Perelman School of Medicine
University of Pennsylvania

Panagiotis Alexiou, PhD
Research Associate

Panagiotis Alexiou

Department of
Pathology and Laboratory Medicine

Division of Neuropathology

613B Stellar Chance Labs
422 Curie Boulevard

Philadelphia, PA 19104-6100

(+1) 267- 277-2661

palexiou@mail.med.upenn.edu

http://www.panalexiou.com/

27 APRIL 2017

CEITEC
Central European Institute of Technology
Brno, Czech Republic

To whom it may concern,

I am writing to apply for the position of 'Head of Bioinformatics Unit' posted recently on the 'nature-jobs' website. Allow me to briefly expose my skills and experience relevant to the position, following with a brief summary of future research directions.

Research Experience

Following my interdisciplinary education in Genetics, Molecular Biology and Bioinformatics, I acquired my PhD in the field of the bioinformatic analysis of microRNA biogenesis and function. The work performed during my PhD introduced me to several subfields of bioinformatics, such as machine learning for microRNA target prediction (classification), the development of relational databases and web-servers, motif analyses, scientific knowledge text mining - as well as the theoretical fields of Small RNA Biology and RNA Binding Protein Biology which would be the unifying paths of my research to date.

I continued my research as a Postdoctoral Researcher at the Perelman School of Medicine at the University of Pennsylvania, joining the Mourelatos Lab, one of the premier laboratories worldwide in small RNA biology and RNA binding protein research. During my time at the University of Pennsylvania, I have collaborated directly with biomedical researchers working extensively with Next Generation Sequencing (NGS) data. I have developed techniques and algorithms for the analysis of NGS data in the fields of piRNA biology (both biogenesis and function), microRNA biology, and RNA Binding Proteins function in neurodegeneration. Specifically pertaining to NGS analysis, I have helped develop both an NGS analysis suite (CLIP-Seq-Tools) and a modern perl object-oriented framework for NGS analysis (GenOO).

Track record of scientific productivity

During my career I have produced 24 publications, which have collectively been cited over 3000 times in the literature (h-index 17, i-10 index 18). The online resources I helped produce during my PhD serve approximately 500 unique users daily.

My work has lead to a number of discoveries including the function of Piwi piRNA complexes in germline development, a pre-miRNA surveillance system of quality control of miRNA synthesis and the effects of TAF15 and FUS on the neuronal transcriptome. Some of these publications are in journals of the highest impact in the field such as Molecular Cell, Nature Structural & Molecular Biology, Genes & Development and Nature, a paper that was highlighted in a commentary in Developmental Cell and has been met by overwhelming acceptance by the piRNA Biology community.

An overview of my publications can be found in the attached Curriculum Vitae.

Communication skills

I pride myself in concise, direct and accurate communication. My career to date has always been in radically interdisciplinary environments, collaborating day to day with Biochemists, Computer Scientists, Bioinformaticians, Geneticists, Medical Doctors and so on. I often and happily take the role of inter-discipline translator between colleagues coming from different backgrounds. I believe my communication (and teaching) skills to be some of my strongest points as a scientist.

Organizational and leadership capabilities

During the PhD and Postdoctoral phases of my career I have had some opportunities to use my organizational and leadership capabilities. I had pivotal roles in setting up and maintaining the hardware of the two labs of which I was member. Additionally I organized code development and review with coworkers around the agile/xp development model and the pair programming approach.

As a teaching assistant and substitute lecturer, I developed and taught an elective course on Bioinformatics for last year engineering students and MSc students. Due to financial constraints of the university, I was tasked with single handedly teaching and overviewing a student lab of 60 students, including 3 students doing their theses in my field. I managed to lead all 3 of these students to produce code fit for inclusion in publications from my lab.

Proposed Future Work

My career up to now has been focused on finding ways to efficiently implement novel technologies that spearhead innovation in every subfield of research I have been involved with. I am a proponent of open software and scientific resource development, having produced web-tools that are being used by thousands of users per month, open-source tools for the analysis of sequencing data, and even a programming package that codifies genomic and NGS entities into programming objects, allowing for rapid and robust prototyping of future services. My proposed future work involves the development of tools for the classification and exploration of NGS data. Specifically, I am proposing the development of a machine-learning trained function classification system for NGS sequencing data of RNA Binding Proteins. Please find attached a brief overview of my research plan.

I believe I have the knowledge, expertise, understanding, and will, to tackle the issues that will face biomedical research in the following years - and I hope that I will be given this opportunity at your Research Institute.

Sincerely,

Panagiotis Alexiou, Phd

Panagiotis Alexiou

Research Associate
Perelman School of Medicine
University of Pennsylvania, USA

Education & Research

University of Aberdeen, UK — *BSc Genetics with Industrial Placement*

1999 - 2004

Universiteit van Amsterdam, Netherlands — *MSc Molecular Cell Biology and Bioinformatics*

2004 - 2006

Aristotelian University of Thessaloniki, Greece — *PhD*
BSRC Alexander Fleming, Greece — *PhD Research*

2007-2011

University of Pennsylvania, USA — *Postdoctoral Research*

2011 - 2016

University of Pennsylvania, USA — *Research Associate*

2016 - Present

References

Zissimos Mourelatos, MD
Professor of Pathology and Laboratory Medicine
University of Pennsylvania Perelman School of Medicine
Department of Pathology and Laboratory Medicine
Email: mourelaz@uphs.upenn.edu

Artemis Hatzigeorgiou, PhD
Professor of Bioinformatics
University of Thessaly, Department of Electrical and Computer Engineering
Email: arhatzig@inf.uth.gr

Theodore Dalamagas, PhD
Senior Researcher
ATHENA Research Center
Email: dalamag@imis.athena-innovation.gr

Contact

Pathology and Laboratory Medicine
Division of Neuropathology
613B Stellar Chance Labs
422 Curie Boulevard
Philadelphia, PA 19104-6100
(+1) 267- 277-2661
palexiou@mail.med.upenn.edu
http://www.panalexiou.com/

Research Focus & Skills

RNA-binding protein biology

NGS sequencing data analysis

Programming (perl, golang)

Statistics (R)

Bioinformatics Algorithms

Grants & Scholarships

2007-2010
BSRC Alexander Fleming PhD funding scholarship

2010-2012
IKYDA personnel exchange program grant (Greece-Germany) in collaboration with Martin Luther University Halle-Wittenberg

Teaching

2010-2011
Graduate Teaching Assistant
University of Thessaly, Department of Electrical and Computer Engineering
Bioinformatics (HY501)

2010-2011
Substitute Lectures
National and Kapodistrian University of Athens, Department of Informatics and Communications,
MSc Information Technologies in Medicine and Biology

Publications

Citations	3008
h-index	17
i10-index	18

List in chronological order (* denotes equal contribution):

1. Zhang L, Volinia S, Bonome T, Calin GA, Greshock J, Yang N, Liu CG, Giannakakis A, **Alexiou P**, Hasegawa K, Johnstone CN, Megraw MS, Adams S, Lassus H, Huang J, Kaur S, Liang S, Sethupathy P, Leminen A, Simossis VA, Sandaltzopoulos R, Naomoto Y, Katsaros D, Gimotty PA, DeMichele A, Huang Q, Bützow R, Rustgi AK, Weber BL, Birrer MJ, Hatzigeorgiou AG, Croce CM, Coukos G (2008) Genomic and Epigenetic Alterations Deregulate microRNA Expression in Human Epithelial Ovarian Cancer, **Proc Natl Acad Sci U S A** ,105(19) ,7004-9. (cited: 500)
2. Maragkakis M, Reczko M, Simossis VA, **Alexiou P**, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG. (2009) DIANA-microT web server: elucidating microRNA functions through target prediction., **Nucleic Acids Res.** ,37 W273-6. (cited: 462)
3. Papadopoulos GL, **Alexiou P**, Maragkakis M, Reczko M, Hatzigeorgiou AG. (2009) DIANA-mirPath: Integrating human and mouse microRNAs in pathways., **Bioinformatics**. ,2009 Aug 1;25(15):1991-3. (cited: 252)
4. Maragkakis M*, **Alexiou P***, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA, Sethupathy P, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG. (2009) Accurate microRNA target prediction correlates with protein repression levels. **BMC Bioinformatics** ,2009 Sep 18;10:295. (cited: 318)
5. **Alexiou P**, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG (2009) Lost in translation: an assessment and perspective for computational microRNA target identification., **Bioinformatics** , 25(23):3049-55. (cited: 285)
6. **Alexiou P**, T. Vergoulis, M. Gleditzsch, G. Prekas, T. Dalamagas, M. Megraw, I. Grosse, T. Sellis, A.G. Hatzigeorgiou (2009) miRGen 2.0: a database of microRNA genomic information and regulation, **Nucleic Acids Research** ,Nucleic Acids Res. 2010 January; 38(Database issue): D137–D141. (cited: 138)
7. **Alexiou P**, Maragkakis M, Papadopoulos GL, Simmosis VA, Zhang L, Hatzigeorgiou AG (2010) The DIANA-mirExTra web server: from gene expression data to microRNA function., **PLoS ONE** ,5(2):e9171. (cited: 69)
8. Marotta D, Karar J, Jenkins WT, Kumanova M, Jenkins KW, Tobias JW, Baldwin D, Hatzigeorgiou A, **Alexiou P**, Evans SM, Alarcon R, Maity A, Koch C, Koumenis C. (2011) In vivo profiling of hypoxic gene expression in gliomas using the hypoxia marker EF5 and laser-capture microdissection. **Cancer Res** ,71(3):779-89. (cited: 33)
9. **Alexiou P,** Maragkakis M, Hatzigeorgiou AG (2011) Online resources for microRNA analysis. **Journal of Nucleic Acid Investigation**, 2(1). (cited: 7)
10. Maragkakis M, Vergoulis T, **Alexiou P**, Reczko M, Plomaritou K, Gousis M, Kourtis K, Koziris N, Dalamagas T, Hatzigeorgiou AG. (2011) DIANA-microT Web server upgrade supports Fly and Worm miRNA target prediction and bibliographic miRNA to disease association. **Nucleic Acids Res.** W145-8. (cited: 80)
11. Vergoulis T, Vlachos IS, **Alexiou P**, Georgakilas G, Maragkakis M, Reczko M, Gerangelos S, Koziris N, Dalamagas T, Hatzigeorgiou AG. (2011) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. **Nucleic Acids Res.** 2011 Dec 1. (cited: 401)
12. Reczko M, Maragkakis M, **Alexiou P**, Grosse I, Hatzigeorgiou AG (2012) Functional microRNA targets in protein coding sequences **Bioinformatics** 28 (6), 771-776 (cited: 188)
13. Reczko M, Maragkakis M, **Alexiou P**, Papadopoulos GL, Hatzigeorgiou AG. (2012) Accurate microRNA target prediction using detailed binding site accessibility and machine learning on proteomics data. **Frontiers in Bioinformatics and Computational Biology**. (cited: 19)
14. Vourekas A, Zheng Q, **Alexiou P**, Maragkakis M, Kirino Y, Gregory BD, Mourelatos Z. (2012) Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. **Nat Struct Mol Biol.** 19(8):773-81. (cited: 103)
15. Honda S, Kirino Y, Maragkakis M, **Alexiou P**, Ohtaki A, Murali R, Mourelatos Z, Kirino Y. (2013) Mitochondrial protein BmPAPI modulates the length of mature piRNAs. **RNA**. 2013 Oct;19(10):1405-18(cited: 26)
16. Ibrahim F, Maragkakis M*, **Alexiou P***, Maronski MA, Dichter MA, Mourelatos Z. (2013) Identification of in vivo, conserved, TAF15 RNA binding sites reveals the impact of TAF15 on the neuronal transcriptome. **Cell Rep.** 3(2):301-8. (cited: 13)
17. Nakaya T, **Alexiou P***, Maragkakis M*, Chang A, Mourelatos Z. (2013) FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. **RNA** 19(4):498-509. (cited: 49)

18. Liu X, Zheng Q, Vrettos N, Maragkakis M, **Alexiou P**, Gregory BD, Mourelatos Z. (2014) A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis. **Mol Cell.** 55(6):868-79. (cited: 29)

19. Vourekas A, Zheng K, Fu Q, Maragkakis M, **Alexiou P**, Ma J, Pillai RS, Mourelatos Z, Wang PJ. (2015) The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. **Genes Dev.** 29(6):617-29. (cited: 22)

20. Maragkakis M*, **Alexiou P***, Mourelatos M. (2015) GenOO: A Modern Perl Framework for High Throughput Sequencing analysis. **bioRxiv** doi:http://dx.doi.org/10.1101/019265. (cited: 1)

21. Maragkakis M*, **Alexiou P***, Nakaya T, Mourelatos M. (2016) CLIPSeqTools—a novel bioinformatics CLIP-seq analysis suite. **RNA**. 22(1):1-9. (cited: 5)

22. Vourekas A*, **Alexiou P***, Vrettos N, Maragkakis M, Mourelatos Z. (2016) Sequence-dependent but not sequence-specific piRNA adhesion traps and anchors mRNAs to the germ plasm. **Nature** 531(7594):390-4. (cited: 8)
this publication was highlighted in: Voronina E (2016) mRNAs Hit a Sticky Wicket. Dev Cell;37(1):9-10

23. Vrettos N, Maragkakis M, **Alexiou P**, Mourelatos Z. (2017) Kc167, a widely used Drosophila cell line, contains an active primary piRNA pathway. **RNA** 23 (1), 108-118

24. **Alexiou P***, Maragkakis M, Mourelatos Z, Vourekas A* (2017). cCLIP-Seq: Retrieval of chimaeric reads from HITS-CLIP (CLIP-Seq) libraries. **Methods in Molecular Biology** (in press)

Next Generation Sequencing based classification and exploration of RNA Binding Protein function.

Background

The completion of the Human Genome Project has ushered in a new era of genetic and medical research by defining the human genome in its entirety. The contribution of perhaps even greater importance is the accompanying technological developments in high throughput sequencing that now allow researchers to sequence human genomes in a few days for as little as 1,000 €. High throughput Next Generation Sequencing (NGS) technologies are continuously lowering the cost of biological molecules sequencing through parallelization, making NGS a widely used tool for genomic analysis and creating an amazing trove of sequencing data openly available to researchers. (NGS reviewed in (van Dijk et al. 2014))

RNA Binding Proteins (RBPs) take part in many physiological cellular functions (splicing, regulation of translation, mRNA decay, transposon suppression etc) and their misregulation has been implicated with several diseases (neurodegenerative diseases, cancer etc). A number of NGS techniques are being used to identify the binding sites of RBPs, HITS-CLIP (high-throughput sequencing of RNA that is isolated after ultraviolet (UV) irradiation-induced crosslinking and immunoprecipitation) (Chi et al. 2009; Moore et al. 2014) individual-nucleotide resolution iCLIP (König et al. 2010) photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) (Hafner et al. 2010) and so on., have been used to identify RBP binding sites on RNA molecules, as well as the small RNA load and targets of RBPs that utilize small RNA 'driver' sequences. Currently, there are hundreds such experiments performed and publicly deposited per year; this rate bound to accelerated increase in the foreseeable future due to reducing costs and ubiquitousness of sequencing machines.

Machine learning systems are currently used to recognize patterns in biomedical images, predict protein structures, classify RNAs as potential targets of RBPs, and many other applications. Recent developments in Machine Learning, such as Deep Learning (reviewed in (LeCun et al. 2015)), are creating new opportunities to harness big data and identify patterns from disparate sources. Conventional machine learning techniques are limited by the need of the development of domain specific features, a process that can introduce bias, is time consuming, and requires extended knowledge of the training sets. Deep learning type techniques instead use several layers of automatically created representation to learn features from raw data in a time efficient manner.

My proposed research focus is the use of Machine Learning techniques to interrogate a large number of available NGS RBP datasets in order to classify, and eventually predict, RBP function and interplay. Conceptually this proposal can be broken down in three aims:

1. Standardization of NGS analyses for RBPs
2. Classification of RBPs
3. Prediction and Exploration of RBP function

Aim 1: Standardization of NGS analysis for RBPs.

The first focus of my research will be to build a series of modular standardized tools for NGS analysis of RBPs. The first steps towards this direction have already been done with the development of ClipSeqTools (Maragkakis et al. 2016) - a suite for the analysis of NGS reads. Additional modules pertaining to secondary structure, target sequence motifs, and relative positioning between samples are already under way for the next installment of the tool. I am currently working towards a framework that deals with RBPs that use small RNAs as drivers (such as piRNAs or miRNAs). User friendly and standardized analyses by themselves are useful for the growing community, but are also the base on which further aims will be built on.

Previously, I have been heavily involved with the development of coding tools used for NGS data analysis (GenOO (Maragkakis et al. 2015), ClipSeqTools (Maragkakis et al. 2016)), which were in turn applied on several different types of RBPs, helping elucidate the biological functions of said RBPs. Specifically, I was heavily involved in the identification of piRNA biogenesis and function in mouse testis, based on CLIP-Seq and RNA-Seq analysis of the mouse MIWI and MILI RBPs (Vourekas et al. 2012). The identification of the role of a mitochondrial RBP (BmPAPI) in the biogenesis and maturation of piRNAs using NGS (Honda et al. 2013).The identification of a mechanism that performs 'quality control' on microRNA precursors, identified via a novel method for circularization and NGS of small RNA precursors (Liu et al. 2014). The identification of the roles of the RBP MOV10L1 in the processing of piRNA precursors in which the use of CLIP-Seq data allowed us to create a model that explained the determination of initiation of piRNA processing (Liu et al. 2014; Vourekas et al. 2015). My latest large published project, identified the mechanism by which an RBP (Aubergine) specifies germ cells in fly embryos using small RNAs (piRNAs) as non-specific drivers (Vourekas et al. 2016). This project by itself involved various analyses such as small RNA targeting prediction using chimeric reads, coverage of mRNAs and piRNA precursors by CLIP reads, enrichment of targeting in localization categories and a total of over 30 NGS samples. I have also worked on two RBPs of the FET family (TAF15 and FUS) associated with ALS (Amyotrophic Lateral Sclerosis or Lou Gehrig's disease) using NGS techniques (mainly CLIP-Seq and RNA-Seq). For TAF15 (Ibrahim et al. 2013) we identified conserved binding sites that affected neuronal transcription. For FUS (Nakaya et al. 2013) we identified conserved intron regions, bound by FUS in patients, that affect splicing of target genes. Both of these projects involved analyses including motif finding, conservation analysis, functional analysis, localization on mRNAs, intron/exon localization.

Aim 2: Classification of RBPs.

Using the hundreds of available datasets we will classify proteins based on their binding profiles. The aim of this endeavour will be the development of sets of features that will classify experiments in groups of close similarity based on their binding profiles and the identity of their target genes. Essentially, RBPs that bind in similar ways, and bind similar genes will be clustered together based on the sets of features that are deemed most important by the Machine Learning classifier.

Previously, I have been involved in the development of Machine Learning classifiers for the prediction of miRNA target sites with high precision (Manolis Maragkakis et al. 2009; Reczko et al. 2012; Reczko et al. 2011). I am currently developing a Machine Learning classifier based on chimeric reads (Vourekas et al. 2016) for the classification of piRNA binding sites across several species.

Aim 3: Prediction and Exploration of RBP function.

Finally, the products of Aim 1 and 2 will allow the development of a tool that allows researchers to predict functional characteristics of their RBP based on an NGS experiment. Briefly, a standardized pipeline (Aim 1) will be used to extract features (defined by Aim 2). Features will then be used to compare and contrast the RBP of choice against the background of all RBPs used for training (Aim 2). The new RBP will be classified, and significant differences against the background will be reported to the user.

Previously, I have been involved in the development of highly used Web servers (~ 2000 unique users/week currently) that allow users to explore miRNA related data such as experimentally verified targets (Vergoulis et al. 2011), predicted targets (M. Maragkakis et al. 2009; Maragkakis et al. 2011), miRNA transcripts and their regulatory regions (Alexiou, Vergoulis, et al. 2010) and so on (Alexiou, Maragkakis, et al. 2010; Papadopoulos et al. 2009).

Conclusion

The pace of technological innovation is ever accelerating in our field, opening new horizons and opportunities. The harnessing of big data using machine learning, as applied on the field of RNA biology, is a promising field that will create novel biological knowledge, as well as infrastructure - accelerating the research process for future experiments.

References

1. Alexiou, P., Vergoulis, T., et al., 2010. miRGen 2.0: a database of microRNA genomic information and regulation. *Nucleic acids research*, 38(Database issue), pp.D137–41.
2. Alexiou, P., Maragkakis, M., et al., 2010. The DIANA-mirExTra web server: from gene expression data to microRNA function. *PloS one*, 5(2), p.e9171.
3. Chi, S.W. et al., 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, 460(7254), pp.479–486.
4. van Dijk, E.L. et al., 2014. Ten years of next-generation sequencing technology. *Trends in genetics: TIG*, 30(9), pp.418–426.
5. Hafner, M. et al., 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, 141(1), pp.129–141.
6. Honda, S. et al., 2013. Mitochondrial protein BmPAPI modulates the length of mature piRNAs. *RNA* , 19(10), pp.1405–1418.
7. Ibrahim, F. et al., 2013. Identification of in vivo, conserved, TAF15 RNA binding sites reveals the impact of TAF15 on the neuronal transcriptome. *Cell reports*, 3(2), pp.301–308.
8. König, J. et al., 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*, 17(7), pp.909–915.
9. LeCun, Y., Bengio, Y. & Hinton, G., 2015. Deep learning. *Nature*, 521(7553), pp.436–444.
10. Liu, X. et al., 2014. A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis. *Molecular cell*, 55(6), pp.868–879.
11. Maragkakis, M. et al., 2009. Accurate microRNA target prediction correlates with protein repression levels. *BMC bioinformatics*, 10, p.295.
12. Maragkakis, M. et al., 2016. CLIPSeqTools--a novel bioinformatics CLIP-seq analysis suite. *RNA* , 22(1), pp.1–9.
13. Maragkakis, M. et al., 2009. DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic acids research*, 37(Web Server issue), pp.W273–6.
14. Maragkakis, M. et al., 2011. DIANA-microT Web server upgrade supports Fly and Worm miRNA target prediction and bibliographic miRNA to disease association. *Nucleic acids research*, 39(Web Server issue), pp.W145–8.
15. Maragkakis, M. et al., 2015. *GenOO: A Modern Perl Framework for High Throughput Sequencing analysis*, Available at: http://dx.doi.org/10.1101/019265.
16. Moore, M.J. et al., 2014. Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nature protocols*, 9(2), pp.263–293.
17. Nakaya, T. et al., 2013. FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. *RNA* , 19(4), pp.498–509.
18. Papadopoulos, G.L. et al., 2009. DIANA-mirPath: Integrating human and mouse microRNAs in pathways. *Bioinformatics* , 25(15), pp.1991–1993.
19. Reczko, M. et al., 2011. Accurate microRNA Target Prediction Using Detailed Binding Site Accessibility and Machine Learning on Proteomics Data. *Frontiers in genetics*, 2, p.103.
20. Reczko, M. et al., 2012. Functional microRNA targets in protein coding sequences. *Bioinformatics* , 28(6), pp.771–776.
21. Vergoulis, T. et al., 2011. TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic acids research*, p.gkr1161.
22. Vourekas, A. et al., 2012. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nature structural & molecular biology*, 19(8), pp.773–781.
23. Vourekas, A. et al., 2016. Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature*, 531(7594), pp.390–394.
24. Vourekas, A. et al., 2015. The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. *Genes & development*, 29(6), pp.617–629.