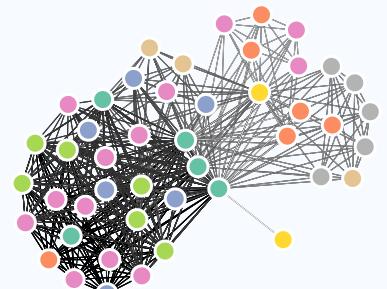


# MIRNA VAZQUEZ ROSAS LANDA

I have made visualizations viewed by hundreds of thousands of people<sup>1</sup>, sped up query times for 25 terabytes of data by an average of 4,800 times<sup>2</sup>, and built packages for R<sup>3</sup> that let you do magic<sup>4</sup>.



## EDUCATION

2017  
|  
2012

### PhD in Sciences

National Autonomous University of Mexico, Ecology Institute.  
 Mexico City, MX

- **Title:** Patterns of diversification and adaptation of the Vibrionaceae family. The case of Cuatro Cienegas.
- **Advisor:** Dr. Valeria Souza Saldivar.
- **Description:** I examined the evolutionary footprints of local adaptation in the genomes of bacteria isolated from a highly oligotrophic environment in Cuatro Ciénegas (CCB), Mexico. I obtained 200 axenic isolates of Vibrionaceae, a group of bacteria with a cosmopolitan distribution, from several ponds around the CCB aquatic system. I identified the patterns of genetic variation among Vibrionaceae isolates, looking for signals of adaptation to structured (Water column) and non-structured (Sediments) environments within the ponds, as well; I analyzed whether the recombination patterns were associated with these environments or constrained to the lineage history. Since several members of Vibrionaceae possess pathogenic relatives, I also evaluate the hypothesis that the members of Vibrionaceae isolated from Cuatro Ciénegas (CCB) had no pathogenic genomic elements, under the assumption that bacteria from CCB rarely get into contact with other bacteria, like oceanic Vibrionaceae.

2011  
|  
2007

### B.Sc. in Biology

National Autonomous University of Mexico, FES Iztacala  
 Mexico City, MX

## RESEARCH EXPERIENCE

current  
|  
2020

### Postdoctoral fellow

The University of Texas at Austin, [Marine Microbial Ecology Lab](#)  
 Austin, USA.

- Currently studying the ecology and evolution of microbes by understanding their metabolism, using metagenomics.
- **Advisor:** Brett Baker<sup>5</sup>

Current  
|  
2015

### Guest Research

The University of Toronto. Cell and System Biology Department  
 Toronto, CA

- Realized the pan-genome analysis of the genomes that I used for my PhD thesis.
- **Advisor:** David Guttman

View this CV online with links at  
[https://github.com/mirnavazquez/MVRL\\_cv/CV\\_](https://github.com/mirnavazquez/MVRL_cv/CV_/)

## CONTACT

-  [mirna@gmail.com](mailto:mirna@gmail.com)
-  [MirnaVRL](#)
-  [MirnaVRL](#)
-  [MirnaVRL](#)
-  [ORCID](#)
-  [Google Scholar](#)

## LANGUAGE SKILLS



Made with the R package  
[pagedown](#).

The source code is available on  
[github.com/hstrayer/cv](https://github.com/hstrayer/cv).

Last updated on 2021-08-18.

Current   2014	<ul style="list-style-type: none"> <li>● <b>Guest Research</b> National Laboratory of Genomics for Biodiversity           <ul style="list-style-type: none"> <li>• I learned how to assembly and annotate genomes.</li> <li>• Advisor: Enrique Ibarra-Laclette</li> </ul> </li> </ul>	Irapuato, MX
Current   2011	<ul style="list-style-type: none"> <li>● <b>Guest Research</b> Center for Research and Advanced Studies of the National Polytechnic Institute           <ul style="list-style-type: none"> <li>• I designed a bacteria that was capable of sensing the arsenic concentrations in water.</li> <li>• Advisor: Agustino Martínez Antonio</li> </ul> </li> </ul>	Irapuato, MX
2020   2017	<ul style="list-style-type: none"> <li>● <b>Postdoctoral fellow</b> Ecology Institute, A.C. <a href="#">Genomics and transcriptomics lab</a> <ul style="list-style-type: none"> <li>• I studied the interaction among plants-insects-microbes. - I designed a diagnosys system to identify Fusarium kurosho using genomics. - I studied the defense mechanisms of Persea americana var. Hass to the attac of the Mexican Fruit Fly, using transcriptomics. - I analyze the microbiome of the Mexican Fruit Fly.</li> <li>• Advisor: Martín Aluja and Enrique Ibarra-Laclette</li> </ul> </li> </ul>	Xalapa, MX.

## 💼 INDUSTRY EXPERIENCE

Current   2020	<ul style="list-style-type: none"> <li>● <b>Software Engineer</b> RStudio           <ul style="list-style-type: none"> <li>• Helping make programming web applications with R easier and more beautiful on the Shiny team</li> </ul> </li> </ul>	Remote
2016   2016	<ul style="list-style-type: none"> <li>● <b>Data Journalist - Graphics Department</b> New York Times           <ul style="list-style-type: none"> <li>• Reporter with the graphics desk covering topics in science, politics, and sport.</li> <li>• Work primarily done in R, Javascript, and Adobe Illustrator.</li> </ul> </li> </ul>	New York, New York
2015   2015	<ul style="list-style-type: none"> <li>● <b>Engineering Intern - User Experience</b> Dealer.com           <ul style="list-style-type: none"> <li>• Built internal tool to help analyze and visualize user interaction with back-end products.</li> </ul> </li> </ul>	Burlington, VT
2015   2015	<ul style="list-style-type: none"> <li>● <b>Data Science Intern</b> Dealer.com           <ul style="list-style-type: none"> <li>• Worked with the product analytics team to help parse and visualize large stores of data to drive business decisions.</li> </ul> </li> </ul>	Burlington, VT

I have worked in a variety of roles ranging from journalist to software engineer to data scientist. I like collaborative environments where I can learn from my peers.

2015  
|  
2014

- **Data Artist In Residence**  
Conduce 📍 Carpinteria, CA
  - Envisioned, prototyped and implemented visualization framework in the course of one month.
  - Constructed training protocol for bringing third parties up to speed with new protocol.
  
- **Software Engineering Intern**  
Conduce 📍 Carpinteria, CA
  - Incorporated d3.js to the company's main software platform.

2014  
|  
2014

## 💻 TEACHING EXPERIENCE

2020

- **Javascript for Shiny Users**  
RStudio::conf 2020
  - Served as TA for two day workshop on how to leverage Javascript in Shiny applications
  - Lectured on using R2D3 package to build interactive visualizations.<sup>6</sup>

2019  
|  
2019

- **Data Visualization Best Practices**  
DataCamp
  - Designed from bottom up course to teach best practices for scientific visualizations.
  - Uses R and ggplot2.
  - In top 10% on platform by popularity.

2019  
|  
2019

- **Improving your visualization in Python**  
DataCamp
  - Designed from bottom up course to teach advanced methods for enhancing visualization.
  - Uses python, matplotlib, and seaborn.

2018  
|  
2017

- **Advanced Statistical Learning and Inference**  
Vanderbilt Biostatistics Department 📍 Nashville, TN
  - TA and lectured
  - Topics covered from penalized regression to boosted trees and neural networks
  - Highest level course offered in department

2018  
|  
2018

- **Advanced Statistical Computing**  
Vanderbilt Biostatistics Department 📍 Nashville, TN
  - TA and lectured
  - Covered modern statistical computing algorithms
  - 4th year PhD level class

I am passionate about education. I believe that no topic is too complex if the teacher is empathetic and willing to think about new methods of approaching task.

2017  
|  
2017

- **Statistical Computing in R**  
Vanderbilt Biostatistics Department  Nashville, TN
  - TA and lectured
  - Covered introduction to R language for statistics applications
  - Graduate level class

## ↗ SELECTED DATA SCIENCE WRITING

- 2019 ● **Using AWK and R to Parse 25tb<sup>8</sup>**  
[LiveFreeOrDichotomize.com](#)
  - Story of parsing large amounts of genomics data.
  - Provided advice for dealing with data much larger than disk.
  - Reached top of HackerNews.
- 2018 ● **Classifying physical activity from smartphone data<sup>9</sup>**  
[RStudio Tensorflow Blog](#)
  - Walk through of training a convolutional neural network to achieve state of the art recognition of activities from accelerometer data.
  - Contracted article.
- 2018 ● **The United States of Seasons<sup>10</sup>**  
[LiveFreeOrDichotomize.com](#)
  - GIS analysis of weather data to find the most 'seasonal' locations in United States
  - Used Bayesian regression methods for smoothing sparse geospatial data.
- 2017 ● **A year as told by fitbit<sup>11</sup>**  
[LiveFreeOrDichotomize.com](#)
  - Analyzing a full years worth of second-level heart rate data from wearable device.
  - Demonstrated visualization-based inference for large data.
- 2017 ● **MCMC and the case of the spilled seeds<sup>12</sup>**  
[LiveFreeOrDichotomize.com](#)
  - Full Bayesian MCMC sampler running in your browser.
  - Coded from scratch in vanilla Javascript.
- 2017 ● **The Traveling Metallurgist<sup>13</sup>**  
[LiveFreeOrDichotomize.com](#)
  - Pure javascript implementation of traveling salesman solution using simulated annealing.
  - Allows reader to customize the number and location of cities to attempt to trick the algorithm.
- I regularly blog about data science and visualization on my blog [LiveFreeOrDichotomize](#).<sup>7</sup>

## SELECTED PRESS (ABOUT)

- 2017 | 2017
- **Great paper? Swipe right on the new Tinder for preprints app<sup>14</sup>**  
Science
    - Story of the app Papr<sup>15</sup> made with Jeff Leek and Lucy D'Agostino McGowan.

2017 | 2017

  - **Swipe right for science: Papr app is Tinder for preprints<sup>16</sup>**  
Nature News
    - Second press article for app Papr.

2016 | 2016

  - **The Deeper Story in the Data<sup>17</sup>**  
University of Vermont Quarterly
    - Story on my path post graduation and the power of narrative.

## SELECTED PRESS (BY)

- 2016 | 2016
- **The Great Student Migration<sup>18</sup>**  
The New York Times
    - Most shared and discussed article from the New York Times for August 2016.

2016 | 2016

  - **Wildfires are Getting Worse, The New York Times<sup>19</sup>**  
The New York Times
    - GIS analysis and modeling of fire patterns and trends
    - Data in collaboration with NASA and USGS

2016 | 2016

  - **Who's Speaking at the Democratic National Convention?<sup>20</sup>**  
The New York Times
    - Data scraped from CSPAN records to figure out who talked and past conventions.

2016 | 2016

  - **Who's Speaking at the Republican National Convention?<sup>21</sup>**  
The New York Times
    - Used same data scraping techniques as Who's Speaking at the Democratic National Convention?

2016  
|  
2016

- **A Trail of Terror in Nice, Block by Block<sup>22</sup>**  
The New York Times
  - Led research effort to put together story of 2016 terrorist attack in Nice, France in less than 12 hours.
  - Work won Silver medal at Malofiej 2017, and gold at Society of News and Design.

## SELECTED PUBLICATIONS, POSTERS, AND TALKS

2020

- **Building a software package in tandem with machine learning methods research can result in both more rigorous code and more rigorous research**  
ENAR 2020
  - Invited talk in Human Data Interaction section.
  - How and why building an R package can benefit methodological research

2020

- **Stochastic Block Modeling in R, Statistically rigorous clustering with rigorous code<sup>23</sup>**  
RStudio::conf 2020
  - Invited talk about new sbmR package<sup>24</sup>.
  - Focus on how software development and methodological research can improve both benefit when done in tandem.
- **PheWAS-ME: A web-app for interactive exploration of multimorbidity patterns in PheWAS<sup>25</sup>**  
Bioinformatics
  - Manuscript detailing application for the exploration of multimorbidity patterns in PheWAS analyses
  - See landing page<sup>26</sup> for more information.

2019  
|  
2019

- **Charge Reductions Associated with Shortening Time to Recovery in Septic Shock<sup>27</sup>**  
Chest
  - Authored with Wesley H. Self, MD MPH; Dandan Liu, PhD; Stephan Russ, MD, MPH; Michael J. Ward, MD, PhD, MBA; Nathan I. Shapiro, MD, MPH; Todd W. Rice, MD, MSc; Matthew W. Semler, MD, MSc.
- **Multimorbidity Explorer | A shiny app for exploring EHR and biobank data<sup>28</sup>**  
RStudio::conf 2019
  - Contributed Poster. Authored with Yaomin Xu.
- **Taking a network view of EHR and Biobank data to find explainable multivariate patterns<sup>29</sup>**  
Vanderbilt Biostatistics Seminar Series
  - University wide seminar series.

- 2019
- **Patient-specific risk factors independently influence survival in Myelodysplastic Syndromes in an unbiased review of EHR records**  
Under-Review (copy available upon request.)
    - Bayesian network analysis used to find novel subgroups of patients with Myelodysplastic Syndromes (MDS).
    - Analysis done using method built for my dissertation.
- 2019
- **Patient specific comorbidities impact overall survival in myelofibrosis**  
Under-Review (copy available upon request.)
    - Bayesian network analysis used to find robust novel subgroups of patients with given genetic mutations.
    - Analysis done using method built for my dissertation.
- 2018  
|  
2018
- **R timelineViz: Visualizing the distribution of study events in longitudinal studies**  
Under-Review (copy available upon request.)
    - Authored with Alex Sunderman of the Vanderbilt Department of Epidemiology.
- 2017  
|  
2017
- **Continuous Classification using Deep Neural Networks<sup>30</sup>**  
Vanderbilt Biostatistics Qualification Exam
    - Review of methods for classifying continuous data streams using neural networks
    - Successfully met qualifying examination standards
- 2015  
|  
2015
- **Asymmetric Linkage Disequilibrium: Tools for Dissecting Multiallelic LD**  
Journal of Human Immunology
    - Authored with Richard Single, Vanja Paunic, Mark Albrecht, and Martin Maiers.
- 2015  
|  
2015
- **An Agent Based Model of Mysis Migration<sup>31</sup>**  
International Association of Great Lakes Research Conference
    - Authored with Brian O'Malley, Sture Hansson, and Jason Stockwell.
- 2015  
|  
2015
- **Declines of Mysis diluviana in the Great Lakes**  
Journal of Great Lakes Research
    - Authored with Peter Euclide and Jason Stockwell.

## 🔗 LINKS

- 1: <https://www.nytimes.com/interactive/2016/08/26/us/college-student-migration.html>
- 2: [https://livefreeordichotomize.com/2019/06/04/using\\_awk\\_and\\_r\\_to\\_parse\\_25tb/](https://livefreeordichotomize.com/2019/06/04/using_awk_and_r_to_parse_25tb/)
- 3: <https://github.com/nstrayer/shinysense>
- 4: <http://nickstrayer.me/dataDayTexas/>
- 5: <https://sites.utexas.edu/baker-lab/>
- 6: [http://nickstrayer.me/js4shiny\\_r2d3/slides](http://nickstrayer.me/js4shiny_r2d3/slides)

- 7: <https://livefreeordichotomize.com/>
- 8: [https://livefreeordichotomize.com/2019/06/04/using\\_awk\\_and\\_r\\_to\\_parse\\_25tb/](https://livefreeordichotomize.com/2019/06/04/using_awk_and_r_to_parse_25tb/)
- 9: <https://blogs.rstudio.com/tensorflow/posts/2018-07-17-activity-detection/>
- 10: <https://livefreeordichotomize.com/2018/02/12/the-united-states-of-seasons/>
- 11: <https://livefreeordichotomize.com/2017/12/27/a-year-as-told-by-fitbit/>
- 12: <https://livefreeordichotomize.com/2017/10/14/mcmc-and-the-case-of-the-spilled-seeds/>
- 13: <https://livefreeordichotomize.com/2017/09/25/the-traveling-metallurgist/>
- 14: <https://www.sciencemag.org/news/2017/06/great-paper-swipe-right-new-tinder-preprints-app>
- 15: <https://jhubiostatistics.shinyapps.io/papr/>
- 16: <https://www.nature.com/news/swipe-right-for-science-papr-app-is-tinder-for-preprints-1.22163>
- 17: <https://www.uvm.edu/uvmnews/news/deeper-story-data>
- 18: <https://www.nytimes.com/interactive/2016/08/26/us/college-student-migration.html?smid=pl-share>
- 19: <https://www.nytimes.com/interactive/2016/07/25/us/wildfire-seasons-los-angeles.html>
- 20: <https://www.nytimes.com/2016/07/26/upshot/democrats-may-not-be-unified-but-their-convention-speakers-are.html>
- 21: <https://www.nytimes.com/2016/07/19/upshot/whos-not-speaking-how-this-republican-convention-differs.html?smid=pl-share>
- 22: <https://www.nytimes.com/interactive/2016/07/14/world/europe/trail-of-terror-france.html>
- 23: [http://nickstrayer.me/rstudioconf\\_sbm](http://nickstrayer.me/rstudioconf_sbm)
- 24: <https://tbilab.github.io/sbmR/>
- 25: <https://academic.oup.com/bioinformatics/advance-article-abstract/doi/10.1093/bioinformatics/btaa870/5922817?redirectedFrom=fulltext>
- 26: [https://prod.tbilab.org/phewas\\_me\\_info/](https://prod.tbilab.org/phewas_me_info/)
- 27: <https://www.ncbi.nlm.nih.gov/pubmed/30419234>
- 28: [http://nickstrayer.me/rstudioconf19\\_me-poster/](http://nickstrayer.me/rstudioconf19_me-poster/)
- 29: [http://nickstrayer.me/biostat\\_seminar/](http://nickstrayer.me/biostat_seminar/)
- 30: [http://nickstrayer.me/qualifying\\_exam/](http://nickstrayer.me/qualifying_exam/)
- 31: <https://www.semanticscholar.org/paper/An-Agent-Based-Model-of-the-Diel-Vertical-Migration-Strayer-Stockwell/40493c78e8ecf22bd882d17ec99fd913ec4b9820>