# Chase Clark

University of Wisconsin-Madison, School of Pharmacy

chasemc ChasingMicrobes chasingmicrobes@gmail.com hirechase chasemc.github.io

## Education

| | | |
|---|---|---|
| 2015 | **PhD Pharmacognosy**<br>Chicago, IL | University of Illinois Chicago |
| 2008 | **BS Biochemistry**<br>Mount Berry, GA | Berry College |
| 2010 | **Principia Consortium Study Abroad**<br>Glasgow, Glasgow | University of Glasgow |

## Employment (Only relevant positions listed)

Jun 2021 - present **Computation and Informatics in Biology and Medicine (CIBM) Postdoctoral Fellow** — UW-Madison
Remote, Chicago, IL

Sep 2020 - present **Postdoctoral Research Associate** — University of Wisconsin-Madison
Remote, Chicago, IL

May 2019 - Aug 2020 **NIH F31 Fellow** — University of Illinois at Chicago
Chicago, IL, US

Aug 2015 - Aug 2020 **Ph.D. Student** — University of Illinois at Chicago
Chicago, IL, US

- Natural product drug discovery, specifically bacterial-derived antibiotics
- Conceived, learned to code, and built a bioinformatics program (R Shiny app) for rapid bacterial metabolomics analyses
- Isolated and cultured 1000's of bacterial isolates and performed natural product isolation and structure elucidation

Mar 2013 - Aug 2015 **Research and Development Technician** — Deerland Probiotics & Enzymes
Kennesaw, GA, US

- Head of method development and identification of raw materials to meet CFR 21, FDA requirements.
- Assisted in the design and results analysis of three clinical studies
- Presented monthly projects to CEO, VP Sales and VP Science & Technology
- Developed identifications for over 140 materials, from enzymes and botanicals to minerals, additives, and probiotics.
- Overhauled the Method Development Department establishing inter-departmental responsibilities, improving and creating new SOPs, streamlining identification requests by establishing a formal request process and interactive database
- Strengthened and created new relationships with third-party vendors, labs, and universities.
- Controlled budget of method development for purchase of standards, consumables, equipment and third-party testing.
- Mentored interns from Kennesaw Mountain High School magnet program

Aug 2008 - May 2012 **Bonner Scholar** — Bonner Foundation
Rome, GA, US

- Four-year community service scholarship requiring weekly and summer commitments.

## Data Science (select, not comprehensive)

**Bioinformatics**: metabolomics () • genomics () • metagenomics ()
**Cloud Computing**: remote server computing () • high-performance and high-throughput computing (Open science grid, HTCondor, AWS, etc)
**Communication**: presentations (public speaking, technical writing,peer-reviewed publications, SOPs) • writing (customer-facing documents, peer-reviewed publications, SOPs) • reports (Rmarkdown, Jupyter) • visualization (Base R, ggplot2, plotly, leaflet, cytoscape, networkx, gephi, etc)
**Development**: full-stack design (Flask, Django, Shiny, Dash)
**Programming**: general (R, Python) • database (Neo4j, SQL) • automation (Nextflow, Make) • containers (Docker) • containers (Singularity) • pipeline (Nextflow, Make)
**Software Development**: source control (Git) • CI/CD/DevOps (Github, Travis, Azure, etc.) • automated testing ()
**Statistics**: machine learning (scikit-learn, keras, tensorflow, Neo4j, etc) • experimental design (DOE, etc)

## Lab Science (select, not comprehensive)

- **microbial genomics and transcriptomics**
- **targeted and untargeted small molecule mass spectrometry and NMR**
- **biochemistry and biosynthetic gene cluster discovery and characterization**
- **method development**

## Select Individual Software Projects

| | | |
|---|---|---|
| 2021-2023 | **SocialGene**<br>Repository-scale genomic graph database | Python, Django, Nextflow, Rust, High-throughput computing |
| 2019 | **electricShine**<br>Shiny packaging with Electron | R, Javascript, Node/Electron |
| 2019 | **mzEasy**<br>Shiny app for converting and visualizing mass spectrometry data | R, Shiny |
| 2019 | **mzPlotter**<br>Automated summary of of LC-MS/MS data with interactive Rmarkdown reports | R, Rmarkdown |
| 2019 | **mzFromImage**<br>Predict a mass spectrum's values from a static image file | R, Shiny |
| 2019 | **mgfparse**<br>Low-dependency R package for quickly and efficiently parsing mgf files into R | R |
| 2016-2023 | **IDBac**<br>MALDI protein and small molecule bioinformatics platform | R, Shiny, Node/Electron |

## Select Team Software Projects

| | | |
|---|---|---|
| 2024-present | **code4np.github.io**<br>Co-creator of and contributor to a blog on computational pharmacognosy | R, Python, Quarto |
| 2021 | **metaBenchmarks**<br>Benchmark metagenomic profiling/binning software | Nextflow |
| 2020-2023 | **Autometa**<br>Automated binning pipeline for single metagenomes | Python, Nextflow |

## Talks

2024-03-14 **Gordon Research Conference on Marine Natural Products**
- A Multi-Repository Scale Genomic and Chemical Search Engine to Enable the Discovery, Production, and Function of Natural Products (selected GRS talk)

2024-03-10 **Gordon Research Seminar (GRS) on Marine Natural Products**
- A Multi-Repository Scale Genomic and Chemical Search Engine to Enable the Discovery, Production, and Function of Natural Products

2023-09-21 **Chicago R User Group**
- How to Write Reproducible Code for Data Science

2023-08-10 **Chicago Python User Group**
- Reproducibly Building and Analyzing Knowledge Graphs for Drug Discovery with Nextflow, Neo4j and Python

2023-03-07 **Computation and Informatics in Biology and Medicine Seminars**
- SocialGene: Large Scale Knowledge Graphs for Microbial Based Drug Discovery

2023-01-30 **Northeastern University Department of Chemistry and Chemical Biology**
- Information-Rich Platforms for Natural Product Antibiotic Drug Discovery and Microbial Characterization

2023-01-30 **2023 ASP Younger Members Symposium**
- SocialGene: A Large Scale Search Engine to Find Metagenomic BGCs in Free-Living Organisms

2022-06-23 **National Library of Medicine (NLM) T15 Training Conference**
- Large Scale Analysis of Protein Homology for Microbial Drug Discovery

2022-03-09 **Gordon Research Conference on Marine Natural Products**
- Towards Near-Instant, Repository-Scale Searching for Homologous BGCs with Socialgene (Selected alternate speaker, didn't present)

2021-11-02 **Computation & Informatics in Biology & Medicine Seminars BMI 915**
- New Computational Tools for Top-Down and Bottom-Up Natural Product Drug Discovery

2020-06-01 **St. Jude National Graduate Student Symposium (cancelled due to Covid19)**
- IDBac: Bioinformatics Software for Microbial Drug Discovery Prioritization and Culturomics Characterization

2020-05-08 **May Institute 2020: Future developers meeting**
- Protein MS isn't the only MS… Programming in R for Metabolomics Mass Spectrometry

2020-02-25 **Gordon Research Conference on Marine Natural Products**
- IDBac: Bioinformatics Software for Microbial Drug Discovery Prioritization and Culturomics Characterization

2020-02-23 **Gordon Research Seminar (GRS) on Marine Natural Products**
- IDBac: Bioinformatics Software for Microbial Drug Discovery Prioritization and Culturomics Characterization (selected GRS talk)

2019-08-23 **R/Pharma, Harvard University**
- Your Missing Step in Reproducible R Programming: Continuous Deployment

2019-04-27 **satRdays Chicago**
- Bioinformatics in R

2019-01-23 **Chicago R User Group**
- Single Function Lightning Talks: "lengths(), not length()"

2018-08-15 **R/Pharma, Harvard University**
- IDBac: A New Paradigm in Developing Microbial Libraries for Drug Discovery

2018-01-27 **Chicago R User Group**
- IDBac: A Shiny App to Analyze Bacterial Fingerprints and Aid in the Discovery of Potential New Antibiotics

2017-03-09 **Center for Biomolecular Sciences, UIC**
- IDBac: A Proteomic & Chemometric Pipeline for Rapid Bacterial Characterization

2016-11-03 **UIC Specialized Metabolite Community**
- Rational Design of Bacterial Strain Libraries for Drug Discovery

2016-07-22 **Chicago Mass Spec Day**
- Use of MALDI-MS to Create 'Smart' Libraries for Drug-Lead Discovery

2015-04-15 **SSAOAC Annual Meeting**
- HPLC-DAD Method for Trace Detection of Benzoic and Sorbic Acids in High Protein Matrices

## Awards

2021-present **Computation and Informatics in Biology and Medicine (CIBM) Training Program (T32)**
$100,000.00

2020 **St. Jude National Graduate Student Symposium (cancelled due to Covid19)**
One of 40 participants selected from 1,000 invite-only applications

2019-2020 **NIH Predoctoral Individual National Research Service Award (F31)**
$90,000.00

2018 **R/Pharma, Travel Award**
$500

2017 **W.E. van Doren Scholar**
$1,000

2017 **American Society of Pharmacognosy Student Travel Award**
$600

2017 **University of Illinois at Chicago Provost/Deiss Award**
$2,500

2008-2012 **Bonner Scholar**
Scholarship for students with financial need "in exchange for weekly commitment to intensive and meaningful service with a local community organization over the four years"

2008-2012 **Academic Scholarship**

2011 **McCaleb, Hubert Scholarship**

2011 **Frank Plummer Scholarship**

2010 **Griggs Academic Scholarship**

2009 **John R. Bertrand Scholarship**

2008 **Sierra Club College Scholarship**

2008 **Cherokee Area Eagle Scout of the Year Scholarship**

## Publications

1. Rees, E. R., Uppal, S., Clark, C. M., Lail, A. J., Waterworth, S. C., Roesemann, S. D., Wolf, K. A., & Kwan, J. C. (2023). Autometa 2: A versatile tool for recovering genomes from highly-complex metagenomic communities. *Citations: 0.* 10.1101/2023.09.01.555939
2. Clark, C. M., Hernandez, A., Mullowney, M. W., Fitz-Henley, J., Li, E., Romanowski, S. B., Pronzato, R., Manconi, R., Sanchez, L. M., & Murphy, B. T. (2022). Relationship between bacterial phylotype and specialized metabolite production in the culturable microbiome of two freshwater sponges. *ISME Communications. Citations: 5*. 10.1038/s43705-022-00105-8
3. Clark, C. M., Nguyen, L., Pham, V. C., Sanchez, L. M., & Murphy, B. T. (2022). Automated microbial library generation using the bioinformatics platform IDBac. *Molecules. Citations: 2.* 10.3390/molecules27072038
4. Leao, T. F., Clark, C. M., Bauermeister, A., Elijah, E. O., Gentry, E. C., Husband, M., Oliveira, M. F., Bandeira, N., Wang, M., & Dorrestein, P. C. (2021). Quick-start infrastructure for untargeted metabolomics analysis in GNPS. *Nature Metabolism. Citations: 12.* 10.1038/s42255-021-00429-0
5. Elfeki, M., Mantri, S., Clark, C. M., Green, S. J., Ziemert, N., & Murphy, B. T. (2021). Evaluating the distribution of bacterial natural product biosynthetic genes across lake huron sediment. *ACS Chemical Biology. Citations: 4*. 10.1021/acschembio.1c00653

6. Clark, C. M., Murphy, B. T., & Sanchez, L. M. (2020). A call to action: The need for standardization in developing open-source mass spectrometry-based methods for microbial subspecies discrimination. *mSystems. Citations: 2.* 10.1128/msystems.00813-19
7. Clark, C. M., Costa, M. S., Conley, E., Li, E., Sanchez, L. M., & Murphy, B. T. (2019). Using the open-source MALDI TOF-MS IDBac pipeline for analysis of microbial protein and specialized metabolite data. *Journal of Visualized Experiments. Citations: 9.* 10.3791/59219
8. Braesel, J., Clark, C. M., Kunstman, K. J., Green, S. J., Maienschein-Cline, M., Murphy, B. T., & Eustáquio, A. S. (2019). Genome sequence of marine-derived streptomyces sp. Strain F001, a producer of akashin a and diaza-quinomycins. *Microbiology Resource Announcements. Citations: 2.* 10.1128/mra.00165-19
9. Costa, M. S., Clark, C. M., Ómarsdóttir, S., Sanchez, L. M., & Murphy, B. T. (2019). Minimizing taxonomic and natural product redundancy in microbial libraries using MALDI-TOF MS and the bioinformatics pipeline IDBac. *Journal of Natural Products. Citations: 16.* 10.1021/acs.jnatprod.9b00168
10. Clark, C. M., Costa, M. S., Sanchez, L. M., & Murphy, B. T. (2018). Coupling MALDI-TOF mass spectrometry protein and specialized metabolite analyses to rapidly discriminate bacterial function. *Proceedings of the National Academy of Sciences. Citations: 66.* 10.1073/pnas.1801247115