

次世代シーケンサーデータの解析手法
第11回統合データ解析環境Galaxy
ウェブ資料

大田達郎*、寺田朋子、清水謙多郎、門田 幸二*

①Galaxyの2005年の論文の引用回数は、②1,467回

W1 : Google Scholar

The screenshot shows a web browser window with the Google Scholar search results for the paper "Galaxy: a platform for interactive large-scale genome analysis". The search bar contains the text "Galaxy: a platform for interactive large-scale genome analysis". The search results show the title, authors (B Giardine, C Riemer, RC Hardison, R Burhans...), the journal (Genome Res., 2005), and the abstract. A red arrow with the number 2 points to the citation count "引用元 1467".

Galaxy: a platform for interactive large-scale genome analysis [HTML] cshlp.org Full text @ UTokyo

B Giardine, C Riemer, RC Hardison, R Burhans... - Genome Res., 2005 - genome.cshlp.org

Abstract Accessing and analyzing the exponentially expanding genomic sequence and functional data pose a challenge for biomedical researchers. Here we describe an interactive system, Galaxy, that combines the power of existing genome annotation databases with a simple Web portal to enable users to search remote resources, combine data from independent queries, and visualize the results. The heart of Galaxy is a flexible ...

引用元 1467 関連記事 全 21バージョン 引用 保存 その他

この検索の最上位の結果を表示しています。 検索結果をすべて見る

①

Giardine et al., *Genome Res.*, 15: 1451–1455, 2005

①Galaxyの2010年の論文の引用回数は、②2,517回

W1 : Google Scholar

ウェブ 画像 もっと見る... ログイン

Google Galaxy: a comprehensive approach for supporting acc

Scholar 期間指定なし

[HTML] Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences [HTML] biomedcen... Full text @ UTokyo

J Goecks, A Nekrutenko... - Genome ..., 2010 - genomebiology.biomedcentral.com

Abstract Increased reliance on computational approaches in the life sciences has revealed grave concerns about how accessible and reproducible computation-reliant results truly are. Galaxy <http://usegalaxy.org>, an open web-based platform for genomic research, addresses these problems. Galaxy automatically tracks and manages data provenance and provides support for capturing the context and intent of computational methods. Galaxy Pages are ...

引用元 2517 関連記事 全 38 バージョン 引用 保存 その他

この検索の最上位の結果を表示しています。 検索結果をすべて見る

Google Scholar について プライバシー 規約 フィードバックを送信



Goecks et al., *Genome Biol.*, 11: 128, 2010

W2: BioStarのGalaxy版

ライフサイエンスQAを知っているヒトには馴染みのある見た目ですね

The screenshot shows a web browser window displaying the BioStar website. The address bar shows the URL <https://biostar.usegalaxy.org/>. The page features a navigation bar with buttons for "Latest", "Home", and "Log In". Below the navigation bar is a search bar with the placeholder text "Live search: start typing...". The main content area displays a list of questions and answers. Each question is preceded by a red box containing the number of answers (0) and the word "answers". The questions listed are:

- stringtie output from multiple samples into an FPKM matrix (written 13 hours ago by vaughandy • 10 answers)
- negative and positive fold change for a set of unigenes with similar gene ids in a single condition (written 14 hours ago by huge_ashes • 0 answers)
- Sequence logo from a text file? (written 15 hours ago by o3964fn4n • 0 answers)
- What Galaxy tools can preprocess VCF file before running the SNPEff annotation process? (written 16 hours ago by msprindzhuk • 0 answers)
- what is thye meaning of too low aQual in Htseq (written 16 hours ago by sandy.mbt • 0 answers)
- Bacterial Sequence for Variant Analysis (written 22 hours ago by npurkavastha2010 • 0 answers, updated 16 hours ago by m_hermt • 0 answers)

Traffic: 109 users visited in the last hour

W3-1 : ToolShed

① ToolShedのトップ画面。② 2017年3月22日現在、4,719個のツールを利用可能

4719 valid tools on Mar 22, 2017

Repositories by Category

search repository name, description

Name	Description	Repositories
Assembly	Tools for working with assemblies	87
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	44
Combinatorial Selections	Tools for combinatorial selection	7
Computational chemistry	Tools for use in computational chemistry	28
Constructive Solid Geometry	Tools for constructing and analyzing 3-dimensional shapes and their properties	11
Convert Formats	Tools for converting data formats	73
Data Export	Tools for exporting data to various destinations	1

W3-1 : ToolShed

①カテゴリーごとに分かれているようだ。たとえば②Assemblyプログラムは87個あるのだろう

Galaxy Tool Shed

4719 valid tools on Mar 22, 2017

Search

- Search for valid tools
- Search for workflows

Valid Galaxy Utilities

- Tools
- Custom datatypes
- Repository dependency definitions
- Tool dependency definitions

All Repositories

- Browse by category

Available Actions

- Login to create a repository

Repositories by Category

search repository name, description

Name	Description	Repositories
Assembly	Tools for working with assemblies	87
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	44
Combinatorial Selections	Tools for combinatorial selection	7
Computational chemistry	Tools for use in computational chemistry	28
Constructive Solid Geometry	Tools for constructing and analyzing 3-dimensional shapes and their properties	11
Convert Formats	Tools for converting data formats	73
Data Export	Tools for exporting data to various destinations	1

W3-1 : ToolShed

①Assemblyをクリック。反応がない、、、が10秒くらい待つとページが切り替わる

The screenshot shows the Galaxy Tool Shed interface. The browser address bar displays <https://toolshed.g2.bx.psu.edu/>. The page title is "Galaxy Tool Shed" and the navigation menu includes "Repositories", "Groups", "Help", and "User". A status bar at the top left indicates "4719 valid tools on Mar 22, 2017".

The main content area is titled "Repositories by Category" and features a search input field with the placeholder text "search repository name, description". Below this is a table listing various repository categories:

Name	Description	Repositories
Assembly	Tools for working with assemblies	87
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	44
Combinatorial Selections	Tools for combinatorial selection	7
Computational chemistry	Tools for use in computational chemistry	28
Constructive Solid Geometry	Tools for constructing and analyzing 3-dimensional shapes and their properties	11
Convert Formats	Tools for converting data formats	73
Data Export	Tools for exporting data to various destinations	1

A red lightning bolt with the number "1" is positioned over the "Assembly" link in the first row of the table.

W3-1 : ToolShed

① Assemblyのカテゴリに属する87個のプログラムがリストアップされているのだろう。
② abyssというアセンブラから表示されているので、アルファベット順なのだろうと妄想

Galaxy Tool Shed

4719 valid tools on Mar 22, 2017

Search

- Search for valid tools
- Search for workflows

Valid Galaxy Utilities

- Tools
- Custom datatypes
- Repository dependency definitions
- Tool dependency definitions

All Repositories

- Browse by category

Available Actions

- Login to create a repository

Repositories in Category Assembly

search repository name, description

Name	Synopsis	Type	Metadata Revision
abyss	Assembly By Short Sequences - a de novo, parallel, paired-end sequence assembler	Unrestricted	1 (2017-)
abyss tool		Unrestricted	0 (2012-)
abyss toolsuite	This suite contains Abyss and Abyss-PE config files and wrappers for Galaxy	Unrestricted	0 (2011-)

W3-1 : ToolShed

Velvetアセンブラがあるかなと妄想しながら①ページ下部に移動。確かに②velvet関連プログラムが存在する

The screenshot shows the Galaxy Tool Shed interface. The browser address bar displays <https://toolshed.g2.bx.psu.edu/>. The page title is "Galaxy Tool Shed". The navigation menu includes "Repositories", "Groups", "Help", and "User".

On the left sidebar, there are sections for "Search", "Valid Galaxy Utilities", "All Repositories", and "Available Actions".

The main content area displays a list of tools. A red box highlights the following tools:

Tool Name	Description	Access	Version
trinityrnaseq_protocol	from RNA-seq using the Trinity platform for reference generation and analysis	Unrestricted	30 (2015-01-01)
velvetoptimiser	Auto optimise a genomic velvet assembly	Unrestricted	0 (2011-01-01)
velvetoptimiser	An updated wrapper for the Velvet Optimiser	Unrestricted	0 (2016-01-01)
velvet_toolsuite	Velvet assembler (a different version than in galaxy-central)	Unrestricted	0 (2011-01-01)

A red arrow with the number "2" points to the highlighted tools. Another red arrow with the number "1" points to the bottom of the page, indicating the scroll action.

W3-1 : ToolShed

①Browse by categoryを押すと、元の画面に戻ります

The screenshot shows the Galaxy Tool Shed interface. The browser address bar displays <https://toolshed.g2.bx.psu.edu/>. The page title is "Galaxy Tool Shed" and the navigation menu includes "Repositories", "Groups", "Help", and "User".

On the left sidebar, under "Valid Galaxy Utilities", there is a link for "Browse by category" which is highlighted with a red arrow and a circled number 1. Other sidebar links include "Search for valid tools", "Search for workflows", "Tools", "Custom datatypes", "Repository dependency definitions", "Tool dependency definitions", "All Repositories", and "Login to create a repository".

The main content area is titled "Repositories by Category" and features a search input field with the placeholder text "search repository name, description". Below this is a table listing various tool categories and the number of repositories in each.

Name	Description	Repositories
Assembly	Tools for working with assemblies	87
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	45
Combinatorial Selections	Tools for combinatorial selection	7
Computational chemistry	Tools for use in computational chemistry	28
Constructive Solid Geometry	Tools for constructing and analyzing 3-dimensional shapes and their properties	11
Convert Formats	Tools for converting data formats	73
Data Export	Tools for exporting data to various destinations	1

W4-1 : Galaxy main

Galaxy mainのトップ画面。①Login or Register、②Registerを選択して、ユーザ登録する

The screenshot shows the Galaxy main page in a web browser. The address bar displays <https://usegalaxy.org/>. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Help, and Login or Register. A red arrow labeled '1' points to the 'Login or Register' link. Below this, a dropdown menu is open, showing 'Login' and 'Register' options, with a red arrow labeled '2' pointing to 'Register'. The main content area features a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a "Want help? Get answers." section with the Biostars logo and the text "GALAXY EXPLAINED". On the left, a "Tools" sidebar lists various categories like Get Data, Lift-Over, Collection Operations, etc. On the right, there is a "Login" section with a search bar for datasets and an "Unnamed history (empty)" section with a message: "ヒストリーは空です。 You can [load your own data](#) or [get data from an external source](#)".

W4-2: Register

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'Login or Register'. The 'Create account' form is highlighted in yellow and contains the following fields:

- Email address:** [text input]
- Password:** [text input] with a **Strength** indicator below it.
- Confirm password:** [text input]
- Public name:** [text input] with a note: "Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes ('.', '_', '-')."
- Subscribe to mailing list:** [checkbox] with a link to "See all Galaxy project mailing lists."

The right sidebar shows the **History** section with a search bar and a message: "历史信息は空です。 You can [load your own data](#) or [get data from an external source](#)."

W4-2: Register

①必要事項を入力して…画面下部に移動できないので画面サイズを大きくする

The screenshot shows the Galaxy web interface with the 'Create account' form. The form fields are highlighted with a red box, and a red arrow with the number 1 points to the form area. The form fields are:

- Email address:** kadota@iu.a.u-tokyo.ac.jp
- Password:** [masked] with a strength indicator (Strength)
- Confirm password:** [masked]
- Public name:** agribio_T_T_desu

Below the form, there is a note: "Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes ('.', '_', '-')."

There is also a checkbox for "Subscribe to mailing list:" and a link to "See all Galaxy project mailing lists."

W4-2: Register

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays the 'Create account' form with the following fields and options:

- Email address:**
- Password:**
Strength:
- Confirm password:**
- Public name:**
Your public name is an identifier that will be used to generate addresses for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes ('.', '_', '-').
- Subscribe to mailing list:**
See [all Galaxy project mailing lists](#).

The 'Submit' button at the bottom of the form is highlighted with a red arrow and a circled '1', indicating the step to click.

The left sidebar contains a 'Tools' section with a search bar and a list of tool categories: Get Data, Lift-Over, Collection Operations, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, NGS: QC and manipulation, NGS: DeepTools, NGS: Mapping, NGS: RNA Analysis, NGS: SAMtools, NGS: BamTools, NGS: Picard, NGS: VCF Manipulation, NGS: Peak Calling, and NGS: Variant Analysis.

The right sidebar contains a 'History' section with a search bar and a message: 'Unnamed history (empty)'. A blue information box states: 'ヒストリーは空です。 You can [load your own data](#) or [get data from an external source](#)'.

W4-3: Public name

①Public nameはlowercase letters (小文字)で書けと言われていたのを忘れていましたm(_ _)m。②泣き顔(T_T)を入れる遊びはダメということですねw

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', 'Login or Register', and a grid icon. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'Text Manipulation', and 'NGS: QC and manipulation'. The main content area features a 'Create account' form with the following fields:

- Email address:** kadota@iu.a.u-tokyo.ac.jp
- Password:** [Redacted]
- Confirm password:** [Redacted]
- Public name:** agribio_T_T_desu

A red error message box is overlaid on the form, stating: "Public name must contain only lowercase letters, numbers, '.', '_' and '-'. It also has to be between 3 and 255 characters in length." A red arrow labeled '1' points to this message. Another red arrow labeled '2' points to the 'Public name' field. The 'History' panel on the right shows 'Unnamed history (empty)' and a blue information box: "ヒストリーは空です。 You can [load your own data](#) or [get data from an external source](#)".

W4-3: Public name

- ① 該当部分を小文字に変えて、
- ② ×を押…せないようなので…
- ③ 画面サイズをさらに広げる

Galaxy

Analyze Data Workflow Shared Data Visualization Help Login or Register

Tools

search tools

Get Data
Lift-Over
Collection Operations
Text Manipulation
Datamash
Convert Formats
Filter and Sort
Join, Subtract and Group
Fetch Alignments/Sequences
NGS: QC and manipulation
NGS: DeepTools
NGS: Mapping
NGS: RNA Analysis
NGS: SAMtools
NGS: BamTools
NGS: Picard
NGS: VCF Manipulation
NGS: Peak Calling
NGS: Variant Analysis

History

search datasets

Unnamed history
(empty)

ヒストリーは空です。 You can [load your own data](#) or [get data from an external source](#)

Create account

Email address:
kadota@iu.a.u-tokyo.ac.jp

Password:
.....
Strength

Confirm password:
.....

Public name:
agribio_t_t_desu

Your public name is an identifier that will be used to generate [general access](#) for information you share publicly. Public names must be at least three characters in length and contain only lower-case letters, numbers, dots, underscores, and dashes ('.', '_', '-').

Subscribe to mailing list:

Public name must contain only lowercase letters, numbers, '.', '_' and '-'. It also has to be between 3 and 255 characters in length.

W4-3: Public name

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area is the 'Create account' form. A red error box at the top of the form states: "Public name must contain only lowercase letters, numbers, '.', '_' and '-'. It also has to be between 3 and 255 characters in length." The 'Public name' field contains the text 'agribio_t_t_desu'. Below the form is a 'Submit' button, which is highlighted with a red circle containing the number '1' and a red arrow pointing to it. The left sidebar contains a 'Tools' menu with various categories like 'Get Data', 'Text Manipulation', and 'NGS: QC and manipulation'. The right sidebar shows a 'History' section with an 'Unnamed history' entry that is empty.

W4-4: ログインできた

①kadota@iu.a.u-tokyo.ac.jpさんとしてログイン状態となったようだ。②e-mailが送られたようなので、それをクリックしてverifyしないといけないようですね

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'Login or Register'. A green notification box in the center contains the following text: 'Now logged in as kadota@iu.a.u-tokyo.ac.jp. Verification email has been sent to your email address. Please verify it by clicking the activation link in the email. Please check your spam/trash folder in case you cannot find the message. Return to the home page.' Two red arrows with the number '1' point to the top of the notification box, and two red arrows with the number '2' point to the text 'Please check your spam/trash folder...'. On the right side, the 'History' panel is visible, showing 'Unnamed history (empty)' and a blue information box with the text: 'ヒストリーは空です. You can load your own data or get data from an external source'. The left sidebar lists various tool categories such as 'Get Data', 'Text Manipulation', and 'NGS: QC and manipulation'.

確かに①ユーザagribio_t_t_desuさんの
アドレス宛にメールがきた。②をクリック

W4-5: メールがすぐきた

2017/05/15 (月) 11:37

GA Galaxy Activation <activate@galaxyproject.org>

Galaxy Account Activation

宛先 kadota@iu.a.u-tokyo.ac.jp

Hello agribio_t_t_desu,



In order to complete the activation process for kadota@iu.a.u-tokyo.ac.jp begun on 05/15/17 at usegalaxy.org, please click on the following link to verify your account:

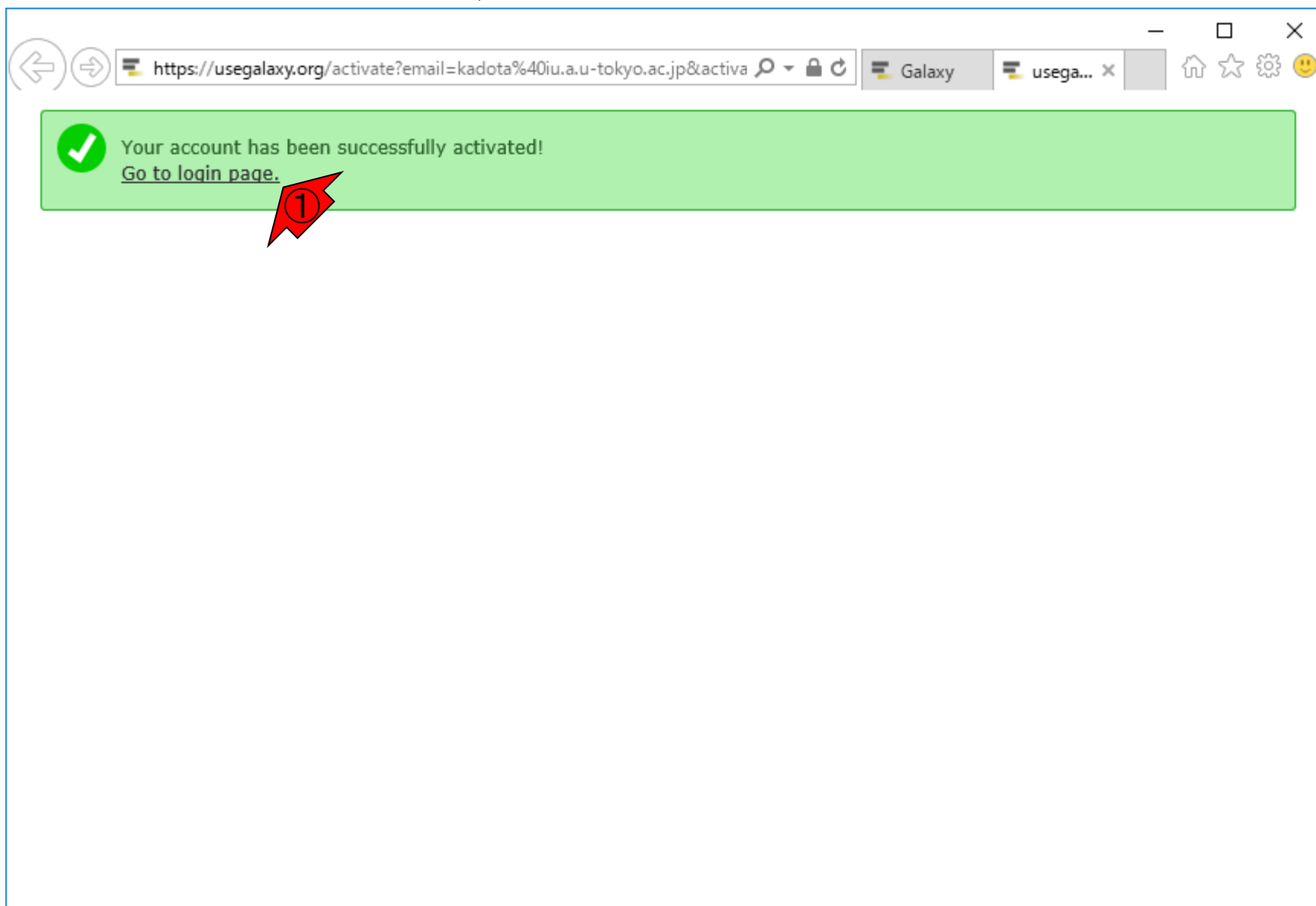
https://usegalaxy.org/activate?email=kadota%40iu.a.u-tokyo.ac.jp&activation_token=a20c86776f6b265216d86c44f915d303c669ed05



By clicking on the above link and opening a Galaxy account you are also confirming that you have read and agreed to Galaxy's Terms and Conditions for use of this service (<https://usegalaxy.org/static/terms.html>). This includes a quota limit of one account per user. Attempts to subvert this limit by creating multiple accounts or through any other method may result in termination of all associated accounts and data.

いわれるがままに、①ログインページに行ってみよう

W4-6: アカウントactivated



W4-7: ログインページ?!

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/root>. The main content area displays a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a large banner for the "iSCB ISMB ECCB 2017 • PRAGUE" conference, featuring a DNA double helix logo and the text "Making Galaxy work for you Register now ISMB/ECCB 2017 Tutorial".

The left sidebar contains a "Tools" section with a search bar and a list of tool categories: Get Data, Send Data, Lift-Over, Collection Operations, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, NGS: QC and manipulation, NGS: DeepTools, NGS: Mapping, NGS: RNA Analysis, NGS: SAMtools, NGS: BamTools, NGS: Picard, NGS: VCF Manipulation, and NGS: Peak Calling.

The right sidebar shows a "History" section with a search bar and a message: "ヒストリーは空です. You can [load your own data](#) or [get data from an external source](#)".

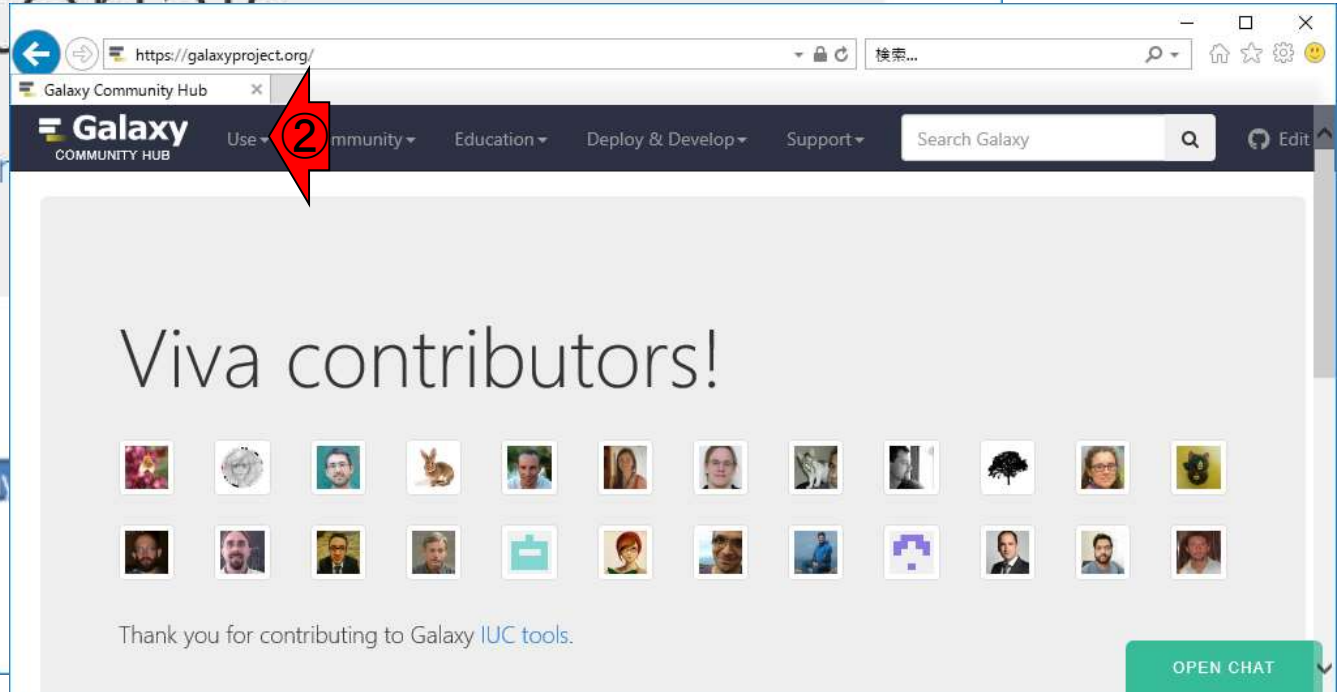
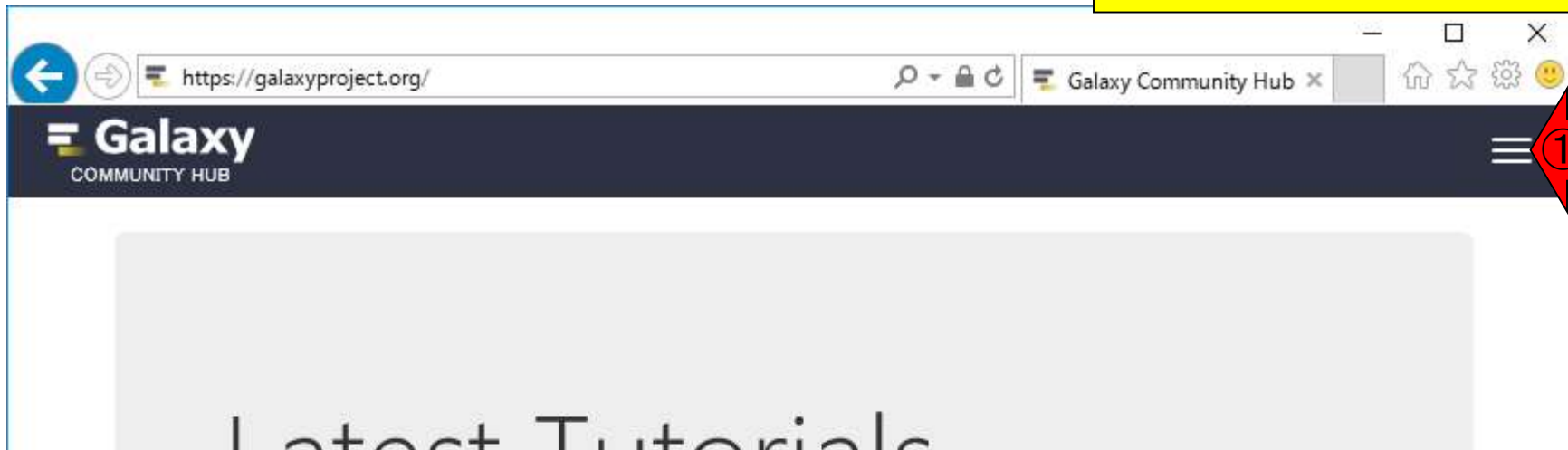
W4-8: ログインしてるのね

①Userをクリックすると、
②Logged in as …となっているからですかね

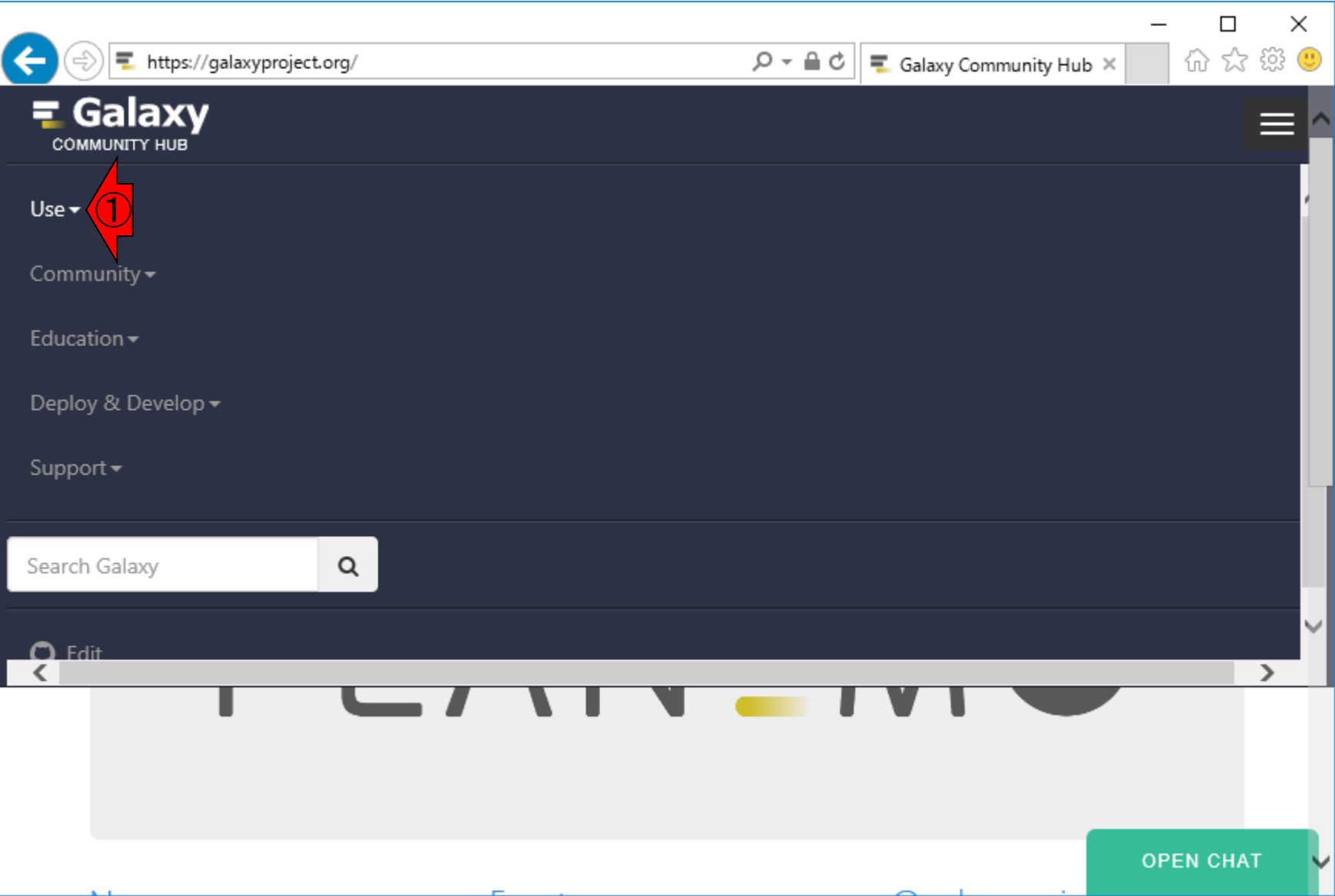
The screenshot shows the Galaxy web interface at <https://usegalaxy.org/root>. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow labeled '①' points to the 'User' dropdown menu. The dropdown menu is open, showing the user is logged in as 'kadota@iu.a.u-t' (indicated by a red arrow labeled '②'). Other options in the menu include 'Preferences', 'Custom Builds', 'Logout', 'Saved Histories', 'Saved Datasets', and 'Saved Pages'. The main content area features a large heading 'Running Your Own Understanding how Galaxy works' and a sub-heading 'An in-depth tutorial'. The left sidebar contains a 'Tools' section with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', etc.

W5-0: 大元から...

Galaxyの大元のページ。①をクリック。右下のようになっている場合は、②Useを押すと次のスライドとほぼ同じになります



W5-0: 大元から...



W5-0: 大元から...

The screenshot shows a web browser window with the URL <https://galaxyproject.org/>. The page title is "Galaxy COMMUNITY HUB". A dark navigation menu is open, listing several categories: "Use", "Community", "Education", "Deploy & Develop", and "Support". Under the "Use" category, "Main Galaxy" is highlighted with a red arrow and a circled "1". Other items in the "Use" menu include "Public Servers" and "Cloud". Below the navigation menu, the main content area features the text "and transparent computational biomedical research." followed by a bullet point: "• **Accessible:** Users without programming experience can easily specify parameters and run tools and workflows." At the bottom right, there is a green button labeled "OPEN CHAT". The browser's address bar at the bottom shows the URL <https://galaxyproject.org/main/>.

main Galaxyにたどり着きます。②ページ下部に移動

W5-1 : main Galaxy

The screenshot shows a web browser window with the URL <https://galaxyproject.org/main/>. A red arrow labeled '1' points to the address bar. The page header features the 'Galaxy COMMUNITY HUB' logo and a hamburger menu icon. A red arrow labeled '2' points to the vertical scrollbar on the right side of the page. The main content area includes the title 'Galaxy Main public site', a paragraph of introductory text, a 'Note' box, and sections for 'Status of the public site' and 'Resouces available to main site'. A green 'OPEN CHAT' button is located in the bottom right corner.

Galaxy
COMMUNITY HUB

Galaxy Main public site

The main Galaxy site at <http://usegalaxy.org> is an installation of the Galaxy software combined with many common tools and data; this site has been available since 2007 for anyone to analyze their data free of charge. The site provides substantial CPU and disk space, making it possible to analyze large datasets. The site supports thousands of users and hundreds of thousands of jobs per month (see [Project Statistics](#)). It is sustained by [TACC](#) hardware using allocation generously provided by the [CyVerse](#) project.

Anyone can use the public servers, with or without an account, but Galaxy user accounts are simple to create (email, password, user name and go!). With an account, data quotas are increased and full functionality across sessions opens up, such as naming, saving, sharing, and publishing Galaxy objects (Histories, Workflows, Datasets, Pages).

Note: *Galaxy's Terms and Conditions* (see below) specifically declare a "one-account per user" requirement.

Status of the public site

The status page with the current state of the Main Server, Test Server, and ToolShed is available at <http://status.galaxyproject.org/>.

Resouces available to main site

OPEN CHAT

W5-1 : main Galaxy

①のあたりまで移動して、②User data and job quotasを眺める。③利用可能なデータ量が1アカウントにつき250GBまで、同時に実行できるジョブが6つまでという制限があることがわかる。④作成できるアカウントは1人1つなので、複数のアカウントを作成するのはルール違反

Galaxy
COMMUNITY HUB

② User data and job quotas

Jobs that have low mamory and CPU requirements are subject to the following limits:

Maximum total accounts per user	1 registered/unregistered
Maximum total user data on server	250GB for registered users, 5GB for unregistered users
Maximum concurrent jobs	6 for registered users, 1 for unregistered users

Tools utilizing multiple CPUs and over 8 Gb of RAM are subject to stricted limits:

Resource	Per-resource job concurrency quotas
Increased memory tools	1 registered/unregistered
Galaxy cluster	2 registered, unregistered not allowed
TACC Stampede	4 registered, unregistered not allowed
Galaxy cluster test/development	1 registered, unregistered not allowed
TACC Stampede test/development	1 registered, unregistered not allowed

OPEN CHAT

Galaxy mainの基本操作画面は、縦に3分割された構成

W5-2: Galaxyの画面構成

The screenshot displays the Galaxy web interface at <https://usegalaxy.org/>. The interface is divided into three vertical sections:

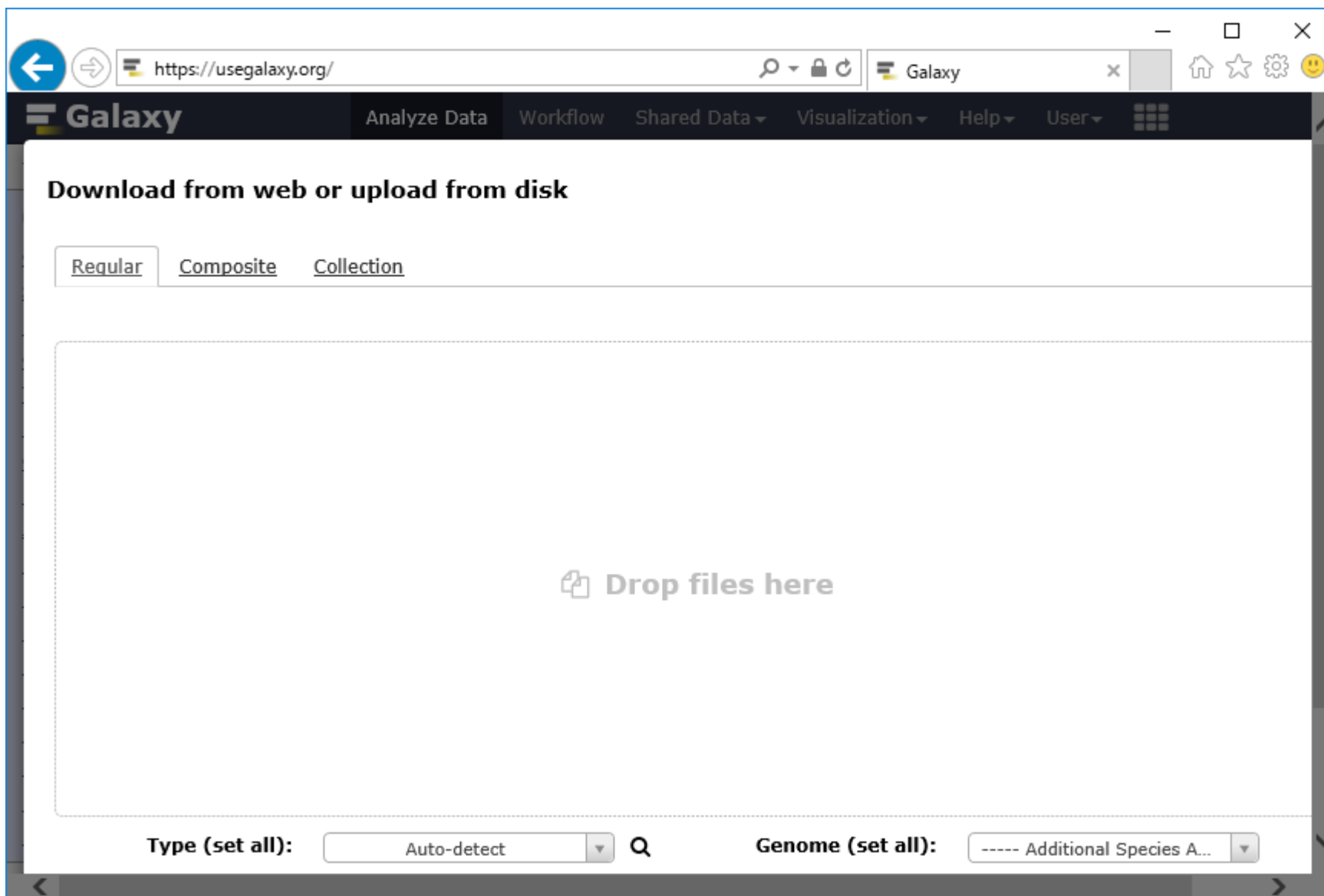
- Tools (Left):** A sidebar containing a search bar and a list of tool categories such as "Get Data", "Send Data", "Lift-Over", "Collection Operations", "Text Manipulation", "Datamash", "Convert Formats", "Filter and Sort", "Join, Subtract and Group", "Fetch Alignments/Sequences", "NGS: QC and manipulation", "NGS: DeepTools", "NGS: Mapping", "NGS: RNA Analysis", "NGS: SAMtools", "NGS: BamTools", "NGS: Picard", and "NGS: VCF Manipulation".
- Main Content (Center):** A large area displaying a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a promotional banner for the "iSCB ISMB ECCB 2017 • PRAGUE JULY 21 TO JULY 25" with the text "Making Galaxy work for you" and "Register now".
- History (Right):** A sidebar titled "History" with a search bar and a message: "Unnamed history (empty)". A blue information box contains the text: "ヒストリーは空です。 You can [load your own data](#) or [get data from an external source](#)".

①を押し、デスクトップ上のgzip圧縮FASTQファイル(DDR024501sub_1.fastq.gz)のアップロードを行う

W5-3: アップロード①

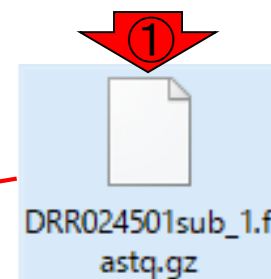
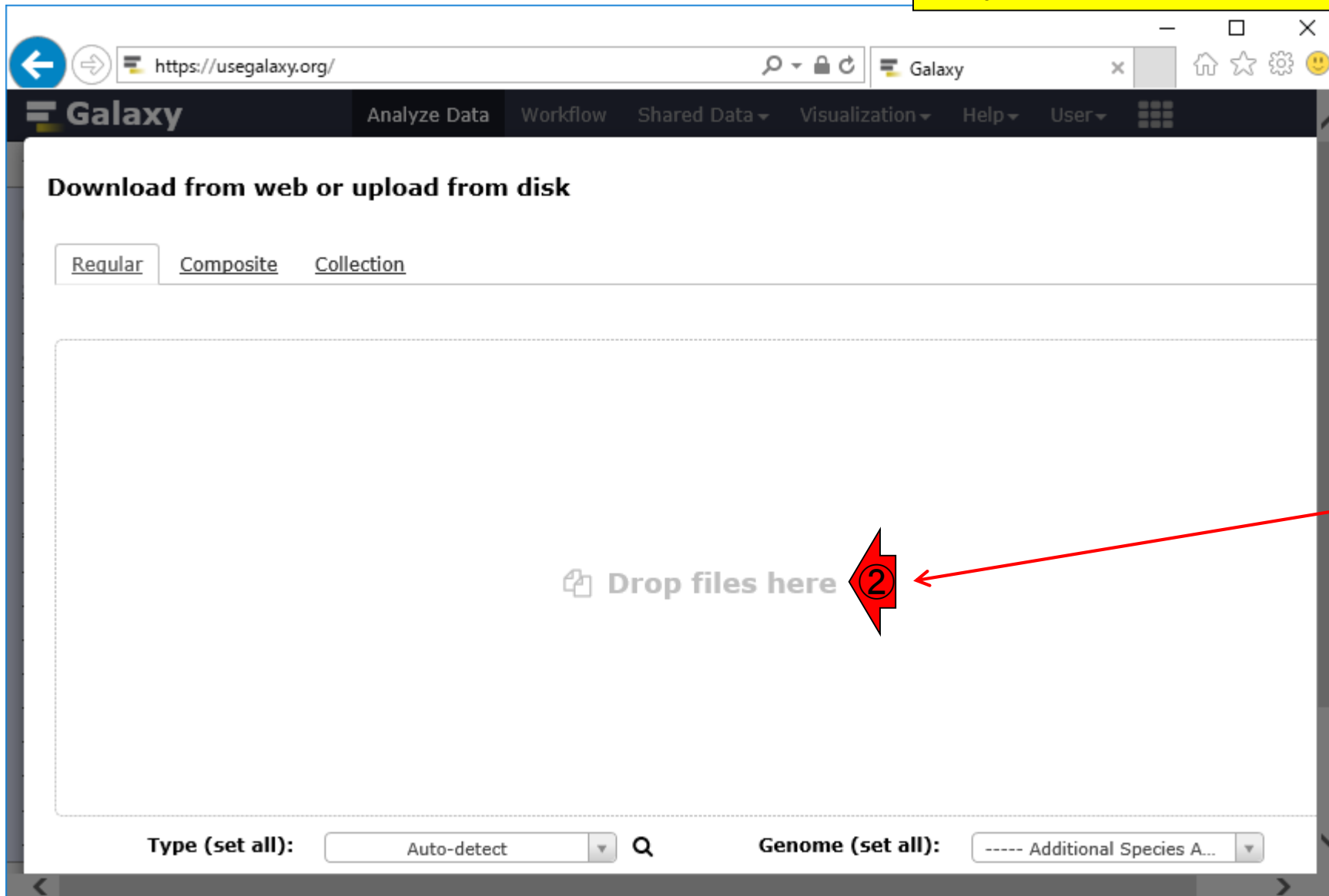
The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists various tool categories such as 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main content area features a 'Try Galaxy on the Cloud' banner with the text 'Now you can have a personal Galaxy within the infinite Universe'. The right sidebar shows the 'History' section, which is currently empty. A red circle with the number '1' and an arrow points to the upload icon (a red square with a white arrow) in the top navigation bar. A tooltip box is visible over the upload icon, containing the text 'Download from URL or upload files from disk'. A blue information box in the History section contains the text: 'ヒストリーは空です。 You can load your own data or get data from an external source'.

W5-3: アップロード2



W5-3: アップロード3

デスクトップ上にある①gzip圧縮FASTQファイル(DDR024501sub_1.fastq.gz)を② Drop files hereのところにドラッグ&ドロップ



W5-3: アップロード4

Galaxy

Analyze Data Workflow Shared Data Visualization Help User

Download from web or upload from disk

Regular Composite Collection

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fastq.gz	56.2 MB	Auto-dete...	----- Additional Sp...		0%

Type (set all): Auto-detect Genome (set all): ----- Additional Species A...

W5-3: アップロード5

追加で他のファイルもアップロードしたい場合は、同様な手順で行うのだろう。ここでは①のファイルのみアップロードしたいので、②Startボタンを押せばいいようだが見当たらない

Galaxy

Analyze Data Workflow Shared Data Visualization Help User

Download from web or upload from disk

Regular Composite Collection

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fastq.gz	56.2 MB	Auto-dete...	----- Additional Sp...	⚙️	0%

Type (set all): Auto-detect Genome (set all): ----- Additional Species A...

W5-3: アップロード6

①ページ下部に移動したら②
Startボタンが見つかったので押す

The screenshot shows the Galaxy web interface. At the top, there are navigation tabs: Regular, Composite, and Collection. Below them, a message states: "You added 1 file(s) to the queue. Add more files or click 'Start' to proceed." A table lists the uploaded file:

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fast q.gz	56.2 MB	Auto-dete...	----- Additional Sp...	⚙️	0%

At the bottom of the interface, there are controls for file type and genome selection, and a row of buttons: Choose local file, Choose FTP file, Paste/Fetch data, Pause, Reset, Start, and Close. A red arrow labeled '1' points to the bottom of the page, and another red arrow labeled '2' points to the 'Start' button.

W5-4: アップロード中...

The screenshot shows the Galaxy web interface. At the top, there is a navigation bar with 'Regular', 'Composite', and 'Collection' tabs. Below this, a message says 'Please wait...1 out of 1 remaining.' A table lists the upload progress for a file named 'DRR024501sub_1.fastq.gz'. The file size is 56.2 MB. The 'Status' column shows a green progress bar at 47%. A red arrow with the number '1' points to the 'Status' column header. At the bottom, there are controls for file selection and upload actions.

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fastq.gz	56.2 MB	Auto-dete...	----- Additional Sp...	⚙	47%

Type (set all): Auto-detect Genome (set all): ----- Additional Species A...

Choose local file Choose FTP file Paste/Fetch data Pause Reset Start Close

W5-5: アップロード完了

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The interface includes tabs for 'Regular', 'Composite', and 'Collection'. A table lists the uploaded file with the following details:

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fast q.gz	56.2 MB	Auto-dete...	---- Additional Sp...	⚙️	100% ✓

A red arrow with the number '1' points to the 'Status' column of the table. Below the table, there are controls for 'Type (set all):' (Auto-detect) and 'Genome (set all):' (---- Additional Species A...). At the bottom, there are buttons for 'Choose local file', 'Choose FTP file', 'Paste/Fetch data', 'Pause', 'Reset', 'Start', and 'Close'. The bottom of the interface shows a breadcrumb trail: 'NGS: VCF Manipulation'.

W5-6: Close

The screenshot shows the Galaxy web interface. At the top, the browser address bar displays <https://usegalaxy.org/>. Below the address bar, there are tabs for 'regular', 'Composite', and 'Collection'. The main content area features a table with the following columns: Name, Size, Type, Genome, Settings, and Status. A single row is highlighted in green, representing the file 'DRR024501sub_1.fastq.gz' with a size of 56.2 MB. The 'Type' is 'Auto-dete...', the 'Genome' is '---- Additional Sp...', and the 'Status' is '100%'. Below the table, there are controls for 'Type (set all):' (Auto-detect) and 'Genome (set all):' (---- Additional Species A...). At the bottom, there are several buttons: 'Choose local file', 'Choose FTP file', 'Paste/Fetch data', 'Pause', 'Reset', 'Start', and 'Close'. A red arrow labeled '1' points to the right side of the interface, and another red arrow labeled '2' points to the 'Close' button.

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fastq.gz	56.2 MB	Auto-dete...	---- Additional Sp...	⚙️	100% ✓

Type (set all): Genome (set all):

W6-1: アップロード後

アップロード後。gzip圧縮状態でアップロードしたオリジナルのファイルサイズは56.1MBであったが、①186.06MBとなっていることがわかる。②gzip非圧縮状態になった(展開された)ようだ

The screenshot shows the Galaxy web interface. The main content area displays a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)."

On the right side, the "History" panel shows a search bar for datasets. Under "Unnamed history", one item is listed: "1:" with a file name "DRR024501sub_1.fastq". The size of this dataset is shown as "186.06 MB". A red arrow labeled "1" points to the size, and another red arrow labeled "2" points to the dataset name.

At the bottom of the main content area, there is a "Want help? Get answers." banner with the Biostars logo and the text "GALAXY EXPLAINED".

W6-2: データを表示

①がデータ表示ボタン(クリック)。②デカすぎるので最初の1MB分しか表示されないようだ

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' section with various categories like 'Get Data', 'Send Data', 'Lift-Over', etc. The main content area displays a large dataset with a warning message: 'This dataset is large and only the first megabyte is shown below.' A red arrow labeled '2' points to this message. Below the message is a preview of FASTQ data. On the right, the 'History' panel shows a search bar and a list of datasets. The first dataset, '1: DRR024501sub_1.fastq', is highlighted in green, and a red arrow labeled '1' points to its 'show' icon.

```
@DRR024501.1 M00278:15:000000000-A2RK1:1:1101:
ATGNATCGAAACAGTATTTACAAGATTTGCATACTGAAATTGAAGC
+DRR024501.1 M00278:15:000000000-A2RK1:1:1101:
???#55<?B<ABBB?BFCAFFHHHHHHFFDGHGHHHHFFHHHHHH
@DRR024501.2 M00278:15:000000000-A2RK1:1:1101:
GTCNGAACACATGAATGGTGAACGGCGCTGAACITTTACGGGACG
+DRR024501.2 M00278:15:000000000-A2RK1:1:1101:
???#55??DD?B?DDDFCCFFHHHHHDHHHHHHHHHHFCGHEHHH
@DRR024501.3 M00278:15:000000000-A2RK1:1:1101:
CAANGATAACAATCATTATCATGAACCTAATGCCGTTCTGGTGAI
+DRR024501.3 M00278:15:000000000-A2RK1:1:1101:
===#55<=@9@@<5<@E/8>88AEEB.AEAF..AC>5>CFFF9CF@
@DRR024501.4 M00278:15:000000000-A2RK1:1:1101:
AGANAGTATTTATCGTAAAATCTTAATTCATTACACAAGGTGGGGCG
+DRR024501.4 M00278:15:000000000-A2RK1:1:1101:
???#5<<?DDD?BDDDBFFFFFFHHHHHHFHHHHHHHHHH/AFHHEH
@DRR024501.5 M00278:15:000000000-A2RK1:1:1101:
AGTNCGGTTAGTGGGAAGCTGCTAGGCATGTGCACACACCATTGTTI
+DRR024501.5 M00278:15:000000000-A2RK1:1:1101:
?,<#55?,<+<5?<<-A>CCC>CC>CE/AECFHHH?E,7CEGHDBE
@DRR024501.6 M00278:15:000000000-A2RK1:1:1101:
TTGNGTTCCTCAACTCTTTGCCAGTTGTAGTGTATAAGTTTACC
```

W6-3: 編集1

①が編集ボタン。②変数を編集する、の意味が分かりづらいかもしれないが、とりあえずは③ファイル名を変更できたり、④このファイルがFASTQ形式ファイルだと正しく認識されているようだ、という程度でいいだろう

The screenshot displays the Galaxy web interface. On the left is a 'Tools' sidebar with categories like 'Get Data', 'Send Data', and 'Convert Formats'. The main area is titled 'Attributes' and 'Convert Format', with a sub-section 'Permissions' and a heading '変数を編集する' (Edit variables). It contains several input fields: 'Name:' with the value 'DRR024501sub_1.fastq', 'Info:' with 'uploaded fastq file', and 'Database/Build:' with a dropdown menu. A 'Save' button is visible. On the right, a 'History' panel shows 'Unnamed history' with '1 shown' and '186.06 MB'. A dataset entry '1: DRR024501sub_1.fastq' is highlighted in green, with an edit icon (a pencil) next to it. Red arrows with circled numbers point to these elements: ① points to the edit icon in the history panel; ② points to the '変数を編集する' heading; ③ points to the 'Name:' input field; ④ points to the 'Info:' input field.

W6-3: 編集2

もしファイル形式を間違っていて認識していたら、①Datatypeのところを手動で修正する

The screenshot shows the Galaxy web interface. The browser address bar is <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow with the number '1' points to the 'Datatype' tab in the 'Convert Format' section. The 'Datatype' tab is active, showing a form to edit dataset variables. The form includes fields for 'Name' (DRR024501sub_1.fastq), 'Info' (uploaded fastq file), and 'Annotation / Notes'. There is also a 'Database/Build' dropdown menu and 'Save' and 'Auto-detect' buttons. The 'History' panel on the right shows a search bar and a list of datasets, with '1: DRR024501sub_1.fastq' highlighted in green.

W6-4: 削除

① ×ボタンで削除することができるようだ。やってみる

The screenshot shows the Galaxy web interface. On the left is a 'Tools' sidebar with various categories like 'Get Data', 'Send Data', and 'NGS: QC and manipulation'. The main area is divided into three panels. The top panel has tabs for 'Attributes', 'Convert Format', and 'Datatype'. The middle panel, titled 'Permissions', contains a form for editing variables with fields for 'Name' (filled with 'DRR024501sub_1.fastq'), 'Info' (filled with 'uploaded fastq file'), and 'Annotation / Notes'. The bottom panel is the 'History' view, showing a list of datasets. The first dataset, '1: DRR024501sub_1.fastq', is highlighted in green. To its right are icons for checkmark, edit, and delete. A red arrow with the number '1' points to the delete icon (an 'X' in a square).

W6-5: 削除されました

①ほんとに消していいかという確認もなしに消えるんですね…。しかし②ディスク容量は不変なので、ヒストリーパネル上からは消えているだけで、実際にはまだ残っているのでしょうか。Windowsの場合はゴミ箱に移動しただけ、と同じような感じなのでしょう

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', and 'Visualization'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', 'NGS: QC and manipulation', 'NGS: DeepTools', 'NGS: Mapping', 'NGS: RNA Analysis', 'NGS: SAMtools', 'NGS: BamTools', 'NGS: Picard', and 'NGS: VCF Manipulation'. The main content area is divided into three panels: 'Attributes', 'Convert Format', and 'Datatype'. The 'Attributes' panel is active, showing 'Permissions' and a section for editing variables. The variable name is 'DRR024501sub_1.fastq' and the info is 'uploaded fastq file'. The 'History' panel on the right shows a search bar and a list of datasets. The top entry is 'Unnamed history' with '1 deleted' and '186.06 MB'. A red arrow labeled '2' points to the '186.06 MB' value. Below this, a blue information box contains the text: 'ヒストリーは空です。 You can load your own data or get data from an external source'. A red arrow labeled '1' points to this message.

W6-7: Saveしてみる

ファイルが存在しないはずだが、①Saveボタンを押してみる

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'Send Data', and 'NGS: QC and manipulation'. The central panel is titled 'Attributes', 'Convert Format', and 'Datatype', with a sub-section 'Permissions' and a heading '変数を編集する' (Edit variable). It contains form fields for 'Name' (filled with 'DRR024501sub_1.fastq'), 'Info' (filled with 'uploaded fastq file'), and 'Annotation / Notes'. Below these is a 'Database/Build' dropdown menu and a 'Save' button, which is highlighted with a red arrow and the number 1. An 'Auto-detect' button is also visible. The right sidebar shows 'History' with a search bar and a message: 'ヒストリーは空です. You can load your own data or get data from an external source'.

W6-8: updated...

①Attributes updatedなので、やはりヒストリーパネル上では消えているだけなのでしょう

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The navigation menu includes "Analyze Data", "Workflow", "Shared Data", "Visualization", "Help", and "User". On the left, a "Tools" sidebar lists various categories like "Get Data", "Send Data", "Text Manipulation", etc. The main content area shows a dataset named "Attributes updated" with a green checkmark and a red arrow pointing to it. Below the dataset name are tabs for "Attributes", "Convert Format", "Datatype", and "Permissions". The "Attributes" tab is active, showing a form to edit variables. The form includes fields for "Name" (DRR024501sub_1.fastq), "Info" (uploaded fastq file), and "Annotation / Notes". A "Database/Build" dropdown menu is set to "Additional Species Are Below". A "Save" button is at the bottom. On the right, the "History" panel shows "Unnamed history" with "1 deleted" and "186.06 MB". A blue information box in the history panel states: "ヒストリーは空です. You can load your own data or get data from an external source".

W7-1: 再アップロード

W5-3で利用した①アップロードボタンではなく、②load your own dataのところを押してもおそらくいいのだろう。クリックしてみる

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains various tool categories like 'Get Data', 'Send Data', 'Lift-Over', etc. The main content area displays a green message 'Attributes updated' and a form for editing dataset attributes. The 'Name' field contains 'DRR024501sub_1.fastq'. The 'Info' field contains 'uploaded fastq file'. The 'History' panel on the right shows 'Unnamed history' with '1 deleted' and '186.06 MB'. A blue message box in the history panel says 'ヒストリーは空です. You can load your own data or get data from an external source'. A red arrow labeled '1' points to the upload button in the top left, and another red arrow labeled '2' points to the 'load your own data' link in the history panel.

W7-2: Drop files here

W5-6と同じような画面になりました。
このままでは追加アップロードできないようなので、一旦①Reset

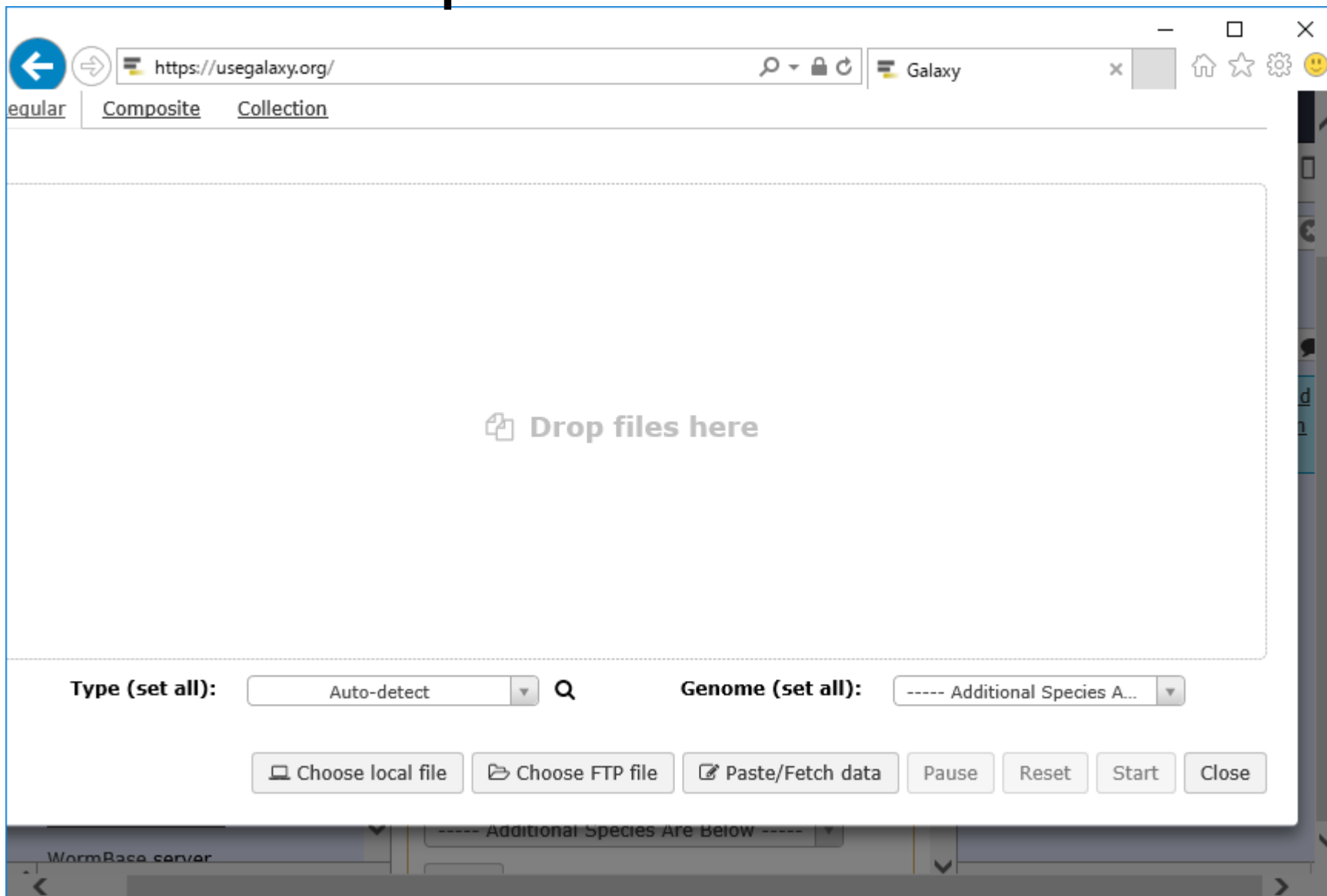
The screenshot shows the Galaxy web interface. At the top, the browser address bar displays 'https://usegalaxy.org/'. Below the address bar, there are tabs for 'regular', 'Composite', and 'Collection'. A table lists the uploaded file 'DRR024501sub_1.fastq.gz' with a size of 56.2 MB, a type of 'Auto-dete...', a genome of '---- Additional Sp...', and a status of '100%'. Below the table, there are controls for 'Type (set all):' and 'Genome (set all):'. At the bottom, there are buttons for 'Choose local file', 'Choose FTP file', 'Paste/Fetch data', 'Pause', 'Reset', 'Start', and 'Close'. A red arrow with the number '1' points to the 'Reset' button.

Name	Size	Type	Genome	Settings	Status
DRR024501sub_1.fastq.gz	56.2 MB	Auto-dete...	---- Additional Sp...	⚙️	100% ✓

Type (set all): Genome (set all):

W7-2: Drop files here

Reset後の状態。この状態から、W5-3と同じ手順でもう一度同じファイルをアップロード



W7-3: アップロード後

① $186.06 \times 2 = 372.12\text{MB}$ と2つ分のファイルサイズになりました。同じファイルですが、上書きではなく別個のものとして認識されているようです

The screenshot shows the Galaxy web interface. The browser address bar displays `https://usegalaxy.org/`. The main navigation bar includes "Galaxy" and menu items: "Analyze Data", "Workflow", "Shared Data", "Visualization", "Help", and "User".

On the left, a "Tools" sidebar is visible with a search bar and various tool categories like "Get Data", "Send Data", "Text Manipulation", etc.

The central panel shows a confirmation message: "Attributes updated" with a green checkmark. Below it are tabs for "Attributes", "Convert Format", "Datatype", and "Permissions". A section titled "変数を編集する" (Edit variables) contains a "Name:" field with the value "DRR024501sub_1.fastq", an "Info:" field with "uploaded fastq file", and an "Annotation / Notes:" field. At the bottom, there is a "Database/Build:" dropdown menu and a "Save" button.

On the right, the "History" panel shows a search bar and a list of datasets under "Unnamed history". The first entry is "372.12 MB" with a red arrow pointing to it and a circled "1". Below it, a dataset named "2: DRR024501sub_1.fastq" is highlighted in green, with a circled "2" next to it.

W7-4: 一旦終了

①を押して一旦終了させてみる。目的は、
②真ん中のパネル表示が再度ログインすると消えるのではないかと想像の確認

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays the 'Attributes updated' message and the 'Edit Dataset' page for 'DRR024501sub_1.fastq'. The 'Workflow' menu item is highlighted with a red arrow labeled '2'. The browser window's close button is also highlighted with a red arrow labeled '1'. The 'History' panel on the right shows the dataset 'DRR024501sub_1.fastq' selected.

W7-5: 再度アクセス

①真ん中のパネル表示が消えています。②右側のヒストリ(履歴)パネルは、 $186.06 \times 2 = 372.12\text{MB}$ と2つ分のファイルサイズになったままです。このことから、W6-4でも行った③×ボタンを押す操作(まだ実際には押さない)は履歴パネルからの表示を消すだけだということを実感

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', and 'Add Data'. A red arrow labeled '1' points to the 'Add Data' dropdown menu. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', etc. The main content area displays a welcome message for Galaxy and a banner for 'GCC 2017 Montpellier' with dates '26 - 30 June France'. On the right, the 'History' panel is visible, showing a search bar and a list of datasets. The first entry is 'Unnamed history' with a size of '372.12 MB'. A red arrow labeled '2' points to this size. Below it, a dataset entry '2: DRR024501sub_1.fastq' is highlighted in green. A red arrow labeled '3' points to the 'x' button in the dataset's action menu.

W7-6: deletedを押してみる

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)."

On the right side, the "History" panel shows a search bar for datasets. Below it, there is an entry for "Unnamed history" with the text "1 shown, 1 deleted". A red arrow with the number "1" points to the "deleted" text. Below this entry, there is a green bar for a dataset named "2: DRR024501sub_1.fastq".

At the bottom of the page, there is a banner for "GCC 2017 Montpellier" with the text "Early registration ends May 15" and "26 - 30 June France".

W7-6: deletedを押してみる

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', etc. The main content area displays a text block about Galaxy: 'Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy start here or consult our help resources. You can install your own Galaxy by following the tutorial and choose from thousands of tools from the Tool Shed.' Below this is a black box with yellow text: 'Looking to learn? New comprehensive tutorials on: Diploid variant calling, Reference based RNAseq, Processing multiple samples, Introduction to NGS technologies, Galaxy 101, parts 1 & 2'. The right sidebar shows the 'History' panel with a search bar and two dataset entries. The top entry is '2: DRR024501sub 1.fastq' and the bottom entry is '1: DRR024501sub 1.fastq'. An orange warning box is overlaid on the top entry, containing the text: 'このデータセットは削除されました', '復元する', and '永久にディスクから削除'. The browser address bar shows 'https://usegalaxy.org/'.

W7-7: 永久に削除

①復元することも、②永久にディスクから削除することもできるのですね。ここでは②削除してみます。③OK

The screenshot shows the Galaxy web interface. The main content area displays a message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a black box with yellow text: "Looking to learn? New comprehensive tutorials on: Diploid variant calling, Reference based RNAseq, Processing multiple samples, Introduction to NGS technologies, Galaxy 101, parts 1 & 2". On the right, the "History" panel shows a dataset "2: DRR024501sub 1.fastq" selected. A confirmation dialog is open, asking "このデータセットは削除されました。復元する (Restore) / 永久にディスクから削除 (Permanently delete from disk)". A red arrow labeled "1" points to the "復元する" option, and another red arrow labeled "2" points to the "永久にディスクから削除" option. A third red arrow labeled "3" points to the "OK" button in the dialog. The dialog also contains the text "Web ページからのメッセージ" and "This will permanently remove the data in your dataset. Are you sure?".

W7-7: 永久に削除

① 永続的に削除されたようです。② ディスク使用量も確かに減りました

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays a welcome message for Galaxy, an open source web-based platform for data intensive biomedical research. Below the main content is a banner for the ISMB/ECCB 2017 Tutorial in Prague, July 21 to July 25, with the text "Making Galaxy work for you" and "Register now".

The History panel on the right shows a list of datasets. The top entry is "Unnamed history" with a size of 186.06 MB. Below it are two datasets, both named "1: DRR024501sub 1.fastg". The first dataset is highlighted in green and has a red arrow labeled "2" pointing to its size. The second dataset is also highlighted in green but has a red arrow labeled "1" pointing to a yellow warning box that says "このデータセットは、永続的にディスクから削除されました" (This dataset has been permanently deleted from the disk).

①消えているので、消すことはできません

W7-8: お前は既に...

The screenshot shows the Galaxy web interface. The main content area displays a banner for "Public Galaxy Servers and still counting" with a "080+" logo. The left sidebar contains a "Tools" panel with a search bar and various tool categories. The right sidebar shows the "History" panel with a search bar and a list of datasets. The top dataset, "2: DRR024501sub_1.fastq", is highlighted in green and has a red arrow pointing to it with the number "1". Below it, a yellow warning box contains the text: "このデータセットは、永続的にディスクから削除されました". A tooltip below the dataset says "Dataset is already deleted". The browser address bar shows "https://usegalaxy.org/".

W7-9: History options

①この歯車マークのところがHistory optionsです

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow with the number '1' points to a gear icon in the History panel header. The History panel on the right shows a search bar, a list of datasets, and a warning message: 'このデータセットは、永続的にディスクから削除されました' (This dataset has been permanently deleted from the disk). The main content area displays a banner for 'GCC 2017 Montpellier' with the text 'Early registration ends May 15' and '26 - 30 June France'.

W7-10: Purge Deleted ...

①Purge Deleted Datasetsは、W7-7で行った「永久にディスクから削除」と同じ作業なのだろうと思いつつ、やってみる。Purgeは、一掃する、の意味

The screenshot shows the Galaxy web interface. The main content area displays a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a logo for "Public Galaxy Servers and still counting" featuring the number "080+" in a stylized font.

On the right side, a dropdown menu is open, showing various actions. The "Purge Deleted Datasets" option is highlighted with a red arrow and a circled "1". Other options in the menu include "Delete Permanently", "Copy Datasets", "Dataset Security", "Resume Paused Jobs", "Collapse Expanded Datasets", "Unhide Hidden Datasets", "Delete Hidden Datasets", "Export Tool Citations", and "Downloads".

The left sidebar contains a "Tools" section with a search bar and various tool categories like "Get Data", "Send Data", "Lift-Over", "Collection Operations", "Text Manipulation", "Datamash", "Convert Formats", "Filter and Sort", "Join, Subtract and Group", "Fetch Alignments/Sequences", "NGS: QC and manipulation", "NGS: DeepTools", "NGS: Mapping", "NGS: RNA Analysis", "NGS: SAMtools", "NGS: BamTools", "NGS: Picard", and "NGS: VCF Manipulation".

The browser address bar shows the URL: https://usegalaxy.org/history/purge_deleted_datasets.

W7-10: Purge Deleted ...

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays a welcome message: "Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#)." Below this is a banner for "Public Galaxy Servers and still counting" with a "080+" logo.

In the top right, a "History" menu is open, showing "HISTORY LISTS" (Saved Histories, Histories Shared with Me) and "CURRENT HISTORY" (Create New). The "Purge Deleted Datasets" option is highlighted in blue.

A confirmation dialog box titled "Web ページからのメッセージ" (Message from the Web Page) is overlaid on the interface. It contains the text: "Really delete all deleted datasets permanently? This cannot be undone." The dialog has two buttons: "OK" and "キャンセル" (Cancel). A red arrow with the number "1" points to the "OK" button.

The browser's address bar shows the URL: https://usegalaxy.org/history/purge_deleted_datasets.

W7-10: Purge Deleted ...

①永続的に削除されたデータセットは0です。妥当ですね

The screenshot shows the Galaxy web interface at <https://usegalaxy.org/>. The main content area displays a green notification box with a checkmark icon and the text "0 datasets have been deleted permanently". A red lightning bolt icon with the number "1" is overlaid on the right side of this notification. The left sidebar contains a "Tools" menu with categories like "Get Data", "Send Data", "Lift-Over", "Collection Operations", "Text Manipulation", "Datamash", "Convert Formats", "Filter and Sort", "Join, Subtract and Group", "Fetch Alignments/Sequences", "NGS: QC and manipulation", "NGS: DeepTools", "NGS: Mapping", "NGS: RNA Analysis", "NGS: SAMtools", "NGS: BamTools", "NGS: Picard", and "NGS: VCF Manipulation". The right sidebar shows the "History" section with a search bar and two dataset entries: "2: DRR024501sub 1.fastq" and "1: DRR024501sub 1.fastq". An orange warning box is present above the second entry, containing the Japanese text: "このデータセットは、永続的にディスクから削除されました".

W8-1: ヒストリー画面

①が消えることを期待して、②のRefresh historyを押してみる

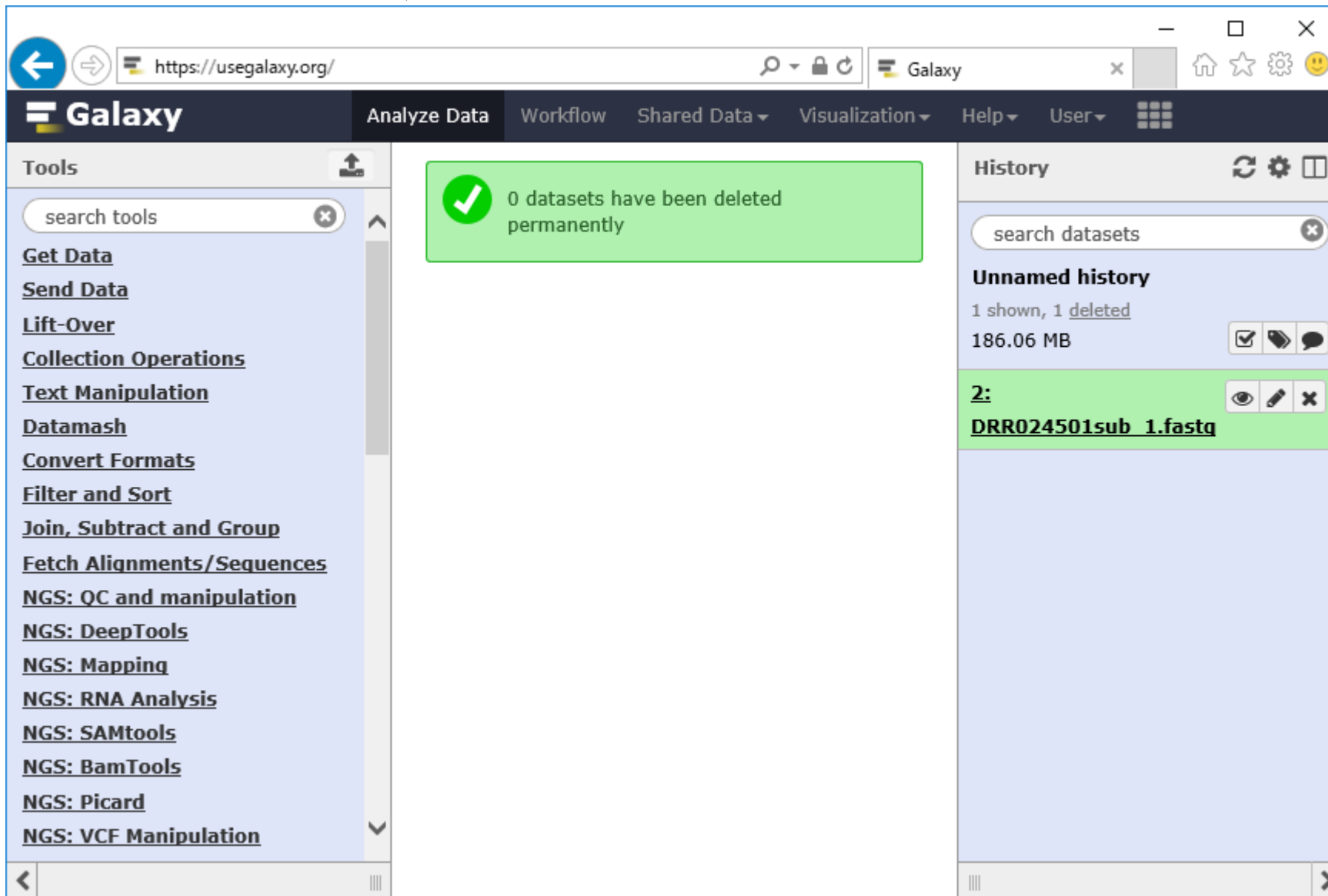
The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow labeled '2' points to the 'Refresh history' button in the History panel. The History panel displays a search bar and a list of datasets. A green notification box at the top states '0 datasets have been deleted permanently'. The dataset list includes 'Unnamed history' (186.06 MB) and two entries for 'DRR024501sub 1.fastq'. The first entry is highlighted in green and has a red arrow labeled '1' pointing to it. A red box highlights the first entry and the warning message above it: 'このデータセットは、永続的にディスクから削除されました'. The second entry is also highlighted in green.

W8-1: ヒストリー画面

変化なし。①hide deletedで、削除されたものを非表示にすることはできません

The screenshot displays the Galaxy web interface. At the top, the browser address bar shows <https://usegalaxy.org/>. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'Send Data', and 'NGS: QC and manipulation'. The central workspace features a green notification box with a checkmark icon and the text '0 datasets have been deleted permanently'. On the right, the 'History' panel is active, showing a search bar and a list of datasets. The top dataset is 'Unnamed history' with '2 shown, [hide deleted](#)' and '186.06 MB'. A red arrow with the number '2' points to the 'hide deleted' link. Below it is a dataset '2: DRR024501sub 1.fastq' with a green background and a warning icon. The warning message reads: 'このデータセットは、永続的にディスクから削除されました'. Below that is another dataset '1: DRR024501sub 1.fastq'.

W8-2: 非表示後



FastQCを用いてクオリティチェックをしてみましょう。①NGS: QC and manipulation

W9-1: クオリティチェック

The screenshot shows the Galaxy web interface. The browser address bar is <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The 'Tools' panel on the left lists various tool categories, with 'NGS: QC and manipulation' highlighted by a red arrow and a circled '1'. A green notification box at the top center displays a checkmark and the text '0 datasets have been deleted permanently'. The 'History' panel on the right shows a search bar for datasets, an 'Unnamed history' section with '1 shown, 1 deleted' and '186.06 MB', and a specific dataset entry '2: DRR024501sub_1.fastq' with view, edit, and delete icons.

W9-2: FastQC

①FastQC。赤枠内のプログラムの表示順は、見る時期によって結構違いますのでご注意ください

The screenshot shows the Galaxy web interface. The browser address bar displays <https://usegalaxy.org/>. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The 'Tools' panel on the left is open, showing a search bar and a list of tool categories. A red circle with the number '1' and an arrow points to the 'New and manipulation' category. A red box highlights the following tools in this category:

- FastQC Read Quality reports
- multiqc aggregate results from bioinformatics analyses into a single report
- Trimmomatic flexible read trimming tool for Illumina NGS data
- Trim Galore! adaptive quality

In the center of the interface, a green notification box states: '0 datasets have been deleted permanently'. The 'History' panel on the right shows an 'Unnamed history' with 1 shown and 1 deleted, totaling 186.06 MB. A dataset named '2: DRR024501sub_1.fastq' is highlighted in green.

W9-3: FastQCの操作画面

① Short read data from your current historyの、②入力ファイルのところに、③アップロードしたファイルが(1つしかない)ので)見えているはず

The screenshot displays the Galaxy web interface for the FastQC tool. The main panel is titled "FastQC Read Quality reports (Galaxy Version 0.69)". It contains several configuration options:

- Short read data from your current history:** A dropdown menu showing "2: DRR024501sub_1.fastq".
- Contaminant list:** A dropdown menu showing "Nothing selected".
- Submodule and Limit specifying file:** A dropdown menu showing "Nothing selected".

A blue "Execute" button is located at the bottom of the configuration panel. Below the configuration, there is a "Purpose" section explaining that FastQC aims to provide a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines.

On the right side, the "History" panel shows "Unnamed history" with "1 shown, 1 deleted" and "186.06 MB". A specific dataset "2: DRR024501sub_1.fastq" is highlighted in green.

W9-4:バージョンの切替え

FastQCのバージョン(2017.07.05現在のデフォルトはver. 0.69)も①Versionsを押すことで切替えられるようです

The screenshot displays the Galaxy web interface for the FastQC tool. The 'Versions' dropdown menu is open, showing options to switch to versions 0.52, 0.63, 0.64, 0.65, 0.67, and 0.68. A red arrow points to the 'Versions' button, which is labeled with a circled '1'. The interface includes a search bar, navigation tabs (Analyze Data, Workflow, Shared Data, Visualization, Help, User), and a history panel on the right showing a dataset named 'DRR024501sub_1.fastq'.

W9-5: オプション

①オプションもいろいろ選べそうですが…よくわからないので見なかったことにする

The screenshot displays the Galaxy web interface for the 'FastQC Read Quality' tool. The 'Options' dropdown menu is open, showing several options: 'Question?', 'Search', 'Share', 'Requirements', and 'See in Tool Shed'. A red arrow points to the 'Options' dropdown. The main content area shows the tool's purpose and execution options.

FastQC Read Quality Versions Options ①

reports (Galaxy Version)

Short read data from [file icon] [folder icon] 2: DRR024501sub_1.fastq

Contaminant list [file icon] [folder icon] Nothing selected

tab delimited file with 2 columns: name and sequence. For example: Illumina Small RNA RT Primer CAAGCAGAAGACGGCATAACGA

Submodule and Limit specifying file [file icon] [folder icon] Nothing selected

a file that specifies which submodules are to be executed (default=all) and also specifies the thresholds for the each submodules warning parameter

Execute

Purpose

FastQC aims to provide a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines. It provides a modular set of analyses which you can use

①の部分が反転されているのでしょうか

W9-6: 1つのファイルなので

The screenshot shows the Galaxy web interface. The main panel displays the configuration for the 'FastQC Read Quality reports (Galaxy Version 0.69)' tool. A red arrow labeled '1' points to the 'Read data from your current history' dropdown menu, which is currently set to '2: DRR024501sub_1.fastq'. The interface includes a left sidebar with tool categories, a central tool configuration panel, and a right sidebar with a history panel showing the selected dataset.

W9-7: 複数のファイルの場合

複数ファイルを同時に実行したい場合は①Multiple datasetsをクリックして指定すればいいのでしょうか

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists tool categories such as 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The central panel displays the 'FastQC Read Quality reports (Galaxy Version 0.69)' tool configuration. A red arrow points to the 'Multiple datasets' button in the 'Show data from your current history' section. The 'Multiple datasets' button is highlighted with a black background. The 'Show data from your current history' section shows a dropdown menu with '2: DRR024501sub_1.fastq' selected. Below this, there is a 'Multiple datasets' button and a dropdown menu with 'Nothing selected'. The 'Submodule and Limit specifying file' section also has a dropdown menu with 'Nothing selected'. The 'Execute' button is visible at the bottom of the configuration panel. The right sidebar shows the 'History' section with a search bar and a list of datasets. The dataset '2: DRR024501sub_1.fastq' is highlighted in green.

とりあえずデフォルトのまま
で実行すべく、①Execute

W9-8: Execute

The screenshot shows the Galaxy web interface. The main panel displays the configuration for the 'FastQC Read Quality reports (Galaxy Version 0.69)' tool. The input is set to '2: DRR024501sub_1.fastq'. The 'Contaminant list' and 'Submodule and Limit specifying file' are both set to 'Nothing selected'. A red arrow with the number '1' points to the 'Execute' button. Below the button, a tooltip reads 'Execute: FastQC (0.65)'. The right-hand panel shows the 'History' section with a search bar and a list of datasets, including '2: DRR024501sub_1.fastq' which is highlighted in green. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'.

W9-9: すぐに画面が

こんな感じで切り替わりました。よくわからないが、①の赤下線部分をよくみると、②右側のヒストリーパネルで③リフレッシュすればいいみたいなことが書いてある

The screenshot shows the Galaxy web interface. On the left is a 'Tools' sidebar with categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main area contains a green notification box with a checkmark icon and the text: '1 job has been successfully added to the queue - resulting in the following datasets: 3: FastQC on data 2: Webpage 4: FastQC on data 2: RawData'. Below this, a red arrow labeled '①' points to a red underlined section: 'You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.' To the right is a 'History' panel with a search bar and a list of datasets. A red arrow labeled '②' points to the top of the history panel, and another red arrow labeled '③' points to the refresh icon in the top right of the history panel. The history list includes '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W9-9: すぐに画面が

①は②で見えているものと同じですね。ちなみに①の部分が灰色のときは実行待ち状態です

The screenshot shows the Galaxy web interface. A green notification box in the center contains a checkmark and the text: "1 job has been successfully added to the queue - resulting in the following datasets:". Below this, a red box highlights two items: "3: FastQC on data 2: Webpage" and "4: FastQC on data 2: RawData". A red arrow labeled "2" points to this box. To the right, the History panel shows a list of datasets. A red box highlights two items: "4: FastQC on data 2: RawData" and "3: FastQC on data 2: Webpage". A red arrow labeled "1" points to this box. The History panel also shows a search bar, "3 shown, 1 deleted", and "186.06 MB". Below the highlighted items, there is a dataset entry "2: DRR024501sub_1.fastq".

W9-10: FastQC実行終了

実行が無事終了すると①の赤枠部分が緑色になります。実行中は黄色、実行に失敗すると赤色になるらしいです。
②ディスク使用量が少し増加しました

The screenshot shows the Galaxy web interface. The main content area displays a green notification box with a checkmark icon, stating: "1 job has been successfully added to the queue - resulting in the following datasets: 3: FastQC on data 2: Webpage 4: FastQC on data 2: RawData". Below this, it provides instructions on how to check the status of queued jobs. The History panel on the right shows a list of datasets, with the two newly created jobs highlighted in green. A red box highlights the two new entries, and a red arrow labeled '1' points to them. Another red arrow labeled '2' points to the 'Unnamed history' section above the new entries.

Tools

search tools

Get Data

Send Data

Lift-Over

Collection Operations

Text Manipulation

Datamash

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

NGS: QC and manipulation

FastQC Read Quality reports

multiqc aggregate results from bioinformatics analyses into a single report

Trimmomatic flexible read trimming tool for Illumina NGS data

Analyze Data Workflow Shared Data Visualization Help User

History

search datasets

Unnamed history

3 shown, 1 hidden

187.46 MB

4: FastQC on data 2: RawData

3: FastQC on data 2: Webpage

2: DRR024501sub_1.fastq

1 job has been successfully added to the queue - resulting in the following datasets:

3: FastQC on data 2: Webpage

4: FastQC on data 2: RawData

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

①RawDataと書いてある
ほうを②データ表示

W10-1: 実行結果を眺める

The screenshot shows the Galaxy web interface. At the top, the browser address bar displays the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fdevte. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. On the left, a 'Tools' sidebar lists various categories like 'Get Data', 'Send Data', and 'NGS: QC and manipulation'. The central area features a green notification box with a checkmark icon, stating: '1 job has been successfully added to the queue - resulting in the following datasets: 3: FastQC on data 2: Webpage 4: FastQC on data 2: RawData'. Below this, it provides instructions on how to check job status. On the right, the 'History' panel shows a list of datasets under 'Unnamed history', including '4: FastQC on data 2: RawData' and '3: FastQC on data 2: Webpage'. A red arrow labeled '1' points to the notification box, and another red arrow labeled '2' points to the 'RawData' entry in the history list.

W10-1: 実行結果を眺め

こんな感じになります。Galaxy Versionは0.69でしたが、内部的に動いているのは、①0.11.5のようですね。中央パネルに表示されているRawDataではわかりにくいので、②Webpageのほうを③データ表示

Galaxy Analyze Data Workflow Shared Data Visualization Help User

Tools

search tools

Get Data

Send Data

Lift-Over

Collection Operations

Text Manipulation

Datamash

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

NGS: QC and manipulation

FastQC Read Quality reports

multiqc aggregate results from bioinformatics analyses into a single report

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim Galore! adaptive quality

```
##FastQC 0.11.5
>>Basic Statistics pass
#Measure Value
Filename DRR024501sub_1_fastq
File type Conventional base calls
Encoding Sanger / Illumina 1.9
Total Sequences 300000
Sequences flagged as poor quality 0
Sequence length 251
%GC 38
>>END_MODULE
>>Per base sequence quality pass
#Base Mean Median Lower Quartile Upper Quartile
1 29.971823333333333 30.0 30.0 30.0
2 30.333043333333332 30.0 30.0 30.0
3 30.46292 30.0 30.0 32.0
4 30.450186666666667 30.0 30.0 30.0
5 30.304846666666666 30.0 30.0 30.0
6 31.86869 33.0 32.0 33.0
7 32.06405 33.0 32.0 33.0
8 32.12523 33.0 33.0 33.0
9 33.91224 35.0 35.0 35.0
10-14 34.13178 35.0 35.0 35.0
15-19 35.74825133333333 36.8 36.8 36.8
20-24 37.345978666666667 38.4 38.4 38.4
25-29 38.30920733333333 39.0 39.0 39.0
30-34 38.31739133333333 39.0 39.0 39.0
35-39 38.076110000000001 39.0 38.8 38.8
40-44 38.083488 39.0 38.4 40.0
45-49 37.871678 39.0 38.0 40.0
50-54 37.668018000000004 39.0 37.0 37.0
```

History

search datasets

Unnamed history

3 shown, 1 deleted

187.46 MB

4: FastQC on data 2: RawData

3: FastQC on data 2: Webpage

2: DRR024501sub_1.fastq

W10-2: htmlレポート

中央パネルが見たことのあるFastQCのhtmlレポートになりました。第6回W4-2と同じような結果です

The screenshot displays the Galaxy web interface. The main panel shows a FastQC report for the dataset 'DRR024501sub_1_fastq', dated 'Wed 5 Jul 2017'. The report summary includes the following items, each with a green checkmark icon:

- [Basic Statistics](#)
- [Per base sequence quality](#)
- [Per tile sequence quality](#)
- [Per sequence quality scores](#)
- [Per base sequence content](#)
- [Per sequence GC content](#)
- [Per base N content](#)
- [Sequence Length Distribution](#)

The right-hand 'History' panel shows a list of datasets:

- 4: FastQC on data 2: RawData
- 3: FastQC on data 2: Webpage
- 2: DRR024501sub_1.fastq

W10-2: htmlレポート

①ちょっとページ下部に移動。②が第6回W4-2では黄色だったが、緑色になっている点が異なる。FastQC ver. 0.11.4(第6回W4-2)と0.11.5(今回)の違いによるのかもしれない

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main content area is divided into three panels: 'Tools', 'Summary', and 'History'. The 'Summary' panel displays a list of metrics with green checkmarks indicating successful completion. The 'History' panel shows a list of datasets, with the top one being '4: FastQC on data 2: RawData'. A red arrow labeled '1' points to the 'History' panel, and another red arrow labeled '2' points to the 'Per base sequence content' metric in the 'Summary' panel.

Tools

search tools

Get Data

Send Data

Lift-Over

Collection Operations

Text Manipulation

Datamash

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

NGS: QC and manipulation

FastQC Read Quality reports

multiqc aggregate results from bioinformatics analyses into a single report

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim Galore! adaptive quality

Summary

- ✓ [Basic Statistics](#)
- ✓ [Per base sequence quality](#)
- ✓ [Per tile sequence quality](#)
- ✓ [Per sequence quality scores](#)
- ✓ [Per base sequence content](#)
- ✓ [Per sequence GC content](#)
- ✓ [Per base N content](#)
- ✓ [Sequence Length Distribution](#)
- ✓ [Sequence Duplication Levels](#)
- ✓ [Overrepresented sequences](#)
- ! [Adapter Content](#)
- ! [Kmer Content](#)

History

search datasets

Unnamed history

3 shown, 1 deleted

187.46 MB

4: FastQC on data 2: RawData

3: FastQC on data 2: Webpage

2: DRR024501sub_1.fastq

W10-3: Basic Statistics

The screenshot shows the Galaxy web interface. The browser address bar displays https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fdevte. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The central 'Summary' panel lists various quality control metrics, each with a green checkmark icon, except for 'Adapter Content' and 'Kmer Content' which have orange warning icons. A red arrow with a circled '1' points to the 'Basic Statistics' link. The right sidebar shows the 'History' panel with a search bar and a list of datasets, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

Galaxy

Analyze Data Workflow Shared Data Visualization Help User

Tools

search tools

Get Data

Send Data

Lift-Over

Collection Operations

Text Manipulation

Datamash

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

NGS: QC and manipulation

FastQC Read Quality reports

multiqc aggregate results from bioinformatics analyses into a single report

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim Galore! adaptive quality

Summary

- ✓ [Basic Statistics](#)
- ✓ [Per base sequence quality](#)
- ✓ [Per tile sequence quality](#)
- ✓ [Per sequence quality scores](#)
- ✓ [Per base sequence content](#)
- ✓ [Per sequence GC content](#)
- ✓ [Per base N content](#)
- ✓ [Sequence Length Distribution](#)
- ✓ [Sequence Duplication Levels](#)
- ✓ [Overrepresented sequences](#)
- ! [Adapter Content](#)
- ! [Kmer Content](#)

History

search datasets

Unnamed history

3 shown, 1 deleted

187.46 MB

4: FastQC on data 2: RawData

3: FastQC on data 2: Webpage

2: DRR024501sub_1.fastq

W10-3: Basic Statistics

①このあたりの結果は第6回W4-2と同じですね。②同じFASTQ形式でもいくつかの亜流があり、これはSanger/Illumina 1.9というもの

The screenshot shows the Galaxy web interface. The main panel displays the results of the 'Basic Statistics' tool. A table lists various measures and their values for the file 'DRR024501sub_1_fastq'. Below the table is a 'Per base sequence quality' plot showing quality scores across the sequence. The right-hand panel shows the 'History' of the tool, listing previous runs.

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

Per base sequence quality plot: Quality scores (Y-axis, 34-40) vs. Position (X-axis, 1-251). The plot shows a blue line representing the mean quality score and yellow bars representing individual quality scores for each base. The quality score starts around 34 and rises to a peak of approximately 38.5 in the middle of the sequence.

W10-4: 結果のダウンロード

The screenshot shows the Galaxy web interface. The main content area displays 'Basic Statistics' for a file named 'DRR024501sub_1_fastq'. Below this, there is a 'Per base sequence quality' plot showing quality scores across the sequence. On the right side, the 'History' panel shows a list of datasets, with '4: FastQC on data 2: Raw' highlighted in green and a red arrow pointing to the 'Raw' link. The browser address bar shows the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fdevte.

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

W10-4: 結果のダウンロード

こんな感じになります。①フロッピーディスクアイコンをクリックすれば、②293.4 KBの③htmlファイルをダウンロードできます。①をクリックして実際にダウンロードします

The screenshot displays the Galaxy web interface. The main content area shows 'Basic Statistics' for a file named 'DRR024501sub_1_fastq'. The statistics table is as follows:

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

Below the statistics is a 'Per base sequence quality' plot showing quality scores across the sequence. The y-axis ranges from 34 to 40. The plot shows a blue line representing the mean quality score and yellow bars representing the distribution of quality scores at each position.

The right-hand 'History' panel shows a list of jobs. The most recent job is '4: FastQC on data 2:'. Below it is a 'RawData' section for job '3: FastQC on data 2: Webpage', which shows a file size of 293.4 KB and a format of 'html'. At the bottom of the history panel, there is a download section for an 'HTML file'.

Red arrows with numbers 1, 2, and 3 point to the following elements:

- ①: The floppy disk icon in the download section of the history panel.
- ②: The file size '293.4 KB' in the 'RawData' section.
- ③: The file name 'HTML file' in the download section.

W10-4: 結果のダウンロード

The screenshot shows the Galaxy web interface with the following components:

- Tools Panel:** Lists various tool categories such as 'Get Data', 'Send Data', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. 'FastQC Read Quality reports' is selected under the NGS category.
- Basic Statistics:** A table showing file and sequence information.

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38
- Per base sequence quality:** A bar chart showing quality scores across the sequence. A red arrow labeled '①' points to a download button in this section.
- History Panel:** Shows a list of jobs, including '4: FastQC on data 2: RawData' and '3: FastQC on data 2: Webpage'.
- Dialog Box:** A yellow-bordered box at the bottom asks: 'usegalaxy.org から FastQC_on_data_2_Webpage.zip (940 KB) を開くか、または保存しますか?' (Do you want to open or save FastQC_on_data_2_Webpage.zip (940 KB) from usegalaxy.org?). Buttons for 'ファイルを開く(O)', '保存(S)', and 'キャンセル(C)' are visible.

W10-5: 解凍すると...

①のような感じになり、②をダブルクリックすると第6回W4-2とほぼ同じ感じになります

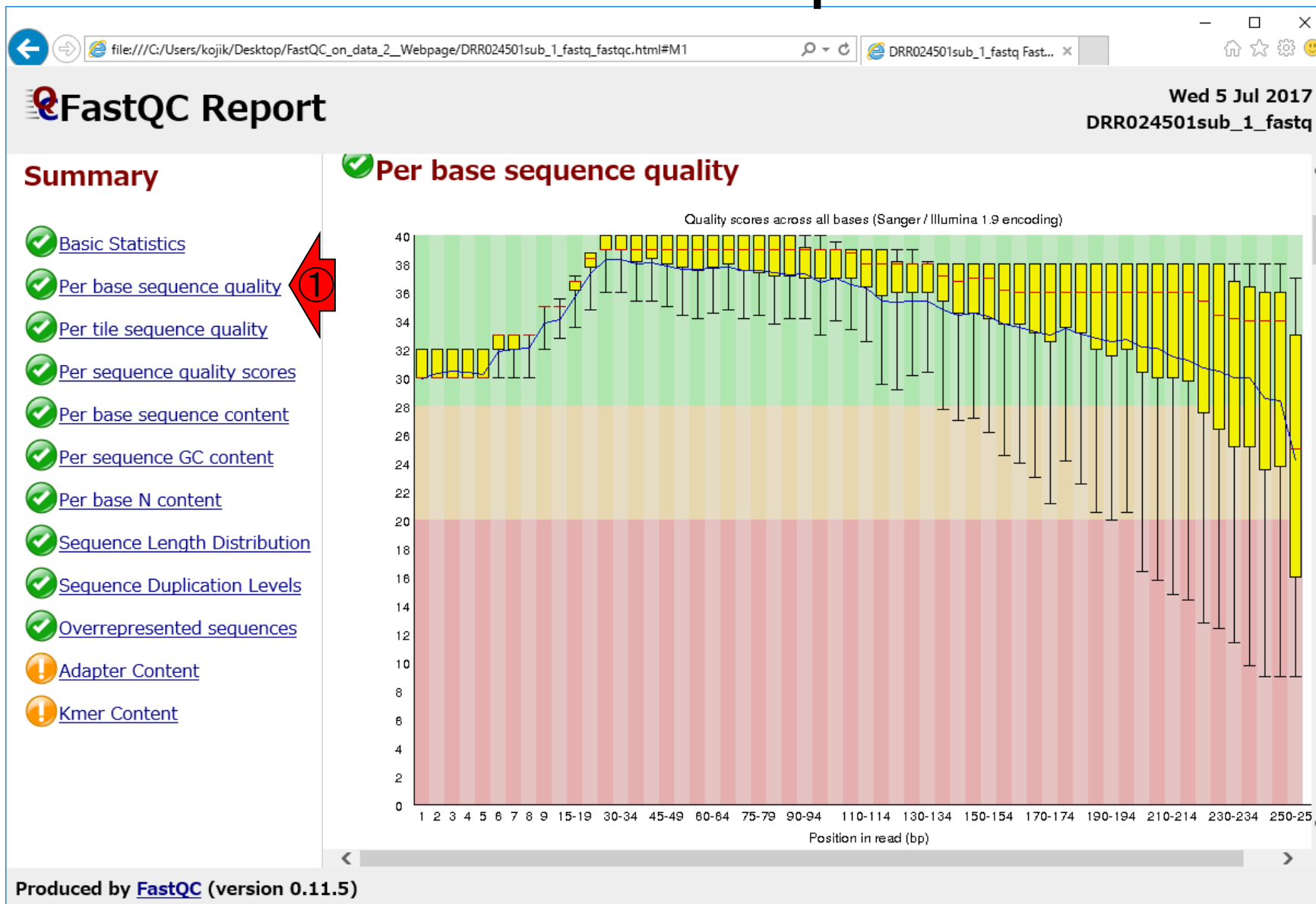
The image shows a Windows File Explorer window and a web browser displaying a FastQC report. The File Explorer window is titled "FastQC_on_data_2_Webpage" and shows a folder containing several files. A red arrow labeled "1" points to the address bar of the File Explorer, which shows the path "C:\Users\Kojik\Desktop\FastQC_on_data_2_Webpage". Another red arrow labeled "2" points to the file "DRR024501sub_1_fastq_fastqc.html" in the file list. The web browser window shows the FastQC report for the file "DRR024501sub_1_fastq_fastqc.html". The report includes a Summary section with a list of metrics, all of which are marked with a green checkmark, indicating that the data is of good quality. The Basic Statistics section includes a table with the following data:

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

The report also includes a section for "Per base sequence quality" which contains a bar chart showing quality scores across all bases. The chart shows a peak in quality around the middle of the sequence, with scores generally between 34 and 40. The report is produced by FastQC (version 0.11.5).

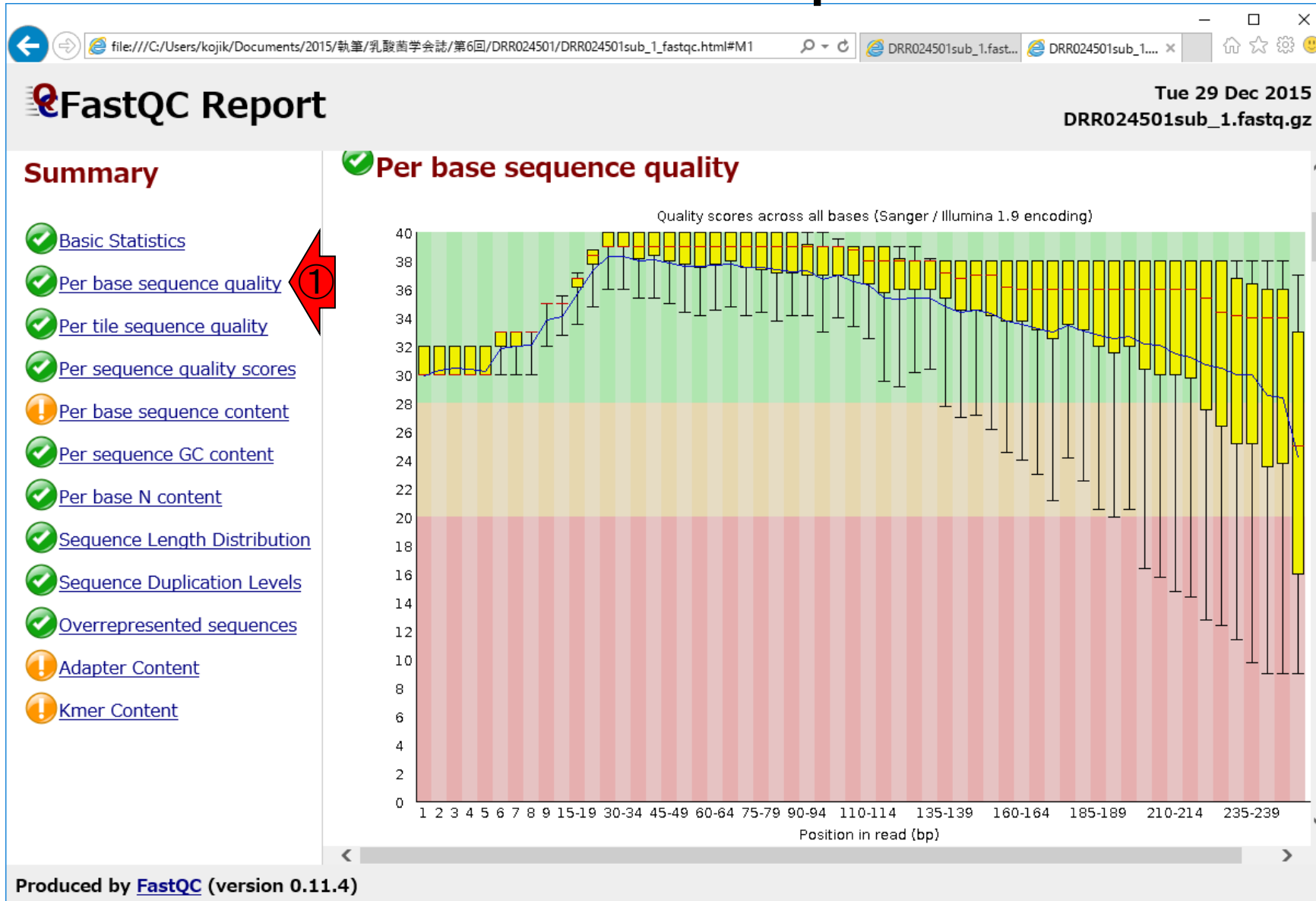
W10-6: Per base sequence...

① Per base sequence quality。
第6回W4-2とほぼ同じですね



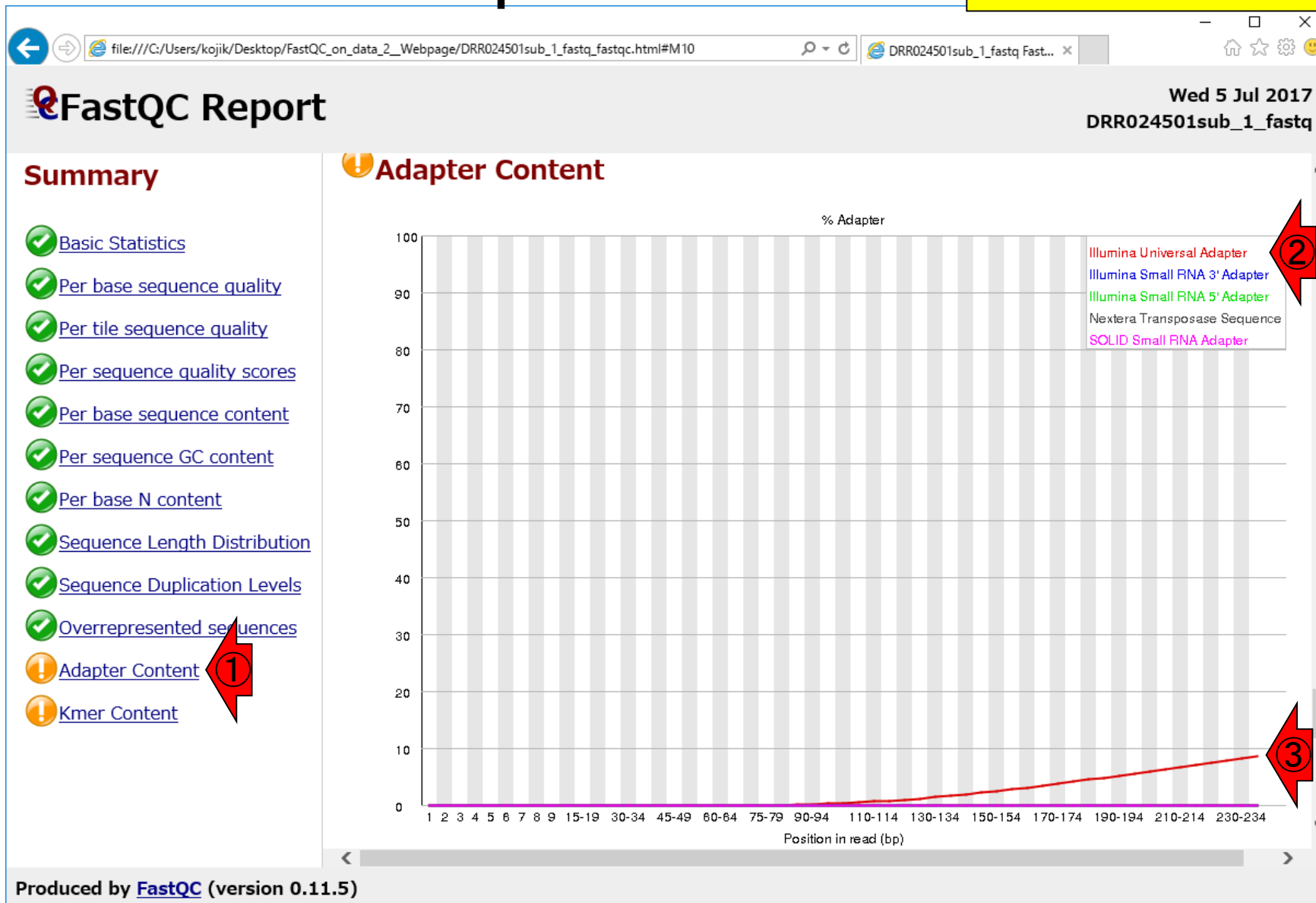
①第6回のhtmlレポートの結果です。同じであることがわかります

W10-7: Per base sequence...



W11-1: Adapter Content

① Adapter Contentの結果として、② Illumina Universal Adapterが③含まれていたことを思い出そう(第6回W4-4)



W11-2: Trimmomatic

Basic Statistics

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

Per base sequence quality

Quality scores

History

- 4: FastQC on data 2:
- 3: FastQC on data 2: Webpage
- 2: Webpage

293.4 KB
フォーマット: html,
データベース: ?

Picked up _JAVA_OPTIONS:
-Djava.io.tmpdir=/galaxy-repl/main/jobdir/016/260/162604
-Xmx7680m -Xms256m

HTML file

W11-2: Trimmomatic

①FastQCと同じ項目内にある、② Trimmomaticを使ってアダプターを除去してみる。③が入力ファイル

The screenshot shows the Galaxy web interface. The 'Tools' sidebar on the left has a search bar and several categories. A red arrow labeled '1' points to the 'NGS: QC and manipulation' category. A red arrow labeled '2' points to the 'Trimmomatic flexible read trimming tool for Illumina NGS data' option. The main panel displays the 'Basic Statistics' tool output for the file 'DRR024501sub_1_fastq'. Below this is the 'Per base sequence quality' tool output, which includes a box plot of quality scores across the sequence length. A red arrow labeled '3' points to the '2: DRR024501sub_1.fastq' entry in the 'History' panel on the right.

Measure	Value
Filename	DRR024501sub_1_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	300000
Sequences flagged as poor quality	0
Sequence length	251
%GC	38

Quality scores

40
38
36
34

W11-3: 入力ファイル形式

①Input FASTQ fileのところ、②No fastqsanger or …となり、③入力ファイルを認識させることができません

The screenshot shows the Galaxy web interface with the Trimmomatic flexible tool configuration. The tool is set to process single-end reads. The input FASTQ file field is empty, and the text "No fastqsanger or fast..." is displayed. The "Perform initial ILLUMINACLIP step?" option is set to "No". The Trimmomatic Operation is set to "1: Trimmomatic Operation" with "Sliding window trimming (SLIDINGWIND...)" selected. The number of bases to average across is 4, and the average quality required is 20. The History panel shows three datasets: "4: FastQC on data 2: RawData", "3: FastQC on data 2: Webpage", and "2: DRR024501sub_1.fastq". Red arrows point to the input FASTQ file field (1), the "No fastqsanger or fast..." text (2), and the dataset "DRR024501sub_1.fastq" (3).

W11-3: 入力ファイル形式

The screenshot shows the Galaxy web interface with the 'Trimmomatic flexible' tool selected. The tool's configuration page is displayed, showing options for 'Single-end or paired-end reads?' (set to 'Single-end') and 'Input Format' (set to 'No fastqsanger or fast...'). A red arrow with the number 1 points to the 'Input Format' dropdown menu. Below the dropdown, there is a note: 'This is a batch mode input field. Separate jobs will be triggered for each dataset selection.' The 'Perform initial ILLUMINACLIP step?' is set to 'No'. The 'Trimmomatic Operation' section shows '1: Trimmomatic Operation' selected, with 'Sliding window trimming (SLIDINGWIND...)' chosen for the operation and 'Number of bases to average across' set to 4. The right sidebar shows the 'History' panel with 'Unnamed history' containing 3 datasets, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W11-3: 入力ファイル形式

認識されない原因は、①のファイル形式が②fastqsangerではないため。③でファイル形式を明示的にfastqsangerに変更する必要がある。③を押す

The screenshot displays the Galaxy web interface for the Trimmomatic flexible tool. The main configuration area shows the following details:

- Tool Name:** Trimmomatic flexible (Galaxy Version 0.36.3)
- Read Type:** Single-end or paired-end reads? (Single-end)
- Input FASTQ file:** No fastqsanger or fast... (This is a batch mode input field. Separate jobs will be triggered for each dataset selection.)
- Perform initial ILLUMINACLIP step?:** Yes (selected), No
- Trimmomatic Operation:** 1: Trimmomatic Operation, Sliding window trimming (SLIDINGWIND...)
- Number of bases to average across:** 4

The History panel on the right shows a list of datasets:

- 4: FastQC on data 2: RawData (highlighted in green, with a red arrow and circled '3' pointing to the 'x' icon)
- 3: FastQC on data 2: Webpage
- 2: DRR024501sub 1.fastq (with a red arrow and circled '1' pointing to the filename)

W11-4: fastqsanger

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow with the number '1' points to the 'Datatype' tab in the 'Attributes', 'Convert Format', and 'Datatype' menu. The 'Datatype' tab is active, showing a form to edit dataset variables. The form includes fields for 'Name' (DRR024501sub_1.fastq), 'Info' (uploaded fastq file), and 'Annotation / Notes'. Below the form are 'Save' and 'Auto-detect' buttons. The 'History' panel on the right shows a list of datasets, with '2: DRR024501sub_1.fastq' selected. The 'Tools' panel on the left is visible, showing various tool categories like 'Get Data', 'Send Data', and 'NGS: QC and manipulation'.

W11-4: fastqsanger

①デフォルトがfastqになっているので、これをfastqsangerに変える

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists various tool categories like 'Get Data', 'Send Data', 'Collection Operations', etc. The main content area is titled 'Datatype' and shows a dialog for changing the data type. The dialog has a title 'データタイプを変更する' and a dropdown menu with 'fastq' selected. A red circle with the number '1' and an arrow points to the dropdown. Below the dropdown is Japanese text: '新しいタイプ: fastq. これは既存のデータセットのデータタイプを変更します。しかしデータセットの中身は変更しません。データセットのタイプの誤判定があったときに使用します。' and a 'Save' button. The right sidebar shows the 'History' panel with a search bar and a list of datasets, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W11-4: fastqsanger

The screenshot shows the Galaxy web interface. The main content area displays a configuration window for the 'fastqsanger' tool. The window has three tabs: 'Attributes', 'Convert Format', and 'Datatype'. The 'Datatype' tab is active, showing a dropdown menu for '新しいタイプ:' (New type) with 'fastq' selected. A red arrow with the number '1' points to the 'fastq' selection. The background shows the Galaxy interface with a 'History' panel on the right listing datasets like '4: FastQC on data 2: RawData'.

W11-4: fastqsanger

The screenshot shows the Galaxy web interface. The main content area is titled 'Attributes Convert Format Datatype Permissions'. Under the 'Datatype' tab, there is a section 'データタイプを変更する' (Change data type) with a dropdown menu '新しいタイプ:' (New type:). The dropdown menu is open, showing a list of data types: fastq, epestrina, eps, equicktandem, est2genome, etandem, excel, fai, fasta, and fastq. A red arrow with the number '1' points to the 'fastq' option in the dropdown. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The right sidebar shows a 'History' panel with a search bar and a list of history items, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W11-5: fastqcssangerではない

The screenshot shows the Galaxy web interface. The main content area is titled "Attributes" and "Convert Format". A dropdown menu is open, showing a list of data types. The current selected type is "fastq". The dropdown list includes:

- fastq
- fastq
- fastq.bz2
- fastq.gz
- fastqcssanger** (highlighted with a red arrow and a circled '1')
- fastqcssanger.bz2
- fastqcssanger.gz
- fastqillumina
- fastqillumina.bz2

The right sidebar shows the "History" panel with a search bar and a list of datasets. The top dataset is "4: FastQC on data 2: RawData" (187.46 MB). Below it are "3: FastQC on data 2: Webpage" and "2: DRR024501sub_1.fastq".

①fastqsangerがあるので、それを選択

W11-6: fastqsanger

The screenshot shows the Galaxy web interface. The browser address bar displays the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fpjbriiggs%2Ffastqsanger. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories such as 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The 'Convert Formats' tab is active, showing a dropdown menu for '新しいタイプ:' (New type:). The dropdown list includes 'fastq', 'fastqillumina', 'fastqillumina.bz2', 'fastqillumina.gz', 'fastqsanger' (highlighted in blue with a red arrow and a circled '1'), 'fastqsanger.bz2', 'fastqsanger.gz', 'fastqsolexa', 'fastqsolexa.bz2', and 'fastqsolexa.gz'. The right sidebar shows the 'History' panel with a search bar and a list of datasets, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W11-7: Save

①fastqsangerになっていることを確認して、②Save

The screenshot shows the Galaxy web interface. The browser address bar displays the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fpjbriggs%2F. The main navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists various tool categories such as 'Get Data', 'Send Data', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The central panel is titled 'Attributes Convert Format Datatype Permissions' and shows a 'Datatype' tab. The main content area is titled 'データタイプを変更する' (Change Data Type). It features a dropdown menu labeled '新しいタイプ:' (New Type) with 'fastqsanger' selected. Below the dropdown is a text box containing the Japanese text: 'これは既存のデータセットのデータタイプを変更します。しかしデータセットの中身は変更しません。データセットのタイプの誤判定があったときに使用します。' (This changes the data type of the existing dataset. However, the contents of the dataset are not changed. Use this when there is a misclassification of the dataset type). A 'Save' button is located at the bottom of the form. Two red arrows with circled numbers point to the 'fastqsanger' dropdown (labeled ①) and the 'Save' button (labeled ②). The right sidebar shows a 'History' panel with a search bar and a list of datasets, including '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W11-8: Save中...

① Save実行中…。この状態が10分以上続くこともあります。以前やったときは1時間以上かかりました

The screenshot shows the Galaxy web interface. At the top, a green notification bar states: "Changed the type of dataset 'DRR024501sub_1.fastq' to fastqsanger". Below this, the dataset details are displayed under the "Attributes" tab. The "Name" field contains "DRR024501sub_1.fastq" and the "Info" field contains "uploaded fastq file". A red arrow with the number "1" points to the "Save" button at the bottom of the dataset details panel. The right sidebar shows a "History" panel with a search bar and a list of datasets. The current dataset, "2: DRR024501sub_1.fastq", is highlighted in yellow. The "Tools" panel on the left is visible, showing various tool categories like "Get Data", "Send Data", and "Text Manipulation".

W11-9: Save終了

The screenshot shows the Galaxy web interface. At the top, a navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A green notification box at the top center states: 'Changed the type of dataset 'DRR024501sub_1.fastq' to fastqsanger'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Text Manipulation', etc. The main content area is divided into tabs: 'Attributes', 'Convert Format', and 'Datatype'. The 'Attributes' tab is active, showing fields for 'Name' (DRR024501sub_1.fastq), 'Info' (uploaded fastq file), and 'Annotation / Notes'. A red arrow labeled '1' points to the 'Save' button at the bottom. The right sidebar shows a 'History' panel with a search bar and a list of datasets. The top entry is '2: DRR024501sub_1.fastq', which is highlighted in green. A red arrow labeled '2' points to the 'Trimmomatic' tool in the left sidebar.

W12-1: Trimmomatic

①の入力ファイルが、②TrimmomaticのInput FASTQ file上で見られるようになりました

The screenshot displays the Galaxy web interface for the Trimmomatic flexible tool. The tool configuration is as follows:

- Trimmomatic flexible** (Version 0.36.3)
- Single-end or paired-end reads?**: Single-end
- Input FASTQ file**: 2: DRR024501sub_1.fastq
- Perform initial ILLUMINACLIP step?**: No
- Trimmomatic Operation**: 1: Trimmomatic Operation
- Select Trimmomatic operation to perform**: Sliding window trimming (SLIDINGWINDOW)
- Number of bases to average across**: 4
- Average quality required**: 20

The right sidebar shows the history with the following entries:

- 4: FastQC on data 2: RawData
- 3: FastQC on data 2: Webpage
- 2: DRR024501sub_1.fastq

Red arrows labeled ① and ② point to the input FASTQ file field and the selected history entry, respectively.

W12-1: Trimmomatic

The screenshot shows the Galaxy web interface for the Trimmomatic flexible tool. The tool configuration is as follows:

- Trimmomatic flexible** (Version 0.36.3)
- Single-end or paired-end reads?**: Single-end
- Input FASTQ file**: 2: DRR024501sub_1.fastq
- Perform initial ILLUMINACLIP step?**: Yes (highlighted with a red arrow and circled '1')
- Trimmomatic Operation**: 1: Trimmomatic Operation
- Select Trimmomatic operation to perform**: Sliding window trimming (SLIDINGWINDOW)
- Number of bases to average across**: 4
- Average quality required**: 20

The right sidebar shows the History panel with the following entries:

- 4: FastQC on data 2: RawData
- 3: FastQC on data 2: Webpage
- 2: DRR024501sub_1.fastq

W12-2: アダプター配列を指定

The screenshot displays the Galaxy web interface for configuring the Trimmomatic flexible tool. The tool is set to process single-end reads from the input FASTQ file '2: DRR024501sub_1.fastq'. The 'Perform initial ILLUMINACLIP step?' is set to 'Yes'. The 'Adapter sequences to use' dropdown menu is highlighted with a red arrow and the number 1, indicating the step to specify the adapter sequence. The selected adapter sequence is 'TruSeq2 (single-ended, for Illumina GAII)'. The 'Maximum mismatch count which will still allow a full match to be performed' is set to 2, and the 'How accurate the match between the two 'adapter ligated' reads must be for PE palindrome read alignment' is set to 30.

W12-2: アダプター配列を指定

The screenshot displays the Galaxy web interface for configuring the Trimmomatic flexible tool. The browser address bar shows the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fpjbbriggs%2Ftrimmomatic. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Help, and User. The left sidebar lists various tool categories such as Get Data, Send Data, Lift-Over, Collection Operations, Text Manipulation, Datamash, Convert Formats, Filter and Sort, Join, Subtract and Group, Fetch Alignments/Sequences, and NGS: QC and manipulation. The central tool configuration area is titled "Trimmomatic flexible read trimming tool for Illumina NGS data (Galaxy Version 0.36.3)". It features a dropdown menu for "Single-end or paired-end reads?" with a search bar. The selected option is "TruSeq3 (single-ended, for MiSeq and HiSeq)", which is highlighted in blue and pointed to by a red arrow labeled "1". Other options in the dropdown include TruSeq2 (single-ended, for Illumina GAII), TruSeq2 (paired-ended, for Illumina GAII), TruSeq3 (paired-ended, for MiSeq and HiSeq), TruSeq3 (additional seqs) (paired-ended, for MiSeq and HiSeq), and Nextera (paired-ended). Below the dropdown, there is a text input field for "Maximum mismatch count which will still allow a full match to be performed" with the value "2". At the bottom, there is another text input field for "How accurate the match between the two 'adapter ligated' reads must be for PE palindrome read alignment" with the value "30". The right sidebar shows a "History" panel with a search bar and a list of datasets. The current dataset is "4: FastQC on data 2: RawData", which is highlighted in green. Other datasets in the history include "3: FastQC on data 2: Webpage" and "2: DRR024501sub_1.fastq".

W12-3: ページ最下部へ

①無事TruSeq3になっていることを確認して、②ページ最下部に移動

The screenshot displays the Galaxy web interface for the Trimmomatic flexible tool. The browser address bar shows the URL: https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fpjbri.... The page title is "Galaxy". The main content area is titled "Trimmomatic flexible read trimming tool for Illumina NGS data (Galaxy Version 0.36.3)". It includes sections for "Single-end or paired-end reads?" (set to Single-end), "Input FASTQ file" (2: DRR024501sub_1.fastq), "Perform initial ILLUMINACLIP step?" (Yes), "Adapter sequences to use" (TruSeq3 (single-ended, for MiSeq and HiSeq)), "Maximum mismatch count which will still allow a full match to be performed" (2), and "How accurate the match between the two 'adapter ligated' reads must be for PE palindrome read alignment" (30). The right sidebar shows the "History" panel with "Unnamed history" (3 shown, 1 deleted, 187.46 MB) and a list of datasets: "4: FastQC on data 2: RawData", "3: FastQC on data 2: Webpage", and "2: DRR024501sub_1.fastq". Two red arrows with circled numbers point to the "TruSeq3" dropdown menu (labeled 1) and the "History" panel (labeled 2).

W12-4: Citations

①ページ最下部。②この場合の引用文献は、GalaxyとTrimmomaticです。お忘れなきよう

The screenshot shows the Galaxy web interface. The main content area displays the Trimmomatic tool page, which includes a description of the tool and a list of citations. A red box highlights a notice that says: "Please kindly acknowledge both this Galaxy tool and the Trimmomatic program if you use it." A red arrow labeled "2" points to this notice. Below the notice is a section titled "Citations" with a "Show BibTeX" button. A red arrow labeled "1" points to the citation list at the bottom of the page, which includes the following text: "Bolger, A. M. and Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. In *Bioinformatics*, 30 (15), pp. 2114-2120. [doi:10.1093/bioinformatics/btu170][Link]"

W12-5: Trimmomatic実行

The screenshot shows the Galaxy web interface for running the Trimmomatic tool. The main panel is titled "Select Trimmomatic operation to perform" and contains the following configuration options:

- Operation: Sliding window trimming (SLIDINGWINDOW)
- Number of bases to average across: 4
- Average quality required: 20

Below the configuration, there are two buttons: "+ Insert Trimmomatic Operation" (indicated by a red arrow labeled 1) and "Execute" (indicated by a red arrow labeled 2).

The "History" panel on the right shows a list of datasets, including "4: FastQC on data 2: rawData", "3: FastQC on data 2: Webpage", and "2: DRR024501sub_1.fastq".

The "What it does" section provides a description of the tool's function and lists the trimming steps it performs:

- ILLUMINA CLIP:** Cut adapter and other illumina-specific sequences from the read
- SLIDING WINDOW:** Perform a sliding window trimming, cutting once the average quality within the window falls below a threshold
- MINLEN:** Drop the read if it is below a specified length

W12-6: 実行待ち状態

The screenshot shows the Galaxy web interface. On the left is a 'Tools' sidebar with categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main area contains a green notification box with a checkmark icon and the text: '1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq'. Below this, it explains that job status can be checked in the History pane. A red arrow with the number '1' points from the notification to the History pane. The History pane on the right shows a list of datasets: '5: Trimmomatic on DRR024501sub_1.fastq' (highlighted in grey), '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'.

W12-7: 実行中

- ①約1分後にリロードしてみると、
- ②実行中(黄色)になっていました

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow labeled '1' points to the user profile icon in the top right. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main content area features a green notification box with a checkmark icon, stating: '1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq'. Below this, it explains that job status can be checked in the History pane. A red arrow labeled '2' points from the notification box to the History pane. The History pane on the right shows a list of jobs: '5: Trimmomatic on DRR024501sub_1.fastq' (highlighted in yellow), '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'. The top job is marked with a gear icon, indicating it is running.

W12-8: 実行終了

①さらに約1分後にリロードしてみると、②実行終了(緑色)になっていました。定期的なリロードによるチェックは重要かもしれません。③ディスクサイズも増えてますね

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. A red arrow labeled '1' points to the top right corner. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main content area features a green notification box with a checkmark icon, stating: '1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq'. Below this, it explains that job status can be checked in the History pane. A red arrow labeled '2' points to this notification. The right sidebar shows the 'History' panel with a search bar and a list of jobs. The top job is '5: Trimmomatic on DRR024501sub_1.fastq', which is highlighted in green. A red arrow labeled '3' points to its size, '347.05 MB'. Below it are other jobs: '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'.

W12-9: 結果のダウンロード

①のところをクリック。基本的にW10-4と同じ作業です

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar contains a 'Tools' panel with a search bar and various tool categories like 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. The main content area features a green notification box with a checkmark icon, stating: '1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq'. Below this, it explains that job status can be checked in the History pane. The History pane on the right shows a list of datasets: '5: Trimmomatic on DRR024501sub_1.fastq' (347.05 MB), '4: FastQC on data 2: RawData', '3: FastQC on data 2: Webpage', and '2: DRR024501sub_1.fastq'. A red arrow with the number '1' points to the '5: Trimmomatic on DRR024501sub_1.fastq' entry in the history list.

W12-9: 結果のダウンロード

①ファイルサイズ。②Trimmomaticの出力ファイルはfastqsanger形式になってますね。③下部に移動

The screenshot shows the Galaxy web interface. A green notification box in the center states: "1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq". Below this, it explains that the job status will change from 'running' to 'finished' upon completion. In the History panel on the right, a job entry for "5: Trimmomatic on DRR024501sub_1.fastq" is visible, showing a size of 159.6 MB and the output format as "fastqsanger". Three red arrows with numbers 1, 2, and 3 point to the file size, the output format, and the job name in the history panel, respectively.

W12-9: 結果のダウンロード

The screenshot shows the Galaxy web interface. A central green notification box contains the following text:

✓ 1 job has been successfully added to the queue - resulting in the following datasets:

5: Trimmomatic on DRR024501sub_1.fastq

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

The History pane on the right shows an entry for "Unnamed history" with a size of 347.05 MB. The job details for this entry are:

フォーマット: **fastqsanger**, データベース: ?

Picked up _JAVA_OPTIONS:
 -Djava.io.tmpdir=/galaxy-repl/main/jobdir/016/260/162609;
 -Xmx7680m -Xms256m
 TrimmomaticSE: Started with arguments:
 -threads 1 -phred33
 fastq_in.fastqsanger
 fastq_out.fastqsanger
 ILLUMINACLIP:/galaxy/main/deps,

At the bottom of the history entry, there is a 'Save' icon (a floppy disk) and a red arrow labeled '②' pointing to it. Another red arrow labeled '①' points to the job details area.

W12-9: 結果のダウンロード

①で示すファイル名、およびfastqsangerという拡張子で保存されます。zip圧縮されているわけではないので、ダウンロード後は普通にテキストエディタで開けます。約160MBあるので、ダウンロードにそれなりに時間はかかる

The screenshot shows the Galaxy web interface. A green notification box in the center states: "1 job has been successfully added to the queue - resulting in the following datasets: 5: Trimmomatic on DRR024501sub_1.fastq". Below this, it provides instructions on how to check the status of the job. On the right, the History panel shows the dataset details: "フォーマット: fastqsanger, データベース: ?" and lists system parameters like Java options and TrimmomaticSE arguments. At the bottom, a file download dialog is open, asking to open or save the file "Galaxy5-[Trimmomatic_on_DRR024501sub_1.fastq].fastqsanger (159 MB)". A red arrow with the number 1 points to the file name in the dialog.

W13-1 : FastQC

The screenshot shows the Galaxy web interface. The browser address bar displays the URL: `https://usegalaxy.org/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fpjbri...`. The navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'.

Tools Panel (Left):

- search tools
- Get Data
- Send Data
- Lift-Over
- Collection Operations
- Text Manipulation
- Datamash
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Fetch Alignments/Sequences
- Read quality and manipulation (marked with red arrow 1)
- FastQC Read Quality reports
- multiqc aggregate results from bioinformatics analyses into a single report
- Trimmomatic flexible read trimming tool for Illumina NGS data (marked with red arrow 2)

Message Box (Center):

✓ 1 job has been successfully added to the queue - resulting in the following datasets:

5: Trimmomatic on DRR024501sub_1.fastq

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History Panel (Right):

search datasets

Unnamed history
4 shown, 1 deleted

347.05 MB

フォーマット: **fastqsanger**, データベース: ?

Picked up _JAVA_OPTIONS:
-Djava.io.tmpdir=/galaxy-repl/main/jobdir/016/260/162609;
-Xmx7680m -Xms256m
TrimmomaticSE: Started with arguments:
-threads 1 -phred33
fastq_in.fastqsanger
fastq_out.fastqsanger
ILLUMINA_CLIP:/galaxy/main/deps,
@DRR024501.1 M00278:15:000000000-A2RK1
ATGNA

W13-2: Execute

①最後に作成したTrimmomatic実行後のFASTQファイルがデフォルトで見えているので、そのまま②Execute

The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The left sidebar lists tool categories: 'Get Data', 'Send Data', 'Lift-Over', 'Collection Operations', 'Text Manipulation', 'Datamash', 'Convert Formats', 'Filter and Sort', 'Join, Subtract and Group', 'Fetch Alignments/Sequences', and 'NGS: QC and manipulation'. Under 'NGS: QC and manipulation', 'FastQC Read Quality reports' is selected. The central panel shows the tool configuration for 'FastQC Read Quality reports (Galaxy Version 0.69)'. It has three input fields: 'Short read data from your current history' (selected as '5: Trimmomatic on DRR0245...'), 'Contaminant list' (selected as 'Nothing selected'), and 'Submodule and Limit specifying file' (selected as 'Nothing selected'). Below these is an 'Execute' button. The right sidebar shows the 'History' panel with 'Unnamed history' (4 shown, 1 deleted, 347.05 MB) and a log window showing the execution details for 'TrimmomaticSE'.

W13-3: 実行待ち

①実行待ち状態(灰色)であることが、このあたりでわかります

The screenshot shows the Galaxy web interface. A central green notification box contains the following text:

- ✓ 1 job has been successfully added to the queue - resulting in the following datasets:
- 6: FastQC on data 5: Webpage
- 7: FastQC on data 5: RawData

Below the notification, it states: "You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered." A red arrow with the number "1" points from the notification to the "data 5: Webpage" entry in the History panel.

The History panel on the right shows:

- search datasets
- Unnamed history
- 6 shown, 1 deleted
- 347.05 MB
- data 5: Webpage
- 5: Trimmomatic on
- DRR024501sub 1.fastq
- 159.6 MB
- フォーマット: fastqsanger,
- データベース: ?

The job details for "5: Trimmomatic on" are shown in a light green box:

```
Picked up _JAVA_OPTIONS:  
-Djava.io.tmpdir=/galaxy-repl/main/jobdir/016/260/162609:  
-Xmx7680m -Xms256m  
TrimmomaticSE: Started with arguments:  
-threads 1 -phred33  
fastq_in.fastqsanger  
fastq_out.fastqsanger  
ILLUMINACLIP:/galaxy/main/deps,
```


W13-4: 実行終了

①緑色になり、FastQCの実行が終了したっぽいので、②ページ上部に移動

The screenshot shows the Galaxy web interface. A central green notification box contains a checkmark and the text: "1 job has been successfully added to the queue - resulting in the following datasets: 6: FastQC on data 5: Webpage 7: FastQC on data 5: RawData". Below this, it explains that job status can be checked in the History pane. A red arrow labeled "1" points to the notification box. On the right, the History pane shows a dataset named "Webpage" with a green background, indicating completion. A red arrow labeled "2" points to the top of this dataset entry. The History entry details include: "5: Trimmomatic on DRR024501sub_1.fastq", "159.6 MB", "フォーマット: fastqsanger", "データベース: ?", and a log snippet showing Java options and TrimmomaticSE arguments.

①こんな感じで無事
終了していますね

W13-4: 実行終了

The screenshot shows the Galaxy web interface. A central green notification box contains the following text:

✓ 1 job has been successfully added to the queue - resulting in the following datasets:

- 6: FastQC on data 5: Webpage
- 7: FastQC on data 5: RawData

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

A red arrow with the number '1' points from the notification box to the History pane on the right. The History pane shows a list of datasets:

- 7: FastQC on data 5: RawData
- 6: FastQC on data 5: Webpage
- 5: Trimmomatic on DRR024501sub_1.fastq

The '7: FastQC on data 5: RawData' and '6: FastQC on data 5: Webpage' entries are highlighted with a red box. Below them, the details for '5: Trimmomatic on DRR024501sub_1.fastq' are visible, including file size (159.6 MB), format (fastqsanger), and database (?).

W13-5: 結果のダウンロード

①をクリックして、②zip
ファイルをダウンロード

The screenshot shows the Galaxy web interface. A central green notification box contains the following text:

✓ 1 job has been successfully added to the queue - resulting in the following datasets:

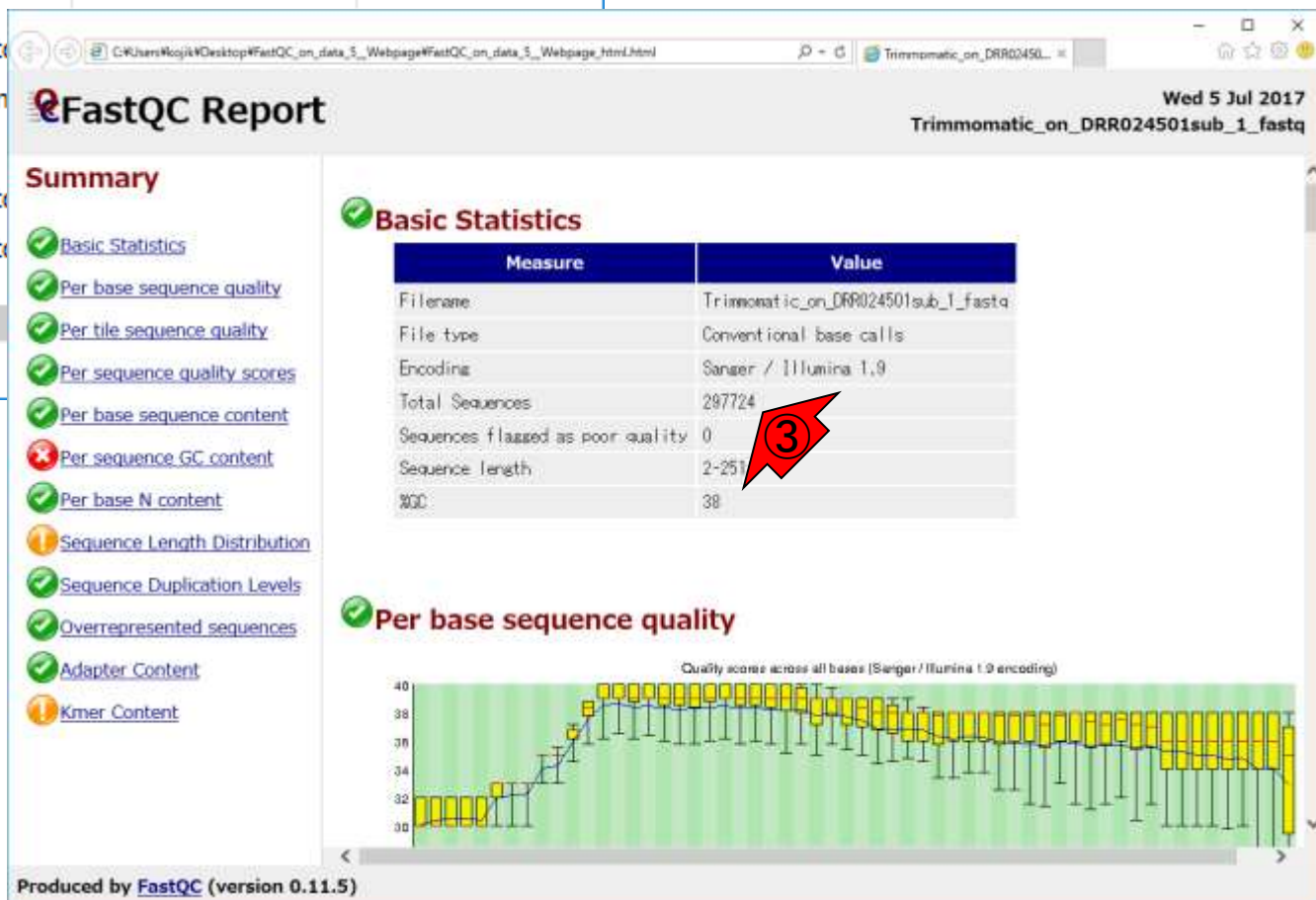
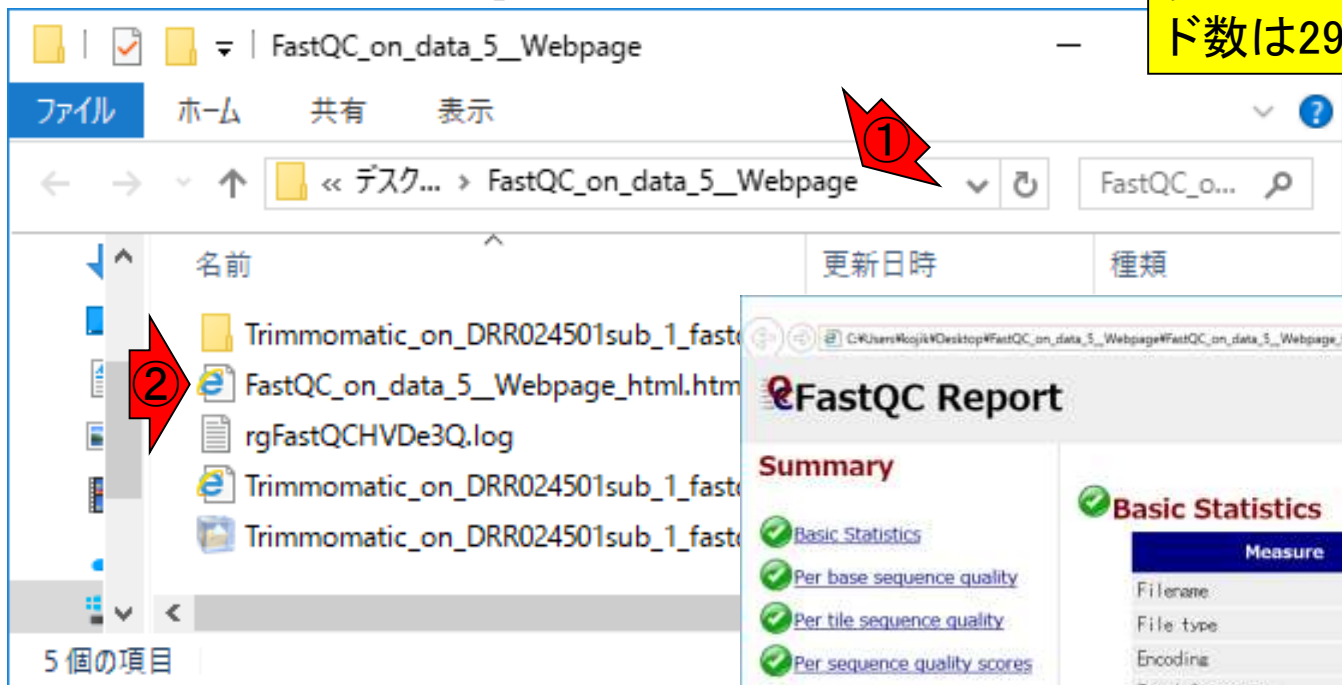
- 6: FastQC on data 5: Webpage
- 7: FastQC on data 5: RawData

You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

The History pane on the right shows a list of datasets. A red arrow labeled '1' points to the entry '7: FastQC on data 5: RawData'. Below it, another red arrow labeled '2' points to the download icon (a floppy disk) for the entry '6: FastQC on data 5: Webpage'. The details for this entry show a size of 294.0 KB and a format of 'html'. Below the details, there is a section for 'Picked up _JAVA_OPTIONS:' with the following values: '-Djava.io.tmpdir=/galaxy-repl/main/jobdir/016/261/162610 -Xmx7680m -Xms256m'. At the bottom of the entry, there is a 'HTML file' label and a download icon.

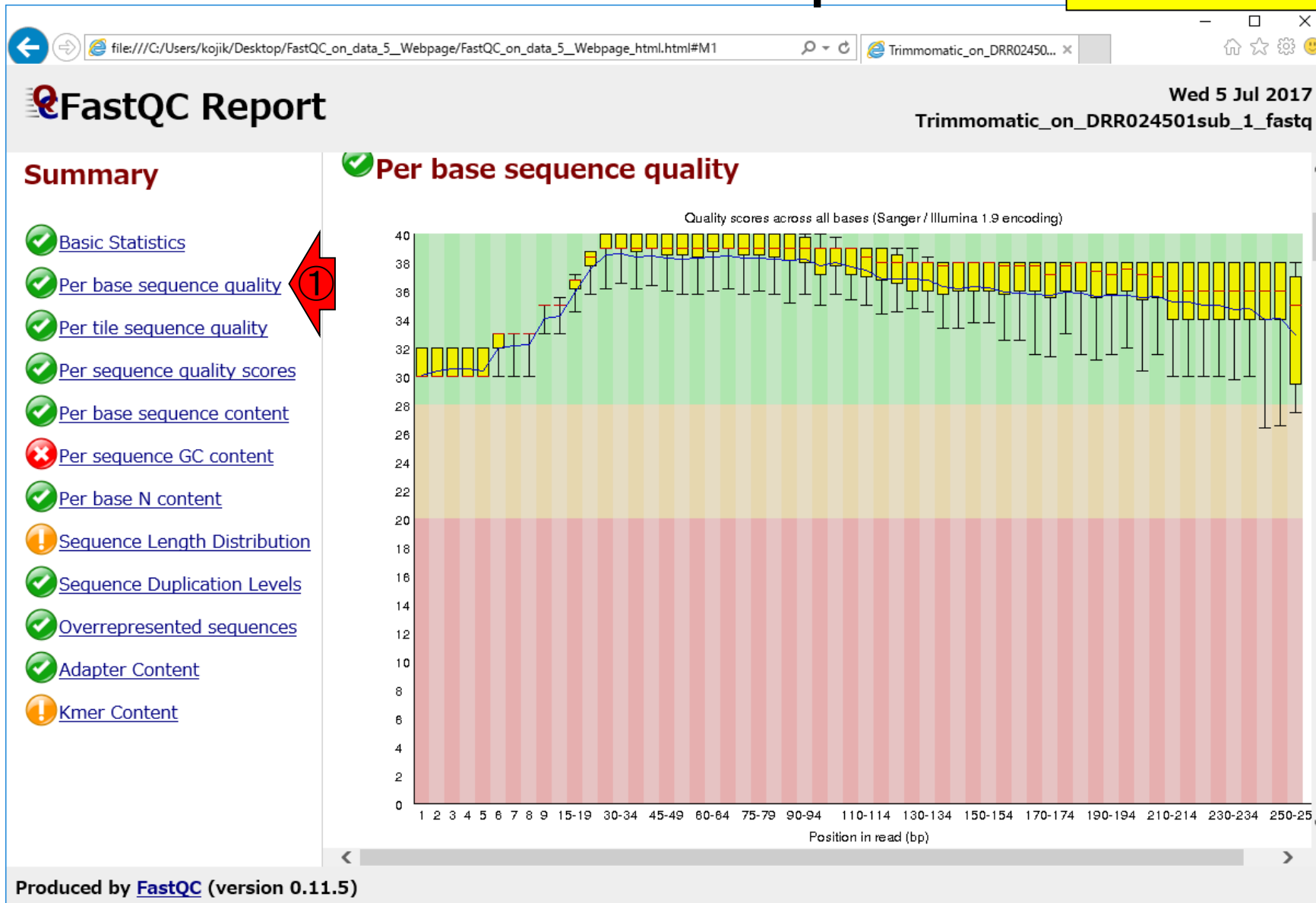
W13-6: 解凍すると...

zipファイルを解凍すると、①のような感じになる。②をダブルクリックするとTrimmomatic実行後のFastQC結果が見られます。③リード数は297,724となったようです



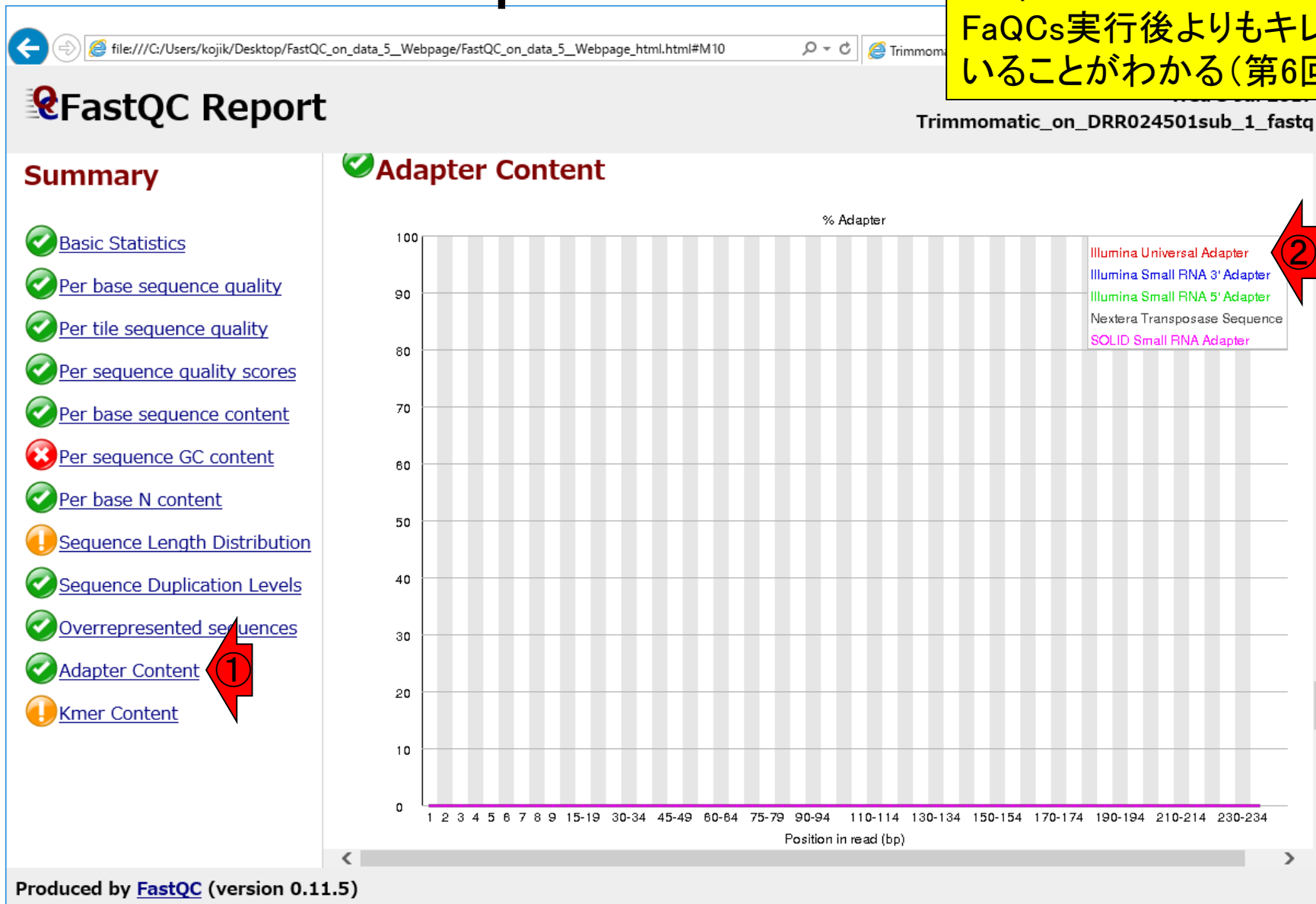
W13-7: Per base sequence.

① Per base sequence qualityをクリックして、クオリティスコア分布を眺めている。綺麗ですね



W13-8: Adapter Content

①Adapter Contentの結果として、W11-1で見えていた②Illumina Universal Adapterがなくなっていることがわかる。FaQCs実行後よりもキレイになくなっていることがわかる(第6回W6-2)



W14-1: NGSハンズオン講習会

①JST-NBDCが主導する②NGSハンズオン講習会の内容は、Linux環境での実習が中心

https://biosciencedbc.jp/human/human-resources/workshop

NBDC National Bioscience Database Center

バイオサイエンスデータベースセンター

English サイトマップ

ホーム NBDCについて 研究開発 公募情報 採用情報 イベント 人材支援 アクセス リンク

Home > 人材支援 > 支援 > 講習会

バイオインフォマティクス人材育成のための講習会

NGS解析

ライフサイエンス分野の研究現場においては、莫大・多様な研究データが産出され、取り扱うデータ量が飛躍的に増えています。一方で、それらのデータを整備・活用するための人材は不足している状況です。そのような研究現場の状況を踏まえ、NBDC運営委員会人材育成分科会において、研究データを整備・活用するバイオインフォマティクス人材を育成するためのカリキュラムに基づき実施した講習会です。本カリキュラムは、対応が急務であると思われる次世代シーケンサデータに焦点をあてた内容となっております。

- 人材育成分科会で策定したカリキュラム
 - バイオインフォマティクス人材育成カリキュラム (次世代シーケンサ)
 - カリキュラムで習得できる技能
 - カリキュラム フロー図
- H29年度NGSハンズオン講習会 (2017年8月28日～9月1日)
- H28年度NGSハンズオン講習会 (2016年7月19日～8月4日)
- H27年度NGSハンズオン講習会 (2015年7月22日～8月6日)
- H26年度NGS速習コース講習会 (2014年9月1日～12日)