

## **The Lifelines DEEP Project (“LLDEEP”): Access to Data**

### **General principles**

Funded by a Top Institute Food and Nutrition Wageningen grant GH001 the University Medical Center Groningen, the LLDEEP microbiome project (“Project” or “LLDEEP”) has deeply phenotyped 1500 participants of the Dutch Lifelines cohort. The Project is led by a consortium (hereinafter also referred to as the Consortium) from the University Medical Center Groningen (“UMCG”).

The primary goal of the LLDEEP microbiome project is to get insight in the relations between microbiome variation, genetic variation, metabolic and phenotype variation. The Consortium is willing to consider applications from third party researchers for access to the anonymized sequence data generated by the Project.

Access to Project Data will be granted to qualified researchers for approved use and will be governed by the provisions laid out in the informed consent of the LLDEEP cohort, the Data Access Procedure and the terms of the Data Access Agreement. A qualified researcher refers to a senior investigator who is employed or legitimately affiliated with an academic, non-profit or government institution and who has a track record in the field.

Access to Project Data is available by application to the LLDEEP Data Access Committee. Researchers granted access to project data must feedback the results of their research to LLDEEP, after publication in accordance with the LLDEEP publication policy set forth in the Data Access Agreement. Access is conditional upon availability of data and on signed agreement by the researcher(s) and the responsible employing institution to abide by the policies and conditions related to publication, data ownership, data return, intellectual property rights, data disposal, ethical approval, confidentiality and commercialization referred herein.

### **Description of the available data**

This study includes microbiome sequencing data for stool samples from 1,135 LifeLines-DEEP participants (474 males and 661 females) from the general population of the northern part of the Netherlands. The cohort comprised predominantly Dutch participants; 93.7% had both parents born in the Netherlands. The study contains four different datasets: (1) MGS sequencing data: the gut microbiome was analyzed using paired-end shotgun sequencing on a HiSeq2000, generating an average of 3.0Gb of data (about 32.3 million reads) per sample; (2) 16S rRNA sequencing data: hypervariable region V4 of 16S rRNA gene was amplified using forward primer 515F [GTGCCAGCMGCCGCGGTAA] and reverse primer 806R [GGACTACHVGGGTWTCTAAT]. The region was sequenced using Illumina MiSeq paired-ends; (3) Age data per sample; and (4) Gender data per sample. The detailed methods please see the description in the publications.

Access to data is available by application to the Lifelines Data Access Committee. Access to data will be granted to qualified investigators for appropriate use.

### **Related Publications:**

PMID: 27126040

PMID: 26358192

### **Data Access Committee**

Applications for access to Project Data must be made to the LLDEEP Data Access Committee (DAC). The LLDEEP DAC comprises representatives of the members of the LLDEEP Consortium.

## Data Access Procedure

### 1. Application for Access

The handling of applications, the issue of Project Data and any associated operations, administration and audits will be performed by the UMCG on behalf of the Consortium. Applications for Access to Project Data must be submitted using the Application Form for Access to LLDEEP Data. The information disclosed in the application will be treated as confidential and will only be disclosed to the persons evaluating the application. All incoming applications will be documented, including any conjunct agreements with the applicants. Unless explicitly consented for by the applicant, this information will not be used for purposes other than evaluation of the application. In cases where an application contains classified plans, know-how or trade secrets, the applicant may request the conclusion of a separate confidentiality agreement.

### 2. Multiple applications

Applicants agree to use the Project Data for the approved purpose and project described in the application; use of the data for a new purpose or project will require a new application and approval. The DAC will consider applications that include named collaborators, but each Institution must sign a separate Data Access Agreement. In the event an applicant wishes to share the data with additional collaborators not previously approved, these additional collaborators must make a separate application for access to the Project Data. In the event two or more applications overlap, the DAC may propose the respective applicants to align their applications.

### 3. Assessment criteria

Applications which seek to reserve Project Data for unspecified research goals will not be taken into consideration. Applications for access to Project Data must be Specific, Measurable, Attainable, Resourced and Timely (SMART). Specifically, the LLDEEP DAC will assess each application to determine whether:

- i. it has been submitted by a qualified researcher or researchers, who is employed by or legitimately affiliated with a recognized research institution that can provide institutional responsibility for appropriate research governance;
- ii. its proposed use of the Project Data is in accordance with and meets the objectives of the LLDEEP Project and the objectives of the associated cohorts and collections;
- iii. its proposed use of the Project Data constitutes research in the context of the individual sample donor consent process, and is likely to be understood as such by the individual sample donors;
- iv. it would breach any of the ethical permissions or restrictions in the consent form;
- v. all required ethico-legal approvals, restrictions and commitments for the proposed use of the Project Data have been obtained and adhered to;
- vi. the proposed research has any adverse potential impact, specifically whether the proposed research could affect minorities;
- vii. the nature of the funding of the application and the applicant;
- viii. the research proposal has been peer-reviewed, and, if not, whether the proposal satisfies applicable scientific standards;
- ix. the requested data are, quantitatively and qualitatively, suitable and not excessive for the applicant's proposed research;
- x. there are any similar applications pending or granted;
- xi. the proposed use of the Project Data has the potential to produce information that will enable identification of the individual sample donors.
- xii. the applicant has warranted that he has obtained all required ethic-legal approvals, restrictions and commitments, without limitation his data subject consent and approval of his pertinent medical ethical review board, for his proposed use of his data in combination with the Project Data, if applicable.

#### 4. *Access decision*

The DAC will usually decide on an application within 5 working days after it has received all pertinent information. However this may be extended to six weeks should the application turn out to be legally complicated. If the DAC grants an application, additional agreements may be made regarding authorships along the lines set forth in the Publication Policy that will be sent upon application. Rejections of applications will be motivated.

#### **Mode of Access to Data; restrictions and authentication**

The Consortium will make the anonymized Project Data available by means of a custom method agreed upon by both parties. Issue of Project Data may be subject to restrictions pursuant to data subject's or ethics committee's consent. Prior to issue of Project Data applicants will be subject to authentication. The issue of Project Data and the Data Access Agreement will be administered by UMCG.

#### **Data Access Conditions**

Access to Project Data is conditional on prior receipt by the Consortium of the Data Access Agreement sent upon receipt of the application, signed and dated by the applicant(s) and the responsible employing Institution(s).

## Data Access Agreement

This agreement governs the terms on which access will be granted to the Project Data generated by the LLDEEP Project. In signing this agreement, You are agreeing to be bound by the terms and conditions of access set out in this agreement. For the sake of clarity, the terms of access set out in this agreement apply both to the User and the User's Institution (as defined below). User Institution and User are referred to within the agreement as "You" and "Your" and shall be construed accordingly.

### Definitions:

"Consortium" refers to the LLDEEP Consortium from the University Medical Center Groningen (UMCG).

"Project" refers to the LLDEEP project as described in TIFN GH001 Grant;

"Project Data" refers to the data generated by the Project.

"Data Subject" refers to the person whose sample has been used to generate Project Data.

"User" refers to a researcher who has successfully applied for access to Project Data.

"Publications" refers to, without limitation, articles published in print journals, electronic journals, reviews, books, posters and other written and verbal presentations of research.

"User Institution" refers to the organization with which the User is employed, affiliated or enrolled.

### Terms and Conditions:

In signing this Agreement:

1. You agree to use Project Data only for the advancement of scientific and medical research.
2. You agree not to use, and not to allow use of, Project Data or any part thereof for the creation of products or services for sale or for any commercial purpose, whether sale by You or by third parties.
3. You acknowledge that the Project Data You receive is anonymized and that the Consortium will under no circumstances provide You with any means to identify any Data Subject. You agree not to use Project Data to identify or contact individual Data Subjects. You agree to preserve, at all times, the confidentiality of information and Project Data. In particular, You undertake not to use, or attempt to use Project Data to compromise or otherwise infringe the confidentiality of information on Data Subjects and their right to privacy.
4. You agree to disclose Project Data only to those of Your employees who strictly need access to it for the purpose of your approved study (Authorized Recipients), and who are bound by a duty of confidentiality which is at least as strict as provided for in this agreement.
5. You agree to ensure that any Authorized Recipient who has access to Project Data is advised of the confidentiality of Project Data. Upon request by the Consortium, you agree to furnish the Consortium the names of all Authorized Recipients who have gained access to Project Data. You agree that You are responsible for all breaches of confidentiality and of the terms of this agreement by your Authorized Recipients.
6. You agree to store Project Data on a computer with adequate security controls that prevent unauthorized access to or loss of Project Data and to maintain appropriate control over the Project Data. You represent and warrant that You have in place, and You agree that You will keep having in place, state of the art technical and

organizational security measures preventing unauthorized access and loss of Project Data or other forms of unlawful processing of Project Data, including without limitation, physical security measures, access controls, security and privacy technologies, security checks in relation to personnel, security incident response management and audit arrangements. You agree to notify forthwith the administrator of the Project, UMCG, of any violations of the foregoing.

7. You represent and warrant that all required ethico-legal approvals, restrictions and commitments, without limitation data subject consent and approval of the pertinent medical ethical review board, for Your proposed use of Your data in combination with the Project Data, have been obtained and adhered to.
8. In the event the Project Data is issued to You subject to any Data Subject's restrictions, you hereby assume the obligation to abide by these restrictions as Your own obligation towards the Data Subject. You agree that Data Subjects can enforce such restrictions as well as any pertinent provisions of this agreement as third-party beneficiaries and, in the event they choose to do so, that You will not object to Data Subjects being represented by an association or other bodies if they so wish. You agree that a Data Subject who has suffered damage(s) as a result of any violation by You of pertinent provisions in this Agreement or of applicable legislation, is entitled to receive compensation from You for the damage(s) suffered. You agree that, in the event of such a violation, the Data Subject may bring an action before a court within the Data Subject's jurisdiction against You.
9. You agree not to attempt to link Project Data to other information or archive data available for the data sets provided to You under this agreement, even if access to that data has been formally granted to you, or it is freely available without restriction, without the required consents and approvals. You agree to retain control over Project Data, and NOT TO DISTRIBUTE (individual-level or aggregated) Project Data provided under this agreement, in any form, to any entity or individual.
10. You agree not to transfer, share or disclose Project Data, in whole or in part, to or with anybody else, whether or not he or she is affiliated with Your institution, unless he or she is a named and approved collaborator under the LLDEEP Data Access Procedure.
11. You agree to use Project Data for the purpose and project described in your application (nr. \_\_\_\_\_ approved by the LLDEEP Data Access Committee; use of the data for a new purpose or project or by another researcher will require a new application and approval. Substantive modification of the original application requires submission of an amendment to the original application.
12. You agree to, when analyses result in newly estimated variables at an individual level, submit these new variables for archiving with the Consortium. This should be done no later than that the related paper will be submitted for publication.
13. You accept that Project Data will be reissued from time to time, with suitable versioning. If the reissue is at the request of sample donors and/or subject to other ethical scrutiny, You will destroy earlier versions of the Project Data.
14. You agree to destroy the Project Data after the resulting paper has been published. You agree to abide by the terms outlined in the Consortium 'Publication Policy' as set out in *Schedule 1*.
15. You agree to acknowledge in any work based in whole or in part on Project Data, the published paper from which the Project Data derives, the version of the data, and the role of the Consortium and the relevant primary collectors and their funders, as per the wording provided in the Publication Policy set forth in *Schedule 1*.
16. You accept that the Consortium, the original Project Data generators and the funders of the Project Data:
  - a) bear no legal responsibility whatsoever for the accuracy, completeness or comprehensiveness of the Project Data; and

- b) accept no liability for direct, indirect, consequential or incidental, damages or losses arising from any use of Project Data by anyone, including without limitation any clinical use and any commercial use thereof, or arising from the unavailability of, or break in access to, the Project Data for whatever reason.
17. You understand and acknowledge that Project Data are protected by copyright, database right and other intellectual property and proprietary rights, and that duplication, sale, licensing or any other form of disposing or distribution or making available, of all or part of the Project Data on any media is not permitted.
  18. You acknowledge that the Project Data are made available to you under this agreement on the title of un-assignable, non-exclusive "loan" ("leen"). You recognize that nothing in this agreement shall operate to transfer to You any rights, including but not limited to intellectual property rights and database rights, in or relating to Project Data.
  19. You accept that this agreement will terminate immediately upon any breach of this agreement by You and You will be required to destroy any Project Data held.
  20. You accept that it may be necessary for the Consortium or its appointed agent to alter the terms of this agreement from time to time in order to address new concerns. In this event, the Consortium or its appointed agent will contact You to inform You of any changes and You agree that Your continued use of Project Data shall be dependent on the parties entering into a new version of the agreement.
  21. You agree that You will submit a report to the Consortium Data Access Committee, if requested, on completion of your study of the Project Data. Such report will include the data resulting from your study of the Project Data. The Consortium Data Access Committee agrees to treat the report and all information, data, results, and conclusions contained within such report as confidential information.
  22. You accept that Project Data may be protected by and subject to national and international laws and that You are responsible for ensuring compliance with any such applicable law. The Consortium Data Access Committee reserves the right to request and inspect Your data security and management documentation to ensure the adequacy of data protection measures in countries that have no national laws comparable to that which pertain in the EAA.
  23. This agreement shall be construed, interpreted and governed by the laws of the Netherlands and shall be subject to the exclusive jurisdiction of the district court of Utrecht.

## **SCHEDULE 1**

### *Publication Policy*

Publications involving Project Data that do not fall under use case 3 as specified below, should be submitted to the LLDEEP Consortium for review, including review of proper authorship allocation and proper acknowledgments, before submission. The LLDEEP Consortium will differentiate between the following three usecases:

*Use case 1.* Primary papers by the Project.

*Use case 2.* Papers which include analyses using Lifelines DEEP Data. The " Lifelines DEEP Project" will be listed as a co-author. The "Lifelines DEEP Project" banner will expand in PubMed or MEDLINE to a full list of individuals in the Lifelines DEEP Project (this list will be provided by the Consortium). Specific individuals of the Project, who have contributed substantially to the analyses or the paper, may be listed explicitly as well (to be agreed upon by the Applicant and the Consortium).

*Use case 3.* Papers which include analyses of LLDEEP Data, made available through our open-access web site. For lookups of specific variants, as part of, for example, disease studies, the Project will make available variant lists across the microbiome that can be accessed. For similar queries where the open-access data is not sufficient, the Consortium can be contacted to see if Project Data can be used. In general, no co-authorship will be needed, although it is emphasized that the open-access Data is made available on an "as is" basis and subject to the condition that any use thereof shall not in any way inhibit or restrict the free use thereof by any thirdparty.

For each of the above use cases, authors who use Project Data must:

1. acknowledge the LLDEEP Consortium, TIFN and Broad using the following wording:

*" We thank the participants and the staff of LifeLines-DEEP for their collaboration. Funding for the project was provided by the Top Institute Food and Nutrition Wageningen grant GH001. The sequencing was carried out in collaboration with the Broad Institute."*

2. acknowledge, by reference, the funders of the sample collections from which the Project Data have been derived:
  - a. The LifeLines Cohort Study (<http://www.lifelines.nl>)
3. cite the relevant primary LLDEEP publication