

ゲノム情報解析基礎(第1回)

¹東京大学 情報学環・学際情報学府

²東京大学・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ユニット

³東京大学・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。

The screenshot shows a web browser window with the URL www.iu.a.u-tokyo.ac.jp/index.shtml. The page is titled "東京大学大学院 アグリバイオ Agricultural Bioinformatics".

① A red arrow points to the "教育プログラム" (Education Program) link in the left sidebar navigation menu.

② A red arrow points to the "各講義のページ" (Individual Lecture Pages) link in the "教育プログラム" section.

③ A red arrow points to the "ゲノム情報解析基礎" (Genome Information Analysis Basics) link in the "基礎" (Basic) section of the "各講義のページ" section.

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	農学生命情報科学特別演習			
セミナー・ 討論形式 研究指導	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

ゲノム情報解析基礎

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。④「ゲノム情報解析基礎」のページ。

東京大学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ユニット
Agricultural Bioinformatics Research Unit

ホーム > 教育プログラム > 各講義のページ > ゲノム情報解析基礎

2. ゲノム情報解析基礎

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。本科目では、データ解析環境（主にR）・公共塩基配列データベース全般、そしてメタゲノム解析に関する実習を含む講義を行います。

担当教員

門田幸二（農大・農・アグリバイオ / 准教授）
栗玉悠一（国立遺伝学研究所 / 特任研究員）
森 富史（国立遺伝学研究所 / 助教）

参考図書

藤澤幸・編、よくわかるバイオインフォマティクス入門、講談社、2018
功重秀雄・小野温徳・監修、生命科学データベース・ウェブツール、MEDSL、2018
功重秀雄、生命科学者のためのデータ解析実践指南、MEDSL、2019
長田道忠、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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（2020年度）

- 2020年04月07日（PC使用）
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
[Rで]塩基配列解析
[Rで]塩基配列解析のサブ
RStudio
- 2020年04月14日（PC使用）
講師：栗玉 悠一
- 2020年04月21日（PC使用）
講師：森 富史
- 2020年04月28日（PC使用）
講師：森 富史

ゲノム情報解析基礎

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。④「ゲノム情報解析基礎」のページ。⑤講義資料。

東京大学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ユニット
Agricultural Bioinformatics Research Unit

ホーム > 教育プログラム > 各講義のページ > ゲノム情報解析基礎

2. ゲノム情報解析基礎

授業の目標・概要

担当教員

門田幸二 (専任、農学・農学系、アグリバイオ / 准教授)
児玉悠一 (国立遺伝学研究所 / 特任研究員)
森宙史 (国立遺伝学研究所 / 助教)

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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 (2020年度)

- 2020年04月07日 (PC使用)
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日 (PC使用)
講師：児玉 悠一
- 2020年04月21日 (PC使用)
講師：森 宙史
- 2020年04月28日 (PC使用)
講師：森 宙史

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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 (2020年度)

- 2020年04月07日 (PC使用)
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日 (PC使用)
講師：児玉 悠一
- 2020年04月21日 (PC使用)
講師：森 宙史
- 2020年04月28日 (PC使用)
講師：森 宙史

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

Introduction1

- 04月13日月曜日(17:15-20:30)、門田幸二(かどた こうじ)
 - バイオインフォマティクスのイントロダクション
 - データ解析環境RStudioの基本的な利用法
- 04月14日火曜日(17:15-20:30)、児玉悠一(こだま ゆういち)
 - 生命科学における公共データベース全般
- 04月21日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用
- 04月28日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用

Introduction2

講義予定。全てPC使用予定です。2004年の設立当初からのアグリバイオの全体方針は「Rをなるべく利用」です。①教育プログラム、②各講義のページから見られる、③赤枠内の講義科目のうち…

The screenshot shows the website interface with several red arrows and boxes highlighting specific areas:

- Arrow 1:** Points to the '教育プログラム' (Education Program) link in the left sidebar.
- Arrow 2:** Points to the '各講義のページ' (Lecture Pages) link in the top navigation bar.
- Arrow 3:** Points to the '各講義のページ' (Lecture Pages) section, which is enclosed in a red border.

The '各講義のページ' section contains the following content:

(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	農学生命情報科学特別演習			
セミナー・討論形式研究指導	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	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			
講義・実習を一体化				
基礎	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・実習を一体化				

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

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。カテゴリーと

カテゴリー	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として利用した様々なツ
方法論	「基礎」の科目を土台として、手法、質量分析法など。データ解析やモデル選択、分子シミュレーション(加)について解説し

お知らせ

- ▶ 2019年度の授業スケジュールは2019年4月5日(水)から2019年12月10日(水)までです。受講希望者は2019年10月15日(火)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」を事務局までご連絡ください。



東京大学
THE UNIVERSITY OF TOKYO

Introduction3

講義予定。全てPC使用予定です。2004年の設立当初からのアグリバイオの全体方針は「Rをなるべく利用」です。①教育プログラム、②各講義のページから見られる、③赤枠内の講義科目のうち…少なくとも8科目でデータ解析環境Rを利用予定。

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

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(科目名をクリックすると各講義のページに移動します)に分けられます。カテゴリと

カテゴリー	目的
基礎	主にバイオインフォマティクスを利用した様々なツールの利用法、質量分析法など
方法論	「基礎」の科目を土台とした、システム生物学概論やモデル選択、分子シミュレーション(加)について解説します。

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
	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	
方法論 講義・実習を 一体化	フィールドインフォマティクス			
	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		
基礎 講義・実習を 一体化	生物配列解析基礎	バイオスタティスティクス基礎論		

Introduction4

講義予定。全てPC使用予定です。2004年の設立当初からのアグリバイオの全体方針は「Rをなるべく利用」です。①教育プログラム、②各講義のページから見られる、③赤枠内の講義科目のうち…少なくとも8科目でデータ解析環境Rを利用予定。プログラミング言語としては、④「生物配列解析基礎」でPerlの基礎を…

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

ようこそ!!
アグリバイオ
教育研究センター

お知らせ

- ▶ 2019年度の授業スケジュール
2019年4月5日現在、受講希望者はまだ少ないです。
- ▶ 2019年度の受講料
講義日当日に受講登録をする場合は16:30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。

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
▼ 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(科目名をクリックすると各講義のページに移動します) に分けられます。カテゴリーと

カテゴリー	目的
基礎	主にバイオインフォマティクス。生命科学のためのデータ解析環境Rを利用した様々なツールの活用。
方法論	「基礎」の科目を土台として、手法、質量分析法などを用いた解析やモデル選択、分子シミュレーション(加)について解説します。

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		

Introduction5

講義予定。全てPC使用予定です。2004年の設立当初からのアグリバイオの全体方針は「Rをなるべく利用」です。①教育プログラム、②各講義のページから見られる、③赤枠内の講義科目のうち…少なくとも8科目でデータ解析環境Rを利用予定。プログラミング言語としては、④「生物配列解析基礎」でPerlの基礎を、そして⑤「農学生命情報科学特論I」でPythonをみっちり教える予定です。昨年度までとは講義内容が異なりますのでご注意ください。

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

ようこそ!!
アグリバイオ
教育研究センター

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。

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
▼ 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(科目名をクリック)

カテゴリー	目的
基礎	主にバイオインフォマティクス。生命科学のためのデータ解析環境Rを利用した様々なツールの活用。
方法論	「基礎」の科目を土台として、手法、質量分析法などを用いた解析やモデル選択、分子シミュレーション(加)について解説します。

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)

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

Introduction6

バイオインフォマティクスのリテラシー（活用能力）として、**教えたことと実際に教えられることには多少乖離があります。**

■ 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。

Introduction7

バイオインフォマティクスのリテラシー（活用能力）として、教えたことと実際に教えられることには多少乖離があります。バイオインフォ全般については、①の参考書をご覧ください。

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019



お知らせ

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（2020年度）

- 2020年04月07日（PC使用）
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日（PC使用）
講師：児玉 悠一
- 2020年04月21日（PC使用）
講師：森 宙史
- 2020年04月28日（PC使用）
講師：森 宙史

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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（2020年度）

- 2020年04月07日（PC使用）
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日（PC使用）
講師：児玉 悠一
- 2020年04月21日（PC使用）
講師：森 宙史
- 2020年04月28日（PC使用）
講師：森 宙史

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

(Rで)塩基配列解析1

2. ゲノム情報解析基礎 | アグリバイ | ×

保護されていない通信 | iu.a.u-tokyo.ac.jp/lectures/AG02/

東京大学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ユニット
Agricultural Bioinformatics Research Unit

ホーム > 教育プログラム > 各講義のページ > 2. ゲノム情報解析基礎

2. ゲノム情報解析基礎

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。本科目では、データ解析環境（主にR）・公共塩基配列データベース全般、そしてメタゲノム解析に関する実習を併せて行います。

担当教員

門田幸二（准大・農・アグリバイオ / 准教授）
児玉悠一（国立遺伝学研究所 / 特任研究員）
森宙史（国立遺伝学研究所 / 助教）

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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（2020年度）

- 2020年04月07日（PC使用）
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日（PC使用）
講師：児玉 悠一
- 2020年04月21日（PC使用）
講師：森 宙史
- 2020年04月28日（PC使用）
講師：森 宙史

参考図書

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
坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSI、2019
長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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（2020年度）

- 2020年04月07日（PC使用）
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
(Rで)塩基配列解析
(Rで)塩基配列解析のサブ
RStudio
- 2020年04月14日（PC使用）
講師：児玉 悠一
- 2020年04月21日（PC使用）
講師：森 宙史
- 2020年04月28日（PC使用）
講師：森 宙史

(Rで)塩基配列解析2

①をクリック。こんな感じになります。②は、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。

(Rで)塩基配列解析

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

(Rで)塩基配列解析

(last modified 2020/03/06, 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](#) ([Windows2019.10.09版](#)と[Macintosh2020.03.02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](#)(Windowsの2020.03.06版の[PPTX](#)と[PDF](#); [Macintosh2019.03.12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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は[申請要項](#) ([Word](#)、[PDF](#))をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップページ](#)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください

(Rで)塩基配列解析3

①をクリック。こんな感じになります。②は、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、③R本体とパッケージ群のインストール済み。④のリンク先は…

(Rで)塩基配列解析

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

(Rで)塩基配列解析

(last modified 2020/03/06, 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 \(Windows2019.10.09版とMacintosh2020.03.02版\)](#)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](#)(Windowsの2020.03.06版のPPTXとPDF ; [Macintosh2019.03.12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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は[申請要項 \(Word、PDF\)](#)をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップページ](#)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

サブ1



(Rで)塩基配列解析

(last modified 2020/03/06, 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 \(Windows2019.10.09版とMacintosh2020.03.02版\)](#)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](#)(Windowsの2020.03.06版のPPTXとPDF ; [Macintosh2019.03.12版](#))で習ってください。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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は[申請要項 \(Word、PDF\)](#)をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップページ](#)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください

サブ2

ちなみに①サブページのリンク先は、②のリンク先と同じ。最近は、クラウド解析環境(ウェブツール)でのNGS解析の教材を提供しています。

The screenshot shows a web browser window with the URL ru.a.u-tokyo.ac.jp/lectures/AG02/. The page title is '2. ゲノム情報解析基礎 | アグリバイ'. The main content area is titled '2. ゲノム情報解析基礎' and includes sections for '授業の目標・概要', '担当教員', '参考図書', 'お知らせ', and '講義日程 (2020年度)'. A red arrow points to the link '(Rで)塩基配列解析のサブ RStudio' in the '参考図書' section.

参考図書

- 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
- 坊農秀雅、生命科学者のためのデータ解析実践道場、MEDSi、2019
- 長田直樹、進化で読み解くバイオインフォマティクス入門、森北出版、2019

お知らせ

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 (2020年度)

- 2020年04月07日 (PC使用)
講師：門田 幸二
講義資料PDF(最終更新：2020.03.xx)
[\(Rで\)塩基配列解析](#)
[\(Rで\)塩基配列解析のサブ RStudio](#)
- 2020年04月14日 (PC使用)
講師：児玉 悠一
- 2020年04月21日 (PC使用)
講師：森 宙史
- 2020年04月28日 (PC使用)
講師：森 宙史

②

サブ3

ちなみに①サブページのリンク先は、②のリンク先と同じ。最近は、クラウド解析環境(ウェブツール)でのNGS解析の教材を提供しています。③サブの見た目は、さきほどのメインのページと同じです。④下矢印キーを押してページ下部に移動。

(Rで)塩基配列解析のサブ

(last modified 2020/02/23, since 2018)

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。

What's new?

- 2020年も[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。受講ガイダンスは、2月14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。東京大学の学生以外で受講を希望される方は、[申請要項](#)([Word形式](#)、[PDF形式](#))をご確認の上、2020年3月19日(木)までにお申込みください。2020年3月31日(火)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外で受講を希望される方は、[申請要項](#)が必要になります。ご注意ください。(2020/02/23) **NEW**
- [日本乳酸菌学会誌](#)のNGS関連連載の[第14回分原稿PDF](#)を公開しました。ウェブ資料も公開しました。(2020/12/22)



サブ4

ちなみに①サブページのリンク先は、②のリンク先と同じ。最近は、クラウド解析環境(ウェブツール)でのNGS解析の教材を提供しています。③サブの見た目は、さきほどのメインのページと同じです。④下矢印キーを押してページ下部に移動。このあたり(項目名が並んでいる最後のほう)。

(Rで)塩基配列解析のサブ

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html

- 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 2018/05/10)
- 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](#) (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 2019/04/05)
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2019/12/23)
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2019/11/26)



サブ5

①がクラウド解析環境(ウェブツール Galaxy)でNGS解析を行う解説記事。

(Rで)塩基配列解析のサブ × +

← → ↻ ⓘ 保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html ゲスト

- 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 2019/04/05)
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2019/12/23)
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2019/11/26)

[トップページへ](#)

サブ6

①がクラウド解析環境(ウェブツール Galaxy)でNGS解析を行う解説記事。②のタイトルにあるDDBJ(DNA Data Bank of Japan)を含む公共データベースの話が...

(Rで)塩基配列解析のサブ × +
保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html

- 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 2019/04/05)
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2019/12/23)
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2019/11/26)

[トップページへ](#)



サブ7

①がクラウド解析環境(ウェブツール Galaxy)でNGS解析を行う解説記事。②のタイトルにあるDDBJ(DNA Data Bank of Japan)を含む公共データベースの話が…、③次回の予定講義内容。

- 04月07日火曜日(17:15-20:30)、門田幸二
 - バイオインフォマティクスのイントロダクション
 - データ解析環境RStudioの基本的な利用法
- 04月14日火曜日(17:15-20:30)、児玉悠一(こだま ゆういち)
 - 生命科学における公共データベース全般
- 04月21日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用
- 04月28日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用



サブ8

①がクラウド解析環境(ウェブツール Galaxy)でNGS解析を行う解説記事。②のタイトルにあるDDBJ(DNA Data Bank of Japan)を含む公共データベースの話が…、③次回の予定講義内容。④の一部である、アセンブリのアルゴリズムやゲノムアノテーション関連の話が…

(Rで)塩基配列解析のサブ

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html

- 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 2018/05/10)
- 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 2019/04/05)
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2019/12/23)
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2019/11/26)



[トップページへ](#)

サブ9

①がクラウド解析環境(ウェブツール Galaxy)でNGS解析を行う解説記事。②のタイトルにあるDDBJ(DNA Data Bank of Japan)を含む公共データベースの話が…、③次回の予定講義内容。④の一部である、アセンブリのアルゴリズムやゲノムアノテーション関連の話が、⑤で行われる予定(谷澤先生担当分)。

東京大学大学院農学系研究科
アグリバイオインフォマティクス
Agricultural Bioinformatics

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
▼ スケジュール

ようこそ!! アグリバイオインフォマティクス教育研究ユニット

プログラム概要

本プログラムで開講する講義科目は(カテゴリーと目的)に分けられます。カテゴリーと目的

カテゴリー	目的
基礎	主にバイオインフォマティクスを利用した様々なツールの利用法、質量分析法など
方法論	「基礎」の科目を土台として、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、モデリングと分子シミュレーション

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」を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	モデリングと分子シミュレーション
	オーム情報解析	機能ゲノム学	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	

サブ10

(Rで)塩基配列解析のサブ × +

← → ↻ ⓘ 保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html ゲスト

- 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 2019/04/05)
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2019/12/23)
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2019/11/26)

[トップページへ](#)



サブ11

⑥に関連して、メタゲノム解析の話が、
⑦で行われる予定。ここでは、dada2と
phyloseqという二つのRパッケージを利用
予定ですが…

- 04月07日火曜日(17:15-20:30)、門田幸二(かした ゆうじ)
 - バイオインフォマティクスのイントロダクション
 - データ解析環境RStudioの基本的な利用法
- 04月14日火曜日(17:15-20:30)、児玉悠一(こだま ゆういち)
 - 生命科学における公共データベース全般
- 04月21日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用
- 04月28日火曜日(17:15-20:30)、森 宙史(もり ひろし)
 - メタゲノム解析の基礎と応用



サブ12

⑥に関連して、メタゲノム解析の話が、⑦で行われる予定。ここでは、dada2とphyloseqという二つのRパッケージを利用予定ですが、⑧のほうの、⑨の手順通りに作業を行えば、これらのパッケージも自動的にインストールされます。

(Rで)塩基配列解析

(last modified 2020/03/06, 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 \(Windows2019.10.09版とMacintosh2020.03.02版\)](#)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](#)(Windowsの2020.03.06版のPPTXとPDF ; [Macintosh2019.03.12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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は[申請要項 \(Word、PDF\)](#)をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップページ](#)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

基本的な利用法1

アグリバイオでは、赤枠で示す多くの科目でデータ解析環境Rを利用予定なので...

The screenshot shows the website's navigation structure. Red arrows and boxes indicate the following steps:

- ①** Click on the "教育プログラム" (Education Program) menu item in the left sidebar.
- ②** Click on "各講義のページ" (Each Lecture Page) in the top navigation bar.
- ③** Click on a specific lecture topic, such as "農学生命情報科学特論 IV" (Advanced Topics in Agricultural Life Information Science IV), which is highlighted with a red box.

The "各講義のページ" (Each Lecture Page) section contains the following content:

(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	農学生命情報科学特別演習			
セミナー・討論形式 研究指導	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		

基本的な利用法2

アグリバイオでは、赤枠で示す多くの科目でデータ解析環境Rを利用予定なので、以降は、④で提供している、⑤「基本的な利用法」とほぼ同じ内容を伝授していきます。

(Rで)塩基配列解析

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

(Rで)塩基配列解析 ④

(last modified 2020/03/06, 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 \(Windows2019.10.09版とMacintosh2020.03.02版\)](#)に⑤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](#)(Windowsの2020.03.06版のPPTXとPDF ; [Macintosh2019.03.12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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は[申請要項 \(Word、PDF\)](#)をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップページ](#)へまでに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください

Contents

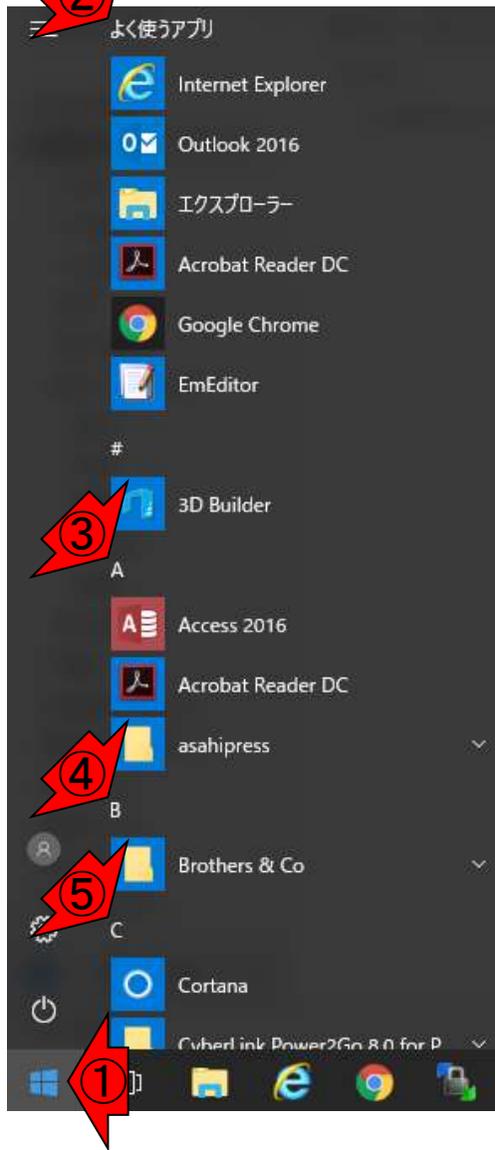
- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

RStudioの起動1

Windows10のデスクトップのスクショ。貸与PCのヒトは、おそらくデスクトップ上に①RStudioのアイコンがあるので、ダブルクリックで起動

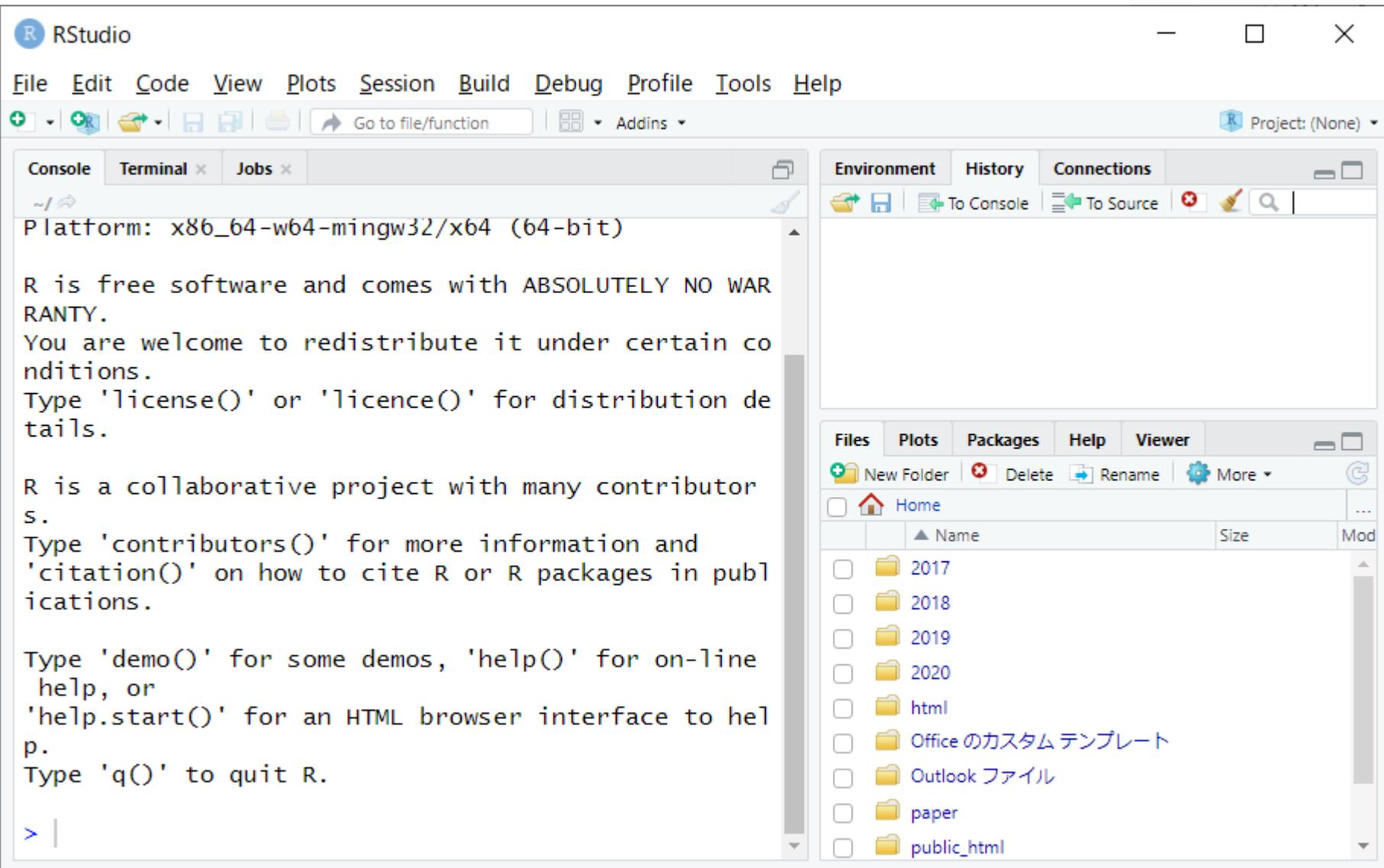


RStudioの起動2



Windows10の画面左下のあたりを表示。
②のあたりに着目して下のほうをずっと眺めていくと、③A、④B、⑤Cのようにアルファベット順になっていることがわかる。RからはじまるところにRStudioがあるはずですが。デスクトップアイコンを間違えて消してしまうヒトがいるので、その場合の対処法です。Excelについても同様に対処してください。

RStudioの起動3



Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

見栄えの統一1

RStudio起動後の見え方はヒトによって異なる。赤枠内が①こんな感じになっているヒトもいれば、

The screenshot shows the RStudio interface. The left pane contains the Console with the following text:

```
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 shows the Environment pane, which is currently empty. A red box highlights this area, and a red arrow with the number 1 points to it. Below the Environment pane is the Files pane, which shows a file explorer view of the Home directory with the following files and folders:

Name	Size	Mod
2017		
2018		
2019		
2020		
html		
Officeのカスタム テンプレート		
Outlook ファイル		
paper		
public_html		

見栄えの統一2

RStudio起動後の見え方はヒトによって異なる。赤枠内が①こんな感じになっているヒトもいれば、こんな感じでゴチャゴチャ文字が見えているヒトもいると思います。

The screenshot shows the RStudio interface with several panes. The top pane is the source editor, showing R code for installing BiocManager. The bottom-left pane is the console, showing help text for the `install.packages()` function. The bottom-right pane is the environment pane, showing the output of the `BiocManager::install()` function. The code in the source editor and the output in the environment pane are highlighted with red boxes.

```
5 install.packages("BiocManager") #BiocManager
6
7 #本番
8 BiocManager::install(param, update=F) #paramで
9
10 #後処理(ウェブサイト上で見えているバージョンと
11 BiocManager::valid(param) #TRUEはOK
12
```

```
BiocManager::install(param, update=F)
#後処理(ウェブサイト上で見えているバージョンと
BiocManager::valid(param) #TRUEはOK
packageVersion(param) #バージョン情報
library(rbamtools)
2λ1.5
```

```
~/
type `contributors()` for more information and
`citation()` on how to cite R or R packages in publ
ications.

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

> |
```

見栄えの統一3

RStudio起動後の見え方はヒトによって異なる。赤枠内が①こんな感じになっているヒトもいれば、こんな感じでゴチャゴチャ文字が見えているヒトもいると思います。何をもちごちゃごちゃしていると言っているのかがそもそもわからんというヒトもいることは承知の上。これは、主に持込PCでこのようになっているヒト向けの話。

The screenshot shows the RStudio interface with several panes. The top pane is the source editor, showing R code for installing BiocManager. The right pane is the Environment pane, showing the loaded packages. The bottom pane is the Console, showing the help text for the 'install.packages()' function. The bottom right pane is the File Explorer, showing the file system structure.

```
5 install.packages("BiocManager") #BiocManager
6
7 #本番
8 BiocManager::install(param, update=F) #paramで
9
10 #後処理(ウェブサイト上で見えているバージョンと
11 BiocManager::valid(param) #TRUEはOK
12
```

Environment

```
BiocManager
#後処理
BiocManager::valid(param) #TRUEはOK...
packageVersion(param) #バージョン情報...
library(rbamtools)
2λ1.5
```

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

Name	Size	Mod
.Rhistory	11.7 KB	Oct
2017		
2018		
2019		
2020		
html		
identity	528 B	Sep
Officeのカスタム テンプレート		
Outlook ファイル		

Console

```
~/
type `contributors()` for more information and
`citation()` on how to cite R or R packages in publ
ications.

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

> |
```


見栄えの統一5

①×を押すと、Saveするかどうかが聞かれるので②Don't Save。この意味がわかっていて、Saveしたいヒトは自己責任で自由に。こんな感じになります。左下の画面が上のほうに伸びて、左側全体を占拠したような感じになっていることが（直前のスライドとの比較でなんとなく）わかります。

The screenshot shows the RStudio interface. The console pane on the left contains the following text:

```
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 environment pane on the right shows the following objects:

```
BiocManager::install(param, update=...
#後処理(ウェブサイト上で見えているページ...
BiocManager::valid(param) #TRUEはOK...
packageVersion(param) #バージョン情報...
library(rbamtools)
2^1.5
```

The file explorer pane at the bottom right shows a file list:

Name	Size	Mod
.Rhistory	11.7 KB	Oct
2017		
2018		
2019		
2020		
html		
Officeのカスタムテンプレート		
Outlook ファイル		
paper		

見栄えの統一6

①×次に、①を押して、赤枠内にごちゃごちゃ書かれている(記載事項はヒトによって異なる)のをすっきりさせます。

The screenshot shows the RStudio interface. The left pane is the Console, displaying the R startup message. The right pane is the Environment pane, which is currently empty. A red box highlights the text in the Environment pane, and a red arrow points to it. The text in the Environment pane is:

```
BiocManager::install(param, update=...
#後処理(ウェブサイト上で見えているページ...
BiocManager::valid(param) #TRUEはOK...
packageVersion(param) #バージョン情報...
library(rbamtools)
2^1.5
```

The Environment pane also shows a file browser view of the Home directory, listing files and folders such as .Rhistory, 2017, 2018, 2019, 2020, html, Officeのカスタムテンプレート, Outlook ファイル, and paper.

見栄えの統一7

①×次に、①を押して、赤枠内にごちゃごちゃ書かれている(記載事項はヒトによって異なる)のをすっきりさせます。こんな感じになります。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main window is divided into several panes:

- Console:** Displays the R startup message, including the platform information (x86_64-w64-mingw32/x64 (64-bit)) and the license information. The text is clean and well-formatted.
- Environment:** Shows the current environment with tabs for Environment, History, and Connections.
- Files:** Shows a file explorer view of the Home directory, listing folders for years (2017-2020), html, Office templates, Outlook files, paper, and public_html.

```
~/  
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.  
  
> |
```

RStudioの終了

普通のソフトウェアと同様、①で終了。もし①×を押した後に、「Save workspace …」的なことを聞かれたら②Don't Save。

The screenshot shows the RStudio application window. The top right corner has a red arrow labeled '①' pointing to the 'X' button. A dialog box titled 'Quit R Session' is open in the center, asking 'Save workspace image to ~/.RData?'. The dialog has three buttons: 'Save', 'Don't Save', and 'Cancel'. A red arrow labeled '②' points to the 'Don't Save' button. The background shows the RStudio interface with a console window displaying the R startup message and a file browser showing a directory listing.

```
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.

> |
```

	Name	Size	Mod
<input type="checkbox"/>	2017		
<input type="checkbox"/>	2018		
<input type="checkbox"/>	2019		
<input type="checkbox"/>	2020		
<input type="checkbox"/>	html		
<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		

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

Console画面の見栄え1

おそらくまだ、貸与PCと持込PCで、① Console画面の②赤枠部分の文字の大きさや見栄えが異なっているヒトがいらっしゃると思います。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

Console Terminal x Jobs x

Environment History Connections

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size	Mod
<input type="checkbox"/>	2017		
<input type="checkbox"/>	2018		
<input type="checkbox"/>	2019		
<input type="checkbox"/>	2020		
<input type="checkbox"/>	html		
<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		

```
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.

> |
```

Console画面の見栄え2

おそらくまだ、貸与PCと持込PCで、① Console画面の②赤枠部分の文字の大きさや見栄えが異なっているヒトがいらっしゃると思います。RStudioのGUI画面を③の方向で右下に広げるなどしてみてください。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main workspace is divided into several panes: Environment, History, Connections, Files, Plots, Packages, Help, and Viewer. The Console pane is active and displays the following text:

```
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.

> |
```

A red box highlights the text in the Console pane. A red arrow labeled '1' points to the top-left corner of the RStudio window. A red arrow labeled '2' points to the top-right corner of the Console pane. A red arrow labeled '3' points to the bottom-right corner of the RStudio window.

Console画面の見栄え3

おそらくまだ、貸与PCと持込PCで、① Console画面の②赤枠部分の文字の大きさや見栄えが異なっているヒトがいらっしゃると思います。RStudioのGUI画面を③の方向で右下に広げるなどしてみてください。RStudioのGUI画面を③の方向で右下に広げて、全体を縮小した結果。②の部分の見栄えが自分のGUI画面と同じになって安心しているヒトもいると思います。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Console Terminal Jobs

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

> |

Environment

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

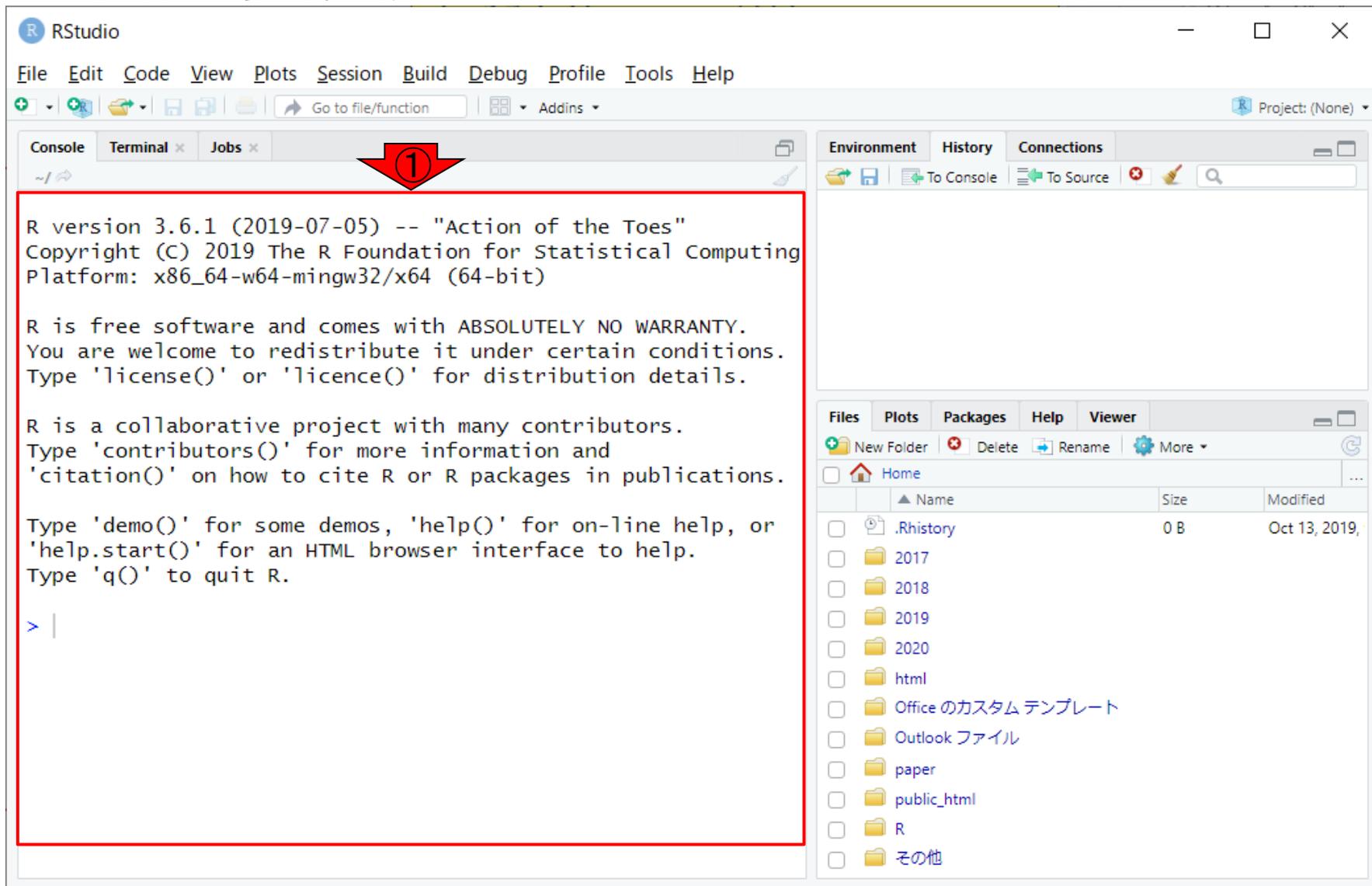
	Name	Size	Modified
<input type="checkbox"/>	.Rhistory	0 B	Oct 13, 2019,
<input type="checkbox"/>	2017		
<input type="checkbox"/>	2018		
<input type="checkbox"/>	2019		
<input type="checkbox"/>	2020		
<input type="checkbox"/>	html		
<input type="checkbox"/>	Officeのカスタム テンプレート		
<input type="checkbox"/>	Outlook ファイル		
<input type="checkbox"/>	paper		
<input type="checkbox"/>	public_html		
<input type="checkbox"/>	R		
<input type="checkbox"/>	その他		

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

設定変更1

まずは①赤枠内の文字サイズの変更方法から。



The screenshot shows the RStudio interface. The console window is highlighted with a red border, and a red arrow with the number 1 points to it. The console text is as follows:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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-hand side of the interface shows the Environment, History, and Connections panels, and a file explorer window displaying the contents of the Home directory.

Name	Size	Modified
.Rhistory	0 B	Oct 13, 2019,
2017		
2018		
2019		
2020		
html		
Officeのカスタム テンプレート		
Outlook ファイル		
paper		
public_html		
R		
その他		

設定変更2

まずは①赤枠内の文字サイズの変更方法から。②Toolsメニューの、③Global Options。

The screenshot shows the RStudio interface with the following elements:

- Terminal:** A red box labeled ① highlights the text in the terminal window, which includes R version information and usage instructions.
- Tools Menu:** A red arrow labeled ② points to the 'Tools' menu in the top navigation bar. The menu is open, showing options like 'Install Packages...', 'Check for Package Updates...', 'Version Control', 'Shell...', 'Terminal', 'Jobs', 'Addins', 'Keyboard Shortcuts Help', 'Modify Keyboard Shortcuts...', 'Project Options...', and 'Global Options...'. A red arrow labeled ③ points to the 'Global Options...' option.
- File Explorer:** A file explorer window is visible at the bottom right, showing a directory structure with folders like '.Rhistory', '2017', '2018', '2019', '2020', 'html', 'Office のカスタム テンプレート', 'Outlook ファイル', 'paper', 'public_html', 'R', and 'その他'.

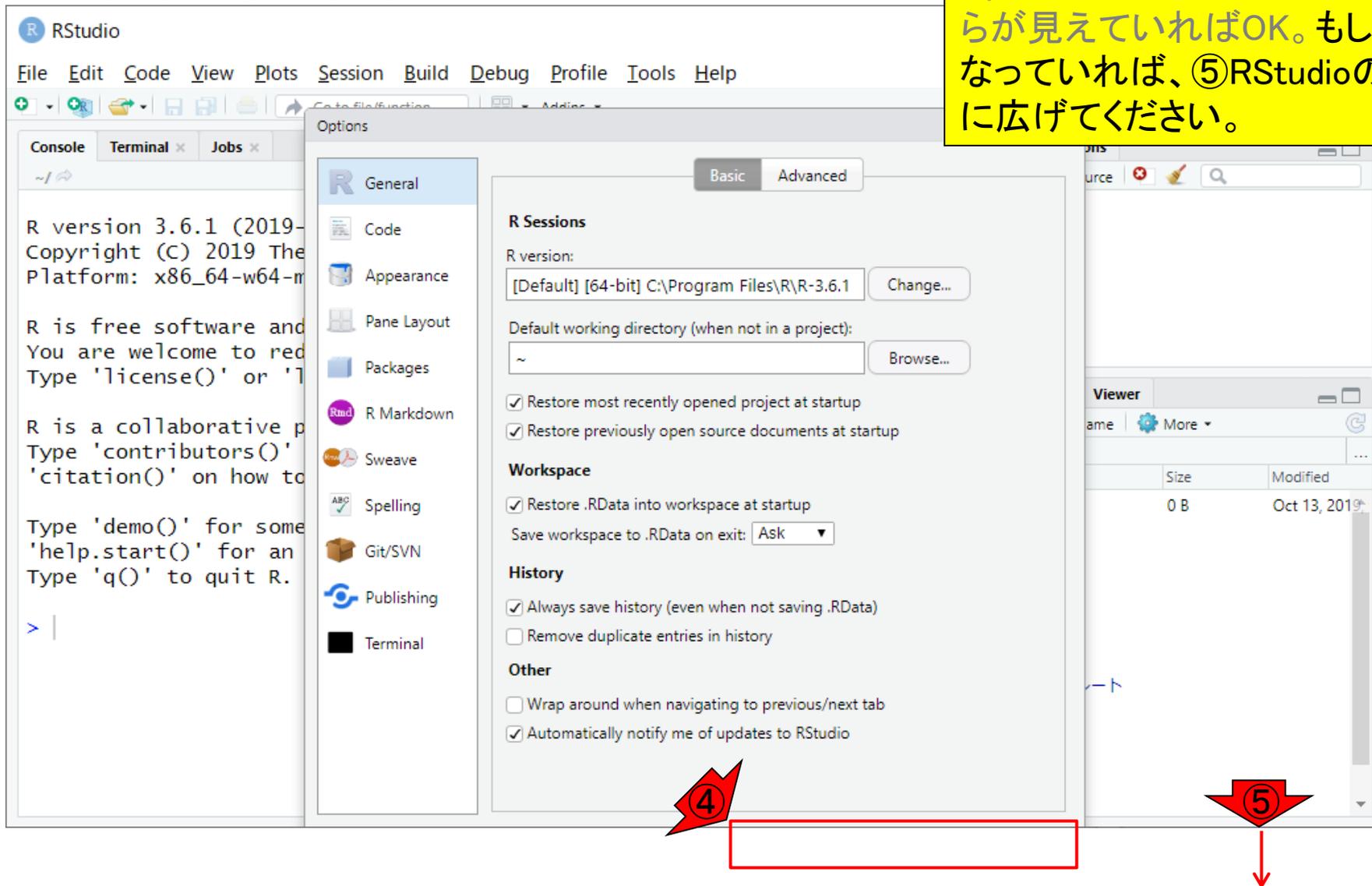
設定変更3

まずは①赤枠内の文字サイズの変更方法から。②Toolsメニューの、③Global Options。こんな感じになります。④これらが見えていればOK。

The screenshot shows the RStudio Options dialog box with the 'Basic' tab selected. The 'General' category is chosen in the left sidebar. The 'R Sessions' section includes fields for 'R version' (set to '[Default] [64-bit] C:\Program Files\R\R-3.6.1') and 'Default working directory' (set to '~'). The 'Workspace' section has checkboxes for 'Restore most recently opened project at startup' and 'Restore previously open source documents at startup', both checked. The 'History' section has 'Always save history (even when not saving .RData)' checked. The 'Other' section has 'Automatically notify me of updates to RStudio' checked. A red arrow with the number '4' points to the 'OK' button at the bottom right of the dialog box.

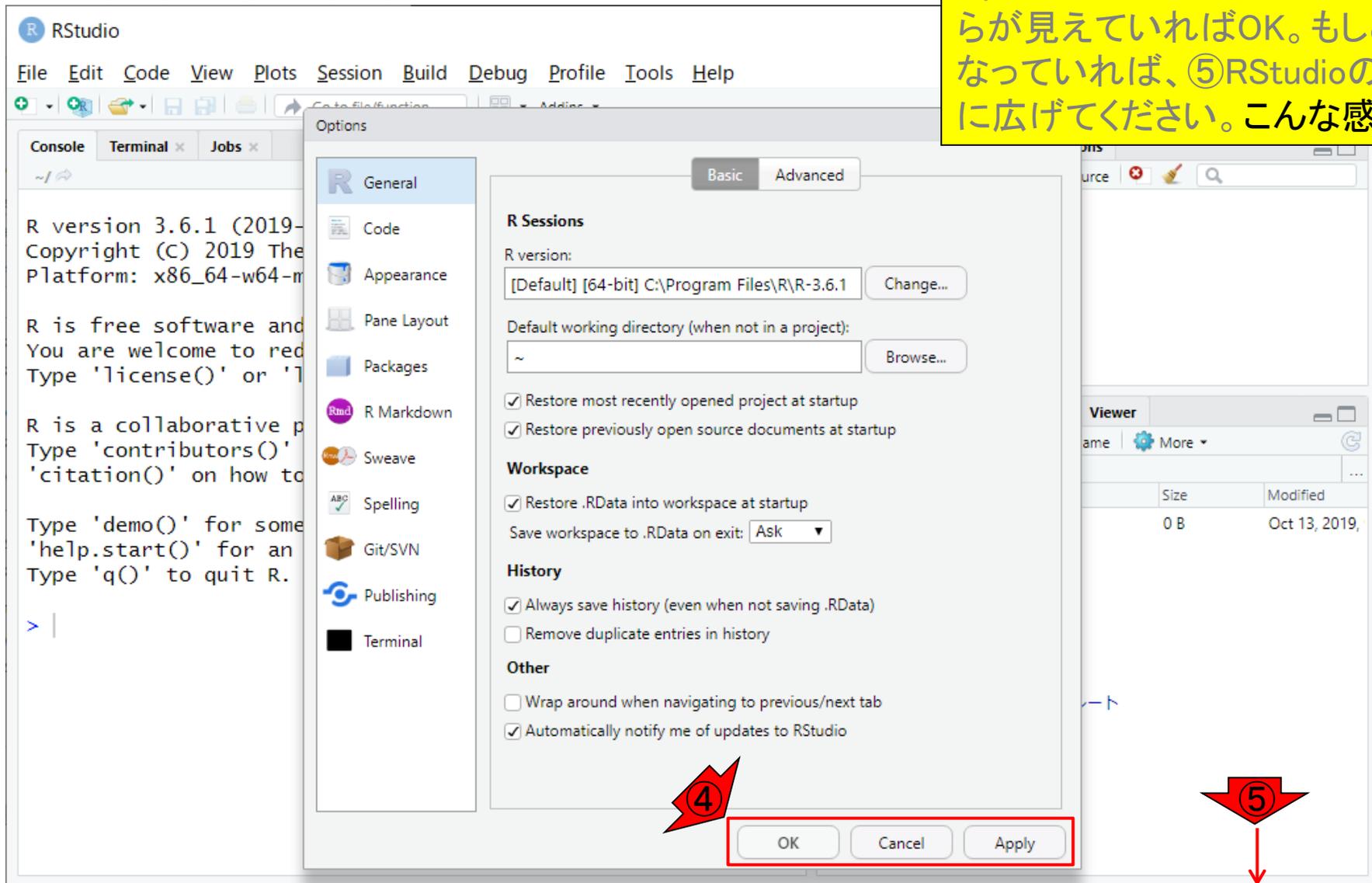
設定変更4

まずは①赤枠内の文字サイズの変更方法から。②Toolsメニューの、③Global Options。こんな感じになります。④これらが見えていればOK。もしこんな感じになっていれば、⑤RStudioのGUI画面を下に広げてください。



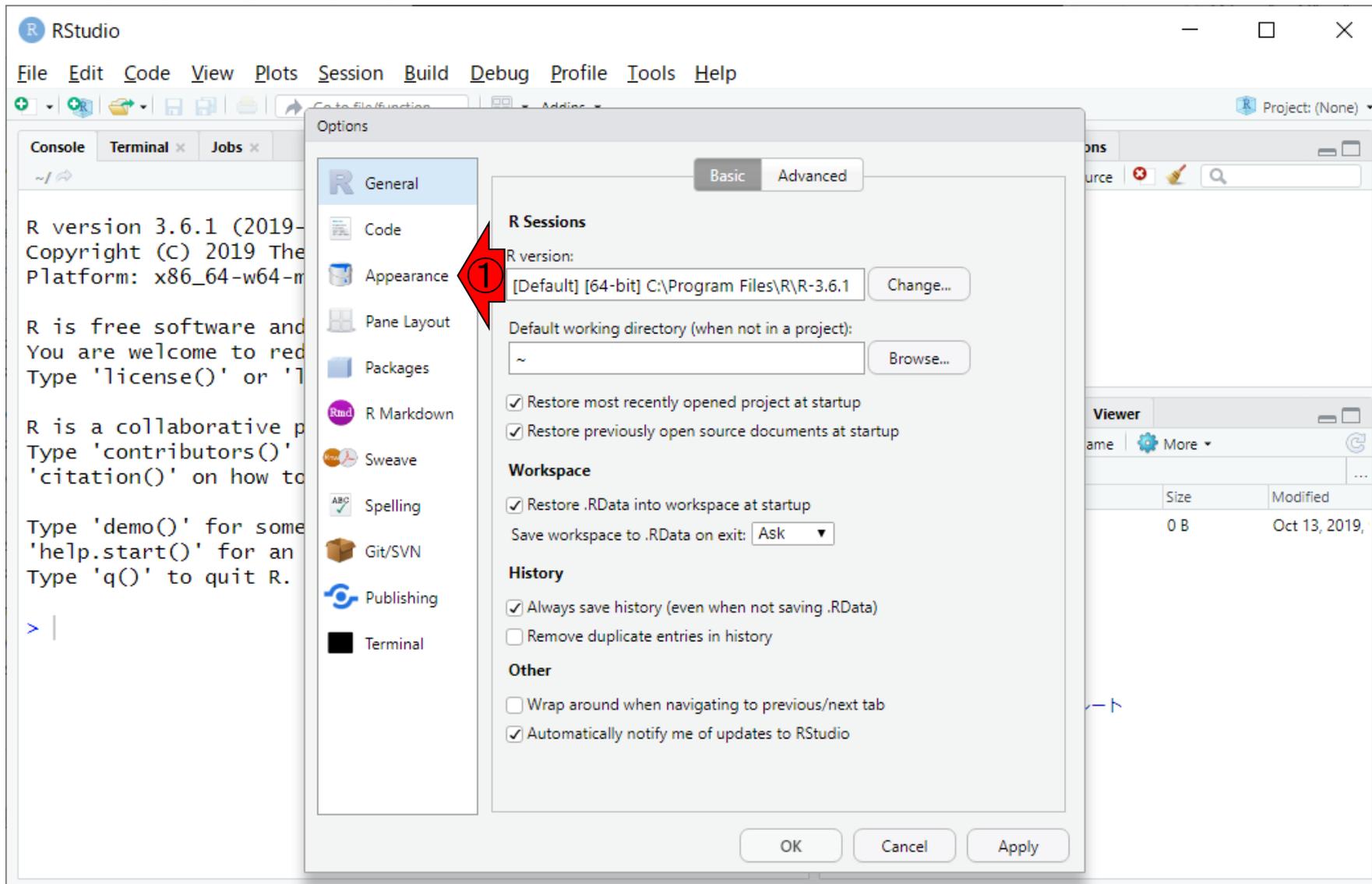
設定変更5

まずは①赤枠内の文字サイズの変更方法から。②Toolsメニューの、③Global Options。こんな感じになります。④これらが見えていればOK。もしこんな感じになっていれば、⑤RStudioのGUI画面を下に広げてください。こんな感じです。



設定変更6

本題の文字サイズ変更は、① Appearanceのところ可能です。



設定変更7

本題の文字サイズ変更は、① Appearanceのところ可能です。こんな感じになります。② Editor font sizeのところサイズ変更を行います。確かデフォルトは10だったと思いますが、私はすでに12に変更済みなのでこんな感じになっています。

The screenshot shows the RStudio Options dialog box with the 'Appearance' tab selected. The 'Editor font size' is set to 12. The 'Editor theme' is set to 'Material'. The 'Code' editor shows a snippet of R code. The 'Viewer' pane shows a table with columns 'Size' and 'Modified'.

```
# plotting of R object
plot <- function(x, y, ...)
{
  if (is.function(x) &&
      is.null(attr(x, "class")))
  {
    if (missing(y))
      y <- NULL

    # check for ylab argument
    hasylab <- function(...)
      !all(is.na(
        pmatch(names(list(...)),
                 "ylab")))

    if (hasylab(...))
      plot.function(x, y, ...)

    else
      plot.function(
        x, y,
        ylab = paste(
          deparse(substitute(x)),

```

Size	Modified
0 B	Oct 13, 2019,

設定変更8

本題の文字サイズ変更は、① Appearanceのところ可能です。こんな感じになります。② Editor font sizeのところサイズ変更を行います。確かデフォルトは10だったと思いますが、私はすでに12に変更済みなのでこんな感じになっています。③例えば14に変更して、④ Applyすると…

The screenshot shows the RStudio Options dialog box with the 'Appearance' tab selected. The 'Editor font size' is set to 14. The 'Apply' button is highlighted with a red arrow labeled 4. The background shows the RStudio interface with a code editor and a terminal window.

Annotations in the image:

- ① Points to the 'Appearance' tab in the Options dialog.
- ② Points to the 'Editor font size' dropdown menu.
- ③ Points to the '14' value in the 'Editor font size' dropdown.
- ④ Points to the 'Apply' button at the bottom of the dialog.

```
# plotting of R
plot <- function(x, y) {
  if (is.function(x) &&
      is.null(attr(x, "
  {
    if (missing(y))
      y <- NULL

  # check for ylab argument
  hasylab <- function(x) {
    !all(is.na(
      pmatch(names(1:length(x)),
              "ylab")))
  }

  if (hasylab(...))
    plot.function(x, y)
  else
    plot.function(x, y,
```

設定変更9

本題の文字サイズ変更は、① Appearanceのところ可能です。こんな感じになります。② Editor font sizeのところサイズ変更を行います。確かデフォルトは10だったと思いますが、私はすでに12に変更済みなのでこんな感じになっています。③例えば14に変更して、④ Applyすると、こんな感じになります。⑤ Console画面と呼ばれる赤枠部分の文字が大きくなったのがわかります。こんな感じで、見やすい大きさに変更してください。⑥OK。

The screenshot shows the RStudio Options dialog box with the 'Appearance' tab selected. The 'Editor font size' is set to 14. A red arrow labeled '5' points to the 'Console' tab in the background. Another red arrow labeled '6' points to the 'OK' button at the bottom of the dialog. The console output is visible in the background, showing R version 3.6.1 and some code snippets.

```
R version 3.6.1 (2019-10-10)
Copyright (c) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64

R is free software; you can redistribute it and/or modify it under the
terms of the GNU General Public License. You are welcome to share it
under the terms of the GNU General Public License.

R is a collaborative project with many contributors.

Type 'contributors()' for more information and
'help.start()' for an overview of R help, or
'help.start()' for an overview of R help, or
'help.start()' for an overview of R help.

# plotting of R
plot <- function(x, y) {
  if (is.function(x) && !is.null(attr("ylab"))) {
    if (missing(y)) {
      y <- NULL
    }
  }

  # check for ylab argument
  hasylab <- function(x) {
    !all(is.na(c(
      pmatch(names(list(...)),
        "ylab")))
  )

  if (hasylab(...)) {
    plot.function(x, y, ...)
  } else {
    plot.function(x, y, ...)
  }
}
```

ここまでで、「字が小さすぎて読めない！」問題が概ね解決。

設定変更10

The screenshot shows the RStudio interface. The console window displays the following text:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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.
```

The file explorer window shows the following table of files and folders:

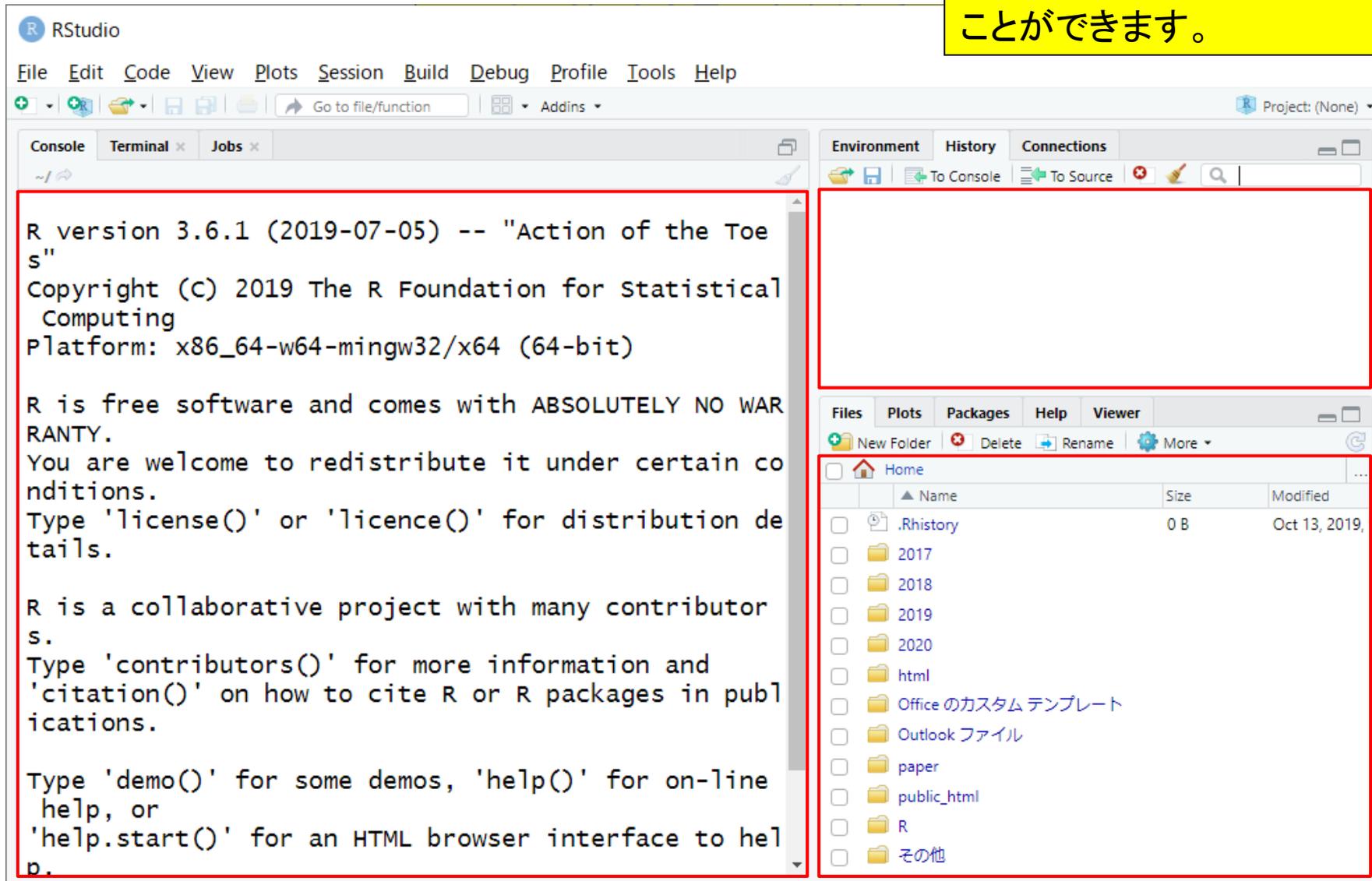
	Name	Size	Modified
<input type="checkbox"/>	.Rhistory	0 B	Oct 13, 2019,
<input type="checkbox"/>	2017		
<input type="checkbox"/>	2018		
<input type="checkbox"/>	2019		
<input type="checkbox"/>	2020		
<input type="checkbox"/>	html		
<input type="checkbox"/>	Officeのカスタム テンプレート		
<input type="checkbox"/>	Outlook ファイル		
<input type="checkbox"/>	paper		
<input type="checkbox"/>	public_html		
<input type="checkbox"/>	R		
<input type="checkbox"/>	その他		

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

枠の幅1

RStudioは、ここで見えているだけでも画面が3分割されていることがわかります。それぞれの画面の高さや幅を変更することができます。



枠の幅2

例えば、①や②のところにカーソルを持っていき、右に移動させると…

The screenshot shows the RStudio interface. The console pane on the left contains the following text:

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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.
```

The environment pane on the right shows a file explorer view of the Home directory. The file list includes:

Name	Size	Modified
.Rhistory	0 B	Oct 13, 2019
2017		
2018		
2019		
2020		
html		
Officeのカスタム テンプレート		
Outlook ファイル		
paper		
public_html		
R		
その他		

Red arrows labeled 1 and 2 point to the vertical scrollbar in the console pane, indicating the width of the scrollbar.

枠の幅3

例えば、①や②のところにカーソルを持っていき、右に移動させると…③
Console(こんそーる)画面の幅を広くすることができます。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

Console Terminal x Jobs x

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

> |
```

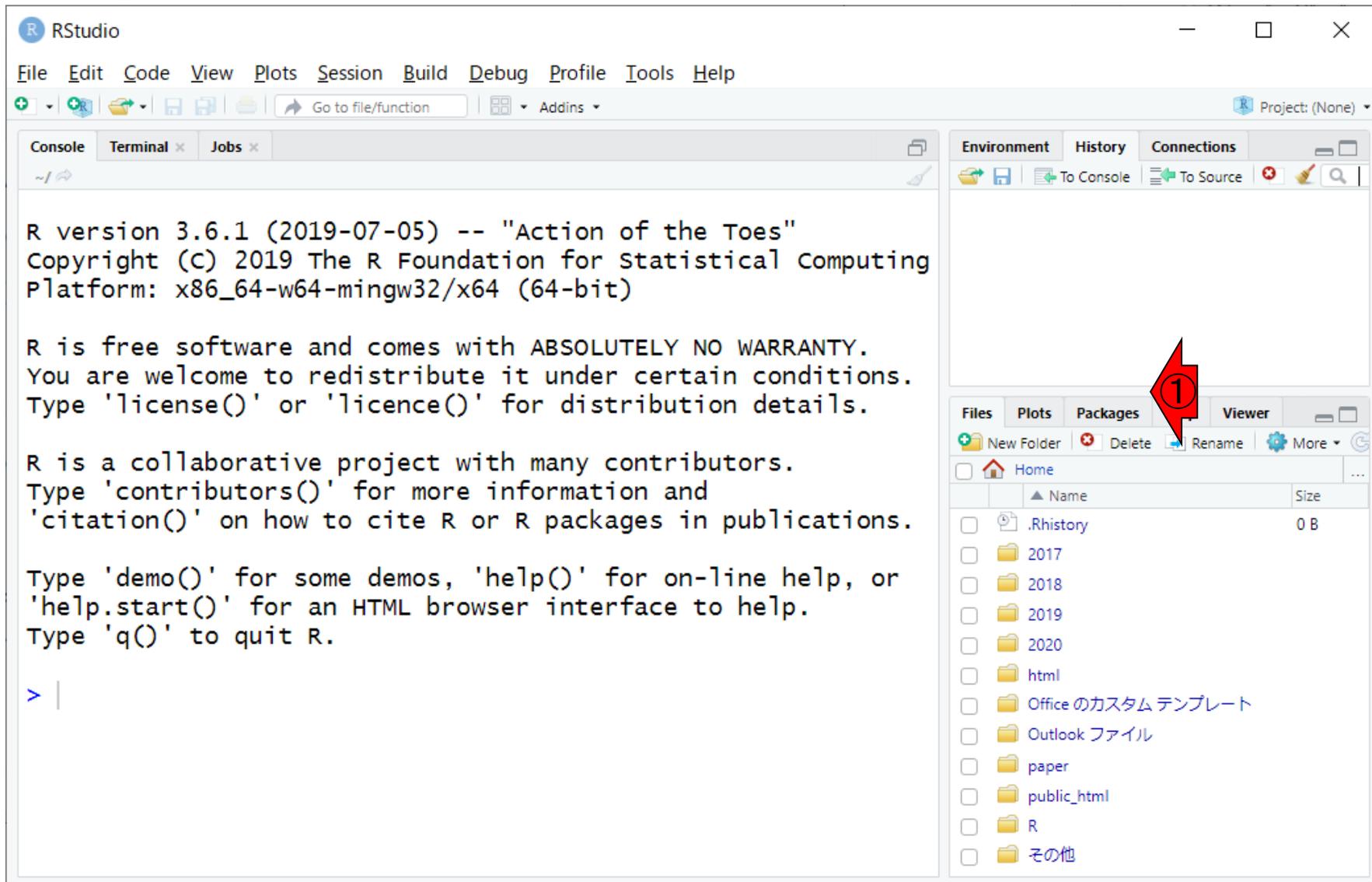
Environment History Connections

Files Plots Packages Help Viewer

Name	Size
.Rhistory	0 B
2017	
2018	
2019	
2020	
html	
Officeのカスタム テンプレート	
Outlook ファイル	
paper	
public_html	
R	
その他	

枠の幅4

①の部分も同様にして上下させることで
…。



The screenshot shows the RStudio interface. The console window on the left displays the R startup message and instructions. The file explorer on the right shows the 'Home' directory with a list of folders and files. A red arrow with the number 1 points to the 'Packages' tab in the file explorer.

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

> |
```

Name	Size
.Rhistory	0 B
2017	
2018	
2019	
2020	
html	
Officeのカスタム テンプレート	
Outlook ファイル	
paper	
public_html	
R	
その他	

枠の幅5

①の部分も同様にして上下させることで…右側の画面の上下幅を変更することができます。状況に応じて見やすいようにいじってください。

The screenshot shows the RStudio interface. The console on the left displays the R startup message and instructions. The file browser on the right shows a directory structure. A red arrow with a circled '1' points to the separator between the 'Plots' and 'Packages' tabs in the file browser panel.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

Console Terminal x Jobs x

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

> |
```

Environment History Connections

Files Plots Packages Viewer

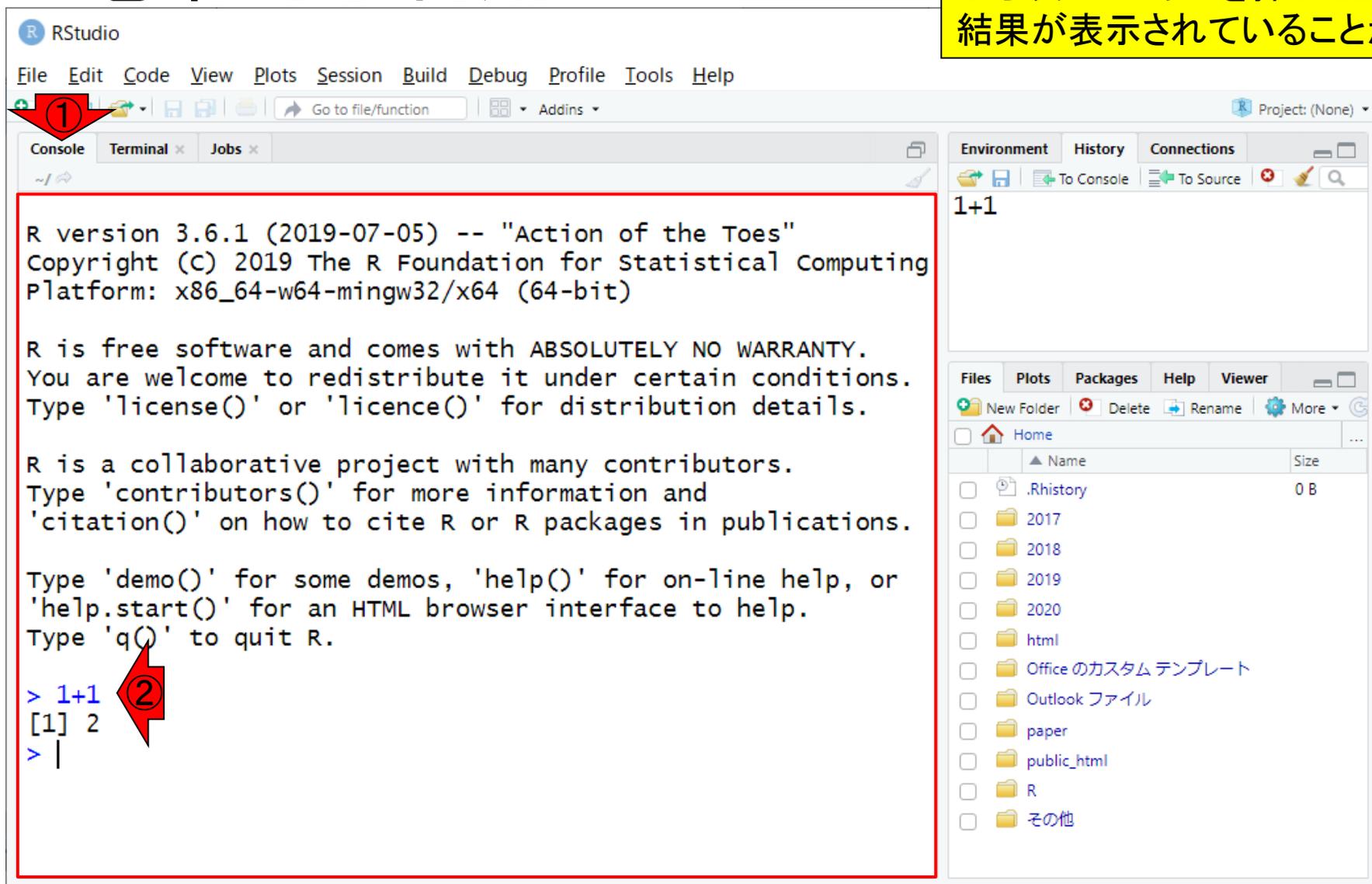
New Folder Delete Rename More

Home

	Name	Size
<input type="checkbox"/>	.Rhistory	0 B
<input type="checkbox"/>	2017	
<input type="checkbox"/>	2018	
<input type="checkbox"/>	2019	
<input type="checkbox"/>	2020	
<input type="checkbox"/>	html	
<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"/>	その他	

電卓的な利用1

(字が小さすぎて見辛いでしょうが…)①
Console画面上で、②「1+1」と打ち込んでからリターンキーを押したところ。2という結果が表示されていることがわかる。



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Environment History Connections

To Console To Source

1+1

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

Name	Size
.Rhistory	0 B
2017	
2018	
2019	
2020	
html	
Officeのカスタム テンプレート	
Outlook ファイル	
paper	
public_html	
R	
その他	

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

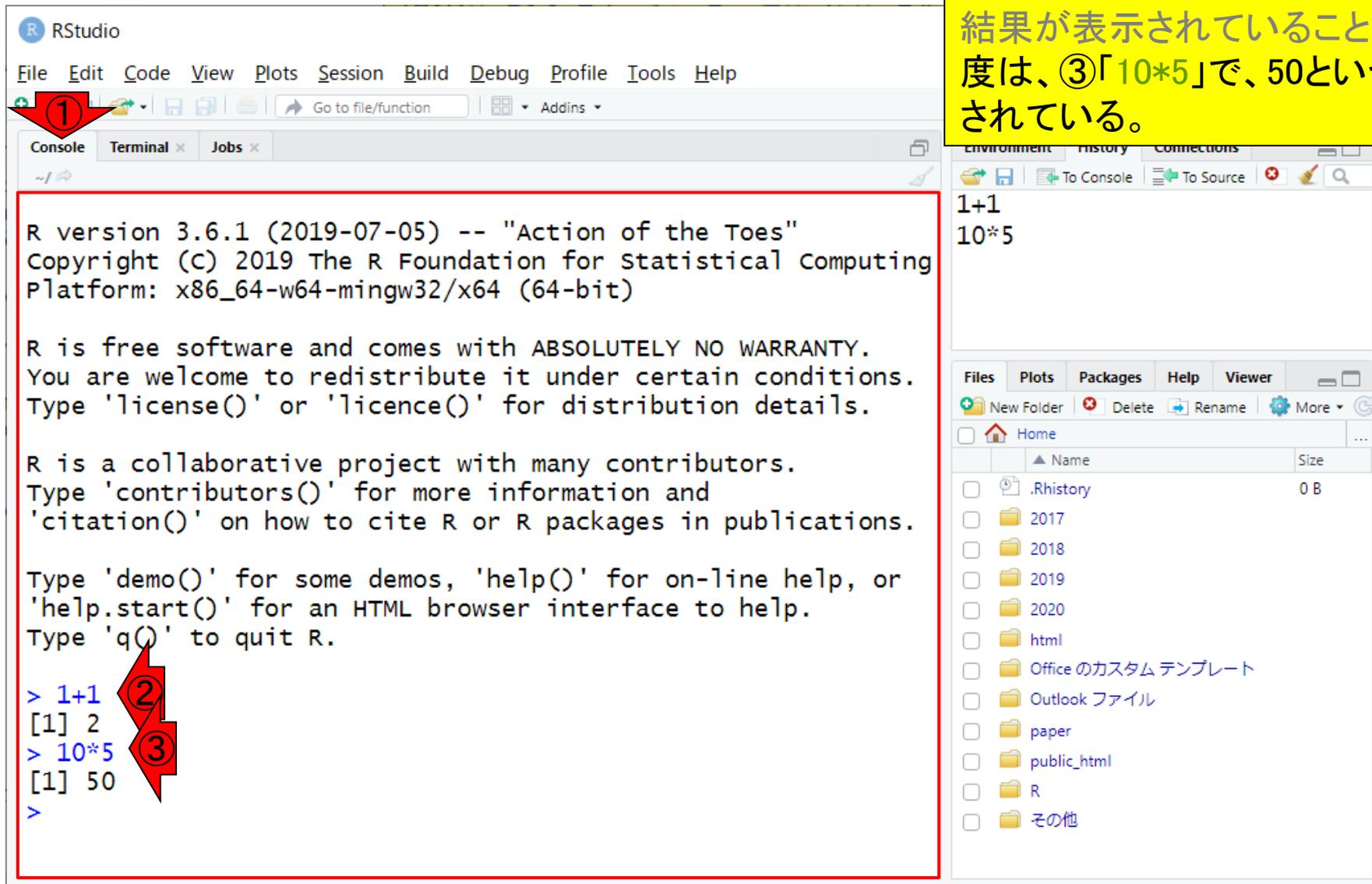
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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.

> 1+1
[1] 2
> |
```

電卓的な利用2

(字が小さすぎて見辛いでしょうが…)①
Console画面上で、②「1+1」と打ち込んでからリターンキーを押したところ。2という結果が表示されていることがわかる。今度は、③「10*5」で、50という結果が表示されている。



The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main console window displays the R startup message and user input/output. A red box highlights the console output, and red arrows with numbers 1, 2, and 3 point to specific elements: 1 points to the console window, 2 points to the input '1+1', and 3 points to the output '[1] 50'. A file explorer window is open in the bottom right, showing a directory structure with folders for years (2017-2020), html, Office templates, Outlook files, paper, public_html, R, and others.

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.

> 1+1
[1] 2
> 10*5
[1] 50
>
```

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

プロンプト (>)

ここまでの作業で、コマンドを実行すると、その結果が表示されるとともに、①「>」が表示されます。これを「プロンプト」と言います。これが見えていれば、「直前のコマンド実行が完了している」と解釈してよいです。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Console Terminal x Jobs x

```
R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (C) 2019 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.
```

```
> 1+1
[1] 2
> 10*5
[1] 50
> ①
```

Environment History Connections

To Console To Source

```
1+1
10*5
```

Files Plots Packages Help Viewer

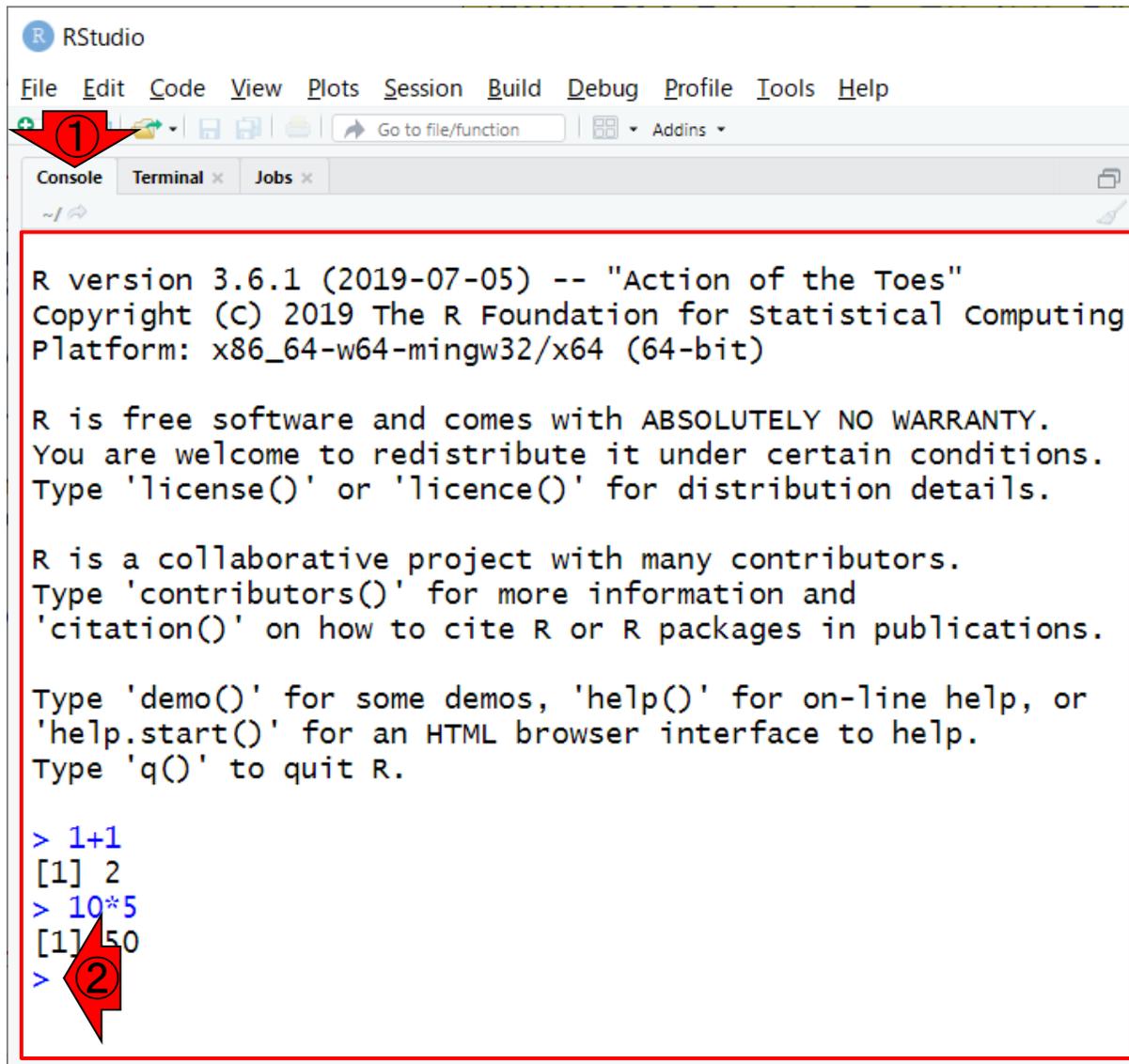
New Folder Delete Rename More

Home

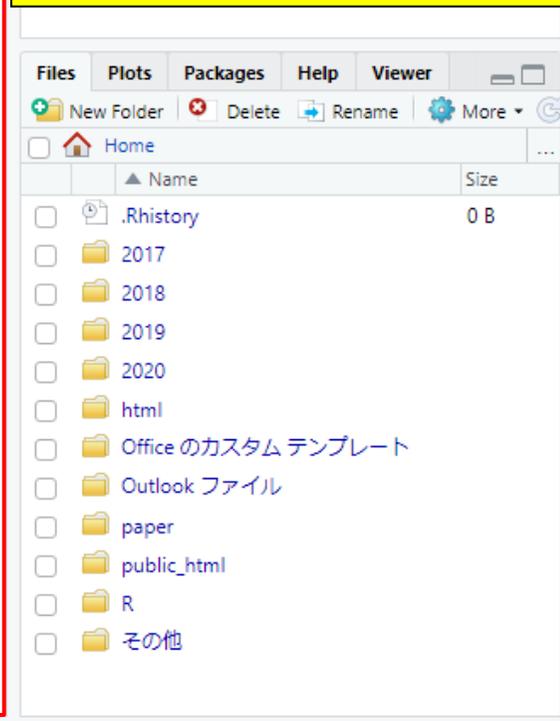
Name	Size
.Rhistory	0 B
2017	
2018	
2019	
2020	
html	
Officeのカスタム テンプレート	
Outlook ファイル	
paper	
public_html	
R	
その他	

Console画面クリア1

このあと、一行一行コマンドを打ち込んでいくと、常に①Console画面上の下部である、②のあたりに入力をするようになります。しかし、これだと講義室で教わる側がスクリーン下部になり見えづらくなります。それゆえ、スクリーン下部が見えづらくなる状況を避けるべく、必要に応じて①Console画面をクリアー(Clear Console)して、上部でコマンド入力できるようにするやり方を示します。



```
RStudio  
File Edit Code View Plots Session Build Debug Profile Tools Help  
Go to file/function Addins  
Console Terminal Jobs  
~ /  
R version 3.6.1 (2019-07-05) -- "Action of the Toes"  
Copyright (C) 2019 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.  
  
> 1+1  
[1] 2  
> 10*5  
[1] 50  
>
```



Console画面クリア2

①を押すだけ。②Editメニューから、③ Clear Console。④キーボード上で「Ctrl」キーを押しながら「L」キーを押すのもよい。

The screenshot shows the RStudio interface with the console window open. The console contains the following text:

```
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.

> 1+1
[1] 2
> 10*5
[1] 50
>
```

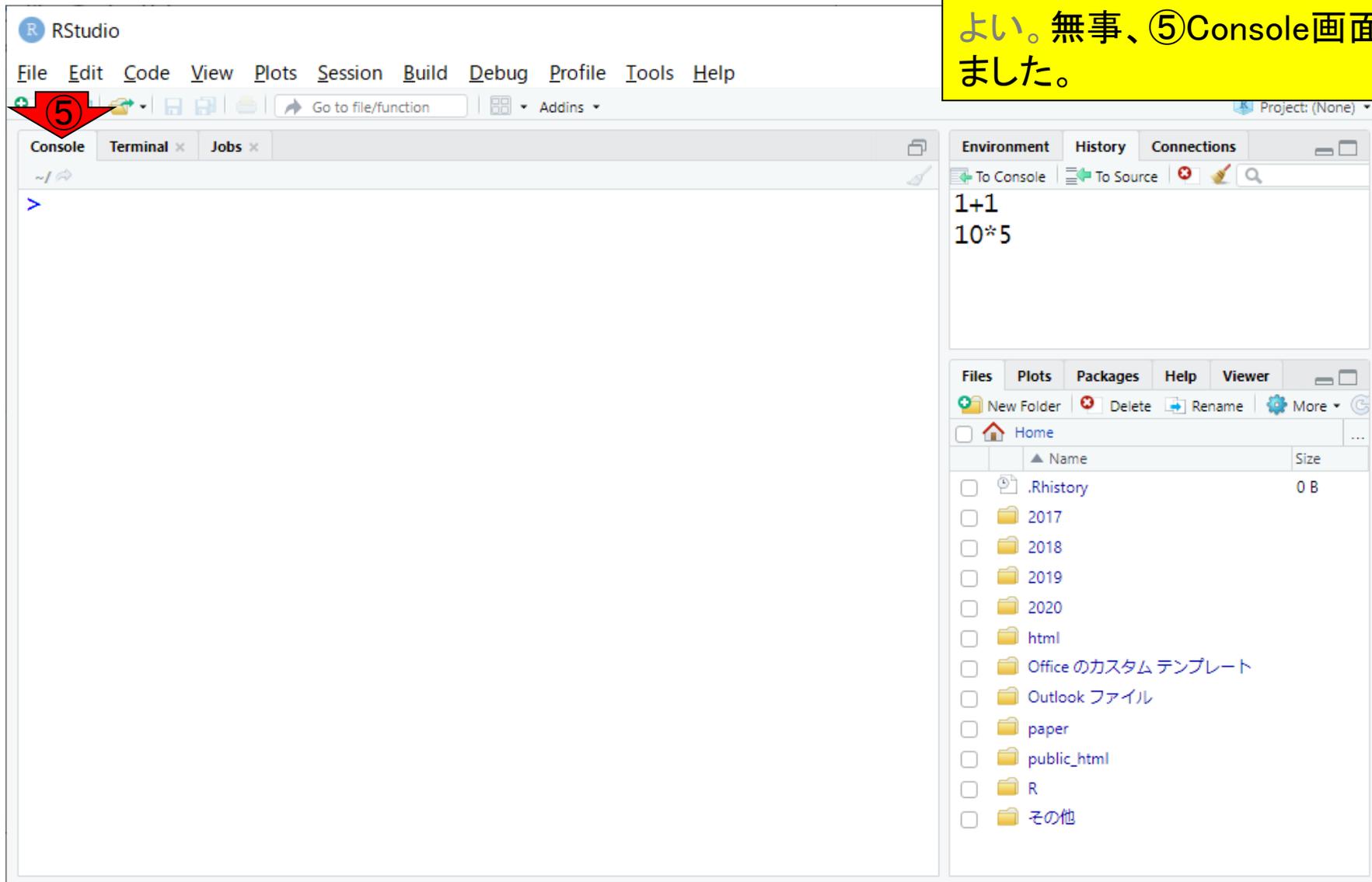
Annotations on the screenshot:

- ②: Points to the Edit menu.
- ③: Points to the 'Clear Console' option in the Edit menu.
- ④: Points to the 'Clear Console' option in the Edit menu.
- ①: Points to the 'Ctrl+L' keyboard shortcut in the Edit menu.

The console also shows the output of the commands: `> 1+1` resulting in `[1] 2` and `> 10*5` resulting in `[1] 50`.

Console画面クリア3

①を押すだけ。②Editメニューから、③Clear Console。④キーボード上で「Ctrl」キーを押しながら「L」キーを押すのもよい。無事、⑤Console画面がクリアされました。

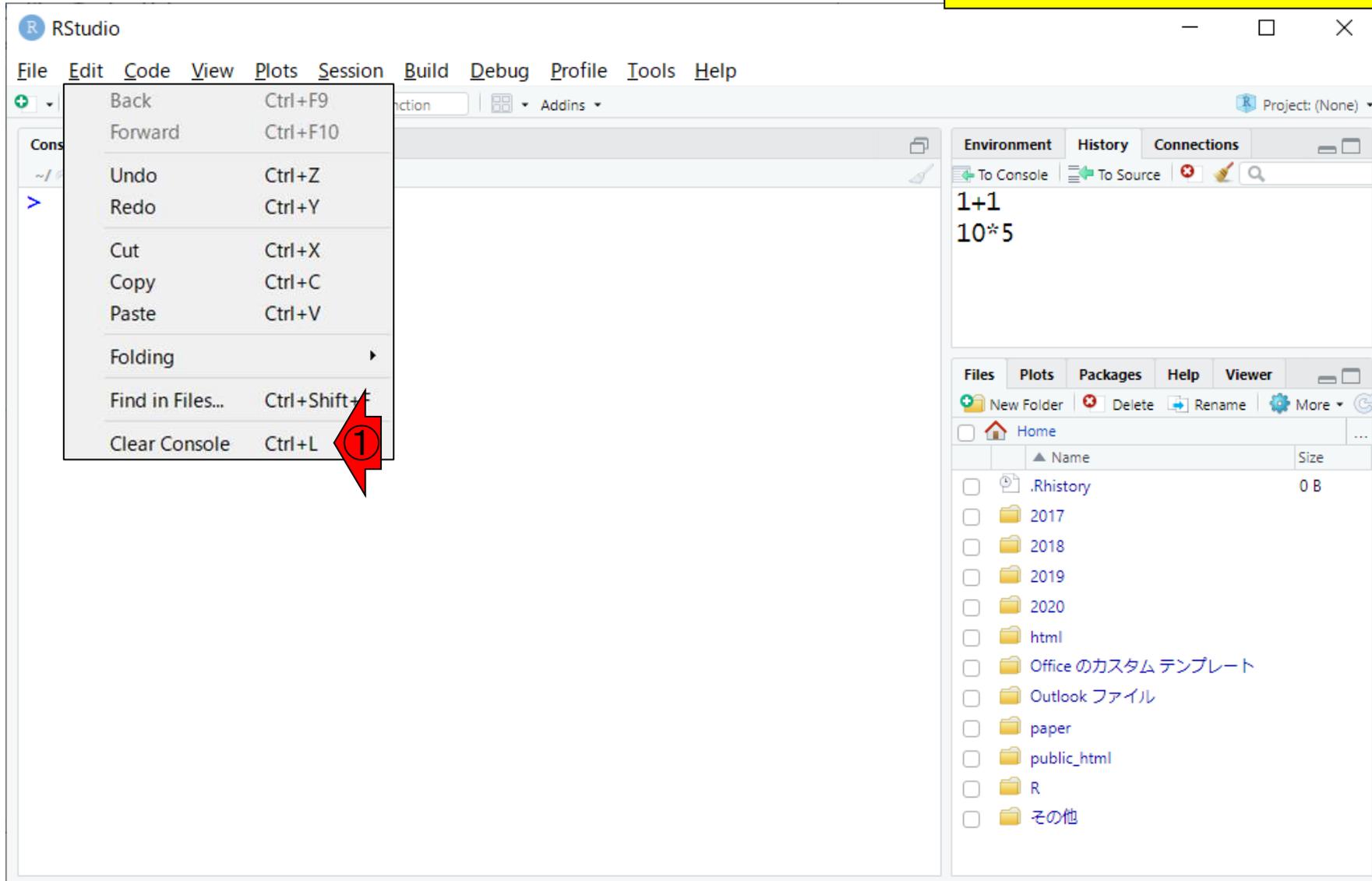


Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

計算の中断1

さきほどの①「Ctrl + L」は、「Ctrlキーを押しながらLキーを押す」という意味で使いました。



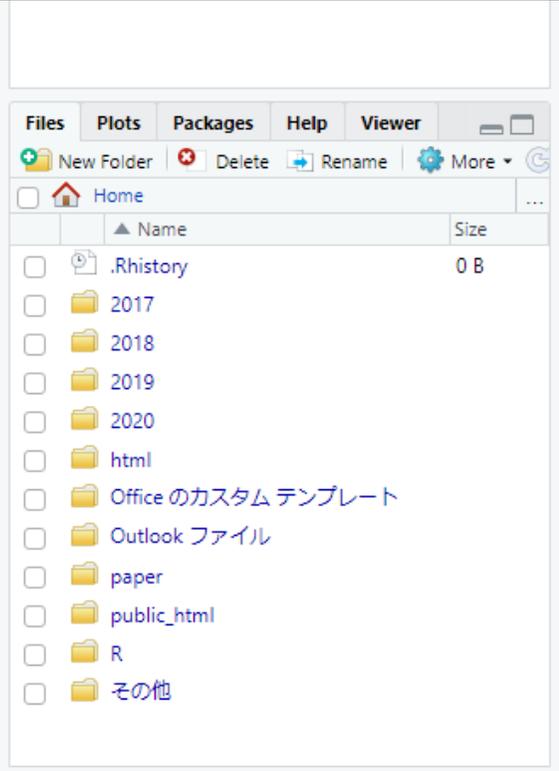
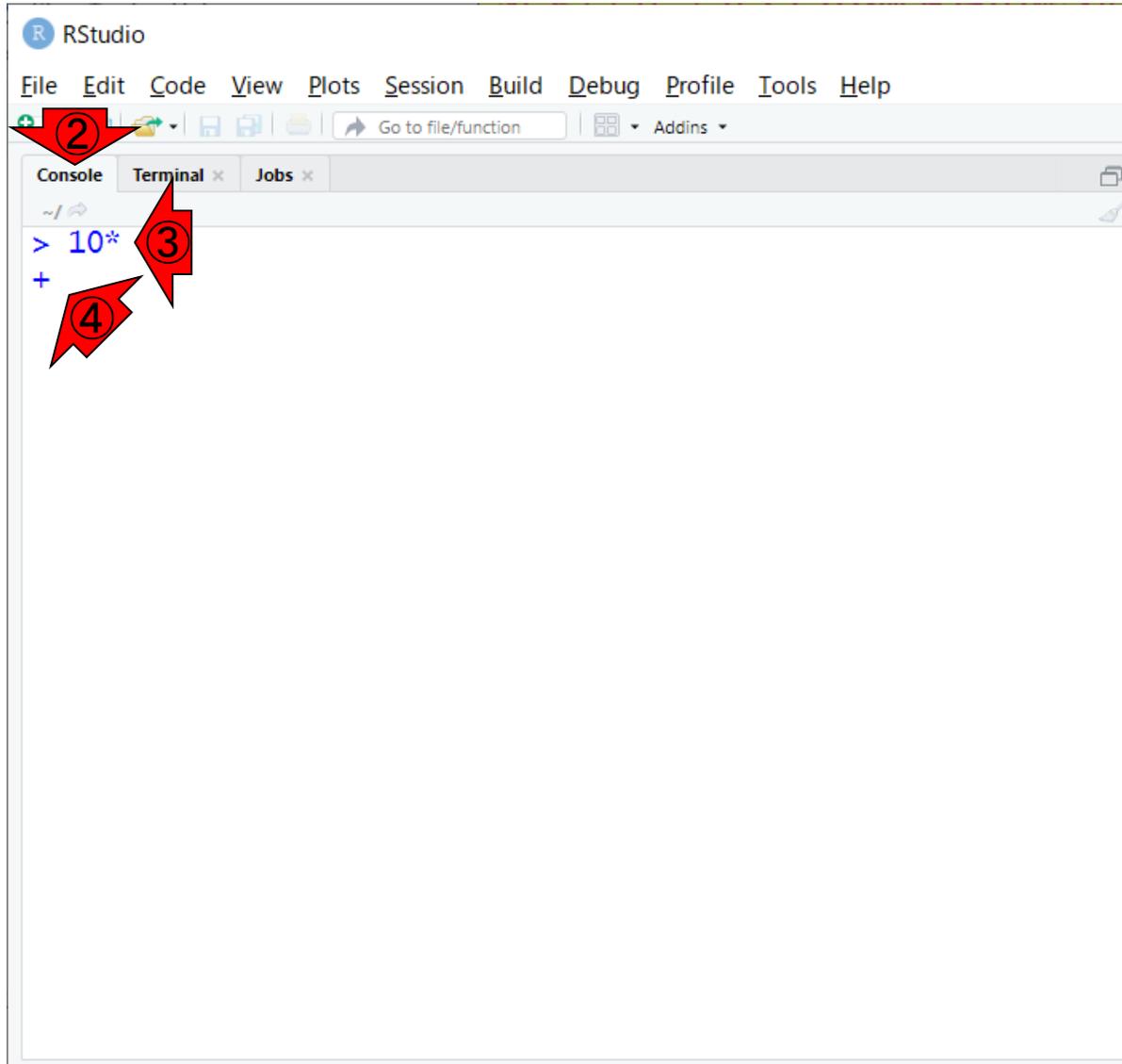
計算の中断2

さきほどの①「Ctrl + L」は、「Ctrlキーを押しながらLキーを押す」という意味で使いました。次に②Console画面上で、コマンド入力が不完全な状態のときに表示される「+」とその解決策を示します。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for saving, opening, and navigating. The main console area on the left shows a prompt character '>' and a red arrow pointing to it with the number '2'. The right-hand pane is divided into several sections: 'Environment', 'History', and 'Connections' at the top; 'Files', 'Plots', 'Packages', 'Help', and 'Viewer' below that. The 'Files' section shows a file explorer view of the user's home directory, listing folders such as .Rhistory, 2017, 2018, 2019, 2020, html, Office のカスタム テンプレート, Outlook ファイル, paper, public_html, R, and その他.

計算の中断3

さきほどの①「Ctrl + L」は、「Ctrlキーを押しながらLキーを押す」という意味で使いました。次に②Console画面上で、コマンド入力が不完全な状態のときに表示される「+」とその解決策を示します。例として、「10*5」と入力するつもりだったが、意図的に間違えて③「10*」まででリターンキーを押してみましょう。④のようになればOK。



計算の中断4

非推奨ですが、どうにかしてこのままの状態
で10*5を計算したい場合は①「5」と
打ち込んでリターンキーを押せばよい。

The screenshot shows the RStudio interface. The console window on the left contains the text `> 10*` followed by a cursor. A red arrow with the number 1 inside points to the cursor. The Environment pane on the right shows the current session state with `1+1`, `10*5`, and `10*` listed. The Files pane at the bottom right shows a file explorer view of the user's home directory.

計算の中断5

非推奨ですが、どうにかしてこのままの状態
で10*5を計算したい場合は①「5」と
打ち込んでリターンキーを押せばよい。
②無事「10*5」の計算結果である50が表
示され、③プロンプトが表示されているこ
とがわかります。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Console Terminal x Jobs x

```
> 10*  
+ 5  
[1] 50  
>
```

Environment History Connections

To Console To Source

```
1+1  
10*5  
10*  
5
```

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size
<input type="checkbox"/>	.Rhistory	0 B
<input type="checkbox"/>	2017	
<input type="checkbox"/>	2018	
<input type="checkbox"/>	2019	
<input type="checkbox"/>	2020	
<input type="checkbox"/>	html	
<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"/>	その他	

計算の中断7

①もう一回同じ状況にする。「プロンプト(>)」ではなく、②「+」になってしまったときの推奨手順は、③「Esc」キー。または「Ctrl + C」または「Ctrl + Z」。④「Esc」キーを押した結果。確かにプロンプト(>)が出ていますね。

RStudio interface showing a calculation interruption. The console displays the following sequence of commands and output:

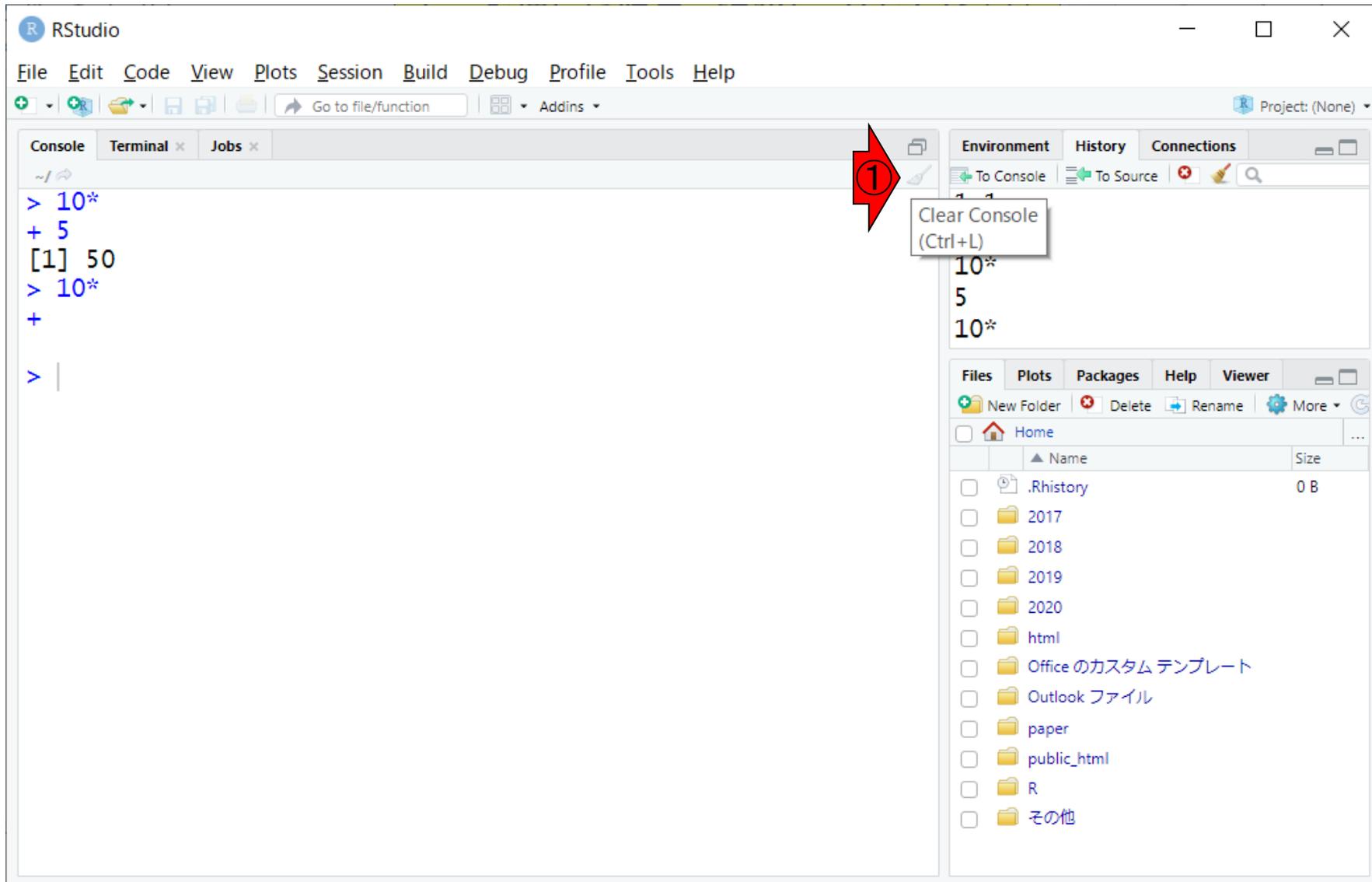
```
> 10*  
+ 5  
[1] 50  
> 10*  
+  
> |
```

A red arrow with the number 4 points to the prompt > |.

The Environment pane on the right shows the following variables:

- 1+1
- 10*5
- 10*
- 5
- 10*

計算の中断8



The screenshot shows the RStudio interface. The console pane on the left contains the following text:

```
> 10*  
+ 5  
[1] 50  
> 10*  
+  
> |
```

The Environment pane on the right shows a tooltip for the 'Clear Console (Ctrl+L)' button. The tooltip text is:

Clear Console
(Ctrl+L)

The Environment pane also displays the following text:

```
10*  
5  
10*
```

The Files pane on the right shows a file explorer view with the following items:

Name	Size
.Rhistory	0 B
2017	
2018	
2019	
2020	
html	
Officeのカスタム テンプレート	
Outlook ファイル	
paper	
public_html	
R	
その他	

計算の中断9

The screenshot shows the RStudio interface. The console window is empty, indicating that the previous session has been cleared. The environment window shows the history of the current session, listing the following expressions: `1+1`, `10*5`, `10*`, `5`, and `10*`. The file explorer on the right shows the current directory structure, including folders for years (2017-2020), `html`, `Officeのカスタム テンプレート`, `Outlook ファイル`, `paper`, `public_html`, `R`, and `その他`.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

Console Terminal x Jobs x

Environment History Connections

To Console To Source

1+1
10*5
10*
5
10*

Files Plots Packages Help Viewer

New Folder Delete Rename More

Home

	Name	Size
<input type="checkbox"/>	.Rhistory	0 B
<input type="checkbox"/>	2017	
<input type="checkbox"/>	2018	
<input type="checkbox"/>	2019	
<input type="checkbox"/>	2020	
<input type="checkbox"/>	html	
<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"/>	その他	

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

項目と例題

① (Rで)塩基配列解析のページでは、基本的な塩基配列解析から、NGSデータ取得、マッピング、統計解析、作図などができます。このウェブページは、サンプルデータと解析例を徹底的に充実させています。項目数が非常に多いですが、慣れです。項目だけのものや古いものも多数あります。

(Rで)塩基配列解析

x +

← → ↻ 保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#

(Rで)塩基配列解析

①

(last modified 2020/03/03, 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](#) ([Windows2019.10.09版](#)と[Macintosh2020.03.02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](#) ([Windows2019.03.12版](#)と [Macintosh2019.03.12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)

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3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。東京大学の学生以外で受講を希望される方は申請要項([Word](#)、[PDF](#))をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年3月31日(火)までに通知します。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですのでご注意ください。(2020/02/23) **NEW**

- [はじめに](#) (last modified 2018/08/04)
- [過去のお知らせ](#) (last modified 2020/03/03) **NEW**
- [インストール | について](#) (last modified 2019/10/09)
- インストール | R本体とRStudio | 最新版 | [Win用](#) (last modified 2019/10/08)推奨
- インストール | R本体とRStudio | 最新版 | [Mac用](#) (last modified 2020/03/03)推奨 **NEW**
- インストール | R本体 | 過去版 | [Win用](#) (last modified 2015/03/22)
- インストール | R本体 | 過去版 | [Mac用](#) (last modified 2015/03/22)
- [インストール | Rパッケージ | について](#) (last modified 2018/11/13)
- インストール | Rパッケージ | [必要最小限プラスアルファ](#) (last modified 2020/03/02)推奨 **NEW**
- インストール | Rパッケージ | [個別\(2018年11月以降\)](#) (last modified 2018/11/12)
- インストール | Rパッケージ | [個別\(2018年11月以前\)](#) (last modified 2018/11/12)

[トップページ](#)へ

Apr 13, 2020

項目と例題

①(Rで)塩基配列解析のページでは、基本的な塩基配列解析から、NGSデータ取得、マッピング、統計解析、作図などができます。このウェブページは、サンプルデータと解析例を徹底的に充実させています。項目数が非常に多いですが、慣れです。項目だけのものや古いものも多数あります。②をクリック。

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_g

- [インストール | Rパッケージ | 個別\(2018年11月以降\)](#) (last modified 2018/11)
- [インストール | Rパッケージ | 個別\(2018年11月以前\)](#) (last modified 2018/11)
- [基本的な利用法](#) (last modified 2019/03/12)
- [サンプルデータ](#) (last modified 2019/09/06)
- [イントロ | 一般 | ランダムに行を抽出](#) (last modified 2014/07/17)
- [イントロ | 一般 | 任意の文字列を行の最初に挿入](#) (last modified 2014/07/17)
- [イントロ | 一般 | 任意のキーワードを含む行を抽出\(基礎\)](#) (last modified 2019/04/24)
- [イントロ | 一般 | ランダムな塩基配列を生成](#) (last modified 2014/06/16)
- [イントロ | 一般 | 任意の長さの可能な全ての塩基配列を作成](#) (last modified 2015/02/19)
- [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 | について](#) (last modified 2019/09/06)
- [イントロ | 一般 | 指定した範囲の配列を取得 | Biostring](#) (last modified 2019/09/06)
- [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 | について](#) (last modified 2019/09/06)
- [イントロ | 一般 | 翻訳配列\(translate\)を取得 | Biostrings](#) ② (last modified 2015/09/12)
- [イントロ | 一般 | 翻訳配列\(translate\)を取得 | seqinr\(Charn_2005\)](#) (last modified 2015/03/09)
- [イントロ | 一般 | 相補鎖\(complement\)を取得](#) (last modified 2019/03/10)
- [イントロ | 一般 | 逆相補鎖\(reverse complement\)を取得](#) (last modified 2019/03/10)
- [イントロ | 一般 | 逆鎖\(reverse\)を取得](#) (last modified 2019/03/10)

[トップページへ](#)

項目と例題

こんな感じになります。①この項目は、塩基配列を入力として読み込んで、その翻訳配列であるアミノ酸配列を出力する際に利用します。

Intro | General | Get translation (translate) | 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存
```

[トップページへ](#)

項目と例題

こんな感じになります。①この項目は、塩基配列を入力として読み込んで、その翻訳配列であるアミノ酸配列を出力する際に利用します。大抵の項目は、いくつかの例題を提供しています。例えば、②が例題1で...

イントロ | 一般 | 翻訳配列(translate)を取得 | 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。

「②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存
```

[トップページへ](#)

項目と例題

こんな感じになります。①この項目は、塩基配列を入力として読み込んで、その翻訳配列であるアミノ酸配列を出力する際に利用します。大抵の項目は、いくつかの例題を提供しています。例えば、②が例題1で、③が例題2だと解釈します。この画面は6行分ほどページ下部に移動させただけです。④クリックする項目を間違えて困った人は、「トップページへ」で一番上からやり直してください。

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_g

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存
```

2. (multi-)FASTA形式ファイル(sample4.fasta)の場合 :

配列中にACGT以外のものが存在するためエラーが出る例です。4番目の配列(つまりgene_4)の17番目のポジションがNなので妥当です。

```
in_f <- "sample4.fasta" #入力ファイル名を指定してin_fに格納
out_f <- "hoge2.fasta" #出力ファイル名を指定してout_fに格納
```

[トップページへ](#)

Apr 13, 2020

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

入力ファイルの保存1

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_translate_bio...

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存
    
```

2. (multi-)FASTA形式ファイル(sample4.fasta)の場合 :

配列中にACGT以外のものが存在するためエラーが出る例です。4番目の配列(つまりgene_4)の17番目のポジションがNなので妥当です。

```

in_f <- "sample4.fasta"    #入力ファイル名を指定してin_fに格納
out_f <- "hoge2.fasta"    #出力ファイル名を指定してout_fに格納
    
```

[トップページへ](#)

入力ファイルの保存2

①例題1をやります。②デスクトップ上に、(名前は何でもよいですが)hogeというフォルダを作成し、そこに入力ファイルとして用いる③sample1.fastaを保存すべく...

(Rで)塩基配列解析

← ① 保護されていない通信 | [iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_translate_bio...](#)

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"

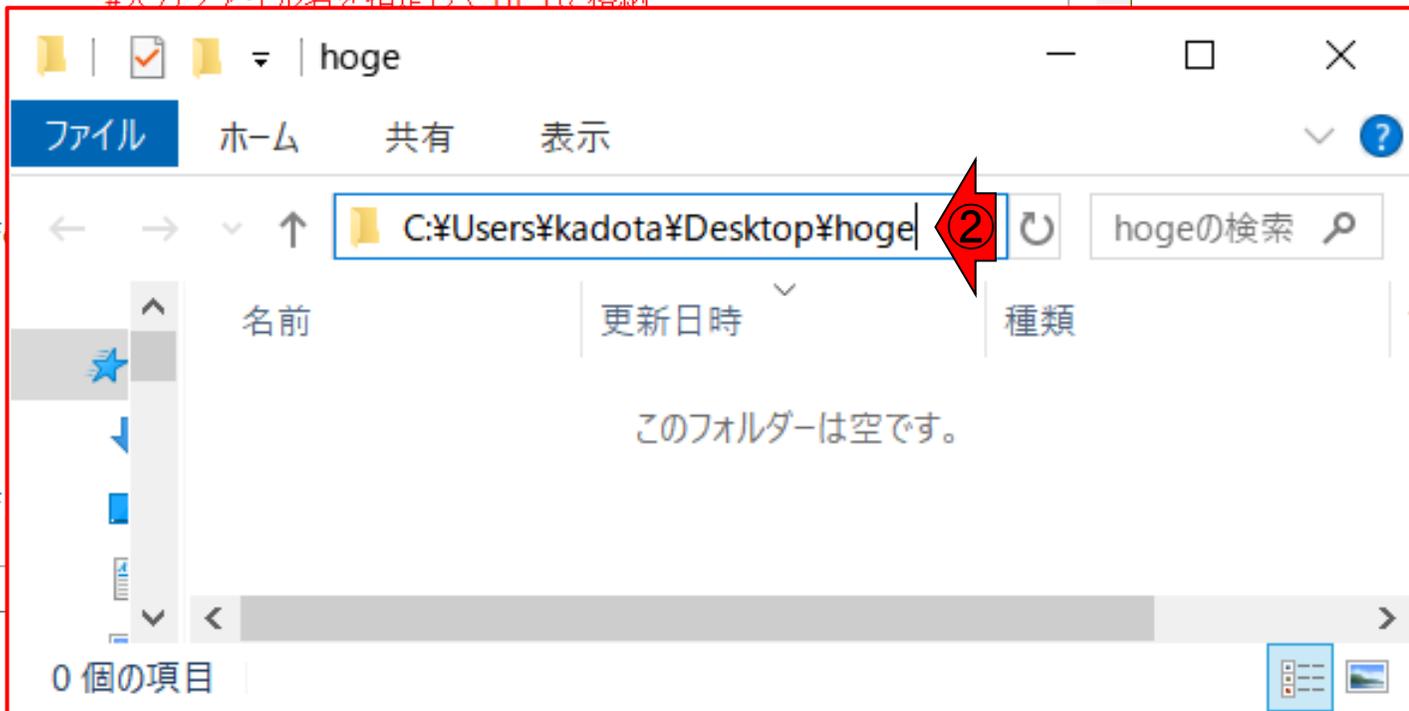
#必要なパッケージをロード
library(Biostrings)

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, f
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



入力ファイルの保存3

①例題1をやります。②デスクトップ上に、(名前は何でもよいですが)hogeというフォルダを作成し、そこに入力ファイルとして用いる③sample1.fastaを保存すべく、右クリックで④保存。

1. FASTA形式ファイル(sample1.fasta)の場合・

```
multi-FASTAではないsingle-FASTAの場合  
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, from="fasta")  
fasta  
  
#本番  
fasta <- translate(fasta)  
fasta  
  
#ファイルに保存  
writeXStringSet(fasta, file=out_f, format="fasta", width=50)
```

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

に格納
(に格納)
ルの読み込み
#fastaの中身を指定したファイル名で保存
トップページへ

入力ファイルの保存4

①例題1をやります。②デスクトップ上に、(名前は何でもよいですが)hogeというフォルダを作成し、そこに入力ファイルとして用いる③sample1.fastaを保存すべく、右クリックで④保存。⑤デスクトップ、⑥hogeを選択して、⑦保存。⑧ときどき勝手に.txtが追加されることがあるので注意してください。sample1.fastaが正解

1. FASTA形式ファイル(sample1.fasta)の場合・

multi-FASTAではないsingle-FAS

```
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(i
fasta
```

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

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



入力ファイルの保存5

①例題1をやります。②デスクトップ上に、(名前は何でもよいですが)hogeというフォルダを作成し、そこに入力ファイルとして用いる③sample1.fastaを保存すべく、右クリックで④保存。⑤デスクトップ、⑥hogeを選択して、⑦保存。⑧ときどき勝手に.txtが追加されることがあるので注意してください。sample1.fastaが正解。こんな感じになれば無事ダウンロードできているはずです。⑨×。

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_g

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(
out_f <- "hoge1.fasta" #出力ファイル名を指定してout_f(

#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存
```

[トップページへ](#)

sample1.fasta

すべて表示

9

入力ファイルの保存6

①例題1をやります。②デスクトップ上に、(名前は何でもよいですが)hogeというフォルダを作成し、そこに入力ファイルとして用いる③sample1.fastaを保存すべく、右クリックで④保存。⑤デスクトップ、⑥hogeを選択して、⑦保存。⑧ときどき勝手に.txtが追加されることがあるので注意してください。sample1.fastaが正解。こんな感じになれば無事ダウンロードできているはず。⑨×。⑩hogeフォルダの中身が、⑪のようになればOK。

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"

#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")

#本番
fasta <- translate(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(

ファイル | ホーム | 共有 | 表示

C:\Users\kadota\Desktop\hoge

名前	更新日時	種類
sample1.fasta	2020/03/04 16:43	FASTA ファイル

1 個の項目

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

作業ディレクトリの変更1

①hogeフォルダ(ディレクトリ)内にある、
②sample1.fastaが目的のファイルだが、
現実には、②のファイル名と同じものが別のフォルダにあるかもしれません。**作業ディレクトリの変更という作業は、RStudio上で作業する場所(ディレクトリ or フォルダ)を宣言すること。これによって、以降の作業を③「ファイル名のみ」を与えることで済ませられるので便利です。**

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"

#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, f
fasta

#本番
fasta <- translate(fasta)
fasta

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



作業ディレクトリの変更2

私の環境では、①の場所を作業ディレクトリに設定することになります。正確には、「C:¥Users¥...」という①の全ての情報(フルパス、といいます)を指定することで、作業ディレクトリがどこであっても②を読み込むことはできます。そういうやり方を教えるヒトもいると思います。

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"

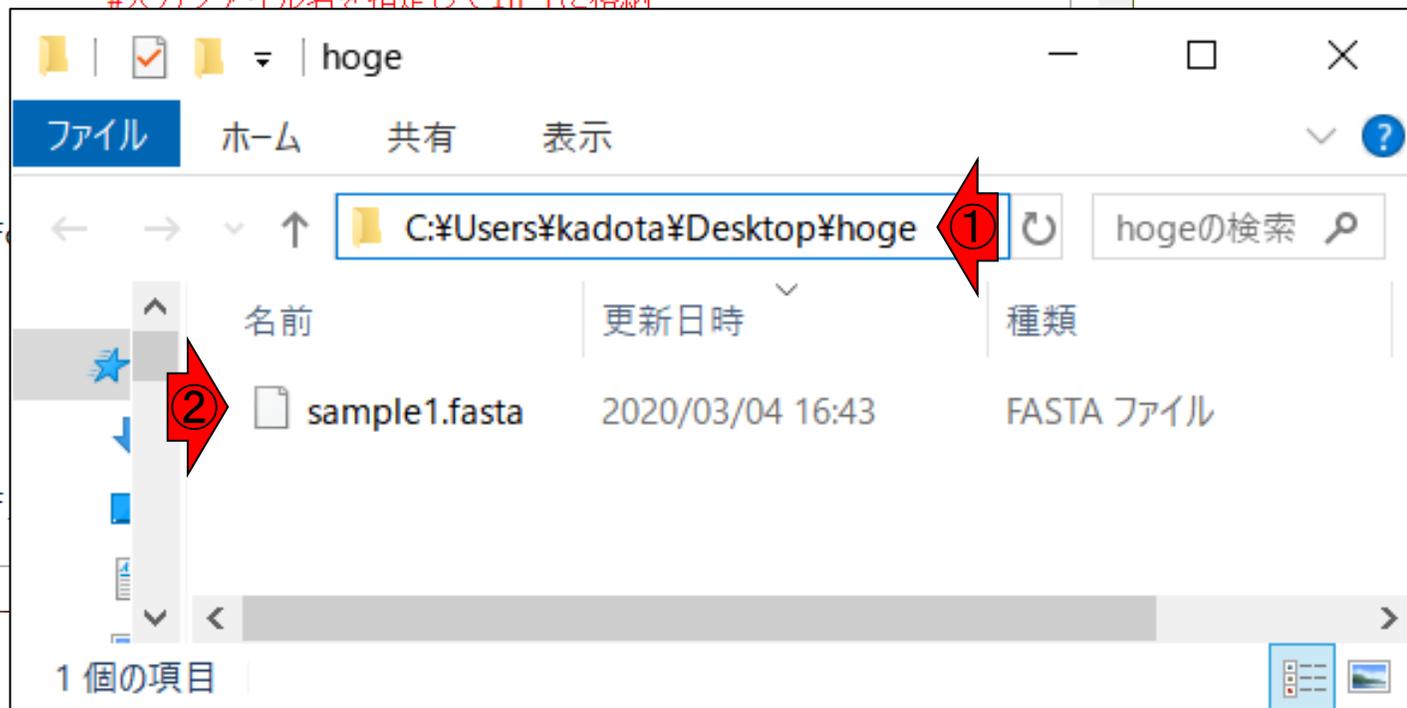
#必要なパッケージをロード
library(Biostrings)

#入力ファイルの読み込み
fasta <- readDNAStringSet(in_f, f
fasta

#本番
fasta <- translate(fasta)
fasta

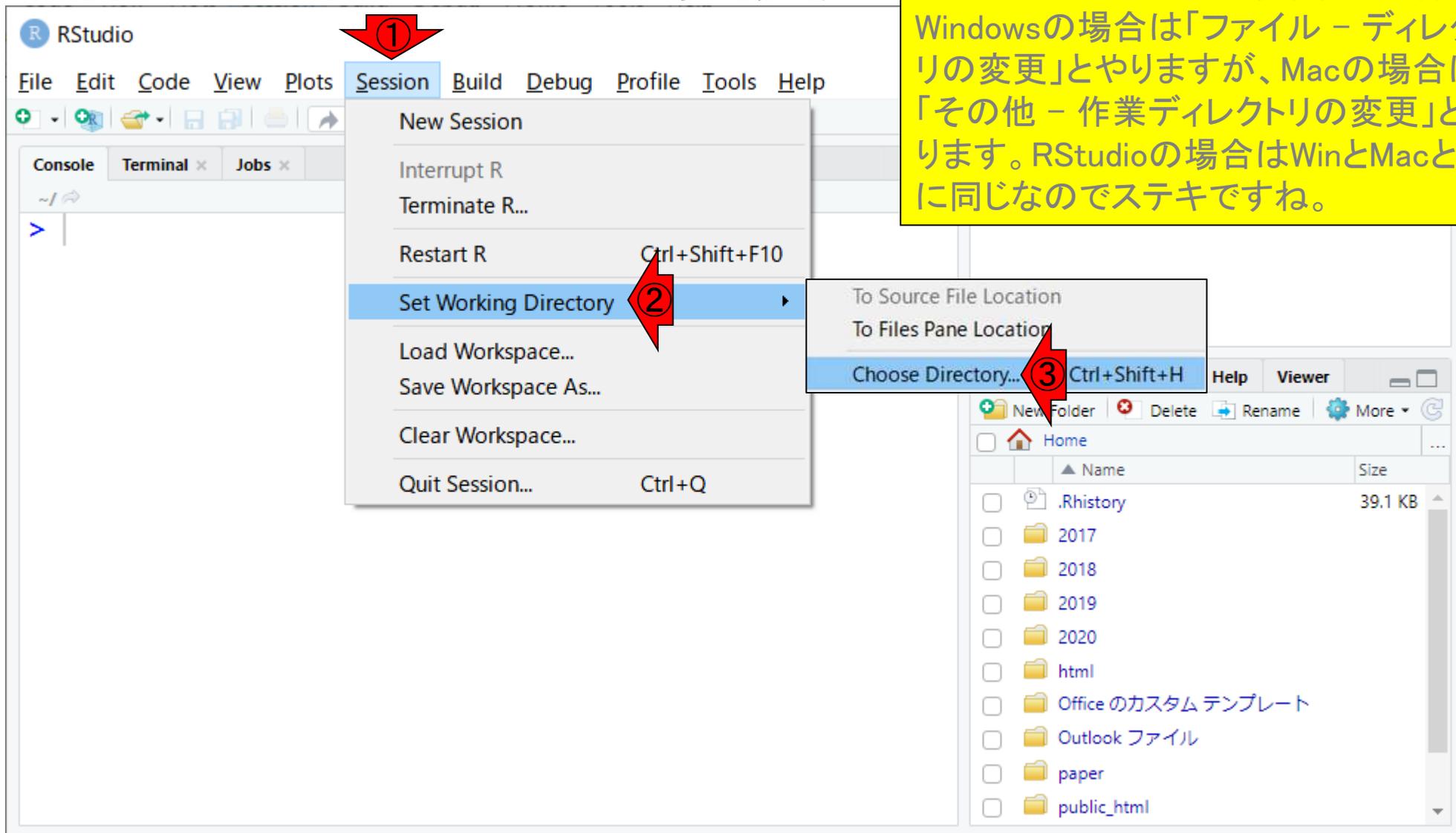
#ファイルに保存
writeXStringSet(fasta, file=out_f
```

#入力ファイル名を指定してin_fに格納



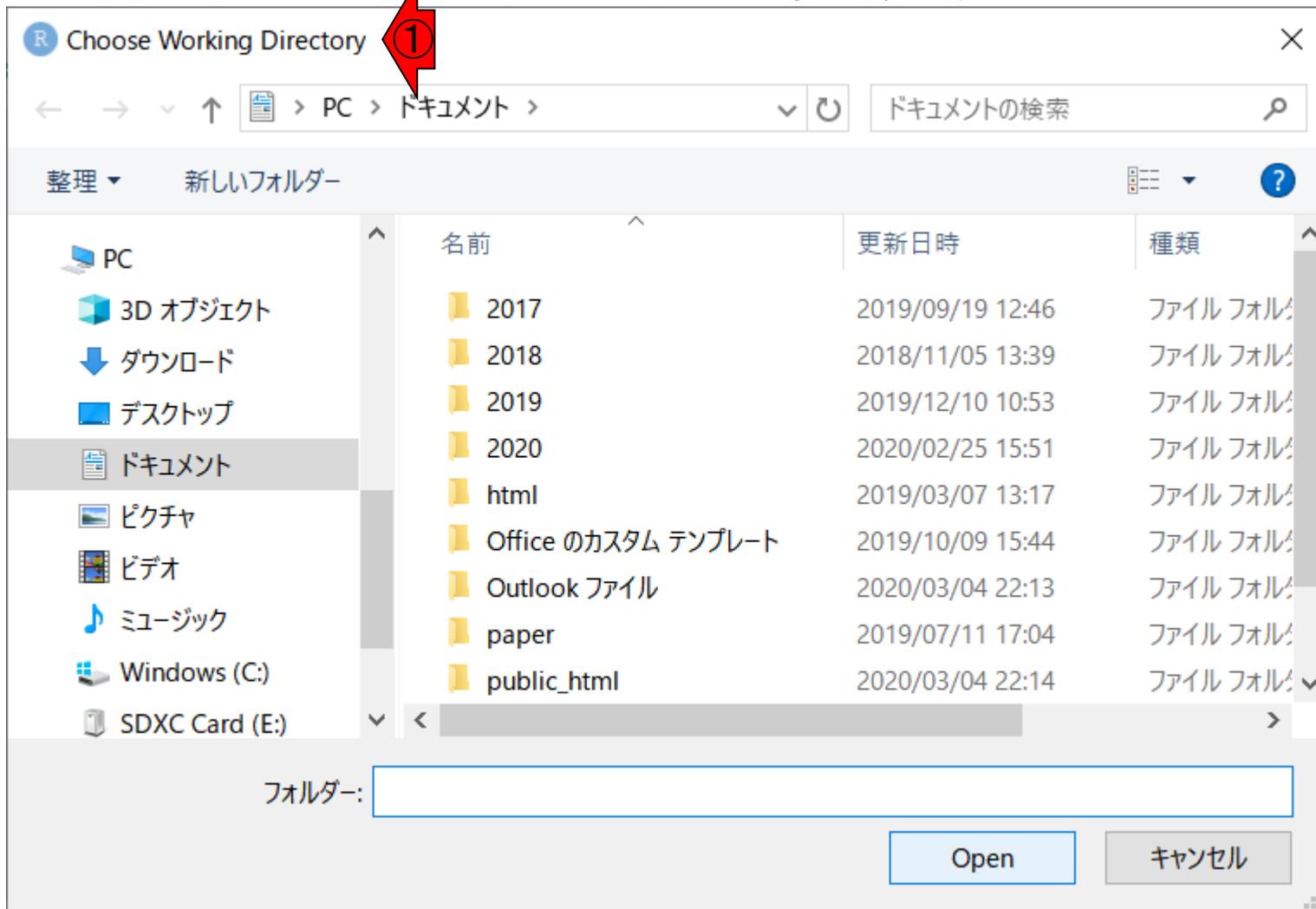
作業ディレクトリの変更3

①Session、②Set Working Directory、③Choose Directory…。RStudioではなくR上で作業ディレクトリの変更を行う際は、Windowsの場合は「ファイル - ディレクトリの変更」とやりますが、Macの場合は「その他 - 作業ディレクトリの変更」とやります。RStudioの場合はWinとMacともに同じなのでステキですね。



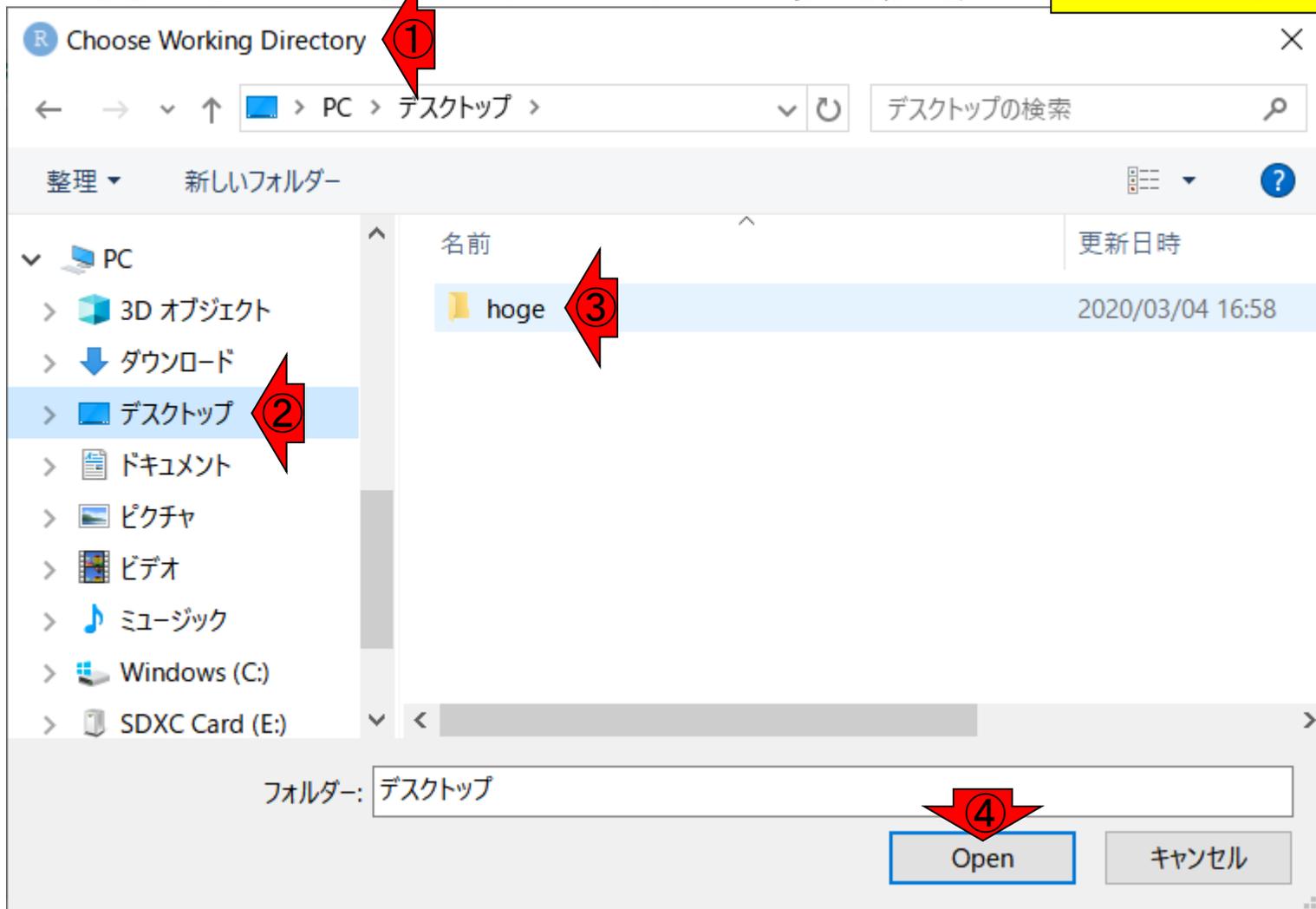
①こんな感じのフォルダ選択画面が出ますので…

作業ディレクトリの変更4



作業ディレクトリの変更5

①こんな感じのフォルダ選択画面が出ますので、②③④のような感じで変更を完了してください。



作業ディレクトリの変更6

The screenshot shows the RStudio interface with the following components:

- Console:** Displays the command `> setwd("C:/Users/kadota/Desktop/hoge")` and a subsequent prompt `> |`.
- Environment:** Shows the command `setwd("C:/Users/kadota/Desk...`.
- Files:** Shows the current directory path `C: > Users > kadota > Desktop > hoge` and a file listing table.

	Name	Size
	..	
	sample1.fasta	23 B

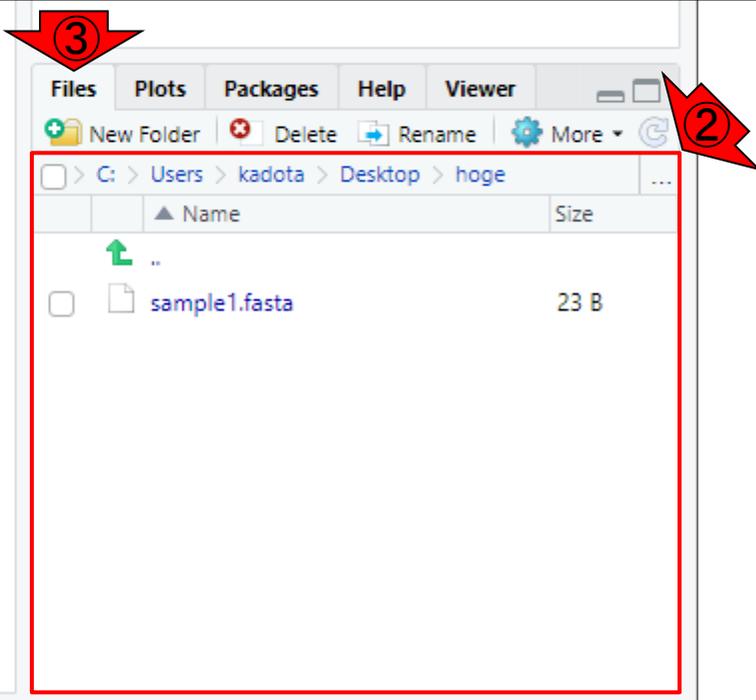
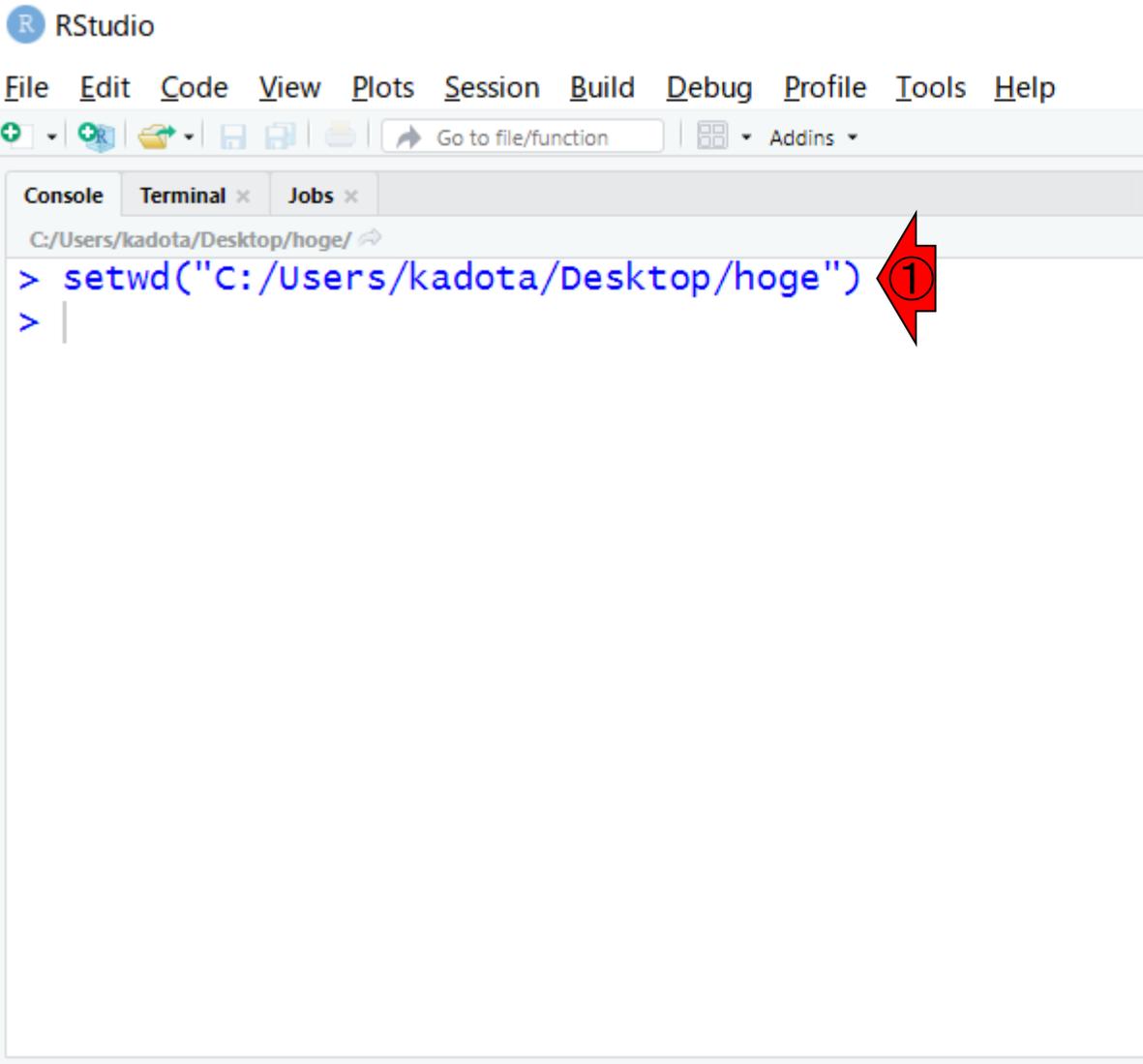
作業ディレクトリの変更7

こんな感じになります。入力したわけでもないのに、①こんなコマンドを実行したことになります。逆にいえば、さきほどの作業は、①のコマンドをコピー実行することでも代用可能だということです。

The screenshot shows the RStudio interface. The console window displays the command `setwd("C:/Users/kadota/Desktop/hoge")` being entered. A red lightning bolt icon with the number 1 is placed over the command. The Environment pane on the right shows the command `setwd("C:/Users/kadota/Desk...`. The Files pane at the bottom right shows the directory `C:/Users/kadota/Desktop/hoge` containing a file named `sample1.fasta` with a size of 23 B.

作業ディレクトリの変更8

こんな感じになります。入力したわけでもないのに、①こんなコマンドを実行したことになっています。逆にいえば、さきほどの作業は、①のコマンドをコピー実行することでも代用可能だということです。次は、②右下の赤枠に注目。自動的に作業ディレクトリに移動し、その中身が見えていることが分かります。もし何らかの拍子で③Filesタブがアクティブになっていない場合は、ここをクリック。



貸与PCのヒトは、①の部分がstudentになっていると思います。

作業ディレクトリの変更9

The screenshot shows the RStudio interface. The console window on the left contains the command `setwd("C:/Users/kadota/Desktop/hoge")`. The file explorer window on the right shows the current directory `C:/Users/kadota/Desktop/hoge` containing a file named `sample1.fasta`. A red arrow with the number 1 points to the 'Packages' tab in the file explorer window.

作業ディレクトリの変更10

貸与PCのヒトは、①の部分がstudentになっていると思います。②赤枠内の任意のフォルダをクリックできますが、移動先のフォルダの中身が見られるだけです。作業ディレクトリの変更はできませんのでご注意ください。

The screenshot shows the RStudio environment. The console on the left contains the command `setwd("C:/Users/kadota/Desktop/hoge")`. On the right, the Environment pane shows the command `setwd("C:/Users/kadota/Desk...`. Below the Environment pane, a file explorer window is open, displaying the path `C: > Users > kadota > Desktop > hoge` in a red box with a red arrow and the number 2 pointing to it. The file explorer shows a folder named `..` and a file named `sample1.fasta` with a size of 23 B.

作業ディレクトリの変更11

貸与PCのヒトは、①の部分がstudentになっていると思います。②赤枠内の任意のフォルダをクリックできますが、移動先のフォルダの中身が見られるだけです。作業ディレクトリの変更はできませんのでご注意ください。作業ディレクトリの確認は、②ではなく③で行います。

The screenshot shows the RStudio interface. The terminal window on the left displays the command `setwd("C:/Users/kadota/Desktop/hoge")` being entered. A red arrow with the number 3 points to the terminal. The file explorer on the right shows the directory path `C: > Users > kadota > Desktop > hoge` highlighted with a red box and a red arrow with the number 2. The file explorer also shows a file named `sample1.fasta` with a size of 23 B.

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

入力ファイルの確認1

①hogeフォルダ内の、②sample1.fastaをクリックしてみましょう。

The screenshot shows the RStudio interface. The console window on the left contains the command `setwd("C:/Users/kadota/Desktop/hoge")` and a cursor. The Environment pane on the right shows the command `setwd("C:/Users/kadota/Desk...`. The Files pane at the bottom right shows the directory `C:/Users/kadota/Desktop/hoge` with a file named `sample1.fasta` (23 B). Red arrows with circled numbers 1 and 2 point to the directory path and the file name, respectively.

入力ファイルの確認2

①hogeフォルダ内の、②sample1.fastaをクリックしてみましょう。こんな感じになります。

The screenshot shows the RStudio interface. The main editor window displays the contents of 'sample1.fasta':

```
1 |>kadota
2 | AGTGACGGTCTT
3 |
```

The console window shows the command `setwd("C:/Users/kadota/Desktop/hoge")` being executed. The file explorer window on the right shows the directory `C:\Users\kadota\Desktop\hoge` with a file named `sample1.fasta` (23 B) highlighted by a red arrow labeled ②. Another red arrow labeled ① points to the 'hoge' folder in the path.

入力ファイルの確認3

①hogeフォルダ内の、②sample1.fastaをクリックしてみましょう。こんな感じになります。赤枠が、③sample1.fastaの中身に相当します。

The screenshot shows the RStudio interface. At the top, the menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main editor window displays a file named 'sample1.fasta' with the following content:

```
1 |>kadota
2 | AGTGACGGTCTT
3 |
```

A red box highlights the content of the file. The console window at the bottom shows the command `setwd("C:/Users/kadota/Desktop/hoge")` being executed. The file explorer on the right shows the directory structure `C:\Users\kadota\Desktop\hoge` with a file named `sample1.fasta` (23 B) highlighted. Red arrows with circled numbers 1, 2, and 3 point to the file explorer, the file name, and the editor content respectively.

入力ファイルの確認4

①hogeフォルダ内の、②sample1.fastaをクリックしてみましょう。こんな感じになります。赤枠が、③sample1.fastaの中身に相当します。④行番号情報です。1行が長いデータの場合に、④を眺めることで行の区切りを把握します。

The screenshot shows the RStudio interface with several components:

- File Explorer (Bottom Right):** Shows the file system path `C:\Users\kadota\Desktop\hoge`. The file `sample1.fasta` is selected and highlighted with a red arrow labeled ②.
- Environment Panel (Top Right):** Shows the current working directory set to `setwd("C:/Users/kadota/Desktop/hoge")`.
- Source Editor (Middle):** Displays the content of `sample1.fasta` with a red box around it labeled ③. The content is:

```
1 >kadota
2 AGTGACGGTCTT
3
```

Red arrows labeled ④ point to the line numbers 1, 2, and 3.
- Terminal (Bottom Left):** Shows the command `> setwd("C:/Users/kadota/Desktop/hoge")` being executed.
- File Panel (Top Left):** Shows the file `sample1.fasta` is open, with a red arrow labeled ③ pointing to the file icon.

入力ファイルの確認5

The screenshot shows the RStudio interface. The top-left pane displays a FASTA file named 'sample1.fasta' with the following content:

```
1 >kadota
2 AGTGACGGTCTT
3
```

A red arrow with the number 1 points to the cursor position at the end of the first line. The bottom-left pane shows the console with the command `setwd("C:/Users/kadota/Desktop/hoge")` entered. The bottom-right pane shows a file explorer view of the directory `C:/Users/kadota/Desktop/hoge`, listing the file `sample1.fasta` with a size of 23 B.

今はカーソルが①の位置にあるので、②の位置に変更しただけ。

入力ファイルの確認6

The screenshot shows the RStudio interface. The top-left pane displays a text editor with a file named 'sample1.fasta'. The content is as follows:

```
1 >kadota
2 AGTGACGGTCTT|
3
```

Red arrows with circled numbers 1 and 2 point to the first and second lines of the FASTA header, respectively.

The bottom-left pane shows the terminal window with the following command and prompt:

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

The bottom-right pane shows the file explorer view for the directory 'C:/Users/kadota/Desktop/hoge'. It lists the following files:

Name	Size
..	
sample1.fasta	23 B

入力ファイルの確認7

今はカーソルが①の位置にあるので、②の位置に変更しただけ。③の拡大図が示すように、①のカーソル位置「1:1」から②のカーソル位置「2:13」に変更されていることも分かります。そこそこのエディタには普通についている機能ではありませんが、ゲノム情報を解析する場合は、例えば②の位置が何番目の塩基に相当するのかを知りたい場合があるので便利。

```
1 >kadota
2 AGTGACGGTCTT|
3
```

2:13

Text File

Console Terminal x Jobs x

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

Files Plots Packages Help Viewer

New Folder Delete Rename More

C: > Users > kadota > Desktop > hoge

Name	Size
..	
sample1.fasta	23 B

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

①赤枠で見えているのが「FASTA形式」です。

FASTA形式1

The screenshot shows the RStudio interface with a text editor window open to a file named 'sample1.fasta'. The file content is as follows:

```
1 >kadota
2 AGTGACGGTCTT
3
```

The first two lines are enclosed in a red box, and a red arrow labeled '1' points to this box. The terminal window at the bottom shows the command:

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

The file explorer on the right shows the current directory is 'C:/Users/kadota/Desktop/hoge' and contains a file named 'sample1.fasta' with a size of 23 B.

FASTA形式2

①赤枠で見えているのが「FASTA形式」です。最初に②「大なり記号(>)」、その後③任意の記述内容(この場合はkadota)からなる1行の「description行」があり、その次の行に④塩基 or アミノ酸配列がある(複数行になってもよい)ような形式です。

The screenshot shows the RStudio interface. The main editor window displays a file named 'sample1.fasta' with the following content:

```
1 >kadota
2 AGTGACGGTCTT
3
```

Red arrows with circled numbers point to specific parts of the FASTA format: arrow 2 points to the greater-than sign (>), arrow 3 points to the sequence identifier 'kadota', and arrow 4 points to the sequence 'AGTGACGGTCTT'. The terminal window at the bottom shows the command `setwd("C:/Users/kadota/Desktop/hoge")` being executed. The file explorer on the right shows the current directory containing the file 'sample1.fasta' with a size of 23 B.

FASTA形式3

①赤枠で見えているのが「FASTA形式」です。最初に②「大なり記号(>)」、その後③任意の記述内容(この場合はkadota)からなる1行の「description行」があり、その次の行に④塩基 or アミノ酸配列がある(複数行になってもよい)ような形式です。⑤の部分は任意に変更可能なので、例えばdescription1にしてみると、⑥何か変更したことを意味するアスタリスク(*)が付加されます。

The screenshot shows the RStudio interface with a text editor open to 'sample1.fasta'. The editor contains the following text:

```
1 >description1
2 AGTGACGGTCTT
3
```

Red arrows point to the description line (labeled 5) and the asterisk in the filename (labeled 6). Below the editor is a terminal window with the command:

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

To the right, a file explorer window shows the directory 'C:/Users/kadota/Desktop/hoge/' containing the file 'sample1.fasta' with a size of 23 B.

FASTA形式4

①赤枠で見えているのが「FASTA形式」です。最初に②「大なり記号(>)」、その後③任意の記述内容(この場合はkadota)からなる1行の「description行」があり、その次の行に④塩基 or アミノ酸配列がある(複数行になってもよい)ような形式です。⑤の部分は任意に変更可能なので、例えばdescription1にしてみると、⑥何か変更したことを意味するアスタリスク(*)が付加されます。⑦を押して保存すると…

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main editor window displays a file named 'sample1.fasta*' with the following content:

```
1 >description1
2 AGTGACGGTCTT
3
```

A red arrow labeled '7' points to the end of the first line. The status bar at the bottom of the editor shows '1:14' and 'Text File'. Below the editor is a console window with the following text:

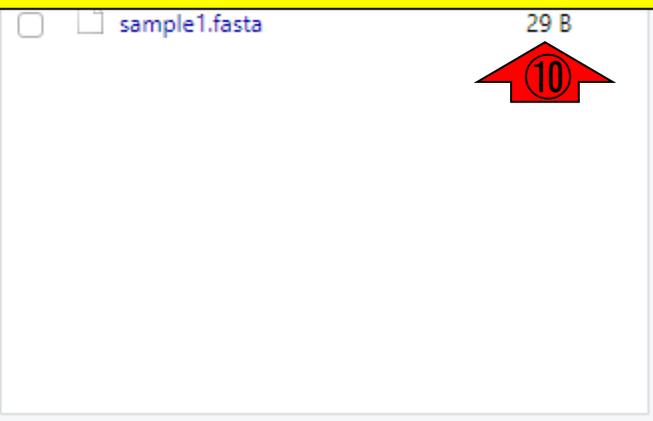
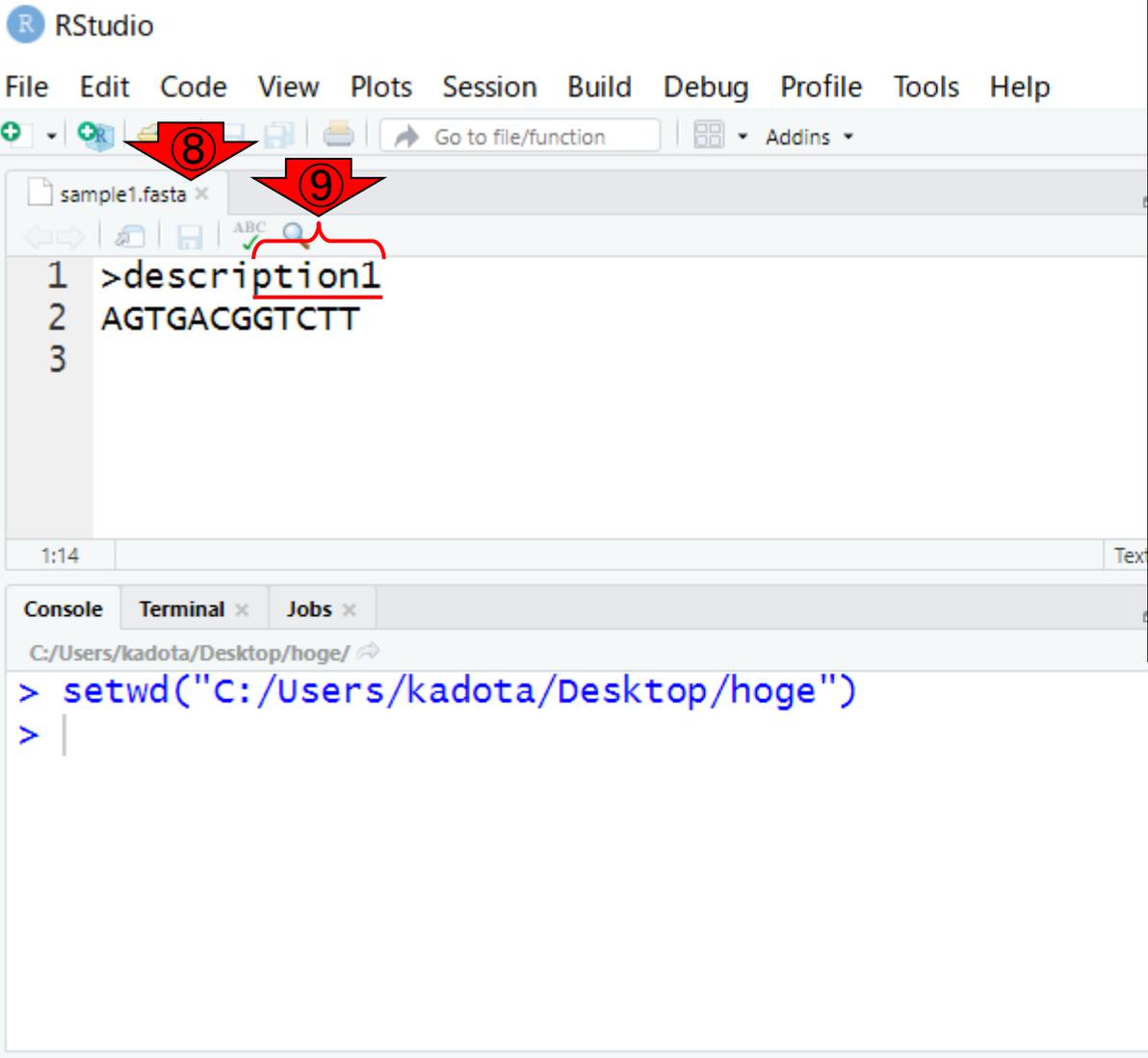
```
> setwd("C:/Users/kadota/Desktop/hoge")
> |
```

The screenshot shows a Windows File Explorer window. The address bar displays the path 'C: > Users > kadota > Desktop > hoge'. The main area shows a list of files and folders:

Name	Size
..	
sample1.fasta	23 B

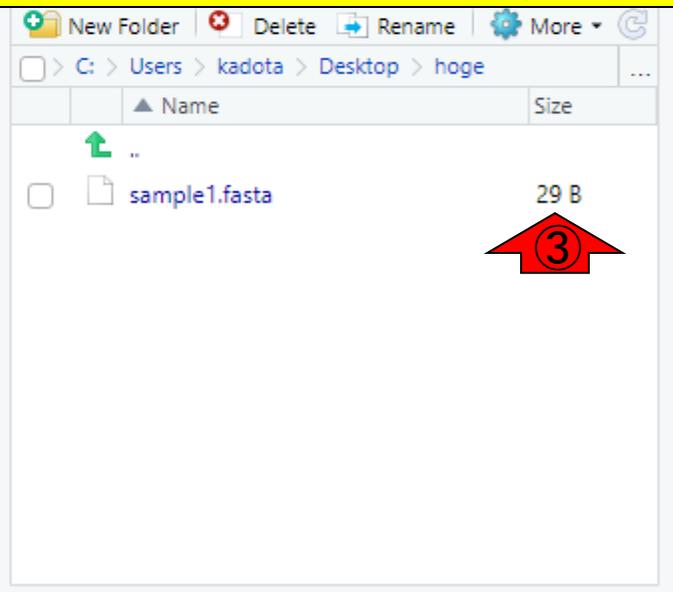
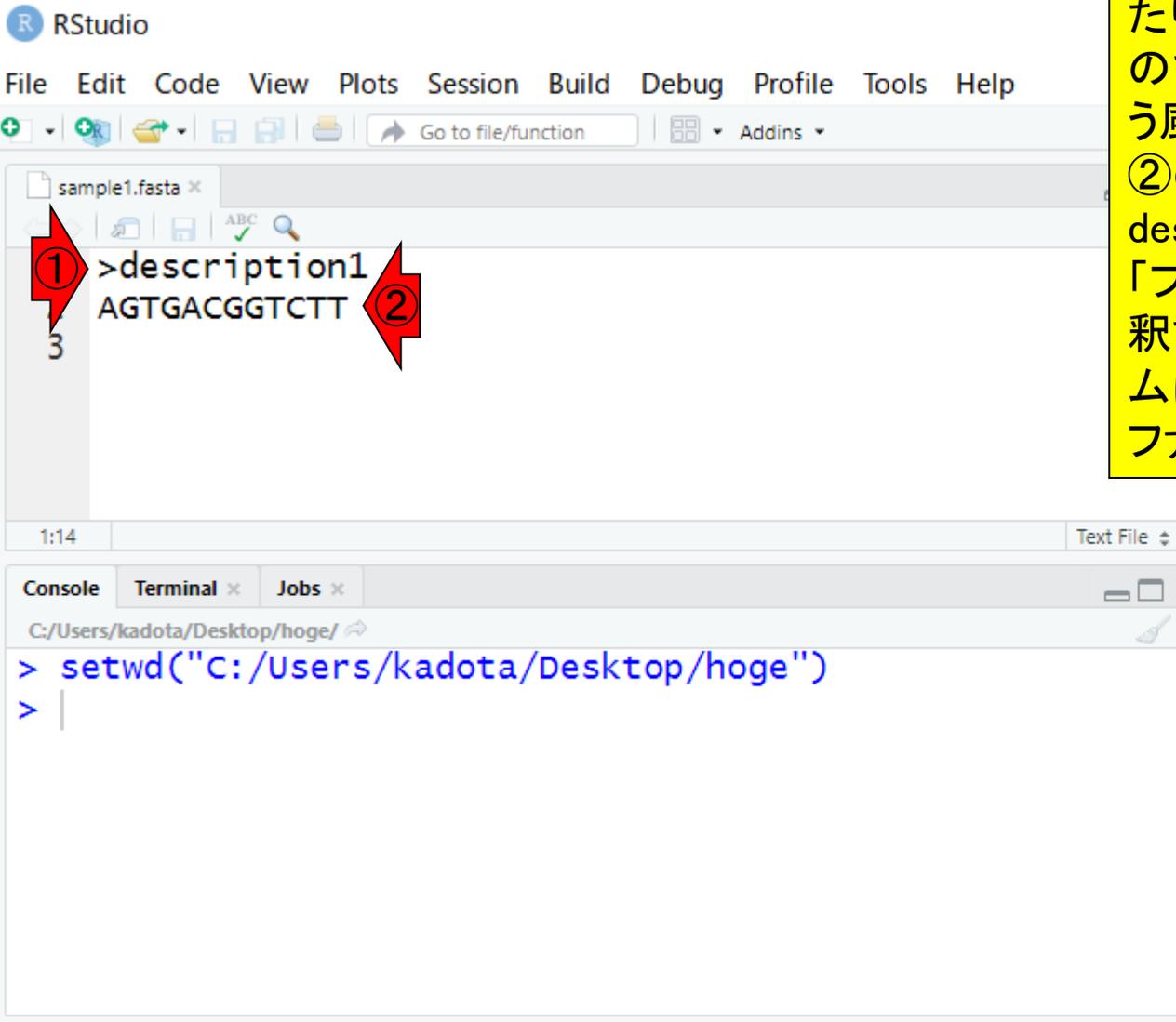
FASTA形式5

①赤枠で見えているのが「FASTA形式」です。最初に②「大なり記号(>)」、その後③任意の記述内容(この場合はkadota)からなる1行の「description行」があり、その次の行に④塩基 or アミノ酸配列がある(複数行になってもよい)ような形式です。⑤の部分は任意に変更可能なので、例えばdescription1にしてみると、⑥何か変更したことを意味するアスタリスク(*)が付加されます。⑦を押して保存すると、⑧*が消え、⑨kadotaの6文字からdescription1の12文字になったので増えた6文字分だけ、⑩ファイルサイズが大きくなっている(23 → 29 bytes)ことが分かります。



FASTA形式6

①description行の文字数は13、②塩基配列の文字数は12。1文字あたり1 byteと計算します。Windowsの場合は1行あたりの改行コードに2 bytes分使いますので、③ $(13 + 12 + 2 + 2) = 29$ bytesという風になります。ゲノム情報の場合は、②の塩基配列情報のほうが圧倒的に①description行の情報よりも多いので、③「ファイルサイズ \approx ゲノムサイズ」と解釈する場合があります。実際、ヒトゲノムは30億塩基対と言われますが、そのファイルサイズは約3GBです。



赤枠がRのエディタ画面。①×を押して編集を終了させると…

FASTA形式7

The screenshot shows the RStudio interface. The main editor window displays a FASTA file named 'sample1.fasta' with the following content:

```
1 >description1
2 AGTGACGGTCTT
3
```

A red arrow with the number '1' points to the close button (X) in the top right corner of the editor window. The terminal window at the bottom shows the command:

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

The file browser on the right shows the current directory 'C:/Users/kadota/Desktop/hoge' containing the file 'sample1.fasta' (29 B).

FASTA形式8

赤枠がRのエディタ画面。①×を押して編集を終了させると、こんな感じになります。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main workspace is divided into several panes:

- Console:** Shows the command `> setwd("C:/Users/kadota/Desktop/hoge")` and a prompt `> |`.
- Environment:** Shows the command `setwd("C:/Users/kadota/Desk...`.
- Files:** Shows the directory `C:/Users/kadota/Desktop/hoge` with a file `sample1.fasta` of size 29 B.

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

コピペ実行1

さきほどdescription行部分を変更した① sample1.fastaを入力として、その翻訳配列を②hoge1.fastaというファイル名で出力する一連のスク립トをコピペ実行します。

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_translate_bio...

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
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存
```

[トップページへ](#)

コピペ実行2

さきほど description 行部分を変更した① sample1.fasta を入力として、その翻訳配列を② hoge1.fasta というファイル名で出力する一連のスク립トをコピペ実行します。赤枠内を全選択して(反転させて)...

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_translate_bio...

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存
```

[トップページへ](#)

コピペ実行3

さきほど description 行部分を変更した① sample1.fasta を入力として、その翻訳配列を② hoge1.fasta というファイル名で出力する一連のスク립トをコピペ実行します。赤枠内を全選択して(反転させて)、
②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。

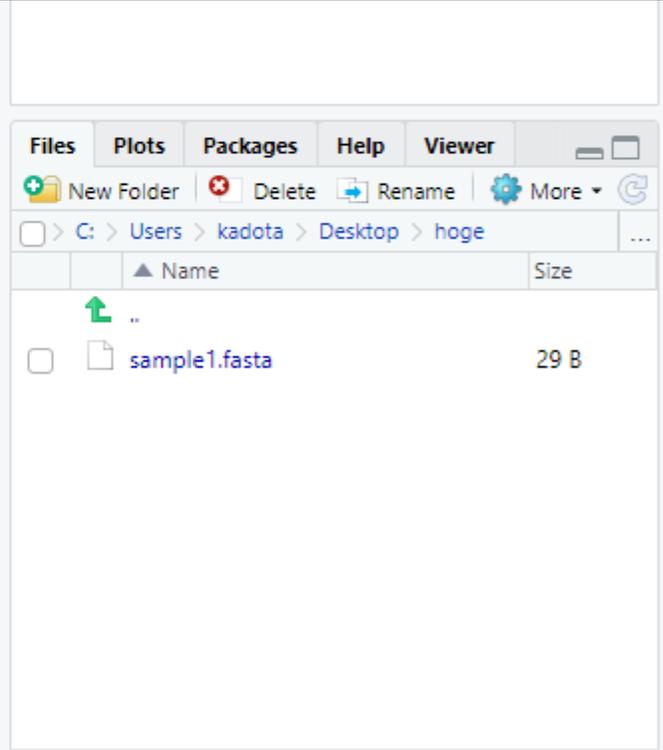
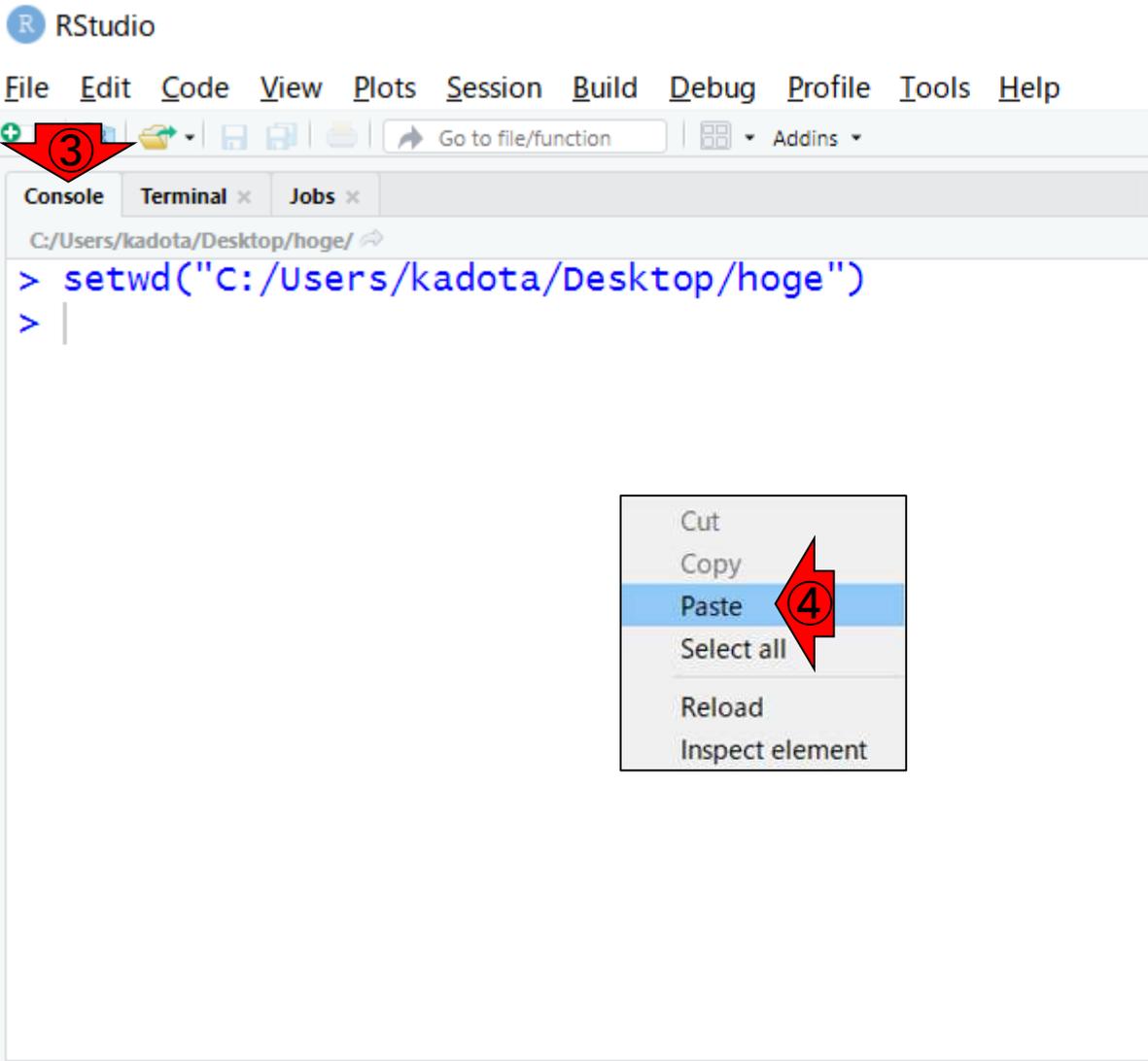
```
in f <- "sample1.fasta" #入力ファイル名を指定してin fに格納
out_f <- "hoge1.fasta"
#必要なパッケージをロード
library(Biostrings)
#入力ファイルの読み込み
fasta <- readDNAStrings
fasta
#本番
fasta <- translate(fasta) #アミノ酸配列に翻訳した結果をfastaに格納
fasta #確認してるだけです
#ファイルに保存
writeXStringSet(fasta, file=out_f, format="fasta", width=50)#fastaの中身を指定したファイル名で保存
```

コピー(C)	Ctrl+C
Google で「in_f <- "sample1.fasta" #入力ファイル名を...」を検索(S)	
印刷(P)...	Ctrl+P
検証(I)	Ctrl+Shift+I

[トップページへ](#)

コピペ実行4

さきほど description 行部分を変更した① sample1.fasta を入力として、その翻訳配列を② hoge1.fasta というファイル名で出力する一連のスクリプトをコピペ実行します。赤枠内を全選択して(反転させて)、②コピー、③Consoleがアクティブな状態であることを確認して、④右クリックでペースト。(うまくいかないヒトは「Ctrl + V」もトライしてください。)

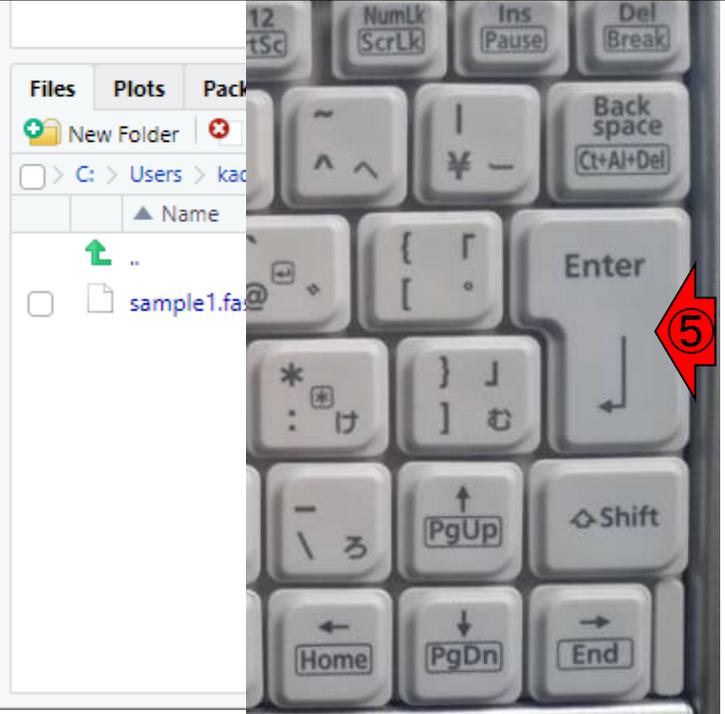


コピペ実行5

さきほどdescription行部分を変更した① sample1.fastaを入力として、その翻訳配列を②hoge1.fastaというファイル名で出力する一連のスクリプトをコピペ実行します。赤枠内を全選択して(反転させて)、②コピー、③Consoleがアクティブな状態であることを確認して、④右クリックでペースト。(うまくいかないヒトは「Ctrl + V」もトライしてください。)こんな感じになればOK。⑤リターンキーを押す。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Console Terminal x Jobs x
C:/Users/kadota/Desktop/hoge/
library(Biostrings)
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
#本番
fasta <- translate(fasta)                #アミノ酸配列
#確認してるだけ
fasta
#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存
```



コピペ実行6

実行完了後の状態。①赤枠内がこんな感じになればOK。出力ファイルである、②hoge1.fastaが作成されていることも分かります。もし①の画面になっていて、②が見られない場合は、③リロードしてみてください。

The screenshot shows the RStudio interface with the following components:

- Console:** Contains R code and its output. A red box highlights the output of the `translate` function, which shows a `DNAStringSet` instance with a width of 12 and a sequence of `AGTGACGGTCTT`. Below this, there are comments in Japanese and the code for writing the result to a file named `hoge1.fasta`.
- Environment:** Shows the current environment with variables `fasta` and `writeXStringSet`.
- Files:** A file explorer window is open, showing the directory `C:\Users\kadota\Desktop\hoge`. It contains two files: `sample1.fasta` (29 B) and `hoge1.fasta` (21 B). Red arrows point to the `hoge1.fasta` file and the file explorer window.

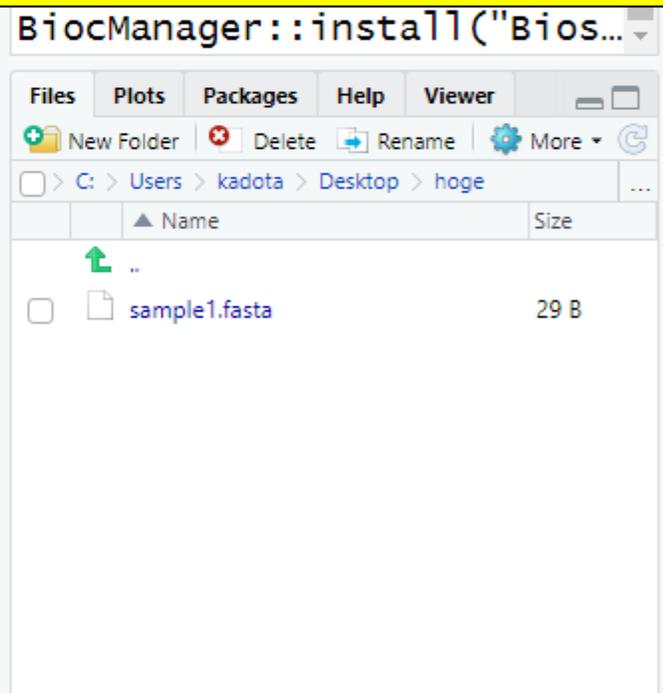
```
> #アミノ酸配列に
> fasta <- translate(fasta) #アミノ酸配列に
#翻訳した結果をfastaに格納
> fasta #確認してるだけ
#確認してるだけ
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", width=50)#fastaの中身を指定したファイル名で保存
>
>
```

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B

コピペ実行7

実行完了後の状態。①赤枠内がこんな感じになればOK。出力ファイルである、②hoge1.fastaが作成されていることも分かります。もし①の画面になっていて、②が見られない場合は、③リロードしてみてください。④のような感じで**エラー**が出たヒトは、**前提条件を満たしていない**(この場合はBiostringsパッケージのインストールができてない)ことを意味します。しかるべき作業を行ってください。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Console Terminal x Jobs x
C:/Users/kadota/Desktop/hoge/
> setwd("C:/Users/kadota/Desktop/hoge")
> in_f <- "sample1.fasta" #入力ファイル
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
library(Biostrings) でエラー:
'Biostrings' という名前のパッケージはありません
>
> #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
readDNAStringSet(in_f, format = "fasta") でエラー:
関数 "readDNAStringSet" を見つけることができませんでした
> fasta #確認してるだけ
です
```



Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

出力ファイルの確認1

①赤枠内の、②の③DNA塩基配列情報、および④の⑤アミノ酸(Amino Acids)配列情報からもなんとなく分かりますが…

RStudio interface showing R code execution and file explorer. The console output is as follows:

```
> #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
#確認してるだけ
A AAStringSet instance of length 1
width seq names
[1] 4 SDGL description1
>
> #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存
>
>
```

The file explorer shows the following files:

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B

出力ファイルの確認2

①赤枠内の、②の③DNA塩基配列情報、および④の⑤アミノ酸(Amino Acids)配列情報からもなんとなく分かりますが、出力ファイルである⑥hoge1.fastaをクリックして見られる情報は…

```
fasta <- translate(fasta) ...
fasta #確認してるだけです
#ファイルに保存
writeXStringSet(fasta, fil...

A DNASTringSet instance of length 1
width seq          names
[1] 12 AGTGACGGTCTT description1
>
> #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
A AAStringSet instance of length 1
width seq          names
[1] 4 SDGL description1
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", w
idth=50)#fastaの中身を指定したファイル名で保存
>
>
```

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B

出力ファイルの確認3

①赤枠内の、②の③DNA塩基配列情報、および④の⑤アミノ酸(Amino Acids)配列情報からもなんとなく分かりますが、出力ファイルである⑥hoge1.fastaをクリックして見られる情報は、⑦こんな感じ。

The screenshot shows the RStudio interface. The top-left pane displays the contents of 'hoge1.fasta' with a red border around it. The top-right pane shows the R console with the following code and comments:

```
tasta <- translate(tasta) ...
fasta #確認してるだけです
#ファイルに保存
writeXStringSet(fasta, fil...
```

The bottom-left pane shows the R console output:

```
A AASCTINGSet instance of length 1
width seq          names
[1]      4 SDGL     description1
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", w
idth=50)#fastaの中身を指定したファイル名で保存
>
> |
```

The bottom-right pane shows a file explorer view of the 'hoge' directory on the desktop. It contains two files: 'sample1.fasta' (29 B) and 'hoge1.fasta' (21 B). A red arrow points to 'hoge1.fasta' with the number 6.

出力ファイルの確認4

①赤枠内の、②の③DNA塩基配列情報、および④の⑤アミノ酸(Amino Acids)配列情報からもなんとなく分かりますが、出力ファイルである⑥hoge1.fastaをクリックして見られる情報は、⑦こんな感じ。⑧得られるアミノ酸配列SDGLは、⑨で見えているものと同じですね。

The screenshot shows the RStudio interface with three main components: a script editor, a console, and a file explorer.

- Script Editor (hoge1.fasta):** Contains the following code:

```
1 |>description1
2 |SDGL
3 |
```

Red arrow ⑧ points to the amino acid sequence 'SDGL' on line 2.
- Console:** Shows the execution of the following R code:

```
A AASCTingSet instance of length 1
width seq          names
[1]      4 SDGL     description1
>
> #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存
>
> |
```

Red arrow ⑨ points to the output 'SDGL' in the console.
- File Explorer:** Shows the directory 'C:\Users\kadota\Desktop\hoge' containing two files:

Name	Size
sample1.fasta	29 B
hoge1.fasta	21 B

Red arrow ⑥ points to the 'hoge1.fasta' file.

出力ファイルの確認5

①赤枠内の、②の③DNA塩基配列情報、および④の⑤アミノ酸(Amino Acids)配列情報からもなんとなく分かりますが、出力ファイルである⑥hoge1.fastaをクリックして見られる情報は、⑦こんな感じ。⑧得られるアミノ酸配列SDGLは、⑨で見えているものと同じですね。⑩sample1.fastaもクリックして、⑪エディタ上に表示。⑫hoge1.fastaのタブと行き来させれば入出力の関係が理解しやすい。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main editor window displays two tabs: 'hoge1.fasta' and 'sample1.fasta'. The 'hoge1.fasta' tab is active, showing the following content:

```
1 >description1
2 AGTGACGGTCTT
3 |
```

The console window at the bottom shows the following R code and output:

```
>
> #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存
>
> |
```

The console output shows:

```
A AASCTINGSet instance of length 1
  width seq          names
[1]     4 SDGL        description1
```

The screenshot shows a file explorer window with the following content:

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B

A red arrow labeled ⑩ points to the 'sample1.fasta' file.

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

全体像の把握1

アノテーションファイル(annotation.txt)中の第1列目に対して、リストファイル(genelist1.txt)中の文字列と一致する行を抜き出して、hoge1.txtというファイル名で出力したい。

入力:アノテーションファイル(annotation.txt)

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene2	hoge02	hohinu	membrane
4	gene3	hoge03	agribio	endoplasmic
5	gene4	hoge04	genesis	endoplasmic
6	gene5	hoge05	kamo	membrane
7	gene6	hoge06	netteba	humei
8	gene7	hoge07	tebasaki	nuclear
9	gene8	hoge08	biiru	nuclear
10	gene9	hoge09	nihonshu	nuclear
11	gene10	hoge10	agene1	membrane
12	gene11	hoge11	iyaaaa	endoplasmic

出力:hoge1.txt



	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene7	hoge07	tebasaki	nuclear
4	gene9	hoge09	nihonshu	nuclear

入力:リストファイル(genelist1.txt)

	A
1	gene1
2	gene7
3	gene9

全体像の把握2

アノテーションファイル(annotation.txt)中の第1列目に対して、リストファイル(genelist1.txt)中の文字列と一致する行を抜き出して、hoge1.txtというファイル名で出力したい。①を押して一番上に移動。

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_translate_bio...

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存
```

[トップページ](#)

全体像の把握3

アノテーションファイル(annotation.txt)中の第1列目に対して、リストファイル(genelist1.txt)中の文字列と一致する行を抜き出して、hoge1.txtというファイル名で出力したい。①を押して一番上に移動。ここから②下矢印キーを押して…

(Rで)塩基配列解析

(last modified 2020/03/03, 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](#) ([Windows2019.10.09版](#)と[Macintosh2020.03.02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されています。初心者の方は[基本的な利用法](#) ([Windows2019.03.12版](#)と[Macintosh2019.03.12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](#))

- 2020年の[アグリバイオインフォマティクス教育研究プログラム](#)の受講ガイダンスは、2020年4月3日(Fri.)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。東京大学の学生以外で受講を希望される方は[申請要項](#) ([Word](#)、[PDF](#))をご確認の上、2020年3月19日(木)までにお申込みください。合格者には2020年[トップ](#)までにご案内いたします。昨年度から引き続き受講される場合も、東京大学の学生以外の方は申請が必要ですので



全体像の把握4

アノテーションファイル(annotation.txt)中の第1列目に対して、リストファイル(genelist.txt)中の文字列と一致する行を抜き出して、hoge1.txtというファイル名で出力したい。①を押して一番上に移動。ここから②下矢印キーを押して、③さきほど行った項目の上部にある、④をクリック。

(Rで)塩基配列解析

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

- [インストール | Rパッケージ | 個別\(2018年11月以前\)](#) (last modified 2018/11)
- [基本的な利用法](#) (last modified 2019/03/12)
- [サンプルデータ](#) (last modified 2019/09/06)
- [イントロ | 一般 | ランダムに行を抽出](#) (last modified 2014/07/17)
- [イントロ | 一般 | 任意の文字列を行の最初に挿入](#) (last modified 2014/07/17)
- [イントロ | 一般 | 任意のキーワードを含む行を抽出\(基礎\)](#) ④ (last modified 2019/04/24)
- [イントロ | 一般 | ランダムな塩基配列を生成](#) (last modified 2014/06/16)
- [イントロ | 一般 | 任意の長さの可能な全ての塩基配列を作成](#) (last modified 2015/02/19)
- [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 | について](#) (last modified 2019/09/06)
- [イントロ | 一般 | 指定した範囲の配列を取得 | Biostring](#) (last modified 2019/09/06)
- [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 | について](#) (last modified 2019/09/06)
- [イントロ | 一般 | 翻訳配列\(translate\)を取得 | Biostrings](#) ③ (last modified 2015/09/12)
- [イントロ | 一般 | 翻訳配列\(translate\)を取得 | seqinr\(Charlet 2005\)](#) (last modified 2015/03/09) [トップページへ](#)
- [イントロ | 一般 | 相補鎖\(complement\)を取得](#) (last modified 2019/03/10)

こんな感じになります。①例題1が一番上になるように、5行分ほど下部に移動。

全体像の把握5

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_grep_kiso

イントロ | 一般 | 任意のキーワードを含む行を抽出(基礎)

例えばタブ区切りテキストファイルが手元があり、この中からリストファイル中の文字列を含む行を抽出するやり方を示します。Linux (UNIX)のgrepコマンドのようなものであり、perlのハッシュのようなものです。

① 「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. 目的のタブ区切りテキストファイル([annotation.txt](#))中の第1列目をキーとして、リストファイル([genelist1.txt](#))中のものが含まれる行全体を出力したい場合：

```

in_f1 <- "annotation.txt" #入力ファイル名を指定してin_f1に格納(アノテーションファイル)
in_f2 <- "genelist1.txt" #入力ファイル名を指定してin_f2に格納(リストファイル)
out_f <- "hoge1.txt" #出力ファイル名を指定してout_fに格納
param <- 1 #アノテーションファイル中の検索したい列番号を指定

#入力ファイルの読み込み
data <- read.table(in_f1, header=TRUE, sep="\t", quote="") #in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
dim(data) #オブジェクトdataの行数と列数を表示

#本番
obj <- is.element(as.character(data[,param]), keywords) #条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out) #オブジェクトoutの行数と列数を表示

#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F) #outの中身を指定したファイル名で保

```

[トップページへ](#)

Apr 13, 2020

全体像の把握6

こんな感じになります。①例題1が一番上になるように、5行分ほど下部に移動。こんな感じ。

1. 目的のタブ区切りテキストファイル([annotation.txt](#))中の第1列目をキーとして、リストファイル([genelist1.txt](#))中のものが含まれる行全体を出力したい場合：

```
in_f1 <- "annotation.txt"
in_f2 <- "genelist1.txt"
out_f <- "hoge1.txt"
param <- 1

#入力ファイルの読み込み
data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2)
dim(data)

#本番
obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,]
dim(out)

#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保
```

[トップページへ](#)

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

入力ファイルの保存1

こんな感じになります。①例題1が一番上になるように、5行分ほど下部に移動。こんな感じ。②が入力ファイルであり、③と④からダウンロード可能なので、右クリックで「デスクトップ上にあるhogeフォルダ」に保存。

(Rで)塩基配列解析

保護されていない通信 | iu.as^③...jp/~kadota/r_seq.html#intro_g

1. 目的のタブ区切りテキストファイル(annotation.txt)中の第1列目をキーとして、リストファイル(genelist1.txt)中のものが含まれる行全体を出力したい場合：

^④
in_f1 <- "annotation.txt"
in_f2 <- "genelist1.txt"
out_f <- "hoge1.txt"
param <- 1



#入力ファイル名を指定してin_f1に格納(アノテーションファイル)
#入力ファイル名を指定してin_f2に格納(リストファイル)
#出力ファイル名を指定してout_fに格納
#アノテーションファイル中の検索したい列番号を指定

#入力ファイルの読み込み

data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
dim(data) #オブジェクトdataの行数と列数を表示

#本番

obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out) #オブジェクトoutの行数と列数を表示

#ファイルに保存

write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#out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2

正しく保存できれば、RStudio上の① hogeフォルダは、②のように見えてい
と思います。③必要に応じてリロードも
行ってください。

The screenshot shows the RStudio interface with the following components:

- Editor:** Contains a FASTA file with the following content:

```
1 >description1
2 AGTGACGGTCTT
3 |
```
- Console:** Shows the execution of `writeXStringSet` and its output:

```
A AASCTingSet instance of length 1
  width seq          names
[1]    4 SDGL        description1
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", w
idth=50)#fastaの中身を指定したファイル名で保存
>
> |
```
- Environment/History:** Shows the R code being executed:

```
tasta <- translate(tasta) ...
fasta #確認してるだけです
#ファイルに保存
writeXStringSet(fasta, fil...
```
- Files Panel:** Shows the file explorer for the 'hoge' directory. It contains a table of files:

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B

入力ファイルの保存3

正しく保存できれば、RStudio上の① hogeフォルダは、②のように見えていると思います。③必要に応じてリロードも行ってください。④⑤⑥⑦を押して全体をすっきりさせます。

The screenshot shows the RStudio interface with several red arrows pointing to specific elements:

- ④: Points to the save icon in the top toolbar.
- ⑤: Points to the 'sample1.fasta' tab in the editor.
- ⑥: Points to the 'To Console' button in the Environment pane.
- ⑦: Points to the 'Yes' button in the 'Confirm Clear History' dialog box.

The console output shows the following commands and results:

```
> description1
AGTGACGGTCTT

A ASCII string instance of length 1
width seq          names
[1]      4 SDGL     description1

> #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存
>
> |
```

The Environment pane shows the following code:

```
tasta <- translate(tasta)
fasta #確認してるだけです
#ファイルに保存
```

The file browser shows the following files:

File Name	Size
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B

入力ファイルの保存4

正しく保存できれば、RStudio上の① hogeフォルダは、②のように見えていると思います。③必要に応じてリロードも行ってください。④⑤⑥⑦を押して全体をすっきりさせます。⑧も押します。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main workspace is divided into several panes. On the left, the Console pane shows the following R code and output:

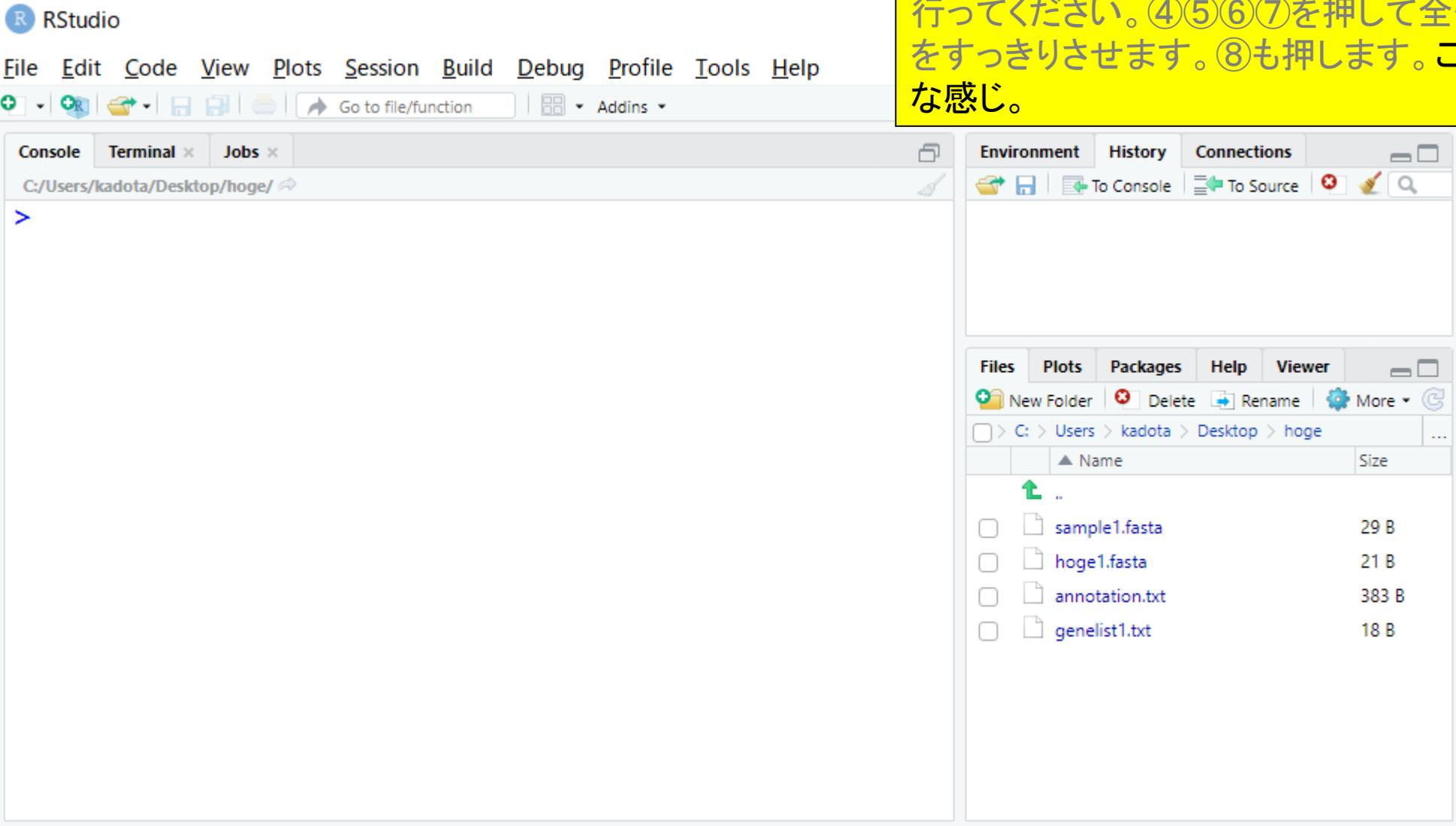
```
>
> #本番
> fasta <- translate(fasta)           #アミノ酸配列に
翻訳した結果をfastaに格納
> fasta                               #確認してるだけ
です
A DNASTringSet instance of length 1
  width seq          names
[1]    12 AGTGACGGTCTT description1
>
> #ファイルに保存
> writeXStringSet(fasta, file=out_f, format="fasta", width=50)#fastaの中身を指定したファイル名で保存
>
> |
```

On the right, the Environment pane shows a tooltip for the 'Clear Console (Ctrl+L)' button, which is highlighted with a red arrow and the number 8. Below the Environment pane is the Files pane, which shows a file explorer view of the project directory. The file explorer shows the following files:

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B

入力ファイルの保存5

正しく保存できれば、RStudio上の① hogeフォルダは、②のように見えていると思います。③必要に応じてリロードも行ってください。④⑤⑥⑦を押して全体をすっきりさせます。⑧も押します。こんな感じ。



入力ファイルの保存6

入力: アノテーションファイル(annotation.txt)

	A	B	C	D
1	gene1	hoge01	plasma_mem	nuclear
2	gene2	hoge02	hohinu	membrane
3	gene3	hoge03	agribio	endoplasmic
4	gene4	hoge04	genesis	endoplasmic
5	gene5	hoge05	kamo	membrane
6	gene6	hoge06	netteba	humei
7	gene7	hoge07	tebasaki	nuclear
8	gene8	hoge08	biiru	nuclear
9	gene9	hoge09	nihonshu	nuclear
10	gene10	hoge10	agene1	membrane
11	gene11	hoge11	iyaaaa	endoplasmic

入力: リストファイル(genelist1.txt)

	A
1	gene1
2	gene7
3	gene9

アノテーションファイル(annotation.txt)中の第1列目に対して、リストファイル(genelist1.txt)中の文字列と一致する行を抜き出して、hoge1.txtというファイル名で出力したい、という目的を達成する上で必要な①2つの入力ファイルが作業ディレクトリ上に存在するのでready-to-analyze。

出力: hoge1.txt

	A	B	C	D
1	gene1	hoge01	plasma_mem	nuclear
2	gene7	hoge07	tebasaki	nuclear
3	gene9	hoge09	nihonshu	nuclear

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

コピー実行1

(Rで)塩基配列解析

保護されていない通信 | iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_grep_kiso

ゲスト

1. 目的のタブ区切りテキストファイル([annotation.txt](#))中の第1列目をキーとして、リストファイル([genelist1.txt](#))中のものが含まれる行全体を出力したい場合：

```

in_f1 <- "annotation.txt"      #入力ファイル名を指定してin_f1に格納(アノテーションファイル)
in_f2 <- "genelist1.txt"      #入力ファイル名を指定してin_f2に格納(リストファイル)
out_f <- "hoge1.txt"          #出力ファイル名を指定してout_fに格納
param <- 1                     #アノテーションファイル中の検索したい列番号を指定

#入力ファイルの読み込み
data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2)  #in_f2で指定したファイルの読み込み
dim(data)                     #オブジェクトdataの行数と列数を表示

#本番
obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,]              #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out)                       #オブジェクトoutの行数と列数を表示

#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保

```

[トップページへ](#)

コピペ実行2

赤枠内をコピー実行。①RStudioのConsole画面上で、②Paste直後の状態。③リターンキーを押す。

The screenshot shows the RStudio interface. The console window contains the following R code with Japanese comments:

```
keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
dim(data) #オブジェクトdataの行数と列数を表示

#本番
obj <- is.element(as.character(ram), keywords)#条件を満たすかどうかを格納
out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out) #オブジェクトoutの行数と列数を表示

#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保存
```

A context menu is open over the code, with the 'Paste' option highlighted. A red arrow labeled '2' points to the 'Paste' option.

A keyboard overlay is shown on the right side of the screenshot, with a red arrow labeled '3' pointing to the 'Enter' key.

Red arrows labeled '1' and '2' point to the top-left corner of the RStudio window and the 'Paste' option in the context menu, respectively.

コピー実行3

赤枠内をコピー実行。①RStudioのConsole画面上で、②Paste直後の状態。③リターンキーを押す。こんな感じになって、④指定した出力ファイル(hoge1.txt)が作成できれば基本成功です。④をクリック。

The screenshot shows the RStudio interface. The console window on the left contains the following R code and its output:

```
> dim(data) #オブジェクトdataの行数と列数を表示
[1] 11 4
>
> #本番
> obj <- is.element(as.character(data[,param]), keywords) #条件を満たすかどうかを判定した結果をobjに格納
> out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
> dim(out) #オブジェクトoutの行数と列数を表示
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F) #outの中身を指定したファイル名で保存
>
>
```

The file explorer window on the right shows the directory structure: C:\Users\kadota\Desktop\hoge. It contains several files, with 'hoge1.txt' highlighted by a red arrow and a circled number 4, indicating the step where the file is created.

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B
hoge1.txt	148 B

出力ファイルの確認1

赤枠内をコピー実行。①RStudioのConsole画面上で、②Paste直後の状態。③リターンキーを押す。こんな感じになって、④指定した出力ファイル(hoge1.txt)が作成されれば基本成功です。④をクリック。⑤のような感じで出力ファイルの中身が見られます。

```
1:1 | Text File
```

geneid	gene name	accession	description	subcellular_location
gene1	hoge01	plasma_mem	nuclear	
gene7	hoge07	tebasaki	nuclear	
gene9	hoge09	nihonshu	nuclear	

```
> dim(out) #オブジェクトoutの行数と列数を表示
[1] 3 4
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保存
```

```
out <- data[obj,] #objがTR...
dim(out) #オブジェクトoutの行..
#ファイルに保存
write.table(out, out_f, se...
```

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B
hoge1.txt	148 B

出力ファイルの確認2

入力: アノテーションファイル(annotation.txt)

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene2	hoge02	hohinu	membrane
4	gene3	hoge03	agribio	endoplasmic
5	gene4	hoge04	genesis	endoplasmic
6	gene5	hoge05	kamo	membrane
7	gene6	hoge06	netteba	humei
8	gene7	hoge07	tebasaki	nuclear
9	gene8	hoge08	biiru	nuclear
10	gene9	hoge09	nihonshu	nuclear
11	gene10	hoge10	agene1	membrane
12	gene11	hoge11	iyaaaa	endoplasmic

入力: リストファイル(genelist1.txt)

	A
1	gene1
2	gene7
3	gene9

赤枠内をコピー実行。①RStudioのConsole画面上で、②Paste直後の状態。③リターンキーを押す。こんな感じになって、④指定した出力ファイル(hoge1.txt)が作成できれば基本成功です。④をクリック。⑤のような感じで出力ファイルの中身が見られます。最初に見せたものと同じ結果が得られていますね。

出力: hoge1.txt

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene7	hoge07	tebasaki	nuclear
4	gene9	hoge09	nihonshu	nuclear

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

Winユーザ向け注意点1

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。

The screenshot shows the RStudio environment with the following components:

- Code Editor:** A file named 'hoge1.txt' is open, containing a table with 5 rows and 4 columns. A red arrow labeled '2' points to the top-left corner of the editor window.
- Console:** Shows the execution of `dim(out)` resulting in `[1] 3 4`, and the execution of `write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)`. Japanese comments explain the output and the purpose of the `write.table` function. A red arrow labeled '1' points to the file explorer.
- File Explorer:** Shows the directory `C:\Users\kadota\Desktop\hoge` containing files like `sample1.fasta`, `hoge1.fasta`, `annotation.txt`, `genelist1.txt`, and `hoge1.txt`. A red arrow labeled '1' points to the `hoge1.txt` file.

Winユーザー向け注意点2

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。先に①hoge1.txtを、②Rエディタ上で開いた状態では、問題なくExcelで開くことができます。

The screenshot shows the RStudio interface. The main editor window displays a text file named 'hoge1.txt' with the following content:

```
1 gene_name accession description subcellular_location
2 gene1 hoge01 plasma_mem nuclear
3 gene7 hoge07 tebasaki nuclear
4 gene9 hoge09 nihonshu nuclear
5
```

The console window shows the following commands and output:

```
> dim(out)
#オブジェクトoutの行数と列数を表示
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保存
>
>
```

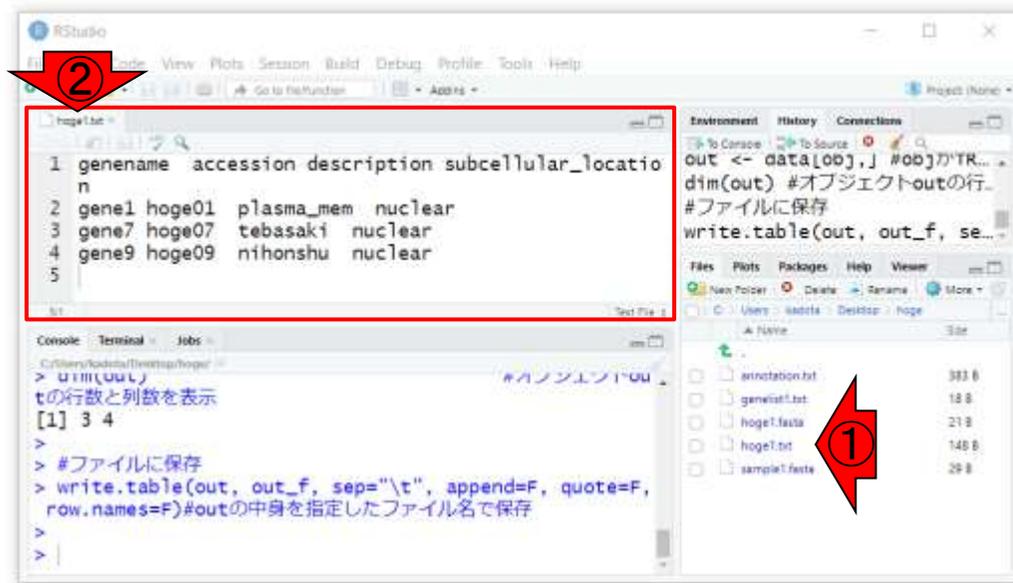
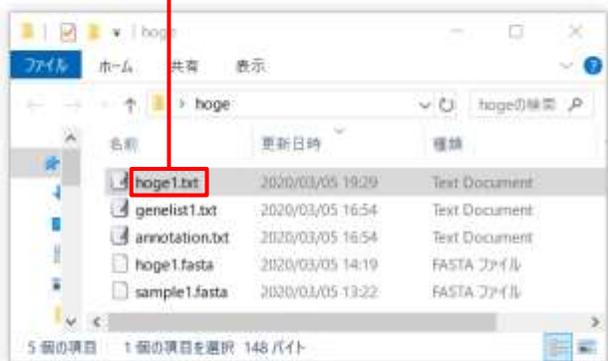
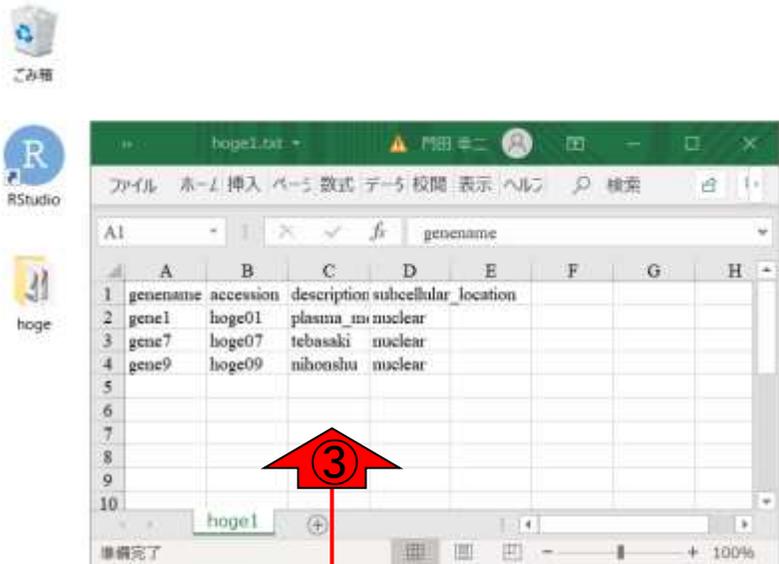
The file explorer on the right shows the directory 'C:\Users\kadota\Desktop\hoge' with the following files:

Name	Size
..	
sample1.fasta	29 B
hoge1.fasta	21 B
annotation.txt	383 B
genelist1.txt	18 B
hoge1.txt	148 B

Red arrows with numbers 1 and 2 point to the 'hoge1.txt' file in the file explorer and the 'hoge1.txt' tab in the RStudio editor, respectively.

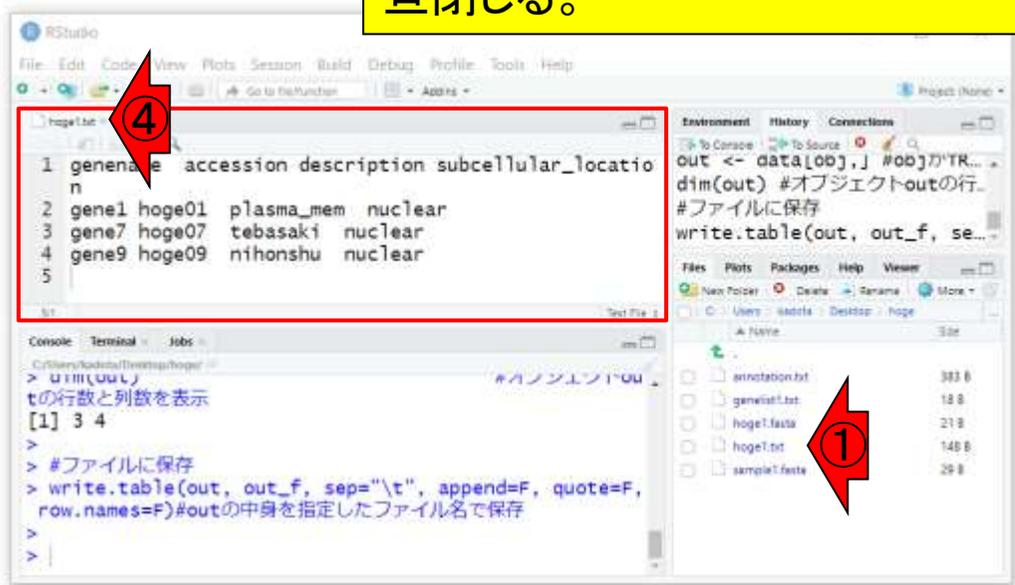
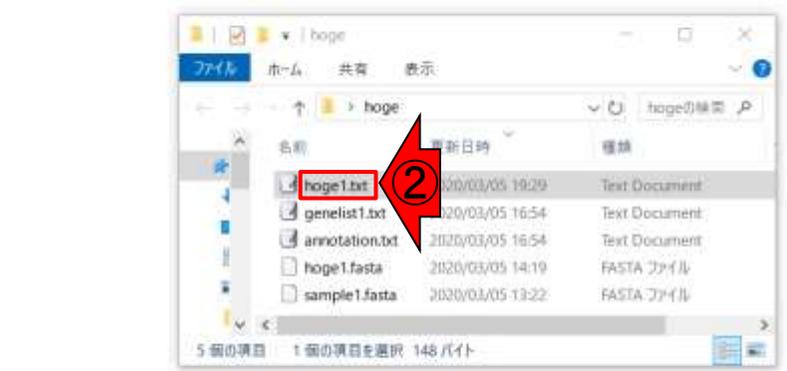
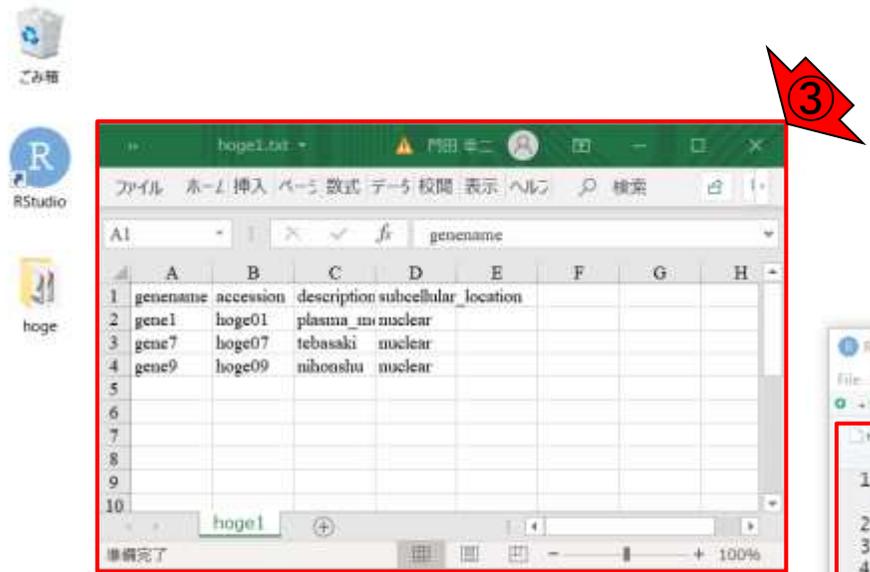
Winユーザー向け注意点3

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。先に①hoge1.txtを、②Rエディタ上で開いた状態では、③問題なくExcelで開くことができます。つまりこういうこと。



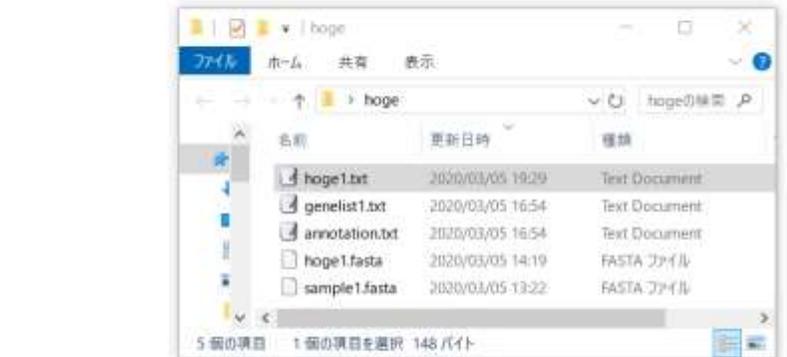
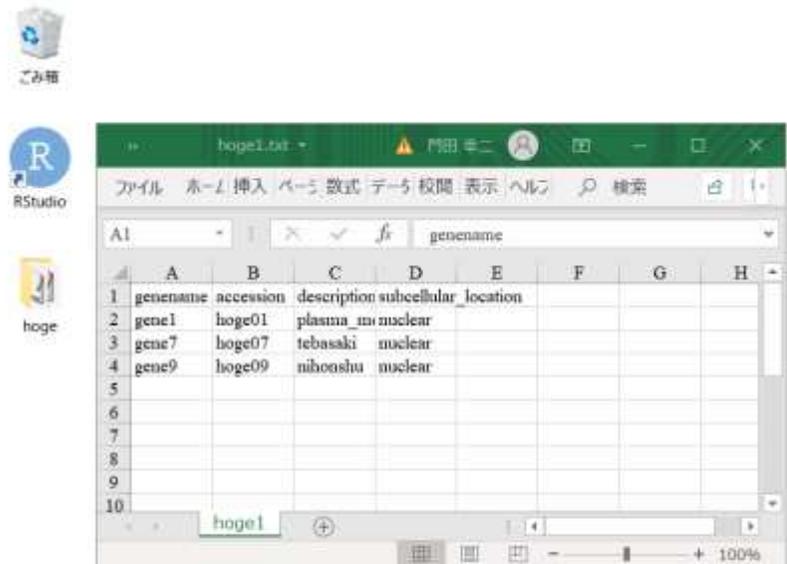
Winユーザ向け注意点4

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。先に①hoge1.txtを、②Rエディタ上で開いた状態では、③問題なくExcelで開くことができます。つまりこういうこと。次に、① or ②を③Excelで開いている状態で、④Rエディタ上で開いているhoge1.txtを一旦閉じる。



Winユーザ向け注意点5

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。先に①hoge1.txtを、②Rエディタ上で開いた状態では、③問題なくExcelで開くことができます。つまりこういうこと。次に、① or ②を③Excelで開いている状態で、④Rエディタ上で開いているhoge1.txtを一旦閉じる。つまりこういうこと。この状態で、再度①hoge1.txtをクリックして開こうとすると…



```

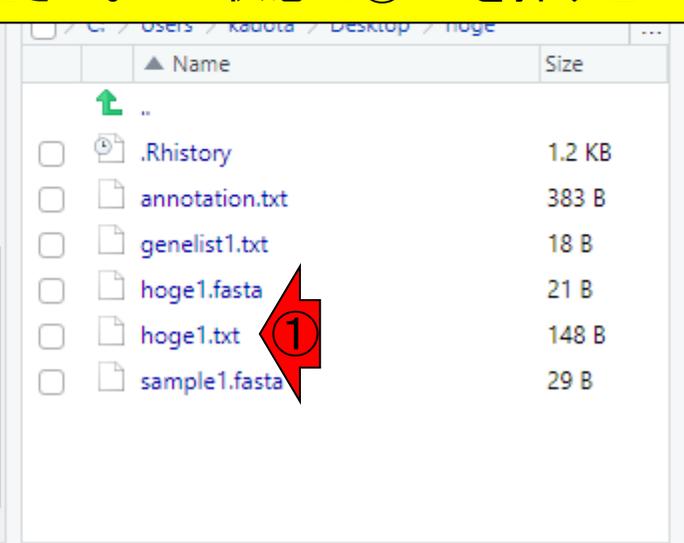
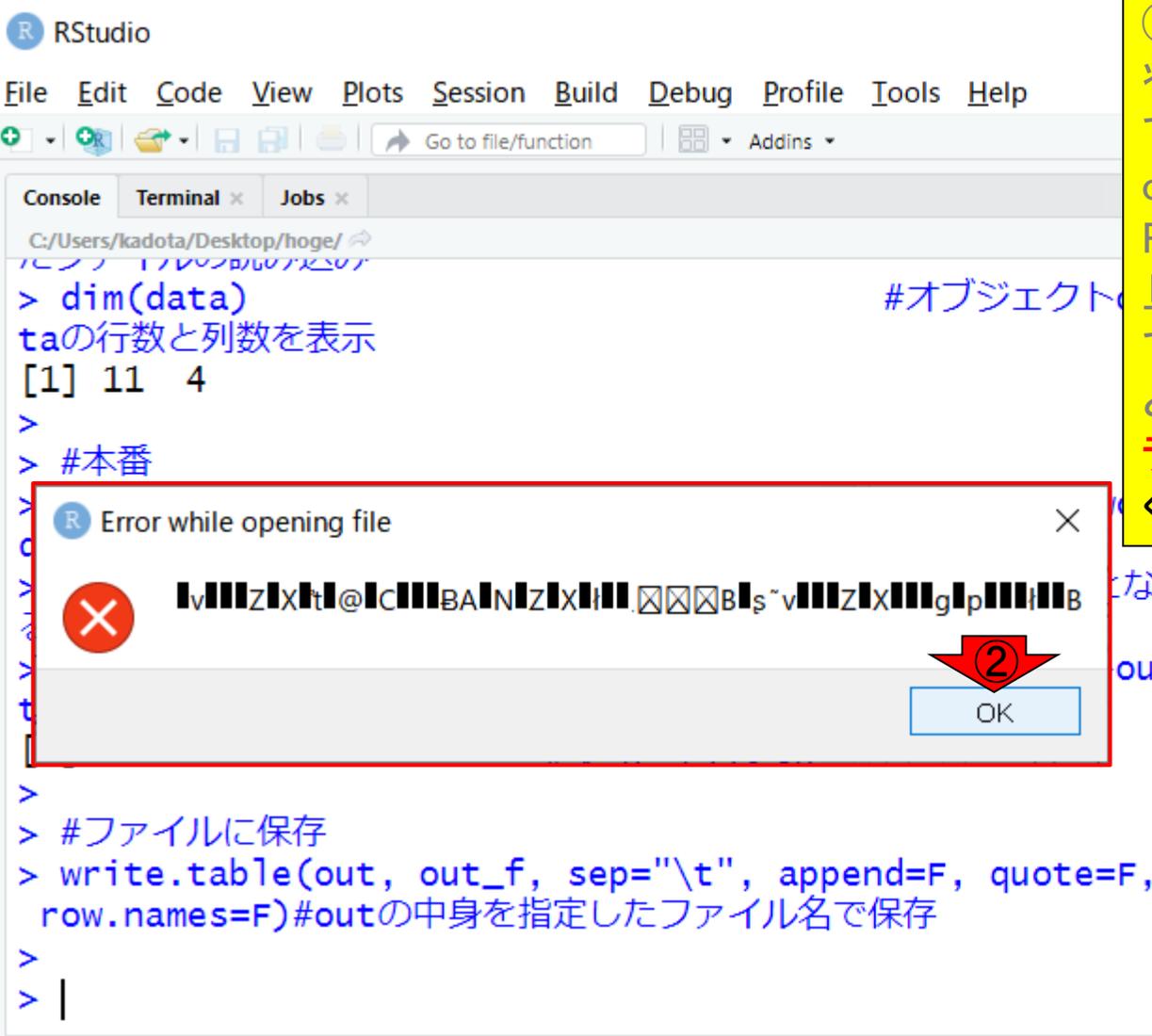
RStudio
File Edit Code View Plots Session Build
Console Terminal Jobs
C:\Users\koki\Desktop\hoge\hoge1.txt
> dim(data)
#オブジェクトdata
taの行数と列数を表示
[1] 11 4
>
> #本番
> obj <- is.element(as.character(data[,param]), keyword)
#条件を満たすかどうかを判定した結果をobjに格納
> out <- data[obj,]
#objがTRUEとなる行のみ抽出した結果をoutに格納
> dim(out)
#オブジェクトoutの行数と列数を表示
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)
#outの中身を指定したファイル名で保存
>
Files Plots Packages Help Viewer
New Folder Delete Rename More
C:\Users\koki\Desktop\hoge
hoge
annotation.txt 383 B
genelists.txt 18 B
hoge1.fasta 21 B
hoge1.txt 148 B
sample1.fasta 29 B

```



Winユーザー向け注意点6

先にExcelで開いているものと同じファイルをRエディタで開こうとすると意味不明なエラーが出るという注意点です。先に①hoge1.txtを、②Rエディタ上で開いた状態では、③問題なくExcelで開くことができます。つまりこういうこと。次に、①or ②を③Excelで開いている状態で、④Rエディタ上で開いているhoge1.txtを一旦閉じる。つまりこういうこと。この状態で、再度①hoge1.txtをクリックして開こうとすると、こんな感じの意味不明なエラーメッセージが出ますのでお気をつけください。この状態で②OKを押すと…



Winユーザ向け注意点7

こんな風に意味不明な状態になりますのでお気をつけください。こうなったら、**RStudioを再起動するのが一番手っ取り早い**です。出力予定ファイルと同一のものをExcelで開いているような場合にも同様の不具合が起こりますのでご注意ください。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main editor area is empty. At the bottom, there are three panels: Console, Terminal, and Jobs. The Console panel shows the following text:

```
C:/Users/kadota/Desktop/hoge/
> dim(out) #オブジェクトoutの行数と列数を表示
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保存
>
> |
```

On the right side, a file explorer window is open, showing the directory C:\Users\kadota\Desktop\hoge. The file list is as follows:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
sample1.fasta	29 B

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

応用1

①例題1は、②の中身をもつ、③リストファイル(hoge1.txt)を読み込んで、④1列目で一致する行を抽出するスクリプトでした。

(Rで)塩基配列解析

← ① iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_grep_kiso ゲスト

1. 目的のタブ区切りテキストファイル([annotation.txt](#))中の第1列目をキーとして、リストファイル([genelist1.txt](#))中のものが含まれる行全体を出力したい場合：

```

in_f1 <- "annotation.txt"
in_f2 <- "genelist1.txt"
out_f <- "hoge1.txt"
param <- 1

#入力ファイルの読み込み
data <- read.table(in_f1, header=T, quote="")#in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2)#in_f2で指定したファイルの読み込み
dim(data)#データオブジェクトdataの行数と列数を表示

#本番
obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,]#条件を満たす行のみ抽出した結果をoutに格納
dim(out)#抽出した結果の行数と列数を表示

#ファイルに書き出す
write.table(out, out_f, row.names=F)#outの中身を指定したファイル名で保存
    
```

②

③

④

④

↑

[トップページ](#)

	A	B	C	D
1	gene1	hoge01	plasma_mem	nuclear
2	gene2	hoge02	hohinu	membrane
3	gene3	hoge03	agribio	endoplasmic
4	gene4	hoge04	genesis	endoplasmic
5	gene5	hoge05	kamo	membrane
6	gene6	hoge06	netteba	humei
7	gene7	hoge07	tebasaki	nuclear
8	gene8	hoge08	biiru	nuclear
9	gene9	hoge09	nihonshu	nuclear
10	gene10	hoge10	agene1	membrane
11	gene11	hoge11	iyaaaa	endoplasmic

応用2

①例題1は、②の中身をもつ、③リストファイル(hoge1.txt)を読み込んで、④1列目で一致する行を抽出するスクリプトでした。応用として、⑤新たに作成したリストファイル(list.txt)を読み込んで、⑥4列目で一致する行の抽出を行います。

1. 目的のタブ区切りテキストファイル(annotation.txt)中の第1列目をキーとして、リストファイル(genelist1.txt)中のものが含まれる行全体を出力したい場合:

```
in_f1 <- "annotation.txt"
in_f2 <- "genelist1.txt"
out_f <- "hoge1.txt"
param <- 1
```

#入力ファイルの読み込み

```
data <- read.table(in_f1, header=T, as.is=T, quote="")
keywords <- readLines(in_f2)
dim(data)
```

#本番

```
obj <- is.element(as.character(data[,param]), keywords)
out <- data[obj,]
dim(out)
```

#ファイルに書き出す

```
write.table(out, out_f, as.is=T, quote="")
```

gene1
gene7
gene9

nuclear
membrane

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene2	hoge02	hohinu	membrane
4	gene3	hoge03	agribio	endoplasmic
5	gene4	hoge04	genesis	endoplasmic
6	gene5	hoge05	kamo	membrane
7	gene6	hoge06	netteba	humei
8	gene7	hoge07	tebasaki	nuclear
9	gene8	hoge08	biiru	nuclear
10	gene9	hoge09	nihonshu	nuclear
11	gene10	hoge10	agene1	membrane
12	gene11	hoge11	iyaaaa	endoplasmic

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene2	hoge02	hohinu	membrane
4	gene3	hoge03	agribio	endoplasmic
5	gene4	hoge04	genesis	endoplasmic
6	gene5	hoge05	kamo	membrane
7	gene6	hoge06	netteba	humei
8	gene7	hoge07	tebasaki	nuclear
9	gene8	hoge08	biiru	nuclear
10	gene9	hoge09	nihonshu	nuclear
11	gene10	hoge10	agene1	membrane
12	gene11	hoge11	iyaaaa	endoplasmic

応用3

①例題1は、②の中身をもつ、③リストファイル(hoge1.txt)を読み込んで、④1列目で一致する行を抽出するスクリプトでした。応用として、⑤新たに作成したリストファイル(list.txt)を読み込んで、⑥4列目で一致する行の抽出を行います。これはRStudioを再起動して、⑦作業ディレクトリの変更まで終了した状態。

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。

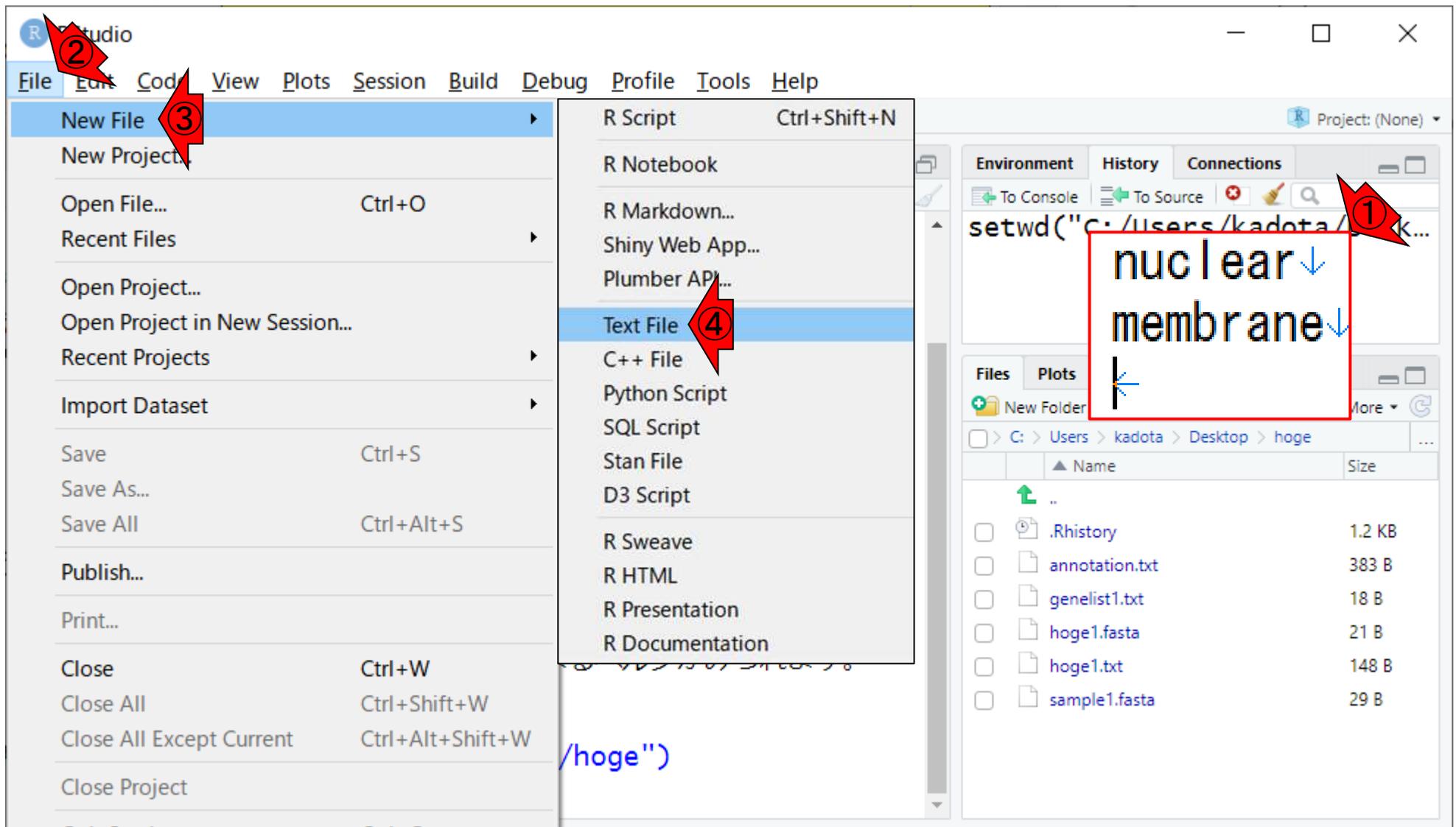
'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

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

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
sample1.fasta	29 B

応用4

まずは①リストファイル(list.txt)を作成
すべく、②File、③New File、④Text File。



応用5

まずは①リストファイル(list.txt)を作成
すべく、②File、③New File、④Test File。
こんな感じになります。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The toolbar contains various icons for file operations and navigation. The main editor window shows a single line of code: `1`. The console window at the bottom displays the following text:

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

The console also contains Japanese text: `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。

The Environment pane on the right shows the current working directory: `setwd("C:/Users/kadota/Desktop/hoge")`. Below it, the Files pane shows a list of files in the directory `C:\Users\kadota\Desktop\hoge`:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
sample1.fasta	29 B

A red box highlights the text `nuclear` and `membrane` in the Environment pane, with a red arrow pointing to the first character of `nuclear`.

応用6

まずは①リストファイル(list.txt)を作成すべく、②File、③New File、④Test File。こんな感じになります。①と同じように抽出したいキーワードを入力。この際、②の部分で最後にリターンキーを「押すか押さないか」でコピー実行時に「警告が出ないか出るか」の違いがでますがそれも含めて楽しんでください。

RStudio interface showing a text editor with the following content:

```
1 nuclear
2 membrane
3
```

The console shows the following text:

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

File explorer window showing a list of files in the directory C:\Users\kadota\Desktop\hoge:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
sample1.fasta	29 B

応用7

まずは①リストファイル(list.txt)を作成すべく、②File、③New File、④Test File。こんな感じになります。①と同じように抽出したいキーワードを入力。この際、②の部分で最後にリターンキーを「押すか押さないか」でコピー実行時に「警告が出ないか出るか」の違いがでますがそれも含めて楽しんでください。③を押して保存。

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main editor window shows a text file named 'Untitled1*' with the following content:

```
1 number
2 membrane
3
```

A red arrow with the number '3' points to the 'New File' icon in the toolbar. The console window at the bottom shows the following R commands and their output:

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

The console also displays help text for the 'demo()' function:

```
'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。
```

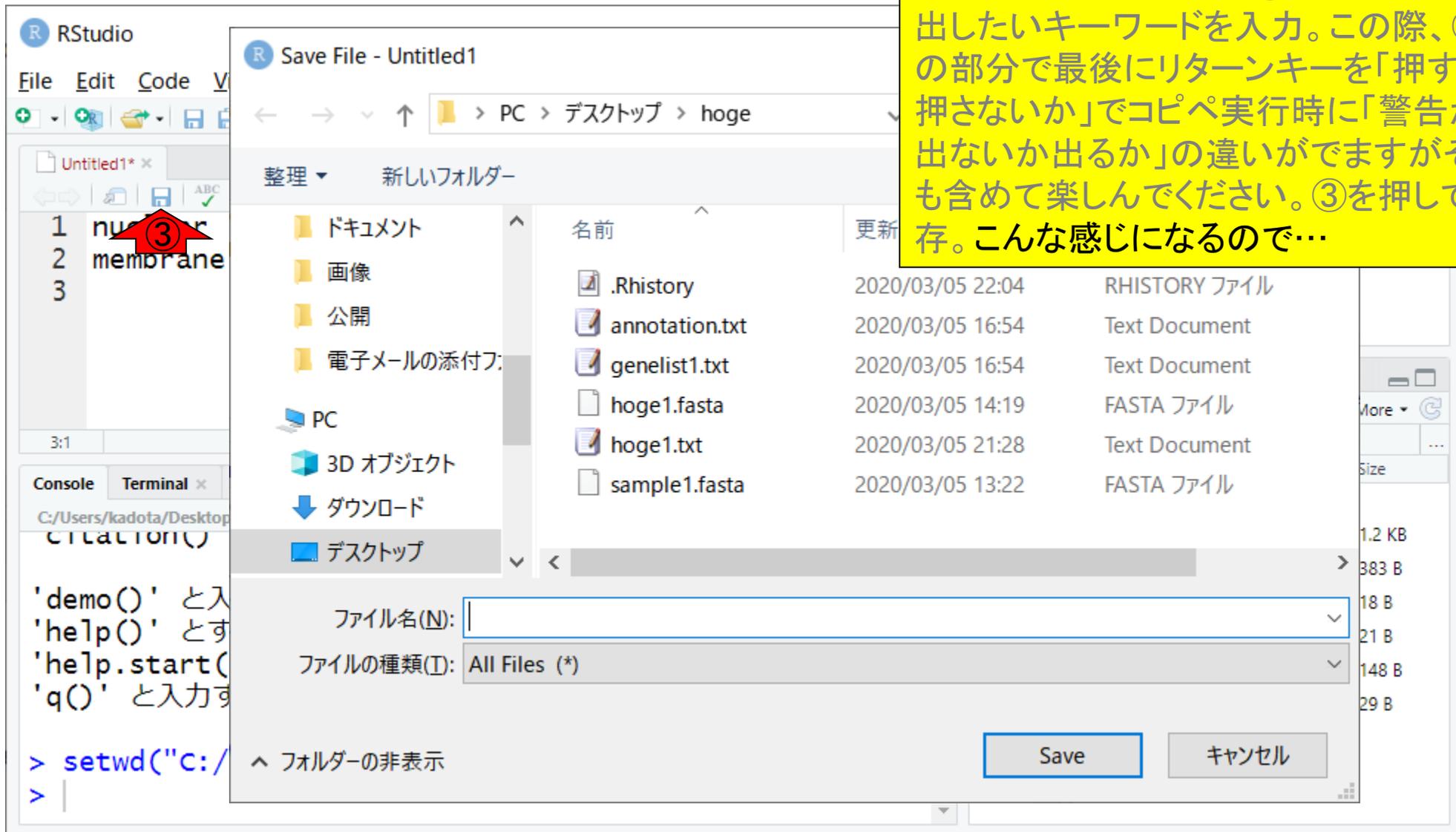
The screenshot shows a file explorer window with the following details:

- Files tab selected
- Buttons: New Folder, Delete, Rename, More
- Path: C:\Users\kadota\Desktop\hoge
- Table listing files and folders:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
sample1.fasta	29 B

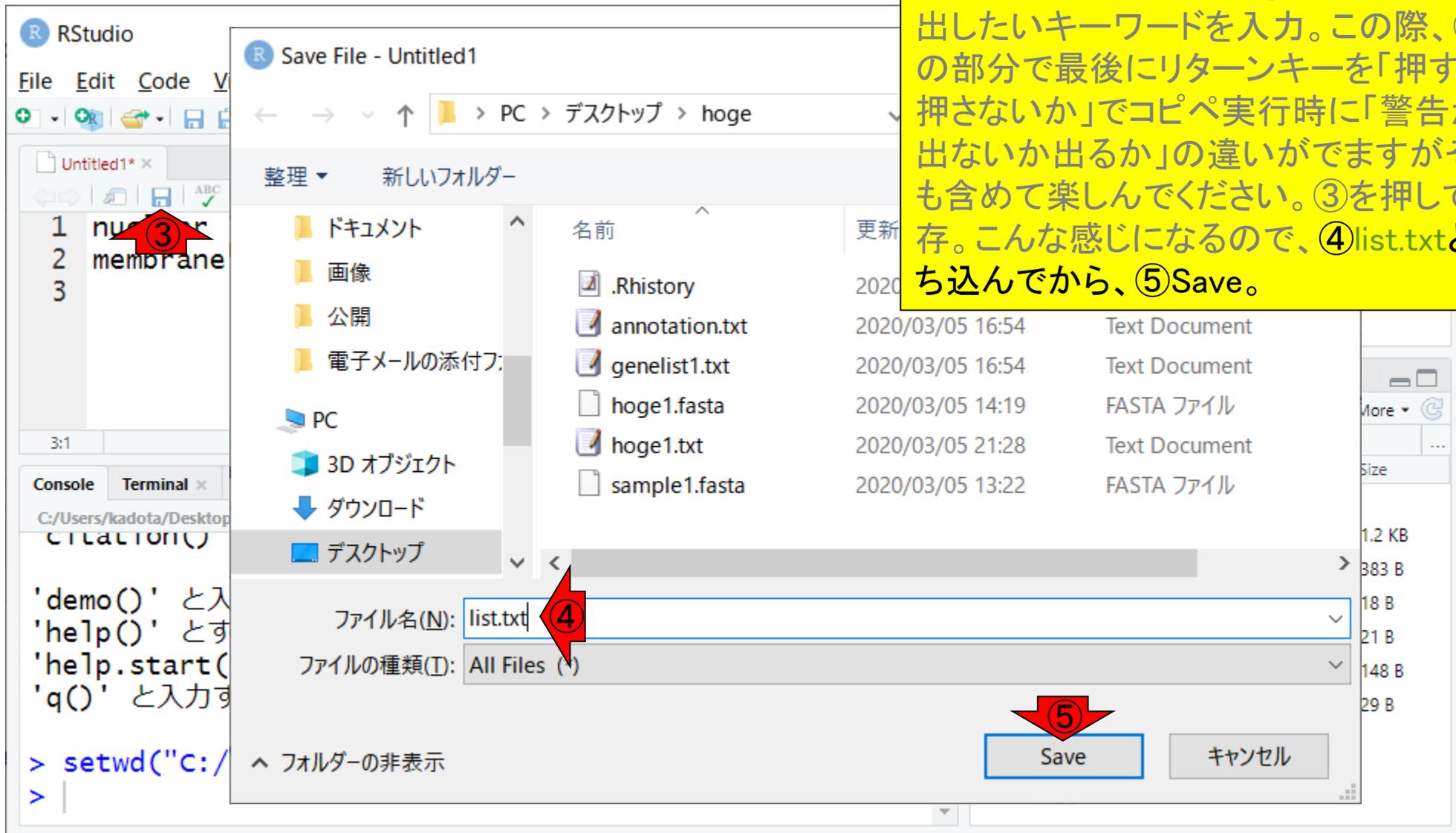
応用8

まずは①リストファイル(list.txt)を作成すべく、②File、③New File、④Test File。こんな感じになります。①と同じように抽出したいキーワードを入力。この際、②の部分で最後にリターンキーを「押すか押さないか」でコピー実行時に「警告が出ないか出るか」の違いがでますがそれも含めて楽しんでください。③を押して保存。こんな感じになるので…



応用9

まずは①リストファイル(list.txt)を作成すべく、②File、③New File、④Test File。こんな感じになります。①と同じように抽出したいキーワードを入力。この際、②の部分で最後にリターンキーを「押すか押さないか」でコピペ実行時に「警告が出ないか出るか」の違いがでますがそれも含めて楽しんでください。③を押して保存。こんな感じになるので、④list.txtと打ち込んでから、⑤Save。



応用10

まずは①リストファイル(list.txt)を作成すべく、②File、③New File、④Test File。こんな感じになります。①と同じように抽出したいキーワードを入力。この際、②の部分で最後にリターンキーを「押すか押さないか」でコピー実行時に「警告が出ないか出るか」の違いがでますがそれも含めて楽しんでください。③を押して保存。こんな感じになるので、④list.txtと打ち込んでから、⑤Save。私の場合は、⑥リロードすると、⑦list.txtが見えました。

RStudio interface showing a text editor with the following content:

```
1 nuclear
2 membrane
3 |
```

The console shows the following text:

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

Windows File Explorer showing a directory listing with the following files:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

Red arrows point to the 'list.txt' file (7) and the 'More' button in the File Explorer toolbar (6).

次は、①例題1のスクリプト(or コード)を
テンプレートとして利用したいので…

応用11

1. 目的のタブ区切りテキストファイル([annotation.txt](#))中の第1列目をキーとして、リストファイル([genelist1.txt](#))中のものが含まれる行全体を出力したい場合：

```
in_f1 <- "annotation.txt"      #入力ファイル名を指定してin_f1に格納(アノテーションファイル)
in_f2 <- "genelist1.txt"       #入力ファイル名を指定してin_f2に格納(リストファイル)
out_f <- "hoge1.txt"           #出力ファイル名を指定してout_fに格納
param <- 1                     #アノテーションファイル中の検索したい列番号を指定

#入力ファイルの読み込み
data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
keywords <- readLines(in_f2)  #in_f2で指定したファイルの読み込み
dim(data)                     #オブジェクトdataの行数と列数を表示

#本番
obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,]              #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out)                       #オブジェクトoutの行数と列数を表示

#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保
```

[トップページへ](#)

応用12

次は、①例題1のスクリプト(or コード)を
テンプレートとして利用したいので、②赤
枠内のコード全体をコピーしておく。

The screenshot shows a web browser window with the URL `iu.a.u-tokyo.ac.jp/~kadota/r_seq.html#intro_general_grep_kiso`. A red arrow labeled '1' points to the back button. Below the address bar, there is a heading: "1. 目的のタブ区切りテキストファイル(annotation.txt)中の第1列目をキーとして、リストファイル(genelist1.txt)中のものが含まれる行全体を出力したい場合:".

The main content is an R script with Japanese comments. A red arrow labeled '2' points to a context menu that is open over the script. The menu items are: "コピー(C) Ctrl+C", "Google で「in_f1 <- \"annotation.txt\" #入力ファイル名を...」を検索(S)", "印刷(P)... Ctrl+P", and "検証(I) Ctrl+Shift+I".

```
in f1 <- "annotation.txt" #入力ファイル名を指定してin f1に格納(アノテーションファイル)
in f2 <- "genelist1.txt" #入力ファイル名を指定してin f2に格納(リストファイル)
out f <- "hoge1.txt" #出力ファイル名を指定してout fに格納
param <- 1
#入力ファイルの読み込み
data <- read.table(in f1)
keywords <- readLines(in f2)
dim(data)
#本番
obj <- is.element(as.character(data[,param]), keywords)#条件を満たすかどうかを判定した結果をobjに格納
out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
dim(out) #オブジェクトoutの行数と列数を表示
#ファイルに保存
write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保
```

[トップページへ](#)

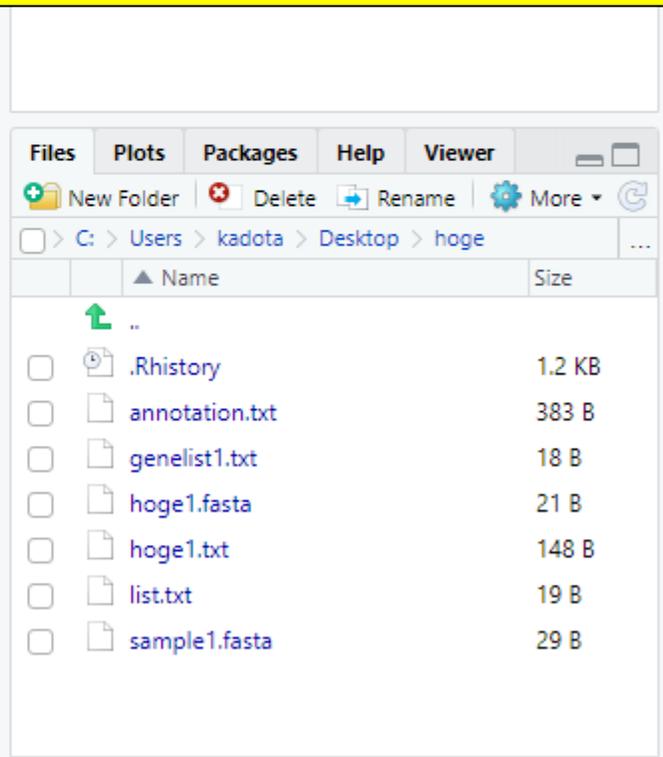
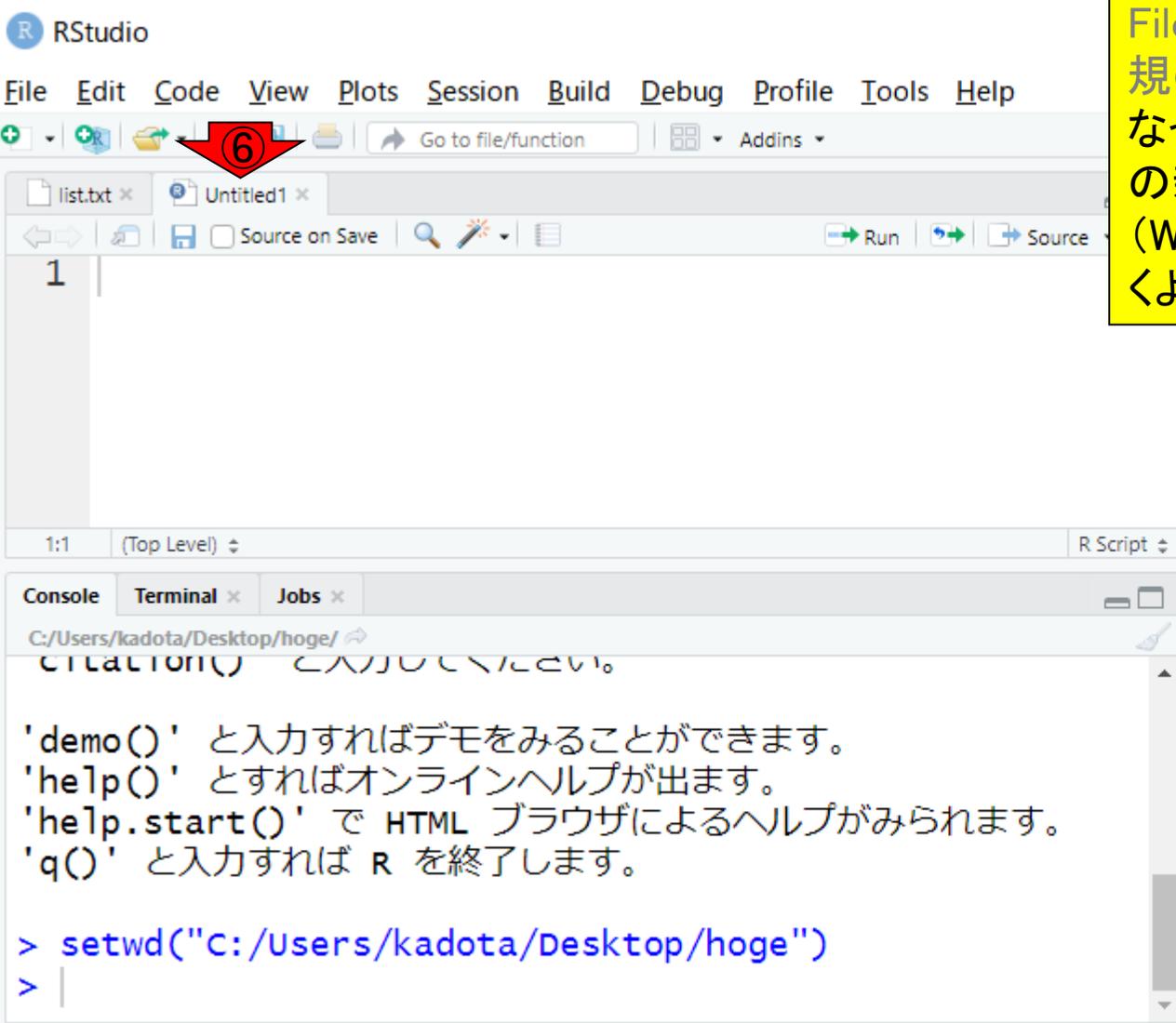
応用13

次は、①例題1のスクリプト(or コード)をテンプレートとして利用したいので、②赤枠内のコード全体をコピーしておく。③ File、④New File、⑤R Scriptとして、「新規のRスクリプト」を開く。

The screenshot shows the R Studio interface. The 'File' menu is open, and the path 'File > New File > R Script' is highlighted with red arrows and numbers 3, 4, and 5. The 'R Script' option is selected. In the background, a file explorer window shows a directory with files like .Rhistory, annotation.txt, and hoge1.fasta. A code editor window shows a snippet of R code: `setwd("C:/Users/kadota/Desktop/hoge"`.

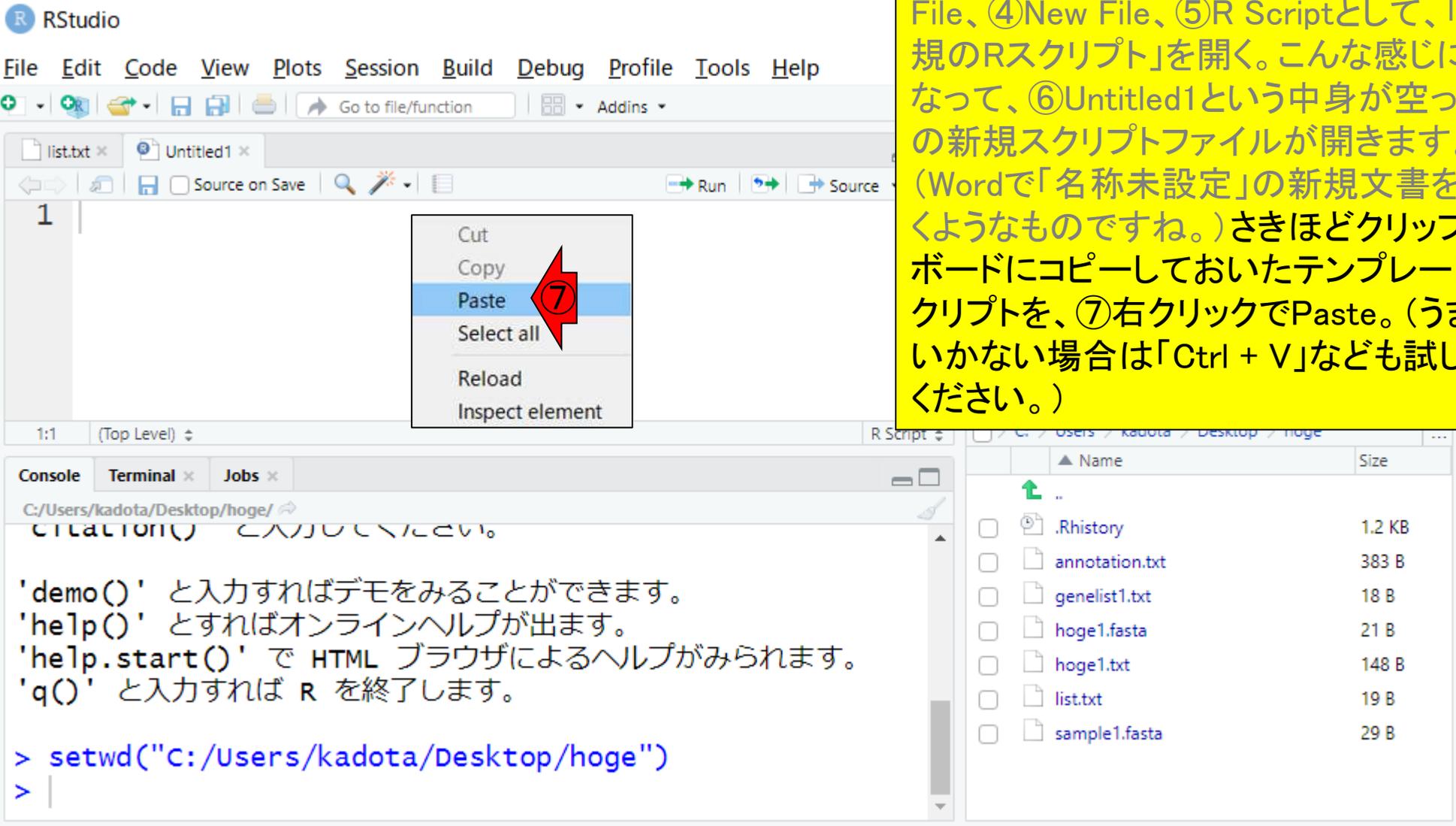
応用14

次は、①例題1のスクリプト(or コード)をテンプレートとして利用したいので、②赤枠内のコード全体をコピーしておく。③File、④New File、⑤R Scriptとして、「新規のRスクリプト」を開く。こんな感じになって、⑥Untitled1という中身が空っぽの新規スクリプトファイルが開きます。(Wordで「名称未設定」の新規文書を開くようなものですね。)



応用15

次は、①例題1のスクリプト(or コード)をテンプレートとして利用したいので、②赤枠内のコード全体をコピーしておく。③File、④New File、⑤R Scriptとして、「新規のRスクリプト」を開く。こんな感じになって、⑥Untitled1という中身が空っぽの新規スクリプトファイルが開きます。(Wordで「名称未設定」の新規文書を開くようなものですね。)さきほどクリップボードにコピーしておいたテンプレートスクリプトを、⑦右クリックでPaste。(うまくいかない場合は「Ctrl + V」なども試してください。)



応用16

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for saving a table to a file.


```

12 obj <- ts.element(as.character(data[,param]), keyv
13 out <- data[obj,]
14 dim(out)
15
16 #ファイルに保存
17 write.table(out, out_f, sep="\t", append=F, quote=
18 '
            
```
- Environment Panel:** Shows the current working directory: `setwd("C:/Users/kadota/Desktop/hoge")`.
- Files Panel:** Displays the contents of the `hoge` directory:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B
- Console:** Shows the execution of `setwd("C:/Users/kadota/Desktop/hoge")` and Japanese instructions for using R help functions:


```

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

'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。

応用17

こんな感じになればOK。今は①エディタ画面と②Console画面の縦幅が非常に狭い。例えば①エディタ画面を広く使いたい場合は、③を押す。

The screenshot shows the RStudio interface with the following components and annotations:

- Editor:** Contains R code for saving a table to a file. A red box highlights this area, and a red arrow labeled '1' points to it.
- Console:** Shows the execution of `demo()` and `setwd()`. A red box highlights this area, and a red arrow labeled '2' points to it.
- Environment/Files:** Shows the current directory and file list. A red arrow labeled '3' points to the maximize button in the top right of the editor area.

```
12 obj <- ts.element(as.character(data[,param]), keyv...
13 out <- data[obj,] #objがTRUE
14 dim(out) #オブジェクト
15
16 #ファイルに保存
17 write.table(out, out_f, sep="\t", append=F, quote=
18 '
```

Console output:

```
> demo()
'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。

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

応用18

こんな感じになればOK。今は①エディタ画面と②Console画面の縦幅が非常に狭い。例えば①エディタ画面を広く使いたい場合は、③を押す。こんな感じになります。もう一度③を押すと元に戻せる。

The screenshot shows the RStudio interface with the following components:

- Script Editor:** Contains R code for reading a table and writing it to a file. A red arrow labeled '3' points to the maximize button in the top right corner of the editor window.
- Environment Pane:** Shows the current working directory as 'C:/Users/kadota/Desktop'.
- Files Pane:** Shows a file explorer view of the 'hoge' directory on the desktop, listing files like 'genelist1.txt', 'hoge1.fasta', and 'hoge1.txt'. A red arrow labeled '1' points to the maximize button in the top right corner of the files pane.
- Console:** Located at the bottom of the interface. A red arrow labeled '2' points to the maximize button in the top right corner of the console window.

応用19

元に戻りました。④はエディタ画面の縦幅を最小にする操作であり、⑤Console画面をを最大にする操作と同じです。以降は特に明記しませんが、見やすいように適宜いじってください。

The screenshot shows the RStudio interface with the following components:

- Code Editor:** Contains R code for reading a file and writing its contents to a table. A red arrow labeled '4' points to the editor's vertical scrollbar.
- Environment/History/Connections:** Shows the current working directory as 'C:/Users/kadota/Desktop/hoge'.
- Files Panel:** Displays a file explorer view of the 'hoge' directory, listing files like '.Rhistory', 'annotation.txt', 'genelist1.txt', 'hoge1.fasta', 'hoge1.txt', 'list.txt', and 'sample1.fasta'. A red arrow labeled '5' points to the file explorer's vertical scrollbar.
- Console:** Shows the output of the 'demo()' function, including instructions on how to use 'demo()', 'help()', 'help.start()', and 'q()'.

応用20

The screenshot shows the RStudio interface. The script editor contains the following R code:

```
1 in_f1 <- "annotation.txt" #入力ファイ
2 in_f2 <- "genelist1.txt" #入力ファ
3 out_f <- "hoge1.txt" #出力ファイ
4 param <- 1 #アノテーシ
5
6 #入力ファイルの読み込み
7 data <- read.table(in_f1, header=TRUE, sep="\t")
8
```

The console shows the following output:

```
> demo()
'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。

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

The file explorer shows the current working directory: C:\Users\kadota\Desktop\hoge. The files listed are:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

応用21

①エディタ画面の上部に移動したところ。
今は、②の部分のファイル名を、③
list.txtに置き換えたい。

The screenshot shows the RStudio interface. The script editor on the left contains the following code:

```
1 in_f1 <- "annotation.txt" #入力ファイ
2 in_f2 <- "genelist1.txt" #入力ファ
3 out_f <- "hoge1.txt" #出力ファイ
4 param <- 1 #アノテーシ
5
6 #入力ファイルの読み込み
7 data <- read.table(in_f1, header=TRUE, sep="\t")
8
```

Red arrows point to specific parts of the code: arrow ① points to the file path in the comment on line 2, arrow ② points to the filename "genelist1.txt" on line 2, and arrow ③ points to the file path in the Environment pane.

The Environment pane on the right shows the current working directory: `setwd("C:/Users/kadota/Desktop/hoge")`. Below this, a search box contains the text "nuclear" and "membrane", with blue arrows pointing to the search results. The Files pane below shows a directory listing:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

The Console pane at the bottom shows the following output:

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

The console also contains Japanese text: `'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。

応用22

①エディタ画面の上部に移動したところ。今は、②の部分のファイル名を、③ list.txtに置き換えたい。また、④1列目ではなく、⑤4列目で一致する行を抽出したいので…

The screenshot shows the RStudio interface with several components:

- Source Editor:** Contains R code for reading files and filtering data. Annotations 1-5 point to specific parts: 1 points to the file path in `setwd()`, 2 points to the file names in `in_f1` and `in_f2`, 3 points to the file name in the `setwd()` path, 4 points to the `param` variable, and 5 points to the `data` variable.
- Environment:** Shows the current working directory as `C:/Users/kadota/Desktop/hoge`.
- Files:** Shows a file explorer view with a search for `nuclear` and `membrane`.
- Console:** Shows the execution of `setwd("C:/Users/kadota/Desktop/hoge")` and a list of instructions for using R help.
- Table:** A table with 12 rows and 4 columns (A, B, C, D) showing gene data. The 4th column (D) contains subcellular locations like 'nuclear' and 'membrane'.

	A	B	C	D
1	gene name	accession	description	subcellular_location
2	gene1	hoge01	plasma_mem	nuclear
3	gene2	hoge02	hohinu	membrane
4	gene3	hoge03	agribio	endoplasmic
5	gene4	hoge04	genesis	endoplasmic
6	gene5	hoge05	kamo	membrane
7	gene6	hoge06	netteba	humei
8	gene7	hoge07	tebasaki	nuclear
9	gene8	hoge08	biiru	nuclear
10	gene9	hoge09	nihonshu	nuclear
11	gene10	hoge10	age ne1	membrane
12	gene11	hoge11	iyaaaa	endoplasmic

応用23

①エディタ画面の上部に移動したところ。
今は、②の部分のファイル名を、③
list.txtに置き換えたい。また、④1列目では
なく、⑤4列目で一致する行を抽出したい
ので、⑥こんな感じで書き替えます。

The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 in_f1 <- "annotation.txt" #入力ファイ
2 in_f2 <- "list.txt" #入力ファイル名:
3 out_f <- "hoge1.txt" #出力ファイ
4 param <- 4 #アノテーシ
5
6 #入力ファイルの読み込み
7 data <- read.table(in_f1, header=TRUE, sep="\t")
8
```

Red arrows labeled '6' point to the file names in lines 2 and 3, and the value '4' in line 4.

The Environment pane shows the command: `setwd("C:/Users/kadota/Desktop/hoge...")`

The Files pane shows the directory structure: `C:\Users\kadota\Desktop\hoge` with files: `..`, `.Rhistory` (1.2 KB), `annotation.txt` (383 B), `genelist1.txt` (18 B), `hoge1.fasta` (21 B), `hoge1.txt` (148 B), `list.txt` (19 B), and `sample1.fasta` (29 B).

The Console shows the following output:

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

The console also contains help text: `'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。`, and `'q()' と入力すれば R を終了します。`

応用24

①エディタ画面の上部に移動したところ。今は、②の部分のファイル名を、③ list.txtに置き換えたい。また、④1列目ではなく、⑤4列目で一致する行を抽出したいので、⑥こんな感じで書き替えます。⑦を押して、(.txtにしても特に問題はありませんが)Rスクリプトであることを明示した拡張子.Rをつけて、⑨Save。

The screenshot shows the RStudio interface on the left and a Windows Explorer window titled "Save File - Untitled1" on the right. The Explorer window is open to the path "PC > デスクトップ > hoge". A table of files is displayed:

名前	更新日時	種類
.Rhistory	2020/03/05 22:04	RHISTORY ファイル
annotation.txt	2020/03/05 16:54	Text Document
genelist1.txt	2020/03/05 16:54	Text Document
hoge1.fasta	2020/03/05 14:19	FASTA ファイル
hoge1.txt	2020/03/05 21:28	Text Document
list.txt	2020/03/05 22:29	Text Document
sample1.fasta	2020/03/05 13:22	FASTA ファイル

At the bottom of the Explorer window, the "ファイル名(N):" field contains "hoge.R" (with a red arrow ⑧ pointing to it) and the "ファイルの種類(I):" dropdown is set to "All Files (*)". The "Save" button (with a red arrow ⑨ pointing to it) is highlighted. In the background RStudio editor, a red arrow ⑦ points to the first column of the script.

応用25

①エディタ画面の上部に移動したところ。今は、②の部分のファイル名を、③ list.txtに置き換えたい。また、④1列目ではなく、⑤4列目で一致する行を抽出したいので、⑥こんな感じで書き替えます。⑦を押して、(.txtにしても特に問題はありませんが)Rスクリプトであることを明示した拡張子.Rをつけて、⑨Save。私の場合は、⑩リロードすると、⑪hoge.Rが見えました。

RStudio interface showing the R script editor and the file explorer.

```
1 in_f1 <- "annotation.txt" #入力ファイル名
2 in_f2 <- "list.txt" #入力ファイル名
3 out_f <- "hoge1.txt" #出力ファイル名
4 param <- 4 #アノテーション数
5
6 #入力ファイルの読み込み
7 data <- read.table(in_f1, header=TRUE, sep="\t", as.is=TRUE)
```

Console output:

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

File Explorer (C:/Users/kadota/Desktop/hoge):

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

応用26

いよいよ変更後のスクリプトの実行。①
赤枠内のコード全体を選択。

The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 in_f1 <- "annotation.txt" #入力ファイル名
2 in_f2 <- "list.txt" #入力ファイル名
3 out_f <- "hoge1.txt" #出力ファイル名
4 param <- 4 #アノテーション
5
6 #入力ファイルの読み込み
7 data <- read.table(in_f1, header=TRUE, sep="\t")
```

The console displays the following text:

```
> 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。

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

The file explorer shows the directory structure:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

応用27

いよいよ変更後のスクリプトの実行。①
赤枠内のコード全体を選択。こんな感じ。

The screenshot shows the RStudio interface. The script editor on the left contains R code with a red box highlighting lines 12 through 18. A red arrow with the number '1' points to the 'Run' button above the script editor. The console at the bottom shows the execution of the highlighted code, including the `setwd` command.

```
12 obj <- ls.element(as.character(data[,param]), keyv...  
13 out <- data[obj,] #objがTRUE  
14 dim(out) #オブジェクト  
15  
16 #ファイルに保存  
17 write.table(out, out_f, sep="\t", append=F, quote=  
18 |
```

Console output:

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

Files panel (C:\Users\kadota\Desktop\hoge):

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

応用28

いよいよ変更後のスクリプトの実行。①
赤枠内のコード全体を選択。こんな感じ。
この状態で②Runを押すと…

The screenshot shows the RStudio interface with the following elements:

- Code Editor:** Contains R code for selecting elements from a data frame and writing them to a table. Lines 12-18 are highlighted in blue. A red box surrounds this code. A red arrow labeled '2' points to the Run button, and another red arrow labeled '1' points to the selected code.
- Run Button:** A red arrow labeled '2' points to the Run button in the toolbar.
- Environment Panel:** Shows the current working directory set to 'C:/Users/kadota/Desktop/hoge...'.
- Files Panel:** Shows a file explorer view of the 'hoge' directory on the desktop, listing files like .Rhistory, annotation.txt, genelist1.txt, hoge.R, hoge1.fasta, hoge1.txt, list.txt, and sample1.fasta.
- Console:** Displays the output of the script, including help text and the command `setwd("C:/Users/kadota/Desktop/hoge")`.

応用29

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。

The screenshot shows the RStudio interface. The source editor contains R code for selecting a column and writing it to a file. The console shows an error message about permissions. The file explorer shows the directory structure.

```
12 obj <- ts.element(as.character(data[,param]), keyv...
13 out <- data[obj,] #objがTRUE
14 dim(out) #オブジェクト
15
16 #ファイルに保存
17 write.table(out, out_f, sep="\t", append=F, quote=
18 |
```

Console output:

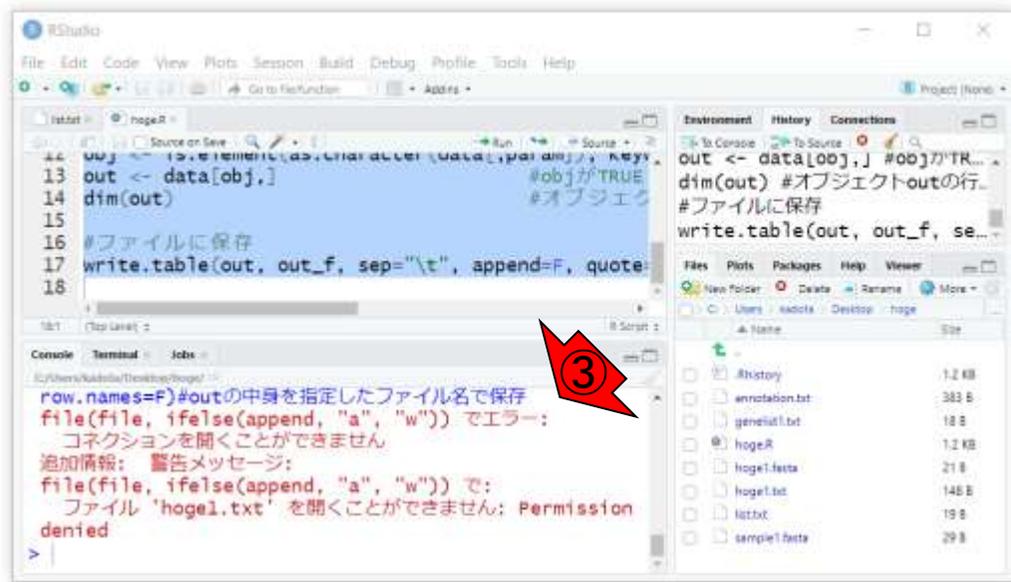
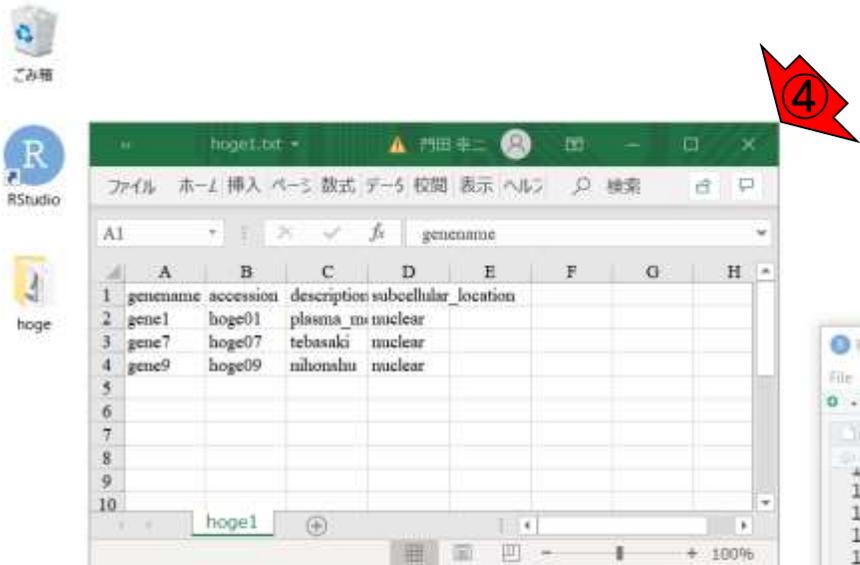
```
C:/Users/kadota/Desktop/hoge/
row.names=F)#outの中身を指定したファイル名で保存
file(file, ifelse(append, "a", "w")) でエラー:
  コネクションを開くことができません
追加情報: 警告メッセージ:
file(file, ifelse(append, "a", "w")) で:
  ファイル 'hoge1.txt' を開くことができません: Permission
denied
> |
```

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B



応用30

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge.1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。



応用31



ごみ箱



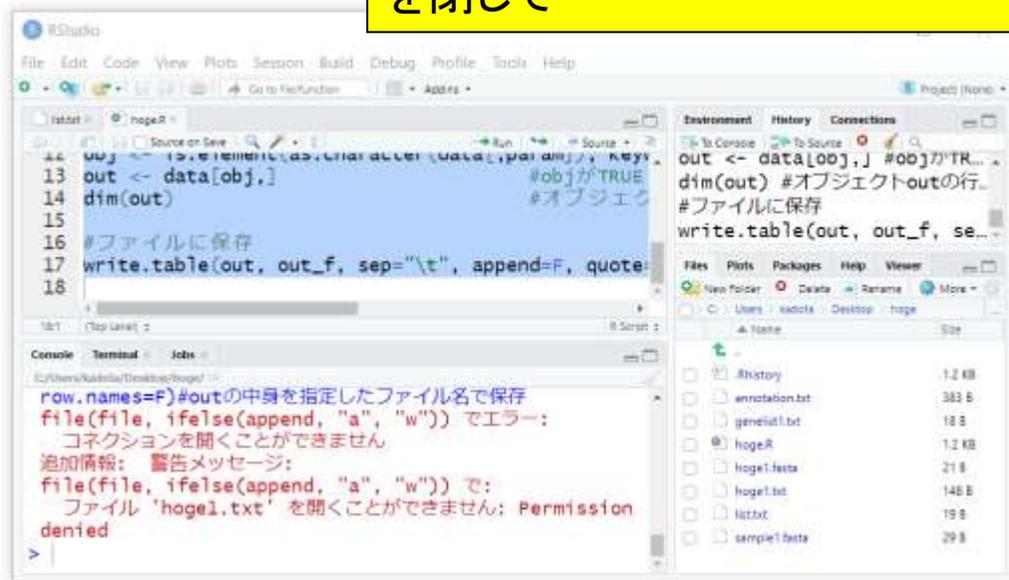
RStudio



hoge



いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge.1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。④Excelで開いているhoge1.txtを閉じて…



応用32

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。④Excelで開いているhoge1.txtを閉じて、⑤Runを押してリトライ。この作業自体はコピペ実行と同じです。

The screenshot shows the RStudio interface. The source editor contains R code for writing a table to a file. A red arrow labeled '5' points to the 'Run' button. The console shows an error message: 'file(file, ifelse(append, "a", "w")) でエラー: コネクションを開くことができません'. The file explorer on the right shows the contents of the 'hoge' directory, including 'hoge.R' and 'hoge1.txt'.

```
12 obj <- rs.element(as.character(data[,param]), K)
13 out <- data[obj,] #objがTRUE
14 dim(out) #オブジェクトの次元
15
16 #ファイルに保存
17 write.table(out, out_f, sep="\t", append=F, quote=)
18
```

Console Output:

```
C:/Users/kadota/Desktop/hoge/
row.names=F)#outの中身を指定したファイル名で保存
file(file, ifelse(append, "a", "w")) でエラー:
 コネクションを開くことができません
追加情報: 警告メッセージ:
file(file, ifelse(append, "a", "w")) で:
ファイル 'hoge1.txt' を開くことができません: Permission
denied
> |
```

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	148 B
list.txt	19 B
sample1.fasta	29 B

応用33

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge.1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。④Excelで開いているhoge1.txtを閉じて、⑤Runを押してリトライ。この作業自体はコピペ実行と同じです。実行結果。エラーは出なくなりましたね。⑥をクリックして中身をエディタで表示させてみましょう。

The screenshot shows the RStudio interface. The source editor contains the following R code:

```
12 obj <- ts.element(as.character(data[,param]), ke...
13 out <- data[obj,] #objがTRUE
14 dim(out) #オブジェクトの次元
15
16 #ファイルに保存
17 write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)
18
```

The console shows the output of the executed code:

```
> dim(out) #オブジェクトout
tの行数と列数を表示
[1] 7 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)
#outの中身を指定したファイル名で保存
>
```

On the right, a file explorer window shows the contents of the 'hoge' directory:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	265 B
list.txt	19 B
sample1.fasta	29 B

A red arrow with the number 6 points to the 'hoge1.txt' file in the file explorer.

応用34

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge.1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。④Excelで開いているhoge1.txtを閉じて、⑤Runを押してリトライ。この作業自体はコピペ実行と同じです。実行結果。エラーは出なくなりましたね。⑥をクリックして中身をエディタで表示させてみましょう。⑦表示結果。

The screenshot shows the RStudio interface. The top pane displays a script with the following content:

```
1 gene_name accession description subcellular_location
2 gene1 hoge01 plasma_mem nuclear
3 gene2 hoge02 hohinu membrane
4 gene5 hoge05 kamo membrane
5 gene7 hoge07 tebasaki nuclear
6 gene8 hoge08 biiru nuclear
7 gene9 hoge09 nihonku nuclear
```

The bottom-left pane shows the R console with the following commands and output:

```
> dim(out) #オブジェクトout
tの行数と列数を表示
[1] 7 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#outの中身を指定したファイル名で保存
>
```

The bottom-right pane shows a file explorer window with the following files:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	265 B
list.txt	19 B
sample1.fasta	29 B

Red arrows in the original image point to the Run button (7) in the RStudio toolbar and the hoge1.txt file in the file explorer (6).

応用35

いよいよ変更後のスクリプトの実行。①赤枠内のコード全体を選択。こんな感じ。この状態で②Runを押すと、出力予定ファイル(hoge1.txt)をExcelで開いている状態で実行したWinユーザは③のような感じでエラーが出ます。つまり、先に④のようにしていると、③のようになるということ。④Excelで開いているhoge1.txtを閉じて、⑤Runを押してリトライ。この作業自体はコピペ実行と同じです。実行結果。エラーは出なくなりましたね。⑥をクリックして中身をエディタで表示させてみましょう。⑦表示結果。確かに意図通り⑧4列目がnuclear or membraneとなっている行の抽出ができていますね。

The screenshot shows the RStudio interface. The script editor contains a table with 7 columns: gene name, accession, description, and subcellular location. The 4th column contains the words 'nuclear' and 'membrane'. A red arrow labeled '7' points to the 'Run' button in the toolbar. Another red arrow labeled '8' points to the 4th column of the table. The console shows the following R commands and their output:

```
> dim(out)
#オブジェクトout
tの行数と列数を表示
[1] 7 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F,
row.names=F)#outの中身を指定したファイル名で保存
>
```

The screenshot shows a file explorer window for the directory 'C:/Users/kadota/Desktop/hoge/'. The files listed are:

File Name	Size
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	265 B
list.txt	19 B
sample1.fasta	29 B

A red arrow labeled '6' points to the 'hoge1.txt' file.

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

警告メッセージ1

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains a table of gene data with columns: gene name, accession, description, and subcellular location.
- Environment/History/Connections:** Shows the execution of `write.table` and the resulting object `out`.
- Files:** Shows the file explorer with a list of files in the `hoge` directory.
- Console:** Shows the execution of `dim(out)` and `write.table` with Japanese comments. A red arrow (2) points to the 'Console' tab.

The console output is as follows:

```
> dim(out) #オブジェクトoutの行数と列数を表示
[1] 7 4
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F) #outの中身を指定したファイル名で保存
>
```

警告メッセージ2

①を押して、②Console画面を最大化。こんな感じになります。③ページ上部に移動して、コピペ実行したコードの最初のほうが見えるようにします。

The screenshot shows the RStudio interface. The console window is active, displaying the following R code and its output:

```
> dim(data) #オブジェクトdataの行数と列数を表示
[1] 11 4
>
> #本番
> obj <- is.element(as.character(data[,param]), keywords) #条件を満たすかどうかを判定した結果をobjに格納
> out <- data[obj,] #objがTRUEとなる行のみ抽出した結果をoutに格納
> dim(out) #オブジェクトoutの行数と列数を表示
[1] 7 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F) #outの中身を指定したファイル名で保存
>
```

The file explorer window on the right shows the directory structure for the current project:

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	265 B
list.txt	19 B
sample1.fasta	29 B

A red arrow labeled '3' points to the scroll bar of the file explorer window, indicating the action of scrolling to the top of the file list.

警告メッセージ3

①を押して、②Console画面を最大化。こんな感じになります。③ページ上部に移動して、コピペ実行したコードの最初のほうが見えるようにします。こんな感じ。

The screenshot shows the RStudio interface. The console on the left contains the following R code and its output:

```
> in_f1 <- "annotation.txt" #入力ファイル名
#入力ファイル名を指定してin_f1に格納(アノテーションファイル)
> in_f2 <- "list.txt" #入力ファイル名を指定
#入力ファイル名を指定してin_f2に格納(リストファイル)
> out_f <- "hoge1.txt" #出力ファイル名
#出力ファイル名を指定してout_fに格納
> param <- 4 #アノテーション
#アノテーションファイル中の検索したい列番号を指定
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
> keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
> dim(data) #オブジェクトdataの行数と列数を表示
[1] 11 4
```

The right-hand pane shows the 'Files' view of the current project directory, listing files such as `annotation.txt`, `list.txt`, and `hoge1.txt`. A red arrow with the number 3 points to the top of the file list.

警告メッセージ4

①を押して、②Console画面を最大化。こんな感じになります。③ページ上部に移動して、コピペ実行したコードの最初のほうが見えるようにします。こんな感じ。これは、④の部分で改行を入れた、⑤ list.txtを、⑥の部分で読み込ませたもの。特に何のメッセージも表示されていない。

RStudio interface showing a script execution and a file explorer. The script in the console defines variables for input files, output file, and parameters, then reads a table and lines from files. The file explorer shows a directory with files like annotation.txt, list.txt, and sample1.fasta. A red box highlights the words 'nuclear' and 'membrane' in the console output, with arrows pointing to them from the text on the right.

```
> in_f1 <- "annotation.txt" #入力ファイル名
#入力ファイル名を指定してin_f1に格納(アノテーションファイル)
> in_f2 <- "list.txt" #入力ファイル名を指定
#入力ファイル名を指定してin_f2に格納(リストファイル)
> out_f <- "hoge1.txt" #出力ファイル名
#出力ファイル名を指定してout_fに格納
> param <- 4 #アノテーション
#アノテーションファイル中の検索したい列番号を指定
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="") #in_f1で指定したファイルの読み込み
> keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
> dim(data) #オブジェクトdata
[1] 11 4
```

Name	Size
..	
.Rhistory	1.2 KB
annotation.txt	383 B
genelist1.txt	18 B
hoge.R	1.2 KB
hoge1.fasta	21 B
hoge1.txt	265 B
list.txt	19 B
sample1.fasta	29 B

警告メッセージ5

④の部分で改行を入れずに⑤list.txtで実行すると、⑥当該ファイルの読み込み部分で、警告メッセージがでます。この場合は⑧出力ファイルを眺めることで結果には影響しないことがわかる。Rは警告メッセージ後の記述内容が比較的分かりやすいのでよく読むべし。

The screenshot shows the RStudio interface with the following elements:

- Source Editor:** Contains R code for reading files and searching for keywords. Red arrows point to specific lines: ⑤ for `in_f2`, ⑥ for `readLines`, and ⑦ for the warning message.
- Console:** Shows the execution output, including a warning message: `readLines(in_f2) で: 'list.txt' で不完全な最終行が見つかりました`. A red arrow points to this message.
- Files Panel:** Shows a file explorer view of the `hoge` directory. A red arrow points to `list.txt` (⑧). Another red arrow points to the text `nuclear` in the console output (④).

```
> in_f1 <- "annotation.txt" #入力ファイル名
#入力ファイル名を指定してin_f1に格納(アノテーションファイル)
> in_f2 <- "list.txt" #入力ファイル名を指定
#入力ファイル名を指定してin_f2に格納(リストファイル)
> out_f <- "hoge1.txt" #出力ファイル名
#出力ファイル名を指定してout_fに格納
> param <- 4 #アノテーション
#アノテーションファイル中の検索したい列番号を指定
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#in_f1で指定したファイルの読み込み
> keywords <- readLines(in_f2) #in_f2で指定したファイルの読み込み
警告メッセージ:
readLines(in_f2) で: 'list.txt' で不完全な最終行が見つかりました
dim(data) #オブジェクトの
```

Contents

- Introduction、(Rで)塩基配列解析、(Rで)塩基配列解析のサブ
- 基本的な利用法、RStudioの起動、見栄えの統一、終了
- Console画面の見栄え、設定変更、枠の幅、電卓的な利用
- プロンプト(>)、Console画面クリア、計算の中断
- 翻訳配列取得
 - 項目と例題、入力ファイルの保存、作業ディレクトリの変更
 - 入力ファイルの確認、FASTA形式、コピー実行、出力ファイルの確認
- フィルタリング(任意のキーワードを含む行を抽出)
 - 全体像の把握、入力ファイルの保存、コピー実行、出力ファイルの確認
 - Winユーザ向け注意点、応用
- 警告メッセージ、ありがちなミス

ありがちなミス0

この後エラーを意図的に出すので、まずはうまくいく例を必要最小限のコードで示す。

The screenshot shows the RStudio interface. The console pane on the left displays the following R code and its output:

```
> in_f1 <- "annotation.txt" #
> in_f2 <- "genelist1.txt" #
> out_f <- "hoge1.txt" #
> param <- 1 #
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#
> keywords <- readLines(in_f2) #
> dim(data) #
[1] 11 4
>
> #本番
> obj <- is.element(as.character(data[,param]), keywords)#
> out <- data[obj,] #
> dim(out) #
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#
> |
```

The environment pane on the right shows the following objects:

```
obj <- is.element(as.char...
out <- data[obj,] #
dim(out) #
#ファイルに保存
```

The Files pane at the bottom right shows the directory structure:

Name	Size
..	
annotation.txt	383 B
genelist1.txt	18 B

ありがちなミス1

①作業ディレクトリの変更を忘れているため、②in_f1で指定した最初のファイルの読み込み段階で、③エラーが出る。作業ディレクトリ内には入力ファイルは存在しないので当然。

The screenshot shows the RStudio interface. The console window displays the following R code and its output:

```
> in_f1 <- "annotation.txt" #
> in_f2 <- "genelist1.txt" #
> out_f <- "hogel.txt" #
> param <- 1 #
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#
file(file, "rt") でエラー: コネクションを開くことができません
追加情報: 警告メッセージ:
file(file, "rt") で:
ファイル 'annotation.txt' を開くことができません: No such file or directory
> keywords <- readLines(in_f2) #
file(con, "r") でエラー: コネクションを開くことができません
追加情報: 警告メッセージ:
file(con, "r") で:
ファイル 'genelist1.txt' を開くことができません: No such file or directory
> dim(data) #
[1] 11 4
>
```

Red arrows point to the following lines in the console:

- Arrow ② points to the line `in_f1 <- "annotation.txt"`.
- Arrow ③ points to the line `#入力ファイルの読み込み`.

The Environment pane on the right shows the following R code:

```
out <- data[obj,] #
dim(out) #
#ファイルに保存
write.table(out, out_f, s...
```

The Files pane on the right shows the file explorer view, with a red arrow ① pointing to the Home directory. The file list includes:

- 2017
- 2018
- 2019
- 2020
- html
- Officeのカスタムテンプレート
- Outlook ファイル
- paper
- public_html

ありがちなミス2

①作業ディレクトリの変更はできているが、②in_f2で指定した最初のファイルが③存在しないので、④その読み込み段階でエラーが出る。

The screenshot shows the RStudio interface. The console window displays the following R code and its output:

```
> in_f1 <- "annotation.txt" #
> in_f2 <- "genelist1.txt" #
> out_f <- "hoge1.txt" #
> param <- 1 #
>
> #入力ファイルの読み込み
> data <- read.table(in_f1, header=TRUE, sep="\t", quote="")#
> keywords <- readLines(in_f2) #
file(con, "r") でエラー: コネクションを開くことができません
追加情報: 警告メッセージ:
file(con, "r") で:
ファイル 'genelist1.txt' を開くことができません: No such file or direc
tory
> dim(data) #
[1] 11 4
>
> #本番
> obj <- is.element(as.character(data[,param]), keywords)#
> out <- data[obj,] #
> dim(out) #
[1] 3 4
```

The file explorer on the right shows the directory structure: C:\Users\kadota\Desktop\hoge. It contains a subdirectory '..' and a file 'annotation.txt'. The file 'genelist1.txt' is missing, which is the cause of the error.

Red arrows with circled numbers point to the following elements:

- ②: The variable `in_f2` in the R code.
- ④: The `readLines(in_f2)` command in the R code.
- ③: The file `annotation.txt` in the file explorer.
- ①: The file explorer window title bar.

ありがちなミス3

①出力予定ファイルが既に作業ディレクトリ内に存在し、それを②Excelなど別のプログラムで開いているWinユーザは、③最後のwrite.table関数のところでエラーが出る。

The screenshot shows the RStudio interface with a file named 'hoge1.txt' open in Excel. The R console shows the following code and output:

```
> #入力ファイルの読み込み
> data <- read.table(in_f
> keywords <- readLines(i
> dim(data)
[1] 11 4
>
> #本番
> obj <- is.element(as.ch
> out <- data[obj,]
> dim(out)
[1] 3 4
>
> #ファイルに保存
> write.table(out, out_f, sep="\t", append=F, quote=F, row.names=F)#
file(file, ifelse(append, "a", "w")) でエラー:
コネクションを開くことができません
追加情報: 警告メッセージ:
file(file, ifelse(append, "a", "w")) で:
ファイル 'hoge1.txt' を開くことができません: Permission denied
>
```

The Excel spreadsheet shows the following data:

gene	accession	description	subcellular	location
gene1	hoge01	plasma_m	nuclear	
gene2	hoge02	hohinu	membrane	
gene5	hoge05	kamo	membrane	
gene7	hoge07	tebasaki	nuclear	
gene8	hoge08	biiru	nuclear	
gene9	hoge09	nihonshu	nuclear	
gene10	hoge10	agenel	membrane	

The RStudio console shows the error message: "file(file, ifelse(append, "a", "w")) でエラー: コネクションを開くことができません" and "追加情報: 警告メッセージ: file(file, ifelse(append, "a", "w")) で: ファイル 'hoge1.txt' を開くことができません: Permission denied".