OMB No. 0925-0001 and 0925-0002 (Rev. 09/17 Approved Through 03/31/2020)

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Marinov, Georgi Kolev

eRA COMMONS USER NAME (credential, e.g., agency login): MARINOV.GEORGI

POSITION TITLE: Postdoctoral Fellow

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Massachusetts Institute of Technology, Cambridge MA	B.S.	06/2008	Biology
California Institute of Technology, Pasadena, CA	Ph.D	06/2014	Biology
Indiana University, Bloomington, IN	Postdoctoral	07/2016	Biology
Stanford University, Stanford CA	Postdoctoral	05/2017-present	Genetics

A. Personal Statement

My interests focus on understanding how the regulation of gene expression is accomplished in time and space through the combined action of transcription factor binding, chromatin regulators, and chromatin folding. I am also particularly interested in using comparative and functional genomics to understand the deeper logic behind these mechanisms and how they developed in the course of evolution of cellular life. The progress of sequencing technology over time has been key to my professional development as it has allowed for ever more powerful techniques to be developed and applied to these and many other questions. As part of the ENCODE Project Consortium I have played key role in developing and establishing protocols and best practices for carrying out and analyzing methods for mapping and quantifying protein-DNA interactions (ChIP-seq) and transcript levels at the bulk and single-cell level (RNA-seq and scRNA-seq), which I have also successfully applied to a wide array of other biological contexts. Most recently, as part of the Greenleaf lab at Stanford, I have developed novel methods for profiling chromatin accessibility using long-read single molecule sequencing, an exciting new area that I am continuing to actively explore. The other main focus of my research interests is charting and understanding the deep evolution of chromatin and gene regulatory mechanisms across the tree of life, questions that I aim to answer using a combination of comparative and functional genomic approaches.

B. Positions and Honors

10/1/2008 – 12/15/2014	Graduate training in the laboratory of Barbara J. Wold, Division of Biology, California Institute of Technology, Pasadena, CA
1/13/2015 – 8/01/2016	Postdoctoral scholar in the laboratory of Michael Lynch, Department of Biology, Indiana University, Bloomington, IN
05/01/2017 – present	Postdoctoral fellow in the laboratories of William J. Greenleaf and Anshul Kundaje, Department of Genetics, Stanford University School of Medicine, CA

Honors

2018	Stanford School of Medicine Dean's Fellowship
2004	Bulgarian National Fellowship for Excellence in Studies
2004	Silver medal at the International Biology Olympiad
2003	Bronze medal at the International Biology Olympiad
2001-2004	First Prize at the National Biology Olympiad of Bulgaria

C. Contributions to Science

1. ***Functional genomic methods for mapping protein-DNA interactions:*** The genome-wide identification of the binding sites of transcription factors and the distribution of histone marks is key to understanding the functioning of gene regulatory networks. During my graduate studies I developed protocols and standards for carrying out ChIP-seq experiments as part of the ENCODE Consortium Project. This work helped establish the stable experimental protocols and the standard ENCODE ChIP-seq processing pipelines that have been used by the consortium ever since.

a. Landt SG*, **Marinov GK***, Kundaje A*, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, Desalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M. 2012. ChIP-seq guidelines and practices used by the ENCODE and modENCODE consortia. Genome Res 22(9):1813–1831.

b. **Marinov GK**, Kundaje A, Park PJ, Wold BJ. 2014. Large-scale quality analysis of published ChIP-seq data. G3 4(2):209–223.

c. **Marinov GK**. 2017. Identification of Candidate Functional Elements in the Genome from ChIP-seq Data. Methods in Mol Biol 1543:19–45.

d. **Marinov GK**. 2017. ChIP-seq for the Identification of Functional Elements in the Human Genome. Methods in Mol Biol 1543:3–18.

2. ***Single-cell studies of the mammalian transcriptome:*** The ability to profile the transcriptome of thousands of individual cells (scRNA-seq) has transformed multiple areas of biology in the last few years, although they are still plagued by issues having to do with detection sparsity and technical noise. As part of my graduate studies I carried out some of the earliest studies in the area, characterizing these issues in depth and developing statistical frameworks for addressing them. Using scRNA-seq I showed that individual cells frequently express numerous genes in monoallelic fashion, and that there is significant variation in splice isoform production at the single cell level.

a. **Marinov GK***, Williams BA*, McCue K, Schroth GP, Gertz J, Myers RM, Wold BJ. 2014. From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. Genome Res 24:496–510..

b. Kim DH, **Marinov GK**, Pepke S, Singer ZS, He P, Williams BA, Schroth GP, Elowitz MB, Wold BJ. 2015. Single-Cell Transcriptome Analysis Reveals Dynamic Changes in lncRNA Expression during Reprogramming. Cell Stem Cell 16(1):88–101

3. ***Functional genomics of organelles:*** Eukaryotic mitochondria and plastids originate from prokaryotic ancestors that became endosymbionts and were subsequently reduced to their present state. An incredible diversity of highly derived patterns of genome organization and mechanisms of gene expression and regulation are observed in organellar genomes. I pioneered the use of functional genomic techniques to study these aspects of organellar biology, directly showing that the TFAM protein serves as a general coating factor in mammalian mitochondria. I also discovered and mapped multiple nuclear transcription factors that associate with mitochondrial DNA in metazoans.

a. Wang YE, **Marinov GK**, Wold BJ, Chan DC. 2013. Genome-wide analysis reveals coating of the mitochondrial genome by TFAM. PLoS ONE 8(8):e74513.

b. **Marinov GK***, Wang YE*, Chan DC, Wold BJ. 2014. Evidence for site-specific occupancy of the mitochondrial genome by nuclear transcription factors. PLoS ONE 9(1):e84713.

4. ***Gene regulation in divergent eukaryotic lineages*****:** A number of eukaryotic lineages exhibit radical deviations from the conventional rules of chromatin organization and gene regulation known from research in metazoans and yeast. The dinoflagellates are the most extreme such case, having, unique among all eukaryotes, lost nucleosomes as a major constituent of chromatin. Using comparative genomics and transcriptomics, I showed that although dinoflagellates have largely replaced nucleosomes as a major packaging factor in their genomes, all species for which data is available possess full and diverse complements of histone genes, though generally highly divergent from those of other eukaryotes.

Analysis of the sequences of these histone proteins and of the presence or absence of other chromatin regulators pinpointed likely conserved and divergent aspects of the histone code and the transcriptional cycle in these organisms. I also carried out similar analyses in nucleomorphs, the most extremely reduced eukaryote genomes, where I also identified major deviations from the conventional histone code.

α. **Marinov GK**, Lynch M. 2015. Diversity and Divergence of Dinofagellate Histone Proteins. G3 6(2):397–422

β. **Marinov GK**, Lynch M. 2016. Conservation and divergence of the histone code in nucleomorphs. Biol Direct 11:18.

5. ***Using massively parallel CRISPR screening to characterize the non-coding genome.*** Large-scale efforts such as the ENCODE Consortium are identifying hundreds of thousands of candidate regulatory elements in mammalian genomes. However, these elements remain to be functionally characterized in detail. To this end, I have been using high-throughput CRISPR screening techniques as part of the latest phase of ENCODE to directly test the function of candidate regulatory elements. Our published efforts in this area have addressed important technical issues having to do with off-target effects in non-coding CRISPR screens, and have shown that, contrary to conventional wisdom, insulator elements marked by association with the CTCF protein are not of vital functional importance, at least in the context of cultured cell lines.

   a. Tycko J*, Wainberg M*, **Marinov GK***, Ursu O, Hess GT, Ego BK, Aradhana, Li A, Truong A, Trevino AE, Spees K, Yao D, Kaplow IM, Greenside PG, Morgens DW, Phanstiel DH, Snyder MP, Bintu L, Greenleaf WJ, Kundaje A, Bassik MC. 2019. Identification and mitigation of pervasive offtarget activity in CRISPR-Cas9 screens for essential non-coding elements. bioRxiv 520569

6. ***Using long-read single-molecule sequencing to map chromatin organization.*** While short-read sequencing has revolutionized genomics since the mid-2000s, it presents a number of limitations that do not allow important biological questions to be addressed. As an example, it has not been possible to directly enumerate the distribution of chromatin states with high resolution and on multikilobase scales, nor have we been able to quantify the extent of dependence/coordination between distal regulatory elements. I have developed a novel method (SMAC-seq) that enables studies of this kind, based on the ability of nanopore sequencing to generate very long reads and to read out DNA modifications directly. I am continuing to actively improve and develop new tools for mapping chromatin organization based on emerging sequencing technologies.

   a. Shipony Z*, **Marinov GK***, Swaffer MP, Sinott–Armstrong NA, Skotheim JM, Kundaje A, Greenleaf WJ. 2018. Long-range single-molecule mapping of chromatin accessibility in eukaryotes. bioRxiv 504662.

(*These authors contributed equally)

Complete List of Published Work: https://scholar.google.com/citations?user=ijL6kD4AAAAJ&hl=en