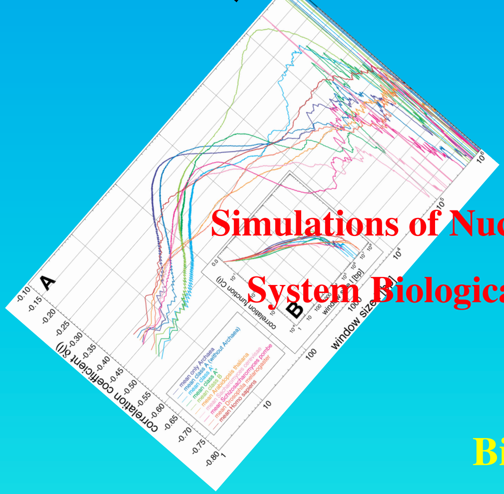
EpiGenSys

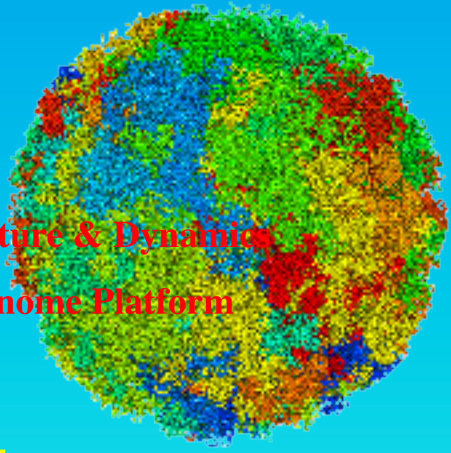
-

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

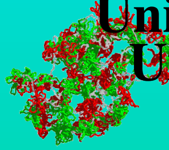
Tobias A. Knoch

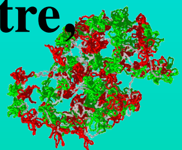
Biophysical Genomics & Erasmus Computing Grid

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

Sir William Dunn School of Pathology, Genome Organization & Function, NWFIII/Biochemistry, System Engeneering 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





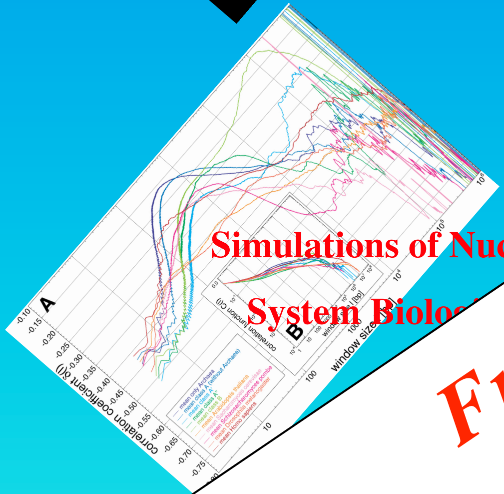
EpiGenSys

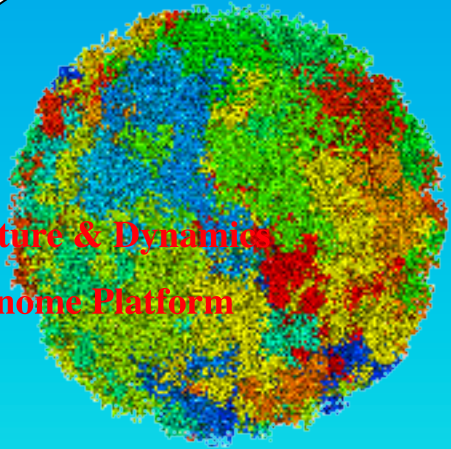
-

Systems Biological Determinatio

of the

Epigenomic Structure F





Nucleosomal

Intra/Inte re

Tra tionship

Simulations of Nucleo Chromosome Architecture & Dynamics

System Biolo ation via the GLOBE 3D Genome Platform

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

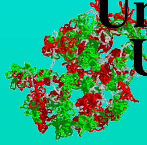
obias A. Knoch

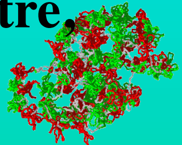
l Genomics & Erasmus Computing Grid

Pete ippe, Gernot Längst, Gero Wedemann, & Frank G. Grosveld

Sir W of Pathology, Genome Organization & Function, NWFIII/Biochemistry, System
Eng nformation Management, Cell Biology & Genetics - Clinical Genetics & Virology

Unive y of Oxford, BioQuant Centre / German Cancer Research Centre,
University of Regensburg, University of Applied Sciences Stralsund,
Erasmus Medical Centre

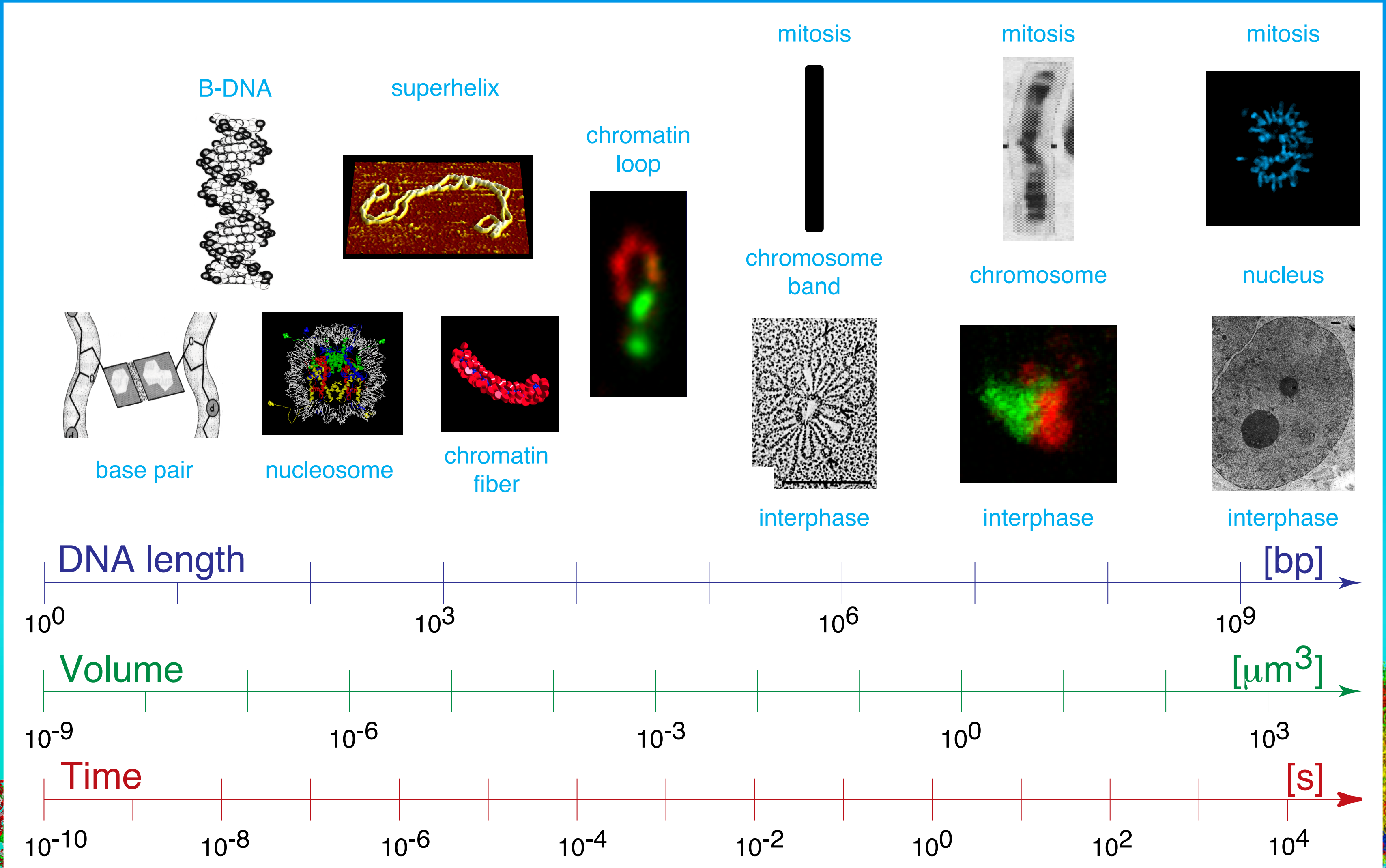




Bundesministerium für Bildung und Forschung
BBSRC bioscience for the future
NWO Nederlandse or Wetenschappelijk Onderzoek
Consorti
ERASysBio ERA-NET for Systems Biology
University Heidelberg
dkfz. German Cancer Research Center
FORD
UR University of Regensburg
fachhochschule stralsund University of Applied Sciences

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?



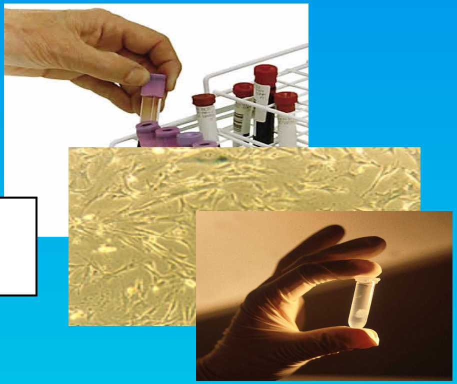
Complexity of e.g. Cytogenetic Diagnostics & Treatment

The process of cytogenetic analysis requires proper patient and sample analysis

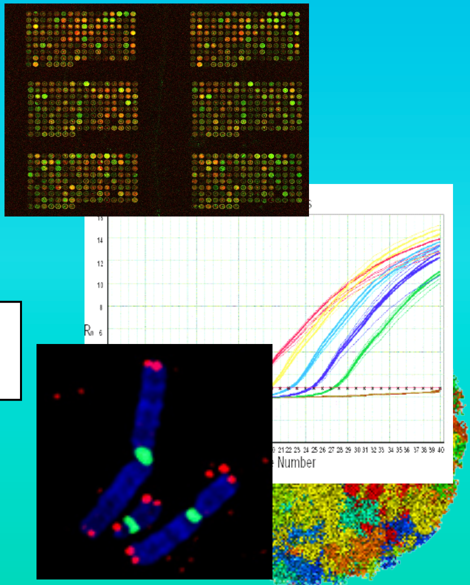
as well asa comprehensive evaluation of the results.

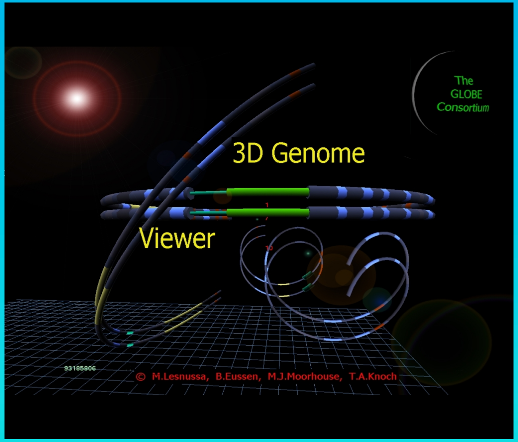
Patient

Sample



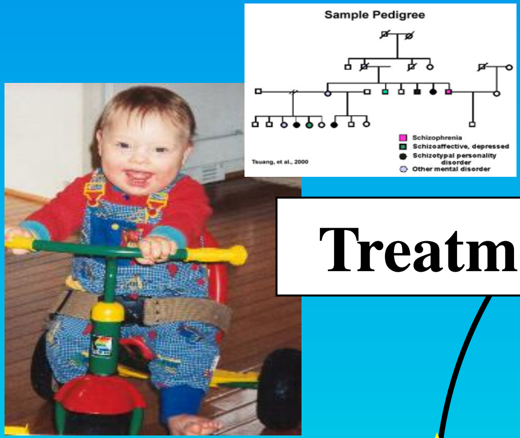
Analysis





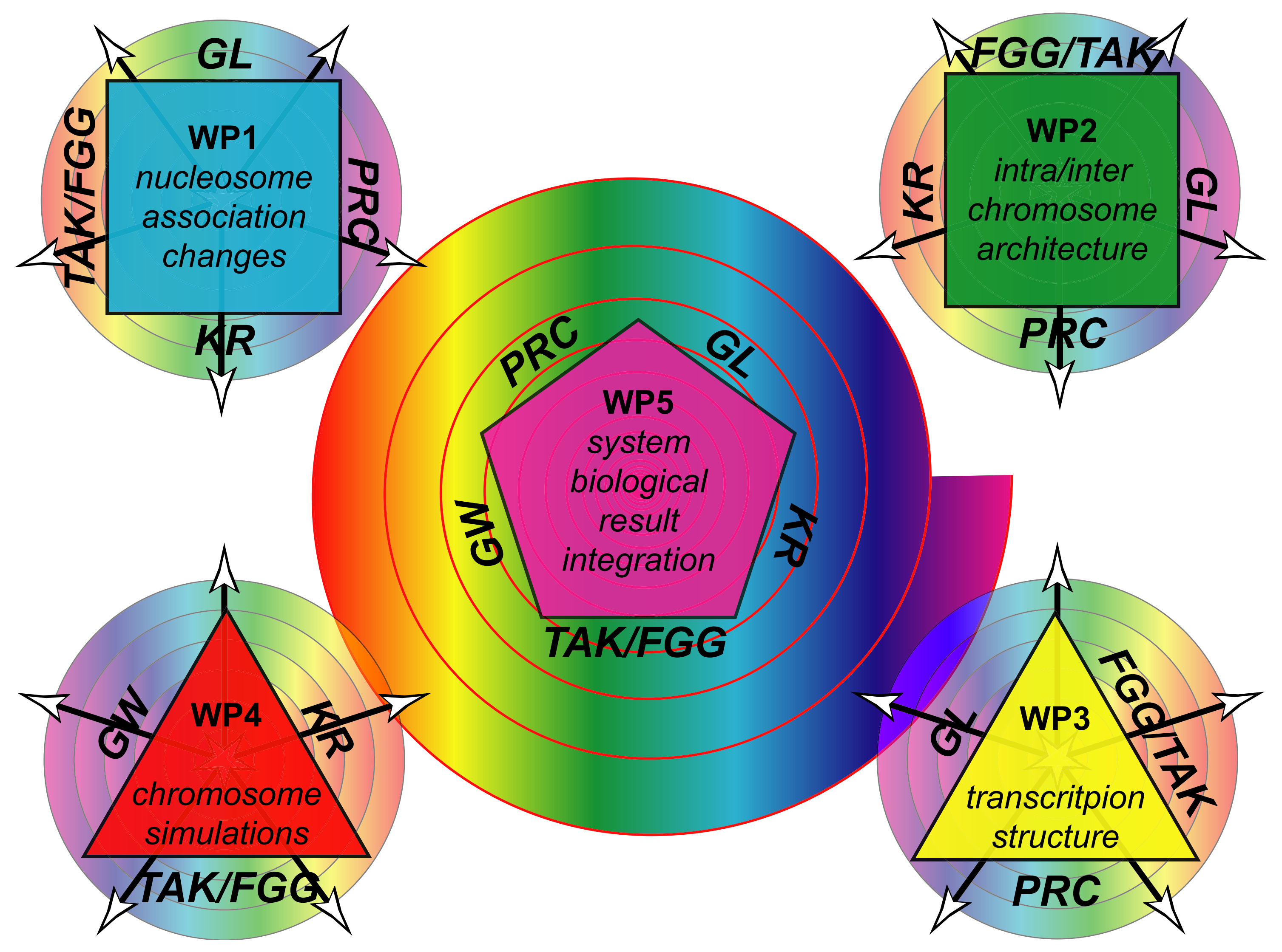
Diagnosis

Treatment



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.



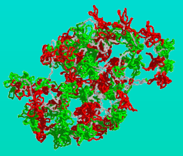
WP1
nucleosome association changes
GL
PRC
KR
TAK/FGG

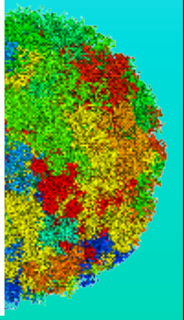
WP2
intra/inter chromosome architecture
FGG/TAK
GL
PRC
KR

WP5
system biological result integration
PRC
GL
KR
TAK/FGG
GW

WP4
chromosome simulations
KR
TAK/FGG
GW

WP3
transcritpion structure
FGG/TAK
PRC
GL





Work Together and Communication in EpiGenSys

The work packages are implemented in such a way that they utilize the established expertise of individual partners (each with their own established network of contacts) while providing maximum benefit to the groups. Maximum output is guarantied by our virtual laboratory communication management.

WP1: Nucleosomal association changes (Längst, Rippe, Wedemann, Knoch/Grosveld; T1-T5)

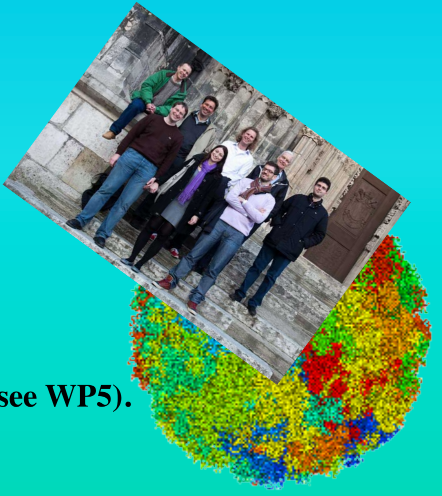
WP2 Intra/inter chromosomal architecture (Grosveld/Knoch, Cook, Rippe, Längst; T1-T3)

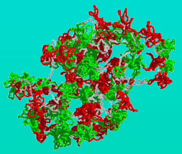
WP3 Transcription structure relationship (Cook, Grosveld/Knoch, Längst; T1-T4)

WP4 Simulations of nucleosomal, chromatin fiber and chromosome architecture and their dynamics (Wedemann, Knoch/Grosveld, Rippe; T1-T3)

WP5 System biological result integration via the GLOBE 3D Genome Platform (Knoch/Grosveld, Cook, Rippe, Längst, Wedemann; T1-T5):

1. Major meetings per year where all participants meet:
    - ❖ Den Haag from 7th to 8th July 2010.
    - ❖ Regensburg from 6th to 8th April 2011.
    - ❖ Vienna from 13th to 15th September 2011.
    - ❖ Heidelberg from 9th to 10th May 2012.
    - ❖ Oxford from 3rd to 5th July 2012.
    - ❖ Oxford from 28th to 30th July 2013.
2. A monthly online conference according to theme.
3. Weekly conferences of the work force related to the specific tasks.
4. Regular work meetings in participant labs with several exchanges.
5. Use of a web-based communication platforms with project database and forum (see WP5).
6. Public conference will be discussed and planned in Oxford in July.

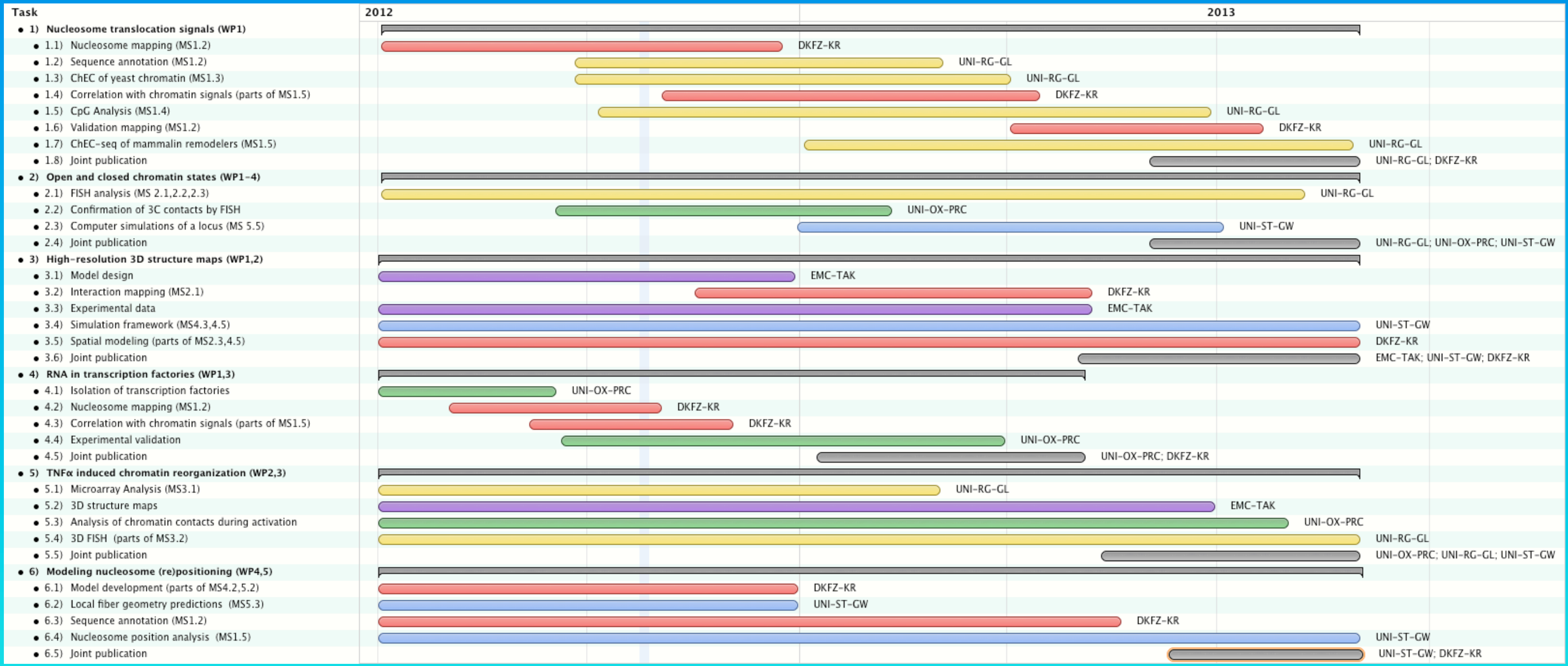




M&D Timeplan in EpiGenSys

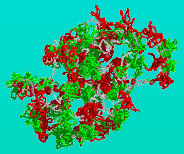
The time plane according to the work packages is currently on track

in respect to the prolongation.

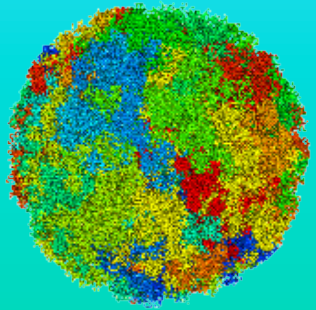


Task	2012–2013
1) Nucleosome translocation signals (WP1)	
1.1) Nucleosome mapping (MS1.2)	DKFZ-KR
1.2) Sequence annotation (MS1.2)	UNI-RG-GL
1.3) ChEC of yeast chromatin (MS1.3)	UNI-RG-GL
1.4) Correlation with chromatin signals (parts of MS1.5)	DKFZ-KR
1.5) CpG Analysis (MS1.4)	UNI-RG-GL
1.6) Validation mapping (MS1.2)	DKFZ-KR
1.7) ChEC-seq of mammalin remodelers (MS1.5)	UNI-RG-GL
1.8) Joint publication	UNI-RG-GL; DKFZ-KR
2) Open and closed chromatin states (WP1-4)	
2.1) FISH analysis (MS 2.1,2.2,2.3)	UNI-RG-GL
2.2) Confirmation of 3C contacts by FISH	UNI-OX-PRC
2.3) Computer simulations of a locus (MS 5.5)	UNI-ST-GW
2.4) Joint publication	UNI-RG-GL; UNI-OX-PRC; UNI-ST-GW
3) High-resolution 3D structure maps (WP1,2)	
3.1) Model design	EMC-TAK
3.2) Interaction mapping (MS2.1)	DKFZ-KR
3.3) Experimental data	EMC-TAK
3.4) Simulation framework (MS4.3,4.5)	UNI-ST-GW
3.5) Spatial modeling (parts of MS2.3,4.5)	DKFZ-KR
3.6) Joint publication	EMC-TAK; UNI-ST-GW; DKFZ-KR
4) RNA in transcription factories (WP1,3)	
4.1) Isolation of transcription factories	UNI-OX-PRC
4.2) Nucleosome mapping (MS1.2)	DKFZ-KR
4.3) Correlation with chromatin signals (parts of MS1.5)	DKFZ-KR
4.4) Experimental validation	UNI-OX-PRC
4.5) Joint publication	UNI-OX-PRC; DKFZ-KR
5) TNFα induced chromatin reorganization (WP2,3)	
5.1) Microarray Analysis (MS3.1)	UNI-RG-GL
5.2) 3D structure maps	EMC-TAK
5.3) Analysis of chromatin contacts during activation	UNI-OX-PRC
5.4) 3D FISH (parts of MS3.2)	UNI-RG-GL
5.5) Joint publication	UNI-OX-PRC; UNI-RG-GL; UNI-ST-GW
6) Modeling nucleosome (re)positioning (WP4,5)	
6.1) Model development (parts of MS4.2,5.2)	DKFZ-KR
6.2) Local fiber geometry predictions (MS5.3)	UNI-ST-GW
6.3) Sequence annotation (MS1.2)	DKFZ-KR
6.4) Nucleosome position analysis (MS1.5)	UNI-ST-GW
6.5) Joint publication	UNI-ST-GW; DKFZ-KR

THE TRUTH IS AND WE HAVE TO ADDMIT:

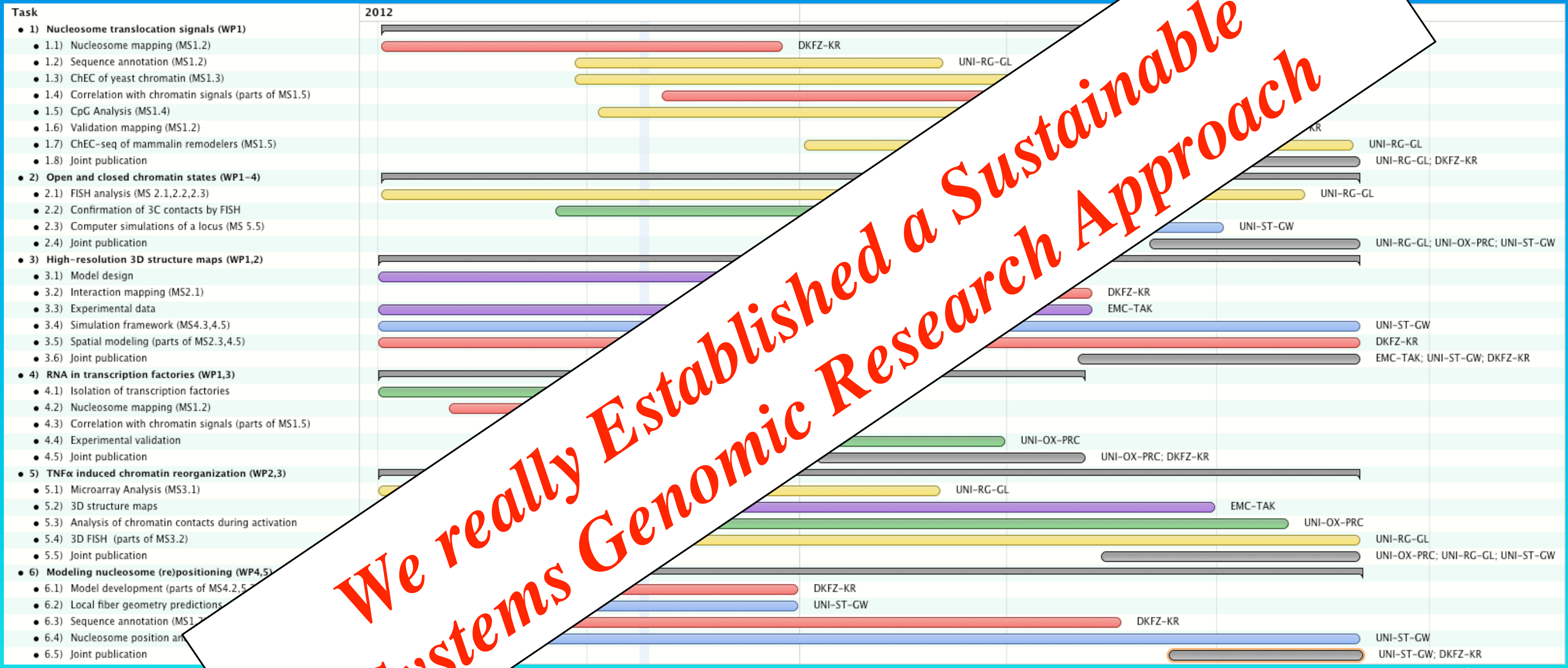


WE PRODUCED AND STILL PRODUCE SUCH A WEALTH OF DATA THAT STILL IN 2-3 YEARS WE WILL ANALYSE AND PUBLISH…



M&D Timeplan in EpiGenSys

The time plane according to the work packages is currently on track in respect to the prolongation.

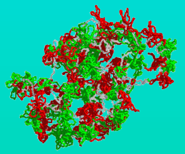


Task

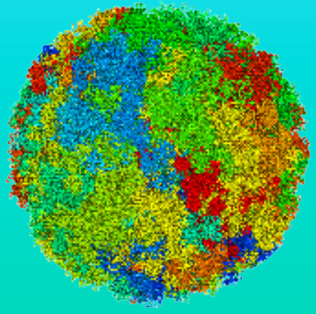
- 1) **Nucleosome translocation signals (WP1)**
  - 1.1) Nucleosome mapping (MS1.2)
  - 1.2) Sequence annotation (MS1.2)
  - 1.3) ChEC of yeast chromatin (MS1.3)
  - 1.4) Correlation with chromatin signals (parts of MS1.5)
  - 1.5) CpG Analysis (MS1.4)
  - 1.6) Validation mapping (MS1.2)
  - 1.7) ChEC-seq of mammalin remodelers (MS1.5)
  - 1.8) Joint publication
- 2) **Open and closed chromatin states (WP1-4)**
  - 2.1) FISH analysis (MS 2.1,2.2,2.3)
  - 2.2) Confirmation of 3C contacts by FISH
  - 2.3) Computer simulations of a locus (MS 5.5)
  - 2.4) Joint publication
- 3) **High-resolution 3D structure maps (WP1,2)**
  - 3.1) Model design
  - 3.2) Interaction mapping (MS2.1)
  - 3.3) Experimental data
  - 3.4) Simulation framework (MS4.3,4.5)
  - 3.5) Spatial modeling (parts of MS2.3,4.5)
  - 3.6) Joint publication
- 4) **RNA in transcription factories (WP1,3)**
  - 4.1) Isolation of transcription factories
  - 4.2) Nucleosome mapping (MS1.2)
  - 4.3) Correlation with chromatin signals (parts of MS1.5)
  - 4.4) Experimental validation
  - 4.5) Joint publication
- 5) **TNFα induced chromatin reorganization (WP2,3)**
  - 5.1) Microarray Analysis (MS3.1)
  - 5.2) 3D structure maps
  - 5.3) Analysis of chromatin contacts during activation
  - 5.4) 3D FISH (parts of MS3.2)
  - 5.5) Joint publication
- 6) **Modeling nucleosome (re)positioning (WP4,5)**
  - 6.1) Model development (parts of MS4.2,5
  - 6.2) Local fiber geometry predictions
  - 6.3) Sequence annotation (MS1
  - 6.4) Nucleosome position an
  - 6.5) Joint publication

We really Established a Sustainable Systems Genomic Research Approach

THE TI ND WE HAVE TO ADDMIT:

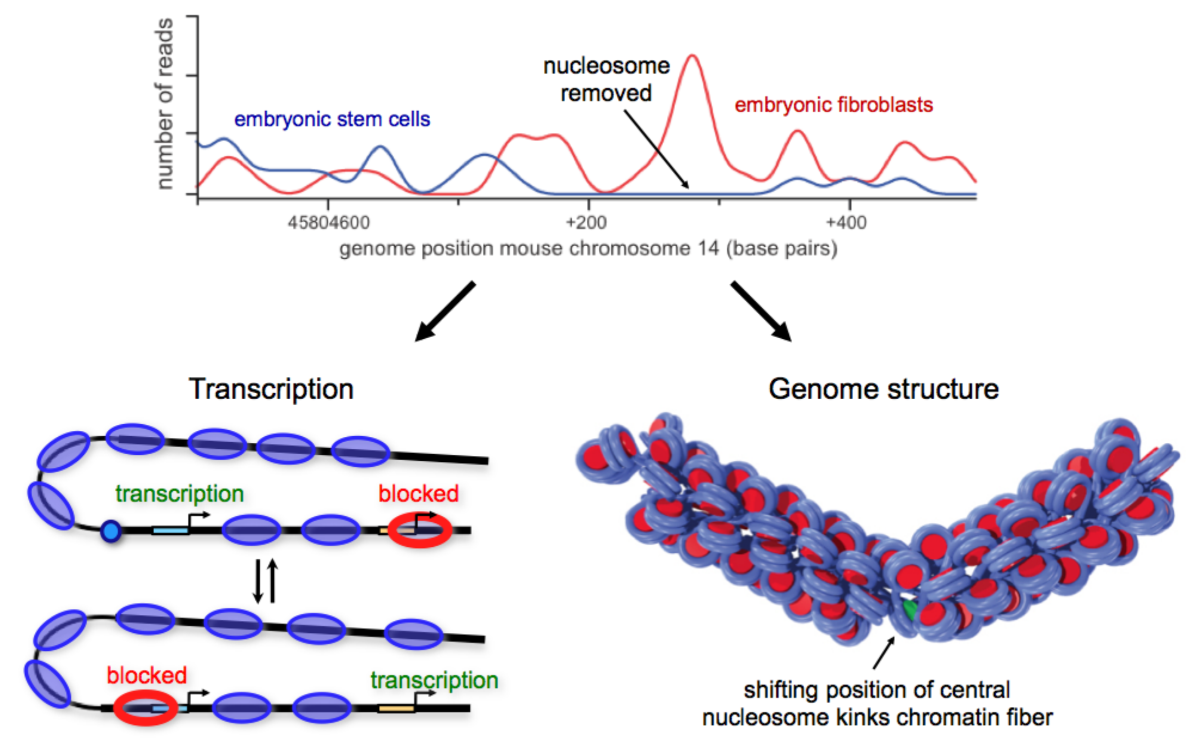


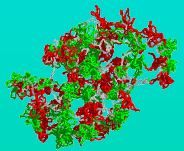
WE PRODUCED AND STILL PRODUCE SUCH A WEALTH OF DATA THAT STILL IN 2-3 YEARS WE WILL ANALYSE AND PUBLISH…

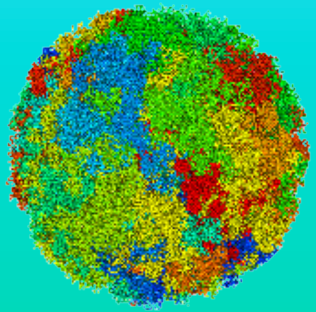


Nucleosomal Association Changes WP1

Nucleosomal association in relation to the DNA sequence, show that nucleosome position and transcription factor binding are linked: Nanog binds to a nucleosome whereas c-Myc binds preferentially in the nucleosomal linker region. For CTCF that is more complicated and after differenciation the association is the opposite.

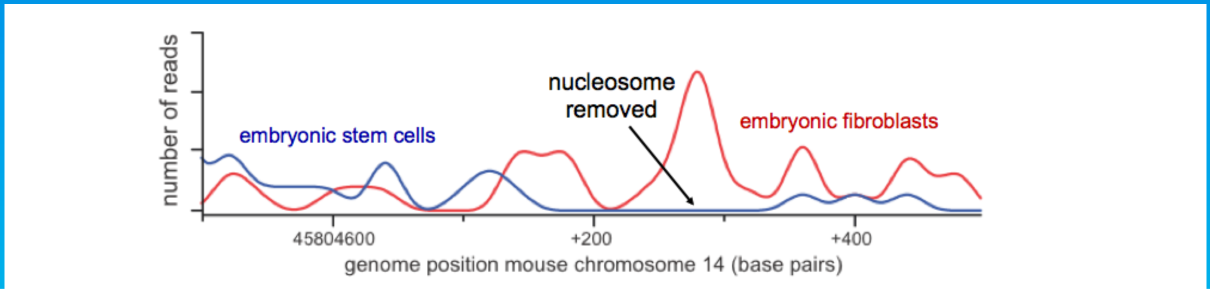


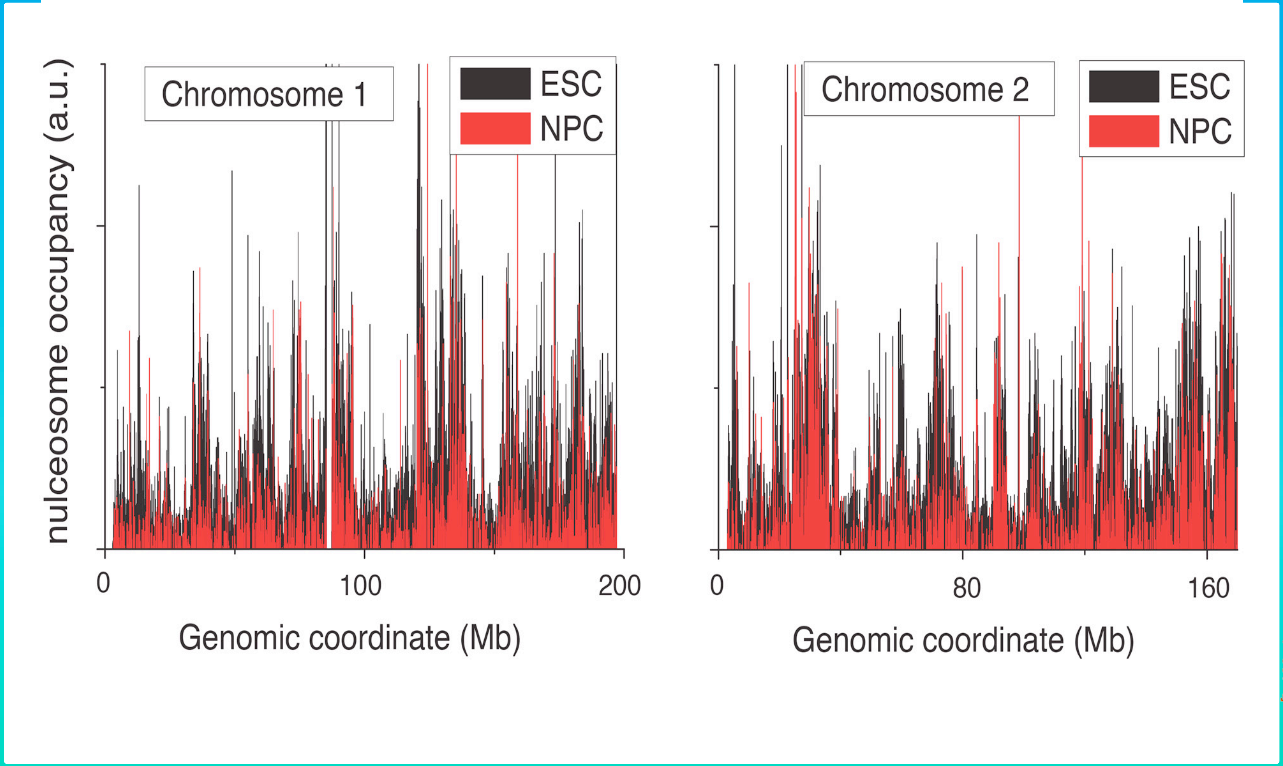


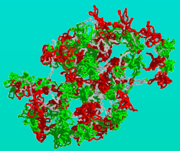


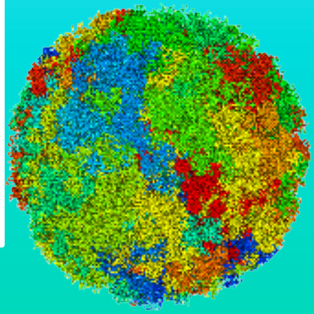
Nucleosomal Association Changes WP1

Nucleosomal association in relation to the DNA sequence, show that nucleosome position and transcription factor binding are linked: Nanog binds to a nucleosome whereas c-Myc binds preferentially in the nucleosomal linker region. For CTCF that is more complicated and after differenciation the association is the opposite.



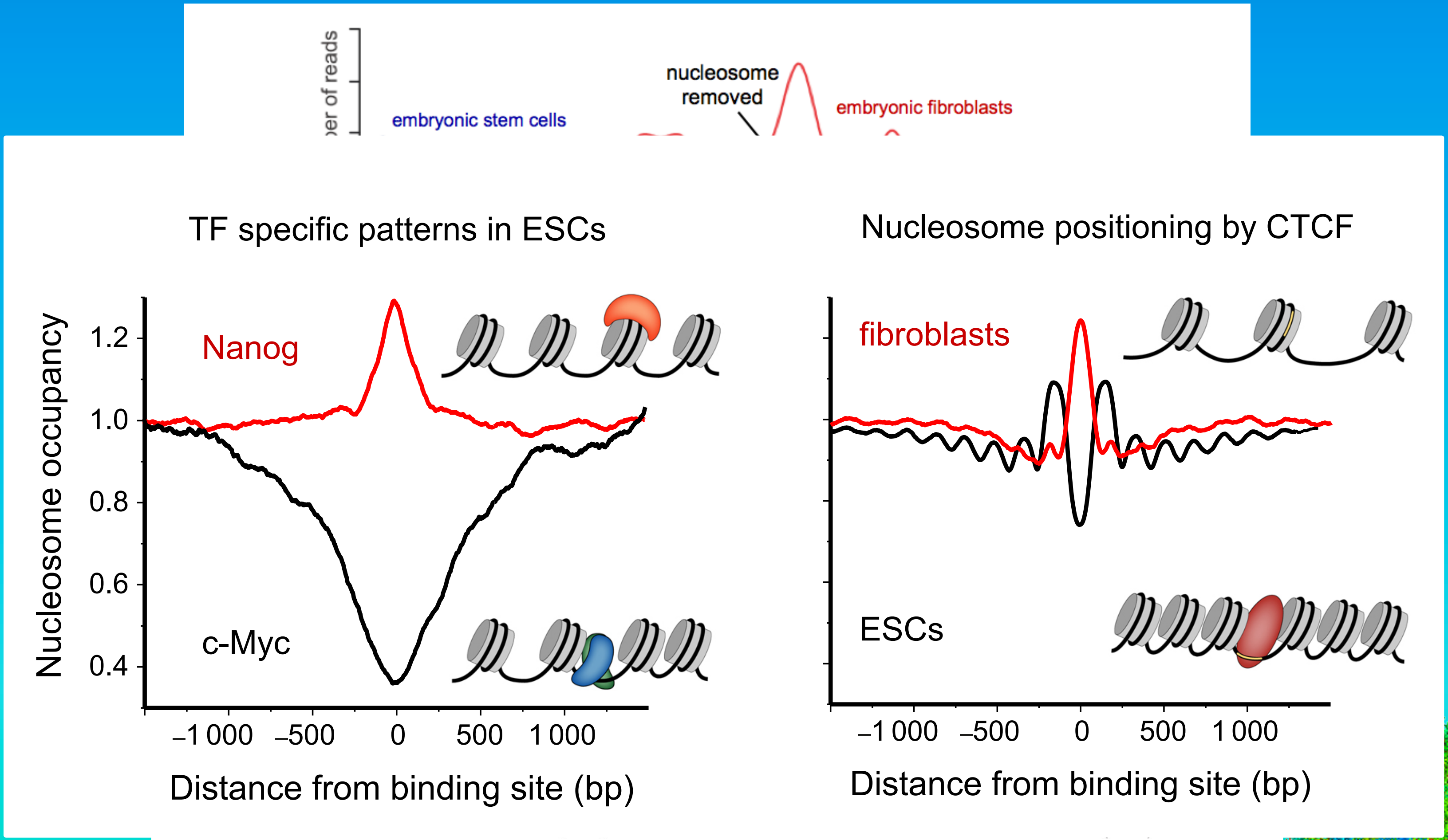


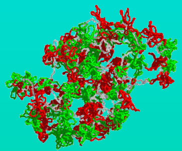


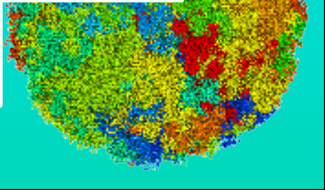


Nucleosomal Association Changes WP1

Nucleosomal association in relation to the DNA sequence, show that nucleosome position and transcription factor binding are linked: Nanog binds to a nucleosome whereas c-Myc binds preferentially in the nucleosomal linker region. For CTCF that is more complicated and after differenciation the association is the opposite.



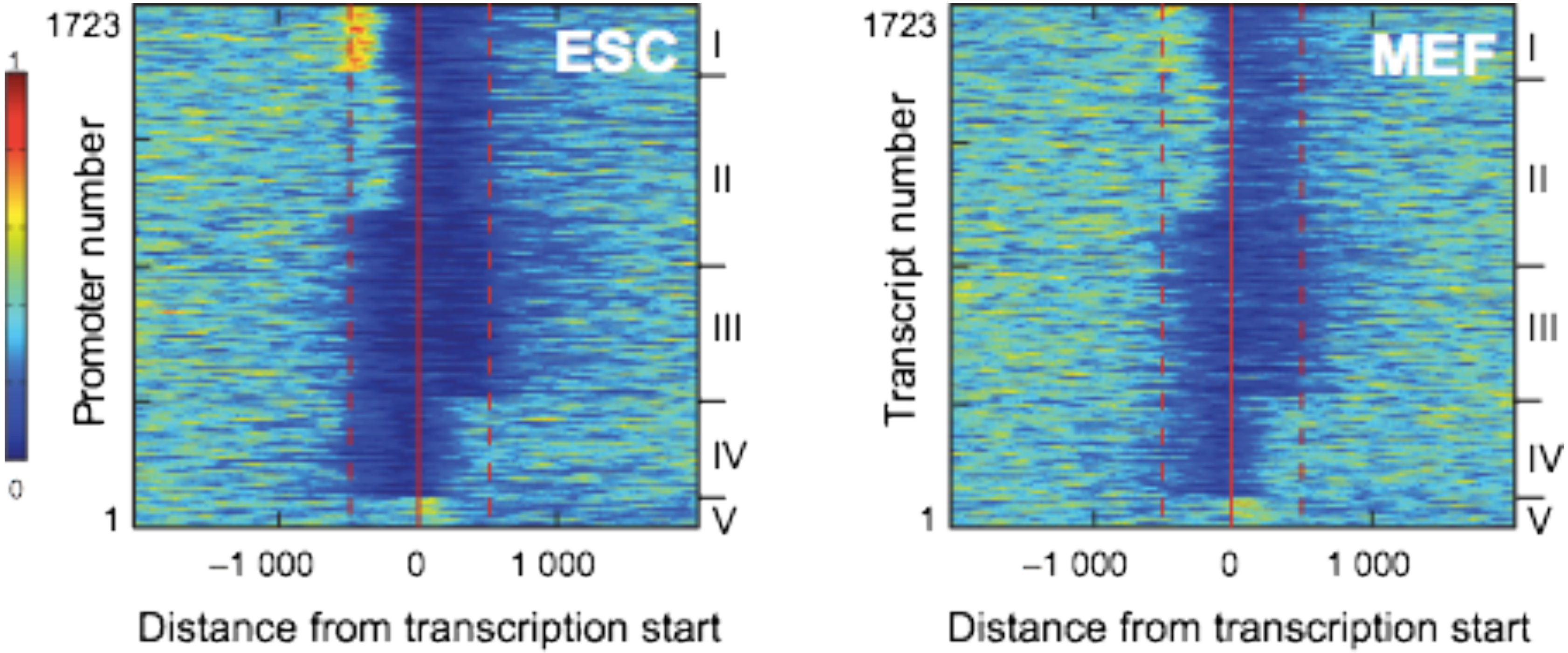


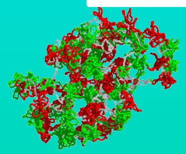


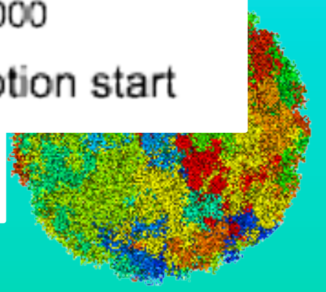
Nucleosomal Association Changes WP1

Nucleosomal association in relation to the DNA sequence, show that nucleosome position and transcription factor binding are linked: Nanog binds to a nucleosome whereas c-Myc binds preferentially in the nucleosomal linker region. For CTCF that is more complicated and after differenciation the association is the opposite.

ESC promoters with H3K4me3 and H3K27me3 marks in cluster 1 become active after removing a nucleosome from position -430.

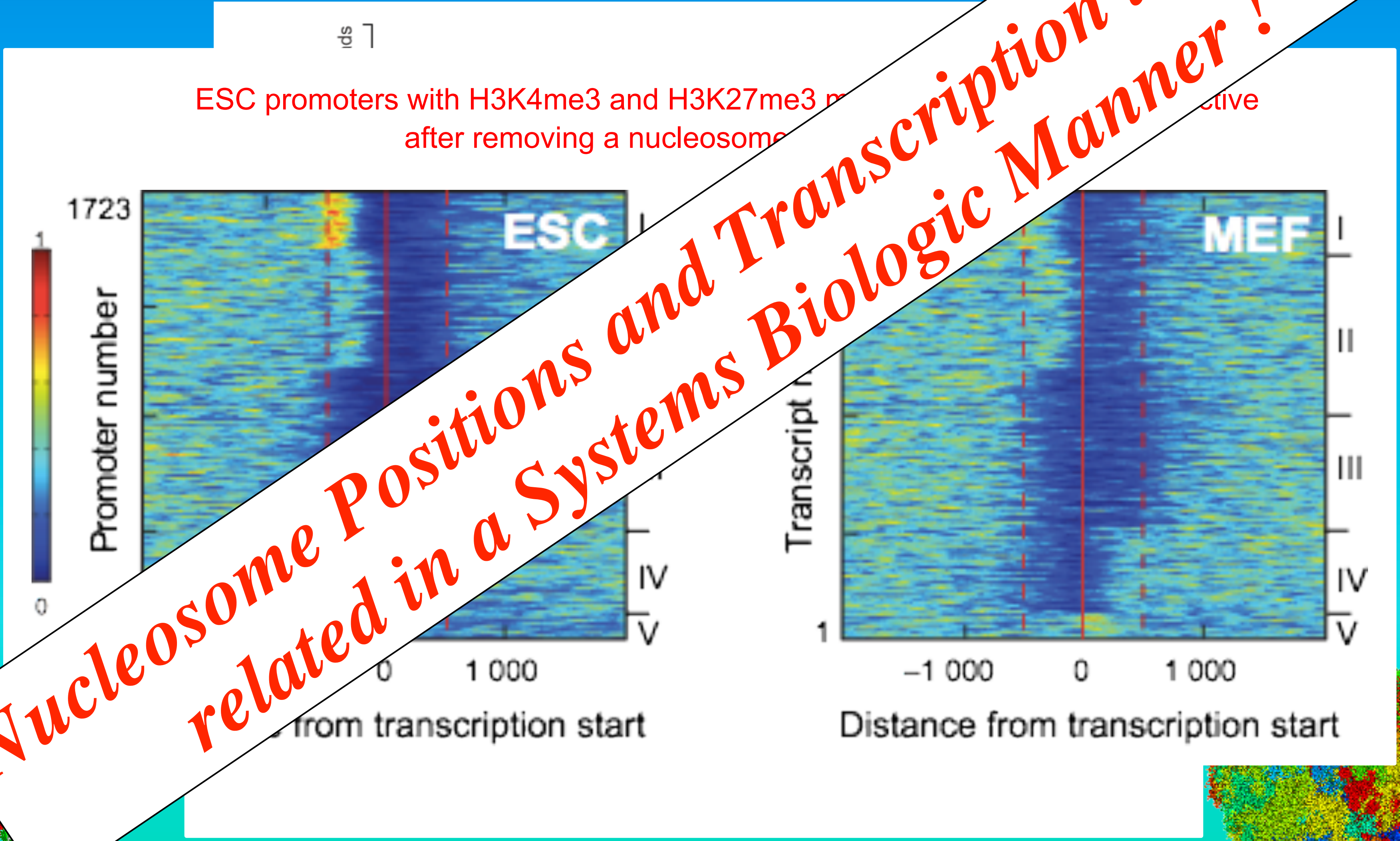






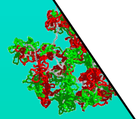
Nucleosomal Association Changes WP1

Nucleosomal association in relation to the DNA sequence, show that nucleosome position and transcription factor binding are linked: Nanog binds to a nucleosome whereas c-Myc binds preferentially in the nucleos linker region. For CTCF that is more complicated and after differenciation the association is the opp



ESC promoters with H3K4me3 and H3K27me3 m ctive after removing a nucleosome

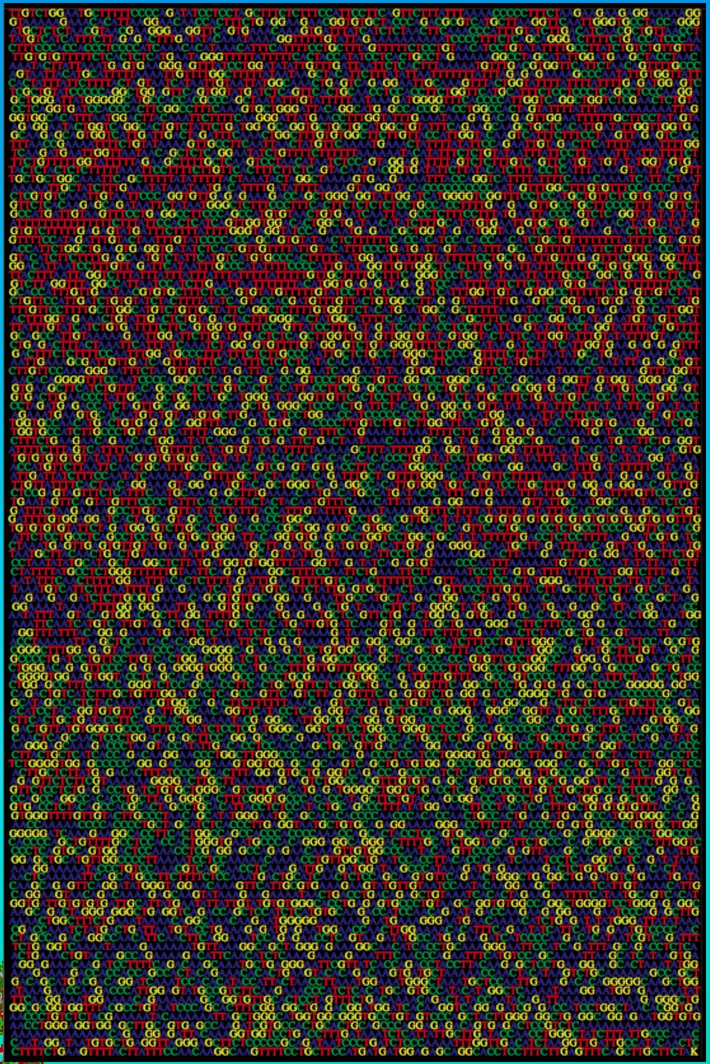
Nucleosome Positions and Transcription is Tightly related in a Systems Biologic Manner !

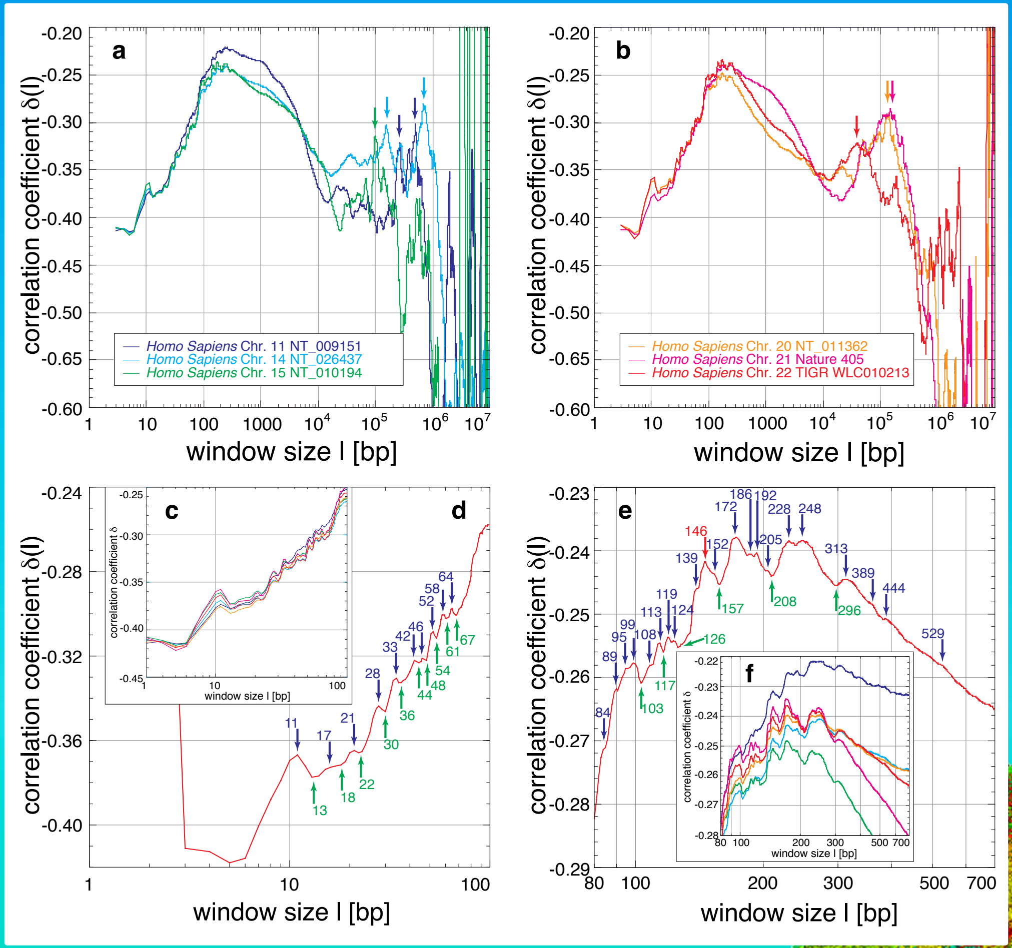


Nucleosomal Positioning Prediction WP1

The DNA sequence is analyzed by the most simplest scaling analysis to find unprejudiced patters as e.g. nucleosome positions as well as chromatin loops and rosettes. The analysis is done using our grid infrastructures and here especially our volunteer grid.

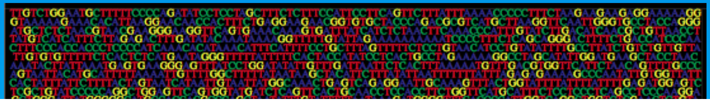
Bundesministerium für Bildung und Forschung
BBSRC
NWO
EpiGenSys Consortium
Systems Structure - Function
ERASysBio
European Comission
Erasmus MC
BioQuant
University Heidelberg
University of Oxford
dkfz. German Cancer Research Center
MediGRID
fachhochschule stralsund

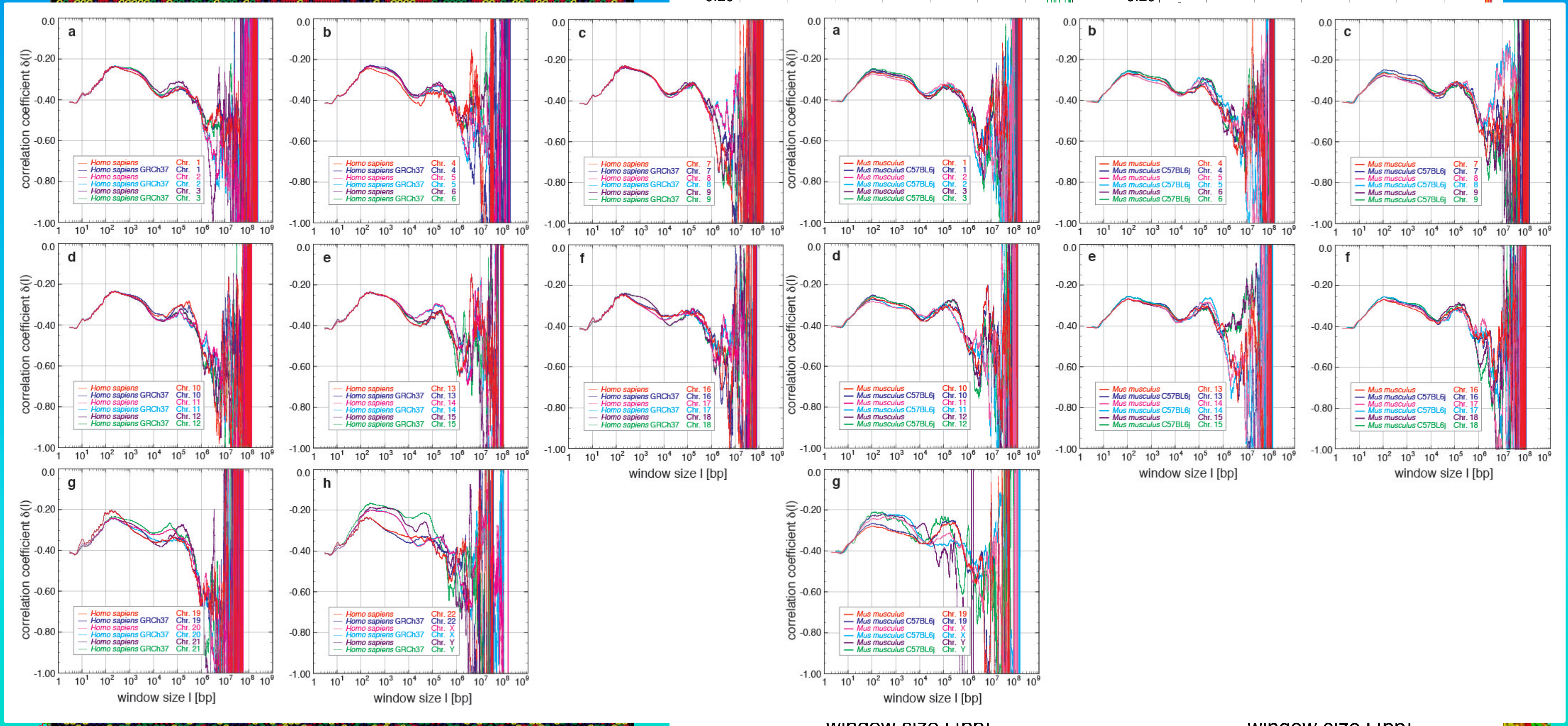


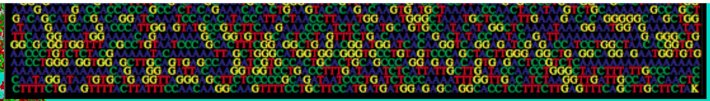


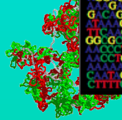
Nucleosomal Positioning Prediction WP1

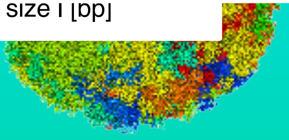
The DNA sequence is analyzed by the most simplest scaling analysis to find unprejudiced patters as e.g. nucleosome positions as well as chromatin loops and rosettes. The analysis is done using our grid infrastructures and here especially our volunteer grid.





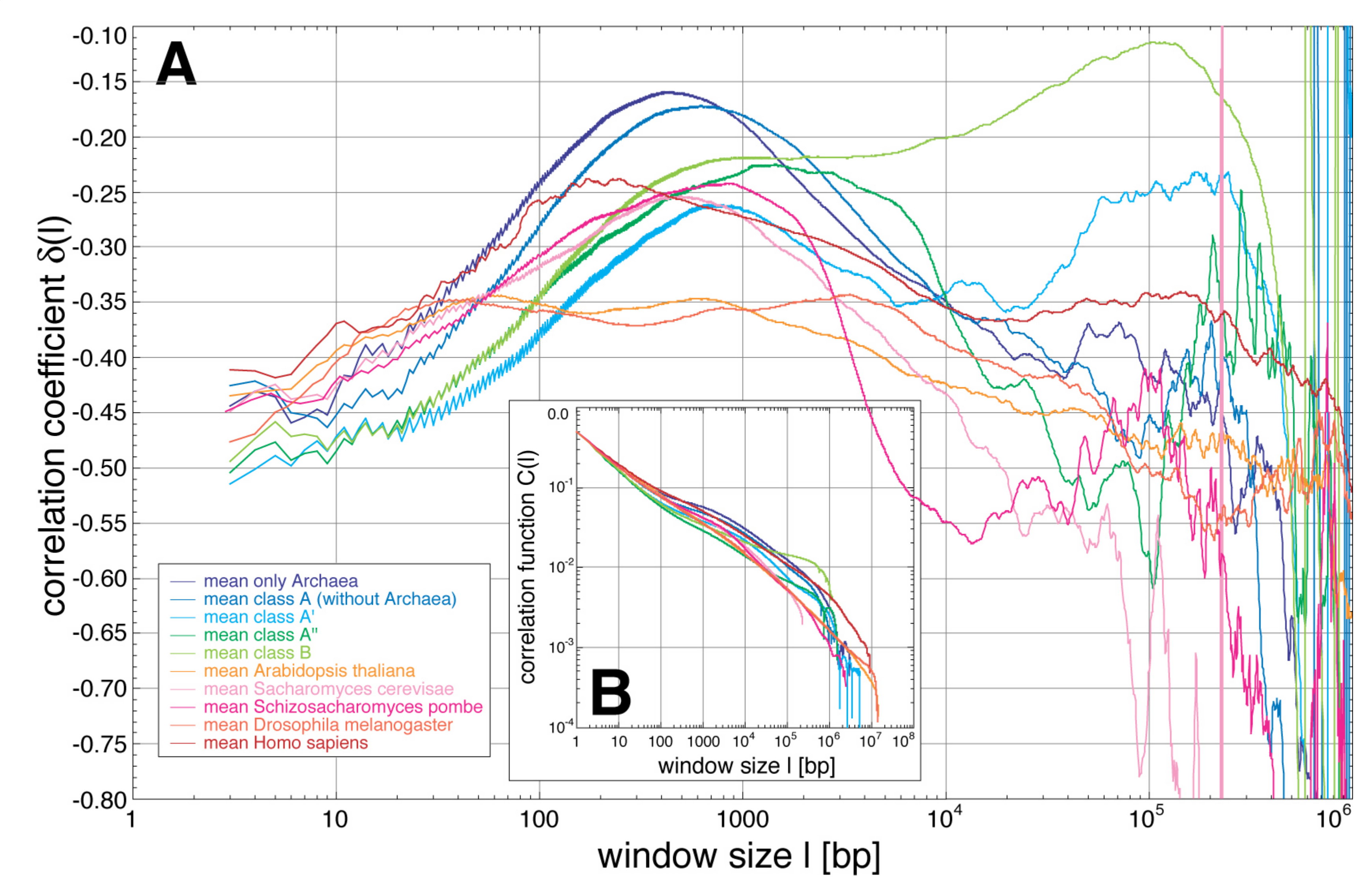






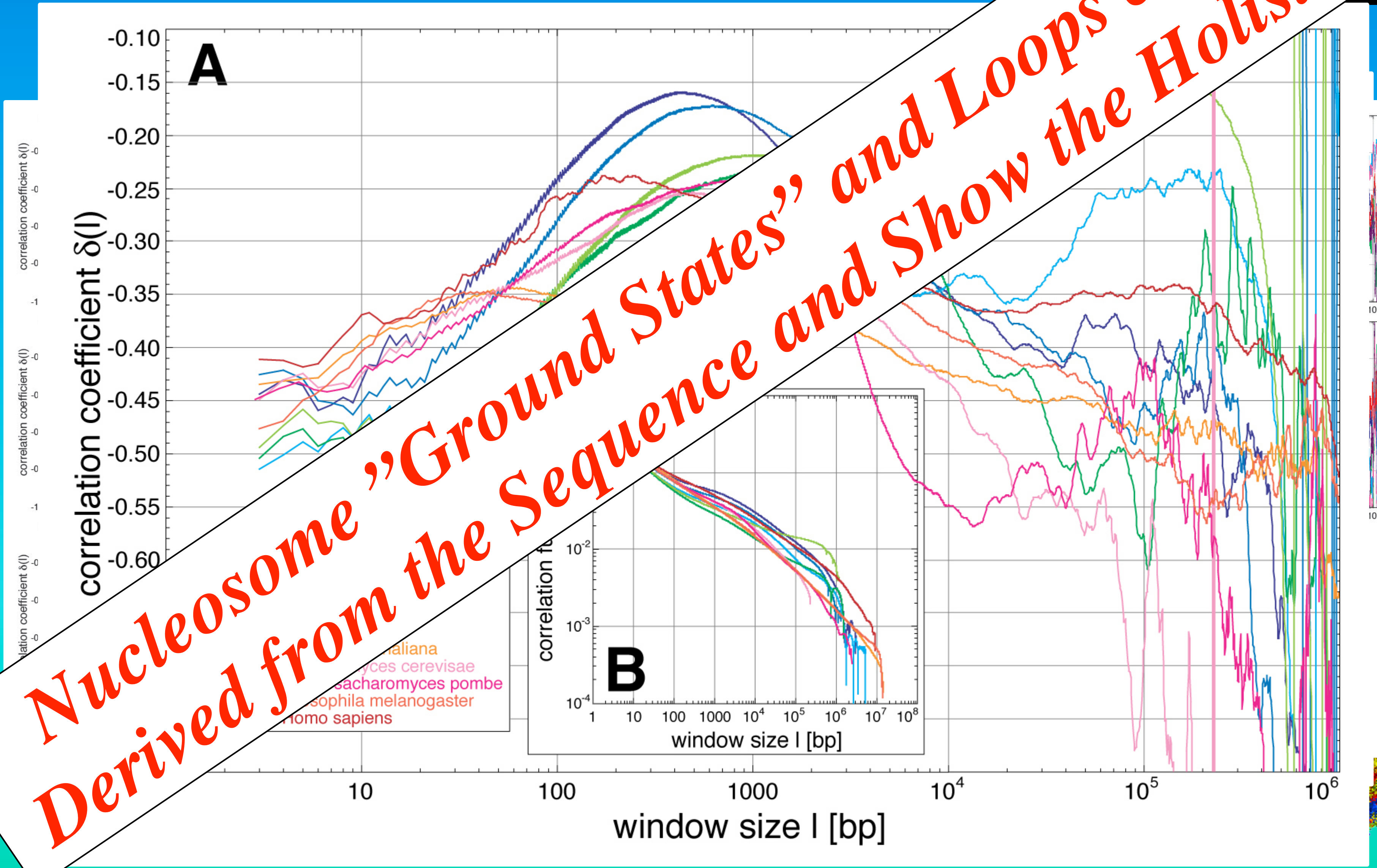
Nucleosomal Positioning Prediction WP1

The DNA sequence is analyzed by the most simplest scaling analysis to find unprejudiced patters as e.g. nucleosome positions as well as chromatin loops and rosettes. The analysis is done using our grid infrastructures and here especially our volunteer grid.



Nucleosomal Positioning Prediction WP1

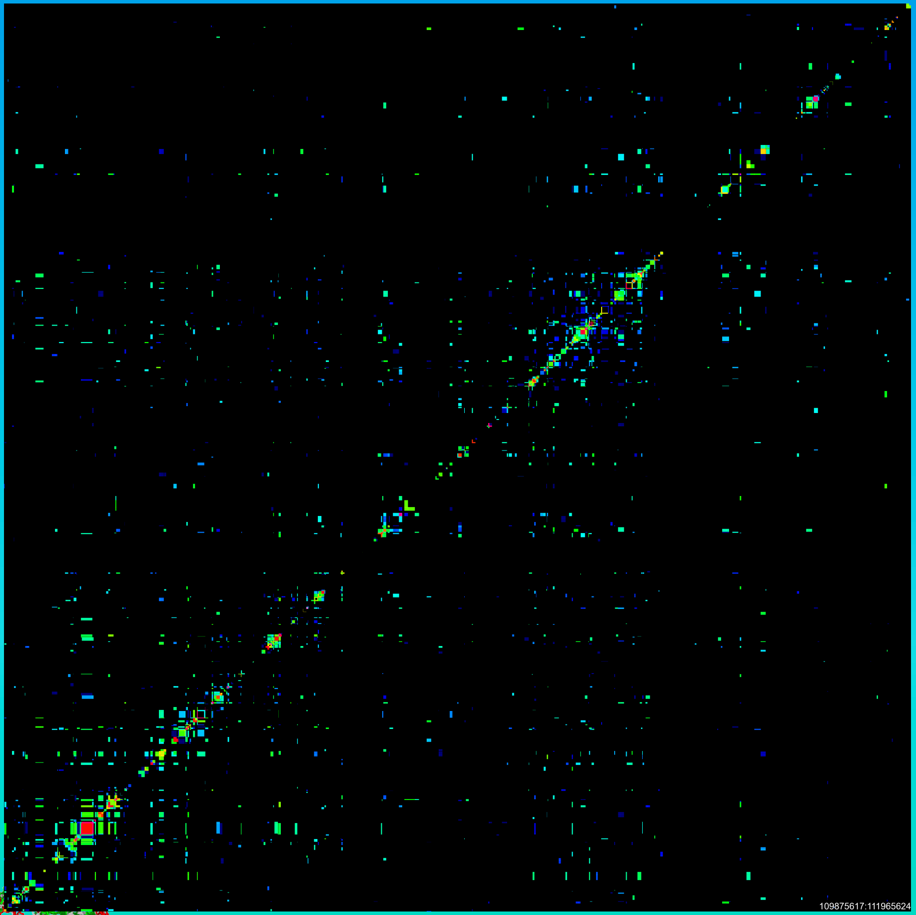
The DNA sequence is analyzed by the most simplest scaling analysis to find unprejudiced patters as e.g. nucleosome positions as well as chromatin loops and rosettes. The analysis is done using our grid infrastructures and here especially our volunteer grid.

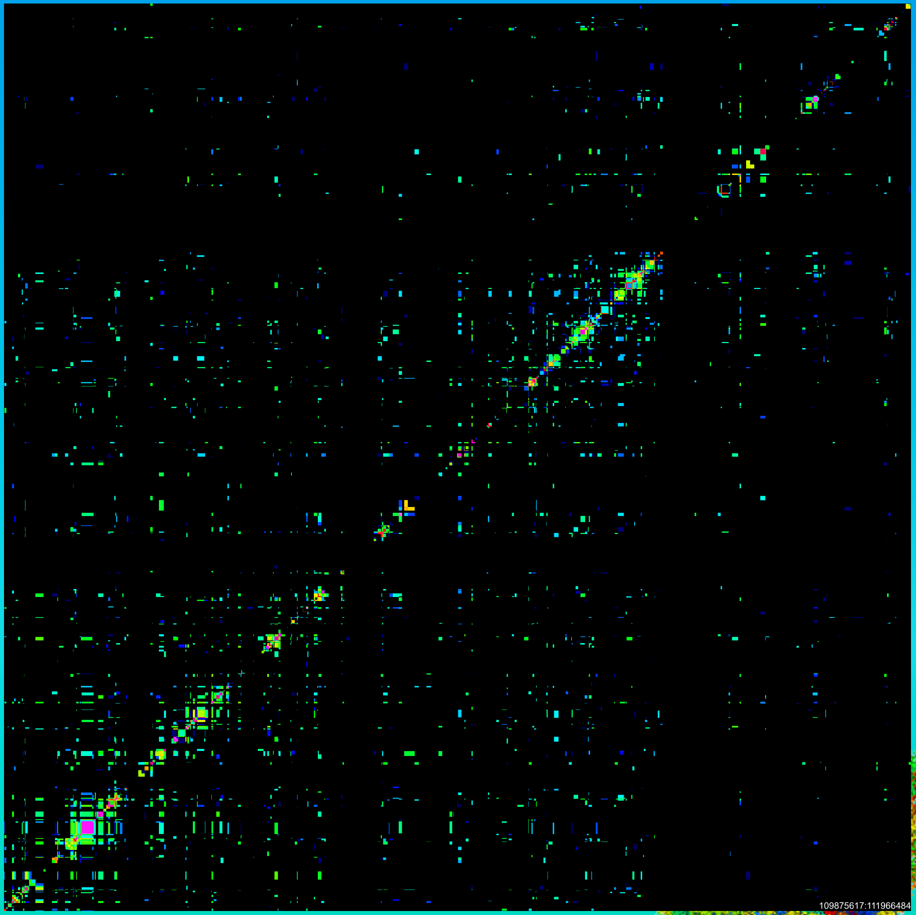


Nucleosome ”Ground States” and Loops can be Derived from the Sequence and Show the Holistics!

Intra/Inter Chromosomal Spatial Architecture WP2

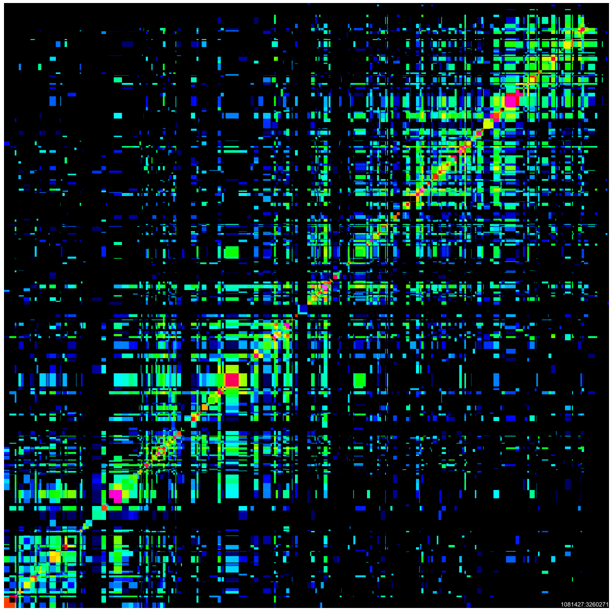
Intra/inter chromosomal contacts are determined using a combination of chromosome conformation capture technology and highest-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure is derived, i.e. its folding into a chromatin fiber and chromatin rosettes.

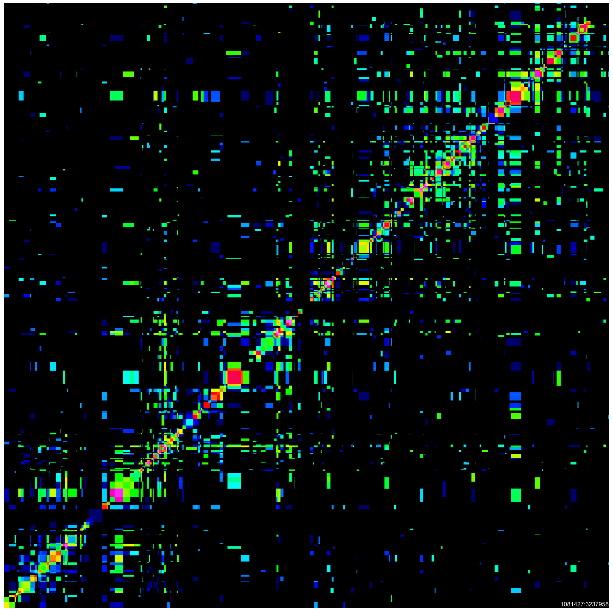


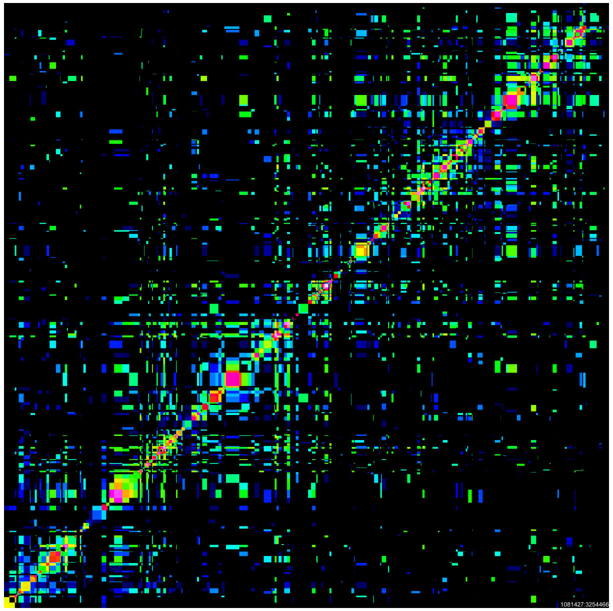


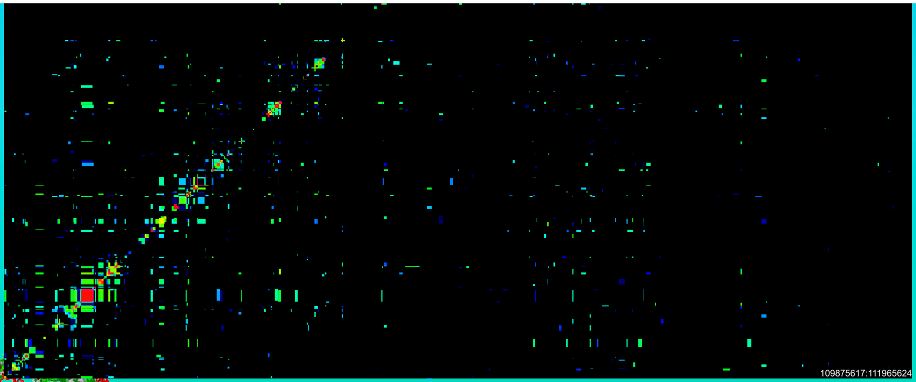
Intra/Inter Chromosomal Spatial Architecture WP2

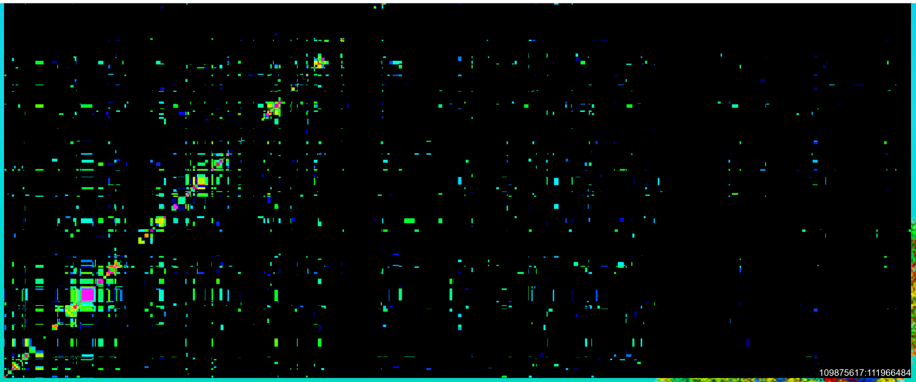
Intra/inter chromosomal contacts are determined using a combination of chromosome conformation capture technology and highest-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure is derived, i.e. its folding into a chromatin fiber and chromatin rosettes.

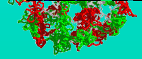


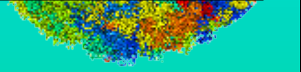






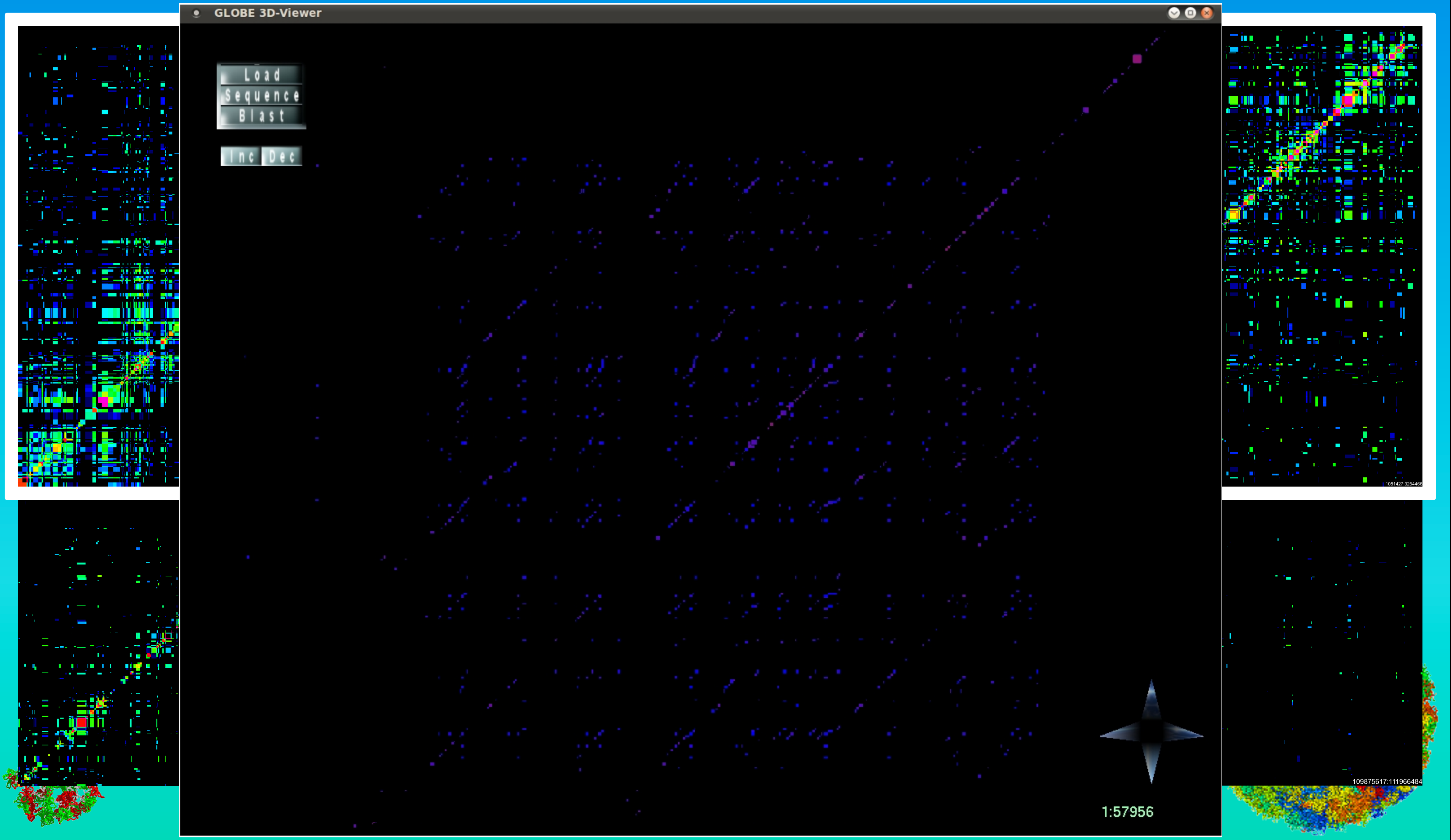






Intra/Inter Chromosomal Spatial Architecture WP2

Intra/inter chromosomal contacts are determined using a combination of chromosome conformation capture technology and highest-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure is derived, i.e. its folding into a chromatin fiber and chromatin rosettes.



Intra/Inter Chromosomal Spatial Architecture WP2

Intra/inter chromosomal contacts are determined using a combination of chromosome conformation ca
technology and highest-throughput deep sequencing. From the interaction maps 3D chromatin conf
and its higher-order structure is derived, i.e. its folding into a chromatin fiber and chromati

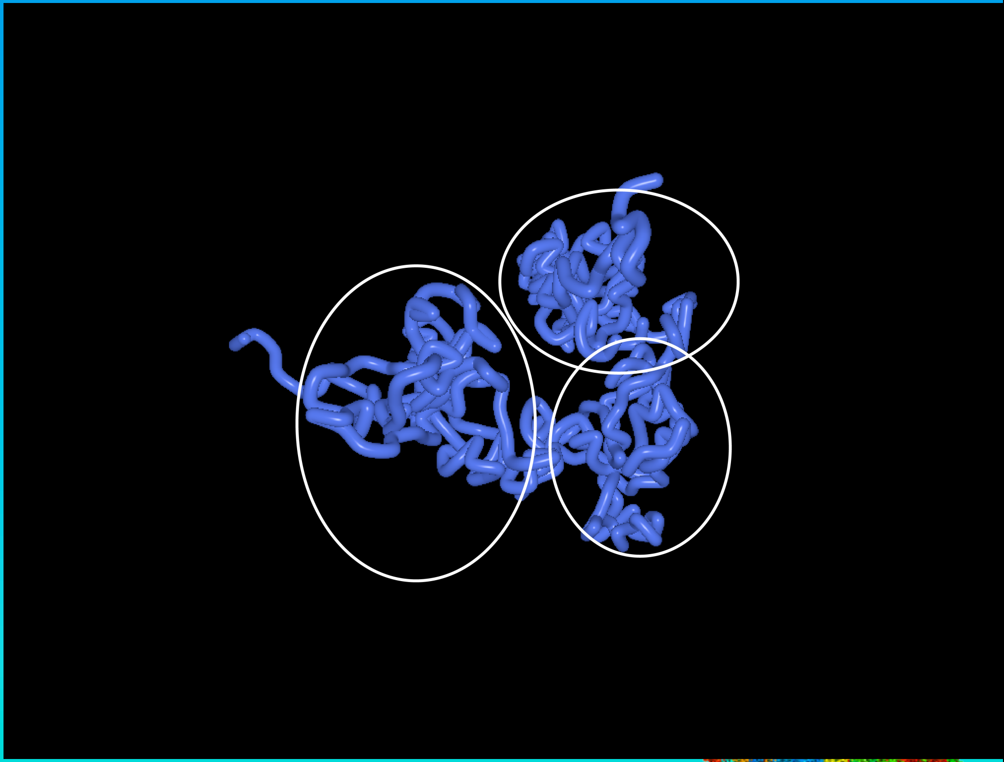
We Reached the Level of the Genomic Uncertainty Principle!

The Conformation of Chromatin and the 3D Architecture
can be derived with Tremendous Resolution !

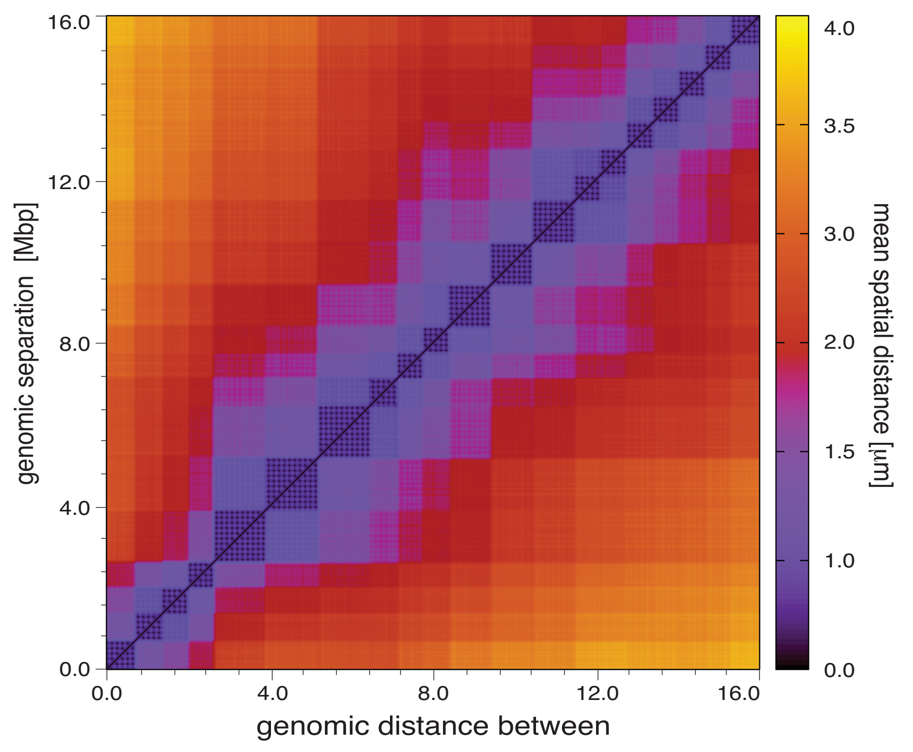
Intra/Inter Chromosomal Interaction Architecture WP2

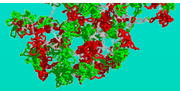
Intra/inter chromosomal contacts are determined using a combination of chromosome conformation capture technology and highest-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its higher-order structure is derived, i.e. its folding into a chromatin fiber and chromatin rosettes.

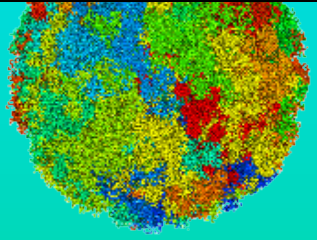
3D Modelling from Interaction Maps



Chromatin Rosettes
Loops 126 kbp





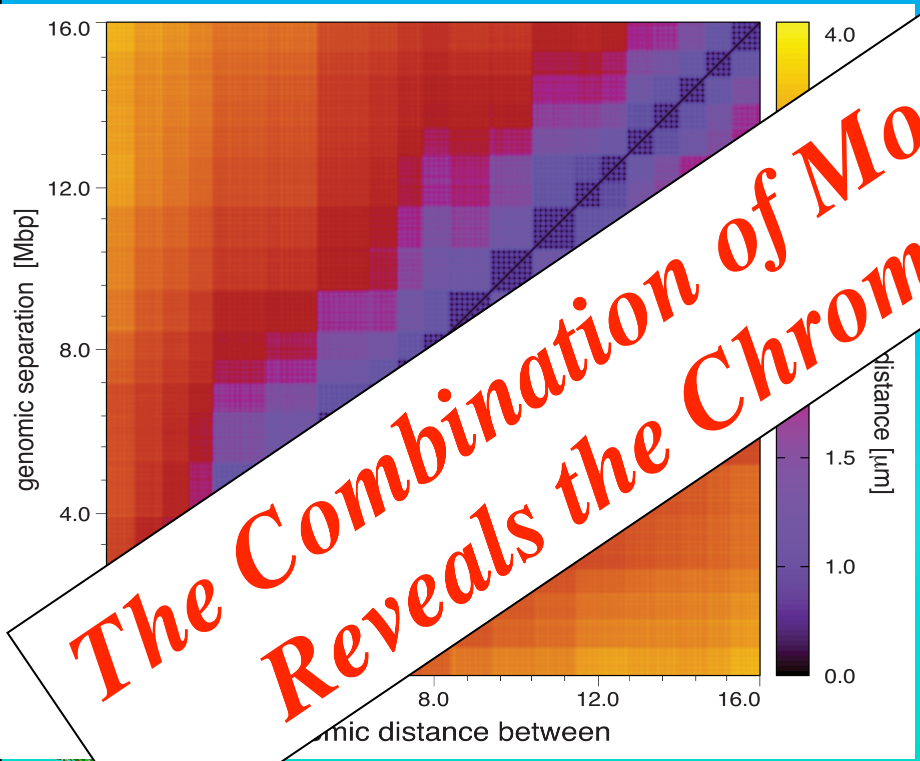


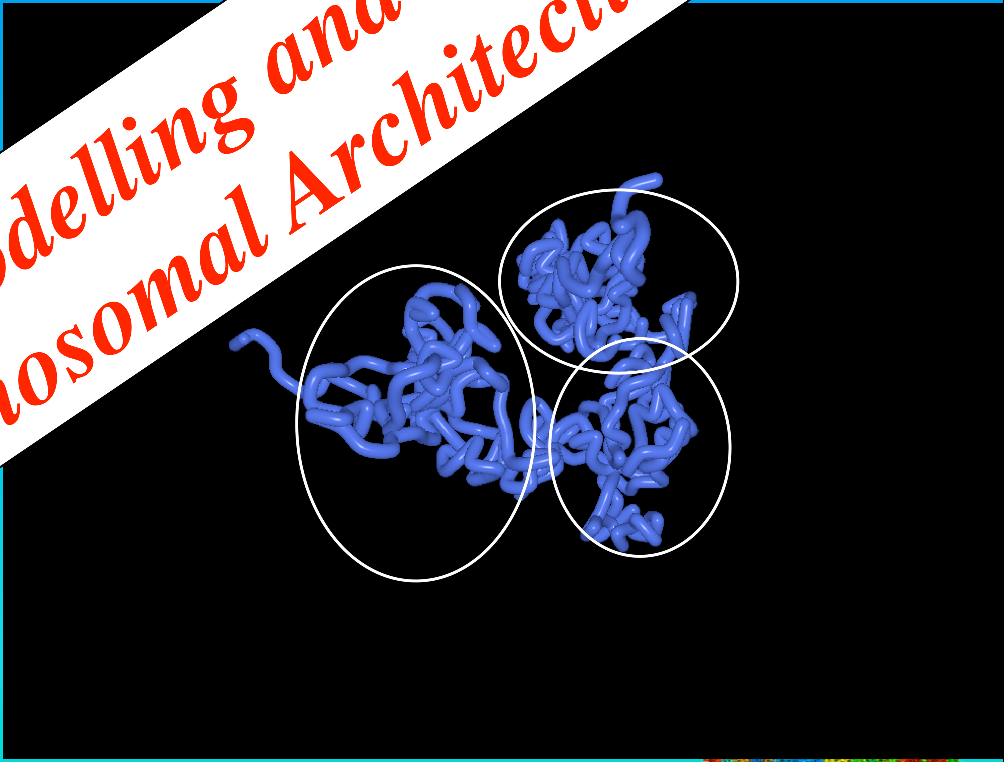
Intra/Inter Chromosomal Interaction Architecture WP2

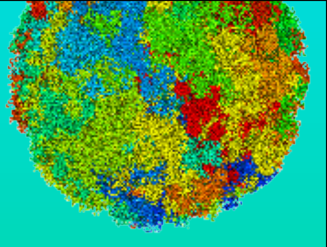
Intra/inter chromosomal contacts are determined using a combination of chromosome conformation capture techn
and highest-throughput deep sequencing. From the interaction maps 3D chromatin conformations and its high
structure is derived, i.e. its folding into a chromatin fiber and chromatin rosettes.

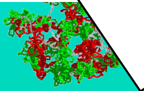
3D M ps

Chromatin Rosettes
Loops 126 kbp





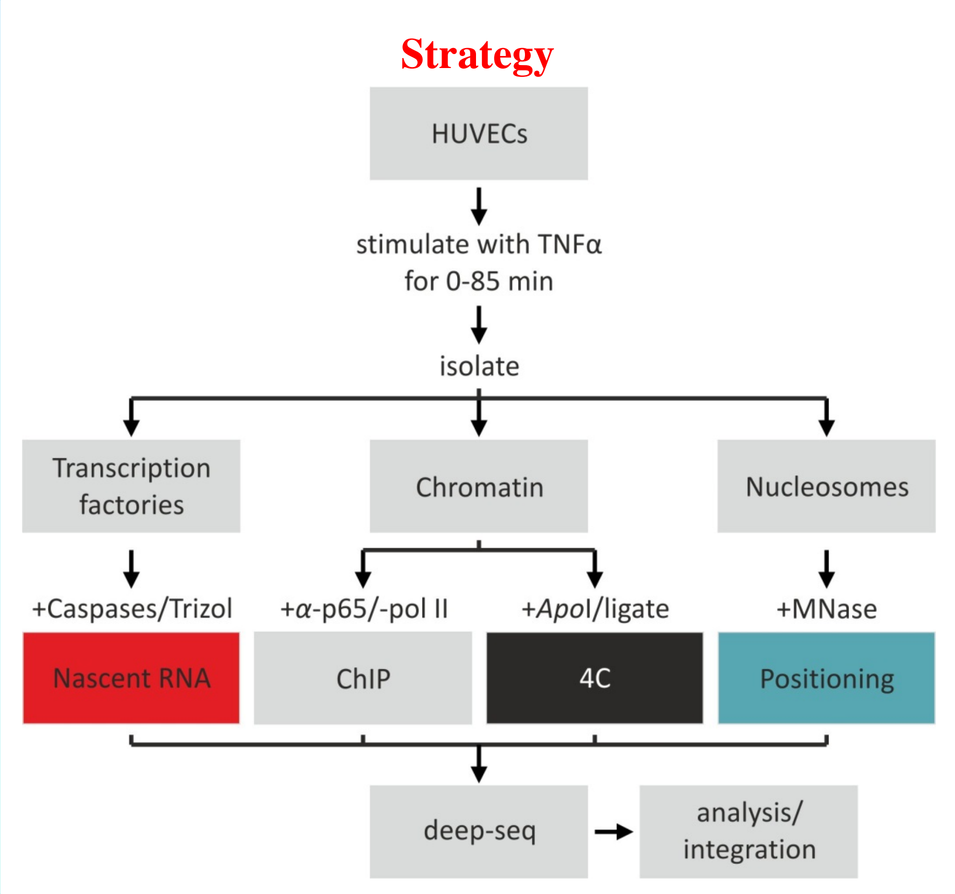




The Combination of Modelling and Experiment Reveals the Chromosomal Architecture !

Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.



Strategy

HUVECs

stimulate with TNFα
for 0-85 min

isolate

Transcription factories

Chromatin

Nucleosomes

+Caspases/Trizol

+α-p65/-pol II

+*Apo*I/ligate

+MNase

Nascent RNA

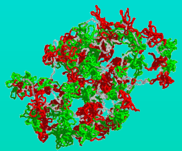
ChIP

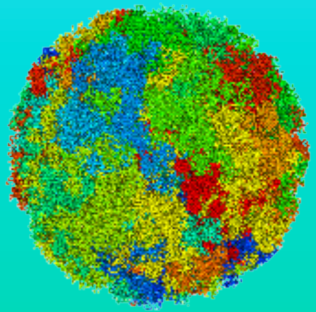
4C

Positioning

deep-seq

analysis/ integration





Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.

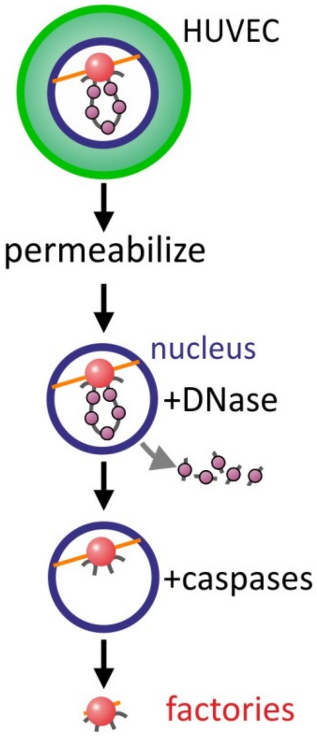
Strategy

somes

lase

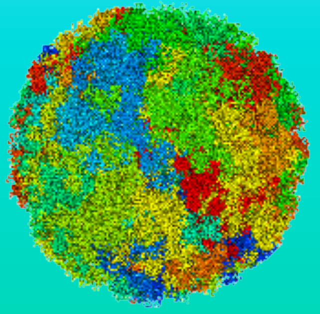
oning

Factory (Nascent RNA)-seq



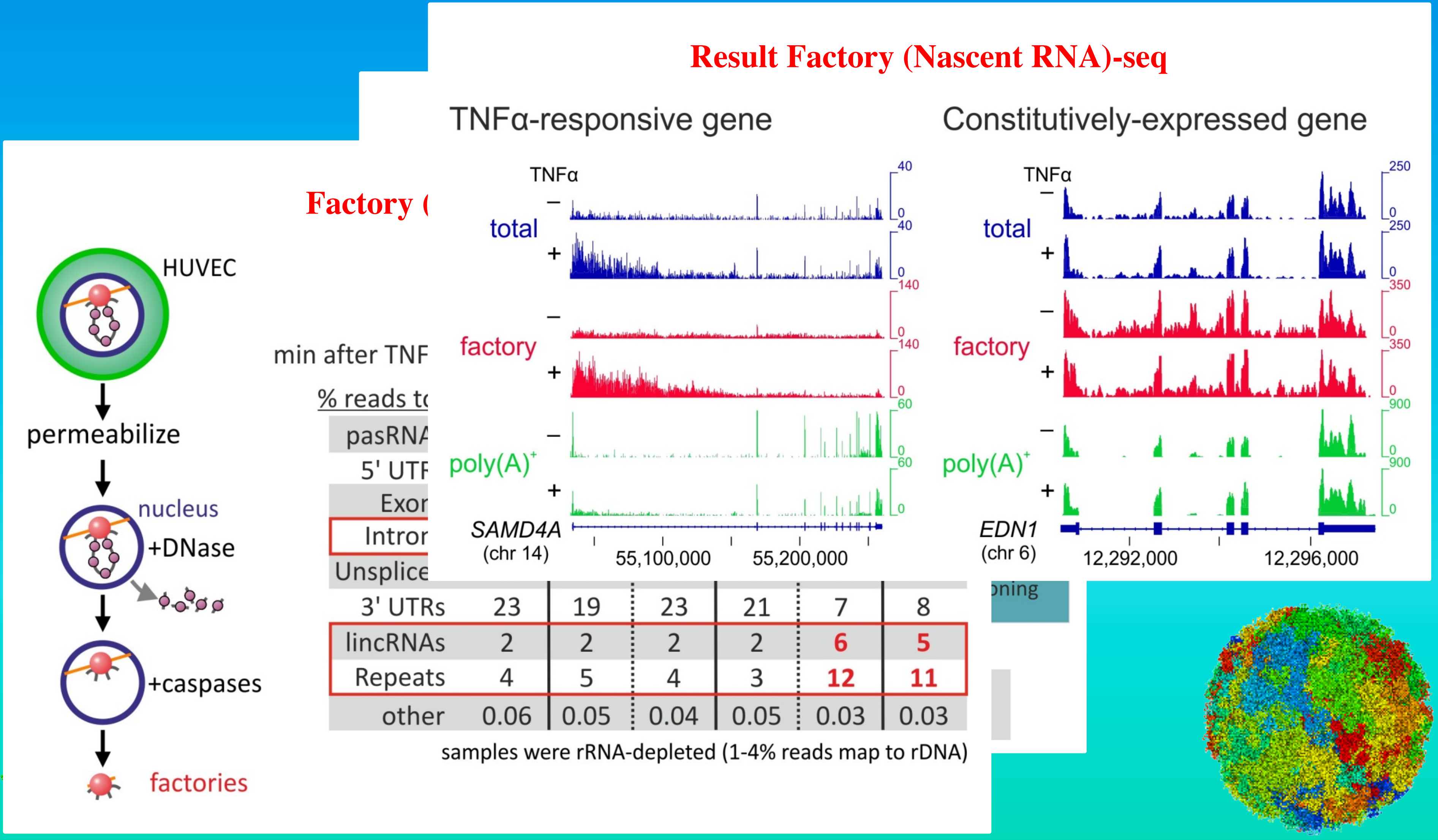
	total		polyA+		'factory'	
min after TNFα	0	30	0	30	0	30
% reads to:						
pasRNAs	37	32	35	34	16	17
5' UTRs	21	18	23	25	7	7
Exons	55	43	**61**	**65**	11	13
Introns	9	12	6	4	**69**	**67**
Unspliced	33	33	31	39	26	26
3' UTRs	23	19	23	21	7	8
lincRNAs	2	2	2	2	**6**	**5**
Repeats	4	5	4	3	**12**	**11**
other	0.06	0.05	0.04	0.05	0.03	0.03

samples were rRNA-depleted (1-4% reads map to rDNA)



Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.



Result Factory (Nascent RNA)-seq

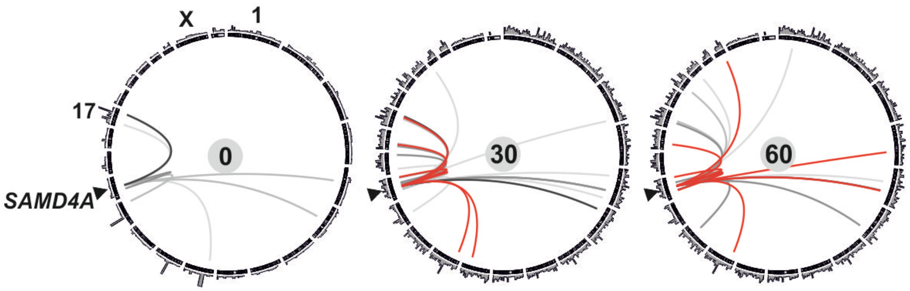
pasRN						
5' UT						
Exo						
Intro						
Unsplice						
3' UTRs	23	19	23	21	7	8
lincRNAs	2	2	2	2	6	5
Repeats	4	5	4	3	12	11
other	0.06	0.05	0.04	0.05	0.03	0.03

samples were rRNA-depleted (1-4% reads map to rDNA)

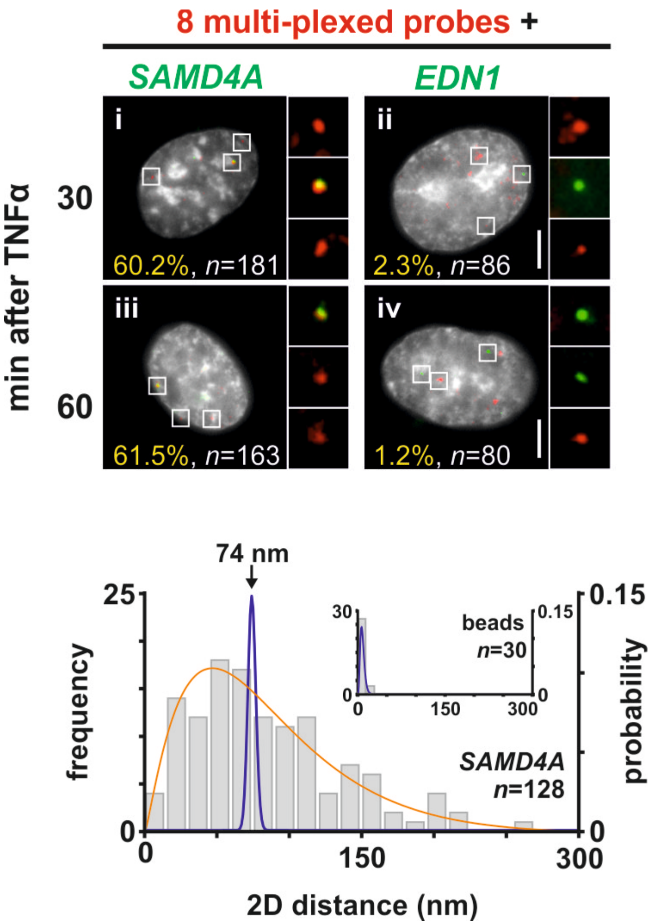
Transcription Structure-Function Relationship WP3

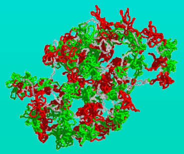
The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.

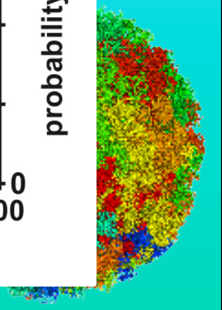
4C Interaction in the SAMD4A Locus



Super-Resolution Microscopy



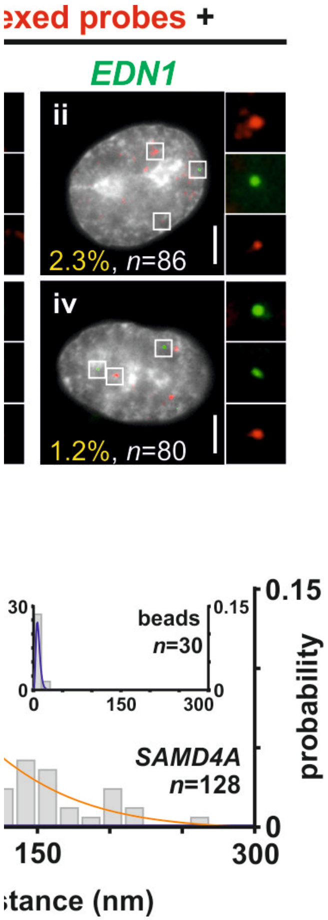




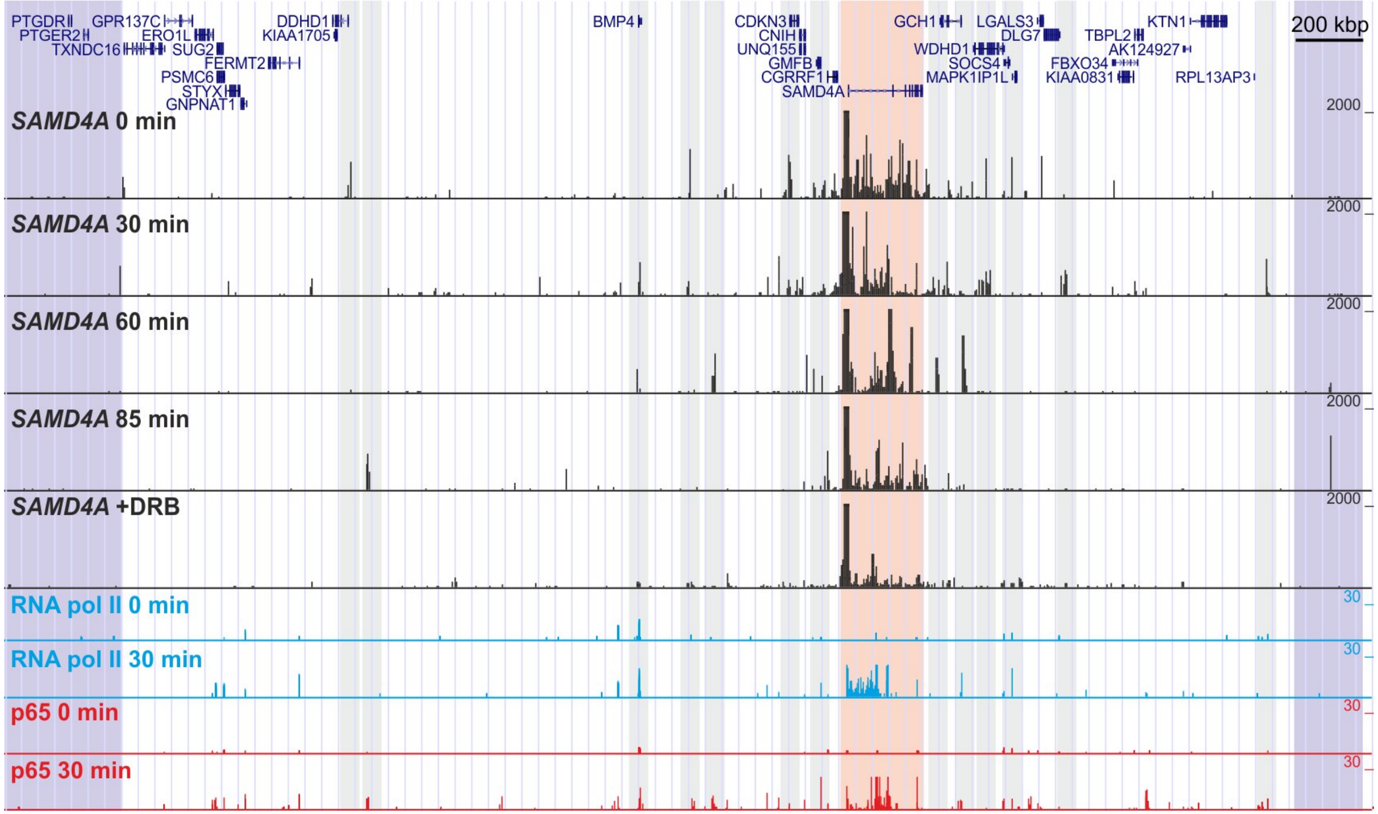
Transcription Structure-Function Relationship WP3

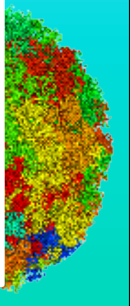
The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.

Super-Resolution Microscopy



Kinetics of 4C

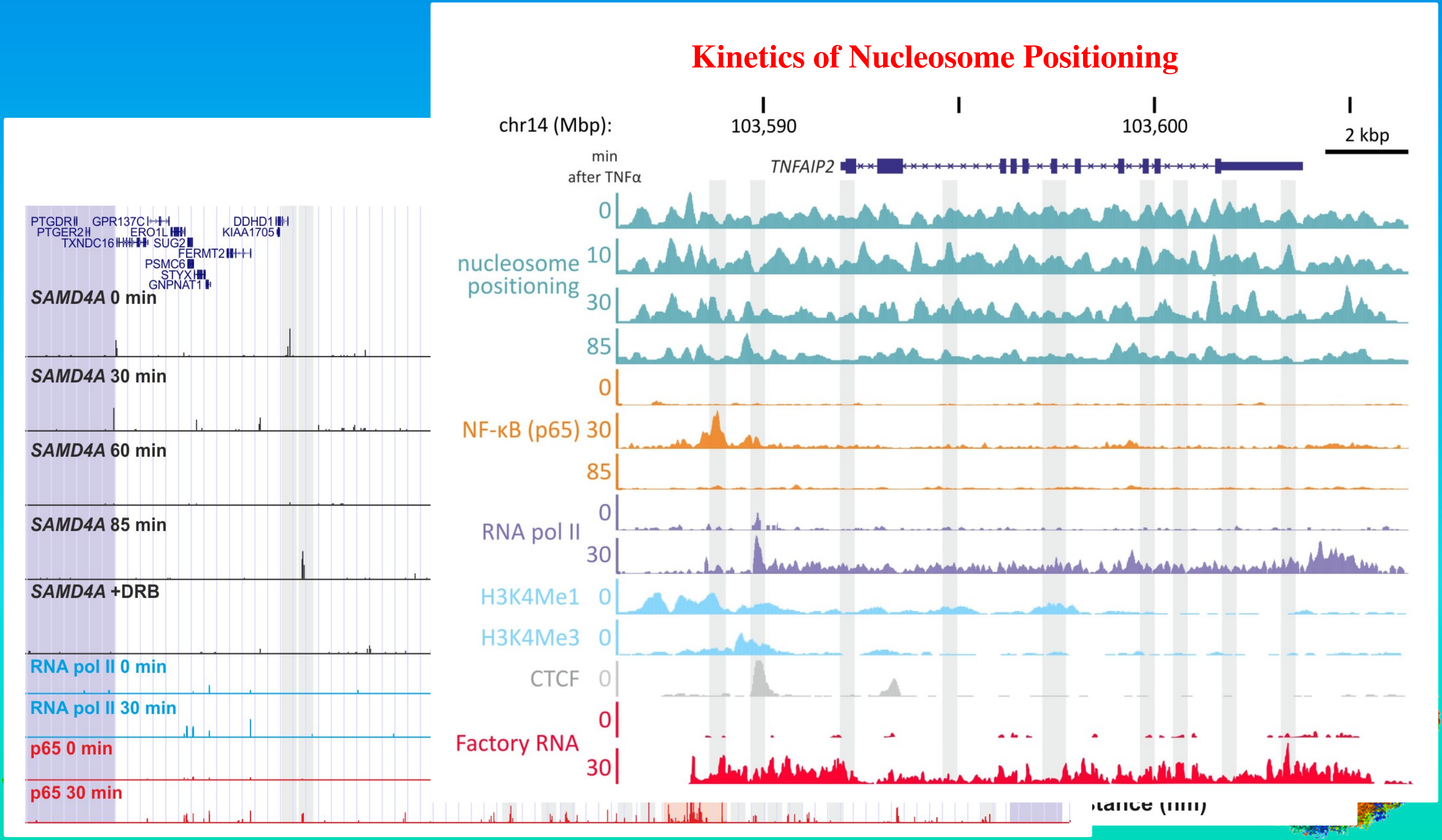




Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.

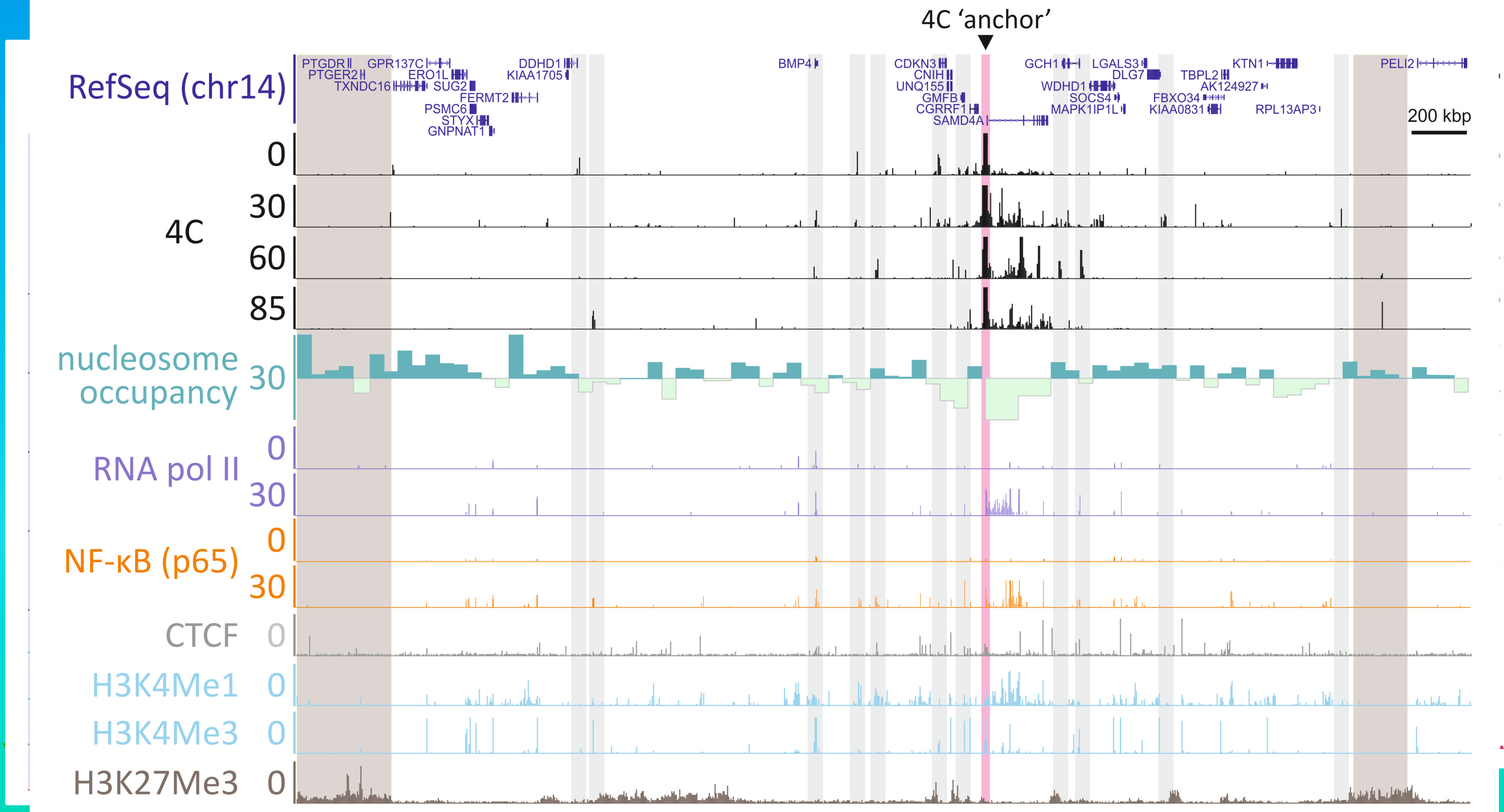
Kinetics of Nucleosome Positioning



Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolution microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail this structure function relationship in several cycles of refinement.

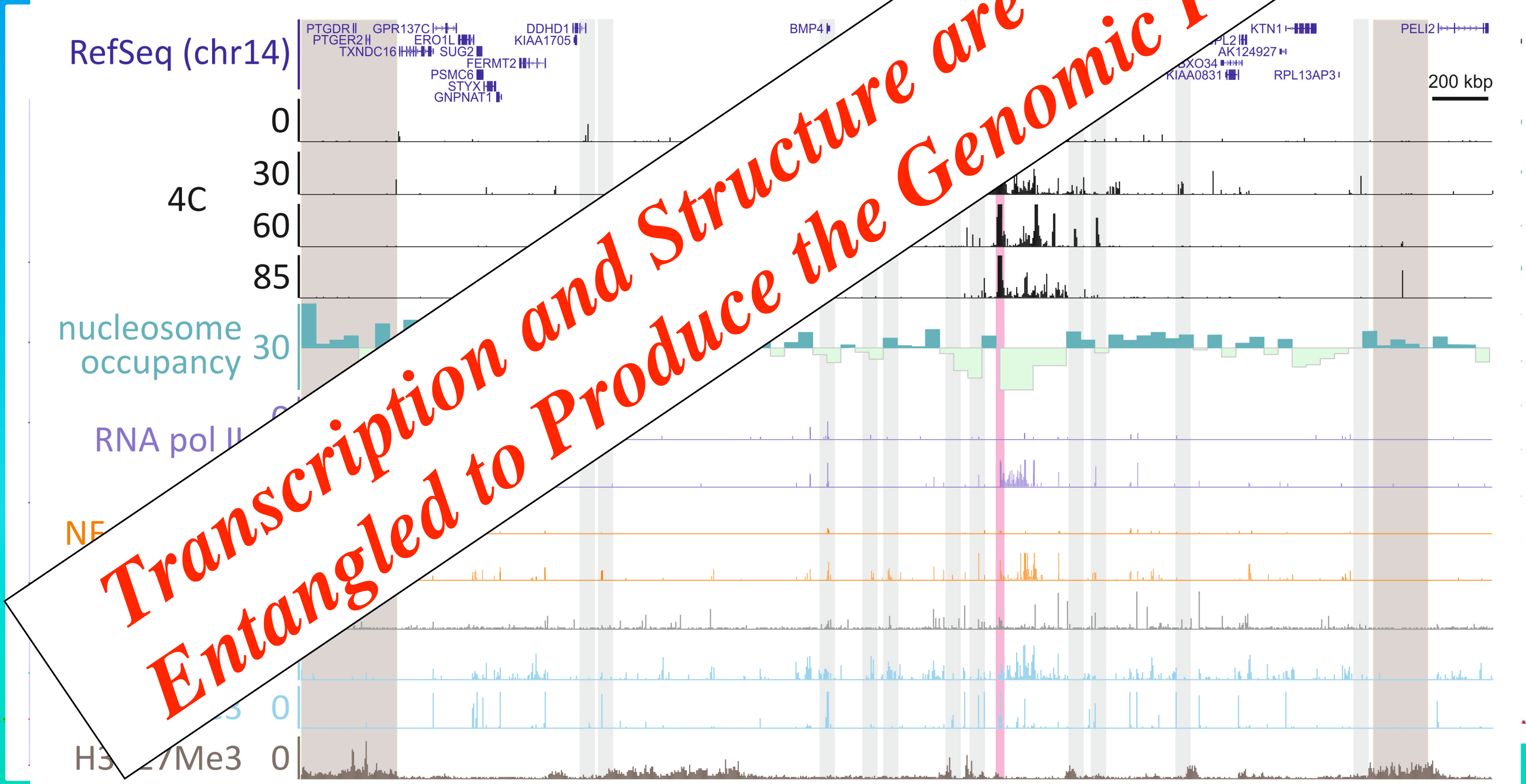
Integration of the Kinetics



Transcription Structure-Function Relationship WP3

The transcription within factories depends on nucleosome position, the chromatin fiber conformation, as well as the 3D architecture of the chromatin fiber. Only the systemic integration of nascent-RNA-seq, super-resolutio microscopy, and 4C (or our novel approach) in combination with computer simulations reveals in detail structure function relationship in several cycles of refinement.

Integration of the Kinetics

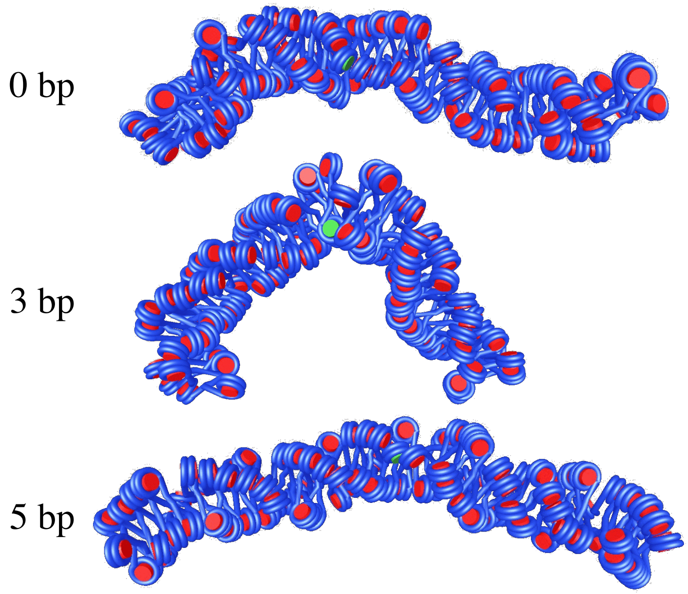


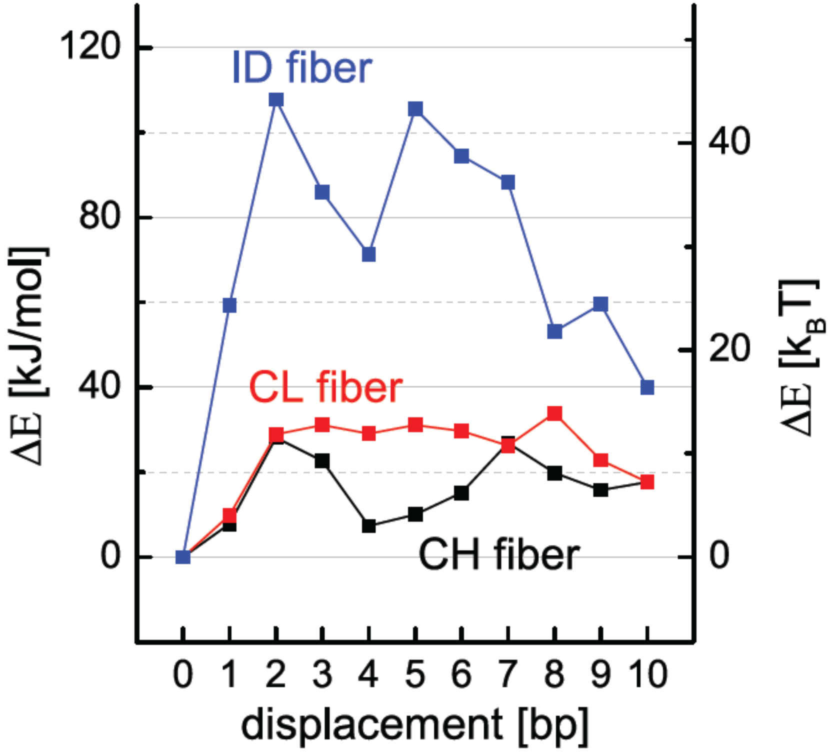
Transcription and Structure are Very Tightly Entangled to Produce the Genomic Function!

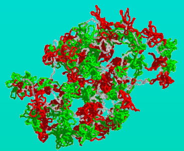
Simulation of Nucleosomes & Chromatin WP4

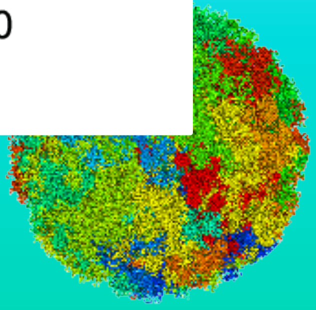
The position of nucleosomes influence greatly the structure of chromatin fibers done on super-computers. Here a dedicated workflow is applied, with which overlapping nucleosome populations can be analyzed and the best positioning of nucleosomes by Monte Carlo simulated annealing can be achieved. For an actual locus in a spherical confinement then a 3D independent nucleosome fiber conformation can be simulated.

Chicken Erythrocytes Snapshots



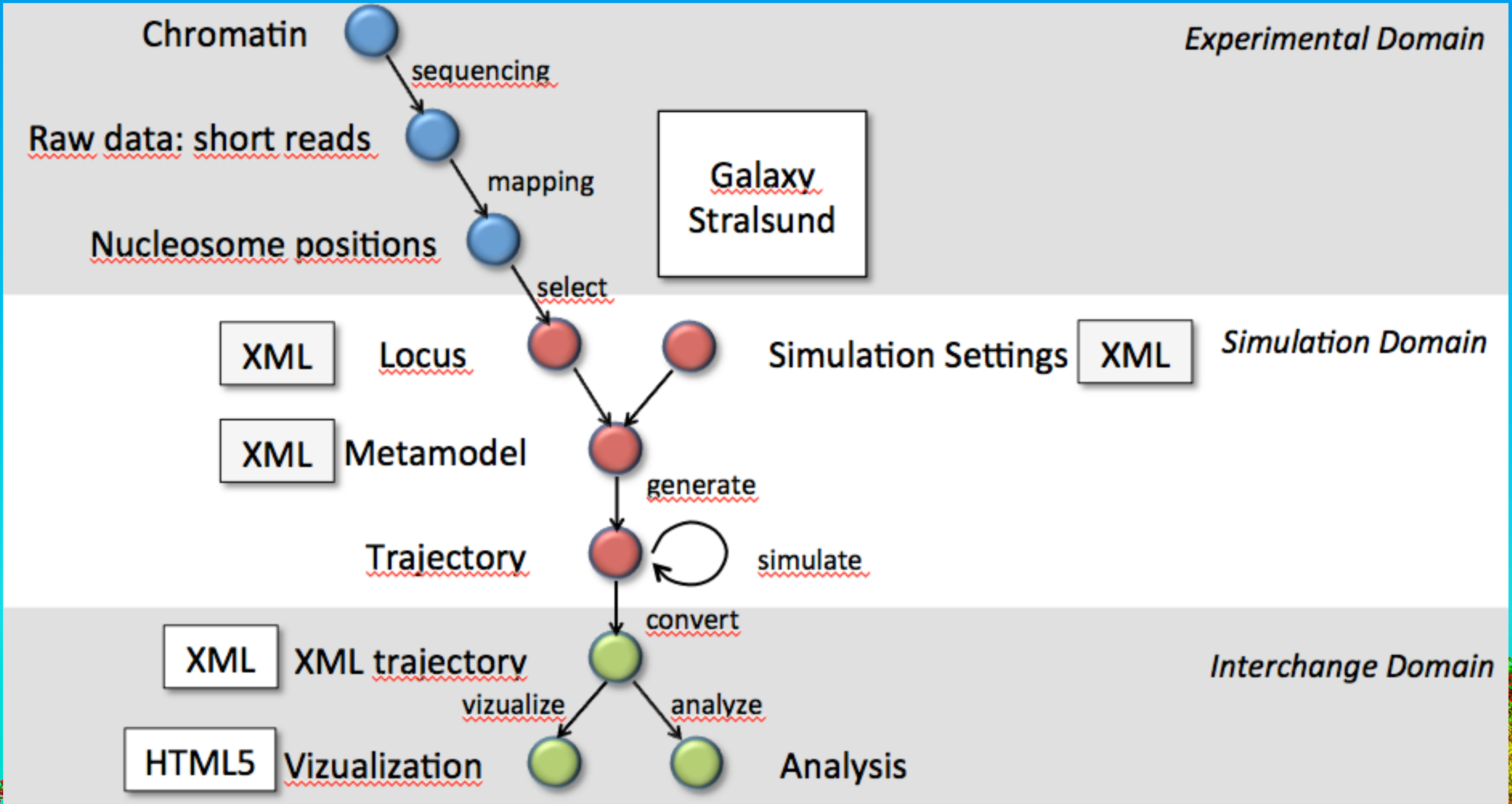






Simulation of Nucleosomes & Chromatin WP4

The position of nucleosomes influence greatly the structure of chromatin fibers done on super-computers. Here a dedicated workflow is applied, with which overlapping nucleosome populations can be analyzed and the best positioning of nucleosomes by Monte Carlo simulated annealing can be achieved. For an actual locus in a spherical confinement then a 3D independent nucleosome fiber conformation can be simulated.



Experimental Domain

Chromatin

sequencing

Raw data: short reads

mapping

Nucleosome positions

Galaxy Stralsund

select

Simulation Domain

XML Locus

Simulation Settings XML

XML Metamodel

generate

Trajectory

simulate

Interchange Domain

convert

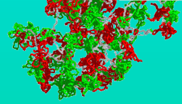
XML XML trajectory

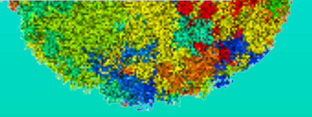
vizualize

analyze

HTML5 Vizualization

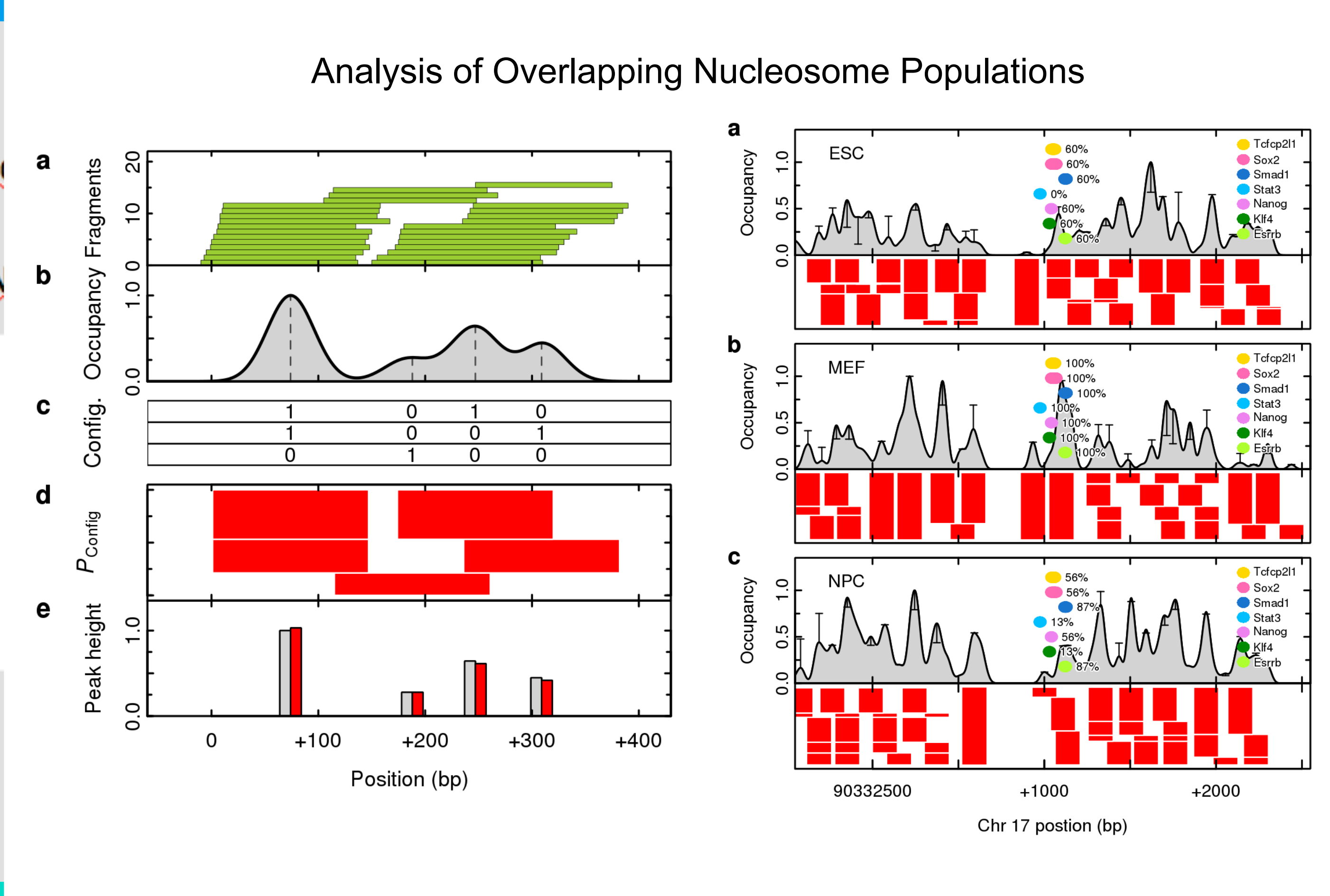
Analysis



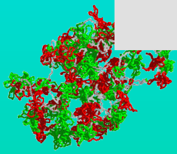


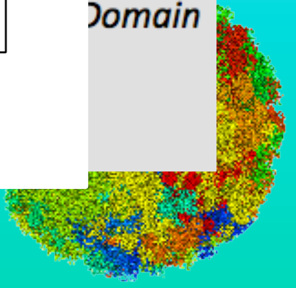
Simulation of Nucleosomes & Chromatin WP4

The position of nucleosomes influence greatly the structure of chromatin fibers done on super-computers. Here a dedicated workflow is applied, with which overlapping nucleosome populations can be analyzed and the best positioning of nucleosomes by Monte Carlo simulated annealing can be achieved. For an actual locus in a spherical confinement then a 3D independent nucleosome fiber conformation can be simulated.



Analysis of Overlapping Nucleosome Populations

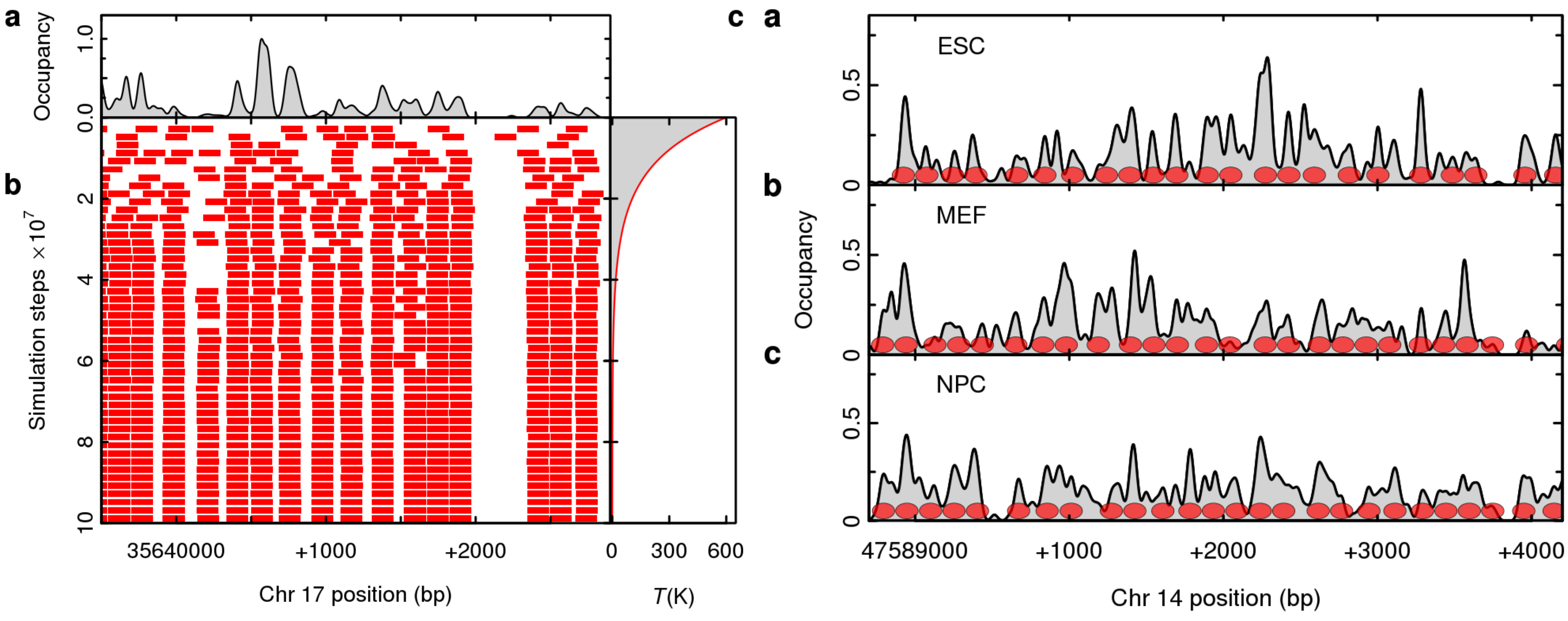


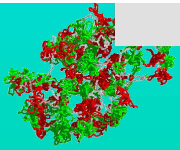


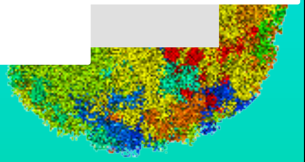
Simulation of Nucleosomes & Chromatin WP4

The position of nucleosomes influence greatly the structure of chromatin fibers done on super-computers. Here a dedicated workflow is applied, with which overlapping nucleosome populations can be analyzed and the best positioning of nucleosomes by Monte Carlo simulated annealing can be achieved. For an actual locus in a spherical confinement then a 3D independent nucleosome fiber conformation can be simulated.

Finding an optimal configuration of nucleosomes by simulated annealing.

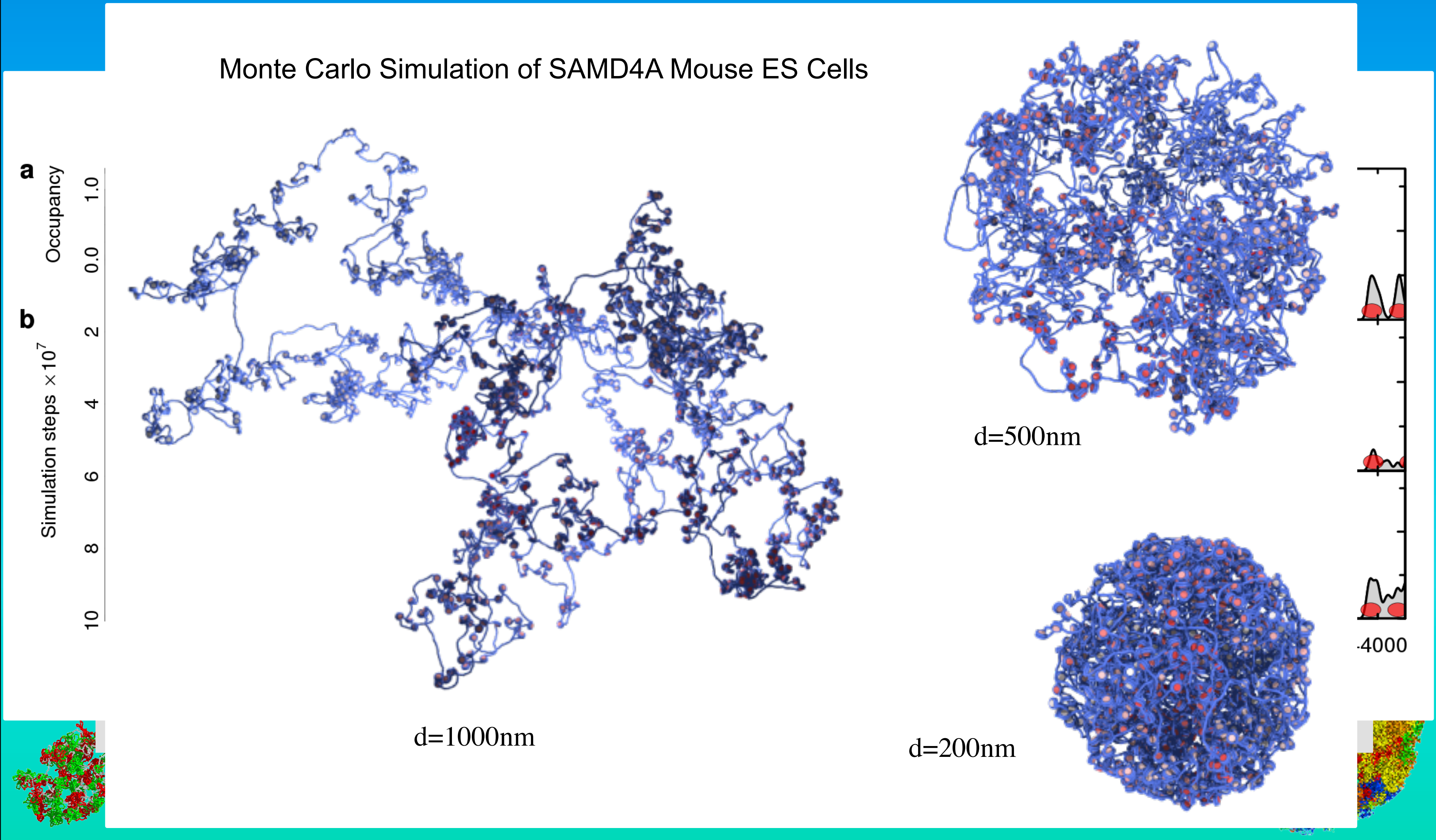






Simulation of Nucleosomes & Chromatin WP4

The position of nucleosomes influence greatly the structure of chromatin fibers done on super-computers. Here a dedicated workflow is applied, with which overlapping nucleosome populations can be analyzed and the best positioning of nucleosomes by Monte Carlo simulated annealing can be achieved. For an actual locus in a spherical confinement then a 3D independent nucleosome fiber conformation can be simulated.

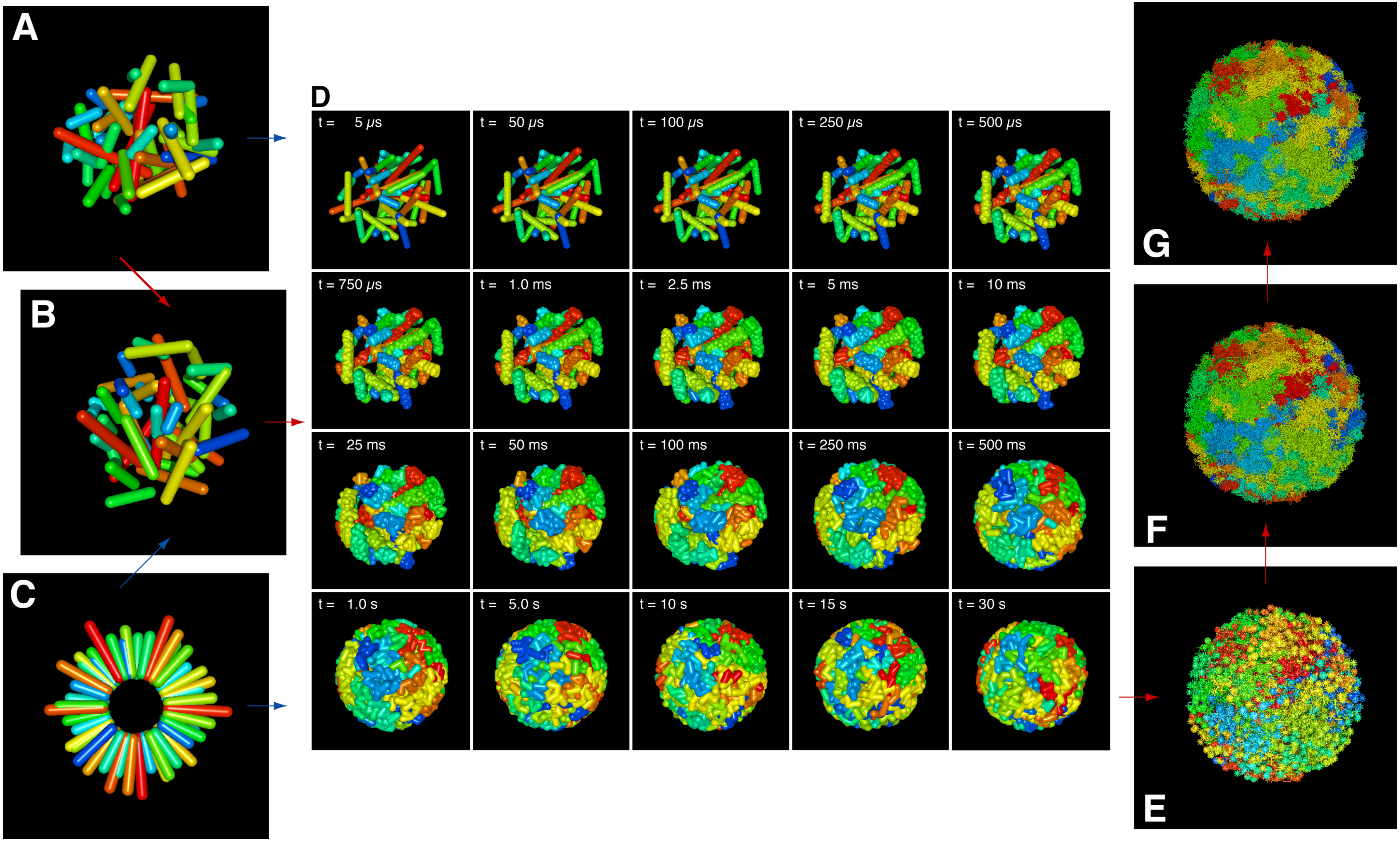


Monte Carlo Simulation of SAMD4A Mouse ES Cells

Simulation of Chromosomes & Nuclei WP4

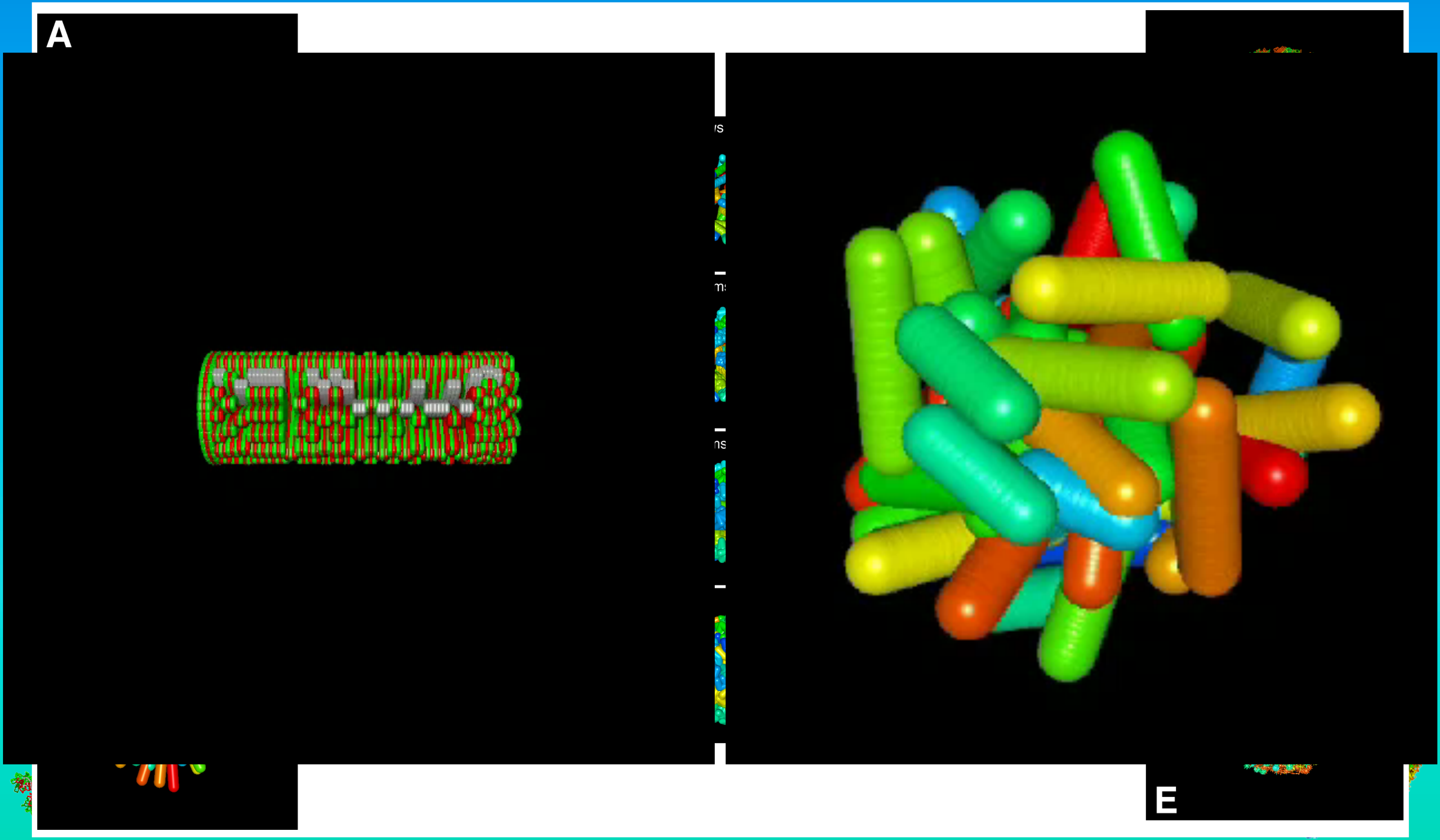
By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches simulate chromosomes and whole nuclei with unprecedented resolution, resulting in novel predictions for the detailed folding of the chromatin fiber with corresponding impact on the experimental evaluation.

Bundesministerium für Bildung und Forschung
BBSRC
NWO
EpiGenSys Consortium
ERASysBio
European Commission
Erasmus MC
BioQuant
University Heidelberg
University of Oxford
dkfz. German Cancer Research Center
MediGRID
fachhochschule stralsund



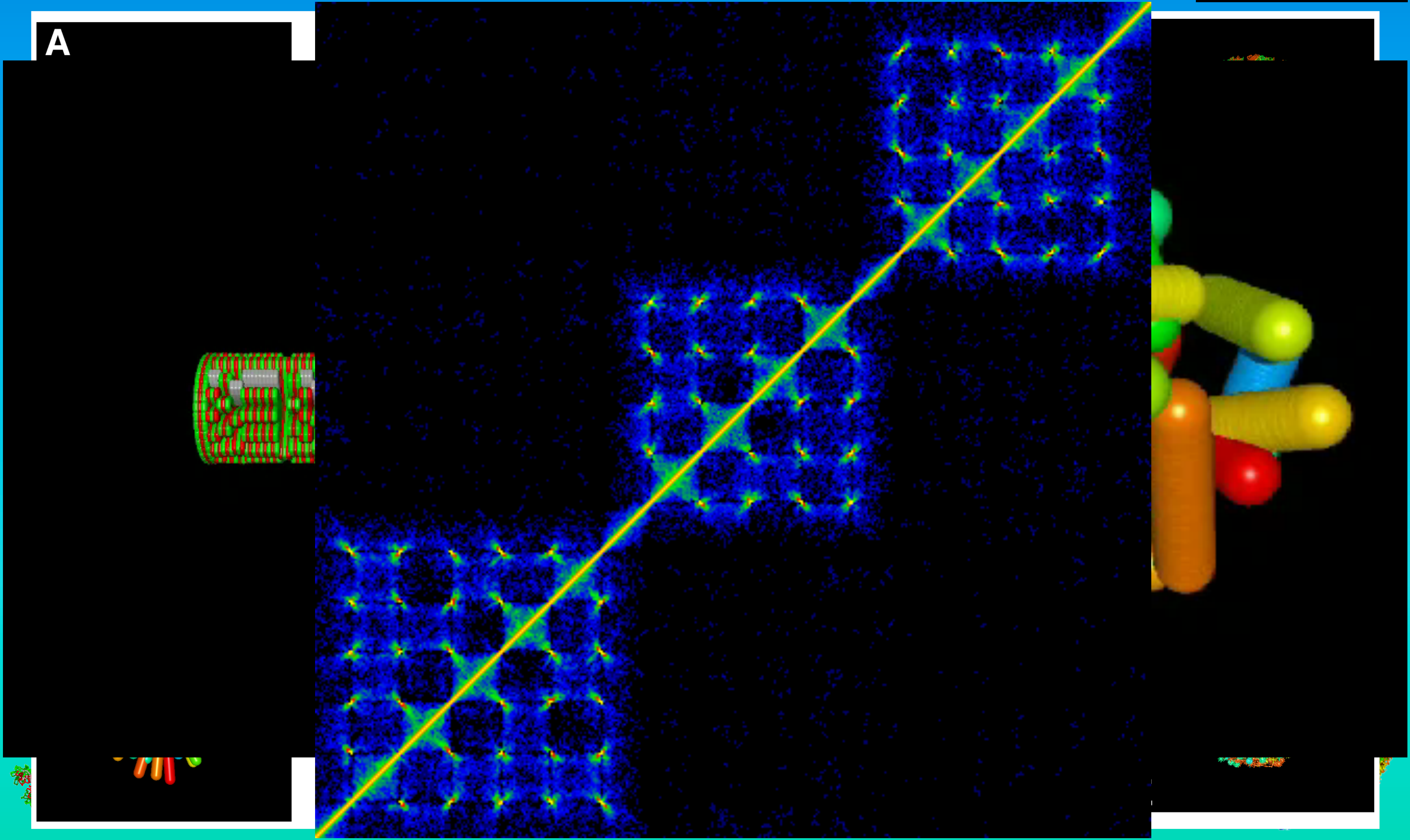
Simulation of Chromosomes & Nuclei WP4

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches simulate chromosomes and whole nuclei with unprecedented resolution, resulting in novel predictions for the detailed folding of the chromatin fiber with corresponding impact on the experimental evaluation.



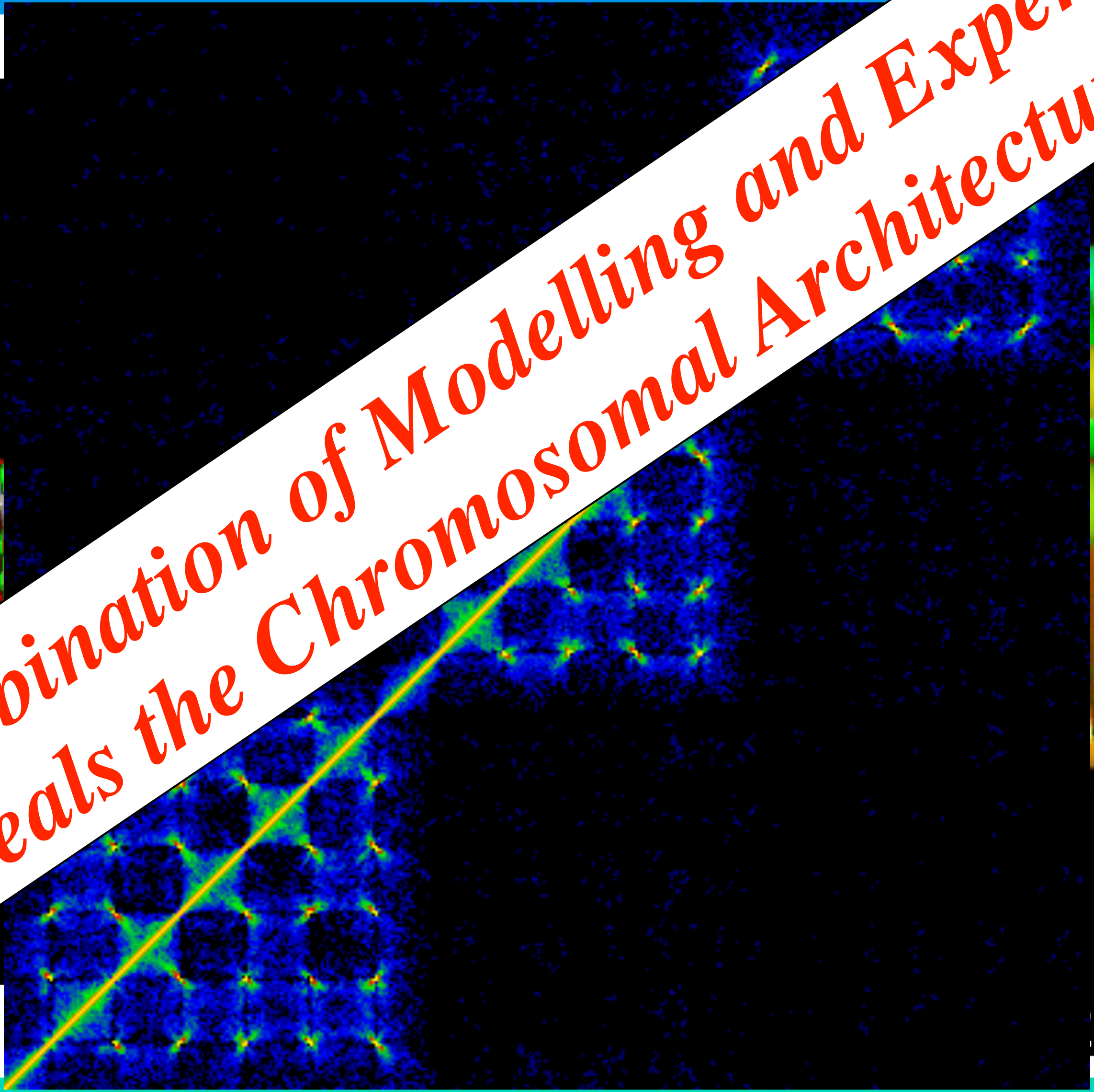
Simulation of Chromosomes & Nuclei WP4

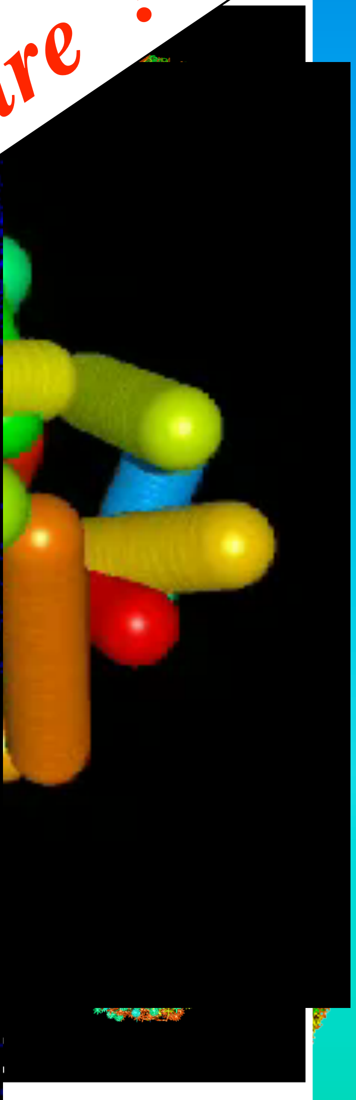
By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches simulate chromosomes and whole nuclei with unprecedented resolution, resulting in novel predictions for the detailed folding of the chromatin fiber with corresponding impact on the experimental evaluation.



Simulation of Chromosomes & Nuclei WP4

By parallel super-computer simulations using novel Monte Carlo and Brownian Dynamics approaches simulate chromosomes and whole nuclei with unprecedented resolution, resulting in novel predictions for the detailed fol
of the chromatin fiber with corresponding impact on the experimental evaluation.



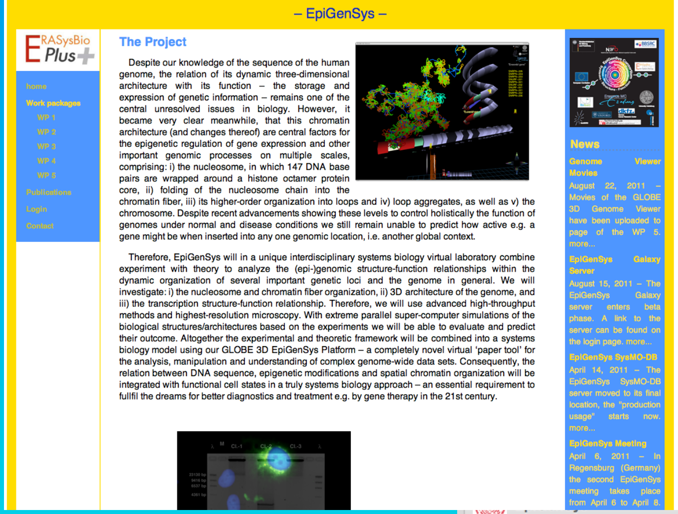


The Combination of Modelling and Experiment Reveals the Chromosomal Architecture !

Systems Biological Result Communcation via Internal and External Communication Windows WP5

Three classic internal and external communication windows were opened: i) the EpiGenSys website communicates with the public, ii) the EpiGenSys Wiki functions is used for exchange, and iii) the EpiGenSys SysMO DB archives in a virtual labnote the data. All allow access to the systems and are the basis for the GLOBE 3D Genome Platform.

EpiGenSys Public Website



System Biological Determination of the Epigenomic Structure – Function Relation
– EpiGenSys –

The Project

Despite our knowledge of the sequence of the human genome, the relation of its dynamic three-dimensional architecture with its function – the storage and expression of genetic information – remains one of the central unresolved issues in biology. However, it became very clear meanwhile, that this chromatin architecture (and changes thereof) are central factors for the epigenetic regulation of gene expression and other important genomic processes on multiple scales, comprising: i) the nucleosome, in which 147 DNA base pairs are wrapped around a histone octamer protein core, ii) folding of the nucleosome chain into the chromatin fiber, iii) its higher-order organization into loops and iv) loop aggregates, as well as v) the chromosome. Despite recent advancements showing these levels to control holistically the function of genomes under normal and disease conditions we still remain unable to predict how active e.g. a gene might be when inserted into any one genomic location, i.e. another global context.

Therefore, EpiGenSys will in a unique interdisciplinary systems biology virtual laboratory combine experiment with theory to analyze the (epi-)genomic structure-function relationships within the dynamic organization of several important genetic loci and the genome in general. We will investigate: i) the nucleosome and chromatin fiber organization, ii) 3D architecture of the genome, and iii) the transcription structure-function relationship. Therefore, we will use advanced high-throughput methods and highest-resolution microscopy. With extreme parallel super-computer simulations of the biological structures/architectures based on the experiments we will be able to evaluate and predict their outcome. Altogether the experimental and theoretic framework will be combined into a systems biology model using our GLOBE 3D EpiGenSys Platform – a completely novel virtual 'paper tool' for the analysis, manipulation and understanding of complex genome-wide data sets. Consequently, the relation between DNA sequence, epigenetic modifications and spatial chromatin organization will be integrated with functional cell states in a truly systems biology approach – an essential requirement to fullfil the dreams for better diagnostics and treatment e.g. by gene therapy in the 21st century.

News

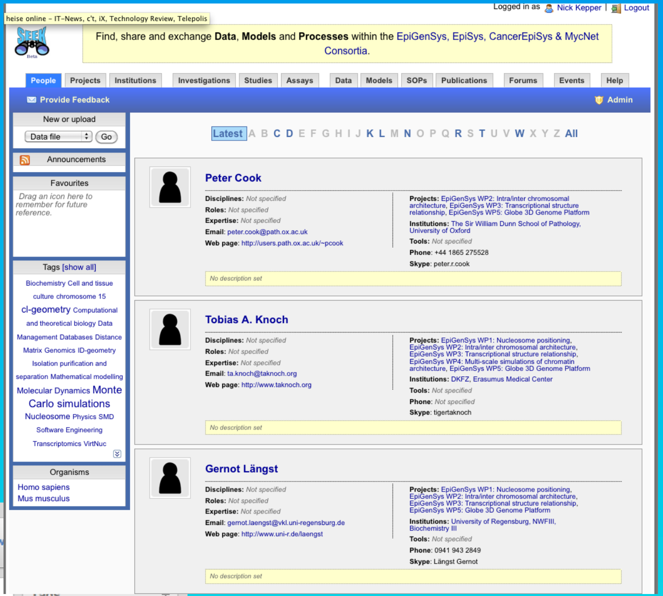
Genome Viewer Movies
August 22, 2011 – Movies of the GLOBE 3D Genome Viewer have been uploaded to page of the WP 5. more...

EpiGenSys Galaxy Server
August 15, 2011 – The EpiGenSys Galaxy server enters beta phase. A link to the server can be found on the login page. more...

EpiGenSys SysMO-DB
April 14, 2011 – The EpiGenSys SysMO-DB server moved to its final location, the "production usage" starts now. more...

EpiGenSys Meeting
April 6, 2011 – In Regensburg (Germany) the second EpiGenSys meeting takes place from April 6 to April 8.

EpiGenSys SysMO DB



Find, share and exchange **Data**, **Models** and **Processes** within the EpiGenSys, EpiSys, CancerEpiSys & MycNet Consortia.

Drag an icon here to remember for future reference.

Peter Cook

Disciplines: *Not specified*
Roles: *Not specified*
Expertise: *Not specified*
Email: peter.cook@path.ox.ac.uk
Web page: http://users.path.ox.ac.uk/~pcook
Projects: EpiGenSys WP2: Intra/inter chromosomal architecture, EpiGenSys WP3: Transcriptional structure relationship, EpiGenSys WP5: Globe 3D Genome Platform
Institutions: The Sir William Dunn School of Pathology, University of Oxford
Tools: *Not specified*
Phone: +44 1865 275528
Skype: peter.r.cook

No description set

Tobias A. Knoch

Disciplines: *Not specified*
Roles: *Not specified*
Expertise: *Not specified*
Email: ta.knoch@taknoch.org
Web page: http://www.taknoch.org
Projects: EpiGenSys WP1: Nucleosome positioning, EpiGenSys WP2: Intra/inter chromosomal architecture, EpiGenSys WP3: Transcriptional structure relationship, EpiGenSys WP4: Multi-scale simulations of chromatin architecture, EpiGenSys WP5: Globe 3D Genome Platform
Institutions: DKFZ, Erasumus Medical Center
Tools: *Not specified*
Phone: *Not specified*
Skype: tigertaknoch

No description set

Gernot Längst

Disciplines: *Not specified*
Roles: *Not specified*
Expertise: *Not specified*
Email: gernot.laengst@vkl.uni-regensburg.de
Web page: http://www.uni-r.de/laengst
Projects: EpiGenSys WP1: Nucleosome positioning, EpiGenSys WP2: Intra/inter chromosomal architecture, EpiGenSys WP3: Transcriptional structure relationship, EpiGenSys WP5: Globe 3D Genome Platform
Institutions: University of Regensburg, NWFIII, Biochemistry III
Tools: *Not specified*
Phone: 0941 943 2849
Skype: Längst Gernot

No description set

EpiGenSys Wiki

Edited Yesterday at 8:27 AM by epigensys...

Epi-genomic structure-function

Despite our knowledge of the sequence of the human genome, the relation of its dynamic three-dimensional architecture with its function – the storage and expression of genetic information – remains one of the central unresolved issues in biology.

The EpiGenSys consortium will combine experiment with theory to analyze the epi-genomic structure-function relationships. We will investigate: i) the nucleosome and chromatin fiber organization, ii) 3D architecture of the genome, and iii) the transcription structure-function relationship.

Groupleader: T.A. Knoch, F. Grosveld, P. Cook, K. Rippe, G. Wedemann, G. Längst

- Project Overview - EpiGenSys Meetings - Login for Members - Funding agencies - Publications - NEWS - Contact -

Work Packages

WP1 Nucleosomal association changes
WP2 Intra/inter chromosomal architecture
WP3 Transcriptional structure relationship
WP4 Simulations of nucleosomal, chromatin fiber and chromosome architecture and dynamics
WP5 System biological result integration via the GLOBE 3D Genome Platform

WP1 Nucleosomal association changes
May 14, 2011 3:30 PM

WP5 System biological result integration via the GLOBE 3D Genome Platform
May 14, 2011 2:56 PM

WP4 Simulations of nucleosomal, chromatin fiber and chromosome architecture and dynamics
May 14, 2011 2:55 PM

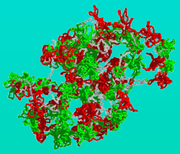
WP3 Transcriptional structure relationship
May 14, 2011 2:54 PM

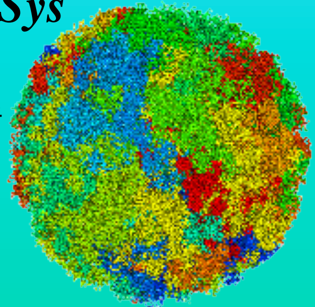
WP2 Intra/inter chromosomal architecture
May 14, 2011 2:53 PM

No items tagged with hot.

Publications
Yesterday at 8:57 AM

Joined with:
EpiGenSys
CancerEpiSys
EpiSys
MycNet



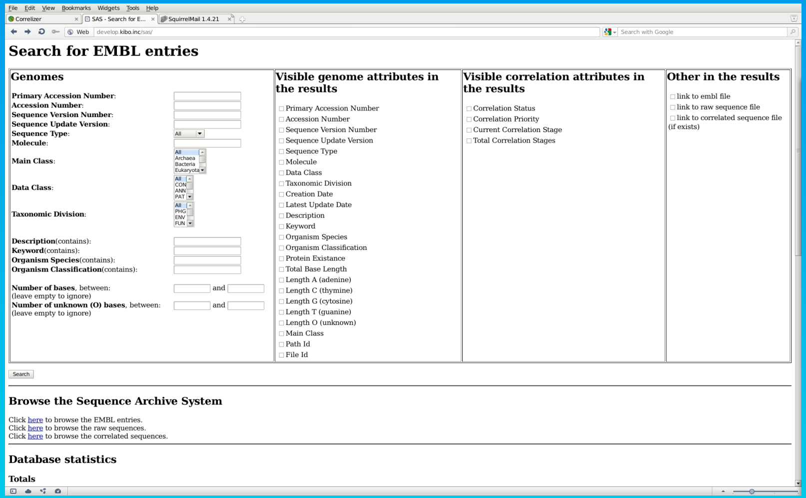


EpiGenSys International Large-Scale High-Throughput Data Analysis Centers and Grid Gateway WP5

The GLOBE 3D Genome Platform consists of individual components: i) the Sequence Archiving System stores complete genome sequences, ii) the Galaxy platform allows the analysis of genomes, and iii) for data analysis a gateway allows access to super-computer and grid access as our own established volunteer grid with a high public impact.

Bundesministerium für Bildung und Forschung
BBSRC bioscience for the future
NWO Nederlandse Organisatie voor Wetenschappelijk Onderzoek
EpiGenSys Consortium
Systems - Structure - Function
European Comission
ERASysBio ERA-NET for Systems Biology
Erasmus MC Eramus University Medical Center Rotterdam
BioQuant MODEL base of LIFE
University Heidelberg
MediGRID
University of Oxford
dkfz. German Cancer Research Center
University of Regensburg
fachhochschule stralsund University of Applied Sciences

EpiGenSys Sequence Archiving Systems of Completely Sequenced Genomes



Search for EMBL entries

Genomes

Primary Accession Number:
Accession Number:
Sequence Version Number:
Sequence Update Version:
Sequence Type: All
Molecule:
Main Class: All, Archaea, Bacteria, Eukaryota
Data Class: All, CON, ANN, PAT
Taxonomic Division: All, PHG, ENV, FUN
Description(contains):
Keyword(contains):
Organism Species(contains):
Organism Classification(contains):
Number of bases, between: and
(leave empty to ignore)
Number of unknown (O) bases, between: and
(leave empty to ignore)

Visible genome attributes in the results

- ☐ Primary Accession Number
- ☐ Accession Number
- ☐ Sequence Version Number
- ☐ Sequence Update Version
- ☐ Sequence Type
- ☐ Molecule
- ☐ Data Class
- ☐ Taxonomic Division
- ☐ Creation Date
- ☐ Latest Update Date
- ☐ Description
- ☐ Keyword
- ☐ Organism Species
- ☐ Organism Classification
- ☐ Protein Existance
- ☐ Total Base Length
- ☐ Length A (adenine)
- ☐ Length C (thymine)
- ☐ Length G (cytosine)
- ☐ Length T (guanine)
- ☐ Length O (unknown)
- ☐ Main Class
- ☐ Path Id
- ☐ File Id

Visible correlation attributes in the results

- ☐ Correlation Status
- ☐ Correlation Priority
- ☐ Current Correlation Stage
- ☐ Total Correlation Stages

Other in the results

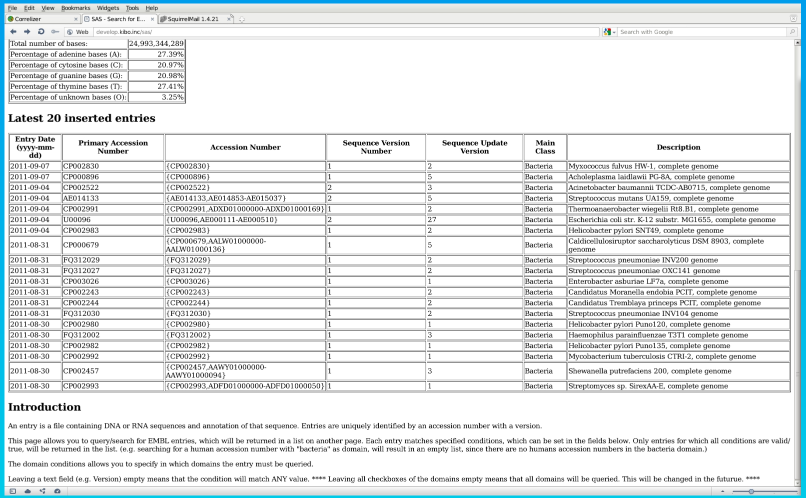
- ☐ link to embl file
- ☐ link to raw sequence file
- ☐ link to correlated sequence file (if exists)

Browse the Sequence Archive System

Click here to browse the EMBL entries.
Click here to browse the raw sequences.
Click here to browse the correlated sequences.

Database statistics

Totals



Total number of bases:	24,993,344,289
Percentage of adenine bases (A):	27.39%
Percentage of cytosine bases (C):	20.97%
Percentage of guanine bases (G):	20.98%
Percentage of thymine bases (T):	27.41%
Percentage of unknown bases (O):	3.25%

Latest 20 inserted entries

Entry Date (yyyy-mm-dd)	Primary Accession Number	Accession Number	Sequence Version Number	Sequence Update Version	Main Class	Description
2011-09-07	CP002830	{CP002830}	1	2	Bacteria	Myxococcus fulvus HW-1, complete genome
2011-09-07	CP000896	{CP000896}	1	5	Bacteria	Acholeplasma laidlawii PG-8A, complete genome
2011-09-04	CP002522	{CP002522}	2	3	Bacteria	Acinetobacter baumannii TCDC-AB0715, complete genome
2011-09-04	AE014133	{AE014133,AE014853-AE015037}	2	5	Bacteria	Streptococcus mutans UA159, complete genome
2011-09-04	CP002991	{CP002991,ADXD01000000-ADXD01000169}	1	2	Bacteria	Thermoanaerobacter wiegelii Rt8.B1, complete genome
2011-09-04	U00096	{U00096,AE000111-AE000510}	2	27	Bacteria	Escherichia coli str. K-12 substr. MG1655, complete genome
2011-09-04	CP002983	{CP002983}	1	2	Bacteria	Helicobacter pylori SNT49, complete genome
2011-08-31	CP000679	{CP000679,AALW01000000-AALW01000136}	1	5	Bacteria	Caldicellulosiruptor saccharolyticus DSM 8903, complete genome
2011-08-31	FQ312029	{FQ312029}	1	2	Bacteria	Streptococcus pneumoniae INV200 genome
2011-08-31	FQ312027	{FQ312027}	1	2	Bacteria	Streptococcus pneumoniae OXC141 genome
2011-08-31	CP003026	{CP003026}	1	1	Bacteria	Enterobacter asburiae LF7a, complete genome
2011-08-31	CP002243	{CP002243}	1	2	Bacteria	Candidatus Moranella endobia PCIT, complete genome
2011-08-31	CP002244	{CP002244}	1	2	Bacteria	Candidatus Tremblaya princeps PCIT, complete genome
2011-08-31	FQ312030	{FQ312030}	1	2	Bacteria	Streptococcus pneumoniae INV104 genome
2011-08-30	CP002980	{CP002980}	1	1	Bacteria	Helicobacter pylori Puno120, complete genome
2011-08-30	FQ312002	{FQ312002}	1	3	Bacteria	Haemophilus parainfluenzae T3T1 complete genome
2011-08-30	CP002982	{CP002982}	1	1	Bacteria	Helicobacter pylori Puno135, complete genome
2011-08-30	CP002992	{CP002992}	1	1	Bacteria	Mycobacterium tuberculosis CTRI-2, complete genome
2011-08-30	CP002457	{CP002457,AAWY01000000-AAWY01000094}	1	3	Bacteria	Shewanella putrefaciens 200, complete genome
2011-08-30	CP002993	{CP002993,ADFD01000000-ADFD01000050}	1	1	Bacteria	Streptomyces sp. SirexAA-E, complete genome

Introduction

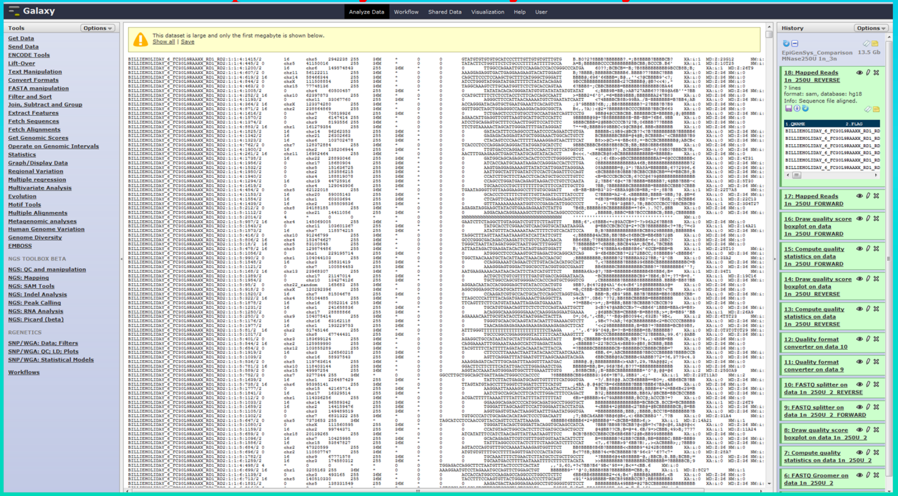
An entry is a file containing DNA or RNA sequences and annotation of that sequence. Entries are uniquely identified by an accession number with a version.

This page allows you to query/search for EMBL entries, which will be returned in a list on another page. Each entry matches specified conditions, which can be set in the fields below. Only entries for which all conditions are valid/true, will be returned in the list. (e.g. searching for a human accession number with "bacteria" as domain, will result in an empty list, since there are no humans accession numbers in the bacteria domain.)

The domain conditions allows you to specify in which domains the entry must be queried.

Leaving a text field (e.g. Version) empty means that the condition will match ANY value. **** Leaving all checkboxes of the domains empty means that all domains will be queried. This will be changed in the futuure. ****

EpiGenSys Galaxy Analysis



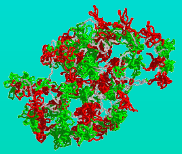
This dataset is large and only the first megabyte is shown below.

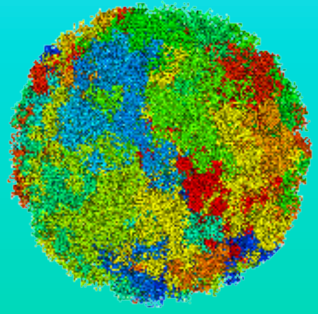
EpiGenSys_Comparison MNase250U 1n_3n 13.5 Gb

18: Mapped Reads 1n_250U_REVERSE
? lines
format: sam, database: hg18
Info: Sequence file aligned.

```
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
BILLIEHOLIDAY_6_FC30L9RAAXX_RD1_RD
```

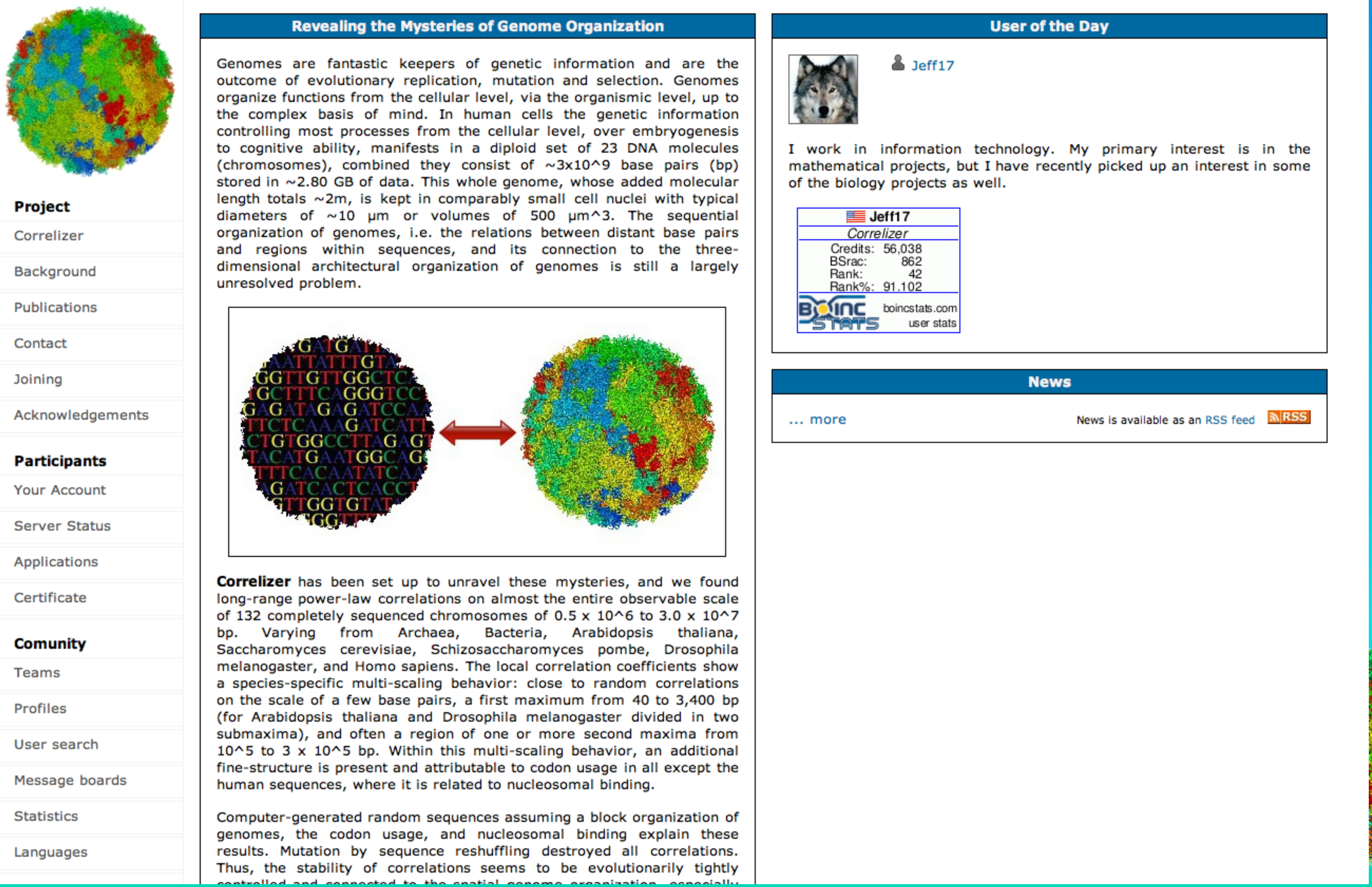
17: Mapped Reads 1n_250U_FORWARD
16: Draw quality score boxplot on data 1n_250U_FORWARD
15: Compute quality statistics on data 1n_250U_FORWARD
14: Draw quality score boxplot on data 1n_250U_REVERSE
13: Compute quality statistics on data 1n_250U_REVERSE
12: Quality format converter on data 10
11: Quality format converter on data 9
10: FASTQ splitter on data 1n_250U_2_REVERSE
9: FASTQ splitter on data 1n_250U_2_FORWARD
8: Draw quality score boxplot on data 1n_250U_2
7: Compute quality statistics on data 1n_250U_2
6: FASTQ Groomer on data 1n_250U_2





EpiGenSys International Large-Scale High-Throughput Data Analysis Centers and Grid Gateway WP5

The GLOBE 3D Genome Platform consists of individual components: i) the Sequence Archiving System stores complete genome sequences, ii) the Galaxy platform allows the analysis of genomes, and iii) for data analysis a gateway allows access to super-computer and grid access as our own established volunteer grid with a high public impact.



Project
- Correlizer
- Background
- Publications
- Contact
- Joining
- Acknowledgements

Participants
- Your Account
- Server Status
- Applications
- Certificate

Comunity
- Teams
- Profiles
- User search
- Message boards
- Statistics
- Languages

Revealing the Mysteries of Genome Organization

Genomes are fantastic keepers of genetic information and are the outcome of evolutionary replication, mutation and selection. Genomes organize functions from the cellular level, via the organismic level, up to the complex basis of mind. In human cells the genetic information controlling most processes from the cellular level, over embryogenesis to cognitive ability, manifests in a diploid set of 23 DNA molecules (chromosomes), combined they consist of ~3x10^9 base pairs (bp) stored in ~2.80 GB of data. This whole genome, whose added molecular length totals ~2m, is kept in comparably small cell nuclei with typical diameters of ~10 µm or volumes of 500 µm^3. The sequential organization of genomes, i.e. the relations between distant base pairs and regions within sequences, and its connection to the three-dimensional architectural organization of genomes is still a largely unresolved problem.

Correlizer has been set up to unravel these mysteries, and we found long-range power-law correlations on almost the entire observable scale of 132 completely sequenced chromosomes of 0.5 x 10^6 to 3.0 x 10^7 bp. Varying from Archaea, Bacteria, Arabidopsis thaliana, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster, and Homo sapiens. The local correlation coefficients show a species-specific multi-scaling behavior: close to random correlations on the scale of a few base pairs, a first maximum from 40 to 3,400 bp (for Arabidopsis thaliana and Drosophila melanogaster divided in two submaxima), and often a region of one or more second maxima from 10^5 to 3 x 10^5 bp. Within this multi-scaling behavior, an additional fine-structure is present and attributable to codon usage in all except the human sequences, where it is related to nucleosomal binding.

Computer-generated random sequences assuming a block organization of genomes, the codon usage, and nucleosomal binding explain these results. Mutation by sequence reshuffling destroyed all correlations. Thus, the stability of correlations seems to be evolutionarily tightly

User of the Day

Jeff17

I work in information technology. My primary interest is in the mathematical projects, but I have recently picked up an interest in some of the biology projects as well.

Jeff17
Correlizer

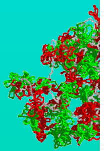
Credits:	56,038
BSrac:	862
Rank:	42
Rank%:	91.102

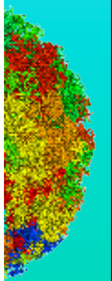
BOINC STATS boincstats.com user stats

News

... more

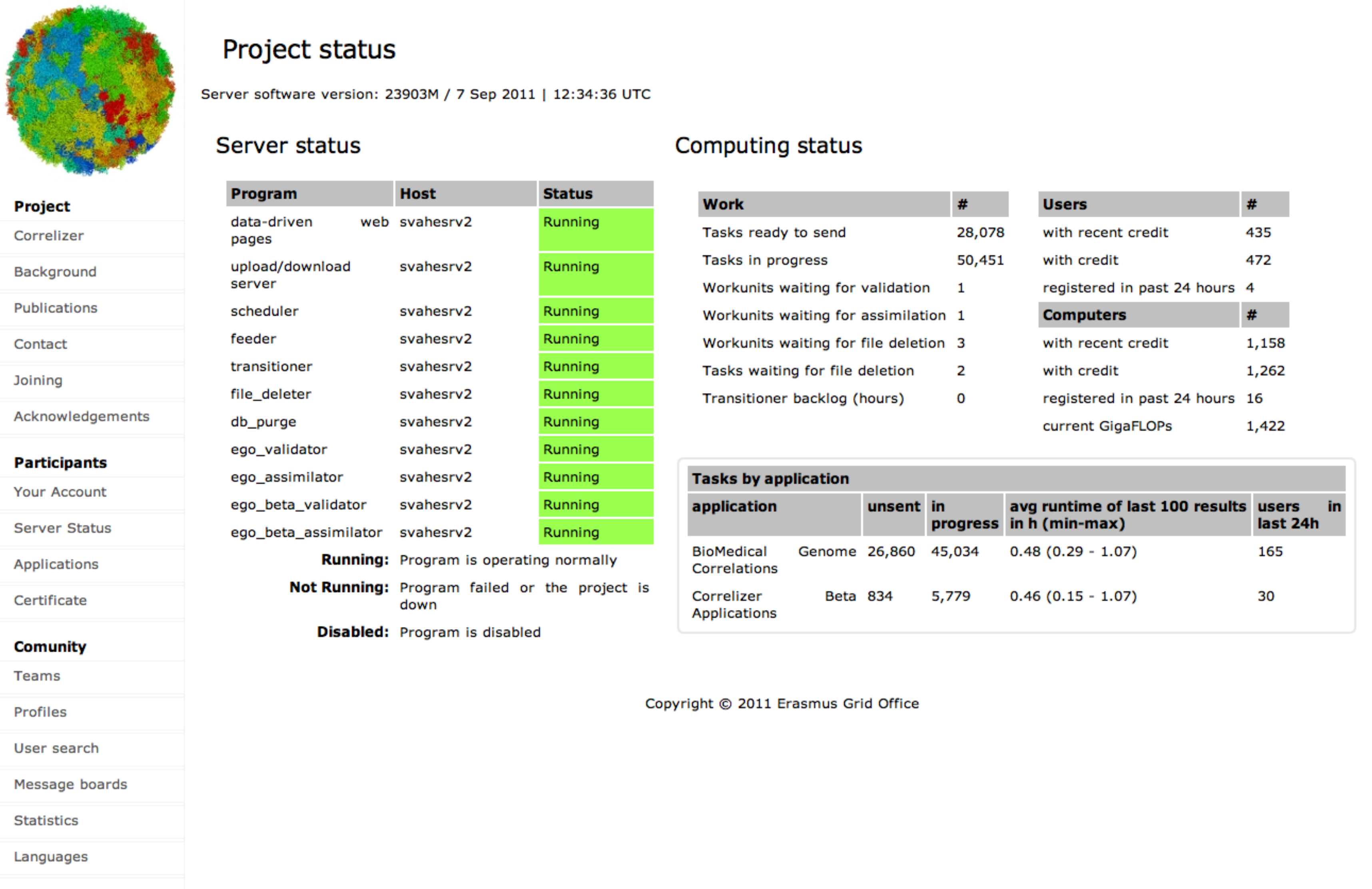
News is available as an RSS feed RSS





EpiGenSys International Large-Scale High-Throughput Data Analysis Centers and Grid Gateway WP5

The GLOBE 3D Genome Platform consists of individual components: i) the Sequence Archiving System stores complete genome sequences, ii) the Galaxy platform allows the analysis of genomes, and iii) for data analysis a gateway allows access to super-computer and grid access as our own established volunteer grid with a high public impact.



Project
- Correlizer
- Background
- Publications
- Contact
- Joining
- Acknowledgements

Participants
- Your Account
- Server Status
- Applications
- Certificate

Comunity
- Teams
- Profiles
- User search
- Message boards
- Statistics
- Languages

Project status

Server software version: 23903M / 7 Sep 2011 | 12:34:36 UTC

Server status

Program	Host	Status
data-driven web pages	svahesrv2	Running
upload/download server	svahesrv2	Running
scheduler	svahesrv2	Running
feeder	svahesrv2	Running
transitioner	svahesrv2	Running
file_deleter	svahesrv2	Running
db_purge	svahesrv2	Running
ego_validator	svahesrv2	Running
ego_assimilator	svahesrv2	Running
ego_beta_validator	svahesrv2	Running
ego_beta_assimilator	svahesrv2	Running

Running: Program is operating normally

Not Running: Program failed or the project is down

Disabled: Program is disabled

Computing status

Work	#
Tasks ready to send	28,078
Tasks in progress	50,451
Workunits waiting for validation	1
Workunits waiting for assimilation	1
Workunits waiting for file deletion	3
Tasks waiting for file deletion	2
Transitioner backlog (hours)	0

Users	#
with recent credit	435
with credit	472
registered in past 24 hours	4
Computers	**#**
with recent credit	1,158
with credit	1,262
registered in past 24 hours	16
current GigaFLOPs	1,422

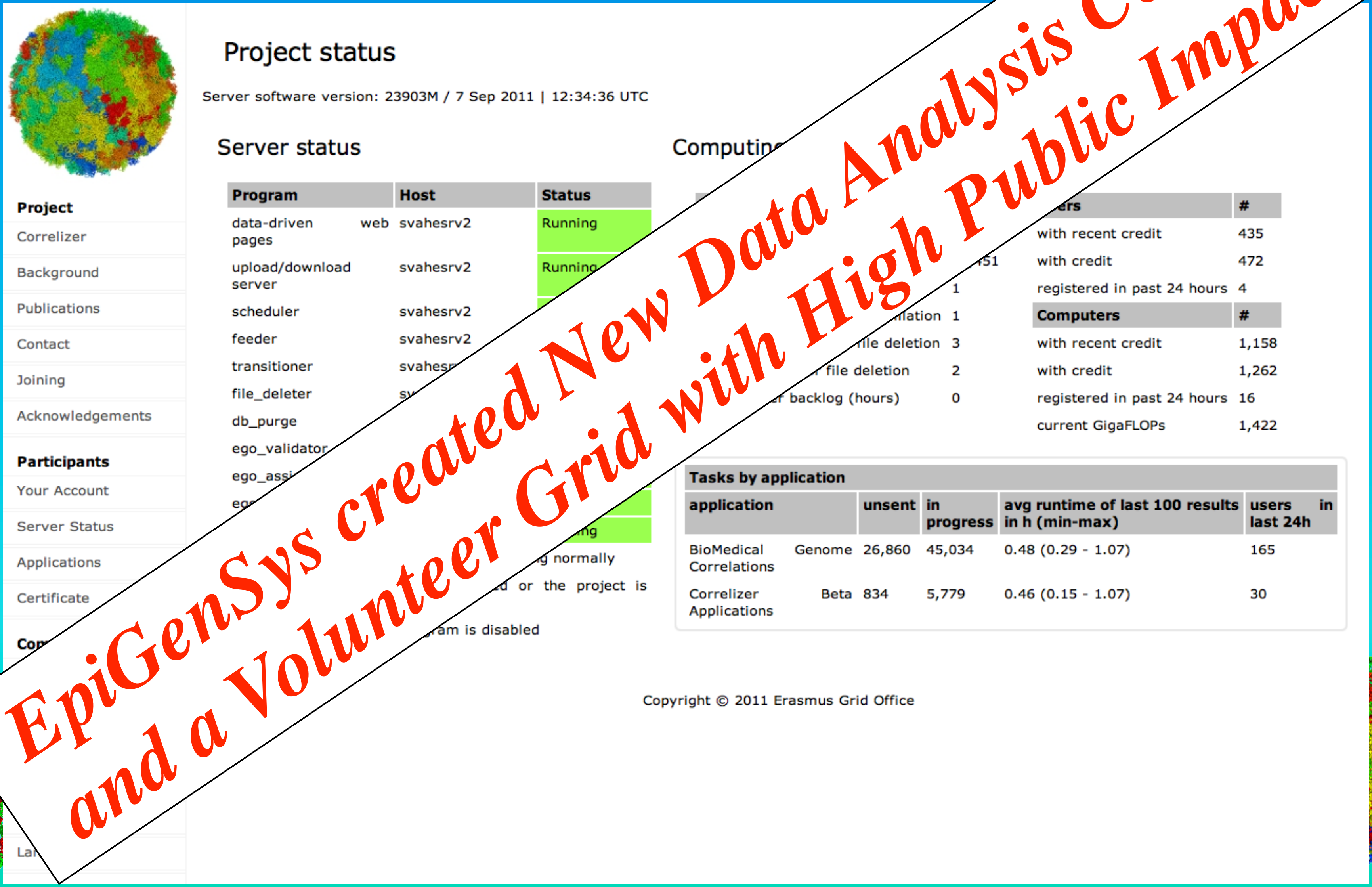
Tasks by application

application	unsent	in progress	avg runtime of last 100 results in h (min-max)	users in last 24h
BioMedical Genome Correlations	26,860	45,034	0.48 (0.29 - 1.07)	165
Correlizer Beta Applications	834	5,779	0.46 (0.15 - 1.07)	30

Copyright © 2011 Erasmus Grid Office

EpiGenSys International Large-Scale High-Throughput Data Analysis Centers and Grid Gateway WP5

The GLOBE 3D Genome Platform consists of individual components: i) the Sequence Archiving System stores co
genome sequences, ii) the Galaxy platform allows the analysis of genomes, and iii) for data analysis a gatew
access to super-computer and grid access as our own established volunteer grid with a high public i



Project status

Server software version: 23903M / 7 Sep 2011 | 12:34:36 UTC

Project
Correlizer
Background
Publications
Contact
Joining
Acknowledgements

Participants
Your Account
Server Status
Applications
Certificate

Com

La

Server status

Program	Host	Status
data-driven web pages	svahesrv2	Running
upload/download server	svahesrv2	Running
scheduler	svahesrv2	
feeder	svahesrv2	
transitioner	svahes	
file_deleter	sv	
db_purge		
ego_validator		
ego_assi		
eg		
		ng

g normally
d or the project is
am is disabled

Computin

451
1
lation 1
ile deletion 3
file deletion 2
backlog (hours) 0

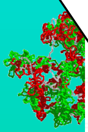
ers	#
with recent credit	435
with credit	472
registered in past 24 hours	4
Computers	**#**
with recent credit	1,158
with credit	1,262
registered in past 24 hours	16
current GigaFLOPs	1,422

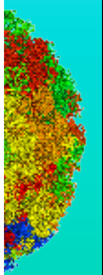
Tasks by application

application	unsent	in progress	avg runtime of last 100 results in h (min-max)	users in last 24h
BioMedical Genome Correlations	26,860	45,034	0.48 (0.29 - 1.07)	165
Correlizer Beta Applications	834	5,779	0.46 (0.15 - 1.07)	30

Copyright © 2011 Erasmus Grid Office

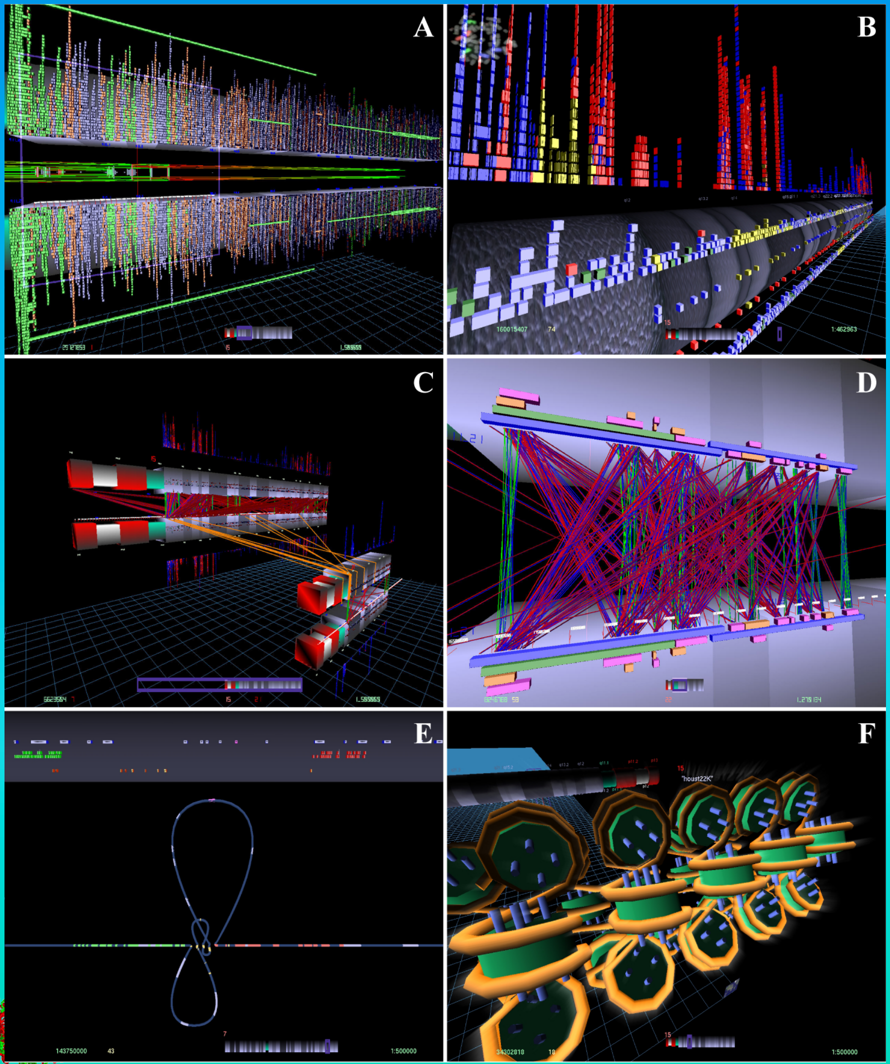
EpiGenSys created New Data Analysis Centers and a Volunteer Grid with High Public Impact!

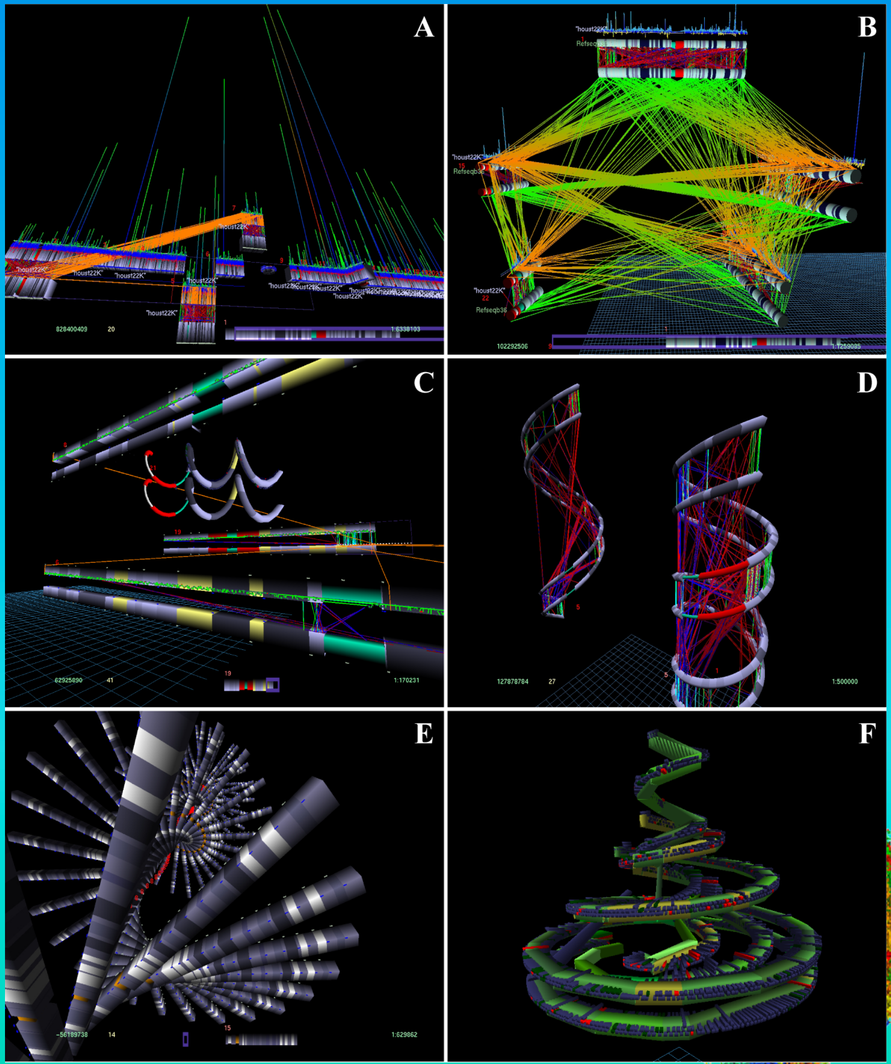




EpiGenSys Systems Biological Result Integration via the GLOBE 3D Genome Platform WP5

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.

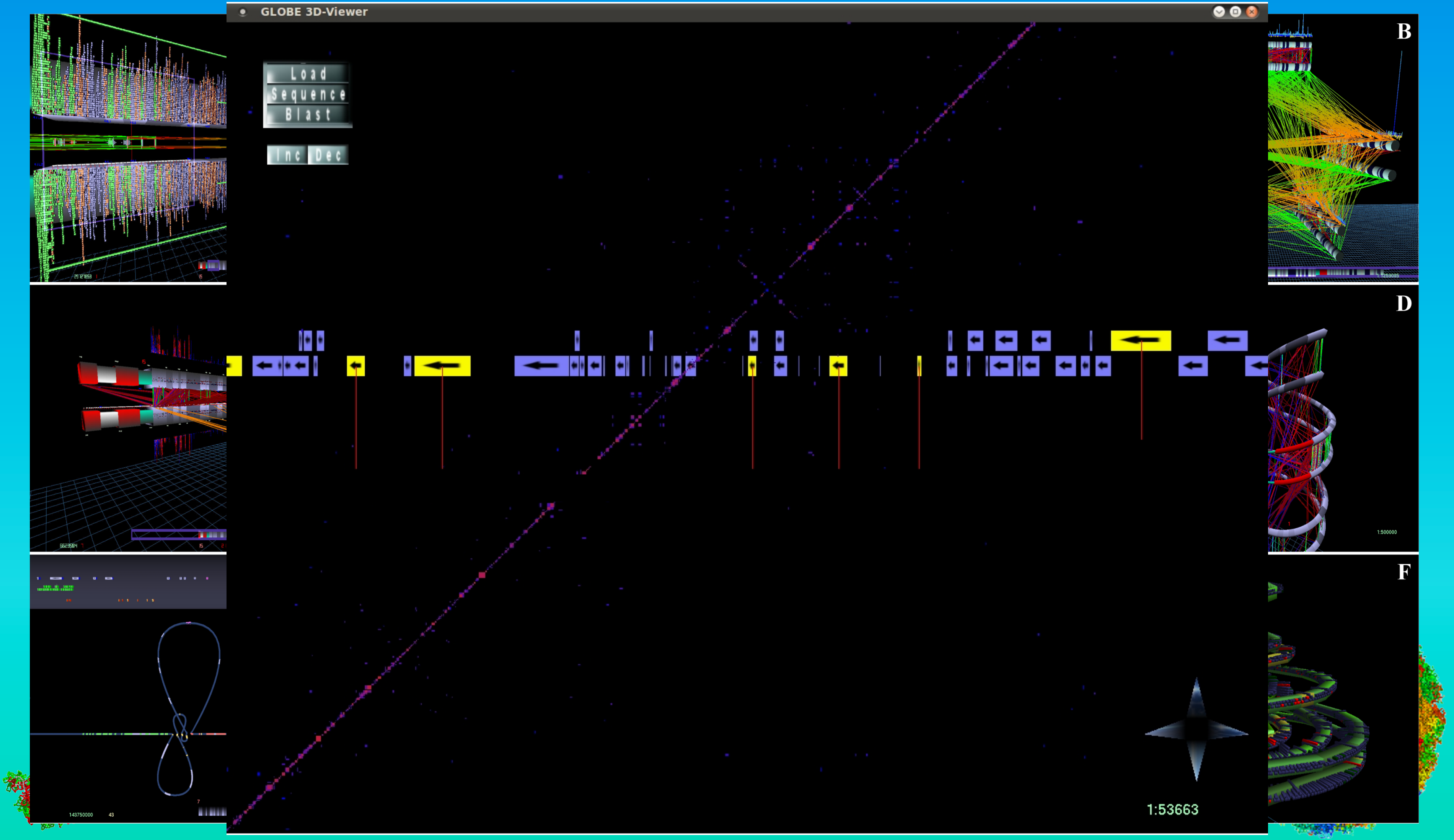




EpiGenSys Systems Biological Result Integration via the GLOBE 3D Genome Platform WP5

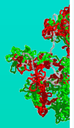
All results will be integrated using our GLOBE 3D Genome Platform, established for analysis, manipulation and understanding of multi-dimensional complex genome wide data. Thus in reiterative cycles between experiments and simulations a systems biological/medical genome model will be achieved.

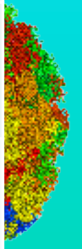
Bundesministerium für Bildung und Forschung
BBSRC
NWO Nederlandse Organisatie voor Wetenschappelijk Onderzoek
EpiGenSys Consortium
ERASysBio ERA-NET for Systems Biology
European Commission
BioQuant
Erasmus MC
University Heidelberg
University of Oxford
dkfz. German Cancer Research Center
UR University of Regensburg
MedGRID
fachhochschule stralsund



Result Dissemination of EpiGenSys

Dissemination Kind	Current Status	Initial Projection
Research Articles	20 (17 in prep.) => >32	~15
Reviews	4 (4 in prep.) => >8 ========= 24 (21 in prep.) => >45	N.N. ======== ~15
IP-Rights	~3-4 in the end	~5
Conference Contributions	>70 latest estimate	N.N.
New Scientific “Relations”	~30	~20
New Industry “Relations”	~8	~10

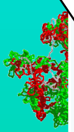


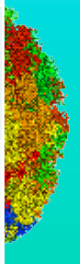


Result Dissemination of EpiGenSys

Dissemination Kind	**Current Status**	**on**
Research Articles	20 (17 in prep.)	~15
Reviews	>45	N.N. ======= ~15
IP-Rights	3-4 in the end	~5
Confe	>70 latest estimate	N.N.
ons”	~30	~20
N stry “Relations”	~8	~10

EpiGenSys is already and will be much more successfull than we previously thought !

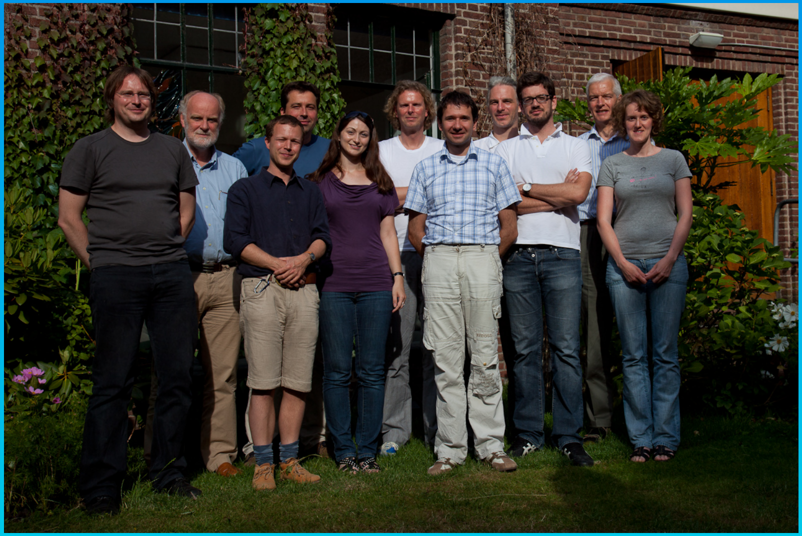




Acknowledgements

Thanks go to all the lab local lab members, 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.



Bundesministerium für Bildung und Forschung

BBSRC bioscience for the future

NWO
Nederlandse Organisatie voor Wetenschappelijk Onderzoek

EpiGenSys Consortium
Systems - Structure - Function

ERASysBio ERA-NET for Systems Biology

European Comission

BioQuant MODEL base of LIFE

Erasmus MC
Eramus University Medical Center Rotterdam

University Heidelberg

MediGRID

UNIVERSITY OF OXFORD

dkfz. German Cancer Research Center

fachhochschule stralsund University of Applied Sciences

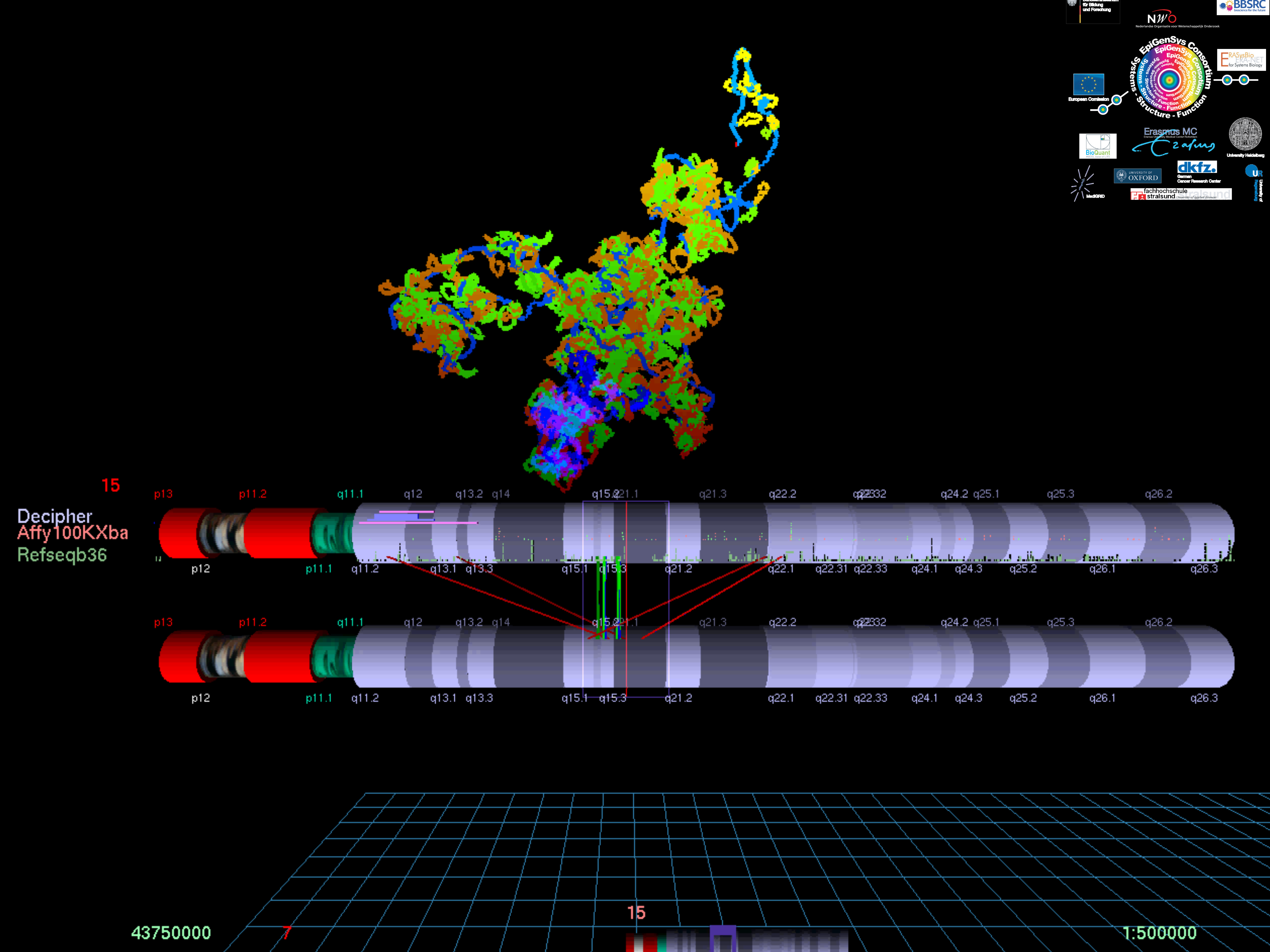
UR University of Regensburg

Acknowledgements

Thanks go to all the lab local lab members, those people who supported this work in the last deca
institutions providing their infrastructure, and the national and international computing infr

Special thanks go to the reviewers, the EraSysBio Plus initiative and the national and E

Towards a Holistic Genome Understanding by a Systems Biological/Medical Model for Research, Diagnostics and Treatment!


15
Decipher
Affy100KXba
Refseqb36
p13
p11.2
q11.1
q12
q13.2
q14
q15.2
q21.1
q21.3
q22.2
q23.2
q24.2
q25.1
q25.3
q26.2
p12
p11.1
q11.2
q13.1
q13.3
q15.1
q15.3
q21.2
q22.1
q22.31
q22.33
q24.1
q24.3
q25.2
q26.1
q26.3
43750000
7
15
1:500000

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 and Dynamics, and Systems Biological/Medical Result Integration via the GLOBE 3D Genome Platform

Knoch, T. A.,
Cook, P. R., Rippe, K, Längst, G, Wedemann, G., & Grosveld, F. G.

EpiGenSys Final Meeting and Evaluation Conference, Hotel Angleterre,
Berlin, 27th June, 2013.

Abstract

Although the sequence of the human genome is known, 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 time. It has become clear that genomes are tremendous co-evolutionary and interwoven molecular storage machines able to manipulate and fabricate information: the genetic information is coded in and along (macro-) molecules that are continually being modified spatially and temporally through a multi-dimensional interaction and regulatory network. Thus, changes in chromatin architecture are central factors in the epigenetic regulation of gene expression and other important genomic processes on length and time scales ranging from 10^{-9} to 10^{-5} m and 10^{-10} to 10^{5} s. Thus, a full understanding of structure-function relationships requires knowledge not only of the linear base-pair composition, but also of its structural and dynamic organization.

Therefore, in "EpiGenSys" we have established an unique consortium with the aim of achieving a major breakthrough in the determination and understanding of the relation between DNA sequence, epigenetic modification and spatial chromatin organization. Using truly systems approaches, we studied the following:

i) The investigation of nucleosome and chromatin fiber structure and remodelling.
ii) The determination of intra/inter chromosomal interactions and chromosomal organization.
iii) The analysis of transcriptional states and their relation to the underlying structure.
iv) The simulation of nucleosomes, chromatin fibers, and chromosomes to provide theoretical insight.
v) The integration of i) to iv) into a systems biological model.

We strongly believe that by addressing these fundamental questions using our systems approach to refine our models through reiterative cycling between experiment and theory, we were able to address a central issue of the genomic era - the way structure influences gene activity (and vice versa).

Thus, we consider the results achieved to be major breakthroughs in all those above mentioned points. We also anticipate that our results have great impact on wider research communities, and on the development of methods

useful for disease diagnosis and treatment e.g. as envisioned by stem cell and gene therapy. This is shown by the number of publications, contributions to conferences and also the still ongoing patenting process.
Thus, we think to have made and still make a considerable contribution to systems biology in the genomic sector and in general, as well as contribute to e-Science, e-Health, e-Learning, as well as e-Commerce and increase the 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, multilism genotype-phenotype, 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 quasi 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, polymer model, analytic mathematical model, Brownian Dynamics, Monte Carlo, fluorescence *in situ* hybridization (FISH), targeted chromatin capture (T2C) confocal laser scanning microscopy, fluorescence correlation spectroscopy, spatial precision distance microscopy, super-resolution microscopy, two dimensional fluorescence correlations spectroscopy (2D-FCS) 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, human ecology, e-human grid ecology, society, social systems, e-social challenge, inverse tragedy of the commons, grid phenomenon, micro-sociality, macro-sociality, autopoietic tragedy of social sub-systems, micro subsystems, macro subsystems, micro operationality, macro operationality, grid psychology micro riskmanagement, macro riskmanagement.

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.

Gil-Parado, S., Fernández-Montalván, A., Assfalg-Machleidt, I., Popp, O., Bestvater, F., Holloschi, A., **Knoch, T. A.**, Auerswald, E. A., Welsh, K., Reed, J. C., Fritz, H., Fuentes-Prior, P., Spiess, E., Salvesen, G. & Machleidt, W. Ionomycin-activated calpain triggers apoptosis: A probable role for Bcl-2 family members. *J. Biol. Chem.* 277(30), 27217-27226, 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.

Westphal, G., van den Berg-Stein, S., Braun, K., **Knoch, T. A.**, Dümmerling, M., Langowski, J., Debus, J. & Friedrich, E. Detection of the NGF receptors TrkaA and p75NTR and effect of NGF on the growth characteristics of human tumor cell lines. *J. Exp. Clin. Canc. Res.* 21(2), 255-267, 2002.

Westphal, G., Niederberger, E., Blum, C., Wollman, Y., **Knoch, T. A.**, Dümmerling, M., Rebel, W., Debus, J. & Friedrich, E. Erythropoietin Receptor in Human Tumor Cells: Expression and Aspects Regarding Functionality. *Tumori* 88(2), 150-159, 2002.

Gil-Parado, S., Popp, O., **Knoch, T. A.**, Zahler, S., Bestvater, F., Felgenträger, M., Holoshi, A., Fernández-Montalván, A., Auerswald, E. A., Fritz, H., Fuentes-Prior, P., Machleidt, W. & Spiess, E. Subcellular localization and subunit interactions of over-expressed human μ-calpain. *J. Biol. Chem. 278(18),* 16336-15346, 2003.

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.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Yassene, M., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. Grid-basierte Services für die elektronische Patientenakte der Zukunft. *E- HEALTH-COM - Magazin für Gesundheitstelematik und Telemedizin*, 4(2), 61-63, 2007.

de Zeeuw, L. V., **Knoch, T. A.,** van den Berg, J. & Grosveld, F. G. Erasmus Computing Grid - Het bouwen van een 20 TeraFLOP virtuelle supercomputer. *NIOC proceedings 2007 - het perspective of lange termijn*. editor Frederik, H. NIOC, Amsterdam, The Netherlands, 52-59, 2007.

Rauch, J., **Knoch, T. A.**, Solovei, I., Teller, K. Stein, S., Buiting, K., Horsthemke, B., Langowski, J., Cremer, T., Hausmann, M. & Cremer, C. Lightoptical precision measurements of the Prader- Willi/Angelman Syndrome imprinting locus in human cell nuclei indicate maximum condensation changes in the few hundred nanometer range. *Differentiation* 76(1), 66-82, 2008.

Sax, U., Weisbecker, A., Falkner, J., Viezens, F., Mohammed, Y., Hartung, M., Bart, J., Krefting, D., **Knoch, T. A.** & Semler, S. C. Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft. *Telemedizinführer Deutschland 2008*, editor Jäckel, A. Deutsches Medizinforum, Minerva KG, Darmstadt, ISBN 3-937948-06-6, ISBN-13 9783937948065, 47-51, 2008.

Drägestein, K. A., van Capellen, W. A., van Haren, J. Tsibidis, G. D., Akhmanova, A., **Knoch, T. A.**, Grosveld, F. G. & Galjart, N. Dynamic behavior of GFP-CLIP-170 reveals fast protein turnover on microtubule plus ends. *J. Cell Biol. 180(4)*, 729-737, 2008.

Jhunjhunwala, S., van Zelm, M. C., Peak, M. M., Cutchin, S., Riblet, R., van Dongen, J. J. M., Grosveld, F. G., **Knoch, T. A.**[+] & Murre, C.[+] The 3D-structure of the Immunoglobulin Heavy Chain Locus: implications for long-range genomic interactions. *Cell 133(2)*, 265-279, 2008.

Krefting, D., Bart, J., Beronov, K., Dzhimova, O., Falkner, J., Hartung, M., Hoheisel, A., **Knoch, T. A.**, Lingner, T., Mohammed, Y., Peter, K., Rahm, E., Sax, U., Sommerfeld, D., Steinke, T., Tolxdorff, T., Vossberg, M., Viezens, F. & Weisbecker, A. MediGRID - Towards a user friendly secured grid infrastructure. *Future Generation Computer Systems 25(3)*, 326-336, 2008.

Knoch, T. A., Lesnussa, M., Kepper, F. N., Eussen, H. B., & Grosveld, F. G. The GLOBE 3D Genome Platform - Towards a novel system-biological paper tool to integrate the huge complexity of genome organization and function. *Stud. Health. Technol. Inform. 147*, 105-116, 2009.

Knoch, T. A., Baumgärtner, V., de Zeeuw, L. V., Grosveld, F. G., & Egger, K. e-Human Grid Ecology: Understanding and approaching the Inverse Tragedy of the Commons in the e-Grid Society. *Stud. Health. Technol. Inform. 147*, 269-276, 2009.

Dickmann, F., Kaspar, M., Löhnardt, B., **Knoch, T. A.,** & Sax, U. Perspectives of MediGRID. *Stud. Health. Technol. Inform. 147*, 173-182, 2009.

Knoch, T. A., Göcker, M., Lohner, R., Abuseiris, A. & Grosveld, F. G. Fine-structured multi-scaling long-range correlations in completely sequenced genomes - features, origin and classification. *Eur. Biophys. J. 38(6)*, 757-779, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *LNCS 5745*, 150-159, 2009.

Dickmann, F., Kaspar, M., Löhnhardt, B., Kepper, N., Viezens, F., Hertel, F., Lesnussa, M., Mohammed, Y., Thiel, A., Steinke, T., Bernarding, J., Krefting, D., **Knoch, T. A.** & Sax, U. Visualization in health-grid environments: a novel service and business approach. *Grid economics and business models - GECON 2009 Proceedings, 6th international workshop, Delft, The Netherlands.* editors Altmann, J., Buyya, R. & Rana, O. F., GECON 2009, LNCS 5745, Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-03863-1, 150-159, 2009.

Estrada, K.[*], Abuseiris, A.[*], Grosveld, F. G., Uitterlinden, A. G., **Knoch, T. A.**[+] & Rivadeneira, F.[+] GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics 25(20), 2750-2752*, 2009.

de Wit, T., Dekker, S., Maas, A., Breedveld, G., **Knoch, T. A.,** Langeveld, A., Szumska, D., Craig, R., Bhattacharya, S., Grosveld, F. G.[+] & Drabek, D. Tagged mutagenesis of efficient minos based germ line transposition. *Mol. Cell Biol 30(1)*, 66-77, 2010.

Kepper, N., Schmitt, E., Lesnussa, M., Weiland, Y., Eussen, H. B., Grosveld, F. G., Hausmann, M. & **Knoch T. A.,** Visualization, Analysis, and Design of COMBO-FISH Probes in the Grid-Based GLOBE 3D Genome Platform. *Stud. Health Technol. Inform. 159*, 171-180, 2010.

Kepper, N., Ettig, R., Dickmann, F., Stehr, R., Grosveld, F. G., Wedemann, G. & **Knoch, T. A.** Parallel high-performance grid computing: capabilities and opportunities of a novel demanding service and business class allowing highest resource efficiency. *Stud. Health Technol. Inform. 159*, 264-271, 2010.

Skrowny, D., Dickmann, F., Löhnhardt, B., **Knoch, T. A.** & Sax, U. Development of an information platform for new grid users in the biomedical field. *Stud. Health Technol. Inform. 159*, 277-282, 2010.

Knoch, T. A., Baumgärtner, V., Grosveld, F. G. & Egger, K. Approaching the internalization challenge of grid technologies into e-Society by e-Human "Grid" Ecology. *Economics of Grids, Clouds, Systems, and Services – GECON 2010 Proceedings*, 7th International Workshop, Ischia, Italy, editors Altman, J., & Rana, O. F., Lecture Notes in Computer Science (LNCS) 6296, Springer Berlin Heidelberg New York, ISSN 0302-9743, ISBN-10 3-642-15680-0, ISBN-13 978-3-642-15680-9, 116-128, 2010.

Dickmann, F., Brodhun, M., Falkner, J., **Knoch, T. A.** & Sax, U. Technology transfer of dynamic IT outsourcing requires security measures in SLAs. *Economics of Grids, Clouds, Systems, and Services – GECON 2010 Proceedings*, 7th International Workshop, Ischia, Italy, editors Altman, J., & Rana, O. F., Lecture Notes in Computer Science (LNCS) 6296, Springer Berlin Heidelberg New York, ISSN 0302-9743, ISBN-10 3-642-15680-0, ISBN-13 978-3-642-15680-9, 1-115, 2010.

Knoch, T. A. Sustained Renewability: approached by systems theory and human ecology. *Renewable Energy 2*, editors M. Nayeripour & M. Keshti, Intech, ISBN 978-953-307-573-0, 21-48, 2011.

Kolovos, P., **Knoch, T. A.**, F. G. Grosveld, P. R. Cook, & Papantonis, A. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics and Chromatin 5(1)*, 1-8, 2012.

Dickmann, F., Falkner, J., Gunia, W., Hampe, J., Hausmann, M., Herrmann, A., Kepper, N., **Knoch, T. A.**, Lauterbach, S., Lippert, J., Peter, K., Schmitt, E., Schwardmann, U., Solodenko, J., Sommerfeld, D., Steinke, T., Weisbecker, A. & Sax, U. Solutions for Biomedical Grid Computing - Case Studies from the D-Grid Project Services@MediGRID. *JOCS 3(5)*, 280-297, 2012.

Estrada, K., Abuseiris, A., Grosveld, F. G., Uitterlinden, A. G., **Knoch, T. A.** & Rivadeneira, F. GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Dissection of the complex genetic architecture of human stature and osteoporosis.* cummulative dissertation, editor Estrada K., Erasmus Medical Center, Erasmus University Rotterdam, Rotterdam, The Netherlands, ISBN 978-94-6169-246-7, 25-30, 1st June 2012.

van de Corput, M. P. C., de Boer, E., **Knoch, T. A.**, van Cappellen, W. A., Quintanilla, A., Ferrand, L., & Grosveld, F. G. Super-resolution imaging reveals 3D folding dynamics of the β-globin locus upon gene activation. *J. Cell Sci. 125 (Pt 19)*, 4630-4639, 2012.

da Silva, P. S. D., Delgado Bieber, A. G., Leal, I. R., **Knoch, T. A.**, Tabarelli, M., Leal, I. R., & Wirth, R. Foraging in highly dynamic environments: leaf-cutting ants adjust foraging trail networks to pioneer plant availability. *Entomologia Experimentalis et Applicata 147*, 110-119. 2013.

Zuin, J., Dixon, J. R., van der Reijden, M. I. J. A., Ye, Z., Kolovos, P., Brouwer, R. W. W., van de Corput, M. P. C., van de Werken, H. J. G., **Knoch, T. A.**, van IJcken, W. F. J., Grosveld, F. G., Ren, B. & Wendt, K. S. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *PNAS 111(3)*, 9906-1001, 2014.

Kolovos, P., Kepper, N., van den Werken, H. J. G., Lesnussa, M., Zuin, J., Brouwer, R. W. W., Kockx, C. E. M., van IJcken, W. F. J., Grosveld, F. G. & **Knoch, T. A.** Targeted Chromatin Capture (T2C): A novel high resolution high throughput method to detect genomic interactions and regulatory elements. *Epigenetics & Chromatin 7:10*, 1-17, 2014.

Diermeier, S., Kolovos, P., Heizinger, L., Schwartz, U., Georgomanolis, T., Zirkel, A., Wedemann, G., Grosveld, F. G., **Knoch, T. A.**, Merkl, R., Cook, P. R., Längst, G. & Papantonis, A. TNFα signalling primes chromatin for NF-kB binding and induces rapid and widespread nucleosome repostioning. *Genome Biology 15(12)*, 536-548, 2014.

Knoch, T. A., Wachsmuth, M., Kepper, N., Lesnussa, M., Abuseiris, A., A. M. Ali Imam, Kolovos, P., Zuin, J., Kockx, C. E. M., Brouwer, R. W. W., van de Werken, H. J. G., van IJcken, W. F. J., Wendt, K. S. & Grosveld, F. G. The detailed 3D multi-loop aggregate/rosette chromatin architecture and functional dynamic organization of the human and mouse genomes. *bioXriv preprint*, 16.07.2016.

Kolovos, P., Georgomanolis, T., Koeferle, A., Larkin, J. D., Brant, J., Nikolić, M., Gusmao, E. G., Zirkel, A., **Knoch, T. A.**, van IJcken, W. F. J., Cook, P. R., Costa, I. G., Grosveld, F. G. & Papantonis, A. Binding of nuclear kappa-B to non-canonical consensus sites reveals its multimodal role during the early inflammatory response. *Genome Research* 26(11), 1478-1489, 2016.

Wachsmuth, M., **Knoch, T. A.** & Rippe, K. Dynamic properties of independent chromatin domains measured by correlation spectroscopy in living cells. *Epigenetics & Chromatin 9:57*, 1-20, 2016.

Knoch, T. A., Wachsmuth, M., Kepper, N., Lesnussa, M., Abuseiris, A., A. M. Ali Imam, Kolovos, P., Zuin, J., Kockx, C. E. M., Brouwer, R. W. W., van de Werken, H. J. G., van IJcken, W. F. J., Wendt, K. S. & Grosveld, F. G. The detailed 3D multi-loop aggregate/rosette chromatin architecture and functional dynamic organization of the human and mouse genomes. *Epigenetics & Chromatin 9:58*, 1-22, 2016.