2014-08-20
ゲノム支援拡大班会議（神戸ポートピアホテル）

DDBJ
DNA Data Bank of Japan

Togo picture gallery by DBCLS is Licensed under a Creative Commons 表示 2.1 日本 (c)

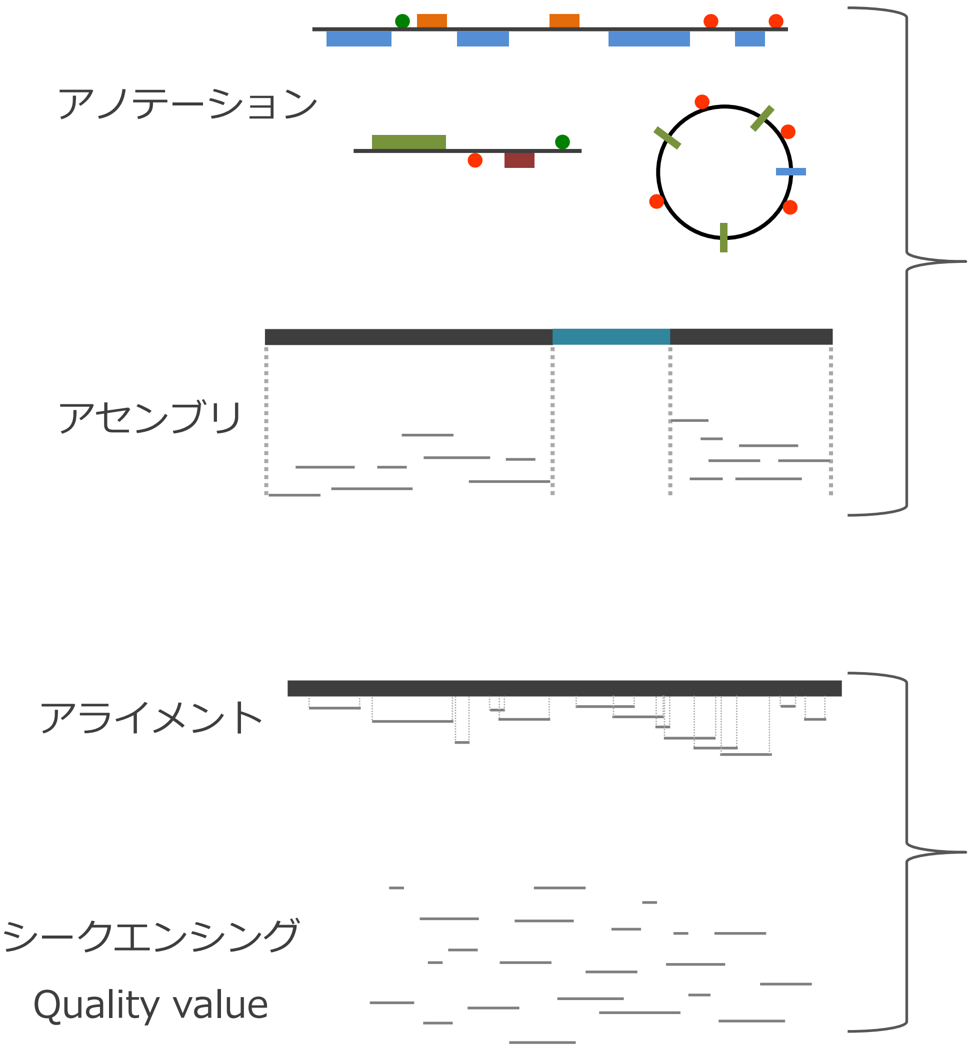
BioProject, BioSample, DRA へのデータ登録

古屋典子
Noriko Furuya, PhD

国立遺伝学研究所　DDBJ センター、アノテータ
senior curator, DDBJ center, National Institute of Genetics

INSDC*の塩基配列データベースには、大きく2種類

* **INSDC:**
DDBJ center, EMBL-EBI, NCBI



アノテーション

アセンブリ

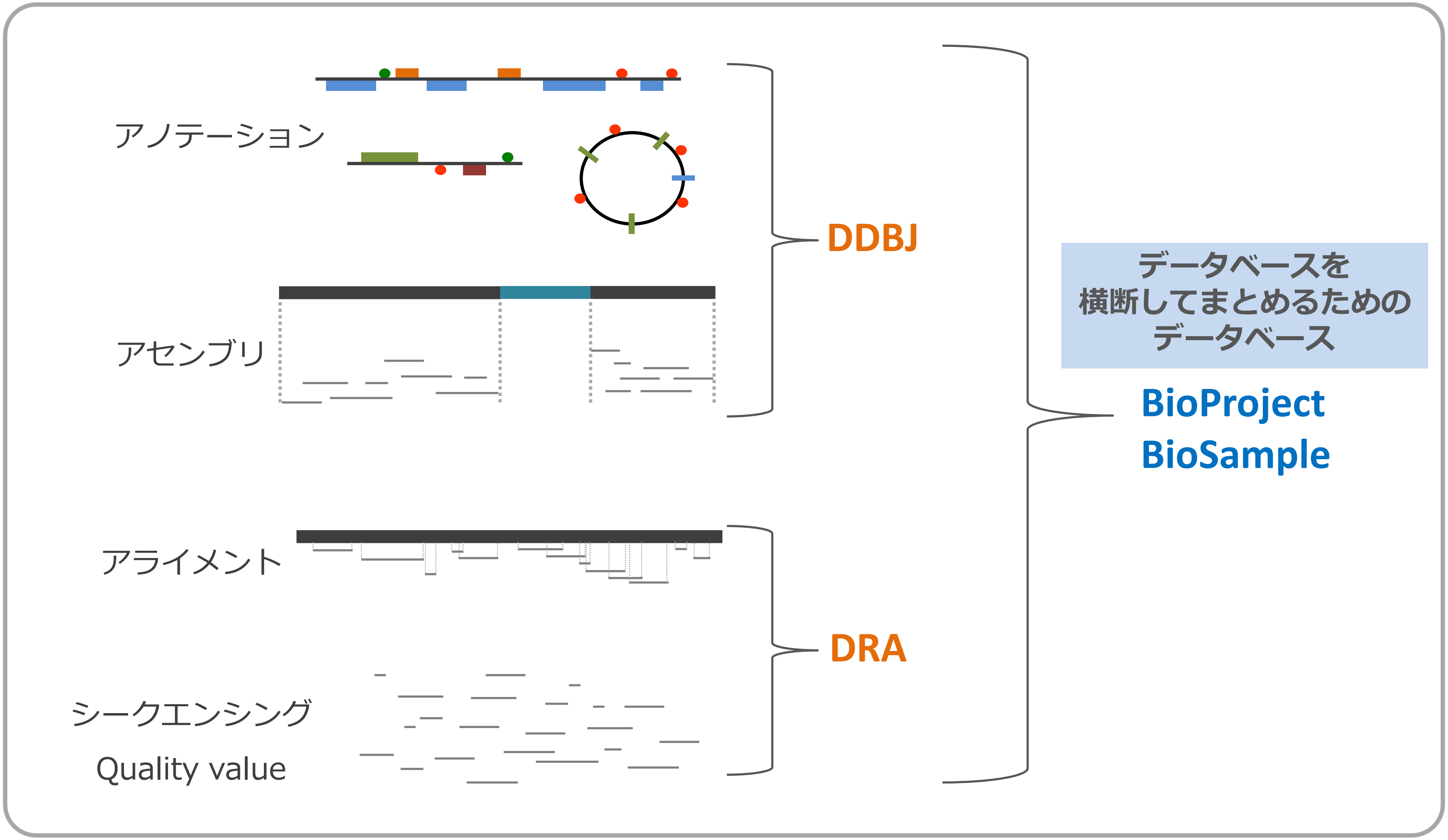
DDBJ / ENA / GenBank

アライメント

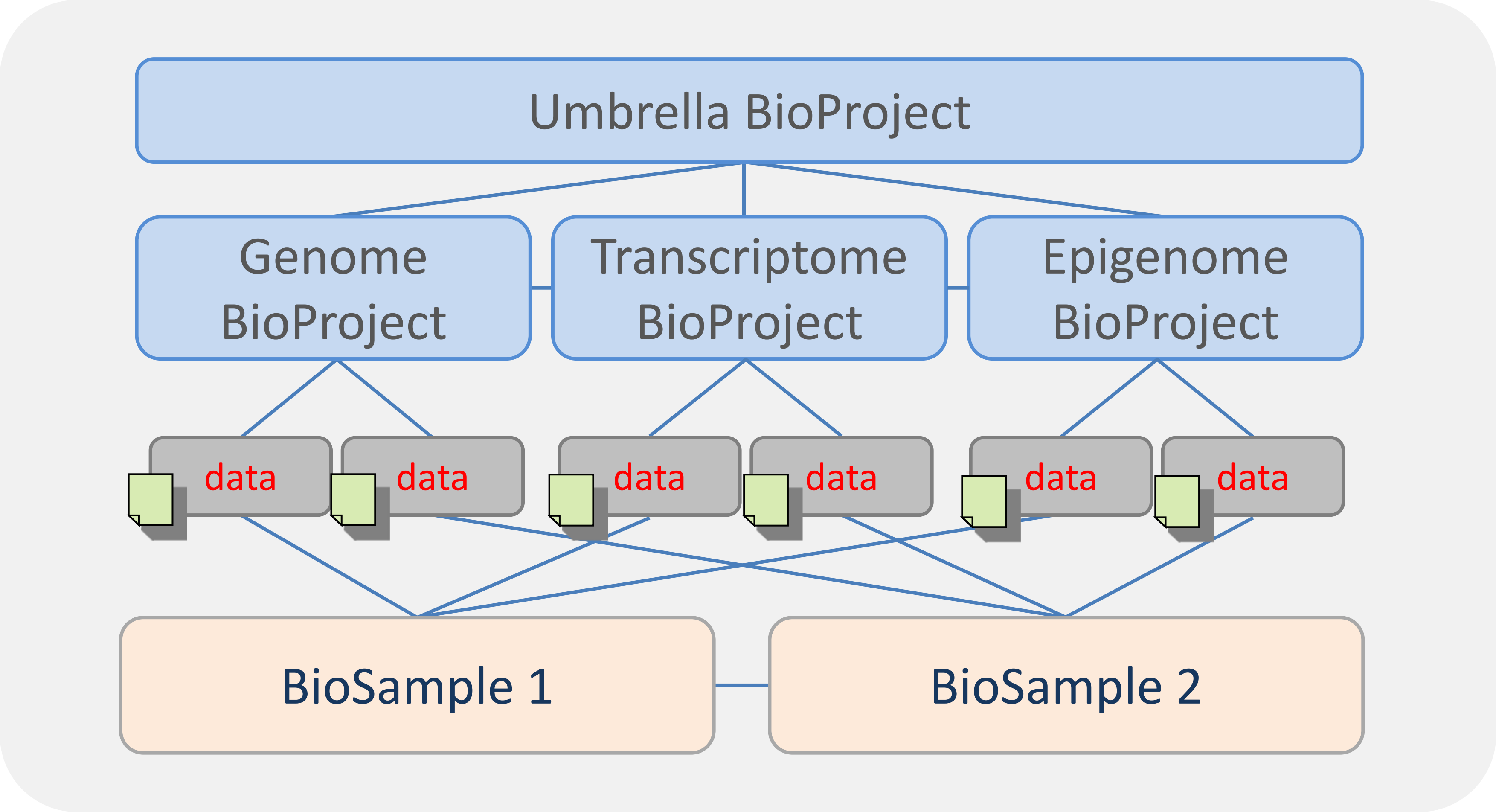
シークエンシング
Quality value

DRA / ENA / SRA
(Sequence Read Archive)

DDBJセンターの中には、DDBJ と DRA の2種類

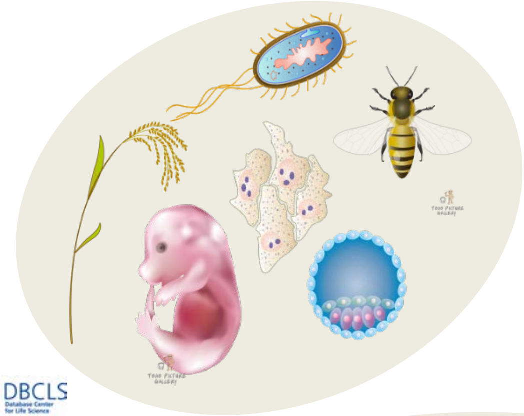


BioProject と BioSample は、「プロジェクト」と「サンプル」という切り口で、多様なデータを組織化する



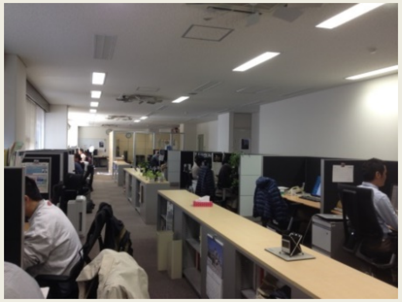
同じIDを引用しデータを関連付ける

DRA は、次世代シーケンサ由来の1次データを格納する



```
@HWI-
ST132:683:D1RDYACXX:6:2316:20294:
100841#12/2
GATGTTCAAGGACAAGAAGCACCCAGC
ACTGGAGAAAGAT
+
FBGIIIIIGGGHIIEJHFEBEHGGHGHEHFHF
HDDDDD=1
```

DDBJ
DNA Data Bank of Japan



DDBJ Sequence Read Archive (DRA)

※ ヒト由来データの
制限公開は、JGAへ。
http://trace.ddbj.nig.ac.jp/jga/index.html

DRA 登録には、メタデータとデータファイルが必要

メタデータ

metadata = "data about data"

- ＊サブミッション情報　Submission
  Submitter, Organization etc.
- ＊研究情報　Study (=BioProjct)
  Title, Purpose, Grant etc.
- ＊サンプル情報　Sample (=BioSample)
  Scientific name, Location etc.
- ＊実験情報　Experiment
  platform, spot length etc.
- ＊ラン情報　Run
  file name, MD5 value etc.

データファイル

シークエンスデータ　もしくは　アライメントデータ

シークエンスデータ

Roche 454 : .sff もしくは .fastq
Illumina : .fastq もしくは .qseq
Applied Biosystems : .csfasta と .QV.qual
Ion Torrent : .sff もしくは .fastq
Helicos Heliscope : .sms もしくは .fastq
Complete Genomics : .fastq
Pacific Biosciences : .fastq

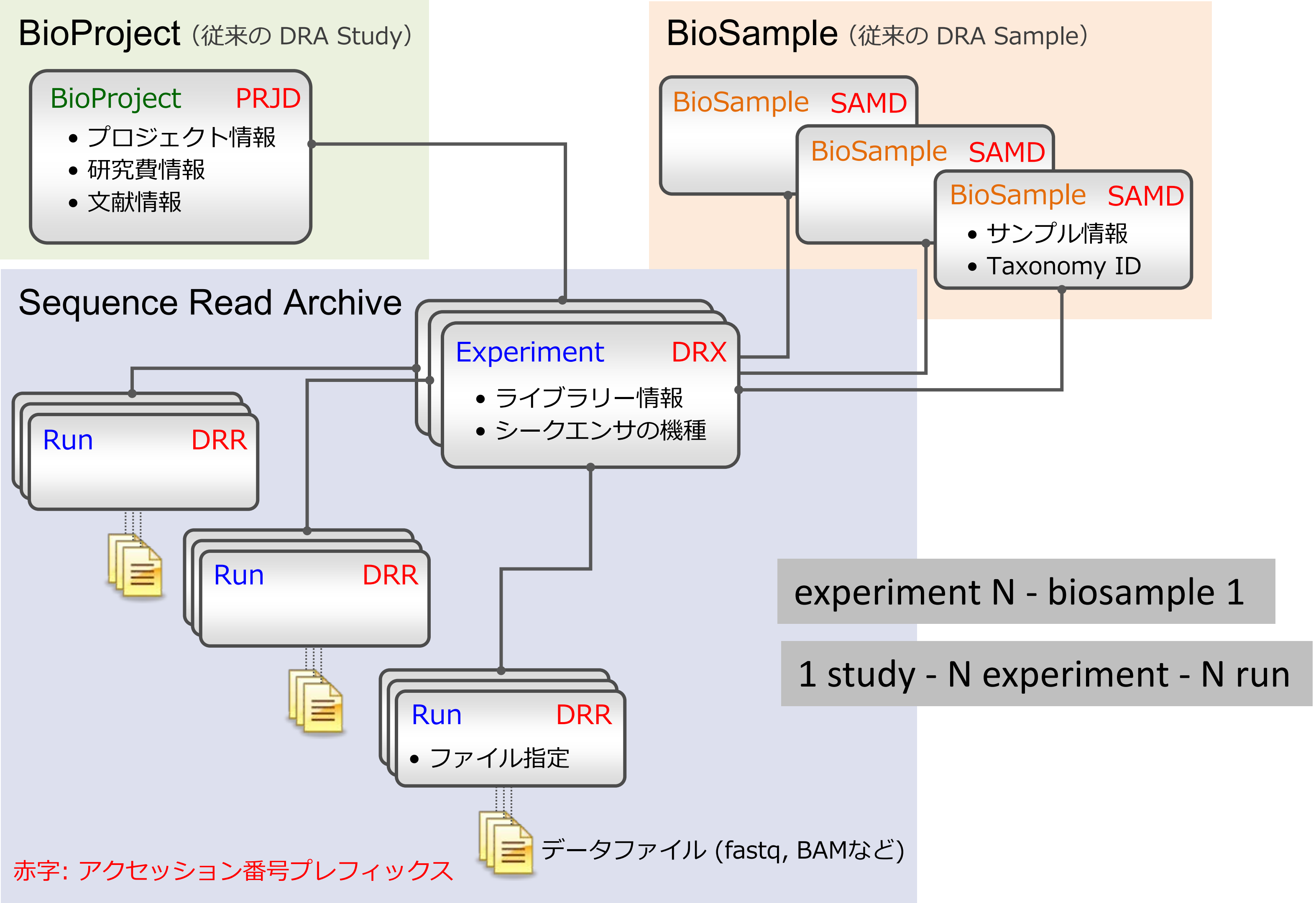
アライメントデータ（3点セット）

1) BAM
2) リファレンス配列
3) SN-リファレンス配列の対応表

データベースで公開する

.sra

メタデータは、複数のオブジェクトで構成される



BioProject（従来の DRA Study）

BioProject PRJD
- プロジェクト情報
- 研究費情報
- 文献情報

BioSample（従来の DRA Sample）

BioSample SAMD

BioSample SAMD

BioSample SAMD
- サンプル情報
- Taxonomy ID

Sequence Read Archive

Experiment DRX
- ライブラリー情報
- シークエンサの機種

Run DRR

Run DRR

Run DRR
- ファイル指定

データファイル (fastq, BAMなど)

赤字: アクセッション番号プレフィックス

experiment N - biosample 1

1 study - N experiment - N run

アクセッション番号は各オブジェクトに割り振られる

メタデータは、機械処理に適した XML ファイルで管理

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<EXPERIMENT_SET>
  <EXPERIMENT accession="DRX0xxxxxx" center_name=“NIG" alias=“furuya-0001_Experiment_0001">
    <TITLE>WGS of Musa balbisiana var. liukiuensis</TITLE>
    <STUDY_REF accession="DRP00xxxx" refcenter=“NIG" refname=“furuya-0001_Study_0001"/>
    <DESIGN>
      <DESIGN_DESCRIPTION></DESIGN_DESCRIPTION>
      <SAMPLE_DESCRIPTOR accession="DRS0xxxxxx" refcenter=“NIG" refname=“furuya-0001_Sample_0001"/>
      <LIBRARY_DESCRIPTOR>
        <LIBRARY_STRATEGY>WGS</LIBRARY_STRATEGY>
        <LIBRARY_SOURCE>GENOMIC</LIBRARY_SOURCE>
        <LIBRARY_SELECTION>RANDOM</LIBRARY_SELECTION>
        <LIBRARY_LAYOUT>
          <PAIRED/>
        </LIBRARY_LAYOUT>
        <POOLING_STRATEGY>multiplexed libraries</POOLING_STRATEGY>
      </LIBRARY_DESCRIPTOR>
      <SPOT_DESCRIPTOR>
        <SPOT_DECODE_SPEC>
          <SPOT_LENGTH>250</SPOT_LENGTH>
          <READ_SPEC>
            <READ_INDEX>0</READ_INDEX>
            <READ_CLASS>Technical Read</READ_CLASS>
            <READ_TYPE>Adapter</READ_TYPE>
            <BASE_COORD>1</BASE_COORD>
          </READ_SPEC>
          <READ_SPEC>
            <READ_INDEX>1</READ_INDEX>
            <READ_CLASS>Application Read</READ_CLASS>
            <READ_TYPE>Forward</READ_TYPE>
            <BASE_COORD>5</BASE_COORD>
          </READ_SPEC>
        </SPOT_DECODE_SPEC>
      </SPOT_DESCRIPTOR>
    </DESIGN>
    <PLATFORM>
      <ILLUMINA>
        <INSTRUMENT_MODEL>Illumina MiSeq</INSTRUMENT_M
      </ILLUMINA>
    </PLATFORM>
    <PROCESSING>
      <PIPELINE>
        <PIPE_SECTION>
          <STEP_INDEX>1</STEP_INDEX>
          <PREV_STEP_INDEX>NIL</PREV_STEP_INDEX>
```

Spot Descriptor には、アダプター配列などの technical read も設定可能

※　<SPOT_DESCRIPTOR> の記載例
http://trace.ddbj.nig.ac.jp/dra/example.html

インターフェース上でメタデータを作成後、
XMLファイルを出し入れすることで、
technical reads を表現することが出来ます。

まずは、DRA ウェブサイト を CHECK！

DRA HP: http://trace.ddbj.nig.ac.jp/dra/index.html

メンテナンスなどのNEWS

登録開始・再開
“D-way” ログイン

News

2014年08月13日 **New**: **DDBJ 夏季休業のお知らせ (8/14-15)** less...

2014年8月14日(木)～15日(金)、DRAおよびBioProject、BioSampleの窓口対応は休業となります。D-way 登録システム、DRA および BioProject、BioSample 検索システムは引き続きご利用可能です。

2014年08月06日: **登録システムがリード長が一定ではない fastq ファイルに対応** less...

DRA 登録システムがリード長が一定ではない fastq ファイルに対応いたしました。マニュアル

・リード長が揃っている fastq ファイル: Run の file type で fastq を選択
・リード長が一定ではない fastq ファイル: Run の file type で generic_fastq を選択

DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System などの次世代シークエンサからの出力データのためのデータベースです。 DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり， NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと，運営されています。 従来のキャピラリ式シークエンサからの出力データは DDBJ Trace Archive にご登録ください。

Databases
Nucleotide Sequence Database
Sequence Read Archive
Trace Archive
Omics Archive
BioProject
BioSample
Japanese Genotype-phenotype Archive

Resources
getentry
ARSA
TXSearch
BLAST
Vector Screening System
ClustalW
Read Annotation Pipeline
MiGAP

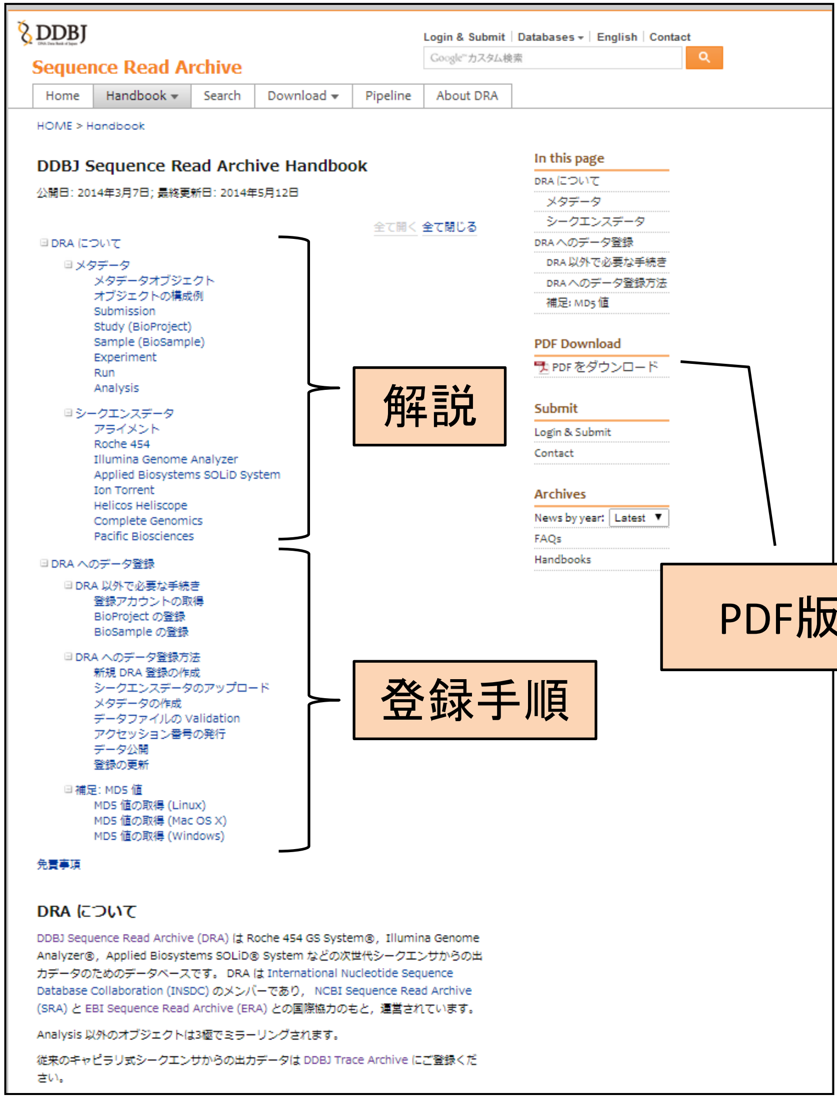
DDBJ Information
DDBJ RSS
DDBJ on Twitter
DDBJ on Youtube
DDBJ on slideshare
DDBJing
DDBJ FTP Site

DRASearch

DRA登録マニュアル

登録開始前に、マニュアルをご一読ください

DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html



DDBJ

Login & Submit | Databases ▾ | English | Contact

Googleカスタム検索

Sequence Read Archive

Home | Handbook ▾ | Search | Download ▾ | Pipeline | About DRA

HOME > Handbook

DDBJ Sequence Read Archive Handbook

公開日: 2014年3月7日; 最終更新日: 2014年5月12日

全て開く 全て閉じる

- DRA について
  - メタデータ
    - メタデータオブジェクト
    - オブジェクトの構成例
    - Submission
    - Study (BioProject)
    - Sample (BioSample)
    - Experiment
    - Run
    - Analysis
  - シークエンスデータ
    - アライメント
    - Roche 454
    - Illumina Genome Analyzer
    - Applied Biosystems SOLiD System
    - Ion Torrent
    - Helicos Heliscope
    - Complete Genomics
    - Pacific Biosciences
- DRA へのデータ登録
  - DRA 以外で必要な手続き
    - 登録アカウントの取得
    - BioProject の登録
    - BioSample の登録
  - DRA へのデータ登録方法
    - 新規 DRA 登録の作成
    - シークエンスデータのアップロード
    - メタデータの作成
    - データファイルの Validation
    - アクセッション番号の発行
    - データ公開
    - 登録の更新
  - 補足: MD5 値
    - MD5 値の取得 (Linux)
    - MD5 値の取得 (Mac OS X)
    - MD5 値の取得 (Windows)

免責事項

DRA について

DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System などの次世代シークエンサからの出力データのためのデータベースです。 DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり, NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと, 運営されています。

Analysis 以外のオブジェクトは3極でミラーリングされます。

従来のキャピラリ式シークエンサからの出力データは DDBJ Trace Archive にご登録ください。

In this page
- DRA について
  - メタデータ
  - シークエンスデータ
- DRA へのデータ登録
  - DRA 以外で必要な手続き
  - DRA へのデータ登録方法
  - 補足: MD5 値

PDF Download

PDF をダウンロード

Submit

Login & Submit

Contact

Archives

News by year: Latest ▾

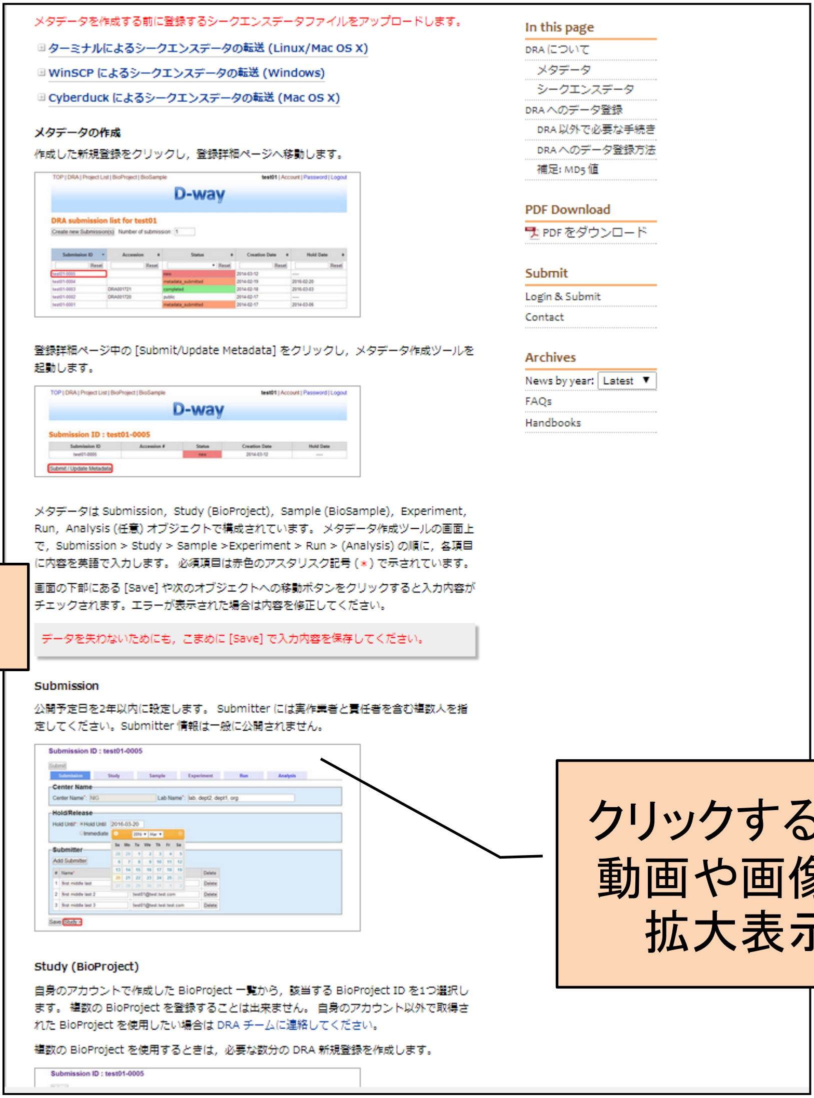
FAQs

Handbooks

解説

登録手順

PDF版



メタデータを作成する前に登録するシークエンスデータファイルをアップロードします。

- ターミナルによるシークエンスデータの転送 (Linux/Mac OS X)
- WinSCP によるシークエンスデータの転送 (Windows)
- Cyberduck によるシークエンスデータの転送 (Mac OS X)

メタデータの作成

作成した新規登録をクリックし, 登録詳細ページへ移動します。

TOP | DRA | Project List | BioProject | BioSample　test01 | Account | Password | Logout

D-way

DRA submission list for test01

Create new Submission(s)　Number of submission : 1

Submission ID	Accession	Status	Creation Date	Hold Date
test01-0005		new	2014-03-12	----
test01-0004		metadata_submitted	2014-02-19	2016-02-20
test01-0003	DRA001721	completed	2014-02-18	2016-03-03
test01-0002	DRA001720	public	2014-02-17	----
test01-0001		metadata_submitted	2014-02-17	2014-03-06

登録詳細ページ中の [Submit/Update Metadata] をクリックし, メタデータ作成ツールを起動します。

TOP | DRA | Project List | BioProject | BioSample　test01 | Account | Password | Logout

D-way

Submission ID : test01-0005

Submission ID	Accession #	Status	Creation Date	Hold Date
test01-0005		new	2014-03-12	----

Submit / Update Metadata

メタデータは Submission, Study (BioProject), Sample (BioSample), Experiment, Run, Analysis (任意) オブジェクトで構成されています。 メタデータ作成ツールの画面上で, Submission > Study > Sample >Experiment > Run > (Analysis) の順に, 各項目に内容を英語で入力します。 必須項目は赤色のアスタリスク記号 (*) で示されています。

画面の下部にある [Save] や次のオブジェクトへの移動ボタンをクリックすると入力内容がチェックされます。エラーが表示された場合は内容を修正してください。

データを失わないためにも, こまめに [Save] で入力内容を保存してください。

Submission

公開予定日を2年以内に設定します。 Submitter には実作業者と責任者を含む複数人を指定してください。Submitter 情報は一般に公開されません。

Submission ID : test01-0005

Submit

Submission | Study | Sample | Experiment | Run | Analysis

Center Name

Center Name*: NIG　Lab Name*: lab, dept2, dept1, org

Hold/Release

Submitter

Save　Study >

Study (BioProject)

自身のアカウントで作成した BioProject 一覧から, 該当する BioProject ID を1つ選択します。 複数の BioProject を登録することは出来ません。 自身のアカウント以外で取得された BioProject を使用したい場合は DRA チームに連絡してください。

複数の BioProject を使用するときは, 必要な数分の DRA 新規登録を作成します。

Submission ID : test01-0005

In this page
- DRA について
  - メタデータ
  - シークエンスデータ
- DRA へのデータ登録
  - DRA 以外で必要な手続き
  - DRA へのデータ登録方法
  - 補足: MD5 値

PDF Download

PDF をダウンロード

Submit

Login & Submit

Contact

Archives

News by year: Latest ▾

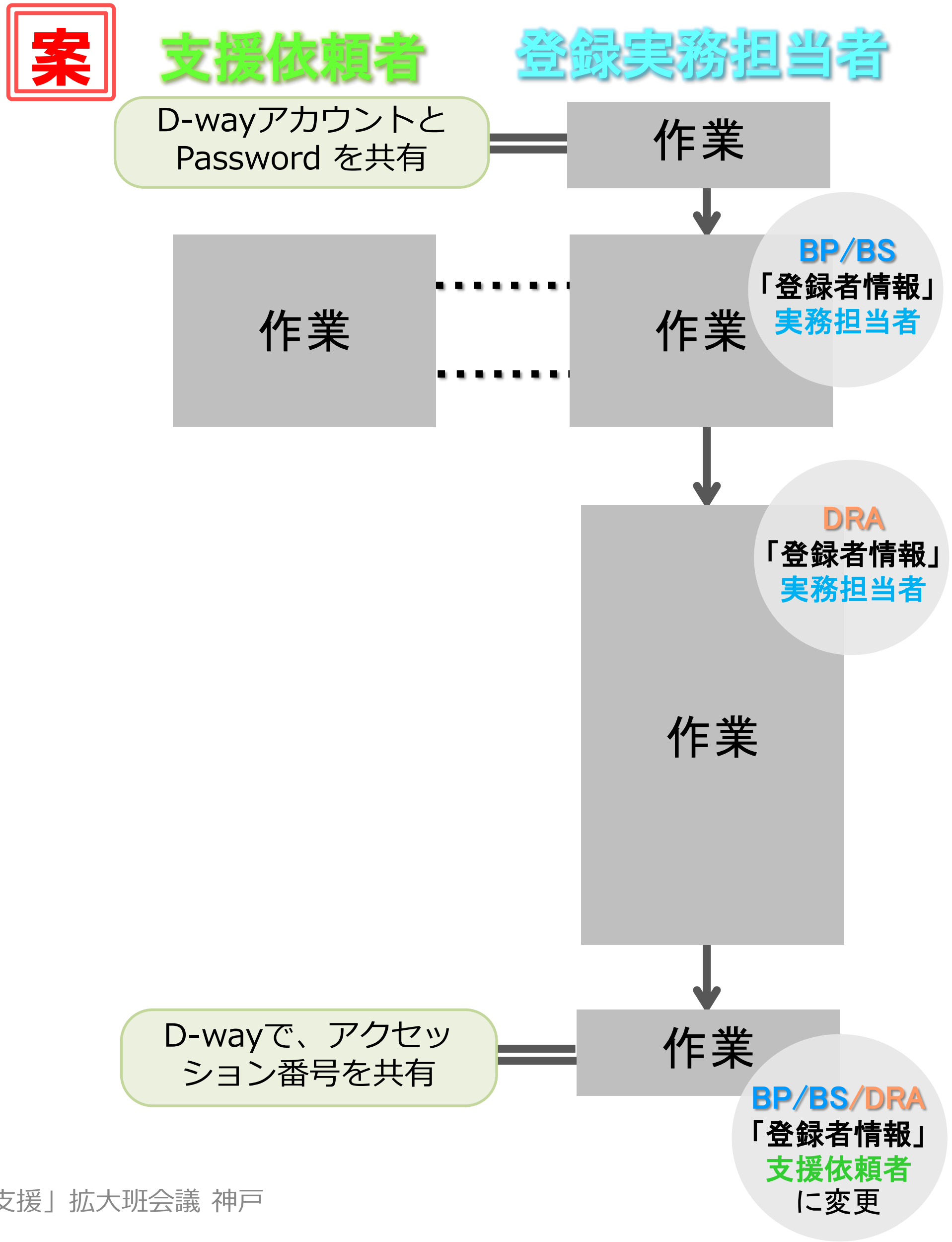
FAQs

Handbooks

クリックすると、
動画や画像が
拡大表示

DRA登録の流れ

1. 新規 D-way アカウントの作成
↓
2. BioProject の登録
↓
3. BioSample の登録
↓
4. 新規 DRA Submission の作成
↓
5. データファイルの転送
↓
6. メタデータの作成
↓
7. Validation 成功
↓
8. アクセッション番号の受領



案

支援依頼者

登録実務担当者

D-wayアカウントと Password を共有

作業

作業

作業

BP/BS 「登録者情報」 実務担当者

作業

DRA 「登録者情報」 実務担当者

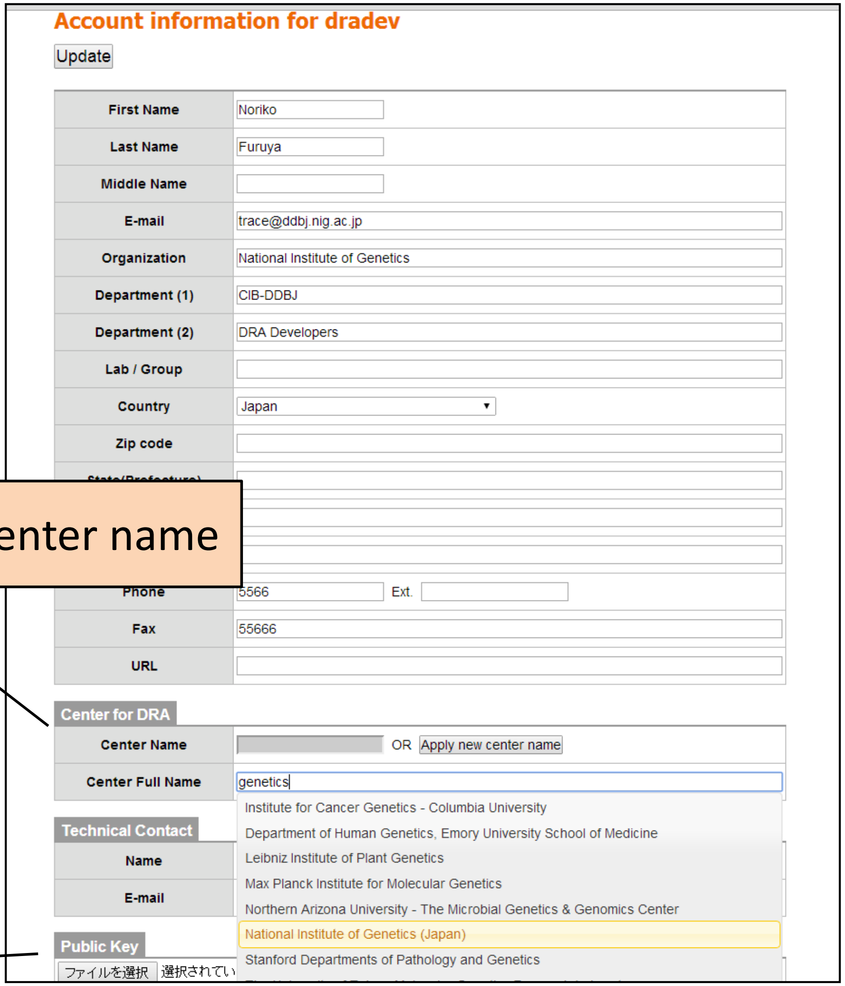
D-wayで、アクセッション番号を共有

作業

BP/BS/DRA 「登録者情報」 支援依頼者 に変更

1. 新規 D-way アカウントの作成 (1)

登録アカウントHandbook: http://trace.ddbj.nig.ac.jp/book/account.html
D-way: https://trace.ddbj.nig.ac.jp/D-way/



Account information for dradev

First Name	Noriko
Last Name	Furuya
Middle Name	
E-mail	trace@ddbj.nig.ac.jp
Organization	National Institute of Genetics
Department (1)	CIB-DDBJ
Department (2)	DRA Developers
Lab / Group	
Country	Japan
Zip code	
Phone	5566 Ext.
Fax	55666
URL	

Center for DRA

Center Name	OR Apply new center name
Center Full Name	genetics

Institute for Cancer Genetics - Columbia University
Department of Human Genetics, Emory University School of Medicine
Leibniz Institute of Plant Genetics
Max Planck Institute for Molecular Genetics
Northern Arizona University - The Microbial Genetics & Genomics Center
National Institute of Genetics (Japan)
Stanford Departments of Pathology and Genetics

Technical Contact

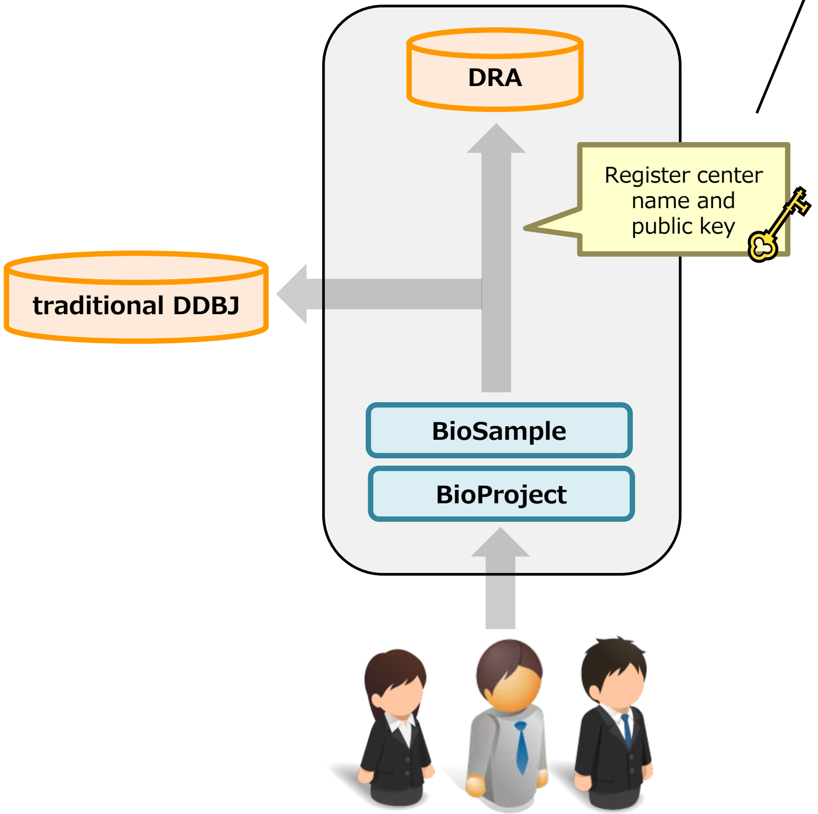
Name	
E-mail	

Public Key

Center name

公開鍵の登録

DRA に登録するためには、”Center name” と ”公開鍵”の登録が必要



Submission Portal “D-way”

1. 新規 D-way アカウントの作成 (2)

公開鍵/秘密鍵: http://trace.ddbj.nig.ac.jp/book/account.html#DRA_への登録権限を追加

DRA への登録権限を追加

目次:

DDBJ Sequence Read Archive (DRA) へデータを登録するため，アカウントに center name と公開鍵を登録します。 Center name は DRA が組織に対して割り振っている略号です。 公開鍵はユーザの認証に用いられます。

公開鍵/秘密鍵ペアの生成

⊟ **Windows**

Putty Key Generator を起動します

"PuTTY Key Generator" をインストールし，起動します。 下のように設定し [Generate] をクリックします。

PuTTY Key Generator

File Key Conversions Help

Key

No key.

Actions

Generate a public/private key pair [Generate]

Load an existing private key file [Load]

Save the generated key [Save public key] [Save private key]

Parameters

Type of key to generate:

○ SSH-1 (RSA) ● SSH-2 RSA ○ SSH-2 DSA

Number of bits in a generated key: 1024

ウィンドウのなかでマウスポインターをランダムに動かして鍵を生成させます。生成される公開鍵と秘密鍵を保存します。

« 閉じる

⊞ **Mac OS X**

⊞ **Unix**

Center name と公開鍵の登録

In this page

- DDBJ 登録アカウント
- ウェブから取得した登録用アカウント
- DRA への登録権限を追加
- 秘密鍵ファイルの変換

PDF Download

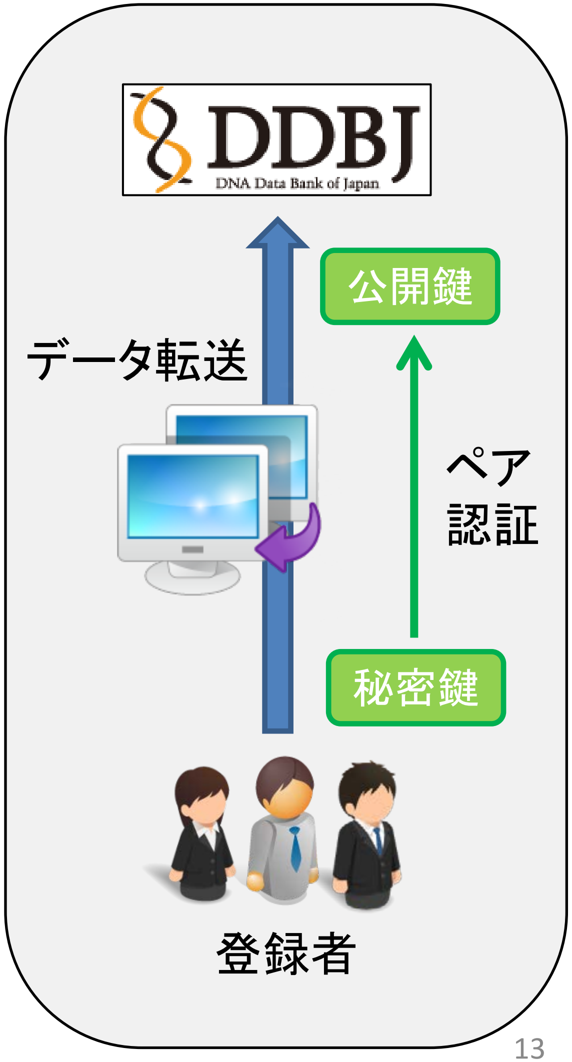
- PDF をダウンロード

Submit

- Login & Submit
- Contact

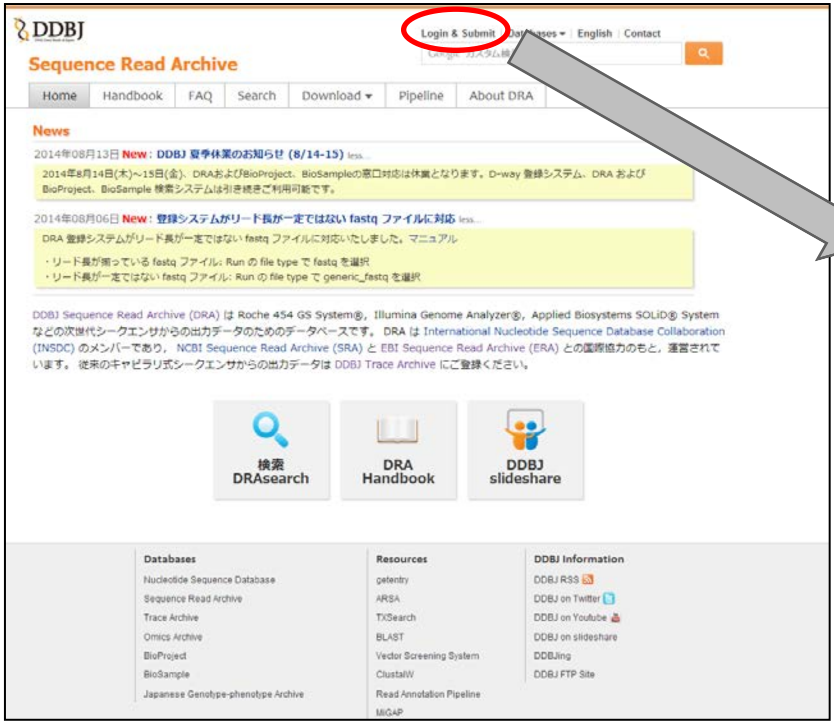
Archives

- News by year: Latest ▼
- FAQs
- Handbooks



2. BioProject の登録 (1)

BioProject Handbook: http://trace.ddbj.nig.ac.jp/bioproject/submission.html
D-way: https://trace.ddbj.nig.ac.jp/D-way/



DDBJ

Login & Submit | Databases | English | Contact

Sequence Read Archive

Home | Handbook | FAQ | Search | Download | Pipeline | About DRA

News

2014年08月13日 **New**：**DDBJ 夏季休業のお知らせ (8/14-15)** less...

2014年8月14日(木)～15日(金)、DRAおよびBioProject、BioSampleの窓口対応は休業となります。D-way 登録システム、DRA および BioProject、BioSample 検索システムは引き続きご利用可能です。

2014年08月06日 **New**：**登録システムがリード長が一定ではない fastq ファイルに対応** less...

DRA 登録システムがリード長が一定ではない fastq ファイルに対応いたしました。マニュアル

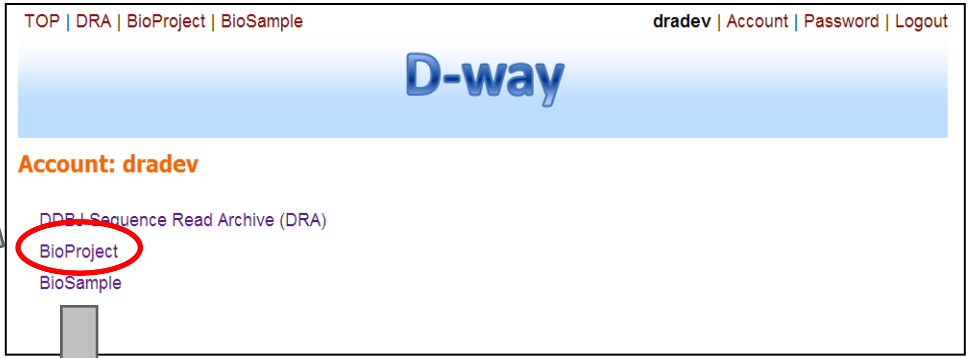
・リード長が揃っている fastq ファイル: Run の file type で fastq を選択
・リード長が一定ではない fastq ファイル: Run の file type で generic_fastq を選択

DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System などの次世代シークエンサからの出力データのためのデータベースです。 DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり， NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと，運営されています。 従来のキャピラリ式シークエンサからの出力データは DDBJ Trace Archive にご登録ください。

Databases
Nucleotide Sequence Database
Sequence Read Archive
Trace Archive
Omics Archive
BioProject
BioSample
Japanese Genotype-phenotype Archive

Resources
getentry
ARSA
TXSearch
BLAST
Vector Screening System
ClustalW
Read Annotation Pipeline
MiGAP

DDBJ Information
DDBJ RSS
DDBJ on Twitter
DDBJ on Youtube
DDBJ on slideshare
DDBJing
DDBJ FTP Site

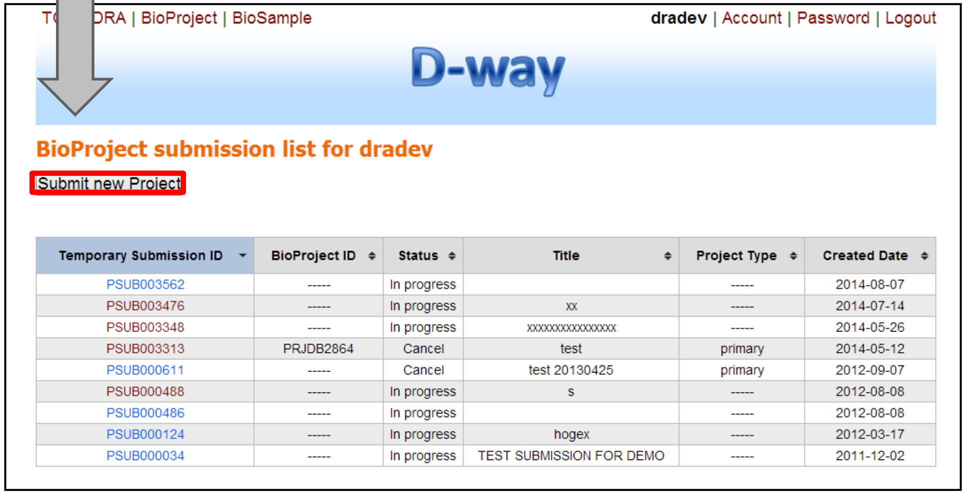


TOP | DRA | BioProject | BioSample

dradev | Account | Password | Logout

Account: dradev

DDBJ Sequence Read Archive (DRA)
BioProject
BioSample



T DRA | BioProject | BioSample

dradev | Account | Password | Logout

BioProject submission list for dradev

Submit new Project

Temporary Submission ID	BioProject ID	Status	Title	Project Type	Created Date
PSUB003562	-----	In progress		-----	2014-08-07
PSUB003476	-----	In progress	xx	-----	2014-07-14
PSUB003348	-----	In progress	xxxxxxxxxxxxxxxx	-----	2014-05-26
PSUB003313	PRJDB2864	Cancel	test	primary	2014-05-12
PSUB000611	-----	Cancel	test 20130425	primary	2012-09-07
PSUB000488	-----	In progress	s	-----	2012-08-08
PSUB000486	-----	In progress		-----	2012-08-08
PSUB000124	-----	In progress	hogex	-----	2012-03-17
PSUB000034	-----	In progress	TEST SUBMISSION FOR DEMO	-----	2011-12-02

[Submit new Project] ボタンから、
新規作成

2. BioProject の登録 (2)

Submitter > General info > Project type > Target > Publication > Overview の順。
赤色のアスタリスク記号 (*)は、必須項目。英語入力。

BioProject Temporary Submission ID: PSUB000488

Project Description

Project title*: Retinal transcriptome profiling at transcription start sites after axonal injury

Description*: In order to comprehensively study changes in gene transcription at transcription start sites. Using CAGE data, we attempted to determine dynamic changes in the regulation of the transcriptional network mediating RGC death after axonal injury.

rovide enough information in the description (more than 100 characters) for other users to interpret the data.

nts to DDBJ staff:

ject

nitiative which is already registered in the BioProject. ○Yes ●No

nks

ion URL

Grants

Agency	Agency abbreviation	Grant ID	Grant title
Japanese Ministry of Educa	MEXT	221S0002	Genome Science

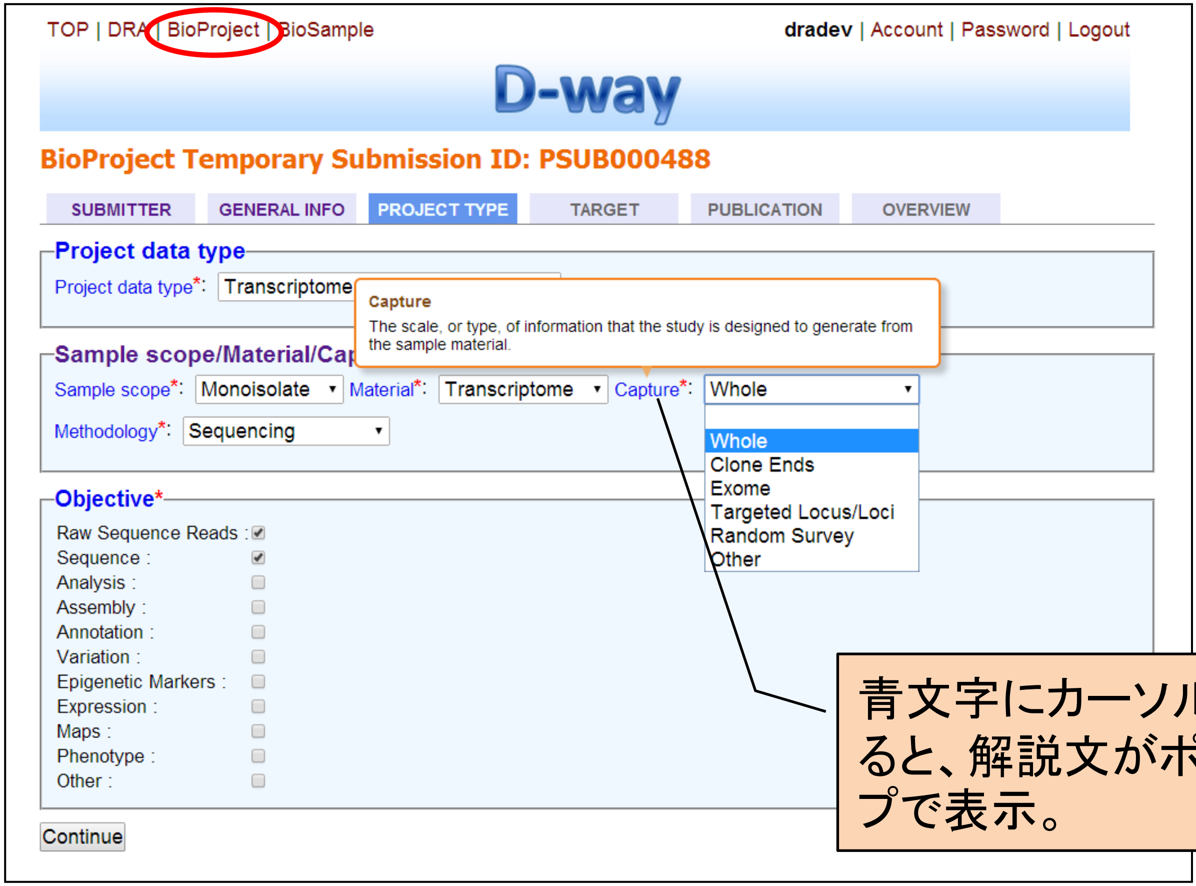
Consortium

Grant

Agency : Japanese Ministry of Education, Culture, Sports, Science and Technology
Agency abbreviation : MEXT
Grant ID : 221S0002
Grant title: Genome Science

Description:
研究の目的やゴールについて、100文字以上、記述

2. BioProject の登録 (3)



BioProject Temporary Submission ID: PSUB000488

Project data type

Project data type*: Transcriptome

Capture

The scale, or type, of information that the study is designed to generate from the sample material.

Sample scope/Material/Cap

Sample scope*: Monoisolate
Material*: Transcriptome
Capture*: Whole

Methodology*: Sequencing

Objective*

Raw Sequence Reads : ☑
Sequence : ☑
Analysis : ☐
Assembly : ☐
Annotation : ☐
Variation : ☐
Epigenetic Markers : ☐
Expression : ☐
Maps : ☐
Phenotype : ☐
Other : ☐

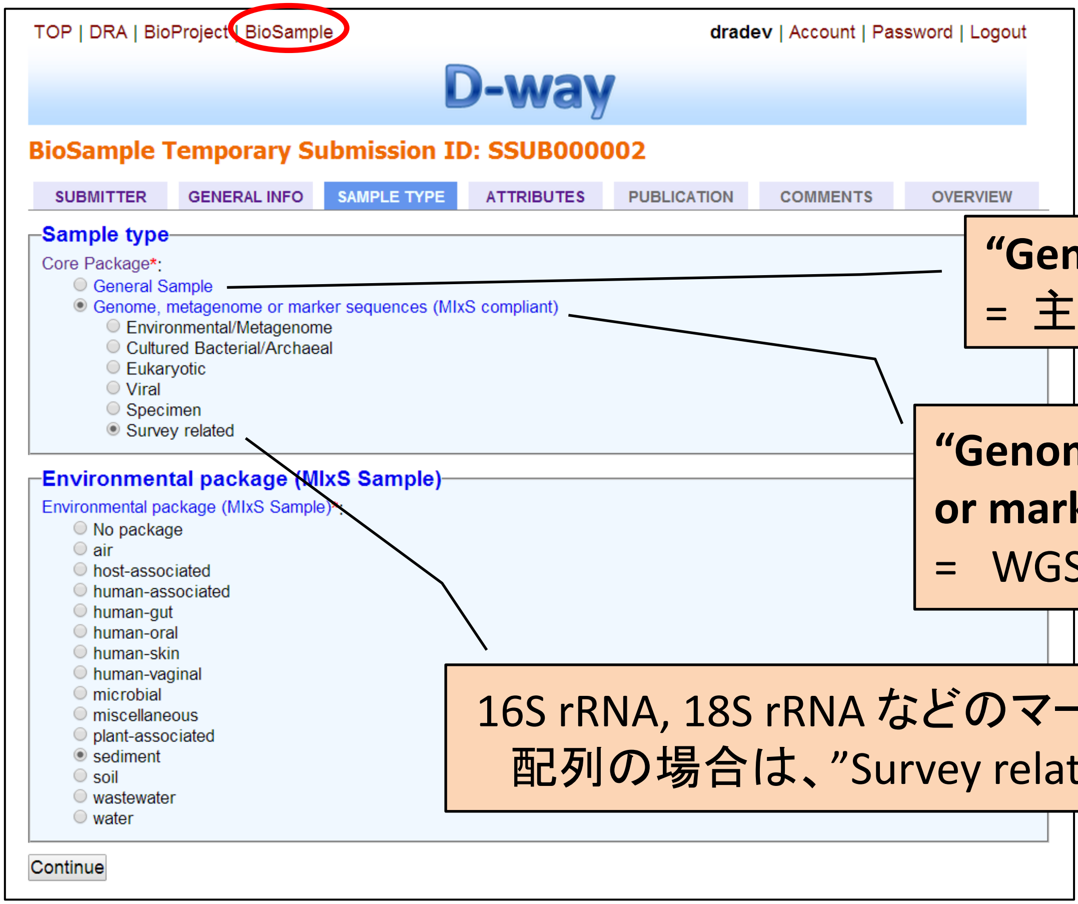
青文字にカーソルを当てると、解説文がポップアップで表示。

最後に [Submit]ボタンをクリックした後、D-way上で修正は出来ません。アノテータにメール (bioproject@ddbj.nig.ac.jp) で知らせてください。
アノテータが査定を行ってから、BioProject IDを発行するので、少しお時間がかかります。

3. BioSample の登録 (1)

BioSample Handbook: http://trace.ddbj.nig.ac.jp/biosample/submission.html
D-way: https://trace.ddbj.nig.ac.jp/D-way/

[Submit new Sample] ボタンから、新規作成。
左から右に向かって、順にタブを移動。



BioSample Temporary Submission ID: SSUB000002

Sample type

Core Package*:

- ○ General Sample
- ● Genome, metagenome or marker sequences (MIxS compliant)
  - ○ Environmental/Metagenome
  - ○ Cultured Bacterial/Archaeal
  - ○ Eukaryotic
  - ○ Viral
  - ○ Specimen
  - ● Survey related

Environmental package (MIxS Sample)

Environmental package (MIxS Sample)*:

- ○ No package
- ○ air
- ○ host-associated
- ○ human-associated
- ○ human-gut
- ○ human-oral
- ○ human-skin
- ○ human-vaginal
- ○ microbial
- ○ miscellaneous
- ○ plant-associated
- ● sediment
- ○ soil
- ○ wastewater
- ○ water

"General Sample"
= 主に、transcriptome

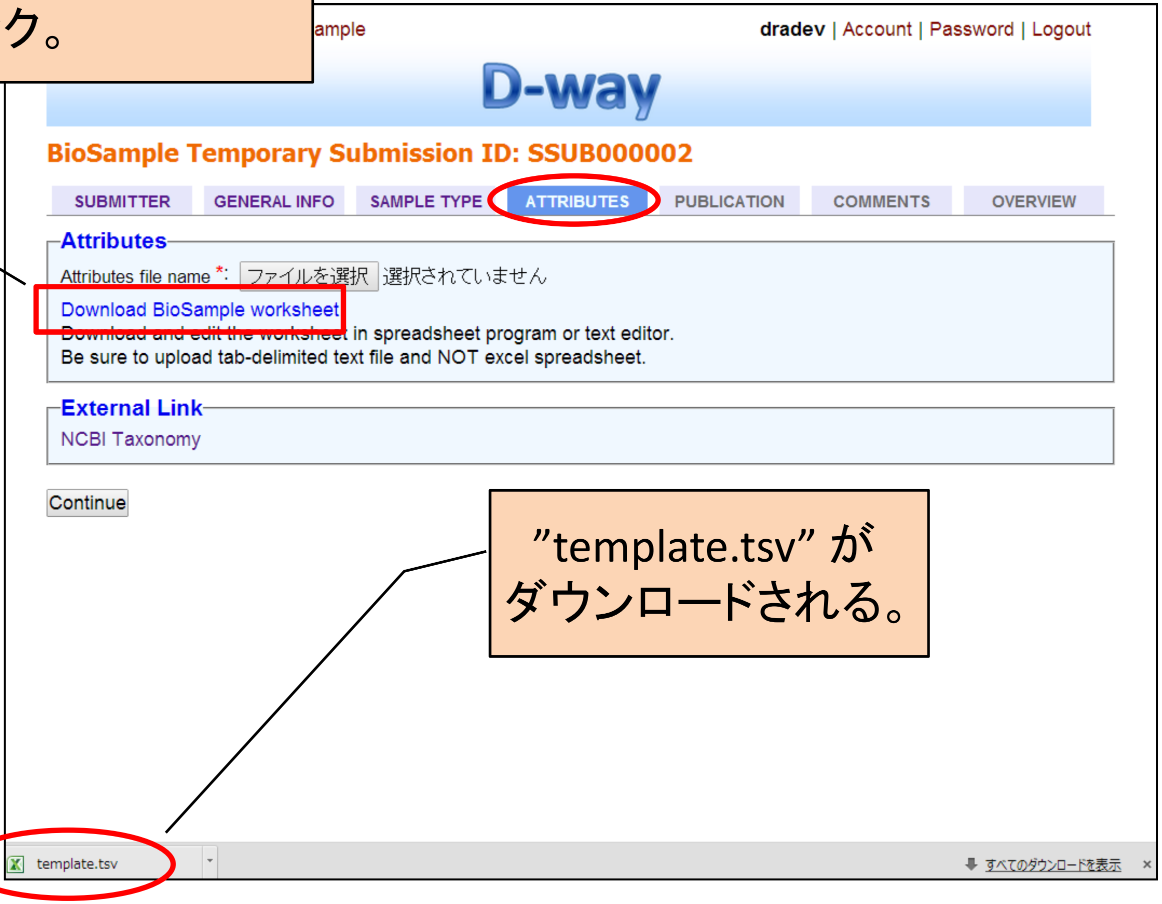
"Genome, metagenome or marker sequences"
= WGSやメタゲノムなど

16S rRNA, 18S rRNA などのマーカー配列の場合は、"Survey related"

3. BioSample の登録 (2)

BioSample Handbook: http://trace.ddbj.nig.ac.jp/biosample/submission.html
D-way: https://trace.ddbj.nig.ac.jp/D-way/

[Download BioSample worksheet] をクリック。



BioSample Temporary Submission ID: SSUB000002

Attributes

Attributes file name *: ファイルを選択 選択されていません

Download BioSample worksheet

Download and edit the worksheet in spreadsheet program or text editor.
Be sure to upload tab-delimited text file and NOT excel spreadsheet.

External Link

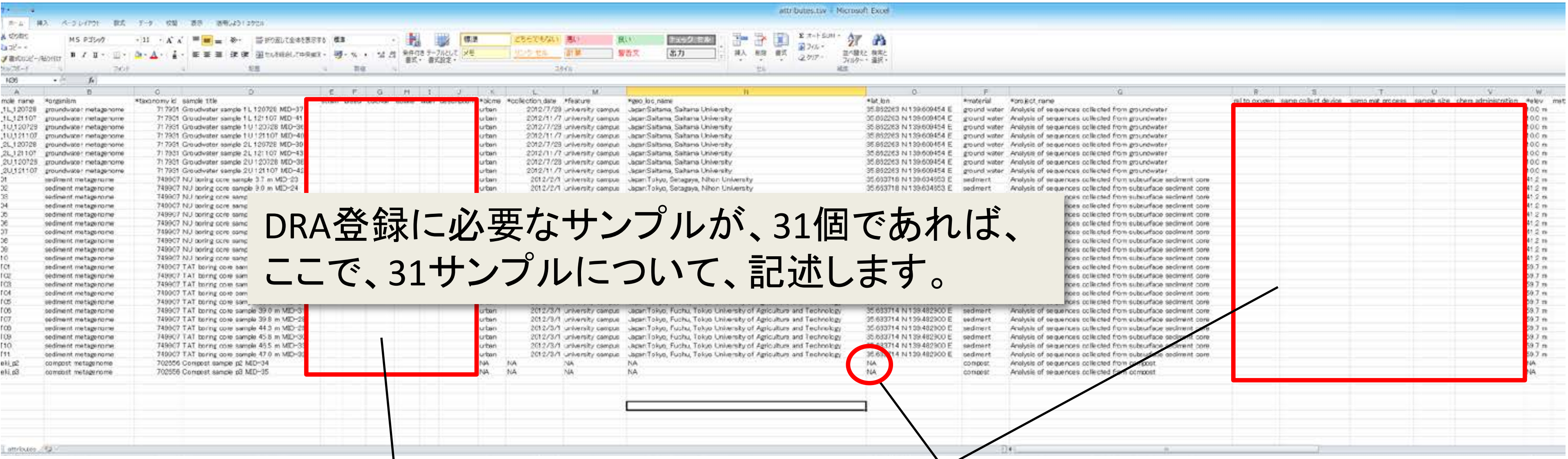
NCBI Taxonomy

”template.tsv” がダウンロードされる。

3. BioSample の登録 (3)

BioSample Handbook: http://trace.ddbj.nig.ac.jp/biosample/submission.html
D-way: https://trace.ddbj.nig.ac.jp/D-way/

TSV（タブ区切り）ファイルは、Excelなどの表計算ソフトで開くと、編集が便利です。
アスタリスク記号 (*)は、必須項目。

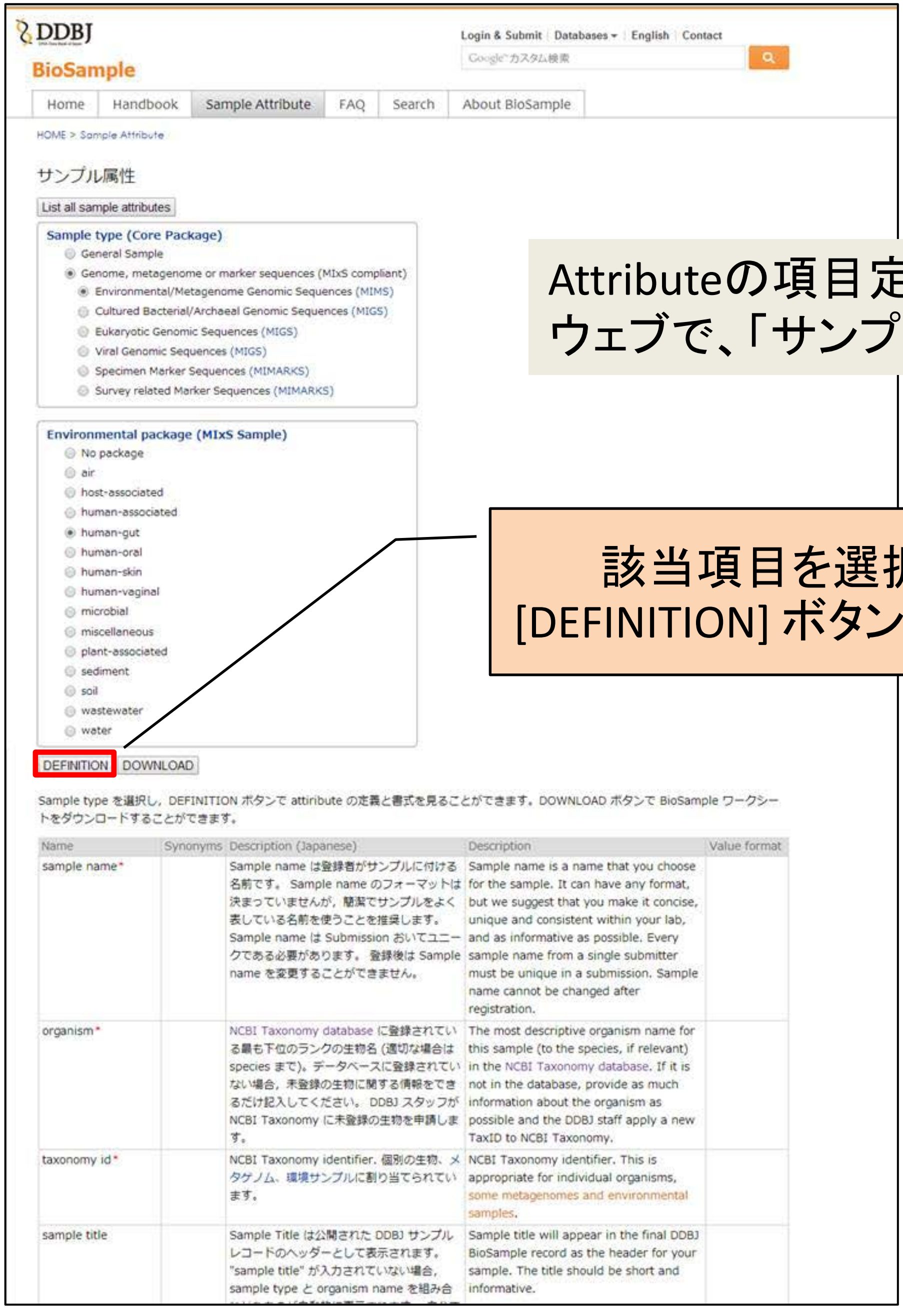


DRA登録に必要なサンプルが、31個であれば、
ここで、31サンプルについて、記述します。

必須でない空欄項目列も、
削除せずに残す

必須項目で該当しない場
合は、"NA"などを記入

BioSample --- Sample Attribute: http://trace.ddbj.nig.ac.jp/biosample/attribute.html



サンプル属性

List all sample attributes

Sample type (Core Package)

- ○ General Sample
- ● Genome, metagenome or marker sequences (MIxS compliant)
  - ● Environmental/Metagenome Genomic Sequences (MIMS)
  - ○ Cultured Bacterial/Archaeal Genomic Sequences (MIGS)
  - ○ Eukaryotic Genomic Sequences (MIGS)
  - ○ Viral Genomic Sequences (MIGS)
  - ○ Specimen Marker Sequences (MIMARKS)
  - ○ Survey related Marker Sequences (MIMARKS)

Environmental package (MIxS Sample)

- ○ No package
- ○ air
- ○ host-associated
- ○ human-associated
- ● human-gut
- ○ human-oral
- ○ human-skin
- ○ human-vaginal
- ○ microbial
- ○ miscellaneous
- ○ plant-associated
- ○ sediment
- ○ soil
- ○ wastewater
- ○ water

DEFINITION　DOWNLOAD

Sample type を選択し，DEFINITION ボタンで attiribute の定義と書式を見ることができます。DOWNLOAD ボタンで BioSample ワークシートをダウンロードすることができます。

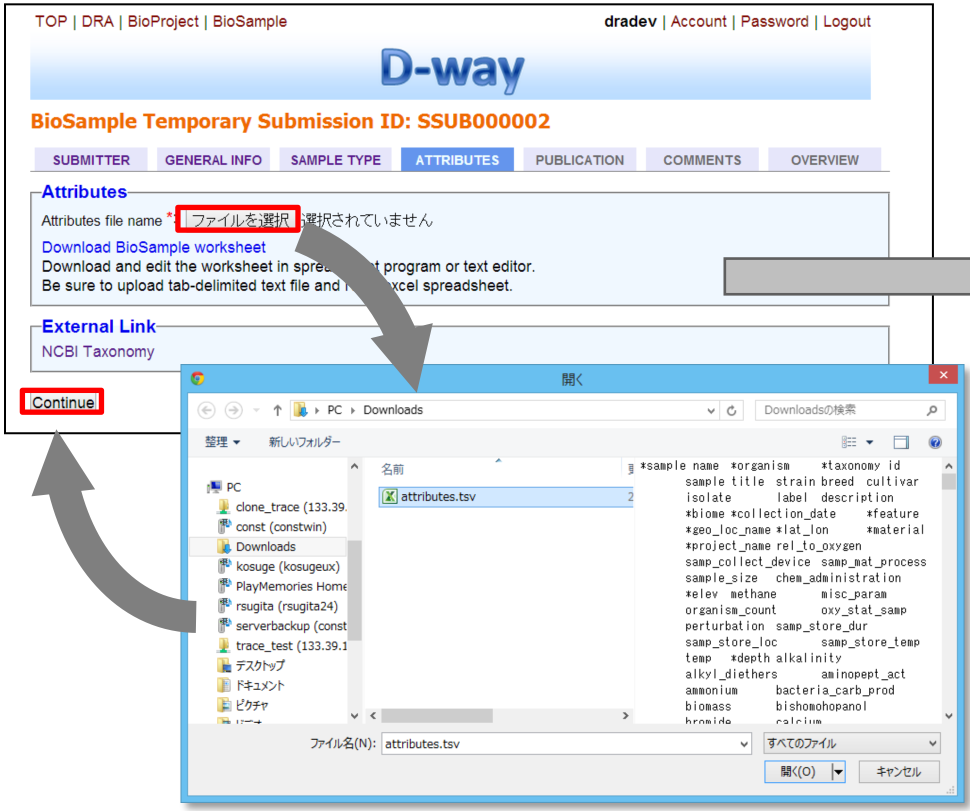
Name	Synonyms	Description (Japanese)	Description	Value format
sample name*		Sample name は登録者がサンプルに付ける名前です。 Sample name のフォーマットは決まっていませんが，簡潔でサンプルをよく表している名前を使うことを推奨します。 Sample name は Submission おいてユニークである必要があります。 登録後は Sample name を変更することができません。	Sample name is a name that you choose for the sample. It can have any format, but we suggest that you make it concise, unique and consistent within your lab, and as informative as possible. Every sample name from a single submitter must be unique in a submission. Sample name cannot be changed after registration.	
organism*		NCBI Taxonomy database に登録されている最も下位のランクの生物名 (適切な場合は species まで)。データベースに登録されていない場合，未登録の生物に関する情報をできるだけ記入してください。 DDBJ スタッフが NCBI Taxonomy に未登録の生物を申請します。	The most descriptive organism name for this sample (to the species, if relevant) in the NCBI Taxonomy database. If it is not in the database, provide as much information about the organism as possible and the DDBJ staff apply a new TaxID to NCBI Taxonomy.	
taxonomy id*		NCBI Taxonomy identifier. 個別の生物、メタゲノム、環境サンプルに割り当てられています。	NCBI Taxonomy identifier. This is appropriate for individual organisms, some metagenomes and environmental samples.	
sample title		Sample Title は公開された DDBJ サンプルレコードのヘッダーとして表示されます。 "sample title" が入力されていない場合，sample type と organism name を組み合	Sample title will appear in the final DDBJ BioSample record as the header for your sample. The title should be short and informative.	

Attributeの項目定義が分からない時は、ウェブで、「サンプル属性」を参照。

該当項目を選択後、
[DEFINITION] ボタンをクリック

サンプル条件ごとに、
異なる項目が表示

3. BioSample の登録 (4)



BioSample Temporary Submission ID: SSUB000002

Attributes

Attributes file name * ファイルを選択 選択されていません

Download BioSample worksheet

Download and edit the worksheet in spre… program or text editor.
Be sure to upload tab-delimited text file and … xcel spreadsheet.

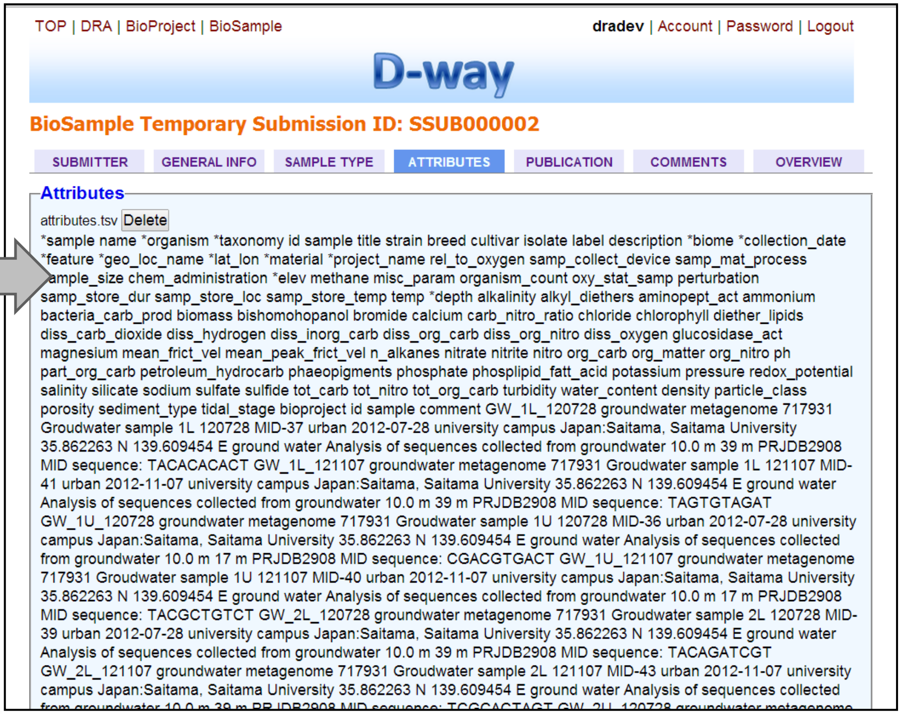
External Link

NCBI Taxonomy

attributes.tsv

ファイル名(N): attributes.tsv

```
*sample name  *organism      *taxonomy id
      sample title  strain breed  cultivar
      isolate       label  description
      *biome *collection_date     *feature
      *geo_loc_name *lat_lon      *material
      *project_name rel_to_oxygen
      samp_collect_device  samp_mat_process
      sample_size   chem_administration
      *elev  methane       misc_param
      organism_count       oxy_stat_samp
      perturbation  samp_store_dur
      samp_store_loc       samp_store_temp
      temp   *depth alkalinity
      alkyl_diethers       aminopept_act
      ammonium      bacteria_carb_prod
      biomass       bishomohopanol
      bromide       calcium
```



BioSample Temporary Submission ID: SSUB000002

Attributes

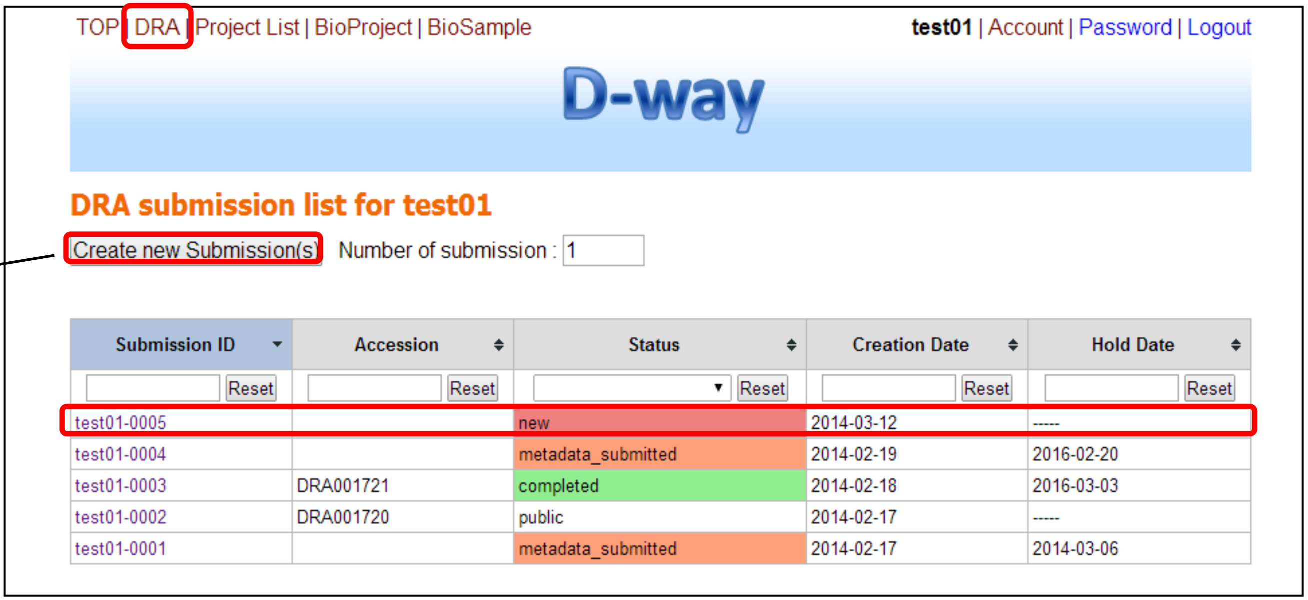
attributes.tsv Delete

*sample name *organism *taxonomy id sample title strain breed cultivar isolate label description *biome *collection_date *feature *geo_loc_name *lat_lon *material *project_name rel_to_oxygen samp_collect_device samp_mat_process sample_size chem_administration *elev methane misc_param organism_count oxy_stat_samp perturbation samp_store_dur samp_store_loc samp_store_temp temp *depth alkalinity alkyl_diethers aminopept_act ammonium bacteria_carb_prod biomass bishomohopanol bromide calcium carb_nitro_ratio chloride chlorophyll diether_lipids diss_carb_dioxide diss_hydrogen diss_inorg_carb diss_org_carb diss_org_nitro diss_oxygen glucosidase_act magnesium mean_frict_vel mean_peak_frict_vel n_alkanes nitrate nitrite nitro org_carb org_matter org_nitro ph part_org_carb petroleum_hydrocarb phaeopigments phosphate phosplipid_fatt_acid potassium pressure redox_potential salinity silicate sodium sulfate sulfide tot_carb tot_nitro tot_org_carb turbidity water_content density particle_class porosity sediment_type tidal_stage bioproject id sample comment GW_1L_120728 groundwater metagenome 717931 Groudwater sample 1L 120728 MID-37 urban 2012-07-28 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 39 m PRJDB2908 MID sequence: TACACACACT GW_1L_121107 groundwater metagenome 717931 Groudwater sample 1L 121107 MID-41 urban 2012-11-07 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 39 m PRJDB2908 MID sequence: TAGTGTAGAT GW_1U_120728 groundwater metagenome 717931 Groudwater sample 1U 120728 MID-36 urban 2012-07-28 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 17 m PRJDB2908 MID sequence: CGACGTGACT GW_1U_121107 groundwater metagenome 717931 Groudwater sample 1U 121107 MID-40 urban 2012-11-07 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 17 m PRJDB2908 MID sequence: TACGCTGTCT GW_2L_120728 groundwater metagenome 717931 Groudwater sample 2L 120728 MID-39 urban 2012-07-28 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 39 m PRJDB2908 MID sequence: TACAGATCGT GW_2L_121107 groundwater metagenome 717931 Groudwater sample 2L 121107 MID-43 urban 2012-11-07 university campus Japan:Saitama, Saitama University 35.862263 N 139.609454 E ground water Analysis of sequences collected from groundwater 10.0 m 39 m PRJDB2908 MID sequence: TCGCACTAGT GW_2U_120728 groundwater metagenome

最後に [Submit]ボタンをクリックした後、D-way上で修正は出来ません。アノテータにメール (biosample@ddbj.nig.ac.jp) で知らせてください。
アノテータが査定を行ってから、BioSample ID を発行するので、少しお時間がかかります。

4. 新規 DRA Submission の作成

DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html#DRA_へのデータ登録方法



TOP | DRA | Project List | BioProject | BioSample

test01 | Account | Password | Logout

D-way

DRA submission list for test01

Create new Submission(s)　Number of submission : 1

Submission ID	Accession	Status	Creation Date	Hold Date
Reset	Reset	▼ Reset	Reset	Reset
test01-0005		new	2014-03-12	-----
test01-0004		metadata_submitted	2014-02-19	2016-02-20
test01-0003	DRA001721	completed	2014-02-18	2016-03-03
test01-0002	DRA001720	public	2014-02-17	-----
test01-0001		metadata_submitted	2014-02-17	2014-03-06

クリックして、
新規 DRA Submission
を作成

DRA 登録のステータス一覧

ステータス	状態
new	メタデータの投稿前
metadata_submitted	メタデータが投稿された
data_validating	データファイルの Validation 中
data_error	データファイルの Validation エラー
submission_validated	メタデータとデータファイルの Validation が完了
completed	アクセッション番号が発行された
confidential	非公開
public	公開

5. データファイルの転送

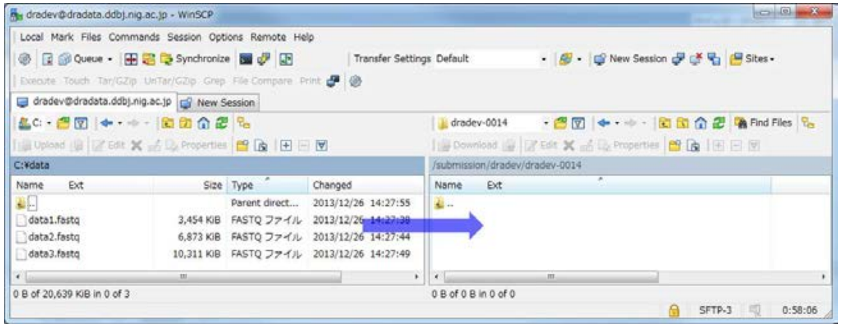
DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html#シークエンスデータのアップロード

SCP（Secure Copy）転送

sshの機能を使ってセキュリティの高いファイル転送を行う。
認証情報と、やり取りされるデータとの両方が、暗号化されてネットワーク上を流れる。

Windows

[WinSCP (http://winscp.net/eng/download.php)](http://winscp.net/eng/download.php)



dradev@dradata.ddbj.nig.ac.jp - WinSCP

C:¥data

Name	Ext	Size	Type	Changed
..			Parent direct...	2013/12/26 14:27:55
data1.fastq		3,454 KiB	FASTQ ファイル	2013/12/26 14:27:38
data2.fastq		6,873 KiB	FASTQ ファイル	2013/12/26 14:27:44
data3.fastq		10,311 KiB	FASTQ ファイル	2013/12/26 14:27:49

0 B of 20,639 KiB in 0 of 3

dradev-0014

/submission/dradev/dradev-0014

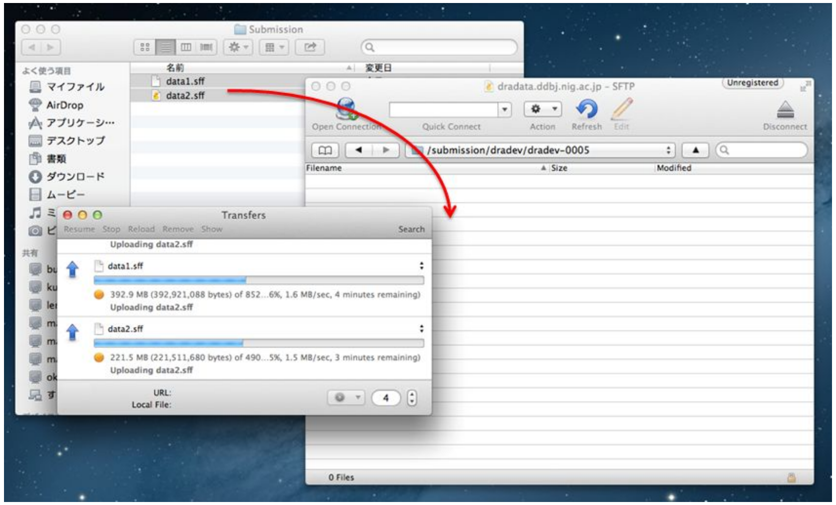
Name	Ext
..	

0 B of 0 B in 0 of 0

SFTP-3 0:58:06

Mac OS X

[Cyberduck (http://cyberduck.ch)](http://cyberduck.ch)



data1.sff
data2.sff

dradata.ddbj.nig.ac.jp - SFTP

/submission/dradev/dradev-0005

0 Files

Uploading data2.sff

data1.sff
392.9 MB (392,921,088 bytes) of 852…6%, 1.6 MB/sec, 4 minutes remaining)
Uploading data2.sff

data2.sff
221.5 MB (221,511,680 bytes) of 490…5%, 1.5 MB/sec, 3 minutes remaining)
Uploading data2.sff

Linux / Mac OS X

ターミナル

```
 $ scp <Your Files> <D-way Login ID>@dradata.ddbj.nig.ac.jp:~/<Submission ID>
        Enter passphrase for key '/home/you/.ssh/id_rsa':
$ ssh <D-way Login ID>@dradata.ddbj.nig.ac.jp
          :
          :
          :
```

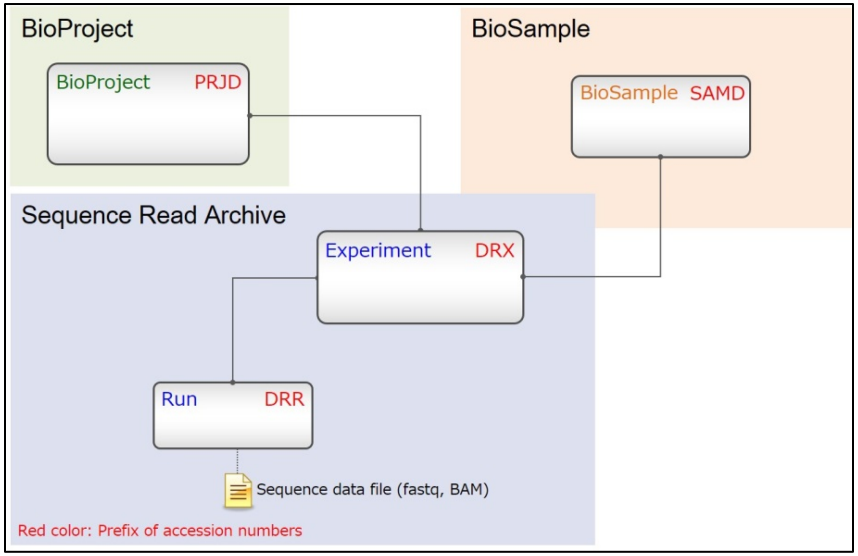
[新DRA登録システム（2014-05-12開始）](#)

technical reads を残したまま登録したい場合は、XMLの修正が必要になりました。(P. 8 参照)

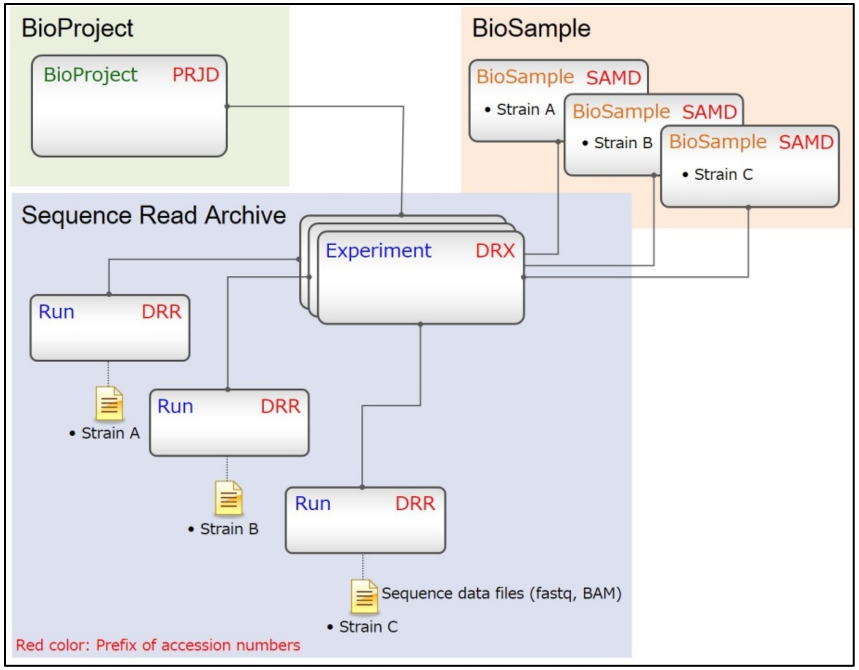
6. メタデータの作成 --- object の構成 ---

DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html#メタデータの作成

最もシンプルなケース

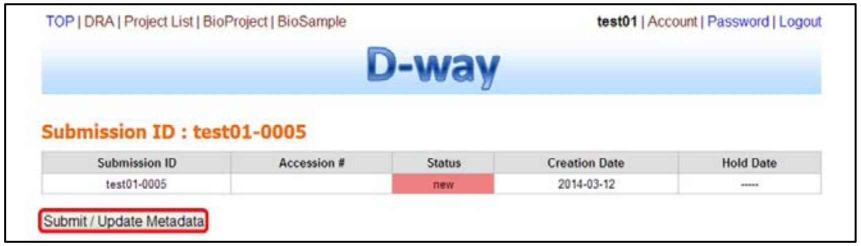


3つの菌株の比較ゲノム解析のケース



6. メタデータの作成　--- Submission ---

Submission > Study > Sample > Experiment > Run > (Analysis) の順に、英語で入力
赤色のアスタリスク記号 (*)は、必須項目

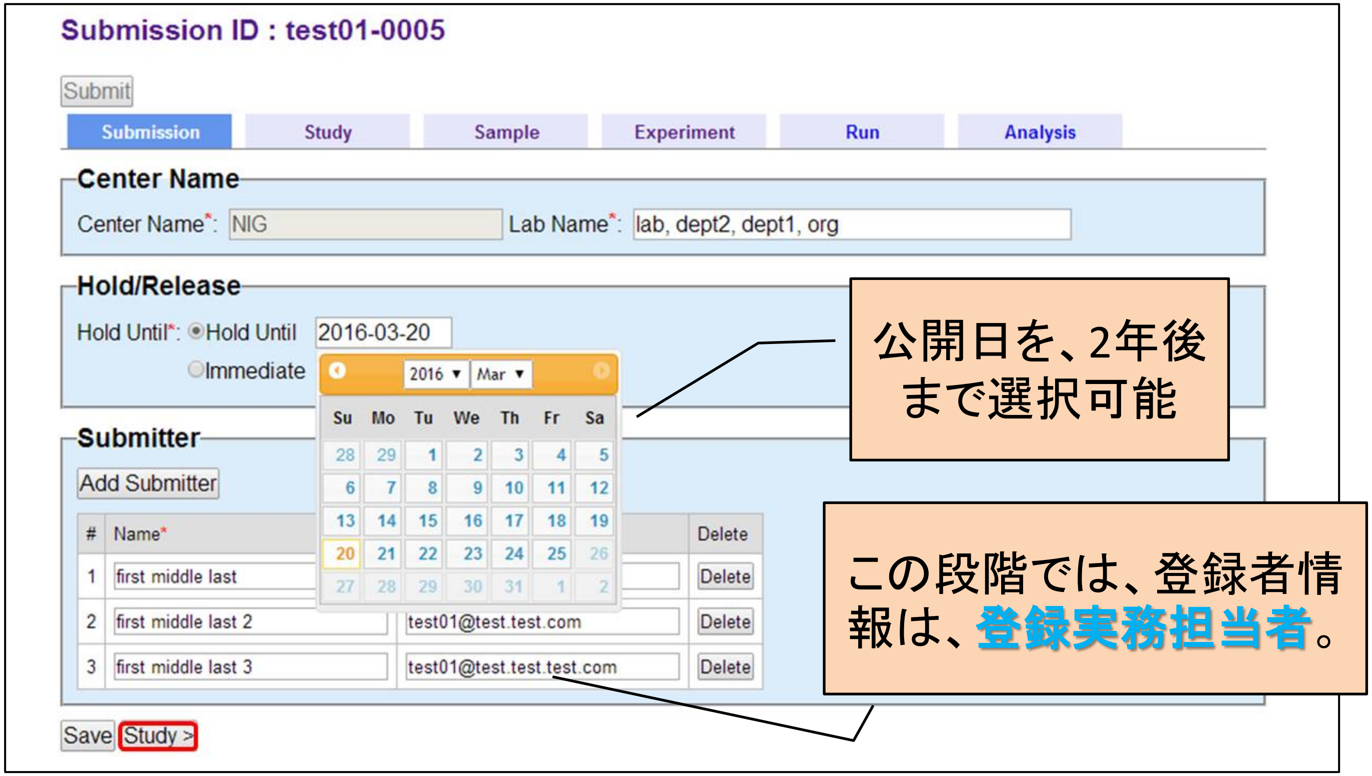


TOP | DRA | Project List | BioProject | BioSample　　test01 | Account | Password | Logout

D-way

Submission ID : test01-0005

Submission ID	Accession #	Status	Creation Date	Hold Date
test01-0005		new	2014-03-12	-----



Submission ID : test01-0005

Center Name

Center Name*: NIG　Lab Name*: lab, dept2, dept1, org

Hold/Release

Hold Until*: ● Hold Until 2016-03-20
○ Immediate

2016 ▾　Mar ▾

Su	Mo	Tu	We	Th	Fr	Sa
28	29	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2

Submitter

#	Name*		Delete
1	first middle last		Delete
2	first middle last 2	test01@test.test.com	Delete
3	first middle last 3	test01@test.test.test.com	Delete

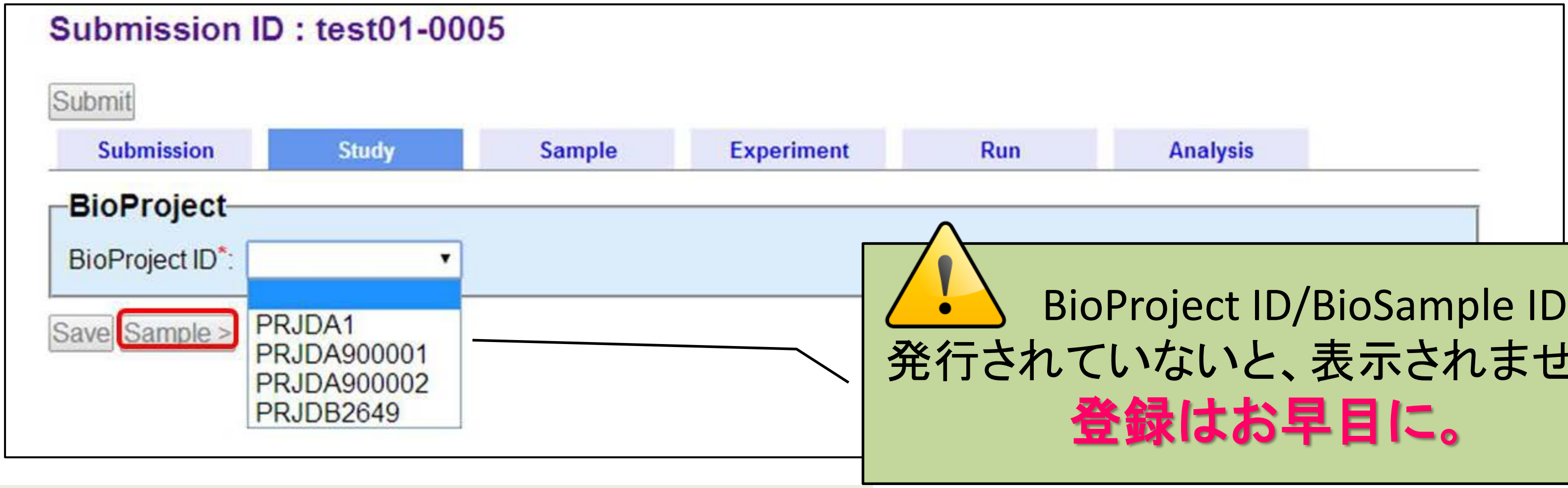
公開日を、2年後まで選択可能

この段階では、登録者情報は、**登録実務担当者**。

2014-08-20

6. メタデータの作成 --- Study/Sample ---

該当する BioProject ID を 1つ 選択する。



Submission ID : test01-0005

BioProject

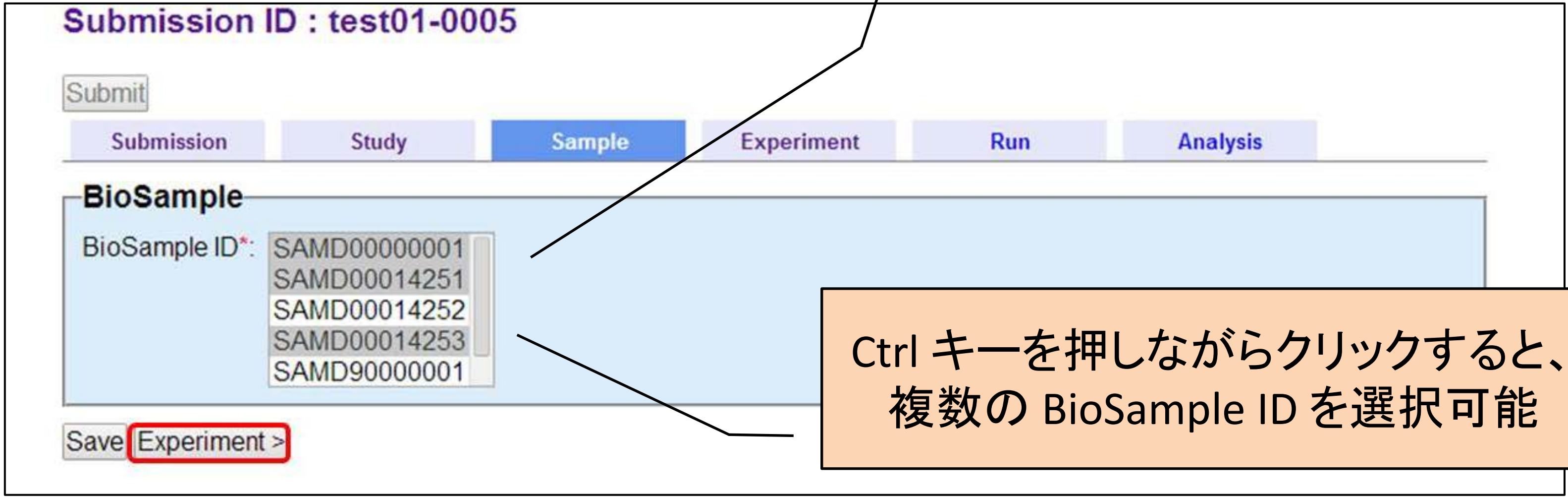
BioProject ID*:

PRJDA1
PRJDA900001
PRJDA900002
PRJDB2649

BioProject ID/BioSample ID が発行されていないと、表示されません。

登録はお早目に。

該当する BioSample ID を 1つ以上 選択する。



Submission ID : test01-0005

BioSample

BioSample ID*:

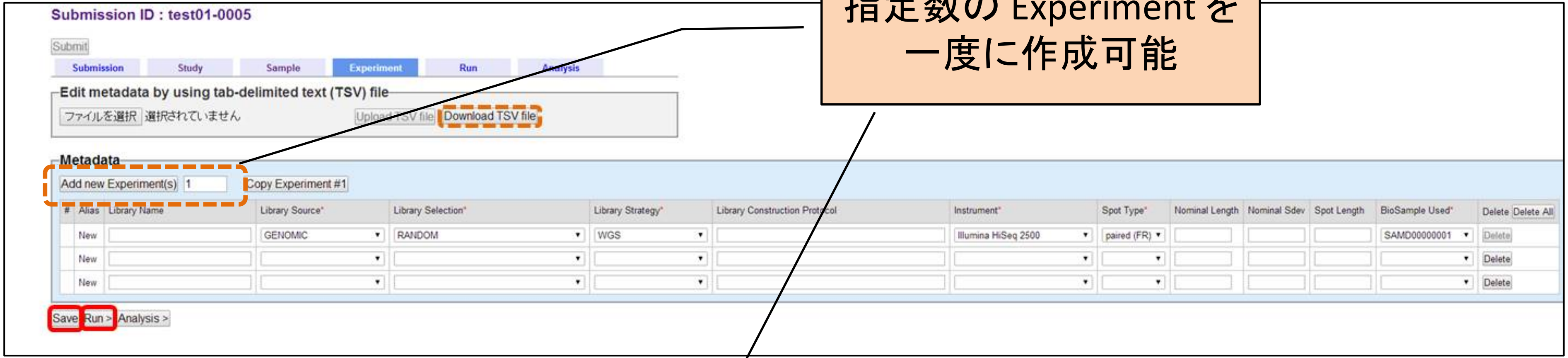
SAMD00000001
SAMD00014251
SAMD00014252
SAMD00014253
SAMD90000001

Ctrl キーを押しながらクリックすると、
複数の BioSample ID を選択可能

6. メタデータの作成　--- Experiment (1) ---

Experimentのメタデータを、順に設定する。
横長な画面に注意。

指定数の Experiment を
一度に作成可能



Submission ID : test01-0005

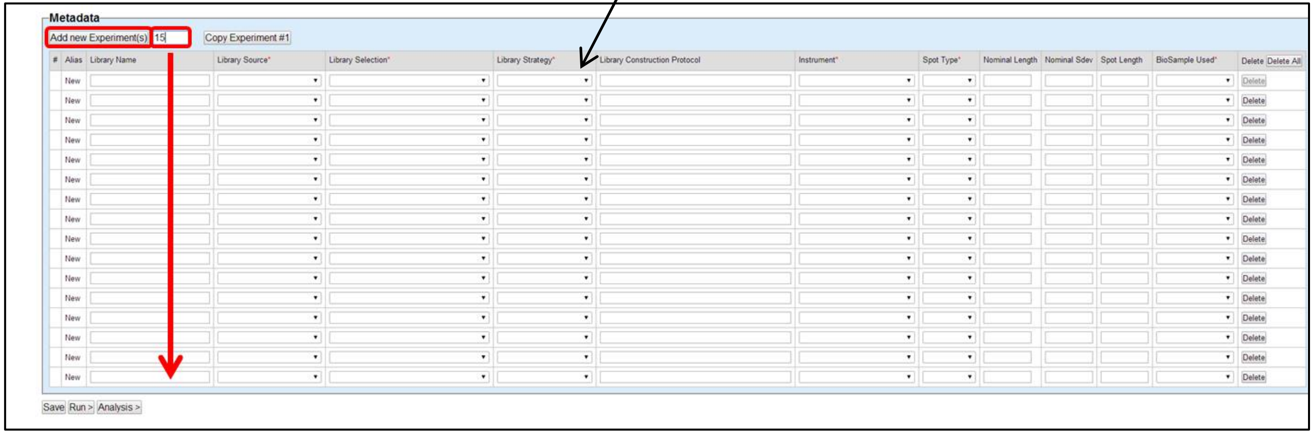
Edit metadata by using tab-delimited text (TSV) file

ファイルを選択　選択されていません

Metadata

Add new Experiment(s) 1

#	Alias	Library Name	Library Source*	Library Selection*	Library Strategy*	Library Construction Protocol	Instrument*	Spot Type*	Nominal Length	Nominal Sdev	Spot Length	BioSample Used*	Delete Delete All
	New		GENOMIC	RANDOM	WGS		Illumina HiSeq 2500	paired (FR)				SAMD00000001	Delete
	New												Delete
	New												Delete



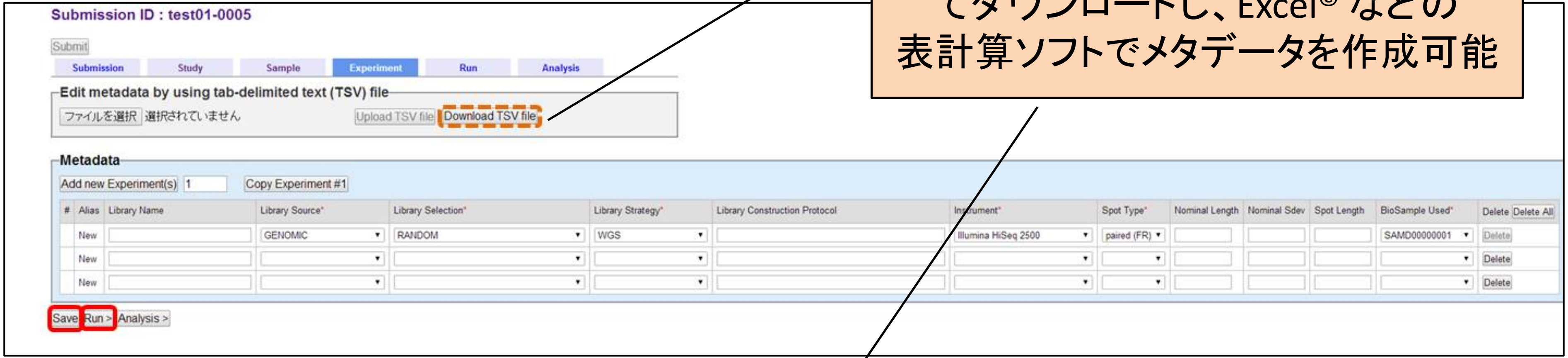
Metadata

Add new Experiment(s) 15

#	Alias	Library Name	Library Source*	Library Selection*	Library Strategy*	Library Construction Protocol	Instrument*	Spot Type*	Nominal Length	Nominal Sdev	Spot Length	BioSample Used*	Delete Delete All
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete
	New												Delete

6. メタデータの作成　--- Experiment (2) ---

タブ区切りテキストファイル (.tsv) としてダウンロードし、Excel© などの表計算ソフトでメタデータを作成可能

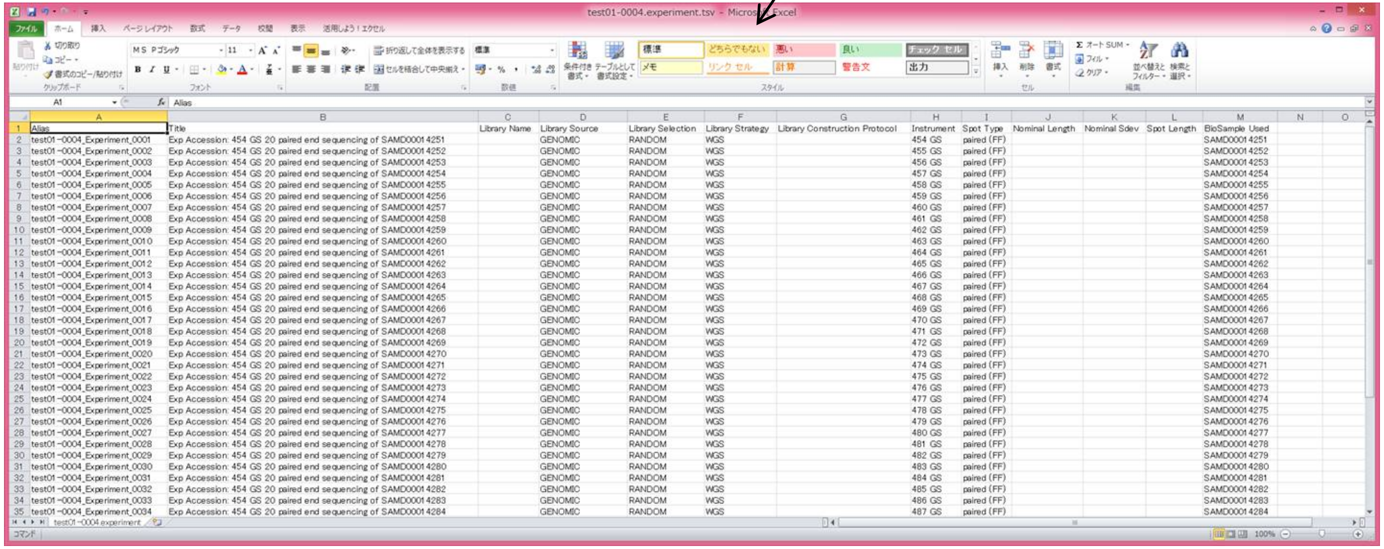


Submission ID : test01-0005

Edit metadata by using tab-delimited text (TSV) file

Metadata

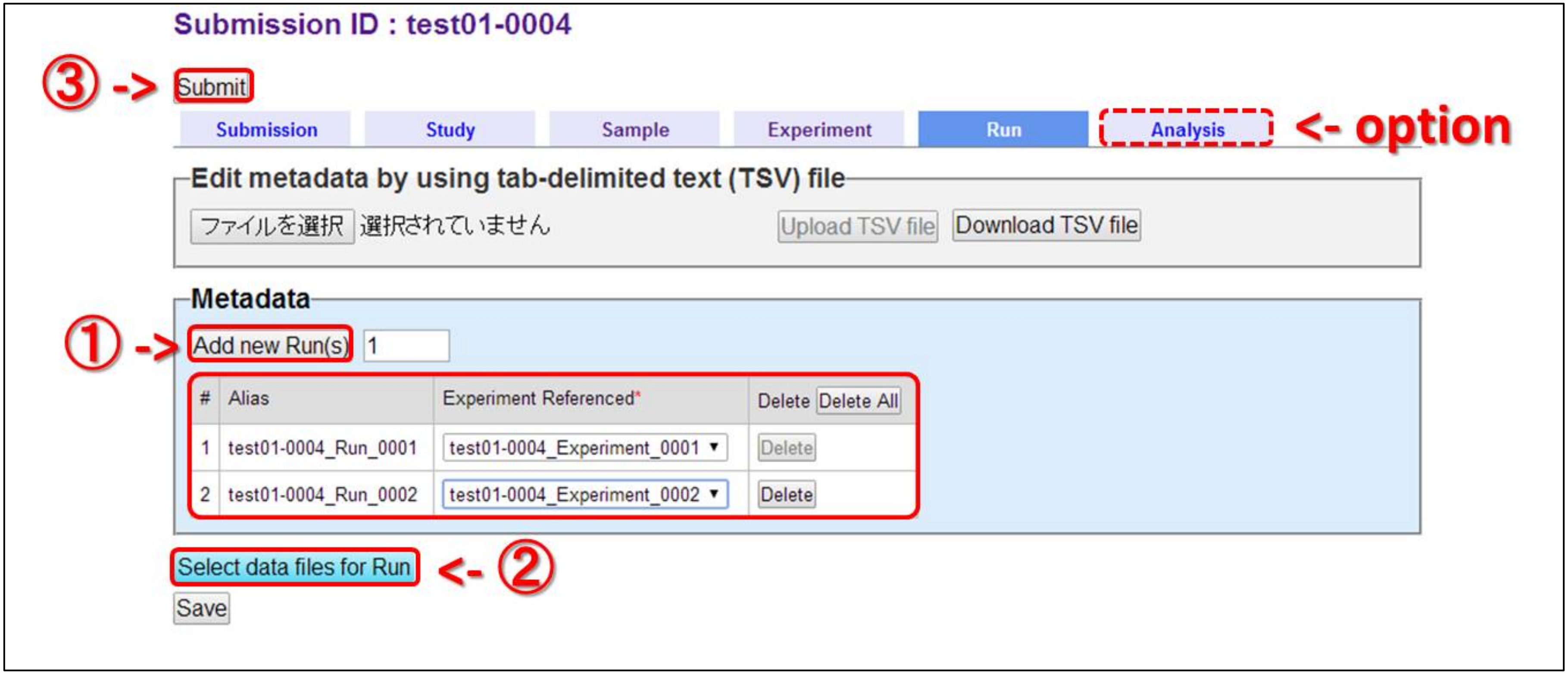
#	Alias	Library Name	Library Source*	Library Selection*	Library Strategy*	Library Construction Protocol	Instrument*	Spot Type*	Nominal Length	Nominal Sdev	Spot Length	BioSample Used*	Delete Delete All
	New		GENOMIC	RANDOM	WGS		Illumina HiSeq 2500	paired (FR)				SAMD00000001	Delete
	New												Delete
	New												Delete



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Alias	Title	Library Name	Library Source	Library Selection	Library Strategy	Library Construction Protocol	Instrument	Spot Type	Nominal Length	Nominal Sdev	Spot Length	BioSample Used
2	test01-0004_Experiment_0001	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014251		GENOMIC	RANDOM	WGS		454 GS	paired (FF)				SAMD00014251
3	test01-0004_Experiment_0002	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014252		GENOMIC	RANDOM	WGS		455 GS	paired (FF)				SAMD00014252
4	test01-0004_Experiment_0003	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014253		GENOMIC	RANDOM	WGS		456 GS	paired (FF)				SAMD00014253
5	test01-0004_Experiment_0004	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014254		GENOMIC	RANDOM	WGS		457 GS	paired (FF)				SAMD00014254
6	test01-0004_Experiment_0005	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014255		GENOMIC	RANDOM	WGS		458 GS	paired (FF)				SAMD00014255
7	test01-0004_Experiment_0006	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014256		GENOMIC	RANDOM	WGS		459 GS	paired (FF)				SAMD00014256
8	test01-0004_Experiment_0007	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014257		GENOMIC	RANDOM	WGS		460 GS	paired (FF)				SAMD00014257
9	test01-0004_Experiment_0008	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014258		GENOMIC	RANDOM	WGS		461 GS	paired (FF)				SAMD00014258
10	test01-0004_Experiment_0009	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014259		GENOMIC	RANDOM	WGS		462 GS	paired (FF)				SAMD00014259
11	test01-0004_Experiment_0010	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014260		GENOMIC	RANDOM	WGS		463 GS	paired (FF)				SAMD00014260
12	test01-0004_Experiment_0011	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014261		GENOMIC	RANDOM	WGS		464 GS	paired (FF)				SAMD00014261
13	test01-0004_Experiment_0012	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014262		GENOMIC	RANDOM	WGS		465 GS	paired (FF)				SAMD00014262
14	test01-0004_Experiment_0013	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014263		GENOMIC	RANDOM	WGS		466 GS	paired (FF)				SAMD00014263
15	test01-0004_Experiment_0014	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014264		GENOMIC	RANDOM	WGS		467 GS	paired (FF)				SAMD00014264
16	test01-0004_Experiment_0015	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014265		GENOMIC	RANDOM	WGS		468 GS	paired (FF)				SAMD00014265
17	test01-0004_Experiment_0016	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014266		GENOMIC	RANDOM	WGS		469 GS	paired (FF)				SAMD00014266
18	test01-0004_Experiment_0017	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014267		GENOMIC	RANDOM	WGS		470 GS	paired (FF)				SAMD00014267
19	test01-0004_Experiment_0018	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014268		GENOMIC	RANDOM	WGS		471 GS	paired (FF)				SAMD00014268
20	test01-0004_Experiment_0019	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014269		GENOMIC	RANDOM	WGS		472 GS	paired (FF)				SAMD00014269
21	test01-0004_Experiment_0020	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014270		GENOMIC	RANDOM	WGS		473 GS	paired (FF)				SAMD00014270
22	test01-0004_Experiment_0021	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014271		GENOMIC	RANDOM	WGS		474 GS	paired (FF)				SAMD00014271
23	test01-0004_Experiment_0022	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014272		GENOMIC	RANDOM	WGS		475 GS	paired (FF)				SAMD00014272
24	test01-0004_Experiment_0023	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014273		GENOMIC	RANDOM	WGS		476 GS	paired (FF)				SAMD00014273
25	test01-0004_Experiment_0024	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014274		GENOMIC	RANDOM	WGS		477 GS	paired (FF)				SAMD00014274
26	test01-0004_Experiment_0025	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014275		GENOMIC	RANDOM	WGS		478 GS	paired (FF)				SAMD00014275
27	test01-0004_Experiment_0026	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014276		GENOMIC	RANDOM	WGS		479 GS	paired (FF)				SAMD00014276
28	test01-0004_Experiment_0027	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014277		GENOMIC	RANDOM	WGS		480 GS	paired (FF)				SAMD00014277
29	test01-0004_Experiment_0028	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014278		GENOMIC	RANDOM	WGS		481 GS	paired (FF)				SAMD00014278
30	test01-0004_Experiment_0029	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014279		GENOMIC	RANDOM	WGS		482 GS	paired (FF)				SAMD00014279
31	test01-0004_Experiment_0030	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014280		GENOMIC	RANDOM	WGS		483 GS	paired (FF)				SAMD00014280
32	test01-0004_Experiment_0031	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014281		GENOMIC	RANDOM	WGS		484 GS	paired (FF)				SAMD00014281
33	test01-0004_Experiment_0032	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014282		GENOMIC	RANDOM	WGS		485 GS	paired (FF)				SAMD00014282
34	test01-0004_Experiment_0033	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014283		GENOMIC	RANDOM	WGS		486 GS	paired (FF)				SAMD00014283
35	test01-0004_Experiment_0034	Exp Accession: 454 GS 20 paired end sequencing of SAMD00014284		GENOMIC	RANDOM	WGS		487 GS	paired (FF)				SAMD00014284

6. メタデータの作成 --- Run (1) ---

Run のメタデータを、順に設定する。



Submission ID : test01-0004

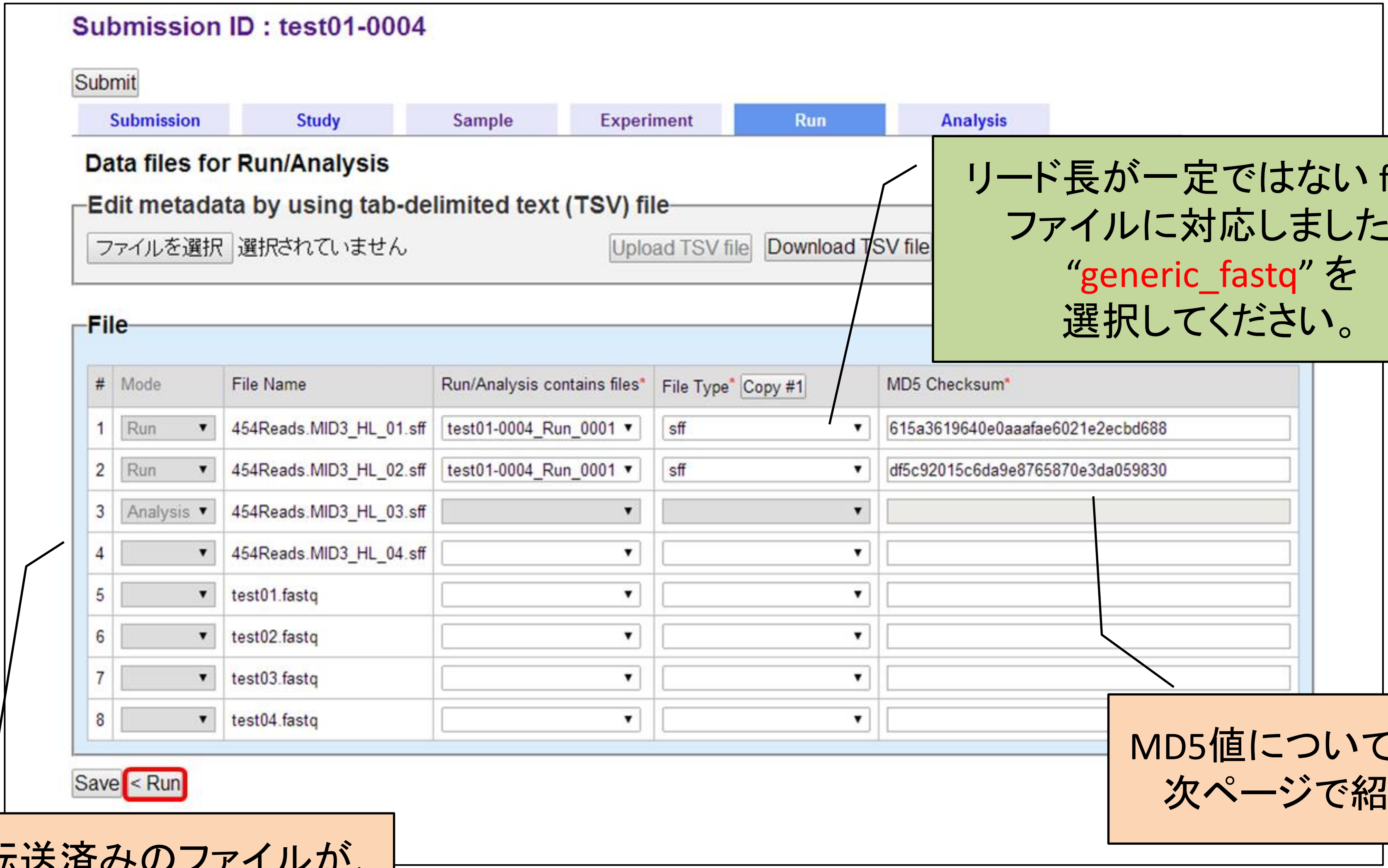
Add new Run(s): 1

#	Alias	Experiment Referenced*	Delete
1	test01-0004_Run_0001	test01-0004_Experiment_0001 ▾	Delete
2	test01-0004_Run_0002	test01-0004_Experiment_0002 ▾	Delete

① を設定後、② [Select data files for Run] のページへ JUMP！

6. メタデータの作成 --- Run (2) ---

② [Select data files for Run] で、データファイルの詳細を設定する。



Submission ID : test01-0004

Data files for Run/Analysis

Edit metadata by using tab-delimited text (TSV) file

File

#	Mode	File Name	Run/Analysis contains files*	File Type* Copy #1	MD5 Checksum*
1	Run	454Reads.MID3_HL_01.sff	test01-0004_Run_0001	sff	615a3619640e0aaafae6021e2ecbd688
2	Run	454Reads.MID3_HL_02.sff	test01-0004_Run_0001	sff	df5c92015c6da9e8765870e3da059830
3	Analysis	454Reads.MID3_HL_03.sff			
4		454Reads.MID3_HL_04.sff			
5		test01.fastq			
6		test02.fastq			
7		test03.fastq			
8		test04.fastq			

リード長が一定ではない fastq ファイルに対応しました。
“generic_fastq” を
選択してください。

MD5値については、
次ページで紹介

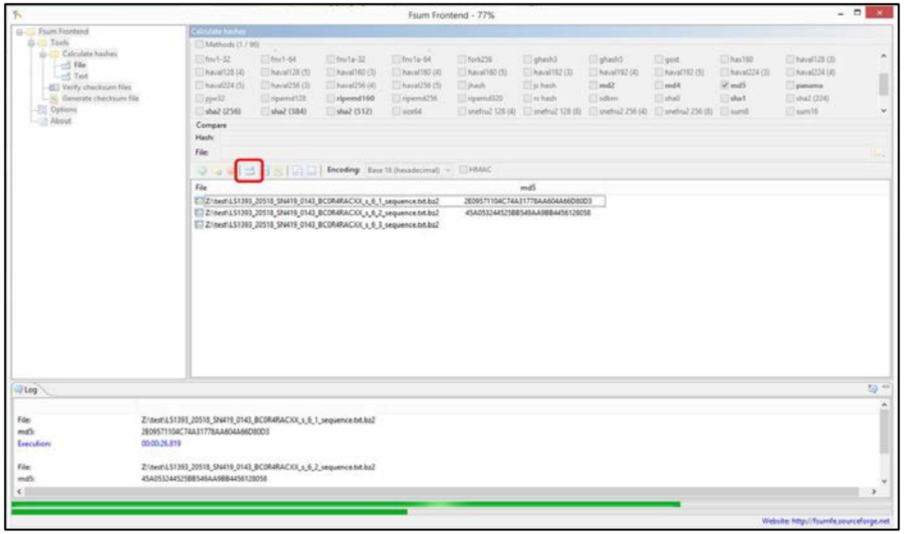
転送済みのファイルが、
自動表示される

6. メタデータの作成 --- Run (3) ---

補足: MD5 値 http://trace.ddbj.nig.ac.jp/dra/submission.html#補足__MD5_値

MD5 (Message Digest Algorithm 5)
32桁の英数字から成るハッシュ関数。ファイルが破損していると、ファイルのMD5値が変化する。
DRA では，到着したファイルの MD5 値の一致をチェックすることで、ファイルの破損がないかどうか確認している。

Windows
[Fsum Frontend (http://fsumfe.sourceforge.net/)](http://fsumfe.sourceforge.net/)



```
File                                                              md5
Z:\test\LS1393_20518_SN419_0143_BC0R4RACXX_s_6_1_sequence.txt.bz2  2E09571104C74A31778AA604A66D80D3
Z:\test\LS1393_20518_SN419_0143_BC0R4RACXX_s_6_2_sequence.txt.bz2  45A0532445258B549AA9BB4456128058
Z:\test\LS1393_20518_SN419_0143_BC0R4RACXX_s_6_3_sequence.txt.bz2
```

```
File:       Z:\test\LS1393_20518_SN419_0143_BC0R4RACXX_s_6_1_sequence.txt.bz2
md5:        2E09571104C74A31778AA604A66D80D3
Execution:  00:00:26.819

File:       Z:\test\LS1393_20518_SN419_0143_BC0R4RACXX_s_6_2_sequence.txt.bz2
md5:        45A0532445258B549AA9BB4456128058
```

Mac OS X
ターミナル

```
$ md5 file1 file2
  9F6E6800CFAE7749EB6C486619254B9C file1
  B636E0063E29709B6082F324C76D0911 file2
```

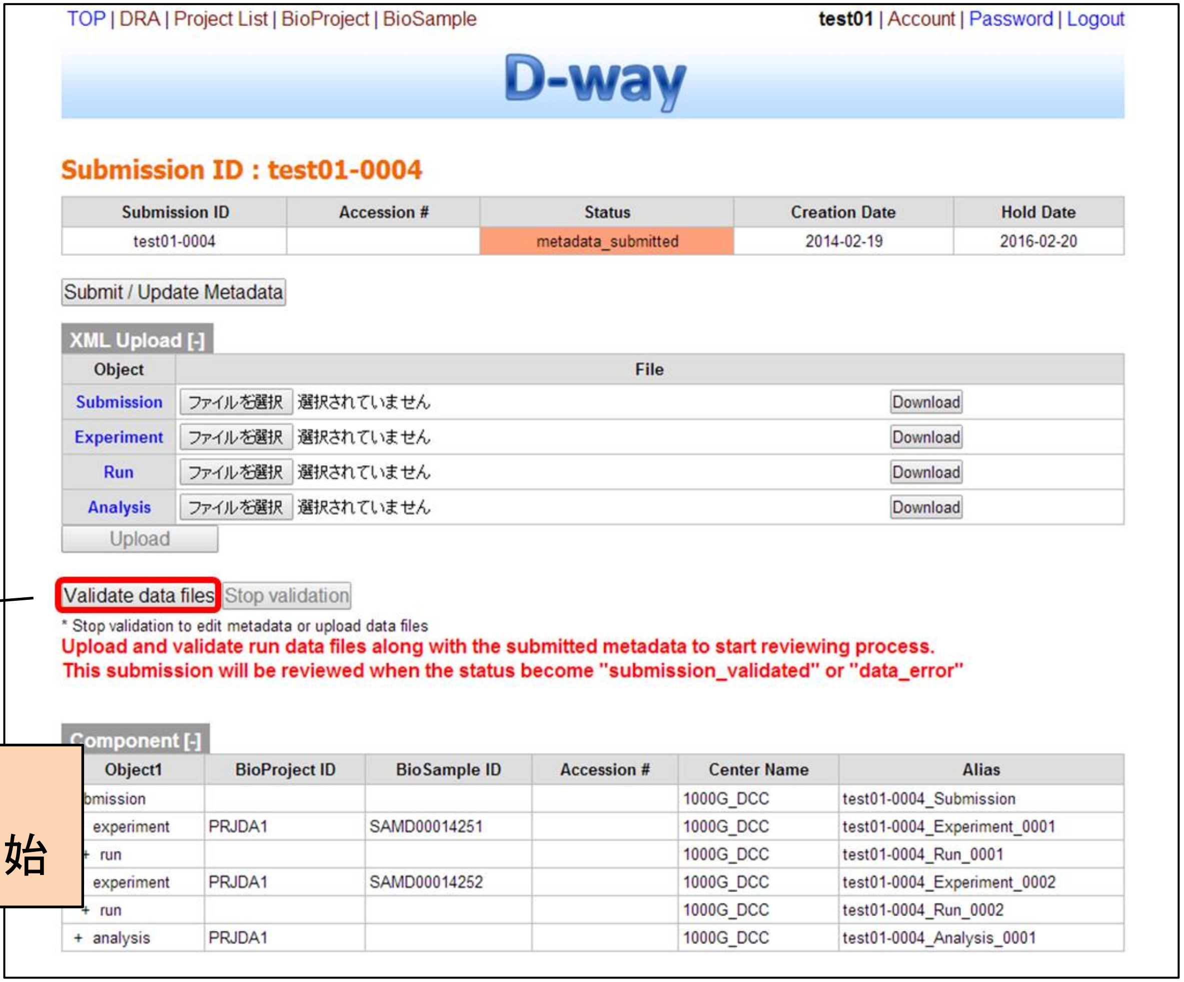
Linux
ターミナル

```
$ md5sum file1 file2
  9F6E6800CFAE7749EB6C486619254B9C file1
  B636E0063E29709B6082F324C76D0911 file2
```

7. Validation 成功 (1)

DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html#データファイルの_Validation

メタデータとデータファイルの整合性を検証し、アーカイブ用 SRA ファイルを作成する作業。
ファイルサイズが大きかったり、混雑していると、長時間かかることがあるので注意。



Submission ID : test01-0004

Submission ID	Accession #	Status	Creation Date	Hold Date
test01-0004		metadata_submitted	2014-02-19	2016-02-20

Object	File	
Submission	ファイルを選択 選択されていません	Download
Experiment	ファイルを選択 選択されていません	Download
Run	ファイルを選択 選択されていません	Download
Analysis	ファイルを選択 選択されていません	Download

* Stop validation to edit metadata or upload data files

Upload and validate run data files along with the submitted metadata to start reviewing process.
This submission will be reviewed when the status become "submission_validated" or "data_error"

Object1	BioProject ID	BioSample ID	Accession #	Center Name	Alias
bmission				1000G_DCC	test01-0004_Submission
experiment	PRJDA1	SAMD00014251		1000G_DCC	test01-0004_Experiment_0001
+ run				1000G_DCC	test01-0004_Run_0001
experiment	PRJDA1	SAMD00014252		1000G_DCC	test01-0004_Experiment_0002
+ run				1000G_DCC	test01-0004_Run_0002
+ analysis	PRJDA1			1000G_DCC	test01-0004_Analysis_0001

クリックして、
Validation を開始

7. Validation 成功 (2)

Validation で、"data_error" が発生したら・・・

[Stop validation] をクリックして Validation 処理を停止した後、メタデータを修正、もしくは、データファイルを再アップロードし、再度 validation を開始します。

＜エラーログの例＞

spot length をメタデータでは400 と設定したが、実際のデータは 500 だったケース

```
fastq-load.2.3.5 err: data excessive while validating formatter within short read archive module - cumulative length of reads
data in file(s): 500 is greater than spot length declared in experiment: 400 in spot 'M00424:28:000000000-
A2G79:1:1101:18351:2171‘
fastq-load.2.3.5 warn: data excessive while validating formatter within short read archive module -
file="Cxxx_L001_R2_001.fastq" line="5" spot_name="M00424:28:000000000-A2G79:1:1101:18351:2171“
fastq-load.2.3.5 warn: data excessive while validating formatter within short read archive module - bad spot
M00424:28:000000000-A2G79:1:1101:18351:2171
```

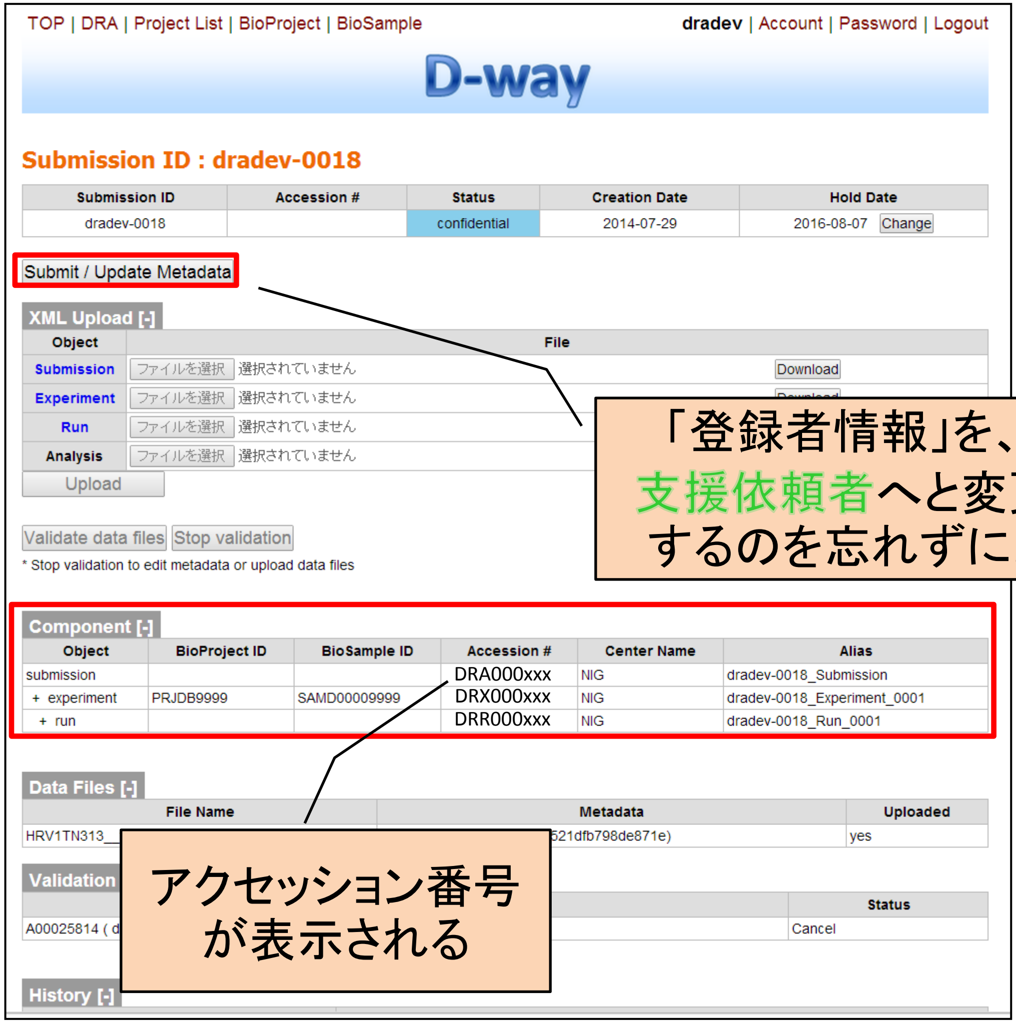
お困りの際は、DRAチーム（trace@ddbj.nig.ac.jp）へご連絡ください。

Validation に成功し、ステータスが “submission_validated” になると・・・

アノテータが、査定を開始します。そのままお待ちください。

8. アクセッション番号の受領

DRA Handbook: http://trace.ddbj.nig.ac.jp/dra/submission.html#アクセッション番号の発行



Submission ID : dradev-0018

Submission ID	Accession #	Status	Creation Date	Hold Date
dradev-0018		confidential	2014-07-29	2016-08-07 Change

Object	File	
Submission	ファイルを選択 選択されていません	Download
Experiment	ファイルを選択 選択されていません	
Run	ファイルを選択 選択されていません	
Analysis	ファイルを選択 選択されていません	

* Stop validation to edit metadata or upload data files

Object	BioProject ID	BioSample ID	Accession #	Center Name	Alias
submission			DRA000xxx	NIG	dradev-0018_Submission
+ experiment	PRJDB9999	SAMD00009999	DRX000xxx	NIG	dradev-0018_Experiment_0001
+ run			DRR000xxx	NIG	dradev-0018_Run_0001

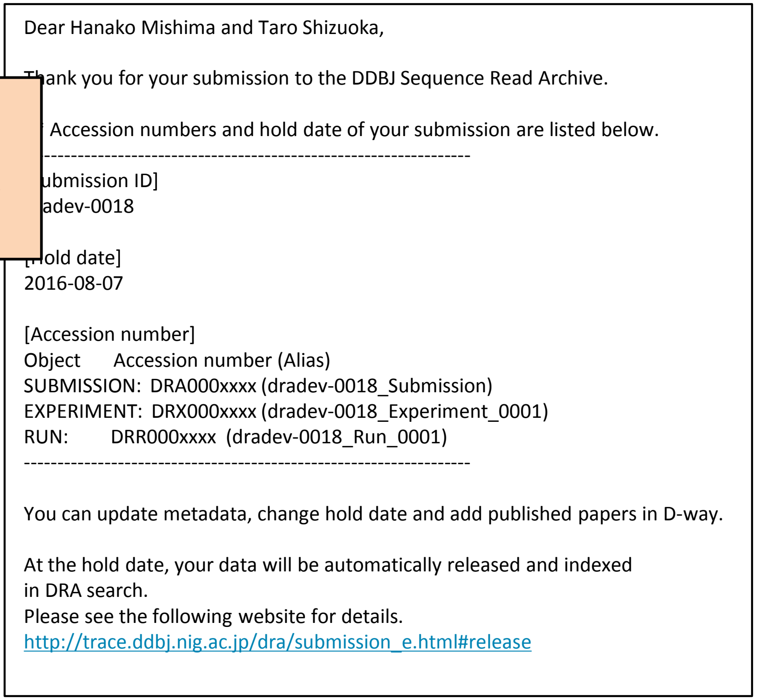
File Name	Metadata	Uploaded
HRV1TN313_	521dfb798de871e)	yes

	Status
A00025814 (d	Cancel

「登録者情報」を、支援依頼者へと変更するのを忘れずに。

アクセッション番号が表示される

アクセッション番号は、メールおよび D-way 上でお知らせします。


Dear Hanako Mishima and Taro Shizuoka,

Thank you for your submission to the DDBJ Sequence Read Archive.

Accession numbers and hold date of your submission are listed below.

----

ubmission ID]
adev-0018

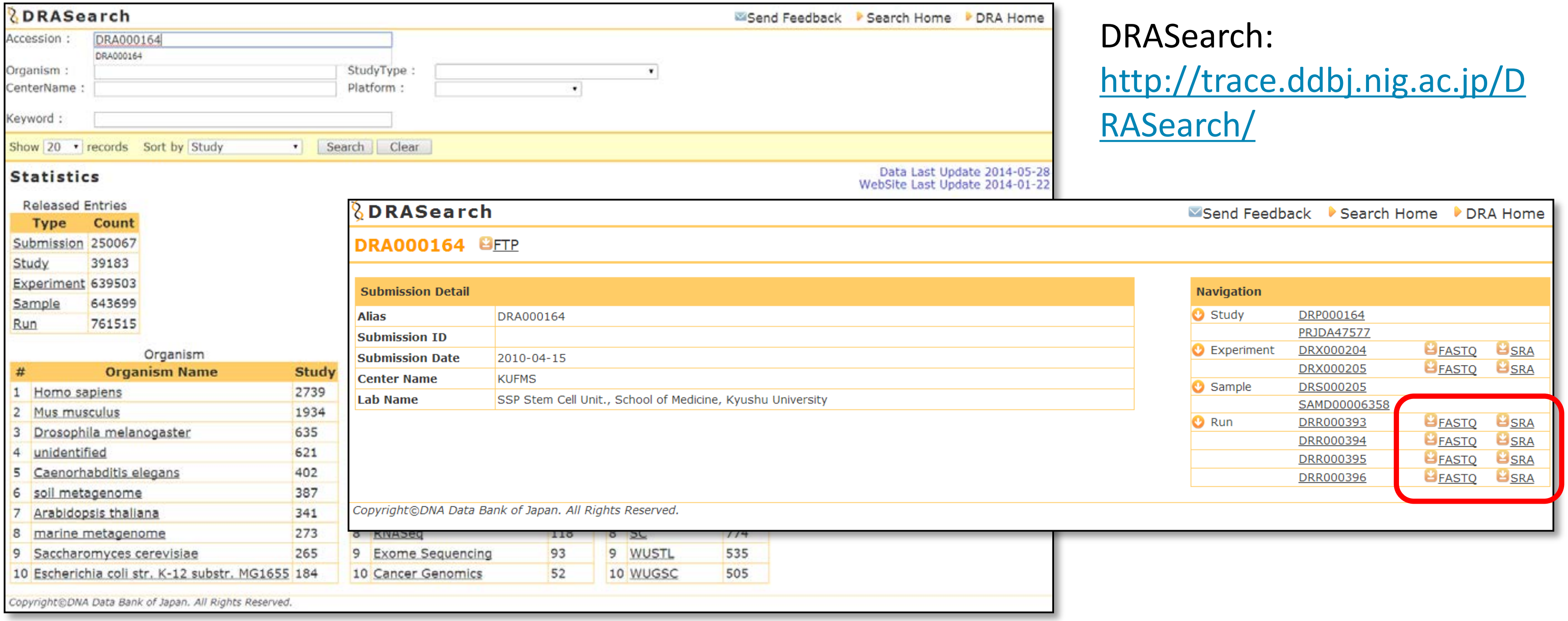
old date]
2016-08-07

[Accession number]
Object　Accession number (Alias)
SUBMISSION: DRA000xxxx (dradev-0018_Submission)
EXPERIMENT: DRX000xxxx (dradev-0018_Experiment_0001)
RUN:　DRR000xxxx (dradev-0018_Run_0001)

You can update metadata, change hold date and add published papers in D-way.

At the hold date, your data will be automatically released and indexed in DRA search.
Please see the following website for details.
http://trace.ddbj.nig.ac.jp/dra/submission_e.html#release

公開済みデータは、DRASearch からダウンロード



DRASearch

Accession : DRA000164

Organism :

StudyType :

CenterName :

Platform :

Keyword :

Show 20 records Sort by Study

Statistics

Data Last Update 2014-05-28
WebSite Last Update 2014-01-22

Released Entries

Type	Count
Submission	250067
Study	39183
Experiment	639503
Sample	643699
Run	761515

Organism

#	Organism Name	Study
1	Homo sapiens	2739
2	Mus musculus	1934
3	Drosophila melanogaster	635
4	unidentified	621
5	Caenorhabditis elegans	402
6	soil metagenome	387
7	Arabidopsis thaliana	341
8	marine metagenome	273
9	Saccharomyces cerevisiae	265
10	Escherichia coli str. K-12 substr. MG1655	184

8	RNASeq	118
9	Exome Sequencing	93
10	Cancer Genomics	52

8	SC	774
9	WUSTL	535
10	WUGSC	505

Copyright©DNA Data Bank of Japan. All Rights Reserved.

DRASearch

DRA000164 FTP

Submission Detail

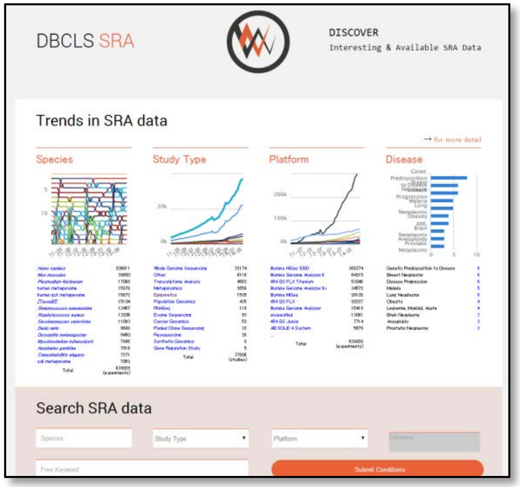
Alias	DRA000164
Submission ID	
Submission Date	2010-04-15
Center Name	KUFMS
Lab Name	SSP Stem Cell Unit., School of Medicine, Kyushu University

Navigation

Study	DRP000164 PRJDA47577		
Experiment	DRX000204	FASTQ	SRA
	DRX000205	FASTQ	SRA
Sample	DRS000205 SAMD00006358		
Run	DRR000393	FASTQ	SRA
	DRR000394	FASTQ	SRA
	DRR000395	FASTQ	SRA
	DRR000396	FASTQ	SRA

Copyright©DNA Data Bank of Japan. All Rights Reserved.

DRASearch:
http://trace.ddbj.nig.ac.jp/DRASearch/



DBCLS SRA

DISCOVER
Interesting & Available SRA Data

Trends in SRA data

Search SRA data

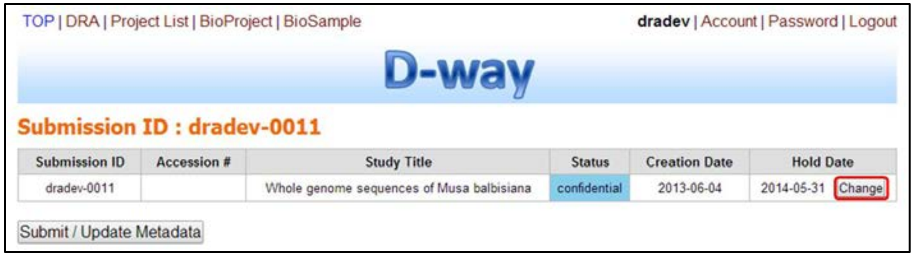
DBCLS SRA (http://sra.dbcls.jp/) もどうぞご利用ください。

支援依頼者の皆様へのお願い

☑ ご登録は、お早目に。一定期間、非公開に出来ます。

☑ 公開日の変更は、D-wayから、ご自身で。

DDBJ データ公開原則: http://www.ddbj.nig.ac.jp/sub/hold_date-j.html

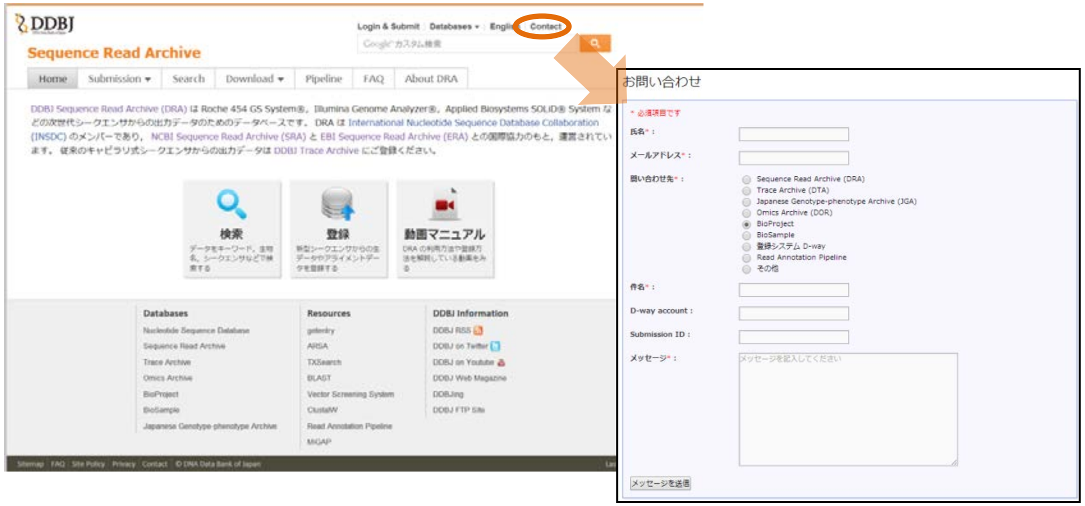


Submission ID : dradev-0011

Submission ID	Accession #	Study Title	Status	Creation Date	Hold Date
dradev-0011		Whole genome sequences of Musa balbisiana	confidential	2013-06-04	2014-05-31 Change

☑ 論文が受理されたら、文献情報を 追加して下さい。

PubMed ID もしくは DOI を BioProjectチーム へ連絡



DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System などの次世代シークエンサからの出力データのためのデータベースです。DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり、NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと、運営されています。従来のキャピラリ式シークエンサからの出力データは DDBJ Trace Archive にご登録ください。

お問い合わせ

* 必須項目です

氏名* :

メールアドレス* :

問い合わせ先* :
- ○ Sequence Read Archive (DRA)
- ○ Trace Archive (DTA)
- ○ Japanese Genotype-phenotype Archive (JGA)
- ○ Omics Archive (DOR)
- ● BioProject
- ○ BioSample
- ○ 登録システム D-way
- ○ Read Annotation Pipeline
- ○ その他

件名* :

D-way account :

Submission ID :

メッセージ* : メッセージを記入してください