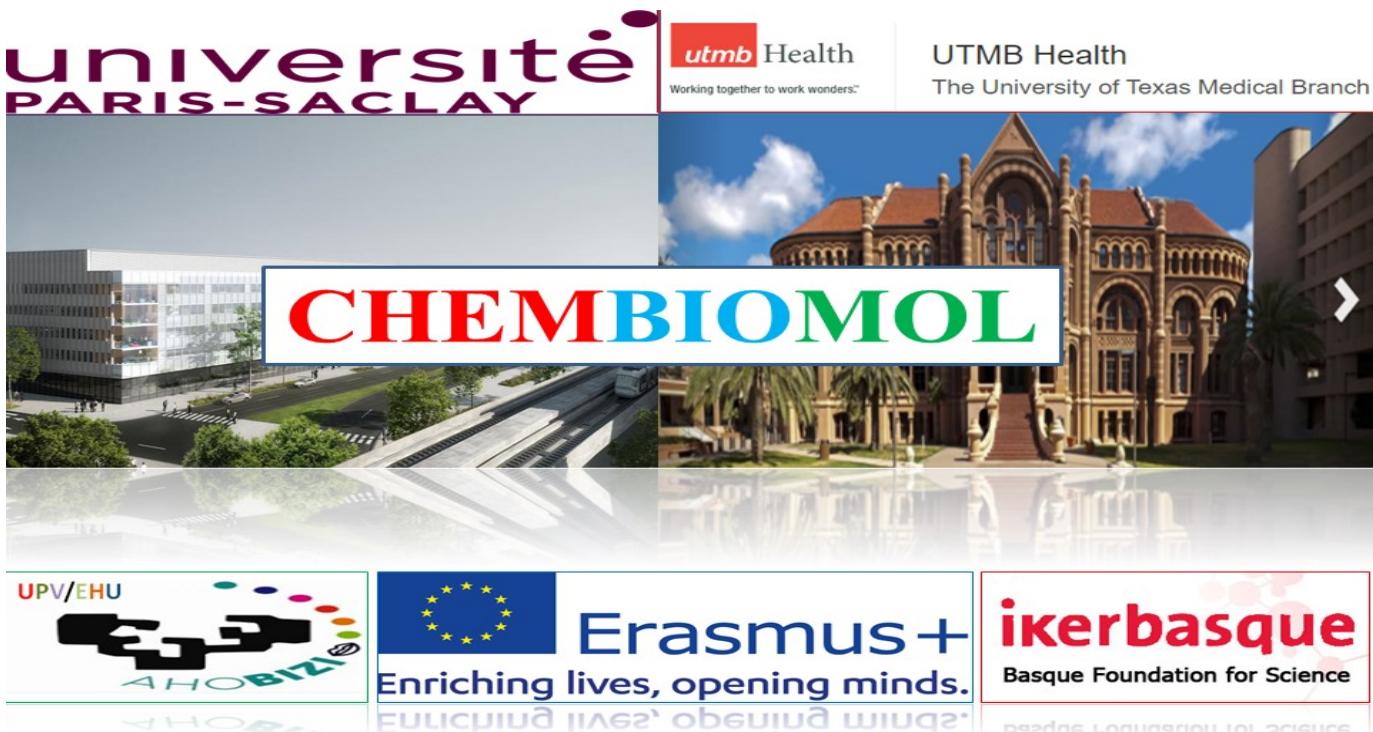




CHEMBIOMOL-07: Chem. Biol., Med.
Chem., & Mol. Biol. Congress, Paris,
France-Galveston, USA, 2021



NAIF.PTML Approach to Artificial Intelligence (AI) Driven Chromosomics in Synthetic Biology

Viviana F. Quevedo-Tumailli^{a, b} and Humberto González-Díaz ^{c,d,e,*}

^a RNASA-IMEDIR, Department of Computer Science, Faculty of Informatics,
University of A Coruña (UDC), 15071, A Coruña, Spain.

^bUniversidad Estatal Amazónica UEA, Puyo, Pastaza, Ecuador.

^c Department of Organic and Inorganic Chemistry,
University of Basque Country UPV/EHU, 48940 Leioa, Spain.

^d BIOFISIKA: Basque Center for Biophysics,

University of Basque Country UPV/EHU, 48940 Leioa, Spain.

^e IKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Biscay, Spain.

Doi: 10.3390/mol2net-07-11821

Abstract. A new term called Chromosomics appeared, probably for the first time, in the research conducted by Deakin *et al.* in the paper entitled Chromosomics: Bridging the Gap between Genomes and Chromosomes. Chromosomics emerges from the necessity of going beyond the information about the sequences of DNAs (gene), RNAs, and Proteins (Genomics, Rnomics, Proteomics). Chromosomics focus them on the study of the formation and structure of chromosomes and the effect of this structure over biological process. Consequently, it also focus on the mechanisms, kinetics, and dynamics, of DNA folding, DNA-interaction with Histones, the assembly, disposition, distribution, organization, and packing of all the genome in chromosomes and the ultimate effect of this processes over the outcomes of biological process. It also studies the relative orientation of gene within the chromosome, the interaction between them inside the Chromosome of inter-chromosomes, the interaction between chromosomes in the cell. One of the ultimate goals of Chromosomics is the study of the relationship of all this information related to chromosomes with Genomics, Rnomics, and Proteomics information and their final contribution to the outcome of biological process. Chromosomics should answer questions like in which degree a determined Genomics, Rnomics, and Proteomics information pre-determines and/or is determined by chromosomes structure.

Very importantly, Chromosomics have work to do with the effect of changes of building blocks of chromosomes on their structure/function relationships and the ultimate effect on pathogenesis. It could help to discover new biomarkers of genetic/chromosomal diseases. This approach can be used also in Drug Discovery to determine the effect of new drugs over known pathogen eukaryote microorganisms like plasmodium, leishmania, etc. given the available information about their Chromosomics, Genomics, Proteomics, and pre-clinical assays of drugs. Chromosomics experts may play an important role in the same sense but in inverse direction. We refer to determining the effect over biological process of intentional changes of building blocks of chromosomes (gene, histones). On our opinion different kind of changes and different levels of the process may be carry out. For instance, at a first level, we can change the chemical structure of the building blocks by replacing natural nitrogen bases of DNA or natural aminoacids of histones by other natural or artificial elements (other nitrogen bases, aminoacids, or other building blocks). At a second level we can change the orientation of genes inside the chromosome to see the effect of these changes. At a third level we can chemically modified the 3D structure of histones, etc. All these kind of actions may led us to enter in the field of Synthetic biology with the aim of designing for instance new microorganisms (bacteria, yeast, etc.) with optimal capabilities for producing/metabolizing different substances. This in turn may be of importance for Biotechnology and Environmental biotechnology. However, important bioethics and regulatory affairs consideration emerge of this field that has to be taken into consideration carefully.

Experimental Chromosomics may use a direct approach for determination of all this information. It starts by the *ab initio* experimental approaching to the problem. This begins by the direct study of the building blocks of chromosomes (nucleotides, nucleic acids, gene, histone proteins). Next, it can study the determination of mechanism of the assembly of these blocks to determine the structure and position of chromosomes. This implies the study of the interaction of genes with histones to form structural units called nucleosomes and in turn the packing chromatin or chromosomes. In so doing, it may use biophysics techniques like Electronic Microscopy (EM), CryoEM, X-Ray, Nuclear Magnetic Resonance (NMR), etc.

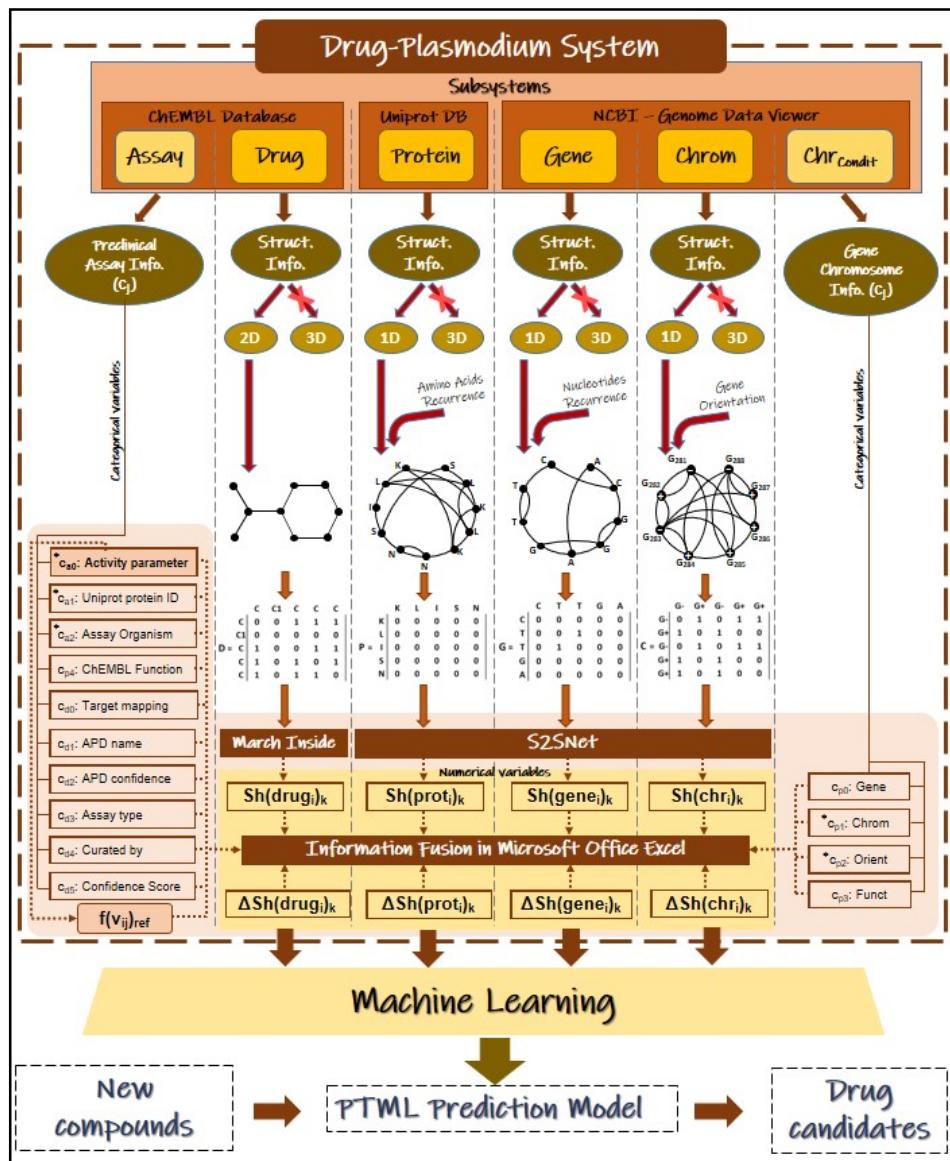


Figure 1. NIFPTML Chromosomics models.

It is also interesting how the use of Big Data analysis tools could lead us to new discoveries in this area. However, the amount of information to be taking into consideration here is vast and comes from many diverse sources. Consequently, on our opinion, Artificial Intelligence and Machine Learning (AI/ML) algorithms are called to play an important role here. In particular the algorithms of the family Artificial Intelligence and Network Information Fusion with Perturbation Theory Machine Learning (NAIF.PTML = N + IF + PT + AI/ML) developed by us may play an interesting role. They has an N-

stage devoted to represent as complex networks and molecular graph and encode numerically all the information related to the structure of DNA, RNA, Proteins, and Chromosomes. Next we can use the IF-stage to curate and fusion all information about the previous biological assays reported for the properties and behaviors of these networks. Next, in the AI/ML phase we can train/validate predictive models useful to predict the desired outcomes of the new systems.

References

1. Claussen, U. Chromosomics. *Cytogenet. Genome Res.* 2005, 111, 101-106.
2. Deakin, J.; Potter, S.; O'Neill, R.; Ruiz-Herrera, A.; Cioffi, Eldridge, M.; Fukui, K.; Marshall, J.; Griffin, D.; Grutzner, F.; Kratochvil, L.; Miura, I.; Rovatsos, M.; Srikulnath, K.; Wapstra, E.; Ezaz, T.; Chromosomics: Bridging the Gap between Genomes and Chromosomes. MDPI.
3. Traut, W.; Vogel, H.; Glöckner, G.; Hartmann, E.; Heckel, D.G. High-throughput sequencing of a single chromosome: A moth W chromosome. *Chromosom. Res.* 2013, 21, 491–505.
4. Skinner, B.M.; Sargent, C.A.; Churcher, C.; Hunt, T.; Herrero, J.; Loveland, J.E.; Dunn, M.; Louzada, S.; Fu, B.; Chow, W.; et al. The pig X and Y Chromosomes: Structure, sequence, and evolution. *Genome Res.* 2016, 26, 130–139.
5. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* 2015, 13, 278–289.
6. Gu, L.; Walters, J.R. Evolution of sex chromosome dosage compensation in animals: A beautiful theory, undermined by facts and bedeviled by details. *Genome Biol. Evol.* 2017, 9, 2461–2476.
7. Deakin, J.E.; Hore, T.A.; Koina, E.; Graves, J.A.M. The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet.* 2008, 4, e1000140.
8. Livernois, A.M.; Waters, S.A.; Deakin, J.E.; Graves, J.A.M.; Waters, P.D. Independent evolution of transcriptional inactivation on sex chromosomes in birds and mammals. *PLoS Genet.* 2013, 9, e1003635.
9. Quevedo-Tumailli, V. F.; Ortega-Tenezaca, B.; Gonzalez-Diaz, H., Chromosome Gene OrientationInversion Networks (GOINs) of Plasmodium Proteome. *J Proteome Res* 2018, 17, 1258-1268.
10. Quevedo-Tumailli V, Ortega-Tenezaca B, González-Díaz H. IFPTML Mapping of Drug Graphs with Protein and Chromosome Structural Networks vs. Pre-Clinical Assay Information for Discovery of Antimalarial Compounds. *Int J Mol Sci.* 2021 Dec 2;22(23):13066.