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

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

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, embryogenesis, 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 as a comprehensive evaluation of the results.
The Complexity of Cytogenetic Diagnostics

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

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).
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 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 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!
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
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 of 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
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 of 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

---

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 (Seattle, 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 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 (Seattle, 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 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

E-commerce:
- 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
The NEW Standard for Genomic Research and Diagnostics!

Ready to be Valorized!!!

2007

Basic Corporate Licenses

2008

White Label Products & Advertising

2009

e Commerce

Content Sale
The 3D Genome Viewer
The NEW Standard for Genomic Research and Diagnostics!
Ready to be Valorized!!!

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!
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!
Approaching the Sequential and Three-Dimensional Organization

Of

Genomes

Knoch, T. A.


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 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 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 holistically 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, autofluorescent 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**


