

次世代シーケンサーデータの解析手法  
第13回RNA-seq解析(その1)  
ウェブ資料

寺田朋子、坂本光央、清水謙多郎、門田 幸二\*

- ①ゲノム配列決定論文。
- ②フルテキストを閲覧

# W1 : *L. rhamnosus* GG

Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein.

Proc Natl Acad Sci U S A. 2009 Oct 6;106(40):17193-8. doi: 10.1073/pnas.0908876106. Epub 2009 Sep 17.

**Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein.**

Kankainen M<sup>1</sup>, Paulin L, Tynkkynen S, von Ossowski J, Reunanen J, Partanen P, Satokari R, Vesterlund S, Hendrickx AP, Lebeer S, De Keersmaecker SC, Vanderleyden J, Hämäläinen T, Laukkanen S, Salovuori N, Ritari J, Alatalo E, Korpela R, Mattila-Sandholm T, Lassig A, Hatakka K, Kinnunen KT, Karjalainen H, Saxelin M, Laakso K, Surakka A, Palva A, Salusjärvi T, Auvinen P, de Vos WM.

**Abstract**  
To unravel the biological function of the widely used probiotic bacterium *Lactobacillus rhamnosus* GG, we compared its 3.0-Mbp genome sequence with the similarly sized genome of *L. rhamnosus* LC705, an adjunct starter culture exhibiting reduced binding to mucus. Both genomes demonstrated high sequence identity and synteny. However, for both strains, genomic islands, 5 in GG and 4 in LC705, punctuated the colinearity. A significant number of strain-specific genes were predicted in these islands (80 in GG and 72 in LC705). The GG-specific islands included genes coding for bacteriophage components, sugar metabolism and transport, and exopolysaccharide biosynthesis. One island only found in *L. rhamnosus* GG

**Full text links**  
PNAS Free Full Text  
PMC FREE Full text

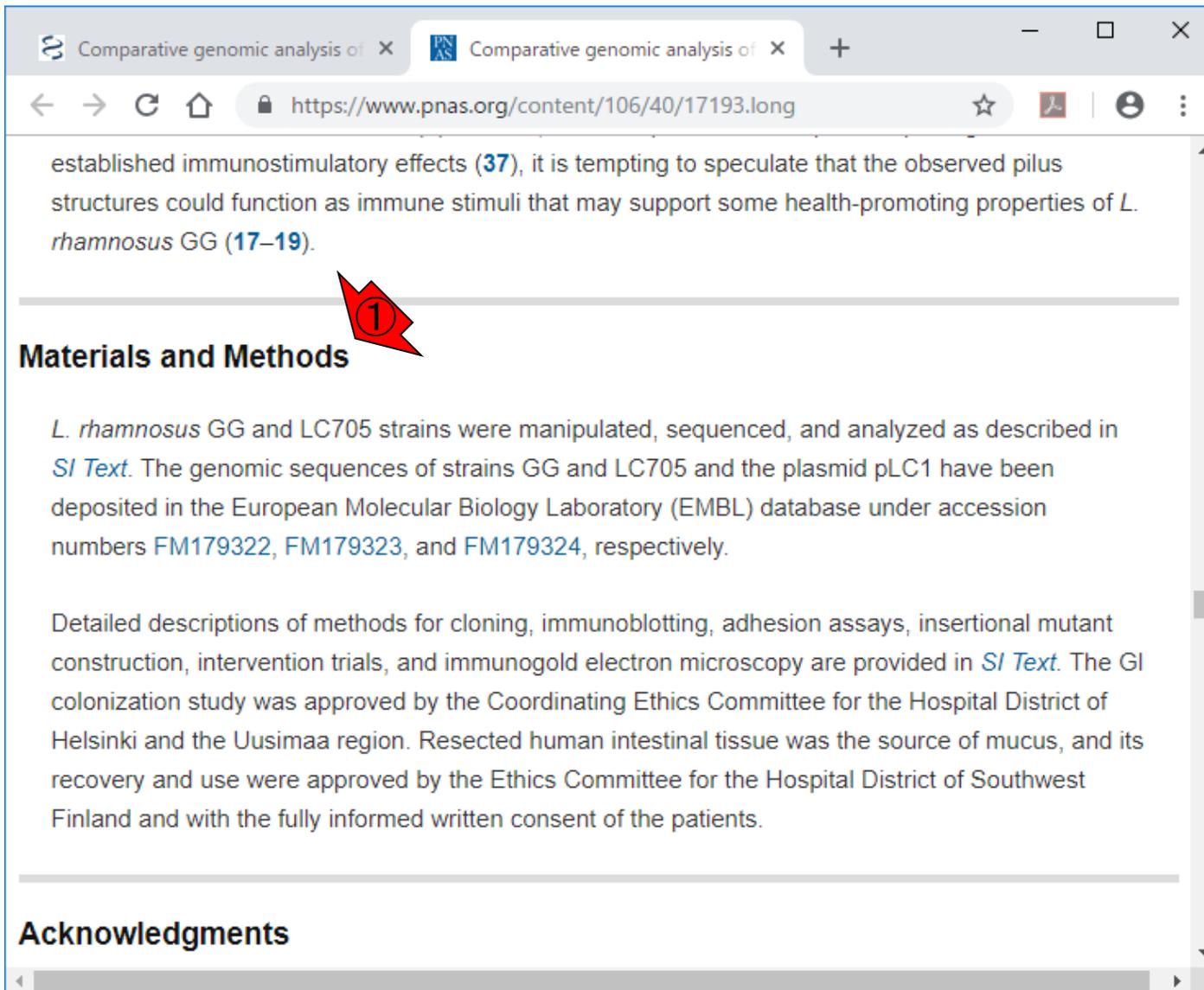
# W1 : *L. rhamnosus* GG

①Table 1で、②*L. rhamnosus* GGの基本情報が見られる。

Table 1.  
General genomic features of *L. rhamnosus* GG and LC705 and selected *Lactobacillus* spp.

Organism	<i>L. rhamnosus</i> GG	<i>L. rhamnosus</i> LC705	<i>L. casei</i> ATCC334	<i>L. plantarum</i> WCFS1
Genome size, Mbp	3.01	3.03	2.92	3.35
No. genes	2,944	2,992	2,771	3,057
Plasmids	0	1	1	3
rRNA operons	5	5	5	5
tRNA genes	57	61	59	70
GC content, %	47	47	47	44
Coding efficiency, %	85	85	82	84

# W1 : *L. rhamnosus* GG



Comparative genomic analysis of x Comparative genomic analysis of x +

https://www.pnas.org/content/106/40/17193.long

established immunostimulatory effects (37), it is tempting to speculate that the observed pilus structures could function as immune stimuli that may support some health-promoting properties of *L. rhamnosus* GG (17–19).

## Materials and Methods

*L. rhamnosus* GG and LC705 strains were manipulated, sequenced, and analyzed as described in *SI Text*. The genomic sequences of strains GG and LC705 and the plasmid pLC1 have been deposited in the European Molecular Biology Laboratory (EMBL) database under accession numbers FM179322, FM179323, and FM179324, respectively.

Detailed descriptions of methods for cloning, immunoblotting, adhesion assays, insertional mutant construction, intervention trials, and immunogold electron microscopy are provided in *SI Text*. The GI colonization study was approved by the Coordinating Ethics Committee for the Hospital District of Helsinki and the Uusimaa region. Resected human intestinal tissue was the source of mucus, and its recovery and use were approved by the Ethics Committee for the Hospital District of Southwest Finland and with the fully informed written consent of the patients.

## Acknowledgments

# W1 : *L. rhamnosus* GG

① *L. rhamnosus* GGのAccession番号は、②FM179322。②のリンク先をクリック

established immunostimulatory effects (37), it is tempting to speculate that the observed pilus structures could function as immune stimuli that may support some health-promoting properties of *L. rhamnosus* GG (17–19).

## Materials and Methods

*L. rhamnosus* GG and LC705 strains were **①** cultured, sequenced, and analyzed as described in *SI Text*. The genomic sequences of strains GG and LC705 and the plasmid pLC1 have been deposited in the European Molecular Biology Laboratory (EMBL) database under accession numbers **②** FM179322, FM179323, and FM179324, respectively.

Detailed descriptions of methods for cloning, immunoblotting, adhesion assays, insertional mutant construction, intervention trials, and immunogold electron microscopy are provided in *SI Text*. The GI colonization study was approved by the Coordinating Ethics Committee for the Hospital District of Helsinki and the Uusimaa region. Resected human intestinal tissue was the source of mucus, and its recovery and use were approved by the Ethics Committee for the Hospital District of Southwest Finland and with the fully informed written consent of the patients.

## Acknowledgments

# W2: FM179322

①ここに飛ばされます。②赤枠部分のデフォルトの表示は、③GenBank形式です。

Comparative genomic analysis of x Lactobacillus rhamnosus GG who x

https://www.ncbi.nlm.nih.gov/nucleotide/FM179322

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank Send to

## Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103)

GenBank: FM179322.1

[FASTA](#) [Graphics](#)

Go to: [v]

LOCUS FM179322 3010111 bp DNA circular BCT 27-FEB-2015

DEFINITION Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103).

ACCESSION FM179322

VERSION FM179322.1

DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)

KEYWORDS complete genome.

SOURCE Lactobacillus rhamnosus GG

## W2: FM179322

Comparative genomic analysis of x Lactobacillus rhamnosus GG who x +

https://www.ncbi.nlm.nih.gov/nuccore/FM179322

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank **1** Send to:

Format

- Summary
- GenBank
- GenBank (full)
- FASTA
- FASTA (text)
- Graphics
- ASN.1
- Revision History
- Accession List
- GI List

**Lactobacillus rhamnosus GG whole genome sequence, strain GG**

3010111 bp DNA circular BCT 27-FEB-2015  
Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC)

ACCESSION FM179322  
VERSION FM179322.1  
DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)  
KEYWORDS complete genome.  
SOURCE Lactobacillus rhamnosus GG

# W2: FM179322

Comparative genomic analysis of Lactobacillus rhamnosus GG who

https://www.ncbi.nlm.nih.gov/nucleotide/FM179322.1?report=fas...

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank **1** Send to:

**Format**

- Summary
- GenBank
- GenBank (full)
- FASTA
- FASTA (text) 2**
- Graphics
- ASN.1
- Revision History
- Accession List
- GI List

**Lactobacillus rhamnosus GG whole genome sequence, strain GG**

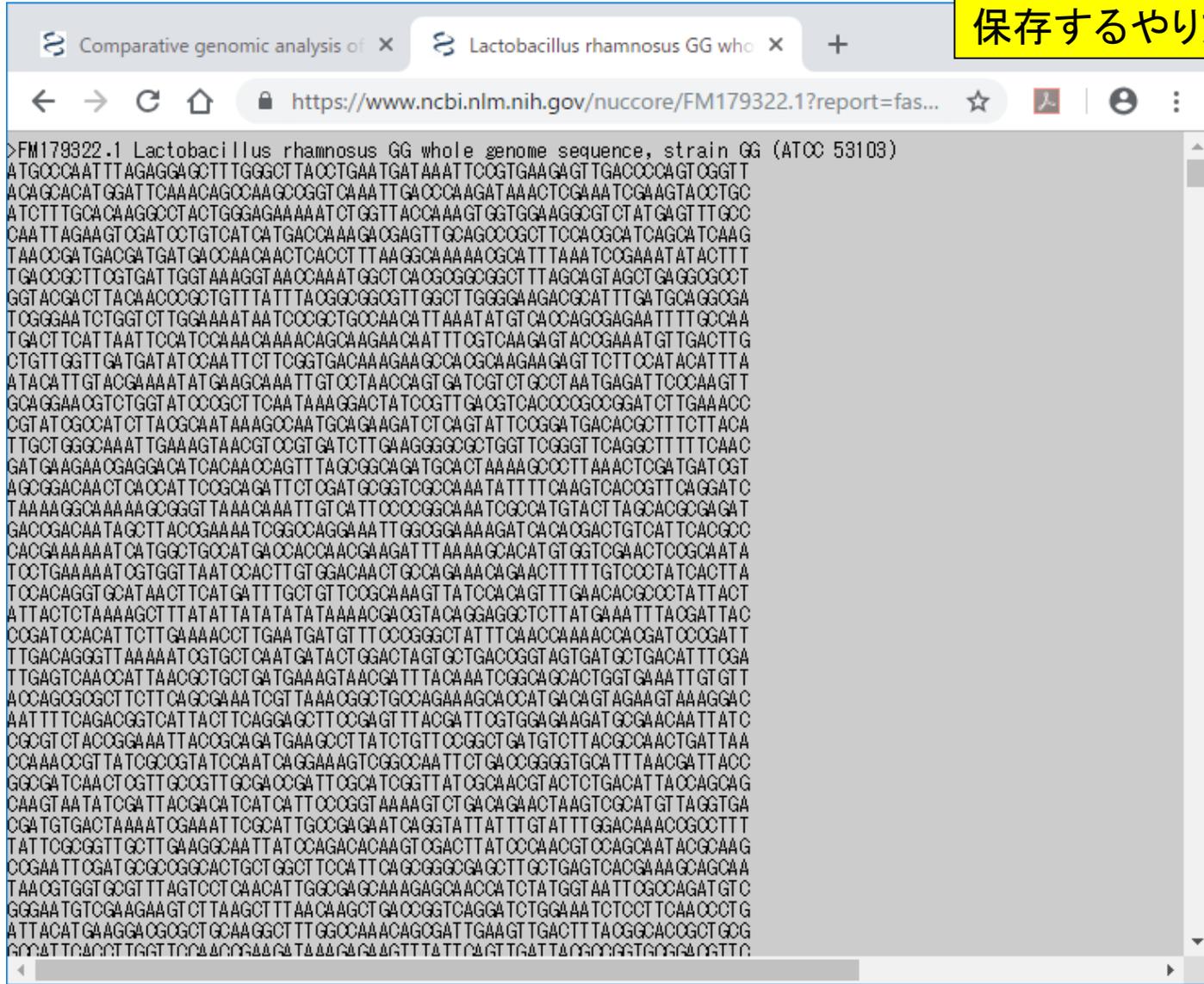
3010111 bp DNA circular BCT 27-FEB-2015  
Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC)

ACCESSION FM179322  
VERSION FM179322.1  
DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)  
KEYWORDS complete genome.  
SOURCE Lactobacillus rhamnosus GG

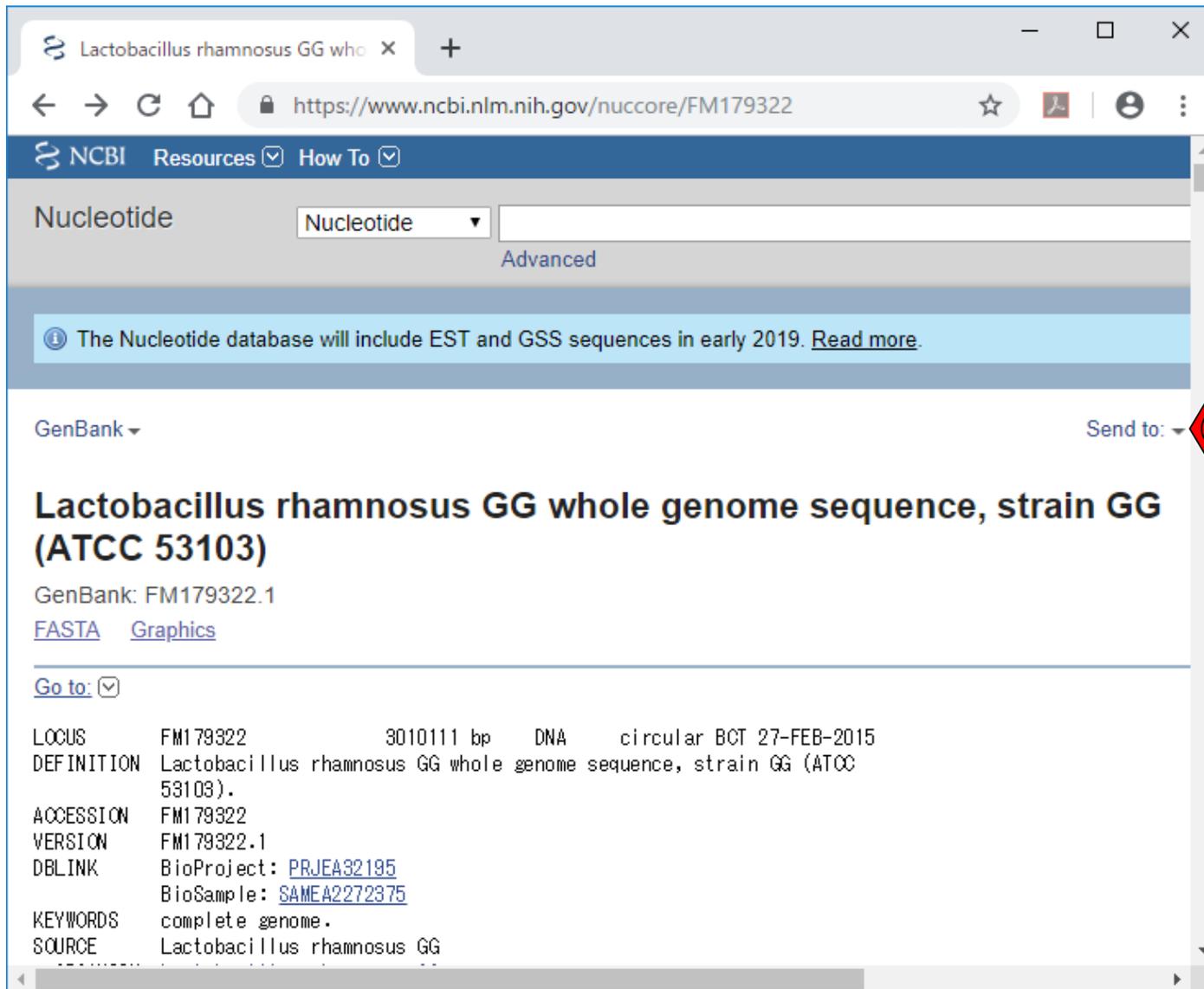
www.ncbi.nlm.nih.gov を待機しています...

# W2: FM179322

FASTA形式に切り替わります。ただ、ここでは表示がFASTA形式なだけで、右クリックで保存できたりするわけではない。FASTA形式で保存するやり方は次のスライドで示します。



# W2: FM179322



Lactobacillus rhamnosus GG who x +

https://www.ncbi.nlm.nih.gov/nucleotide/FM179322

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank Send to ①

## Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103)

GenBank: FM179322.1

[FASTA](#) [Graphics](#)

Go to: [v]

LOCUS FM179322 3010111 bp DNA circular BCT 27-FEB-2015

DEFINITION Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103).

ACCESSION FM179322

VERSION FM179322.1

DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)

KEYWORDS complete genome.

SOURCE Lactobacillus rhamnosus GG

# W2: FM179322

Lactobacillus rhamnosus GG who x +

https://www.ncbi.nlm.nih.gov/nuccore/FM179322

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank **Lactobacillus rhamnosus GG whole genome (ATCC 53103)**

GenBank: FM179322.1

[FASTA](#) [Graphics](#)

Go to:

**Send to:**

- Complete Record
- Coding Sequences
- Gene Features

Choose Destination

- File
- Clipboard
- Collections
- Analysis Tool

LOCUS FM179322 3010111 bp DNA circular BCT 27-FEB-2015

DEFINITION Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103).

ACCESSION FM179322

VERSION FM179322.1

DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)

KEYWORDS complete genome.

SOURCE Lactobacillus rhamnosus GG

# W2: FM179322

①Send to、②File、③がFASTA形式になっていることを確認して、④Create File。

GenBank

## Lactobacillus rhamnosus GG whole genome (ATCC 53103)

GenBank: FM179322.1

[FASTA](#) [Graphics](#)

Go to:

LOCUS FM179322 3010111 bp DNA circular  
DEFINITION Lactobacillus rhamnosus GG whole genome sequence, strain ATCC 53103.  
ACCESSION FM179322  
VERSION FM179322.1  
DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)  
KEYWORDS complete genome.  
SOURCE Lactobacillus rhamnosus GG

Send to: **①**

- Complete Record
- Coding Sequences
- Gene Features

Choose Destination

- File **②**
- Clipboard
- Collections
- Analysis Tool

Download 1 item.

Format **③**

FASTA

Show GI

Create File **④**



# W2: FM179322

①Send to、②File、③がFASTA形式になっていることを確認して、④Create File。こんな感じになります。⑤デフォルトはsequence.fastaなので、⑥FM179322.fastaというファイル名で、⑦保存。

The screenshot shows the NCBI Nucleotide database page for **Lactobacillus rhamnosus GG whole genome (ATCC 53103)**. The page includes the GenBank accession number **FM179322.1** and links for **FASTA** and **Graphics**. The FASTA section is partially visible, showing the following information:

LOCUS FM179322 3010111 bp DNA circular  
DEFINITION Lactobacillus rhamnosus GG whole genome sequence, str. ATCC 53103.  
ACCESSION FM179322  
VERSION FM179322.1  
DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)  
KEYWORDS complete genome.  
SOURCE Lactobacillus rhamnosus GG

Overlaid on the right is a Windows file save dialog titled "名前を付けて保存" (Save with name). The dialog shows the current directory as "PC > デスクトップ" (Desktop). The file name field contains "FM179322.fasta" (marked with a red arrow and the number 6), and the file type is set to "FASTA ファイル (\*.fasta) (\*.fasta)". The "保存(S)" (Save) button is highlighted with a red arrow and the number 7.

# W3: FM179322

①GenBank形式で再表示。原著論文のTable 1ではゲノムサイズが3.01Mbpと書かれていたが、②正確には3,010,111 bpであることがわかる。また、③ATCC 53103という菌株であることもわかる。

Comparative genomic analysis of x Lactobacillus rhamnosus GG who x +

https://www.ncbi.nlm.nih.gov/nuccore/FM179322

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank 1 Send to:

## Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103)

GenBank: FM179322.1

[FASTA](#) [Graphics](#)

Go to: 2

LOCUS FM179322 3010111 bp DNA circular BCT 27-FEB-2015 3

DEFINITION Lactobacillus rhamnosus GG whole genome sequence, strain GG (ATCC 53103).

ACCESSION FM179322

VERSION FM179322.1

DBLINK BioProject: [PRJEA32195](#)  
BioSample: [SAMEA2272375](#)

KEYWORDS complete genome.

SOURCE Lactobacillus rhamnosus GG

# W3: FM179322

①少しページ下部に移動したところ。②  
FEATURESのところの、③source部分を眺めること  
で、Taxonomy IDが568703であることもわかる。

Comparative genomic analysis of Lactobacillus rhamnosus GG who

https://www.ncbi.nlm.nih.gov/nucore/FM179322

Hendrickx,A.P., Lebeer,S., De Keersmaecker,S.C., Vanderleyden,J., Hamalainen,T., Laukkanen,S., Salovuori,N., Ritari,J., Alatalo,E., Korpela,R., Mattila-Sandholm,T., Lassig,A., Hatakka,K., Kinnunen,K.T., Karjalainen,H., Saxelin,M., Laakso,K., Surakka,A., Palva,A., Salusjarvi,T., Auvinen,P. and de Vos,W.M.

TITLE Comparative genomic analysis of Lactobacillus rhamnosus GG reveals pili containing a human- mucus binding protein

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 106 (40), 17193-17198 (2009)

PUBMED [19805152](#)

REFERENCE 2 (bases 1 to 3010111)

AUTHORS Paulin,L.G.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-2008) Paulin L.G., Institute of Biotechnology, University of Helsinki, PObox 56 (Viikinkaari 4), 00014 Helsinki, FINLAND

FEATURES Location/Qualifiers

source	1..3010111 /organism="Lactobacillus rhamnosus GG" /mol_type="genomic DNA" /strain="GG (ATCC 53103)" /culture_collection="ATCC:53103" /db_xref="taxon:568703"
gene	1..1350 /gene="dnaA" /locus_tag="LGG_00001"
CDS	1..1350 /gene="dnaA" /locus_tag="LGG_00001" /codon_start=1 /transl_table=11 /product="Chromosomal replication initiator protein dnaA" /protein_id="CARR5896.1"



# W4: 注意点

Taxonomy IDを辿っていくと、別グループの原著論文に行きついてしまい混乱する例を示します。  
①568703をクリック

Comparative genomic analysis of Lactobacillus rhamnosus GG who

https://www.ncbi.nlm.nih.gov/nuccore/FM179322

Hendrickx,A.P., Lebeer,S., De Keersmaecker,S.C., Vanderleyden,J., Hamalainen,T., Laukkanen,S., Salovuori,N., Ritari,J., Alatalo,E., Korpela,R., Mattila-Sandholm,T., Lassig,A., Hatakka,K., Kinnunen,K.T., Karjalainen,H., Saxelin,M., Laakso,K., Surakka,A., Palva,A., Salusjarvi,T., Auvinen,P. and de Vos,W.M.

TITLE Comparative genomic analysis of Lactobacillus rhamnosus GG reveals pili containing a human- mucus binding protein

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 106 (40), 17193-17198 (2009)

PUBMED [19805152](#)

REFERENCE 2 (bases 1 to 3010111)

AUTHORS Paulin,L.G.

TITLE Direct Submission

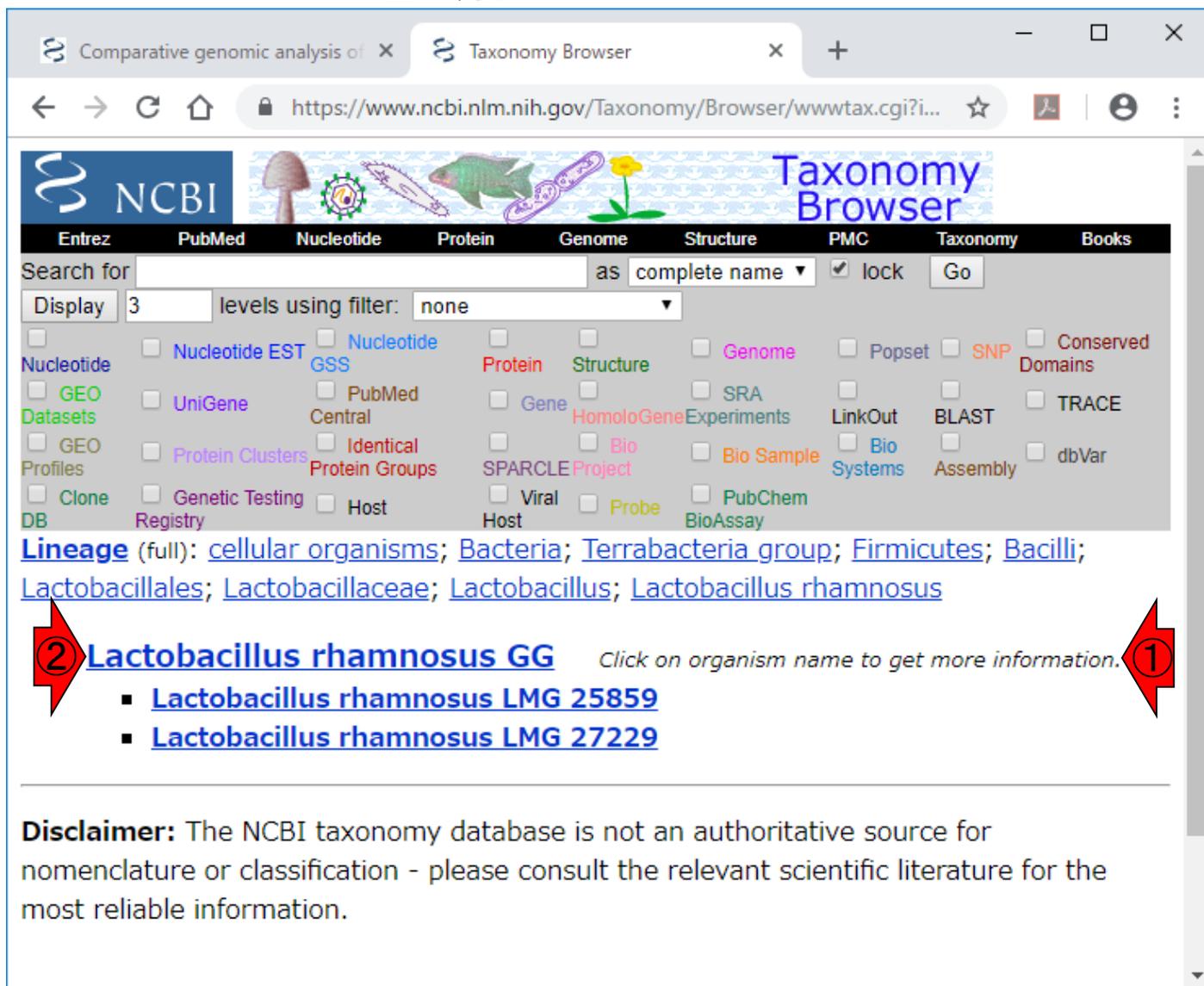
JOURNAL Submitted (02-JUL-2008) Paulin L.G., Institute of Biotechnology, University of Helsinki, PObox 56 (Viikinkaari 4), 00014 Helsinki, FINLAND

FEATURES

Location/Qualifiers	
source	1..3010111 /organism="Lactobacillus rhamnosus GG" /mol_type="genomic DNA" /strain="GG (ATCC 53103)" /culture_collection="ATCC 53103" /db_xref="taxon:568703" ①
gene	1..1350 /gene="dnaA" /locus_tag="LGG_00001"
CDS	1..1350 /gene="dnaA" /locus_tag="LGG_00001" /codon_start=1 /transl_table=11 /product="Chromosomal replication initiator protein dnaA" /protein_id="CARR5896.1"

# W4: 注意点

こんなページに飛びます。①と  
ことなので、②の部分をクリック



Comparative genomic analysis of Taxonomy Browser

https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?i...

NCBI Taxonomy Browser

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Books

Search for \_\_\_\_\_ as complete name  lock Go

Display 3 levels using filter: none

Nucleotide  Nucleotide EST  Nucleotide GSS  Protein  Structure  Genome  Popset  SNP  Conserved Domains

GEO Datasets  UniGene  PubMed Central  Gene  HomoloGene  SRA  LinkOut  BLAST  TRACE

GEO Profiles  Protein Clusters  Identical Protein Groups  SPARCLE Project  Bio  Bio Sample  Bio Systems  Assembly  dbVar

Clone DB  Genetic Testing Registry  Host  Viral Host  Probe  PubChem BioAssay

**Lineage** (full): [cellular organisms](#); [Bacteria](#); [Terrabacteria group](#); [Firmicutes](#); [Bacilli](#); [Lactobacillales](#); [Lactobacillaceae](#); [Lactobacillus](#); [Lactobacillus rhamnosus](#)

**②** [Lactobacillus rhamnosus GG](#) *Click on organism name to get more information.* **①**

- [Lactobacillus rhamnosus LMG 25859](#)
- [Lactobacillus rhamnosus LMG 27229](#)

**Disclaimer:** The NCBI taxonomy database is not an authoritative source for nomenclature or classification - please consult the relevant scientific literature for the most reliable information.

# W4: 注意点

こんなページに飛びます。現在は①の論文から辿ってきたFM179322 → ②NCBI:txid568703を眺めています。③ページ下部に移動。

Search for  as  complete name  lock

Display  levels using filter:

## Lactobacillus rhamnosus GG

Taxonomy ID: 568703 (for references in articles please use NCBI:txid568703)

Scientific name: **Lactobacillus rhamnosus GG** ②

equivalent: **Lactobacillus rhamnosus str. GG**

Inherited blast name: **firmicutes**

Rank: **no rank**

Genetic code: [Translation table 11](#) (Bacterial, Archaeal and Plant Plastid)

Other names:

synonym: **Lactobacillus rhamnosus ATCC 53103**

synonym: **Lactobacillus rhamnosus LMG 18243**

Entrez records		
Database name	Subtree links	Direct links
Nucleotide	<a href="#">68</a>	<a href="#">68</a>
Protein	<a href="#">11,630</a>	<a href="#">11,630</a>
Structure	<a href="#">14</a>	<a href="#">14</a>
Genome	<a href="#">1</a>	<a href="#">1</a>
Popset	<a href="#">3</a>	<a href="#">3</a>
GEO Datasets	<a href="#">83</a>	<a href="#">83</a>
PubMed Central	<a href="#">294</a>	<a href="#">294</a>
Gene	<a href="#">6,306</a>	<a href="#">6,306</a>
SRA Experiments	<a href="#">54</a>	<a href="#">52</a>
Identical Protein Groups	<a href="#">3,520</a>	<a href="#">3,520</a>
Bio Project	<a href="#">9</a>	<a href="#">9</a>
Bio Sample	<a href="#">58</a>	<a href="#">56</a>



# W4: 注意点

①このあたりまでページ下部に移動。②の論文から辿ってきたはずなのに、なぜか③ Referenceとして提示されているものが、④異なる論文であることがわかります。④をクリック

The screenshot shows a web browser window with the URL <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?...>. The page content includes:

- Comments and References:** A section with a red arrow labeled ③ pointing to the heading. Below it is the text: "genome sequence" and "Determination of the DNA genome sequence of this strain has been or is being determined either in whole or in part."
- Reference:** A red arrow labeled ④ points to a blue link icon next to the citation: "Morita H et al. (2009)". The full citation text is: "Morita, H., Toh, H., Oshima, K., Murakami, M., Taylor, T.D., Igimi, S., and Hattori, M. "Complete genome sequence of probiotic Lactobacillus rhamnosus ATCC 53103." J. Bacteriol. (2009) 191:7630-7631."
- External Information Resources (NCBI LinkOut):** A table with three columns: LinkOut, Subject, and LinkOut Provider. A red arrow labeled ① points to the top of this table.

LinkOut	Subject	LinkOut Provider
<a href="#">Lactobacillus rhamnosus</a>	meta-databases	<a href="#">BacDive</a>
<a href="#">Lactobacillus rhamnosusGG</a>	organism-specific	<a href="#">BioCyc</a>
<a href="#">GOLD: Go0001197</a>	organism-specific	<a href="#">Genomes On Line Database</a>
<a href="#">Show Biotic Interactions</a>	taxonomy/phylogenetic	<a href="#">Global Biotic Interactions</a>
<a href="#">2 records from this provider</a>	organism-specific	<a href="#">Integrated Microbial Genomes</a>
<a href="#">OMA</a>	taxonomy/phylogenetic	<a href="#">OMA Browser: Orthologous Matrix</a>

②

# W5: 別の論文

①同一菌株に対する別グループからのゲノム配列決定論文。②フルテキストを閲覧

The screenshot shows a web browser window displaying a PubMed article. The browser's address bar shows the URL: <https://www.ncbi.nlm.nih.gov/pubmed?cmd=Retrieve&dopt=A...>. The page header includes the NCBI logo and navigation links for 'Resources' and 'How To'. The main content area features the PubMed logo and a search bar. Below the search bar, there are options for 'Format: Abstract' and 'Send to'. The article title is 'Complete genome sequence of the probiotic Lactobacillus rhamnosus ATCC 53103.' with a red arrow labeled '1' pointing to the citation information: 'J Bacteriol. 2009 Dec;191(24):7630-1. doi: 10.1128/JB.01287-09. Epub 2009 Oct 9.' To the right of the article title, there are 'Full text links' for 'JB' (J Bacteriology) and 'PMC' (PubMed Central), both marked as 'FREE' and 'FINAL VERSION'. A red arrow labeled '2' points to these links. Below the title, the authors are listed: 'Morita H<sup>1</sup>, Toh H, Oshima K, Murakami M, Taylor TD, Igimi S, Hattori M.' There is also an 'Author information' section. The abstract text begins: 'Lactobacillus rhamnosus is a facultatively heterofermentative lactic acid bacterium and is frequently isolated from human gastrointestinal mucosa of healthy individuals. L. rhamnosus ATCC 53103, isolated from a healthy human intestinal flora, is one of the most widely used and well-documented probiotics. Here, we report the finished and annotated genome sequence of this organism.' At the bottom, there are identifiers: 'PMID: 19820099', 'PMCID: PMC2786603', and 'DOI: 10.1128/JB.01287-09'. There are also social media icons for Facebook, Twitter, and Google+.

# W5: 別の論文

Complete genome sequence of : x Complete Genome Sequence of : x +

← → ↻ 🏠 🔒 https://jlb.asm.org/content/191/24/7630.long ☆ 📄 👤 ⋮

Stay Informed with SSPA!  
Sign up to receive research and updates from  
actively-sponsored journals like ASM

SSPA Scientific Society  
Publisher Alliance  
BY SCIENTISTS FOR SCIENTISTS

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

GENOME ANNOUNCEMENTS

## Complete Genome Sequence of the Probiotic *Lactobacillus rhamnosus* ATCC 53103

Hidetoshi Morita, Hidehiro Toh, Kenshiro Oshima, Masaru Murakami, Todd D. Taylor, Shizunobu Igimi, Masahira

DOI: 10.1128/JB.01287-09

Article Info & Metrics PDF

PDF Help

Article Info & Metrics PDF

ABSTRACT

# W5: 別の論文

①このあたりまで移動。②このゲノムの基本情報が書かれている。

The complete genome sequence of *Lactobacillus rhamnosus* ATCC 53103 was determined by a whole-genome shotgun strategy with the Sanger method. Genomic libraries containing 2-kb inserts were constructed and sequenced, and 39,936 sequences were generated, giving 8.6-fold coverage from both ends of the genomic clones. Sequence reads were assembled with the Phred-Phrap-Consed program (2). Remaining gaps between contigs were closed by direct sequencing of clones. Prediction and annotation of protein-coding genes were performed as described previously (5).

The genome of *L. rhamnosus* ATCC 53103 consists of a circular 3,005,051-bp chromosome containing 2,834 predicted protein-coding genes and has no plasmid. Of all predicted protein-coding genes, we could assign 1,939 (68%) to known functions, 610 (22%) as conserved hypothetical genes, and 285 (10%) as novel hypothetical genes. This strain has a relatively high number of proteins involved in carbohydrate and amino acid metabolism and transport and defense mechanisms, compared with other sequenced intestinal lactobacilli. The genome encodes 28 complete phosphoenolpyruvate-carbohydrate phosphotransferase-type transporter systems (PTSs) and 25 putative glycosyl hydrolases, which classified into 12 different carbohydrate-active enzyme families (<http://www.cazy.org/>). Of the 12 alpha-L-fucosidase (GH29) and alpha-mannosidase (GH38) are not found in other sequenced intestinal lactobacilli. Of the 28 PTSs, 12 are encoded by genes adjacent to glycosyl hydrolase genes and transcriptional regulator genes, allowing localized transcriptional control. This organism carries 22 multidrug ABC transporters, eight antimicrobial peptide ABC transporters, and seven beta-lactamases, suggesting its broad range of antibiotic resistance. The genome contains 17 complete two-component

# W5: 別の論文

- ①また少しページ下部に移動。
- ②Accession番号はAP011548。

contain imperfect repeats consisting of serine and alanine. The genes for both proteins could encode mucin-like cell surface adhesives, because both genes are located adjacent to glycosyltransferase genes (7). The presence of genes encoding proteins for a diverse number of fermentable sugars, a variety of cell surface adherence proteins, bacteriocin biosynthetic proteins (LRHM\_2289 to LRHM\_2312), and bile salt hydrolase (LRHM\_0484) is likely to contribute to the organisms' gastric survival and promote interactions with the intestinal mucosa and microbiota.

During the preparation of this article, the genome sequence (3,010,111 bp) of *L. rhamnosus* GG, the original strain of *L. rhamnosus* ATCC 53103, was deposited in a public database (GenBank accession no. FM179322) (3). The genome of *L. rhamnosus* ATCC 53103 is 5 kb shorter than that of *L. rhamnosus* GG. Furthermore, an alignment analysis of both genome sequences shows that the 8.9-kb region (genome coordinates 618415 to 627294) of *L. rhamnosus* ATCC 53103 is inverted.

---

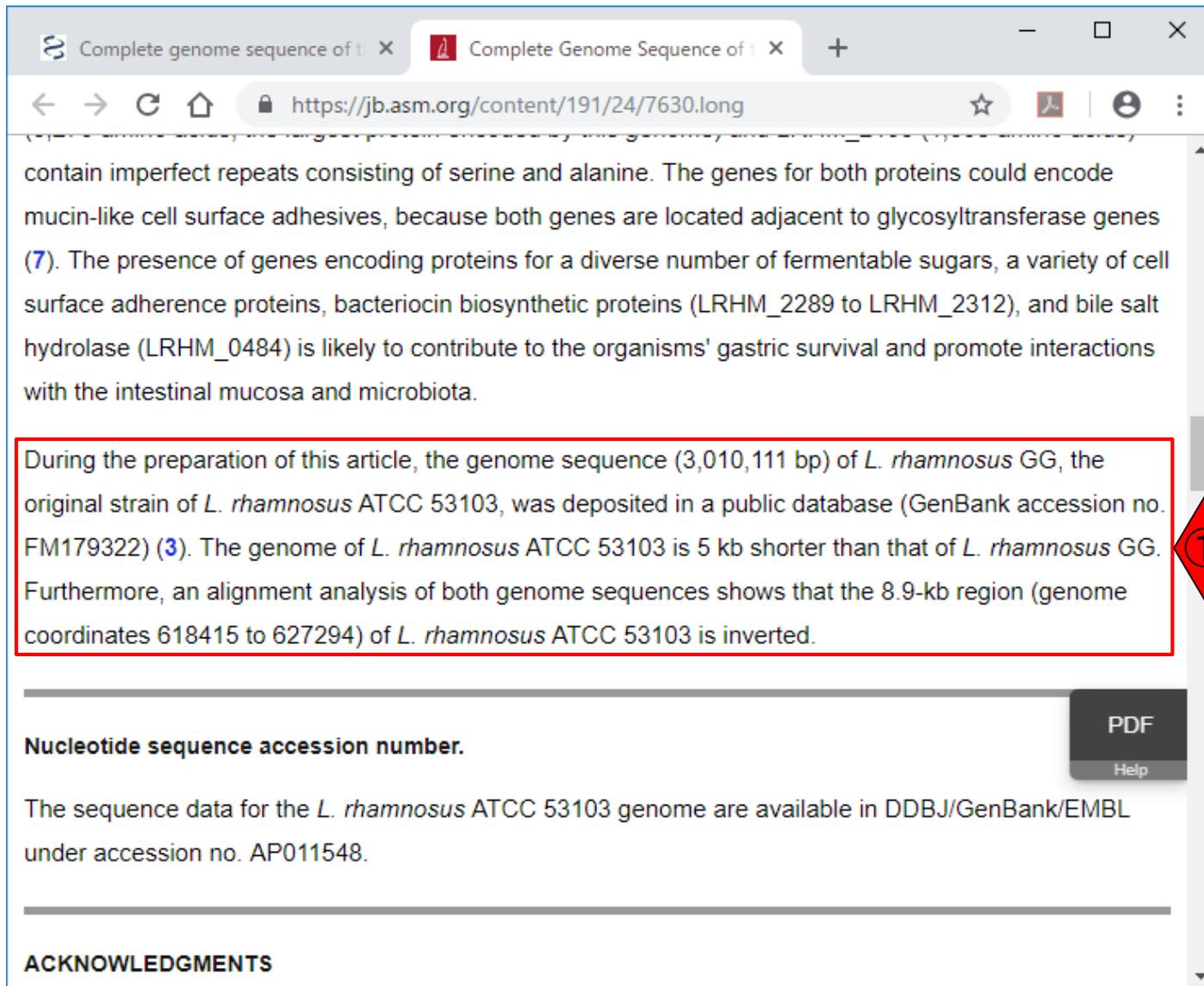
**Nucleotide sequence accession number.**

The sequence data for the *L. rhamnosus* ATCC 53103 genome are available in DDBJ/GenBank/EMBL under accession no. AP011548.

---

**ACKNOWLEDGMENTS**

# W5: 別の論文



Complete genome sequence of t x Complete Genome Sequence of t x +

https://jb.asm.org/content/191/24/7630.long

contain imperfect repeats consisting of serine and alanine. The genes for both proteins could encode mucin-like cell surface adhesives, because both genes are located adjacent to glycosyltransferase genes (7). The presence of genes encoding proteins for a diverse number of fermentable sugars, a variety of cell surface adherence proteins, bacteriocin biosynthetic proteins (LRHM\_2289 to LRHM\_2312), and bile salt hydrolase (LRHM\_0484) is likely to contribute to the organisms' gastric survival and promote interactions with the intestinal mucosa and microbiota.

During the preparation of this article, the genome sequence (3,010,111 bp) of *L. rhamnosus* GG, the original strain of *L. rhamnosus* ATCC 53103, was deposited in a public database (GenBank accession no. FM179322) (3). The genome of *L. rhamnosus* ATCC 53103 is 5 kb shorter than that of *L. rhamnosus* GG. Furthermore, an alignment analysis of both genome sequences shows that the 8.9-kb region (genome coordinates 618415 to 627294) of *L. rhamnosus* ATCC 53103 is inverted.

**Nucleotide sequence accession number.**

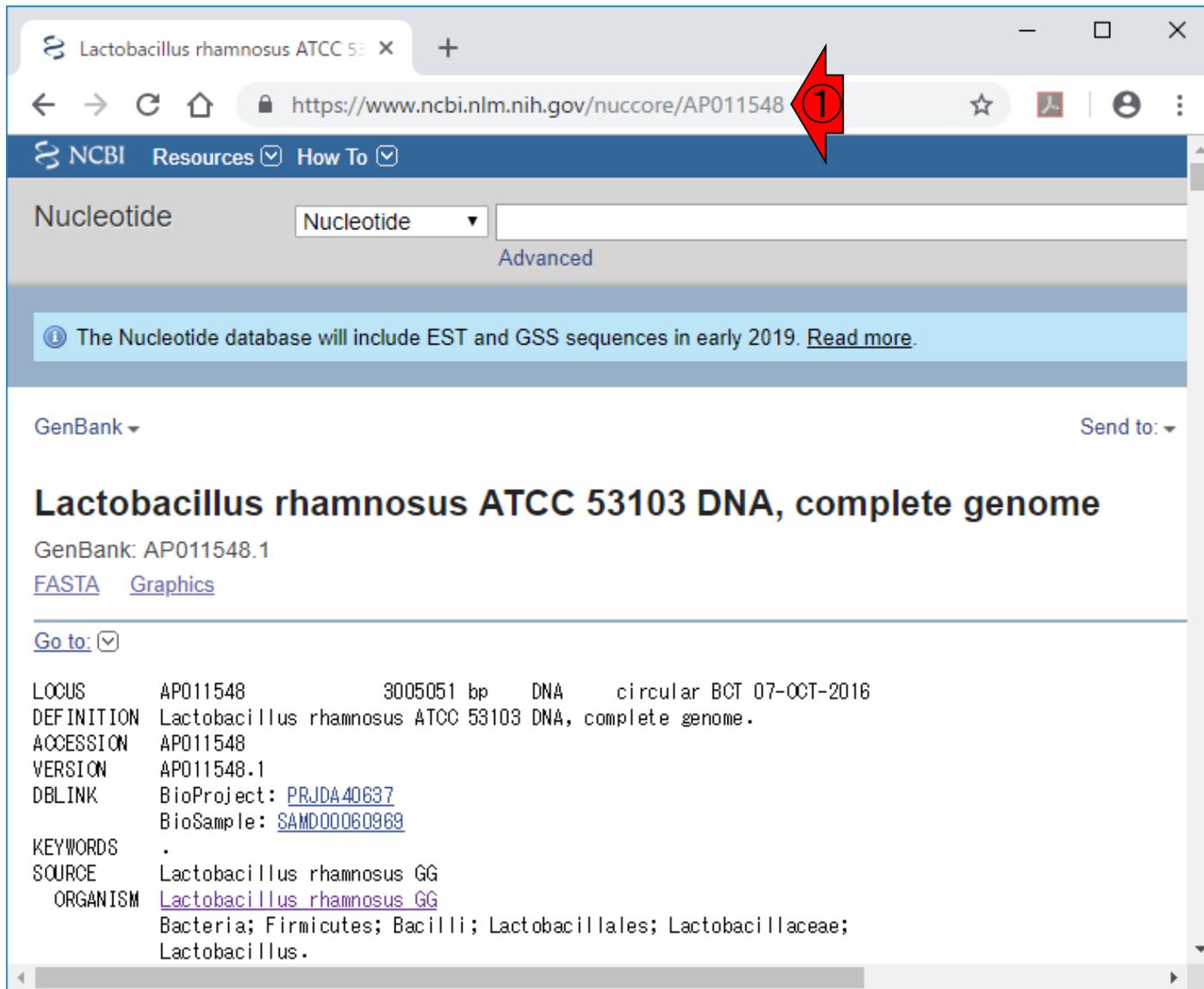
The sequence data for the *L. rhamnosus* ATCC 53103 genome are available in DDBJ/GenBank/EMBL under accession no. AP011548.

**ACKNOWLEDGMENTS**

PDF  
Help

①

# W6: AP011548.fasta



The screenshot shows a web browser window displaying the NCBI Nucleotide database page for the accession number AP011548. The browser's address bar shows the URL <https://www.ncbi.nlm.nih.gov/nucleotide/AP011548>. A red arrow with the number 1 points to this URL. The page content includes the title "Lactobacillus rhamnosus ATCC 53103 DNA, complete genome" and the GenBank accession number AP011548.1. Below the title, there are links for "FASTA" and "Graphics". The "FASTA" link is selected, and the sequence information is displayed in a table format.

LOCUS	AP011548	3005051 bp	DNA	circular	BCT 07-OCT-2016
DEFINITION	Lactobacillus rhamnosus ATCC 53103 DNA, complete genome.				
ACCESSION	AP011548				
VERSION	AP011548.1				
DBLINK	BioProject: <a href="#">PRJDA40637</a>				
	BioSample: <a href="#">SAM000060969</a>				
KEYWORDS	.				
SOURCE	Lactobacillus rhamnosus GG				
ORGANISM	<a href="#">Lactobacillus rhamnosus GG</a>				
	Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus.				

# W6: AP011548.fasta

Lactobacillus rhamnosus ATCC 53103 DNA, complete genome  
 3005051 bp DNA circular BCT 07-OCT-2016  
 Lactobacillus rhamnosus ATCC 53103 DNA, complete genome.

BioSample: [SAM000060969](#)

KEYWORDS: .  
 SOURCE: Lactobacillus rhamnosus GG  
 ORGANISM: [Lactobacillus rhamnosus GG](#)  
 Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae;  
 Lactobacillus.

www.ncbi.nlm.nih.gov を待機しています...



# W7 : Gepard

Gepard: a rapid and sensitive tool

https://www.ncbi.nlm.nih.gov/pubmed/17309896

NCBI Resources How To

PubMed.gov  
US National Library of Medicine  
National Institutes of Health

PubMed [Search]

Advanced

Format: Abstract Send to

Bioinformatics. 2007 Apr 15;23(8):1026-8. Epub 2007 Feb 19.

**Gepard: a rapid and sensitive tool for creating dotplots on genome scale.**

Krumsiek J<sup>1</sup>, Arnold R, Rattei T.

Author information

**Abstract**

Gepard provides a user-friendly, interactive application for the quick creation of dotplots. It utilizes suffix arrays to reduce the time complexity of dotplot calculation to  $\Theta(m \cdot \log n)$ . A client-server mode, which is a novel feature for dotplot creation software, allows the user to calculate dotplots and color them by functional annotation without any prior downloading of sequence or annotation data.

**AVAILABILITY:** Both source codes and executable binaries are available at <http://mips.gsf.de/services/analysis/gepard>

PMID: 17309896 DOI: [10.1093/bioinformatics/btr](https://doi.org/10.1093/bioinformatics/btr)

[Indexed for MEDLINE]

Full text links  
OXFORD ACADEMIC

Save items  
Add to Favorites

Similar articles  
JDotter: a Java interface to multiple do [Bioinformatics. :  
PhyloGena--a user-friendly system for [Bioinformatics. :  
Idiographica: a general-purp web applic. [Bioinformatics. :  
Review VEGA, the genome browser w [Brief Bioinform. :]



# W7: Gepard

The screenshot shows a web browser window with the address bar containing the URL `cube.univie.ac.at/gepard`. A red arrow with the number '1' points to the URL. The page header features the 'Cube' logo (Computational Systems Biology) and a search bar. Below the header is a 'NAVIGATION' menu. The main content area is divided into three sections: 'GENOME PAIR RAPID DOTTER (GEPARD)', 'CUBE NEWS', and 'LATEST PUBLICATIONS'. The 'GENOME PAIR RAPID DOTTER (GEPARD)' section contains a paragraph describing the tool and a reference: Krumsiek J, Arnold R, Rattei T. Gepard: A rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 2007; 23(8): 1026-8. PMID: 17309896. The 'CUBE NEWS' section lists three news items with dates: 'Happy new year 2019' (01.01.19), 'New paper in Nature Biotechnology' (21.12.18), and 'Hiring: Open group leader position at CUBE' (05.12.18). The 'LATEST PUBLICATIONS' section lists two items: 'Minimum Information about an Uncultivated Virus Genome (MIUViG)' and 'Insecticidal Toxicity of Involves the'.

# W7: Gepard

①論文自体は若干古いですが、②バクテリアゲノム程度ならこれで十分です。③ちょっと下に移動。

Genome Pair Rapid Dotter (gepa x +)

← → ↻ 🏠 ⓘ 保護されていない通信 | cube.univie.ac.at/gepard ☆ 👤 ⋮

**Cube**   
Computational Systems Biology

Search 

☰ NAVIGATION

## GENOME PAIR RAPID DOTTER (GEPARD)

Gepard (German: "cheetah", Backronym for "GEnome PAir - Rapid Dotter") allows the calculation of dotplots even for large sequences like chromosomes or bacterial genomes. Reference: Krumsiek J, Arnold R, Rattei T. Gepard: A rapid and sensitive method for creating dotplots on genome scale. *Bioinformatics* 2007;22(8): 1026-8. PMID: 17309896

**USE CASES**

Local comparison two of nucleotide or amino acid sequences from user-specified files. Batch dotplot functionality provided by command line access to Gepard.

**FEATURES**

### CUBE NEWS

- » [Happy new year 2019](#)  
01.01.19
- » [New paper in Nature Biotechnology](#)  
21.12.18
- » [Hiring: Open group leader position at CUBE](#)  
05.12.18

### LATEST PUBLICATIONS

- » [Minimum Information about an Uncultivated Virus Genome \(MIUViG\).](#)
- » [Insecticidal Toxicity of Involves the](#)

②このあたりまで移動。GepardはJavaプログラムなので、③Javaをインストールしておきましょう。

# W7: Gepard

Genome Pair Rapid Dotter (gepard) x +

保護されていない通信 | cube.univie.ac.at/gepard

gepard guarantees the privacy of all input data, does not store user data remotely and does not contain any form of malware.

## SYSTEM REQUIREMENTS

Gepard requires the Java Runtime Environment Version 5.0 or later (<http://www.java.com/download/>). It has been tested on the following operating systems:

- Microsoft Windows XP and later
- Linux/Un\*x systems
- MacOS 10.x

## DOWNLOAD

Please download the **jar file** if you want to run gepard on your computer (Java needs to be installed). On MacOS, navigate to your Downloads folder and open the jar file via right-click. Confirm the execution of this program.

## SOURCE CODE

The source code is available in our GitHub repository [gepard](#).

## TUTORIAL

An offline version of the tutorial is included in the download package

Javaセットアップ - 完了

Java  
ORACLE

✓ Javaが正常にインストールされました

Javaの更新が可能になると、プロンプトが表示されます。必ず更新をインストールして、パフォーマンスおよびセキュリティの最新の改善を反映させてください。  
[更新設定の詳細](#)

閉じる(C)

# W7: Gepard

①jar fileをクリックしてダウンロードしましょう。これがGepardプログラムの実体です。Javaプログラムの場合、Javaさえ予めインストールしておけば、「ダウンロード完了 = インストール完了」なのがよいところ

The screenshot shows a web browser window with the URL [cube.univie.ac.at/gepard](http://cube.univie.ac.at/gepard). The page content includes:

- SYSTEM REQUIREMENTS**  
Gepard requires the Java Runtime Environment Version 5.0 or later (<http://www.java.com/download/>). It has been tested on the following operating systems:
  - Microsoft Windows XP and later
  - Linux/Un\*x systems
  - MacOS 10.x
- DOWNLOAD** (indicated by a red arrow with a circled 1)  
Please download the [jar file](#) if you want to run gepard on your computer (java needs to be installed). On MacOS, navigate to your Downloads folder and open the jar file via right-click. Confirm the execution of this program.
- SOURCE CODE**  
The source code is available in our GitHub repository [gepard](#).
- TUTORIAL**  
An offline version of the tutorial is included in the download package

On the right side of the browser window, there is a sidebar with a search bar and a list of articles, including "The Genetic Transformation of Chlamydia pneumoniae". Below the search bar is a "CONTACT" section with the following information:

- Phone: +43 1 4277 76681
- Fax: +43 1 4277 876681
- Email: [contact.cube@univie.ac.at](mailto:contact.cube@univie.ac.at)

# W7: Gepard

基本どこでもよいと思いますが、ここでは①デスクトップ上に、②Gepard-1.40.jarファイルを、③保存しました。

The image shows a browser window displaying the Gepard website (cube.univie.ac.at/gepard) and a Windows File Explorer dialog box. The website content includes system requirements, download instructions, and source code information. The File Explorer dialog is titled "名前を付けて保存" (Save with name) and shows the file "Gepard-1.40.jar" being saved to the Desktop. Red arrows with numbers 1, 2, and 3 point to the Desktop location, the filename, and the Save button, respectively.

Genome Pair Rapid Dotter (gepa x +)

保護されていない通信 | cube.univie.ac.at/gepard

deglycosylation in *Lamium galeobdolon*.

» The Genetic Transformation of *Chlamydia pneumoniae*.

### SYSTEM REQUIREMENTS

Gepard requires the Java Runtime Environment Version 5.0 or later (<http://www.java.com/download/>). It has been tested on the following operating systems:

- Microsoft Windows XP and later
- Linux/Un\*x systems
- MacOS 10.x

### DOWNLOAD

Please download the [jar file](#) if you want to run gepard on your computer (java needs to be installed). On MacOS, navigate to your Downloads folder and open the jar file via right-click. Confirm the execution of the program.

### SOURCE CODE

The source code is available in our GitHub repository [gepard](#).

### TUTORIAL

An offline version of the tutorial is included in the download package

名前を付けて保存

PC > デスクトップ ①

整理 新しいフォルダー

名前	更新日時
hoge	2018/12/23 14:20
第13回	2019/01/29 15:22

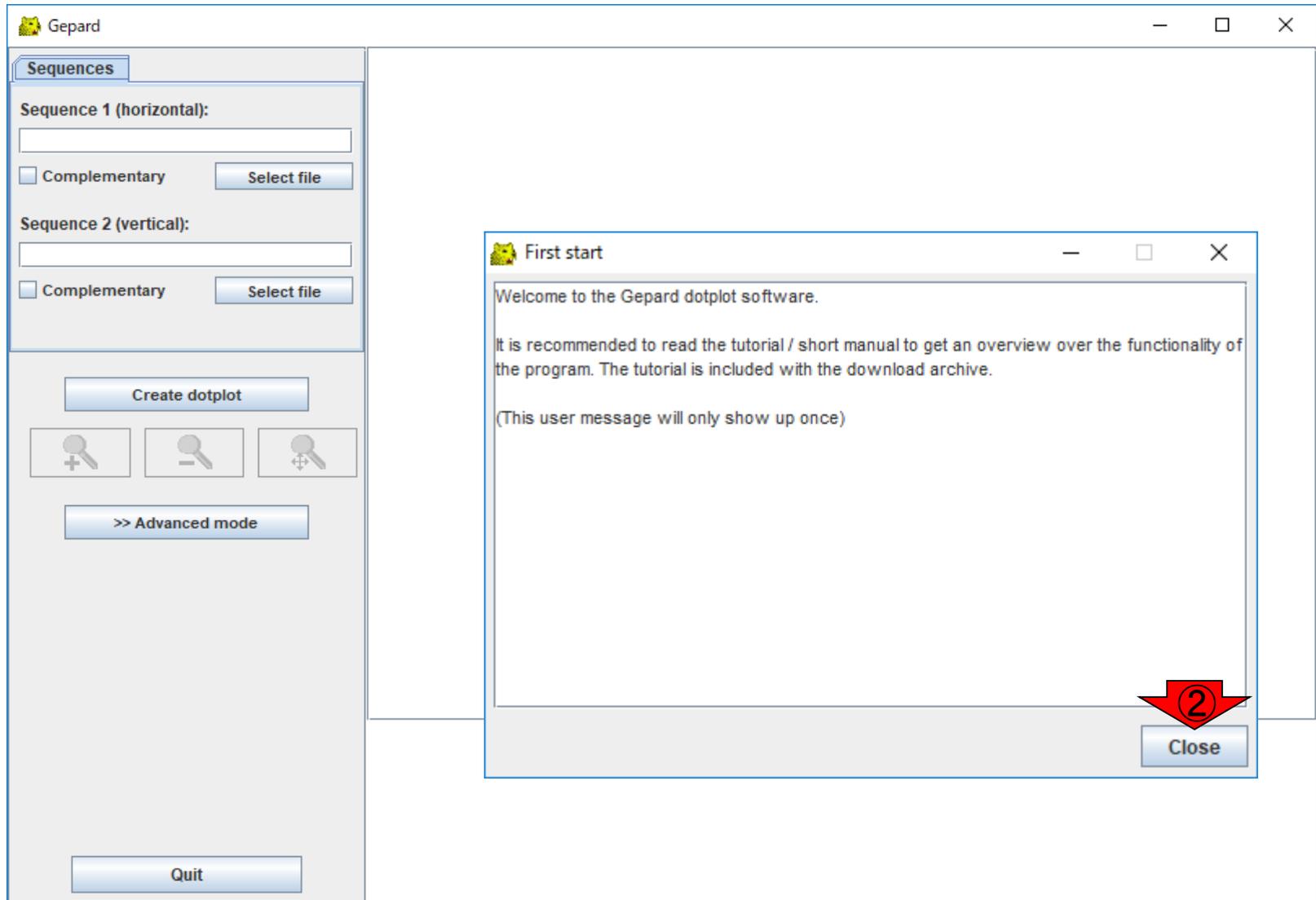
ファイル名(N): Gepard-1.40.jar ②

ファイルの種類(T): Executable Jar File

保存(S) ③ キャンセル

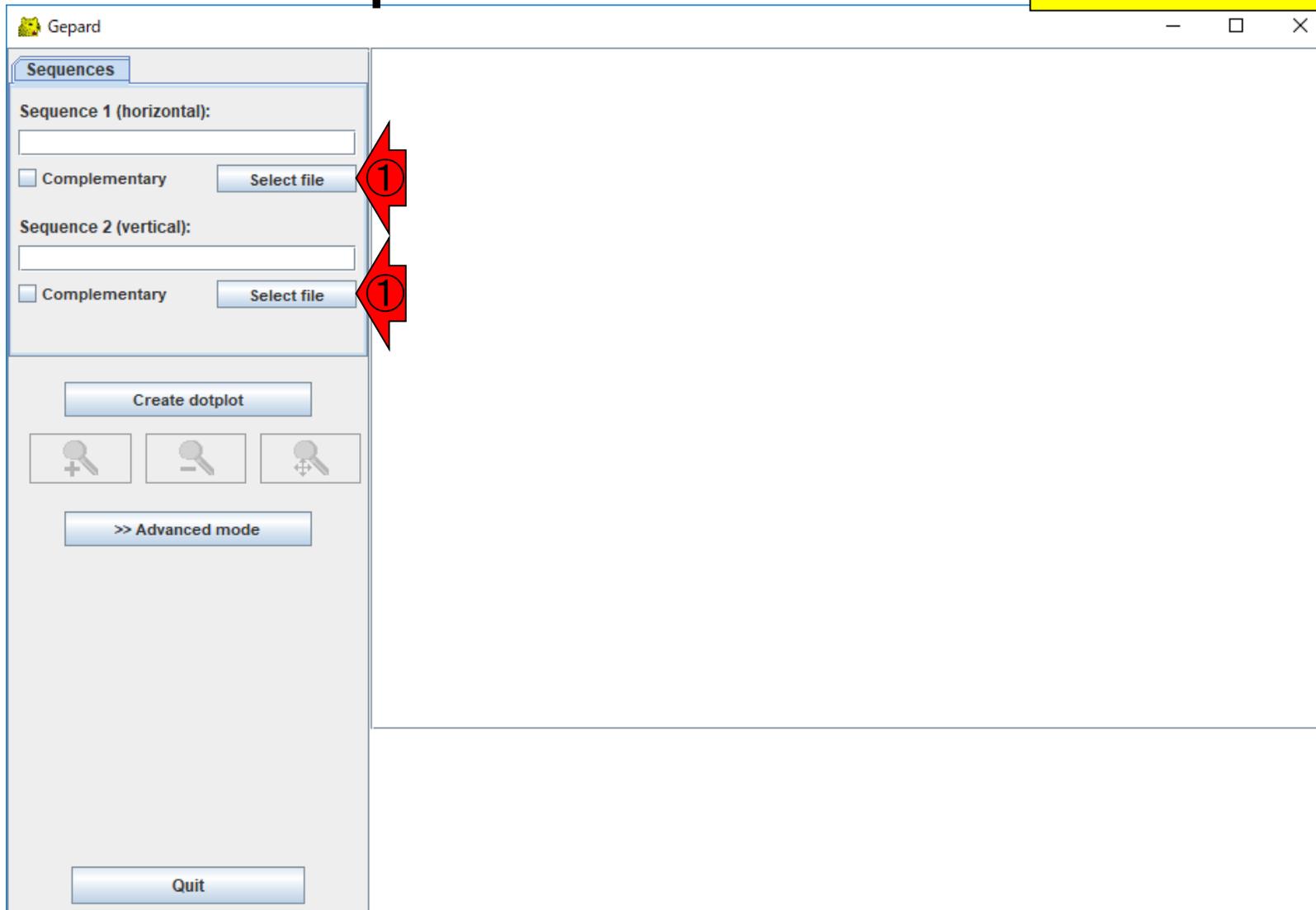
# W8: Gepard実行

①保存した場所にあるGepard-1.40.jarをダブルクリックして起動。②Close



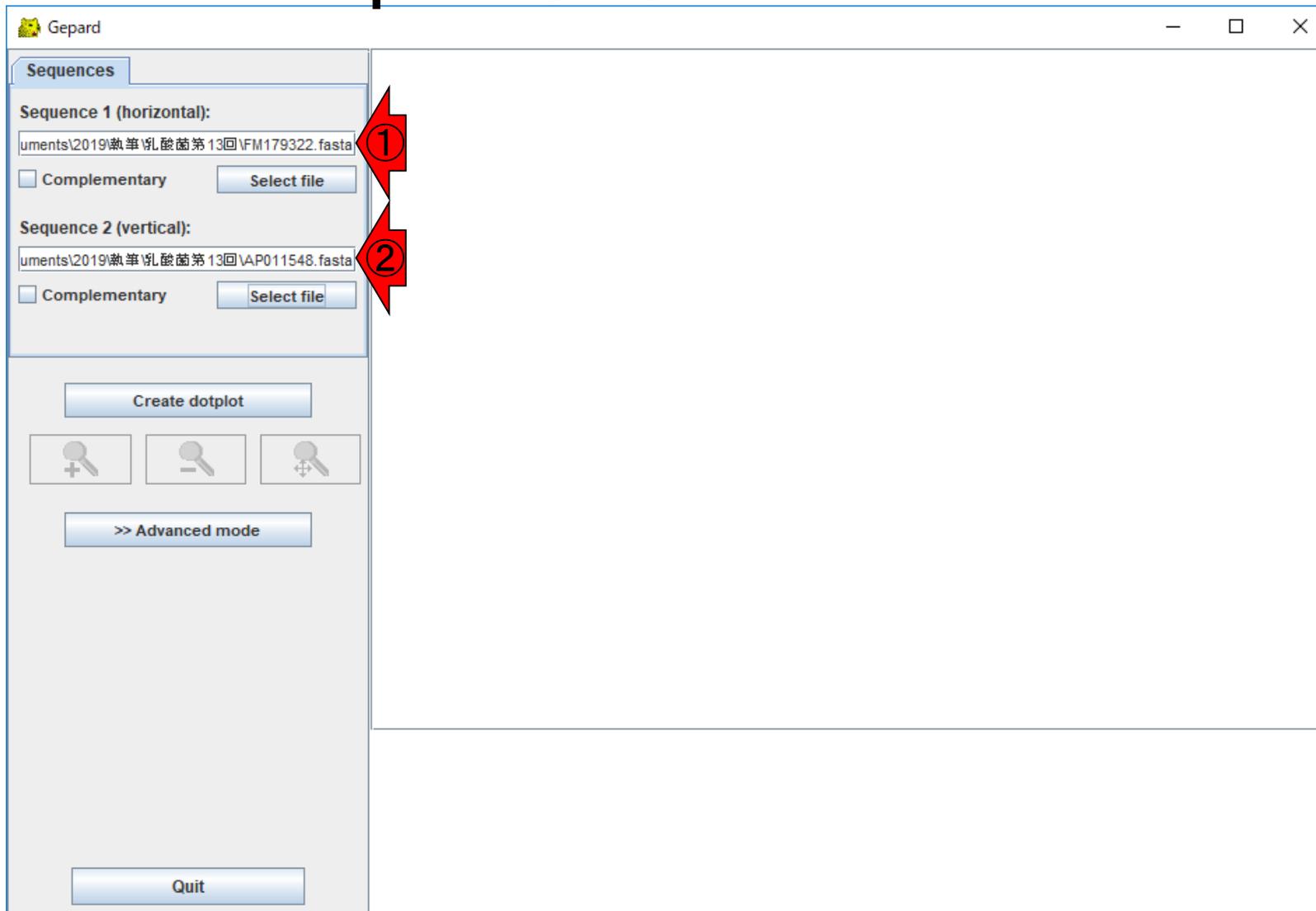
# W8: Gepard実行

① 予め作成しておいた比較したい2つのファイル(FM179322.fastaとAP011548.fasta)をそれぞれ指定。



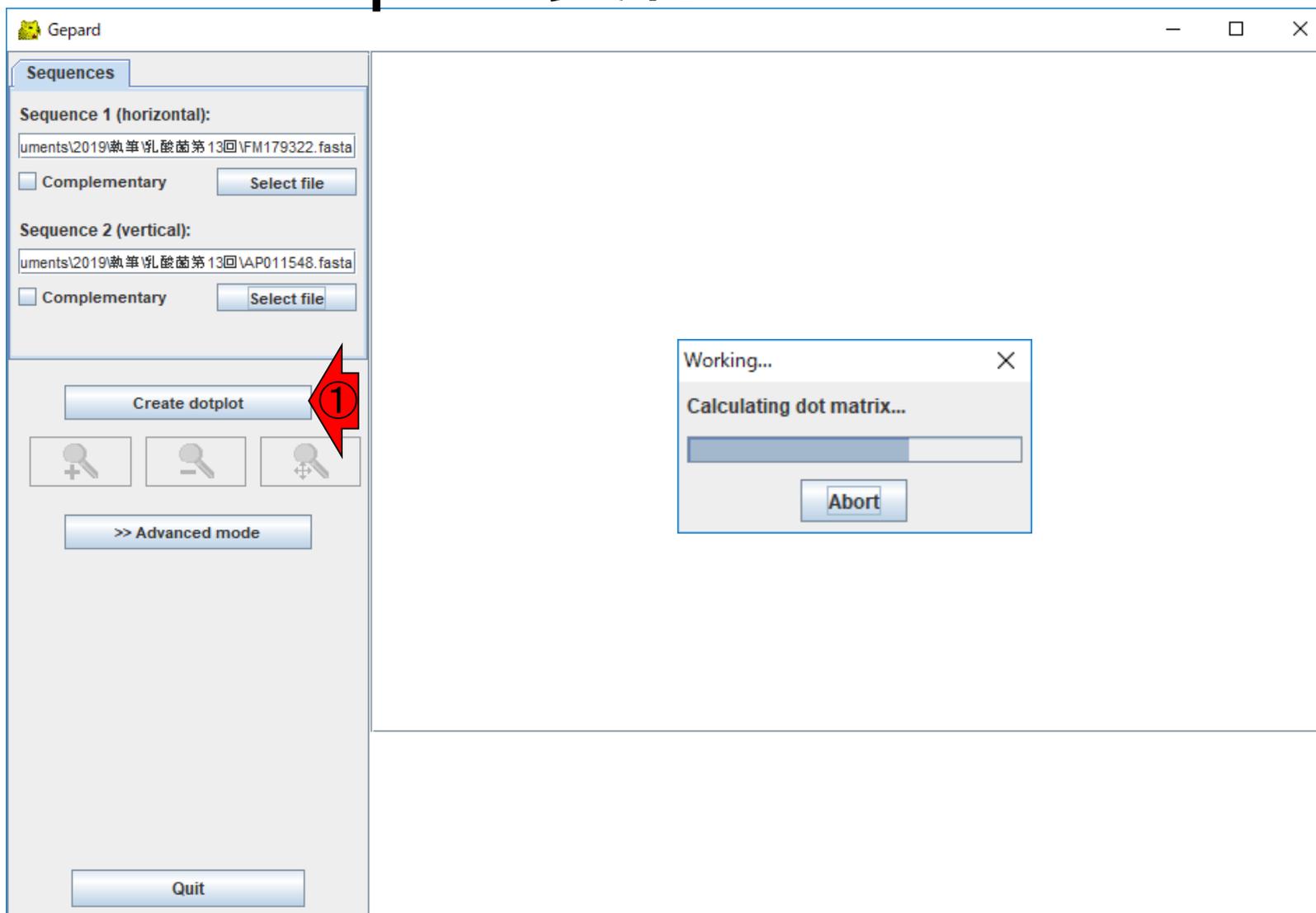
# W8: Gepard実行

ここでは、①FM179322.fastaと②AP011548.fastaのように指定しました。



# W8: Gepard実行

① Create dotplotを押して、数秒経過したところ。



# W8: Gepard実行

FM179322.1 Lactobacillus rhamnosus GG whole gen... vs. AP011548.1 Lactobacillus rhamnosus ATCC 53103 D... - Gepard

**Sequences**

Sequence 1 (horizontal):  
uments\2019\執筆\乳酸菌第13回\FM179322.fasta  
 Complementary

Sequence 2 (vertical):  
uments\2019\執筆\乳酸菌第13回\AP011548.fasta  
 Complementary

FM179322.1 Lactobacillus ... vs. AP011548.1 Lactobacillus ...  
Zoom: 8234 : 1  
Word length: 10  
Window size: 0  
Matrix: DNA  
GC ratio seq1: 0.4669  
GC ratio seq2: 0.4669  
Program: Gepard (1.40 final)

Hor : -  
Vert : -

Position: 3010110, 798698





# W9: Advanced mode

FM179322.1 Lactobacillus rhamnosus GG whole gen... vs. AP011548.1 Lactobacillus rhamnosus ATCC 53103 D... - Gepard

**Sequences**

Sequence 1 (horizontal):  
uments\2019\執筆\乳酸菌第13回\FM179322.fasta  
 Complementary

Sequence 2 (vertical):  
uments\2019\執筆\乳酸菌第13回\AP011548.fasta  
 Complementary

**①**

FM179322.1 Lactobacillus ... vs. AP011548.1 Lactobacillus ...  
Zoom: 8234 : 1  
Word length: 10  
Window size: 0  
Matrix: DNA  
GC ratio seq1: 0.4669  
GC ratio seq2: 0.4669  
Program: Gepard (1.40 final)

Hor : -  
Vert : -

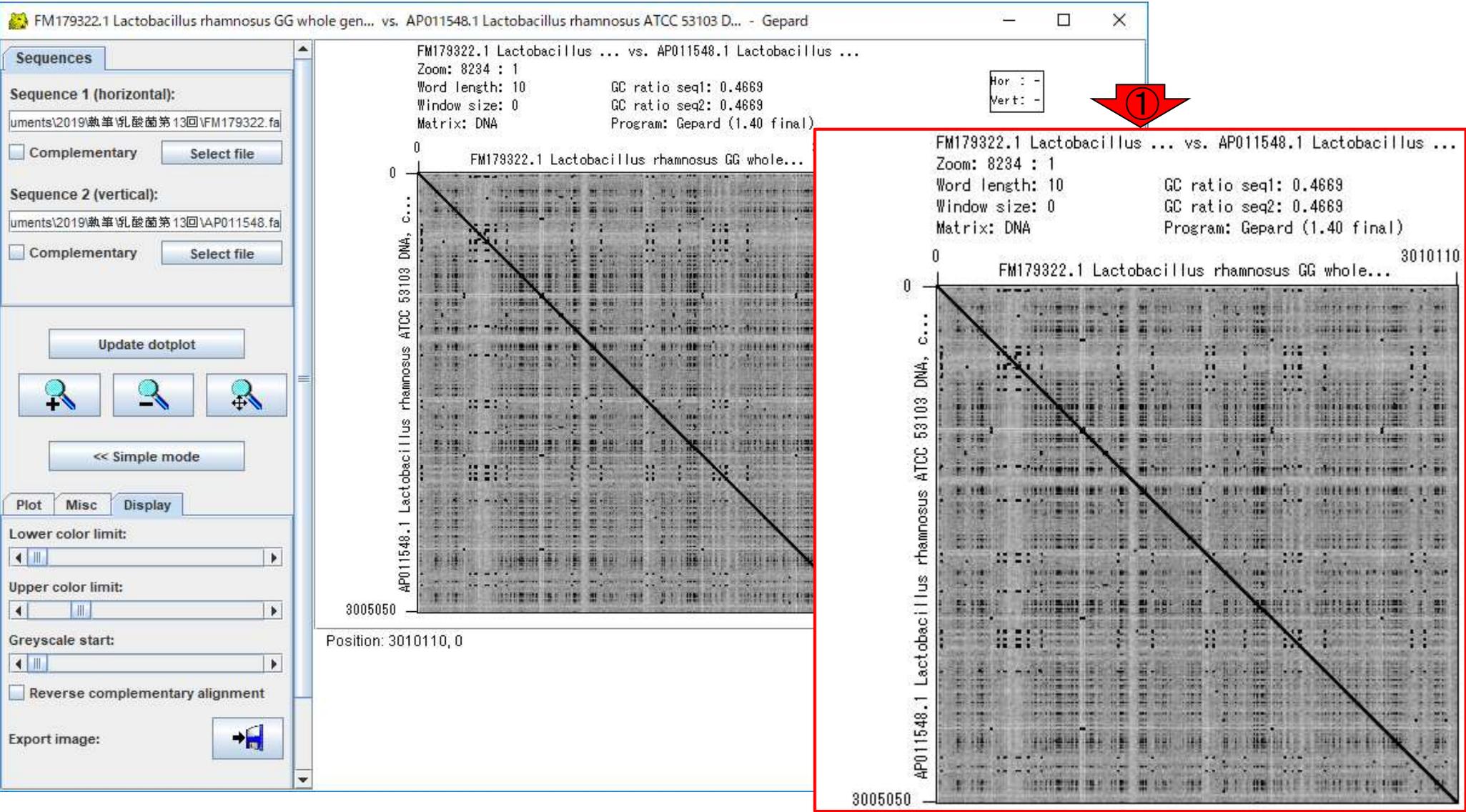
Position: 3010110, 378764

# W9: Advanced mode

①Export imageで、ファイルに保存することができます。ここでは、② Fig1a.pngとして保存しました。③OK

The screenshot displays a dotplot comparison of two DNA sequences: FM179322.1 Lactobacillus rhamnosus GG whole genome and AP011548.1 Lactobacillus rhamnosus ATCC 53103 DNA. The dotplot shows a strong diagonal correlation, indicating high sequence similarity. The left sidebar contains controls for sequence selection, zooming, and display options. The 'Export image' button is highlighted with a red arrow and the number 1. The '保存' (Save) dialog box is open, showing the file name 'Fig1a' and the file type 'PNG image files (.png)', both highlighted with a red box and the number 2. The 'Export' dialog box is also open, showing the message 'Dotplot successfully exported' and the 'OK' button highlighted with a red arrow and the number 3.

# W9: Advanced mode



①Plotタブで、任意の描画領域を指定可能です。

# W9: Advanced mode

The screenshot displays the Gepard software interface for sequence comparison. The main window title is "FM179322.1 Lactobacillus rhamnosus GG whole gen... vs. AP011548.1 Lactobacillus rhamnosus ATCC 53103 D... - Gepard".

**Sequences Panel (Left):**

- Sequence 1 (horizontal):** uments\2019\熟単乳酸菌第13回\FM179322.fa
- Sequence 2 (vertical):** uments\2019\熟単乳酸菌第13回\AP011548.fa
- Buttons: Complementary, Select file
- Update dotplot button
- Navigation icons: +, -, and a combined +/- icon
- << Simple mode button

**Plot Panel (Bottom Left):**

- Coordinates:
  - Start 1: 0
  - Stop 1: 3010110
  - Start 2: 0
  - Stop 2: 3005050
- Auto zoom:  Small plots:
- Zoom: 8234
- Auto params:
- Word length: 10
- Window size: 0
- Auto matrix:

**Plot Area (Center):**

- Text: FM179322.1 Lactobacillus ... vs. AP011548.1 Lactobacillus ...
- Parameters: Zoom: 8234 : 1, Word length: 10, Window size: 0, Matrix: DNA, GC ratio seq1: 0.4669, GC ratio seq2: 0.4669, Program: Gepard (1.40 final)
- Plot axes: Horizontal axis (0 to 3010110) labeled "FM179322.1 Lactobacillus rhamnosus GG whole..."; Vertical axis (0 to 3005050) labeled "AP011548.1 Lactobacillus rhamnosus ATCC 53103 DNA, c...".
- Position: 0, 1778544

A red arrow with the number "1" points to the "Plot" tab in the bottom left panel.

# W9: Advanced mode

①ここでは2つの配列ともに、領域 [610000, 640000 bp]としました。② Update dotplot。

FM179322.1 Lactobacillus rhamnosus GG whole gen... vs. AP011548.1 Lactobacillus rhamnosus ATCC 53103 D... - Gepard

**Sequences**

Sequence 1 (horizontal):  
uments\2019\熱帯乳酸菌第13回\FM179322.fa  
 Complementary

Sequence 2 (vertical):  
uments\2019\熱帯乳酸菌第13回\AP011548.fa  
 Complementary

②

**Plot** Misc Display

Coordinates

Start 1:	610000
Stop 1:	640000
Start 2:	610000
Stop 2:	640000

①

Auto zoom  Small plots  
Zoom: 8234

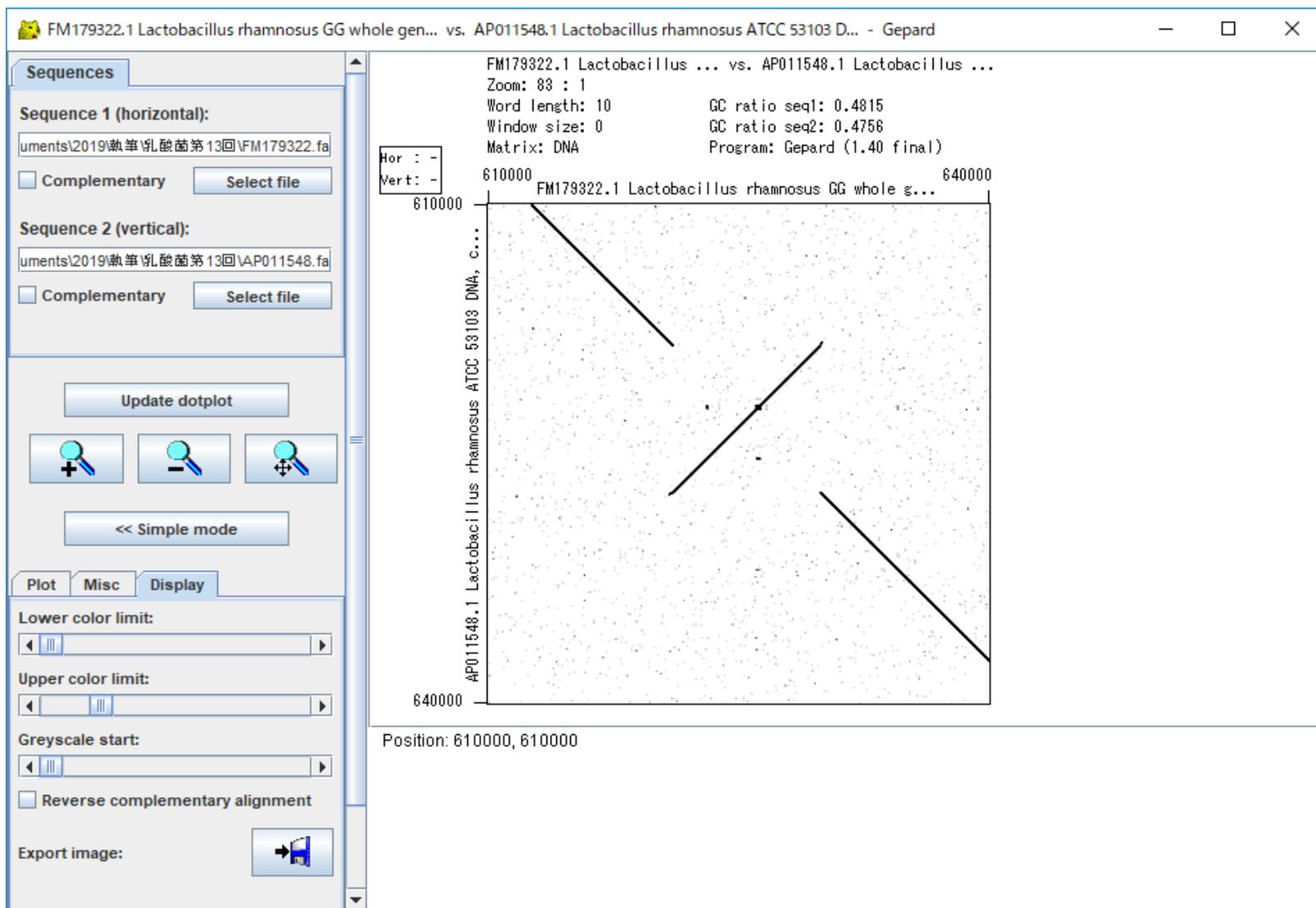
Auto params  
Word length: 10  
Window size: 0  
 Auto matrix

FM179322.1 Lactobacillus ... vs. AP011548.1 Lactobacillus ...  
Zoom: 8234 : 1  
Word length: 10  
Window size: 0  
Matrix: DNA  
GC ratio seq1: 0.4669  
GC ratio seq2: 0.4669  
Program: Gepard (1.40 final)

0 3010110  
FM179322.1 Lactobacillus rhamnosus GG whole...  
0 3005050  
AP011548.1 Lactobacillus rhamnosus ATCC 53103 DNA, c...  
Position: 0, 3005050

①確かに逆位になっている領域がありました。

# W9: Advanced mode



①Export image、②Fig1b.png  
として保存。③OK

# W9: Advanced mode

FM179322.1 Lactobacillus rhamnosus GG whole gen... vs. AP011548.1 Lactobacillus rhamnosus ATCC 53103 D... - Gepard

Sequences

Sequence 1 (horizontal):  
uments\2019\執筆乳酸菌第13回\FM179322.fa  
 Complementary

Sequence 2 (vertical):  
uments\2019\執筆乳酸菌第13回\AP011548.fa  
 Complementary

Update dotplot

Zoom: 83 : 1  
Word length: 10  
Window size: 0  
Matrix: DNA  
Program: Gepard

FM179322.1 Lactobacillus rhamnosus ... vs. AP011548.1 Lactobacillus rhamnosus ...  
GC ratio se  
GC ratio se  
Program: Ge

Hor : -  
Vert: -

610000  
610000  
FM179322.1 Lactobacillus rhamno

AP011548.1 Lactobacillus rhamnosus ATCC 53103 DNA, c...

640000

Position: 610000, 610000

保存

保存: Desktop

hoge  
第13回  
Fig1a.png

ファイル名(N): Fig1b.png

ファイルのタイプ(T): PNG image files (.png)

保存 取消

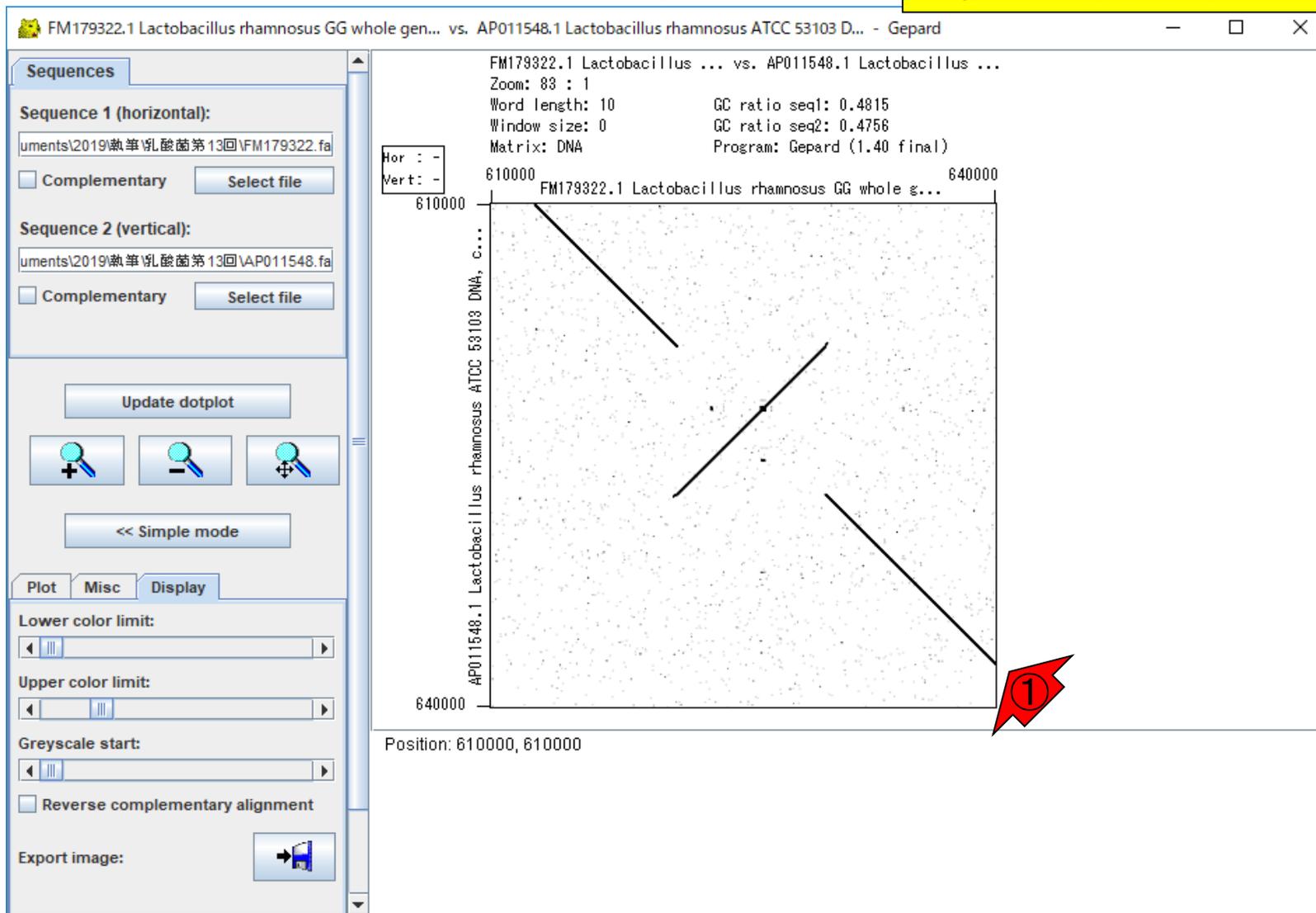
Export

Dotplot successfully exported

OK

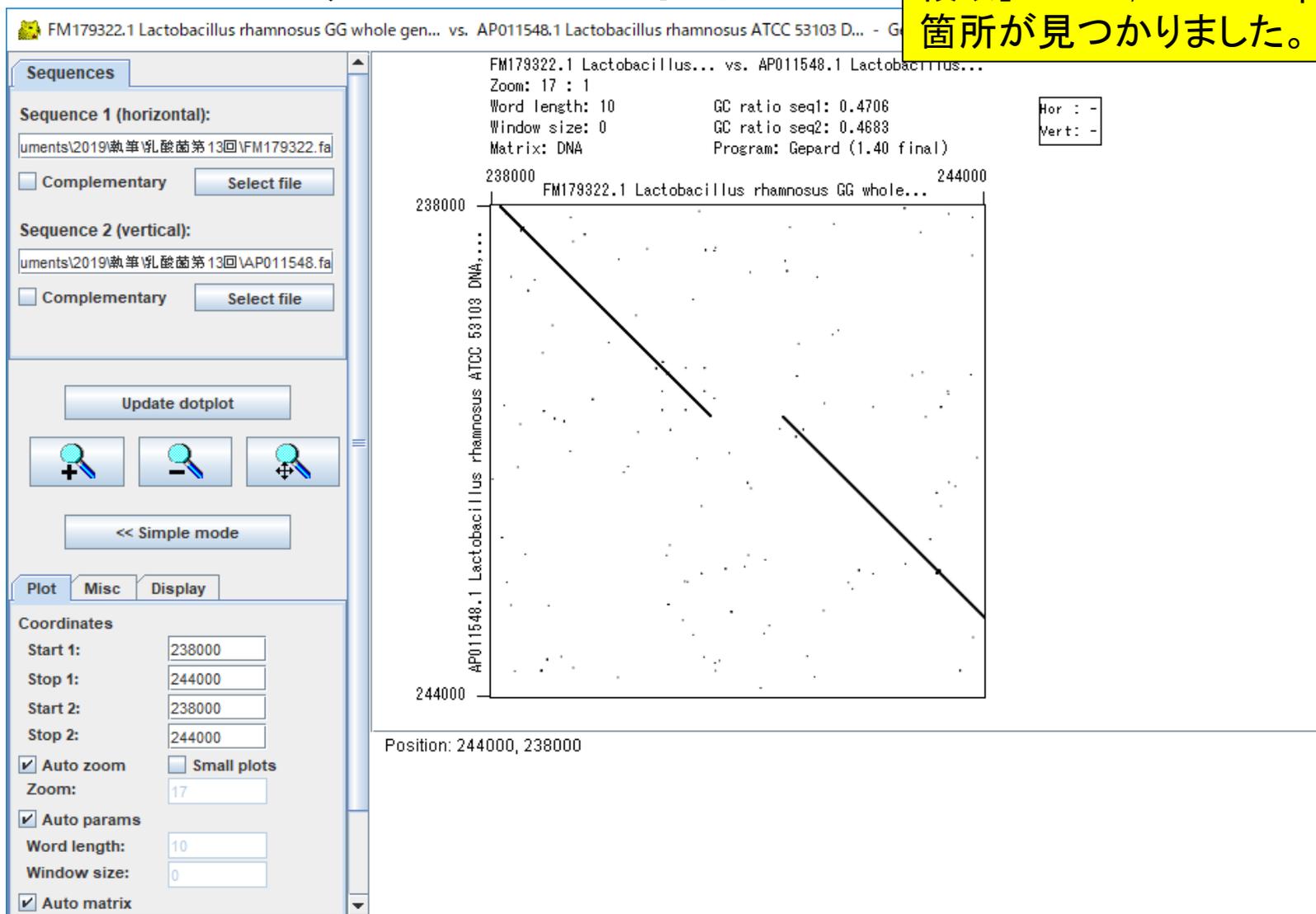
# W10: ずれの理由

①メインの直線が若干右上側にずれている理由は、FM179322にはあるがAP011548にはない領域が存在するためだと解釈すればよい。



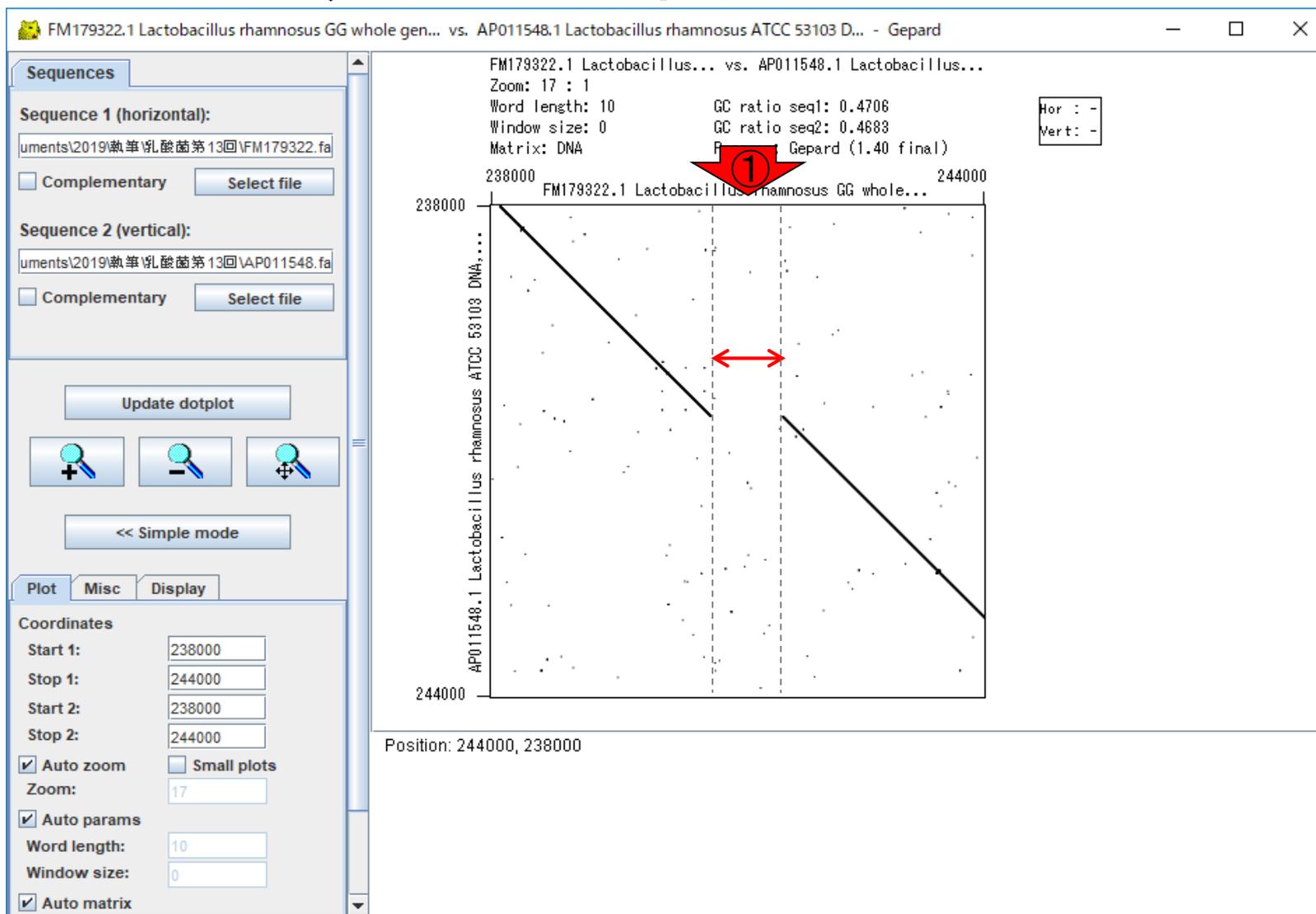
# W10: ずれの理由

ここではいろいろと領域を試行錯誤して、右上にずれ始めた箇所を探しました。結果として、領域[238000, 244000 bp]で説明がつけられる箇所が見つかりました。



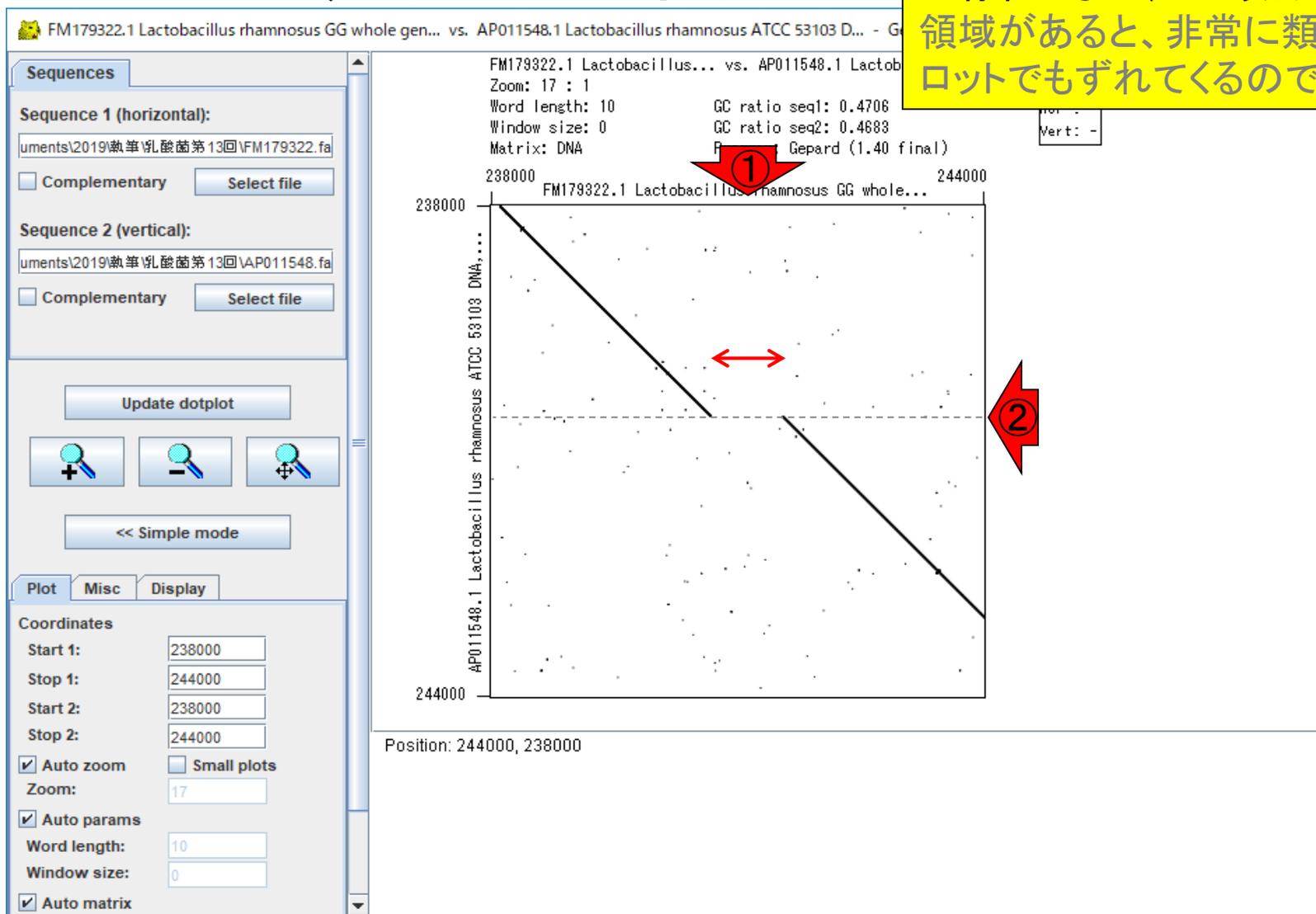
# W10: ずれの理由

①FM179322中に存在する赤矢印の長さに対応する部分の塩基配列が...



# W10: ずれの理由

①FM179322中に存在する赤矢印の長さに対応する部分の塩基配列が…、②AP011548には存在しない、という風に解釈します。こういう領域があると、非常に類似した配列のドットプロットでもずれてくるのです。



# W11 : Fig. 2a

## seqinrでドットプロット

[seqinr](#)パッケージ(Charif et al., *Bioinformatics*, 2005)中のdotPlot関数を用いてドットプロットを作成。

- W11 : Fig2a  
配列k同士のドットプロットです

```
seq1 <- "ACTCGTAGTCTATCATACGA"      #1つめの塩基配列を指定(配列k)
seq2 <- "ACTCGTAGTCTATCATACGA"      #2つめの塩基配列を指定(配列k)
out_f <- "Fig2a.png"                 #出力ファイル名を指定してout_fに格納
param_fig <- c(300, 300)              #ファイル出力時の横幅と縦幅を指定(単位はピクセル)

#必要なパッケージをロード
library(seqinr)                       #パッケージの読み込み

#ファイルに保存(pngファイル)
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])#出力ファイルの各
par(mar=c(0, 0, 0, 0))                #下、左、上、右の順で余白(行)を指定
dotPlot(s2c(seq1), s2c(seq2), xlabel="", ylab="")#プロット
dev.off()                               #おまじない
```

①

# W11 : Fig. 2a

こんな感じ。①作業ディレクトリ上に、  
②Fig2a.pngができていないはず。

## seqinrでドットプロット

[seqinr](#)パッケージ(Charif et al., *Bioinformatics*, 2005)中のdotPlot関数を用いてドットプロットを作成。

- W11 : Fig2a  
配列k同士のドットプロットです

```
seq1 <- "ACTCGTAGTCTATCATACGA" #1つめの塩基配列を指定(配列k)  
seq2 <- "ACTCGTAGTCTATCATACGA" #2つめの塩基配列を指定(配列k)  
out_f <- "Fig2a.png" #出力ファイル名を指定し  
param_fig <- c(300, 300) #ファイル出力時の横幅と高さ
```

```
#必要なパッケージをロード  
library(seqinr)
```

```
#ファイルに保存(pngファイル)  
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])  
par(mar=c(0, 0, 0, 0)) #下、左、上、右の順で余白を指定  
dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="") #プロット  
dev.off() #おまじない
```

```
RGui (64-bit)  
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ  
R Console  
> getwd()  
[1] "C:/Users/kadota/Desktop"  
> seq1 <- "ACTCGTAGTCTATCATACGA" #1つめの塩基配列を指定(配列k)  
> seq2 <- "ACTCGTAGTCTATCATACGA" #2つめの塩基配列を指定(配列k)  
> out_f <- "Fig2a.png" #出力ファイル名を指定し  
> param_fig <- c(300, 300) #ファイル出力時の横幅と高さ  
>  
> #必要なパッケージをロード  
> library(seqinr) #パッケージの読み込み  
>  
> #ファイルに保存(pngファイル)  
> png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])  
> par(mar=c(0, 0, 0, 0)) #下、左、上、右の順で余白を指定  
> dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="") #プロット  
> dev.off() #おまじない  
null device  
1  
> |
```

# W11 : Fig. 2a

実物です。同一配列同士のドットプロットなので、左下端から右上端にかけて、①キレイに直線状のドットが描かれていることがわかります。

## seqinrでドットプロット

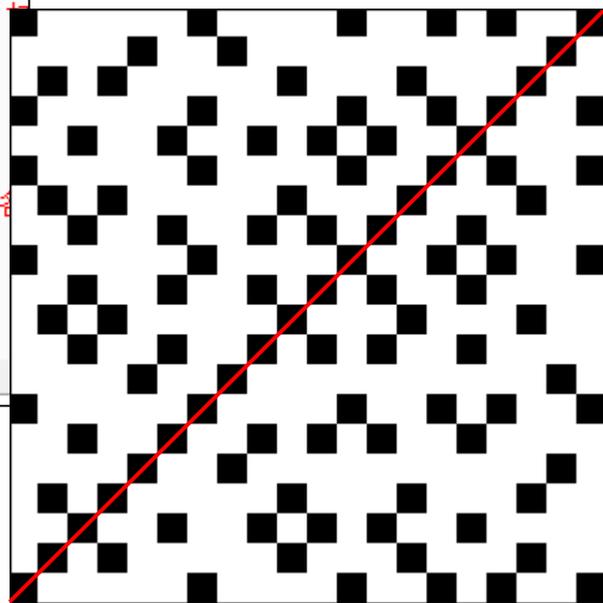
[seqinr](#)パッケージ(Charif et al., *Bioinformatics*, 2005)中のdotPlot関数を用いてドットプロットを作成。

- W11 : Fig2a  
配列k同士のドットプロットです

```
seq1 <- "ACTCGTAGTCTATCATACGA" #1つめの塩基配列を指定(配列k)
seq2 <- "ACTCGTAGTCTATCATACGA" #2つめの塩基配列を指定(配列k)
out_f <- "Fig2a.png" #出力ファイル名を指定してout_fに格納
param_fig <- c(300, 300) #ファイル出力時の横幅と縦幅を指定(単位はピクセル)

#必要なパッケージをロード
library(seqinr) #パッケージの読み込み

#ファイルに保存(pngファイル)
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2]) #出力ファイルの名前
par(mar=c(0, 0, 0, 0)) #下、左、上、右の順で余白(行)を指定
dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="") #プロット
dev.off() #おまじない
```



# W12: Fig. 2b

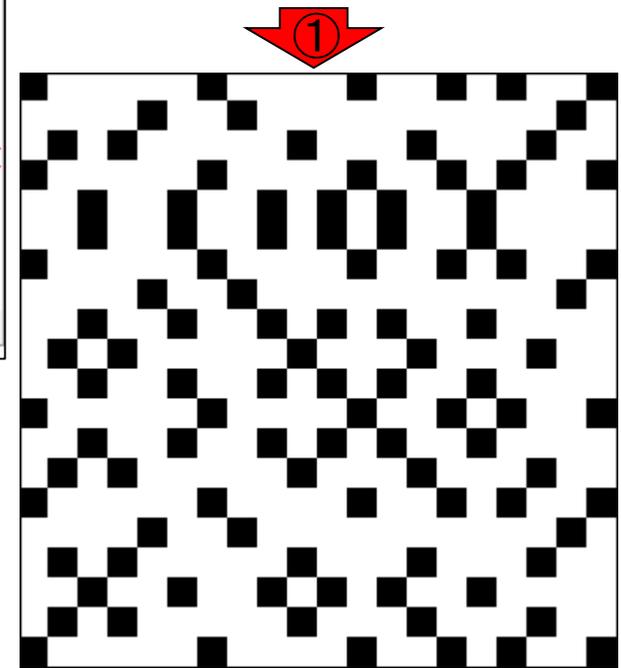
- W12 : Fig2b

配列kとのドットプロットです

```
seq1 <- "ACTCGTAGTCTATCATACGA"      #1つめの塩基配列を指定(配列k)
seq2 <- "ACTCGACTATCTCTATACGA"      #2つめの塩基配列を指定(配列1)
out_f <- "Fig2b.png"                  #出力ファイル名を指定してout_fに格納
param_fig <- c(300, 300)              #ファイル出力時の横幅と縦幅を指定(単位はピクセル)

#必要なパッケージをロード
library(seqinr)                        #パッケージの読み込み

#ファイルに保存(pngファイル)
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])#出力ファイルの各種
par(mar=c(0, 0, 0, 0))                #下、左、上、右の順で余白(行)を指定
dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="")#プロット
dev.off()                              #おまじない
```



# W12: Fig. 2b

- W12 : Fig2b  
配列kとのドットプロットです

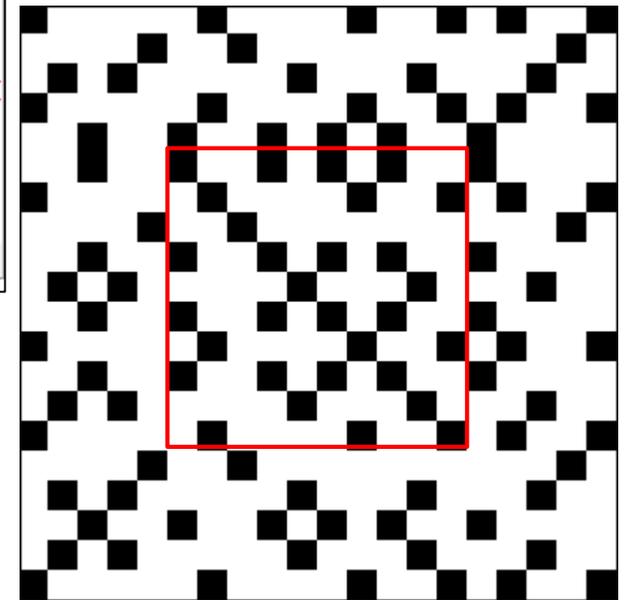
```

seq1 <- "ACTCGTAGTCTATCATACGA"      #1つめの塩基配列を指定(配列k)
seq2 <- "ACTCGACTATCTGATTACGA"      #2つめの塩基配列を指定(配列1)
out_f <- "Fig2b.png"                 #出力ファイル名を指定してout_fに格納
param_fig <- c(300, 300)             #ファイル出力時の横幅と縦幅を指定(単位はピクセル)

#必要なパッケージをロード
library(seqinr)                       #パッケージの読み込み

#ファイルに保存(pngファイル)
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])#出力ファイルの各種
par(mar=c(0, 0, 0, 0))               #下、左、上、右の順で余白(行)を指定
dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="")#プロット
dev.off()                              #おまじない

```



# W12: Fig. 2b

確かに反転させることで、赤枠の領域内のドットが「左下端から右上端」ではなく「左上端から右下端」に切り替わっていることがわかる。

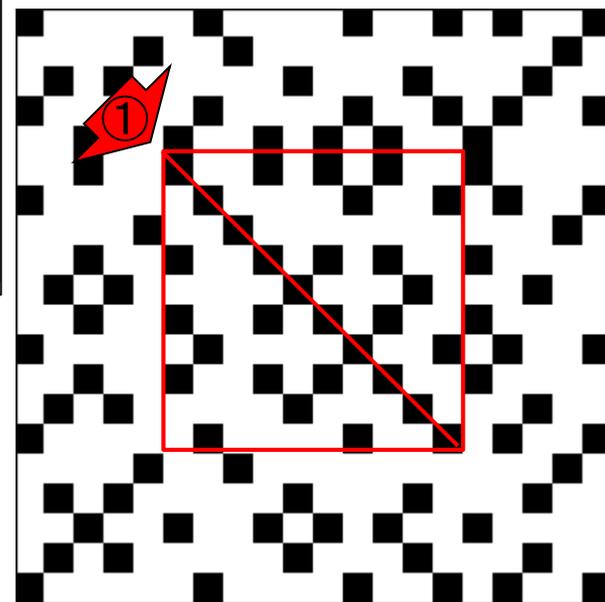
## W12 : Fig2b

配列kとlのドットプロットです

```
seq1 <- "ACTCGTAGTCTATCATACGA"      #1つめの塩基配列を指定(配列k)
seq2 <- "ACTCGACTATCTGATTACGA"      #2つめの塩基配列を指定(配列l)
out_f <- "Fig2b.png"                 #出力ファイル名を指定してout_fに格納
param_fig <- c(300, 300)              #ファイル出力時の横幅と縦幅を指定(単位はピクセル)

#必要なパッケージをロード
library(seqinr)                       #パッケージの読み込み

#ファイルに保存(pngファイル)
png(out_f, pointsize=13, width=param_fig[1], height=param_fig[2])#出力ファイルの各種
par(mar=c(0, 0, 0, 0))               #下、左、上、右の順で余白(行)を指定
dotPlot(s2c(seq1), s2c(seq2), xlab="", ylab="")#プロット
dev.off()                             #おまじない
```



# W13: Ensembl Bacteria

①Ensembl Bacteriaの、②L.  
*rhamnosus* GGのページ。③と④から  
得られるファイルと同じものが…

Lactobacillus rhamnosus GG - En - x +

← → ↻ 🏠 ⓘ 保護されていない通信 | bacteria.ensembl.org/Lactobacillus\_rhamnosus\_gg/Info/Index ☆ 👤 | 🧑 Login/Register

**EnsemblBacteria** ① MMR | BLAST | Tools | Downloads | Documentation | More ▾  🔍

**Lactobacillus rhamnosus GG** ②

*Lactobacillus rhamnosus* GG

Provider [European Nucleotide Archive](#) | Taxonomy ID [568703](#)

e.g. [is1](#) or [FM179322:17235-18251](#) or [synthetase](#)

**About *Lactobacillus rhamnosus* GG**

📄 Information and statistics

**Genome assembly:** [ASM2650v1](#)

📄 More information and statistics

📄 Download DNA sequence (FASTA) ③

📄 Display your data in Ensembl Bacteria

**Gene annotation**

What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

📄 More about this genebuild

📄 Download genes, cDNAs, ncRNA, proteins - FASTA - GFF3 ④

🔑 Update your old Ensembl IDs

**Comparative genomics**

What can I find? Gene families based on HAMAP and PANTHER classification.

📄 More about comparative analyses

📄 Phylogenetic overview of gene families

**Variation**

This species currently has no variation database. However you can process your own variants using the Variant Effect Predictor.

[Variant Effect Predictor](#) **Ve!P**

Ensembl Bacteria release 42 - Dec 2018 © EMBL-EBI

# W13: Ensembl Bacteria

①Ensembl Bacteriaの、②L.  
*rhamnosus* GGのページ。③と④から  
得られるファイルと同じものが⑤です

Ensembl Bacteria

Lactobacillus rhamnosus GG

Lactobacillus rhamnosus GG

Provider [European Nucleotide Archive](#) | Taxonomy ID [568703](#)

Search Lactobacillus rhamnosus GG...

e.g. [is1](#) or [FM179322:17235-18251](#) or [synthetase](#)

About *Lactobacillus rhamnosus* GG

- Information and statistics

Genome assembly: [ASM2650v1](#)

- More information and statistics
- Download DNA sequence (FASTA)
- Display your data in Ensembl Bacteria

Comparative genomics

What can I find? Gene families based on HAMAP and PANTHER classification.

- More about comparative analyses
- Phylogenetic overview of gene families

What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

- More about this genebuild
- Download genes, cDNAs, ncRNA, proteins - FASTA - GFF3
- Update your old Ensembl IDs

Variation

This species currently has no variation database. However you can process your own variants using the Variant Effect Predictor.

[Variant Effect Predictor](#)

Ensembl Bacteria release 42 - Dec 2018 © EMBL-EBI

• W13: [Ensembl Bacteria : Cunningham et al., Nucleic Acids Res., 2019](#)

- [Lactobacillus rhamnosus GG](#)

オリジナルはファイル名が長いので短くしています。念のためここでもダウンロードできるようにしています。

- release42のゲノム配列ファイル: [ASM2650v1.fa](#) (約3MB)

- release42のアノテーションファイル: [ASM2650v1.gff3](#) (約2MB)

# W14: SRA

公共DBの1つである①SRA上で、②「*Lactobacillus RNA-seq*」で、③キーワード検索。

The screenshot shows the NCBI SRA homepage. A search bar at the top contains the text "Lactobacillus RNA-seq" and a "Search" button. Three red arrows with circled numbers point to specific elements: arrow 1 points to the "SRA" header, arrow 2 points to the search input field, and arrow 3 points to the "Search" button. Below the header, there are three columns of links: "Getting Started", "Tools and Software", and "Related Resources".

Home - SRA - NCBI  
https://www.ncbi.nlm.nih.gov/sra/

NCBI Resources How To Sign in to NCBI

SRA SRA Lactobacillus RNA-seq Search

**SRA**  
Sequence Read Archive (SRA) makes biological sequence data available to the research community to enhance reproducibility and allow for new discoveries by comparing data sets. The SRA stores raw sequencing data and alignment information from high-throughput sequencing platforms, including Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD System®, Helicos Heliscope®, Complete Genomics®, and Pacific

**Getting Started**  
[How to Submit](#)  
[Log in to SRA \(for updating and troubleshooting submissions\)](#)  
[Log in to Submission Portal \(for submitting sequence data\)](#)  
[SRA Documentation](#)  
[Download Guide](#)  
[SRA Fact Sheet \(.pdf\)](#)

**Tools and Software**  
[Download SRA Toolkit](#)  
[SRA Toolkit Documentation](#)  
[SRA-BLAST](#)  
[SRA Run Browser](#)  
[SRA Run Selector](#)

**Related Resources**  
[Submission Portal](#)  
[Trace Archive](#)  
[dbGaP Home](#)  
[BioProject](#)  
[BioSample](#)

# W14: SRA

①688個もリストアップされているが、②プロジェクトとしては14個しかないようだ。②をクリック

Access  
Controlled (32)  
Public (656)

Source  
DNA (215)  
RNA (473)

Other  
aligned data (51)

Clear all  
[Show additional filters](#)

Summary ▾ 20 per page ▾  
Send to ▾ Filters: [Manage Filters](#)

View results as an expanded interactive table using the RunSelector. [Send results to Run selector](#)

**Search results**  
Items: 1 to 20 of 688

1. [BACTERIA](#)  
1 LS454 (454 GS) run: 80,474 spots, 20.1M bases, 11.7Mb downloads  
Accession: SRX5295446

2. [BACTERIA](#)  
1 LS454 (454 GS) run: 79,901 spots, 20M bases, 11.8Mb downloads  
Accession: SRX5295445

3. [BACTERIA](#)  
1 LS454 (454 GS) run: 80,017 spots, 20M bases, 11.6Mb downloads  
Accession: SRX5295444

4. [BACTERIA](#)  
1 LS454 (454 GS) run: 80,080 spots, 20M bases, 11.7Mb downloads  
Accession: SRX5295443

**Results by taxon**  
Top Organisms [\[Tree\]](#)  
Lactobacillus plantarum (139)  
Lactobacillus acidophilus (56)  
Lactobacillus (472)  
Lactobacillus animalis (47)  
Homo sapiens (41)  
All other taxa (352)  
More...

**Search in related databases**

Database	Access		all
	public	controlled	
BioSample	<a href="#">11</a>		<a href="#">11</a>
BioProject	<a href="#">14</a>		<a href="#">1</a> <a href="#">15</a>
dbGaP		<a href="#">2</a>	<a href="#">2</a>
GEO Datasets	<a href="#">6</a>		<a href="#">6</a>

**Find related data**

# W14: SRA

①確かに14個になった。②プロジェクト数は、実質的にsubmission数ということなのだろう。③データの種別が発現データ(Transcriptome or Gene expression)であることが明記されている、④をクリックしてみる。

NCBI Resources How To Sign in to NCBI

BioProject (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NOT bioproject\_gap[filter] Search

Create alert Advanced Browse by Project attributes Help

Project Types Display Settings: Primary submission (14) Summary, 20 per page, Sorted by Default order Send to: Filters: Manage Filters

Data Types Other (2) Transcriptome (10) See also 198 genomes matching your organism search

Project Data SRA (14) GEO DataSets (6) Search results Items: 14

Scope Monoisolate (2) Multi-isolate (7) Multi-species (4) Other (1)

Organism Groups Bacteria (12) Fungi (1) Other vertebrates (1)

Clear all Show additional filters

1. RNA-seq of *Lactobacillus plantarum* LY-78  
Project data type: Raw sequence reads  
Scope: Multispecies  
Heilongjiang Bayi Agricultural University  
Accession: PRJNA483366 ID: 483366

2. How *Lactobacillus plantarum* Shapes its Transcriptome in Response to Contrasting Habitats  
Taxonomy: *Lactobacillus plantarum*  
Project data type: Transcriptome or Gene expression  
Scope: Multiisolate  
della Pianta e degli Alimenti - Di.S.S.P.A, Dipartimento di Scienze del Suolo, Università  
Accession: PRJNA475032 ID: 475032

3. Lactobacillus rhamnosus strain:SCT-10-10-60

Find related data Database: Select Find items

Search details ("Lactobacillus"[Organism] OR Lactobacillus[All Fields]) AND RNA-seq[All Fields] AND bioproject\_sra[filter] NOT bioproject\_gap[filter] Search See more...

Recent activity Turn Off Clear Your browsing activity is empty.

# W14: SRA

①原著論文がある場合は、ここに表示されるようだ。②ページ下部に移動。

How Lactobacillus plantarum Shapes its Transcriptome in Response to Contrasting Habitats

Accession: PRJNA475032 ID: 475032

Triplets of Lactobacillus plantarum strains were isolated from nine contrasting habitats. [More...](#)

Accession	PRJNA475032; GEO: GSE115448
Data Type	Transcriptome or Gene expression
Scope	Multisolate
Organism	<b>Lactobacillus plantarum</b> [Taxonomy ID: 1590] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus plantarum
Publications	Filannino P <i>et al.</i> , "How Lactobacillus plantarum shapes its transcriptome in response to contrasting habitats.", <i>Environ Microbiol</i> , 2018 Oct 5;20(10):3700-3716
Submission	Registration date: 7-Jun-2018 della Pianta e degli Alimenti - Di.S.S.P.A, Dipartimento di Scienze del Suolo, Università degli Studi di Bari "Aldo Moro"
Relevance	Industrial

See Genome Information for Lactobacillus plantarum

NAVIGATE ACROSS  
309 additional projects are related by organism.

Send to:

Related information

- BioSample
- Genome
- GEO DataSets
- PubMed
- SRA
- Taxonomy

Recent activity

- How Lactobacillus plantarum Shapes its Transcriptome BioProject [Turn Off](#) [Clear](#)
- (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject [See more...](#)

# W14: SRA

① サンプル数は36ということなのだろう。のちにGalaxy上で実行するにはちょっと多すぎるので、他を探すこととする。1つ前のページに移動。

How Lactobacillus plantarum She x +

https://www.ncbi.nlm.nih.gov/bioproject/475032

transcriptome in response to contrasting habitats.", *Environ Microbiol*, 2018 Oct 5;20(10):3700-3716

Submission Registration date: 7-Jun-2018  
della Pianta e degli Alimenti - Di.S.S.P.A, Dipartimento di Scienze del Suolo, Università degli Studi di Bari "Aldo Moro"

Relevance Industrial

(Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject  
See more...

**Project Data:**

Resource Name	Number of Links
SEQUENCE DATA	
SRA Experiments	36
PUBLICATIONS	
PubMed	1
OTHER DATASETS	
BioSample	36
GEO DataSets	1

▼ GEO Data Details

Parameter	Value
Data volume, Supplementary Mbytes	1

▼ SRA Data Details

Parameter	Value
Data volume, Gbases	144
Data volume, Mbytes	57383



# W15: *L. rhamnosus* GG

NCBI Resources How To Sign in to NCBI

BioProject BioProject (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NOT bioproject\_g Search

Create alert Advanced Browse by Project attributes Help

Project Types  
Primary submission (14)

Data Types  
Other (2)  
Transcriptome (10)

Project Data  
SRA (14)  
GEO DataSets (6)

Scope  
Monoisolate (2)  
Multi-isolate (7)  
Multi-species (4)  
Other (1)

Organism Groups  
Bacteria (12)  
Fungi (1)  
Other vertebrates (1)

Display Settings: Summary, 20 per page, Sorted by Default order

Send to:

See also 198 genomes matching your organism search

Filters: Manage Filters

Find related data  
Database: Select  
Find items

Search details  
(("Lactobacillus"[Organism] OR Lactobacillus[All Fields]) AND RNA-seq[All Fields]) AND bioproject\_sra[filter] NOT bioproject\_gap[filter]

Search See more...

Recent activity  
Turn Off Clear  
Your browsing activity is empty.

Search results  
Items: 14

- [RNA-seq of \*Lactobacillus plantarum\* LY-78](#)
- 1. Project data type: Raw sequence reads  
Scope: Multispecies  
Heilongjiang Bayi Agricultural University  
Accession: PRJNA483366 ID: 483366
- [How \*Lactobacillus plantarum\* Shapes its Transcriptome in Response to Contrasting Habitats](#)
- 2. Taxonomy: *Lactobacillus plantarum*  
Project data type: Transcriptome or Gene expression  
Scope: Multiisolate  
della Pianta e degli Alimenti - Di.S.S.P.A, Dipartimento di Scienze del Suolo, Università  
Accession: PRJNA475032 ID: 475032
- [Lactobacillus rhamnosus strain:SCT-10-10-60](#)

Clear all  
Show additional filters

# W15: *L. rhamnosus*

①このあたりまで移動。②*L. rhamnosus* GGの、③RNA-seqデータを発見したのでクリック。実際にはこのRNA-seqデータに着目した後に、リファレンスゲノム配列を取得しているのだが、細かいことは気にしない

(Lactobacillus RNA-seq) AND bio × +

https://www.ncbi.nlm.nih.gov/bioproject/?term=(Lactobacillus%20RNA-seq)%20AND%20...

- [Bilophila wadsworthia worsens high-fat diet-induced metabolic impairments in inflammation dependent and independent manners \[bacterial gene expression\]](#)  
Taxonomy: *uncultured bacterium*  
Project data type: Transcriptome or Gene expression  
Scope: Multiisolate  
ILTOO Pharma  
Accession: PRJNA445875 ID: 445875
- [RNA-seq analysis of \*Lactobacillus\* at acidic stress](#) ③  
6. Taxonomy: *Lactobacillus rhamnosus* GG  
Project data type: Transcriptome or Gene expression ②  
Scope: Multiisolate  
Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea I  
Accession: PRJNA419802 ID: 419802
- [Lactobacillus plantarum](#)  
7. Global transcriptional responses to oxidative stress conditions in *lactobacillus* plantarum  
Taxonomy: *Lactobacillus plantarum*  
Project data type: Transcriptome or Gene expression  
Scope: Multiisolate  
College of Food Science and Nutritional Engineering, China Agricultural University  
Accession: PRJNA387288 ID: 387288
- [Candida albicans](#)  
8. Antifungal defense of probiotic *Lactobacillus rhamnosus* GG is mediated by blocking adhesion and nutrient depletion  
Taxonomy: *Candida albicans*  
Project data type: Transcriptome or Gene expression  
Scope: Multiisolate  
MBT, Fraunhofer IGB

# W15: *L. rhamnosus* GG

こんな感じのページに飛びます。先ほどの例では①のあたりに原著論文情報があったが、ここでは見当たらない。

RNA-seq analysis of Lactobacillus at acidic stress    Accession: PRJNA419802    ID: 419802

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 <b>Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University</b>
Relevance	Industrial

**Project Data:**

**Related information**

- BioSample
- Genome
- GEO DataSets
- SRA
- Taxonomy

**Recent activity**

- RNA-seq analysis of Lactobacillus at acidic s: BioProject
- SRA Links for BioProject (Select 475032) (36) SRA
- How Lactobacillus plantarum Shapes its Transcriptom: BioProject
- (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject

See more...



NAVIGATE ACROSS  
74 additional projects are related by organism.

See Genome Information for Lactobacillus rhamnosus

# W15: *L. rhamnosus* GG

①このGSE107337というIDを頼りにして、のちに原著論文の確認を行いました。

RNA-seq analysis of Lactobacillus at acidic stress    Accession: PRJNA419802    ID: 419802

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 <b>Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University</b>
Relevance	Industrial

**Project Data:**

**Related information**

- BioSample
- Genome
- GEO DataSets
- SRA
- Taxonomy

**Recent activity**

- RNA-seq analysis of Lactobacillus at acidic s: BioProject
- SRA Links for BioProject (Select 475032) (36) SRA
- How Lactobacillus plantarum Shapes its Transcriptom: BioProject
- (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO: BioProject

See more...

# W15: *L. rhamnosus* GG

①この部分を見て、実験デザインがシンプルで、発現変動遺伝子検出まで行うことができるデータセットであることを知り、詳細に調べていく決断をしました。

RNA-seq analysis of Lactobacillus at acidic stress    Accession: PRJNA419802    ID: 419802

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions  
Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University
Relevance	Industrial

**Project Data:**

**Related information**

- BioSample
- Genome
- GEO DataSets
- SRA
- Taxonomy

**Recent activity**

- RNA-seq analysis of Lactobacillus at acidic s: BioProject
- SRA Links for BioProject (Select 475032) (36) SRA
- How Lactobacillus plantarum Shapes its Transcriptom: BioProject
- (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO: BioProject

See more...

See Genome Information for Lactobacillus rhamnosus

①

NAVIGATE ACROSS

74 additional projects are related by organism.

# W16: GSE107337

RNA-seq analysis of Lactobacillus at acidic stress    Accession: PRJNA419802    ID: 419802

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 <b>Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University</b>
Relevance	Industrial

**Project Data:**

**Related information**

- BioSample
- Genome
- GEO DataSets
- SRA
- Taxonomy

**Recent activity**

- RNA-seq analysis of Lactobacillus at acidic s: BioProject
- SRA Links for BioProject (Select 475032) (36) SRA
- How Lactobacillus plantarum Shapes its Transcriptom: BioProject
- (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject

①

# W16: GSE107337

①このあたりまで移動。②実験数も③サンプル数も9。とりあえず②のほうを見る

RNA-seq analysis of Lactobacillus × +

https://www.ncbi.nlm.nih.gov/bioproject/419802

Organism	Lactobacillus rhamnosus GG [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University
Relevance	Industrial

**Project Data:**

Resource Name	Number of Links
SEQUENCE DATA	
SRA Experiments	9
OTHER DATASETS	
BioSample	9
GEO DataSets	1

▼ GEO Data Details

Parameter	Value
Data volume, Supplementary Mbytes	1

▼ SRA Data Details

Parameter	Value
Data volume, Gbases	9
Data volume, Mbytes	5634

RNA-seq analysis of Lactobacillus at acidic s' BioProject

SRA Links for BioProject (Select 475032) (36) SRA

How Lactobacillus plantarum Shapes its Transcriptom BioProject

(Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject

See more...

# W16: GSE107337

赤枠内の、①と②の情報から、トータル9サンプルからなるこのデータセットは、全部で3状態の比較(3群間比較)を行っており、各群につき3反復のデータを取得しているのだろうと判断できる。また、③Illumina MiSeqが用いられていることもわかる。

SRA Links for BioProject (Select 4) × +

https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject\_sra\_all&f

NCBI Resources How To Sign in to NCBI

SRA SRA Search Help

Advanced

Access Public (9)

Source RNA (9)

Clear all Show additional filters

Summary 20 per page Send to: Filters: Manage Filters

Send results to Blast

Find related data Database: Select Find items

Recent activity Turn Off Clear

SRA Links for BioProject (Select 419802) (9) SRA

BioSample for BioProject (Select 419802) (9) BioSample

RNA-seq analysis of Lactobacillus at acidic s' BioProject

SRA Links for BioProject (Select 475032) (36) SRA

How Lactobacillus plantarum Shapes its Transcriptom BioProject

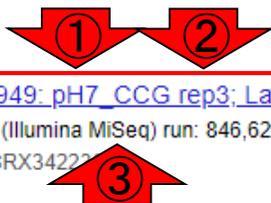
See more...

Links from BioProject

Items: 9

- [GSM2864949: pH7\\_CCG rep3; Lactobacillus rhamnosus GG; RNA-Seq](#)  
1. 1 ILLUMINA (Illumina MiSeq) run: 846,623 spots, 425M bases, 267.3Mb downloads  
Accession: SRX3422367
- [GSM2864948: pH7\\_CCG rep2; Lactobacillus rhamnosus GG; RNA-Seq](#)  
2. 1 ILLUMINA (Illumina MiSeq) run: 2.6M spots, 1.3G bases, 810.6Mb downloads  
Accession: SRX3422368
- [GSM2864947: pH7\\_CCG rep1; Lactobacillus rhamnosus GG; RNA-Seq](#)  
3. 1 ILLUMINA (Illumina MiSeq) run: 3.1M spots, 1.6G bases, 967.5Mb downloads  
Accession: SRX3422367
- [GSM2864946: pH4\\_24h rep3; Lactobacillus rhamnosus GG; RNA-Seq](#)  
4. 1 ILLUMINA (Illumina MiSeq) run: 1.8M spots, 901.5M bases, 564Mb downloads  
Accession: SRX3422366

https://www.ncbi.nlm.nih.gov/portal/utils/pageresolver.fcgi?recordid=5c53dad839a66e5f7aefdeb3



# W16: GSE107337

- [GSM2864949: pH7\\_CCG rep3; Lactobacillus rhamnosus GG; RNA-Seq](#)  
1. 1 ILLUMINA (Illumina MiSeq) run: 846,623 spots, 425M bases, 267.3Mb downloads  
Accession: SRX3422369
- [GSM2864948: pH7\\_CCG rep2; Lactobacillus rhamnosus GG; RNA-Seq](#)  
2. 1 ILLUMINA (Illumina MiSeq) run: 2.6M spots, 1.3G bases, 810.6Mb downloads  
Accession: SRX3422368
- [GSM2864947: pH7\\_CCG rep1; Lactobacillus rhamnosus GG; RNA-Seq](#)  
3. 1 ILLUMINA (Illumina MiSeq) run: 3.1M spots, 1.6G bases, 967.5Mb downloads  
Accession: SRX3422367
- [GSM2864946: pH4\\_24h rep3; Lactobacillus rhamnosus GG; RNA-Seq](#)  
4. 1 ILLUMINA (Illumina MiSeq) run: 1.8M spots, 901.5M bases, 564Mb downloads  
Accession: SRX3422366
- [GSM2864945: pH4\\_24h rep2; Lactobacillus rhamnosus GG; RNA-Seq](#)  
5. 1 ILLUMINA (Illumina MiSeq) run: 3.9M spots, 1.9G bases, 1.2Gb downloads  
Accession: SRX3422365
- [GSM2864944: pH4\\_24h rep1; Lactobacillus rhamnosus GG; RNA-Seq](#)  
6. 1 ILLUMINA (Illumina MiSeq) run: 1.4M spots, 690.4M bases, 428.8Mb downloads  
Accession: SRX3422364
- [GSM2864943: pH4\\_1h rep3; Lactobacillus rhamnosus GG; RNA-Seq](#)  
7. 1 ILLUMINA (Illumina MiSeq) run: 1.8M spots, 883.8M bases, 562.8Mb downloads  
Accession: SRX3422363
- [GSM2864942: pH4\\_1h rep2; Lactobacillus rhamnosus GG; RNA-Seq](#)  
8. 1 ILLUMINA (Illumina MiSeq) run: 1.5M spots, 738.2M bases, 467.2Mb downloads  
Accession: SRX3422362
- [GSM2864941: pH4\\_1h rep1; Lactobacillus rhamnosus GG; RNA-Seq](#)  
9. 1 ILLUMINA (Illumina MiSeq) run: 301,126 spots, 151.2M bases, 94.2Mb downloads  
Accession: SRX3422361



全9サンプル分を表示。ここでも大まかに全体像を理解することができる。これは①通常のpH環境と、酸性ストレス環境下(pH = 4)の比較を行っている。②酸性ストレス条件で1h経過後と、③24h経過後のデータを取得していることがわかる

# W17: 原著論文を探

①GSE107337の原著論文はここでは見られないが、②Registrationから1年以上経過しているため、もしかしたらPubMedで見つかるかもという視点で探す。

RNA-seq analysis of Lactobacillus at acidic stress

Accession: PRJNA419802 ID: 419802

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University
Relevance	Industrial

**Project Data:**

Send to:   
Related information: BioSample, Genome, GEO DataSets, SRA, Taxonomy  
Recent activity: RNA-seq analysis of Lactobacillus at acidic s' BioProject, SRA Links for BioProject (Select 475032) (36) SRA, How Lactobacillus plantarum Shapes its Transcriptom BioProject, (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject

# W17: 原著論文を探す

①まずはこれで検索し、②2件ヒットしたが、違う論文でした。

RNA-seq analysis of Lactobacillus

https://www.ncbi.nlm.nih.gov/bioproject/419802

NCBI Resources How To Sign in to NCBI

BioProject BioProject Search

RNA-seq analysis of Lactobacillus at acidic stress

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions. Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions to identify differentially expressed genes

Accession	PRJNA419802; GEO: GSE107337
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus
Submission	Registration date: 25-Nov-2017 <b>Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University</b>
Relevance	Industrial

Project Data:

PubMed RNA-seq analysis of Lactobacillus at acidic stress

Format: Summary Sort by: Most Recent Send to Filters: Manage Filters

Search results Items: 2

- [Global Transcriptomic Analysis and Function Identification of Malolactic Enzyme Pathway of \*Lactobacillus paracasei\* L9 in Response to Bile Stress.](#)  
Ma X, Wang G, Zhai Z, Zhou P, Hao Y. Front Microbiol. 2018 Aug 23;9:1978. doi: 10.3389/fmicb.2018.01978. eCollection 2018. PMID: 30210466 Free PMC Article Similar articles
- [Transcriptome responses of \*Lactobacillus acetotolerans\* F28 to a short and long term ethanol stress.](#)  
Yang X, Teng K, Zhang J, Wang F, Zhang T, Ai G, Han P, Bai F, Zhong J. Sci Rep. 2017 Jun 1;7(1):2650. doi: 10.1038/s41598-017-02975-8. PMID: 28572611 Free PMC Article Similar articles

Search details: RNA-seq[All Fields] AND ("analysis"[Subheading] OR "analysis"[All Fields]) AND ("lactobacillus"[MeSH Terms] OR "lactobacillus"[All...])

Recent Activity: RNA-seq analysis of Lactobacillus at acidic str PubMed

# W17: 原著論文を探す

RNA-seq analysis of Lactobacillus x +

https://www.ncbi.nlm.nih.gov/bioproject/419802

NCBI Resources How To Sign in to NCBI

BioProject BioProject Search

Display Settings: v

### RNA-seq analysis of Lactobacillus at acidic stress

To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions. Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions to identify differentially expressed genes

Accession	PRJNA419802; GEO: <u>GSE107337</u>
Data Type	Transcriptome or Gene expression
Scope	Multiisolate
Organism	<b>Lactobacillus rhamnosus GG</b> [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus
Submission	Registration date: 25-Nov-2017 <b>Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University</b>
Relevance	Industrial

Project Data:

No items found - PubMed - NCBI x +

https://www.ncbi.nlm.nih.gov/pubmed/?term=GSE107337

NCBI Resources How To Sign in to NCBI

PubMed GSE107337 Search

Article types  
Clinical Trial  
Review  
Customize ...

Text availability  
Abstract  
Free full text  
Full text

Publication dates  
5 years  
10 years  
Custom range...

Species  
Humans  
Other Animals

Clear all  
Show additional filters

[RNA-seq analysis of Lactobacillus at acidic stress](#)  
To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions  
Species: Lactobacillus rhamnosus GG Type: Expression profiling by high throughput sequencing  
Dataset: GSE107337

**Search results**  
Items: 0  
No documents match your search terms

Search details  
(GSE107337[All Fields])  
Search See more...

Recent Activity  
Turn Off Clear

- GSE107337 (0) PubMed
- GEO DataSets for BioProject (Select 419802) (1) GEO DataSets
- RNA-seq analysis of Lactobacillus at acid GEO DataSets
- SRA Links for BioProject (Select 419802) (9) SRA
- Transcriptional Response and Enhanced Intestinal Adherence PubMed

# W17: 原著論文を探す

論文著者情報の手がかりを得るため、①ページ下部に移動して、②GEO DataSetsのリンク先にいく。このやり方でなくても、GSE107337で単純にググるのもよい。

RNA-seq analysis of Lactobacillus × +  
https://www.ncbi.nlm.nih.gov/bioproject/419802

Organism	Lactobacillus rhamnosus GG [Taxonomy ID: 568703] Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus; Lactobacillus rhamnosus; Lactobacillus rhamnosus GG
Submission	Registration date: 25-Nov-2017 Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University
Relevance	Industrial

## Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
SRA Experiments	9
OTHER DATASETS	
BioSample	9
GEO DataSets	1

▼ GEO Data Details

Parameter	Value
Data volume, Supplementary Mbytes	1

▼ SRA Data Details

Parameter	Value
Data volume, Gbases	9
Data volume, Mbytes	5634

- RNA-seq analysis of Lactobacillus at acidic s' BioProject
  - SRA Links for BioProject (Select 475032) (36) SRA
  - How Lactobacillus plantarum Shapes its Transcriptom BioProject
  - (Lactobacillus RNA-seq) AND bioproject\_sra[filter] NO BioProject
- See more...



# W17: 原著論文を探す

RNA-seq analysis of Lactobacillus × +

https://www.ncbi.nlm.nih.gov/gds?Db=gds&DbFrom=bioproject&Cmd=Link&LinkName...

NCBI Resources How To Sign in to NCBI

GEO DataSets GEO DataSets Search Help

Summary Send to

### Links from BioProject

[RNA-seq analysis of Lactobacillus at acidic stress](#) ①

(Submitter supplied) To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions

Organism: Lactobacillus rhamnosus GG

Type: Expression profiling by high throughput sequencing

Platform: GPL24302 9 Samples

Download data: CSV

Series Accession: GSE107337 ID: 200107337

[SRA Run Selector](#)

### Related information

BioProject

SRA

Taxonomy

### Recent activity

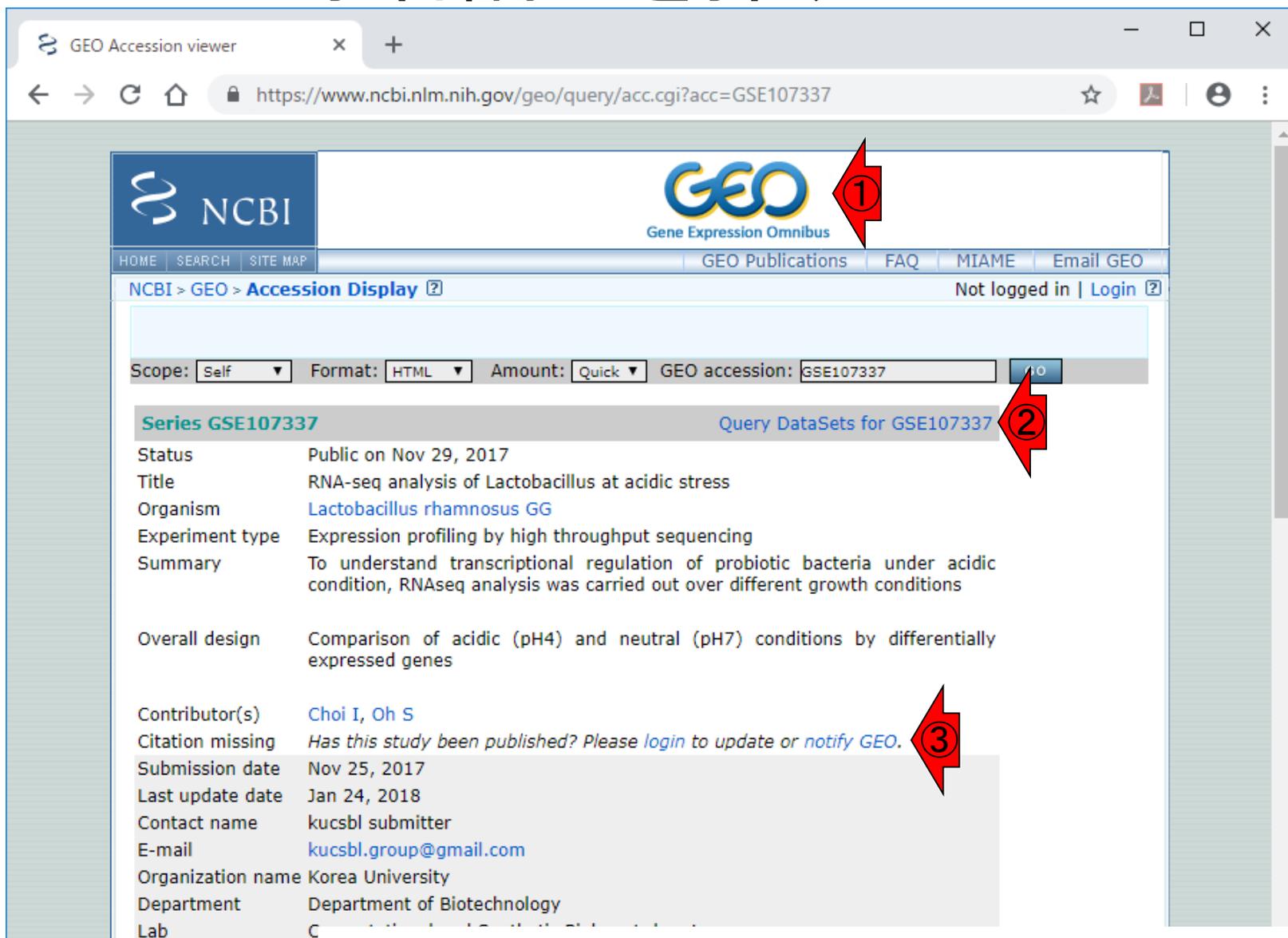
Turn Off Clear

- GEO DataSets for BioProject (Select 419802) (1) GEO DataSets
- RNA-seq analysis of Lactobacillus at acid GEO DataSets
- GSE107337 (0) PubMed
- SRA Links for BioProject (Select 419802) (9) SRA
- Transcriptional Response and Enhanced Intestinal Adh PubMed

See more...

# W17: 原著論文を探す

①GEOの、②GSE107337のページ。ここでもまだ③原著論文情報はない。



NCBI GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

NCBI > GEO > **Accession Display** Not logged in | Login

Scope: Self Format: HTML Amount: Quick GEO accession: GSE107337 **Go**

**Series GSE107337** [Query DataSets for GSE107337](#)

Status	Public on Nov 29, 2017
Title	RNA-seq analysis of Lactobacillus at acidic stress
Organism	<a href="#">Lactobacillus rhamnosus GG</a>
Experiment type	Expression profiling by high throughput sequencing
Summary	To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions
Overall design	Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes
Contributor(s)	<a href="#">Choi I, Oh S</a>
Citation missing	<i>Has this study been published? Please <a href="#">login</a> to update or <a href="#">notify GEO</a>.</i>
Submission date	Nov 25, 2017
Last update date	Jan 24, 2018
Contact name	kucsbl submitter
E-mail	<a href="mailto:kucsbl.group@gmail.com">kucsbl.group@gmail.com</a>
Organization name	Korea University
Department	Department of Biotechnology
Lab	C

①と②の情報を用いて  
PubMed検索することで...

# W17: 原著論文を探す

NCBI GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

NCBI > GEO > Accession Display

Scope: Self Format: HTML Amount: Quick GEO accession: GSE107337 GO

**Series GSE107337** Query DataSets for GSE107337

Status	Public on Nov 29, 2017
Title	RNA-seq analysis of Lactobacillus at acidic stress
Organism	<a href="#">Lactobacillus rhamnosus GG</a> ①
Experiment type	Expression profiling by high throughput sequencing
Summary	To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions
Overall design	Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes
Contributor(s)	<a href="#">Choi I, Oh S</a> ②
Citation missing	Has this study been published? Please <a href="#">login</a> to update or <a href="#">notify GEO</a> .
Submission date	Nov 25, 2017
Last update date	Jan 24, 2018
Contact name	kucsbl submitter
E-mail	<a href="mailto:kucsbl.group@gmail.com">kucsbl.group@gmail.com</a>
Organization name	Korea University
Department	Department of Biotechnology
Lab	Computational and Synthetic Biology Laboratory

# W17: 原著論文を探す

The screenshot shows a web browser window displaying a PubMed search result. The browser's address bar shows the URL: <https://www.ncbi.nlm.nih.gov/pubmed/?term=Lactobacillus+rhamnosus+GG+Choi+I,+Oh+S>. The page header includes the NCBI logo, navigation links for Resources and How To, and a 'Sign in to NCBI' button. The main search area shows 'PubMed' as the selected database and the search term 'Lactobacillus rhamnosus GG Choi I, Oh S'. Below the search bar, there are links for 'Create RSS', 'Create alert', 'Advanced', and 'Help'. The article details include the journal 'J Microbiol Biotechnol', the date '2018 Oct 28;28(10):1604-1613', and the DOI '10.4014/jmb.1807.07033'. The title is 'Transcriptional Response and Enhanced Intestinal Adhesion Ability of *Lactobacillus rhamnosus* GG after Acid Stress.' The authors listed are 'Bang M<sup>1</sup>, Yong CC<sup>1</sup>, Ko HJ<sup>2,3</sup>, Choi IG<sup>2</sup>, Oh S<sup>1</sup>.' There is a section for 'Author information'. The 'Abstract' section begins with 'Lactobacillus rhamnosus GG (LGG) is a probiotic commonly used in fermented dairy products. In this study, RNA-sequencing was performed to unravel the effects of acid stress on LGG. The transcriptomic data revealed that the exposure of LGG to acid at pH 4.5 (resembling the final pH of fermented dairy products) for 1 h or 24 h provoked a stringent-type transcriptomic response wherein stress response- and glycolysis-related genes were upregulated, whereas genes involved in gluconeogenesis, amino acid metabolism, and nucleotide metabolism were suppressed. Notably, the pilus-specific adhesion genes, *spaC*, and *spaF* were significantly upregulated upon exposure to acid-stress. The transcriptomic results were further confirmed via quantitative polymerase chain reaction analysis. Moreover, acid-stressed LGG demonstrated an enhanced mucin-binding ability in vitro, with 1 log more LGG cells ( $p < 0.05$ ) bound to a mucin layer in a 96-well culture plate as compared to the control. The enhanced intestinal binding ability of acid-stressed LGG was confirmed in an animal study, wherein significantly more viable LGG cells ( $\geq 2$  log CFU/g) were observed in the ileum, caecum, and colon of acid-stressed LGG-treated mice as compared with a non-acid-stressed LGG-treated control group. To our knowledge, this is the first report showing that acid stress enhanced the intestine-binding ability of LGG through the induction of pili-related genes.' On the right side of the page, there are sections for 'Full text links' (with a 'FREE Full-Text' button), 'Save items' (with an 'Add to Favorites' button), and 'Similar articles' (listing related papers like 'Lactobacillus rhamnosus GG SpaC pilin subunit binds [Anim Sci J. 2016]', 'Shear-Enhanced Dynamic Adhesion of Lactobacillus rha [Langmuir. 2018]', 'Adhesion and nanomechanics of pili from the probiotic [ACS Nano. 2013]', and 'Adhesive interactions between milk fat  $\zeta$  [Colloids Surf B Biointerfaces...]').

# W17: 原著論文を探す

GSE107337中の記述内容と同じ。  
①Full-Text中でも、GSE107337と明記されていることを確認済み。

The screenshot shows a web browser window displaying a PubMed search result. The search query is "Lactobacillus rhamnosus GG Choi I, Oh S". The search results show a paper titled "Transcriptional Response and Enhanced Intestinal Adhesion Ability of *Lactobacillus rhamnosus* GG after Acid Stress." by Bang M<sup>1</sup>, Yong CC<sup>1</sup>, Ko HJ<sup>2,3</sup>, Choi IG<sup>2</sup>, Oh S<sup>1</sup>. The abstract is visible, starting with "Lactobacillus rhamnosus GG (LGG) is a probiotic commonly used in fermented dairy products. In this study, RNA-sequencing was performed to unravel the effects of acid stress on LGG. The transcriptomic data revealed that the exposure of LGG to acid at pH 4.5 (resembling the final pH of fermented dairy products) for 1 h or 24 h provoked a stringent-type transcriptomic response wherein stress response- and glycolysis-related genes were upregulated, whereas genes involved in gluconeogenesis, amino acid metabolism, and nucleotide metabolism were suppressed. Notably, the pilus-specific adhesion genes, *spaC*, and *spaF* were significantly upregulated upon exposure to acid-stress. The transcriptomic results were further confirmed via quantitative polymerase chain reaction analysis. Moreover, acid-stressed LGG demonstrated an enhanced mucin-binding ability in vitro, with 1 log more LGG cells ( $p < 0.05$ ) bound to a mucin layer in a 96-well culture plate as compared to the control. The enhanced intestinal binding ability of acid-stressed LGG was confirmed in an animal study, wherein significantly more viable LGG cells ( $\geq 2$  log CFU/g) were observed in the ileum, caecum, and colon of acid-stressed LGG-treated mice as compared with a non-acid-stressed LGG-treated control group. To our knowledge, this is the first report showing that acid stress enhanced the intestine-binding ability of LGG through the induction of pili-related genes." On the right side of the page, there are sections for "Full text links" (with a "FREE Full-Text" button and a red arrow pointing to it), "Save items" (with an "Add to Favorites" button), and "Similar articles" (listing related papers like "Lactobacillus rhamnosus GG SpaC pilin subunit binds [Anim Sci J. 2016]").

# W18: ENA

The screenshot shows the ENA website interface. A yellow box at the top right contains the text '①ENA上で、②GSE107337を、③検索'. Three red arrows with circled numbers point to specific elements: arrow ① points to the ENA logo, arrow ② points to the search input field containing 'GSE107337', and arrow ③ points to the 'Search' button. The browser address bar shows 'https://www.ebi.ac.uk/ena'. The website header includes 'EMBL-EBI' and navigation links for 'Services', 'Research', 'Training', and 'About us'. A main navigation bar contains 'Home', 'Search & Browse', 'Submit & Update', 'Software', 'About ENA', and 'Support'. The main content area features a 'European Nucleotide Archive' title, a descriptive paragraph, and a 'Text Search' section with an input field and a 'search' button. A 'Popular' section lists various actions like 'Submit and update' and 'Sequence submissions'. A 'Latest ENA news' section shows a recent update from 06 Nov 2018.

# W18: ENA

GSE107337で検索しているはずですが、検索結果は① PRJNS419802や②SRP125628となっていることがわかります。結論としては問題ない。③ページ下部に移動

The screenshot shows the ENA (European Nucleotide Archive) website. At the top, there is a search bar with a "Search" button and links for "Advanced" and "Sequence". Below the search bar is a navigation menu with "Home", "Search & Browse", "Submit & Update", "Software", "About ENA", and "Support". The main content area displays the study details for PRJNS419802, including the title "RNA-seq analysis of Lactobacillus at acidic stress", the submitting center "Computational and Synthetic Biology Laboratory, Department of Biotechnology, Korea University", and the organism "Lactobacillus rhamnosus GG". A secondary accession number SRP125628 is also listed. The description states: "To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes".

# W18: ENA

①このあたりまで移動すると、②全9サンプルに付随する様々なID情報を一覧できる。原著論文には明記されていないが、③このデータがpaired-endであることもわかる。

Navigation Read Files Portal Attributes Parent Projects

Bulk Download Files (If the downloader app doesn't open, please try using Firefox to launch it.)

Download: 1 - 9 of 9 results in TEXT

Select columns

Showing results 1 - 9 of 9 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098216</a>	<a href="#">SRS2714081</a>	<a href="#">SRX3422361</a>	<a href="#">SRR6322562</a>	568703	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098215</a>	<a href="#">SRS2714083</a>	<a href="#">SRX3422362</a>	<a href="#">SRR6322563</a>	568703	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098214</a>	<a href="#">SRS2714082</a>	<a href="#">SRX3422363</a>	<a href="#">SRR6322564</a>	568703	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098213</a>	<a href="#">SRS2714084</a>	<a href="#">SRX3422364</a>	<a href="#">SRR6322565</a>	568703	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>

# W19: ダウンロード

①赤枠内のリンク先からもFASTQファイルをダウンロードできるが、②Bulk Download Filesを利用することで一括ダウンロードもできる。ここでは②をクリック。

Navigation Read Files Portal Attributes Parent Projects

Bulk Download Files (If the downloader app doesn't open, please try using Firefox to launch it.)

Download:  -  of 9 results in TEXT

Select columns

Showing results 1 - 9 of 9 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098216</a>	<a href="#">SRS2714081</a>	<a href="#">SRX3422361</a>	<a href="#">SRR6322562</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098215</a>	<a href="#">SRS2714083</a>	<a href="#">SRX3422362</a>	<a href="#">SRR6322563</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098214</a>	<a href="#">SRS2714082</a>	<a href="#">SRX3422363</a>	<a href="#">SRR6322564</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098213</a>	<a href="#">SRS2714084</a>	<a href="#">SRX3422364</a>	<a href="#">SRR6322565</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>

# W19:ダウンロード

The screenshot shows a web browser window at <https://www.ebi.ac.uk/ena/data/view/PRJNA419802>. A dialog box asks "Java(TM) Web Launcher を開きますか？" (Do you want to open Java(TM) Web Launcher?). Below the question is a checkbox "このタイプのリンクは常に関連付けられたアプリで開く" (Always open this type of link with the associated application) which is unchecked. There are two buttons: "Java(TM) Web Launcher を開く" (Open Java(TM) Web Launcher) and "キャンセル" (Cancel). A red arrow with the number "1" points to the "開く" button. Below the dialog, a table shows search results. An application verification window is overlaid on the table, displaying "アプリケーションの起動中..." (Starting application...), "アプリケーションを検証しています。" (Verifying application.), and the URL "場所: https://www.ebi.ac.uk". A progress bar and a "取消" (Cancel) button are also visible in the verification window.

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	...
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098216</a>	<a href="#">SRS2714081</a>	<a href="#">SRX3422361</a>	<a href="#">SRR6322562</a>	568703
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098215</a>	<a href="#">SRS2714083</a>	<a href="#">SRX3422362</a>	<a href="#">SRR6322563</a>	568703
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098214</a>	<a href="#">SRS2714082</a>	<a href="#">SRX3422363</a>	<a href="#">SRR6322564</a>	568703
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098213</a>	<a href="#">SRS2714084</a>	<a href="#">SRX3422364</a>	<a href="#">SRR6322565</a>	568703

# W19:ダウンロード

Java(TM) Web Launcher を開きますか？  
 このタイプのリンクは常に関連付けられたアプリで開く

Java(TM) Web Launcher を開く    キャンセル

Requires Java version 8

このアプリケーションを実行しますか。

 **名前:** ENA File Downloader  
**発行者:** European Molecular Biology Laboratory  
**場所:** https://www.ebi.ac.uk

このアプリケーションは、コンピュータおよび個人情報を危険にさらす可能性がある無制限のアクセスで実行されます。上記の場所と発行者を信頼する場合にのみ、このアプリケーションを実行してください。

上記の発行者と場所からのアプリケーションについては、次回から表示しない(D)

 詳細情報(M)    **実行(R)**    取消

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	T
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098216</a>	<a href="#">SRS2714081</a>	<a href="#">SRX3422361</a>	<a href="#">SRR6322562</a>	5
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098215</a>	<a href="#">SRS2714083</a>	<a href="#">SRX3422362</a>	<a href="#">SRR6322563</a>	5
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098214</a>	<a href="#">SRS2714082</a>	<a href="#">SRX3422363</a>	<a href="#">SRR6322564</a>	5
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098213</a>	<a href="#">SRS2714084</a>	<a href="#">SRX3422364</a>	<a href="#">SRR6322565</a>	568703
					GG
					Lactobacillus rhamnosus GG
					Illumina MiSeq
					PAIRED
					File 1 File 2
					File 1 File 2

# W19: ダウンロード

The screenshot shows the ENA FTP Downloader application interface. At the top, there is a 'Local Download Folder' field with a 'Browse' button. Below it is a checkbox for 'Create Subfolder for each accession'. The main area is titled 'Remote Files' and contains a table with columns: Download, Accession, Name, Size, Progress, and MD5 OK. The table lists 14 files with their respective accession numbers and sizes. At the bottom left, a 'Select All' button is highlighted with a red arrow and the number 1. Other buttons include 'Back', 'Start Download', and 'Stop Download'.

Download	Accession	Name	Size	Progress	MD5 OK
<input type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB		
<input type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB		
<input type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB		
<input type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB		
<input type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB		
<input type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB		
<input type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB		
<input type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB		
<input type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB		
<input type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB		
<input type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB		
<input type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB		

# W19: ダウンロード

①全部ダウンロードしても約6GBです。  
次は、②ダウンロードする場所を指定。

ENA FTP Downloader

Local Download Folder:  Browse **②**

Create Subfolder for each accession:

Remote Files

FASTQ X Submitted SRA

Download	Accession	Name	Size	Progress	MD5 OK
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB	<input type="text"/>	
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB	<input type="text"/>	

Deselect All **①** 18 FASTQ files selected. Total size: 5.98 GB

Back Start Download Stop Download

# W19: ダウンロード

ここでは、デスクトップ上に予め作成しておいた①GSE107337を、②選択しました。

The screenshot shows the ENA FTP Downloader interface. A file selection dialog box is open, showing the path 'デスクトップ > GSE107337'. The 'デスクトップ' folder is selected in the left pane. The dialog box also shows a search bar with 'GSE107337の検索' and a 'フォルダーの選択' button. The main window shows a list of files with columns for 'Download', 'Accession', and file size. The status bar indicates '18 FASTQ files selected. Total size: 5.98 GB'.

Download	Accession	File Name	Size
<input checked="" type="checkbox"/>	SRR6322562		
<input checked="" type="checkbox"/>	SRR6322562		
<input checked="" type="checkbox"/>	SRR6322563		
<input checked="" type="checkbox"/>	SRR6322563		
<input checked="" type="checkbox"/>	SRR6322564		
<input checked="" type="checkbox"/>	SRR6322564		
<input checked="" type="checkbox"/>	SRR6322565		
<input checked="" type="checkbox"/>	SRR6322565		
<input checked="" type="checkbox"/>	SRR6322566		
<input checked="" type="checkbox"/>	SRR6322566		
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB

# W19:ダウンロード

ENA FTP Downloader

Local Download Folder: C:\Users\kadota\Desktop\GSE107337 Browse

Create Subfolder for each accession:

Remote Files

FASTQ X Submitted SRA

Download	Accession	Name	Size	Progress	MD5 OK
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB		
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB		
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB		
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB		
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB		
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB		
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB		
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB		
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB		
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB		
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB		
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB		

Deselect All 18 FASTQ files selected. Total size: 5.98 GB

Back Start Download Stop Download

## W19: ダウンロード

Local Download Folder: C:\Users\kadota\Desktop\GSE107337 Browse

Create Subfolder for each accession:

Remote Files

FASTQ X Submitted SRA

Download	Accession	Name	Size	Progress	MD5 OK
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB	<div style="width: 100%;"></div>	
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB	<div style="width: 100%;"></div>	

Deselect All 18 FASTQ files selected. Total size: 5.98 GB

Back Start Download Stop Download

# W20: 成功例

ENA FTP Downloader

Local Download Folder:

Create Subfolder for each a...

Remote Files

FASTQ × Submitted SRA

Downlo...	Accession	Name	Size	Progress	MD5 ...
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB	<div style="width: 100%;"></div>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB	<div style="width: 10%;"></div>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB	<div style="width: 0%;"></div>	
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB	<div style="width: 0%;"></div>	

SRR6322562\_1.fastq.gz downloaded.

無事ダウンロードが完了すると、①指定したフォルダ内に、gzip圧縮FASTQファイル群が計18個分できます。

# W20: 成功例

ENA FTP Downloader

Local Download Folder: C:\Users\kadota\Desktop\GSE107337 **①** Browse

Create Subfolder for each a...

Remote Files

FASTQ × Submitted SRA

Downlo...	Accession	Name	Size	Progress	MD5 ...
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322568	SRR6322568_1.fastq.gz	525.84 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322568	SRR6322568_2.fastq.gz	587.34 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322569	SRR6322569_1.fastq.gz	430.22 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322569	SRR6322569_2.fastq.gz	486.67 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322570	SRR6322570_1.fastq.gz	141.62 MB	<div style="width: 100%;"></div>	✓
<input checked="" type="checkbox"/>	SRR6322570	SRR6322570_2.fastq.gz	159.59 MB	<div style="width: 100%;"></div>	✓

Deselect All 18 files have been successfully downloaded.

Back Start Download Stop Download

# W20: 成功例

こんな感じです。自宅の光無線LAN環境で4時間弱かかっていることが分かります。

ENA FTP Downloader

Local Download Folder: C:\Users\kadota\Desktop

Create Subfolder for each a...

Remote Files

FASTQ ×	Submitted	SRA
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Downloaded files table:

Downlo...	Accession	Name
<input type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz
<input checked="" type="checkbox"/>	SRR6322568	SRR6322568_1.fastq.gz
<input checked="" type="checkbox"/>	SRR6322568	SRR6322568_2.fastq.gz
<input checked="" type="checkbox"/>	SRR6322569	SRR6322569_1.fastq.gz
<input checked="" type="checkbox"/>	SRR6322569	SRR6322569_2.fastq.gz
<input checked="" type="checkbox"/>	SRR6322570	SRR6322570_1.fastq.gz
<input checked="" type="checkbox"/>	SRR6322570	SRR6322570_2.fastq.gz

Deselect All 18 files have been successfully downloaded

Back

GSE107337

名前	更新日時	サイズ	種類
SRR6322562_1.fastq.gz	2019/02/02 9:51	52,570 KB	WinRAR 書庫
SRR6322562_2.fastq.gz	2019/02/02 9:59	58,262 KB	WinRAR 書庫
SRR6322563_1.fastq.gz	2019/02/02 10:03	259,607 KB	WinRAR 書庫
SRR6322563_2.fastq.gz	2019/02/02 10:51	289,949 KB	WinRAR 書庫
SRR6322564_1.fastq.gz	2019/02/02 10:55	312,309 KB	WinRAR 書庫
SRR6322564_2.fastq.gz	2019/02/02 11:36	346,247 KB	WinRAR 書庫
SRR6322565_1.fastq.gz	2019/02/02 11:39	235,891 KB	WinRAR 書庫
SRR6322565_2.fastq.gz	2019/02/02 12:09	262,052 KB	WinRAR 書庫
SRR6322566_1.fastq.gz	2019/02/02 12:30	666,507 KB	WinRAR 書庫
SRR6322566_2.fastq.gz	2019/02/02 12:43	741,892 KB	WinRAR 書庫
SRR6322567_1.fastq.gz	2019/02/02 12:50	308,215 KB	WinRAR 書庫
SRR6322567_2.fastq.gz	2019/02/02 12:54	344,822 KB	WinRAR 書庫
SRR6322568_1.fastq.gz	2019/02/02 13:09	538,457 KB	WinRAR 書庫
SRR6322568_2.fastq.gz	2019/02/02 13:16	601,436 KB	WinRAR 書庫
SRR6322569_1.fastq.gz	2019/02/02 13:23	440,541 KB	WinRAR 書庫
SRR6322569_2.fastq.gz	2019/02/02 13:38	498,347 KB	WinRAR 書庫
SRR6322570_1.fastq.gz	2019/02/02 13:40	145,018 KB	WinRAR 書庫
SRR6322570_2.fastq.gz	2019/02/02 13:41	163,425 KB	WinRAR 書庫

18 個の項目

# W21: 失敗例

①こんな感じで、②失敗することもあります、めげずに③を押してリトライ。この日は何度トライしても同じ失敗が繰り返されました。あまり失敗しまくるようだと、諦めて個別にダウンロードでもよいと思います。

Navigation

Local Download Folder: C:\Users\kadota\Desktop\GSE107337 Browse

Create Subfolder for each accession:

Remote Files

Download: FASTQ X Submitted SRA

Download	Accession	Name	Size	Progress	MD5 OK
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_1.fastq.gz	51.34 MB		X
<input checked="" type="checkbox"/>	SRR6322562	SRR6322562_2.fastq.gz	56.90 MB		X
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_1.fastq.gz	253.52 MB		
<input checked="" type="checkbox"/>	SRR6322563	SRR6322563_2.fastq.gz	283.15 MB		
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_1.fastq.gz	304.99 MB		
<input checked="" type="checkbox"/>	SRR6322564	SRR6322564_2.fastq.gz	338.13 MB		
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_1.fastq.gz	230.36 MB		
<input checked="" type="checkbox"/>	SRR6322565	SRR6322565_2.fastq.gz	255.91 MB		
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_1.fastq.gz	650.88 MB		
<input checked="" type="checkbox"/>	SRR6322566	SRR6322566_2.fastq.gz	724.50 MB		
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_1.fastq.gz	300.99 MB		
<input checked="" type="checkbox"/>	SRR6322567	SRR6322567_2.fastq.gz	336.74 MB		

Deselect All Back Downloading stopped due to an error. Start Download Stop Download

# W22: 失敗した場合の対応策1

https://www.ebi.ac.uk/ena/data/ × +

European Bioinformatics Institute [GB] | https://www.ebi.ac.uk/ena/data/view/PRJNA419802 ☆ 人 | ⋮

Navigation Read Files Portal Attributes Parent Projects

Bulk Download Files ⚠ (If the downloader app doesn't open, please try using Firefox to launch it.)

Download:  -  of 9 results in TEXT

[Select columns](#)

Showing results 1 - 9 of 9 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098216</a>	<a href="#">SRS2714081</a>	<a href="#">SRX3422361</a>	<a href="#">SRR6322562</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098215</a>	<a href="#">SRS2714083</a>	<a href="#">SRX3422362</a>	<a href="#">SRR6322563</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098214</a>	<a href="#">SRS2714082</a>	<a href="#">SRX3422363</a>	<a href="#">SRR6322564</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>
<a href="#">PRJNA419802</a>	<a href="#">SAMN08098213</a>	<a href="#">SRS2714084</a>	<a href="#">SRX3422364</a>	<a href="#">SRR6322565</a>	<a href="#">568703</a>	<a href="#">Lactobacillus rhamnosus GG</a>	Illumina MiSeq	PAIRED	<a href="#">File 1</a> <a href="#">File 2</a>	<a href="#">File 1</a> <a href="#">File 2</a>





# W22: 失敗した場合の対応策

①の部分でSRAMetadb.sqliteというファイルをダウンロードする際に結構時間がかかります。

- W22: SRADB: Zhu et al., BMC Bioinformatics, 2013

原著論文(Bang et al., J Microbiol Biotechnol, 2018)の記述からGSE107227を頼りにSRP125628にたどり着いています。したがって、ここで指定「イントロ | NGS | 配列取得 | FASTQ or SR

```
param <- "SRP125628"
```

```
#必要なパッケージをロード  
library(SRADb)
```

```
#前処理
```

```
#sqlfile <- "SRAMetadb.sqlite"  
sqlfile <- getSRADBFile()  
sra_con <- dbConnect(SQLite(), sqlfile)
```

```
#前処理(実験デザインの全体像を表示)
```

```
hoge <- sraConvert(param, sra_con=sra_con)  
hoge  
apply(hoge, 2, unique)  
getFASTQinfo(in_acc=hoge$run, sra_con=sra_con)
```

```
#本番(FASTQファイルのダウンロード)
```

```
getFASTQfile(hoge$run, sra_con=sra_con)
```

```
RGui (64-bit)  
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ  
R Console  
anyDuplicated, append, as.data.frame, basename, cbind, colMeans,  
colnames, colSums, dirname, do.call, duplicated, eval, evalq,  
Filter, Find, get, grep, grepl, intersect, is.unsorted, lapply,  
lengths, Map, mapply, match, mget, order, paste, pmax, pmax.int,  
pmin, pmin.int, Position, rank, rbind, Reduce, rowMeans,  
rownames, rowSums, sapply, setdiff, sort, table, tapply, union,  
unique, unsplit, which, which.max, which.min  
要求されたパッケージ Rcurl をロード中です  
要求されたパッケージ bitops をロード中です  
Setting options('download.file.method.GEOquery'='auto')  
Setting options('GEOquery.inmemory.gpl'=FALSE)  
>  
> #前処理  
> #sqlfile <- "SRAMetadb.sqlite" #最新でなくてもよく、手元に予めダウンロード  
> sqlfile <- getSRADBFile() #最新のSRAMetadb SQLiteファイルを探す  
URL 'https://s3.amazonaws.com/starbuck1/sradb/SRAMetadb.sqlite.gz' を試して  
Content type 'binary/octet-stream' length 2660684574 bytes (2537.4 MB)  
downloaded 2537.4 MB  
  
Unzipping...
```

# W23: DRA

①DRASearch画面上で、②Accession  
の場所でGSE107337で、③Search

The screenshot shows the DRASearch web interface. A red arrow labeled '1' points to the browser address bar. A second red arrow labeled '2' points to the 'Accession' input field containing 'GSE107337'. A third red arrow labeled '3' points to the 'Search' button. Below the search bar, there are dropdown menus for 'Show' (set to 20 records) and 'Sort by' (set to Study), along with 'Search' and 'Clear' buttons. The page title is 'DRASearch' and it includes links for 'Search Home' and 'DRA Home'. The data last update is noted as 2019-02-01.

### Statistics

Released Entries

Type	Count
<a href="#">Submission</a>	1112526
<a href="#">Study</a>	178800
<a href="#">Experiment</a>	5504151
<a href="#">Sample</a>	4947100
<a href="#">Run</a>	6227253

Organism		
#	Organism Name	Study
1	<a href="#">Homo sapiens</a>	16356
2	<a href="#">Mus musculus</a>	13445
3	<a href="#">soil metagenome</a>	5530
4	<a href="#">Zea mays</a>	3281
5	<a href="#">Populus trichocarpa</a>	3064

Study Type		
#	Study Type	Study
1	<a href="#">Other</a>	69729
2	<a href="#">Whole Genome Sequencing</a>	56641
3	<a href="#">Metagenomics</a>	26156
4	<a href="#">Transcriptome Analysis</a>	24085
5	<a href="#">Population Genomics</a>	802

Center Name		
#	Center Name	Study
1	<a href="#">BioProject</a>	108088
2	<a href="#">GEO</a>	29441
3	<a href="#">DOE - JOINT GENOME INSTITUTE</a>	2590
4	<a href="#">UMIGS</a>	2557
5	<a href="#">JGI</a>	2365

# W23: DRA

The screenshot shows a web browser window with the URL `ddbj.nig.ac.jp/DRAsearch/query?acc=GSE107337&show=20&sort=Study`. The page title is "Result List - DRA Search". The search interface includes fields for "Accession" (GSE107337), "Organism", "CenterName", "Keyword", "StudyType", and "Platform". Below the search fields, there are controls for "Show 20 records" and "Sort by Study". The "Search Results" section is currently empty, displaying "Search Results ( 0 studies )". A red arrow with the number "1" points to the "0" in the search results count. At the bottom of the page, there is a footer with "Website policy | © DNA Data Bank of Japan".

# W24: DRA

①Keywordの場所でGSE107337で、②Searchすると、③3つヒットします。④SRP125628をクリック

Result List - DRA Search

保護されていない通信 | ddbj.nig.ac.jp/DRASearch/query?keyword=GSE107337&show=20

## DRASearch

Search Home DRA Home

Accession :

Organism :  StudyType :

CenterName :  Platform :

Keyword :  ①

Show 20 records Sort by Study Search Clear

Search Results ( 3 records ) ③ ② << < 1 / 1 Page > >>

**Filtered by**  
document type:experiment(1) study(1) submission(1)  
organism:Lactobacillus rhamnosus GG(3)

#	META_FILE	ACCESSION	STUDY	STUDY_TITLE	STUDY_TYPE	ORGANISM	BASES	SUBMITTED	CENTER_NAME
1	<a href="#">SRA633931_submission.xml</a> <?xml version="1.0" encoding="UTF-8"?><SUBMISSION alias="GEO:GSE107337" broker	<a href="#">SRA633931</a>	<a href="#">SRP125628</a>	RNA-seq analysis of Lactobacillus at acidic stress	<a href="#">Transcriptome Analysis</a>	<a href="#">Lactobacillus rhamnosus GG</a>	0		<a href="#">GEO</a>
2	<a href="#">SRA633931_study.xml</a> .org/2001/XMLSchema-instance"><STUDY center_name="GEO" alias="GSE107337" accession="SRP125628"><IDENTIFIERS> <PRIMARY	<a href="#">SRP125628</a>	<a href="#">SRP125628</a>	RNA-seq analysis of Lactobacillus at acidic stress	<a href="#">Transcriptome Analysis</a>	<a href="#">Lactobacillus rhamnosus GG</a>	0		<a href="#">GEO</a>
3	<a href="#">SRA633931_experiment.xml</a> ; Lactobacillus rhamnosus GG; RNA-Seq</TITLE> <STUDY_REF accession="SRP125628" refname="GSE107337	<a href="#">SRX3422361</a> <a href="#">SRX3422362</a> <a href="#">SRX3422363</a> <a href="#">SRX3422364</a> <a href="#">SRX3422365</a> <a href="#">SRX3422366</a> <a href="#">SRX3422367</a> <a href="#">SRX3422368</a>	<a href="#">SRP125628</a>	RNA-seq analysis of Lactobacillus at acidic stress	<a href="#">Transcriptome Analysis</a>	<a href="#">Lactobacillus rhamnosus GG</a>	0		<a href="#">GEO</a>

①こんな感じでSRP125628のページに飛びます。

# W24: DRA

The screenshot shows a web browser window with two tabs: 'Result List - DRA Search' and 'SRP125628 - DRA Search'. The address bar shows the URL 'ddbj.nig.ac.jp/DRASearch/study?acc=SRP125628'. The page title is 'DRASearch' and the main heading is 'SRP125628'. A red arrow with the number 1 points to the 'SRP125628' link. The page is divided into two main sections: 'Study Detail' and 'Navigation'.

Study Detail	
Title	RNA-seq analysis of Lactobacillus at acidic stress
Study Type	Transcriptome Analysis
Abstract	To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes
Description	
Center Name	GEO

Navigation	
Submission	<a href="#">SRA633931</a>
Experiment	<a href="#">SRX3422361</a>
	<a href="#">SRX3422362</a>
	<a href="#">SRX3422363</a>
	<a href="#">SRX3422364</a>
	<a href="#">SRX3422365</a>
	<a href="#">SRX3422366</a>
	<a href="#">SRX3422367</a>
	<a href="#">SRX3422368</a>
	<a href="#">SRX3422369</a>
Sample	<a href="#">SRS2714081</a>
	<a href="#">SRS2714082</a>
	<a href="#">SRS2714083</a>
	<a href="#">SRS2714084</a>
	<a href="#">SRS2714085</a>
	<a href="#">SRS2714086</a>
	<a href="#">SRS2714087</a>
	<a href="#">SRS2714088</a>
	<a href="#">SRS2714089</a>

Website policy | © DNA Data Bank of Japan

## W25: DRA

Result List - DRA Search x DRA Search x +

保護されていない通信 | ddbj.nig.ac.jp/DRASearch/ ☆ 人 | ⋮

**DRASearch** Search Home DRA Home

Accession :  ①

Organism :  StudyType :

CenterName :  Platform :

Keyword :

Show 20 records Sort by Study Search Clear

Data Last Update 2019-02-01

### Statistics

Released Entries

Type	Count
<a href="#">Submission</a>	1112526
<a href="#">Study</a>	178800
<a href="#">Experiment</a>	5504151
<a href="#">Sample</a>	4947100
<a href="#">Run</a>	6227253

Organism			Study Type			Center Name		
#	Organism Name	Study	#	Study Type	Study	#	Center Name	Study
1	<a href="#">Homo sapiens</a>	16356	1	<a href="#">Other</a>	69729	1	<a href="#">BioProject</a>	108088
2	<a href="#">Mus musculus</a>	13445	2	<a href="#">Whole Genome Sequencing</a>	56641	2	<a href="#">GEO</a>	29441
3	<a href="#">soil metagenome</a>	5530	3	<a href="#">Metagenomics</a>	26156	3	<a href="#">DOE - JOINT GENOME INSTITUTE</a>	2590
4	<a href="#">Zea mays</a>	3281	4	<a href="#">Transcriptome Analysis</a>	24085	4	<a href="#">UMIGS</a>	2557
5	<a href="#">Populus trichocarpa</a>	3064	5	<a href="#">Population Genomics</a>	802	5	<a href="#">JGI</a>	2365

# W25: DRA

- ①こんな感じでSRP125628のページに飛びます。
- ②FASTQおよびSRAファイルともにまだ提供されていないことが分かります。

Result List - DRA Search × SRP125628 - DRA Search × +

保護されていない通信 | ddbj.nig.ac.jp/DRASearch/study?acc=SRP125628

**DRASearch** Search Home DRA Home

**SRP125628** ①

Study Detail	
Title	RNA-seq analysis of Lactobacillus at acidic stress
Study Type	Transcriptome Analysis
Abstract	To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions Overall design: Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes
Description	
Center Name	GEO

Navigation	
Submission	<a href="#">SRA633931</a> FTP
Experiment	<a href="#">SRX3422361</a> FASTQ SRA
	<a href="#">SRX3422362</a> FASTQ SRA
	<a href="#">SRX3422363</a> FASTQ SRA
	<a href="#">SRX3422364</a> FASTQ SRA
	<a href="#">SRX3422365</a> FASTQ SRA
	<a href="#">SRX3422366</a> FASTQ SRA
	<a href="#">SRX3422367</a> FASTQ SRA
	<a href="#">SRX3422368</a> FASTQ SRA
	<a href="#">SRX3422369</a> FASTQ SRA
Sample	<a href="#">SRS2714081</a>
	<a href="#">SRS2714082</a>
	<a href="#">SRS2714083</a>
	<a href="#">SRS2714084</a>
	<a href="#">SRS2714085</a>
	<a href="#">SRS2714086</a>
	<a href="#">SRS2714087</a>
	<a href="#">SRS2714088</a>
	<a href="#">SRS2714089</a>

Website policy | © DNA Data Bank of Japan

# W26: 実験デザイン

The screenshot shows a web browser window with the URL <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337>. The page displays the NCBI logo and the GEO Gene Expression Omnibus logo. A navigation menu includes links for HOME, SEARCH, SITE MAP, GEO Publications, FAQ, MIAME, and Email GEO. The breadcrumb trail is NCBI > GEO > Accession Display. The search bar contains the text "Scope: Self", "Format: HTML", "Amount: Quick", and "GEO accession: GSE107337". A "GO" button is next to the search bar. Below the search bar, the page title is "Series GSE107337" and the subtitle is "Query DataSets for GSE107337". The main content area shows the following details:

Status	Public on Nov 29, 2017
Title	RNA-seq analysis of Lactobacillus at acidic stress
Organism	<a href="#">Lactobacillus rhamnosus GG</a>
Experiment type	Expression profiling by high throughput sequencing
Summary	To understand transcriptional regulation of probiotic bacteria under acidic condition, RNAseq analysis was carried out over different growth conditions
Overall design	Comparison of acidic (pH4) and neutral (pH7) conditions by differentially expressed genes
Contributor(s)	<a href="#">Choi I, Oh S</a>
Citation missing	<i>Has this study been published? Please <a href="#">login</a> to update or <a href="#">notify GEO</a>.</i>
Submission date	Nov 25, 2017
Last update date	Jan 24, 2018
Contact name	kucsbl submitter

# W26: 実験デザイン

①このあたりまで移動。②Moreを押してサンプル全ての表示。

GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

City	Seoul
State/province	Seoul
ZIP/Postal code	KS013
Country	South Korea

Platforms (1) [GPL24302](#) Illumina MiSeq (Lactobacillus rhamnosus GG)

Samples (9)  
[More...](#)

<a href="#">GSM2864941</a>	pH4_1h rep1
<a href="#">GSM2864942</a>	pH4_1h rep2
<a href="#">GSM2864943</a>	pH4_1h rep3

**Relations**

BioProject	<a href="#">PRJNA419802</a>
SRA	<a href="#">SRP125628</a>

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	<a href="#">SOFT</a> ?
<a href="#">MINiML formatted family file(s)</a>	<a href="#">MINiML</a> ?
<a href="#">Series Matrix File(s)</a>	<a href="#">TXT</a> ?

Supplementary file	Size	Download	File type/resource
<a href="#">GSE107337_RPKM.csv.gz</a>	21.1 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV
<a href="#">GSE107337_RawCounts.csv.gz</a>	20.8 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV

Raw data are available in SRA  
Processed data is available on Series record

| [NLM](#) | [NIH](#) | [GEO Help](#) | [Disclaimer](#) | [Accessibility](#) |

# W26: 実験デザイン

原著論文の記載内容と、①の名前の付け方から、実験デザインの全体像がわかる。

GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

Platforms (1) [GPL24302](#) Illumina MiSeq (Lactobacillus rhamnosus GG)

Samples (9)  
[Less...](#)

<a href="#">GSM2864941</a>	pH4_1h rep1
<a href="#">GSM2864942</a>	pH4_1h rep2
<a href="#">GSM2864943</a>	pH4_1h rep3
<a href="#">GSM2864944</a>	pH4_24h rep1
<a href="#">GSM2864945</a>	pH4_24h rep2
<a href="#">GSM2864946</a>	pH4_24h rep3
<a href="#">GSM2864947</a>	pH7_CCG rep1
<a href="#">GSM2864948</a>	pH7_CCG rep2
<a href="#">GSM2864949</a>	pH7_CCG rep3

**Relations**

BioProject [PRJNA419802](#)

SRA [SRP125628](#)

**Download family**

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE107337_RPKM.csv.gz</a>	21.1 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV
<a href="#">GSE107337_RawCounts.csv.gz</a>	20.8 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV

Raw data are available in SRA

# W27: 表1

	A	B	C	D	E	F
1						
2		サンプル名	SRR ID	リード数(片側のみ)	ファイルサイズ(forward側)	ファイルサイズ(reverse側)
3		pH4.5_1h rep1	SRR6322562	301,126	53,833,728	59,662,336
4		pH4.5_1h rep2	SRR6322563	1,470,602	265,838,592	296,910,848
5		pH4.5_1h rep3	SRR6322564	1,760,461	319,807,488	354,557,952
6		pH4.5_24h rep1	SRR6322565	1,375,368	241,553,408	268,341,248
7		pH4.5_24h rep2	SRR6322566	3,869,088	682,504,192	759,697,408
8		pH4.5_24h rep3	SRR6322567	1,795,874	315,613,184	353,099,776
9		pH7_CCG rep1	SRR6322568	3,095,834	551,383,040	615,870,464
10		pH7_CCG rep2	SRR6322569	2,570,876	451,117,056	510,308,352
11		pH7_CCG rep3	SRR6322570	846,623	148,500,480	167,350,272
12						

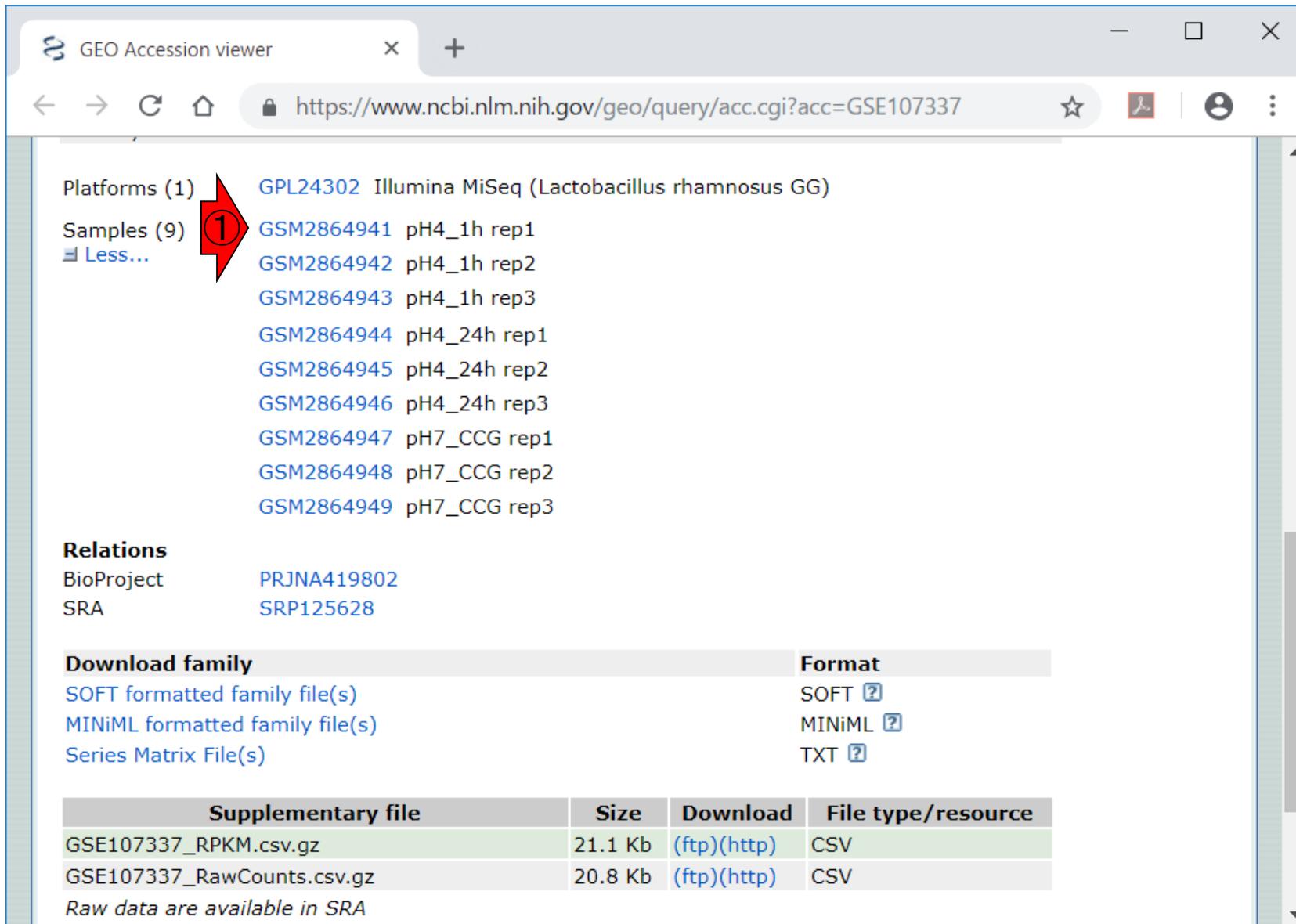
# W27: 表1

①リード数最大と、②最小の間には、10倍以上の違いがある。

	A	B	C	D	E	F
1						
2		サンプル名	SRR ID	リード数(片側のみ)	ファイルサイズ(forward側)	ファイルサイズ(reverse側)
3		pH4.5_1h rep1	SRR6322562	301,126	53,833,728	59,662,336
4		pH4.5_1h rep2	SRR6322563	1,470,602	265,838,592	296,910,848
5		pH4.5_1h rep3	SRR6322564	1,760,461	319,807,488	354,557,952
6		pH4.5_24h rep1	SRR6322565	1,375,368	241,553,408	268,341,248
7		pH4.5_24h rep2	SRR6322566	3,869,088	682,504,192	759,697,408
8		pH4.5_24h rep3	SRR6322567	1,795,874	315,613,184	353,099,776
9		pH7_CCG rep1	SRR6322568	3,095,834	551,383,040	615,870,464
10		pH7_CCG rep2	SRR6322569	2,570,876	451,117,056	510,308,352
11		pH7_CCG rep3	SRR6322570	846,623	148,500,480	167,350,272
12						

W26で見ているものと同じ。  
①GSM2864941をクリック

# W28: GSM2864941



GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

Platforms (1) [GPL24302](#) Illumina MiSeq (Lactobacillus rhamnosus GG)

Samples (9) [GSM2864941](#) pH4\_1h rep1  
[GSM2864942](#) pH4\_1h rep2  
[GSM2864943](#) pH4\_1h rep3  
[GSM2864944](#) pH4\_24h rep1  
[GSM2864945](#) pH4\_24h rep2  
[GSM2864946](#) pH4\_24h rep3  
[GSM2864947](#) pH7\_CCG rep1  
[GSM2864948](#) pH7\_CCG rep2  
[GSM2864949](#) pH7\_CCG rep3

[Less...](#)

**Relations**

BioProject [PRJNA419802](#)  
SRA [SRP125628](#)

**Download family**

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE107337_RPKM.csv.gz</a>	21.1 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV
<a href="#">GSE107337_RawCounts.csv.gz</a>	20.8 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV

Raw data are available in SRA

# W28: GSM2864941

- ①GSM2864941のページに移動しました。
- ②少しページ下部に移動。

NCBI > GEO > **Accession Display** Not logged in | Login

**GEO help:** Mouse over screen elements for information.

Scope:  Format:  Amount:  GEO accession:

**Sample GSM2864941** [Query DataSets for GSM2864941](#)

Status	Public on Nov 29, 2017
Title	pH4_1h rep1
Sample type	SRA
Source name	whole transcriptome
Organism	<a href="#">Lactobacillus rhamnosus GG</a>
Characteristics	ph: 4
Extracted molecule	total RNA
Extraction protocol	TruSeq RNA Sample Preparation Kit (Illumina) was used. RNA libraries were prepared for sequencing using standard Illumina protocols
Library strategy	RNA-Seq
Library source	transcriptomic
Library selection	cDNA
Instrument model	Illumina MiSeq

# W28: GSM2864941

①このあたりまで移動すると、②Data processing  
情報が見つかります。リファレンスゲノムとして、③  
NC\_013198.1が利用されていることがわかります。

GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM2864941

Source name whole transcriptome  
Organism [Lactobacillus rhamnosus GG](#)  
Characteristics ph: 4  
Extracted molecule total RNA  
Extraction protocol TruSeq RNA Sample Preparation Kit (Illumina) was used.  
RNA libraries were prepared for sequencing using standard Illumina protocols  
Library strategy RNA-Seq  
Library source transcriptomic  
Library selection cDNA  
Instrument model Illumina MiSeq

Data processing Read data were aligned on the reference genome sequence.  
DESeq and edgeR package in the Bioconductor/R packages were used for  
making RPKM values.  
Base-calling by the MiSeq platform  
Alignment by Bowtie software  
Using R packages for DEG analysis  
Genome\_build: Lactobacillus rhamnosus GG (ATCC 53103, RefSeq ID:  
NC\_013198.1)  
Supplementary\_files\_format\_and\_content: CSV file of normalized RPKM value  
from edgeR for every samples.

Submission date Nov 25, 2017  
Last update date Nov 29, 2017  
Contact name kucsbl submitter  
E-mail [kucsbl.group@gmail.com](mailto:kucsbl.group@gmail.com)  
Organization name Korea University  
Department Department of Biotechnology

# W29: BowtieとRPKM

①mapping (alignment)にはBowtieが使われており、②RPKMの作成にはedgeRが用いられていると書かれています。ここではDESeqも書かれていますが、別のところではedgeRのみしか書かれていなかったためDESeqは無視しています

GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM2864941

Source name whole transcriptome  
Organism [Lactobacillus rhamnosus GG](#)  
Characteristics ph: 4  
Extracted molecule total RNA  
Extraction protocol TruSeq RNA Sample Preparation Kit (Illumina) was used.  
RNA libraries were prepared for sequencing using standard Illumina protocols  
  
Library strategy RNA-Seq  
Library source transcriptomic  
Library selection cDNA  
Instrument model Illumina MiSeq

Data processing Read data were aligned on the reference genome sequence.  
② DESeq and edgeR package in the Bioconductor/R packages were used for making RPKM values.  
① Base-calling by the MiSeq platform  
Alignment by Bowtie software  
Using R packages for DEG analysis  
Genome\_build: Lactobacillus rhamnosus GG (ATCC 53103, RefSeq ID: NC\_013198.1)  
Supplementary\_files\_format\_and\_content: CSV file of normalized RPKM value from edgeR for every samples.

Submission date Nov 25, 2017  
Last update date Nov 29, 2017  
Contact name kucsbl submitter  
E-mail [kucsbl.group@gmail.com](mailto:kucsbl.group@gmail.com)  
Organization name Korea University  
Department Department of Biotechnology

# W30: CountとRPKM

W26と同じく、①GSE107337のページ。  
②RPKMデータファイルと、③カウントデータファイルをダウンロードしておく。

GEO Accession viewer

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107337

Platforms (1) [GPL24302](#) Illumina MiSeq (Lactobacillus rhamnosus GG)

Samples (9)  
[Less...](#)

- [GSM2864941](#) pH4\_1h rep1
- [GSM2864942](#) pH4\_1h rep2
- [GSM2864943](#) pH4\_1h rep3
- [GSM2864944](#) pH4\_24h rep1
- [GSM2864945](#) pH4\_24h rep2
- [GSM2864946](#) pH4\_24h rep3
- [GSM2864947](#) pH7\_CCG rep1
- [GSM2864948](#) pH7\_CCG rep2
- [GSM2864949](#) pH7\_CCG rep3

**Relations**

BioProject [PRJNA419802](#)

SRA [SRP125628](#)

**Download family**

Download family	Format
<a href="#">SOFT formatted family file(s)</a>	SOFT <a href="#">?</a>
<a href="#">MINiML formatted family file(s)</a>	MINiML <a href="#">?</a>
<a href="#">Series Matrix File(s)</a>	TXT <a href="#">?</a>

Supplementary file	Size	Download	File type/resource
<a href="#">GSE107337_RPKM.csv.gz</a>	21.1 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV
<a href="#">GSE107337_RawCounts.csv.gz</a>	20.8 Kb	<a href="#">(ftp)</a> <a href="#">(http)</a>	CSV

Raw data are available in SRA

Processed data is available on Series record

# W31: カウントデータ

①raw countデータファイルのほう。② averageと書かれていることからわかるように、3状態間比較で、反復データの平均値になっている。だから3列分しかない。

自動保存 GSE107337\_RawCounts.csv

ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 ヘルプ 操作アシスト 共有

E27

	A	pH4_1h_average_rawcount	pH4_24h_average_rawcount	pH7_CCG_average_rawcount	E
1	Gene				
2	LGG_00001	286	676	545	
3	LGG_00002	776	1601	1450	
4	LGG_00003	71	114	91	
5	LGG_00004	201	470	368	
6	LGG_00005	368	1160	1447	
7	LGG_00006	1041	2742	2880	
8	LGG_00007	7	10	2	
9	LGG_00008	73	76	499	
10	LGG_00009	16	45	42	
11	LGG_00010	20	65	57	
12	LGG_00011	2281	2596	2150	
13	LGG_00012	2510	3631	3097	
14	LGG_00013	971	1301	1424	
15	LGG_00014	17	78	16	
16	LGG_00015	82	255	33	

GSE107337\_RawCounts

ScrollLock + 100%

# W32: 記載内容と一致

①1列目が遺伝子名、②2列目がpH4.5の酸ストレスを1時間与えた平均のカウント、③3列目がpH4.5の酸ストレスを24時間与えた平均のカウント、④4列目がpH7のコントロールの平均のカウント、であることを覚えておく。

	①	②	③	④	E
1	Gene	pH4_1h_average_rawcount	pH4_24h_average_rawcount	pH7_CCG_average_rawcount	
2	LGG_00001	286	676	545	
3	LGG_00002	776	1601	1450	
4	LGG_00003	71	114	91	
5	LGG_00004	201	470	368	
6	LGG_00005	368	1160	1447	
7	LGG_00006	1041	2742	2880	
8	LGG_00007	7	10	2	
9	LGG_00008	73	76	499	
10	LGG_00009	16	45	42	
11	LGG_00010	20	65	57	
12	LGG_00011	2281	2596	2150	
13	LGG_00012	2510	3631	3097	
14	LGG_00013	971	1301	1424	
15	LGG_00014	17	78	16	
16	LGG_00015	82	255	33	

# W32: 記載内容と一致

原著論文のp1608の左上のほうでは、① LGG\_02240 (GroES gene) の②24時間ストレスの発現が、③コントロールに比べて、2.53倍変化したと記載されている。

	A	B	C	D	E
2151	LGG_02232	33	95	70	
2152	LGG_02233	227	331	539	
2153	LGG_02234	406	791	1093	
2154	LGG_02235	582	999	1289	
2155	LGG_02236	277	467	647	
2156	LGG_02238	768	674	577	
2157	LGG_02239	5146	18308	5690	
2158	LGG_02240	1106	2444	422	
2159	LGG_02241	87	107	130	
2160	LGG_02242	377	340	540	
2161	LGG_02244	36	116	44	
2162	LGG_02245	19	101	29	
2163	LGG_02246	23	63	8	
2164	LGG_02247	384	531	556	
2165	LGG_02248	245	405	318	
2166	LGG_02249	125	188	227	

# W32: 記載内容と一致

原著論文のp1608の左上のほうでは、① LGG\_02240 (GroES gene) の②24時間ストレスの発現が、③コントロールに比べて、2.53倍変化したと記載されている。倍率変化なので、④5.79倍変化したと書くべきだとは思われるが、⑤log2変換した結果と確かに一致する。論文のTable S2の脚注部分で、確かにlog2変換したものをFold changeと表現している…まあ、妥当

	A	B	C
2151	LGG_02232	33	95
2152	LGG_02233	227	331
2153	LGG_02234	406	791
2154	LGG_02235	582	999
2155	LGG_02236	277	467
2156	LGG_02238	768	674
2157	LGG_02239	5146	18308
2158	LGG_02240	1106	2444
2159	LGG_02241	87	107
2160	LGG_02242	377	340
2161	LGG_02244	36	116
2162	LGG_02245	19	101
2163	LGG_02246	23	63
2164	LGG_02247	384	531
2165	LGG_02248	245	405
2166	LGG_02249	125	188

```
R Console
> 2444/422
[1] 5.791469
> log2(2444/422)
[1] 2.533929
> |
```