

Nationale Forschungsdateninfrastruktur für Mikrobiota- Forschung (NFDI4Microbiota)

Sprecher: Prof. Dr. Konrad Förstner, ZB MED - Informationszentrum
Lebenswissenschaften, foerstner@zbmed.de

<https://doi.org/10.5281/zenodo.3897525>



Mikrobiota, die ökologische Gemeinschaft mikrobieller Spezies (Bakterien, Archaeen, einzellige Eukaryonten und Viren), haben einen starken Einfluss auf zahlreiche Aspekte des menschlichen Lebens, angefangen von der individuellen Gesundheit bis hin zum Klimawandel. Angesichts umfassender globaler Probleme muss sich unser Verständnis der Ökosysteme massiv verbessern, damit wir in der Lage sind, den vom Menschen verursachten Herausforderungen besser zu begegnen. Dabei wurden Mikroben bisher weitgehend ignoriert¹. Eine der größten medizinischen Bedrohungen ist laut UNO die Zunahme von Antibiotikaresistenzen bakterieller Erreger² und die momentane COVID-19-Pandemie zeigt die dramatischen Auswirkungen von viralen Infektionen auf alle Bereiche unseres Lebens. Beide scheinbar voneinander abgekoppelten wichtigen Themen erfordern ein besseres Verständnis der Welt der Mikroorganismen. Darüber hinaus sind unzählige, noch nicht kultivierte Arten potenzielle Quellen für Substanzen mit Relevanz für Biotechnologie und Medizin, die noch erforscht werden müssen. Eine der größten Herausforderungen für das Verständnis der Mikrobiota liegt in der Komplexität der zahlreichen biotischen Wechselwirkungen zwischen spezifischen Organismen eines Mikrobioms und abiotischen Umweltfaktoren. Die Erforschung einzelner Arten und ganzer Mikrobiome, insbesondere die Kartierung und Entschlüsselung der molekularen Wechselwirkungen mit den ihnen zugrunde liegenden Regulationsmechanismen, ist ein entscheidender Schritt zum Verständnis und zur effektiveren Nutzung mikrobieller Arten.

Forschungsbereiche wie z.B. Biomedizin, Landwirtschaft, Meeresforschung und Biotechnologie verfolgen sehr unterschiedliche wissenschaftliche Fragestellungen, teilen jedoch die Notwendigkeit einer effizienten Speicherung, Dokumentation und

¹ Cavicchioli et al., Nature Review Microbiology, 2019, <https://doi.org/10.1038/s41579-019-0222-5>

² <https://www.who.int/who-un/about/amr/en/>

Analyse von Daten zu mikrobiellen Organismen, Mikrobiomen und der molekularen Interaktion ihrer Mitglieder. Während Hochdurchsatzmethoden leicht eine Fülle von Omics-Daten verschiedener molekularer Ebenen (DNA, RNA, Proteine, Metabolite) liefern können, um z.B. die Funktionsweise der einzelnen Arten und Mikroorganismen sowie ihre Beiträge zu menschlichen Gesundheits- und Umweltprozessen zu entschlüsseln, ist die Speicherung, Analyse und Integration dieser Daten mit verfügbarem Wissen (z.B. aus der Literatur) nicht trivial. Häufig verfügen Forscher nicht über die Expertise und/oder Ressourcen, um mit solchen durch Hochdurchsatztechnologien erzeugten Daten umzugehen. Weiterhin ergeben sich aus der zunehmenden Größe sowie der Komplexität der Daten und den erforderlichen Verarbeitungs- und Archivierungsschritten erhebliche Anforderungen.

Das Ziel des NFDI4Microbiota-Konsortiums ist es, die Speicherung, Integration und Analyse von Omics-Daten für die Erforschung mikrobieller Arten und diverser Mikrobiome konsistent, reproduzierbar und für alle Bereiche der Biowissenschaften zugänglich zu machen. Das Konsortium wird Forscherinnen und Forscher mit unterschiedlichen wissenschaftlichen Fragestellungen dabei unterstützen, einzelne mikrobielle Arten und Gemeinschaften sowie die Interaktion zwischen den Arten in ihnen zu verstehen und eigene Daten zugänglich und nachnutzbar zu machen. Zu diesem Zweck wird NFDI4Microbiota IT-Infrastruktur, Analysewerkzeuge und Training für die Gemeinschaft bereitstellen. Zudem sollen neue Standards gemeinsam mit internationalen Partnern entwickelt werden. Das Konsortium wird somit eine effiziente und reproduzierbare Verarbeitung von Omics-Daten ermöglichen, die mit Hochdurchsatz-Analysegeräten erzeugt werden, und dadurch helfen, mikrobielle Arten und ihre Interaktionen auf molekularer Ebene zu entschlüsseln. Zu den Daten gehören Genome, Transkriptome, Proteome und Metabolome für einzelne Spezies sowie die Analoga für ganze mikrobielle Gemeinschaften: Metagenome, Meta-Transkriptome, Meta-Proteome und meta-metabolomische Daten. Darüber hinaus werden derzeit neu entstehende Analyseansätze für Daten aus der Einzelzellsequenzierung und Hochdurchsatz-Bildgebung unterstützt. Zusätzlich wird das Konsortium die Daten interoperabel machen und die Anreicherung dieser Daten durch Metadaten aus Analysegeräten und Datenbanken sowie durch automatisch aus der Literatur extrahiertem Wissen ermöglichen. Um einen nahtlosen Daten-Workflow zu gewährleisten, wird NFDI4Microbiota mit Datengeneratoren wie Sequenzierereinrichtungen zusammenarbeiten und eine direkte Ablage der Messdaten und der damit verbundenen Metadaten in der vom Konsortium bereitgestellten Forschungsdateninfrastruktur fördern. Die Daten werden entsprechend der Wünsche der Forschenden, die die Analyse in Auftrag gegeben haben, durch Analyse-Pipelines ausgewertet. Rohdaten, Metadaten und die Ergebnisse werden zur langfristigen Verfügbarkeit in standardisierter Form in passgenauen Repositorien abgelegt.

NFDI4Microbiota wird die FAIR-Prinzipien (Findable, Accessible, Interoperable, Re-usable) vollständig erfüllen und offene Wissenschaft mit all ihren Facetten fördern. Sensible persönliche Daten werden mit der notwendigen Sorgfalt behandelt. Das Konsortium wird die erforderlichen Spezifikationen zur reichhaltigen Erfassung von Metadaten definieren, die die Proben-Bedingungen beschreiben und eine standardisierte und qualitätskontrollierte Einreichung von Daten erlauben. Die Mitglieder unseres Konsortiums sind federführend an der Entwicklung der International Human Microbiome Standards (IHMS) beteiligt, und NFDI4Microbiota wird zusammen mit der Gemeinschaft ähnliche Standards für andere Mikrobiomquellen fördern. Darüber hinaus wird das Konsortium die Teilnehmer ermutigen, für die eingereichten Datensätze möglichst liberale Lizenzen zu wählen, um rechtliche Hindernisse für die gemeinsame Nutzung von Daten zu vermeiden. Die konsistente Verwaltung der eingereichten Daten sowie die reichhaltige Annotation mit Metadaten bilden den Kern für die leistungsfähige Suche nach Originaldaten und der Ergebnisse sowie für die effiziente Wiederverwendung und den Vergleich durch die Forschungsgemeinschaft. Um nachhaltige Lösungen zu generieren, wird NFDI4Microbiota der guten Softwareentwicklungspraxis folgend Software-Stacks auf der Basis von FLOSS (Free/Libre/Open Source Software) generieren. Alle entwickelte Software wird unter OSI-konformen (Open Source Initiative) Lizenzen veröffentlicht. NFDI4Microbiota wird in ständigem Austausch mit seinem wissenschaftlichen Beirat und Nutzern stehen, um Feedback zu erhalten und die Dienstleistungen an die Bedürfnisse ihrer Forschungsgemeinschaft anzupassen. Darüber hinaus werden etablierte Verbindungen mit internationalen Organisationen wie ELIXIR und dem Projekt International Human Microbiome Standards genutzt, um die Aktivitäten der NFDI4Microbiota auf diese abzustimmen und Synergien zu finden.

Das Konsortium besteht aus 10 etablierten Institutionen, die schon jetzt Infrastruktur für die Mikrobiologieforschungsgemeinschaft bereitstellen, sowie zahlreichen weiteren teilnehmenden Institutionen. Zudem steht NFDI4Microbiota mit mehreren anderen NFDI-Konsortien im engen Austausch (z.B. NFDI4Health, GHGA, NFDI4BioDiversity, NFDI4Agri, NFDI4AIRR) und plant u.a. gemeinsame Standards zu etablieren. NFDI4Microbiota ist Unterzeichner der "Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung"³ und kann zu verschiedenen Querschnittsthemen wie z.B. Training beitragen.

Wir erwarten uns von der NFDI-Konferenz 2020 weiteren Austausch mit anderen Konsortien sowie weitere Diskussion zu Querschnittsthemen.

³ <https://zenodo.org/record/3895209>

Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):

Co-Sprecher/in	Zugehörige Institution
Prof. Dr. Peer Bork Leiter der Structural and Computational Biology Unit bork@embl.de	Europäisches Laboratorium für Molekularbiologie Meyerhofstraße 1 69117 Heidelberg
Prof. Dr. Anke Becker Leiterin Microbial Comparative Genomics, Direktorin von SYNMIKRO anke.becker@synmikro.uni-marburg.de	Philipps-Universität Marburg Biegenstraße 10 35037 Marburg
Prof. Dr. Thomas Clavel Leiter der AG Funktionelle Mikrobiomforschung am Universitätsklinikum tclavel@ukaachen.de	RWTH Aachen Templergraben 55 52062 Aachen
Prof. Dr. Alexander Goesmann Leiter Systembiologie Alexander.Goesmann@Computational.Bio.Uni-Giessen.de	Justus-Liebig-Universität Giessen Ludwigstraße 23 35390 Giessen
Prof. Dr. Manja Marz Leiterin RNA Bioinformatics & High-Throughput Analysis manja@uni-jena.de	Friedrich Schiller Universität Jena Fürstengraben 1 07743 Jena
Prof. Dr. Alice McHardy Leiterin Bioinformatik der Infektionsforschung Alice.McHardy@helmholtz-hzi.de	Helmholtz-Zentrum für Infektionsforschung Inhoffenstraße 7 38124 Braunschweig
Prof. Dr. Jörg Overmann Direktor joerg.overmann@dsmz.de	Deutsche Sammlung von Mikroorganismen und Zellkulturen Inhoffenstraße 7B 38124 Braunschweig
Prof. Dr. Jens Stoye / Prof. Dr. Alexander Sczyrba Leiter Bielefelder Institut für Bioinformatik-Infrastruktur (BIBI) / Leiter Computational Metagenomics, Koordinator de.NBI Cloud jens.stoye@uni-bielefeld.de / asczyrba@techfak.uni-bielefeld.de	Universität Bielefeld Universitätsstraße 27 33615 Bielefeld
Dr. Ulisses Nunes da Rocha Leiter Microbial Systems Data Science ulisses.rocha@ufz.de	Helmholtz-Zentrum für Umweltforschung Permoserstraße 15 04318 Leipzig