EpiGenSys - Systems Biological Determination of the Epigenomic Structure Function Relation

Nucleosomal Association Changes Intra/Inter Chromosomal Architecture Transcriptional Structure Relationship

Simulations of Nucleosomal/Chromatin Fiber/Chromosome Architecture & Dynamics

System Biological/Medical Result Integration via the GLOBE 3D Genome Platform

Tobias A. Knoch, Peter R. Cook, Karsten Rippe, Gernot Längst, Gero Wedemann, & Frank G. Grosveld

Sir William Dunn School of Pathology, Genome Organization & Function, NWFIII/Biochemistry, System Engineering, Cell Biology & Genetics, University of Oxford, BioQuant Centre, German Cancer Research Centre, University of Regensburg, University of Applied Sciences Stralsund, Erasmus Medical Centre
EpiGenSys
- Systems Biological Determination of the Epigenomic Structure Function Relation

Nucleosomal Association Changes
Intra/Inter Chromosomal Architecture
Transcriptional Structure Relationship
Simulations of Nucleosomal / Chromatin Fiber / Chromosome Architecture & Dynamics
System Biological/Medical Result Integration via the GLOBE 3D Genome Platform

Tobias A. Knoch
Biophysical Genomics & Erasmus Computing Grid

Peter R. Cook, Karsten Rippe, Gernot Längst, Gero Wedemann, & Frank G. Grosveld
Sir William Dunn School of Pathology, Genome Organization & Function, NWFIII/Biochemistry, System Engineering and Information Management, Cell Biology & Genetics - Clinical Genetics & Virology

University of Oxford, BioQuant Centre / German Cancer Research Centre
University of Regensburg, University of Applied Sciences Stralsund, Erasmus Medical Centre
EpiGenSys
- Systems Biological Determination of the Epigenomic Structure Function Relation

Nucleosomal
Intra/Interchromosomal Architecture & Dynamics

Simulations of Nucleosomal
Intra/Interchromosomal Architecture & Dynamics
System Biological/Medical Result Integration via the GLOBE 3D Genome Platform

Towards a Holistic Understanding of Genomes!
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 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 a comprehensive evaluation of the results.
EpiGenSys

Systems Biological/Medical Determination of the Epigenomic Structure-Function Relation in:
i) the Beta-Globin locus, ii) the Immuno Globin loci, iii) the SAMD4 region, and iv) the Prader-Willi / Angelmann Syndrom region, in mouse and human active and inactive cell states and their global context.
EpiGenSys

From Sequence to Morphology Towards an Advanced Genome Understanding!

Determination of the Epigenomic Structure-Function Relations in:

i) the Beta-Globin locus,
ii) the Immuno Globin loci,
iii) the SAMD4 region,
iv) the Prader-Willi/ Angelman Syndrome region, in mouse and human active and inactive cell states and their global context.
Nucleosomal Association Changes

Nucleosomal association in relation to the DNA sequence, epigenetic modifications and the activity of ATP-driven chromatin remodelling complexes using high-throughput sequencing. The resulting localization probability maps will be evaluated by a novel combination of analysis and generic data ontologies.
Nucleosomal Association Changes

Nucleosomal association in relation to the DNA sequence, epigenetic modifications and the activity of ATP-driven chromatin remodelling complexes using high-throughput sequencing. The resulting localization probability maps will be evaluated by a novel combination of analysis and generic data ontologies.
Nucleosomal Association Changes

Nucleosomal association in relation to the DNA sequence, epigenetic modifications and the activity of ATP-driven chromatin remodelling complexes using high-throughput sequencing. The resulting location probability maps will be evaluated by a novel combination of analysis and generic data ontologies.

Improved Understanding of Chromatin Morphogenesis & Fiber Formation!
Intra/Inter Chromosomal Architecture

Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure will be derived, i.e. its folding into loops and loop clusters.
Intra/Inter Chromosomal Architecture

Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure will be derived, i.e. its folding into loops and loop clusters.
Intra/Inter Chromosomal Architecture

Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure will be derived, i.e. its folding into loops and loop clusters.
Intra/Inter Chromosomal Architecture

Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure will be derived, i.e. its folding into loops and loop clusters.
Intra/Inter Chromosomal Architecture

Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure will be derived, i.e. its folding into loops and loop clusters.
Intra/inter chromosomal contacts will be determined using a combination of chromosome conformation capture technology and high-throughput deep sequencing. From the interaction maps, 3D chromatin conformations and its higher-order structure will be derived, i.e., its folding into loops and loop clusters.
Transcription rates will be determined by qRT-PCR, RNA and DNA FISH using intronic probes and high-resolution laser scanning and single molecule imaging. Transcription-dependent changes of active and inactive loci compared resulting in the transcription structure-function relationship.
Transcription rates will be determined by qRT-PCR, RNA and DNA FISH using intronic probes and high-resolution laser scanning and single molecule imaging. Transcription-dependent changes of active and inactive loci compared resulting in the transcription structure-function relationship.
Transcription Structure-Function Relationship

Transcription rates will be determined by qRT-PCR, RNA and DNA FISH using intronic probes and high-resolution laser scanning and single molecule imaging. Transcription-dependent changes of active and inactive loci compared resulting in the transcription structure-function relationship.
Transcription Structure-Function Relationship

Transcription rates will be determined by qRT-PCR, RNA and DNA FISH using intronic probes and high-resolution laser scanning and single molecule imaging. Transcription-dependent changes of active and inactive loci compared resulting in the transcription structure-function relationship.
Transcription rates will be determined by qRT-PCR, RNA and DNA FISH using intronic probes and high-resolution laser scanning and single molecule imaging. Transcription-dependent changes of active and inactive loci compared resulting in the transcription structure-function relationship.
Simulation of Nucleosomes, Chromatin, & Nuclei

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches we will simulate nucleosomes, chromatin fibers, chromosomes and whole nuclei with unprecedented resolution, resulting in a virtual multi-scale model of mouse and human genomes.
Simulation of Nucleosomes, Chromatin, & Nuclei

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches we will simulate nucleosomes, chromatin fibers, chromosomes and whole nuclei with unprecedented resolution, resulting in a virtual multi-scale model of mouse and human genomes.
Simulation of Nucleosomes, Chromatin, & Nuclei

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches we will simulate nucleosomes, chromatin fibers, chromosomes and whole nuclei with unprecedented resolution, resulting in a virtual multi-scale model of mouse and human genomes.
Simulation of Nucleosomes, Chromatin, & Nuclei

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches we will simulate nucleosomes, chromatin fibers, chromosomes and whole nuclei with unprecedented resolution, resulting in a virtual multi-scale model of mouse and human genomes.

Determination of the Transcription Structure-Function Relationship!
Systems Biological Result Integration via the GLOBE 3D Genome Platform

All results will be integrated using our GLOBE 3D Genome Platform, established for analysis, manipulation and understanding of multi-dimensional complex genome wide data. Thus in reiterative cycles between experiments and simulations a systems biological/medical genome model will be achieved.
Systems Biological Result Integration via the GLOBE 3D Genome Platform

All results will be integrated using our GLOBE 3D Genome Platform, established for analysis, manipulation and understanding of multi-dimensional complex genome wide data. Thus in reiterative cycles between experiments and simulations a systems biological/medical genome model will be achieved.
Systems Biological Result Integration via the GLOBE 3D Genome Platform

All results will be integrated using our GLOBE 3D Genome Platform, established for analysis, manipulation and understanding of multi-dimensional complex genome wide data. Thus in reiterative cycles between experiments and simulations a systems biological/medical genome model will be achieved.
Systems Biological Result Integration via the GLOBE 3D Genome Platform

All results will be integrated using our GLOBE 3D Genome Platform, established for analysis, manipulation and understanding of multi-dimensional complex genome wide data. Thus in reiterative cycles between experiments and simulations a systems biological/medical genome model will be achieved.

To Bring Together What belongs Systems Biologically/Medically Together!
Acknowledgements

Thanks go to all those people who supported this work in the last decades, the institutions providing their infrastructure, and the national and international computing infrastructures.

Special thanks go to the reviewers, the EraSysBio Plus initiative and the national and EU funding bodies.
Acknowledgements

Thanks go to all those people who supported this work in the last decades, the institutions providing infrastructure, and the national and international computing infrastructures.

Special thanks go to the reviewers, the EraSysBio Plus initiative and the national and EU funding bodies.

Towards a Holistic Genome Understanding by a Systems Biological/Medical Model for Research, Diagnostics and Treatment!
Despite our knowledge of the sequence of the human genome, the relation of its three-dimensional dynamic architecture with its function – the storage and expression of genetic information – remains one of the central unresolved issues of our age. It became very clear meanwhile that this link is crucial for the entire holistic function of the genome on all genomic coding levels from the DNA sequence to the entire chromosomes. To fulfil the dreams for better diagnostics and treatment in the 21st century (e.g. by gene therapy by inserting a gene into a new global context), we propose here in a unique interdisciplinary project to combine experiment with theory to analyze the (epi-)genomic structure function relationships within the dynamic organization of the β-Globin locus, the Immuno Globin loci, and the Tumor Necrosis Factor Alpha regulated SAMD4 region in mouse and human active and inactive cell states, and their global genomic context. The project consists of five work packages (WP1-WP5) and corresponding tasks connected in a system biological approach with iterative use of data, modelling and experiments via a unique data sharing and visualization platform:

In WP1 (Längst, Rippe, Wedemann, Knoch/Grosfeld; T1-T5) to investigate nucleosomal association changes in relation to the DNA sequence and the activity of ATP-driven chromatin remodelling complexes, nucleosome positions will be determined by high-throughput sequencing. The resulting nucleosomal localization probability maps will be evaluated by a novel combination of analysis tools and innovative generic data ontologies. The relation to epigenetic modifications, to the activity of ATP-driven remodelling complexes and compaction degree of nucleosomes will be analysed to understand chromatin morphogenesis and fiber formation. In parallel, in WP2 (Grosfeld/Knoch, Cook, Rippe, Längst; T1-T3) we determine by high-throughput monitoring of intra/inter chromosomal contacts and architecture absolute DNA-DNA interaction probability maps for the individual loci and their global context using a novel chromosome conformation capture approach based on deep sequencing. From these the 3D conformation of the chromatin fiber and its higher-order folding into loops and loop clusters can be derived using algorithms recently developed by us. WP3 (Cook, Grosfeld/Knoch, Längst; T1-T5) focuses on the determination of transcription rates and structure by qRT-PCR, DNA and RNA fluorescence in situ hybridization using intronic probes and high-resolution laser-scanning and single molecule imaging with advanced image analysis tools. Transcription-dependent changes of active and inactive loci as well as rapid synchronous transcription alteration against the unchanged background is one main interest here. This will yield results in a detailed cartography of the structure-transcription-function dependency and its importance.
To rationalize the experimental results theoretically, in **WP4 (Wedemann Knoch/Grosveld, Rippe; T1-T3)** simulations are made of nucleosomal structure, chromatin fiber conformation and chromosomal architecture using parallel and grid super-computers with ~10,000 CPUs. The impact of different nucleosomal positions and epigenetic modifications on the nucleosomal structure and the chromatin fiber conformation will be assessed by novel Monte Carlo approaches. To understand the higher-order architecture Brownian Dynamics simulations of entire cell nuclei with molecular resolution, morphogenetic processes and transcriptional states will be made. This results in a virtual multi-scale model with unseen spatial and time resolution providing novel insights into genome organization. All the resulting virtual architectures will be compared to experiment to prompt again new experiments and **vice versa** in reiterative cycles. In **WP5 (Knoch/Grosveld, Cook, Rippe, Längst, Wedemann; T1-T5)** all partners together will integrate the experimental and theoretic results to achieve a system biological model with existing genome-wide data in the GLOBE 3D Genome Browser – a novel platform developed by us for the analysis, manipulation and understanding of multi-dimensional complex genome wide data in an easy to understand 3D visualization environment. The entire experimental data will be archived, all simulations and analysis on high-performance computing infrastructures (e.g. German MediGRID, Dutch Erasmus Computing Grid) will be controlled by a new management system. Together with a novel generic correlation finder and a process/pathway data base this will result together with our Portable Genome Format (PGF) in a unique system biological platform publicly available to understand genome complexity.

We are strongly convinced, that the reiterative combination of quantitative experiment with theory leads to a virtual system biological model of the (epi-)genomic structure-function relationship with major impact and great valorization opportunities in research, training, diagnosis and treatment due to the uniqueness, novelty, and frontier position (~20 academic, ~10 industry collaborations underway). This is stressed by our i) scientific excellence, ii) interdisciplinary experience, iii) participative (SOP) management, and iv) IP valorization successes, embedded within strong education/training regimes and famous infrastructures. Consequently, our “EpiGenSys” virtual laboratory is a prime example for systems biology combining high-throughput/performance techniques of cell biology, mathematics, physics and informatics to solve one of the most fundamental issues of personalized genomic medicine. Beyond, we believe to make a major contribution to e-Science, e-Health, e-Learning, as well as e-Commerce creating a novel awareness and understanding of genomic complexity within society.

**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 statistical mechanics, genomic uncertainty principle, 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, chromatin cross-linking, chromosome conformation capture (3C), selective high-resolution high-throughput chromosome interaction capture (T2C), 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


