4DNucleome Initiative in Europe

White paper, 29 April 2016

http://www.4DNucleome.eu

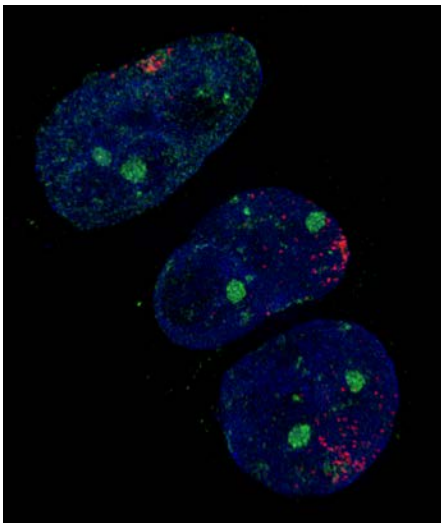
> **"We are looking for the big picture! We expect new ideas for grand challenges in science and technology requiring large scale collaboration and multidisciplinary exploration."**
>
> *Thierry Van der Pyl,*
> *Director for Excellence in Science*

With the above quote, Thierry Van der Pyl called upon the European scientific community to propose grand challenges to be addressed with newly funded FET FLAGSHIPS by 2018. This white-paper is a response to that call and proposes to the European Commission to launch a €1 billion large-scale initiative to decipher the structure-function relationships of the cell nucleus as a complex biological system at all levels from molecules to the entire genomic and epigenomic landscapes as they adapt to environmental changes, cell-cell-interactions, tissue organization and changes during cell reprogramming and during ageing. **This is one of the biggest challenge in the post-genomic era**.

New ideas for grand challenges

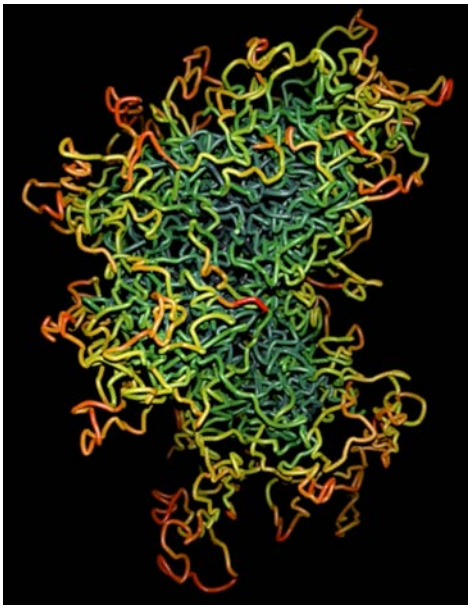
Our genome, our DNA, has taken a central place in our daily life, whether we think about our health, our well-being and longevity, our susceptibility to disease, our aptitude for learning, and our adaptation and responses to diet, drugs and the environment. **Yet despite having successfully sequenced the full human genome, it is not enough to make sense of its function as a whole. Developing the technologies to empower the community to address this major issue represents one of the most significant S&T challenge that the biomedical world will have to deal with in the next decade to come.**

Grand challenge. An exciting emerging view is that the genome offers a lot more than a simple passive library of genetic information. It functions as a powerful information retrieval system that dynamically changes shape and conformation to selectively expose selectively the genetic information (genes) that are required for a particular cell type, while filing away and hiding other information that is not required. To adapt to the diverse aspects of our life such as embryonic development, ageing, response to diet, learning or environmental stress, it is now proposed that distinct changes in the 3D arrangement of our genomes occur. These changes are coordinated through regulation of gene expression. Just as a map of the world is more than a mere list of places and street names, the nucleus cannot be reduced to a string of



OMX Microscopy R. Chaligne_T.Piolot_E.Heard (Institut Curie). Breast cancer cells with the XIST RNA in red and the chromatin marker H3K27me3 in green.

letters. The orchestrated organization of the genome impact all scales of an organism, from setting diverse cell identities, specific functions, cell cooperation, tissue organization, morphogenesis and eventually the maintenance of an entire body. **How the genome functions within the space of the nucleus and its dynamic context has remained an elusive challenge until recently with the advent of new technologies.**



Folding of the Human X chromosome
http://doi.org/10.1038/nature12593

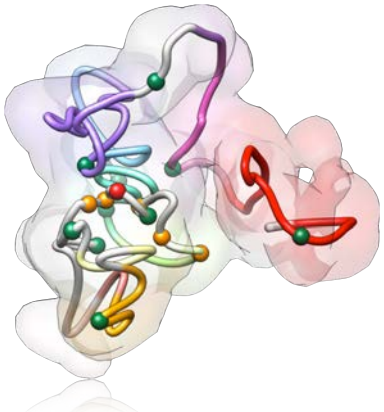
Aims. The 4DNucleome Initiative has as a main goal to be a driving force in the development and exploration of novel technologies that monitor and analyse the spatio-temporal changes in nuclear and chromosome organization. We will further describe the different contexts where changes occur; the forces and influences underlying such changes; the underlying dynamics of epigenomes and DNA-associated proteins that both scaffold genome organization and respond to these changes, and their consequences for genome function, including gene expression, DNA replication and repair, and genomic integrity. In other words, we aim at deciphering the structure-function relationships of the cell nucleus as a complex biological system at all levels from molecules to the entire genomic and epigenomic llandscapes. Thanks to the ongoing development of new technologies, we will be able to fully characterize the **sequence-structure-function relationships** of the genome and its link to the phenotypic manifestation of disease. Together, our community will specifically aim to:

- **Support community driven projects in technology development for studying the dynamic organization of the nucleus.**
- **Generate new paradigms, knowledge and specific large datasets to characterize the sequence-structure-function relationship of the genome in the dynamic nuclear environment.**
- **Translate discoveries/technologies to the public through promoting Public-Private Partnerships.**
- **Develop robust diagnostics of cell status based on nuclear and chromosomal organization.**
- **Strengthen and reinforce the lead held by the 4DNucleome community in Europe, in a competitive international environment.**
- **Link our community to other international initiatives in the US and Japan.**
- **Integrate our insights into other endeavors such as genome editing or medical genomics, in order to address the challenges in those fields.**

Game Changer. International efforts to identify human genetic variation linked to common disease (cancer, diabetes, autoimmunity, cardiovascular and neurological disease) have revealed that most of the genetic variants that affect human health lie far from coding sequences (that is, they affect regulatory elements such as enhancers). We now know that during embryonic development (at least in model organisms such as mouse), the spatial organization of chromosomes has a crucial role in ensuring cell type stability. Laminopathies, which disrupt key structural proteins of the nucleus, are the cause of diverse diseases ranging from muscular dystrophy to progeria (premature ageing). They are a prime example of tissue-specific genetic diseases that reflect changes in the spatial organization of the genome and its dynamics during tissue differentiation.

The time is ripe to exploit and manipulate the spatial organization of the genome to understand disease states in terms of the nuclear landscape. This will arguably leverage the wealth of data that many of the genome projects have collected over the last decade. For example, it has been estimated in several economic studies that the Human Genome Project outcome has returned to society between 10 and 150US$ for each US$ publicly invested between the years 1988 and 2003 (https://doi.org/10.1038/nature.2013.13187). This economical return, however, has been hampered by the limited impact it has had on human health concerns, patient outcomes and the discovery of new drugs or new diagnostics. Indeed, knowledge of the sequence of the human genome is not enough to make an impact in how medicine is practiced today. This will only be achievable by identifying the link between sequence and function to determine in full the nature of the information encoded in the genome. The spatial dynamics of the genome, arguably its most diverse and varied aspect, will help us more accurately develop new biomarkers, therapeutics, and diagnostics for improved medicine. **We envisage that the 4DNucleome Initiative will help us at last understand how the genome functions; it will change the way we edit genomes** (see some very recent exciting examples https://doi.org/10.1016/j.cell.2015.04.004 and https://doi.org/10.1126/science.aad9024)**, synthesise life** (http://doi.org/10.1038/531557a)**, treat patients with targeted genome editing** (https://doi.org/10.1038/nature.2015.18737)**, and help us move towards precision medicine in the near future.**

Science and technology



Folding of the human alpha-globin domain
http://dx.doi.org/10.1038/nsmb.1936

As many other emerging areas in science, ours is a technology driven field. The challenge in the 4DNucleome Initiative includes advanced technologies from four interdisciplinary fields at its core: cell biology, molecular genetics, imaging, and computational modelling. Recent technological advances in high resolution and live microscopy, high-throughput genomics/cell biology approaches and modelling, coupled with increased awareness of the importance of genome organization will soon allow to perform precision analysis of our genomic organization and its dynamic translations from one epigenome to another, as cells differentiate,

age, and respond to the environment. This a perfect time to launch a concerted effort towards characterizing the dynamic organization of the genome, the epigenome, and the rules that govern determination and maintenance of cell types in face of both internal and external stress linked to disease. We can now envisage having a complete 3D atlas in time (4D) of nuclei within the many cell types that form our body. **The huge challenge before us is to take the one dimensional genome sequence provided by the Human Genome Project, decorated with the valuable annotations provided by the ENCODE project, and create an integrated 4D understanding of the complexity of this incredible, living, breathing machine that holds the secret of life.**

The 4DNucleome Initiative is in line with – and will actively contribute to - the promotion of the EU's Digital Single Market Strategy. In particular, by the nature of the field and the technologies involved, the 4DNucleome Initiative can contribute significantly to the development of the European Open Science Cloud and data infrastructure. Therefore, the Initiative will prioritize engaging with the EU over the challenge of promoting interoperability and data sharing among disciplines and infrastructures. Broadly, the field of nucleomics develops and applies four main set of technologies:

Molecular Genomics.
The 4DNucleome community has over the years developed and implemented genomics technologies allowing an integrated investigation of gene expression, epigenetic marks and DNA localization, both for a population of cells and at the single cell level. These technologies rely on the use of high-throughput experimental approaches and next generation sequencing to characterize genome organization and chromatin status at the molecular level. Specifically, our community develops and applies technologies for RNA-Seq, Chromatin Immuno Precipitation (ChIP-Seq), Chromosome Conformation Capture (3C) and all the upcoming possibilities that can be anticipated in engineering combinations of these approaches.

Cell Biology.
The 4DNucleome community has provided many of the existing technologies from multiplex labeling of DNA, RNA and proteins to live cell imaging of chromosomes and the proteins they associate with has revealed some of the basic folding principles of the genome. Cell sorting, cell type identification and cell-cell crosstalk are crucial tools that will be developed in this domain. We should provide profound insights normal cell function, as well as disease. For example, single cell reprogramming requires a rewiring of the nucleus to express different genes, and to reverse the heritable repression of genes that occurs during differentiation. Single cell manipulations, induced pluripotency and tissue or organoid differentiation in vitro are just a few of the many cell-based technologies used in this field.

Modelling.
Classically, the dynamic nucleus has been theoretically studied through computational approaches using ICT technologies including storage, simulation, analysis, integration and dissemination or large datasets of experimental observations. The 4DNucleome Initiative will pay special attention to coordinate these efforts by connecting all available databases of experimental datasets as well as homogenizing the tools to study the 4D nucleome. To that end, it is essential to lead the promotion of the first international deposition database of 3D and 4D representations of the nucleus. Overall, the 4DNucleome Initiative will address such challenges by using integrative modelling approaches combining physics, mathematics, statistics, and computational science.

Imaging.
The recent advent of 3D high-throughput (deep-imaging) and super-resolution imaging (nanoscopy) technologies, and novel DNA labelling technologies allow us to visualize genomic domains in 3D, characterize the dynamics of several genomic loci per cell, and detect the simultaneous positions of multiple genomic loci in single-cells at high-throughput. Specifically, the 4DNucleome Initiative will advance and further evolve methods for deterministic super-resolution (such as STED, GSD, RESOLFT and SSIM) as well as stochastic super-resolution (such as SPDM, SPDMphymod, PALM, FPALM, STORM and dSTORM). Additionally, live imaging of the nucleus will be reveal spatio-temporal dynamics, which can be integrated with molecular information profiling of the same cells.

Importantly, all mentioned technologies will be further developed to seamlessly apply them both at the population-based (millions of cells) and single-cell-based scenarios. Moreover, those technologies will be applied to study the fourth dimension of time either by "snap-shooting" the experiments or by real-time analysis of living cells.

Large-scale collaboration

European scientists play a leading role in this field (9,300 of 22,000 papers published in the field of nuclear organization over the last 5 years are from Europe, about 20% more than those from USA and Canada). European scientists have discovered many of the basic principles underpinning our current understanding of spatial genome organization and dynamics, but Europe's status will be lost without clear community building and collective goal-setting in this field. The interest and expertise in the 4D nucleomics field is clearly represented by the supporting members of this Initiative. Early-stage support has already been secured through specific grants in the H2020 program.

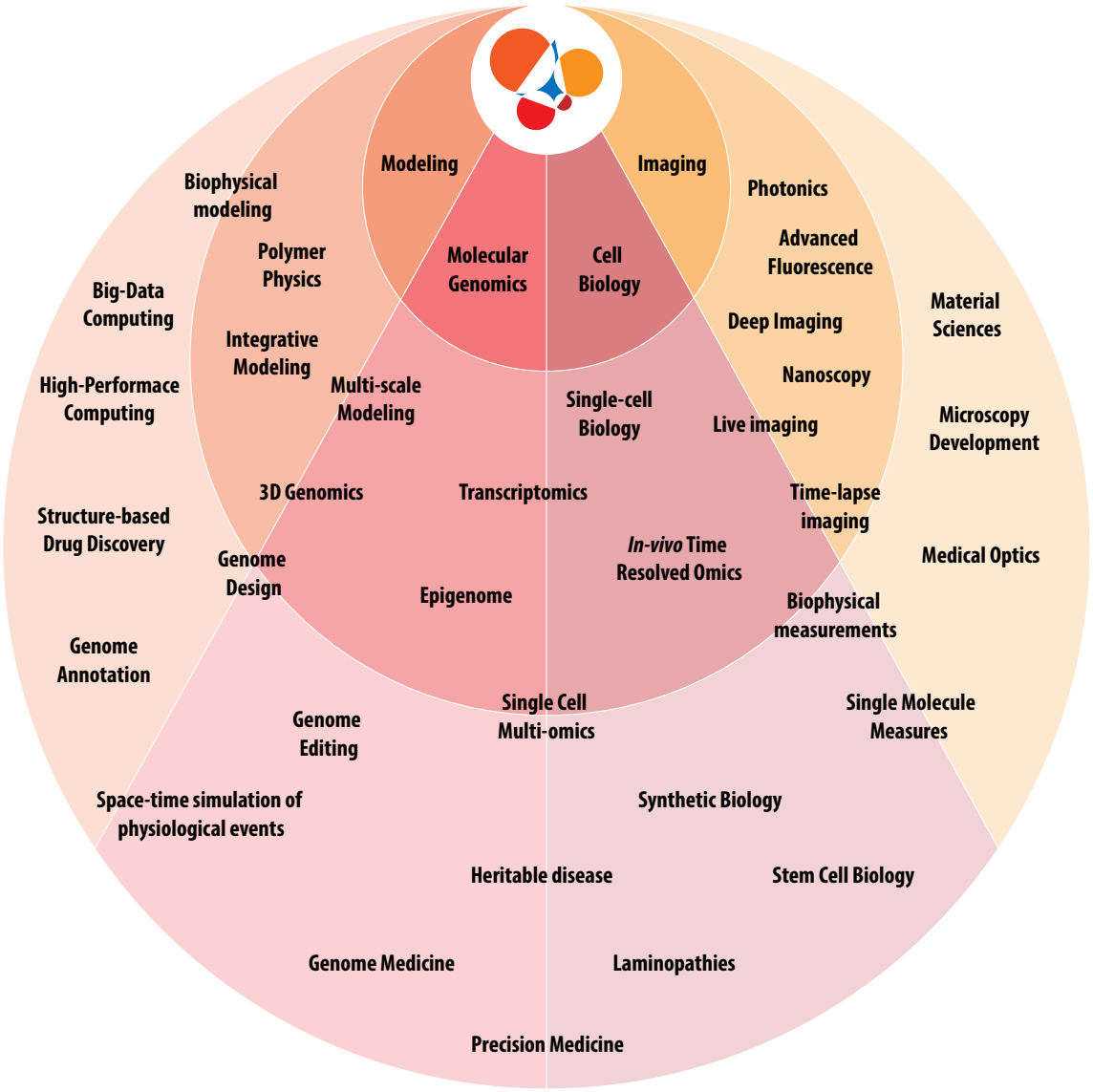
The 4DNucleome Initiative (http://www.4DNucleome.eu) is being supported by a wide research community in Europe (updated support: http://www.4dnucleome.eu/supporters/). This community is also supported by the UE-LIFE (http://eu-life.eu), an alliance of thirteen leading European research institutes in life sciences. We also collaborate with other sister initiatives beyond the UE, including the US (NIH 4DNucleome Initiative, https://commonfund.nih.gov/4dnucleome/index) and the Japan (Hiroshima Initiative, https://doi.org/10.1080/19491034.2015.1022703).

Some of the supporting members attending the final meeting of the EpiGeneSys NoE in Paris.

Several of the EU researchers supporting this Initiative have collaborated through the very successful NoE EpiGeneSys (http://www.epigenesys.eu) and are connected to these two international efforts through collaborations with the research laboratories or even as partners on funded grants. Indeed, the NIH has seen the potential that lies in 4D nucleomics and is currently funding laboratories located outside the US. Here we aim to exploit the collaborative nature of European research (i.e., its capacity to assemble large collaborative research consortia) as this is an aspect in which European science excels. Unfortunately, there is currently no European network through which researchers in the 4D nucleome community can interact to develop technologies and share resources in order to develop new approaches and exploit existing know-how through industrial contacts. The concerted strategy a 4DNucleome Initiative will let Europe take the lead in this challenging endeavor. **We foresee the establishment of a truly international network of excellent scientists linked by a well-defined goal, that of deciphering the organization of the nucleus and its changes over time. Our coordinated efforts will set the standards for this expanding and highly interdisciplinary field of biomedical research.**

Multidisciplinary exploration

Many recent discoveries in Europe have just begun to shed light on the 4D organization of the genome and its relevance to disease. However, there is currently no tool in Europe with which to exploit the potential that these discoveries have opened up for health care and industry. This is not the case in the US for example. We foresee that these discoveries as well as future ones will have a direct impact in many fields including diagnostics and therapeutics by allowing: (i) structure-based engineering of genomes, (ii) the characterization of the effects of new drugs targeting of chromatin remodelers, (iii) the identification of new targets for drug discovery, (iv) the development of new biomarkers of disease, (v) the design of new forms of life in synthetic biology, etc. **All these aspects have direct paths towards innovative biotechnological and medical industries as well as a broader impact on society due to the new ways of performing genome-informed medical treatments or the use of engineered genomes/cells in regenerative medicine.**



Map of impacting fields around the 4DNucleome technology fields.

The 4DNucleome Initiative in Europe proposes a concerted effort (to be carried out in close collaboration with other sister international initiatives in Japan and the US) that will generate means to understand the spatial and dynamic organization of the genome within the nucleus. This effort will require the complementary skills of clinicians, geneticists, biochemists, molecular biologists, biophysicists, microscopists, bioinformaticians and mathematicians. The collective development of technologies, associated analytic tools and the knowledge they generate, along with their application to medical problems, will have a crucial impact on medical diagnosis, genome editing, and on our ability to interpret epigenetic effects on gene expression and phenotypic penetrance. **Altogether, our efforts will place Europe at the forefront of the emerging field of 4D Nucleomics that promises to bring new scientific paradigms to the understanding of the human genome and the many cell types and tissues it encodes.**

http://www.4DNucleome.eu