BIOGRAPHICAL SKETCH

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

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. My main expertise is in developing and then applying novel genomic approaches to discover fundamental principles in gene regulation. Over the last 5 years, we have published 13 peer-reviewed papers that established new and unexpected insights into RNA-mediated silencing as well as computational and experimental tools that will accelerate progress in many areas of research ([**a-d**] and **Section C**). Most notably, we developed a method of non-linear sequence comparison which demonstrated that despite lacking linear homology, repressive lncRNAs are enriched in similar sequence motifs, suggesting a general model for how they encode repression [**a**]. We discovered that repressive lncRNAs exhibit enriched associations with specific proteins and DNA regulatory elements, supporting the general model as well as providing an explanation for why certain genomic regions are hypersensitized to silencing by lncRNAs [**b**]. We made discoveries that provide a molecular explanation for how a single domain in the 5' end of *Xist* is required for both the transcription of *Xist* and for the earliest stages of *Xist*-induced gene silencing, potentially revealing new mechanisms by which RNA elements can both promote and repress transcription [**c**]. We developed new approaches to map, quantify, and manipulate lncRNA-protein interactions [**d** and **Section C**]. These advances provide a foundation for dissecting the molecular mechanisms of lncRNAs throughout eukaryotes and in many forms of human disease.

Training: My main educational objectives as a mentor are to help students learn how to perform high-impact biomedical research by promoting hard work, creativity, rigorous and unbiased experimental design, methodology, analysis, interpretation and reporting of results. Five students have obtained Ph.D.'s in my lab, all in a timely fashion with the skills, credentials and experiences to transition into careers in the biomedical research workforce. All have continued working in research or related fields. Six graduate students and one postdoctoral fellow currently train in my laboratory. A total of seven undergraduates and two post-baccalaurate student have trained or are training in my laboratory. ~70% of trainees in my lab are women or identify as being from groups that have been historically minoritized in science. I currently sit or have sat on ~45 thesis committees. I co-teach multiple short courses for graduate students, the most notable being "Practical RNA-seq", in which students learn RNA-seq starting from the molecular biological aspects through the computational analyses. I participate regularly in workshops designed to build mentoring skills and promote healthy interpersonal interactions, communication, and inclusivity. I strive to empower all of my trainees to pursue personalized career trajectories. Since Janurary of 2020, I have served as Director of Graduate Studies in the Department of Pharmacolgy. In

that capacity, I oversee the development and execution of core coursework and provide career support for the 65+ students in our graduate program.

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) 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. F1000 recommended by Prof. Neil Brockdorff.

c) Trotman J.B.*, Lee D.M.*, Cherney R.E., Inoue, K. Schertzer M.D., Bischoff S.R., Cowley D.O., **Calabrese J.M**. 2020. Elements at the 5′ end of *Xist* harbor SPEN-independent transcriptional antiterminator activity. Nucleic Acids Res., gkaa789, https://doi.org/10.1093/nar/gkaa789. PMID: 32986830. PMC7544216.

d) Weidmann C.A., Mustoe A.M., Jariwala P.B., **Calabrese J.M.**, Weeks K.M. 2021. Analysis of RNA-protein networks with RNP-MaP defines functional hubs on RNA. *Nat Biotechnol*. Mar;39(3):347-356. doi: 10.1038/s41587-020-0709-7. PMID: 33077962; PMC7956044.

*'s denote equal contribution.

Ongoing funded projects I would like to highlight:

R01 GM136819 (Calabrese) 5/01/2020-2/29/2024
Cooperative control of Polycomb Repressive Complexes by long noncoding RNAs, CpG island DNA, and RNA- binding proteins
Project Goals: Determine the mechanisms by which DNA elements called CpG islands, RNA-binding proteins, and chromatin-bound RNAs (most notably pre-mRNAs) cooperate to control Polycomb Repressive Complexes, specifically as way to maintain transcriptional fidelity during development and in stress.
JMC Role: Principal Investigator

R01 GM121806 (Calabrese) 5/15/2022-3/31/2026
Mechanisms of gene silencing by long noncoding RNAs
Project Goals: Determine the mechanisms through which *Xist* and related lncRNAs can initiate gene silencing in the absence of known epigenetic regulatory complexes.
JMC Role: Principal Investigator

B. Positions, Scientific Appointments, and Honors

Positions and Scientific Appointments:
2002 Graduate Student, Phillip Sharp lab, Biology Department, MIT
2008 Postdoctoral Fellow, Terry Magnuson lab, Genetics Department, UNC Chapel Hill
2014 Assistant Professor of Pharmacology, UNC Chapel Hill
2020 Associate Professor of Pharmacology, UNC Chapel Hill
2020 Director of Graduate Studies, Pharmacology Curriculum, UNC Chapel Hill
2022 co-Director, RNA Discovery Center, UNC Chapel Hill

Other experience and professional memberships:
2016 Ad-hoc member, Molecular Genetics B Study Section (NIH)
2017 Ad-hoc member, Graduate Women In Science grant reviewer
2018 Ad-hoc member, CSR Program Evaluation The Role of Anonymization in Peer Review
2018 Ad-hoc member, Review of K99/R00 applications to NIGMS [ZGM1 TWD-8 (KR)]
2019 Ad-hoc member, Breast Cancer Now grant review

2019 Ad-hoc member, Review of K99/R00 applications to NIGMS [ZGM1 TWD-8 (KR)]
2021 Ad-hoc reviewer, Fondation pour la Recherche Médicale

Member, RNA Society
Member, Genetics Society of America
Member, Lineberger Comprehensive Cancer Center, UNC Chapel Hill

Honors and Awards:
1997 Order Sons of Italy Scholarship
2000 Elvehjem Scholarship for Excellence in Biochemistry
2001 Phi Beta Kappa
2009 American Cancer Society Post-doctoral fellowship
2013 UNC Chapel Hill Postdoctoral Research Excellence Award
2015 March of Dimes Basil O'Connor Scholar

C. Contributions to Science (chronological)

***1) Divergent transcription from active promoters*.**

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.

a) Preker, P…Jensen, T.H. et al. 2008. RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters. *Science*. 322:1851-1854

2) Molecular interactions between lncRNAs and target-gene chromatin.

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, that *Xist*, *Airn*, and *Kcnq1ot1* each require the same RNA-binding protein (HNRNPK) to elicit the accumulation of Polycomb in their target chromosomal domains, and that subsets of DNA regulatory elements are hypersensitized to silencing by lncRNAs [**b**]. Understanding the physical nature of the RNA-protein interactions that drive lncRNA-mediated effects, as well as 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. F1000 recommended by Prof. Neil Brockdorff.

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. PMC6633203.

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) Relationships between RNA structure and regulatory function in lncRNAs.

At 18 kilobases in length, *Xist* is one of the longest regulatory RNAs known. When I started my lab in 2014, 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*. We developed and optimized methods to use SHAPE chemistry in living cells [**a**]. We employed this approach to report the first nucleotide-resolution structural map of authentic, full-length *Xist* [**b**]. More recently, we developed a chemical probing approach to identify protein interaction networks on RNAs in living cells [**c**]. Together, [**a,b,c**] represent some of the first methods to map RNA structure and RNA/protein interactions in living cells. We are using these approaches to understand how RNA structure confers function to *Xist* and *Xist*-analogous RNAs, which should provide paradigms relevant to many RNAs in the transcriptome.

e) Smola M.J., **Calabrese J.M**., Weeks K.M. 2015. Detection of RNA-Protein Interactions in Living Cells with SHAPE. *Biochemistry*. 54(46):6867-75. PMID: 17135486. PMC1664724.

f) 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. PMID: 27578869. PMC5027438.

g) Weidmann C.A., Mustoe A.M., Jariwala P.B., **Calabrese J.M.**, Weeks K.M. 2021. Analysis of RNA-protein networks with RNP-MaP defines functional hubs on RNA. *Nat Biotechnol*. Mar;39(3):347-356. doi: 10.1038/s41587-020-0709-7. PMID: 33077962; PMC7956044.

4) Computational methods to identify regulatory function in lncRNAs.

The functions of most lncRNAs are unknown. In contrast to proteins, lncRNAs with similar function often lack linear sequence homology; thus, the identification of function in one lncRNA rarely informs the 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 demonstrated that evolutionarily unrelated lncRNAs can encode similar function through different spatial arrangements of functionally analogous protein-binding domains [**a**, **b**]. The early successes of *k*-mer-based classification schemes suggest it will be possible to computationally identify recurrent relationships between sequence and function in many types of lncRNAs. One could liken these discoveries to finally recognizing a way to understand the different scripts in the Rosetta Stone; understanding the sequence patterns that confer function in well-

characterized lncRNAs will enable the identification and subsequent study of analgous patterns in uncharacterized lncRNAs. We are now investigating ways in which *k*-mer content relates to structure and function in diverse forms of RNA, including those that contain introns. We are also working to incorporate markov-chain statistics and structural predictions into *k*-mer based comparisons, which should improve the power of *k*-mer-based searches as well as make them accessible to a greater subset of 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**. 2019. Non-linear sequence similarity between *Xist* and *Rsx* suggests shared functions of tandem repeat domains. *RNA*. 25(8):1004-1019. PMID: 31097619. PMC6633197.

c) Kirk J.M.*, Sprague D.*, **Calabrese J.M.** Classification of Long Noncoding RNAs by *k*-mer Content. 2021. *Methods Mol Biol.* 2254:41-60. doi: 10.1007/978-1-0716-1158-6_4. PubMed PMID: 33326069; PMC7850294.

5) Mechanisms of lncRNA-induced gene silencing and lncRNA production.

Xist is thought to encode repressive function through the concerted action of at least four RNA-protein interaction domains that are each comprised of discrete tandem repeats. Necessary for the function of these tandem repeats are ubiquitously expressed proteins that also bind tens of thousands of positions in thousands of different RNAs. How ubiquitously expressed proteins are leveraged by *Xist* and potentially other lncRNAs to orchestrate epigenetic silencing remains unclear. With this in mind, we recently completed a study of the Repeat A domain in *Xist*. For >20 years, Repeat A has been considered to be the domain within *Xist* responsible for inducing transcriptional silencing. Our findings indicate that much of what has been assumed about the function of Repeat A needs to be reconsidered. Specifically, we discovered that when expressed as a stand-alone transgene, a 5´ fragment of *Xist* containing Repeat A is unexpectedly unable to induce transcriptional silencing. Yet, in this same context, the transgenic RNA associates with chromatin, binds the silencing factor SPEN, and functions as a transcriptional antiterminator [**a**]. In endogenous contexts, this antiterminator activity may help produce full-length *Xist* RNA while rendering the *Xist* locus resistant to silencing by the same repressive complexes that the lncRNA recruits to other genes. Our data hint at the existence of a broader class of RNA elements that function to promote transcriptional processivity in mammals. Moreover, data from my lab and others support the hypothesis that Repeat A functions in silencing upstream of chromatin-modifying enzmes, by disrupting protein interactions that would normally occur on nascent RNAs of *Xist*-target genes. As part of [**a**], we developed a protocol to measure RNA-protein interactions that facilitates quantitative cross-sample comparisons and provides a simplified workflow relative to CLIP [**b**].

a) Trotman J.B.*, Lee D.M.*, Cherney R.E., Inoue, K. Schertzer M.D., Bischoff S.R., Cowley D.O., **Calabrese J.M**. 2020. Elements at the 5′ end of *Xist* harbor SPEN-independent transcriptional antiterminator activity. Nucleic Acids Res., gkaa789, https://doi.org/10.1093/nar/gkaa789. PMID: 32986830. PMC7544216.

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 (32 peer-reviewed publications):
https://www.ncbi.nlm.nih.gov/myncbi/j%20mauro.calabrese.1/bibliography/public/