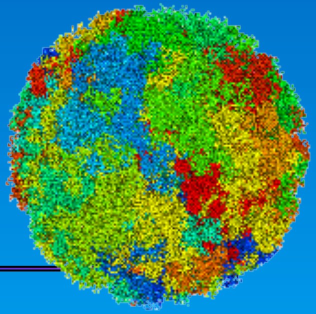


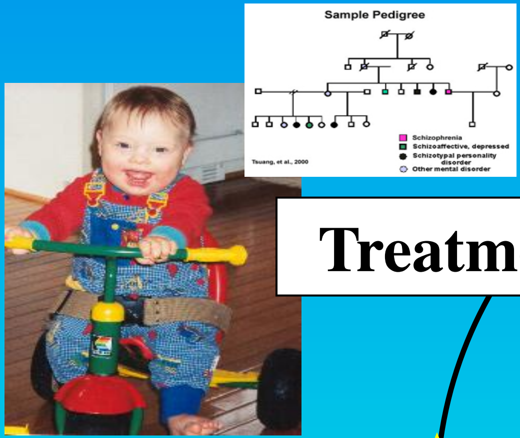
Only a Holistic Genome View will increase the Knowledge of these Complex Processes

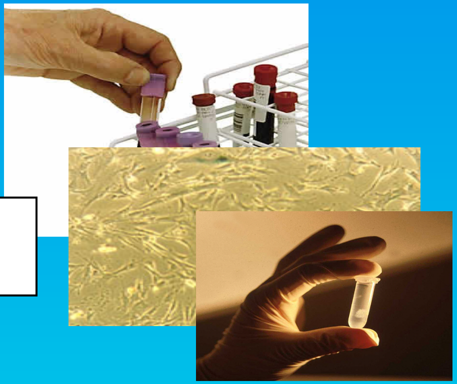


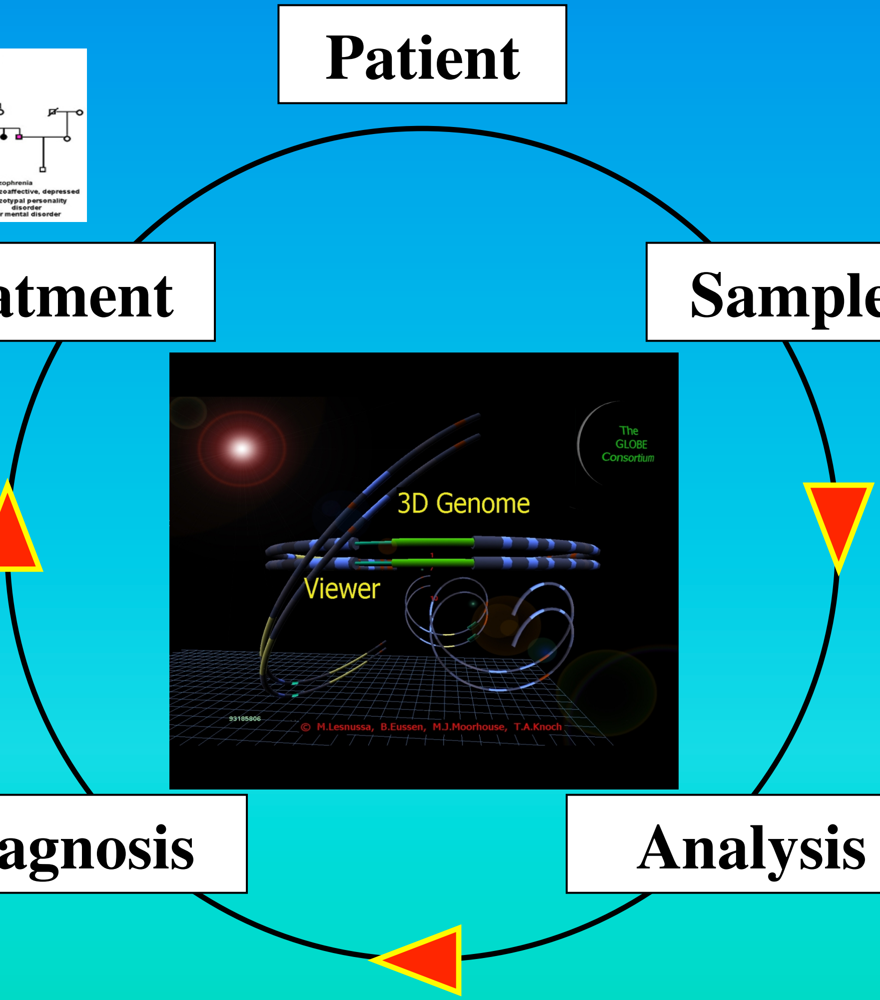
The Complexity of Cytogenetic Diagnostics

The process of cytogenetic analysis requires proper patient and sample analysis
as well asa comprehensive evaluation of the results.









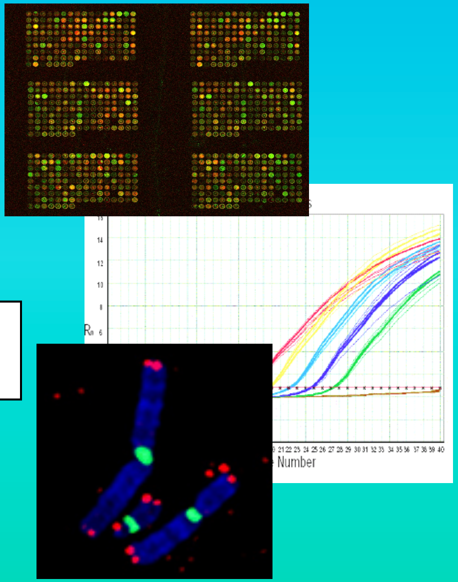
Patient

Sample

Analysis

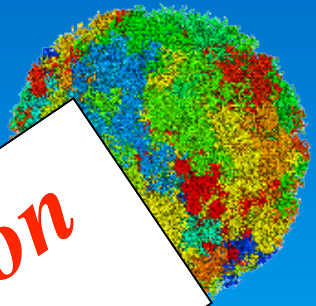
Diagnosis

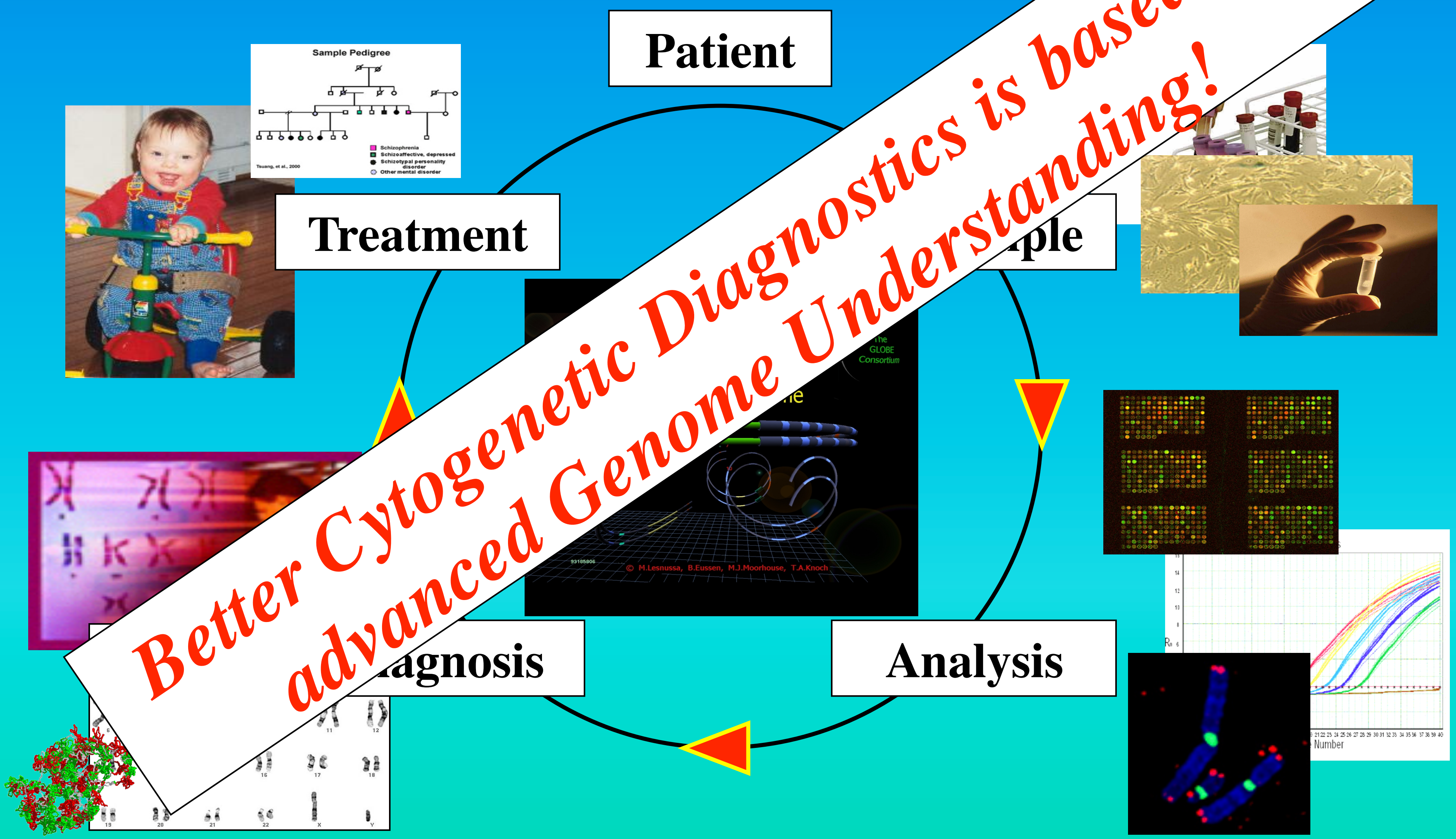
Treatment



The Complexity of Cytogenetic Diagnostics

The process of cytogenetic analysis requires proper patient and sample analysis
as well asa comprehensive evaluation of the results.





Patient

ple

Analysis

agnosis

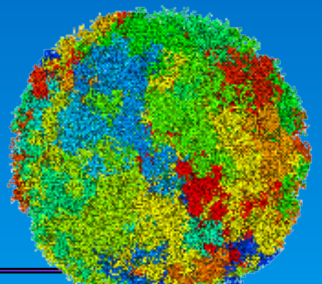
Treatment

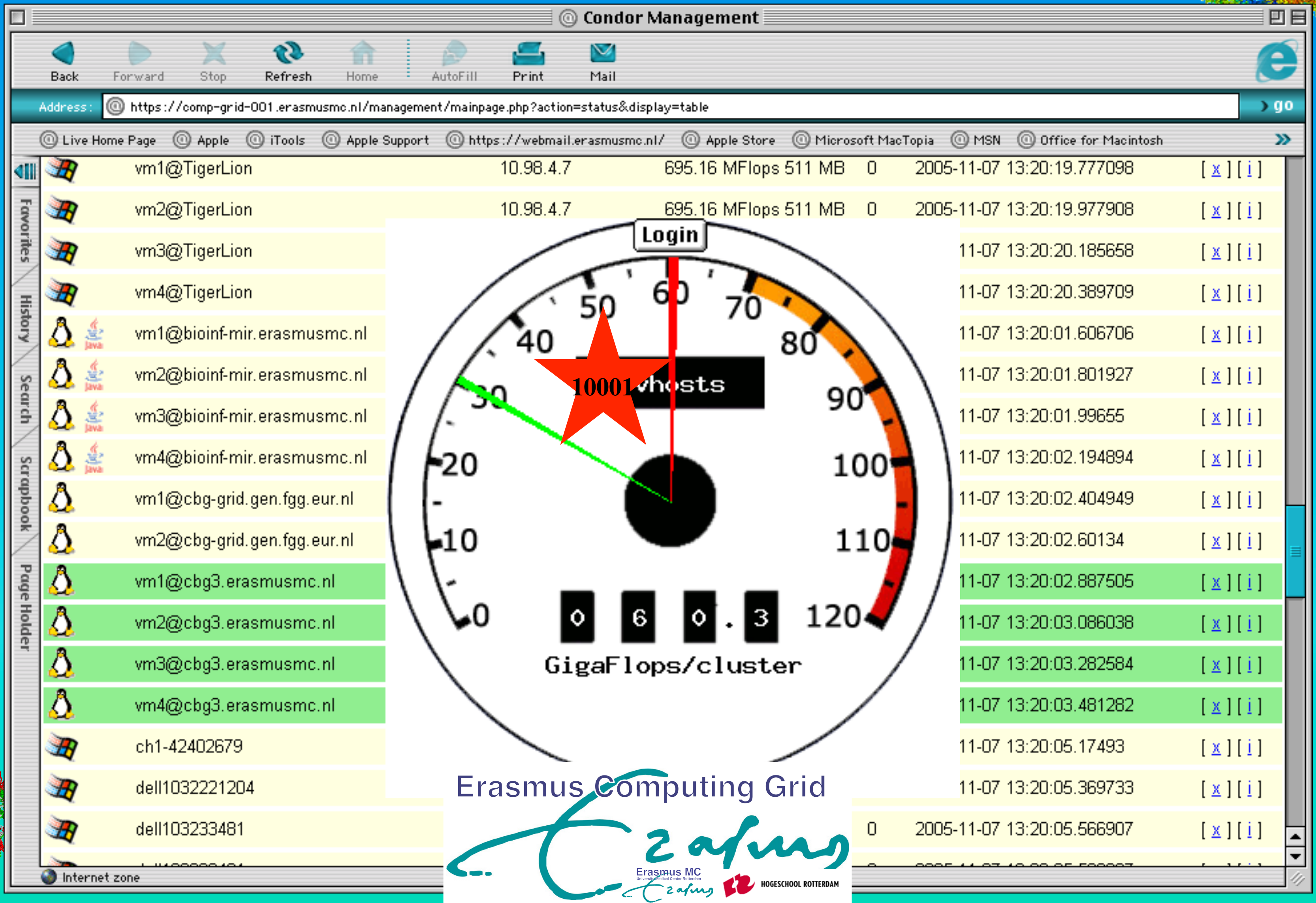
Better Cytogenetic Diagnostics is based on advanced Genome Understanding!

The Erasmus Computing Grid

Building a Super-Computer at Erasmus MC for Free!

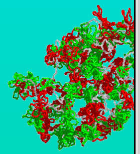
Applications at Erasmus MC: E.g. genome sequence analysis, protein structure simulations, chip/array analysis, epidemiology of viral infections, patient diagnostic image analysis (AMI).





Address: https://comp-grid-001.erasmusmc.nl/management/mainpage.php?action=status&display=table

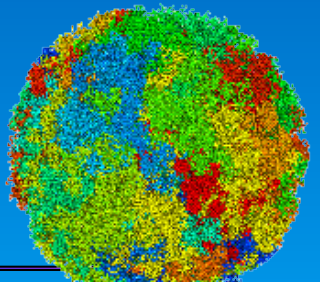
vm1@TigerLion	10.98.4.7	695.16 MFlops 511 MB	0	2005-11-07 13:20:19.777098	[x] [i]
vm2@TigerLion	10.98.4.7	695.16 MFlops 511 MB	0	2005-11-07 13:20:19.977908	[x] [i]
vm3@TigerLion				11-07 13:20:20.185658	[x] [i]
vm4@TigerLion				11-07 13:20:20.389709	[x] [i]
vm1@bioinf-mir.erasmusmc.nl				11-07 13:20:01.606706	[x] [i]
vm2@bioinf-mir.erasmusmc.nl				11-07 13:20:01.801927	[x] [i]
vm3@bioinf-mir.erasmusmc.nl				11-07 13:20:01.99655	[x] [i]
vm4@bioinf-mir.erasmusmc.nl				11-07 13:20:02.194894	[x] [i]
vm1@cbg-grid.gen.fgg.eur.nl				11-07 13:20:02.404949	[x] [i]
vm2@cbg-grid.gen.fgg.eur.nl				11-07 13:20:02.60134	[x] [i]
vm1@cbg3.erasmusmc.nl				11-07 13:20:02.887505	[x] [i]
vm2@cbg3.erasmusmc.nl				11-07 13:20:03.086038	[x] [i]
vm3@cbg3.erasmusmc.nl				11-07 13:20:03.282584	[x] [i]
vm4@cbg3.erasmusmc.nl				11-07 13:20:03.481282	[x] [i]
ch1-42402679				11-07 13:20:05.17493	[x] [i]
dell1032221204				11-07 13:20:05.369733	[x] [i]
dell103233481			0	2005-11-07 13:20:05.566907	[x] [i]

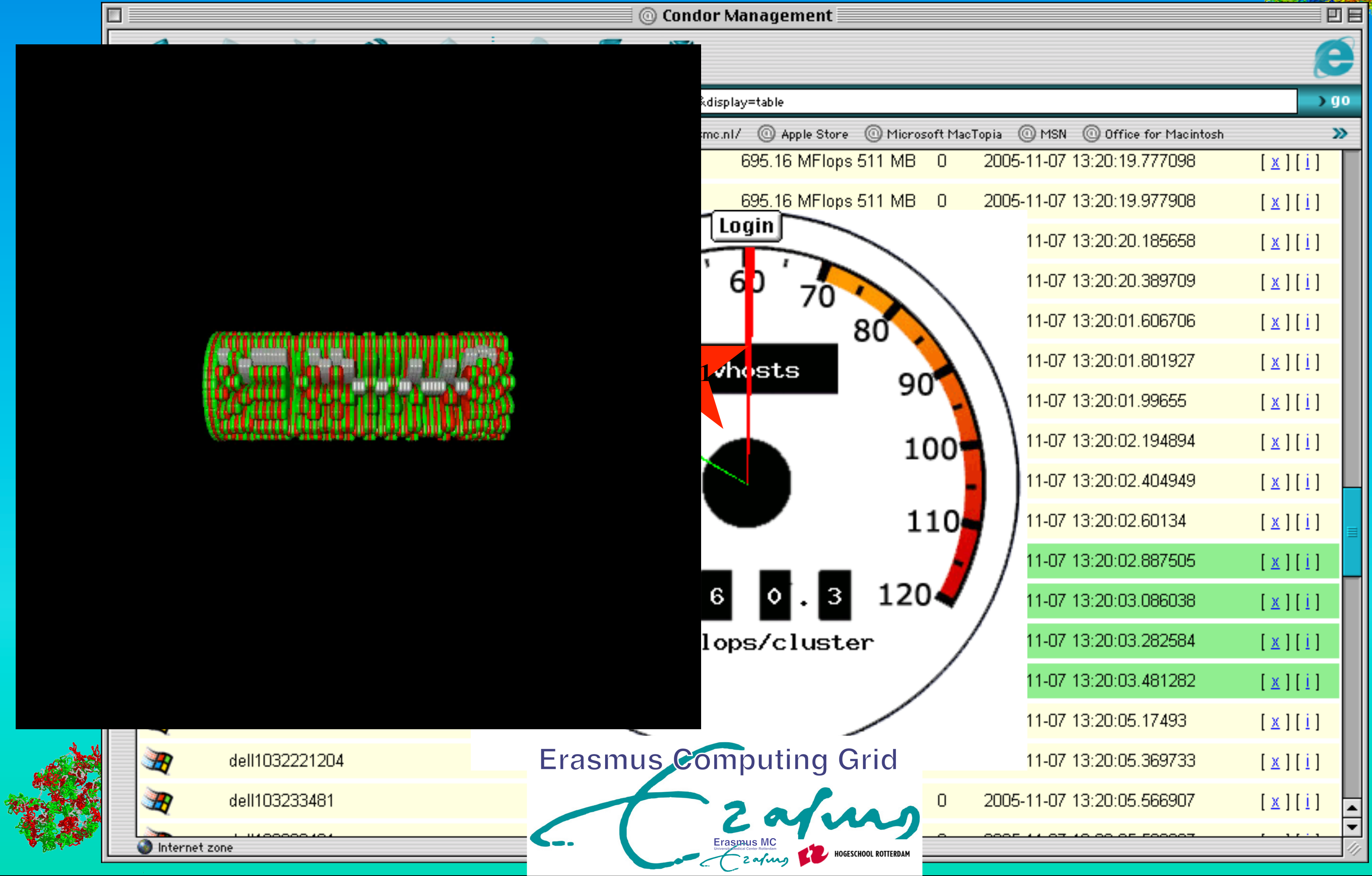


The Erasmus Computing Grid

Building a Super-Computer at Erasmus MC for Free!

Applications at Erasmus MC: E.g. genome sequence analysis, protein structure simulations, chip/array analysis, epidemiology of viral infections, patient diagnostic image analysis (AMI).



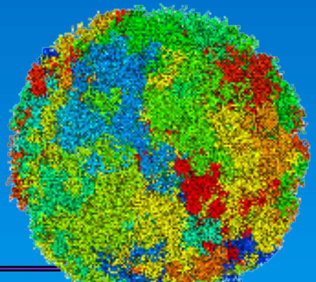


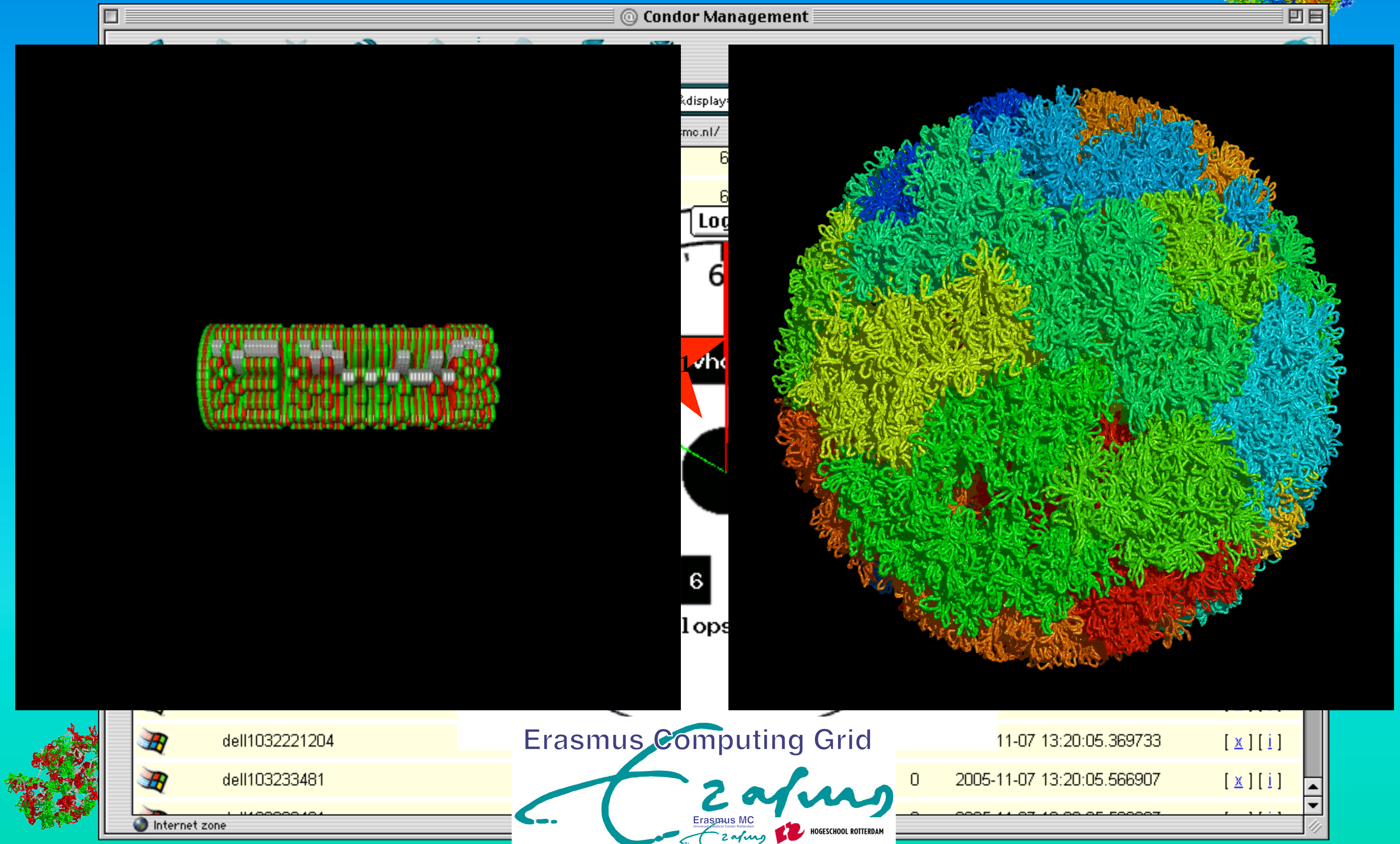
695.16 MFlops 511 MB 0 2005-11-07 13:20:19.777098 [x] [i]
695.16 MFlops 511 MB 0 2005-11-07 13:20:19.977908 [x] [i]
11-07 13:20:20.185658 [x] [i]
11-07 13:20:20.389709 [x] [i]
11-07 13:20:01.606706 [x] [i]
11-07 13:20:01.801927 [x] [i]
11-07 13:20:01.99655 [x] [i]
11-07 13:20:02.194894 [x] [i]
11-07 13:20:02.404949 [x] [i]
11-07 13:20:02.60134 [x] [i]
11-07 13:20:02.887505 [x] [i]
11-07 13:20:03.086038 [x] [i]
11-07 13:20:03.282584 [x] [i]
11-07 13:20:03.481282 [x] [i]
11-07 13:20:05.17493 [x] [i]
dell1032221204 11-07 13:20:05.369733 [x] [i]
dell103233481 0 2005-11-07 13:20:05.566907 [x] [i]

The Erasmus Computing Grid

Building a Super-Computer at Erasmus MC for Free!

Applications at Erasmus MC: E.g. genome sequence analysis, protein structure simulations, chip/array analysis, epidemiology of viral infections, patient diagnostic image analysis (AMI).





dell1032221204		11-07 13:20:05.369733	[x] [i]
dell103233481	0	2005-11-07 13:20:05.566907	[x] [i]

Erasmus Computing Grid

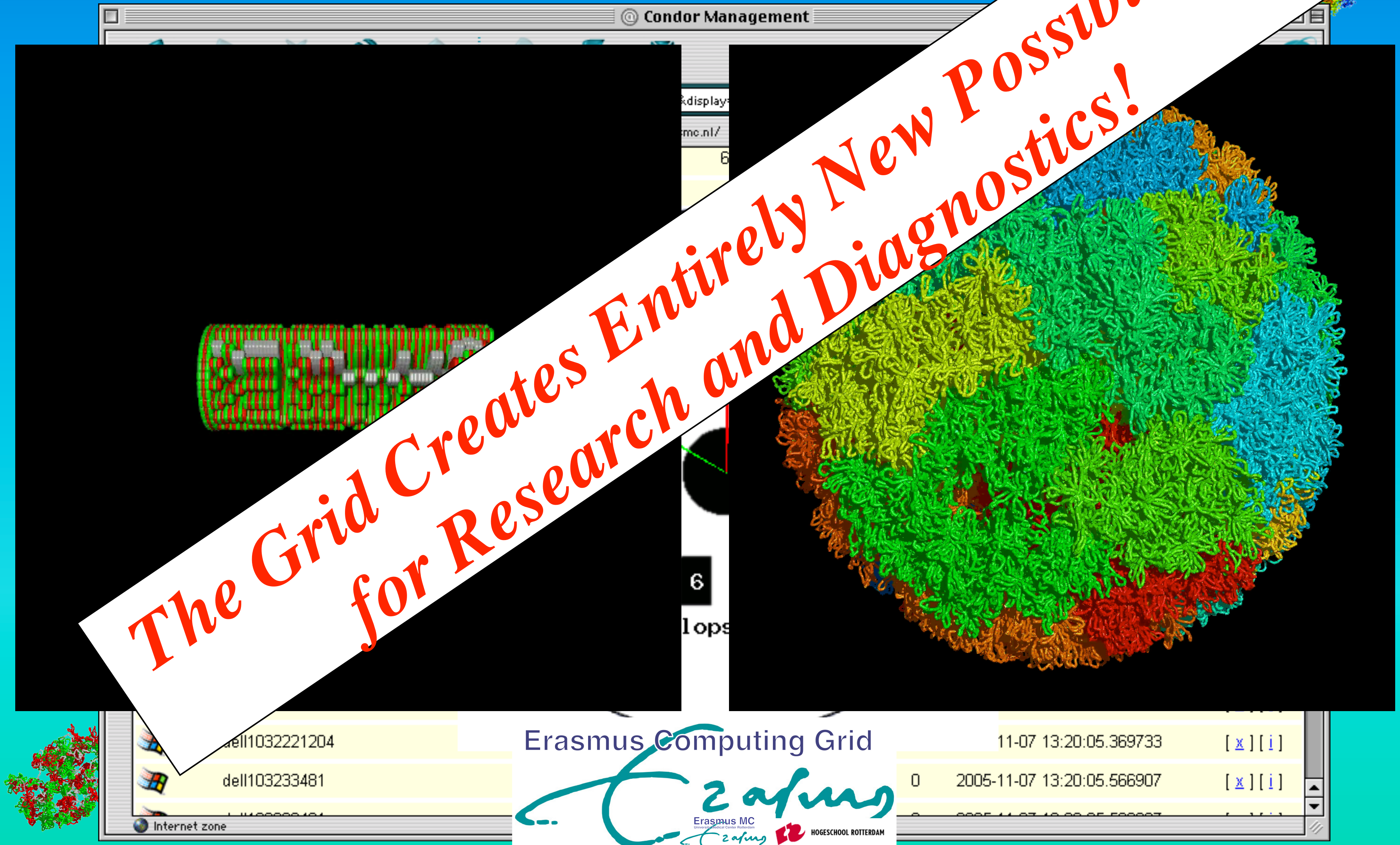
Erasmus MC

HOGESCHOOL ROTTERDAM

The Erasmus Computing Grid

Building a Super-Computer at Erasmus MC for Free!

Applications at Erasmus MC: E.g. genome sequence analysis, protein structure simulations, chip/array analysis, epidemiology of viral infections, patient diagnostic image analysis (AMI).



The Grid Creates Entirely New Possibilities for Research and Diagnostics!

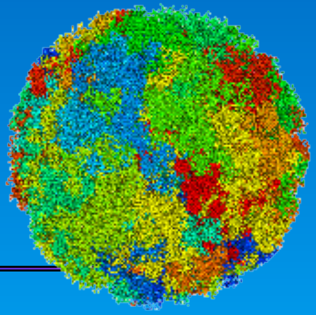
Erasmus Computing Grid

Erasmus MC

HOGESCHOOL ROTTERDAM

The 3D Genome Viewer

To develop an entire novel system-biology oriented genome browser, i.e. integrating the holistic complexity of genome organization in a single easy comprehensible platform has required completely new approaches to represent the genome architecture realistically in combination with the various types experimental data or instant analysis capabilities and annotation.



Paradigm:

- **simplicity, flexibility, customisability, expandability and open-source sustainability**

concerning

- **user accessibility and needs, the in-/output of data sources, system modularity, and integration of future developments**

Content:

- **all spatial scales from the base-pair to nuclear morphology including all other system composites**
- **every type of even large multi-dimensional datasets**

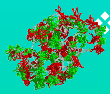
User Handling:

- **intuitive navigation with dynamical semantic resolution and relational arrangement**

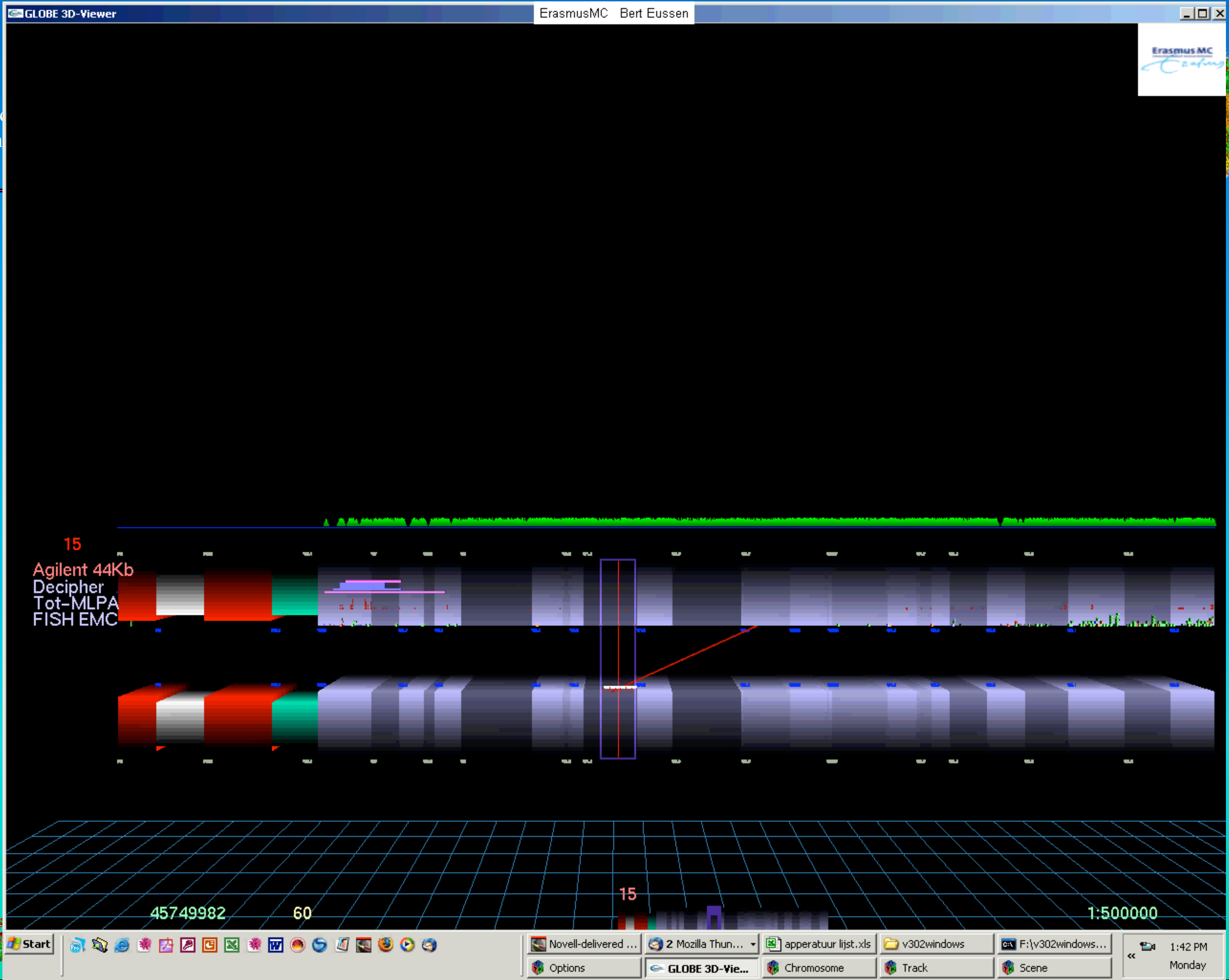
Data Handling:

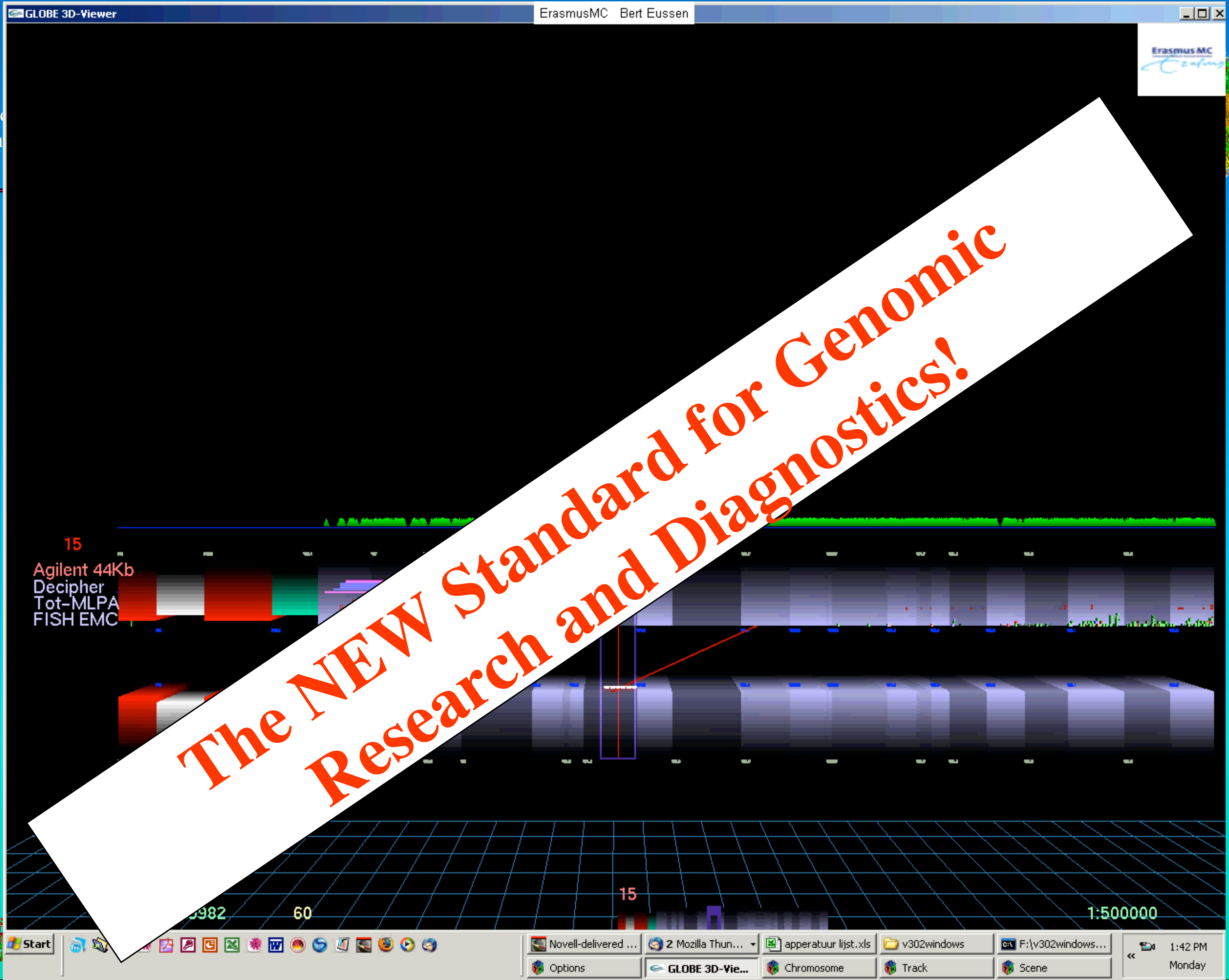
- **real-time interaction with data manipulation or annotation**
- **simple relational or complex correlative multi-dimensional analysis and planning capabilities**

Control Options:



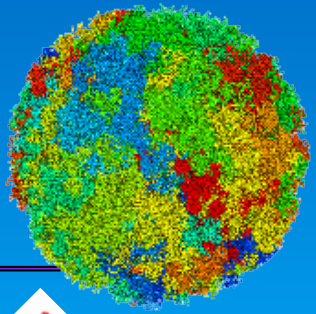
- **front-end access and control of virtual system biological genome simulations**


GLOBE 3D-Viewer
ErasmusMC Bert Eussen
Erasmus MC
15
Agilent 44Kb
Decipher
Tot-MLPA
FISH EMC
45749982
60
15
1:500000
Start
Novell-delivered ...
2 Mozilla Thun...
apperatuur lijst.xls
v302windows
F:\v302windows...
Options
GLOBE 3D-Vie...
Chromosome
Track
Scene
1:42 PM
Monday



The NEW Standard for Genomic Research and Diagnostics!

The 3D Genome Viewer



Due to its broad impact, the GLOBE 3D Genome Viewer addresses different markets and revenue streams in the areas of research, education, diagnostics and industry with different detailed business models and contracting patterns which will guaranty a constant high-profile valorization.

Basic Corporate License:

- 50% science
- 50% industry

White Label Products:

- special customer products

Content Sale

- specially defined data sets
- specially connected data sets

Advertising:

- sponsored free version
- content and user related

eCommerce:

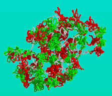
- community portal and referral fees

Companies/Organizations:

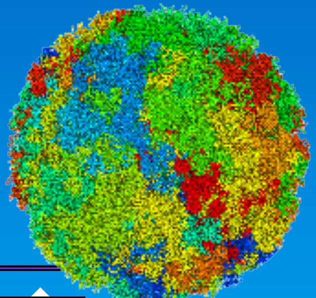
- Invitrogen (Karlsbad, CA, USA)
- Affimetrix (Santa Clara, CA, USA)
- Soft Genetics (State College, PA, USA)
- Combimatrix CMDX (San Antonio, CA, USA)
- Philips (The Netherlands)
- MediGRID (Goettingen, Germany)

Collaborators:

- Prof. Dr. Evan E. Eichler (Seatle, WA, USA)
- Prof. Dr. Stephen W. Scherer (Toronto, Canada)
- Prof. Dr. Michael Hausmann (Heidelberg, Germany)
- Dr. Malte Wachsmuth (Seoul, Korea)
- Dr. Gregory Cox (Karlsbad, CA, USA)
- Dr. Eberhard Schmitt (Jena, Germany)
- Dr. Sabine Baars (Leiden, The Netherlands)



The 3D Genome Viewer



Due to its broad impact, the GLOBE 3D Genome Viewer addresses different markets and revenue streams in the areas of research, education, diagnostics and industry with different detailed business models and contracting patterns which will guaranty a constant high-profile valorization.

Basic Co

- 50% sc
- 50% in

White L

- special

Content

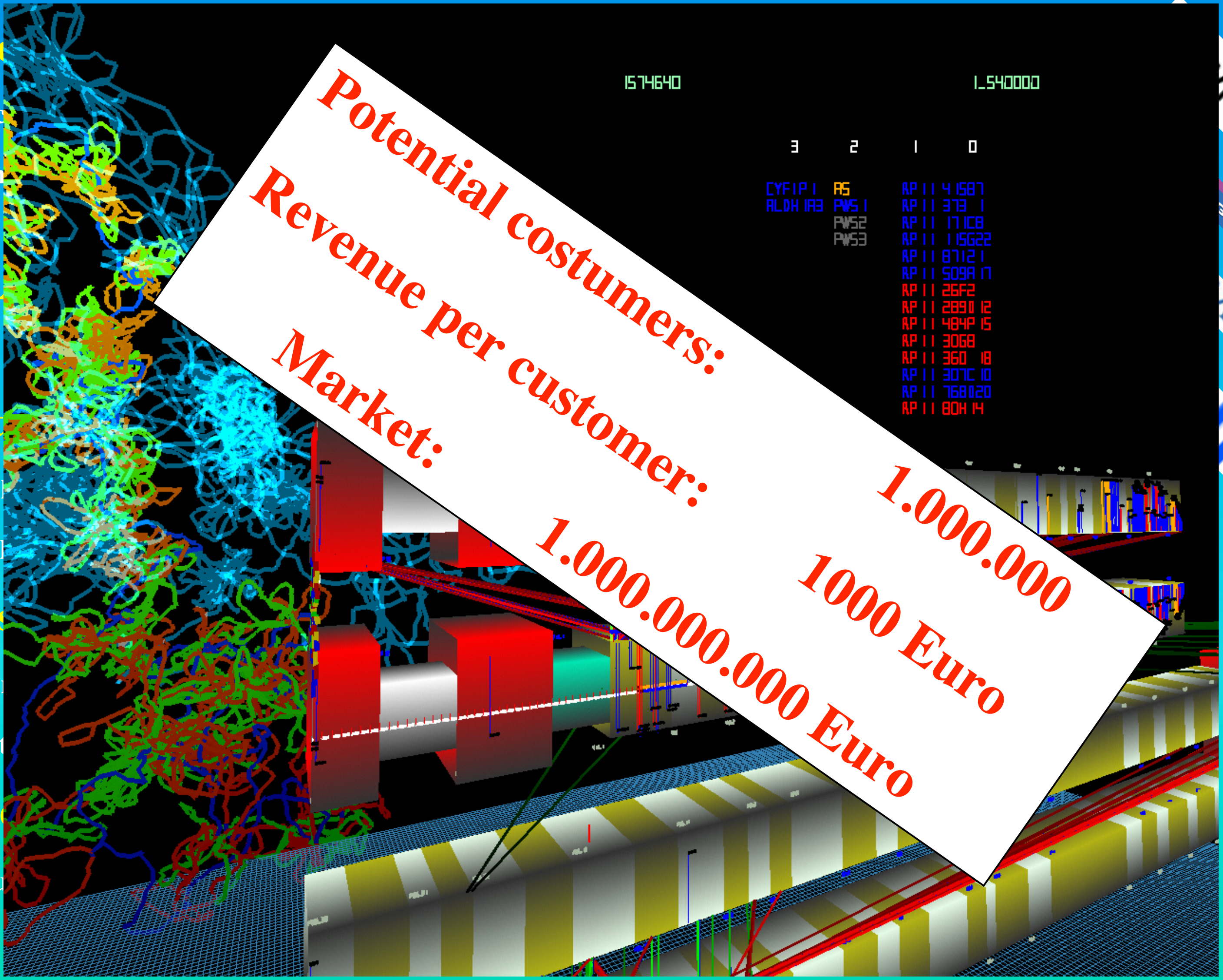
- special
- special

Advertis

- sponso
- content

eComm

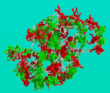
- commu



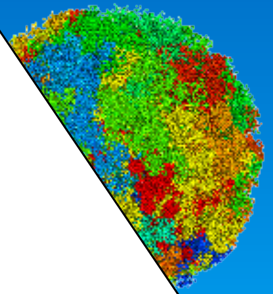
Potential costumers: 1.000.000

Revenue per customer: 1000 Euro

Market: 1.000.000.000 Euro



The 3D Genome Viewer



Due to its broad impact, the GLOBE 3D Genome Viewer addresses different markets and revenue streams
in the areas of research, education, diagnostics and industry with different detailed business models a
contracting patterns which will guaranty a constant high-profile valorization.

Basic C

- 50% sc
- 50% in

White L

- special

Content

- special
- special

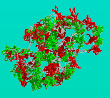
Ad

eC

- co

Potential
Revenu
1.000.000
1000 Euro
000.000.000 Euro

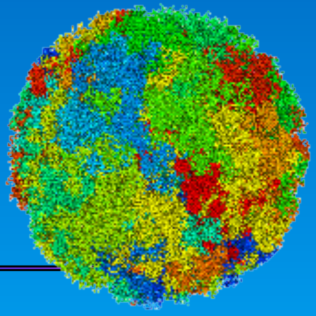
The 3D Genome Viewer allows a Holistic Genome View with Clear Benefits for Research, Diagnostics and Treatment!

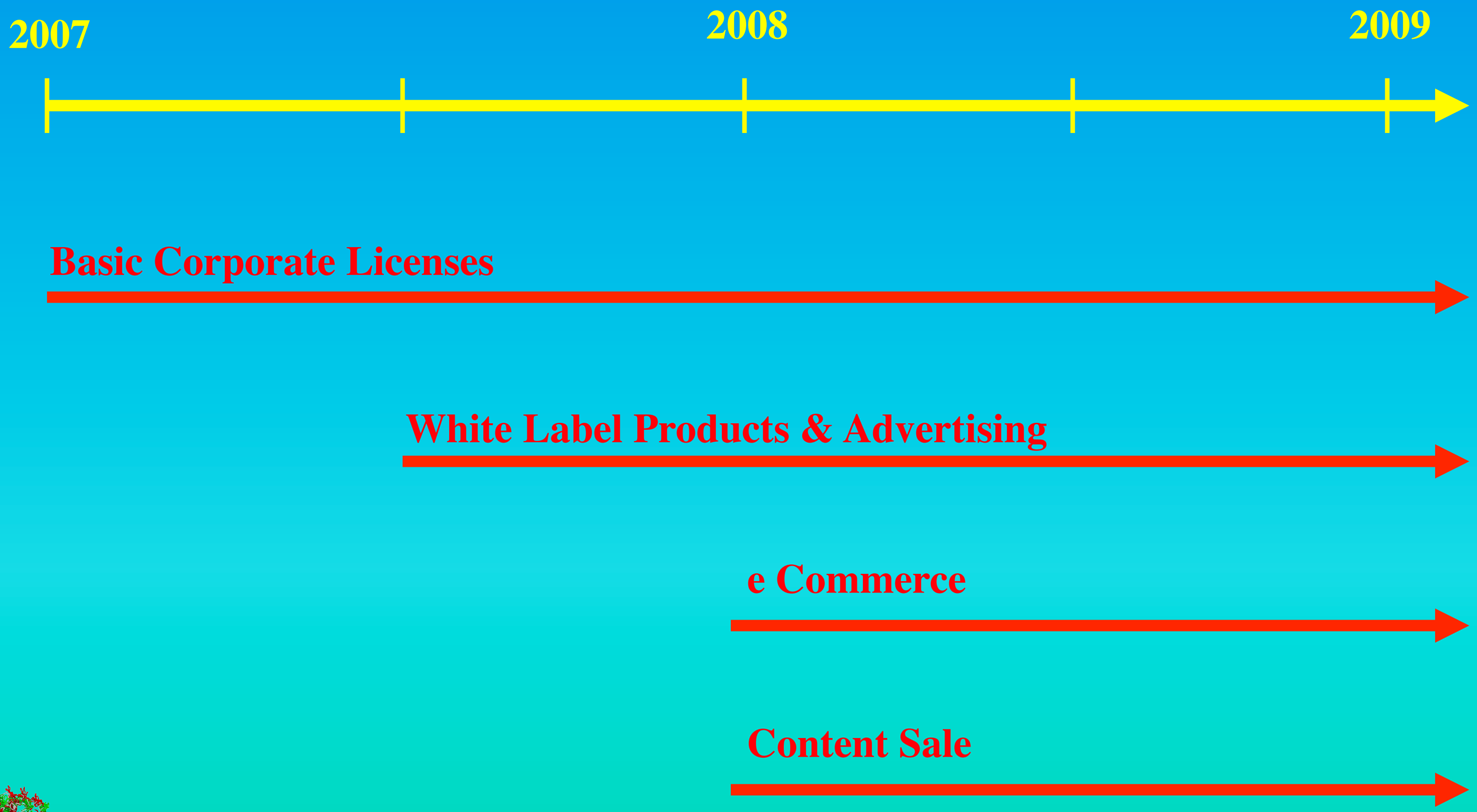


The 3D Genome Viewer

The NEW Standard for Genomic Research and Diagnostics!

Ready to be Valorized!!!





2007

2008

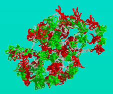
2009

Basic Corporate Licenses

White Label Products & Advertising

e Commerce

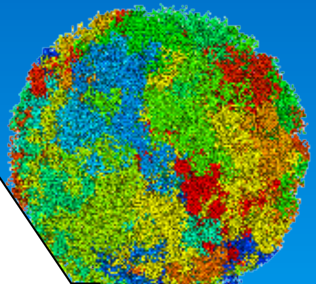
Content Sale



The 3D Genome Viewer

The NEW Standard for Genomic Research and Diagnostics!

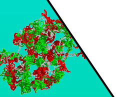
Ready to be Valorized!!!



2007

Basic Co

The 3D Genome Viewer allows a Holistic Genome View with Clear Benefits for Research, Diagnostics and Treatment!



Opportunities

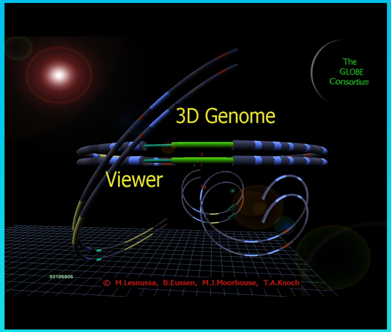
We have started to establish successfully a unique environment for genomic research and diagnostics leading to a holistic understanding of genomes necessary for the improvement of mankind!

The Erasmus Computing Grid:

The Erasmus Computing Grid greatly advances the computing capabilities of the Erasmus Mc, e.g. for:

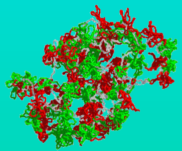
- genomic and proteomic analysis
- epidemiology
- clinical image analysis, e.g. Applied Molecular Imaging (AMI)

The 3D Genome Viewer:

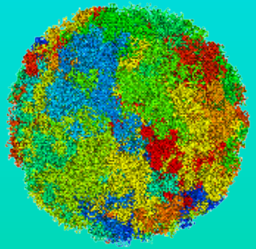


The 3D Genome Viewer greatly advances the integration of multi-dimensional data, e.g.:

- sequential genomic data
- structural genomic data
- clinical diagnosis data



Both the Erasmus Computing Grid and the 3D Genome Viewer have potential for commercialization!



Opportunities

We have started to establish successfully a unique environment for genomic research and diagnostics leading to a holistic understanding of genomes necessary for the improvement of mankind!

The Erasmus Computing Grid:

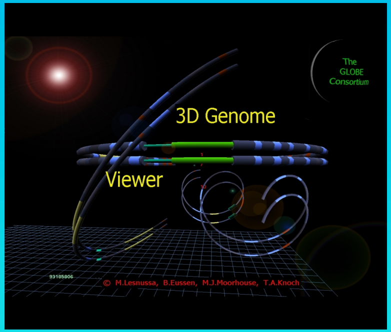
The Erasmus Computing Grid greatly advances the computi

- genomic and proteomic analysis
- epidemiology
- clinical image analysis, e.g. Appli

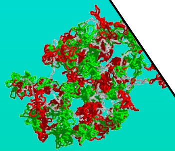
The 3D Genom

The 3D Genc ulti-dimensional data, e.g.:

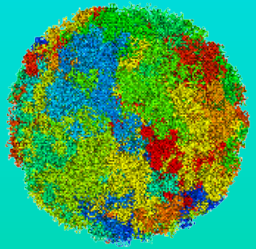
-



Both the Erasmus Computing Grid & the 3D Genome Viewer have great potential for commercialization!



asmus Computing Grid and the 3D Genome Viewer have potential for commercialization!



Approaching the Sequential and Three-Dimensional Organization

Of

Genomes

Knoch, T. A.

EuroBio 2006 - Tenth European Biotech Crossroads, Paris, France, 25th - 27 October, 2006.

Abstract

Genomes are one of the major foundations of life due to their role in information storage, process regulation and evolution. To achieve a deeper unterstanding 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 their sequential organization and (*in vivo*) experiments: The Multi-Loop-Subcompartment model is currently in best agreement with experiment concerning formation and overlap of chromosomes, -territories -arms and subcompartments. Even small model changes lead to pathological recognisable morphologic differences. Review and comparison of experimental to simulated spatial distances between genomic markers even lead to specifics as a chromatin loop size of 63 to 126 kbp. Correlation analyses on the DNA sequence level of completely sequenced Archaea, Bacteria and Eukarya revealed fine-structured positive multi-scaling long-range correlations being again in agreement. Simulations and experiments agree also in the diffusion behaviour of biologically relevant tracers being only moderately obstructed. Thus, the local, global and dynamic genome characteristics are tightly inter-connected and integrated holisticly to fulfill their function.

On this bases we also developed the first system-biology oriented genome browser - the ***GLOBE 3D Genome Viewer*** - necessary to access, present, annotate, and to simulate this holistic complexity in a unique gateway towards a real understanding, educative presentation and curative manipulation planning of genomes. Additionally, we currently set-up the ***Erasmus Computing Grid*** for the enormous computational needs of live-science research and diagnostics. With 3 Tera flops and reaching 20 Tera flops in the first set-up phase the ECG is one of the worlds largest desktop computing grids.

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, 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.

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* 177, 4277-4287, 2004.

Ermler, S., Krunic, 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. & Cziepluch, 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.