A mouse fecal microbial gene catalogue established by Illumina-based sequencing

Xiao, Liang; Sonne, Si Brask; Long, Hua; Chou, Joyce; Glanville, Jacob; Hao, Qin; Li, Xiaoping; Qin, Junjie; Liang, Suisha; Fjære, Even; Licht, Tine Rask; Mortensen, Alicja; Vogel, Ulla; Kotowska, Dorota; Colding, Camilla; Tao, Ma; Kiilerich, Pia; Sørup Tastesen, Hanne; Tremaroli, Valentina; Bäckhed, Fredrik; Madsen, Lise; Ehrlich, Dusko; Doré, Joël; Wang, Jun; Lin, John C.; Kristiansen, Karsten

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

A mouse fecal microbial gene catalogue established by Illumina-based sequencing

Liang Xiao[#1], Si Brask Sonne[#2], Hua Long[3], Joyce Chou[3], Jacob Glanville[3], Qin Hao[2], Xiaoping Li[1], Junjie Qin[1], Suisha Liang[1], Even Fjære[2,4], Tine Rask Licht[5], Alicja Mortensen[5], Ulla Vogel[6], Dorota Kotowska[2], Camilla Colding[2], Ma Tao[2], Pia Kiilerich[2], Hanne Sørup Tastesen[2], Valentina Tremaroli[7], Fredrik Bäckhed[7], Lise Madsen[2,4], Dusko Ehrlich[8], Joël Doré[8], Jun Wang[1], John C. Lin*[3], Karsten Kristiansen*[2]

1. Address BGI-shenzhen, Shenzhen 518083, China 2. Department of Biology, University of Copenhagen, DK-2200 Copenhagen, Denmark 3. Rinat, Pfizer Inc., South San Francisco, California, USA 4. National Institute of Nutrition and Seafood Research (NIFES), Bergen, Norway 5. National food Institute, Technical University of Denmark, DK-2860 Søborg, Denmark 6. National Research Centre for the Working Environment, DK-2100 Copenhagen, Denmark 7. Sahlgrenska University Hospital, S-41345 Gothenburg, Sweden 8. Institut National de la Recherche Agronomique, 78350 Jouy en Josas, France

# These authors contributed equally to this work.

* Corresponding authors: * kk@bio.ku.dk, *john.lin@pfizer.com

Abstract

Background: Evidence has accumulated that the gut microbiota is a pivotal player in the regulation of whole body metabolism. To establish a comprehensive catalogue of microbial genes in the important animal model, the mouse, we collected feces from commonly used mouse strains fed a normal low fat diet or a high fat diet and housed in different laboratories in Europe, USA and China. Both males and females were included in the study.

Results and conclusion: By Illumina-based metagenomic sequencing we obtained 764.08 gigabases of sequencing reads from 154 fecal samples.

1. Comparison of chow and high fat diets showed that a high fat diet increased the diversity of the microbiome.
2. The obesity prone strain C57BL/6J exhibited a lower diversity than the obesity resistant strains.
3. Finally, mice bred and housed in different laboratories harbored significantly different microbiomes indicating that subtle differences in housing conditions exert a profound influence on the composition of the gut microbiome.

Samples collection, sequencing and data analysis

We collected 304 fresh mouse fecal samples from laboratories in Europe, United States and China, among which 154 samples have been sequenced and analyzed.

Instead of rRNA sequencing, we applied metagenomic sequencing to give a better overview of the content, diversity and function of the mouse gut microbiome. This method generated 764 Gb of microbial sequence from the 154 mouse fecal samples, and for each samples generated 4.96 Gb on average. We then used SOAPdenovo to assemble these reads. The assembly procedure is shown in the flowchart in figure 1.

After assembly 73.2% reads were successfully assembled to a total of 6.99 million contigs, with an N50 length of 4.99 kb and a total length of 14.2 Gb (Table 1).

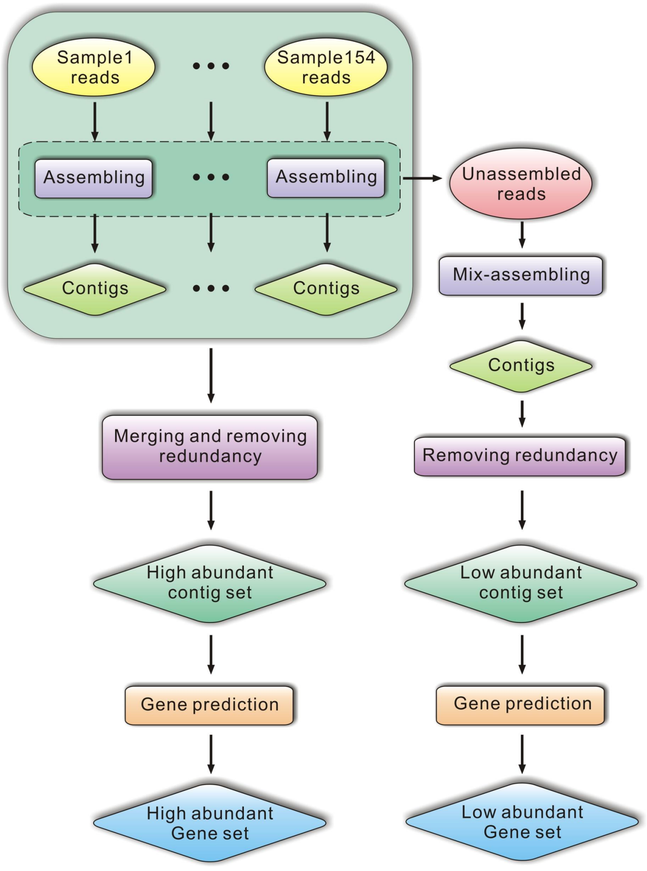

Figure 1

Table 1

Contig set	ctg_num	ctg_length_all	ctg_n50	ctg_n90	ctg_max	ctg_min
Low abundance contig set	455,050	735,874,753	2,249	679	96,971	500
High abundance contig set	664,601	1,657,278,123	6,273	803	418,662	500

A gene catalogue of mouse gut microbiome

We used the MetaGene program to predict ORFs in our contigs and found 17.53 million ORFs with an average length of 726 bp. 50.4% of them appeared to be complete ORFs (Table 2). All the ORFs were mapped to 2.3 million target genes. In figure 2 we show the cumulative number of genes identified across the individuals sequenced. The plateau at the end of the curve indicates that we have identified most of the genes accessible under the conditions used.

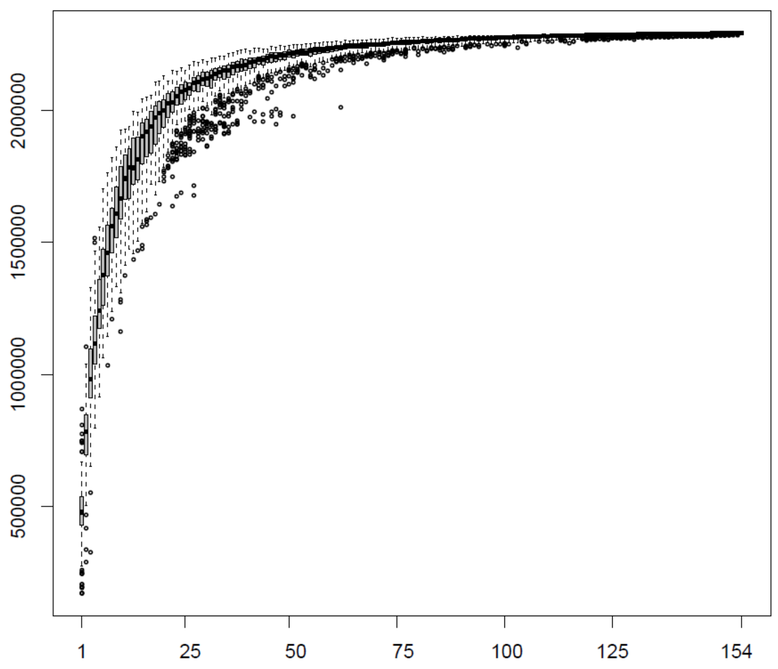

Figure 2

Table 2

Sample	ORFs	Total_length	Average_length	Complete_ORFs	Fragmental_ORFs	Complete	Fragmental
Low abundance ORFs	982,780	645,025,185	656	409,986	572,794	41.72%	58.28%
High abundance ORFs	1,622,031	1,211,482,383	747	879,934	742,097	54.25%	45.75%

Comparison to human gut microbiome

We selected the common genes (exist in above 50% individuals) from both human and mouse gene catalogues and compared them at the gene family level. About one quarter of all gene families were shared between human and mouse (Fig. 3A), most of them involved in metabolism (Fig. 3B).

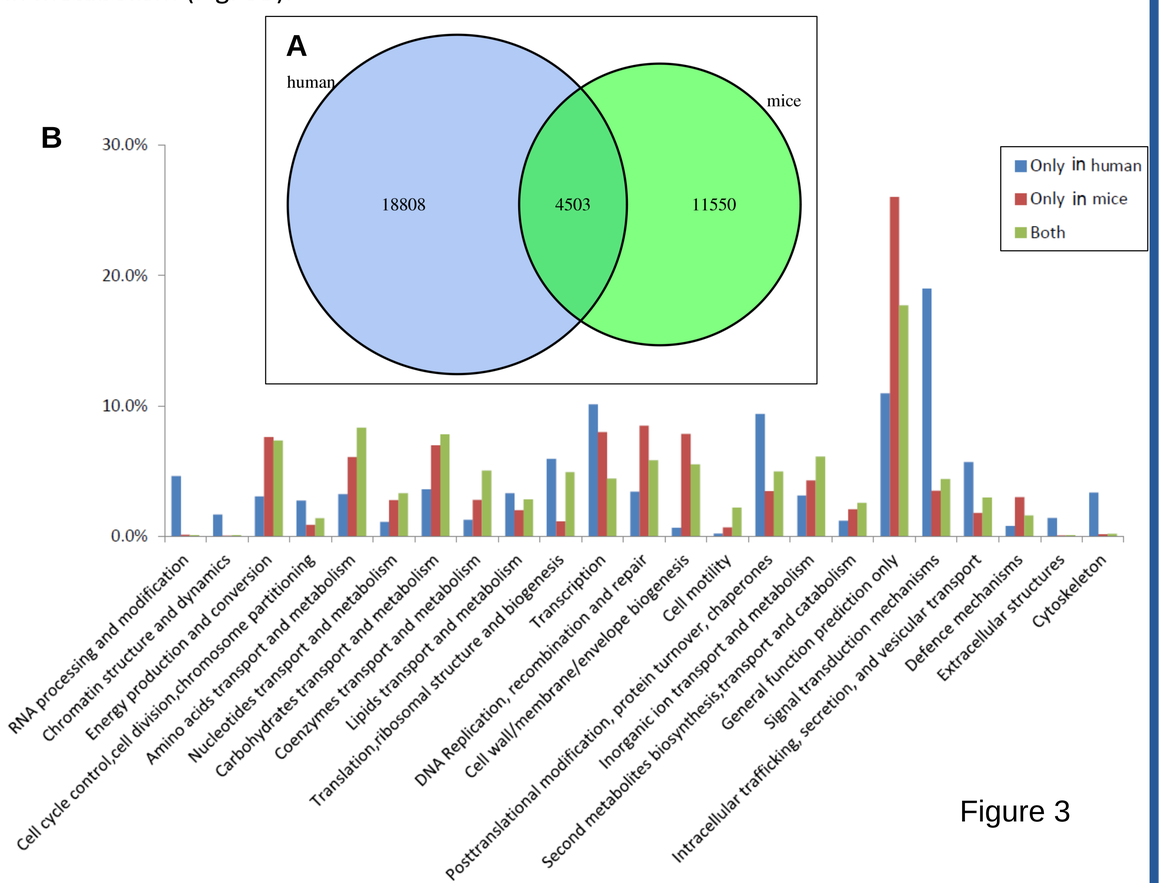

Figure 3

Correlation between mouse gut microbiome and environment, strain and diets

The PCA result showed that environment is the most important factor influencing mouse gut microbiome (Fig. 4A). Simpson index showed that high fat (HF) diet increased diversity (Fig. 4B) and C57BL/6J mice had lower diversity of gut microbiome than BALB/C and all other lines tested (Fig. 4C).

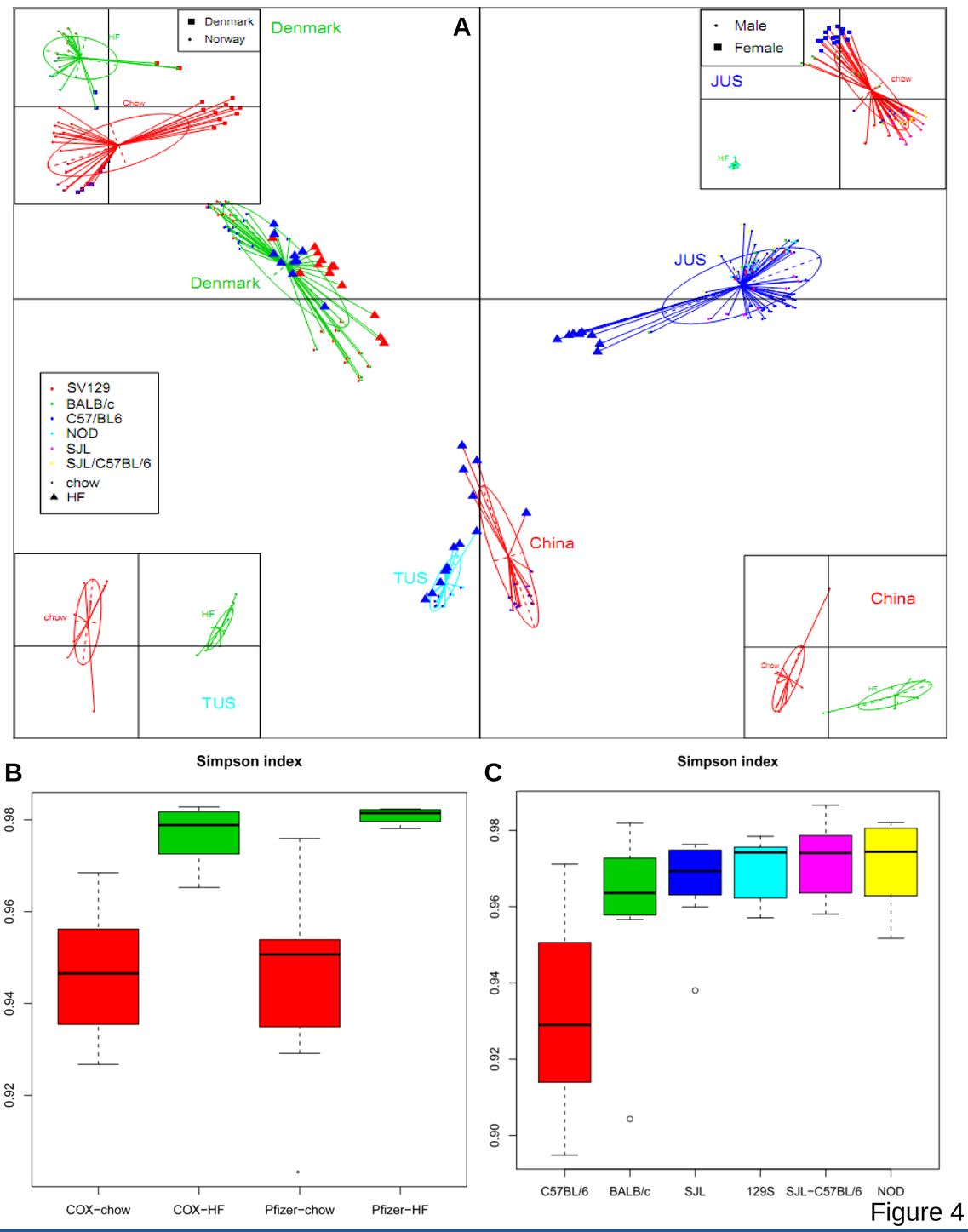

Figure 4