**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: Joseph Mauro Calabrese

eRA COMMONS USER NAME: jmcalabr

POSITION TITLE: Associate Professor of Pharmacology

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE | Completion Date | FIELD OF STUDY |
|---|---|---|---|
| University of Wisconsin-Madison | B.S. | 12/01 | Chemistry/Biochemistry |
| Massachusetts Institute of Technology | Ph.D. | 02/08 | Molecular Biology |
| University of North Carolina at Chapel Hill | Postdoc | 02/14 | Genetics and Genomics |

**A. Personal Statement**

**1. Research:** The long-term goals of my laboratory are to determine the molecular mechanisms through which long noncoding RNAs (lncRNAs) regulate gene expression, and to develop new methods to modulate lncRNA function for therapeutic gain. I have been studying RNA-mediated gene regulation since 2003 and epigenetic regulation governed by *Xist* and related lncRNAs since 2008. My faculty appointment began in March of 2014. Broadly, research in my lab focuses on understanding how intrinsic sequence elements within lncRNAs and the trans-acting factors that bind them cooperate to induce specific epigenetic states at targeted regions within mammalian genomes. My lab employs state-of-the-art genomic, cell biological, biochemical, and computational approaches to achieve our research goals. Our main expertise is in developing and applying novel genomic approaches to discover fundamental principles in gene regulation.

**2. Training and mentorship:** My major goals as a mentor are (1) to teach students and fellows how to perform high-impact, basic research that brings to light new, fundamental principles of biological regulation, and (2) to prepare them to become leaders in their chosen field after leaving my laboratory. In order to improve my own skills as a mentor, I participated in an 8-part “Becoming a better mentor” workshop run by the Center for Faculty Excellence. My laboratory is small (~6 employees) and I am able to have scientific discussions with most of my students daily. Beyond providing guidance in their immediate research needs, my overarching goal in these interactions is to teach trainees how to design rigorous and unbiased experiments, decide upon appropriate methodology, analysis, and interpretation, and how to report results to the public and to scientific peers. A primary goal is to help my trainees obtain their Ph.Ds in a timely fashion with the skills, credentials and experiences they need to transition into careers in the biomedical workforce. We have weekly meetings to discuss progress towards these goals. I engage them in one-on-one discussions about career trajectories and talk through their long-term plans annually. I require my students to present in public forums both within UNC and at locally and (inter)nationally attended meetings. When students have asked it of me, I have encouraged them to obtain training in non-academic scientific positions, such as internships at Google Life Sciences and the Biotechnology Transfer Office at UNC. My first four students each had at least one first-author original research article accepted or submitted within their four years in my lab. They each graduated in less than 6 years of starting their PhD (<5 years in the lab). I am also heavily engaged in graduate education outside of mentoring the trainees in my lab. Foremost, I am the director of graduate studies for the Pharmacology T32 Training Grant (GM135095). In this capacity, I oversee the development and execution of core coursework and facilitate the transition through to Ph.D. for our 40+ students. I have served on the graduate admissions committee; I teach several seminar courses in genetics and epigenetics; I have hosted external speakers for the BCB, GMB, and PHCO student

seminar series; I serve on >20 thesis committees; I designed and teach a month-long module called “Practical RNA-Seq” that teaches the essentials of one entire RNA-seq experiment, from molecular biology to data analysis –the last three years this class has had an average enrollment of ~20 students across all of the biomedical curricula as well as an average of ~8 auditors which include students, technicians, and professors.

---

## B. Positions and Honors

| | |
|---|---|
| 1997 | Order Sons of Italy Scholarship |
| 1998 | Undergraduate Researcher, University of Wisconsin-Madison |
| 2000 | Elvehjem Scholarship for Excellence in Biochemistry |
| 2001 | Phi Beta Kappa |
| 2001 | University of Wisconsin-Madison, Comprehensive Honors |
| 2002 | Tutor, University of Wisconsin-Madison |
| 2002 | Graduate Student, Phillip Sharp lab, MIT |
| 2007 | RNA Society Meeting abstract-selected speaker |
| 2008 | Post-doctoral Fellow, Terry Magnuson lab, UNC Chapel Hill |
| 2008 | UNC Lineberger Comprehensive Cancer Center Postdoctoral Training Award |
| 2008 | UNC UCRF Pilot Project Award |
| 2009 | American Cancer Society Post-doctoral fellowship |
| 2012 | Keystone Symposia on Noncoding RNA abstract-selected speaker |
| 2012 | Keystone Symposia on Noncoding RNA Scholarship recipient |
| 2013 | Keystone Symposia on Noncoding RNA abstract-selected speaker |
| 2013 | UNC Chapel Hill Postdoctoral Research Excellence Award |
| 2014 | Assistant Professor of Pharmacology, UNC Chapel Hill |
| 2014 | UNC Translational and Clinical Sciences Institute Faculty Development Award |
| 2015 | March of Dimes Basil O’Connor Scholar |
| 2016 | CSHL Meeting on Nuclear Organization and Function abstract-selected speaker |
| 2020 | Keystone Symposia on Noncoding RNA invited speaker |

---

## C. Contributions to Science (chronological)

### ***1) Discovering principles of the transcriptional process through genomic approaches*.**

The advent of next-generation sequencing in 2005 allowed small RNA populations in cells to be profiled at an unprecedented level. My graduate advisor Phil Sharp was an early advocate of the technology and encouraged post-doctoral fellow AC Seila and myself to undertake a collaborative study of the small RNA populations in wild-type and *Dicer* knockout cells. In addition to findings described in [**a**], Dr. Seila and I made the unexpected observation that small RNAs that were not *Dicer* dependent were produced in both the sense and anti-sense directions relative to transcription start sites of most expressed genes. Our findings, as well as those co-published by the John Lis and Torben Jensen labs, indicated that transcription initiates divergently at most mammalian promoters ([**b-d**]). These works fundamentally altered our understanding of the mammalian transcriptional process. Divergent transcription is thought to maintain promoters in a state poised for regulation and evidence suggests that divergent initiation can lead to the evolution of new genes.

**a) Calabrese, J.M.***, Seila, A.C.*, Yeo, G.W., and Sharp, P.A. 2007. RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 104(46): 18097-18102. PMC2084302.

**b)** Seila, A.C.*, **Calabrese, J.M.***, Levine, S.S., Yeo, G.W. Rahl, P.B, Young, R.A., and Sharp, P.A. 2008. Divergent transcription from active promoters. *Science* 322:1849-1851. PMC2692996.

**c)** Core, L.J., Waterfall, J.J, Lis, J.T. 2008. Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters. *Science*. 322: 1845-1848. PMC2833333.

**d)** Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., Jensen, T.H. 2008. RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters. *Science*. 322:1851-1854.

## ***2) Discovering principles of lncRNA/genome interactions.***

In 2008 I joined Terry Magnuson's lab as a post-doctoral fellow to study the molecular mechanisms through which the *Xist* lncRNA silences gene expression. Despite its discovery in 1991, the mechanisms by which *Xist* interfaces with target genes on the X chromosome to induce silencing remain poorly defined. To address this gap in knowledge, I developed an experimental platform and analytical approach to determine the effects that *Xist* and other monoallelically expressed lncRNAs exert on chromatin. I made several unexpected observations, the most notable being that regulatory elements along the inactive X harbor an epigenetic signature suggestive of an active role in targeting *Xist* to specific regions of the chromosome [**a**]. We recently extended findings in [**a**] to discover that lncRNA abundance can play a major role in dictating lncRNA repressive potency, and that subsets of DNA regulatory elements are hypersensitized to silencing by lncRNAs [**b**]. Understanding the physical nature of this hypersensitivity and the mechanisms that stabilize or degrade lncRNAs in cells may reveal ways to pharmacologically disrupt or augment lncRNA-meditated silencing in inherited genetic disorders and cancers. Approaches that we developed in [**a, b**] have enabled many additional discoveries ([**c**,**d**], others not listed).

**a) Calabrese, J.M**., Sun, W. Song, L., Mugford, J.W., Williams, L., Yee, D., Starmer, J., Mieczkowski, P., Crawford, G.E., Magnuson, T. 2012. Site-specific silencing of regulatory elements as a mechanism of X-inactivation. *Cell* 151(5): 951-63. PMC3511858.

**b)** Schertzer M.D., Starmer J., Braceros K.C.A., Lee D.M., Salazar G., Justice M., Bischoff S.R., Cowley D.O., Ariel P., Dowen J.M., Zylka M.J., Magnuson T., **Calabrese JM**. lncRNA-induced spread of Polycomb controlled by genome architecture, RNA abundance, and CpG island DNA. *Molecular Cell*. 2019 Aug 8;75(3):523-37. PMID: 31256989. PMC6688959. Featured on the cover of the August 8th 2019 issue.

**c)** Schertzer M.D., Thulson E., Braceros K.C.A., Lee D.M., Hinkle E.R., Murphy R.M., Kim S.O., Vitucci E.C.M., **Calabrese J.M**. A piggyBac-based toolkit for inducible genome editing. 2019. *RNA*. Aug;25(8):1047-58. PMID: 31101683. PMCID in progress.

**d)** Chu C., Zhang Q.C., da Rocha S.T., Flynn R.A., Bharadwaj M., **Calabrese J.M.**, Magnuson T., Heard E., Chang H.Y. 2015. Systematic discovery of *Xist* RNA binding proteins. *Cell*. 161(2):404-16. PMC4425988.

## ***3) Determining relationships between structure and function in regulatory RNAs.***

At 18 kilobases in length, *Xist* is one of the longest regulatory RNAs known. In 2014, when I started my lab, very little was known about the internal structure of *Xist*. To address this gap in knowledge, I initiated a collaboration with UNC Professor Kevin Weeks to use a method developed in his lab called SHAPE-MaP to probe the structural properties of *Xist*. Together, we developed and optimized methods to use SHAPE chemistry to detect RNA-protein interactions in living cells [**a**]. We employed this approach to report the first nucleotide-resolution structural map of authentic, full-length *Xist* [**b**]. We are now interrogating this map to understand how RNA structure confers function to *Xist*, which should provide paradigms relevant to many functional RNAs. Our work in [**a,b**] also established some of the first methods to map RNA structure and RNA/protein interactions in living cells.

**a)** Smola M.J., **Calabrese J.M**., Weeks K.M. 2015. Detection of RNA-Protein Interactions in Living Cells with SHAPE. *Biochemistry*. 2015 54(46):6867-75. PMC1664724.

**b)** Smola, M.J., Christy T.W., Inoue K., Nicholson C., Friedersdorf M., Keene J., Lee D.M., **Calabrese J.M.** [ca], and Weeks K.M. [ca]. 2016. SHAPE reveals transcript-wide interactions, complex structural

domains, and principles of protein interaction across the Xist lncRNA in living cells. Proc Natl Acad Sci U S A. 113(37):10322-7. PMC5027438.

***4) Developing computational methods to identify regulatory function in RNA.***

The functions of most lncRNAs are unknown. In contrast to proteins, lncRNAs with similar function often lack linear sequence homology; thus, identification of function in one lncRNA rarely informs identification of function in others. We developed a sequence comparison method to deconstruct linear sequence relationships in lncRNAs and evaluate similarity based on abundance of short motifs called k-mers. We found that lncRNAs of related function often had similar k-mer profiles despite lacking linear homology, and that k-mer profiles correlated with protein binding and lncRNA subcellular localization. We proposed, and then directly demonstrated, that evolutionarily unrelated lncRNAs can encode similar function through different spatial arrangements of functionally analogous protein-binding domains [**a**, **b**]. Kmer-based classification is a powerful approach to detect recurrent relationships between sequence and function in many types of lncRNAs. One could liken it to finally being able to begin to understand the different scripts in the Rosetta Stone. We are now investigating ways in which k-mer content relates to structure and function in diverse forms of RNA, including in pre-mRNAs. We are also working to incorporate markov-chain statistics into our k-mer based approach, which should simplify the implementation of k-mer based homology searches in the biomedical community.

**a)** Kirk J.M., Kim S.O., Inoue K., Smola M.J., Lee D.M., Schertzer M.D., Wooten J.S., Baker A.R., Sprague D., Collins D.W., Horning C.R., Wang S., Chen Q., Weeks K.M., Mucha P.J., and **Calabrese J.M.** 2018. Functional classification of long non-coding RNAs by k-mer content. *Nature Genetics*. 50(10):1474-1482. PMID: 30224646. PMC6262761.

**b)** Sprague D., Waters S.A., Kirk J.M., Wang J.R., Samollow P.B., Waters P.D., **Calabrese J.M**. Non-linear sequence similarity between *Xist* and *Rsx* suggests shared functions of tandem repeat domains. 2019. *RNA*. E-pub ahead of print. PMID: 31097619. PMCID in progress.

**c)** Kirk J.M., Sprague D., **Calabrese J.M**. Classification of long non-coding RNAs by kmer content. 2019. *Methods in Molecular Biology*. In press.

***5) Mechanisms and cooperativity of functional modules in regulatory RNAs.***

Many of the proteins required for the function of *Xist* and related lncRNAs are RNA-binding proteins that have been implicated in a wide array of biological processes and bind tens of thousands of positions in thousands of different RNAs. How these ubiquitously expressed proteins are leveraged by distinct RNAs to orchestrate specific regulatory functions remains unclear. We recently completed a study of a domain in *Xist* called Repeat-A. For the last 20 years, Repeat-A has been considered to be the core domain within *Xist* that causes the induction of transcriptional silencing. Our findings indicate that much of what has been assumed to be true about the function of Repeat-A in *Xist* may need to be reconsidered. Specifically, we discovered that when expressed as a stand-alone transgene, Repeat-A is unexpectedly unable to induce transcriptional silencing. Yet, in this same context Repeat-A does retain an unexpectedly potent transcriptional antiterminator activity ([**a**] and in preparation). We infer that critical components of the *Xist*-induced silencing pathway remain uncharacterized; our preliminary data suggest that ubiquitously-expressed proteins associated with splicing may play a role. A part of [**a**], we developed a protocol to measure RNA-protein interactions that provides a simplified workflow relative to CLIP [**b**].

**a)** Lee D.M., Trotman, J.B., Cherney R.E., Inoue, K. Schertzer M.D., Bischoff S.R., Cowley D.O., **Calabrese J.M.** A 5' fragment of *Xist* can sequester RNA produced from adjacent genes on chromatin. 2019. *Nucleic Acids Research*. 2019 Jul 26; 47(13):7049-62. PMID: 31114903. PMCID in process.

**b)** Raab J.R., Smith K.N., Spear C.S., Manner C.J., **Calabrese J.M**., Magnuson T. 2018. SWI/SNF remains localized to chromatin in the presence of SCHLAP1. *Nature Genetics*. 51(1):26-29. PMID: 30510238. PMC6339527.

*'s denote equal contribution. [ca] denotes co-corresponding authorship.

Complete list of Published work (25 peer-reviewed publications):
https://www.ncbi.nlm.nih.gov/myncbi/j%20mauro.calabrese.1/bibliography/public/