

Curriculum Vitae - Dr. Milot Mirdita

Laboratory of Machine Learning and Bioinformatics
Seoul National University
1 Gwanak-ro, Gwanak-gu, 08826 Seoul, Korea
mmirdit@snu.ac.kr

Research interests

Open-source methods for metagenomic analysis, fast and sensitive homology searches, clustering, taxonomy, function and structure prediction

Education

07/2017-02/2022 **Dr. rer. nat.** (summa cum laude). Advisor: Johannes Söding,
Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany
University of Göttingen, Göttingen, Germany
04/2014–08/2016 **M.Sc.** in Computer science, LMU, Munich, Germany
10/2010–03/2014 **B.Sc.** in Bioinformatics, LMU & TUM, Munich, Germany

Research experience

Since 07/2022 **Postdoctoral fellow** at Steinegger Lab
Seoul National University, Seoul, Korea
03/2022-05/2022 **Researcher** at Söding Lab
Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany
07/2017-02/2022 **PhD candidate** at Söding Lab
Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany
04/2014-06/2017 **Research engineer** at Söding Lab
Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany
Gene Center Munich, Munich, Germany
07/2011-05/2012 **Research assistant** at Rost Lab (part time)
Technical University of Munich, Munich, Germany

Industry experience

06/2011-04/2016 **Software engineer** (part time)
SpinSoft IT Solutions GmbH, München, Germany

Publications

van Kempen M, Kim S, Tumescheit C, **Mirdita M**, Gilchrist CLM, Soeding J, Steinegger M (2022)
Foldseek: fast and accurate protein structure search. bioRxiv, 2022.02.07.479398
Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S and Steinegger M (2022)
ColabFold – Making protein folding accessible to all. Nature Methods, 19, 679–682
(ISSN 1548-7091, IF 28.467)
Mirdita M, Steinegger M, Breitwieser F, Söding J and Levy Karin E (2021)
Fast and sensitive taxonomic assignment to metagenomic contigs. Bioinformatics,
37, 3029–3031 (ISSN 1367-4803, IF 6.937)

- Zhang R, **Mirdita M**, Levy Karin E, Norroy C, Galiez C and Söding J (2021)
SpacePHARER: Sensitive identification of phages from CRISPR spacers in prokaryotic hosts. *Bioinformatics* 37, 3364–3366 (ISSN 1367-4803, IF 6.937)
- Zhao B, [...], **Mirdita M**, [...], Kurgan L (2021)
DescribePROT: database of amino acid-level protein structure and function predictions. *Nucleic Acids Research* 49 (D1), D298-D308 (ISSN 0305-1048, IF 16.971)
- Bernhofer M, Dallago C, [...], **Mirdita M**, [...], Rost B (2021)
PredictProtein - Predicting Protein Structure and Function for 29 Years. *Nucleic Acids Research* 49 (W1), W535–W540 (ISSN 0305-1048, IF 16.971)
- Aevansson A, [...], **Mirdita M**, [...] (2021)
Going to extremes – a metagenomic journey into the dark matter of life. *FEMS Microbiology Letters* 368 (12), fnab067 (ISSN 0378-1097, IF 2.742)
- Gabler F, Nam SZ, S Till, **Mirdita M**, Steinegger M, Söding J, Lupas AN, Alva V (2020)
Protein sequence analysis using the MPI bioinformatics toolkit. *Current Protocols in Bioinformatics* 72 (1), e108 (ISSN 1934-3396, IF 9.231)
- Levy Karin E, **Mirdita M** and Söding J (2020)
MetaEuk – sensitive, high-throughput gene discovery and annotation for large-scale eukaryotic metagenomics. *Microbiome*, 8, 48. (ISSN: 2049-2618, IF 14.652)
- Steinegger M, Meier M, **Mirdita M**, Vöhringer H, Haunsberger SJ and Söding J (2019)
HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics*, 20, 473. (ISSN: 1471-2105, IF 3.169)
- Mirdita M**, Steinegger M and Söding J (2019)
MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics*, 35, 2856–2858. (ISSN 1367-4803, IF 6.937)
- Steinegger M, **Mirdita M** and Söding J (2019)
Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods*, 16, 603–606. (ISSN 1548-7091, IF 28.467)
- Keasar C, [...], **Mirdita M**, [...], Baker D (2018)
An analysis and evaluation of the WeFold collaborative for protein structure prediction and its pipelines in CASP11 and CASP12. *Scientific Reports* 8 (1), 1-18 (ISSN 2045-2322, IF 4.380)
- Mirdita M**, von den Driesch L, Galiez C, Martin M, Söding J, Steinegger M (2017)
Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Research* 45 (D1), D170-D176 (ISSN 0305-1048, IF 16.971)

Selected Talks (International Meetings)

- Nov 2021 Fast and sensitive taxonomic assignment to metagenomic contigs.
SNU Online Symposium on Bioinformatics for metagenomic analysis, Virtual
- July 2021 Fast and sensitive taxonomic assignment to metagenomic contigs.
ISMB/ECCB 2021, Virtual

Selected Posters (International Meetings)

Hyunjoo J, **Mirdita M**, Sommer HG, Galiez C, Söding J and Steinegger M.

MMseqs2 profile/profile: fast and ultra sensitive searches beyond the twilight zone.
ISMB/ECCB 2021, July 26-30, 2021, Virtual

Steinegger M, **Mirdita M** and Söding J.

New algorithms and tools for large-scale sequence analysis of metagenomics data.
Genome Informatics, November 6-9, 2019, Cold Spring Harbor, USA

Levy Karin E, **Mirdita M** and Söding J.

MetaEuk – sensitive, high-throughput gene discovery and annotation for large-scale eukaryotic metagenomics. Microbiome COSI. ISMB/ECCB. July 21-25, 2019, Basel, Switzerland

Mirdita M, Steinegger M and Söding J. (Poster award)

MMseqs2 desktop and local web server app for fast, interactive sequence searches.
September 7-12, 2018, ECCB, Athens, Greece

Teaching & Workshops

2017-2022 (Co-)supervised multiple interns, B.Sc. and M.Sc. students

Sept. 2020 Workshop Deep dive into metagenomics using MMseqs2. ECCB 2020, Virtual.

Sept. 2018 Workshop Modern and scalable tools for the efficient analysis of very large metagenomic datasets. ECCB 2018, Athens, Greece

2017-2020 Yearly 1.5-day tutorial on protein structure prediction, homology search and metagenomics analysis in the International M.Sc. Curriculum for Molecular Biology, University of Göttingen, Germany

Awards, Fellowships and Achievements

2018 Best poster award, ECCB 2018, Athens, Greece

2016 Best poster award, CASP 2016, Gaeta, Italy

2016 Stefan-Hell scholarship, Max-Planck-Gesellschaft

2012 Hardware donations for HPC, NVIDIA and AMD

2011 Amazon AWS in Education, research grant, Amazon

Open source software

Foldcomp github.com/steineggerlab/foldcomp

ColabFold github.com/sokrypton/ColabFold

Foldseek github.com/steineggerlab/foldseek

Metaeuk github.com/soedinglab/metaeuk

Spacepharer github.com/soedinglab/spacepharer

Plass github.com/soedinglab/plass

MMseqs2 github.com/soedinglab/mmseqs2

HH-suite github.com/soedinglab/hh-suite