

# ゲノム情報解析基礎：第1回

<sup>1</sup>大学院農学生命科学研究科  
アグリバイオインフォマティクス教育研究プログラム  
<sup>2</sup>微生物科学イノベーション連携研究機構  
門田幸二(かどた こうじ)  
kadota@iu.a.u-tokyo.ac.jp  
<http://www.iu.a.u-tokyo.ac.jp/~kadota/>

# 各講義科目へのアクセス

①教育プログラム、②各講義のページ、③「ゲノム情報解析基礎」の場合。ブラウザは、Google Chromeを推奨。



東京大学大学院農学系研究所  
アグリバイオ  
Agricultural Bioinformatics

ようこそ!!  
アグリバイオ  
教育研究ユニット

- + ホーム
- + 本ユニットについて
- + メンバー
- + 教育プログラム
- + 研究フォーラム
- + イベント
- + お問い合わせ
- + リンク

東京大学  
THE UNIVERSITY OF TOKYO

## 教育プログラム

- ▼ プログラム概要
- ▼ 講義について
- ▼ 受講について
- ▼ 各講義のページ
- ▼ スケジュール

### プログラム概要

本プログラムで開講する講義科目は( )に分けられます。カテゴリと

カテゴリ	目的
基礎	主にバイオインフォマティクス。生命科学のためのツールやソフトウェアを利用した様々なツ
方法論	「基礎」の科目を土台として、解析手法、質量分析法など)やモデル選択、分子シミュレーション(加)について解説し

### お知らせ

- ▶ 2019年度の受講希望者は2019年4月5日までに事務局までお越しください。
- ▶ 2019年度受講生募集要項は [こちら](#) (PDF) です。 **NEW!!**
- ▶ 受講に関する質問はまずこちらをごらんください。  
Q & A集(本学の大学院生の方)  
Q & A集(本学の大学院生以外の方)
- ▶ 成績証明書の発行を希望される方は申込用紙「Word形式、PDF形式」事務局までご連絡ください。

## 各講義のページ

(科目名をクリックすると各講義のページに移動します)

**先端トピックス**  
セミナー・討論形式  
研究指導

**農学生命情報科学特別演習**

農学生命情報科学特論 I

農学生命情報科学特論 II

農学生命情報科学特論 III

農学生命情報科学特論 IV

**方法論**  
講義・実習を一体化

生物配列統計学 システム生物学概論 知識情報処理論  
オーム情報解析 機能ゲノム学 分子モデリングと分子シミュレーション  
**フィールドインフォマティクス**

**基礎**  
講義・実習を一体化

ゲノム情報解析基礎 構造バイオインフォマティクス基礎  
生物配列解析基礎 バイオスタティスティクス基礎論

# 講義予定

- 04月08日月曜日(17:15-20:30)
  - バイオインフォマティクスのイントロダクション
  - Rの基本的な利用法、翻訳配列取得、RStudio
- 04月15日月曜日(17:15-20:30)
  - Rの基本的な利用法:コードの解説、関数マニュアルの読み解き方
- 04月22日月曜日(17:15-20:30)
  - Rで塩基配列解析(multi-FASTAファイルの各種解析)
  - Rパッケージの話(CRANとBioconductor)
- 05月13日月曜日(17:15-20:30)
  - ゲノムアノテーション

# 全体方針や制約

Q: どのようなプログラミング言語を学べばよいのでしょうか?  
A: 目的さえ達成できれば何でもよいのではないのでしょうか。Rに特化していることを最大限利用して、主にRで塩基配列解析をやるのが「ゲノム情報解析基礎」。用いるRの基本スキルは同じなので、比較的理解が容易な本科目で慣れておくメリットはあると思われる。もちろん受け止め方は多様。

## ■ 全体方針 (Rを利用)

- アグリバイオの講義全体として、Rとなるべく利用(2004/平成16年度～)

## ■ 主な制約

- 主に大学院講義の枠組みで行っているため、講義として成立させる必要性。
- バイオインフォマティクスを学ぶ上でLinux習得は基本だと思っているが、アグリバイオ固有の問題(受講人数は非常に多いが、講義補助が非常に少ない)のために、講義の枠組みでLinuxスキルを実用上使えるレベルまで引き上げることは事実上不可能。また、多数派を占める実験系の修士1年の学生の多くは、そこまでのレベルを求めている。
- 特に受講人数の多い基礎系科目では、ウェブツールの利用も困難。1つのサイトに同時期にアクセスすることになるため、実行結果が得られないヒトが続出してしまふ。

# ゲノム情報解析基礎

①「ゲノム情報解析基礎」のページ。②講義資料PDF。③本日分の解説からスタート。

The screenshot shows the course page for '2. ゲノム情報解析基礎' (Genome Information Analysis Basics) on the website of the Agricultural Bioinformatics Research Unit at the University of Tokyo. The page includes a navigation menu on the left, a header with the unit's name, and a main content area with sections for '授業の目標・概要' (Course Objectives and Overview), '担当教員' (Instructor), '参考図書' (Reference Books), 'お知らせ' (Notice), and '講義日程 (2019年度)' (Lecture Schedule for 2019). A red box highlights the '2. ゲノム情報解析基礎' section, and a red arrow points from this box to the right-hand side of the image.

This block provides a detailed view of the course page content, with red arrows pointing to specific elements:

- 2. ゲノム情報解析基礎** (2. Genome Information Analysis Basics)
- 授業の目標・概要** (Course Objectives and Overview): 次世代シーケンサーの普及により、ゲノム情報を基盤とした膨大な塩基配列情報を自在に解析するスキルが要求される時代になっています。フリーソフトRを用いた配列決定後の基礎情報取得など、各種配列解析の基本スキル向上を目指した実習を含む講義を行います。
- 担当教員** (Instructor): 門田幸二 (東大・農・アグリバイオ / 准教授)
- 参考図書** (Reference Books): 藤博幸・編、よくわかるバイオインフォマティクス入門、講談社、2018; 坊農秀雅・小野浩雅・監修、生命科学データベース・ウェブツール、MEDSi、2018; 坊農秀雄 著、Dr.Bonoの生命科学データ解析、MEDSi、2017
- お知らせ** (Notice): 講義では、Rの様々なパッケージを利用します。持ち込み用PC利用希望者は [インストール |](#) についてを参考にしてR本体および必要なパッケージ群を必ずインストールしておいてください。
- 講義日程 (2019年度)** (Lecture Schedule for 2019):
  - 2019年04月08日 (PC使用)
    - 講義資料PDF**
    - 学会(国外): ISCB
    - 学会(国内): JSBi
    - QAサイト: Biostar (Parnell et al., PLoS Comput Biol., 2011)
    - QAサイト: SEQanswers (Li et al., Bioinformatics, 2012)
    - 学習教材: バイオインフォマティクス人材育成のための講習会(平成26-29年度)
    - 学習教材: (Rで)塩基配列解析
    - 学習教材: (Rで)塩基配列解析のサブ
  - 2019年04月15日 (PC使用)

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# Introduction

「ゲノム情報解析基礎」の、①本日分のところを拡大表示。最初のうちはリラックスして、スライドを眺めているだけでよい。最初の数十分程度は、無線LAN不調の貸与PCや持込PCへの対応、および遅れてきたヒトへの対応などでバタバタします。そのため、よく聞いていないと後々困る情報にはしていません。

1. 2019年04月08日 (PC使用) ①

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., Bioinformatics, 2012)

学習教材 : バイオインフォマティクス人材育成のための講習会(平成26-29年度)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



# 学会(国外)

①ISCB (International Society for Computational Biology) は、バイオインフォマティクス分野の国際学会。バリバリのバイオインフォマティシャンとして活動することがなければ、基本的には無縁だろう。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB



学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comp)

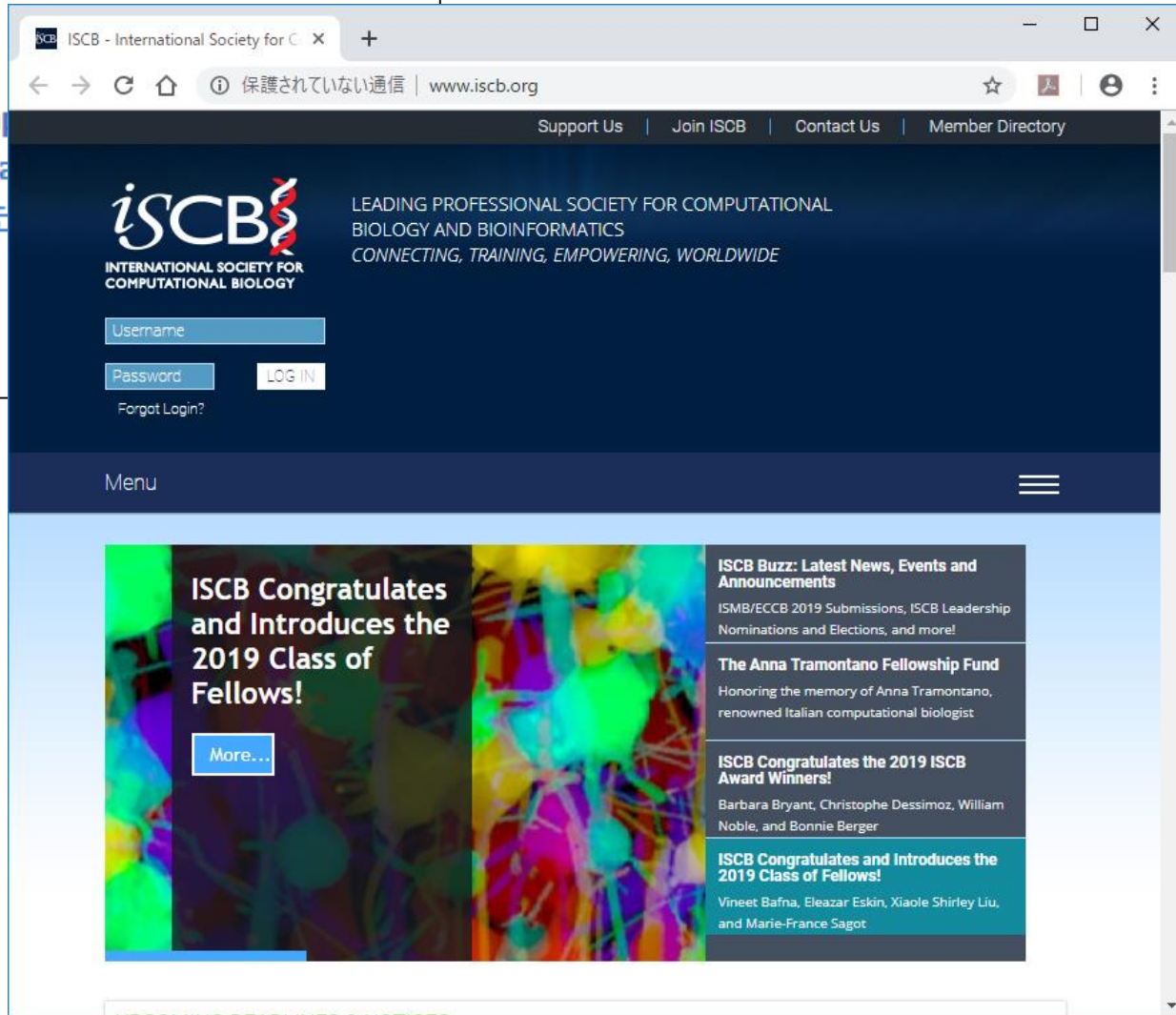
QAサイト : SEQanswers (Li et al., Bioinforma

学習教材 : バイオインフォマティクス人材育成のた

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio





# 学会(国内)

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi



QAサイト : Biostar (Parnell et al.,

QAサイト : SEQanswers (Li et al.,

学習教材 : バイオインフォマティクス

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio

# QAサイト

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : **Biostar (Parnell et al., PLoS Comput Biol., 2011)**

QAサイト : **SEQanswers (Li et al., Bioinformatics, 2012)**

学習教材 : バイオインフォマティクス人材育成のための講習会(平成26-29年度)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio




# QAサイト

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : 

QAサイト : Biostar (Parnell et al.,

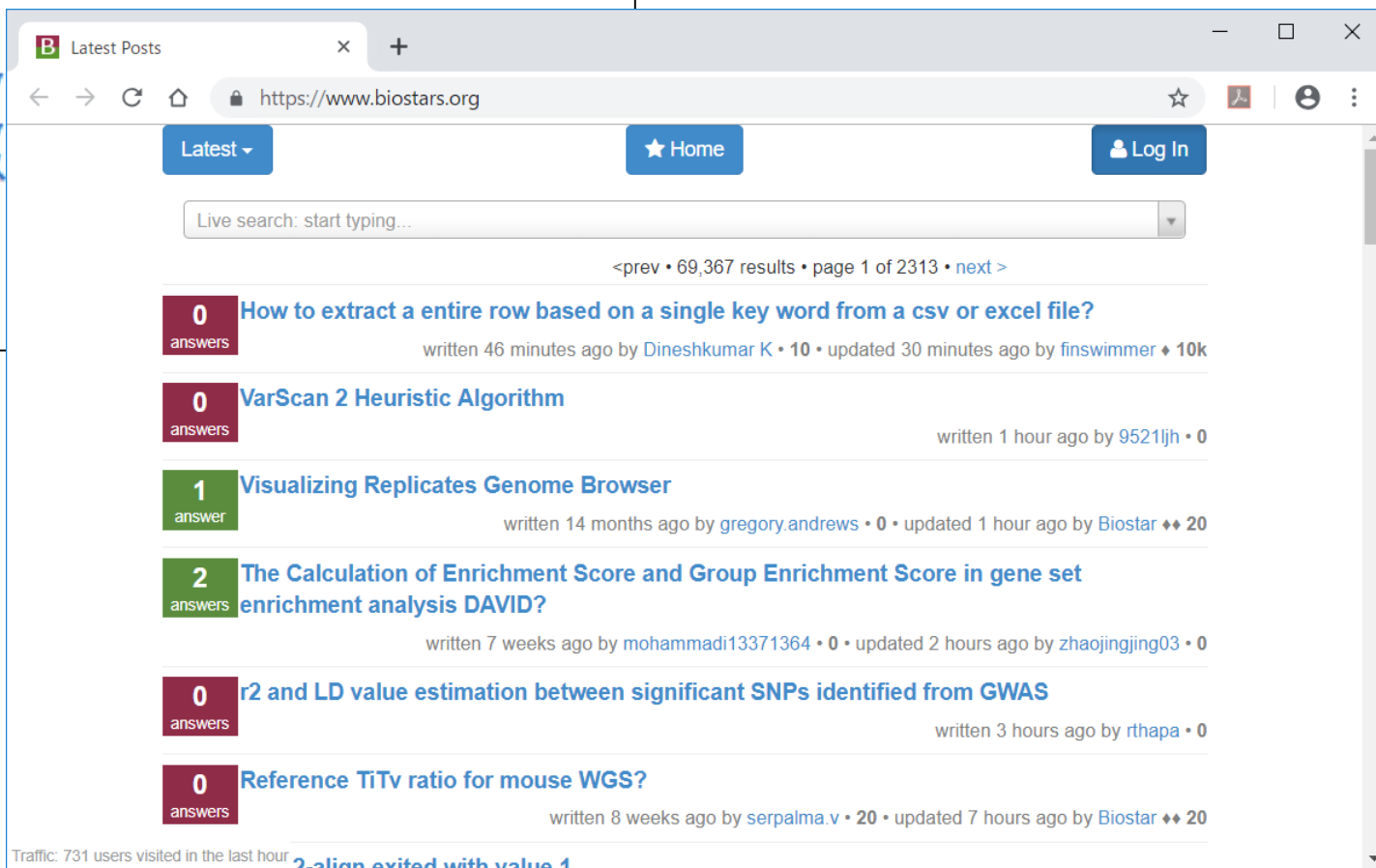
QAサイト : SEQanswers (Li et al.,

学習教材 : バイオインフォマティクス

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



The screenshot shows the 'Latest Posts' page on the Biostars website. The browser address bar displays 'https://www.biostars.org'. The page features a search bar and navigation buttons for 'Latest', 'Home', and 'Log In'. Below the search bar, there are several question entries, each with a colored box indicating the number of answers (0, 1, or 2) and the title of the question. The questions listed are:

- 0 answers: How to extract a entire row based on a single key word from a csv or excel file?
- 0 answers: VarScan 2 Heuristic Algorithm
- 1 answer: Visualizing Replicates Genome Browser
- 2 answers: The Calculation of Enrichment Score and Group Enrichment Score in gene set enrichment analysis DAVID?
- 0 answers: r2 and LD value estimation between significant SNPs identified from GWAS
- 0 answers: Reference TiTv ratio for mouse WGS?

Traffic: 731 users visited in the last hour

# QAサイト

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al.,

QAサイト : SEQanswers (Li et al.,

学習教材 : バイオインフォマティクス

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



Forum	Last Post	Threads	Posts
<b>Introductions</b> (10 Viewing) New here? Stop in and introduce yourself. Where you are, what you work on, etc.	<b>Hi all</b> by <a href="#">Erara</a> 03-05-2019 02:37 AM	2,081	4,132
<b>General</b> (11 Viewing) Any topic/question that does not fit into the subcategories below. If you're unsure of where to put something, ask in here!	<b>Getting rid of PGM Supplies</b> by <a href="#">cheaney</a> 03-07-2019 01:36 PM	1,729	6,699
<b>Core Facilities</b> (1 Viewing)	<b>Fragment Analyzer-Advanced...</b>		

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

①NGSに特化したバイオインフォ人材育成のための講習会。主にNBDCと東大アグリバイオで4年間実施。

リンク先をクリック

# 学習教材

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., Bioinformatics, 2012)

学習教材 : [バイオインフォマティクス人材育成のための講習会\(平成26-29年度\)](#)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio





# 学習教材

①NGSに特化したバイオインフォ人材育成のための講習会。主にNBDCと東大アグリバイオで4年間実施。②ページ下部に移動

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外): ISCB

学会(国内): JSBi

QAサイト: Biostar (P

QAサイト: SEQansw

学習教材: **バイオイン**

学習教材: (Rで)塩基配

学習教材: (Rで)塩基配

RStudio

The screenshot shows a web browser window displaying the NBDC (National Bioscience Database Center) website. The URL is <https://biosciencedbc.jp/human/human-resources/workshop>. The page features the NBDC logo and navigation menus. A red arrow labeled '1' points to the workshop title: **バイオインフォマティクス人材育成のための講習会**. Another red arrow labeled '2' points to the scroll bar on the right side of the page, indicating the location of the workshop details.

講習会 - バイオサイエンスデータベース × +  
← → ↻ 🏠 🔒 <https://biosciencedbc.jp/human/human-resources/workshop> ☆ 人 | 🛡️ | ⋮

**NBDC** - 散在するデータベースを、まとめて、使い易く -  
National Bioscience Database Center

English サイトマップ

文字サイズ変更 大 中 小

Search for... Search

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 NBDCブログ

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** ①

**お知らせ**  
2018年5月 平成30年度以降、NGSハンズオン講習会は開催いたしません。

**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

# 学習教材

①NGSに特化したバイオインフォ人材育成のための講習会。主にNBDCと東大アグリバイオで4年間実施。②ページ下部に移動

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (P

QAサイト : SEQansw

学習教材 : **バイオイン**

学習教材 : (Rで)塩基配

学習教材 : (Rで)塩基配

RStudio

講習会 - バイオサイエンスデータベース × +

← → ↻ 🏠 🔒 https://biosciencedbc.jp/human/human-resources/workshop ☆ 人 | 🛡️ ⋮

## 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 \(2014年9月1日～12日\)](#)

[サイトポリシー](#) [ロゴについて](#)

Copyright© 2011 Japan Science and Technology Agency. All Rights Reserved.



# 学習教材

①NGSに特化したバイオインフォ人材育成のための講習会。主にNBDCと東大アグリバイオで4年間実施。②ページ下部に移動。③このあたりに4年分の膨大な講義資料・映像がある。講義内容が被っているものもあり、中身も古くなっているが、特にNGS解析の実務と絡めたLinuxの基本的な事柄が有用だと思います。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (P

QAサイト : SEQansw

学習教材 : **バイオイン**

学習教材 : (Rで)塩基配

学習教材 : (Rで)塩基配

RStudio

講習会 - バイオサイエンスデータベース

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

## 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)

[サイトポリシー](#) [ロゴについて](#)

Copyright© 2011 Japan Science and Technology Agency. All Rights Reserved.

# H27年度講習会

①NGSに特化したバイオインフォ人材育成のための講習会。主にNBDCと東大アグリバイオで4年間実施。②ページ下部に移動。③このあたりに4年分の膨大な講義資料・映像がある。講義内容が被っているものもあり、中身も古くなっているが、特にNGS解析の実務と絡めたLinuxの基本的な事柄が有用だと思います。例えば、④H27年度講習会のページをクリック。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外): ISCB

学会(国内): JSBi

QAサイト: Biostar (P

QAサイト: SEQansw

学習教材: バイオイン

学習教材: (Rで)塩基配

学習教材: (Rで)塩基配

RStudio

講習会 - バイオサイエンスデータベース ×

← → ↻ 🏠 🔒 https://bioscience.nibdc.ac.jp/

## 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 \(2014年9月1日～12日\)](#)

[サイトポリシー](#) [ロゴについて](#)

Copyright© 2011 Japan Science and Technology Agency. All Rights Reserved.

# H27年度講習会

①H27年度NGSハンズオン講習会のページ。②のあたりに少し見えているが、この年は11日間実施。③講義資料や動画はココ。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (P

QAサイト : SEQansw

学習教材 : バイオイン

学習教材 : (Rで)塩基配

学習教材 : (Rで)塩基配

RStudio

講義会 - バイオサイエンスデータベース × 平成27年度NGSハンズオン講習会 - ×

https://biosciencedbc.jp/human/human-resources/workshop/h27

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

Home ▶ 人材支援 ▶ 支援 ▶ 講習会 ▶ 平成27年度NGSハンズオン講習会

## ① 平成27年度NGSハンズオン講習会

平成27年度は、平成26年度の実績を踏まえ、講義内容の改善等を行い、ハンズオンに特化した、より効果的なNGS講習会を開催しました。

- H27年度概要
- H27年度講義日程・参考資料
- H26年度講習会の情報についてはこちらをご覧ください。
- H27年度実施報告書・講義資料・動画等

● [講習会実施報告書 \(PDF : 2.17MB\)](#) および [受講者アンケート集計結果 \(データ集\) \(PDF : 662KB\)](#)

● [講義資料・動画](#) \*講義資料一覧のファイル名をクリックすると資料ファイル (PDF等) がダウンロードできます。

実施日	実施時間	大項目	項目	レベル	習得技術	担当講師(敬称略)	講義資料・動画(統合TV)
7月22日 (水)	10:30-12:00	PC環境の構築	Bio-Linux8とRのインストール状況確認		・Linux導入 ・R導入 ・NGS解析に必要な	門田 幸二 (東京大学)	事前予習資料一覧(PDF:52KB)
	13:15-14:45						



# H27年度講習会

①H27年度NGSハンズオン講習会のページ。②のあたりに少し見えているが、この年は11日間実施。③講義資料や動画はココ。④実施報告書には、受講生の感想などを全て掲載。多様な感想があることが分かります。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (P

QAサイト : SEQansw

学習教材 : バイオイン

学習教材 : (Rで)塩基配

学習教材 : (Rで)塩基配

RStudio

Home ▶ 人材支援 ▶ 支援 ▶ 講習会 ▶ 平成27年度NGSハンズオン講習会

## 平成27年度NGSハンズオン講習会

平成27年度は、平成26年度の実績を踏まえ、講義内容の改善等を行い、ハンズオンに特化した、より効果的なNGS講習会を開催しました。

- [H27年度概要](#)
- [H27年度講義日程・参考資料](#)
- [H26年度講習会の情報についてはこちらをご覧ください。](#)
- [H27年度実施報告書・講義資料・動画等](#)
- [講習会実施報告書 \(PDF: 2.17MB\) および受講者アンケート集計結果 \(データ集\) \(PDF: 662KB\)](#)
- [講義資料・動画](#) \*講義資料一覧のファイル名をクリックすると資料ファイル (PDF等) がダウンロードできます。

実施日	実施時間	大項目	項目	レベル	習得技術	担当講師(敬称略)	講義資料・動画(統合TV)
7月22日 (水)	10:30-12:00	PC環境の構築	Bio-Linux8とRのインストール状況確認		<ul style="list-style-type: none"> <li>Linux導入</li> <li>R導入</li> <li>NGS解析に必要な</li> </ul>	門田 幸二 (東京大学)	事前予習資料一覧(PDF:52KB)
	13:15-14:45						



# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# 私のスタンス

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comp

QAサイト : SEQanswers (Li et al., Bioinformat

学習教材 : バイオインフォマティクス人材育成のため

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio

私のスタンスは、10年以上におよぶハンズオン講義経験を踏まえ、特に多数派の意見や正論を反映させてきたものです。アグリバイオ大学院講義の枠組み(特に4-5月の基礎科目の実施時期)では、受講生の多数派を占める修士1年の状況(バイオインフォ初学者、他の講義科目との兼ね合い、大学院からのヒトが多く研究室での生活にも慣れなければいけない)を考慮し、講義時間内のみで無理なく理解できるようにゆっくり丁寧に教えます。また、講義に来られない方にも配慮して、講義に出なくてもよいレベルの分かりやすさを追求しています。意欲のある一部のヒトにとっては残念な講義ですので、予めご承知おきください。



# 学習教材

①(Rで)塩基配列解析は、スライド4で述べたアグリバイオの全体方針に従って、主にRでゲノム・トランスクリプトーム解析を行うために作り込んできたウェブサイト。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., Bioinformatics, 2012)

学習教材 : バイオインフォマティクス人材育成のための講習会(平成26-29年度)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



# 学習教材

①(Rで)塩基配列解析は、スライド4で述べたアグリバイオの全体方針に従って、主にRでゲノム・トランスクリプトーム解析を行うために作り込んできたウェブサイト。アグリバイオ貸与PCは、②R本体とパッケージ群のインストール済み。③受講生募集要項の、本科目の「履修上の注意」に相当。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外): ISCB

学会(国内): JSBi

QAサイト: Biostar (Parnell)

QAサイト: SEQanswers (Parnell)

学習教材: バイオインフォマ

学習教材: (Rで)塩基配列解析

学習教材: (Rで)塩基配列解析

RStudio

(Rで)塩基配列解析  
(last modified 2019/03/11, since 2019/03/11)

このウェブページのR関連部分は、[インストール | についての推奨手順 \(Windows2018.11.15版とMacintosh2018.11.27版\)](#)に従ってフリーソフトRと必要なパッケージをインストール済みであるという前提で記述しています。初心者の方は[基本的な利用法 \(Windows2018.12.23版とMacintosh2019.01.15版\)](#)で自習してください。2018年7月に(Rで)塩基配列解析の一部(講習会・書籍・学会誌など)を切り分けて[サブページ](#)に移行しました。(2018/07/18)

What's new? ([過去のお知らせ](#)はこちら)

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) **NEW**
- 細かいところの修正はここに明記してなくても随時行っています。(2019/03/11) **NEW**
- 「インストール | Rパッケージ | [必要最小限プラスアルファ](#)」を更新しました。(2019/03/08) **NEW**
- 「[生命科学データ解析を支える情報技術](#) (監修: 坊農秀雅)」が出版されています。最先端のネタを含むかなり広範な内容を含んでいますので、一通り目次を眺めてみるとよいと思います。Bioconda, Homebrew, Docker, GitHub, EC2, AWSなど聞いたことがある有用そうなものの全体像がわかるというメリットがあると思います。(2019/02/06)

• [はじめに](#) (last modified 2018/08/04) [トップページへ](#)

• [過去のお知らせ](#) (last modified 2019/03/02) **NEW**

# 履修上の注意

2019 年度

アグリバイオインフォマ

教育研究プログラ

③ 受講生募集要項



科目名	ゲノム情報解析基礎 Introduction to Ge
担当教員	門田幸二
実施日	4/8(月), 4/15(月),

## 授業の目標・概要

次世代シーケンサーの普及により、ゲノム情報を基盤とした膨大な塩基配列情報を自在に解析するスキルが要求される時代になっています。フリーソフト R を用いた配列決定後の基礎情報取得など、各種配列解析の基本スキル向上を目指した実習を含む講義を行います。

## 授業計画

以下を予定しています：

- 1) バイオインフォマティクスのイントロダクション
- 2) R の基本的な利用法
- 3) R で塩基配列解析 (multi-FASTA ファイルの各種解析)
- 4) R パッケージの話 (CRAN と Bioconductor)
- 5) ゲノムアノテーション

## 授業の方法

ノート PC を用いた実習形式で授業を進めます。

## 成績評価の方法

出席および小テストなどの内容に基づいて総合的に評価します。

## 参考書など

藤博幸 編、よくわかるバイオインフォマティクス入門、講談社、2018  
坊農秀雄・小野浩雅 監修、生命科学データベース・ウェブツール、MEDSi、2018  
坊農秀雄、生命科学データ解析、MEDSi、2017

④

## 履修上の注意

Windows PC の基本的な操作ができること。

本プログラム所有 PC の利用を基本としますが、個人所有 PC の利用を希望⑤は [http://www.iu.a.u-tokyo.ac.jp/~kadota/r\\_seq.html#about\\_install](http://www.iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#about_install) を参考に  
して R および各種 R パッケージのインストールをしておいて下さい。

① (Rで) 塩基配列解析は、スライド4で述べたアグリバイオの全体方針に従って、主にRでゲノム・トランスクリプトーム解析を行うために作り込んできたウェブサイト。アグリバイオ貸与PCは、②R本体とパッケージ群のインストール済み。③受講生募集要項の④本科目の「履修上の注意」に相当。⑤のURLのリンク先と同じだということです。

⑤

# 学習教材

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外): [ISCB](#)

学会(国内): [JSBi](#)

QAサイト: [Biostar \(Parnell\)](#)

QAサイト: [SEQanswers](#) (①)

学習教材: [バイオインフォマ](#)

学習教材: [\(Rで\)塩基配列解析](#)

学習教材: [\(Rで\)塩基配列解析](#)

[RStudio](#)

(Rで)塩基配列解析  
(last modified 2019/03/11, since 2010)

このウェブページのR関連部分は、[インストール | についての推奨手順 \(Windows2018.11.15版とMacintosh2018.11.27版\)](#)に従ってフリーソフトRと必要なパッケージをインストール済みであるという前提で記述しています。初心者の方は[基本的な利用法 \(Windows2018.12.23版とMacintosh2019.01.15版\)](#)で自習してください。2018年7月に(Rで)塩基配列解析の一部 (講習会・書籍・学会誌など) を切り分けてサブページに移行しました。(2018/07/18)

②

What's new? ([過去のお知らせ](#)はこちら)

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) **NEW**
- 細かいところの修正はここに明記してなくても随時行っています。(2019/03/11) **NEW**
- 「インストール | Rパッケージ | [必要最小限プラスアルファ](#)」を更新しました。(2019/03/08) **NEW**
- 「[生命科学データ解析を支える情報技術](#) (監修: 坊農秀雅)」が出版されています。最先端のネタを含むかなり広範な内容を含んでいますので、一通り目次を眺めてみるとよいと思います。Bioconda, Homebrew, Docker, GitHub, EC2, AWSなど聞いたことがある有用そうなものの全体像がわかるというメリットがあると思います。(2019/02/06)

• [はじめに](#) (last modified 2018/08/04) [トップページへ](#)

• [過去のお知らせ](#) (last modified 2019/03/02) **NEW**



2018年7月に①の内容の一部を、②サブページに移行しています。それが③のこと。②または③をクリック。

# 学習教材

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., Bioinformatics, 2012)

学習教材 : バイオインフォマティクス人材育成のための講習会(平成26-29年度)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



推奨手順 ([Windows2018.11.15版](#)と[Macintosh2018.11.27](#)済みであるという前提で記述しています。初心者の方は[基15版](#)で自習してください。2018年7月に(Rで)塩基配列解析の一部(講習会・書籍・学会誌など)を切り分けてサブページに移行しました。(2018/07/18)

What's new? ([過去のお知らせ](#)はこちら)

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) NEW
- 細かいところの修正はここに明記してなくても随時行っています。(2019/03/11) NEW
- 「インストール | Rパッケージ | [必要最小限プラスアルファ](#)」を更新しました。(2019/03/08) NEW
- 「[生命科学データ解析を支える情報技術](#) (監修: 坊農秀雅)」が出版されています。最先端のネタを含むかなり広範な内容を含んでいますので、一通り目次を眺めてみるとよいと思います。Bioconda, Homebrew, Docker, GitHub, EC2, AWSなど聞いたことがある有用そうなものの全体像がわかるというメリットがあると思います。(2019/02/06)

[はじめに](#) (last modified 2018/08/04) [トップページ](#)

[過去のお知らせ](#) (last modified 2019/03/02) NEW

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# サブページ

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : [ISCB](#)

学会(国内) : [JSBi](#)

QAサイト : [Biostar \(Parnell\)](#)

QAサイト : [SEQanswers \(Li\)](#)

学習教材 : [バイオインフォマテ](#)

学習教材 : [\(Rで\)塩基配列解析](#)

学習教材 : [\(Rで\)塩基配列解析](#)

[RStudio](#)



(Rで)塩基配列解析のサブ (last modified 2019/03/11, since 2010)

ここは、[\(Rで\)塩基配列解析](#)のサブページです。メインは[こちら](#)です。(2018/07/18)

### What's new?

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) **NEW**
- [TCC-GUI](#)の[オンライン版](#)の基本的な利用法は、2019年3月15日予定の[講義資料](#) (の後半部分) に記載しています。このオンライン版はインストールが不要という長所がある一方で、すぐに接続が切れる短所があります。これを回避するローカル版の起動方法 (RStudioのインストール含む) も[講義資料](#) (の最後のほう) に記載しています。(2019/03/03) **NEW**
- 「参考資料 | [講義](#)、[講演資料](#)」を更新しました。(2019/03/02) **NEW**
- 2018年11月に「[よくわかるバイオインフォマティクス入門](#) (藤博幸 編)」という本が出ました。(2018/11/21)
- [TCCのオンラインGUI版 \(TCC-GUI\)](#) を公開しました。volcano plotなど図の作成までやってくれる上に、それがどのようなスクリプトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_ \_)m(2018/10/15)

[トップページ](#)へ

# サブページ

①少しずつページ下部に移動して全体像を眺める。1ページ下に移動したのが、次のスライド。自分が今どこにいるのかわからなくなったら、②を押して一番上に戻しましょう。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html

## (Rで)塩基配列解析のサブ

(last modified 2019/03/11, since 2010)

ここは、[\(Rで\)塩基配列解析](#)のサブページです。メインは[こちら](#)です。(2018/07/18)

### What's new?

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) **NEW**
- [TCC-GUI](#)の[オンライン版](#)の基本的な利用法は、2019年3月15日予定の[講義資料](#)（の後半部分）に記載しています。このオンライン版はインストールが不要という長所がある一方で、すぐに接続が切れる短所があります。これを回避するローカル版の起動方法（RStudioのインストール含む）も[講義資料](#)（の最後のほう）に記載しています。(2019/03/03) **NEW**
- 「参考資料 | [講義](#)、[講演資料](#)」を更新しました。(2019/03/02) **NEW**
- 2018年11月に「[よくわかるバイオインフォマティクス入門](#)（藤博幸 編）」という本が出ました。(2018/11/21)
- [TCC](#)の[オンラインGUI版 \(TCC-GUI\)](#) を公開しました。 volcano plotなど図の作成までやってくれる上に、それがどのようなスクリプトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_ \_)m(2018/10/15)

[トップページへ](#)

• [はじめに](#) (last modified 2018/07/18)

# サブページ

①少しずつページ下部に移動して全体像を眺める。②は私の解説記事や、講習会の講義資料PDF、講義や講演資料PDFをカテゴリ別に時系列(新→古)で並べたリスト。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html

プトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_  
\_)m(2018/10/15)

- [はじめに](#) (last modified 2018/07/18)
- [参考資料 | 書籍、学会誌](#) (last modified 2017/11/13)
- [参考資料 | 講習会](#) (last modified 2017/11/13)
- [参考資料 | 講義、講演資料](#) (last modified 2019/03/02)
- [過去のお知らせ](#) (last modified 2018/11/21)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2017](#) (last modified 2017/08/08)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2016](#) (last modified 2016/08/22)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2015](#) (last modified 2016/08/19)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGS速習コース2014](#) (last modified 2015/02/11)
- [書籍 | トランスクリプトーム解析 | について](#) (last modified 2014/05/12)
- [書籍 | トランスクリプトーム解析 | 2.3.1 RNA-seqデータ\(FASTQファイル\)](#) (last modified 2016/03/17)
- [書籍 | トランスクリプトーム解析 | 2.3.2 リファレンス配列](#) (last modified 2014/04/16)
- [書籍 | トランスクリプトーム解析 | 2.3.3 アノテーション情報](#) (last modified 2017/08/02)
- [書籍 | トランスクリプトーム解析 | 2.3.4 マッピング\(準備\)](#) (last modified 2014/06/20)
- [書籍 | トランスクリプトーム解析 | 2.3.5 マッピング\(本番\)](#) (last modified 2014/06/21)
- [書籍 | トランスクリプトーム解析 | 2.3.6 カウントデータ取得](#) (last modified 2016/02/09)
- [書籍 | トランスクリプトーム解析 | 3.3.1 解析目的別留意点](#) (last modified 2014/04/20)
- [書籍 | トランスクリプトーム解析 | 3.3.2 データの正規化\(基礎編\)](#) (last modified 2014/06/23)
- [書籍 | トランスクリプトーム解析 | 3.3.3 クラスタリング](#) (last modified 2014/04/20)

[トップページへ](#)

# サブページ

①少しずつページ下部に移動して全体像を眺める。②は私の解説記事や、講習会の講義資料PDF、講義や講演資料PDFをカテゴリ別に時系列(新→古)で並べたリスト。例えば③「書籍、学会誌」をクリックすると…

(Rで)塩基配列解析のサブ

x +

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html

プトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_  
\_)m(2018/10/15)

- [はじめに](#) (last modified 2018/07/18)
- [参考資料 | 書籍、学会誌](#) **③** (last modified 2017/11/13)
- [参考資料 | 講習会](#) (last modified 2017/11/13) **②**
- [参考資料 | 講義、講演資料](#) (last modified 2019/03/02) **②**
- [過去のお知らせ](#) (last modified 2018/11/21)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2017](#) (last modified 2017/08/08)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2016](#) (last modified 2016/08/22)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2015](#) (last modified 2016/08/19)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGS速習コース2014](#) (last modified 2015/02/11)
- [書籍 | トランスクリプトーム解析 | について](#) (last modified 2014/05/12)
- [書籍 | トランスクリプトーム解析 | 2.3.1 RNA-seqデータ\(FASTQファイル\)](#) (last modified 2016/03/17)
- [書籍 | トランスクリプトーム解析 | 2.3.2 リファレンス配列](#) (last modified 2014/04/16)
- [書籍 | トランスクリプトーム解析 | 2.3.3 アノテーション情報](#) (last modified 2017/08/02)
- [書籍 | トランスクリプトーム解析 | 2.3.4 マッピング\(準備\)](#) (last modified 2014/06/20)
- [書籍 | トランスクリプトーム解析 | 2.3.5 マッピング\(本番\)](#) (last modified 2014/06/21)
- [書籍 | トランスクリプトーム解析 | 2.3.6 カウントデータ取得](#) (last modified 2016/02/09)
- [書籍 | トランスクリプトーム解析 | 3.3.1 解析目的別留意点](#) (last modified 2014/04/20)
- [書籍 | トランスクリプトーム解析 | 3.3.2 データの正規化\(基礎編\)](#) (last modified 2014/06/23) [トップページへ](#)
- [書籍 | トランスクリプトーム解析 | 3.3.3 クラスタリング](#) (last modified 2014/04/20)



# サブページ

①少しずつページ下部に移動して全体像を眺める。②は私の解説記事や、講習会の講義資料PDF、講義や講演資料PDFをカテゴリ別に時系列(新→古)で並べたリスト。例えば③「書籍、学会誌」をクリックすると、こんな感じになります。時系列なので、①2019年、②2018年、③2017年のような感じになっています。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~ka

## 参考資料 | 書籍、学会誌 NEW

基本的に私門田の個人ページに記載してあるものです。かなり古い講演資料などの情報をもとに勉強されている方もいらっしゃるようですので、ここでは2013年秋以降の情報のみです。大まかな内容についても述べています。

- 寺田朋子, 清水謙多郎, 門田幸二, 「[次世代シーケンサーデータの解析手法: 第13回RNA-seq解析\(その1\)](#)」, [日本乳酸菌学会誌](#), **30(1)**:38-45, 2019.  
内容: 日本乳酸菌学会誌のNGS関連連載の第13回分です。L. rhamnosus GGのRNA-seqデータ取得まで。この菌株の2種類のゲノム配列(FM179322とAP011548)を、Gepardというドットプロット用プログラムを用いて描画し、論文中に記載されている一部の反転領域を確認。反転領域の意味を完全に理解すべく、Rのseqinrパッケージを用いて仮想塩基配列で確認。L. rhamnosus GGのRNA-seqデータ(SRP125628 or GSE107337)の取得と、実験デザインの理解まで。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [第13回RNA-seq解析\(その1\)](#)」の項目をご覧ください。
- 藤博幸 編, よくわかるバイオインフォマティクス入門, 講談社, 2018. [ISBN: 978-4-06-513821-2](#).  
内容: 藤博幸先生によってまとめられたバイオインフォの入門書です。私は第5章NGSデータ概論と第7章トランスクリプトーム解析を担当しました。
- 寺田朋子, 大田達郎, 清水謙多郎, 門田幸二, 「[次世代シーケンサーデータの解析手法: 第12回Galaxy: ヒストリーとワークフロー](#)」, [日本乳酸菌学会誌](#), **29(2)**:79-88, 2018.  
内容: 日本乳酸菌学会誌のNGS関連連載の第12回分です。統合データ解析環境Galaxyの続きの内容です。ヒストリーの共有、ヒストリーの操作。ワークフローの作成、ワークフローの実行、ワークフローの編集と共有などの内容を含みます。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [第12回統合データ解析環境Galaxy](#)」の項目をご覧ください。
- 大田達郎, 寺田朋子, 清水謙多郎, 門田幸二, 「[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](#)」, [日本乳酸菌学会誌](#), **28(3)**:167-175, 2017.  
内容: 日本乳酸菌学会誌のNGS関連連載の第11回分です。Galaxyプロジェクト、解析ソフトとしての特徴、Galaxyの[トップページ](#)、Public Galaxy Serverの基本的な利用法、クオリティチェック(FastQC)、ファイルの型(Trimmomaticを例に)、[リミシング\(Trimmomatic\)](#)などGalaxyのイントロ的な内容です。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [第11回統合データ解析環境Galaxy](#)」の項目をご覧ください。



# サブページ

例えば、①2019年の最新のものは、私が2014年から②日本乳酸菌学会誌上で連載しているNGSデータ解析手法の、③第13回の内容。④大まかな内容も記載している。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html#reference\_materials1

## 参考資料 | 書籍、学会誌 NEW

基本的に私門田の個人ページに記載してあるものです。かなり古い資料などの情報をもとに勉強されている方もいらっしゃるようですので、ここでは2013年秋以降の情報のみです。大まかな内容も③も述べています。

- 寺田朋子, 清水謙多郎, 門田幸二, 「次世代シーケンサーデータの解析手法: 第13回RNA-seq解析(その1)」, 日本乳酸菌学会誌, 29(2):38-45, 2019.

④ 内容: 日本乳酸菌学会誌のNGS関連連載の第13回分です。L. rhamnosus GGのRNA-seqデータ取得まで。この菌株の2種類のゲノム配列(FM179322とAP011548)を、Gepardというドットプロット用プログラムを用いて描画し、論文中に記載されている一部の反転領域を確認。反転領域の意味を完全に理解すべく、Rのseqinrパッケージを用いて仮想塩基配列で確認。L. rhamnosus GGのRNA-seqデータ(SRP125628 or GSE107337)の取得と、実験デザインの理解まで。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | 第13回RNA-seq解析(その1)」の項目をご覧ください。

- 藤博幸 編, よくわかるバイオインフォマティクス入門, 講談社, 2018. ISBN: 978-4-06-513821-2.

内容: 藤博幸先生によってまとめられたバイオインフォの入門書です。私は第5章NGSデータ概論と第7章トランスクリプトーム解析を担当しました。

- 寺田朋子, 大田達郎, 清水謙多郎, 門田幸二, 「次世代シーケンサーデータの解析手法: 第12回Galaxy: ヒストリーとワークフロー」, 日本乳酸菌学会誌, 29(2):79-88, 2018.

内容: 日本乳酸菌学会誌のNGS関連連載の第12回分です。統合データ解析環境Galaxyの続きの内容です。ヒストリーの共有、ヒストリーの操作。ワークフローの作成、ワークフローの実行、ワークフローの編集と共有などの内容を含みます。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | 第12回統合データ解析環境Galaxy」の項目をご覧ください。

- 大田達郎, 寺田朋子, 清水謙多郎, 門田幸二, 「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」, 日本乳酸菌学会誌, 28(3):167-175, 2017.

内容: 日本乳酸菌学会誌のNGS関連連載の第11回分です。Galaxyプロジェクト、解析ソフトとしての特徴、Galaxyのトップページ、Public Galaxy Serverの基本的な利用法、クオリティチェック(FastQC)、ファイルの型(Trimmomaticを例に)、リミシング(Trimmomatic)などGalaxyのイントロ的な内容です。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 |

# サブページ

例えば、①2019年の最新のものは、私が2014年から②日本乳酸菌学会誌上で連載しているNGSデータ解析手法の、③第13回の内容。④大まかな内容も記載している。⑤は本科目の参考書としてもリストアップしている本。時系列にすることで、新情報を見つけやすいメリットがあります。⑥トップページへ移動。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~ka

## 参考資料 | 書籍、学会誌 **NEW**

基本的に私門田の個人ページに記載してあるものです。かなり古い講演資料などの情報をもとに勉強されている方もいらっしゃるようですので、ここでは2013年秋以降の情報のみです。大まかな内容についても述べています。

- 寺田朋子, 清水謙多郎, 門田幸二, 「次世代シーケンサーデータの解析手法: 第13回RNA-seq解析(その1)」, 日本乳酸菌学会誌, **30(1)**:38-45, 2019.

**内容:** 日本乳酸菌学会誌のNGS関連連載の第13回分です。L. rhamnosus GGのRNA-seqデータ取得まで。この菌株の2種類のゲノム配列(FM179322とAP011548)を、Gepardというドットプロット用プログラムを用いて描画し、論文中に記載されている一部の反転領域を確認。反転領域の意味を完全に理解すべく、Rのseqinrパッケージを用いて仮想塩基配列で確認。L. rhamnosus GGのRNA-seqデータ(SRP125628 or GSE107337)の取得と、実験デザインの理解まで。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [第13回RNA-seq解析\(その1\)](#)」の項目をご覧ください。

- 藤博幸 編, よくわかるバイオインフォマティクス入門, 講談社, 2018. ISBN: 978-4-06-513821-2.

**内容:** 藤博幸先生によってまとめられたバイオインフォの入門書です。私は第5章NGSデータ概論と第7章トランスクリプトーム解析を担当しました。

- 寺田朋子, 大田達郎, 清水謙多郎, 門田幸二, 「次世代シーケンサーデータの解析手法: 第12回Galaxy: ヒストリーとワークフロー」, 日本乳酸菌学会誌, **29(2)**:79-88, 2018.

**内容:** 日本乳酸菌学会誌のNGS関連連載の第12回分です。統合データ解析環境Galaxyの続きの内容です。ヒストリーの共有、ヒストリーの操作。ワークフローの作成、ワークフローの実行、ワークフローの編集と共有などの内容を含みます。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [第12回統合データ解析環境Galaxy](#)」の項目をご覧ください。

- 大田達郎, 寺田朋子, 清水謙多郎, 門田幸二, 「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」, 日本乳酸菌学会誌, **28(3)**:167-175, 2017.

**内容:** 日本乳酸菌学会誌のNGS関連連載の第11回分です。Galaxyプロジェクト、解析ソフトとしての特徴、Galaxyの[トップページ](#)、Public Galaxy Serverの基本的な利用法、クオリティチェック(FastQC)、ファイルの型(Trimmomaticを例に)、トリミング(Trimmomatic)などGalaxyのイントロ的な内容です。書籍中のリンク先やウェブ資料などは「書籍 | 日本乳酸菌学会誌 | [Galaxyのイントロ](#)」の項目をご覧ください。

5

6

# サブページ

再びサブページのこのあたりを解説。①講習会のところは、最近のものは②の講義資料PDFが中心です。特に②は、年度間で講義内容が被っているものもあり、中身も古くなっているが、特にNGS解析の実務と絡めたLinuxの基本的な事柄が有用だと思います。スライド14と関連

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~ka

プトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_  
\_)m(2018/10/15)

- [はじめに](#) (last modified 2018/07/18)
- [参考資料 | 書籍、学術誌](#) (last modified 2017/11/13)
- [参考資料 | 講習会](#) (last modified 2017/11/13)
- [参考資料 | 講義、講義資料](#) (last modified 2019/03/02) **NEW**
- [過去のお知らせ](#) (last modified 2018/11/21)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2017](#) (last modified 2017/08/08)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2016](#) (last modified 2016/08/22)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGSハンズオン講習会2015](#) (last modified 2016/08/19)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | NGS速習コース2014](#) (last modified 2015/02/11)
- [書籍 | トランスクリプトーム解析 | について](#) (last modified 2014/05/12)
- [書籍 | トランスクリプトーム解析 | 2.3.1 RNA-seqデータ\(FASTQファイル\)](#) (last modified 2016/03/17)
- [書籍 | トランスクリプトーム解析 | 2.3.2 リファレンス配列](#) (last modified 2014/04/16)
- [書籍 | トランスクリプトーム解析 | 2.3.3 アノテーション情報](#) (last modified 2017/08/02)
- [書籍 | トランスクリプトーム解析 | 2.3.4 マッピング\(準備\)](#) (last modified 2014/06/20)
- [書籍 | トランスクリプトーム解析 | 2.3.5 マッピング\(本番\)](#) (last modified 2014/06/21)
- [書籍 | トランスクリプトーム解析 | 2.3.6 カウントデータ取得](#) (last modified 2016/02/09)
- [書籍 | トランスクリプトーム解析 | 3.3.1 解析目的別留意点](#) (last modified 2014/04/20)
- [書籍 | トランスクリプトーム解析 | 3.3.2 データの正規化\(基礎編\)](#) (last modified 2014/06/23) [トップページへ](#)
- [書籍 | トランスクリプトーム解析 | 3.3.3 クラスタリング](#) (last modified 2014/04/20)



# サブページ

また少しページ下部に移動したところ。①が日本乳酸菌学会誌のNGS連載誌上で連載しているNGSデータ解析手法のウェブ資料のサイト。アグリバイオの講義の枠組みとは無関係なので、ダウンロードに時間がかかる巨大なファイルを取り扱って、計算時間のかかるプログラムをウェブツールを用いて行うやり方など、より実用性の高いものを解説。しかし第6回(2016年3月)で紹介したDDBJ Pipelineは、2019年2月にサービスが終了しました。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~ka

- 書籍 | トランスクリプトーム解析 | [3.3.3 クラスタリング](#) (last modified 2017/07/27)
- 書籍 | トランスクリプトーム解析 | [3.3.4 各種プロット](#) (last modified 2017/07/27)
- 書籍 | トランスクリプトーム解析 | [4.3.1 シミュレーションデータ](#) (last modified 2017/07/27)
- 書籍 | トランスクリプトーム解析 | [4.3.2 データの正規化\(応用編\)](#) (last modified 2017/07/27)
- 書籍 | トランスクリプトーム解析 | [4.3.3 2群間比較](#) (last modified 2014/04/28)
- 書籍 | トランスクリプトーム解析 | [4.3.4 他の実験デザイン\(3群間\)](#) (last modified 2014/04/28)
- **書籍 | [日本乳酸菌学会誌 | について](#) (last modified 2018/09/03)**
- 書籍 | 日本乳酸菌学会誌 | [第1回イントロダクション](#) (last modified 2018/09/03)
- 書籍 | 日本乳酸菌学会誌 | [第2回GUI環境からコマンドライン環境へ](#) (last modified 2018/09/03)
- 書籍 | 日本乳酸菌学会誌 | [第3回Linux環境構築からNGSデータ取得まで](#) (last modified 2017/07/02)
- 書籍 | 日本乳酸菌学会誌 | [第4回クオリティコントロールとプログラムのインストール](#) (last modified 2018/05/10)
- 書籍 | 日本乳酸菌学会誌 | [第5回アセンブル、マッピング、そしてQC](#) (last modified 2017/06/25)
- 書籍 | 日本乳酸菌学会誌 | [第6回ゲノムアセンブリ](#) (last modified 2017/06/21)
- 書籍 | 日本乳酸菌学会誌 | [第7回ロングリードアセンブリ](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第8回アセンブリ後の解析](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第9回ゲノムアノテーションとその可視化、DDBJへの登録](#) (last modified 2017/03/13)
- 書籍 | 日本乳酸菌学会誌 | [第10回DDBJへの塩基配列の登録\(後編\)](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第11回統合データ解析環境Galaxy](#) (last modified 2017/11/13)
- 書籍 | 日本乳酸菌学会誌 | [第12回Galaxy : ヒストリーとワークフロー](#) (last modified 2018/07/04)
- 書籍 | 日本乳酸菌学会誌 | [第13回RNA-seq解析 \(その1\)](#) (last modified 2019/01/31)

[トップページへ](#)



# サブページ

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html#

- 書籍 | トランスクリプトーム解析 | [3.3.3 クラスタリング](#) (last modified 2014/04/20)
- 書籍 | トランスクリプトーム解析 | [3.3.4 各種プロット](#) (last modified 2014/04/27)
- 書籍 | トランスクリプトーム解析 | [4.3.1 シミュレーションデータ\(負の二項分布\)](#) (last modified 2014/04/27)
- 書籍 | トランスクリプトーム解析 | [4.3.2 データの正規化\(応用編\)](#) (last modified 2014/04/27)
- 書籍 | トランスクリプトーム解析 | [4.3.3 2群間比較](#) (last modified 2014/04/28)
- 書籍 | トランスクリプトーム解析 | [4.3.4 他の実験デザイン\(3群間\)](#) (last modified 2014/04/28)
- 書籍 | [日本乳酸菌学会誌](#) | [について](#) (last modified 2018/09/03)
- 書籍 | 日本乳酸菌学会誌 | [第1回イントロダクション](#) (last modified 2018/09/03)
- 書籍 | 日本乳酸菌学会誌 | [第2回GUI環境からコマンドライン環境へ](#) (last modified 2018/09/03)
- 書籍 | 日本乳酸菌学会誌 | [第3回Linux環境構築からNGSデータ取得まで](#) (last modified 2017/07/02)
- 書籍 | 日本乳酸菌学会誌 | [第4回クオリティコントロールとプログラムのインストール](#) (last modified 2018/05/10)
- 書籍 | 日本乳酸菌学会誌 | [第5回アセンブル、マッピング、そしてQC](#) (last modified 2017/06/25)
- 書籍 | 日本乳酸菌学会誌 | [第6回ゲノムアセンブリ](#) (last modified 2017/06/21)
- 書籍 | 日本乳酸菌学会誌 | [第7回ロングリードアセンブリ](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第8回アセンブリ後の解析](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第9回ゲノムアノテーションとその可視化、DDBJへの登録](#) (last modified 2017/03/13)
- 書籍 | 日本乳酸菌学会誌 | [第10回DDBJへの塩基配列の登録\(後編\)](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第11回統合データ解析環境Galaxy](#) (last modified 2017/11/13)
- 書籍 | 日本乳酸菌学会誌 | [第12回Galaxy : ヒストリーとワークフロー](#) (last modified 2018/07/04)
- 書籍 | 日本乳酸菌学会誌 | [第13回RNA-seq解析 \(その1\)](#) (last modified 2019/01/31)



[トップページへ](#)



# サブページ

ちなみに1週間程度かかる①日本乳酸菌学会誌のNGS連載第1-4回分の内容を理解するのが、②H27-29年度NGS講習会の必須予習事項でした。

(Rで)塩基配列解析のサブ

保護されていない通信 | www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq2.html

プトで書かれているかも理解できる仕様になっています。バグレポートやリクエストなどをいただければ幸いですm(\_  
\_)m(2018/10/15)

- [はじめに](#) (last modified 2018/07/18)
- 参考資料 | [書籍、学会誌](#) (last modified 2017/11/13)
- 参考資料 | [講習会](#) (last modified 2017/11/13)
- 参考資料 | [講義、講演資料](#) (last modified 2019/03/02) **NEW**
- [過去のお知らせ](#) (last modified 2018/11/21)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | \*\*NGSハンズオン講習会2017\*\*](#) (last modified 2017/08/08)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | \*\*NGSハンズオン講習会2016\*\*](#) (last modified 2016/08/22)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | \*\*NGSハンズオン講習会2015\*\*](#) (last modified 2016/08/19)
- [バイオインフォマティクス人材育成カリキュラム\(次世代シーケンサ\) | \[NGS速習コース\]\(#\)2014](#) (last modified 2015/02/11)
- [書籍 | \[トランスクリプトーム解析 | \\[について\\]\\(#\\)\]\(#\)](#) (last modified 2014/05/12)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.1 RNA-seqデータ\\\(FASTQファイル\\\)\\]\\(#\\)\]\(#\)](#) (last modified 2016/03/17)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.2 リファレンス配列\\]\\(#\\)\]\(#\)](#) (last modified 2014/04/16)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.3 アノテーション情報\\]\\(#\\)\]\(#\)](#) (last modified 2017/08/02)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.4 マッピング\\\(準備\\\)\\]\\(#\\)\]\(#\)](#) (last modified 2014/06/20)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.5 マッピング\\\(本番\\\)\\]\\(#\\)\]\(#\)](#) (last modified 2014/06/21)
- [書籍 | \[トランスクリプトーム解析 | \\[2.3.6 カウントデータ取得\\]\\(#\\)\]\(#\)](#) (last modified 2016/02/09)
- [書籍 | \[トランスクリプトーム解析 | \\[3.3.1 解析目的別留意点\\]\\(#\\)\]\(#\)](#) (last modified 2014/04/20)
- [書籍 | \[トランスクリプトーム解析 | \\[3.3.2 データの正規化\\\(基礎編\\\)\\]\\(#\\)\]\(#\)](#) (last modified 2014/06/23) [トップページへ](#)
- [書籍 | \[トランスクリプトーム解析 | \\[3.3.3 クラスタリング\\]\\(#\\)\]\(#\)](#) (last modified 2014/04/20)



# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# 基本的な利用法

おさらい。①(Rで)塩基配列解析は、スライド4で述べたアグリバイオの全体方針に従って、主にRでゲノム・トランスクリプトーム解析を行うために作り込んできたウェブサイト。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., Bioinformatics, 2012)

学習教材 : バイオインフォマティクス人材育成のための講習会(平成26-29年度)

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio



# 基本的な利用法

おさらい。①(Rで)塩基配列解析は、スライド4で述べたアグリバイオの全体方針に従って、主にRでゲノム・トランスクリプトーム解析を行うために作り込んできたウェブサイト。後半はRの②基本的な利用法を学びます。③Windows版の内容を中心に話しますが、④Macintosh版についてもWindows版と異なる部分を左側のスクリーンで示します。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外): ISCB

学会(国内): JSBi

QAサイト: Biostar (Parnell)

QAサイト: SEQanswers (Parnell)

学習教材: バイオインフォマ

学習教材: (Rで)塩基配列解析

学習教材: (Rで)塩基配列解析

RStudio

(Rで)塩基配列解析  
(last modified 2019/03/11, since 2010)

このウェブページのR関連部分は、[インストール | についての推奨手順](#) ([Windows2018.11.15版](#)と[Macintosh2018.11.27版](#))に従ってフリーソフトRと必要なパッケージをインストール済みであるという前提で記述しています。初心者の方は[基本的な利用法](#) ([Windows2018.12.23版](#)と[Macintosh2019.01.15版](#))で自習してください。2018年7月に(Rで)塩基配列解析の一部(講習会・書籍・[雑誌](#)など)を切り分け、[ページ](#)に移行しました。(2018/07/18)

What's new? ([過去のお知らせ](#)はこちら)

- 2019年度も[アグリバイオインフォマティクス教育研究プログラム](#)を実施します。例年東大以外の企業の方、研究員、大学院生が2割程度受講しております。受講ガイダンスは、2019年4月5日(Fri.)17:15より東大農学部2号館2階化学第一講義室で開催します。(2019/03/11) **NEW**
- 細かいところの修正はここに明記してなくても随時行っています。(2019/03/11) **NEW**
- 「インストール | Rパッケージ | [必要最小限プラスアルファ](#)」を更新しました。(2019/03/08) **NEW**
- 「[生命科学データ解析を支える情報技術](#) (監修: 坊農秀雅)」が出版されています。最先端のネタを含むかなり広範な内容を含んでいますので、一通り目次を眺めてみるとよいと思います。Bioconda, Homebrew, Docker, GitHub, EC2, AWSなど聞いたことがある有用そうなものの全体像がわかるというメリットがあると思います。(2019/02/06)

[はじめに](#) (last modified 2018/08/04) [トップページへ](#)

[過去のお知らせ](#) (last modified 2019/03/02) **NEW**

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

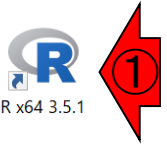
## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# Rの起動



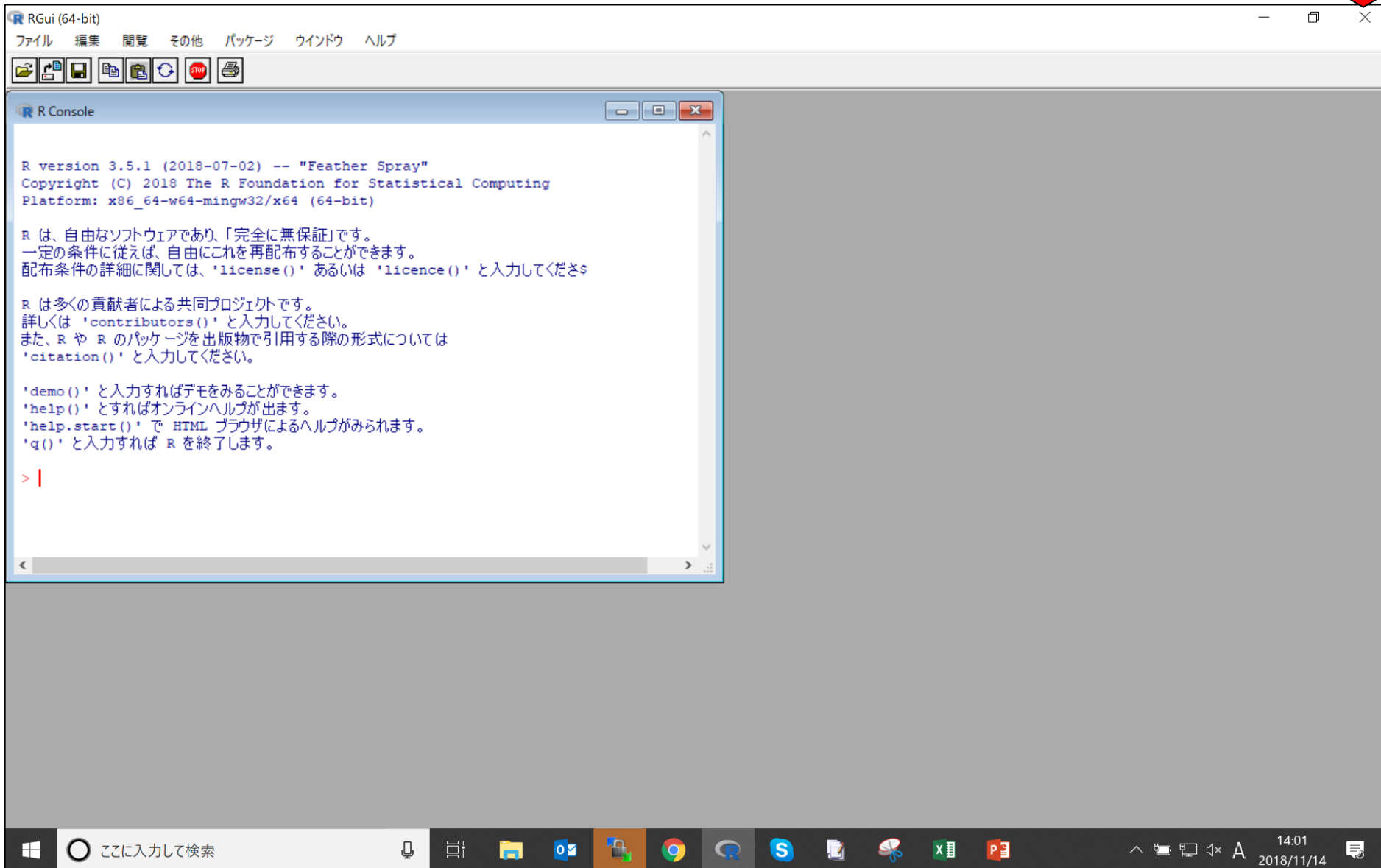
ごみ箱



R x64 3.5.1



# Rの起動



# Rの起動



ごみ箱



R x64 3.5.1

```
RGui (64-bit)
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ
[R Console]
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。

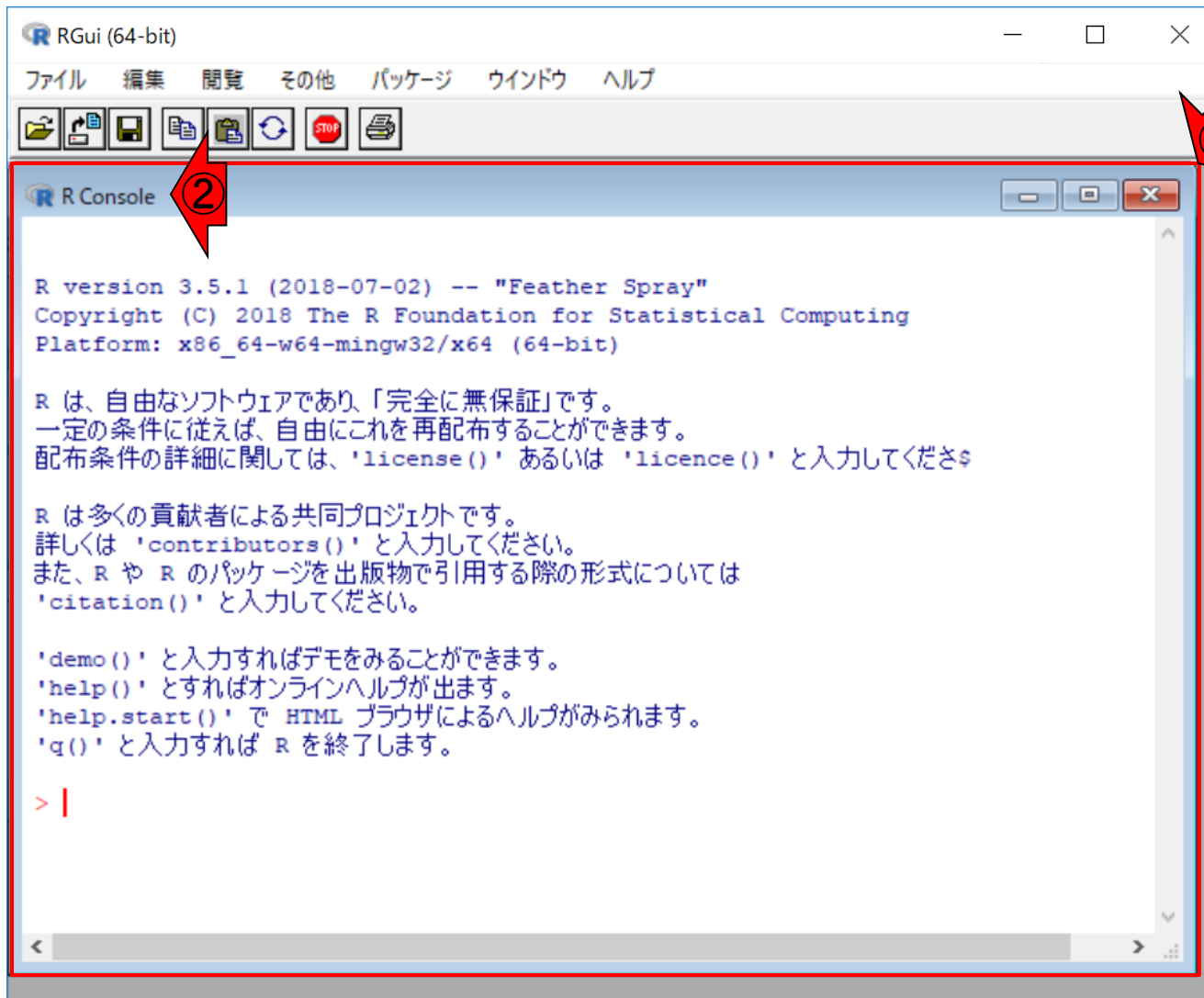
R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

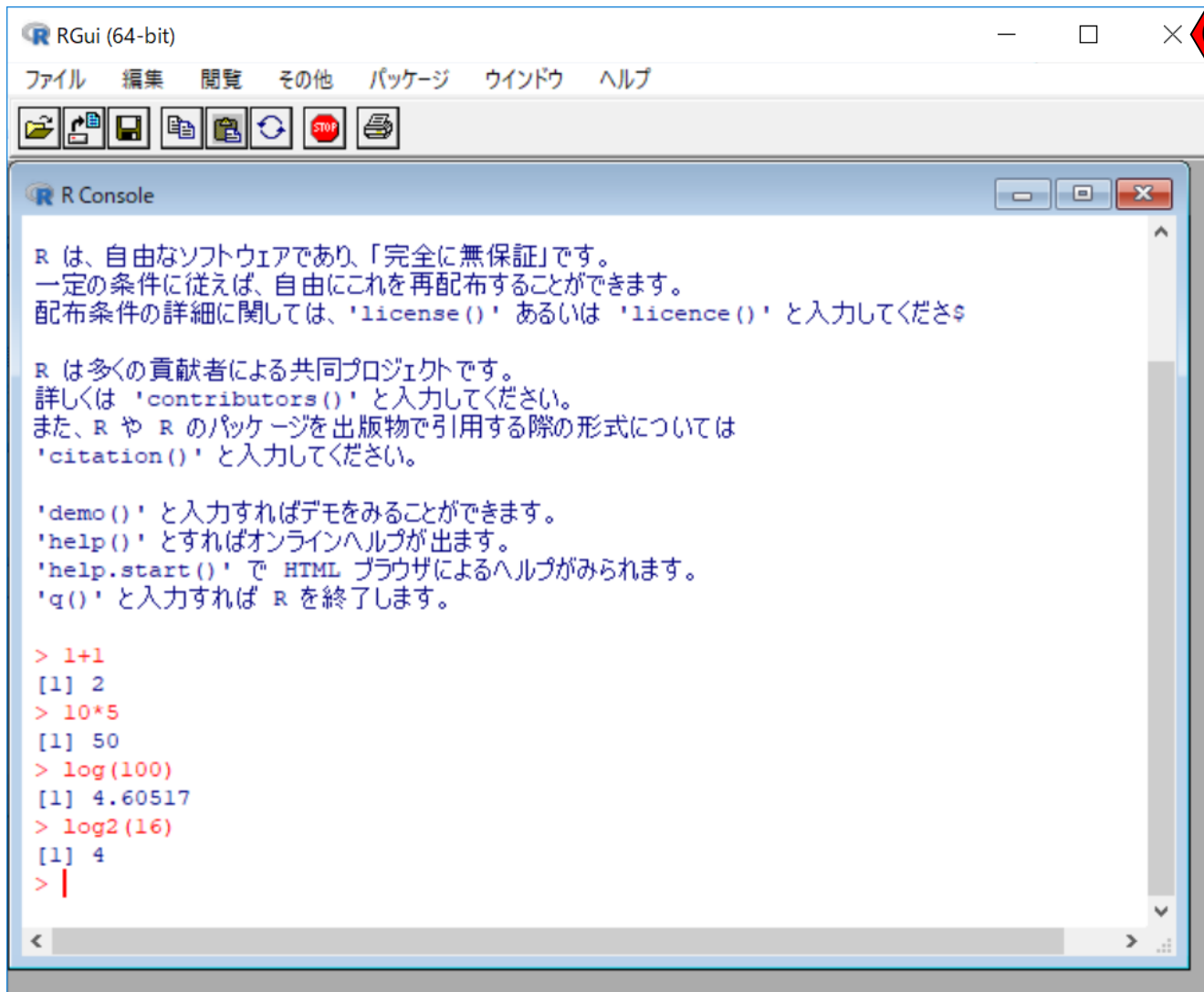
> |
```



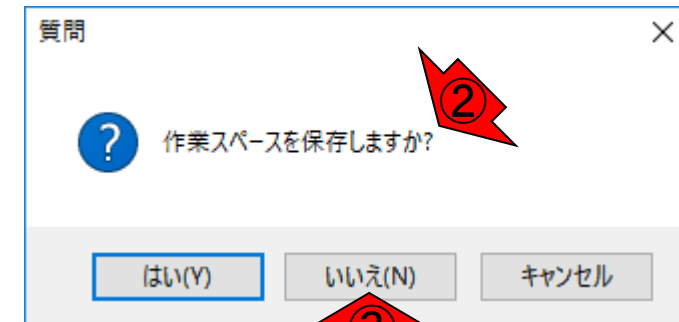
# Rの起動



# Rの終了



通常のソフトウェアと同様、①右上の×ボタンを押せばよい。②「作業スペースを保存しますか?」というダイアログが出る。この意味が不明な最初のうちは③いいえでよい。間違っても「はい」を押してしまっても.Rdataと.Rhistoryという2つのファイルが作成されるだけなので特に問題はない



# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

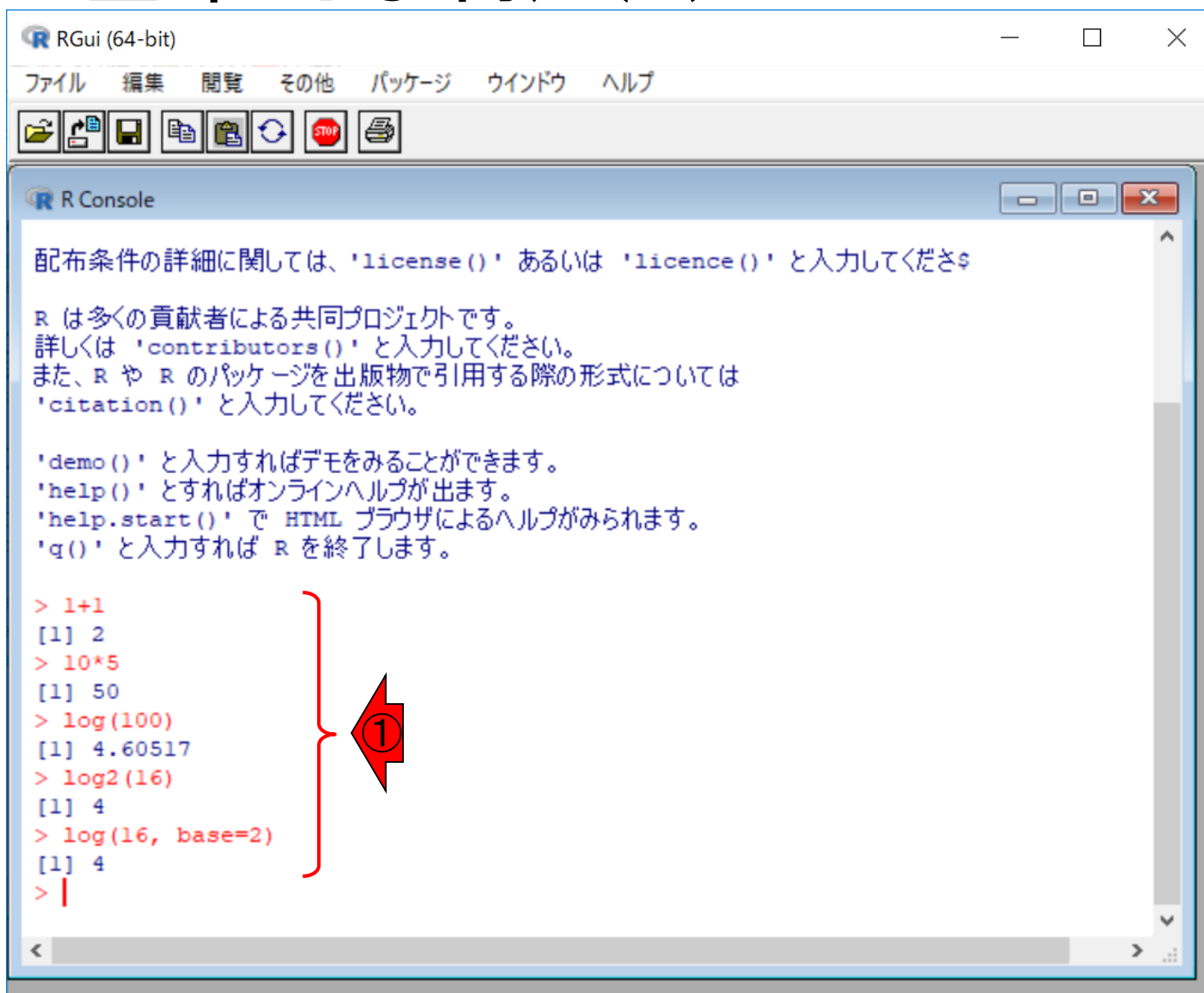
## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# 基本的な利用法



RGui (64-bit)

ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ

R Console

配布条件の詳細に関しては、`'license()'` あるいは `'licence()'` と入力してください。

R は多くの貢献者による共同プロジェクトです。  
詳しくは `'contributors()'` と入力してください。  
また、R や R のパッケージを出版物で引用する際の形式については  
`'citation()'` と入力してください。

`'demo()'` と入力すればデモをみることができます。  
`'help()'` とすればオンラインヘルプが出ます。  
`'help.start()'` で HTML ブラウザによるヘルプがみられます。  
`'q()'` と入力すれば R を終了します。

```
> 1+1
[1] 2
> 10*5
[1] 50
> log(100)
[1] 4.60517
> log2(16)
[1] 4
> log(16, base=2)
[1] 4
> |
```

A red bracket and arrow with the number 1 point to the numerical calculation results in the console.





# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# 翻訳配列取得

- イントロ | 一般 | [任意の位置の塩基を置換](#) (last modified 2013/09/12)
- イントロ | 一般 | [指定した範囲の配列を取得](#) (last modified 2015/04/06)
- イントロ | 一般 | [指定したID\(染色体やdescription\)の配列を取得](#) (last modified 2014/03/10)
- イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | [Biostrings](#) (last modified 2015/09/12)
- イントロ | 一般 | 翻訳配列(translate)を取得(応用) | [seqinr\(Cham 2005\)](#) (last modified 2015/03/09)
- イントロ | 一般 | [相補鎖\(complement\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | [逆相補鎖\(reverse complement\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | [逆鎖\(reverse\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | k-mer解析 | k=1(塩基ごとの出現頻度解析) | [Biostrings](#) (last modified 2016/04/27)
- イントロ | 一般 | k-mer解析 | k=2(2連続塩基の出現頻度解析) | [Biostrings](#) (last modified 2016/01/28)
- イントロ | 一般 | k-mer解析 | k=3(3連続塩基の出現頻度解析) | [Biostrings](#) (last modified 2016/01/28)
- イントロ | 一般 | k-mer解析 | k=n(n連続塩基の出現頻度解析) | [Biostrings](#) (last modified 2016/05/01)
- イントロ | 一般 | Tips | [任意の拡張子でファイルを保存](#) (last modified 2013/09/26)
- イントロ | 一般 | Tips | [拡張子は同じで任意の文字を追加して保存](#) (last modified 2013/09/26)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [公共DBから](#) (last modified 2017/04/11)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [BSgenome](#) (last modified 2015/04/22)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [公共DBから](#) (last modified 2017/04/11)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [BSgenomeとTxDbから](#) (last modified 2015/02/20)

# 翻訳配列取得

①をクリックすると、②のようなページに移動します。ここでは、塩基配列を入力として、その翻訳されたアミノ酸配列を取得するための項目です。クリックする場所を間違えた場合は、③トップページへ

- イントロ | 一般 | [任意の位置の塩基を置換](#) (last modified 2013/09/12)
- イントロ | 一般 | [指定した範囲の配列を取得](#) (last modified 2015/04/06)
- イントロ | 一般 | [指定したID\(染色体やdescription\)の配列を取得](#) (last modified 2014/03/10)
- イントロ | 一般 | [翻訳配列\(translate\)を取得\(基礎\)](#) | [Biostrings](#) ① (last modified 2015/09/12)
- イントロ | 一般 | [翻訳配列\(translate\)を取得\(応用\)](#) | [seqinr\(Cham 2005\)](#) (last modified 2015/03/09)
- イントロ | 一般 | [相補鎖\(complement\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | [逆相補鎖\(reverse complement\)を取](#)
- イントロ | 一般 | [逆鎖\(reverse\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | k-mer解析 | k=1(塩基ごとの出現頻度)
- イントロ | 一般 | k-mer解析 | k=2(2連続塩基の出現頻度)
- イントロ | 一般 | k-mer解析 | k=3(3連続塩基の出現頻度) ②
- イントロ | 一般 | k-mer解析 | k=n(n連続塩基の出現頻度)
- イントロ | 一般 | Tips | [任意の拡張子でファイルを保存](#)
- イントロ | 一般 | Tips | [拡張子は同じで任意の文字を指定](#)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [公共DBから](#)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [BSgenome](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [公共DB](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [BSgenome](#)

ドキュメント (Rで)塩基配列解析

← → ↻ 🏠 ⓘ 保護されていない | www.iu.a.u-tokyo.ac... 🔍 ☆ 👤 ⋮

## イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | Biostrings

[Biostrings](#)パッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するやり方を示します。翻訳のための**遺伝コード(genetic code)**は、Standard Genetic Codeだそうです。もちろん生物種?!によって多少違い(variants)があるようで、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"などいろいろ選べるようです。「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてあるディレクトリに移動し以下をコピー。

### 1. FASTA形式ファイル(sample1.fasta)の場合 :

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta" #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta" #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta #確認してるだけです

#本番
fasta <- translate(fasta) #アミノ酸配列に翻訳した結果をfastaに格納
fasta #確認してるだけです

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50) #fastaの中身を指定したファイル名で保
```

③ [トップページへ](#)

# hogeフォルダの作成

①デスクトップにあるhogeフォルダ中のファイルを解析するやり方として説明します。よく「hogeって何ですか?」と質問されるのですが、特別な意味はありません。嫌ならugeでもなんでも好きな名前を自己責任でつけてください。

RGui (64-bit)

ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ



R Console

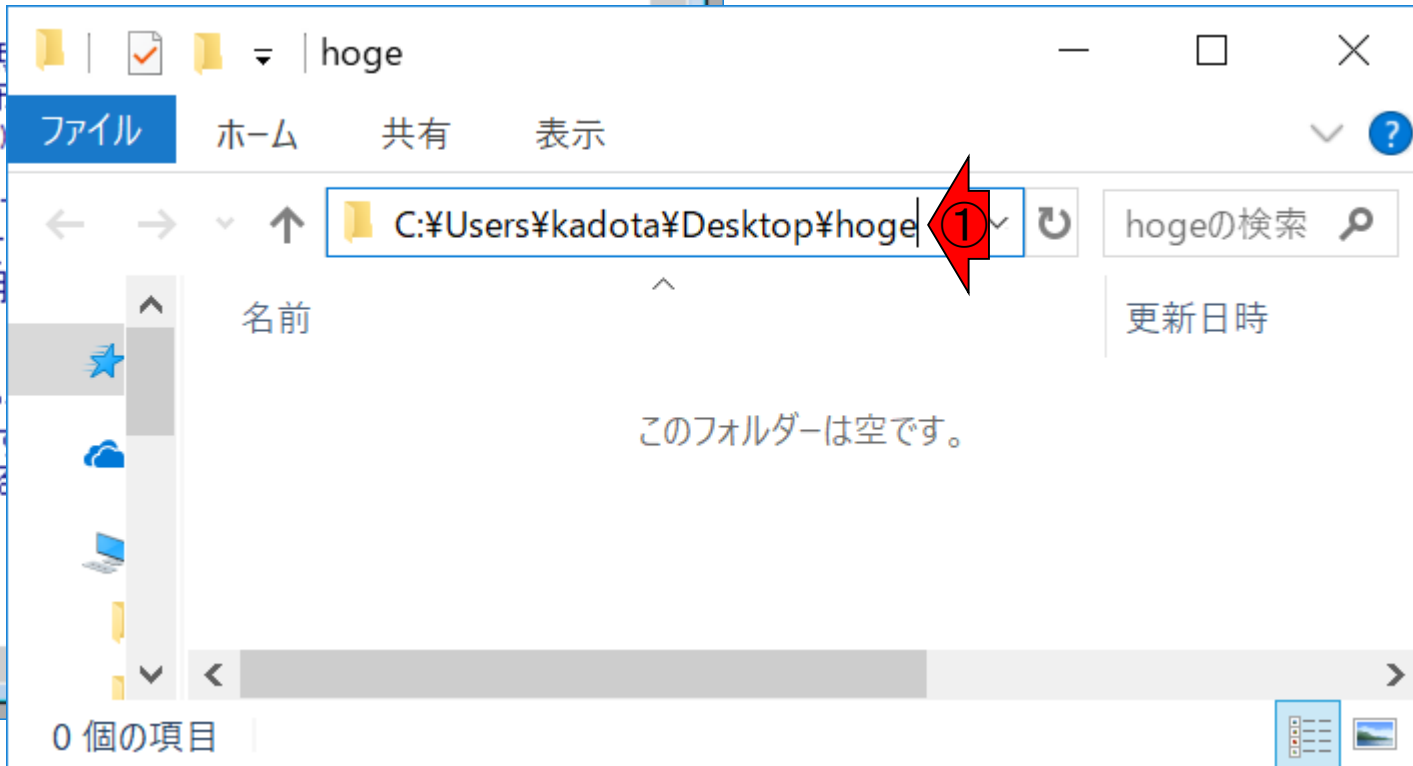
```
R version 3.5.1 (2018-07-02) -- "Feather Spray"  
Copyright (C) 2018 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)
```

R は、自由なソフトウェアであり、「完全に無  
一定の条件に従えば、自由にこれを再配布  
配布条件の詳細に関しては、'license()'

R は多くの貢献者による共同プロジェクトで  
詳しくは 'contributors()' と入力して  
また、R や R のパッケージを出版物で引用  
'citation()' と入力してください。

'demo()' と入力すればデモをみることがで  
'help()' とすればオンラインヘルプが出ま  
'help.start()' で HTML ブラウザによる  
'q()' と入力すれば R を終了します。

```
> |
```





①項目名からもある程度わかるようになっていますが、  
②の部分にもう少し詳細な説明や解説があります。

# 翻訳配列取得

- イントロ | 一般 | [任意の位置の塩基を置換](#) (last modified 2013/09/12)
- イントロ | 一般 | [指定した範囲の配列を取得](#) (last modified 2015/04/06)
- イントロ | 一般 | [指定したID\(染色体やdescription\)の配列を取得](#) (last modified 2014/03/10)
- イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | [Biostrings](#) (last modified 2015/09/12)
- イントロ | 一般 | 翻訳配列(translate)を取得(応用) | [seqinr\(Charif\\_2005\)](#) (last modified 2015/03/09)
- イントロ | 一般 | [相補鎖\(complement\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | [逆相補鎖\(reverse complement\)を取](#)
- イントロ | 一般 | [逆鎖\(reverse\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | k-mer解析 | k=1(塩基ごとの出現頻度)
- イントロ | 一般 | k-mer解析 | k=2(2連続塩基の出現頻度)
- イントロ | 一般 | k-mer解析 | k=3(3連続塩基の出現頻度)
- イントロ | 一般 | k-mer解析 | k=n(n連続塩基の出現頻度)
- イントロ | 一般 | Tips | [任意の拡張子でファイルを作成](#) ②
- イントロ | 一般 | Tips | [拡張子は同じで任意の文字を指定](#)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [公共DBから](#)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [BSgenome](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [公共DB](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [BSgenome](#)

イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | Biostrings ①

Biostringsパッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するやり方を示します。翻訳のための遺伝コード(genetic code)は、Standard Genetic Codeだそうです。もちろん生物種?!によって多少違い(variants)があるようで、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"などいろいろ選べるようです。

「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてあるディレクトリに移動し以下をコピー。

1. FASTA形式ファイル(sample1.fasta)の場合：

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta" #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta" #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta #確認してるだけです

#本番
fasta <- translate(fasta) #アミノ酸配列に翻訳した結果をfastaに格納
fasta #確認してるだけです

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50) #fastaの中身を指定したファイル名で保
```

[トップページ](#)



①1つの項目内には大抵複数の例題があります。  
ここでは、②例題1をやっていきます。

# 翻訳配列取得

- イントロ | 一般 | [任意の位置の塩基を置換](#) (last modified 2013/09/12)
- イントロ | 一般 | [指定した範囲の配列を取得](#) (last modified 2015/04/06)
- イントロ | 一般 | [指定したID\(染色体やdescription\)の配列を取得](#) (last modified 2014/03/10)
- イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | [Biostrings](#) (last modified 2015/09/12)
- イントロ | 一般 | 翻訳配列(translate)を取得(応用) | [seqinr\(Charif\\_2005\)](#) (last modified 2015/03/09)
- イントロ | 一般 | [相補鎖\(complement\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | [逆相補鎖\(reverse complement\)を取](#)
- イントロ | 一般 | [逆鎖\(reverse\)を取得](#) (last modified 2013/06/14)
- イントロ | 一般 | k-mer解析 | k=1(塩基ごとの出現頻度)
- イントロ | 一般 | k-mer解析 | k=2(2連続塩基の出現頻度)
- イントロ | 一般 | k-mer解析 | k=3(3連続塩基の出現頻度)
- イントロ | 一般 | k-mer解析 | k=n(n連続塩基の出現頻度)
- イントロ | 一般 | Tips | [任意の拡張子でファイルを保存](#)
- イントロ | 一般 | Tips | [拡張子は同じで任意の文字を](#)
- イントロ | 一般 | 配列取得 | ゲノム配列 | [公共DBから](#) ②
- イントロ | 一般 | 配列取得 | ゲノム配列 | [BSgenome](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [公共DBから](#)
- イントロ | 一般 | 配列取得 | プロモーター配列 | [BSgenome](#)

イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | Biostrings ①

Biostringsパッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するやり方を示します。翻訳のための遺伝コード(genetic code)は、Standard Genetic Codeだそうです。もちろん生物種?!によって多少違い(variants)があるようで、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"などいろいろ選べるようです。「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてあるディレクトリに移動し以下をコピー。

1. FASTA形式ファイル(sample1.fasta)の場合 :

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta" #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta" #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta #確認してるだけです

#本番
fasta <- translate(fasta) #アミノ酸配列に翻訳した結果をfastaに格納
fasta #確認してるだけです

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50) #fastaの中身を指定したファイル名で保
```

[トップページ](#)

# 入力ファイルの保存

①例題1は、②sample1.fastaというファイルを入力として利用します。赤枠部分のみを拡大表示します。

(Rで)塩基配列解析

保護されていない | www.iu.a.u-tokyo.ac...

## イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | Biostrings

Biostringsパッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するやり方を示します。翻訳のための遺伝コード(genetic code)は、Standard Genetic Codeだそうです。もちろん生物種?!によって多少違い(variants)があるようで、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"などいろいろ選べるようです。

「ファイル」 - 「ディレクトリ」で解析したいファイルを置いてあるディレクトリに移動し以下をコピペ。

### 1. FASTA形式ファイル(sample1.fasta)の場合:

multi-FASTAではないsingle-FASTA形式ファイルで。

```
in_f <- "sample1.fasta"      #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta"      #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNASTringSet(in_f, format="fasta")#in_fで指定したファイルの読み込み
fasta                        #確認してるだけです

#本番
fasta <- translate(fasta)   #アミノ酸配列に翻訳した結果をfastaに格納
fasta                       #確認してるだけです

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50)#fastaの中身を指定したファイル名で保
```

[トップページへ](#)

# 入力ファイルの保存

①sample1.fastaを、先ほどデスクトップ上に作成した②hogeフォルダ内に保存します。

## 1. FASTA形式ファイル(sample1.fasta)①場合:

multi-FASTAではないsingle-FASTA形式ファイルです。

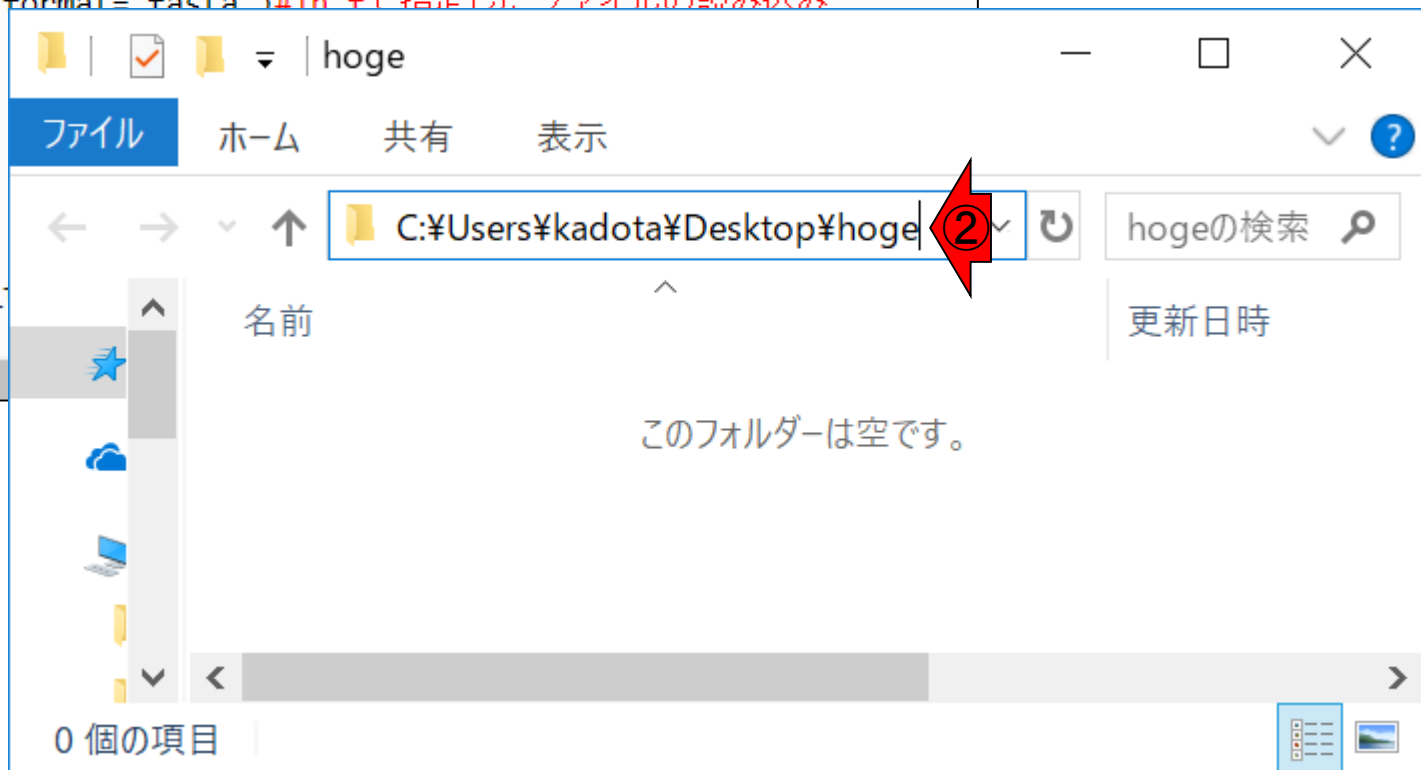
```
in_f <- "sample1.fasta" #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta" #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNASTringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta

#本番
fasta <- translate(fasta)
fasta

#ファイルに保存
writeXStringSet(fasta, file=out_f)
```



# 入力ファイルの保存

①sample1.fasta上で、右クリックで②「...リンク先を保存」。保存先はもちろん③ここ

## 1. FASTA形式ファイル(sample1.fasta)の場合

multi-FASTAではないsingle-FASTA形式

```
in_f <- "sample1.fasta"  
out_f <- "hoge1.fasta"
```

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

```
#入力ファイルの読み込み  
fasta <- readDNASTringSet(in_f,  
fasta)
```

```
#本番  
fasta <- translate(fasta)  
fasta
```

```
#ファイルに保存  
writeXStringSet(fasta, file=out_f)
```

新しいタブで開く(T)  
新しいウィンドウで開く(W)  
シークレット ウィンドウで開く(G)  
名前を付けてリンク先を保存(K)...  
リンクのアドレスをコピー(E)  
検証(I) Ctrl+Shift+I

ファイル ホーム 共有 表示

C:\Users\kadota\Desktop\hoge

hogeの検索

更新日時

このフォルダーは空です。

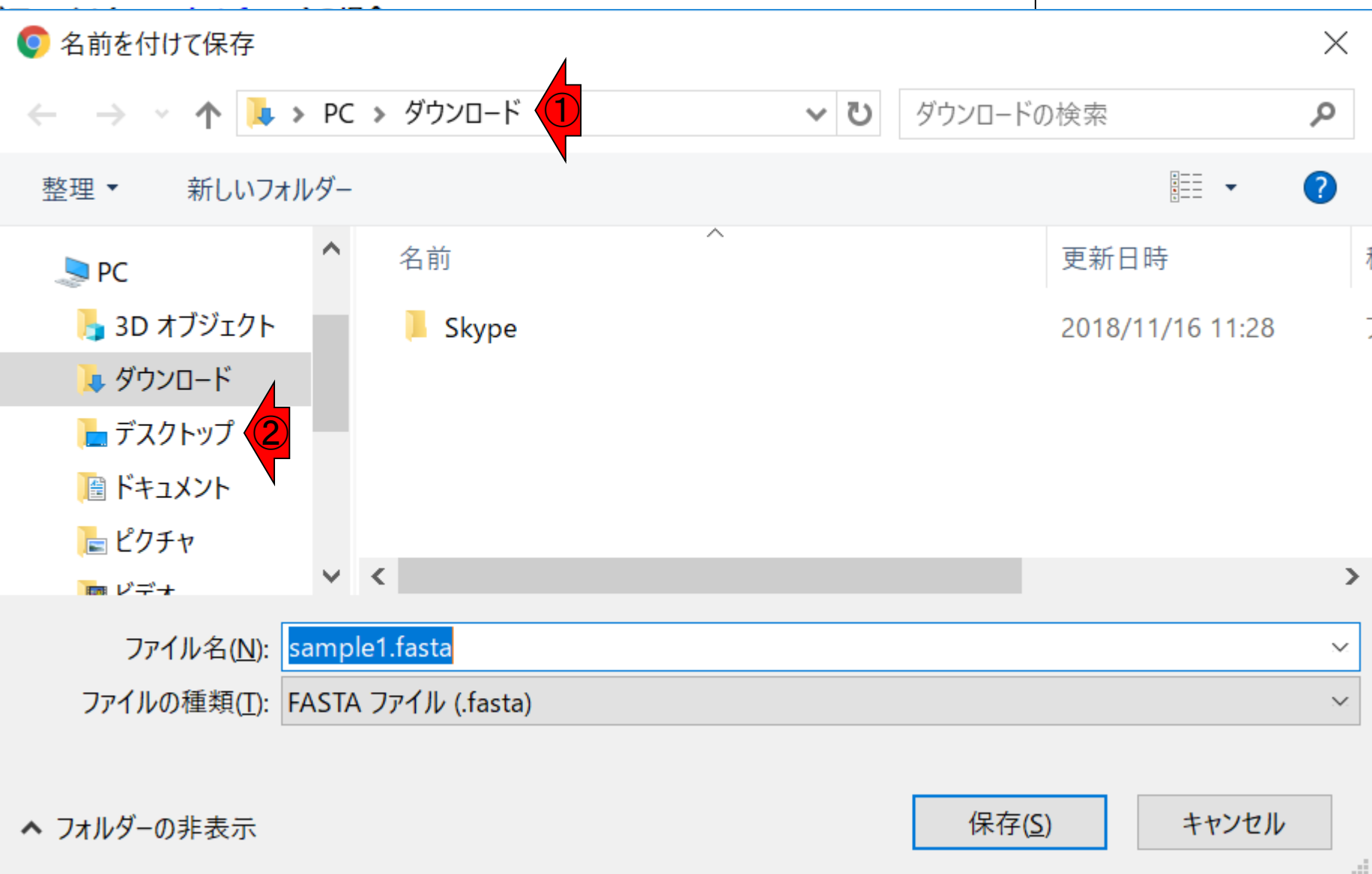
0 個の項目

# 入力ファイルの保存

大抵の場合、デフォルトの保存先は①ダウンロードになっていますが、②デスクトップ上にある...

## 1. FASTA形式

```
multi-FASTA
in_f <- "
out_f <- "
#必要なパッケージ
library(Biostrings)
#入力ファイルを読み込む
fasta <- readLines("input.fasta")
#本番処理
fasta <- writeXStringFromFASTA(fasta, "output.fasta")
#ファイルの書き出し
writeXStringFromFASTA(fasta, "output.fasta")
```

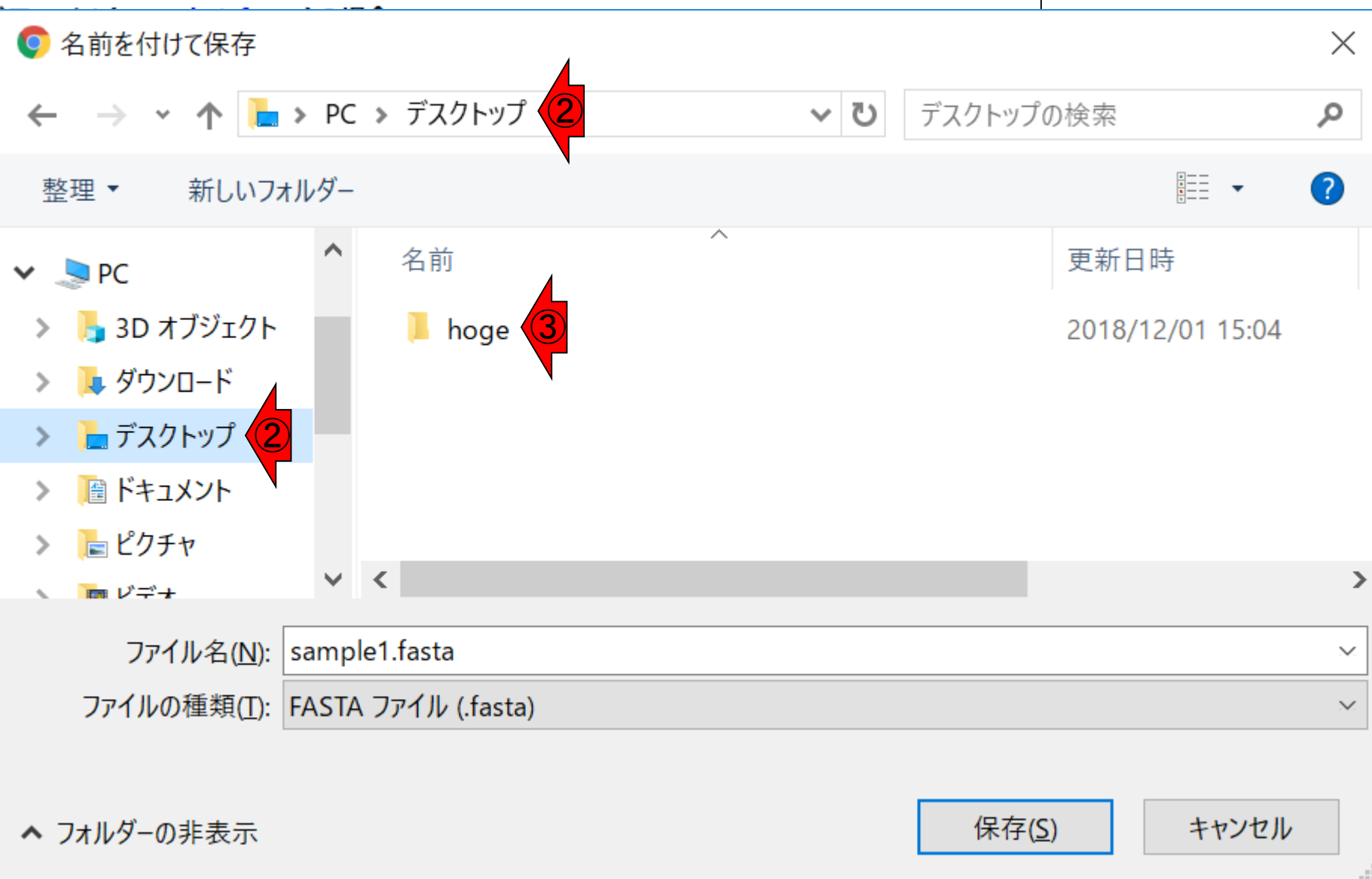


# 入力ファイルの保存

大抵の場合、デフォルトの保存先は①ダウンロードになっていますが、②デスクトップ上にある、③hogeフォルダです！

## 1. FASTA形式

```
multi-FASTA
in_f <- "
out_f <-
#必要なパッケージ
library(Biostrings)
#入力ファイルを読み込む
fasta <- readLines("sample1.fasta")
#本番処理
fasta <- writeXStringFromFASTA(fasta, "sample1.fasta")
#ファイルの書き出し
writeXStringFromFASTA(fasta, "sample1.fasta")
```





# 入力ファイルの保存

## 1. FASTA形式

```
multi-FASTA  
  
in_f <- "  
out_f <- "  
  
#必要なパ  
library(B  
  
#入力ファ  
fasta <- r  
fasta  
  
#本番  
fasta <- t  
fasta  
  
#ファイルは  
writeXStr
```

名前を付けて保存

PC > デスクトップ > hoge

hogeの検索

整理 ▾ 新しいフォルダー

名前 更新日時

検索条件に一致する項目はありません。

ファイル名(N): sample1.fasta

ファイルの種類(T): FASTA ファイル (.fasta)

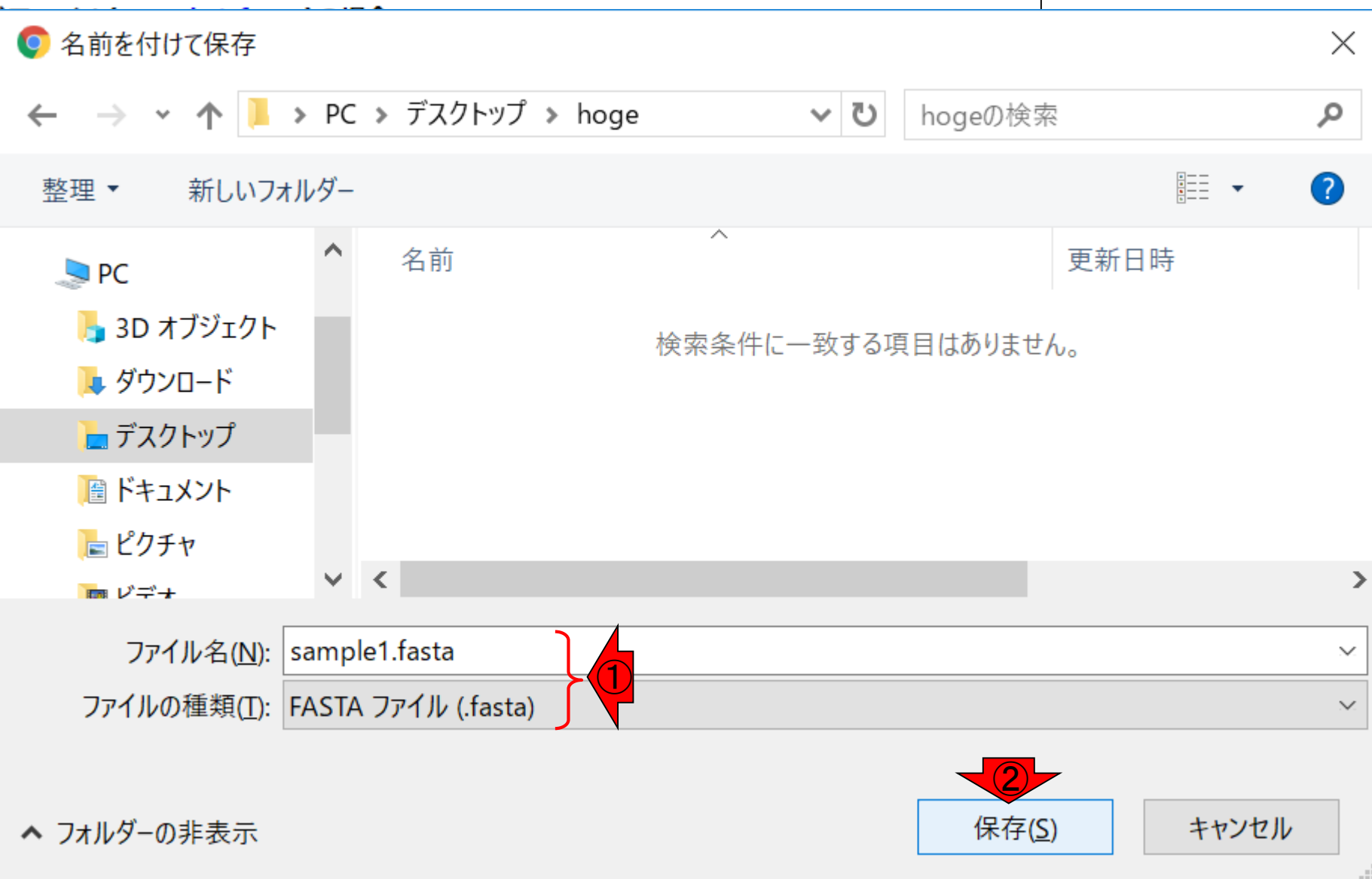
保存(S) キャンセル

# 拡張子に注意

①ときどきファイルの種類欄がテキストファイルと自動判定され(つまり.txtが付加されて)しまうことがあります。sample1.fasta.txtになるなどしたら、sample1.fastaに戻してから②保存してください。

## 1. FASTA形式

```
multi-FASTA
in_f <- "
out_f <- "
#必要なパッケージ
library(Biostrings)
#入力ファイル
fasta <- readDNAStringSet("sample1.fasta")
#本番
fasta <- writeXStringSet(fasta, "sample1.fasta")
#ファイルの書き出し
writeXStringSet(fasta, "sample1.fasta")
```



# 入力ファイルの保存

①こんな感じに見えていれば無事ダウンロードができています。②×。

(Rで)塩基配列解析

保護されていない | www.iu.a.u-tokyo.ac.j...

「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてあるディレクトリに移動し以下をコピー。

## 1. FASTA形式ファイル(sample1.fasta)の場合：

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta"      #入力ファイル名を指定してin_fに格納
out_f <- "hoge1.fasta"       #出力ファイル名を指定してout_fに格納

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

#入力ファイルの読み込み
fasta <- readDNASTringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta                                     #確認してるだけです

#本番
fasta <- translate(fasta)    #アミノ酸配列に翻訳した結果をfastaに格納
fasta                       #確認してるだけです

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50) #fastaの中身を指定したファイルに保存
```

sample1.fasta

①

すべて表示

②

# 入力ファイルの保存

①hogeフォルダ内に、②sample1.fastaが見えていればOK。②の中身は③のような感じです。ダブルクリックで見ると、別のソフト(MEGA)が起動してしまいました。テキストエディタで見てください。

```
>kadota↓  
AGTGACGGTCTT↓
```

## 1. FASTA形式ファイル(sample1.fasta)の場合:

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta"  
out_f <- "hoge1.fasta"
```

#入力ファイル名を指定してin\_fに格納  
#出力ファイル名を指定してout\_fに格納

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

#パッケージの読み込み

```
#入力ファイルの読み込み
```

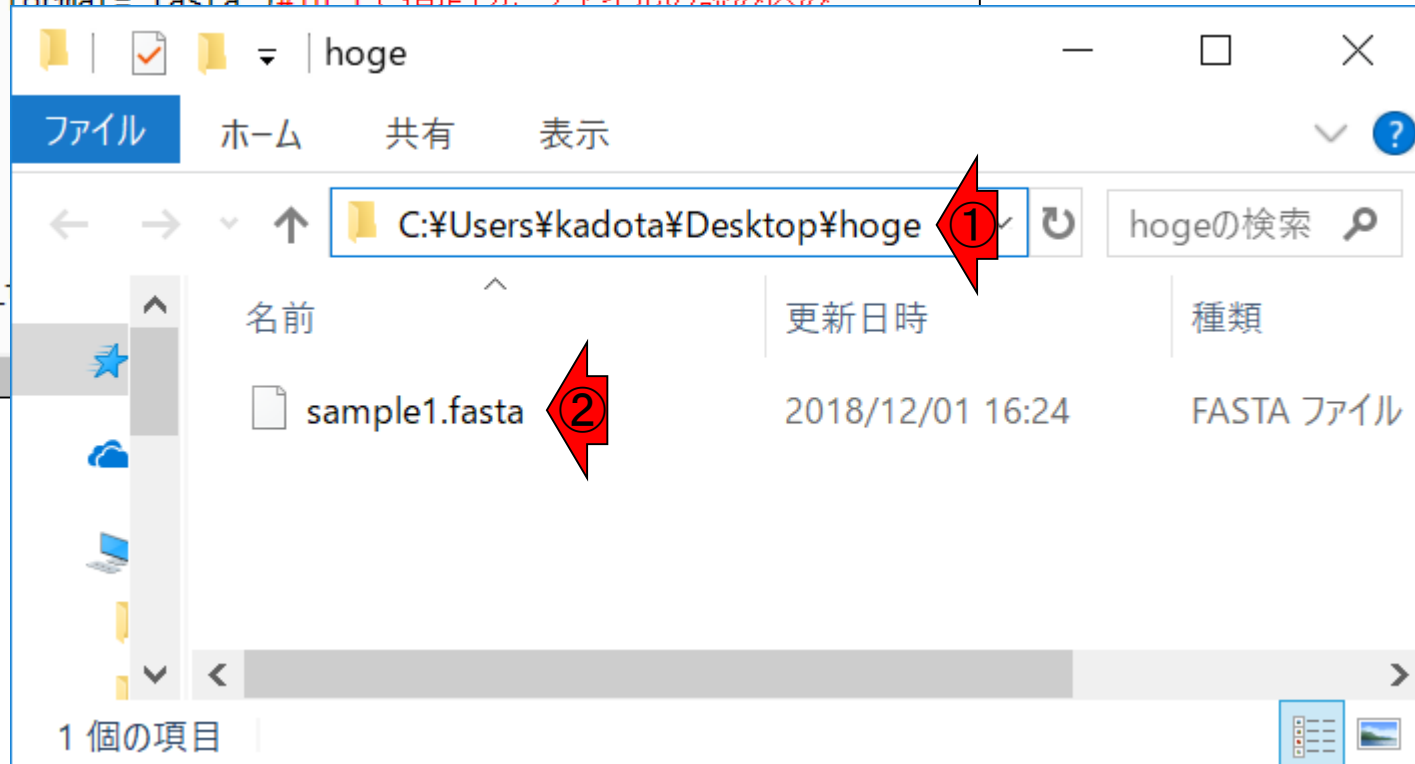
```
fasta <- readDNASTringSet(in_f, format="fasta")#in_fで指定したファイルの読み込み  
fasta
```

```
#本番
```

```
fasta <- translate(fasta)  
fasta
```

```
#ファイルに保存
```

```
writeXStringSet(fasta, file=out_f)
```



# 目的をおさらい

①hogeフォルダ内にある、②sample1.fasta中の③塩基配列に対応する翻訳配列(アミノ酸配列)を得るのが目的です。

## 1. FASTA形式ファイル(sample1.fasta)の場合:

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta"  
out_f <- "hoge1.fasta"
```

②

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

```
#入力ファイルの読み込み
```

```
fasta <- readDNASTringSet(in_f, format="fasta")  
fasta
```

```
#本番
```

```
fasta <- translate(fasta)  
fasta
```

```
#ファイルに保存
```

```
writeXStringSet(fasta, file=out_f)
```

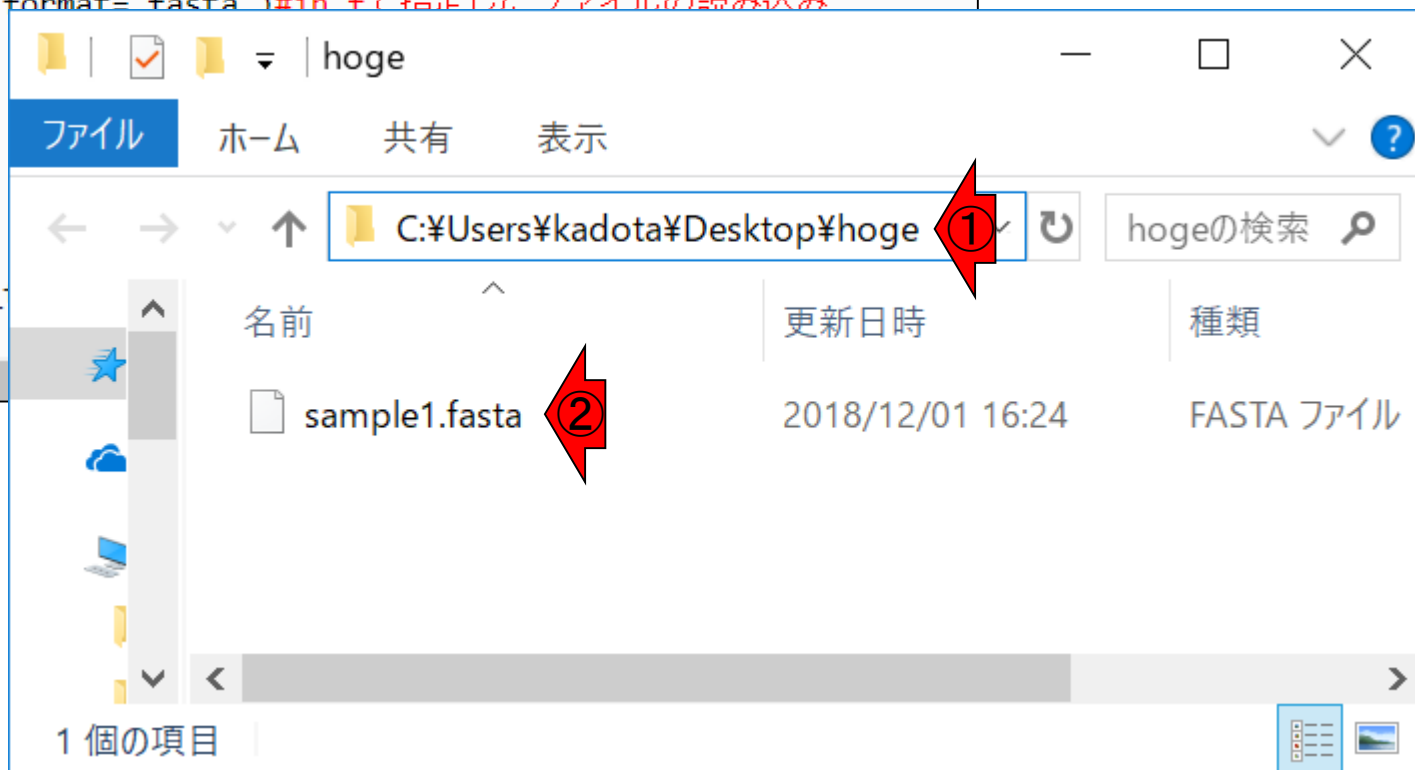
```
#入力ファイル名を指定してin_fに格納  
#出力ファイル名を指定してout_fに格納
```

```
#パッケージの読み込み
```

```
>kadota↓
```

```
AGTGACGGTCTT↓
```

③



# 目的をおさらい

実際には、プログラム実行結果として、①で指定した名前の、翻訳配列を含む出力ファイルが、②hogeフォルダ中に保存されます。

## 1. FASTA形式ファイル(sample1.fasta)の場合:

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta"  
out_f <- "hoge1.fasta"
```



#入力ファイル名を指定してin\_fに格納  
#出力ファイル名を指定してout\_fに格納

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

#パッケージの読み込み

```
#入力ファイルの読み込み
```

```
fasta <- readDNASTringSet(in_f, format="fasta")  
fasta
```

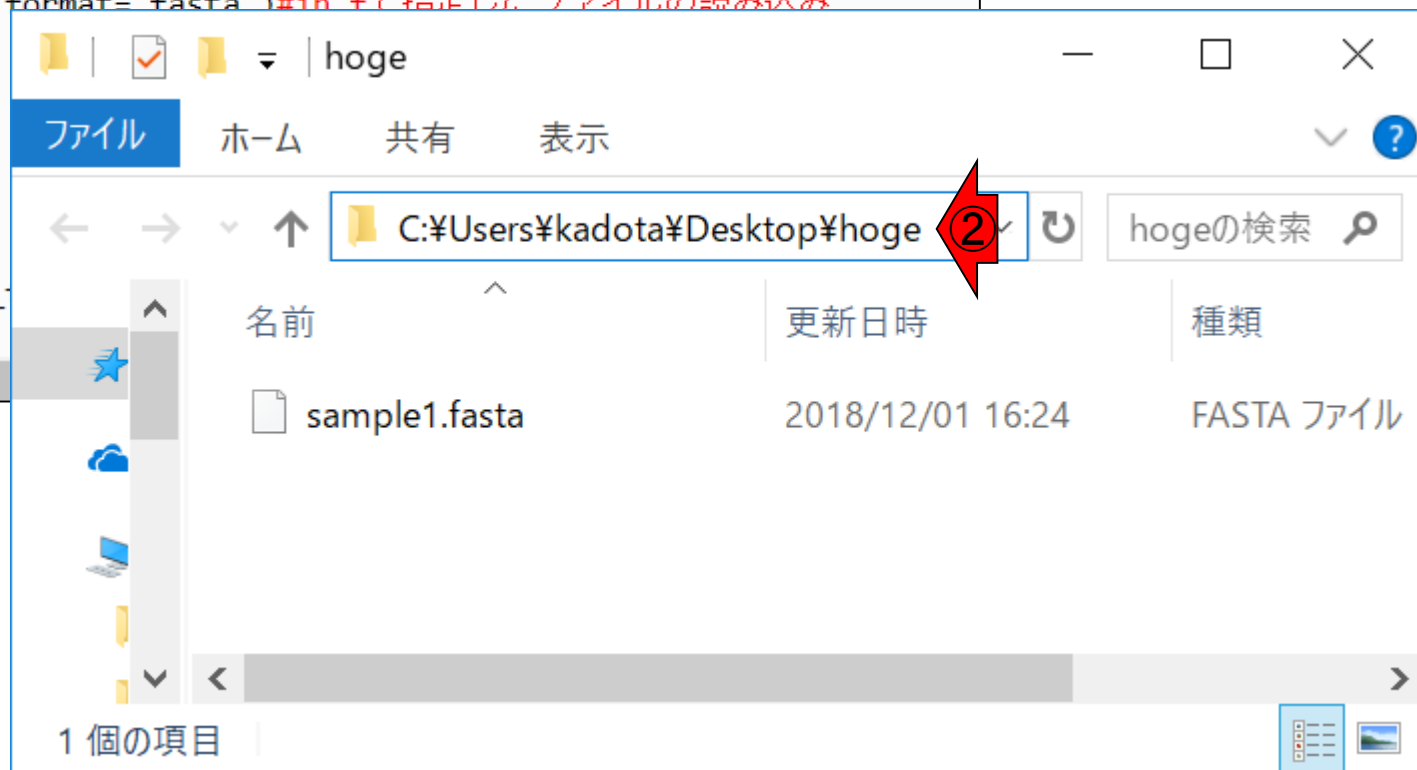
#in\_fで指定したファイルの読み込み

```
#本番  
fasta <- translate(fasta)  
fasta
```

```
#ファイルに保存
```

```
writeXStringSet(fasta, file=out_f)
```

>kadota↓  
AGTGACGGTCTT↓





# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

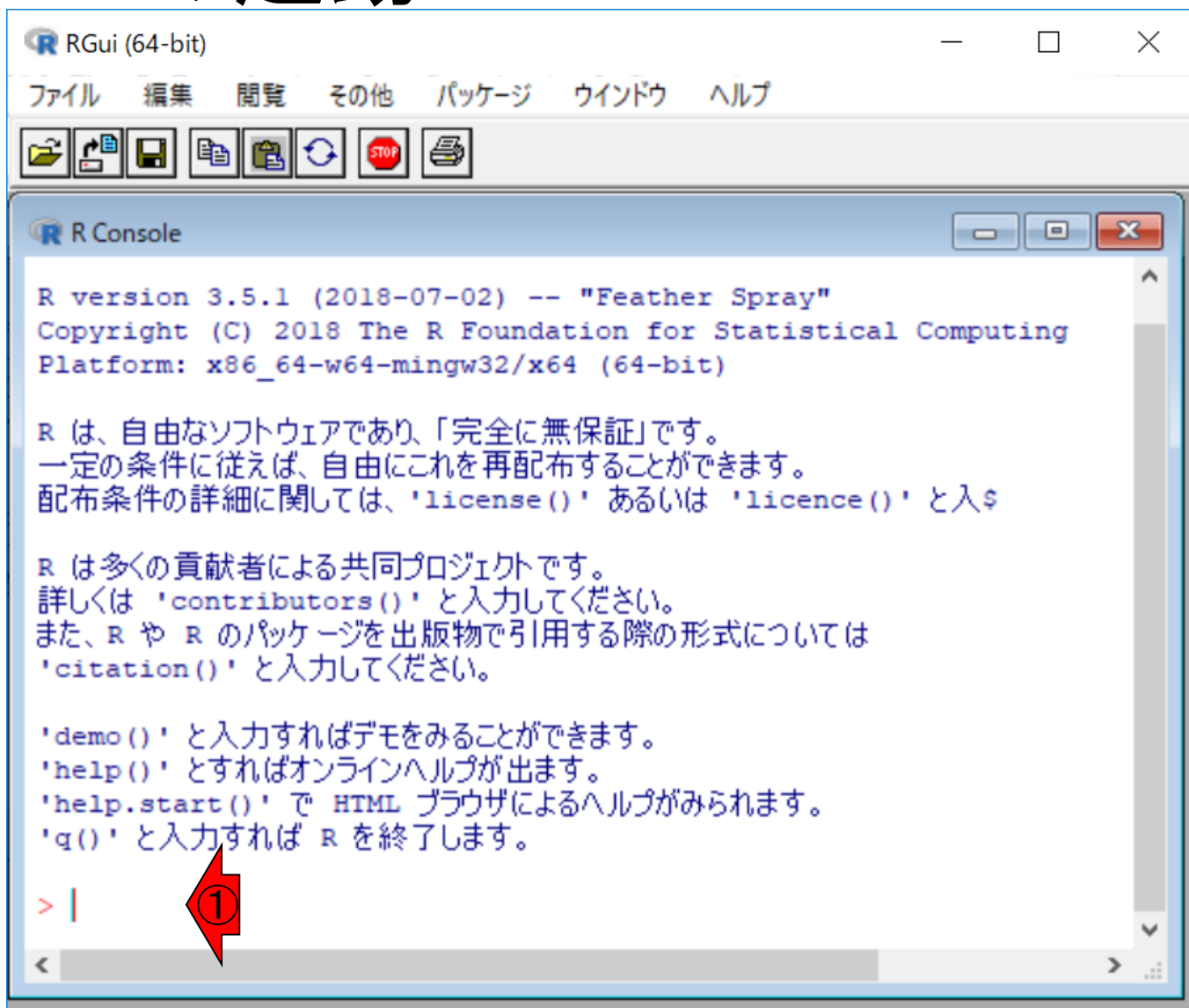
- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

①getwd()と打ち込んで、リターンキーを押す。

# Rの起動



# Rの起動

①このことです。R起動直後のデフォルトの作業ディレクトリは、②ユーザ名kadotaの環境では、「C:/Users/kadota/Documents」です。

```
Platform: x86_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入$

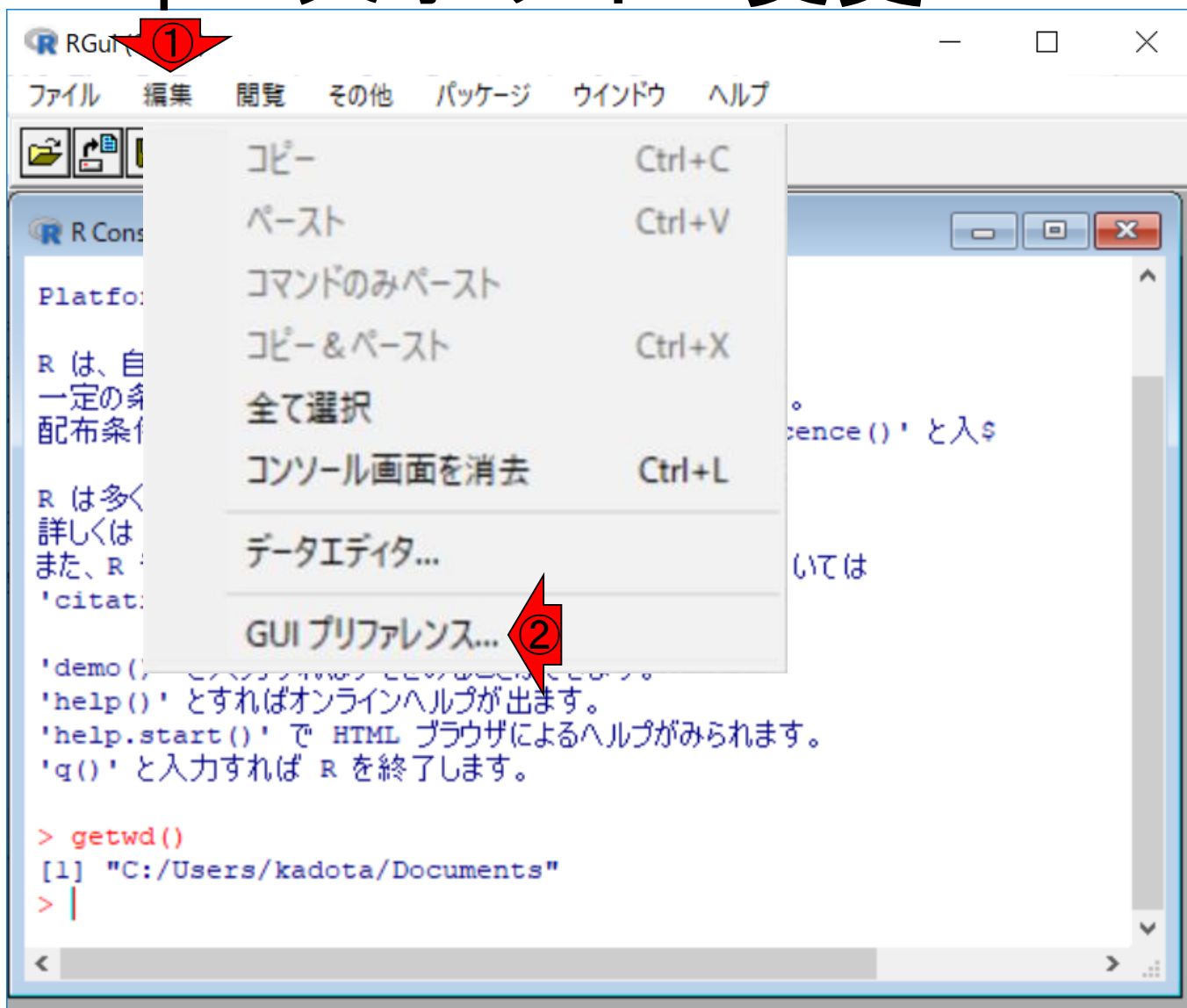
R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

> getwd()
[1] "C:/Users/kadota/Documents"
> |
```

文字サイズを変更したい場合は、①編集、②GUIプリファレンス

# Tips: 文字サイズ変更



# Tips: 文字サイズ変更

文字サイズを変更したい場合は、①編集、②GUIプリファレンス、③sizeのところを14とかにしてお使いください。

The image shows the RGui (64-bit) interface. The 'Edit' menu is open, showing options like 'コピー' (Copy), 'ペースト' (Paste), and 'GUIプリファレンス...' (GUI Preferences...). The 'Rgui 設定エディター' (Rgui Settings Editor) dialog box is open, showing various settings. The 'Font' section is highlighted, with the 'size' dropdown set to 10. The 'Console and Pager Colours' section is also visible, with 'background' set to 'normaltext' and 'white' selected in the color list.

① Edit menu

② GUIプリファレンス...

③ size dropdown

# getwd()

「getwd()」は、現在の作業ディレクトリを表示させるコマンドです。その一方で、②今解析したいファイルは、③デスクトップ上にあるhogeなので、作業ディレクトリをそこに変更する必要があります。

RGui (64-bit)

ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ



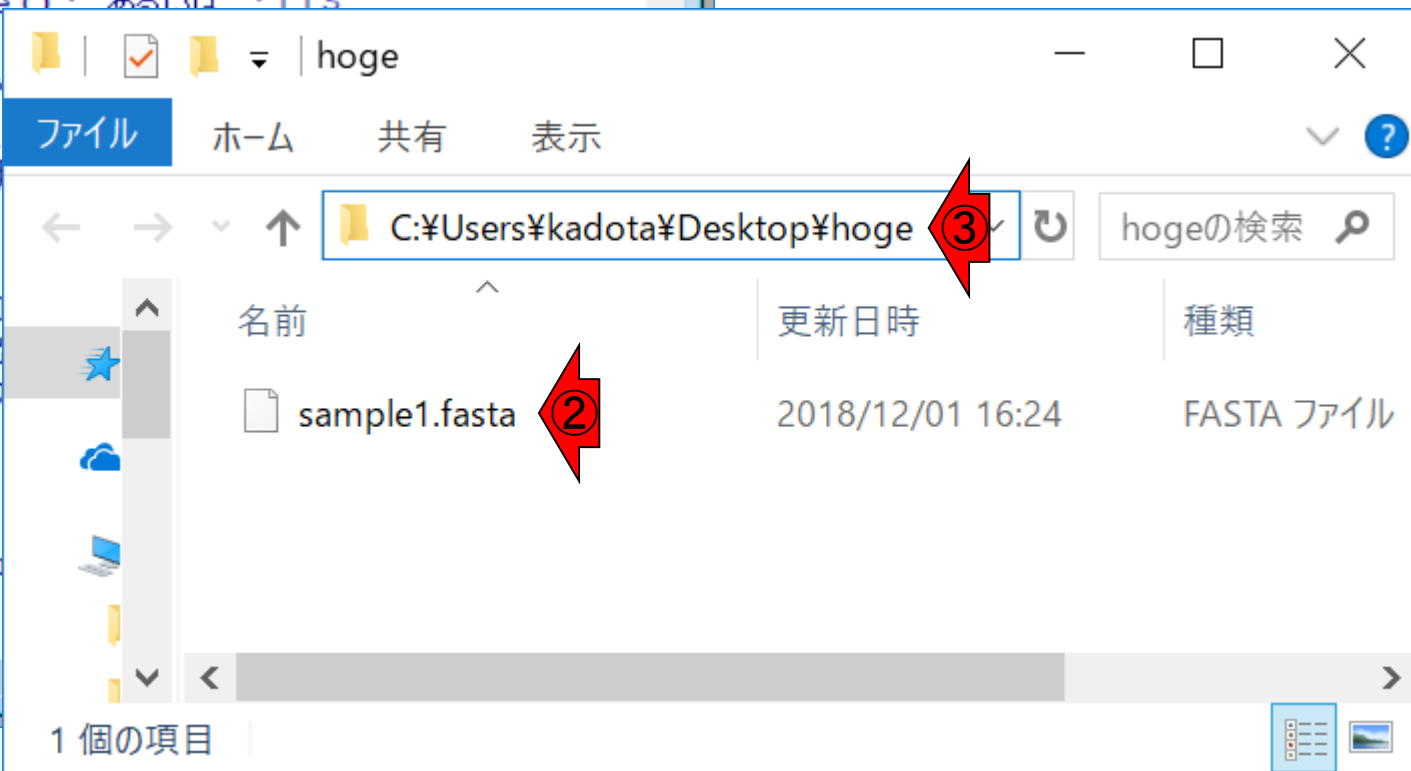
R Console

```
R は、自由なソフトウェアであり、「完全に無保証」で$  
一定の条件に従えば、自由にこれを再配布することがで$  
配布条件の詳細に関しては、'license()' または 'lis
```

```
R は多くの貢献者による共同プロジェクトで  
詳しくは 'contributors()' と入  
また、R や R のパッケージを出版物で引  
'citation()' と入力してください。
```

```
'demo()' と入力すればデモをみることか  
'help()' とすればオンラインヘルプが出  
'help.start()' で HTML ブラウ  
'q()' と入力すれば R を終了します。
```

```
> getwd()  
[1] "C:/Users/kadota/Docu  
> |
```





# 作業ディレクトリの変更

The screenshot shows an R console window titled "R (64-bit)". The menu bar includes "ファイル", "編集", "閲覧", "その他", "パッケージ", "ウインドウ", and "ヘルプ". The "ファイル" (File) menu is open, listing options such as "Rコードのソースを読み込み...", "新しいスクリプト", "スクリプトを開く...", "ファイルの表示...", "作業スペースの読み込み...", "作業スペースの保存..." (with Ctrl+S), "履歴の読み込み...", "履歴の保存...", "ディレクトリの変更..." (highlighted with a blue bar and a red arrow labeled '2'), "印刷..." (with Ctrl+P), "ファイルを保存...", and "終了". The console output shows the command `> getwd()` and the result `[1] "C:/Users/kadota/Documents"`. The console text is in Japanese, including "保証]で\$...", "することがで\$...", "()' あるいは 'li\$...", "。", "してください。", "する際の形\$...", "できます。", "ます。", and "げによるヘルプがみら\$".

# 作業ディレクトリの変更

ユーザkadotaの環境ではこのように見えている。デフォルトは①ドキュメントなので、②の位置がハイライトされている。すぐ上の③デスクトップを選択すると…

フォルダの参照

作業ディレクトリの変更  
C:\Users\kadota\Documents

- PC
- > 3D オブジェクト
- > ダウンロード
- > **デスクトップ**
- > **ドキュメント**
- > ピクチャ
- > ビデオ
- > ミュージック
- > Windows (C:)
- > DVD RW ドライブ (D:)
- > SDXC Card (E:)

フォルダー(E): ドキュメント

新しいフォルダーの作成(N) OK キャンセル

R Console

```
R は、自由なソフトウェアであ  
一定の条件に従えば、自由  
配布条件の詳細に関しては、  
  
R は多くの貢献者による共同  
詳しくは 'contributo  
また、R や R のパッケージ  
'citation()' と入力  
  
'demo()' と入力すれば  
'help()' とすればオンラ  
'help.start()' で  
'q()' と入力すれば R  
  
> getwd()  
[1] "C:/Users/kad  
> |
```

# 作業ディレクトリの変更

- ①の部分がDesktopに切り替わる。
- ②目的のhogeフォルダを選択

RGui (64-bit)

ファイル 編集 閲覧 その他

R Console

```
R は、自由なソフトウェアであり、  
一定の条件に従えば、自由に  
配布条件の詳細に関しては、  
  
R は多くの貢献者による共同  
詳しくは 'contributor' と入力  
また、R や R のパッケージを  
'citation()' と入力すれば  
'demo()' と入力すれば  
'help()' とすればオンラ  
'help.start()' で  
'q()' と入力すれば R  
  
> getwd()  
[1] "C:/Users/kadota/Desktop"  
> |
```

フォルダーの参照

作業ディレクトリの変更  
C:\Users\kadota\Desktop

- PC
  - > 3D オブジェクト
  - > ダウンロード
  - ▼ デスクトップ
    - hoge
  - > ドキュメント
  - > ピクチャ
  - > ビデオ
  - > ミュージック
  - > Windows (C:)
  - > DVD RW ドライブ (D:)

フォルダー(E): デスクトップ

新しいフォルダーの作成(N) OK キャンセル

# 作業ディレクトリの変更

RGui (64-bit)

ファイル 編集 閲覧 その他

R Console

R は、自由なソフトウェアであり、  
一定の条件に従えば、自由に  
配布条件の詳細に関しては、  
R は多くの貢献者による共同  
詳しくは 'contributor'。  
また、R や R のパッケージを  
'citation()' と入力  
'demo()' と入力すれば  
'help()' とすればオンラ  
'help.start()' で  
'q()' と入力すれば R

```
> getwd()  
[1] "C:/Users/kadota/Desktop/hoge"  
> |
```

フォルダーの参照

作業ディレクトリの変更  
C:\Users\kadota\Desktop\hoge

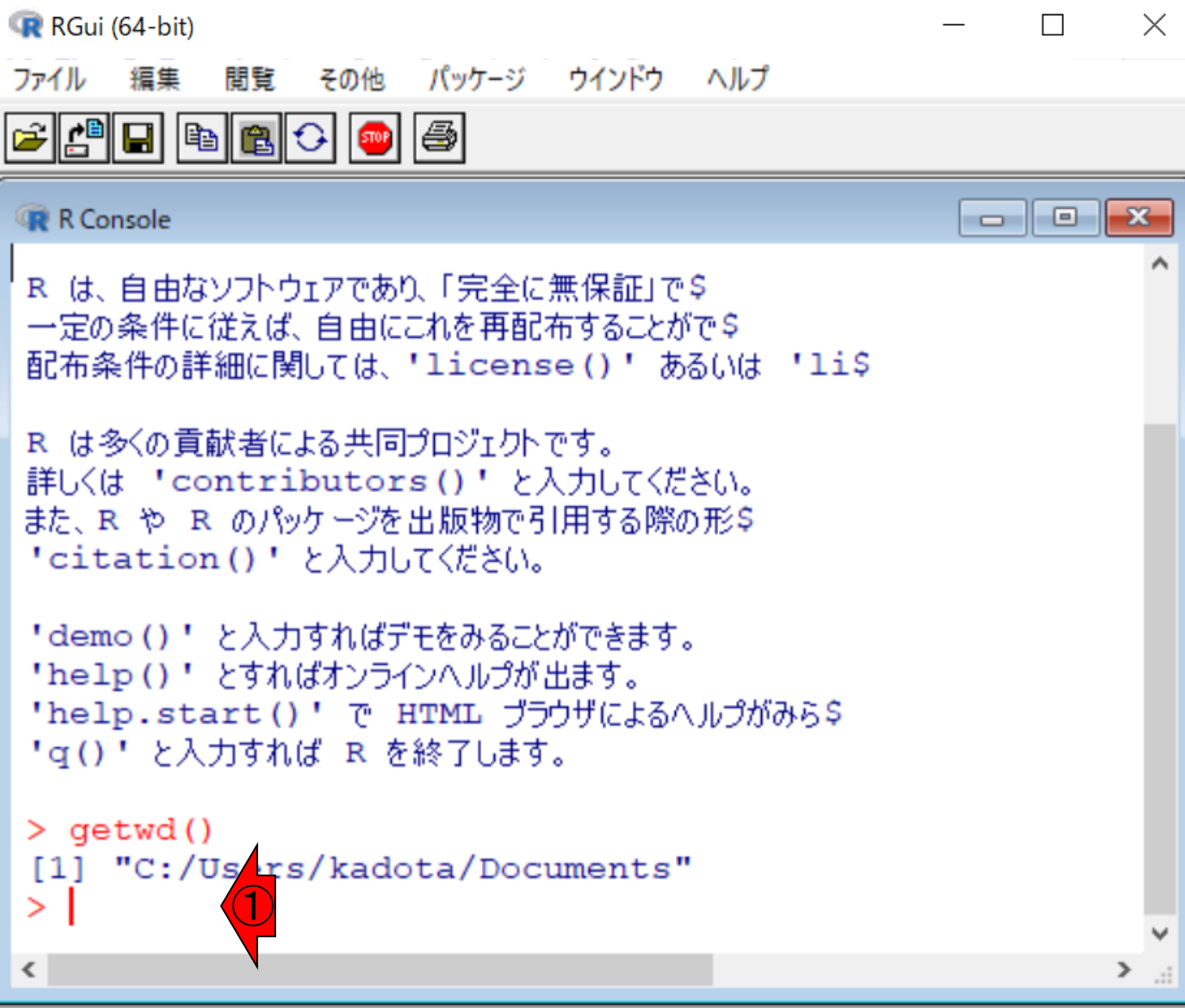
- PC
  - > 3D オブジェクト
  - > ダウンロード
  - ▼ デスクトップ
    - hoge
  - > ドキュメント
  - > ピクチャ
  - > ビデオ
  - > ミュージック
  - > Windows (C:)
  - > DVD RW ドライブ (D:)

フォルダー(E): hoge

新しいフォルダーの作成(N) OK キャンセル

# もう一度getwd()

一見すると、何も変わってなさそうですが、①  
もう一度`getwd()`を実行すれば、作業ディレク  
トリが変更されていることが確認できます。



```
RGui (64-bit)
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ
[Icons: Home, Recent, Save, Print, Refresh, Stop, Print]

R Console
R は、自由なソフトウェアであり、「完全に無保証」で$
一定の条件に従えば、自由にこれを再配布することがで$
配布条件の詳細に関しては、'license()' あるいは 'li$

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形$
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみら$
'q()' と入力すれば R を終了します。

> getwd()
[1] "C:/Users/kadota/Documents"
> |
```

# もう一度getwd()

さきほどと同様に`getwd()`とベタ打ちしてもよいが、キーボードの②上矢印キーを一回押すと、直前に打ち込んだコマンド(この場合は`getwd()`)が表示される。これは打ち込んだのと同じ意味なので、そのままリターンキーを押せばよい。いくつか入力したコマンドがあれば、上矢印キーを押していけば見られます。行き過ぎたら下矢印キーを押していけば戻れます

RGui (64-bit)

ファイル 編集 閲覧 その他 パッケージ ウィンドウ



R Console

```
R は、自由なソフトウェアであり、「完全に無保証」で$  
一定の条件に従えば、自由にこれを再配布することがで$  
配布条件の詳細に関しては、'license()' あるいは 'li$
```

```
R は多くの貢献者による共同プロジェクトです。  
詳しくは 'contributors()' と入力してください。  
また、R や R のパッケージを出版物で引用する際の形$  
'citation()' と入力してください。
```

```
'demo()' と入力すればデモをみることができます。  
'help()' とすればオンラインヘルプが出ます。  
'help.start()' で HTML ブラウザによるヘルプがみら$  
'q()' と入力すれば R を終了します。
```

```
> getwd()  
[1] "C:/Users/kadota/Documents"  
> |
```





# 確認

①こんな感じで、作業ディレクトリが「.../Desktop/hoge」であればOK。当たり前ですが、解析したいディレクトリ(またはフォルダ)を正しく指定できていなければエラーに遭遇します。また、解析したいファイルが存在しない状態でもエラーが出ます。

RGui (64-bit)

ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ



R Console

配布条件の詳細に関しては、`'license()'` あるいは `'li$`

R は多くの貢献者による共同プロジェクトです。

詳しくは `'contributors()'` と入力してください。

また、R や R のパッケージを出版物で引用する際の形\$

`'citation()'` と入力してください。

`'demo()'` と入力すればデモをみることができます。

`'help()'` とすればオンラインヘルプが出ます。

`'help.start()'` で HTML ブラウザによるヘルプがみら\$

`'q()'` と入力すれば R を終了します。

```
> getwd()
```

```
[1] "C:/Users/kadota/Documents"
```

```
> getwd()
```

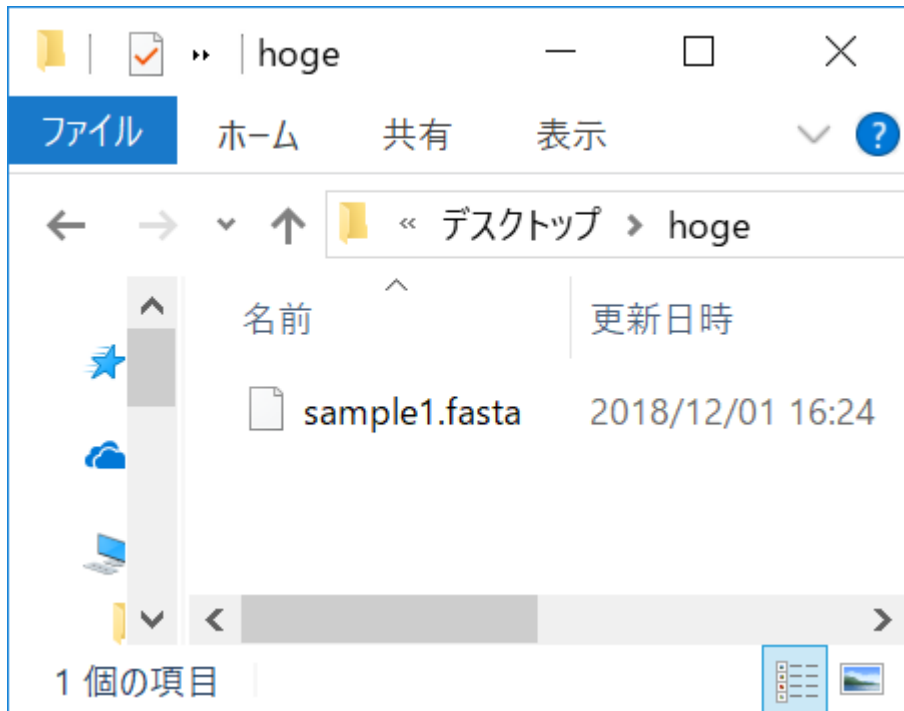
```
[1] "C:/Users/kadota/Desktop/hoge"
```

```
> |
```



①list.files()は、作業ディレクトリの中身を表示するコマンドです。

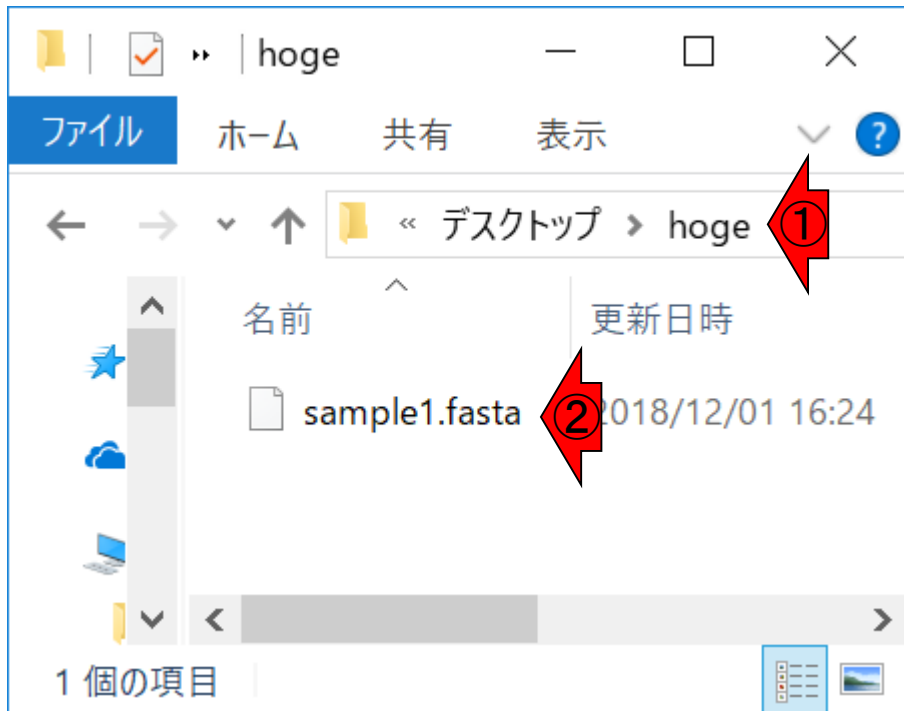
# list.files()でフォルダ内を見る



```
> getwd()
[1] "C:/Users/kadota/Desktop/hoge"
① list.files()
[1] "sample1.fasta"
> |
```

①hogeフォルダの②中身が対応しているのがわかりますね。

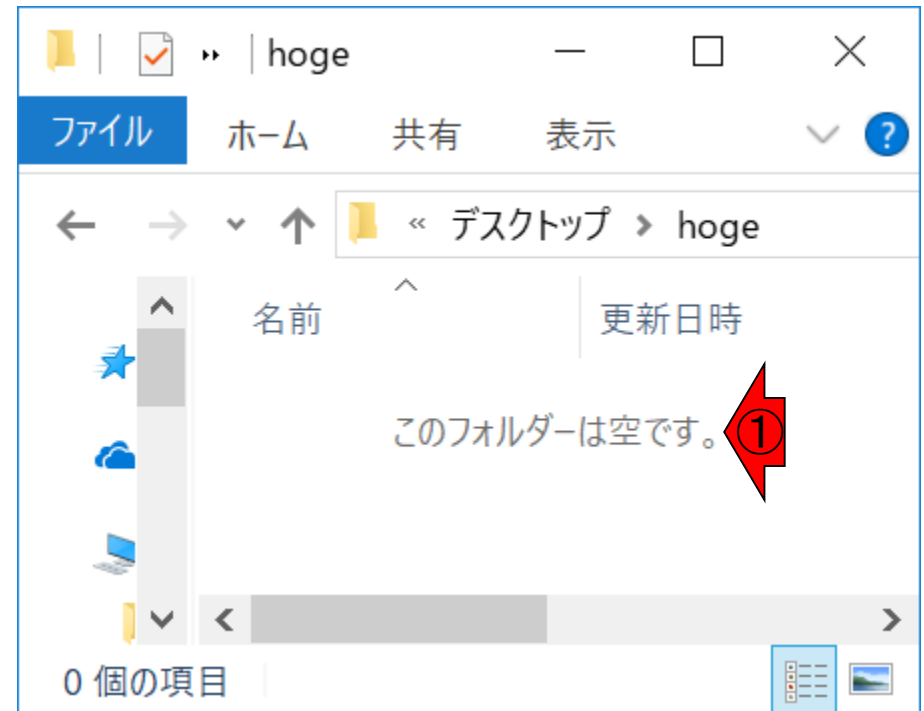
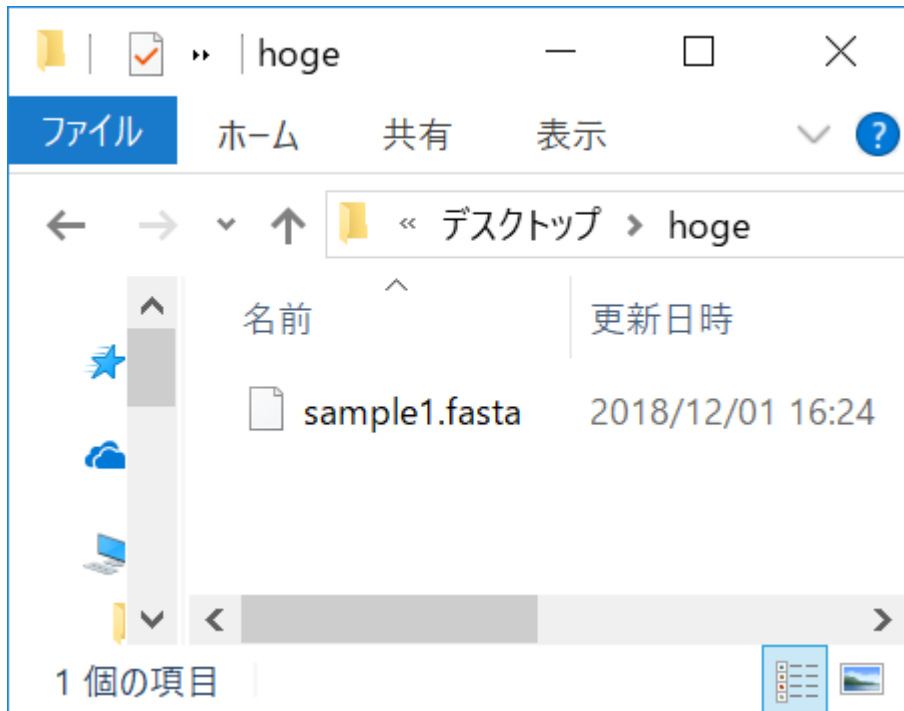
# list.files()でフォルダ内を見る



```
R Console
> getwd()
[1] "C:/Users/kadota/Desktop/hoge"
> list.files()
[1] "sample1.fasta"
> |
```

# list.files()でフォルダ内を見

もし①フォルダの中身が何もないときは、②のように見えるので覚えておきましょう。character(0)は何もないという意味



```
R Console
> getwd()
[1] "C:/Users/kadota/Desktop/hoge"
> list.files()
[1] "sample1.fasta"
> |
```

```
R Console
> getwd()
[1] "C:/Users/kadota/Desktop/hoge"
> list.files()
character(0)
> |
```

A red lightning bolt icon with the number '2' points to the output 'character(0)' in the R Console.

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# 基本はコピペ

- ①一連のコマンド群をコピーして
- ②R Console画面上でペースト。

## イントロ | 一般 | 翻訳配列(translate)を取得(基礎) | Biostrings

Biostringsパッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するやり方を示します。翻訳のための遺伝コード(genetic code)は、Standard Genetic Codeだそうです。もちろん生物種?!によって多少違い(variants)があるようで、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"などいろいろ選べるようです。

「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてあるディレクトリに移動し以下をコピペ。

### 1. FASTA形式ファイル(sample1.fasta)の場合：

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in f <- "sample1.fasta" #入力ファイル名を指定してin fに格納
out f <- "hoge1.fasta" #出力ファイル名を指定してout fに格納

#必要なパッケージをロード
library(Biostrings)

#入力ファイルの読み込み
fasta <- readDNASTringSet("sample1.fasta")
fasta

#本番
fasta <- translate(fasta) #アミノ酸配列を生成
fasta #確認して

#ファイルに保存
writeXStringSet(fasta, file=out f, format="fasta")
```

コピー(C)

①

Google で in\_f <- "sample1.fasta"

印刷(P)...

検証(I)

R Console

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全な条件に従えば、自由にこれを再配布条件の詳細に関しては、'license()' と入力してください。

R は多くの貢献者による共同プロジェクトです。詳しくは 'contributors()' と入力してください。また、R や R のパッケージを出版物に引用する場合は 'citation()' と入力してください。

'demo()' と入力すればデモをみる。'help()' とすればオンラインヘルプを開く。'help.start()' で HTML ブラウザによるヘルプを見る。'q()' と入力すれば R を終了します。

> |

コピー

Ctrl+C

ペースト

Ctrl+V

コマンドのみペースト

コピー & ペースト

Ctrl+X

ウィンドウの消去

Ctrl+L

全て選択

バッファに出力

Ctrl+W

ウィンドウを常にトップに置く



# 基本はコピペ

## イントロ | 一般 | 翻訳配列(translate)を取得

Biostringsパッケージを用いて塩基配列を読み込んでアミノ酸配列に翻訳するコード(genetic code)は、Standard Genetic Codeだそうです。もちろん、"Standard", "SGC0", "Vertebrate Mitochondrial", "SGC1"など、他の遺伝コードも利用可能です。「ファイル」 - 「ディレクトリの変更」で解析したいファイルを置いてください。

### 1. FASTA形式ファイル(sample1.fasta)の場合:

multi-FASTAではないsingle-FASTA形式ファイルです。

```
in_f <- "sample1.fasta" #入力ファイル名を
out_f <- "hoge1.fasta" #出力ファイル名を

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

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
fasta #確認してるだけで

#本番
fasta <- translate(fasta) #アミノ酸配列に翻訳
fasta #確認してるだけで

#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50)
```

R Console

```
> in_f <- "sample1.fasta"
> out_f <- "hoge1.fasta"
>
> #必要なパッケージをロード
> library(Biostrings)
要求されたパッケージ BiocGenerics をロード中です
要求されたパッケージ parallel をロード中です

次のパッケージを付け加えます: 'BiocGenerics'

The following objects are masked from 'package:parallel':

clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
clusterExport, clusterMap, parApply, parCapply, parLapply,
parLapplyLB, parRapply, parSapply, parSapplyLB

The following object is masked from 'package:stats':

xtabs

The following objects are masked from 'package:base':

anyDuplicated, append, as.data.frame, as.vector, cbind,
colnames, do.call, duplicated, eval, evalq, Filter, Find, get,
intersect, is.unsorted, lapply, Map, mapply, match, mget, order,
paste, pmax, pmax.int, pmin, pmin.int, Position, rank, rbind,
Reduce, rep.int, rownames, sapply, setdiff, sort, table, tapply,
union, unique, unlist

要求されたパッケージ IRanges をロード中です
要求されたパッケージ XVector をロード中です

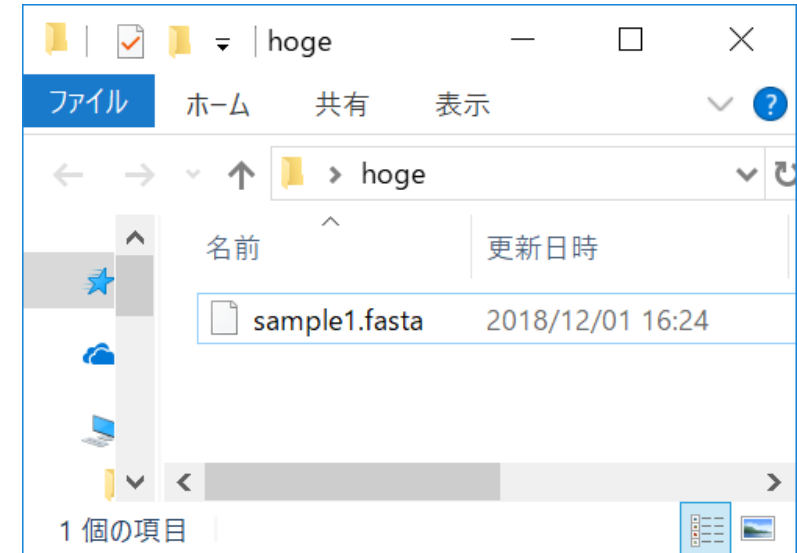
> #入力ファイルの読み込み
> fasta <- readDNAStringSet(in_f, format="fasta") #in_fで指定したファイルの読み込み
> fasta #確認してるだけで
  A DNAStringSet instance of length 1
    width seq          names
[1]    12 AGTGACGGTCTT kadota
>
> #本番
> fasta <- translate(fasta) #アミノ酸配列に翻訳した結果をfasta$
> fasta #確認してるだけで
  A AAStringSet instance of length 1
    width seq          names
[1]     4 SDGL          kadota
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", width=50) #fastaの中身を$
> |
```

エラーなく実行できた場合の全貌。多少見栄えが異なっても、エラーという文字が見えていなければOK。

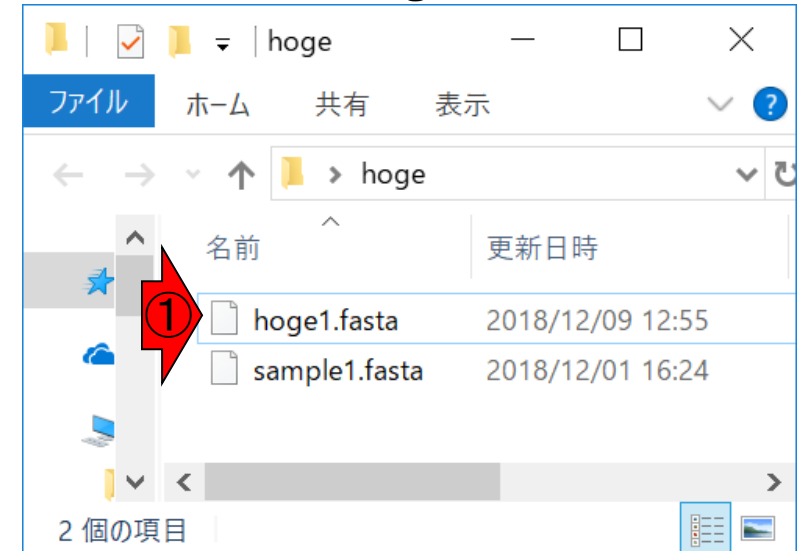
# 実行結果

①出力ファイル名として指定したhoge1.fastaが生成されていることが分かります

## 実行前のhogeフォルダ



## 実行後のhogeフォルダ



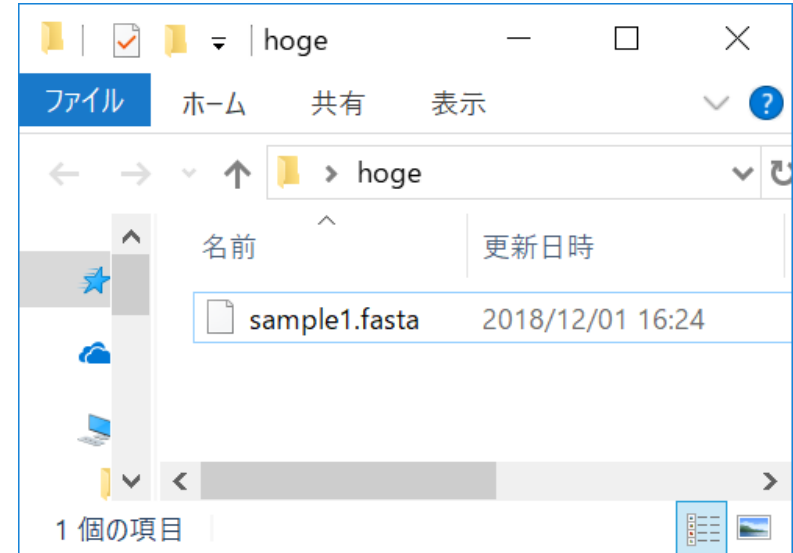
```
R Console
> fasta <- readDNAStringSet(in_f, format="fasta")$
> fasta                                     #確認し$
A DNAStringSet instance of length 1
  width seq          names $
[1]    12 AGTGACGGTCTT kadota
>
> #本番
> fasta <- translate(fasta)                #アミノ$
> fasta                                     #確認し$
A AAStringSet instance of length 1
  width seq          names $
[1]     4 SDGL        kadota
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fast$
> |
```

# 実行結果

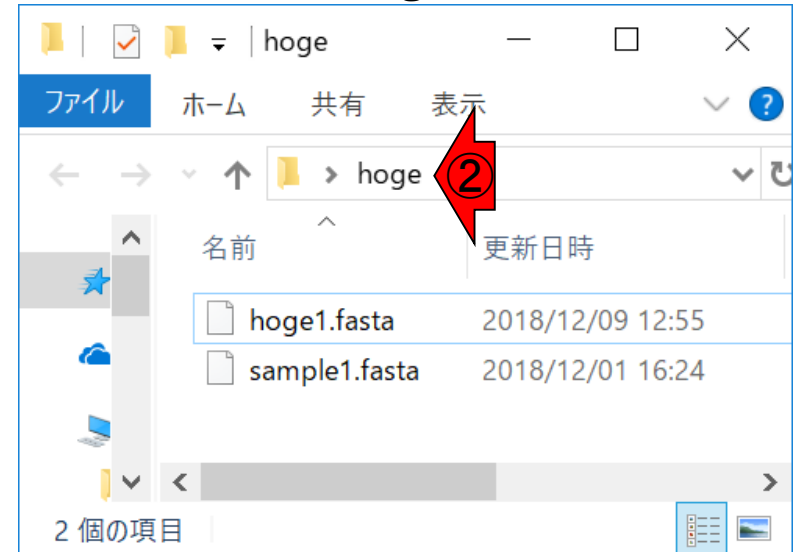
①list.files()で表示される結果と、②「実行後のhogeフォルダの中身は当然同じ

```
R Console
> fasta <- readDNASTringSet(in_f, format="fasta")$
> fasta #確認し$
A DNASTringSet instance of length 1
width seq names
[1] 12 AGTGACGGTCTT kadota
>
> #本番
> fasta <- translate(fasta) #アミノ$
> fasta #確認し$
A AAStringSet instance of length 1
width seq names
[1] 4 SDGL kadota
>
> #ファイルに保存
> writeXString(fasta, file=out_f, format="fast$
> list.files()
[1] "hoge1.fasta" "sample1.fasta"
> |
```

実行前のhogeフォルダ



実行後のhogeフォルダ



# 実行結果

入力ファイル中の塩基配列は、3の倍数の12塩基長、ACGTのみからなるので何のエラーもない。講義では入力ファイル中の塩基をいじって色々試してくださいと言いました。例えば最後の塩基Tを消すと、3の倍数でなくなるのでエラーが出るはずだと言いましたが、SDGになりエラーが出なかったそうです。Biostringsパッケージがバージョンアップしたのでしょうか。

入力: 塩基配列ファイル (sample1.fasta)

```
sample1.fasta - メモ帳
ファイル(F) 編集(E) 書式(O) 表示(V)
ヘルプ(H)
>kadota
AGTGACGGTCTT
```

出力: アミノ酸配列ファイル (hoge1.fasta)

```
hoge1.fasta - メモ帳
ファイル(F) 編集(E) 書式(O) 表示(V)
ヘルプ(H)
>kadota
SDGL
```

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

# インストール

①RStudioは、RのGUI強化版のような位置づけです。他の講義で利用することがあるかもしれませんが、「こういうものもあり、こんな感じで使う」程度は知っておいたほうがよいと思います。慣れていない先生のハンズオン講義は苦行ですが、受講生側がRStudioのGUIに多少でも慣れておけば苦痛は軽減できます。②Download RStudio。

1. 2019年04月08日 (PC使用)

講義資料PDF

学会(国外) : ISCB

学会(国内) : JSBi

QAサイト : Biostar (Parnell et al., PLoS Comput Biol., 2011)

QAサイト : SEQanswers (Li et al., 2011)

学習教材 : バイオインフォマティクス

学習教材 : (Rで)塩基配列解析

学習教材 : (Rで)塩基配列解析のサブ

RStudio

①

Home - RStudio

https://www.rstudio.com

RStudio

Products Resources Pricing About Us Blogs

RStudio

Open source and enterprise-ready professional software for R

Download RStudio ②

Discover Shiny

Discover RStudio Package Manager

Discover RStudio Connect



# インストール

- ①RStudio Desktop Open Source Licenseの、
- ②DOWNLOADをクリック。

Download RStudio - RStudio


https://www.rstudio.com/products/rstudio/download/

RStudio

Products Resources Pricing About Us Blogs

## Choose Your Version of RStudio

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace. [Learn More about RStudio features.](#)



RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License	RStudio Server Pro RStudio Connected Commercial License
FREE	\$995 per year	FREE	\$9,995 per year	\$29,995 per year
<b>DOWNLOAD</b> Learn More	<b>BUY</b> Learn More	<b>DOWNLOAD</b> Learn More	<b>DOWNLOAD</b> Learn More	<b>TALK</b> Learn More
●	●	●	●	●

# インストール

①の中の、Windows版(上)、Mac版(下)のいずれかのインストーラをクリック。

Download RStudio - RStudio

https://www.rstudio.com/products/rstudio/download/#download

RStudio

Products Resources Pricing About Us

**RStudio Desktop 1.1.463** — Release Notes

RStudio requires R 3.0.1+. If you don't already have R, download it here.

Linux users may need to import RStudio's public code-signing key prior to installation, depending on the system's security policy.

### Installers for Supported Platforms

Installers	Size	Date	MD5
<b>RStudio 1.1.463 - Windows Vista/7/8/10</b>	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab64d4
<b>RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)</b>	74.5 MB	2018-10-29	a79032ba4d7daaa86a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	8a6755fa9fae2bafce289df3358aaf63
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-10-29	bc50d6bd34926c1cc3ae4a209d67d649
RStudio 1.1.463 - Ubuntu 16.04+/Debian 9+ (64-bit)	65 MB	2018-10-29	cf d659db18619cc78d1592fefaa7c753
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-10-29	742f0bad60dfaea3281576e14ad6699e
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-10-29	c7303067a0ca99deea7e427b856952d1

### Zip/Tarballs

ユーザー アカウント制御

この不明な発行元からのアプリがデバイスに変更を加えることを許可しますか?

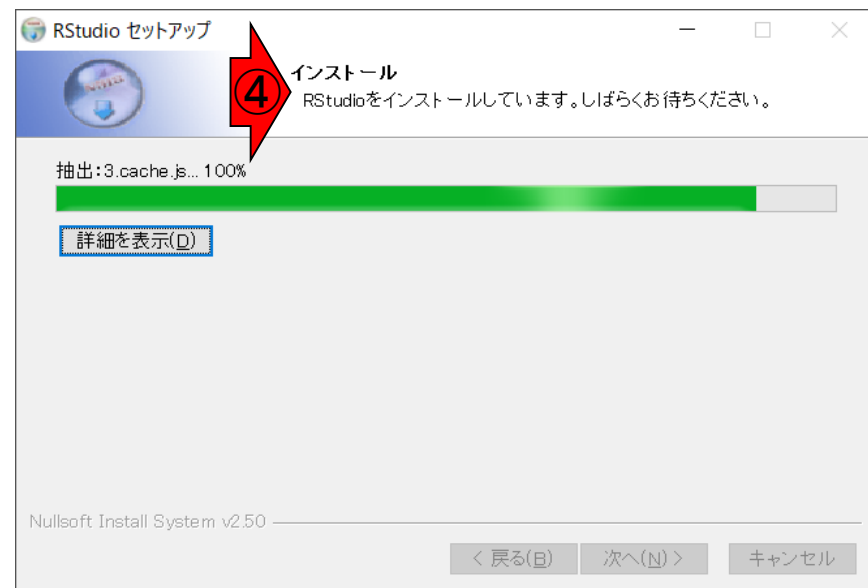
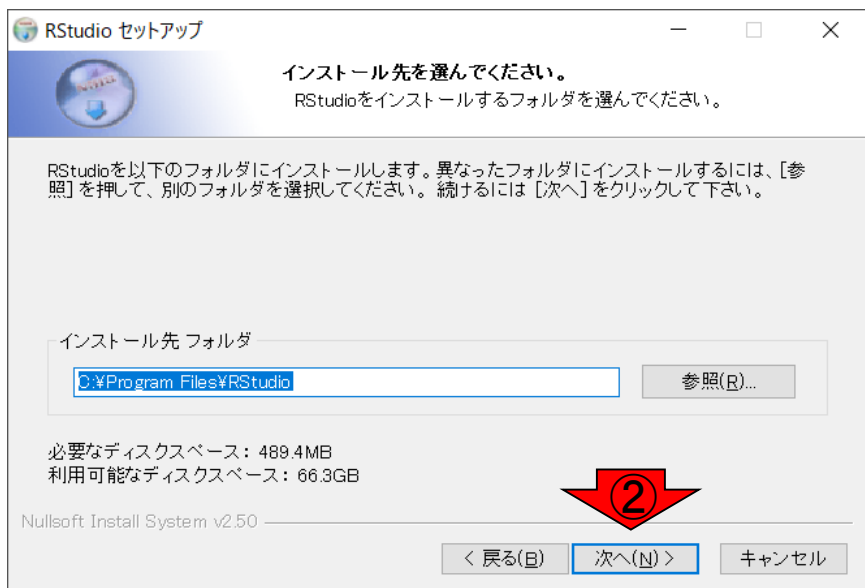
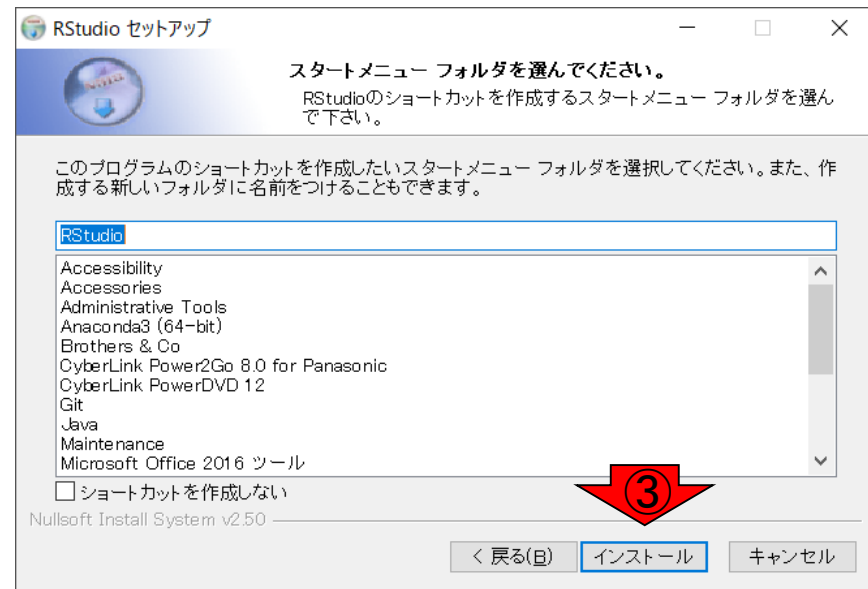
RStudio-1.1.463.exe

発行元: 不明  
ファイルの入手先: このコンピューター上のハードドライブ

詳細を表示

はい いいえ

# インストール



# インストール完了

インストール完了。①完了ボタンをクリックしますが、デスクトップのどこにもRStudioは見当たりません。そこで、②スタートボタンをクリックします。

The screenshot shows a Windows desktop with a taskbar at the bottom. The desktop background is a blue and orange gradient with a globe. On the left side, there are several icons: PC, ネットワーク, ごみ箱, iU, R x64 3.5.2, Oracle VM VirtualBox, Acrobat Reader DC, share, and hoge. In the center, a dialog box titled "RStudio セットアップ" is open. The dialog box contains the following text: "RStudio セットアップ ウィザードは完了しました。 RStudioは、このコンピュータにインストールされました。 ウィザードを閉じるには [完了] を押してください。" At the bottom of the dialog box, there are three buttons: "< 戻る(B)", "完了(F)", and "キャンセル". A red arrow with the number "1" points to the "完了(F)" button. To the right of the dialog box, there is a warning sign with a yellow background and black border. The warning sign contains the following text: "PCの貸出に関する注意 以下の注意事項を厳守してください • PCを汚損しないこと(PC付近では飲食禁止) • 教員が特に指示した場合以外は、ソフトウェア等をインストールしないこと" Below this, there is a "Warning" section with a yellow background and black border. The warning section contains the following text: "Warning ST observe the following rules: • Do not make the PC clean. No eating or drinking near PCs. • Do not install/uninstall software. • Do not change settings. • If you violate these rules, you will not be allowed to attend any lectures of Agricultural Bioinformatics Research Unit." At the bottom of the warning sign, there is a small text: "Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, The University of Tokyo". In the bottom left corner of the desktop, there is a search bar with the text "ここに入力して検索" and a red arrow with the number "2" pointing to it. The taskbar at the bottom shows the Start button, a search bar, and several application icons. The system tray in the bottom right corner shows the time "14:51" and the date "2019/02/07".

# Contents

## ■ Introduction

- 学会(国外と国内)、QAサイト
- 学習教材: バイオインフォマティクス人材育成のための講習会
- 学習教材: (Rで)塩基配列解析のメインページとサブページ

## ■ Rの基本的な利用法

- Rの起動と終了
- 数値計算、(Rで)塩基配列解析のノリに慣れる

## ■ 翻訳配列取得

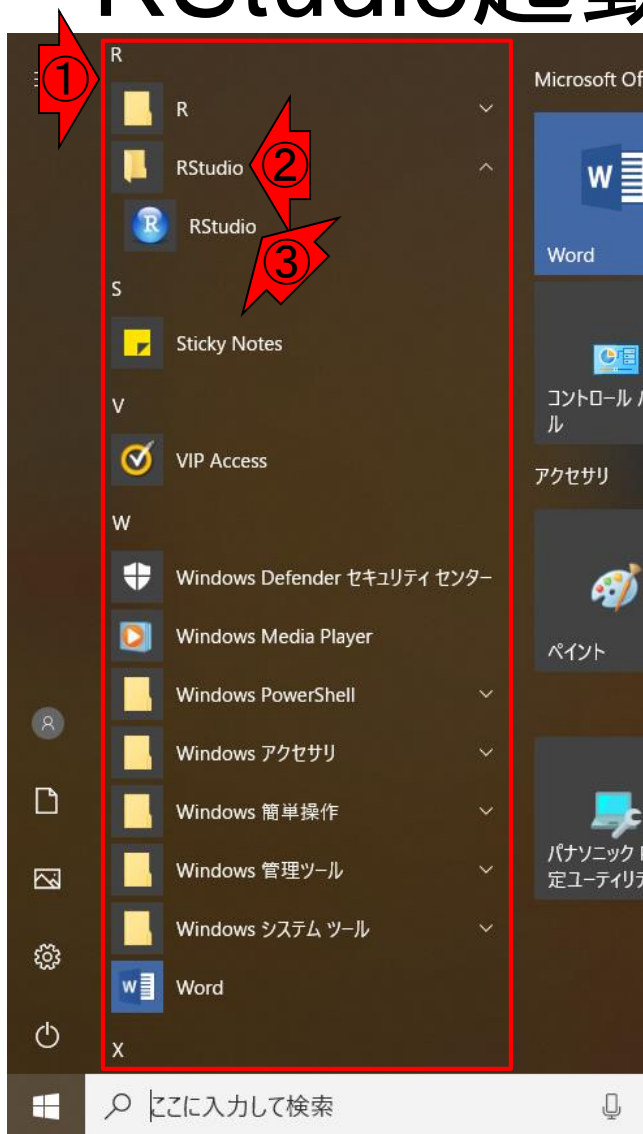
- 準備: フォルダの作成から入力ファイルの保存まで
- 作業ディレクトリの変更と確認: `getwd()`と`list.files()`
- コピペ実行と結果の確認

## ■ RStudio

- インストール、起動してディレクトリの変更、`getwd()`

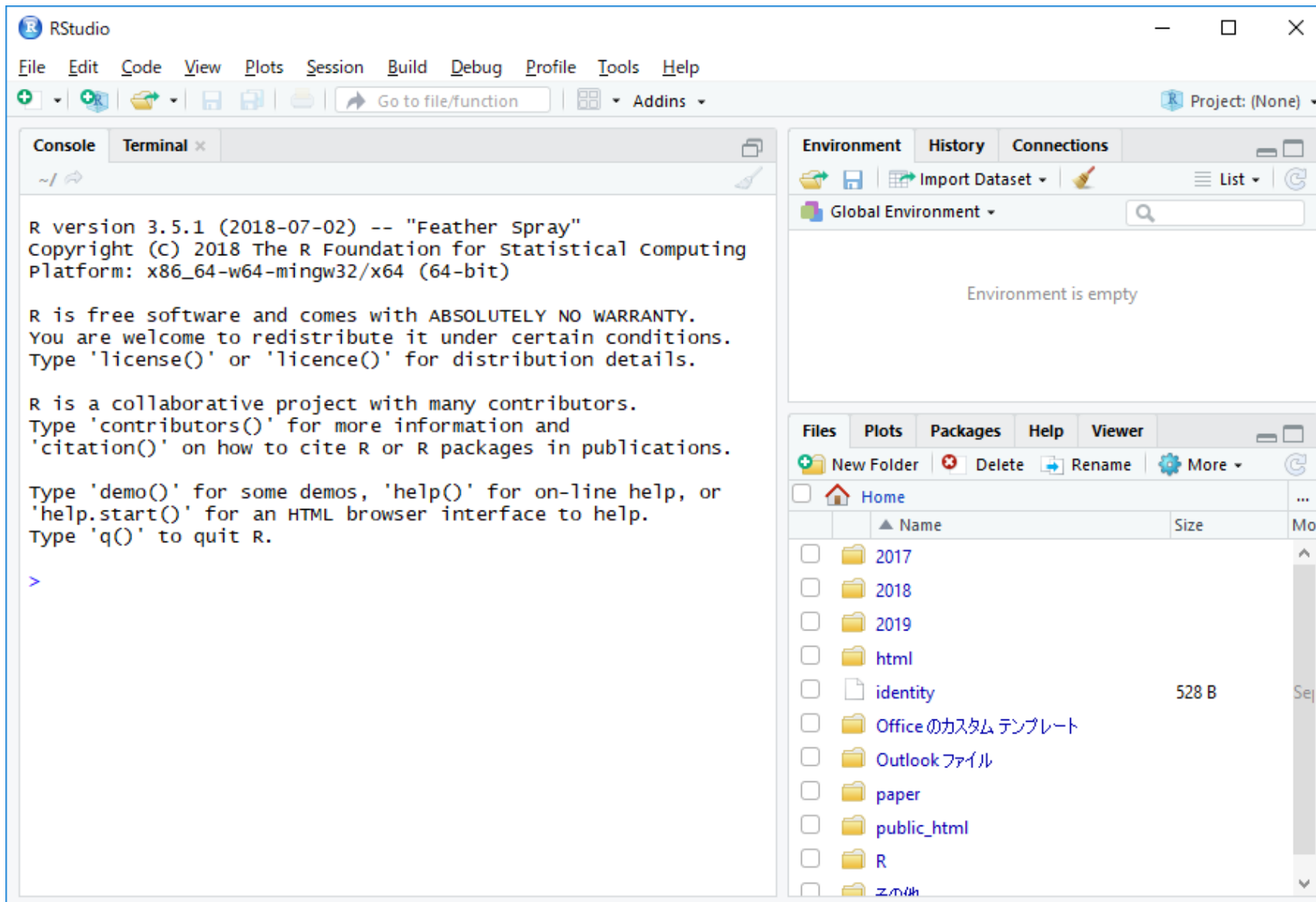
# RStudio起動

- ①スタートメニューが表示されるので、スクロールダウンし、
- ②RStudio、③RStudioをクリックして、RStudioを起動します。





# RStudio起動



The screenshot shows the RStudio application window. The title bar reads "RStudio". The menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The toolbar contains icons for file operations and a search bar labeled "Go to file/function". The "Project: (None)" dropdown is visible in the top right.

The left pane is split into "Console" and "Terminal" tabs. The "Terminal" tab is active, displaying the following text:

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

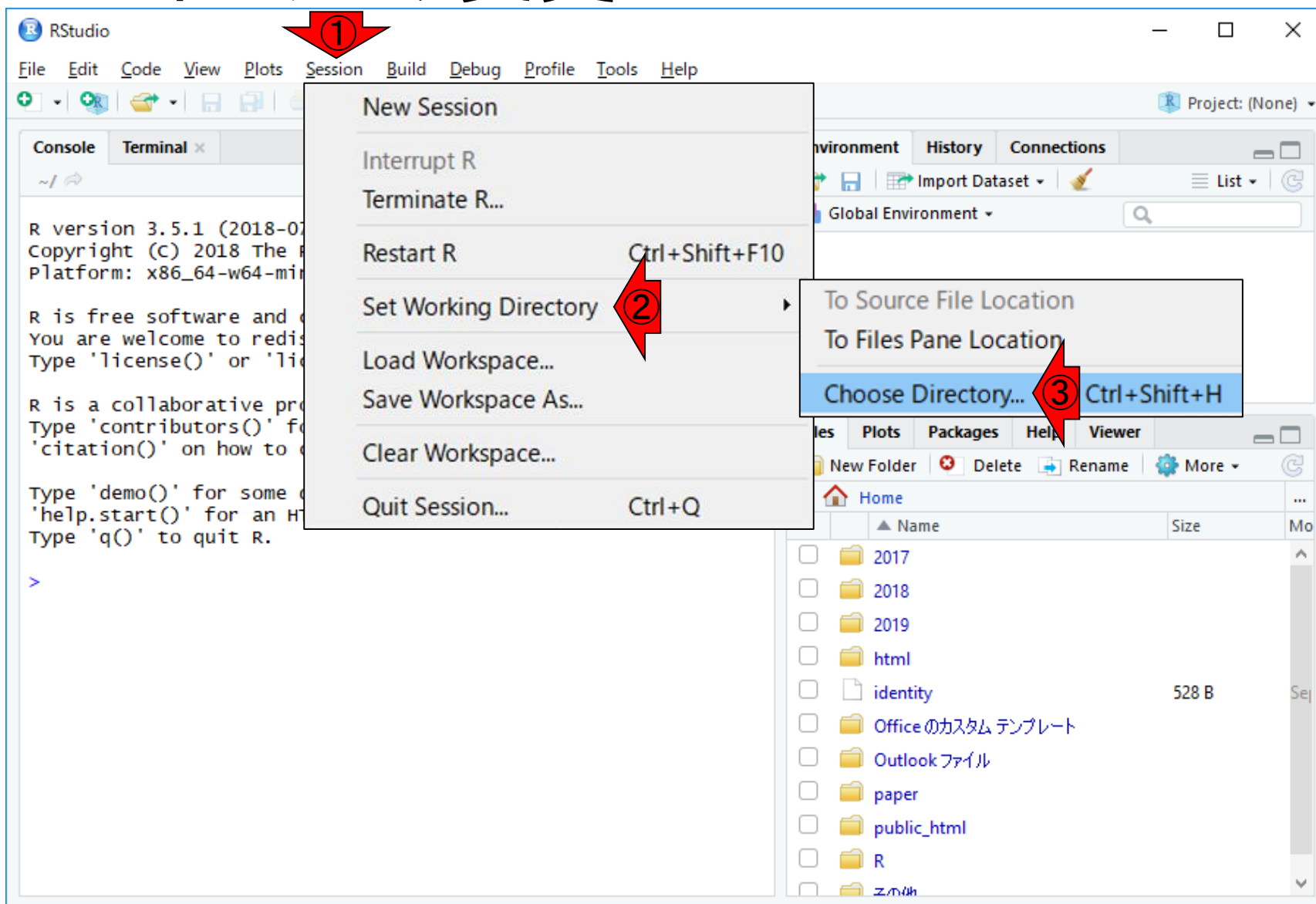
>
```

The right pane is split into "Environment", "History", and "Connections" tabs. The "Environment" tab is active, showing "Global Environment" and "Environment is empty". Below this is a file explorer pane with tabs for "Files", "Plots", "Packages", "Help", and "Viewer". The file explorer shows the "Home" directory with a list of folders and files:

	Name	Size	Mo
<input type="checkbox"/>	2017		^
<input type="checkbox"/>	2018		
<input type="checkbox"/>	2019		
<input type="checkbox"/>	html		
<input type="checkbox"/>	identity	528 B	Se
<input type="checkbox"/>	Officeのカスタム テンプレート		
<input type="checkbox"/>	Outlook ファイル		
<input type="checkbox"/>	paper		
<input type="checkbox"/>	public_html		
<input type="checkbox"/>	R		
<input type="checkbox"/>	エディタ		

# ディレクトリ変更

- ①Session、②Set Working Directory、③Choose Directory。



①デスクトップ上にある、②hogeを選択して、③Open。

# ディレクトリ変更

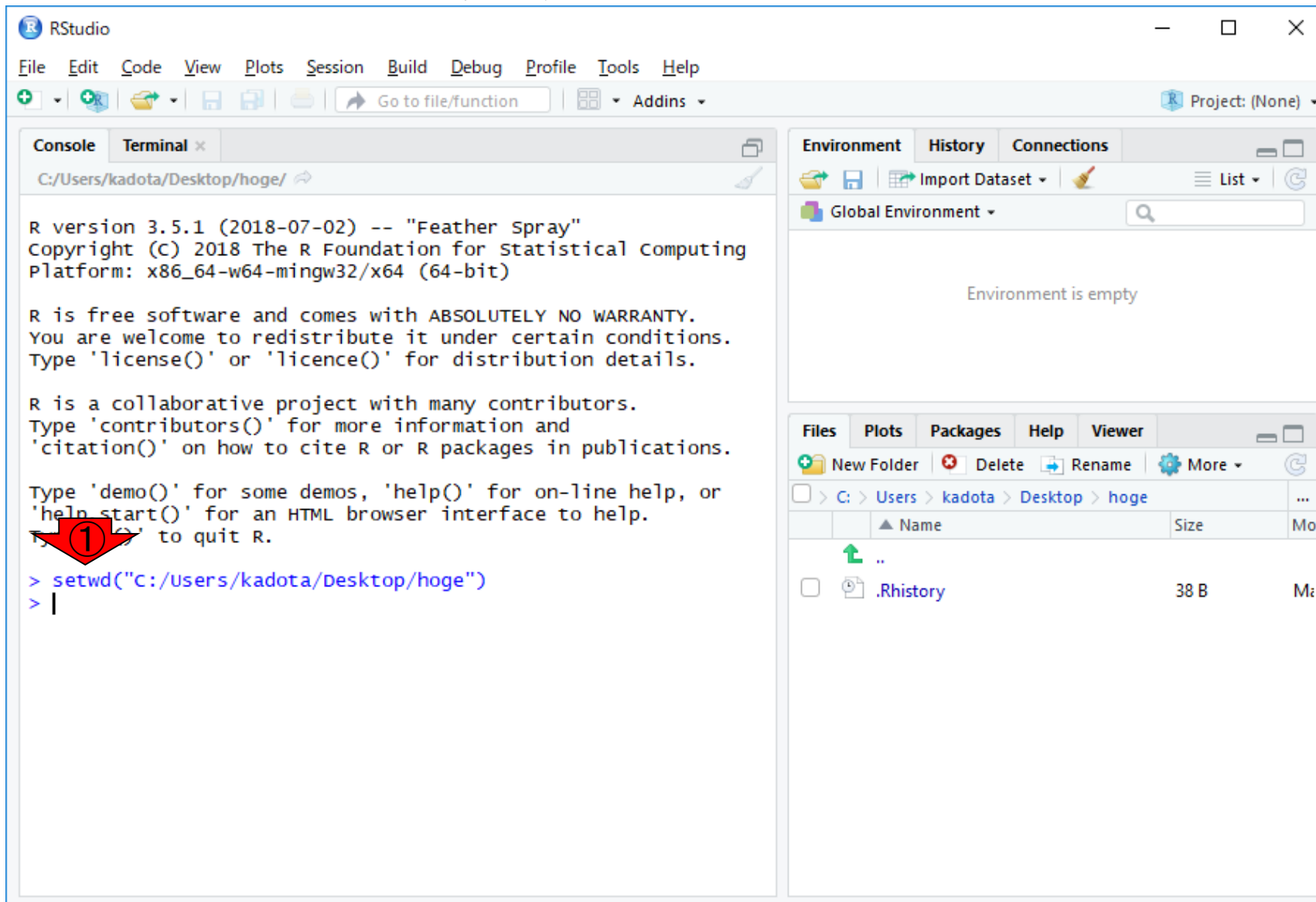
The screenshot shows the RStudio interface with the 'Choose Working Directory' dialog box open. The dialog box displays the current path as 'PC > デスクトップ >' (1). The left sidebar shows a list of folders, with 'デスクトップ' (Desktop) selected. The main area shows a list of folders on the Desktop, with 'hoge' selected (2). The 'フォルダー:' field at the bottom contains 'hoge'. The 'Open' button is highlighted with a red arrow (3).

RStudio Console Output:

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"  
Copyright (C) 2018 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
>
```

こんな感じになります。①内部的にsetwdという関数が動いていることがわかります。

# ディレクトリ変更



The screenshot shows the RStudio interface. The terminal window displays the R version information and the command `setwd("C:/Users/kadota/Desktop/hoge")` being executed. A red arrow with the number 1 points to the `setwd` function call. The file explorer on the right shows the current directory as `C:/Users/kadota/Desktop/hoge` and lists the files `..` and `.Rhistory`.

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
> setwd("C:/Users/kadota/Desktop/hoge")
> |
```

Name	Size	Mo
..		
.Rhistory	38 B	Ma

# ディレクトリ変更

こんな感じになります。①内部的にsetwdという関数が動いていることがわかります。②や③の表示結果から、確かにディレクトリ変更できていることがわかります。

The screenshot shows the RStudio interface. The terminal window displays the R startup messages, including the version (3.5.1) and the current directory (C:/Users/kadota/Desktop/hoge/). A red arrow labeled ① points to the terminal output. The file explorer window shows the current directory (C:/Users/kadota/Desktop/hoge/) and contains a file named .Rhistory. A red arrow labeled ③ points to the file explorer window. A red arrow labeled ② points to the terminal output.

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:/Users/kadota/Desktop/hoge/")
> |
```

Environment: Global Environment (empty)

Files: C:/Users/kadota/Desktop/hoge/

Name	Size	Modified
..		
.Rhistory	38 B	

# getwd

こんな感じになります。①内部的にsetwdという関数が動いていることがわかります。②や③の表示結果から、確かにディレクトリ変更できていることがわかります。getwd()で作業ディレクトリの確認をすべく、④getwまで打つとこんな感じになります。関数名をうろ覚えでも、候補がきっちり表示されるのがR単体とは違うところですね。目的の⑤を選択すると…

RStudio Console Terminal

```
C:/Users/kadota/Desktop/hoge/
```

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:/Users/kadota/Desktop/hoge")
> getwd
```

Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

C: > Users > kadota > Desktop > hoge

Name	Size	Mo
..		
.Rhistory	38 B	Ma

getwd [base]

- getwindowsHandle {utils}
- getwindowsHandles {utils}
- getWindowTitle {utils}
- getwinProgressBar {utils}
- getAnywhere {utils}

getwd()

getwd returns an absolute filepath representing the current working directory of the R process; setwd(dir) is used to set the working directory to dir.

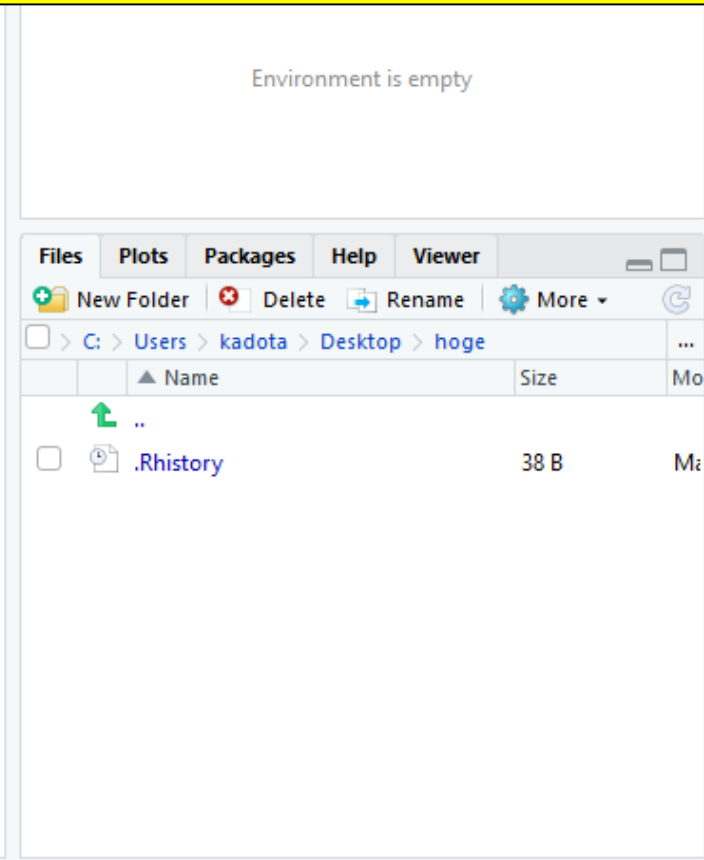
Press F1 for additional help



# getwd

こんな感じになります。①内部的にsetwdという関数が動いていることがわかります。②や③の表示結果から、確かにディレクトリ変更できていることがわかります。getwd()で作業ディレクトリの確認をすべく、④getwまで打つとこんな感じになります。関数名をうろ覚えでも、候補がきっちり表示されるのがR単体とは違うところですね。目的の⑤を選択すると、こんな感じになります。この状態でリターンキーを押すと…

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Console Terminal x
C:/Users/kadota/Desktop/hoge/
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> setwd("C:/Users/kadota/Desktop/hoge")
> getwd()
```



# getwd

こんな感じになります。①内部的にsetwdという関数が動いていることがわかります。②や③の表示結果から、確かにディレクトリ変更できていることがわかります。getwd()で作業ディレクトリの確認をすべく、④getwまで打つとこんな感じになります。関数名をうろ覚えでも、候補がきっちり表示されるのがR単体とは違うところですね。目的の⑤を選択すると、こんな感じになります。この状態でリターンキーを押すと、⑥確かにディレクトリ変更できていることがわかりますが、RStudioの場合は③で見えているので実質的にgetwdは不要ですね。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Console Terminal x
C:/Users/kadota/Desktop/hoge/
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> setwd("C:/Users/kadota/Desktop/hoge")
> getwd()
[1] "C:/Users/kadota/Desktop/hoge"
> |
```

