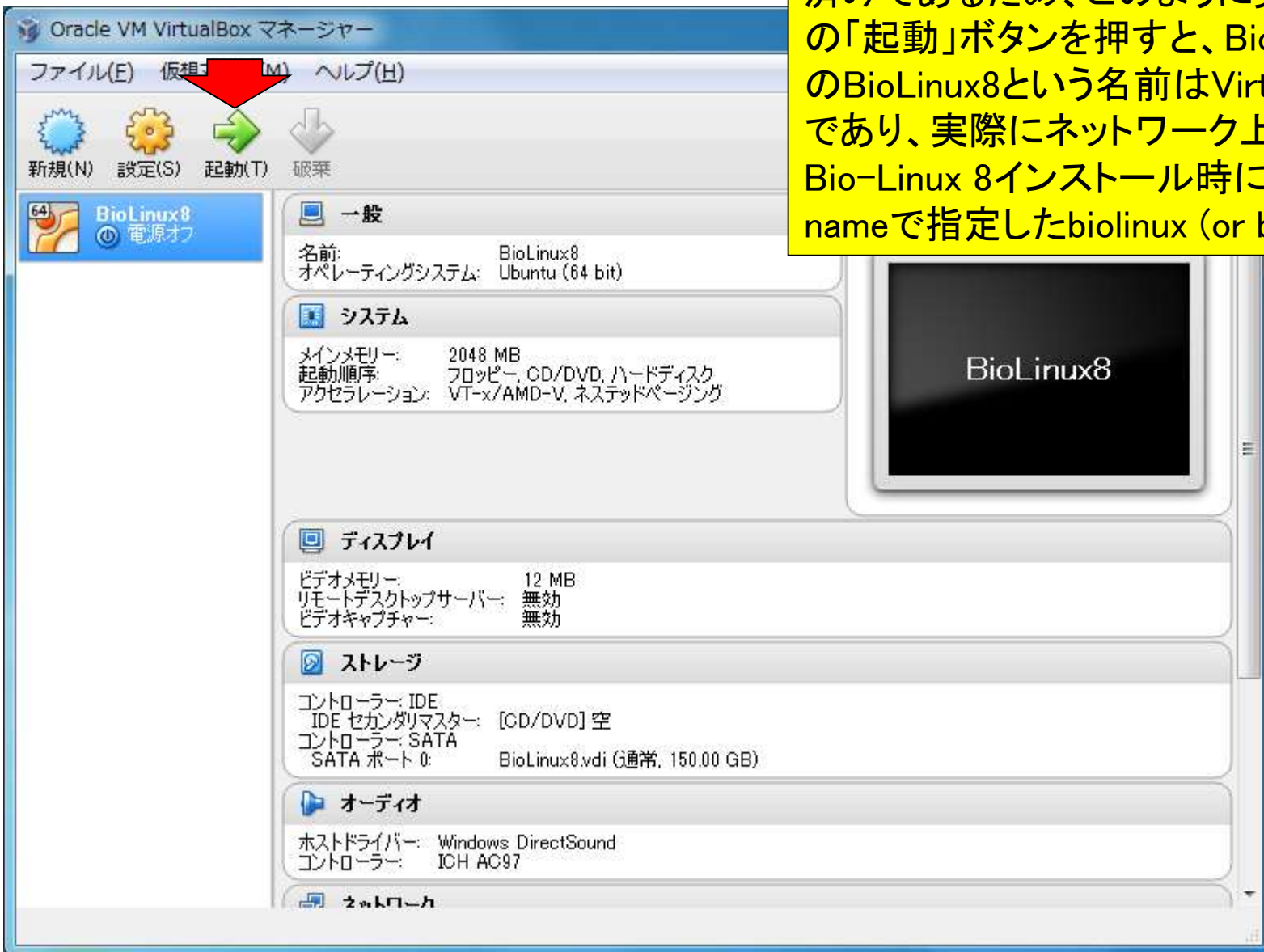


次世代シーケンサーデータの解析手法 第3回Linux環境構築からNGSデータ取得まで W(Windows版)

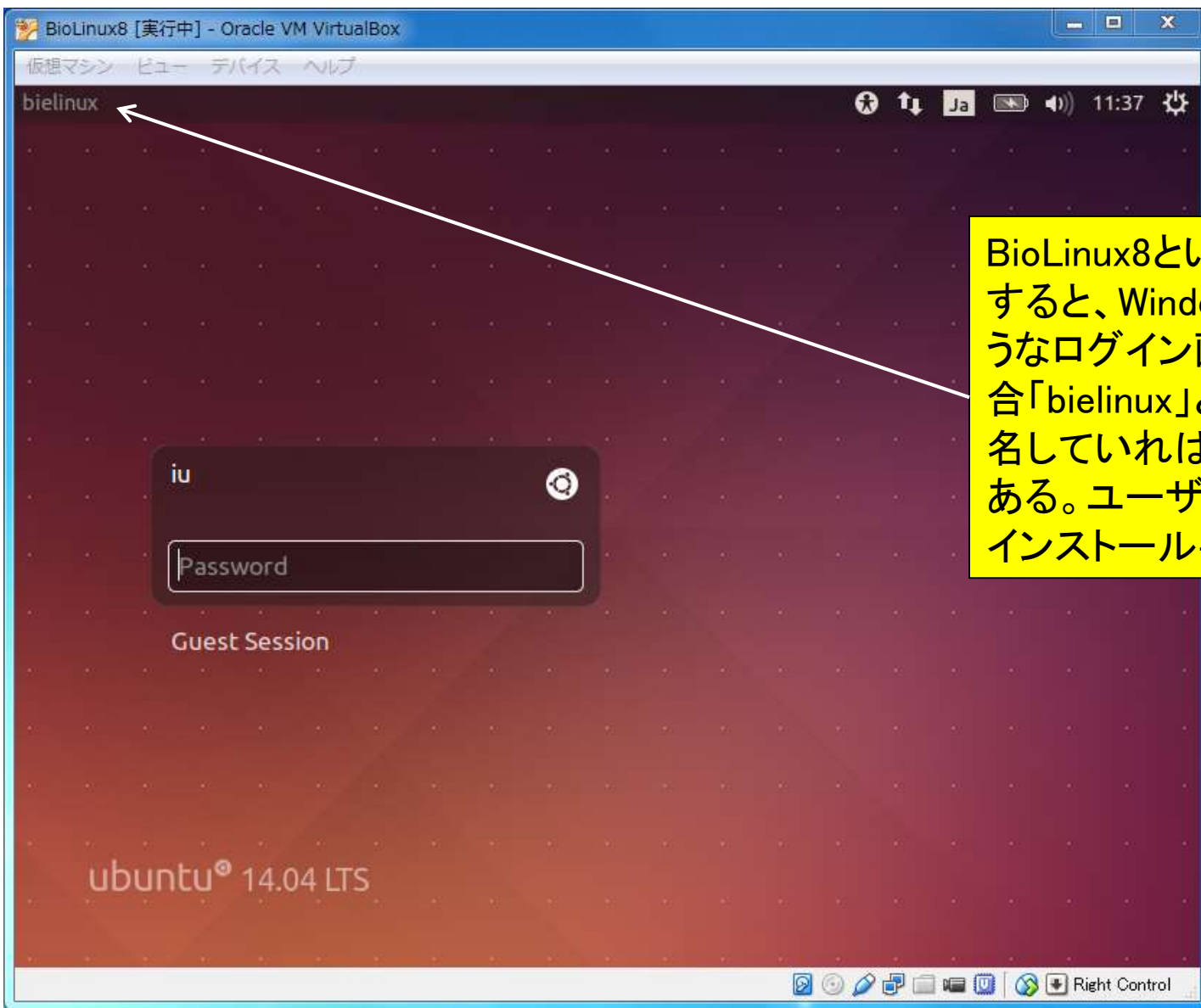
東京大学・大学院農学生命科学研究科
孫 建強、三浦 文、清水 謙多郎、門田 幸二
kadota@bi.a.u-tokyo.ac.jp
<http://www.iu.a.u-tokyo.ac.jp/~kadota/>

ウェブ資料1 (W1)

仮想化ソフトVirtualBox起動直後の画面。BioLinux8という名前の仮想PC本体(仮想マシン環境)を導入済みであるため、このように見られる。赤矢印部分の「起動」ボタンを押すと、BioLinux8が起動する。このBioLinux8という名前はVirtualBox内での識別用であり、実際にネットワーク上で見られる名前はBio-Linux 8インストール時にYour computer's nameで指定したbiolinux (or bielinux)である。



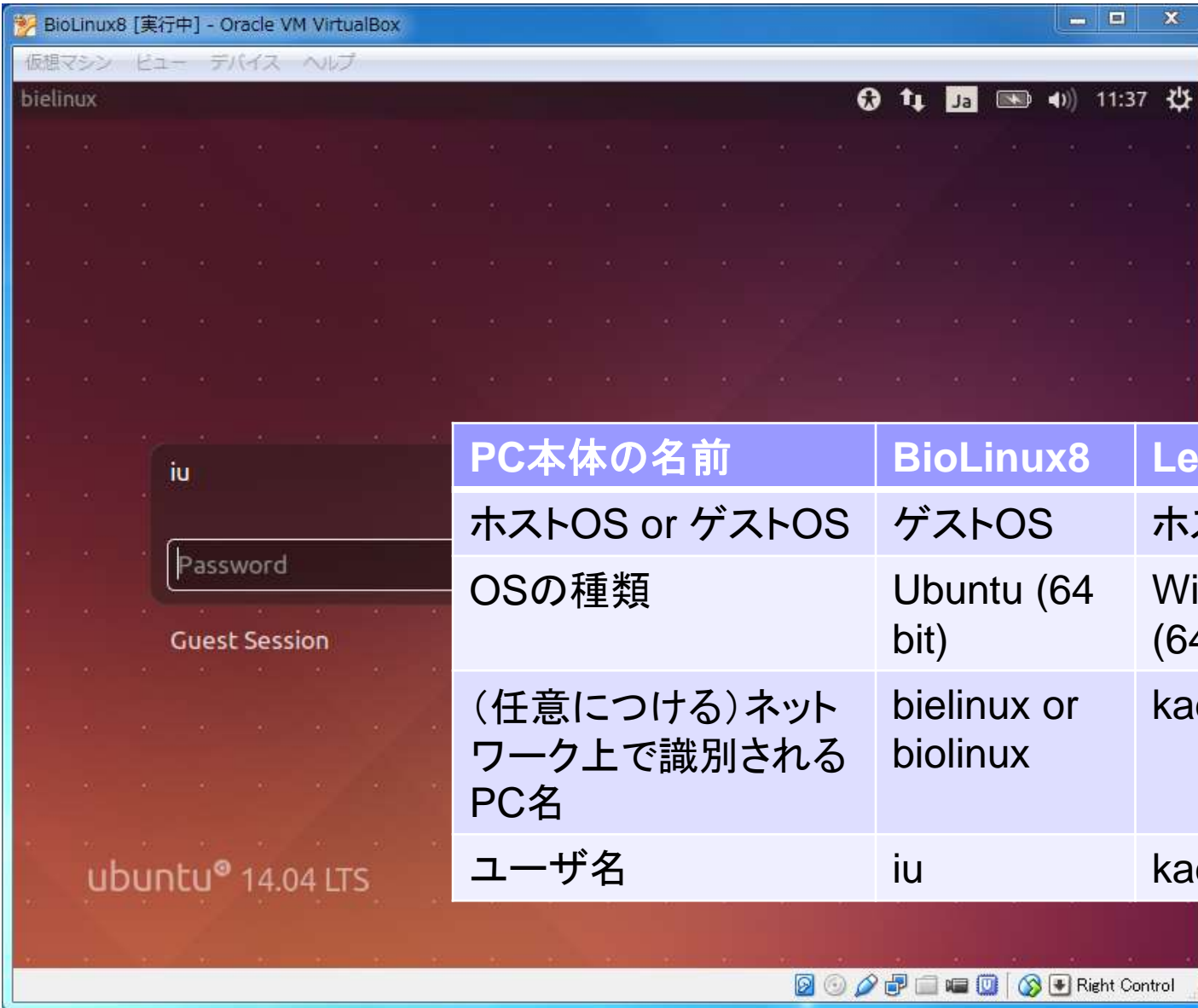
ウェブ資料2-1 (W2-1)



BioLinux8という名前の仮想PC本体を起動すると、WindowsやMacintoshで見られるようなログイン画面が現れる。PC名はこの場合「bielinux」となっているが、「biolinux」と命名していればそのようになっているはずである。ユーザ名はiu。ログインパスワードはインストール手順通りだとpass1409となる。

W2-2

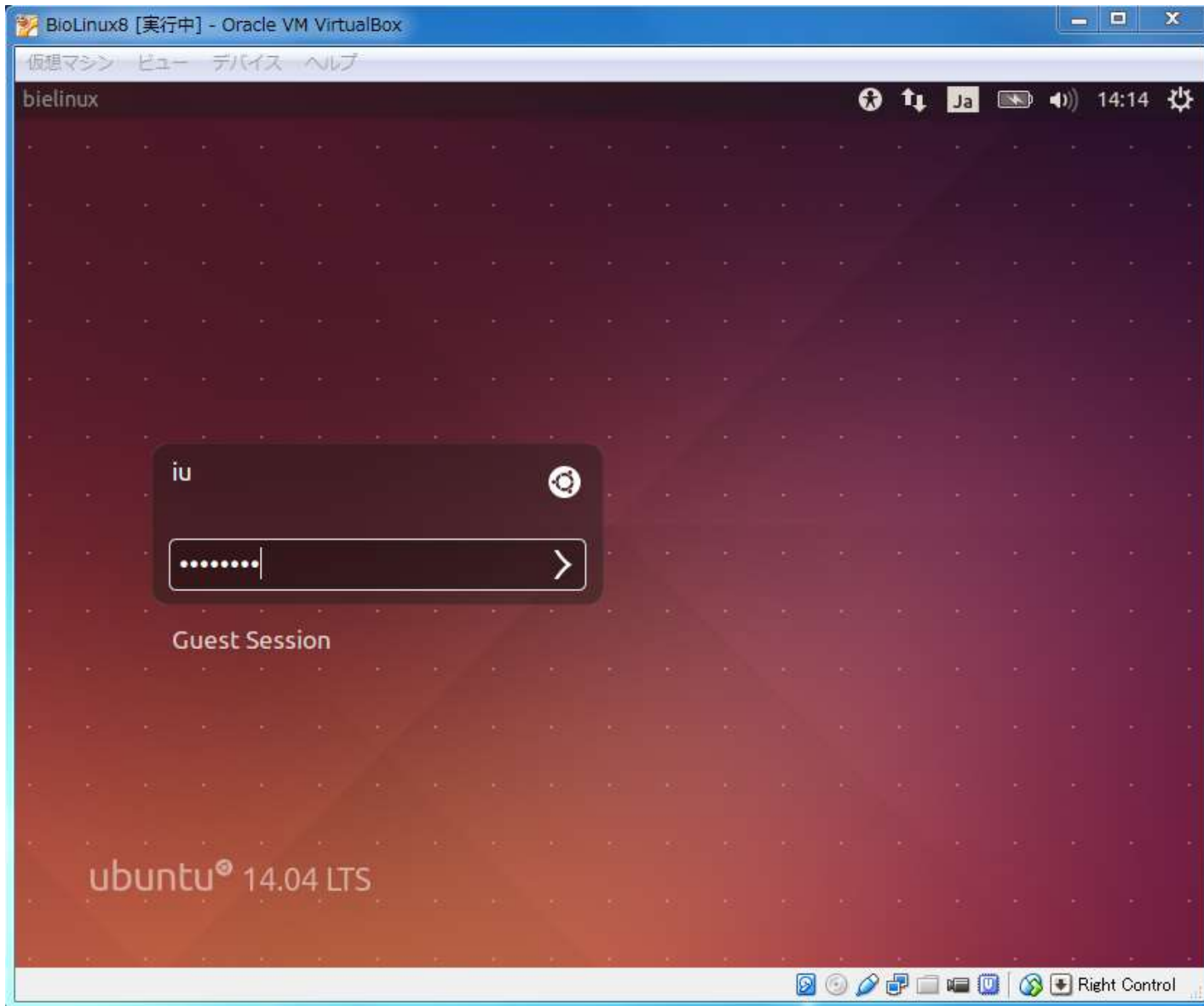
連載第2回で用いたWindows PC (Panasonic Let's note)とMacintosh PC (MacBook)との対応関係としては、概ね以下のように解釈すればよい。



PC本体の名前	BioLinux8	Let's note	MacBook
ホストOS or ゲストOS	ゲストOS	ホストOS	ホストOS
OSの種類	Ubuntu (64 bit)	Windows 7 (64 bit)	Mac OS X (32 bit)
(任意につける)ネットワーク上で識別されるPC名	bielinux or biolinux	kadota-pc	agribio-macbook
ユーザ名	iu	kadota	kadota

W2-3

ログインパスワードを打ち込んで「Enter」キーを押す。



W2-4

ホストOSがWindows環境下での
Bio-Linux 8起動直後の状態。



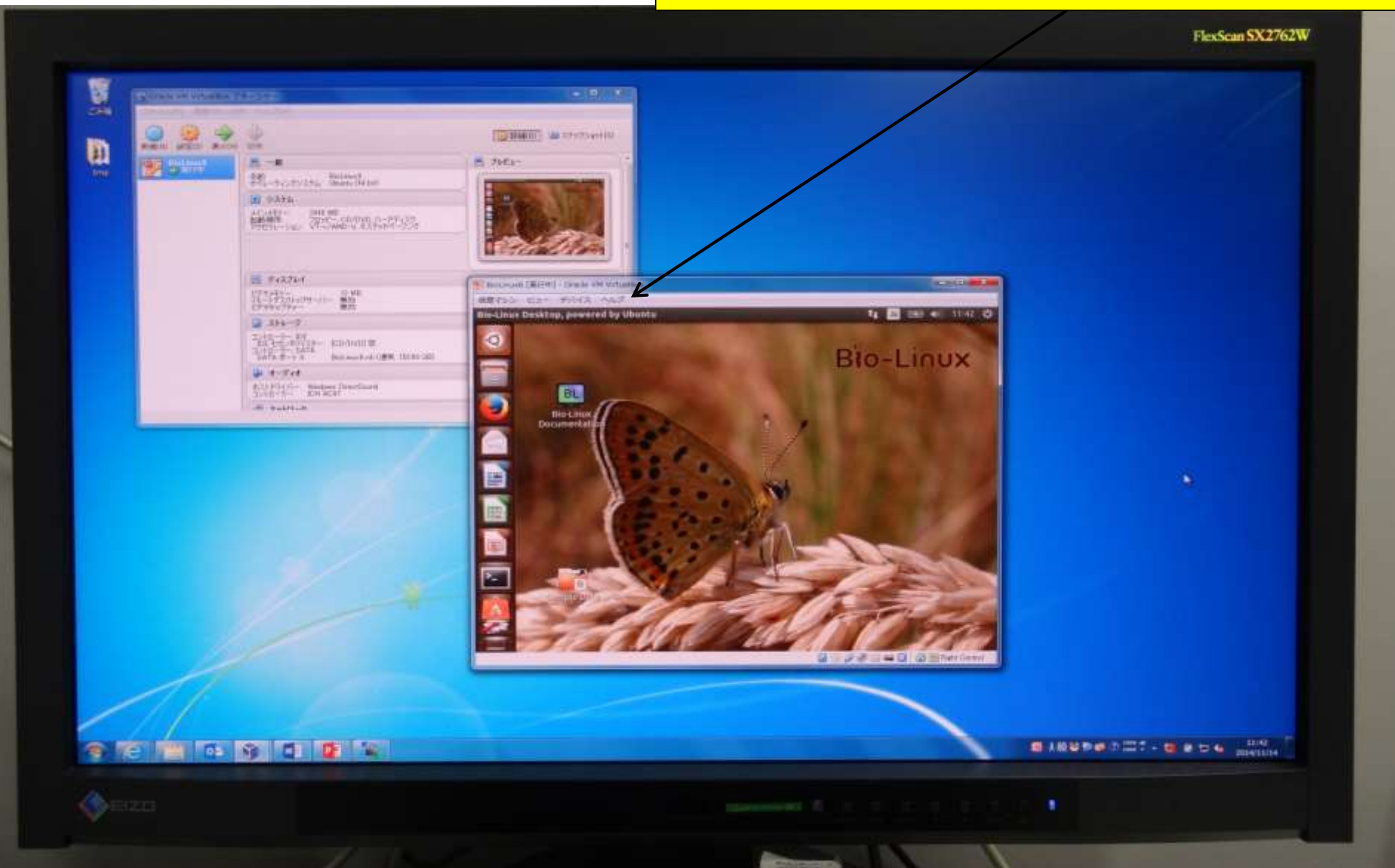
W3-1

点線の赤枠部分に相当するBioLinux8のウィンドウがアクティブでない場合に、ウェブ資料2-4のスクリーンショットを得ることができる。赤枠の範囲外でマウスクリックした状態であるため、BioLinux8ウィンドウのメニューバーが灰色になっている(つまり非アクティブ)になっていることが分かる。



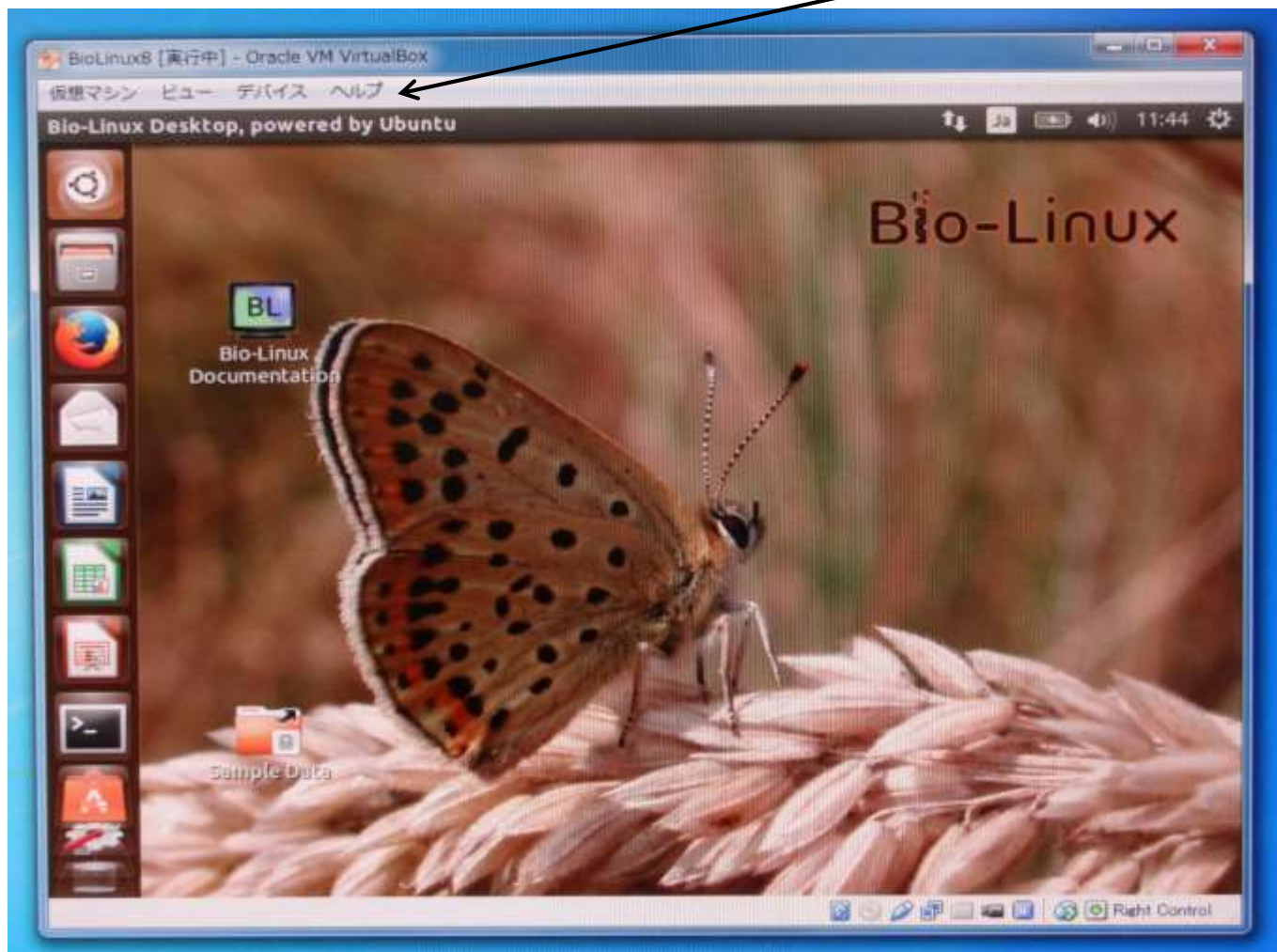
W3-2

BioLinux8のウィンドウがアクティブな状態だとメニューバーが黒色になる。この状態ではウェブ資料2-4のようなスクリーンショットを得られない。デジカメで撮影。



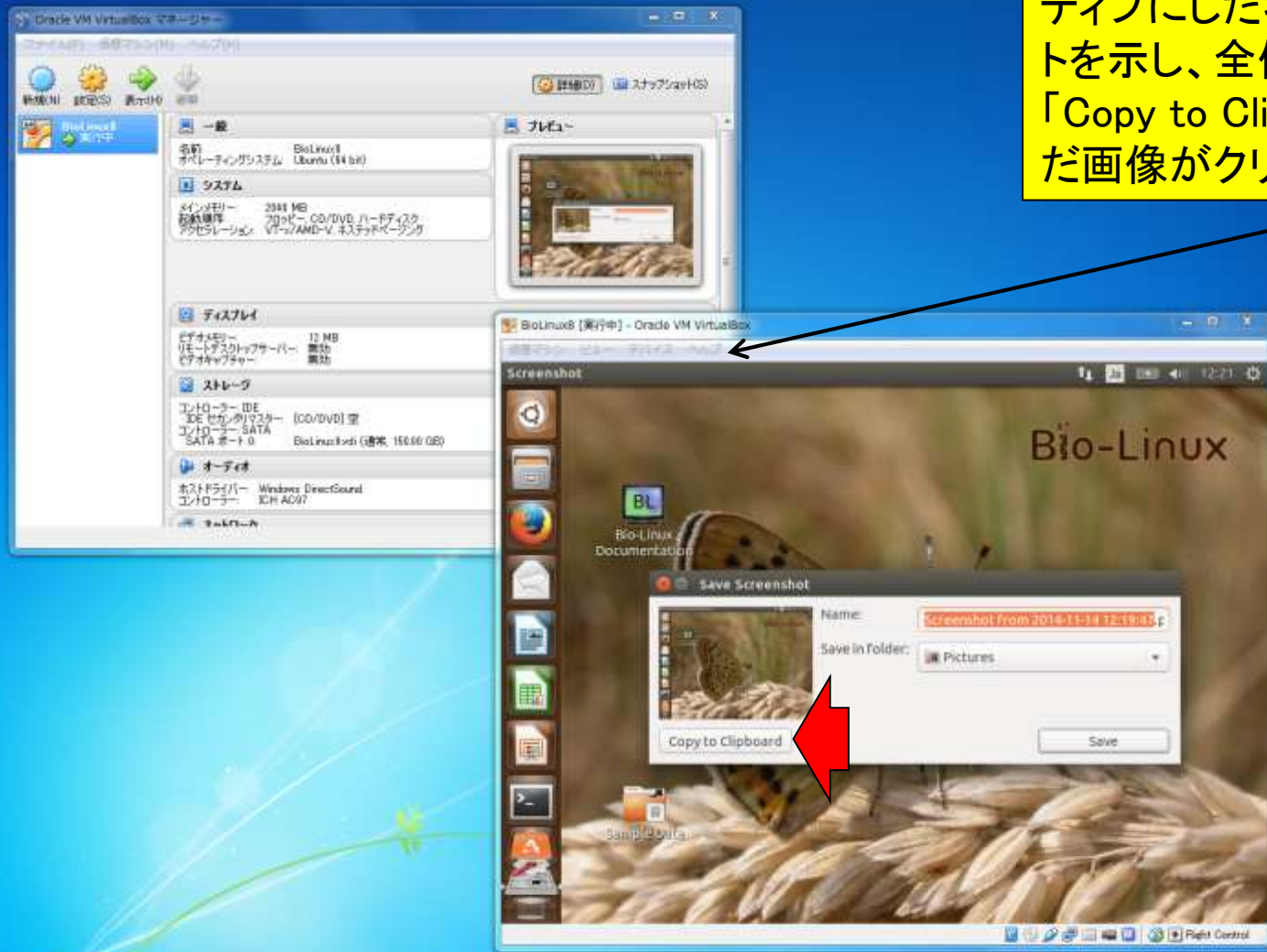
W3-3

BioLinux8ウィンドウがアクティブ状態の拡大図。



W4-1

BioLinux8ウィンドウ(ゲストOS)がアクティブな状態でPrint Screenボタンを押した結果。BioLinux8ウィンドウを非アクティブにした状態で得たスクリーンショットを示し、全体像を俯瞰している。「Copy to Clipboard」を押すと、取り込んだ画像がクリップボードにコピーされる



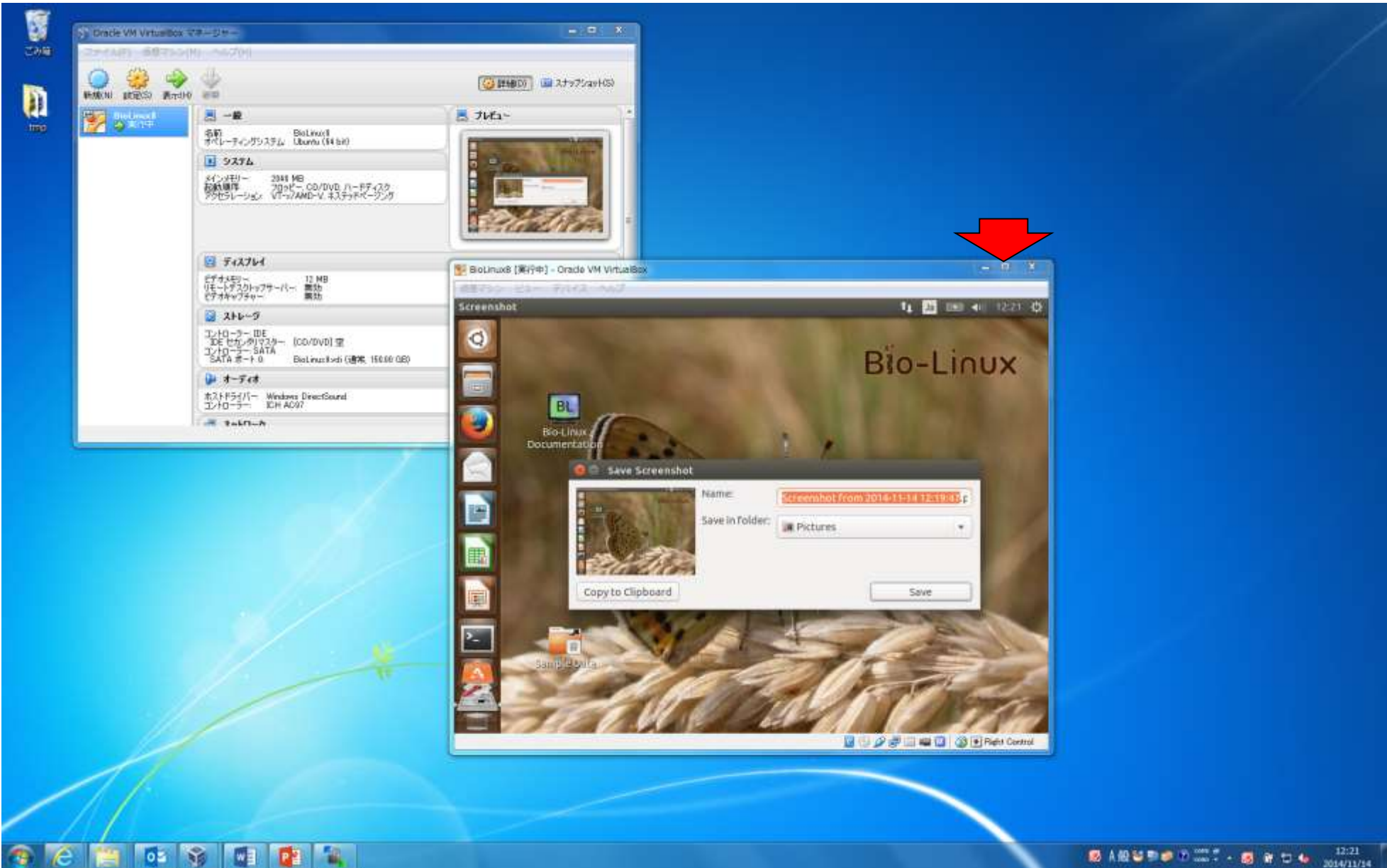
W4-2

クリップボードにコピーされた画像（取り込んだゲストOSのPC画面）。



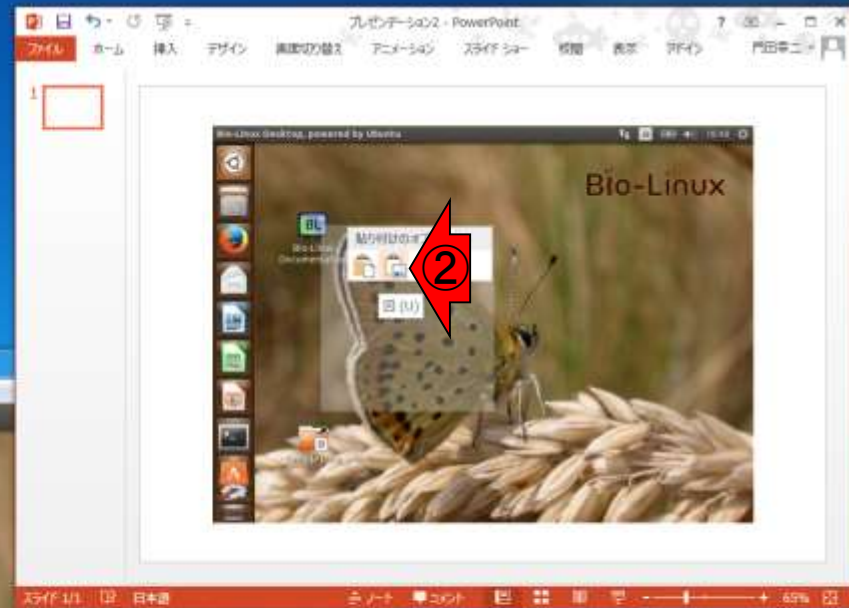
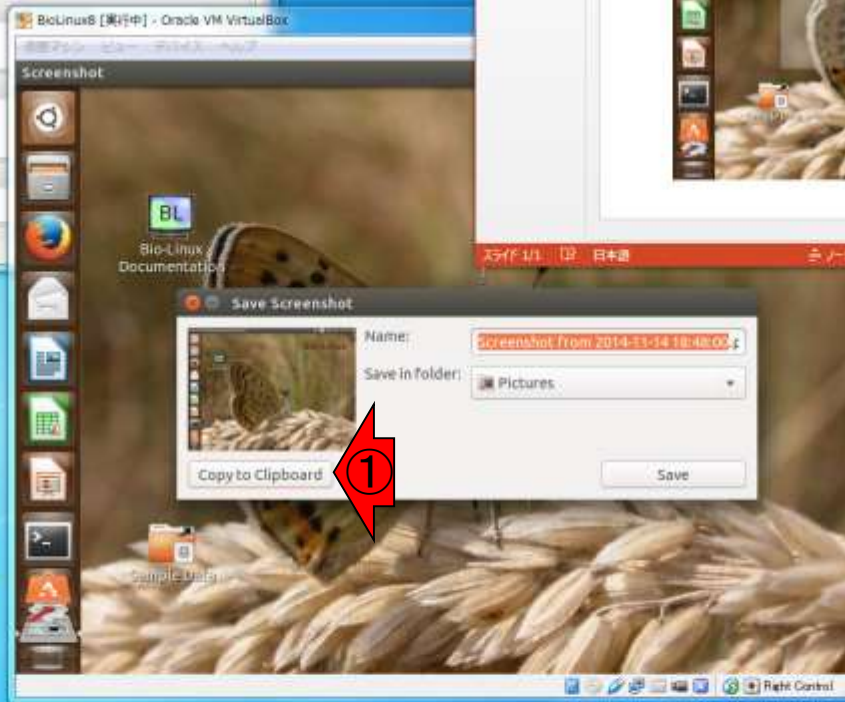
W4-3

画面サイズを最大にすれば、
ほぼ完全なLinux環境となる。



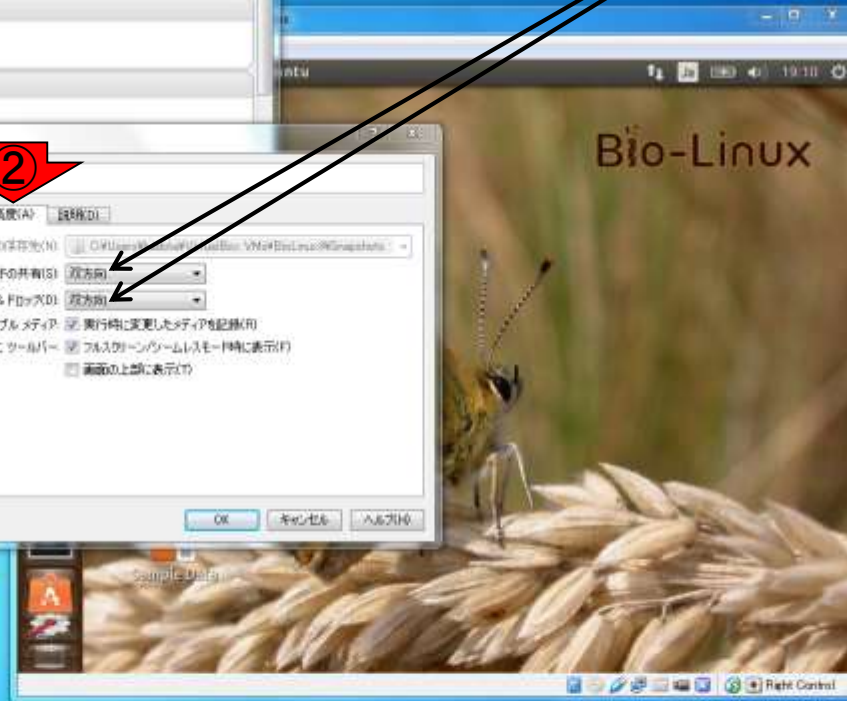
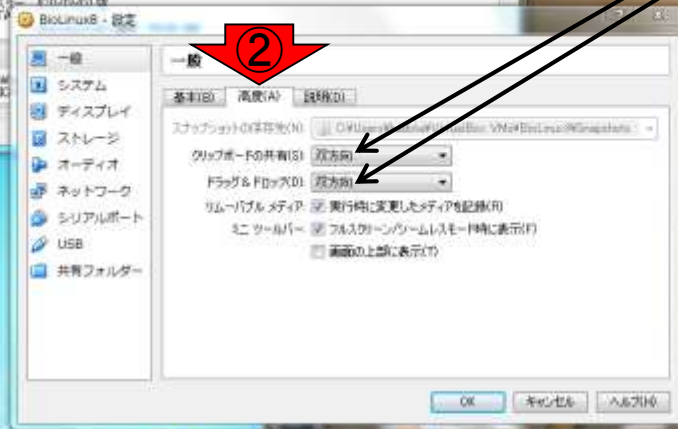
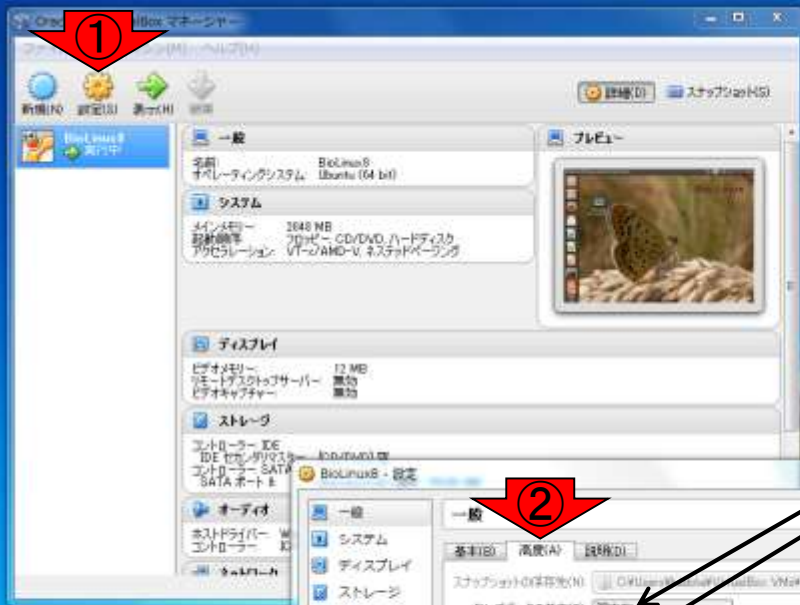
W5-1

ゲストOSのスクリーンショットをクリップボードにコピーし、
ホストOSのMicrosoft PowerPointに張りつけているところ。



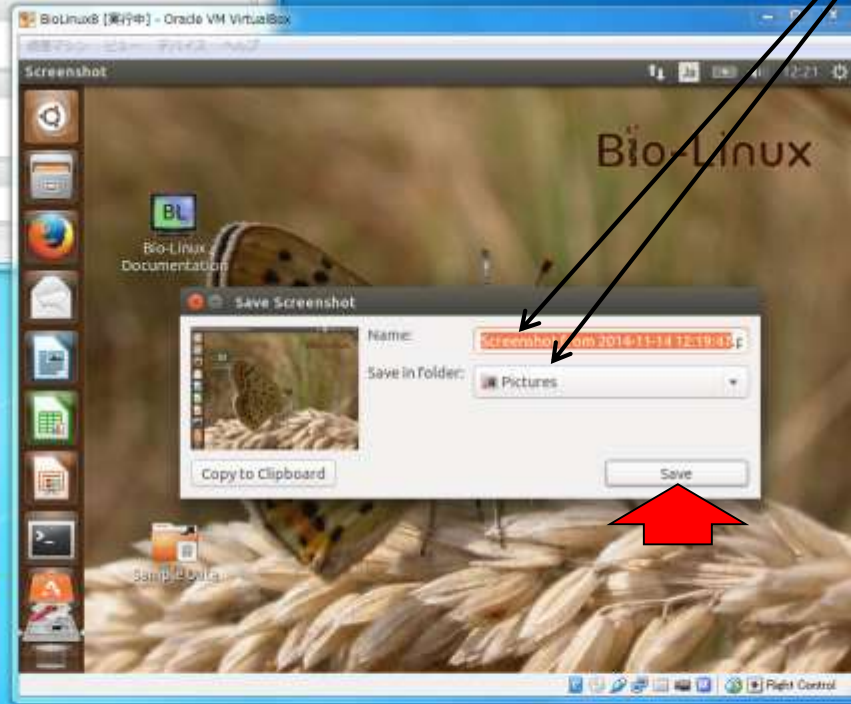
W5-2

ゲストOSのクリップボード情報をホストOSで利用できたのは、デフォルトでは無効になっていた「クリップボードの共有」と「ドラッグ&ドロップ」を、Bio-Linux 8のインストール時に「双方向」に変更していたから。



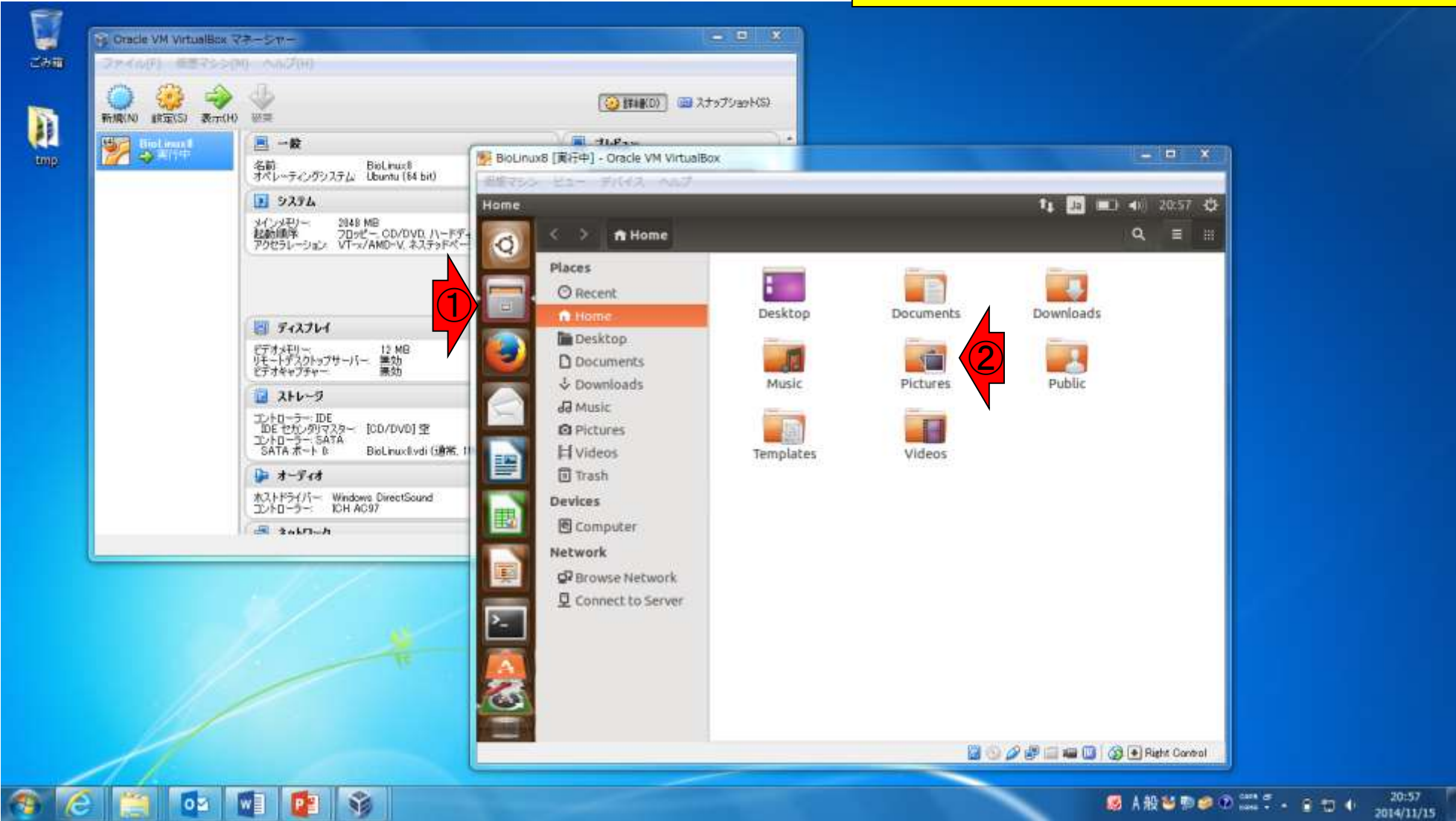
W6-1

Saveボタンを押すと、Nameに記載されているファイル名で、Save in folderに記載されているフォルダ(つまりPicturesフォルダ)に保存される。反転されているファイル名の右のほうをみてみると最後が.pngとなっており、PNG形式で保存されることが分かる。



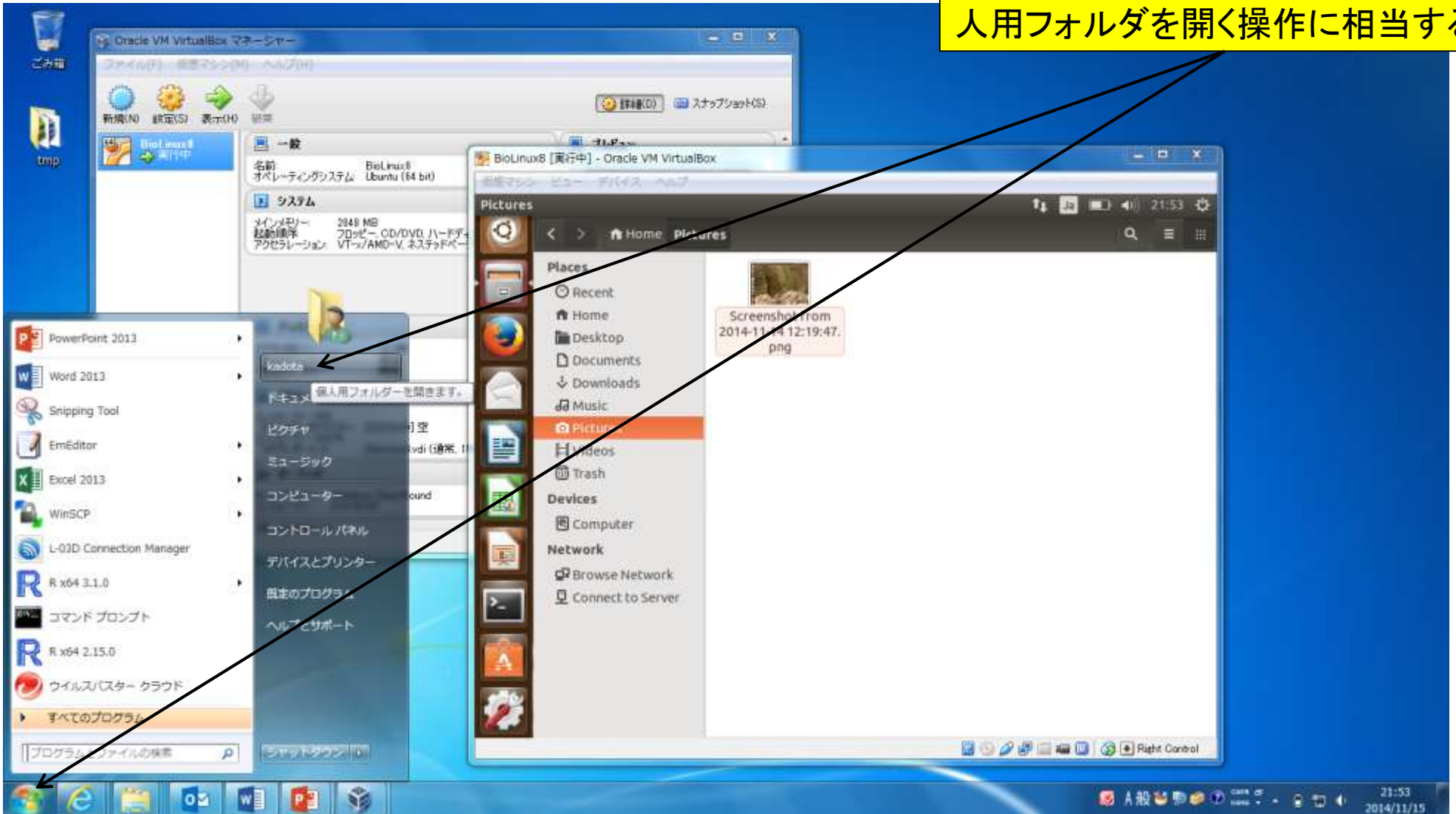
W6-2

①の引き出しみたいなのアイコンをクリックし、
②Picturesフォルダをクリックすると、保存したpngファイルを見ることができる。



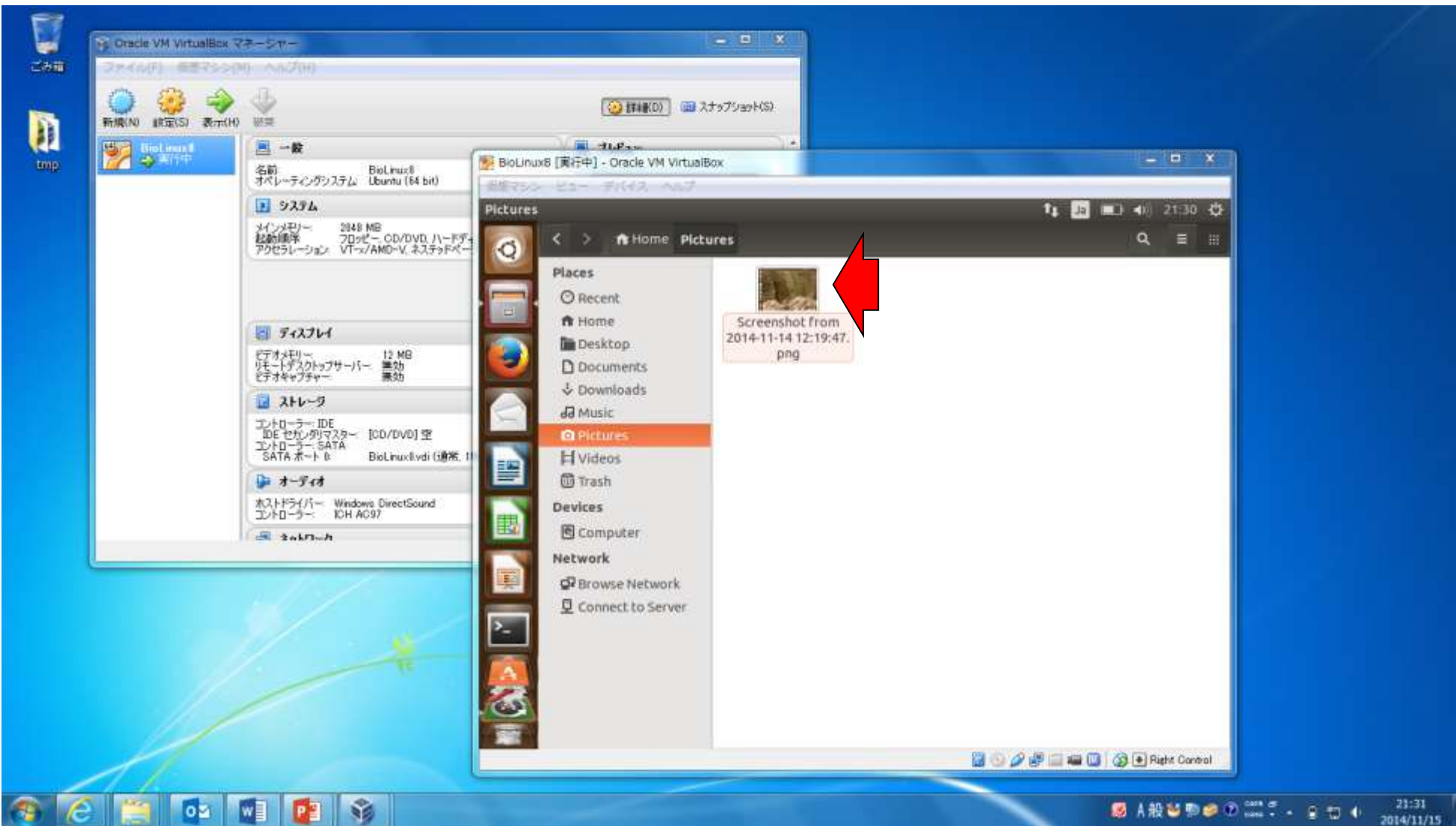
W6-2

①は、ホストOS (Windows 7) の場合、左下のスタートメニューを押して、ログインユーザ(この場合kadota)の個人用フォルダを開く操作に相当する。



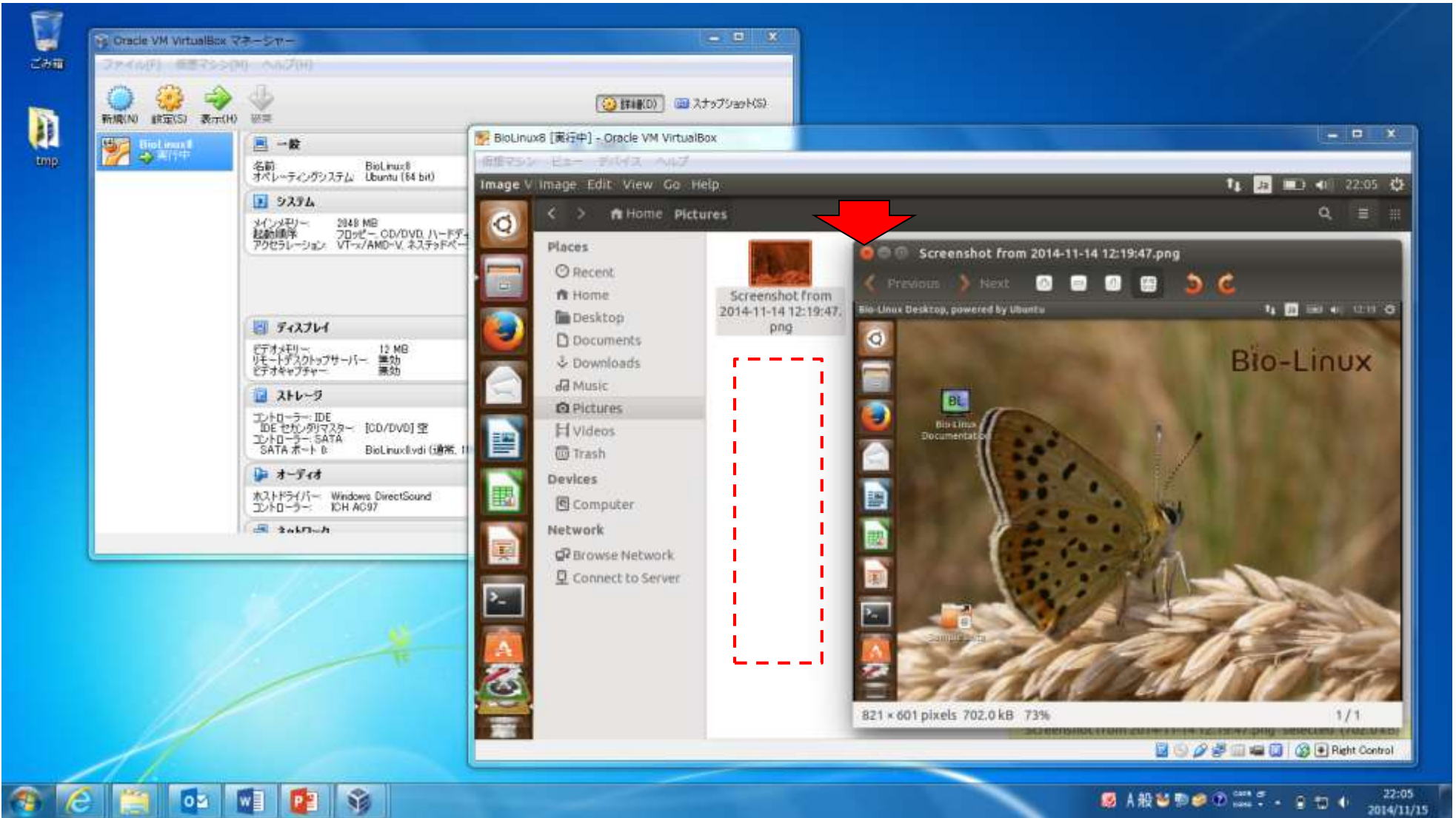
ダブルクリックでpngファイルを開くことができる。

W6-2



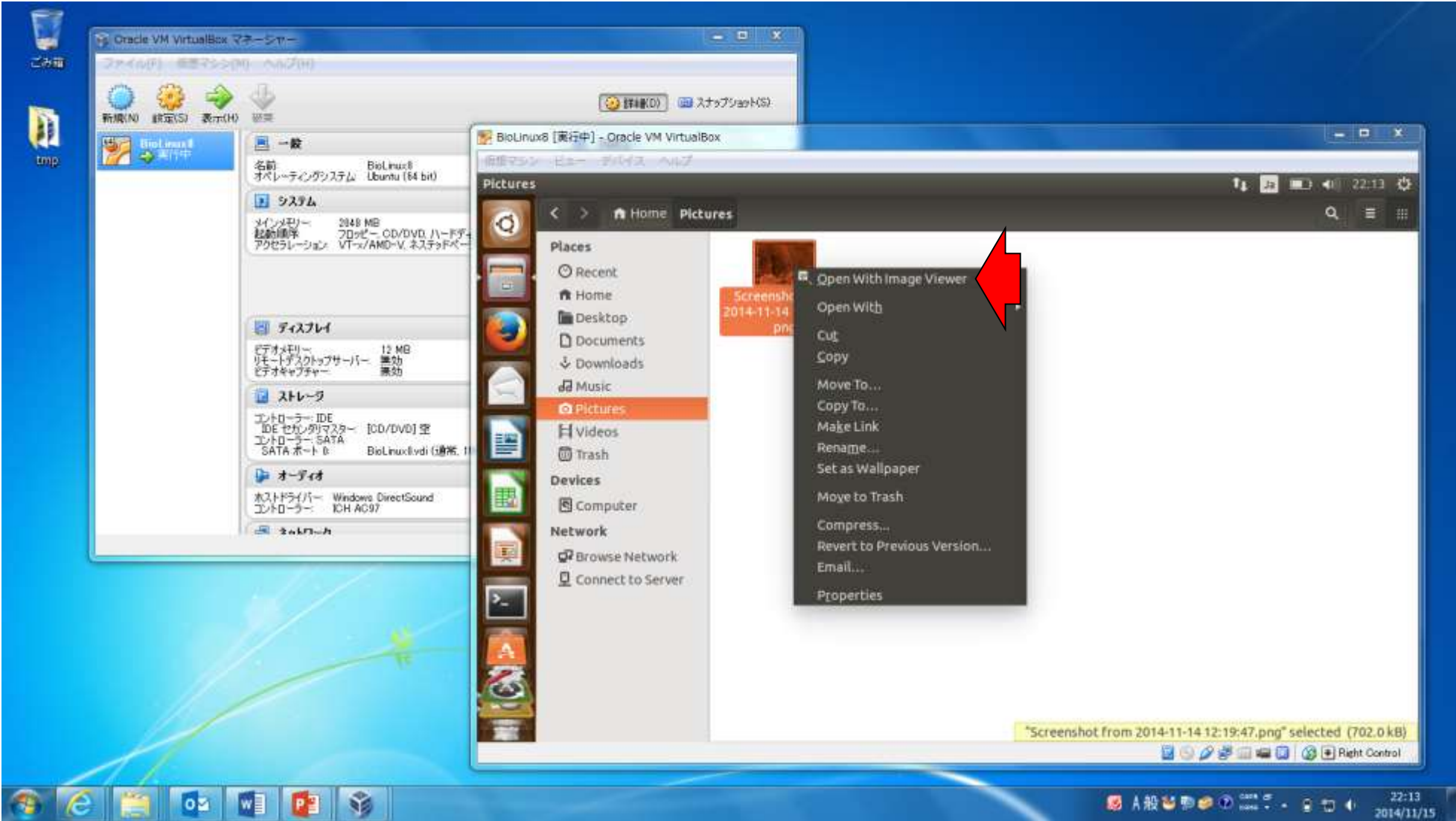
W6-2

× ボタンを押すと閉じることができるが、
やらずに赤枠あたりでクリックすると…。



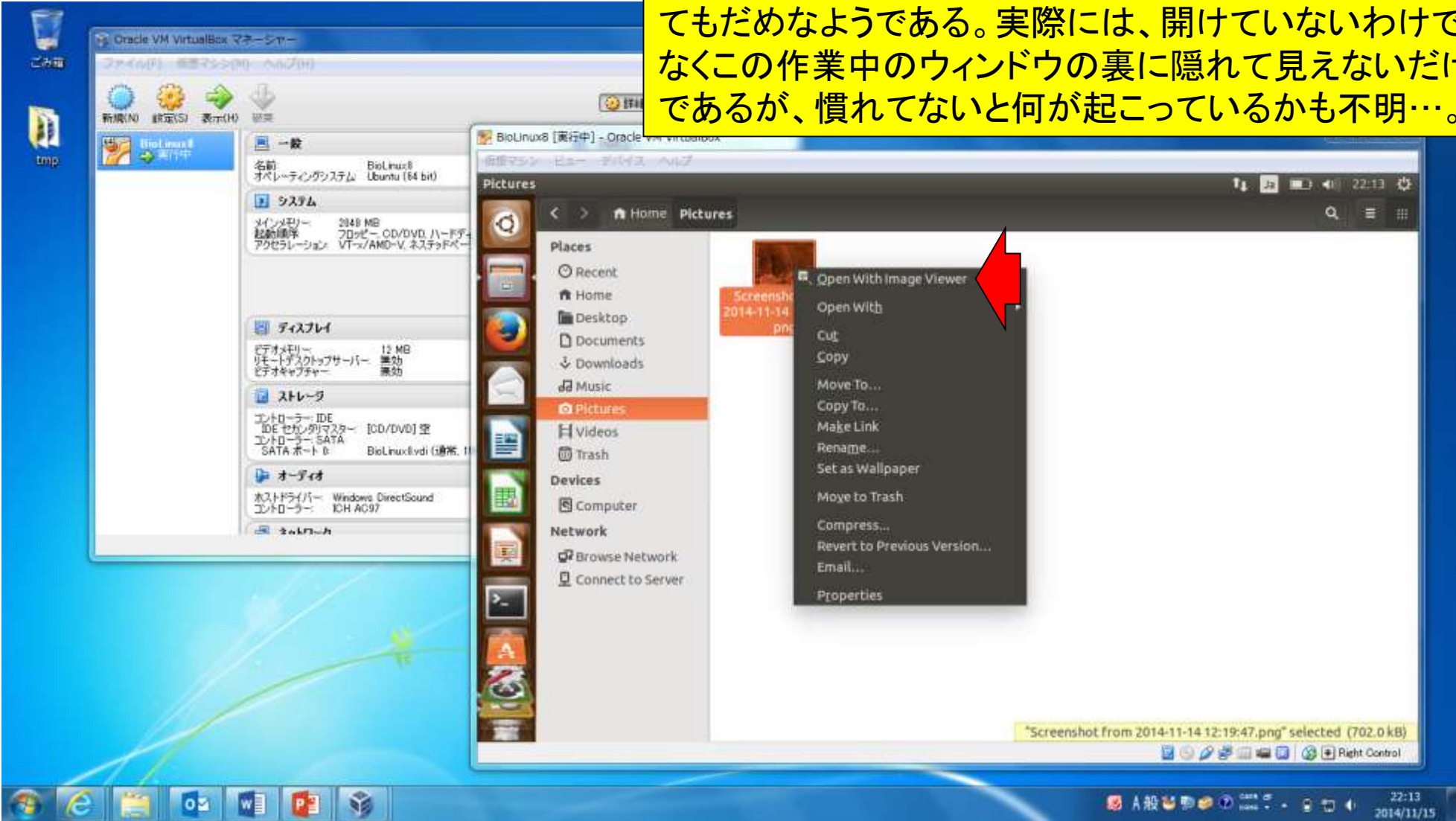
開いていたpngファイルが消える(ように見えて、実はPicturesフォルダの裏に隠れているだけ)。

W6-2



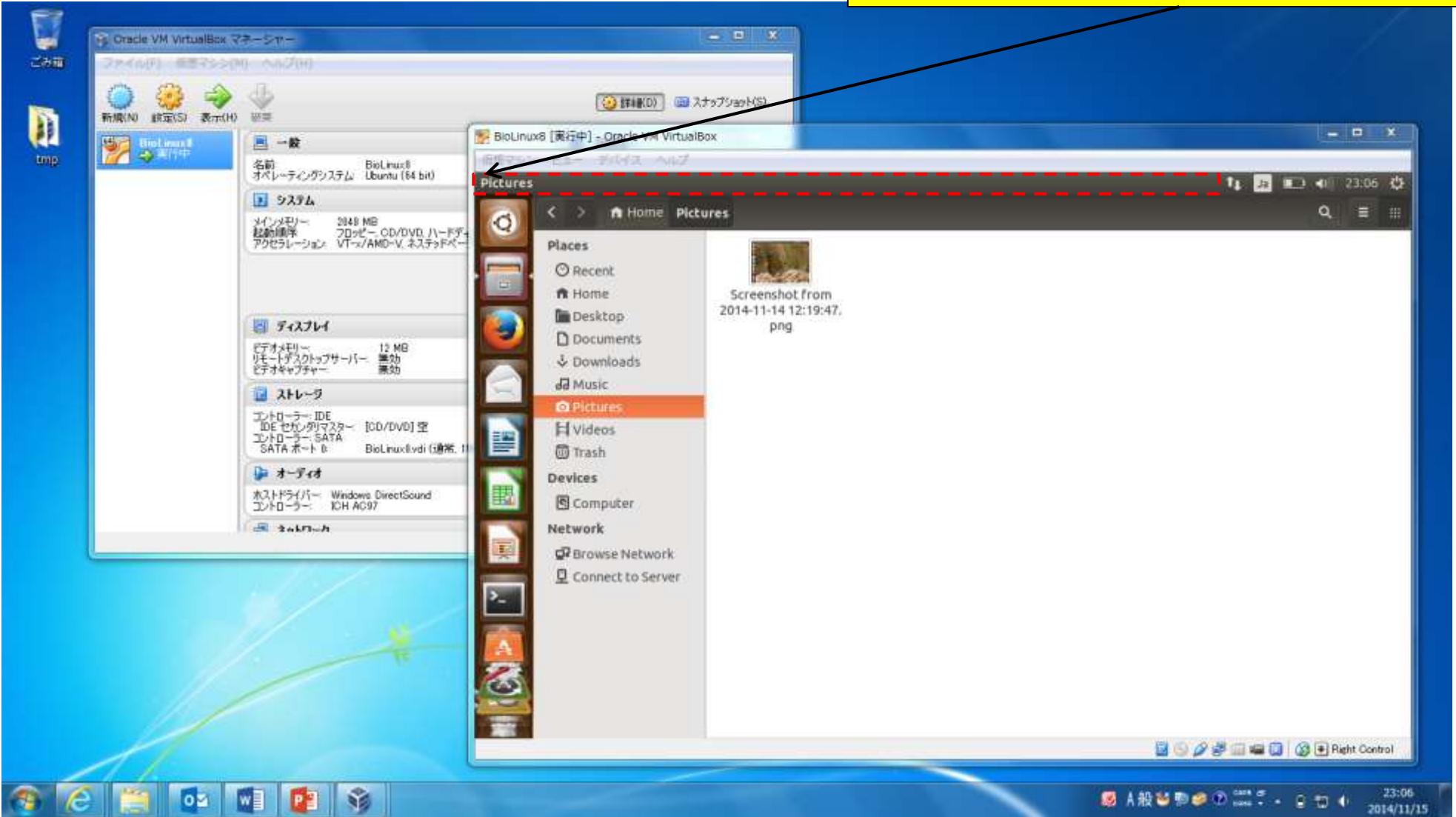
W6-3

もう一度ダブルクリックでpngファイルを開こうとしても開けないことがある。こうなると、例えば右クリックでOpen with Image Viewerを選択するなど、他の手段で開こうとしてもだめなようである。実際には、開けていないわけではなくこの作業中のウィンドウの裏に隠れて見えないだけであるが、慣れてないと何が起きているかも不明…。



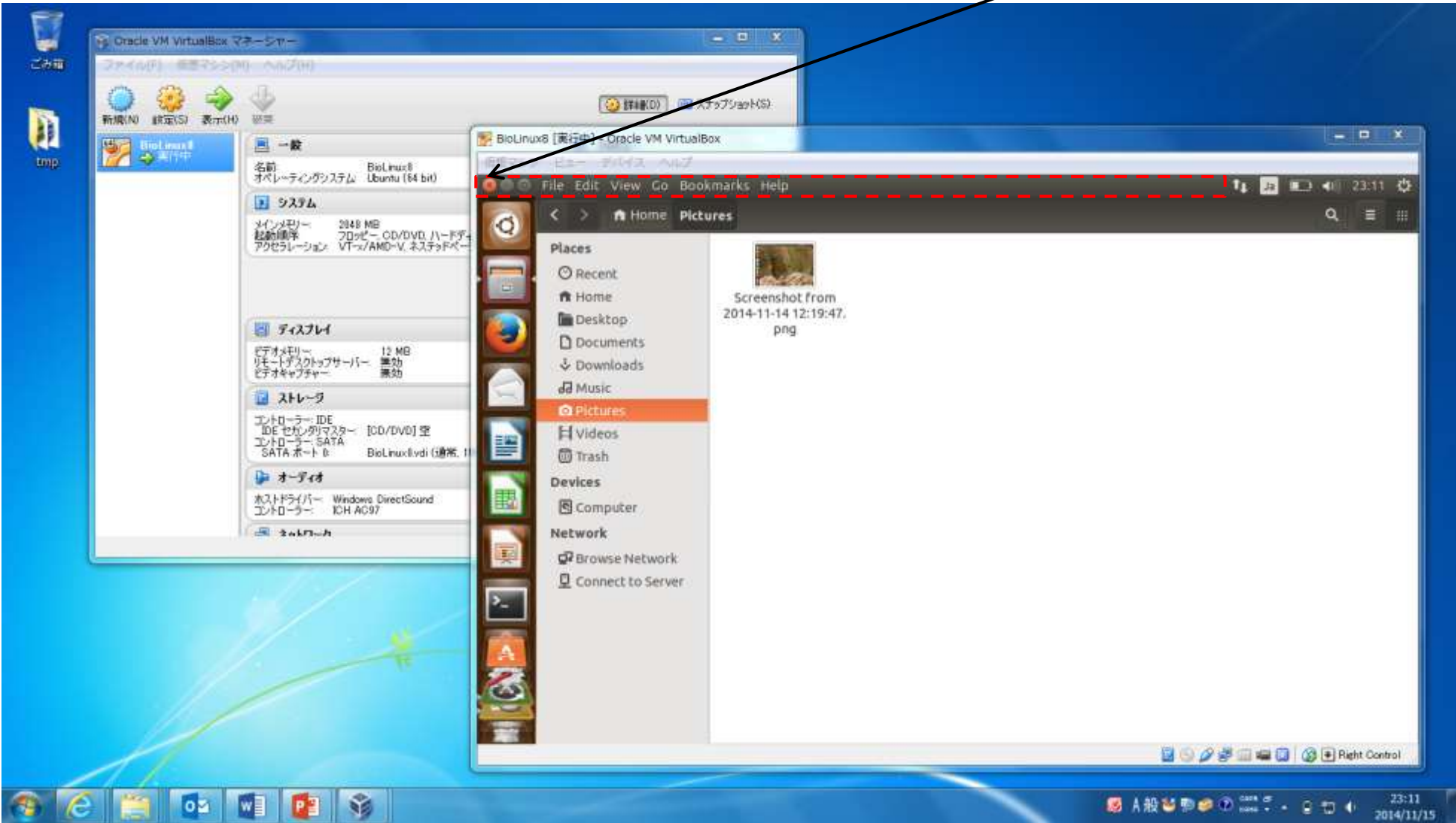
W6-3

もとの画面に戻したい場合は、赤い点線の枠内にカーソルを移動させると、左側の Pictures の P 付近に × ボタンが見られる。



この×ボタンを押すと元の画面に戻る。

W6-3



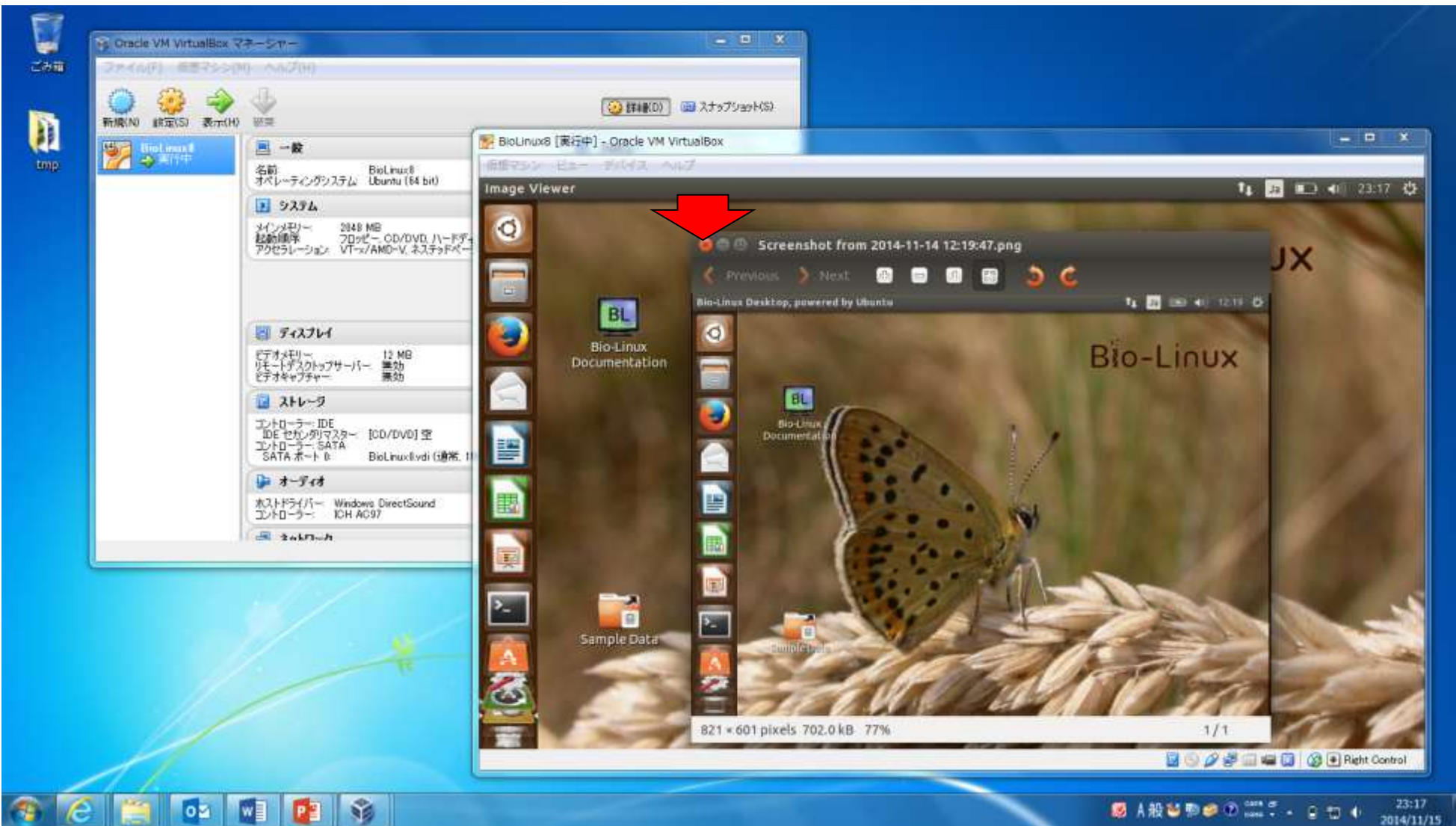
W6-3

この段階で初めて、Image Viewerで画像ファイルを開けていたことに気づく。つまり、さきほどのPicturesフォルダのウィンドウに隠れていただけである。このような状況は、ホストOS上の通常の作業でも遭遇するが、使い慣れているため無意識に対処できる場合が多い。



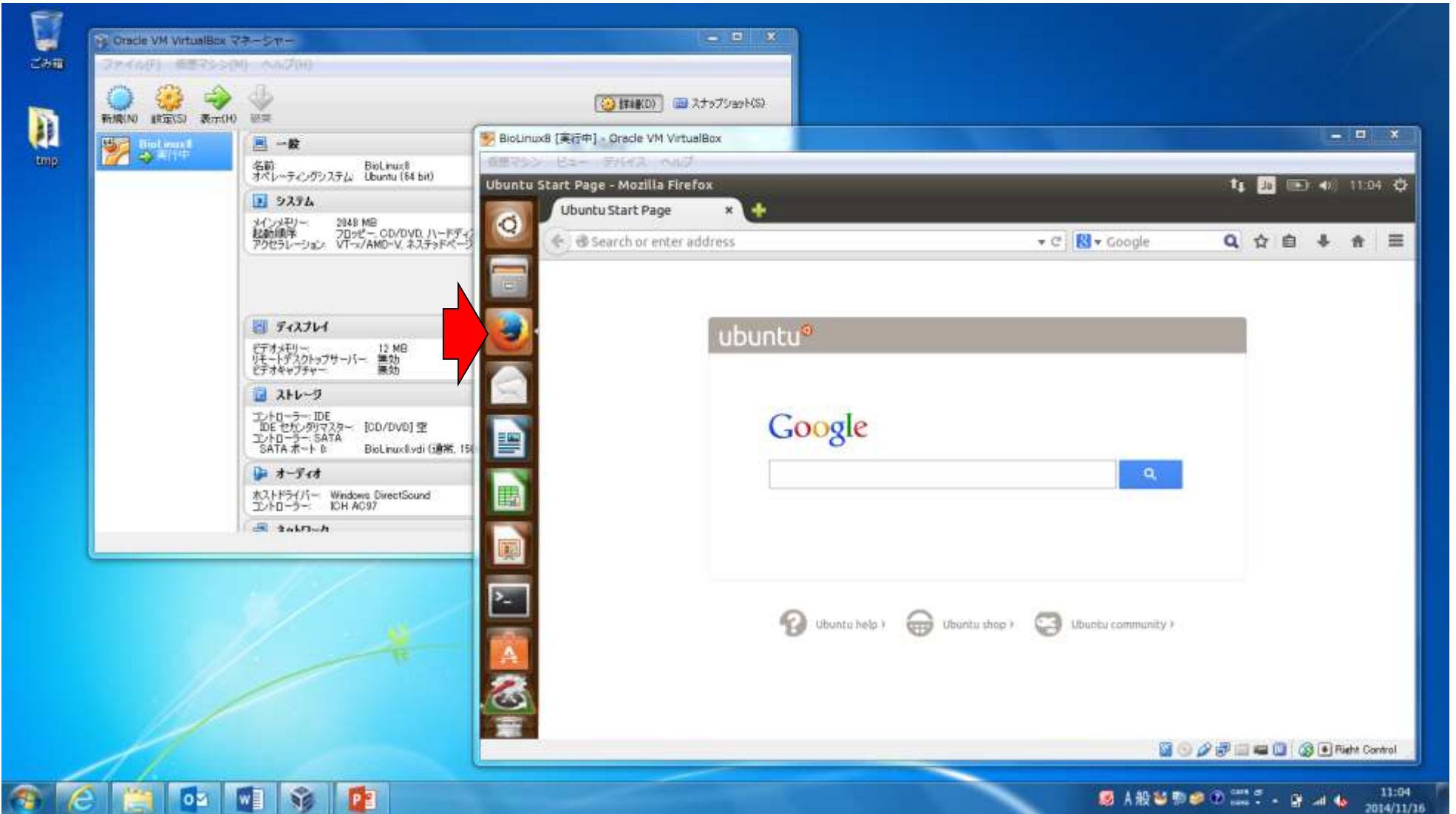
×ボタンを押すと閉じることができる。

W6-3



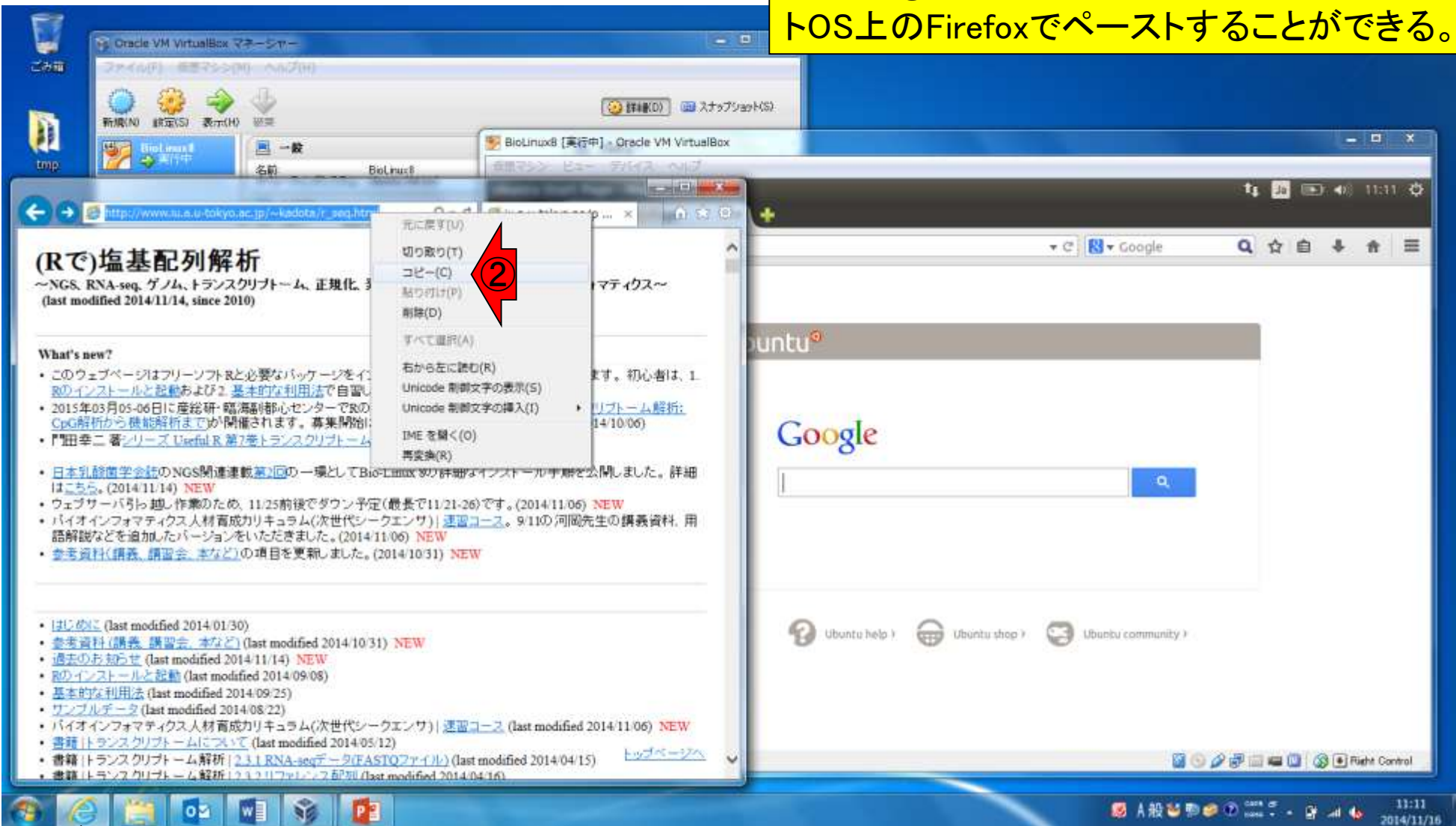
W7-1

ウェブブラウザFirefox起動後の状態。
これはゲストOS上のブラウザである。



W7-1

①ホストOS上のウェブブラウザInternet Explorerを起動し、本連載の参考ウェブページを起動したところ。②URLをクリップボードにコピーして、ゲストOS上のFirefoxでペーストすることができる。



①

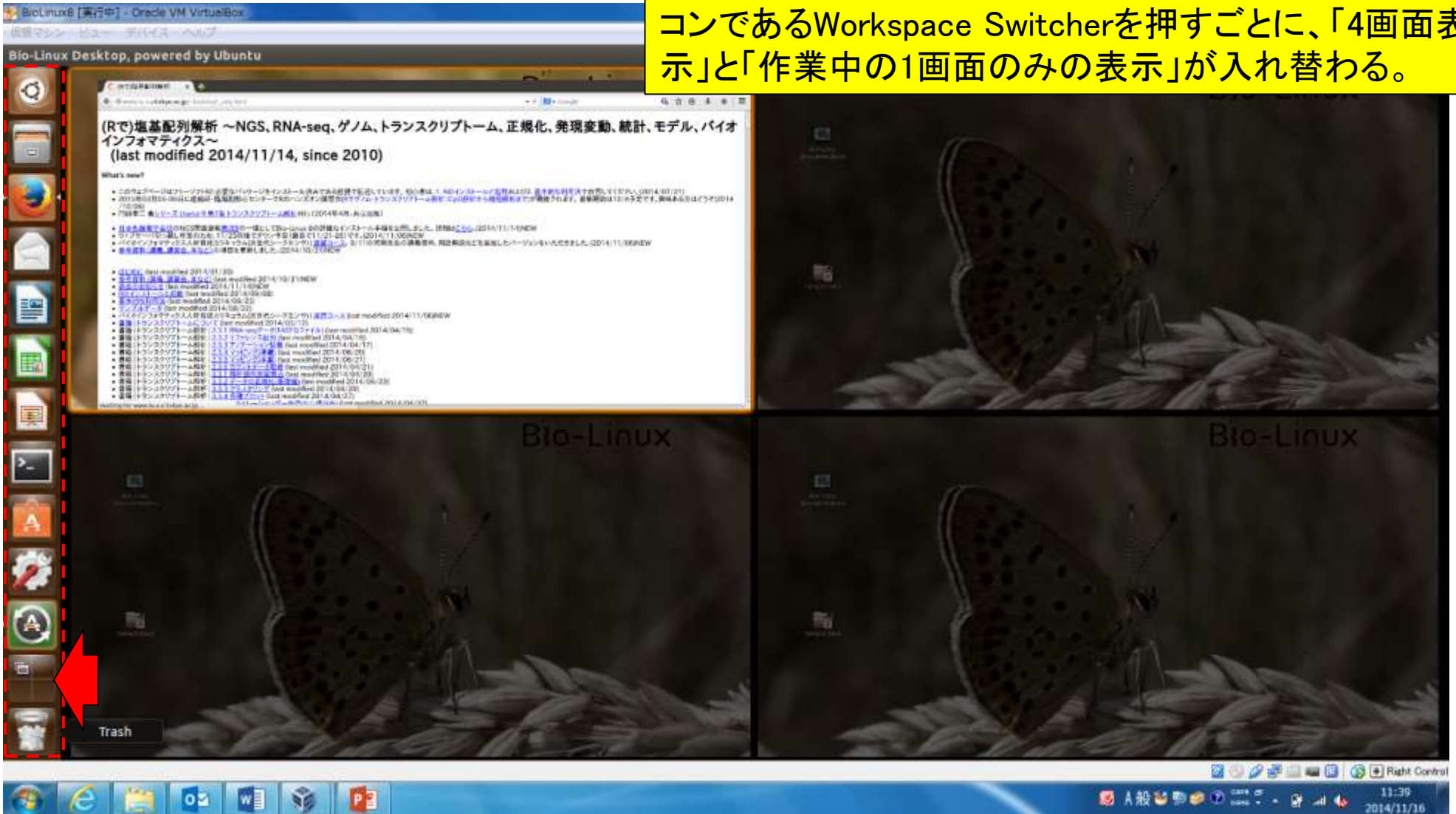
W7-1

クリップボードにコピーしたURL情報を、ゲストOS上のFirefoxでペーストした結果。全画面表示にする。



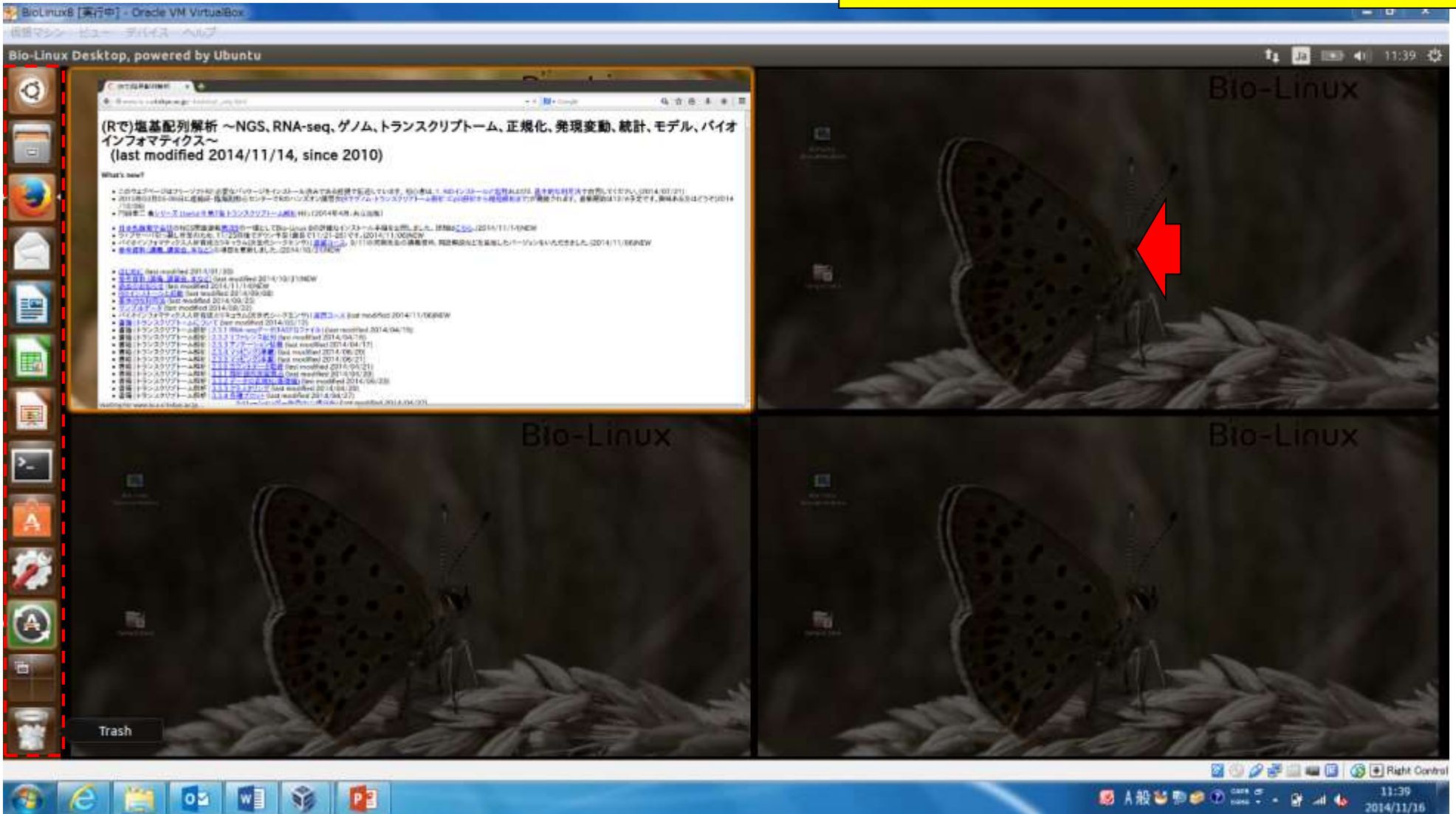
W7-2

赤枠内がメニューバーのようなもの。この枠内にカーソルを移動させると、下のほうにもいくつかのアイコンがあることがわかる。一番下がごみ箱で、下から2番目のアイコンであるWorkspace Switcherを押すごとに、「4画面表示」と「作業中の1画面のみの表示」が入れ替わる。



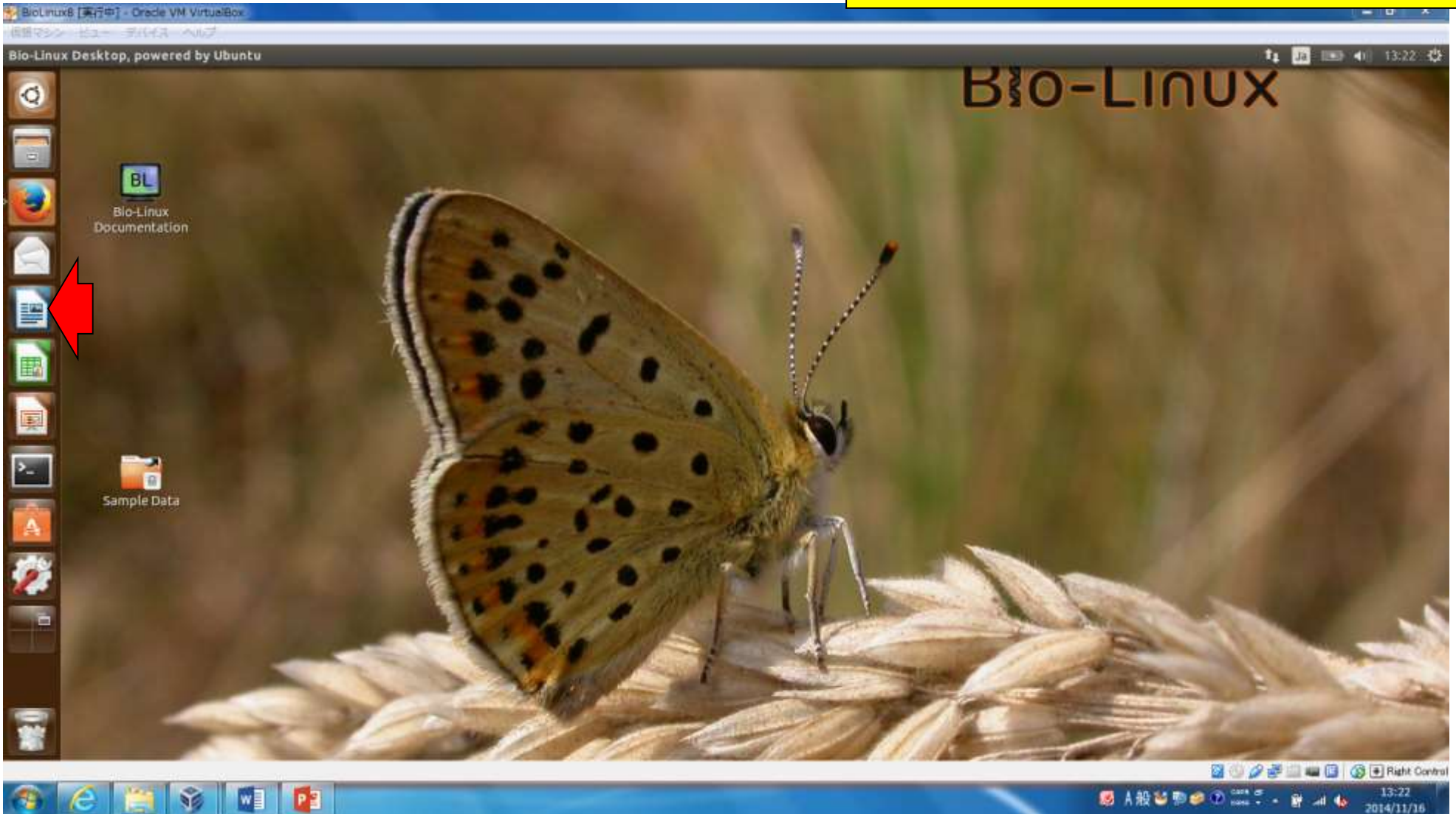
W7-2

右上の画面上でクリックすると、ウェブブラウザが表示されない新しい画面となる。このようにマルチ画面を駆使するのが一般的。

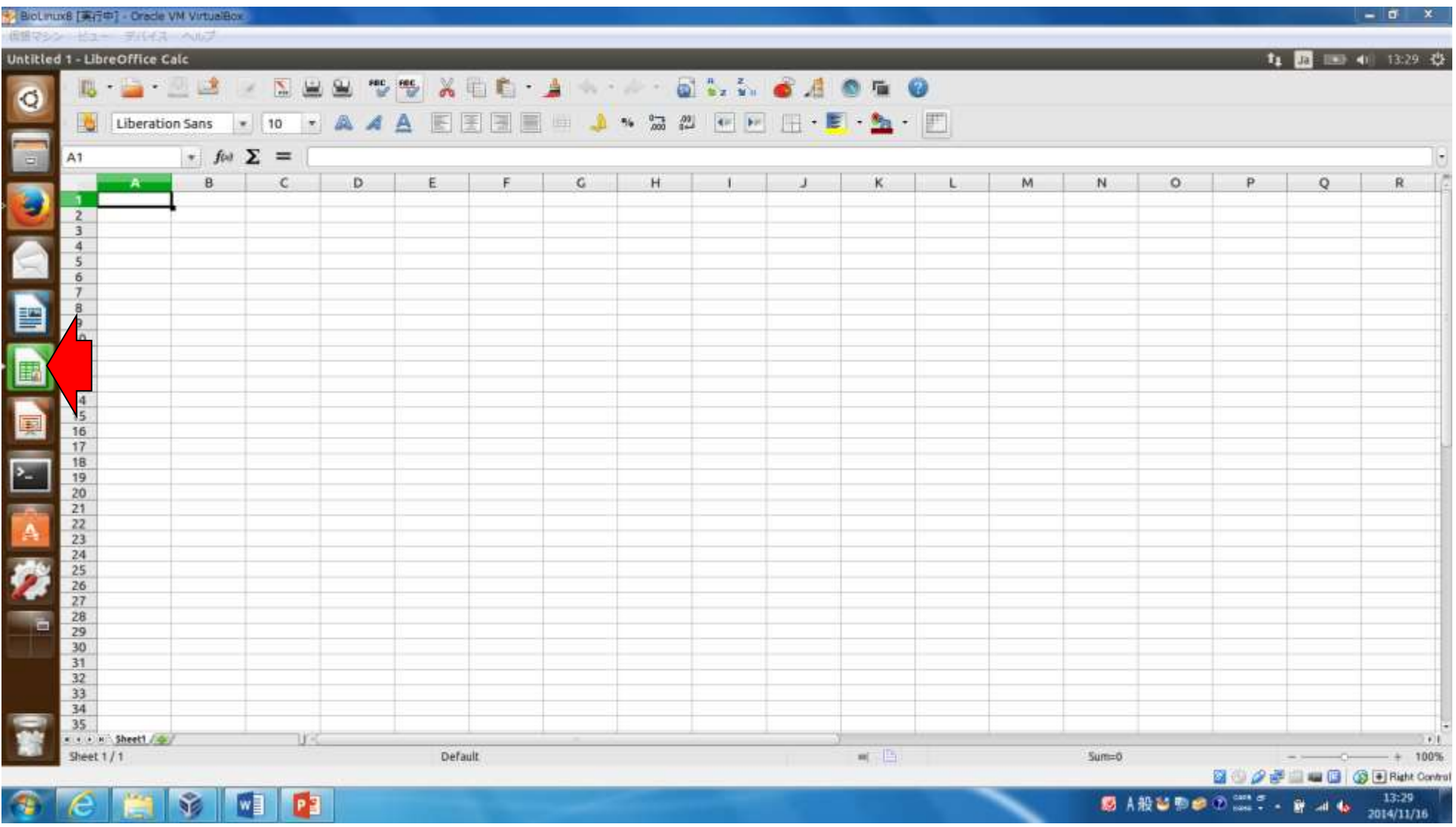


W7-3

4画面分の右上を表示させた結果。テキストエディタ(LibreOffice Writer)を開きたい場合はここをクリック。実際にはしなくてよい。

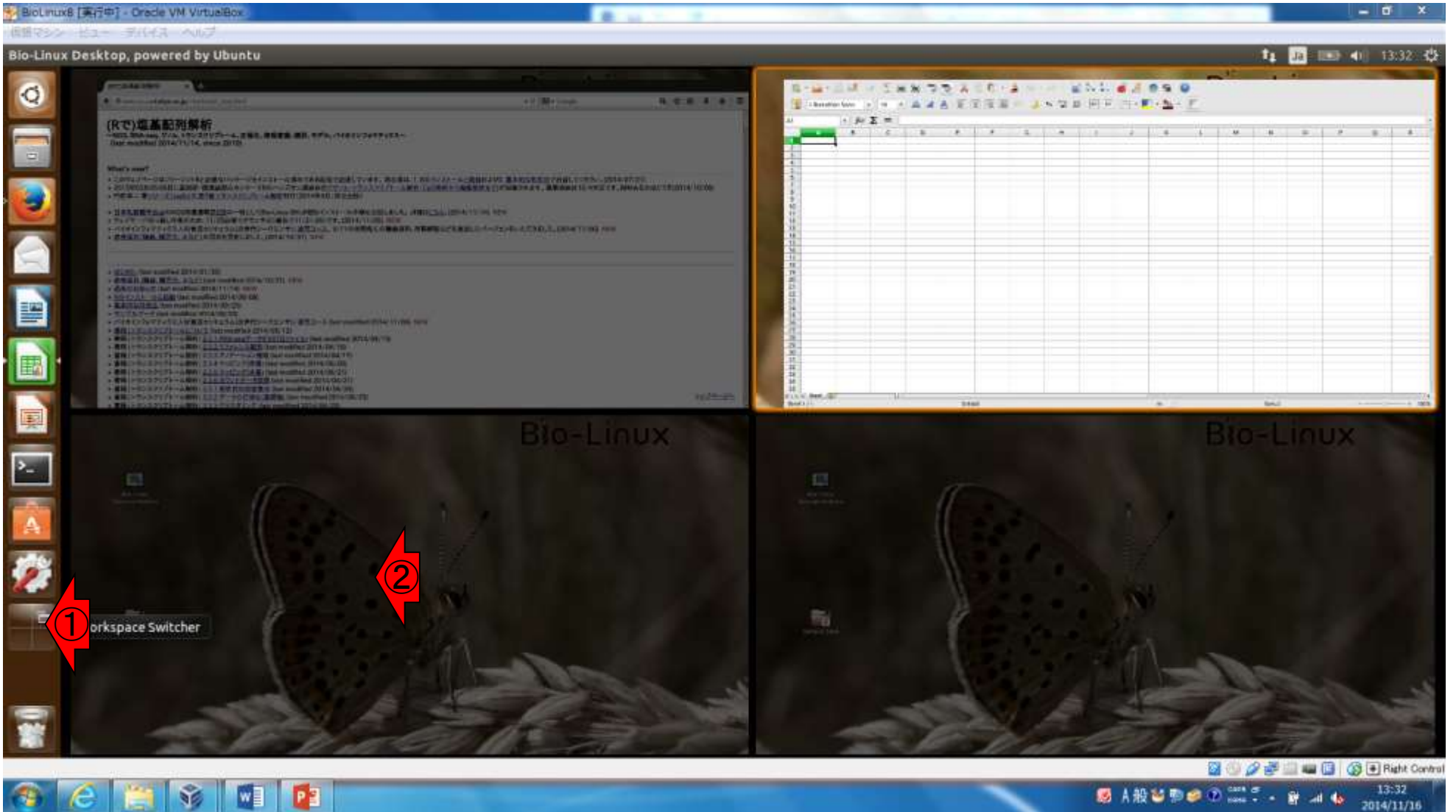


W7-3



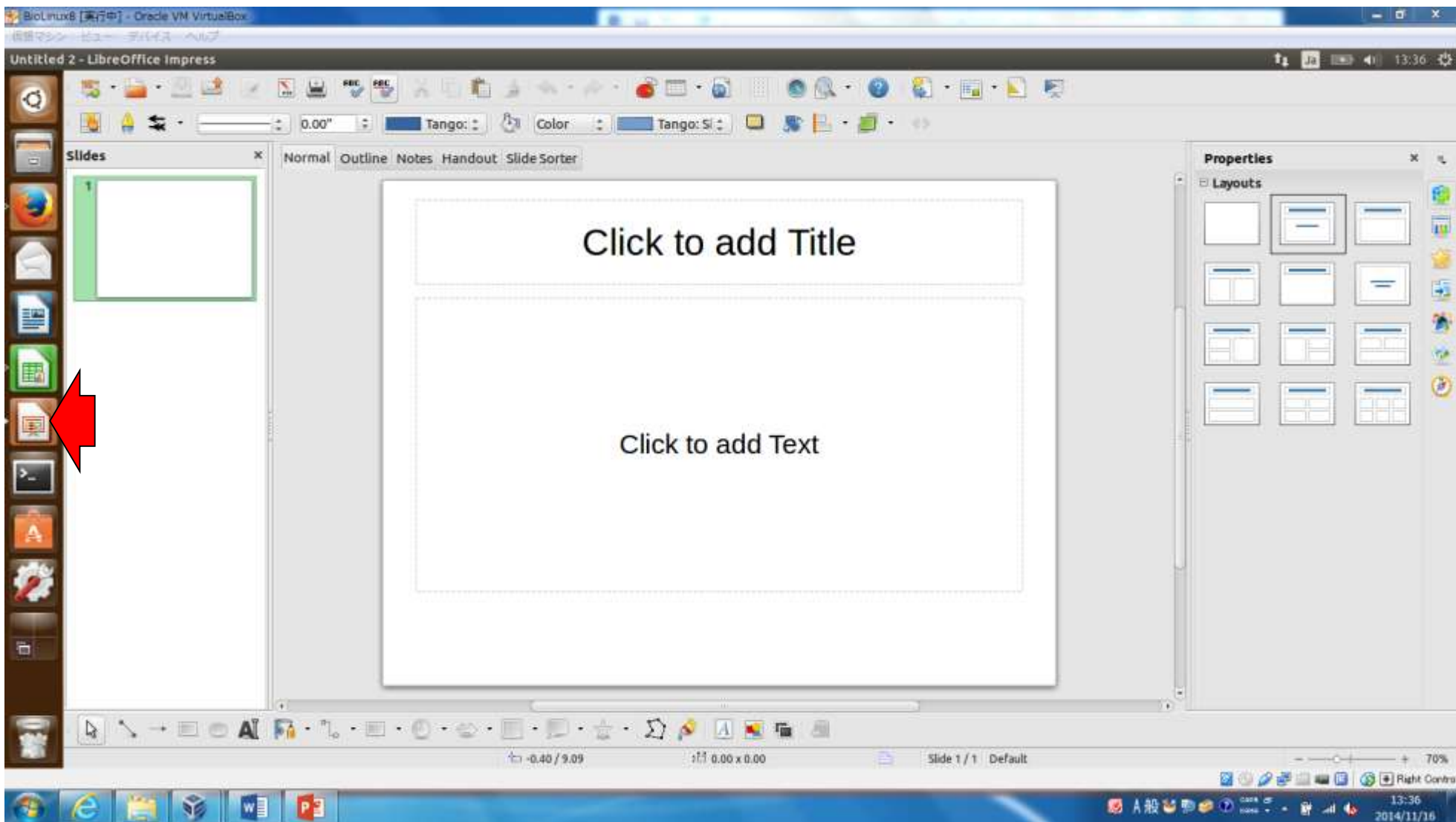
W7-3

①Workspace Switcherアイコンをクリックした結果。②4画面分の左下を表示させる。



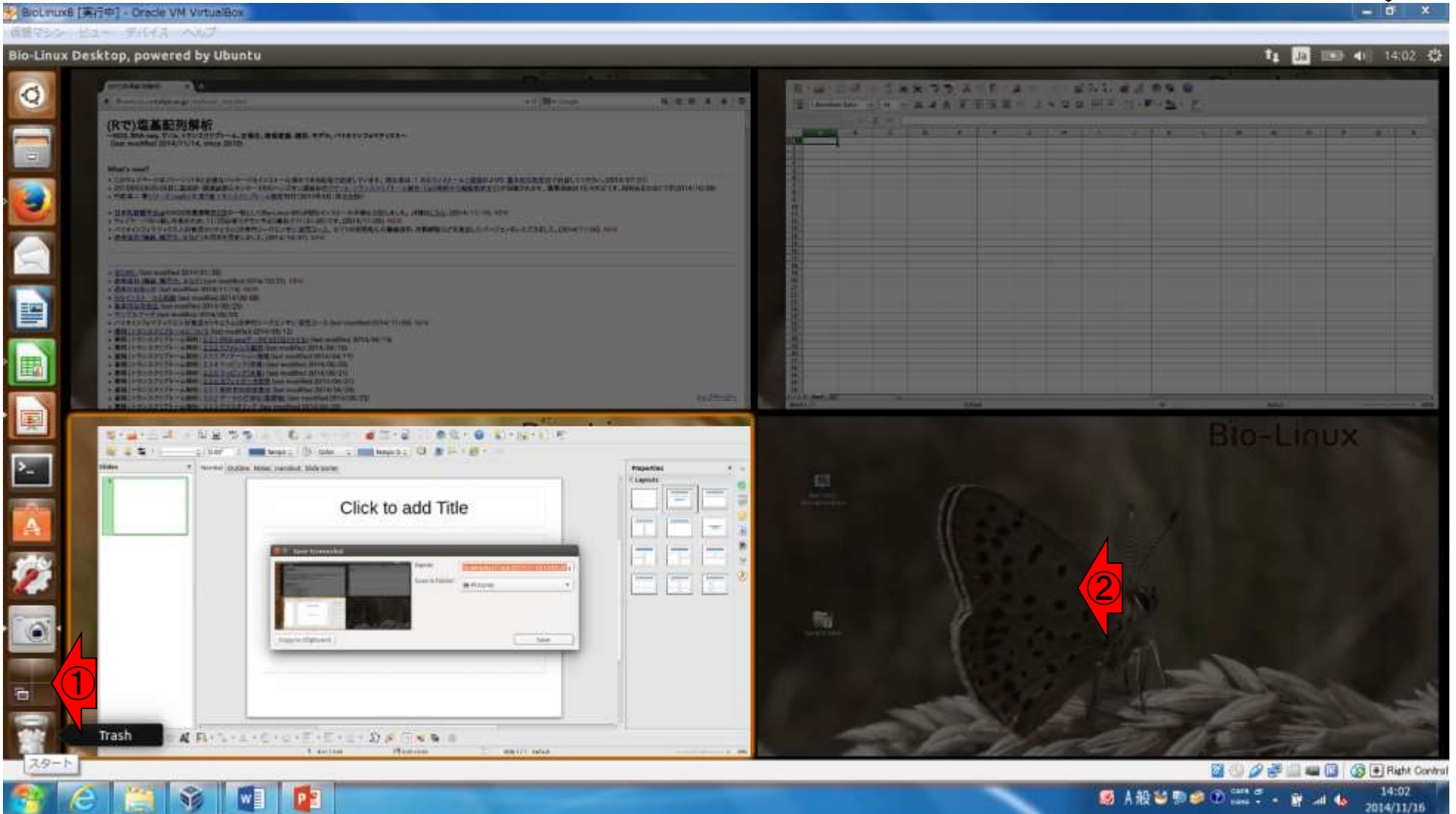
W7-3

4画面分の左下を表示させ、プレゼンテーションソフト(LibreOffice Impress)を起動した結果。



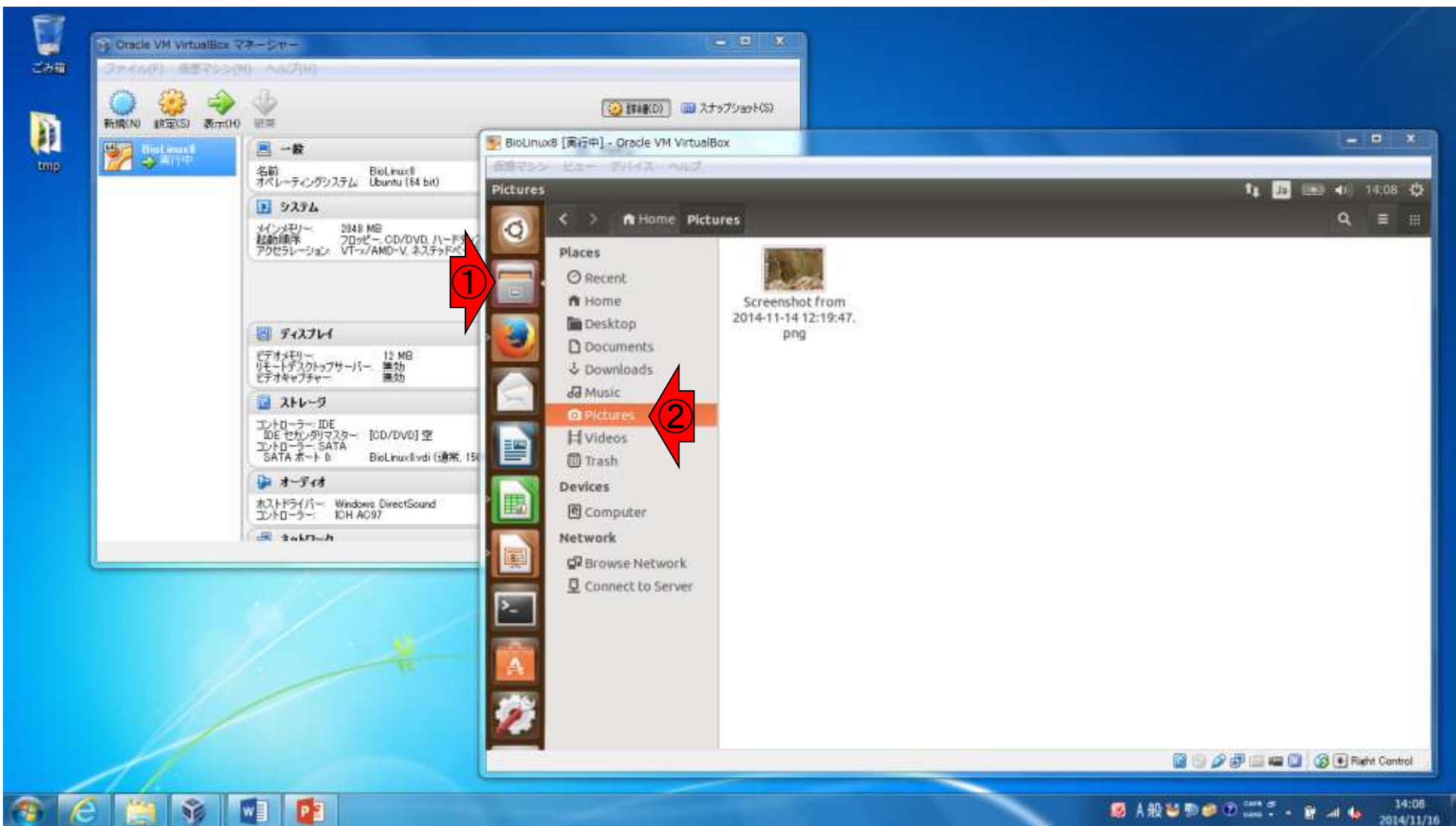
4画面分の右下を表示させ、全画面表示を解除。

W8-1



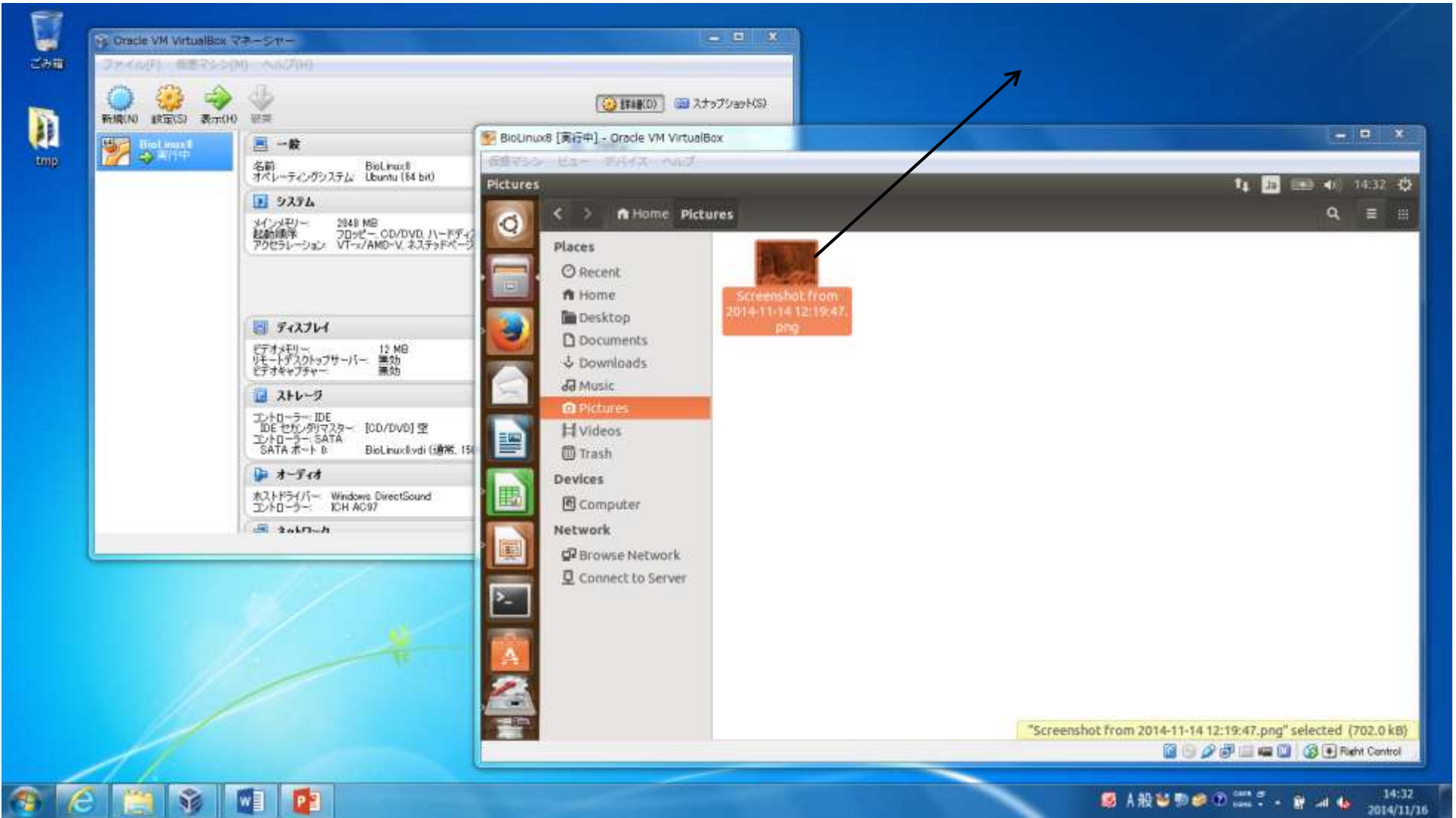
W8-1

- ①フォルダアイコンをクリックし、
- ②Picturesフォルダをクリック。



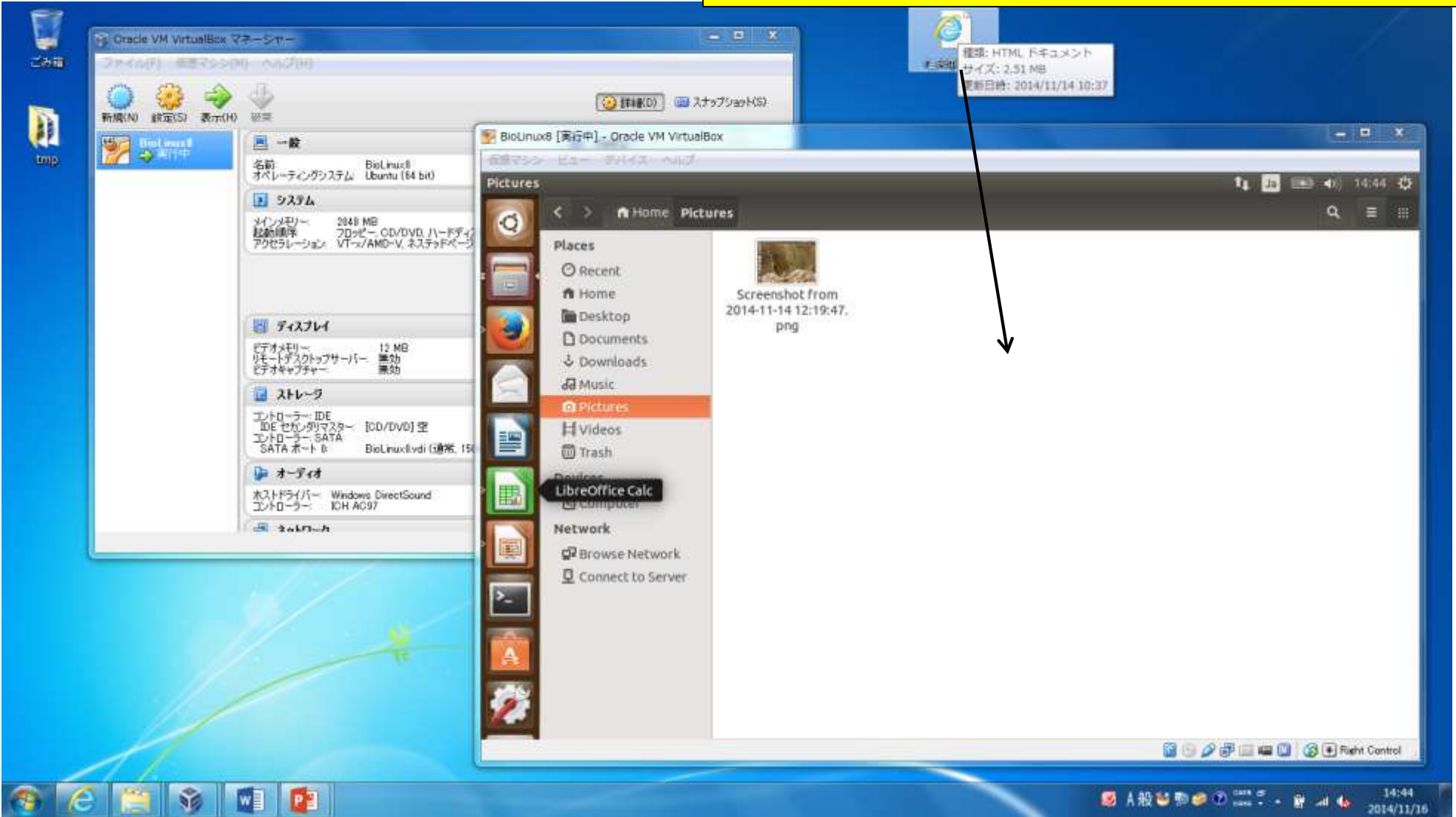
W8-1

ゲストOSのフォルダ内にあるpngファイルを、ホストOSのデスクトップにドラッグ&ドロップすることはできないようだ。

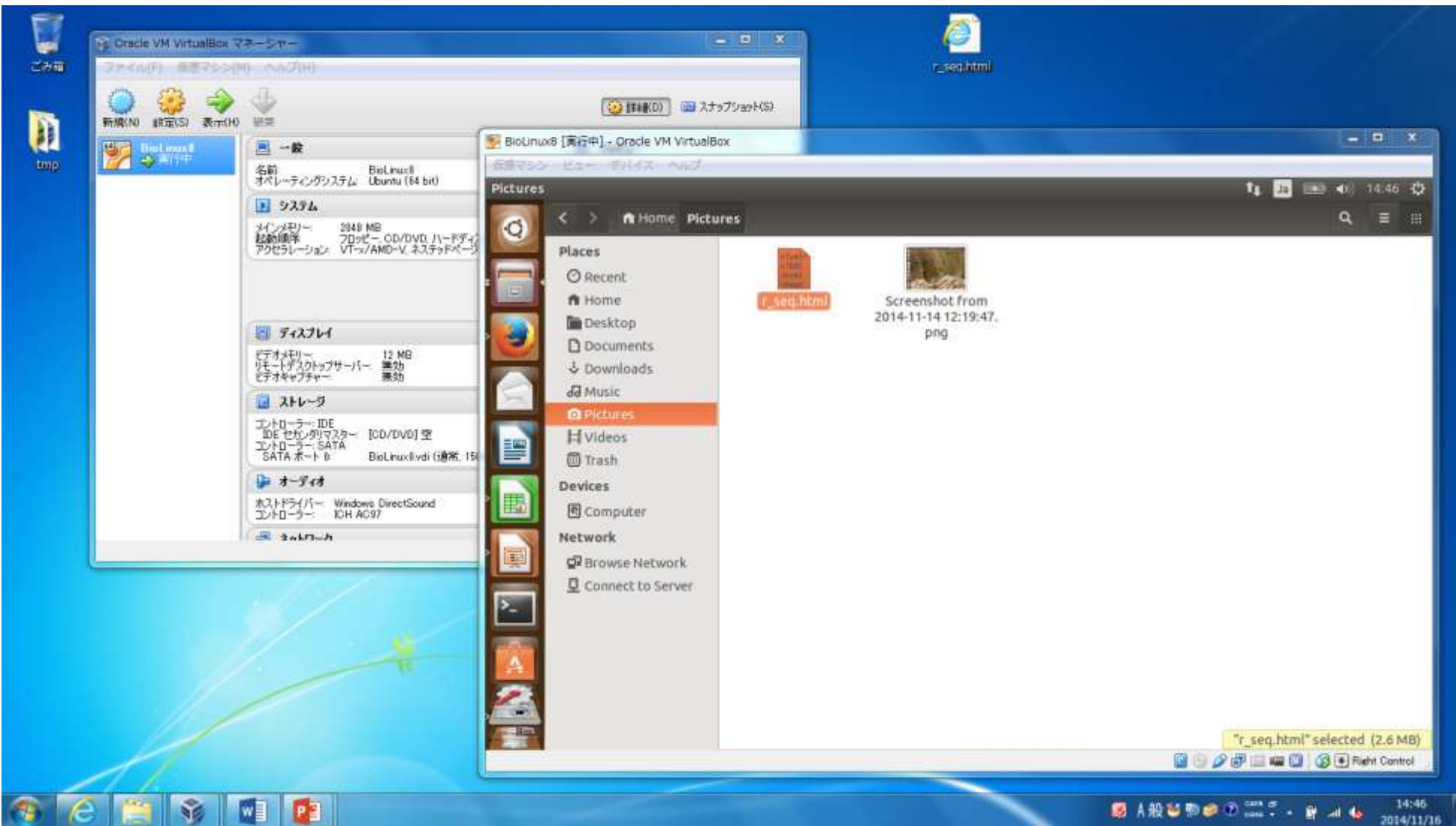


W8-1

ホストOSのデスクトップにあるhtmlファイル(手元にある任意のファイルで試してよい)を、ゲストOSのフォルダ内にドラッグ&ドロップすることはできるようだ。

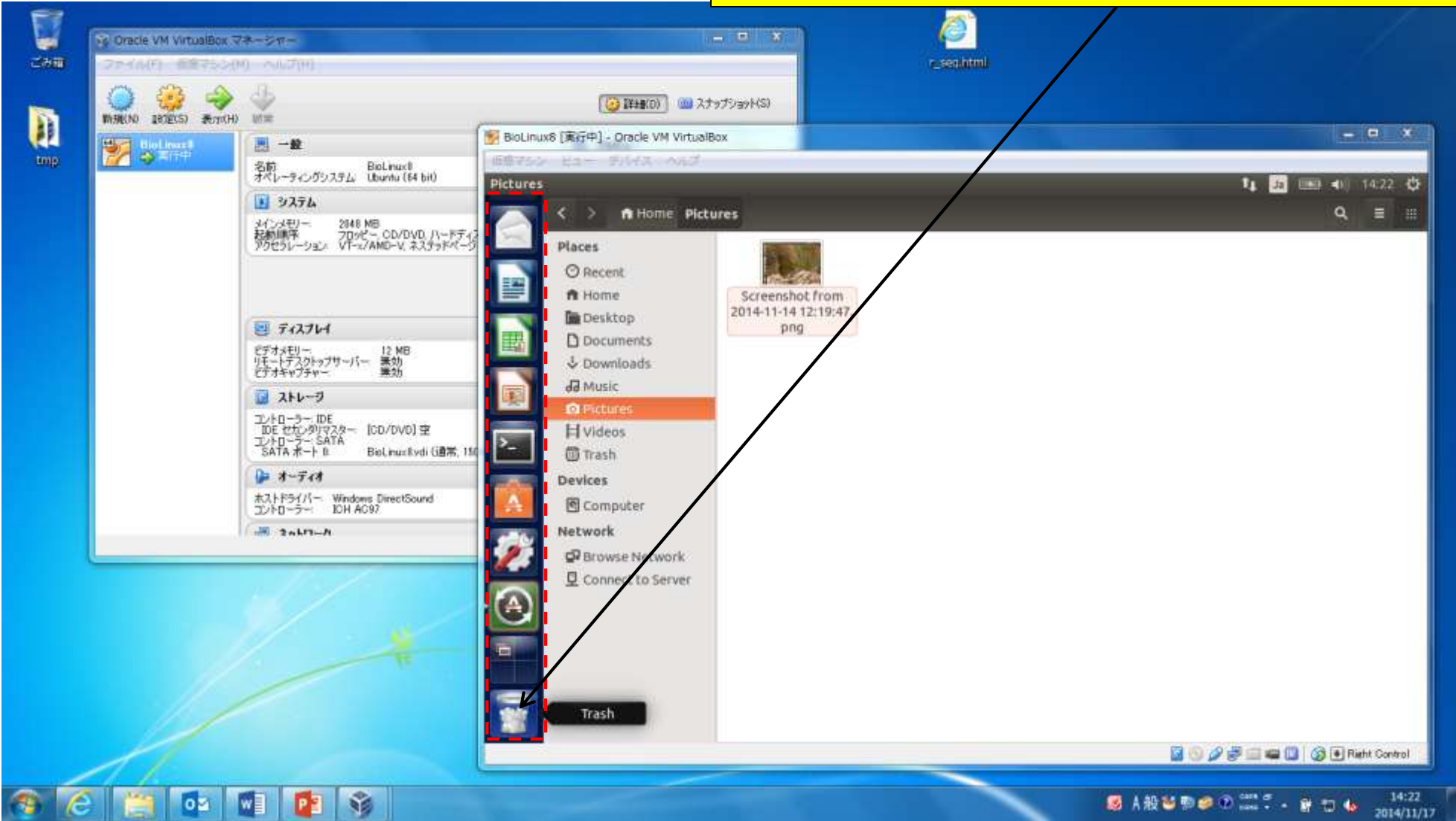


W8-1



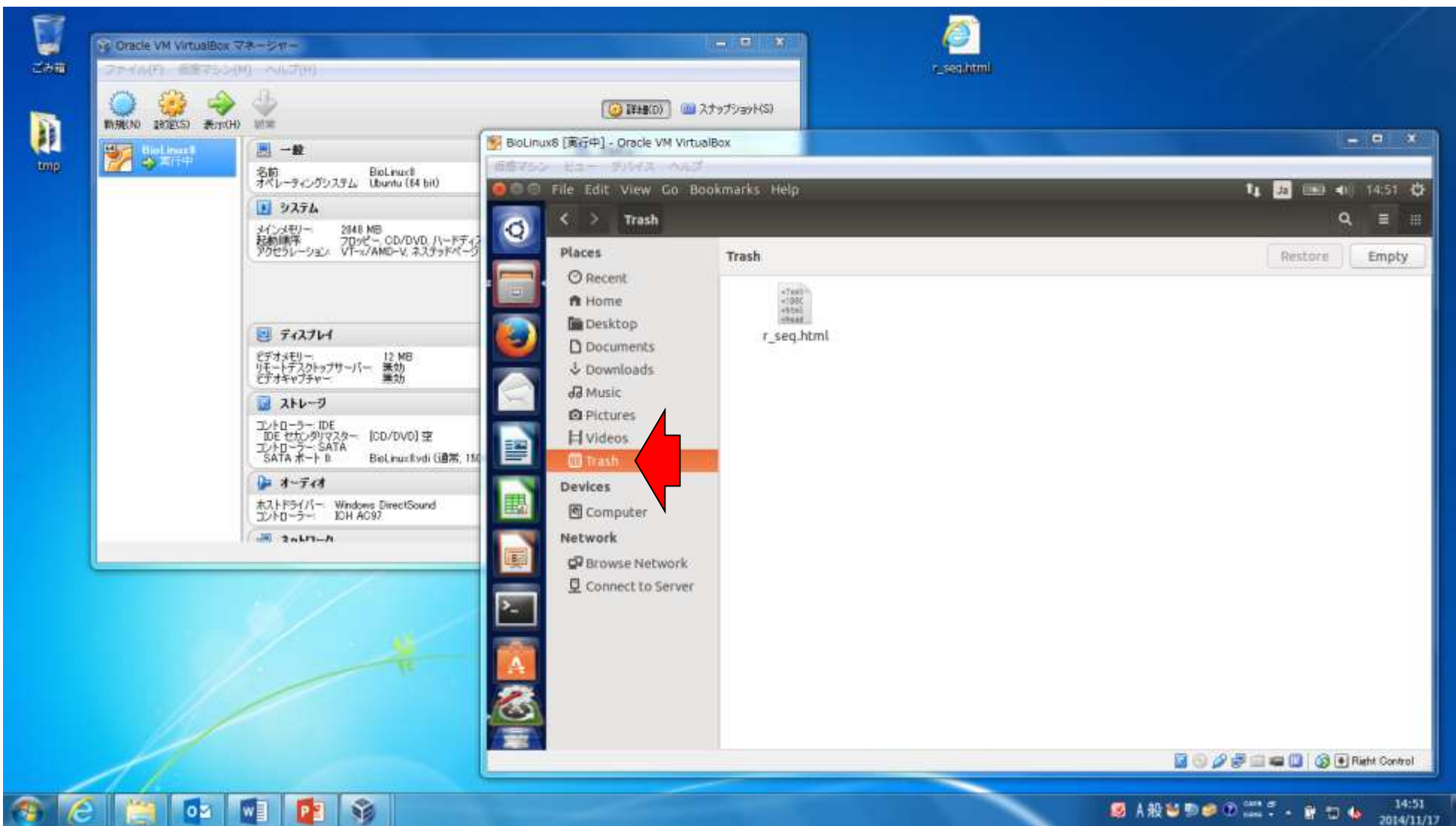
W8-2

ファイルの削除は、通常のGUI環境と基本的に同じ。削除したいファイル(例:r_seq.html)を赤枠アイコンの一番下にあるゴミ箱(Trash)に移動させればよい。



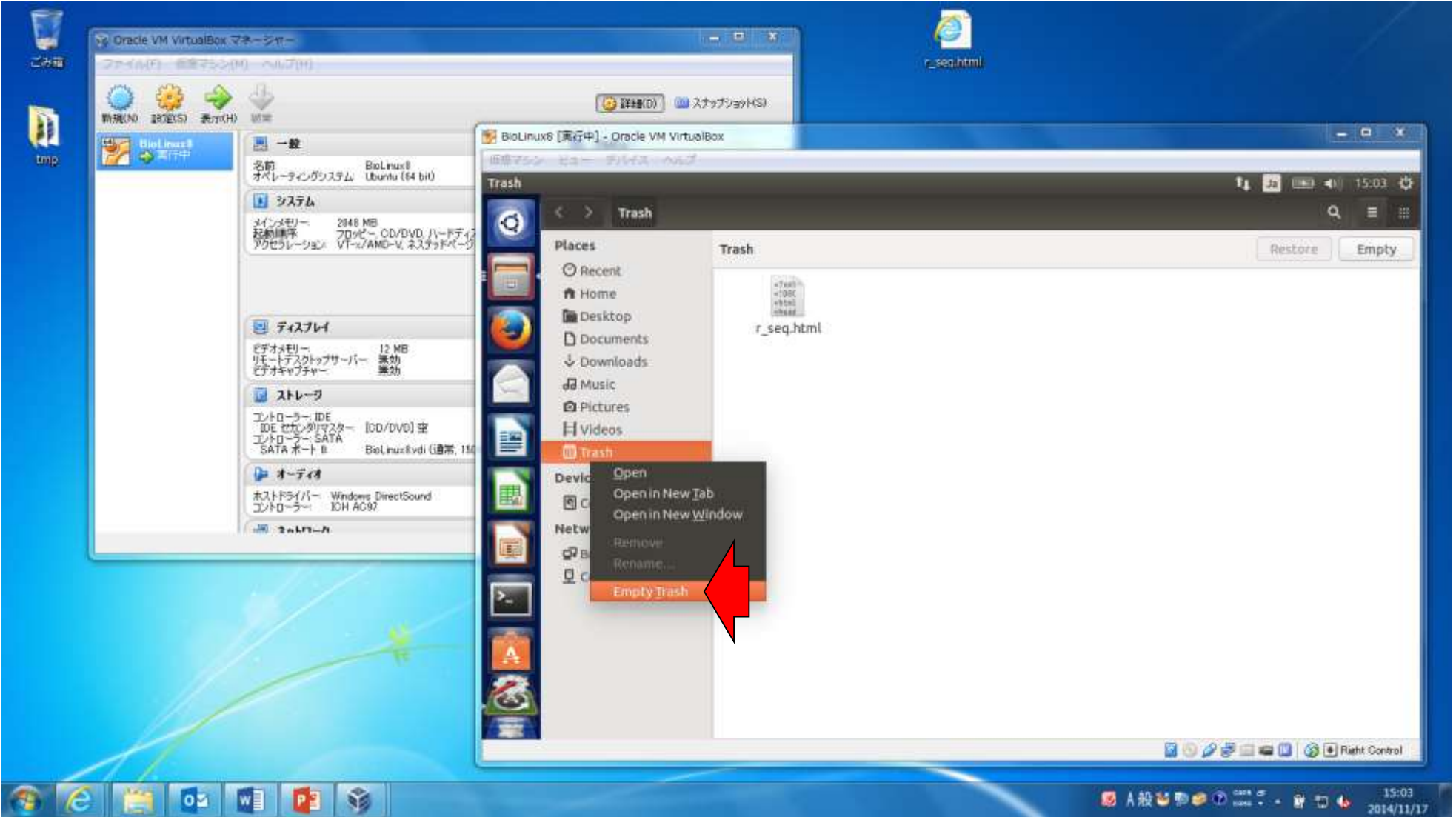
W8-2

ゴミ箱をクリックするなり、赤矢印部分をクリックすると、ごみ箱の中身をみることができる。



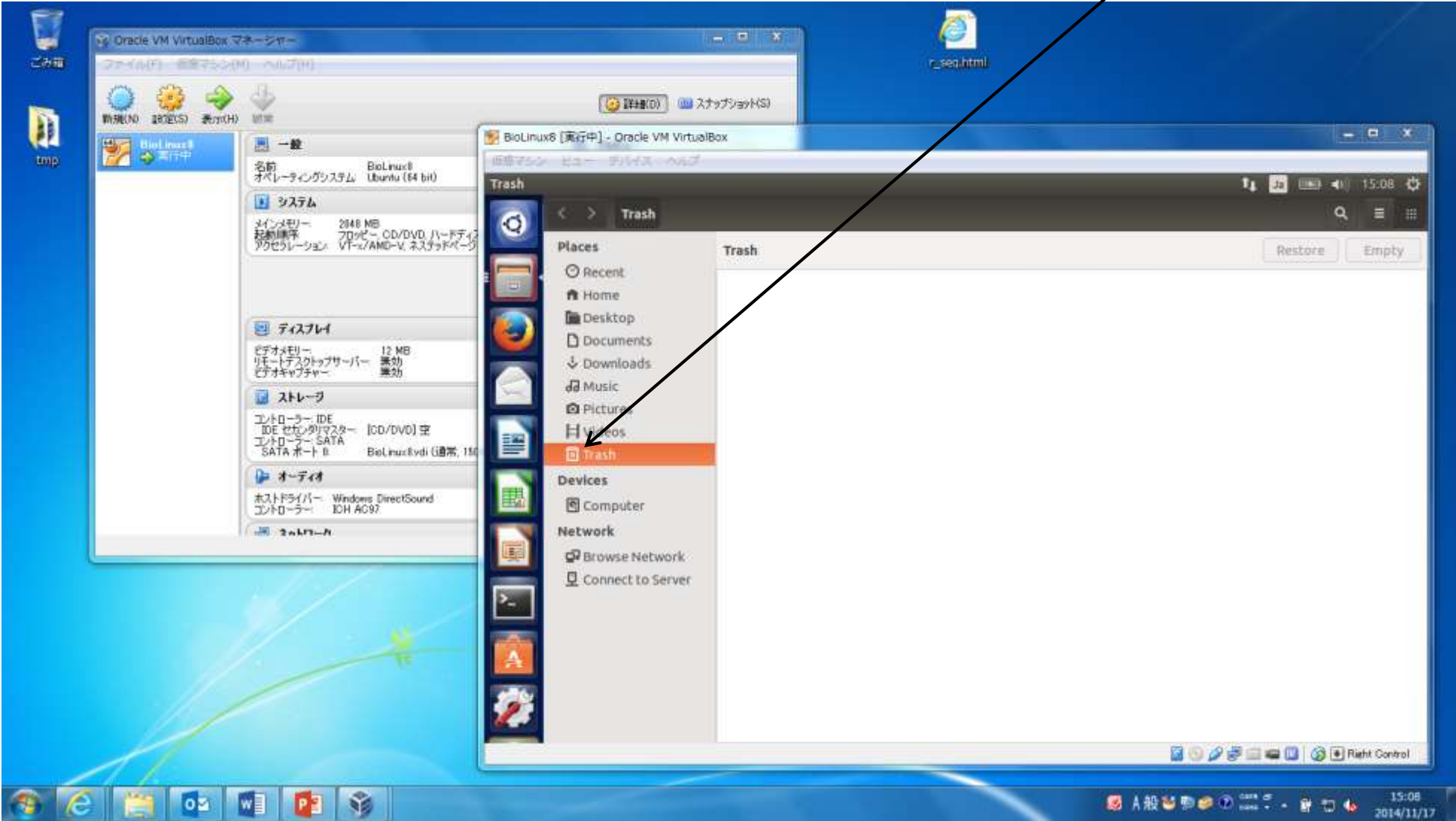
ゴミ箱アイコン上で右クリックし、Empty Trashを選択することでゴミ箱を空にすることができます。

W8-2



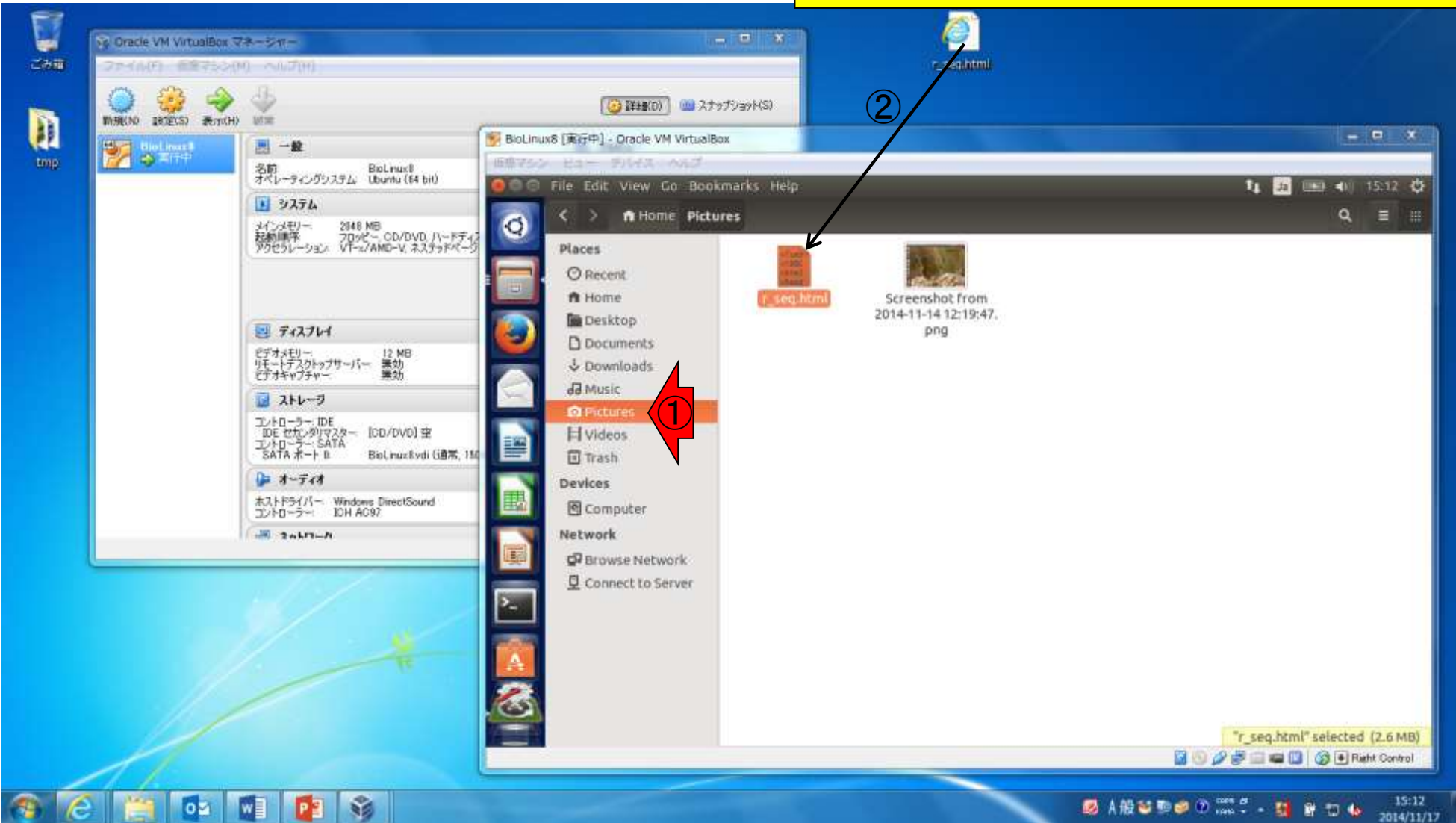
ゴミ箱アイコンがスリムになっていますね。

W8-2



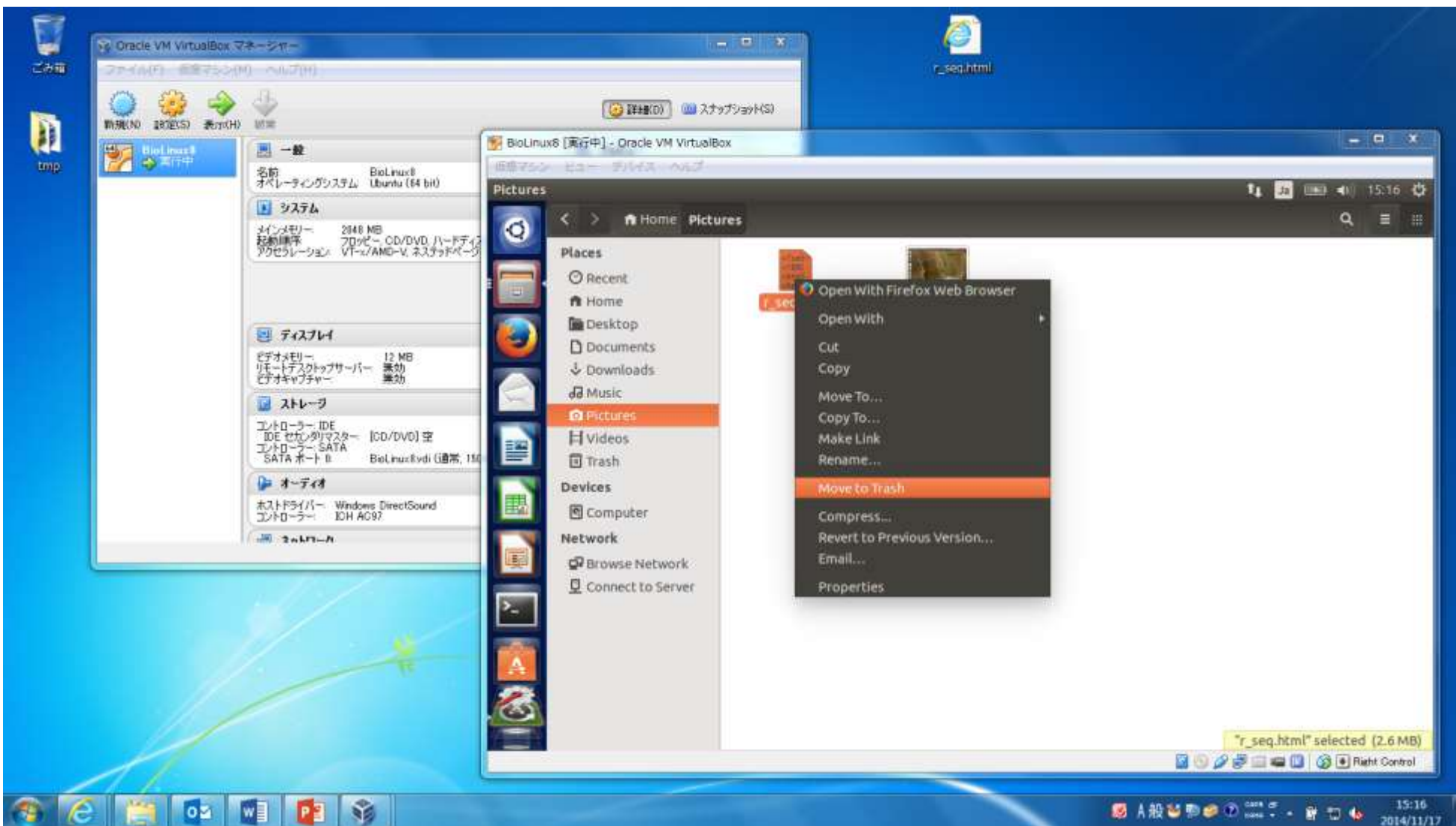
W8-2

もう一度、さきほどのPicturesフォルダに移動して
ホストOSのデスクトップにあるr_seq.htmlをゲ
ストOSのPicturesフォルダにドラッグ&ドロップ



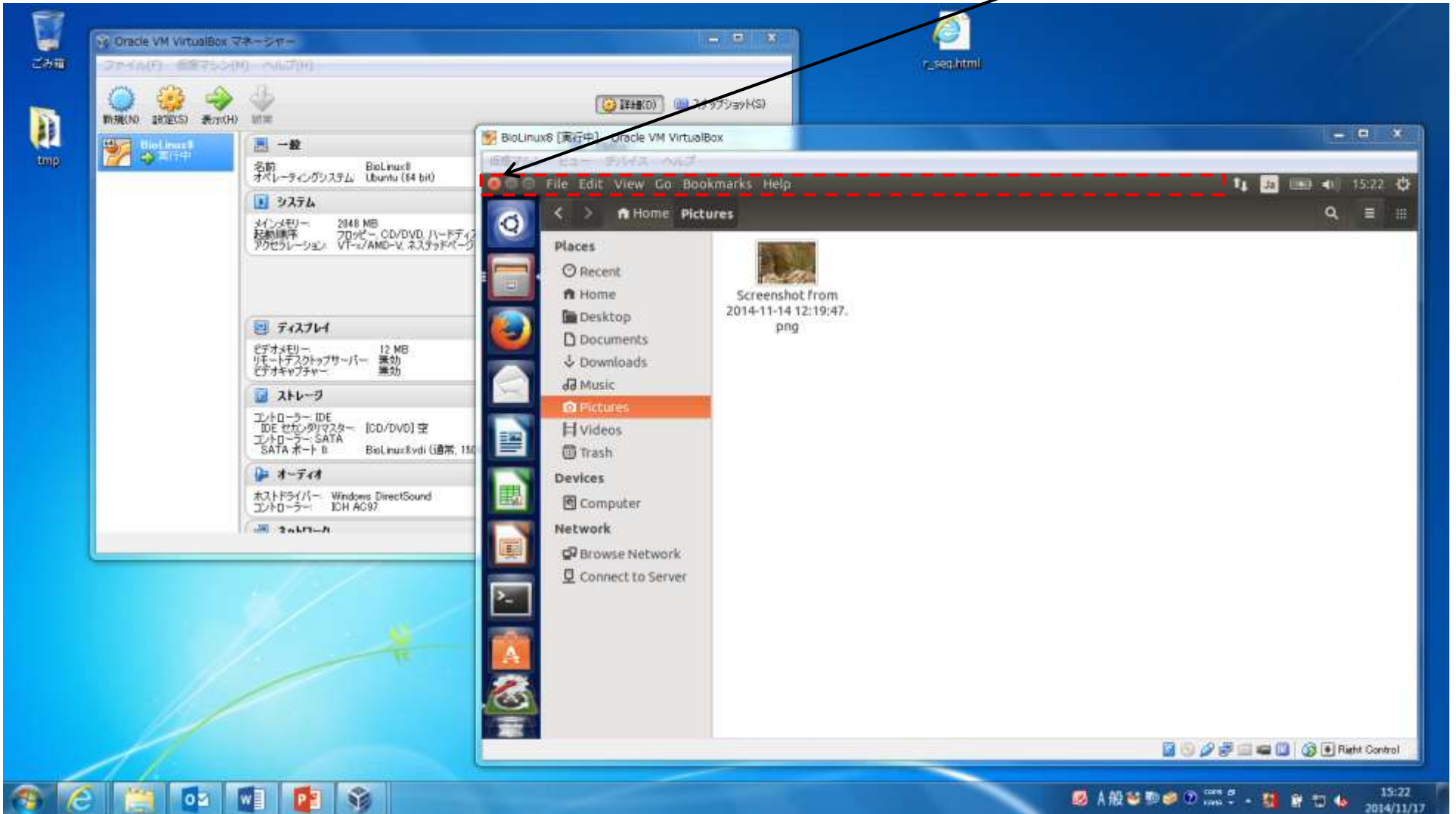
W8-2

削除したいファイルのアイコン上で右クリックして、Move to Trashを選択するのもよい。



W8-3

ターミナルの説明に入る前に、赤枠内にカーソル移動し、一番左側の×ボタンを押してフォルダ画面を閉じる。



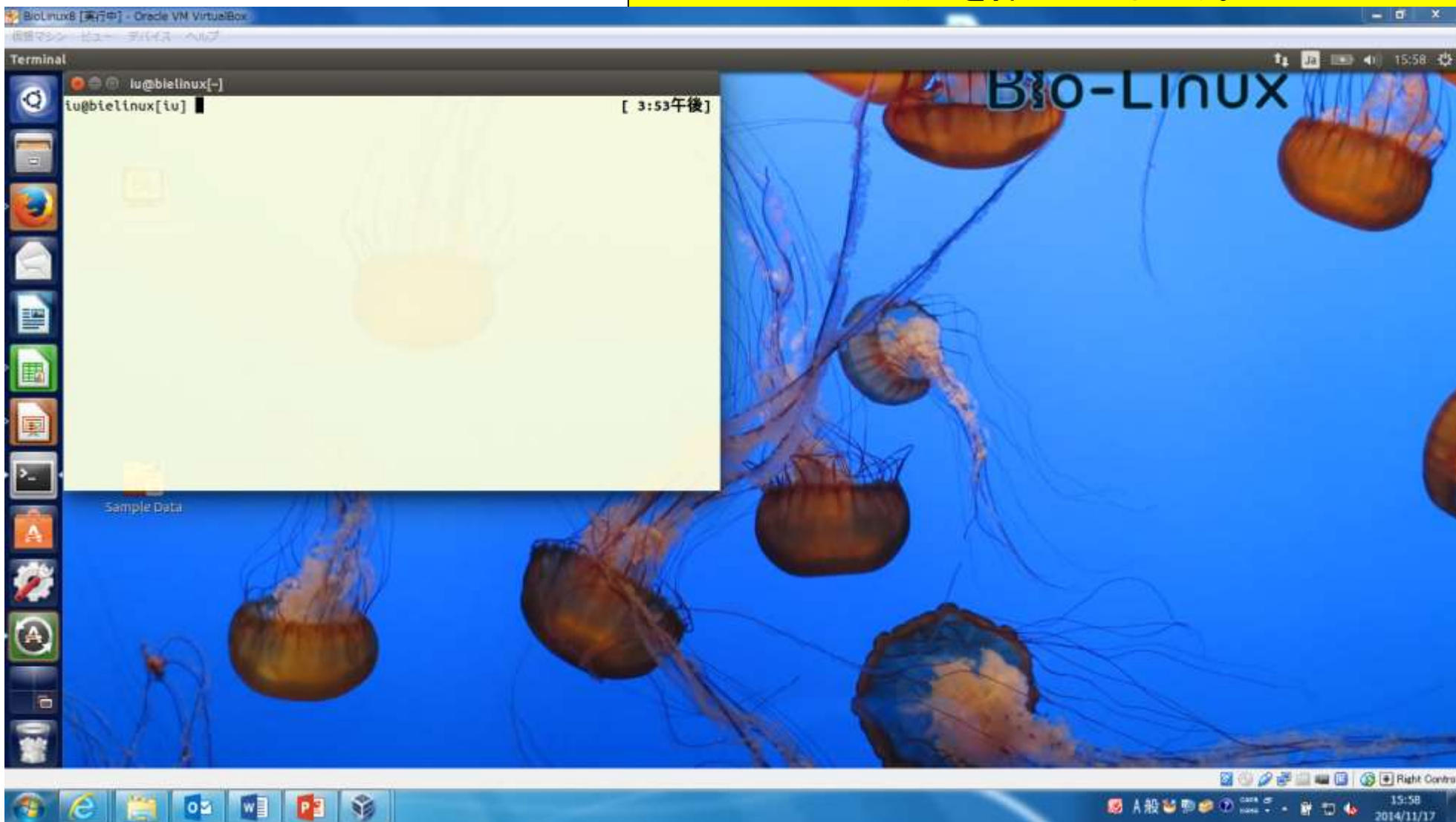
背景が変わっているが、特に気にしなくて良い。
①全画面表示に切り替えて、②ターミナルを起動。

W8-3



W8-3

ターミナル起動後の状態。ホストOSでのスクリーンショット。BioLinux8ウィンドウを非アクティブ状態にして、キーボードのPrint Screenボタンを押しているだけ。



W8-3

ターミナル起動後の状態。ゲストOSでのスクリーンショット。キーボードのPrint Screenボタンを押しているだけ。



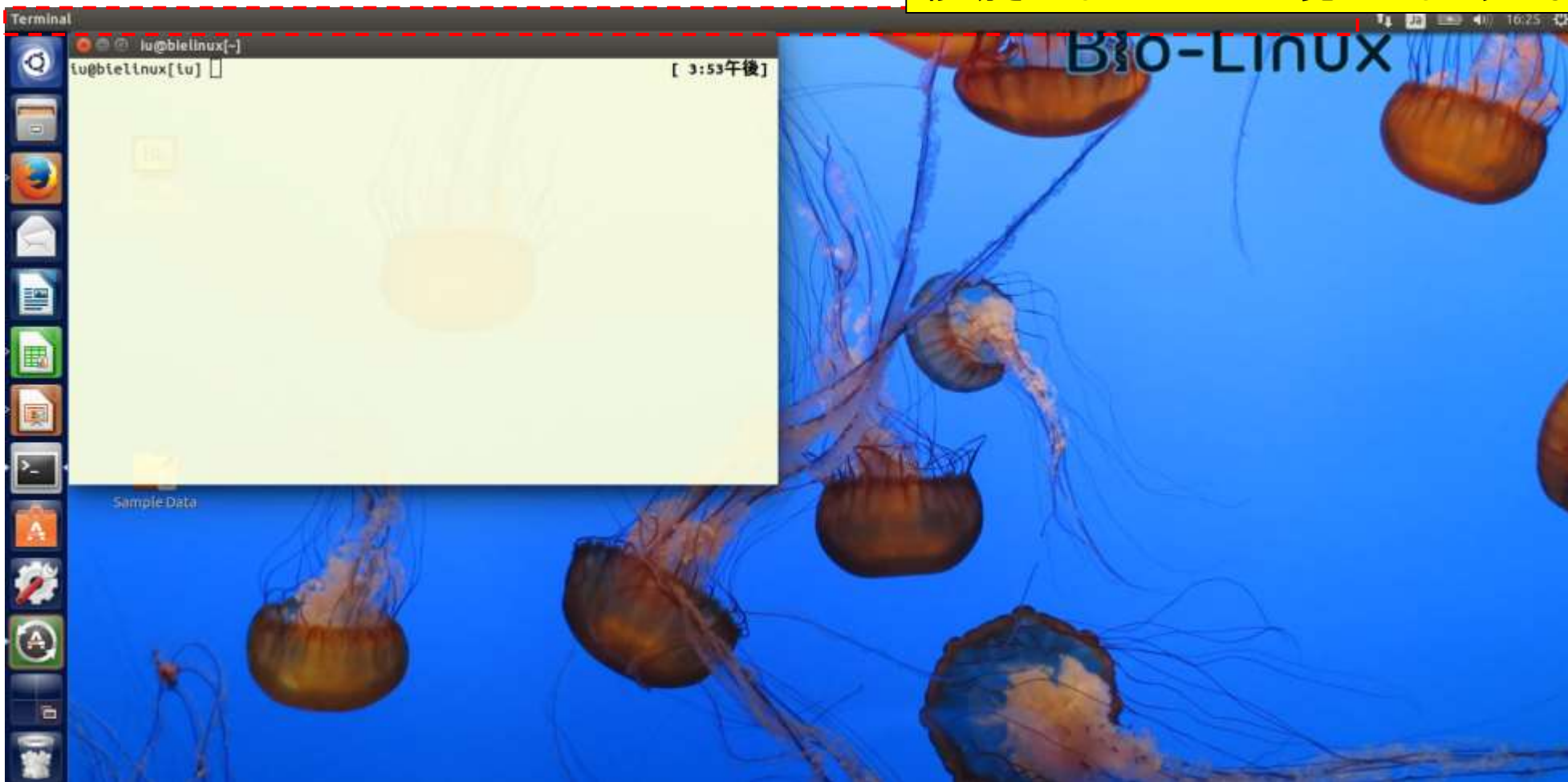
W8-4

ターミナルの文字の大きさを変える作業を行う。赤枠内でクリックし、ターミナルのウィンドウをアクティブにする。
(この画面自体はまだアクティブになっていない状態)



W8-4

ターミナルのウィンドウがアクティブになった状態。この状態でまた赤枠部分にカーソル移動させるとメニューが見られるようになる。

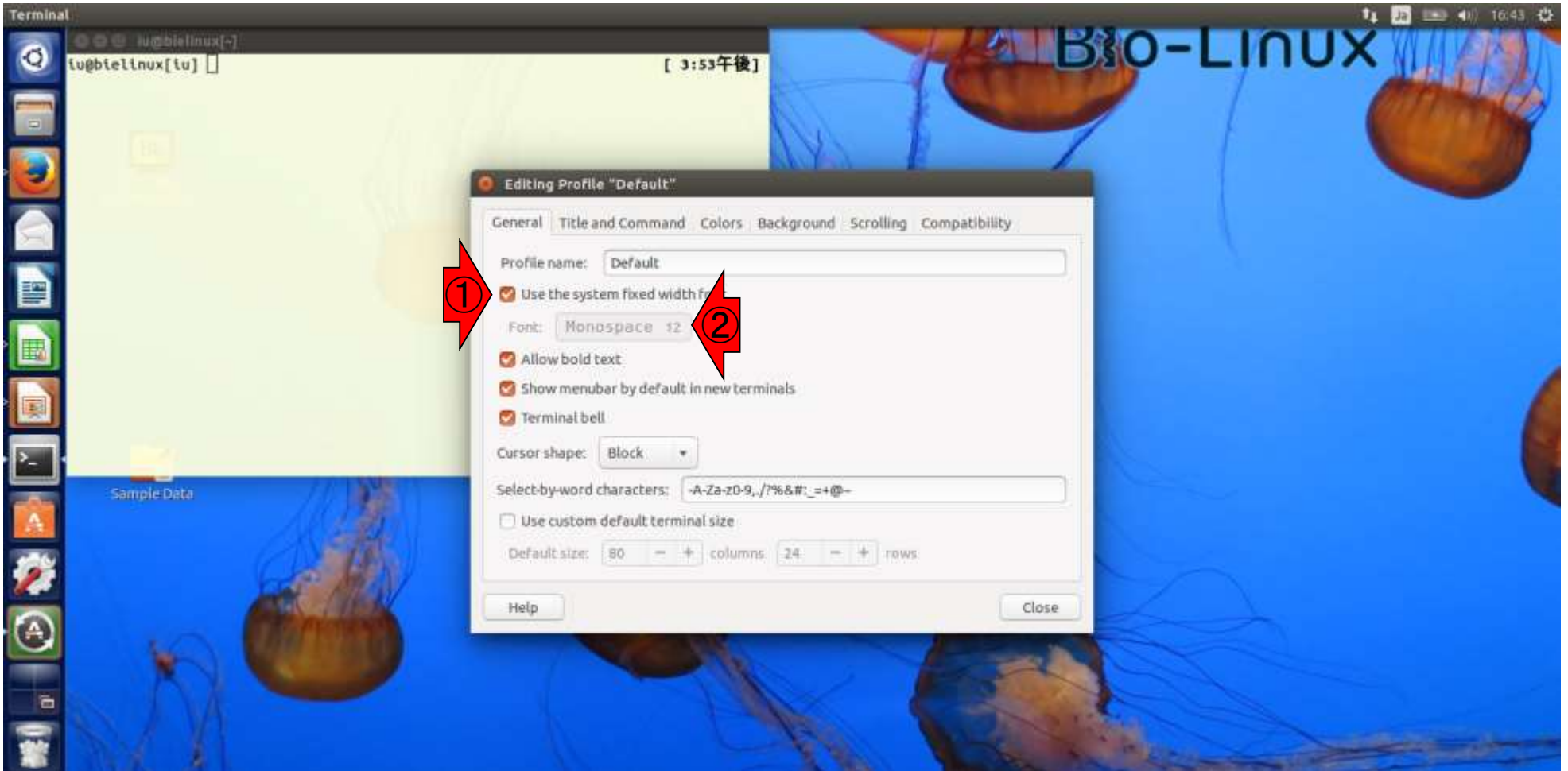


W8-4



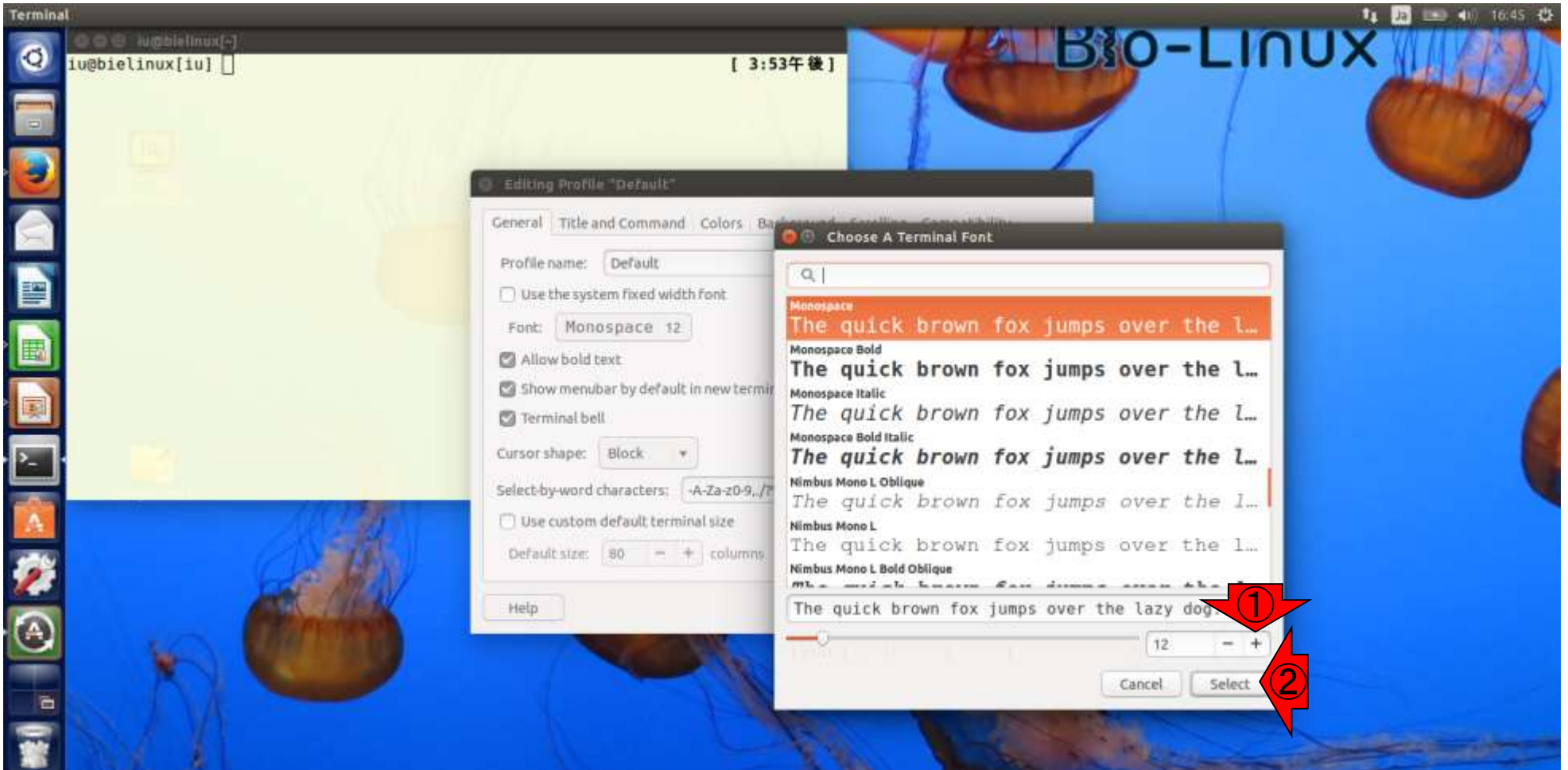
W8-4

①「Use the system fixed width font」のチェックを外し、②「Monospace 12」のあたりをクリック



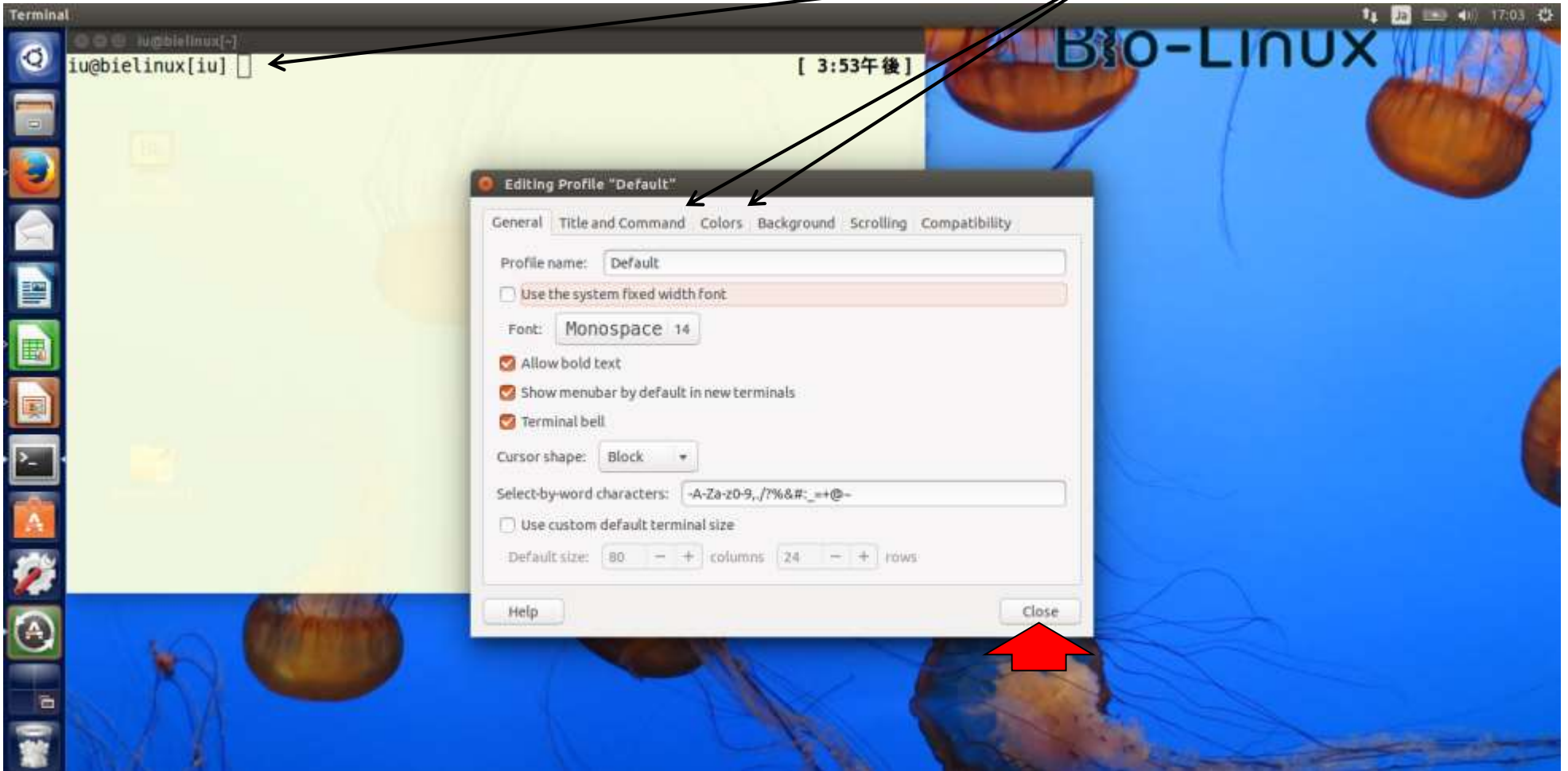
W8-4

フォントを文字の大きさを選択できます。ここでは、①+を2回押して大きさを14にしてから、②Selectボタンを押します。



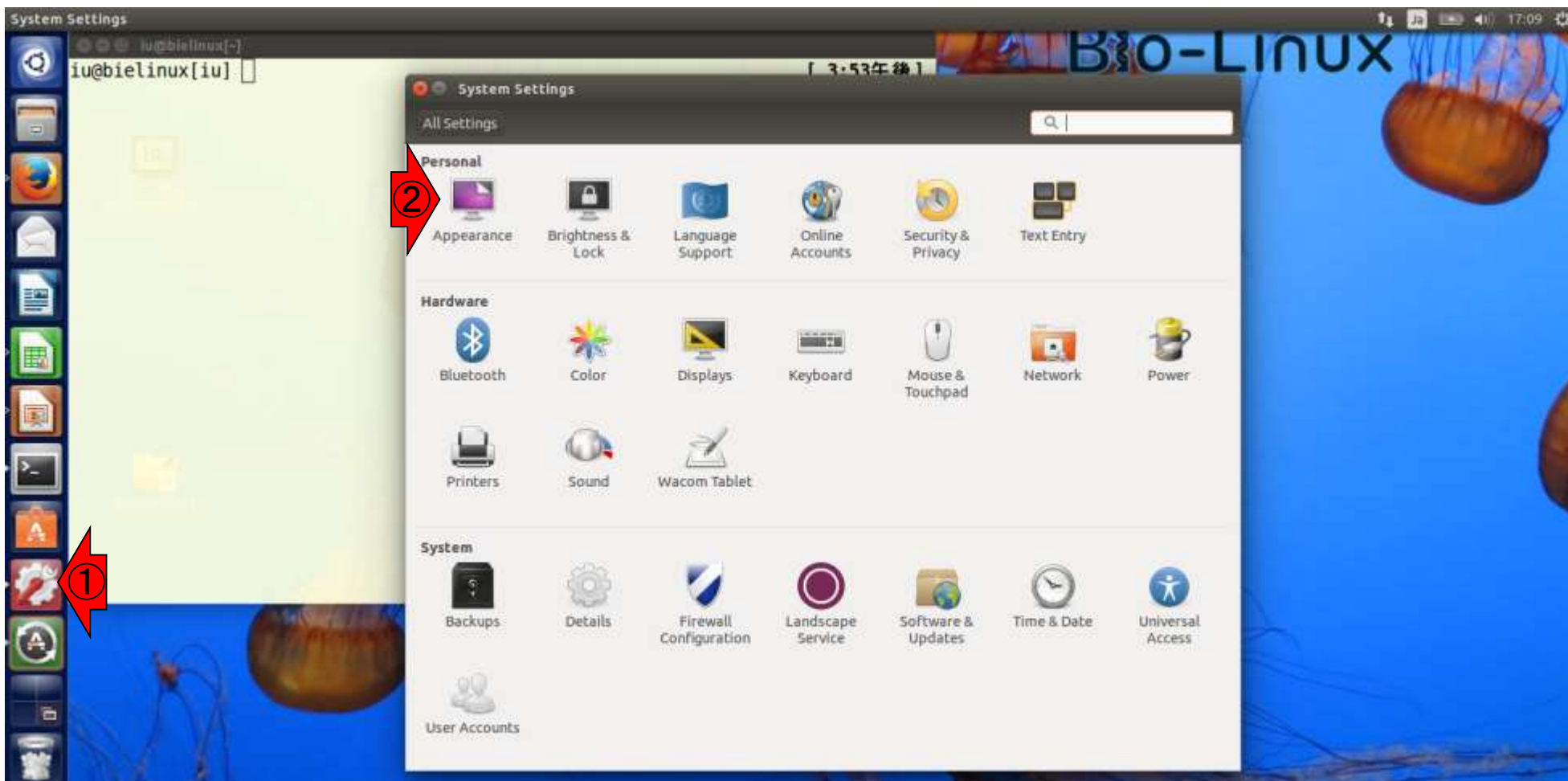
W8-4

確かに文字が大きくなったことがわかります。他にもTitle and CommandやColorsなどいくつか設定変更できることがわかる。



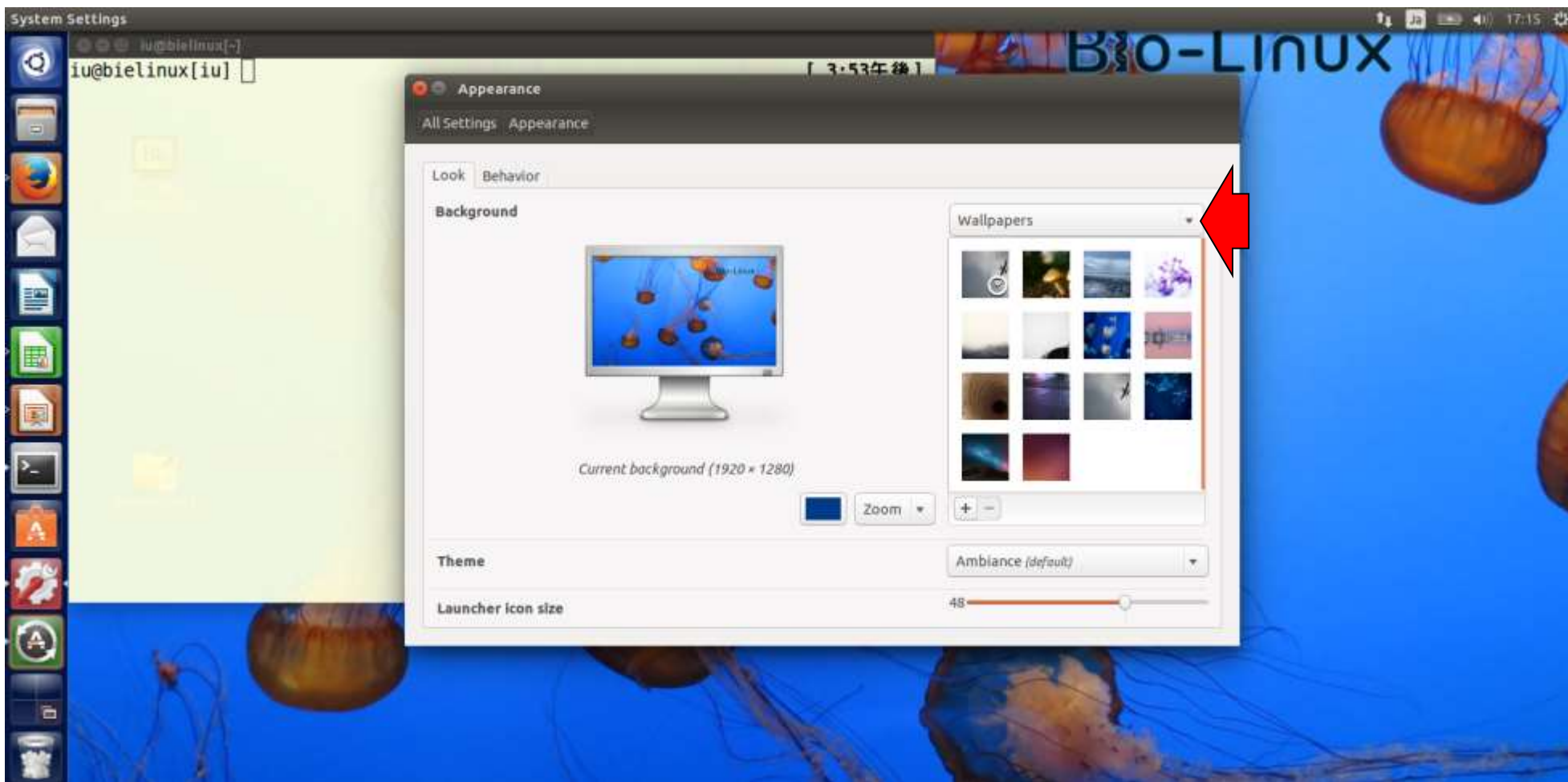
W8-5

ターミナル画面以外にも、Bio-Linux本体の設定も変更できます。①System Settings、②Appearanceをクリック。

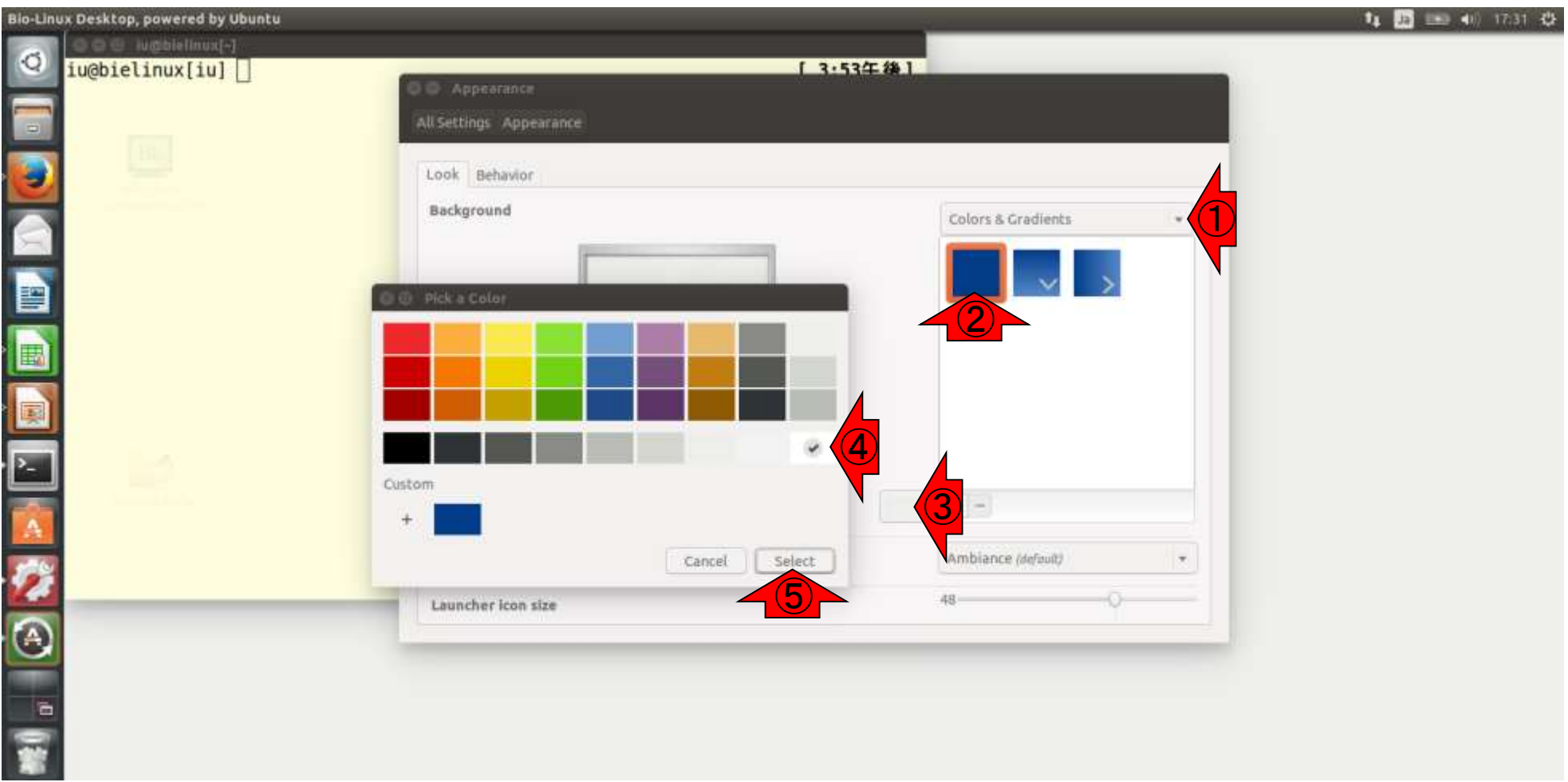


W8-5

「Wallpapers」の中から白っぽいものを選んでよいが、ここでは、「Colors & Gradients」から辿ってゆき、真っ白にする。

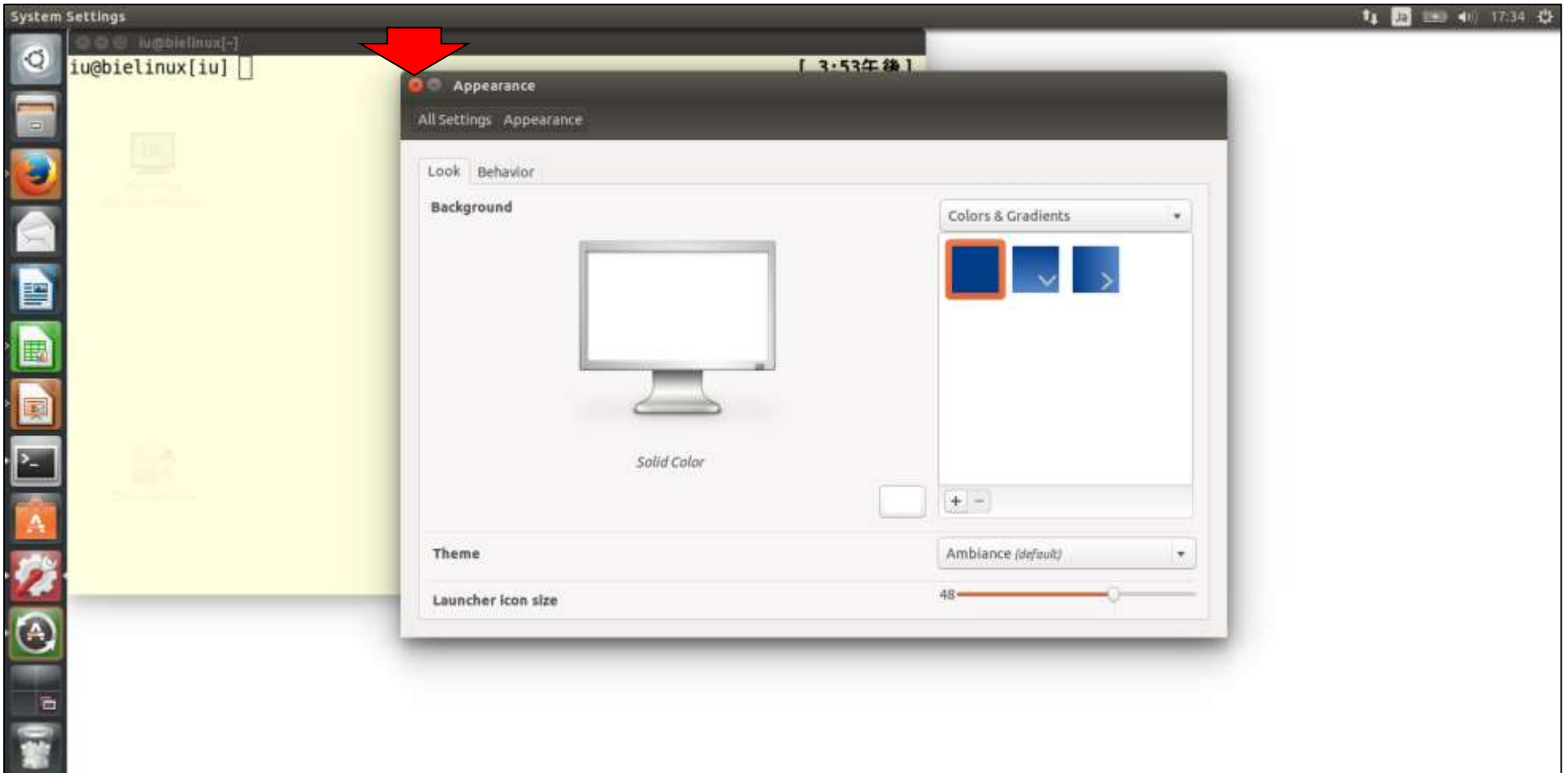


W8-5



背景が白に変更されたことがわかる。

W8-5



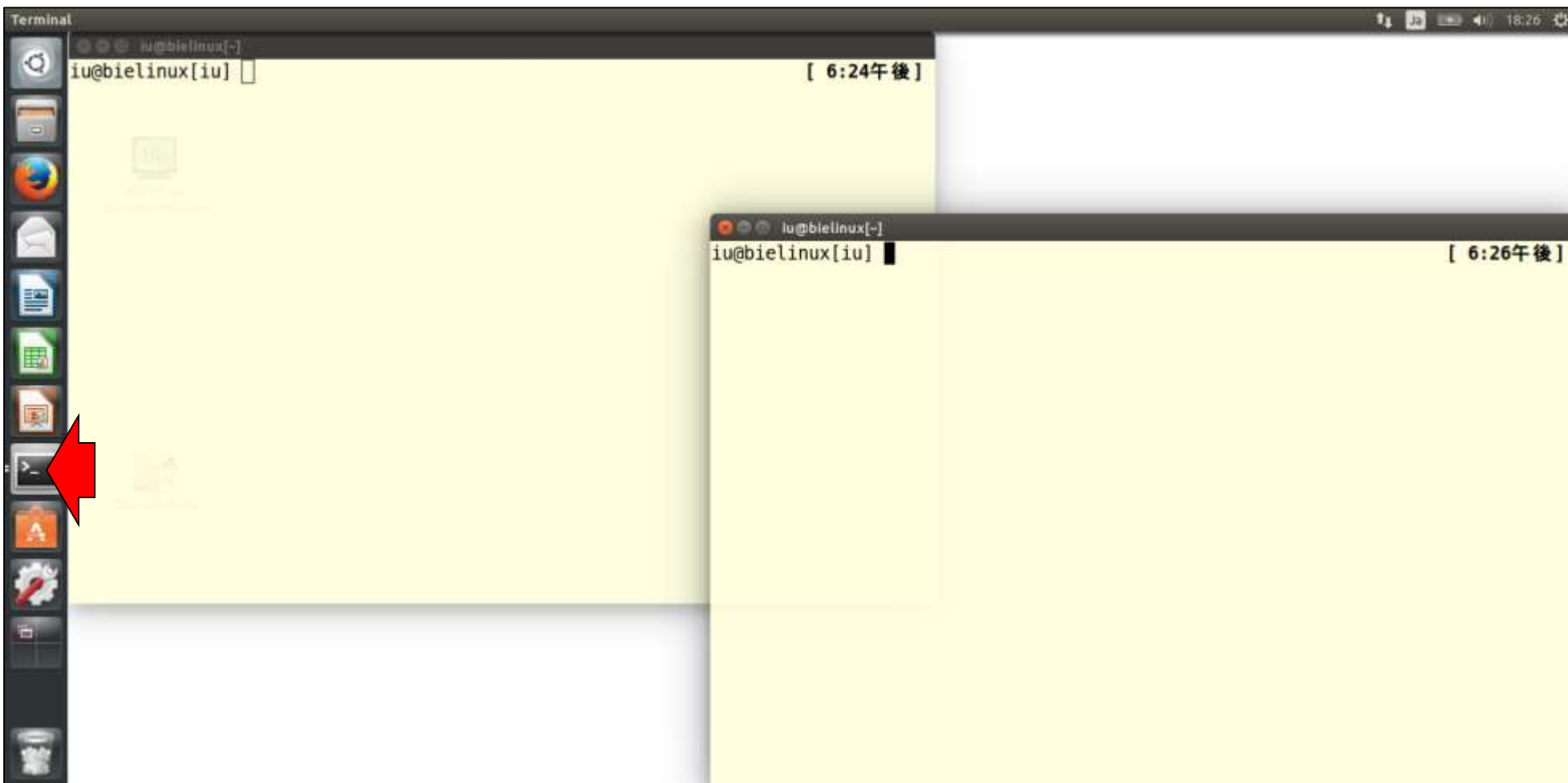
「File」 - 「Open Terminal」で新しいターミナルを開くことができる。

W8-6



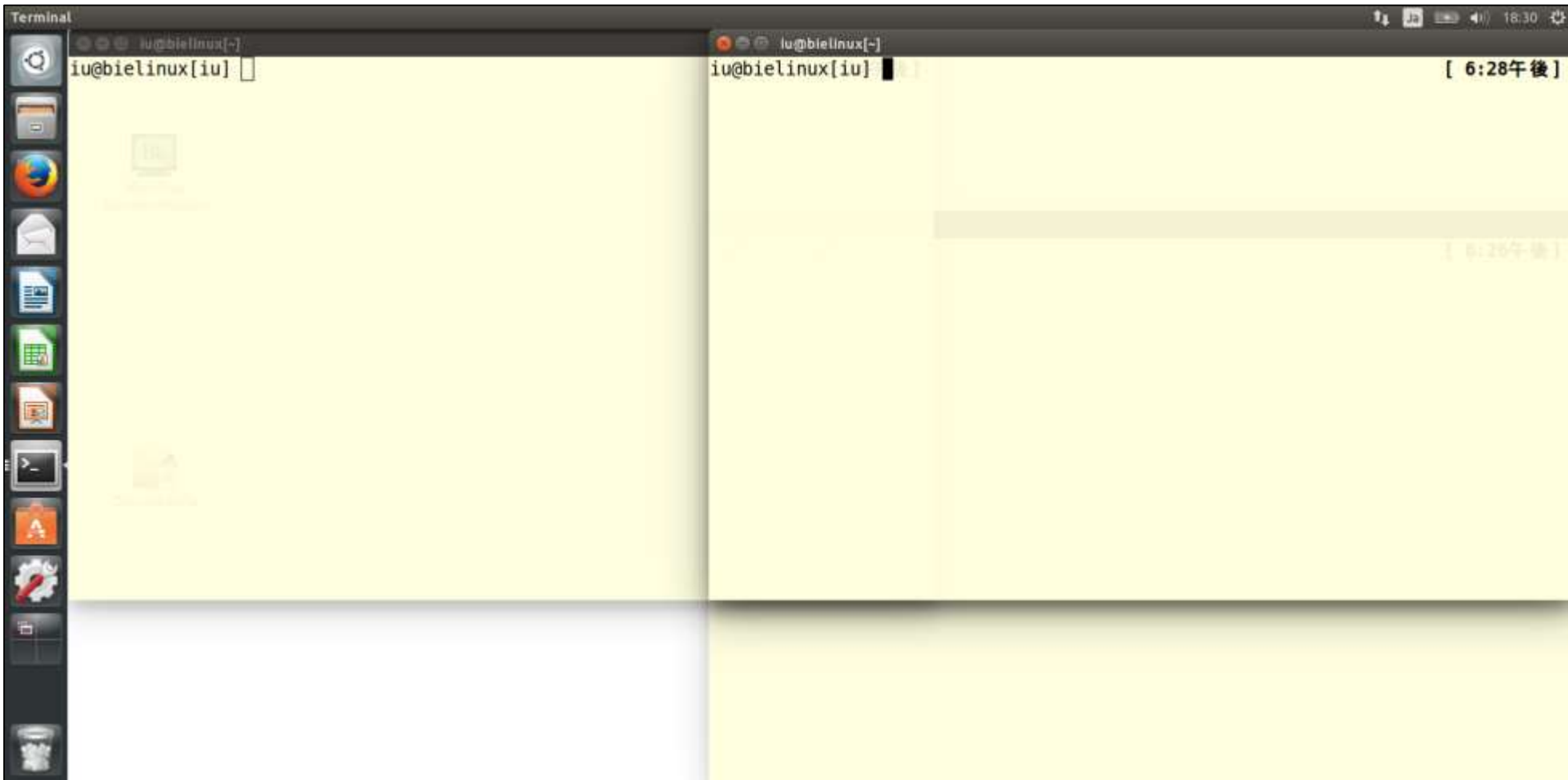
W8-6

赤矢印のターミナルアイコン上で右クリックして、「New Terminal」を選択するのもよい。



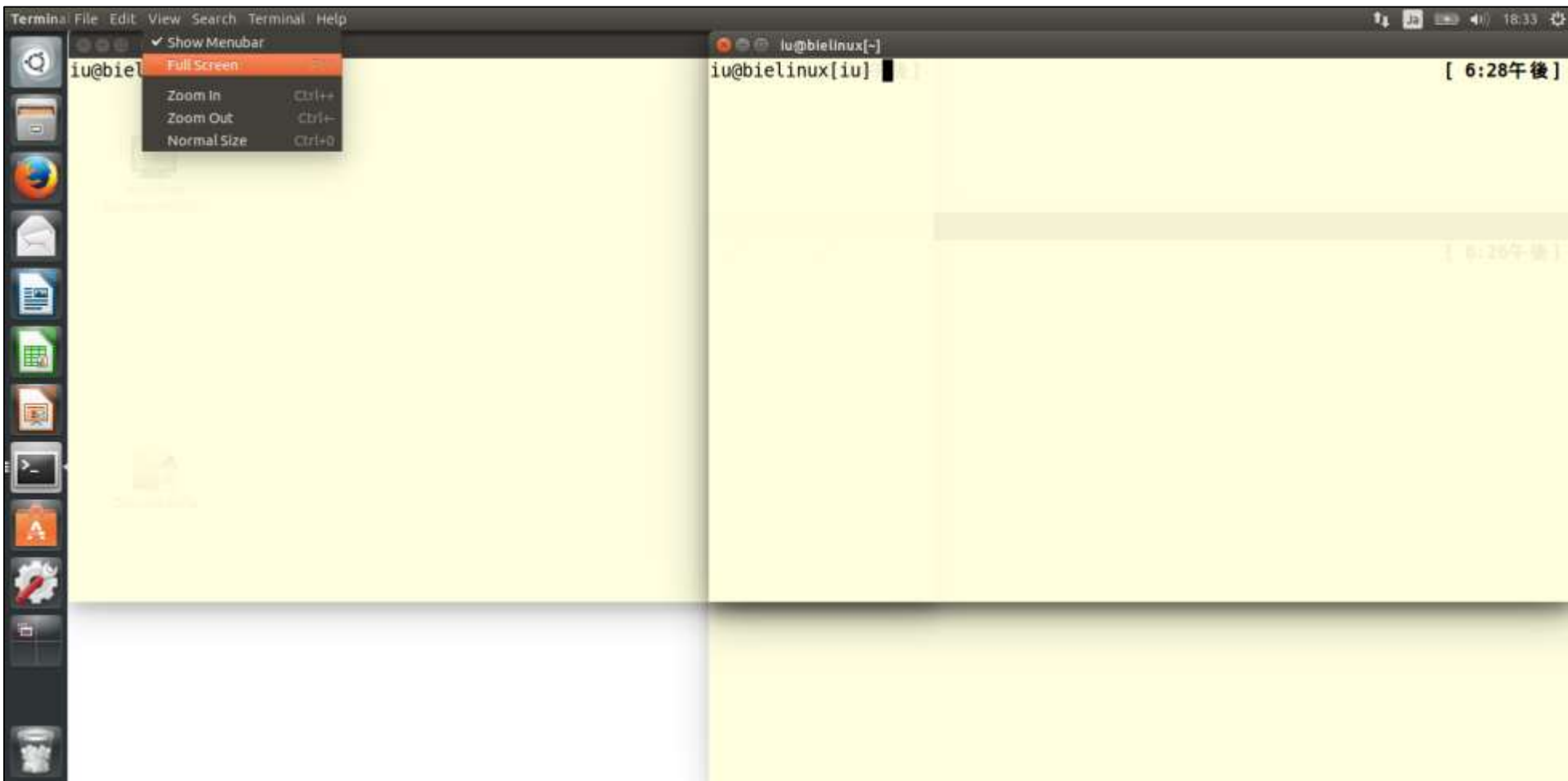
W8-6

このように複数のターミナルを同時に開いて作業を行うのが一般的。



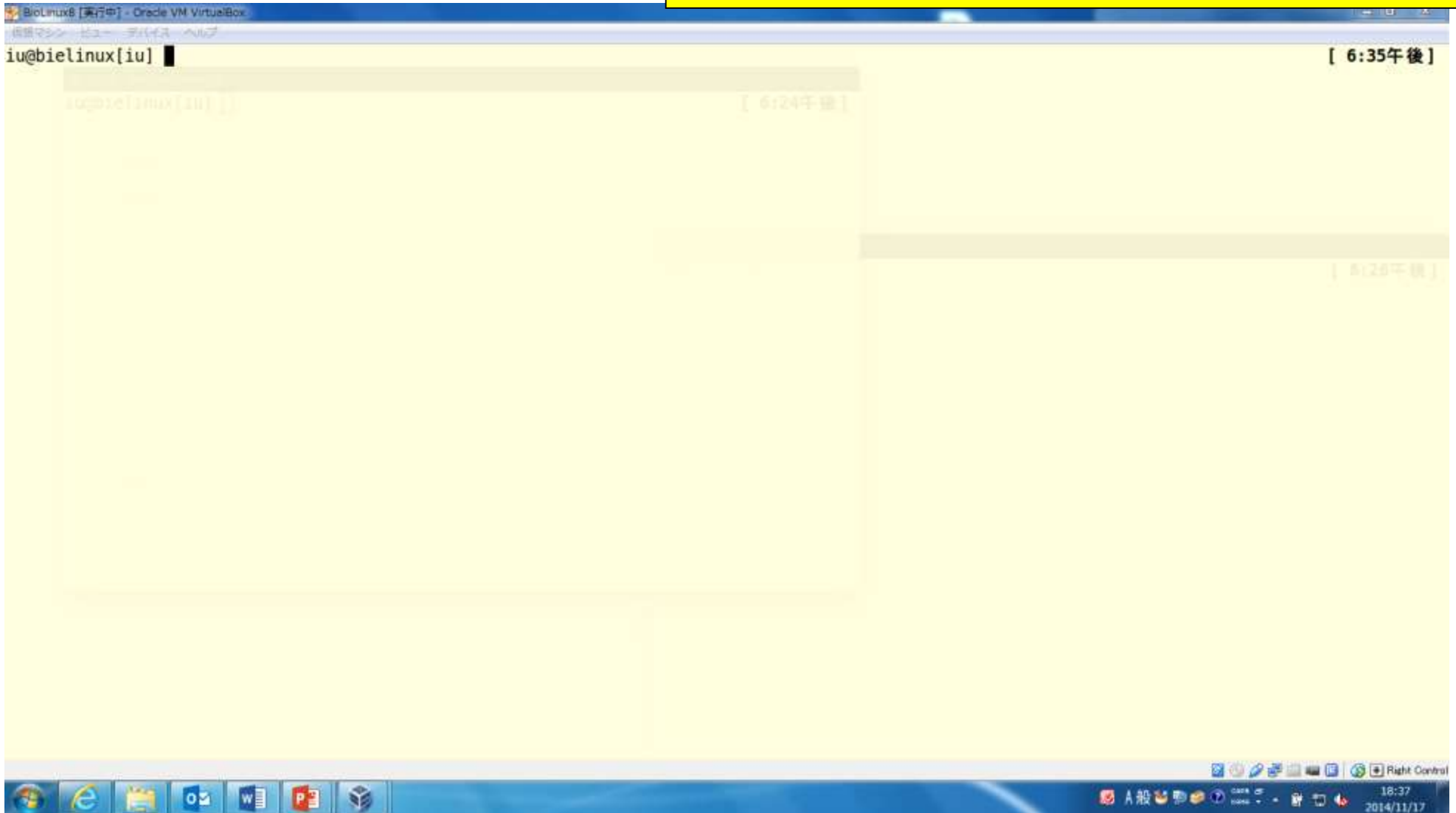
W8-7

メニューバーをいろいろいじってみてもよい。例えば「View」-「Full Screen」。



W8-7

これはホストOS上でスクリーンショットをとったもの。このような状況からどのようにして元に戻せばいいのか？一つの対処法は、右クリックで「Close Window」。



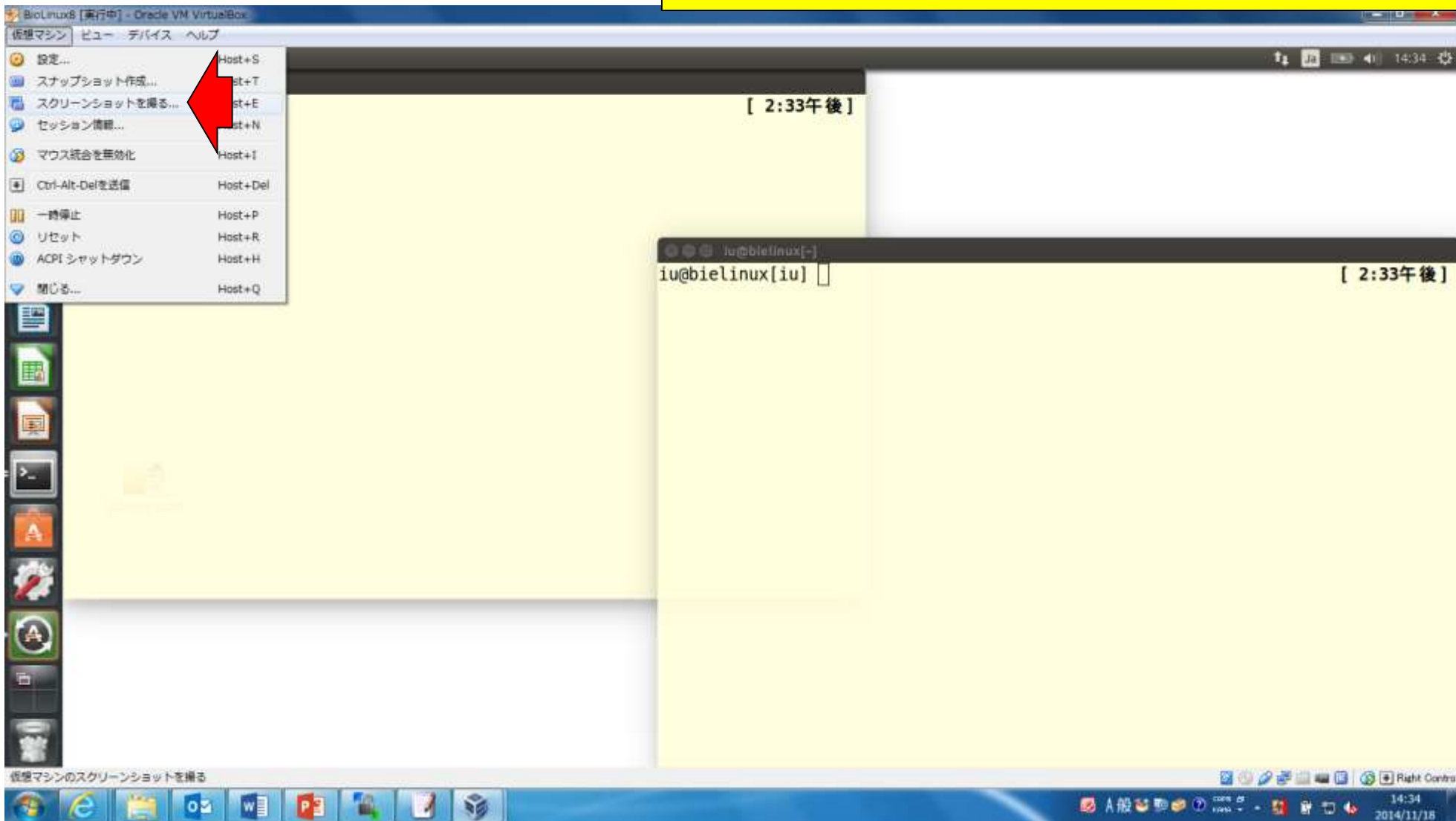
W8-7

該当ターミナルはクローズされますが、一応元の画面に戻ります。



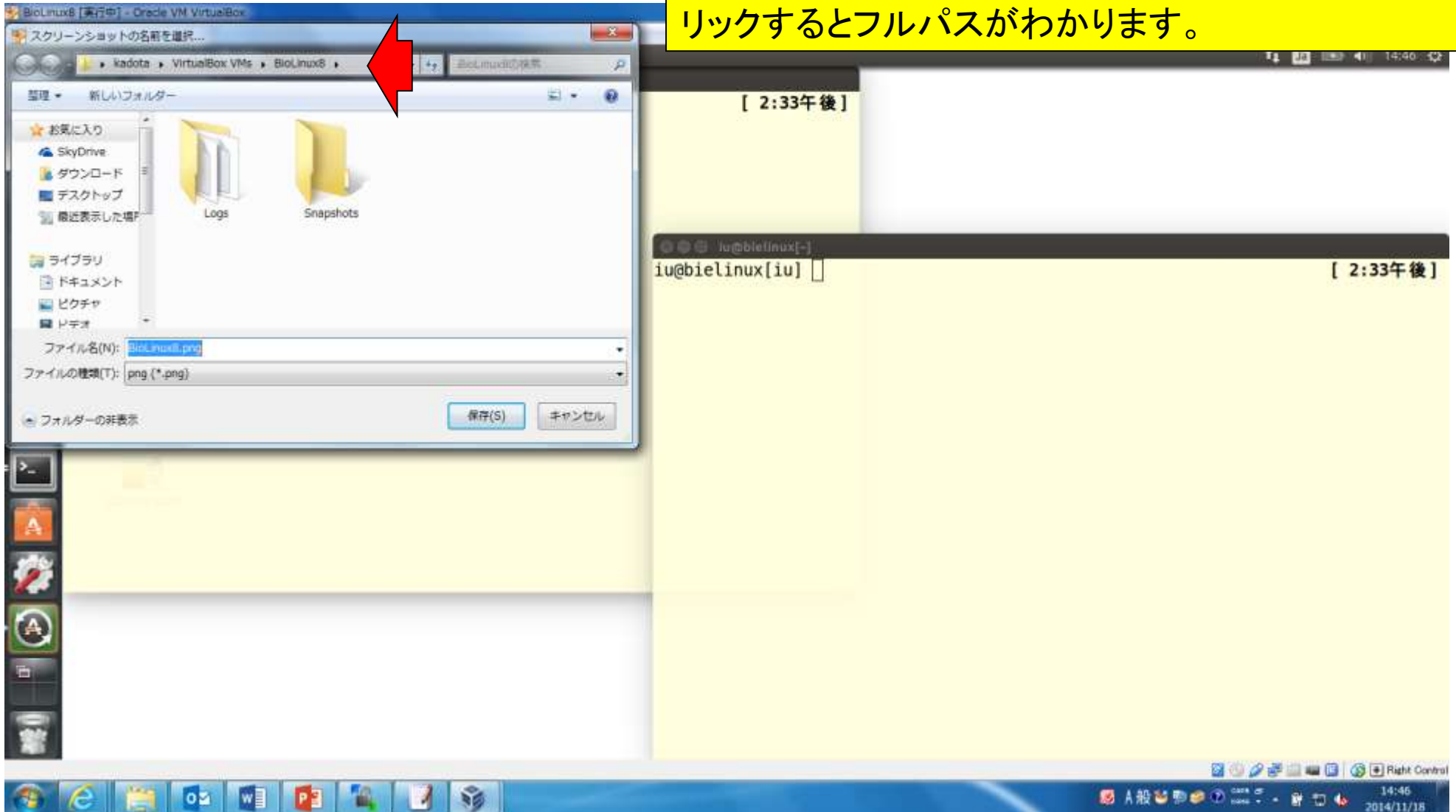
W8-8

「仮想マシン」-「スクリーンショットを撮る」の手順でやると、ゲストOSのスクリーンショットのpng形式ファイルをホストOS中の任意のフォルダに保存することができます。



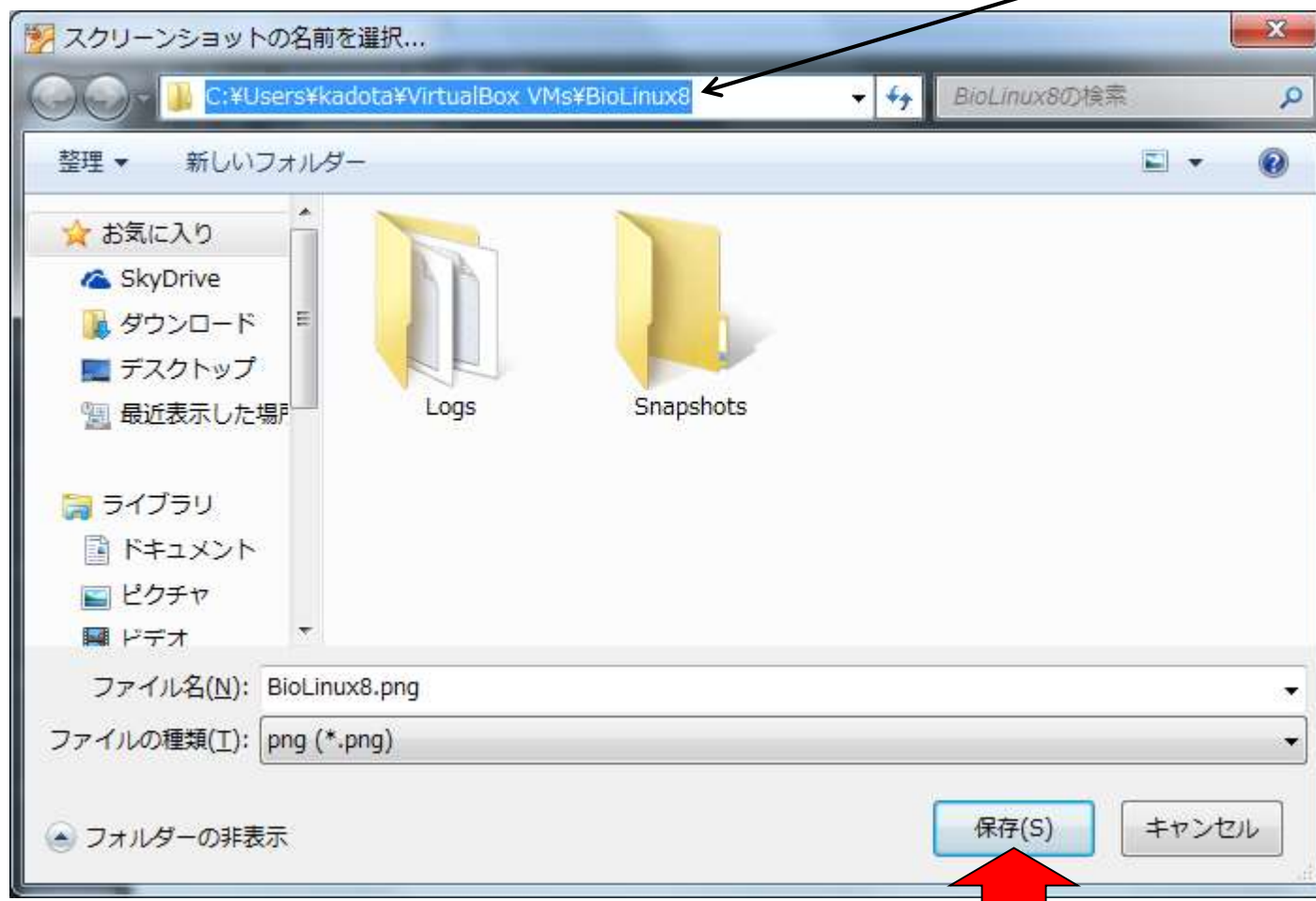
W8-8

デフォルトの保存先は、ホストOSのログインユーザ(この場合kadota)の個人用フォルダ直下にある「VirtualBox VMs」-「BioLinux8」です。赤矢印付近をクリックするとフルパスがわかります。



W8-8

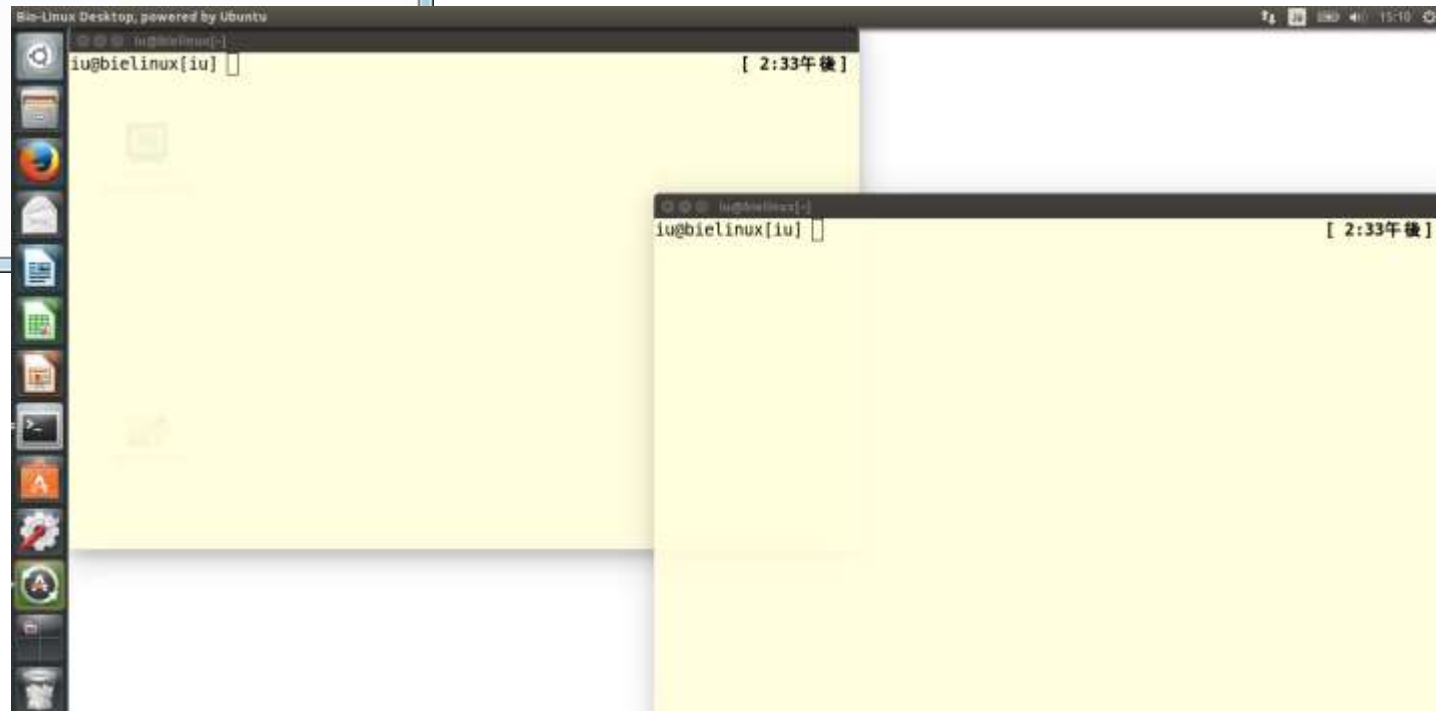
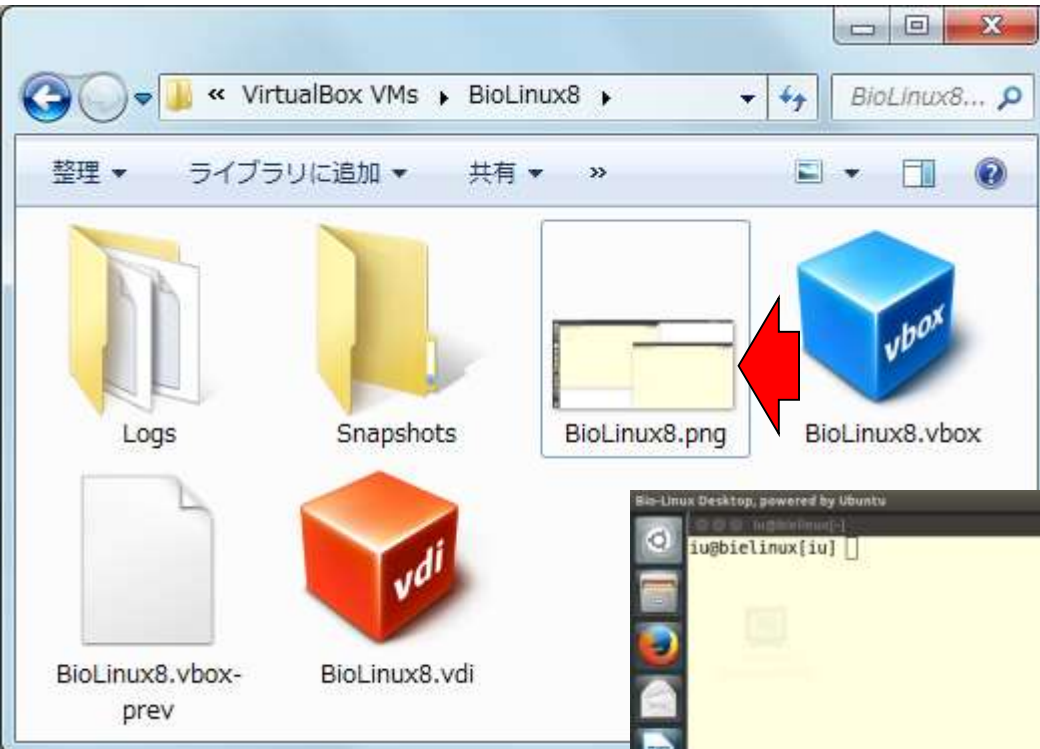
デフォルトの保存先は、ホストOSのログインユーザ(この場合kadota)の個人用フォルダ直下にある「VirtualBox VMs」-「BioLinux8」です。赤矢印付近をクリックするとフルパスがわかります。



W8-8

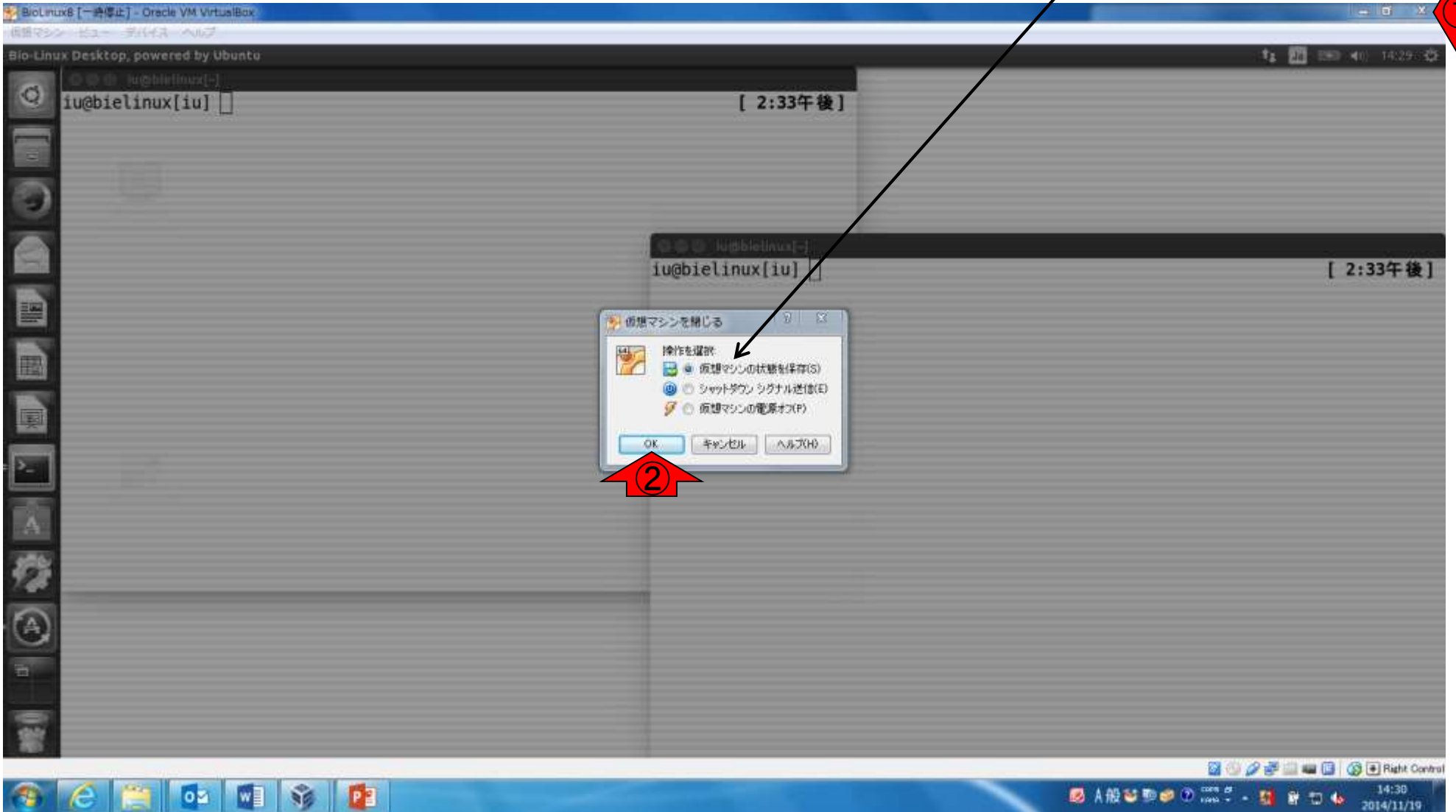


W8-8



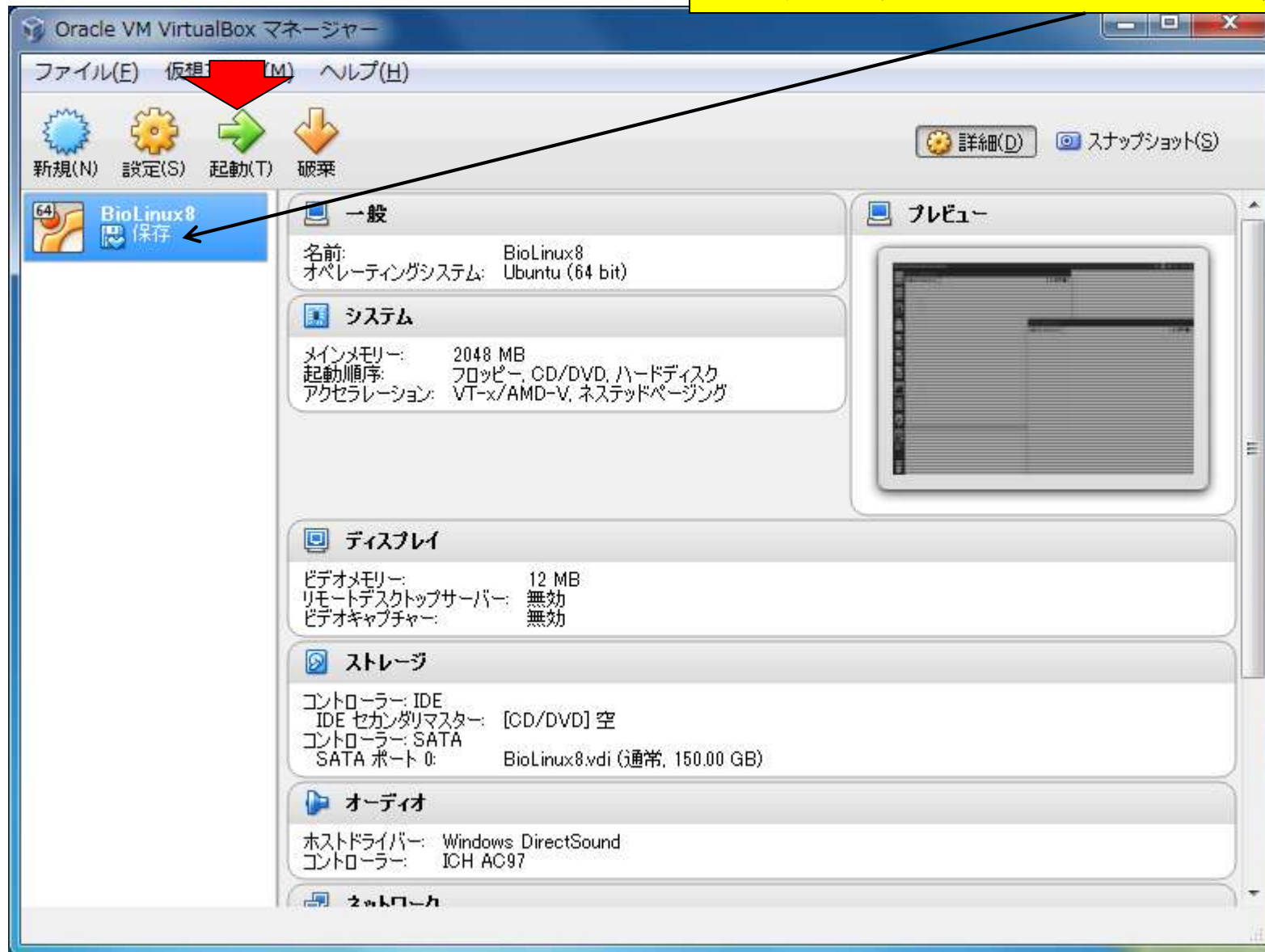
W8-9

ゲストOSの終了時に「仮想マシンの状態を保存」にして
おけば、次回起動時に続きから作業できるので便利。



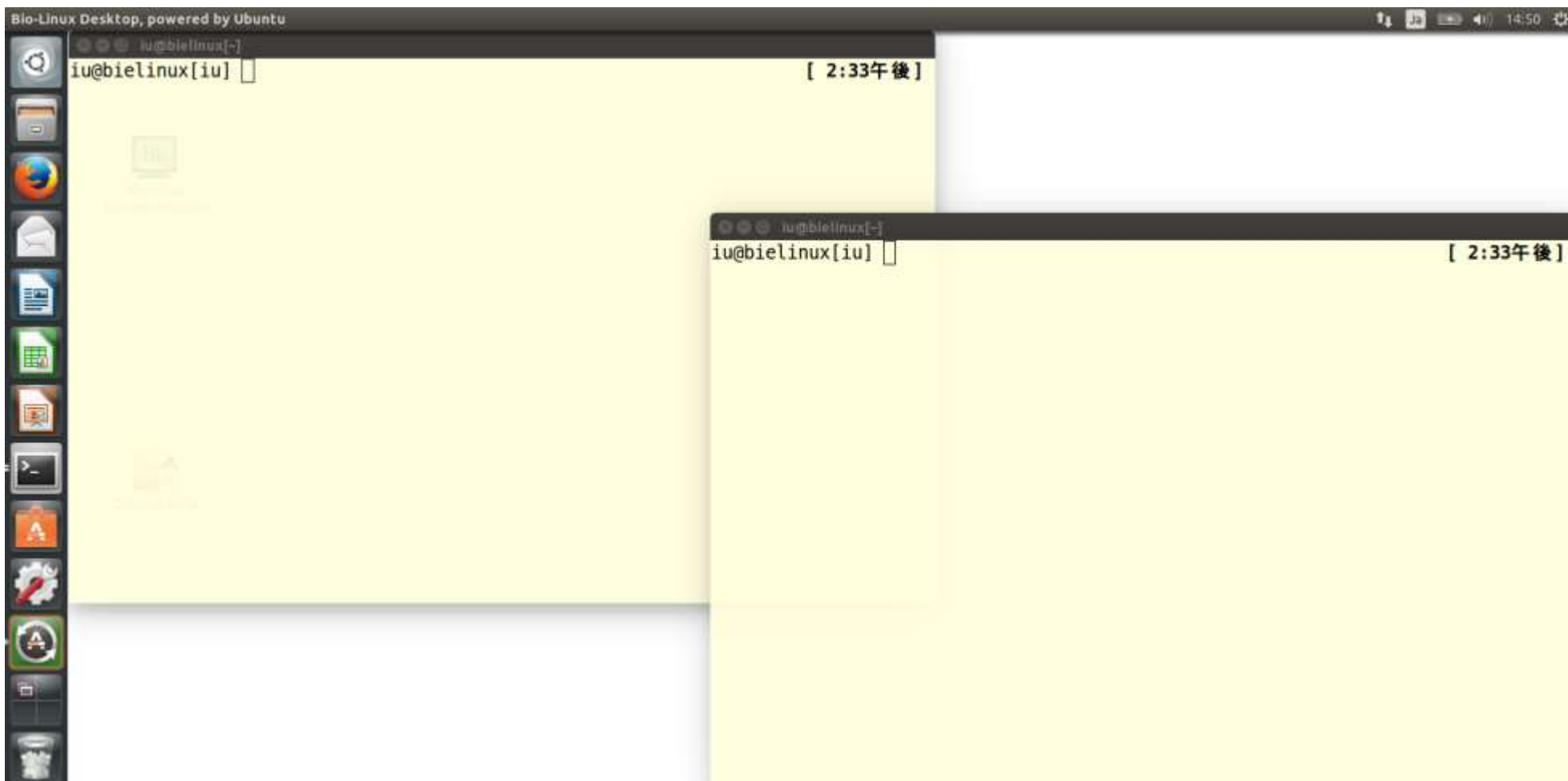
W8-9

ゲストOS終了直後のVirtualBoxマネージャー。起動ボタンを押すと、終了前と同じ(全画面表示でターミナルが2つ立ち上がった)状態になる。



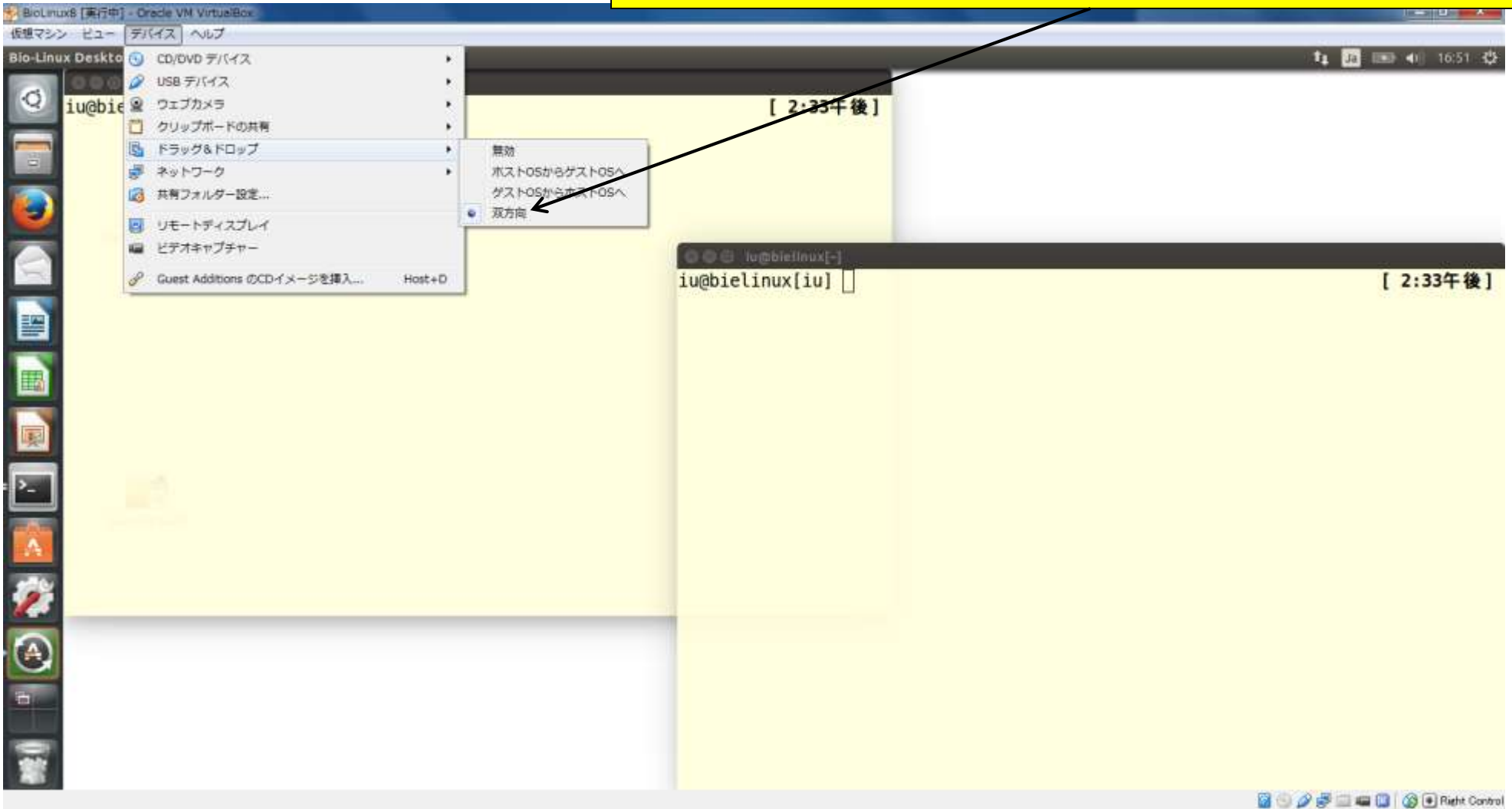
W8-9

起動ボタンを押すと、終了前と同じ(全画面表示でターミナルが2つ立ち上がった)状態になる。



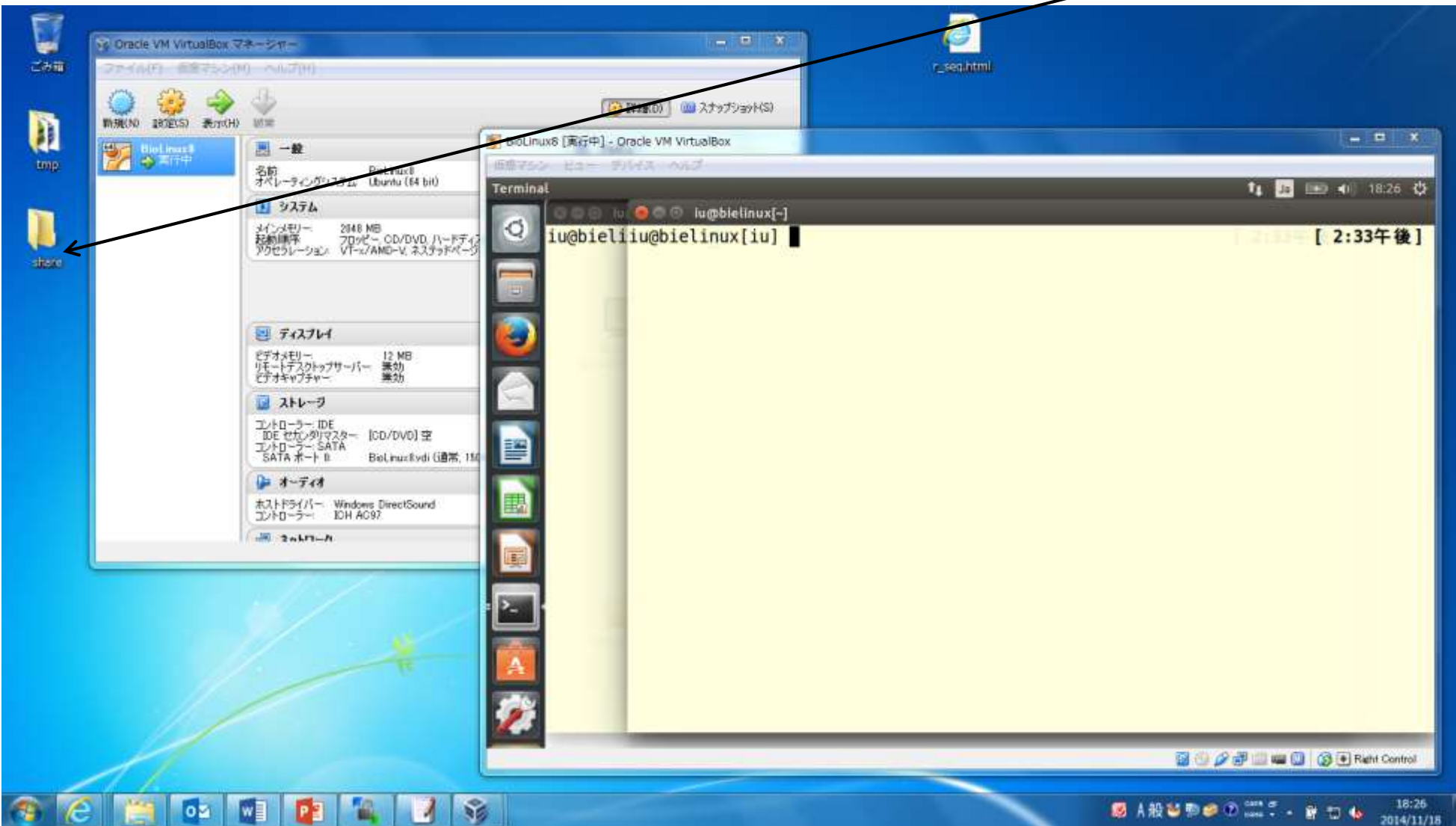
W9-1

VirtualBoxウィンドウメニューの「デバイス」-「ドラッグ&ドロップ」でも設定変更可能。VirtualBox ver. 4.3.18現在、「双方向」を選択していても「ゲストOSからホストOS」は不可能…。



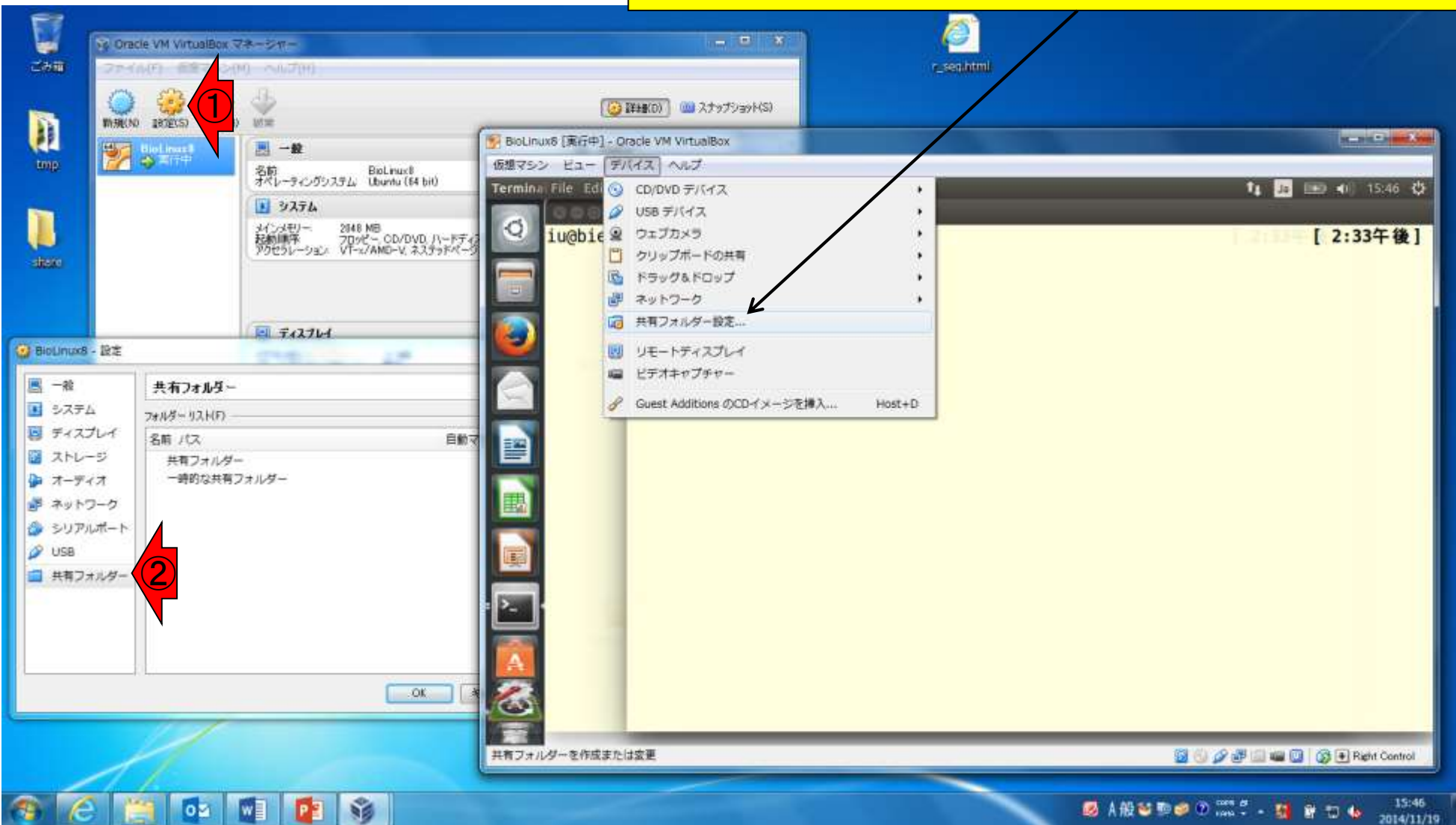
共有フォルダの設定を行う。ホストOSのデスクトップ上にshareフォルダを新規作成。

W9-2-1



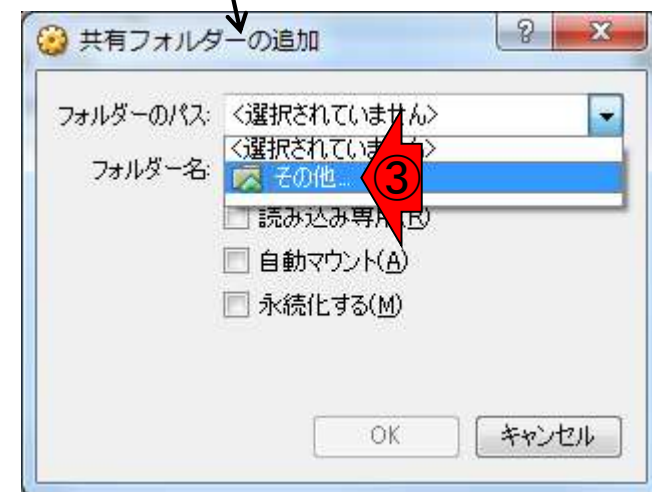
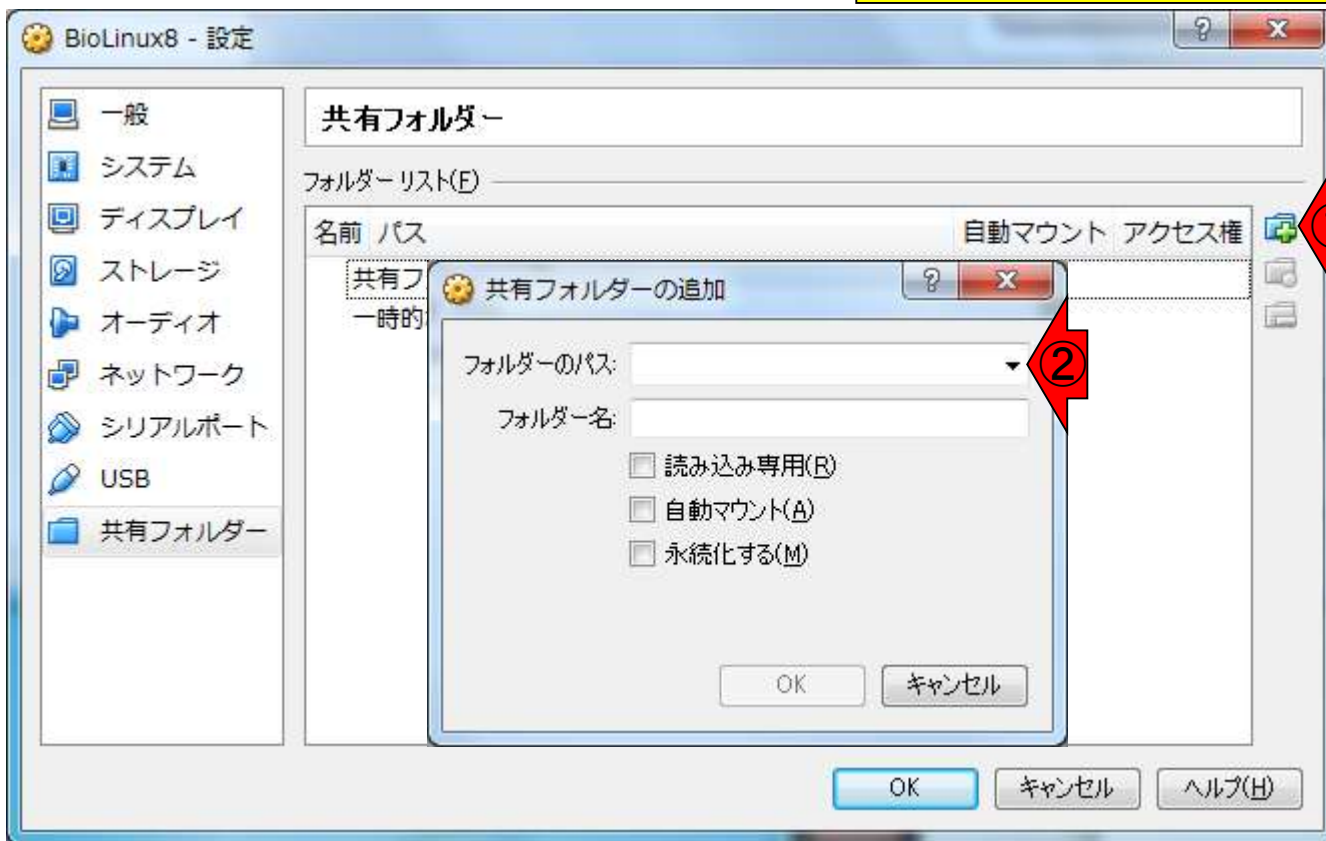
W9-2-2

VirtualBoxマネージャーウィンドウの①設定、②共有フォルダーをクリック。別の手段として、ゲストOS (BioLinux8) ウィンドウ中で「デバイス」-「共有フォルダー設定」でもよい。



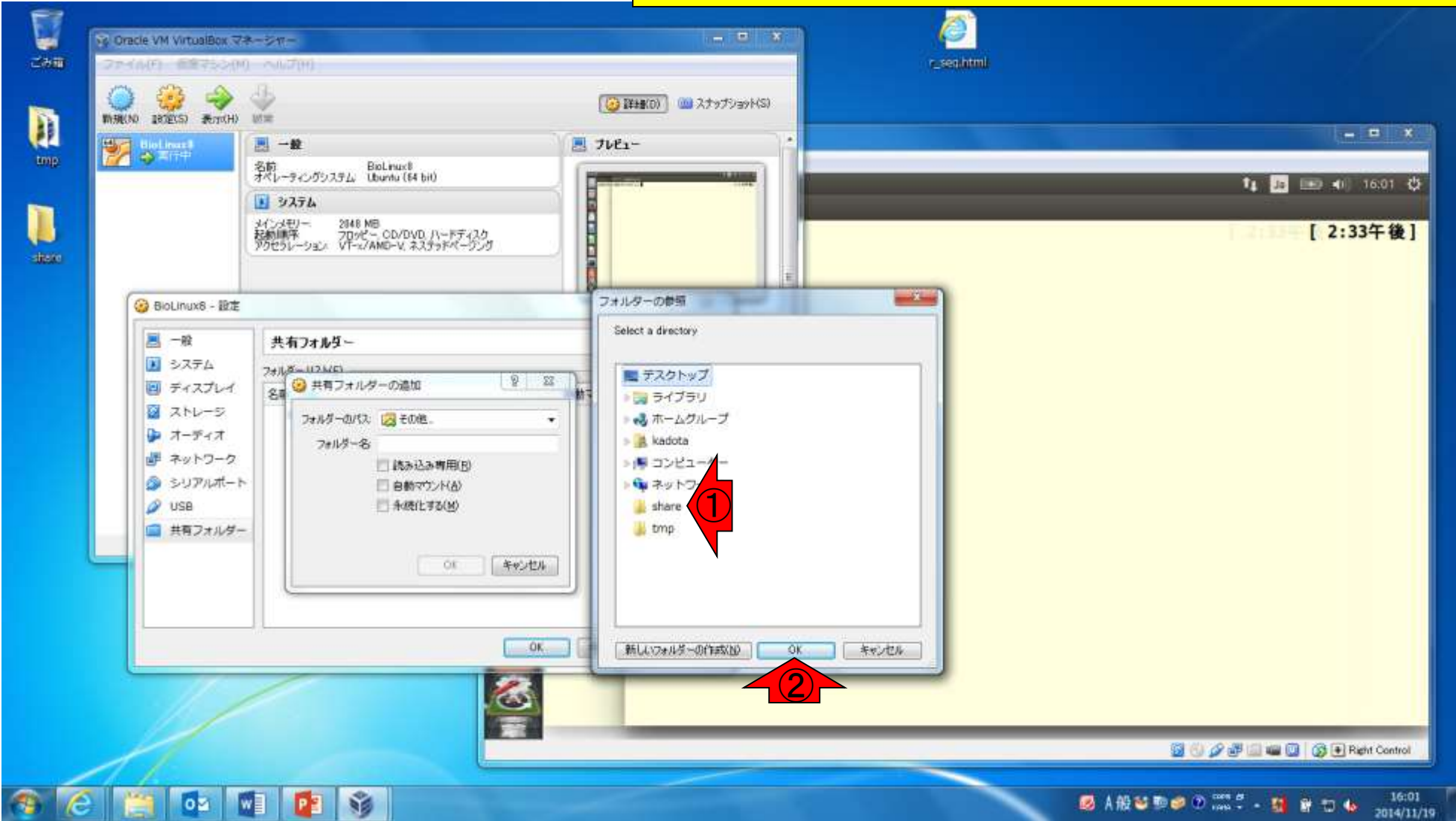
W9-2-2

ホストOSのデスクトップ上に作成したshareフォルダを共有フォルダとするための手続きを行う。「共有フォルダの追加」で、ホストOS上のshareフォルダのパスを指定している。



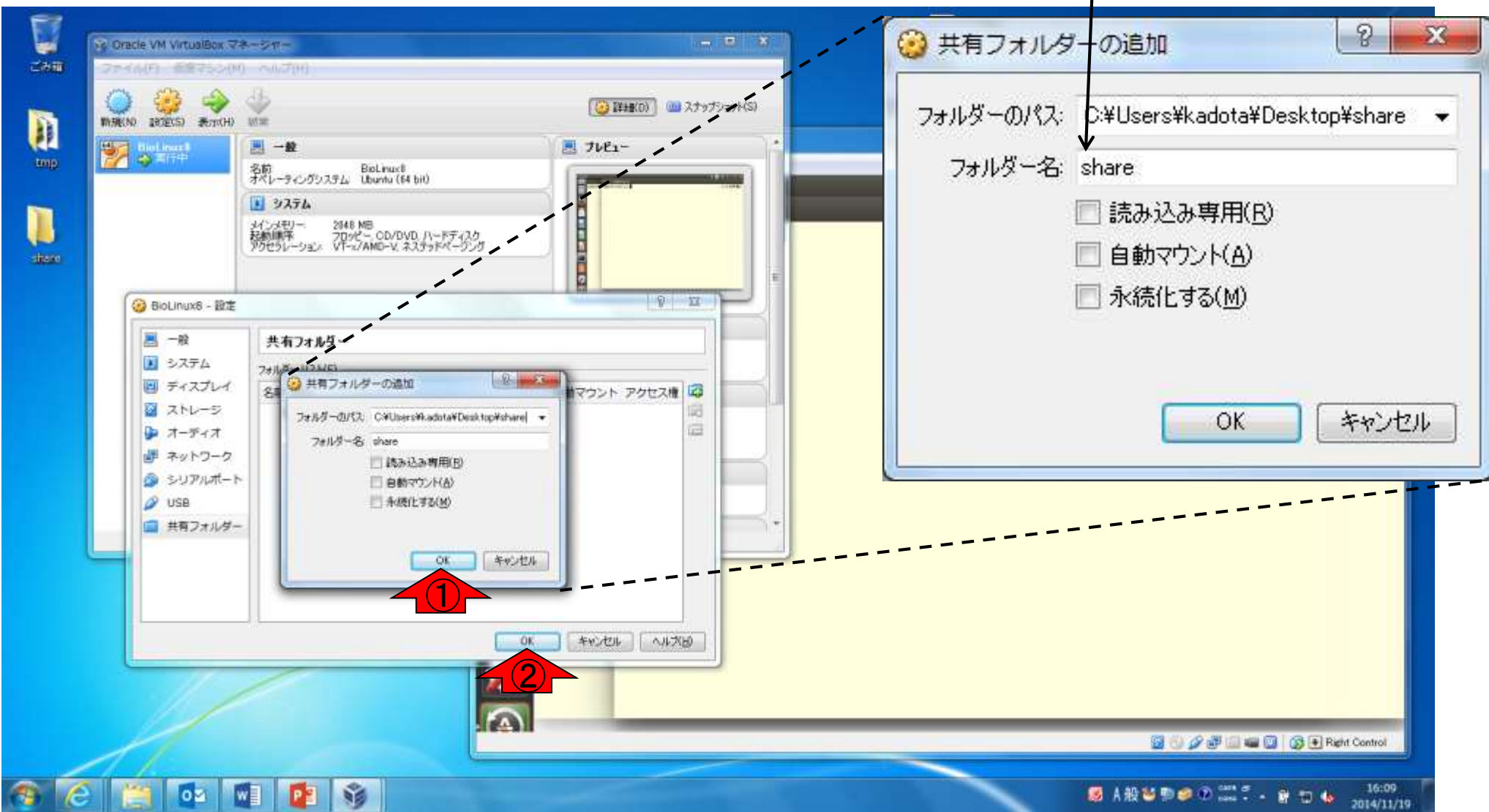
W9-2-2

ホストOSのデスクトップ上に作成したshareフォルダを共有フォルダとするための手続きを行う。「共有フォルダの追加」で、ホストOS上のshareフォルダのパスを指定している。



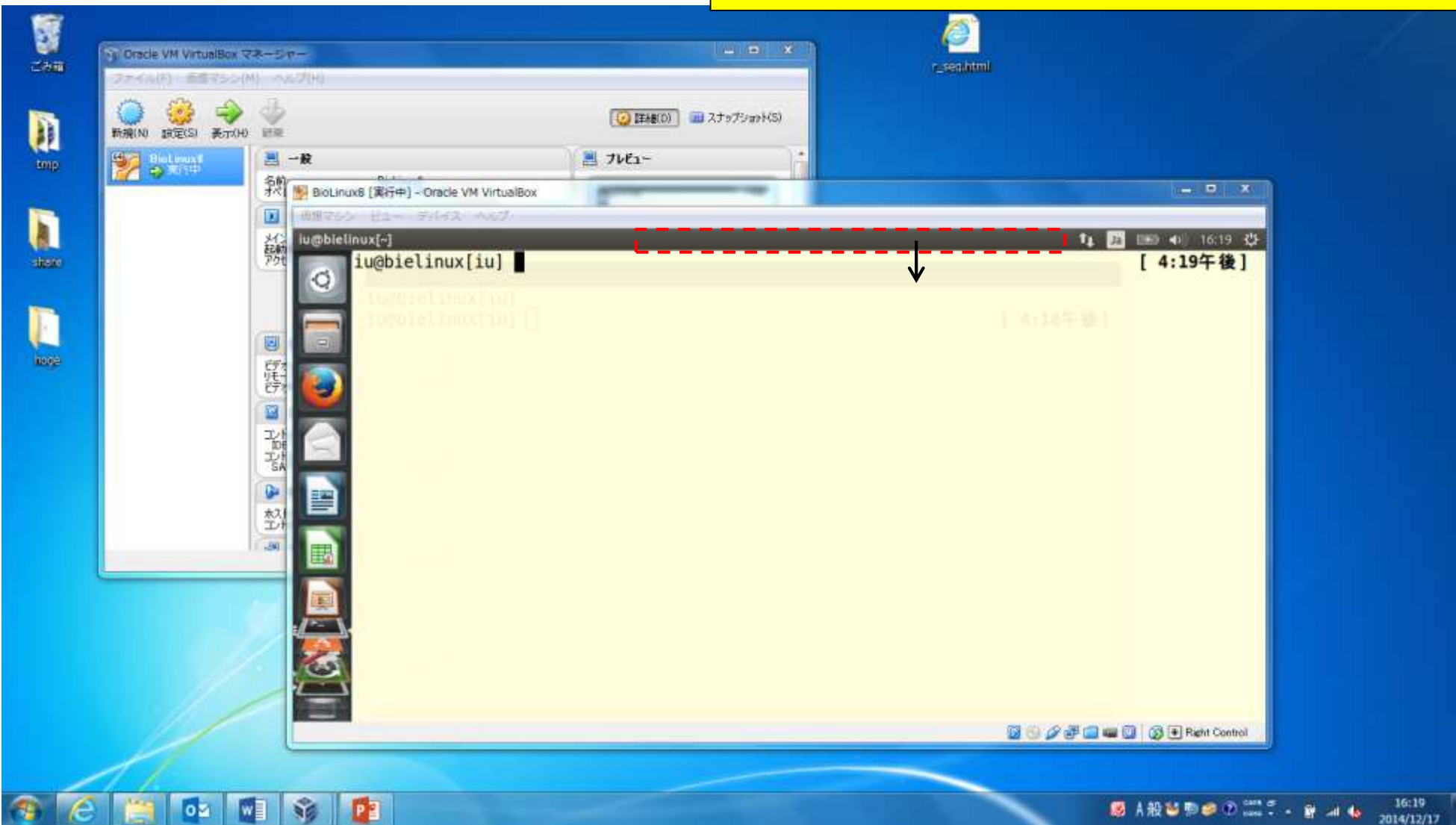
W9-2-2

ホストOS上のshareフォルダのパスを指定すると、「フォルダ名」のところに自動的にshareが入力される。



W9-2-3

Tips。ときどきこのような画面になって他のターミナルが見られなくなったりすることがある。このようになったら、赤枠付近でクリックしたまま下にずらすとよい。



W9-2-3

ゲストOSのターミナル上で、②Desktopに移動し、④mac_shareという名前のフォルダを作成しているところ。それ以外のlsコマンドはただの確認。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bielinux[iu] ls ← ① [10:15午前]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop ← ② [10:16午前]
iu@bielinux[Desktop] ls ← ③ [10:16午前]
Bio-Linux Documentation Sample Data
iu@bielinux[Desktop] mkdir mac_share ← ④ [10:16午前]
iu@bielinux[Desktop] ls ← ⑤ [10:16午前]
Bio-Linux Documentation mac_share Sample Data
iu@bielinux[Desktop] ls -la ← ⑥ [10:16午前]
total 16
drwx----- 4 iu iu 4096 11月 29 10:16 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:16 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
iu@bielinux[Desktop] [10:16午前]
```

現在開いている2つのターミナル画面を「最小化」することでデスクトップを眺めることができる。

W9-2-3

```
iu@bielinux[~/Desktop]
iu@bielinux[iu] ls
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop
iu@bielinux[Desktop] ls
Bio-Linear Documentation Sample Data
iu@bielinux[Desktop] mkdir mac_share
iu@bielinux[Desktop] ls
Bio-Linear Documentation mac_share Sample Data
iu@bielinux[Desktop] ls -la
total 16
drwx----- 4 iu iu 4096 11月 29 10:16 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linear Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:16 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledatal/
iu@bielinux[Desktop]
```

確かにゲストOSのデスクトップ上にmac_shareというフォルダが作成されていることがわかる。

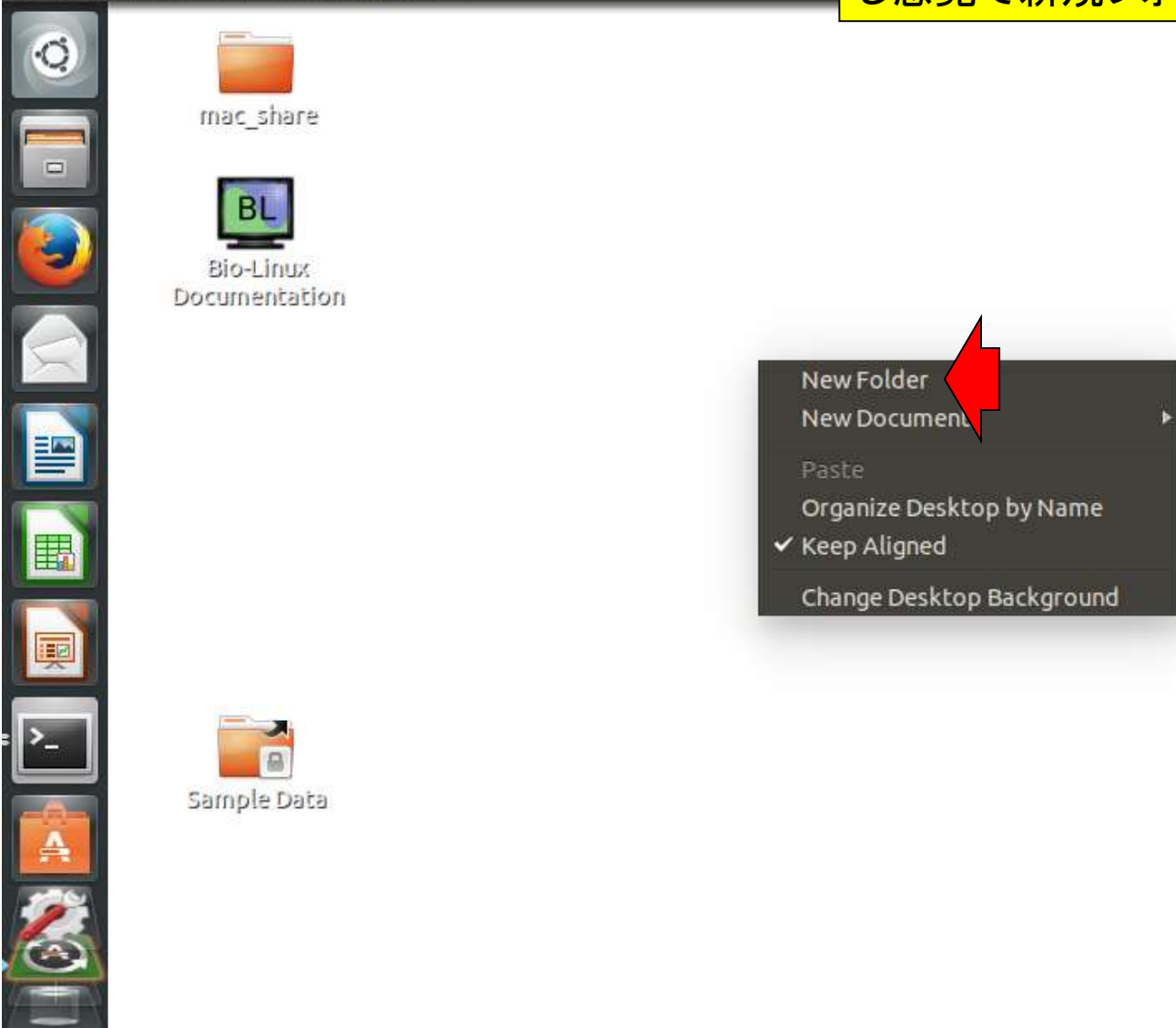
W9-2-3



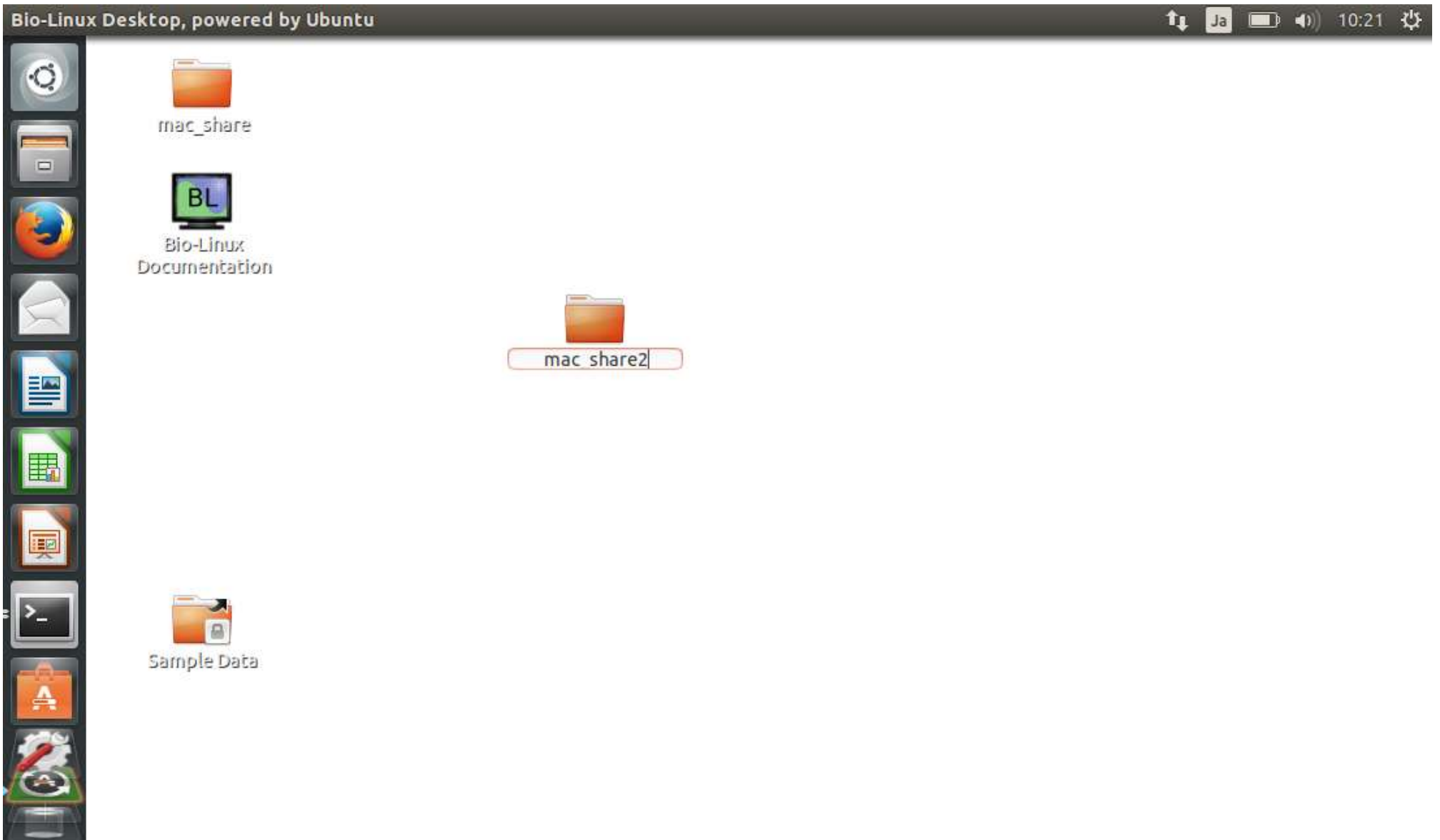
W9-2-3

ターミナル画面上でmkdirコマンドを利用してmac_shareディレクトリを作成したが、もちろんデスクトップ画面上で「右クリック」-「New Folder」で通常のGUI画面と同じ感覚で新規フォルダの作成を行ってもよい。

Bio-Linux Desktop, powered by Ubuntu

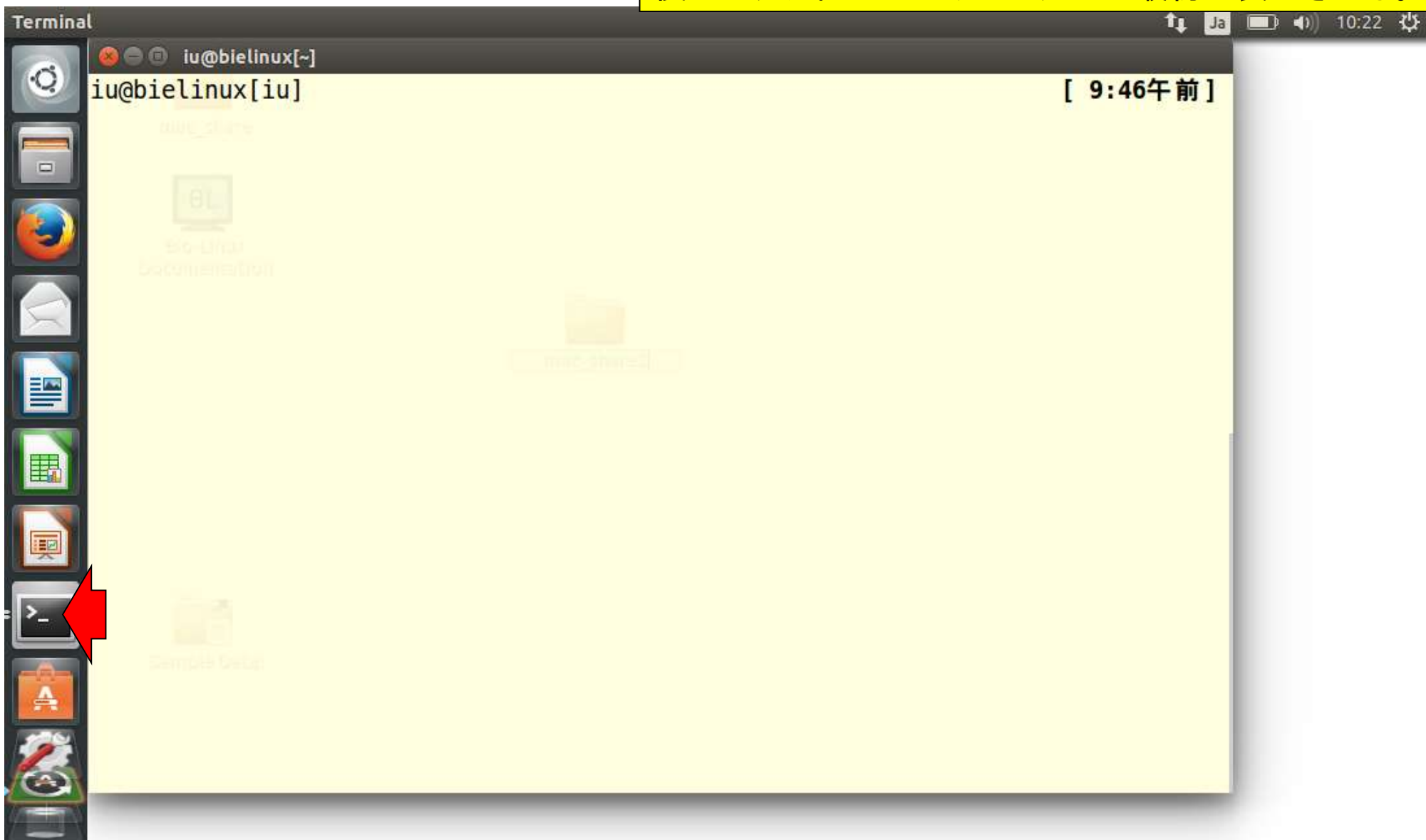


W9-2-3



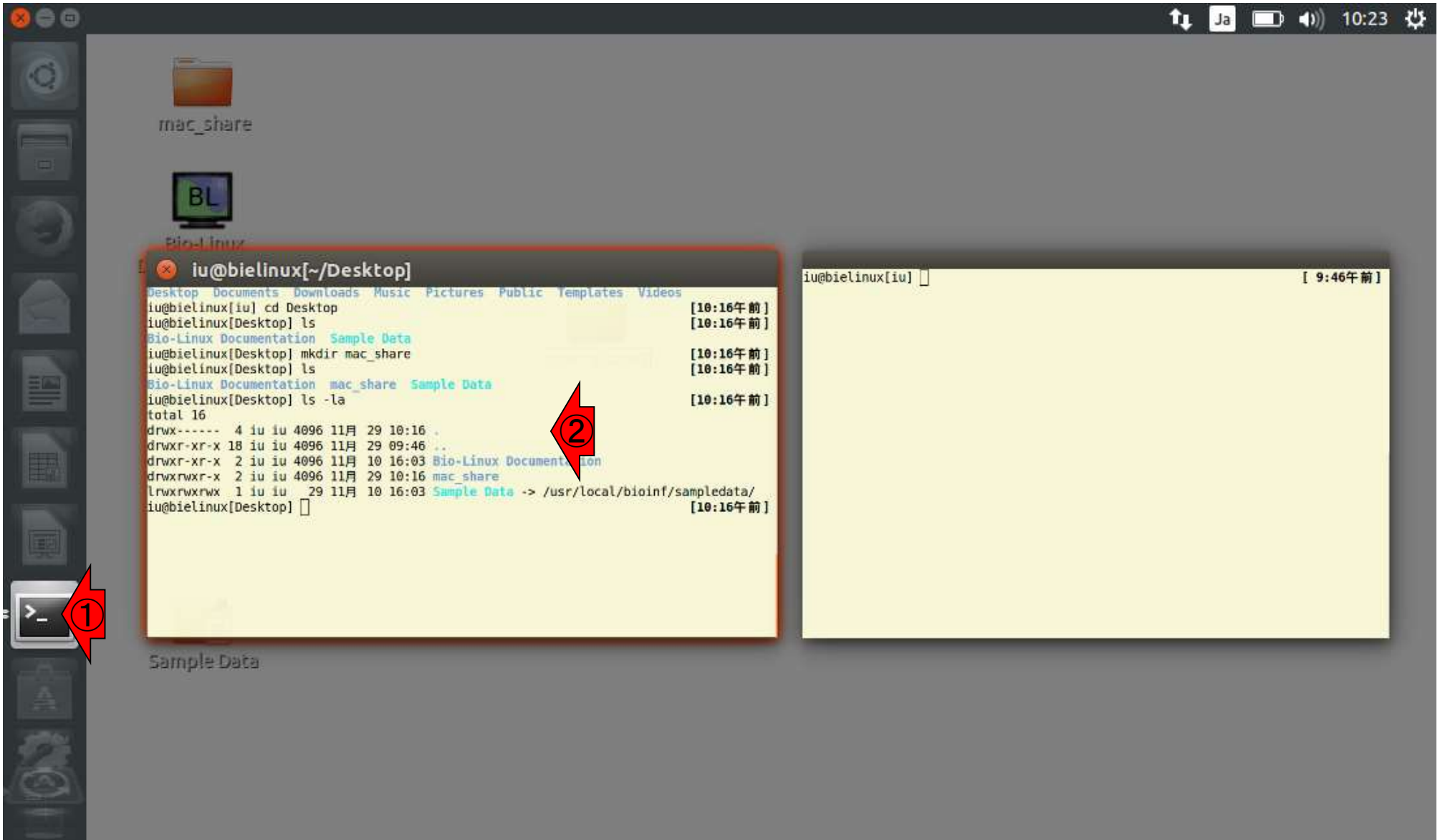
W9-2-4

ターミナルアイコンをクリックすると、最小化されていたターミナルが表示される。ターミナルは2つあったが、最後にアクティブだったターミナルが最初に表示される。



W9-2-4

もう一度ターミナルアイコンをクリックすると、コマンドを打ち込んでいたほうのターミナルも選択可能になる。



W9-2-4

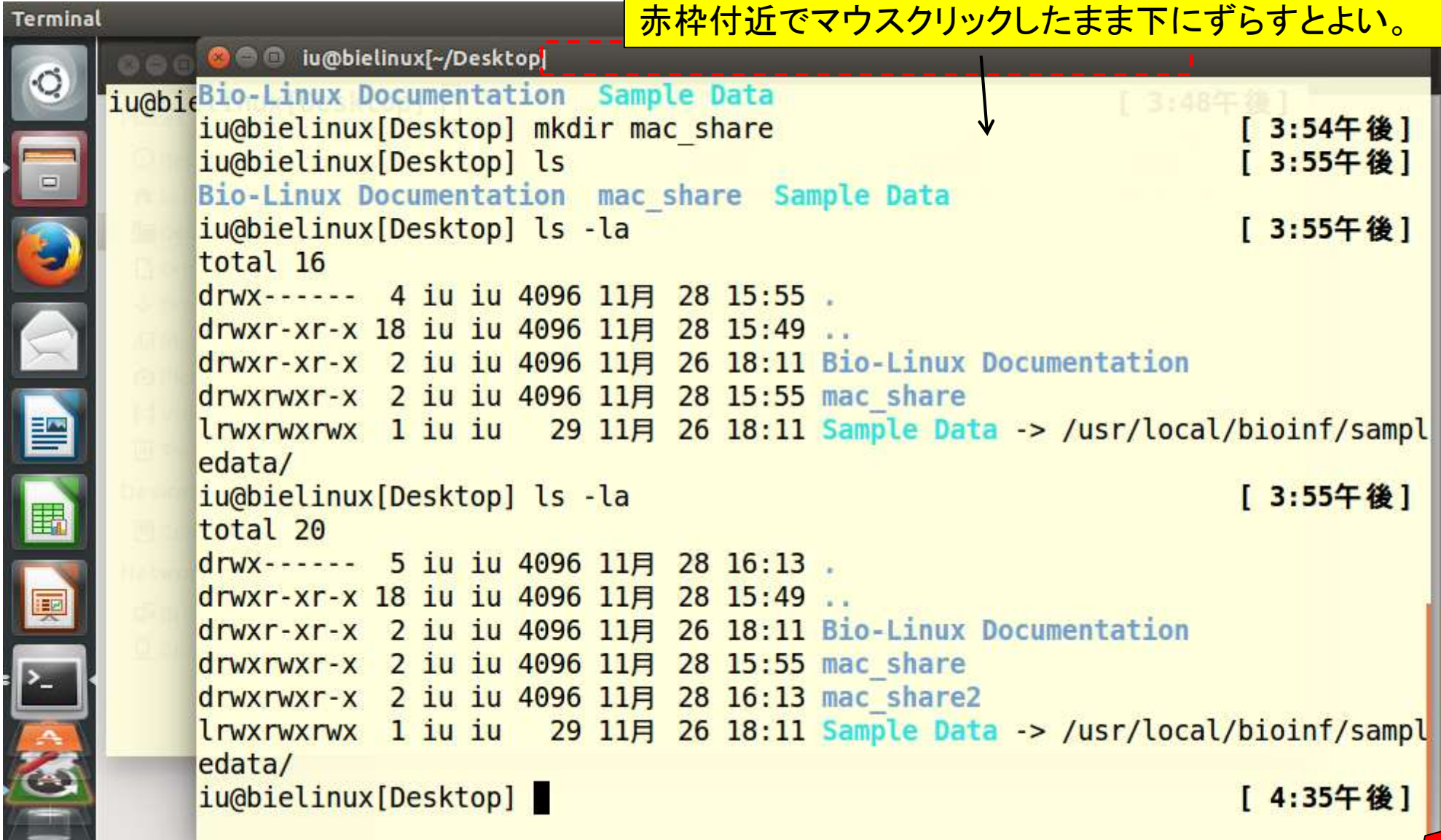
mac_share2フォルダ作成後に再び⑦「ls -la」した結果。コマンドは、改めて打ち込まなくても「上矢印ボタン」を1回押せばよい。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bielinux[iu] cd Desktop ← ② [10:27午前]
iu@bielinux[Desktop] ls ← ③ [10:27午前]
Bio-Linux Documentation Sample Data Untitled Folder
iu@bielinux[Desktop] mkdir mac_share ← ④ [10:27午前]
iu@bielinux[Desktop] ls ← ⑤ [10:27午前]
Bio-Linux Documentation mac_share Sample Data Untitled Folder
iu@bielinux[Desktop] ls -la ← ⑥ [10:27午前]
total 20
drwx----- 5 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] ls -la ← ⑦ [10:27午前]
total 24
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] [10:28午前]
```



W9-2-4

Tips。ターミナルの画面サイズを変えるのは右下が基本だが、右下部分にマウスポインタを合わせてもサイズ変更可能な状態にならないときがある。そのような場合は赤枠付近でマウスクリックしたまま下にずらすとよい。



The screenshot shows a terminal window with the following commands and output:

```
iu@bielinux[~/Desktop]
iu@bielinux[Desktop] mkdir mac_share
iu@bielinux[Desktop] ls
iu@bielinux[Desktop] ls -la
total 16
drwx----- 4 iu iu 4096 11月 28 15:55 .
drwxr-xr-x 18 iu iu 4096 11月 28 15:49 ..
drwxr-xr-x 2 iu iu 4096 11月 26 18:11 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 28 15:55 mac_share
lrwxrwxrwx 1 iu iu 29 11月 26 18:11 Sample Data -> /usr/local/bioinf/sampl
edata/
iu@bielinux[Desktop] ls -la
total 20
drwx----- 5 iu iu 4096 11月 28 16:13 .
drwxr-xr-x 18 iu iu 4096 11月 28 15:49 ..
drwxr-xr-x 2 iu iu 4096 11月 26 18:11 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 28 15:55 mac_share
drwxrwxr-x 2 iu iu 4096 11月 28 16:13 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 26 18:11 Sample Data -> /usr/local/bioinf/sampl
edata/
iu@bielinux[Desktop]
```

A red dashed box highlights the terminal output area. A black arrow points from the tip text to the top of this box. A red arrow at the bottom right points to the terminal's scroll bar area.

W9-2-4

こんな感じになるので、ウィンドウの右下を左クリックボタンを押しながら移動させることでウィンドウサイズを任意に変更できるようになる。

The screenshot shows a Linux desktop environment. In the foreground, a terminal window is open, displaying the output of the `ls -la` command. The output lists several files and directories with their permissions, owner, group, size, and date. The files are: `Bio-Linux Documentation`, `mac_share`, `Sample Data` (a symbolic link to `/usr/local/bioinf/sampledata/`), `.`, `..`, `mac_share2`, and another `Sample Data` (a symbolic link to `/usr/local/bioinf/sampledata/`). The terminal prompt is `iu@bielinux[Desktop]`. In the background, a file manager window is open, showing the same files and directories as the terminal. The file manager window has a title bar with the text `iu@bielinux[~/Desktop]` and a timestamp `[3:48午後]`. The desktop environment includes a sidebar with application icons and a top panel with system status icons and a clock showing `16:46`.

```
iu@bielinux[~/Desktop]
iu@bielinux[Desktop] [ 3:48午後]

drwxr-xr-x  2 iu iu 4096 11月 26 18:11 Bio-Linux Documentation
drwxrwxr-x  2 iu iu 4096 11月 28 15:55 mac_share
lrwxrwxrwx  1 iu iu   29 11月 26 18:11 Sample Data -> /usr/local/bioinf/sampledata/
iu@bielinux[Desktop] ls -la [ 3:55午後]
total 20
drwx-----  5 iu iu 4096 11月 28 16:13 .
drwxr-xr-x 18 iu iu 4096 11月 28 15:49 ..
drwxr-xr-x  2 iu iu 4096 11月 26 18:11 Bio-Linux Documentation
drwxrwxr-x  2 iu iu 4096 11月 28 15:55 mac_share
drwxrwxr-x  2 iu iu 4096 11月 28 16:13 mac_share2
lrwxrwxrwx  1 iu iu   29 11月 26 18:11 Sample Data -> /usr/local/bioinf/sampledata/
iu@bielinux[Desktop] [ 4:45午後]
```

W9-2-5

「sudo mount -t vboxsf share mac_share」という赤矢印部分のコマンドを打ち込む。ホストOSのデスクトップ上にあるshareフォルダをゲストOSのデスクトップ上のmac_shareフォルダとして認識させ、利用可能な状態にする作業を行っている(この操作をマウントという)。すると[sudo] password for iu:と何かを聞かれる状態になる。これはログインユーザiuのパスワードを聞かれているので、この場合pass1409と打ち込む。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bielinux[Desktop] ls
Bio-Linux Documentation Sample Data
iu@bielinux[Desktop] mkdir mac_share
iu@bielinux[Desktop] ls
Bio-Linux Documentation mac_share
iu@bielinux[Desktop] ls -la
total 20
drwx----- 5 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] ls -la ← ⑦
total 24
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share
[sudo] password for iu:
```

[10:27午前]

[10:27午前]

[10:28午前]

W9-2-5

ログインユーザのパスワードを打ち込んだ結果。見ため上、特に何も起こらず、通常のコマンド入力待ち状態になるだけである。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bie Bio-Linux Documentation Sample Data Untitled Folder [ 9:46午前 ]
iu@bielinux[Desktop] mkdir mac_share [10:27午前]
iu@bielinux[Desktop] ls [10:27午前]
Bio-Linux Documentation mac_share Sample Data Untitled Folder
iu@bielinux[Desktop] ls -la [10:27午前]
total 20
drwx----- 5 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] ls -la ← ⑦ [10:27午前]
total 24
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share [10:28午前]
[sudo] password for iu:
iu@bielinux[Desktop] [10:34午前]
```



W9-2-5

⑧もう一度「ls -la」した結果。赤の下線で示すように、「sudo mount -t vboxsf share mac_share」実行前後でmac_shareフォルダの
ところの記述が変わっていることがわかる。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bie | lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampled/
| drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] ls -la ← ⑦ [10:27午前]
total 24
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampled/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share [10:28午前]
[sudo] password for iu:
iu@bielinux[Desktop] ls -la ← ⑧ [10:34午前]
total 20
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxrwx 1 root root 0 11月 28 18:07 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampled/
/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop] [10:35午前]
```

W9-2-5

特に重要なのは赤枠で囲った部分。「drwxrwxrwx」のようになっていればOK。一番左のdはディレクトリの意味。それ以外の9文字のrwxrwxrwxは、mac_shareディレクトリの「所有者、特定のグループ、その他のヒト」を含め全てのヒトに、このディレクトリ中のファイルを読むこと(r)、書き込み(w)、実行(x)の権限を与えるという意味です。もしここがrwxrwxrwxになっていなければ、共有フォルダとしてゲストOS上に作成したユーザiuが、mac_shareディレクトリを介して「ホストOSのデスクトップ上に作成したshareフォルダ」に書き込みが行えないなどの不具合が生じます。Windowsユーザは特にいじっていないければこのデフォルトはrwxrwxrwxとなって問題ないはず。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bielinux[Desktop] ls -la
total 24
drwx----- 6 iu iu 4096 11月
drwxr-xr-x 18 iu iu 4096 11月
drwxr-xr-x 2 iu iu 4096 11月
drwxrwxr-x 2 iu iu 4096 11月
drwxrwxr-x 2 iu iu 4096 11月
lrwxrwxrwx 1 iu iu 29 11月
drwxrwxr-x 2 iu iu 4096 11月
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share
[sudo] password for iu:
iu@bielinux[Desktop] ls -la
total 20
drwx----- 6 iu iu 4096 11月 29 10:27 .
drwxr-xr-x 18 iu iu 4096 11月 29 09:46 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-Linux Documentation
drwxrwxrwx 1 root root 0 11月 28 18:07 mac_share
drwxrwxr-x 2 iu iu 4096 11月 29 10:27 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
drwxrwxr-x 2 iu iu 4096 11月 29 10:21 Untitled Folder
iu@bielinux[Desktop]
```

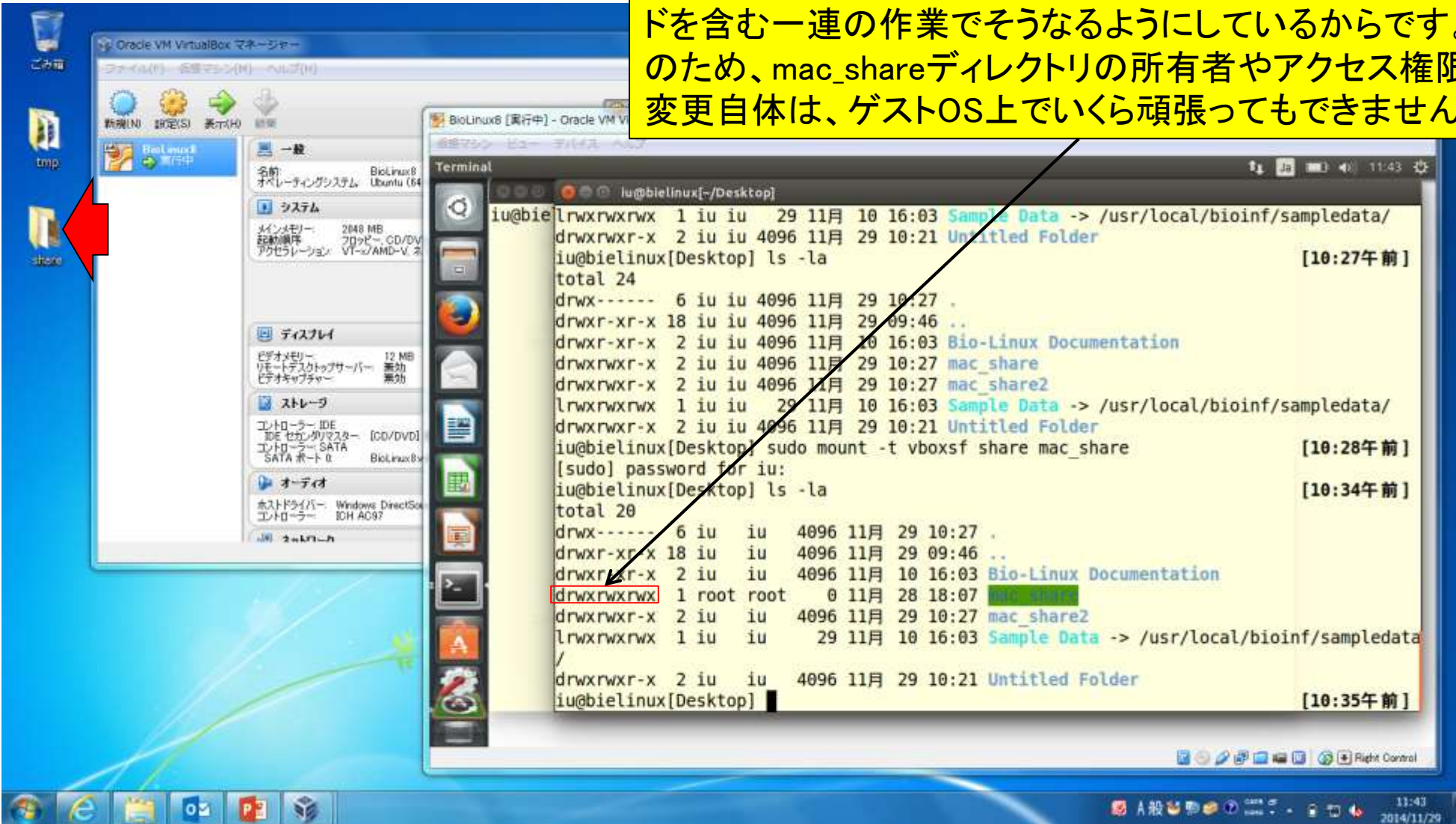
[10:28午前]

[10:34午前]

[10:35午前]

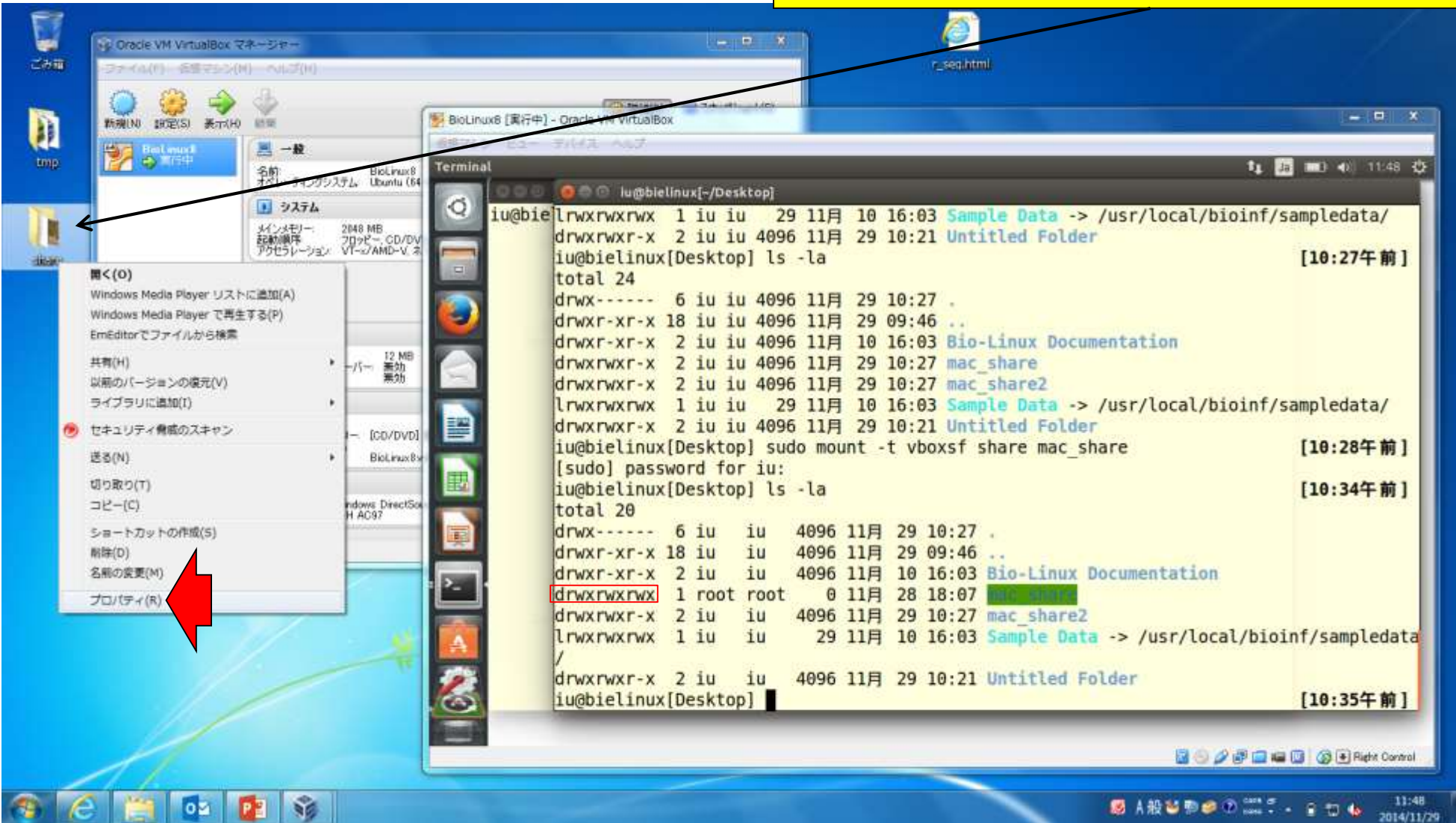
W9-2-5

ゲストOSのデスクトップに作成したmac_shareディレクトリの実体は、ホストOSのデスクトップ上に作成したshareフォルダです。「sudo mount -t vboxsf share mac_share」コマンドを含む一連の作業でそうなるようにしているからです。そのため、mac_shareディレクトリの所有者やアクセス権限の変更自体は、ゲストOS上でいくら頑張ってもできません。



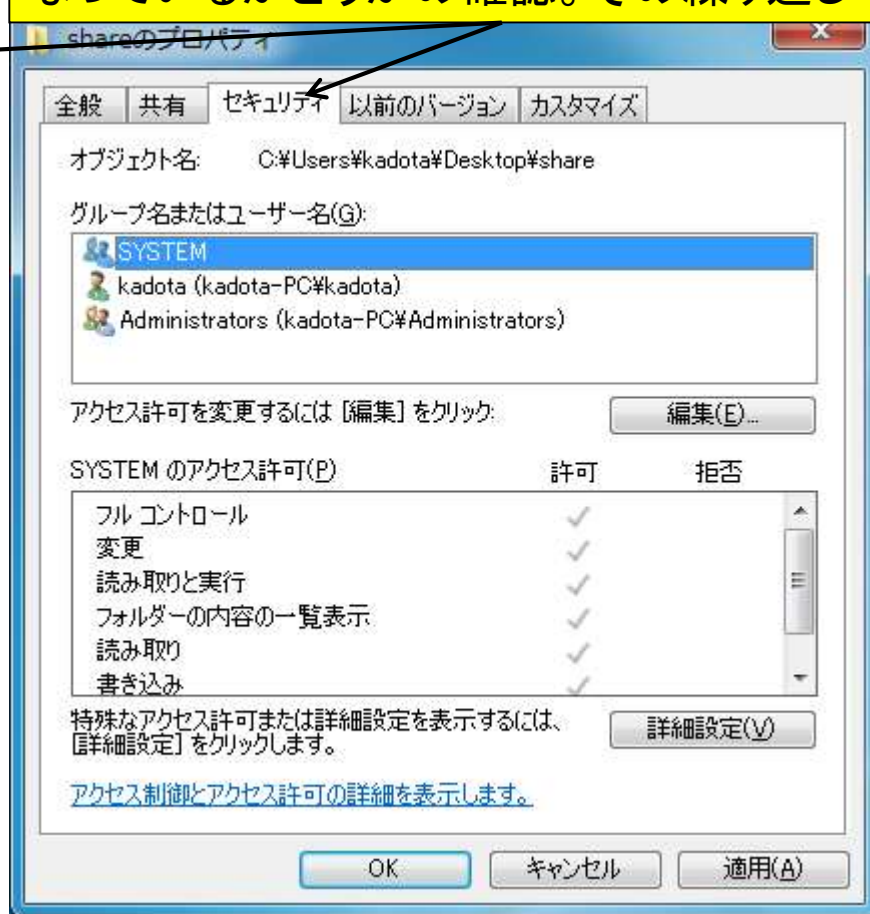
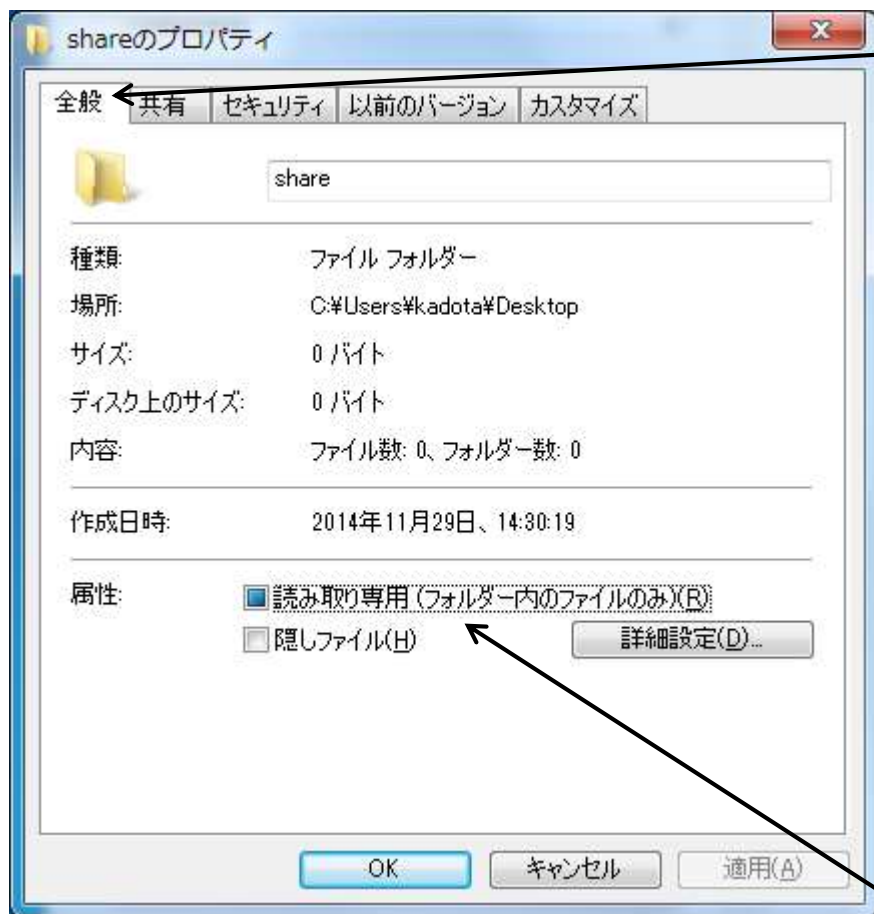
W9-2-6

赤枠部分が「drwxrwxrwx」になっていない場合は、
ホストOSのデスクトップ上にあるshareフォルダ
のアイコン上で右クリックして「プロパティ」を選択。



W9-2-6

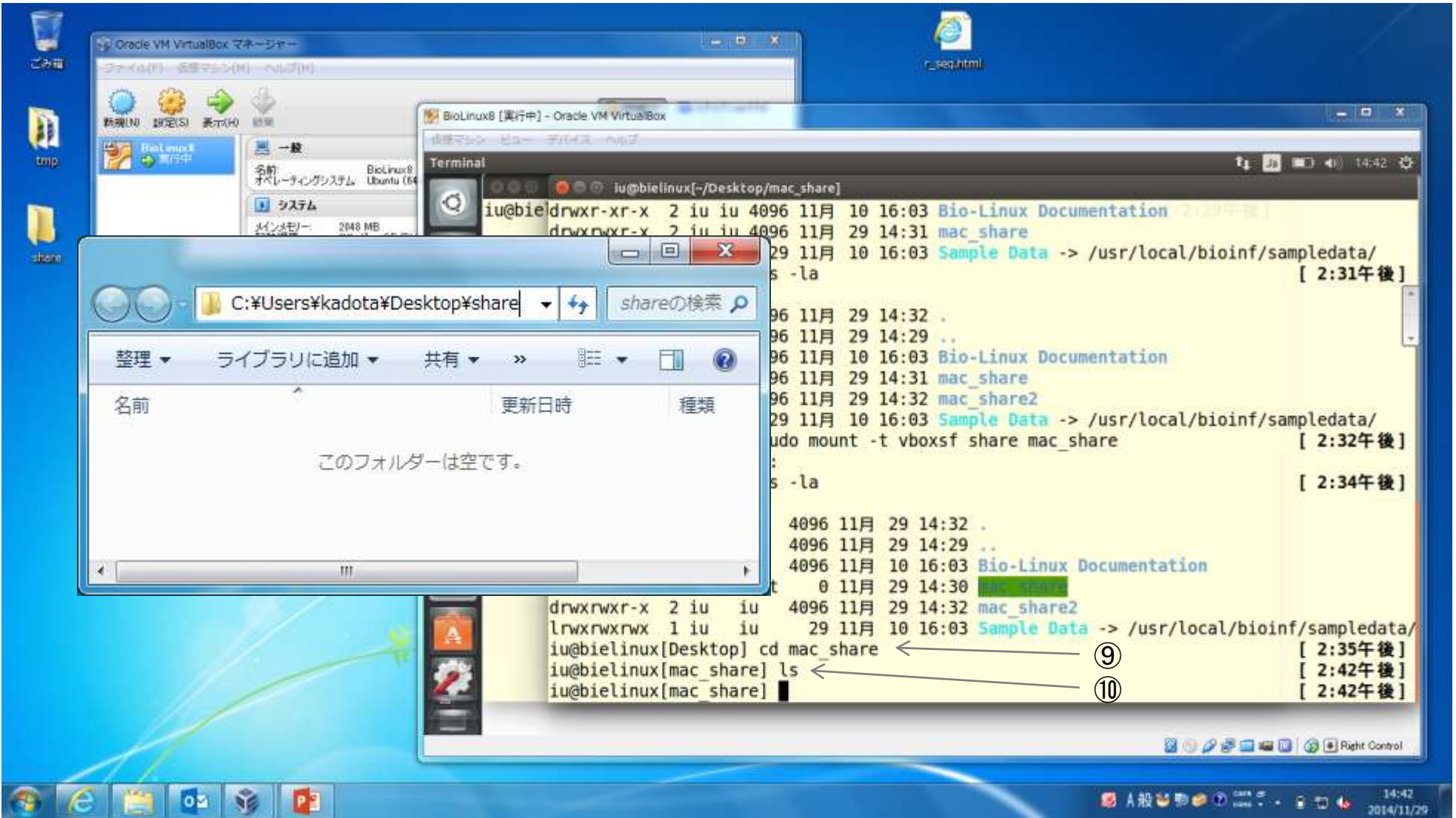
「全般」や「セキュリティ」タブなどの項目中に「drwxrwxrwx」になっていない原因があるはずですので、どうにかして変更しましょう。その後、またターミナル上で「ls -la」で「drwxrwxrwx」となっているかどうかの確認。その繰り返しです。



門田の場合、「drwxrwxrwx」となっている環境でも「読み取り専用」となっていました。うまくいっている状況でわざわざこのチェックを外してハマったことがあります。繰り返しますが、うまくいっているときにはいろいろいじらないほうがいいです。

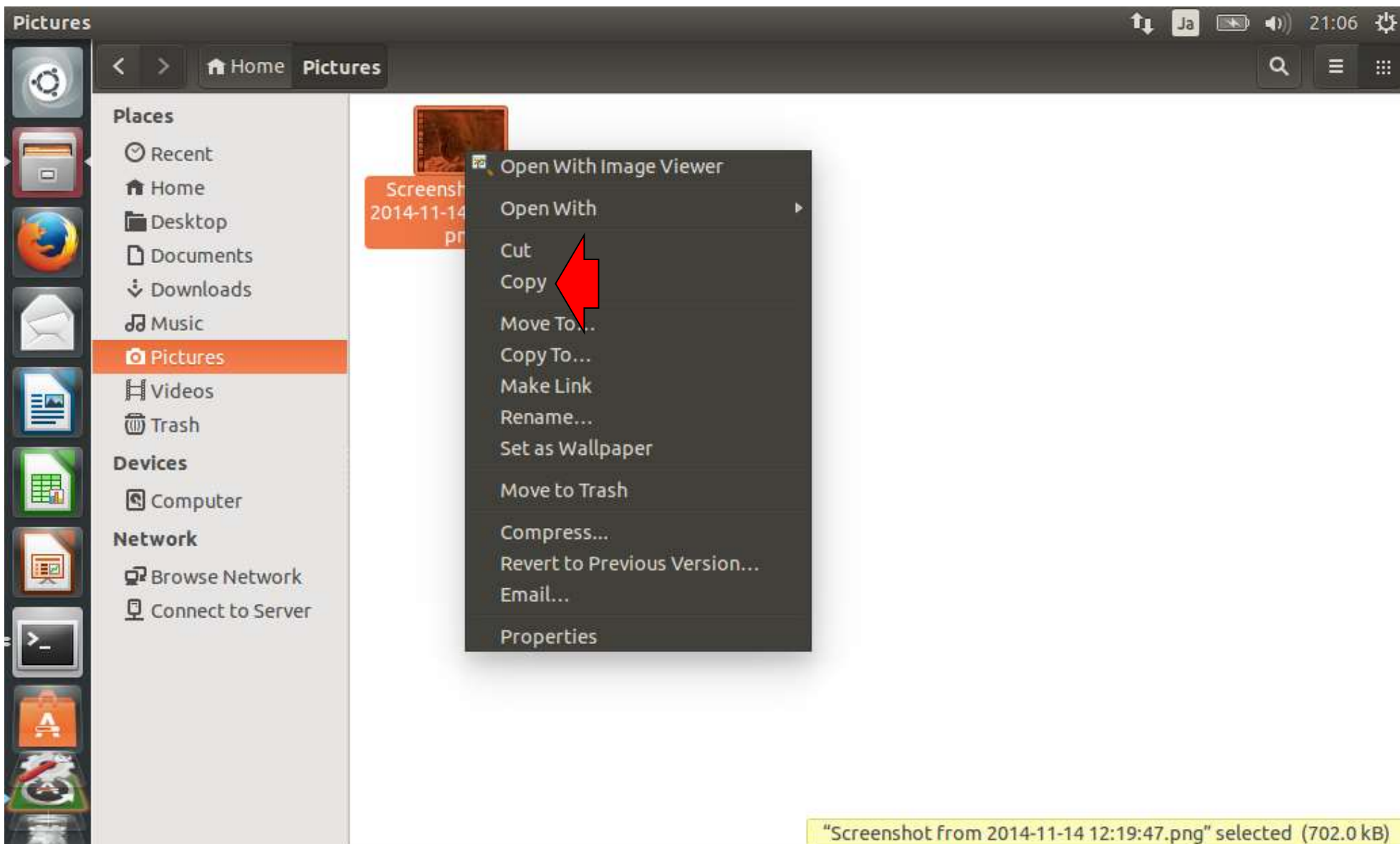
W9-2-7

共有フォルダの状況確認。⑨mac_shareディレクトリに移動し、⑩lsして何もないことを確認。



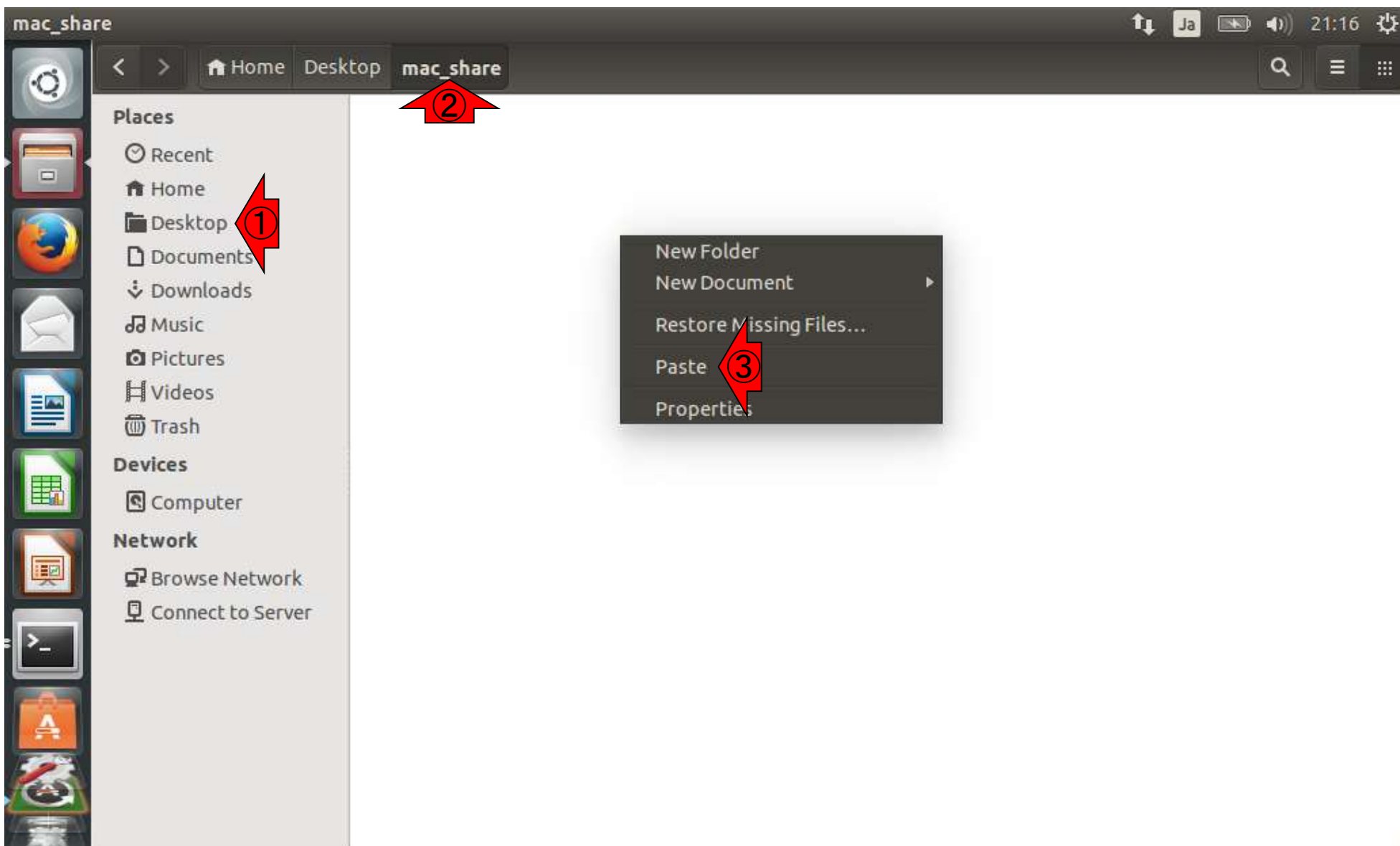
W9-2-8

以前にスクリーンショットを保存していた Pictures フォルダ中の png ファイルをコピー。

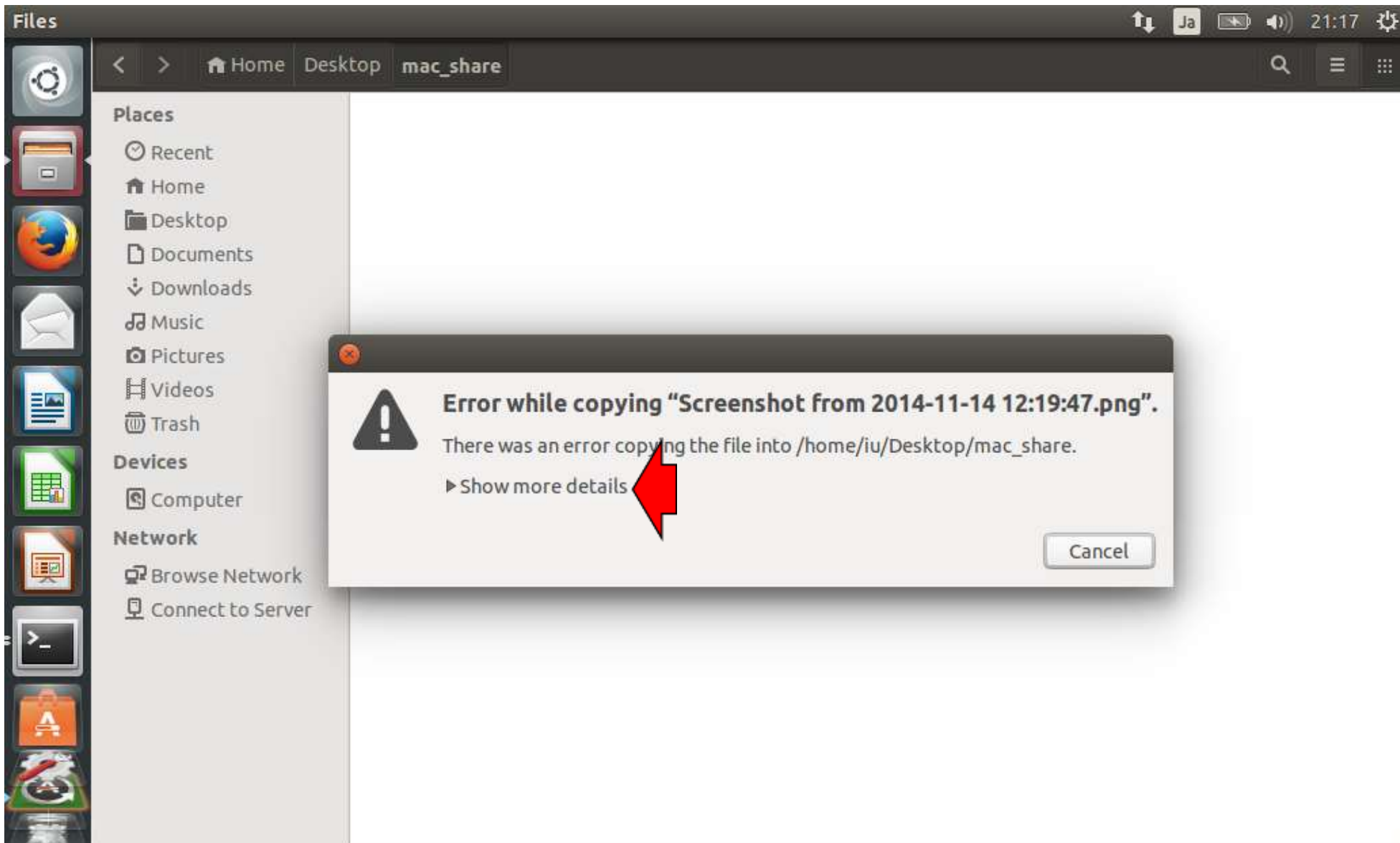


W9-2-8

ゲストOSのデスクトップ上にある
mac_shareフォルダ上でペースト。

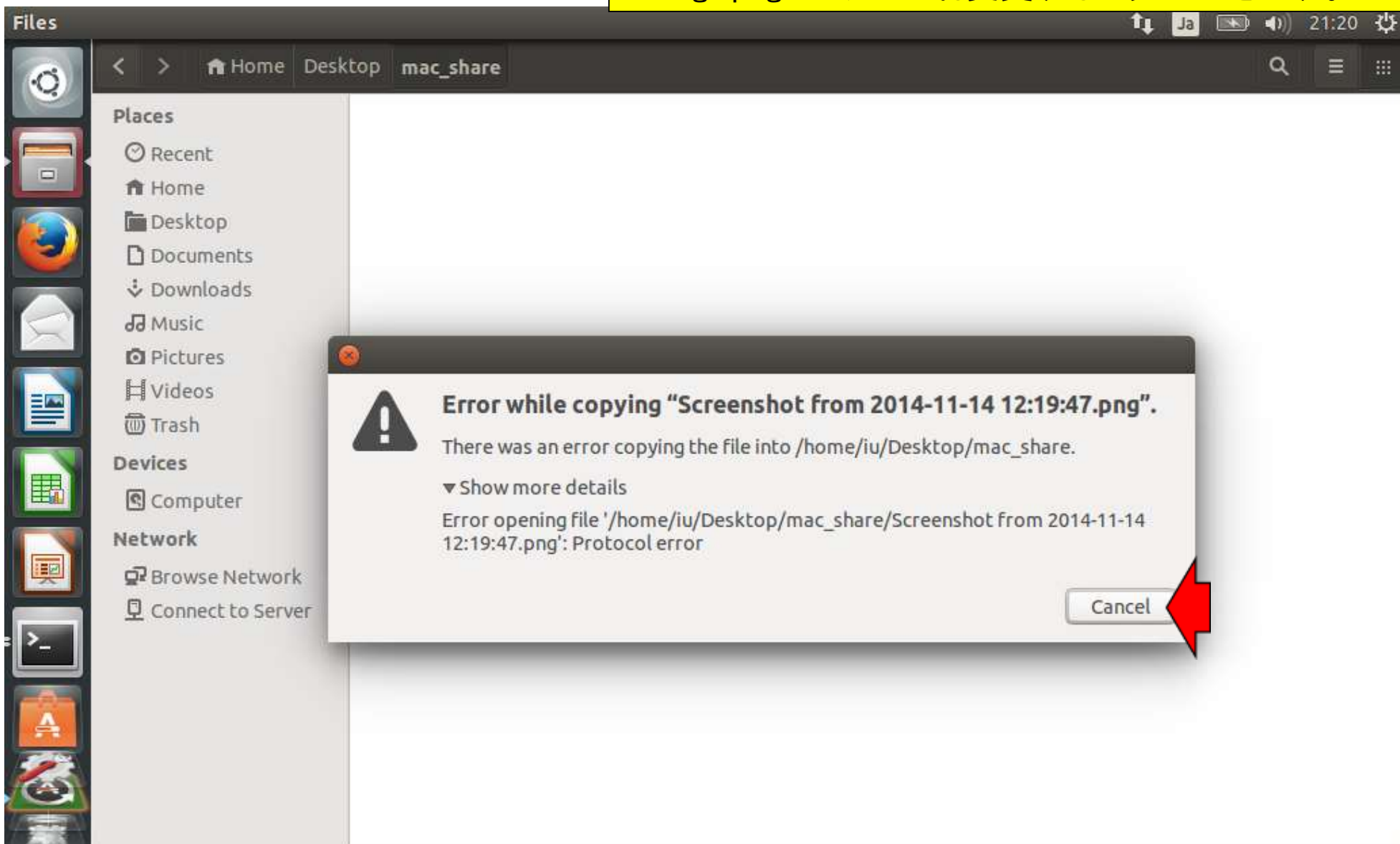


W9-2-8



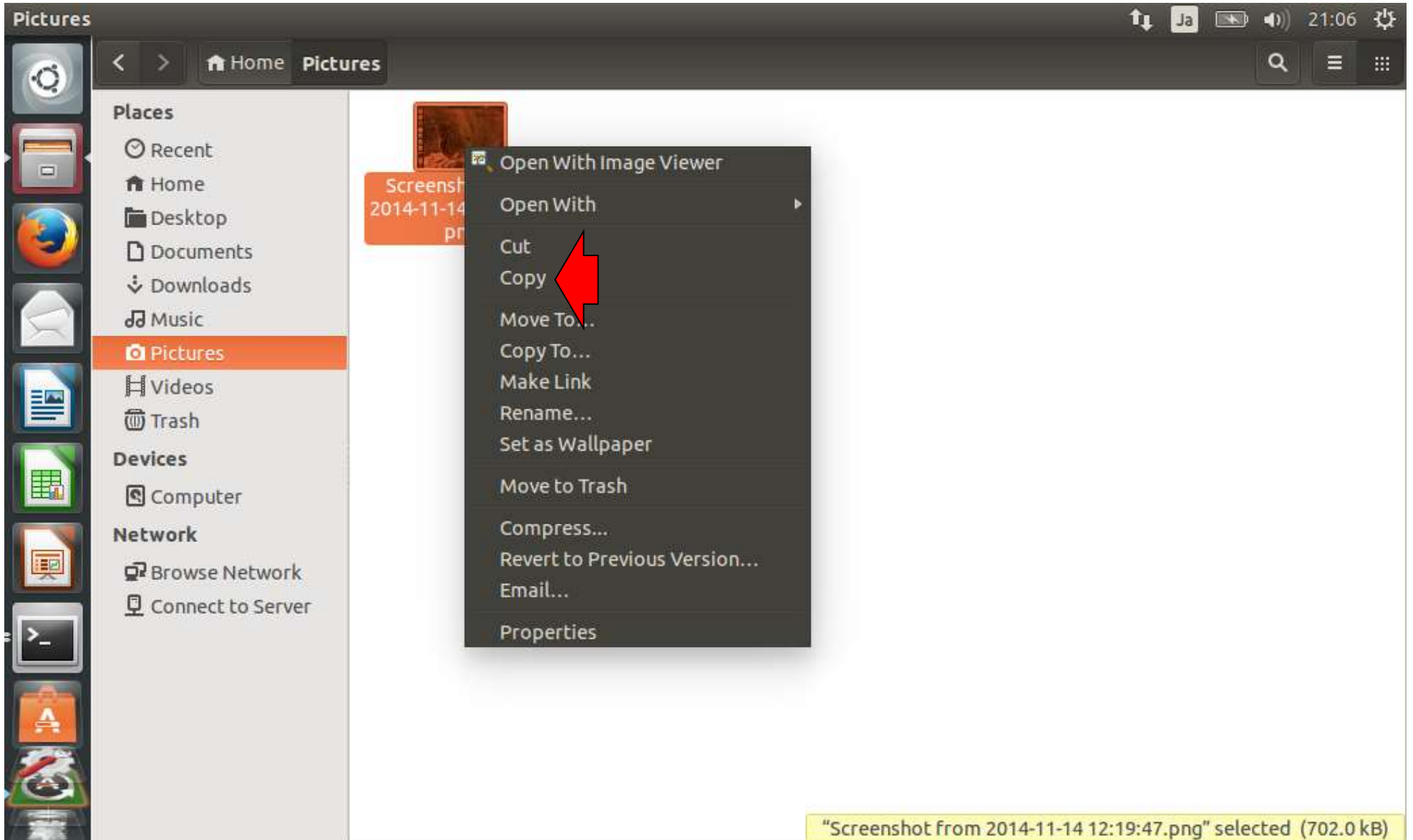
W9-2-8

結論としては、「Screenshot from 2014-11-14 12:19:47.png」というファイル名に問題があるようです。例えばhoge.pngにファイル名変更するとうまくいきます。

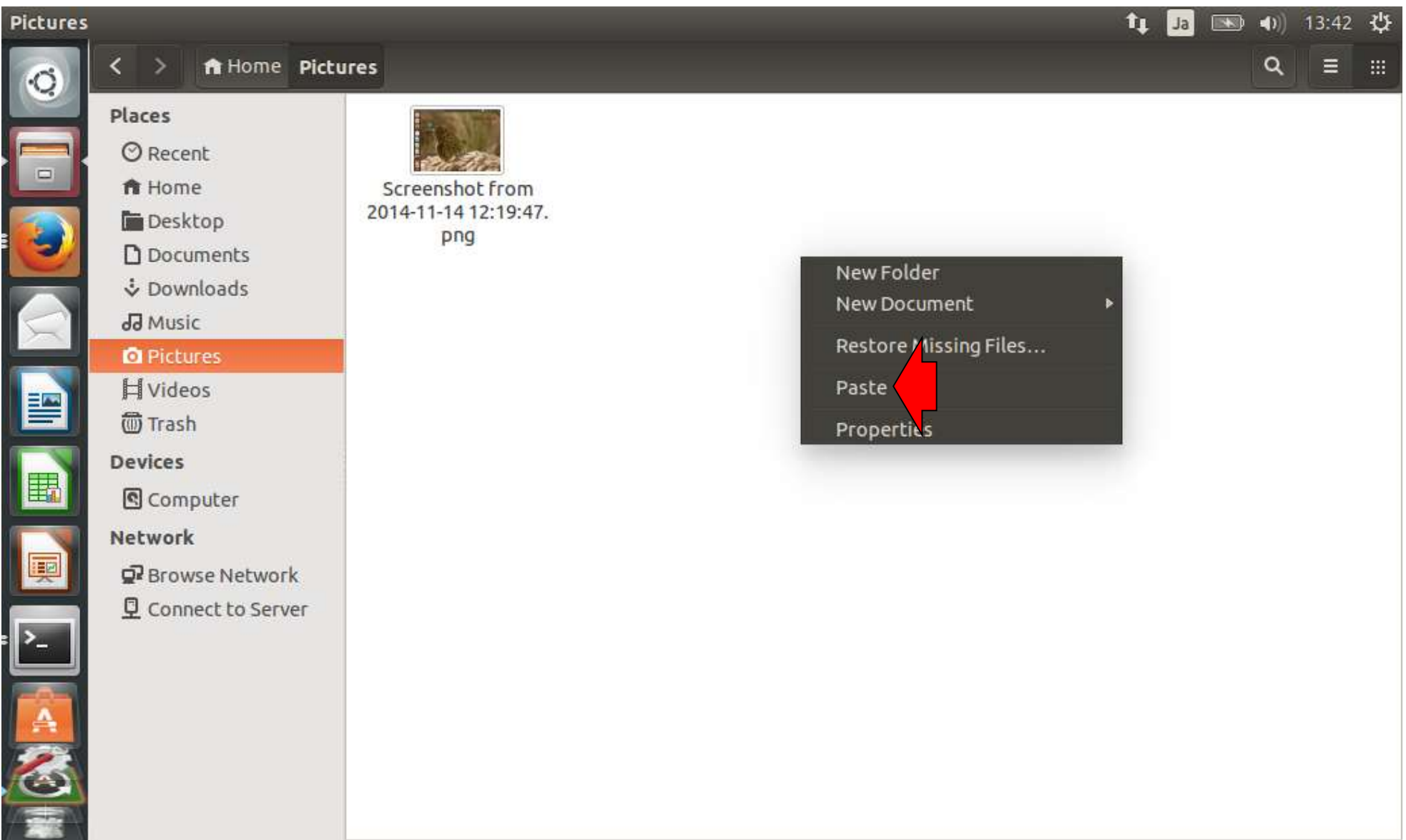


W9-3

「Screenshot from 2014-11-14 12:19:47.png」
のアイコン上で右クリックしてCopy。

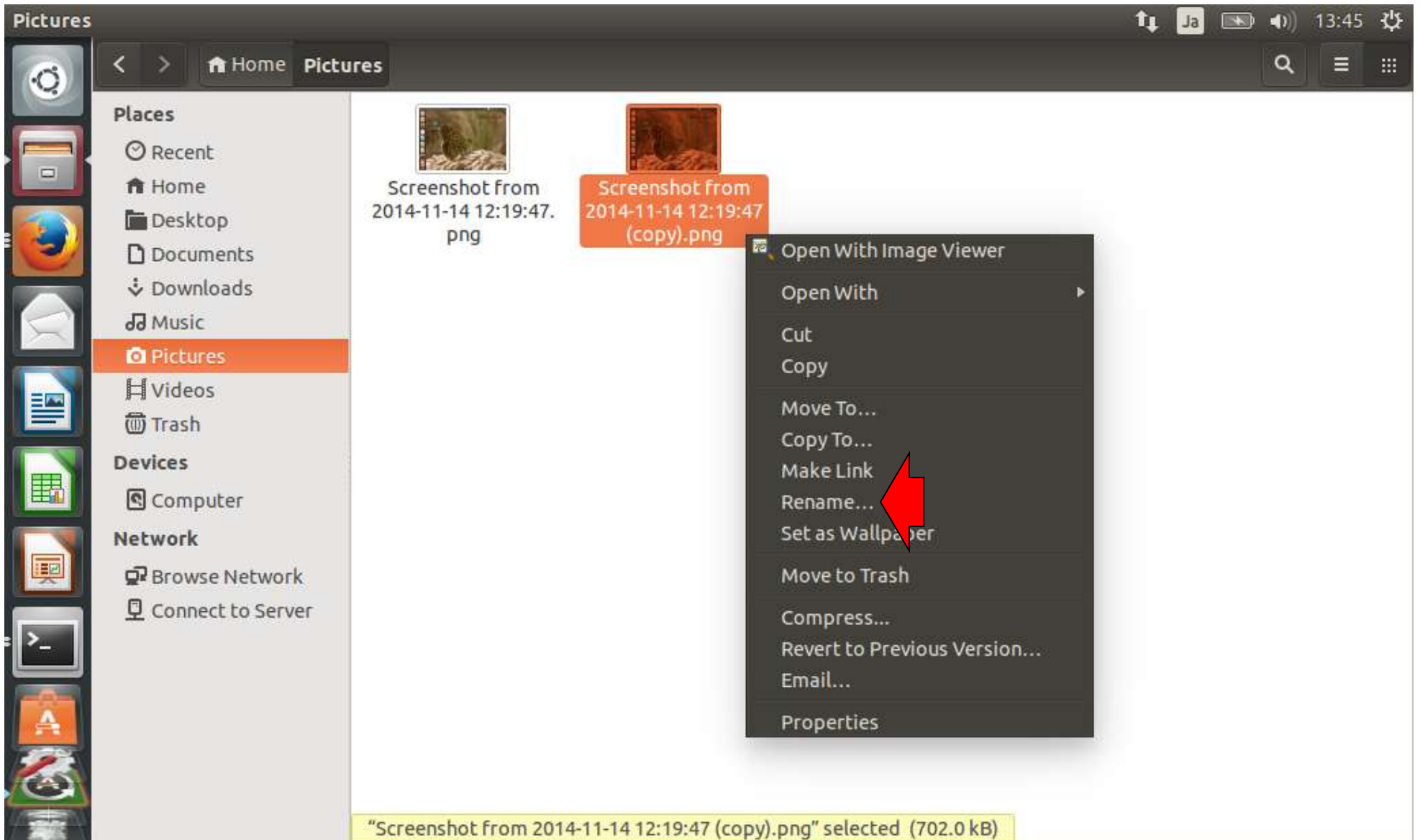


W9-3

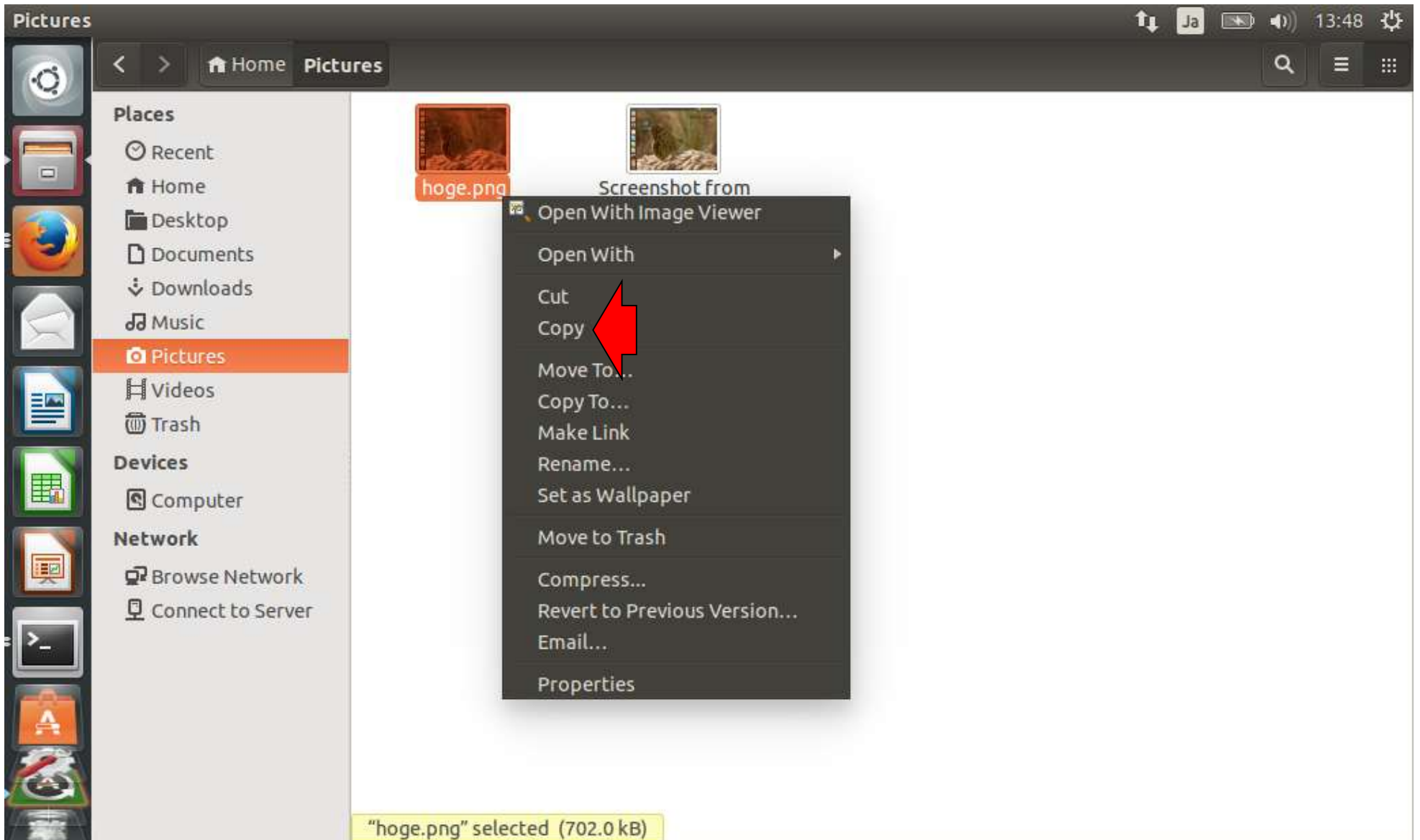


W9-3

作成されたファイルのアイコン上で、再び右クリックしてRenameし、hoge.pngとする。

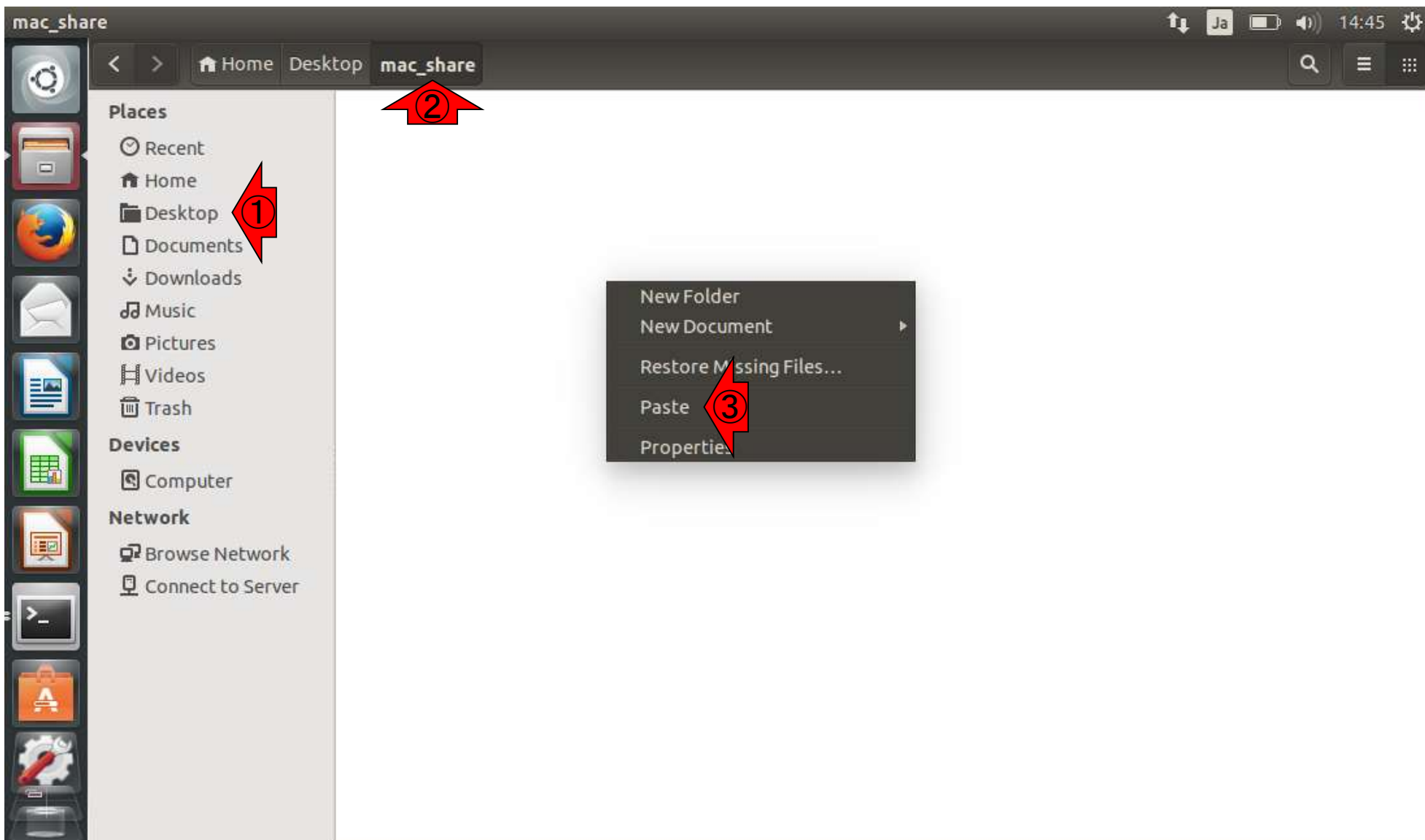


W9-3



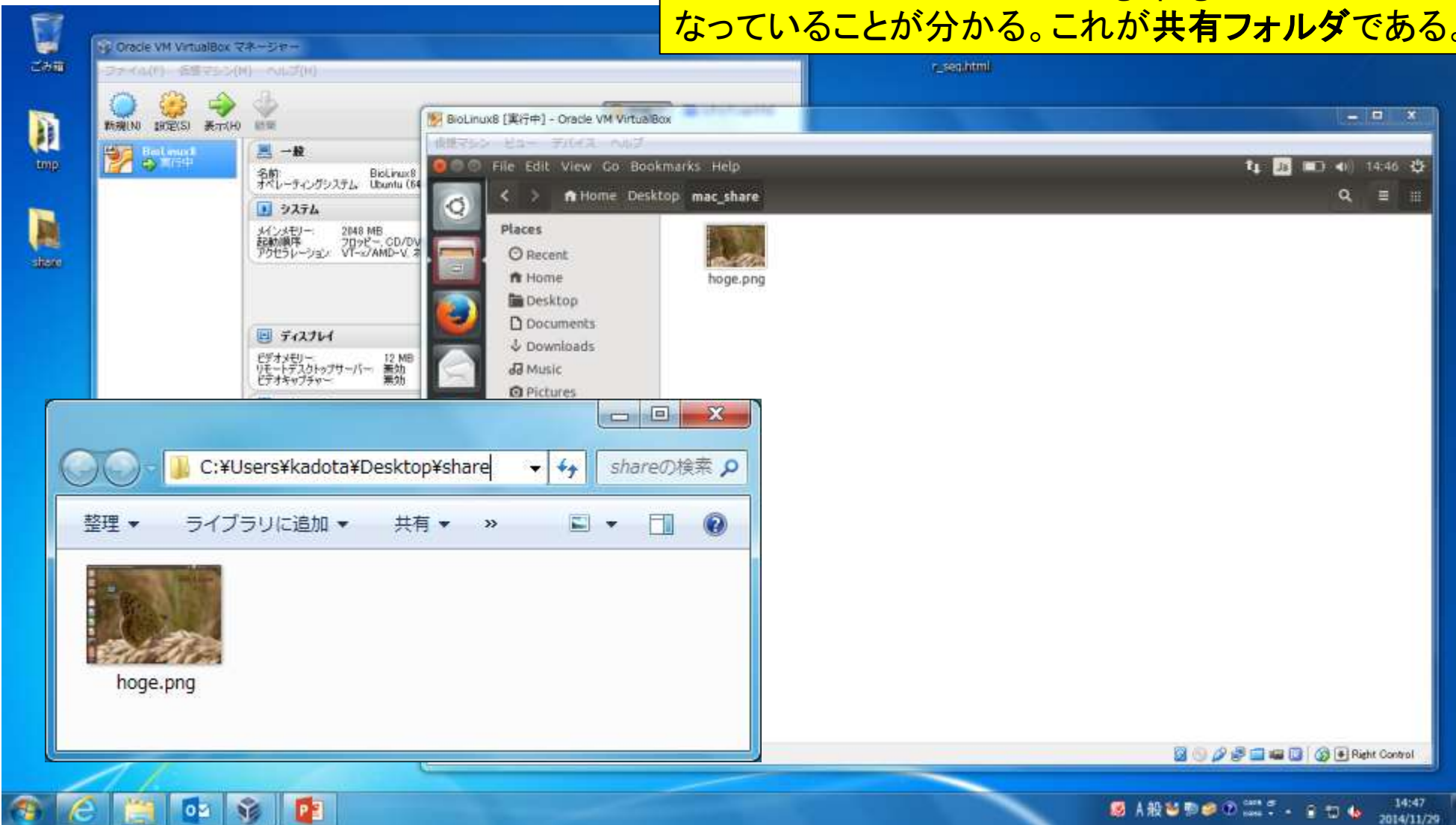
W9-3

ゲストOSのデスクトップ上にあるmac_shareフォルダ上でペースト。



W9-3

ゲストOSのデスクトップ上のmac_shareフォルダにhoge.pngが確かに存在する。また、ホストOSのデスクトップ上のshareフォルダからもhoge.pngが見られるようになっていいることが分かる。これが共有フォルダである。



W9-4

マウントの説明。USBメモリなどを差し込んだときに自動再生というウィンドウが出る。Windowsの場合は、このようにデフォルトで自動的にマウント(オートマウント、とも呼ばれる)してくれる。



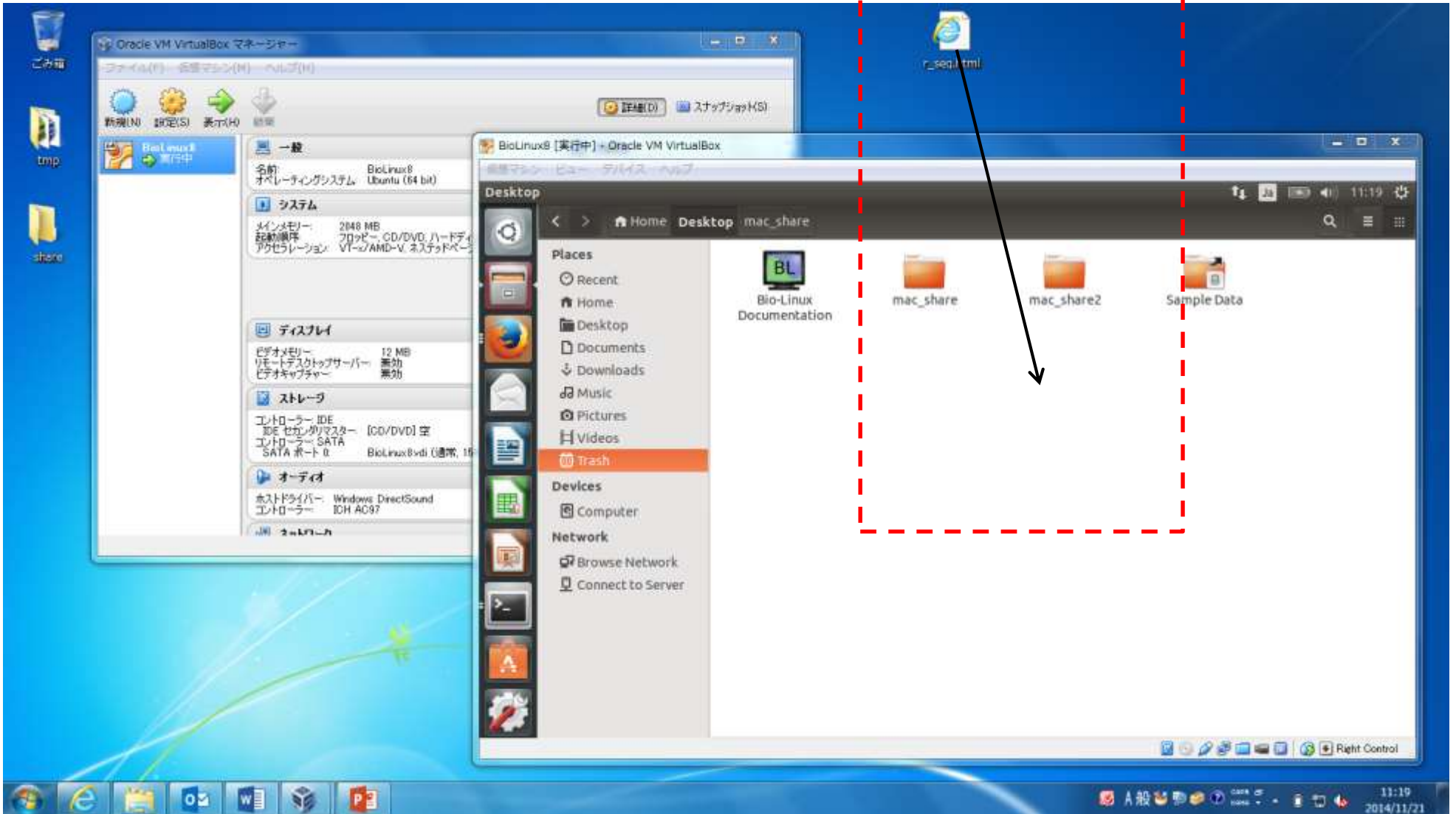
W9-4

USBメモリとPC間で何かデータのやり取りをしている最中に、手でいきなり引き抜くという物理的な取り外しは危険である。そのため、USBメモリを取り外したい場合には、一般にいきなり手で引き抜く(物理的な取り外し)前に、画面右下にあるタスクトレイから目的の「USBメモリの取り出し」を行う。これはPC側からの認識を解除(アンマウント、という)していることに相当し、物理的な取り外しの対比的な用語として、論理的な取り外しという表現がなされる。



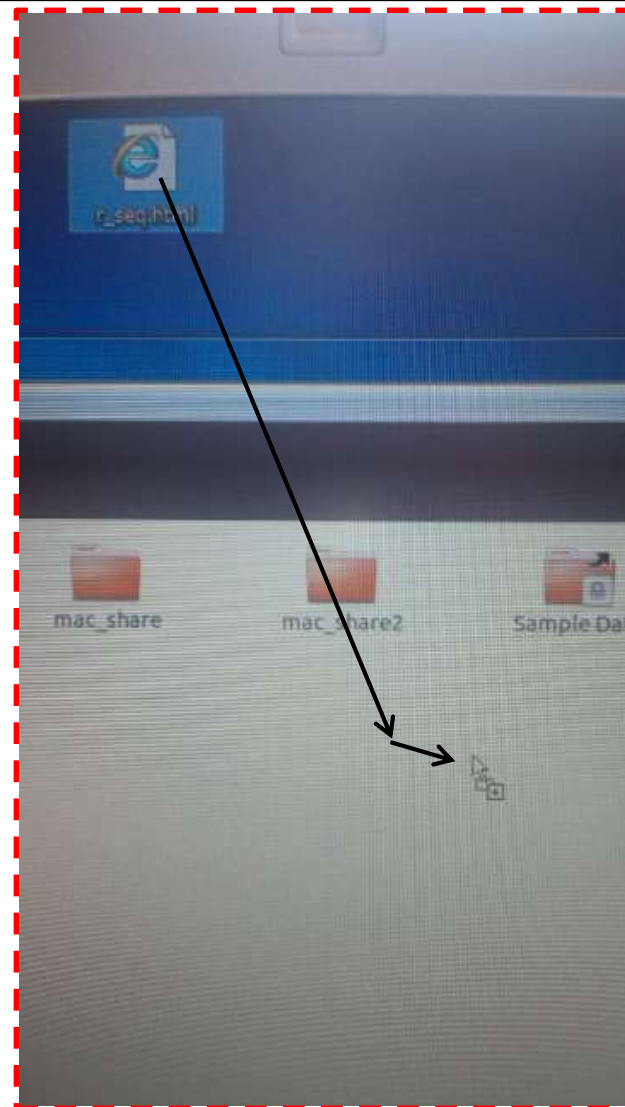
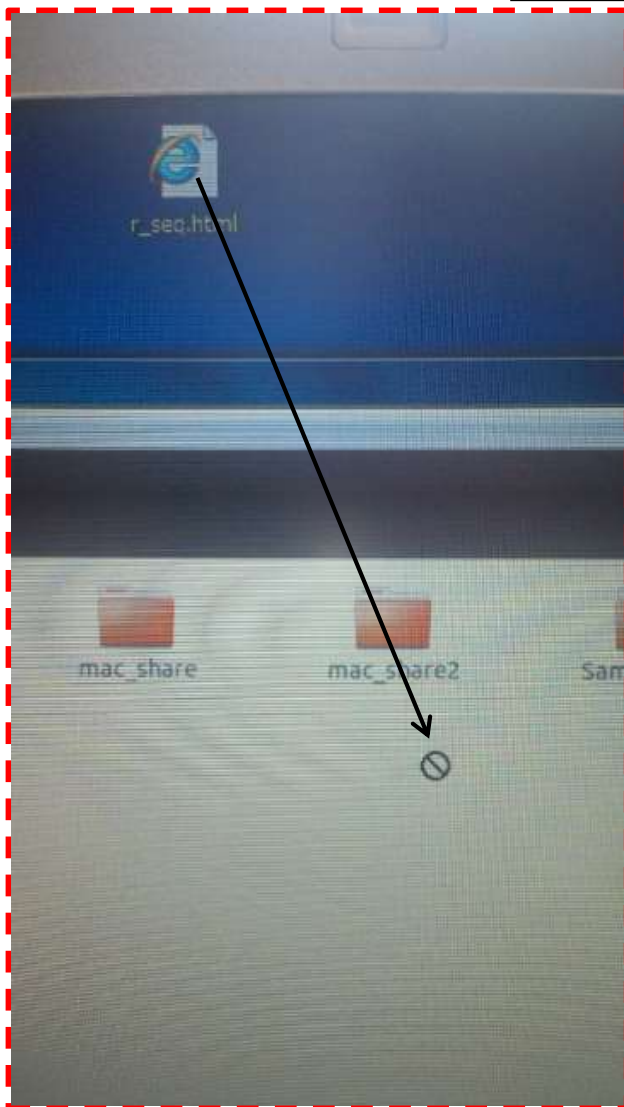
W9-5

ホストOSからゲストOSへドラッグ&ドロップできるときとできないときがあるようだ。



W9-5

ゲストOSのウィンドウ上で、左図のようにマウスカーソルが駐車禁止マークのようになっていると、ドロップ(ファイルを置く)することができないが、ちょっとカーソルの位置をずらしたりすれば右図のように置くことができるマークに変わる。



W9-6

ゲストOSの終了時に「仮想マシンの状態を保存」にせず、「仮想マシンの電源オフ」を選択して終了してみる。

BioLinux8 [一時停止] - Oracle VM VirtualBox

仮想マシン ビュー デバイス ヘルプ

Terminal

```
iu@bielinux[~/Desktop/mac_share]
iu@bielinux[Desktop] ls -la
total 16
drwx----- 4 iu iu 4096 11月 19 16:26 .
drwxr-xr-x 18 iu iu 4096 11月 18 14:28 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 19 16:26 mac_share
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
iu@bielinux[Desktop] ls
total 20
drwx----- 5 iu iu 4096 11月 19 19:41 mac_share
drwxr-xr-x 18 iu iu 4096 11月 18 14:28 ..
drwxr-xr-x 2 iu iu 4096 11月 10 16:03 Bio-linux Documentation
drwxrwxr-x 2 iu iu 4096 11月 19 16:26 mac_share
drwxrwxr-x 2 iu iu 4096 11月 19 19:41 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledata/
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share
[sudo] password for iu:
iu@bielinux[Desktop] cd mac_share
iu@bielinux[mac_share] ls
iu@bielinux[mac_share] ls
r_seq.html
iu@bielinux[mac_share]
```

仮想マシンを閉じる

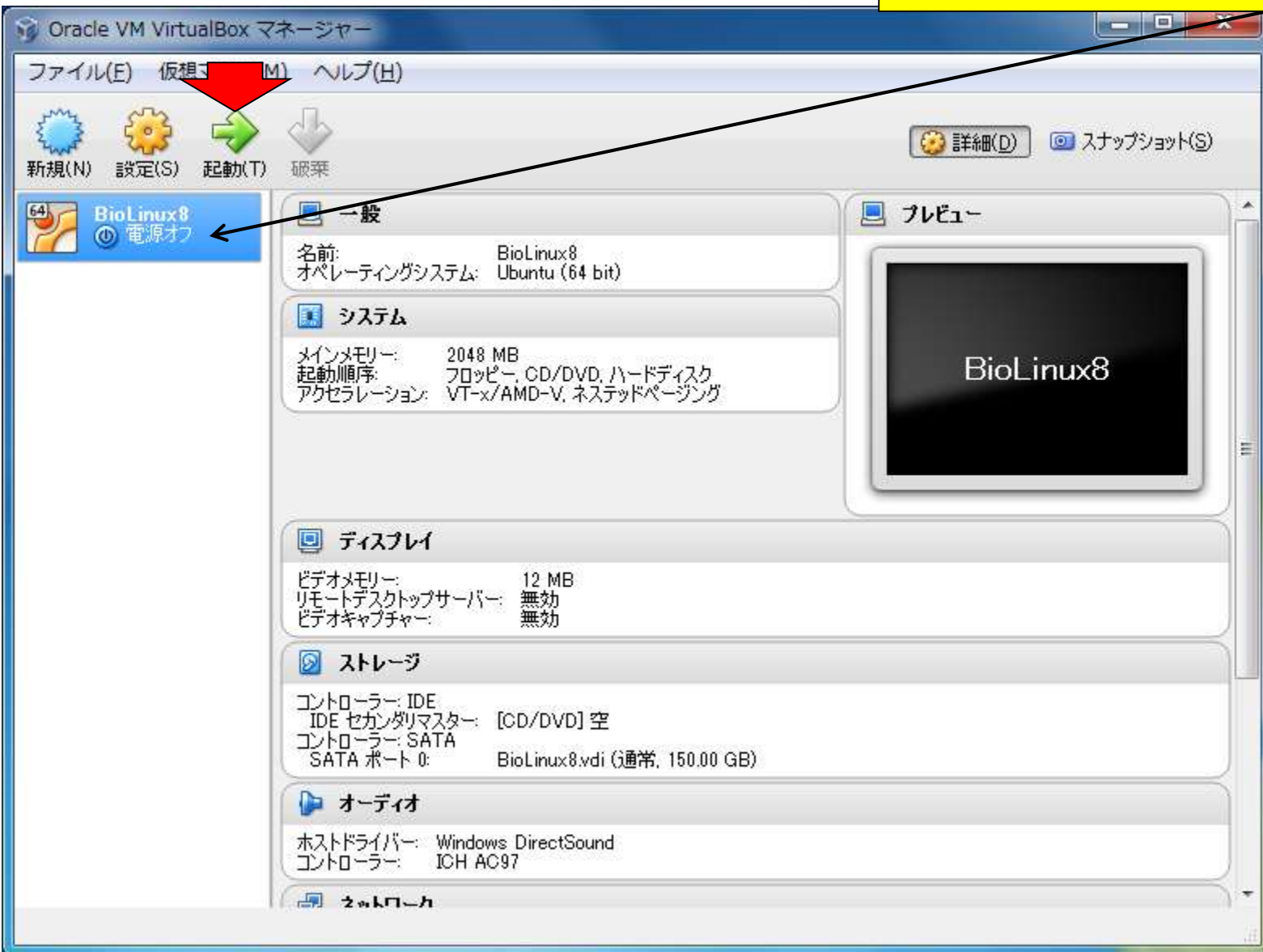
操作を選択:

- 仮想マシンの状態を保存(S)
- シャットダウン シグナル送信(E)
- 仮想マシンの電源オフ(P)

OK キャンセル ヘルプ(H)

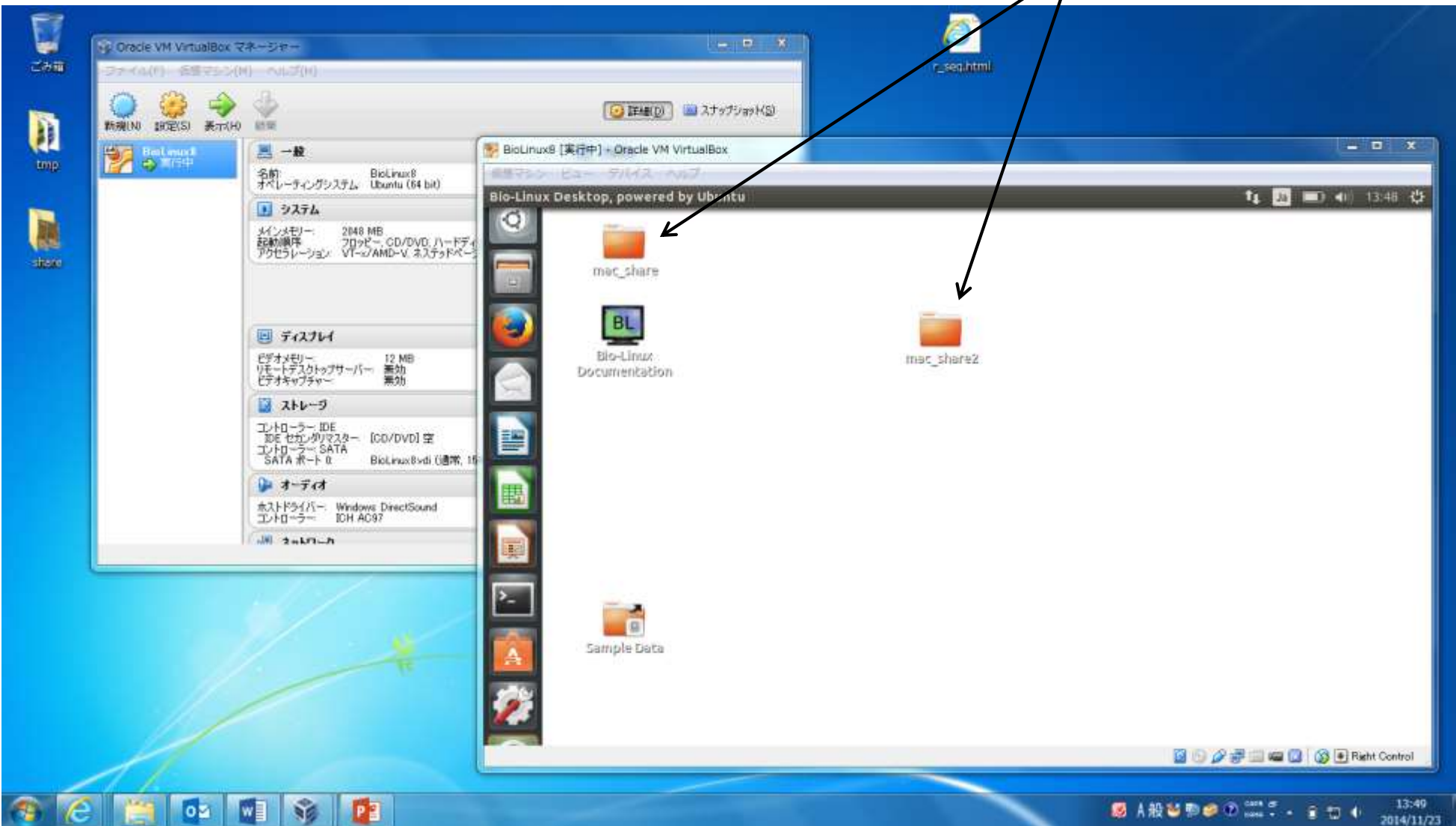
W9-6

ゲストOS終了直後のVirtualBoxマネージャー。
WW8-9と違って、「保存」とはなっておらず、
「電源オフ」になっていることがわかる。



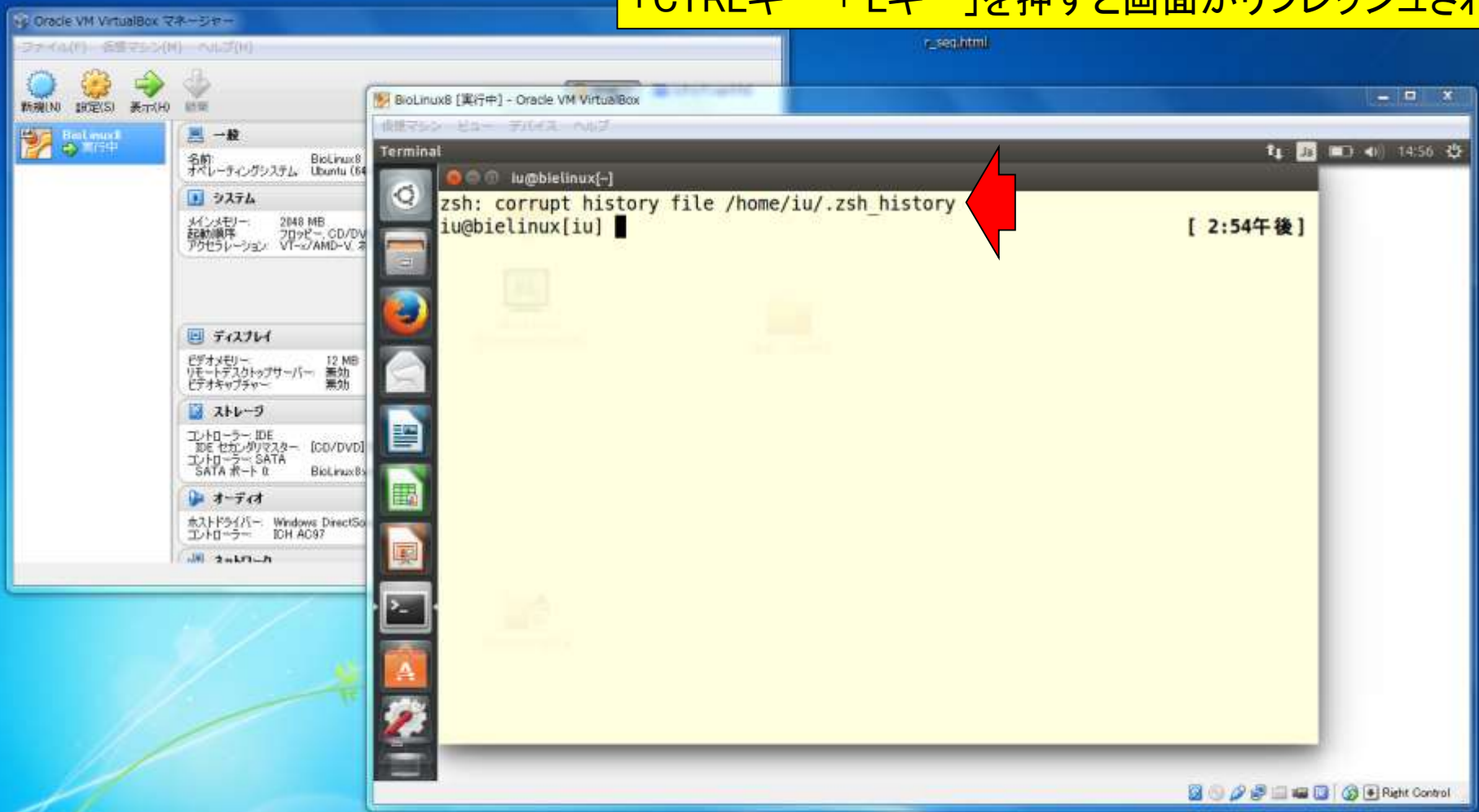
W9-6

BioLinux8起動直後の状態。前回終了前に作成した mac_share や mac_share2 フォルダは残っていることがわかる。



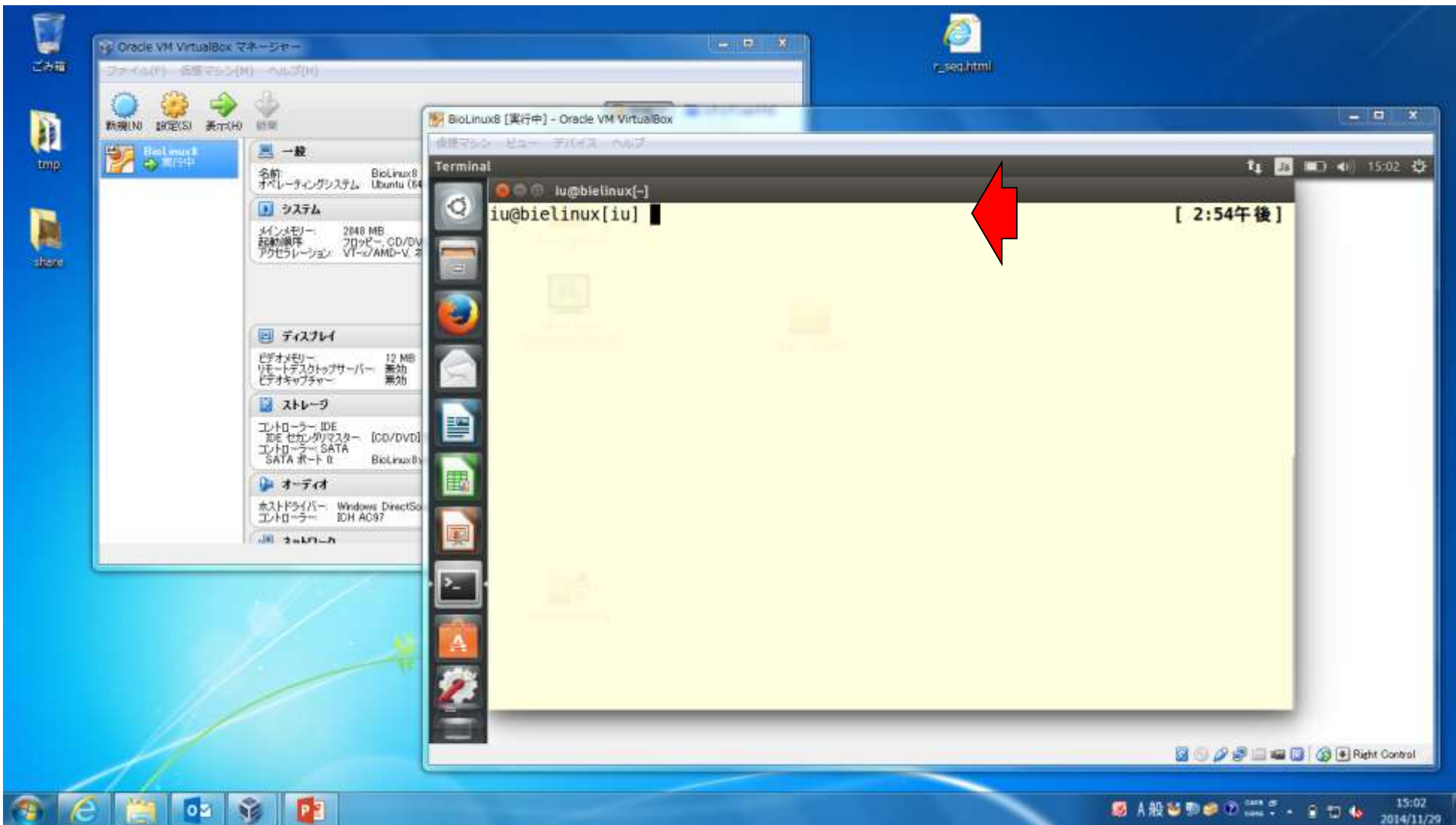
W9-6

ターミナル起動後にこのようなメッセージが出ることもあるが、本質的な問題ではないので初心者のうちは気にしなくてよい。対処法は、連載第4回のW10-2に示している。「CTRLキー + Lキー」を押すと画面がリフレッシュされる。



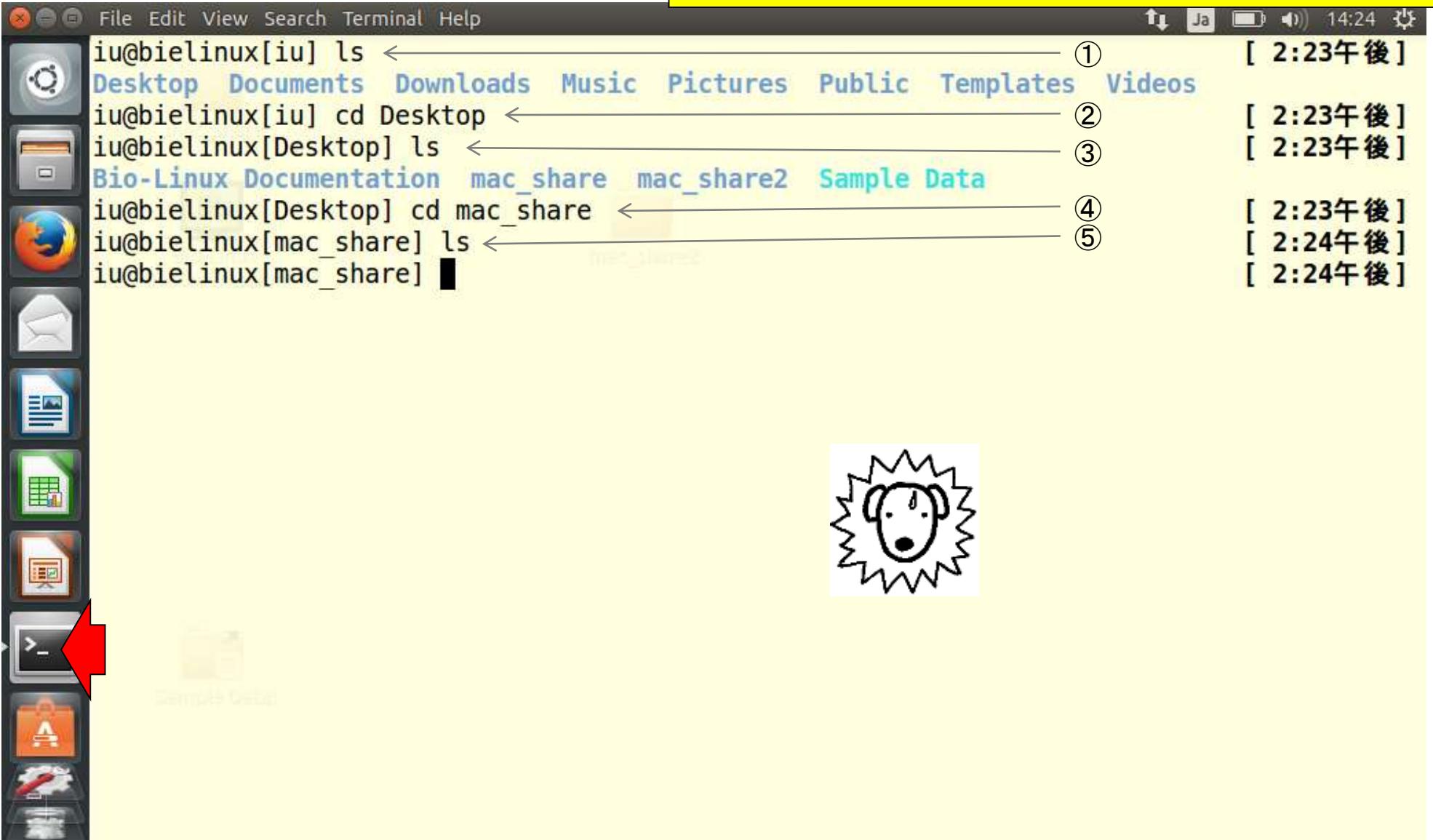
「CTRLキー + Lキー」を押して画面がリフレッシュされた状態。

W9-6



W9-6

ターミナルを起動して前回共有フォルダとして設定した「Desktop」 - 「mac_share」に移動し、lsした結果。共有フォルダ設定が無効になっていることがわかる。



```
iu@bielinux[iu] ls ← ① [ 2:23午後]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop ← ② [ 2:23午後]
iu@bielinux[Desktop] ls ← ③ [ 2:23午後]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share ← ④ [ 2:23午後]
iu@bielinux[mac_share] ls ← ⑤ [ 2:24午後]
iu@bielinux[mac_share] █ [ 2:24午後]
```

The terminal window shows the following sequence of commands and outputs:

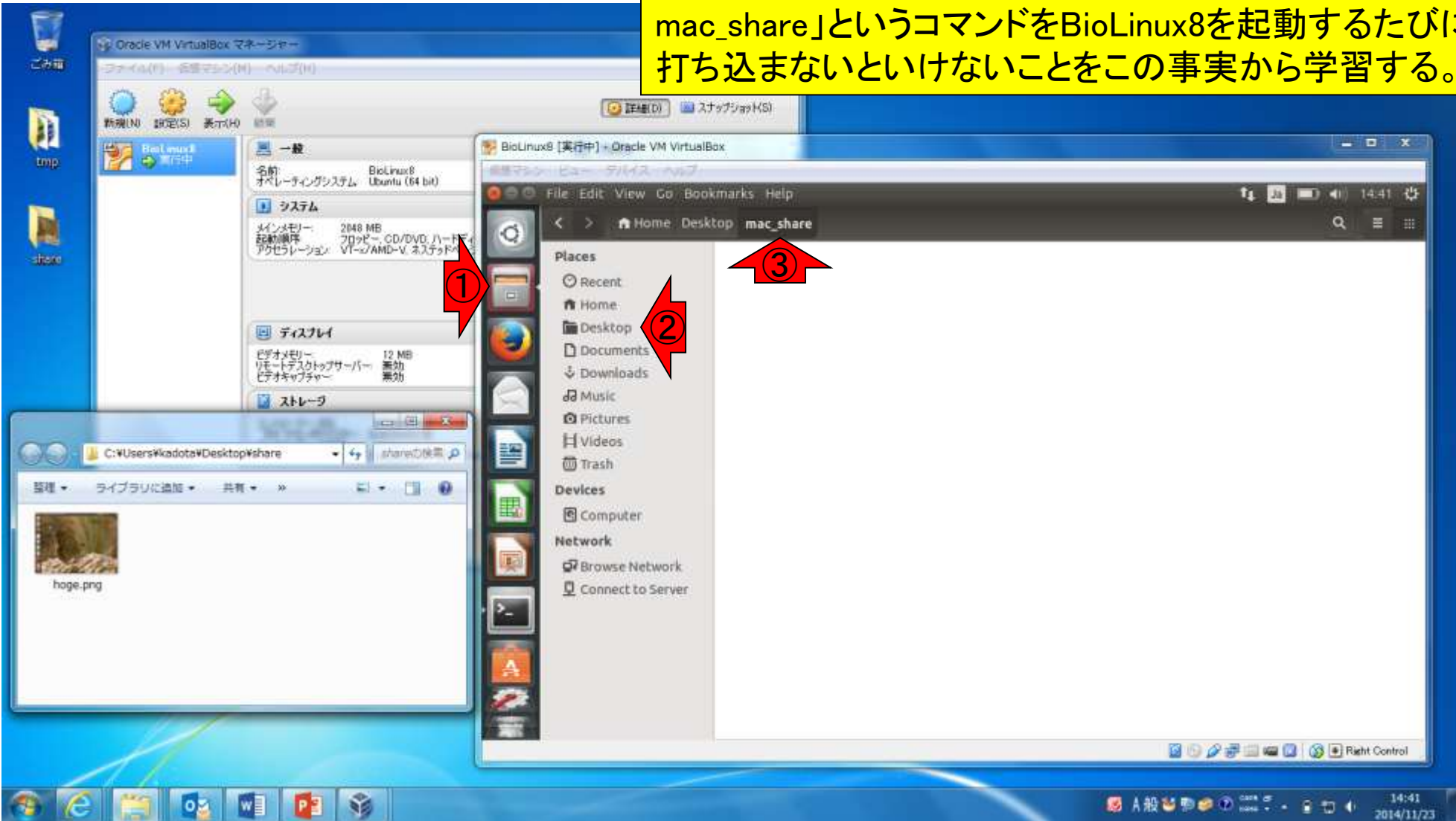
- ① `iu@bielinux[iu] ls` → Output: `Desktop Documents Downloads Music Pictures Public Templates Videos`
- ② `iu@bielinux[iu] cd Desktop`
- ③ `iu@bielinux[Desktop] ls` → Output: `Bio-Linux Documentation mac_share mac_share2 Sample Data`
- ④ `iu@bielinux[Desktop] cd mac_share`
- ⑤ `iu@bielinux[mac_share] ls`

A red arrow points to the terminal icon in the sidebar.



W9-6

ホストOS上の「Desktop」 - 「share」にあるpngファイルが、ゲストOSのGUI画面でも確かにない。共有フォルダを利用したい場合は、W9-2-5の「sudo mount -t vboxsf share mac_share」というコマンドをBioLinux8を起動するたびに打ち込まないといけないことをこの事実から学習する。



W9-6

②までやった段階で、黒矢印部分に表示されていた「共有フォルダ名とそのパス」が見られなくなっている場合は、WW9-2-2の最初のほうと同じようにVirtualBoxマネージャーのほうで、ホストOSのデスクトップ上にあるshareフォルダを共有フォルダとするところからやりなおし。

Oracle VM VirtualBox マネージャー

ファイル(E) 仮想マシン(M) ヘルプ(H)

新規(N) 設定(S) 起動(T) 破棄

BioLinux8 電源オフ

一般 プレビュー

名前: BioLinux8

BioLinux8 - 設定

共有フォルダ

フォルダリスト(E)

名前	パス
共有フォルダ	

自動マウント アクセス権

共有フォルダの追加

フォルダのパス: <選択されていません>

フォルダ名: <選択されていません>

読み込み専用(R)

自動マウント(A)

他のフォルダ

OK キャンセル

フォルダの参照

Select a directory

- デスクトップ
- ライブラリ
- ホームグループ
- kadota
- コンピューター
- ネットワーク
- share
- tmp

新しいフォルダの作成(N) OK キャンセル

W9-6

「sudo mount -t vboxsf share mac_share」という赤矢印部分のコマンドを打ち込む。パスワードを聞かれて打ち込んでも何も表示されないが、遠慮なく打ち込んでリターンキーを押す。ホストOSのデスクトップ上にあるshareフォルダをゲストOSのデスクトップ上のmac_shareフォルダとしてマウントさせているつもりだがエラーが出ていることがわかる。

```
iu@bielinux[iu] ls ←
Desktop Documents Downloads Music P
iu@bielinux[iu] cd Desktop ←
iu@bielinux[Desktop] ls ←
Bio-Linux Documentation mac_share mac_sharez Sample Data
iu@bielinux[Desktop] cd mac_share ← ④ [ 3:06午後]
iu@bielinux[mac_share] ls ← ⑤ [ 3:06午後]
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share ← [ 3:06午後]
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] [ 3:07午後]
```



W9-6

「そんなファイルまたはディレクトリはない」というエラーメッセージが出ている。確かにその通り。赤矢印はゲストOSの「デスクトップ」-「mac_share」フォルダに移動してからmountコマンドを打ち込んでいるが、実際には、1つ上のディレクトリ(つまりデスクトップ)に移動してmac_shareディレクトリを見ることができる状態で打ち込むべきコマンドだから。

```
File Edit View Search Terminal Help
iu@bielinux[iu] ls ←
Desktop Documents Downloads Mu
iu@bielinux[iu] cd Desktop ←
iu@bielinux[Desktop] ls ←
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share ←
iu@bielinux[mac_share] ls ←
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] █
```

③ [3:06午後]
④ [3:06午後]
⑤ [3:06午後]
[3:06午後]
[3:07午後]



W9-6

⑦1つ上のディレクトリ(つまりデスクトップ)に移動して、赤矢印のマウントコマンドを再度実行し、⑨ mac_shareディレクトリに移動して、⑩確かにホストOSのshareフォルダ中に存在するpngファイルを視認。

```
iu@bielinux[iu] ls <----- ① [10:08午後]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop <----- ② [10:09午後]
iu@bielinux[Desktop] ls <----- ③ [10:09午後]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share <----- ④ [10:09午後]
iu@bielinux[mac_share] ls <----- ⑤ [10:09午後]
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share [10:09午後]
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] pwd <----- ⑥ [10:09午後]
/home/iu/Desktop/mac_share
iu@bielinux[mac_share] cd .. <----- ⑦ [10:10午後]
iu@bielinux[Desktop] pwd <----- ⑧ [10:10午後]
/home/iu/Desktop
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share <----- ⑧ [10:10午後]
iu@bielinux[Desktop] cd mac_share <----- ⑨ [10:10午後]
iu@bielinux[mac_share] ls <----- ⑨ [10:11午後]
hoge.png
iu@bielinux[mac_share] █ <----- ⑩ [10:11午後]
```



W9-6

打つべき場所は正しいはずなのに、Protocol error
というエラーが出るときは、VirtualBoxマネージャー
画面の「設定」-「共有フォルダー」でshareフォル
ダを共有フォルダにできているかを再度チェック。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop]
iu@bielinux[iu] ls [ 2:54午後 ]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop [ 3:04午後 ]
iu@bielinux[Desktop] ls [ 3:04午後 ]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share [ 3:04午後 ]
iu@bielinux[mac_share] ls [ 3:04午後 ]
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share [ 3:04午後 ]
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] pwd [ 3:05午後 ]
/home/iu/Desktop/mac_share
iu@bielinux[mac_share] cd .. [ 3:05午後 ]
iu@bielinux[Desktop] pwd [ 3:05午後 ]
/home/iu/Desktop
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share 3:05午後 ]
/sbin/mount.vboxsf: mounting failed with the error: Protocol error
iu@bielinux[Desktop] [ 3:06午後 ]
```



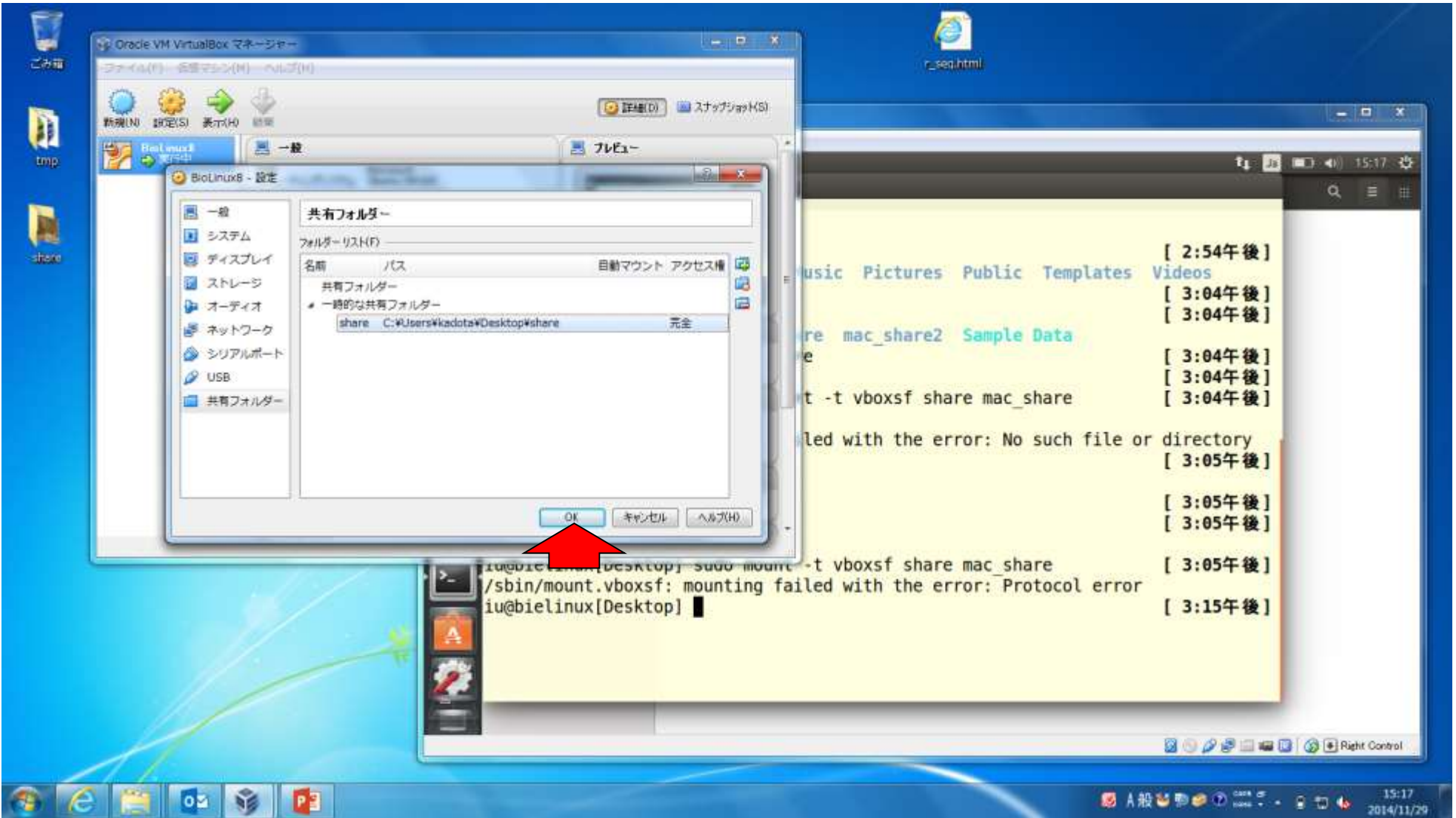
W9-6

打つべき場所は正しいはずなのに、Protocol error
というエラーが出るときは、VirtualBoxマネージャー
画面の「設定」-「共有フォルダー」でshareフォル
ダを共有フォルダにできているかを再度チェック。
これは設定が消えている例。

The screenshot shows the Oracle VM VirtualBox Manager interface. The 'BioLinux8 - 設定' (Settings) window is open, with the '共有フォルダー' (Shared Folders) tab selected. The '共有フォルダー' section is currently empty. An arrow points to the '共有フォルダー' list. In the background, a terminal window shows the following command and output:

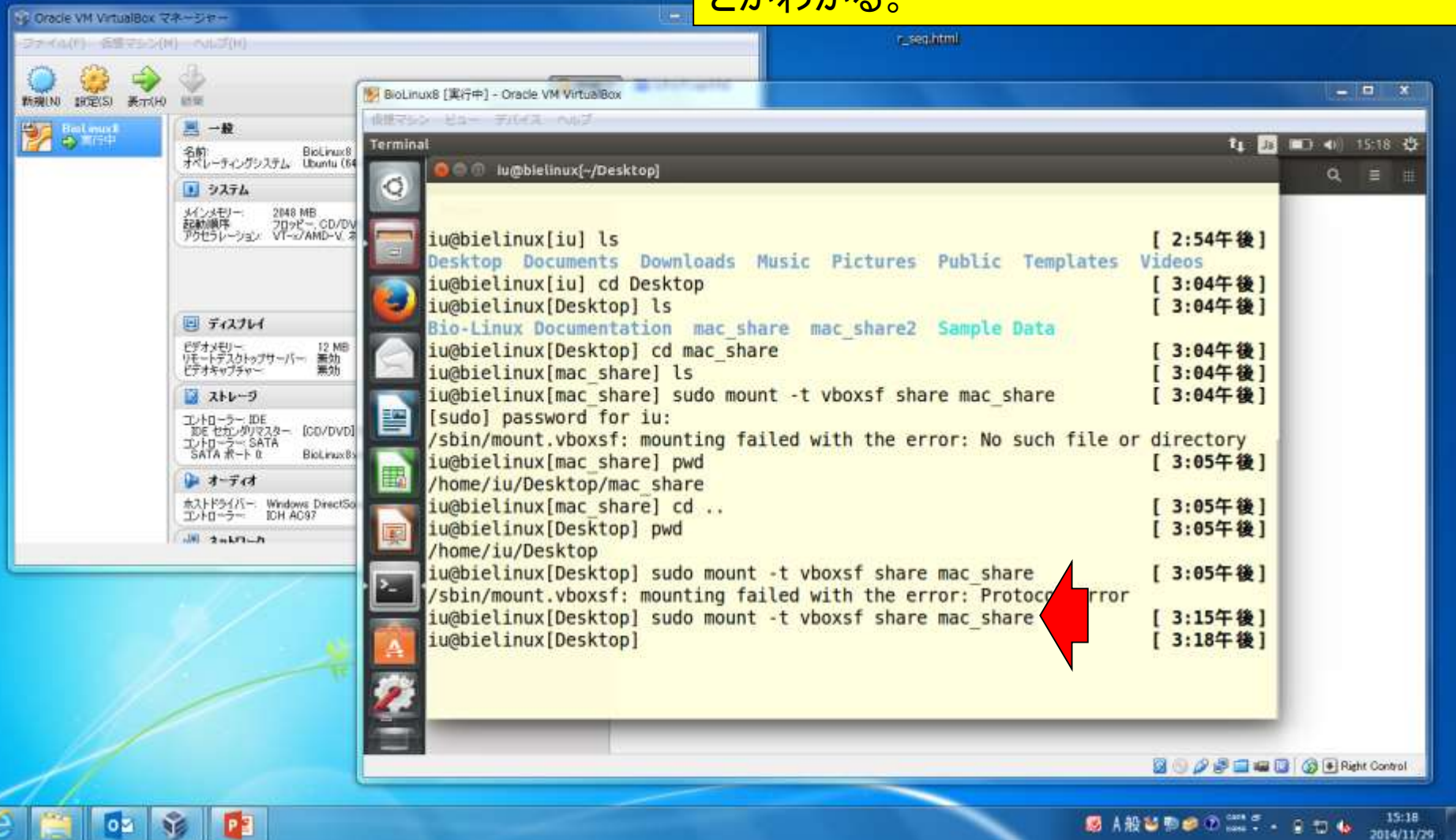
```
iu@bielinux[Desktop]:~$ sudo mount -t vboxsf share mac_share
/bin/mount.vboxsf: mounting failed with the error: Protocol error
iu@bielinux[Desktop]:~$
```

W9-6

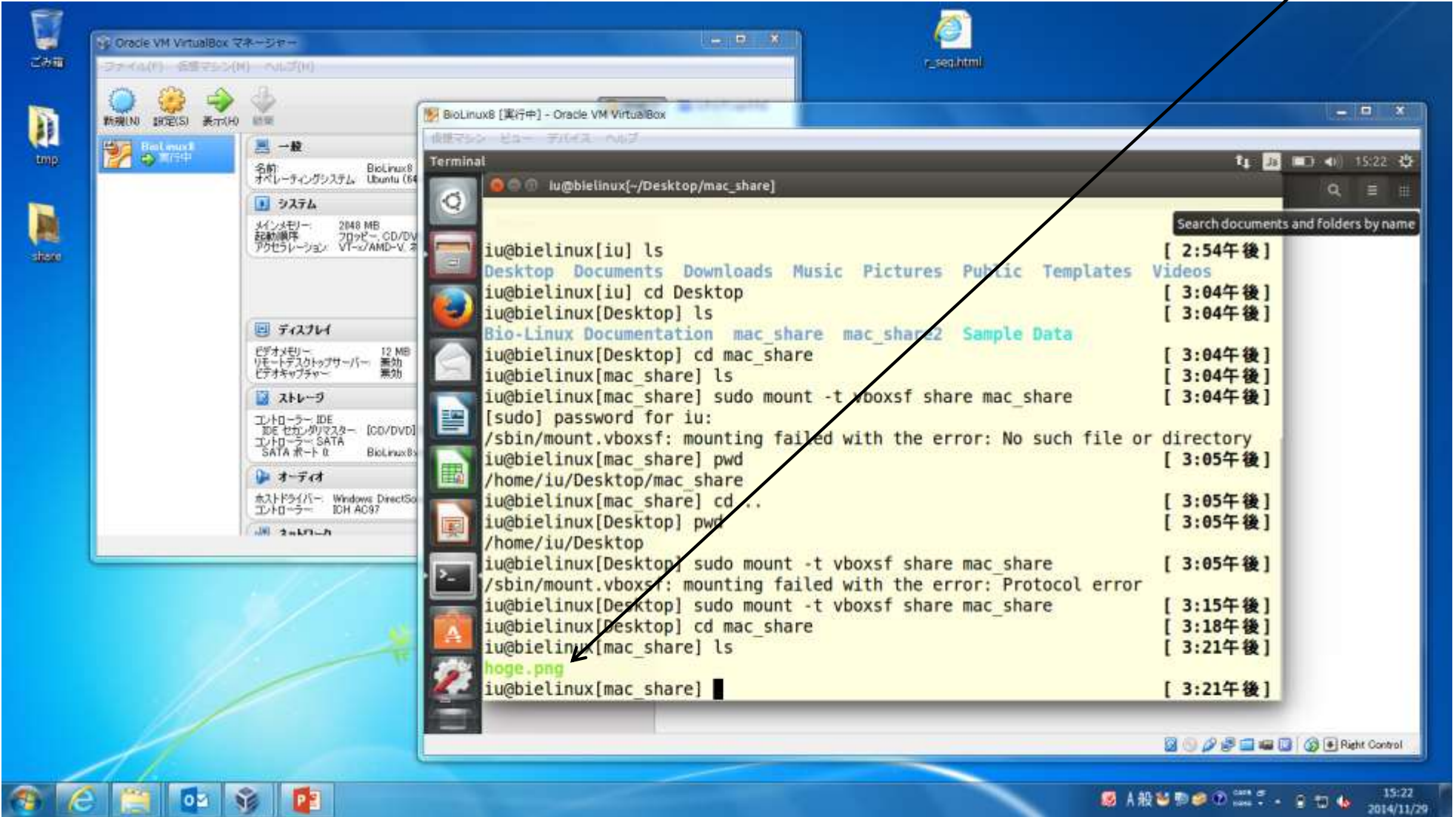


W9-6

共有フォルダ再設定後に「sudo mount -t vboxsf share mac_share」という赤矢印部分のコマンドを打ち込んだ結果。今度はエラーが出なくなっていることがわかる。



W9-6



W10-1

新しいターミナル起動後と似た状態にするやり方:「cd」と打ちこんでホームディレクトリに移動(その後のpwdはただの確認)し、「CTRLキー + Lキー」を押すと画面がリフレッシュされる。

```
iu@bielinux[iu] ls [10:08午後]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop [10:09午後]
iu@bielinux[Desktop] ls [10:09午後]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share [10:09午後]
iu@bielinux[mac_share] ls [10:09午後]
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share [10:09午後]
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] pwd [10:09午後]
/home/iu/Desktop/mac_share
iu@bielinux[mac_share] cd .. [10:10午後]
iu@bielinux[Desktop] pwd [10:10午後]
/home/iu/Desktop
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share [10:10午後]
iu@bielinux[Desktop] cd mac_share [10:10午後]
iu@bielinux[mac_share] ls [10:11午後]
hoge.png
iu@bielinux[mac_share] cd [10:11午後]
iu@bielinux[iu] pwd [11:26午後]
/home/iu
iu@bielinux[iu] [11:26午後]
```



W10-1

新しいターミナル起動後と似た状態にする
やり方:この画面はリフレッシュ後の状態。



W10-1

新しいターミナルの起動。ターミナルのアイコン上で右クリックして「New Terminal」を選択。

The screenshot shows a terminal window with the following content:

```
iu@bielinux[~/Desktop/mac_share]
iu@bielinux[iu] ls
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop
iu@bielinux[Desktop] ls
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd mac_share
iu@bielinux[mac_share] ls
iu@bielinux[mac_share] sudo mount -t vboxsf share mac_share
[sudo] password for iu:
/sbin/mount.vboxsf: mounting failed with the error: No such file or directory
iu@bielinux[mac_share] pwd
/home/iu/Desktop/mac_share
iu@bielinux[mac_share] cd ..
iu@bielinux[Desktop] pwd
/home/iu/Desktop
iu@bielinux[Desktop] sudo mount -t vboxsf share mac_share
iu@bielinux[Desktop] cd mac_share
iu@bielinux[mac_share] ls
```

A context menu is open over the terminal icon in the sidebar, with the following options:

- New Terminal
- Terminal
- Unlock from Launcher
- Quit

A red arrow points to the 'New Terminal' option.

W10-1

新しいターミナルの起動。ターミナルのアイコン上で右クリックして「New Terminal」を選択。画面は新規ターミナル起動後の状態。



W10-2

本文中の図1の詳細説明。赤の下線部分は「ユーザ名@PC名」となっている。赤矢印部分はホームディレクトリ(~)から、カレントディレクトリまでの相対パスを示している。この表示はディレクトリ移動のたびに変遷する。角括弧[]内の文字(iu, Desktop, hoge)はカレントディレクトリの名前を示し、赤矢印の表示と連動している。⑦のpwd実行結果(/home/iu/Desktop/hoge)がカレントディレクトリの絶対パスに相当する。



```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[iu] ls ← ① [10:23午前]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop ← ② [10:24午前]
iu@bielinux[Desktop] ls ← ③ [10:24午前]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] mkdir hoge ← ④ [10:24午前]
iu@bielinux[Desktop] cd hoge ← ⑤ [10:24午前]
iu@bielinux[hoge] ls ← ⑥ [10:24午前]
iu@bielinux[hoge] pwd ← ⑦ [10:24午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] whoami ← ⑧ [10:25午前]
iu
iu@bielinux[hoge] █ [10:29午前]
```

図1 Bio-Linux 8のターミナル画面。PC名はbielinux、ユーザ名はiu。ゲストOSのデスクトップ上にhogeフォルダを作成する④のmkdirコマンドを含め、計8つのコマンドを実行している。

W10-3-1

ホストOSのウェブブラウザIEを開いて、ホストOSのデスクトップのhogeに保存しようとしているところ。ここでは便宜上*L. casei* 12AのURLへのリンク (release 22)を貼っているが、2014年12月現在は release24となっていてファイル名も異なるので注意。Lhaplusのようなフリーの圧縮/解凍ソフトウェアで拡張子gzがついたgzip圧縮ファイルを解凍。

データ取得とチェックサム:

- [Ensembl Genomes: Flicek et al., Nucleic Acids Res., 2014](#)
 - [Lactobacillus casei 12A \(Taxonomy ID: 1051650\): Broadbent et al., BMC Genomics, 2012](#)
 - [Ensembl Bacteria - Lactobacillus casei 12A - Download DNA sequence](#)
[Lactobacillus casei 12a.GCA_000309565.1.22.dna.toplevel.fa.gz](#)
 - [Toplevelの概念を説明しているところ](#)
 - [Release note](#)
 - [Ensembl Bacteria](#)
[Lactobacillus cas](#)
 - [Ensembl Bacteria](#)
 - [Lactobacill](#)
 - [Escherichia](#)
 - ...
 - [Ensembl Fungi](#)
 - [Saccharom](#)
 - [Schizosacc](#)



FTP ディレクトリ /pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/ / ftp.ensemblgenomes.org

エクスプローラーでこの FTP サイトを表示するには、Alt キーを押して、[表示]をクリックし、[エクスプローラーで FTP サイトを開く]をクリックしてください。

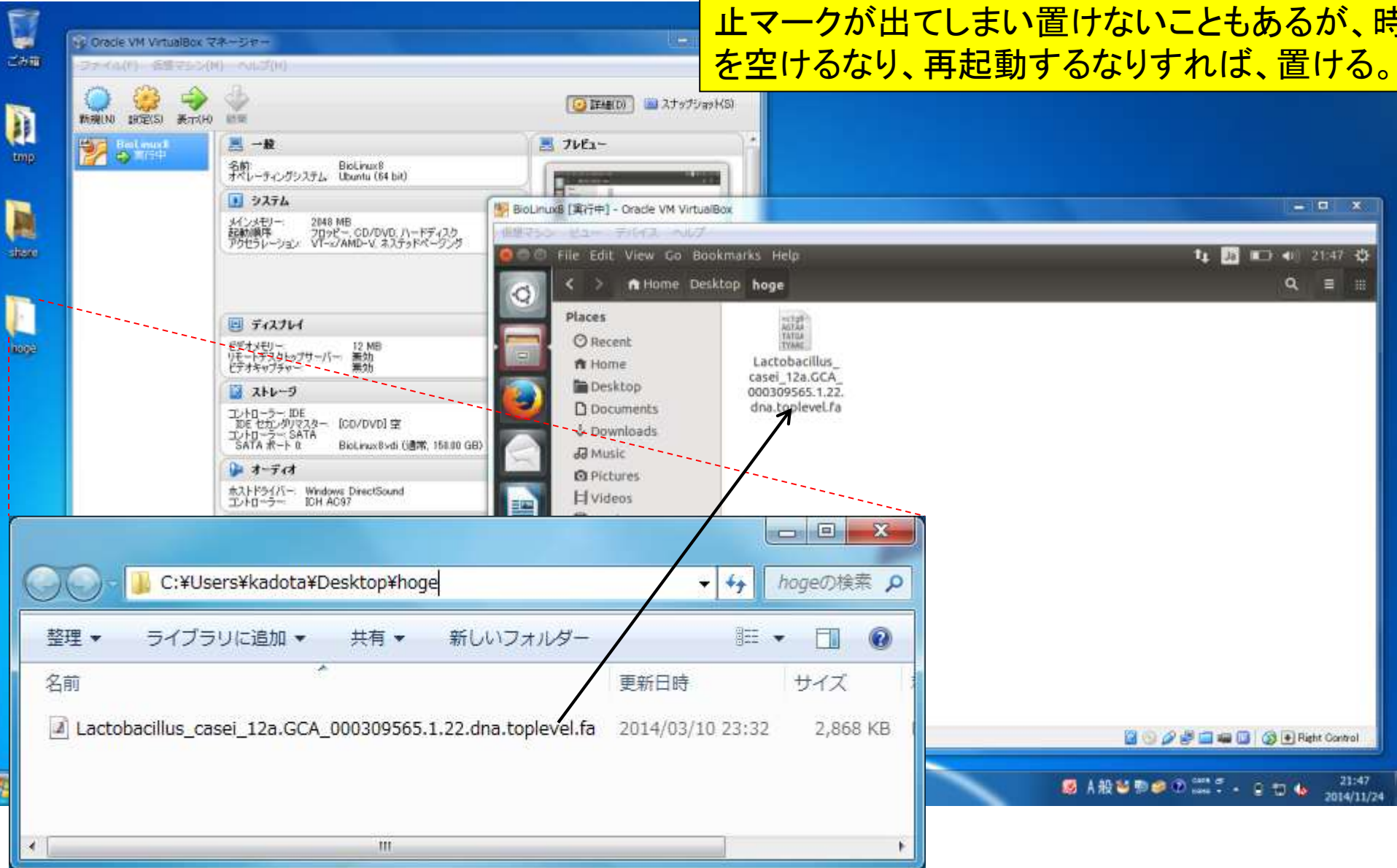
1 階層上のディレクトリへ

04/03/2014 12:00午前	469	CHECKSUMS
03/21/2014 12:00午前	902,847	Lactobacillus casei 12a.GCA_000309565.1.22.dna.genome.fa.gz
03/21/2014 12:00午前	902,847	Lactobacillus casei 12a.GCA_000309565.1.22.dna.toplevel.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus casei 12a.GCA_000309565.1.22.dna.rm.genome.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus casei 12a.GCA_000309565.1.22.dna.rm.toplevel.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus casei 12a.GCA_000309565.1.22.dna.sm.genome.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus casei 12a.GCA_000309565.1.22.dna.sm.toplevel.fa.gz
03/21/2014 12:00午前	3,111	README



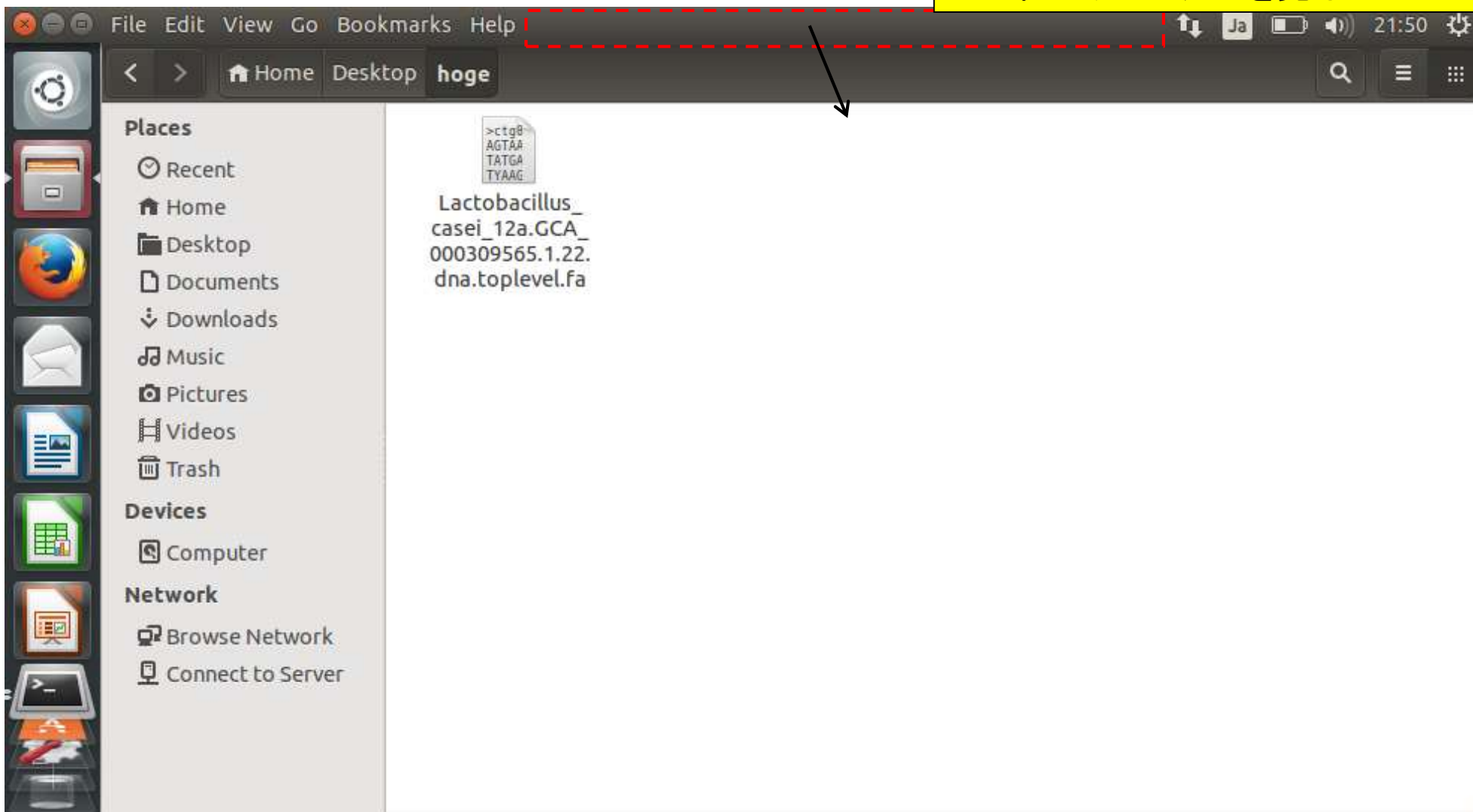
W10-3-2

やり方1: ホストOSのデスクトップのhogeフォルダ中にあるFASTAファイルを、ドラッグ&ドロップでゲストOSのデスクトップのhogeフォルダにコピー。駐車禁止マークが出てしまい置けないこともあるが、時間を空けるなり、再起動するなりすれば、置ける。



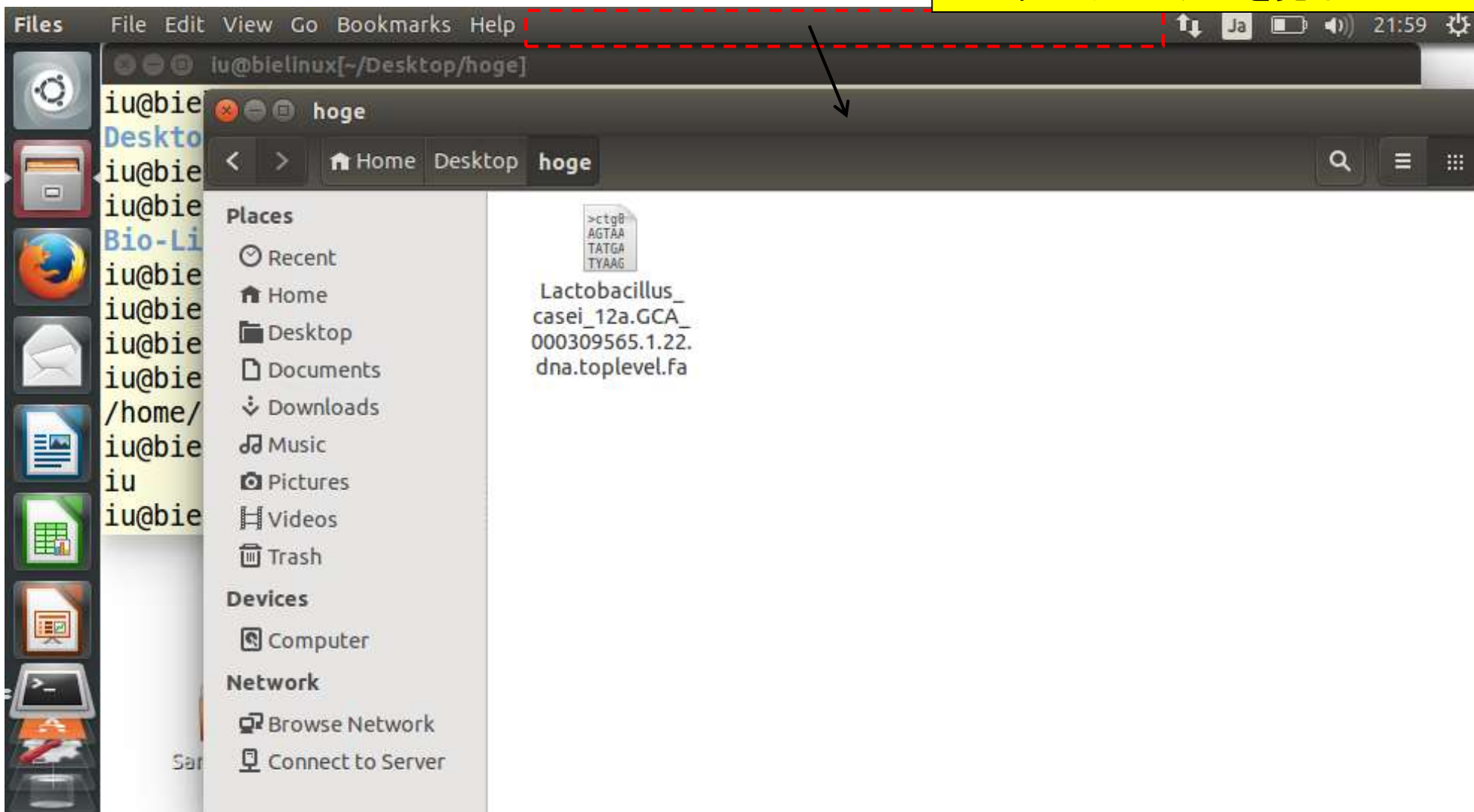
W10-3-2

Tips。赤枠付近でクリックしたまま下にずらすと、hogeフォルダの裏に隠れていた他のターミナルを見ることができる。



W10-3-2

Tips。赤枠付近でクリックしたまま下
にずらすと、hogeフォルダの裏に隠れて
いた他のターミナルを見ることができる。



W10-3-2

ターミナル画面をアクティブにして、もう一度lsすると、確かにFASTAファイルが見られる。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop/hoge]
iu@bielinux[iu] ls [10:23午前]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop [10:24午前]
iu@bielinux[Desktop] ls [10:24午前]
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] mkdir hoge [10:24午前]
iu@bielinux[Desktop] cd hoge [10:24午前]
iu@bielinux[hoge] ls [10:24午前]
iu@bielinux[hoge] pwd [10:24午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] whoami [10:25午前]
iu
iu@bielinux[hoge] ls [10:29午前]
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] [10:00午後]
```

Computer
Network
Browse Network
Connect to Server

W10-4

FASTAファイルを共有フォルダ経由でもおけることを示すべく、一旦FASTAファイルをrmコマンドを用いて削除。「rm L」まで打ってからタブ補完を有効に利用すべし。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop/hoge]
iu@bielinux[iu] ls [10:23午前]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Desktop [10:24午前]
iu@bielinux[Desktop] ls [10:24午前]
Bio-Linux Documentation mac share mac_share2 Sample Data
iu@bielinux[Desktop] mkdir hoge [10:24午前]
iu@bielinux[Desktop] cd hoge [10:24午前]
iu@bielinux[hoge] ls [10:24午前]
iu@bielinux[hoge] pwd [10:24午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] whoami [10:25午前]
iu
iu@bielinux[hoge] ls [10:29午前]
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] rm Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
rm: remove regular file 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel
.fa'?
```

W10-4

「rm: remove regular file '...' ?」は、本当に消してもいいか?と聞いてきているので、yesに相当するyを打ち込んでリターンキーを押す。その後lsすると確かにFASTAファイルが消えていることがわかる。注意点として、この作業はゴミ箱への移動を意味するわけではなく、本当に消えてしまうので、重要なファイルは気をつけるべし。

```
Terminal File Edit View Search Terminal Help
iu@bielinux[~/Desktop/hoge]
iu@bielinux[iu] ls
Desktop Documents Downloads Music
iu@bielinux[iu] cd Desktop
iu@bielinux[Desktop] ls
Bio-Linux Documentation mac_share mac_share2 Sample Data
iu@bielinux[Desktop] mkdir hoge
iu@bielinux[Desktop] cd hoge
iu@bielinux[hoge] ls
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] whoami
iu
iu@bielinux[hoge] ls
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] rm Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
rm: remove regular file 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel
.fa'? y
iu@bielinux[hoge] ls
iu@bielinux[hoge]
```

[10:24午前]

[10:24午前]

[10:24午前]

[10:24午前]

[10:24午前]

[10:24午前]

[10:25午前]

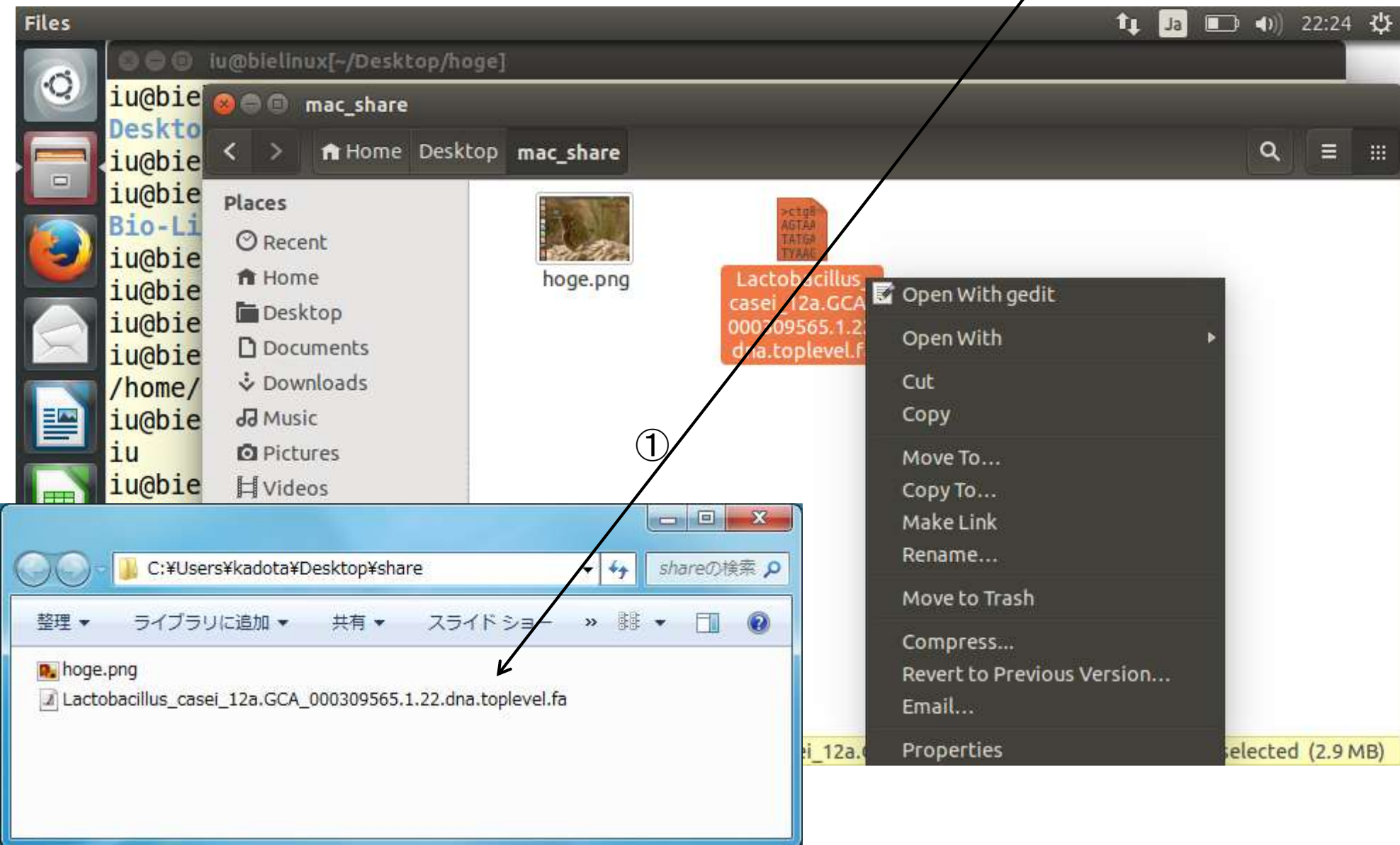
[10:29午前]

[10:13午後]

[10:13午後]

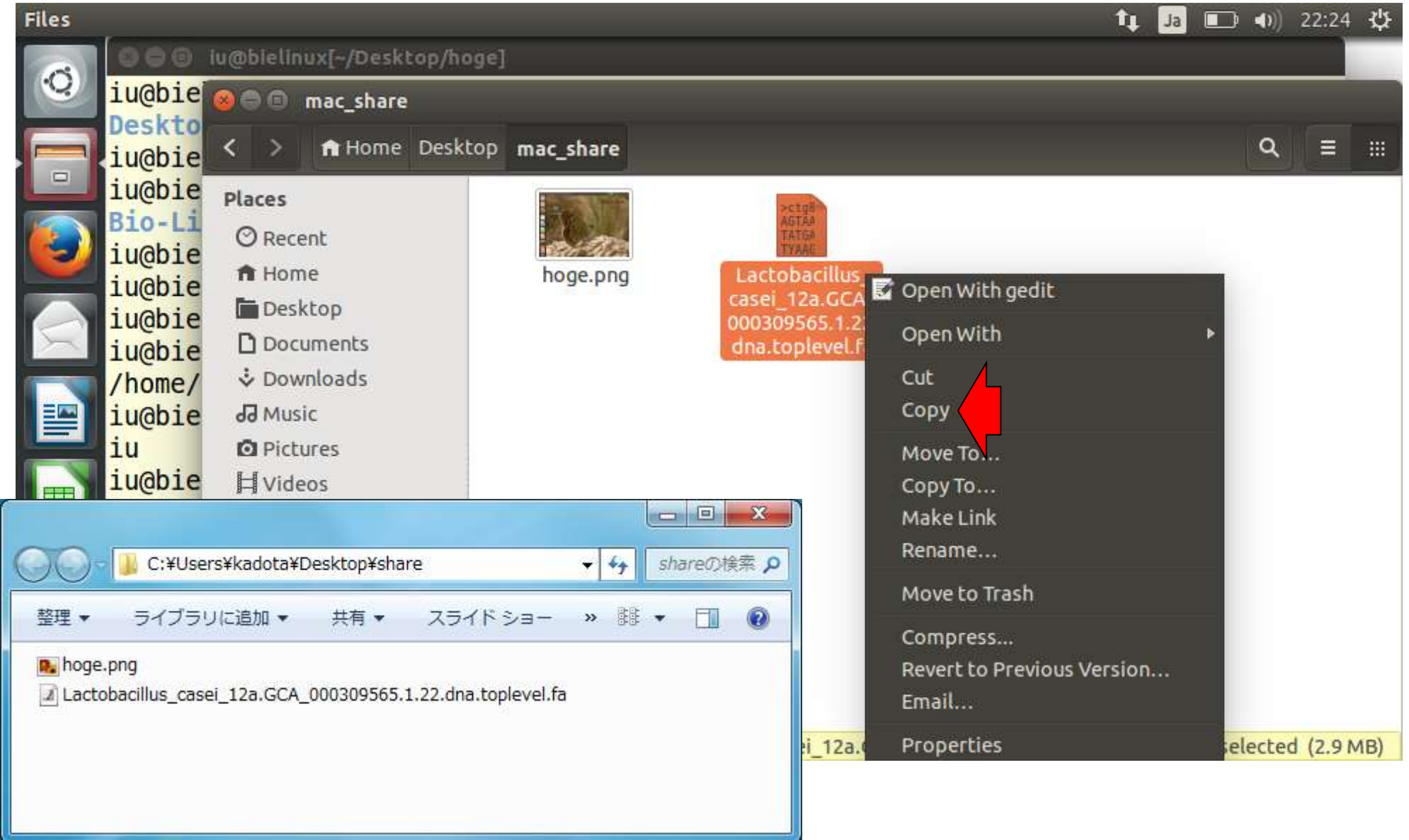
W10-5

やり方2。①共有フォルダとして設定したホストOSのデスクトップ上にあるshareフォルダにFASTAファイルをコピー。



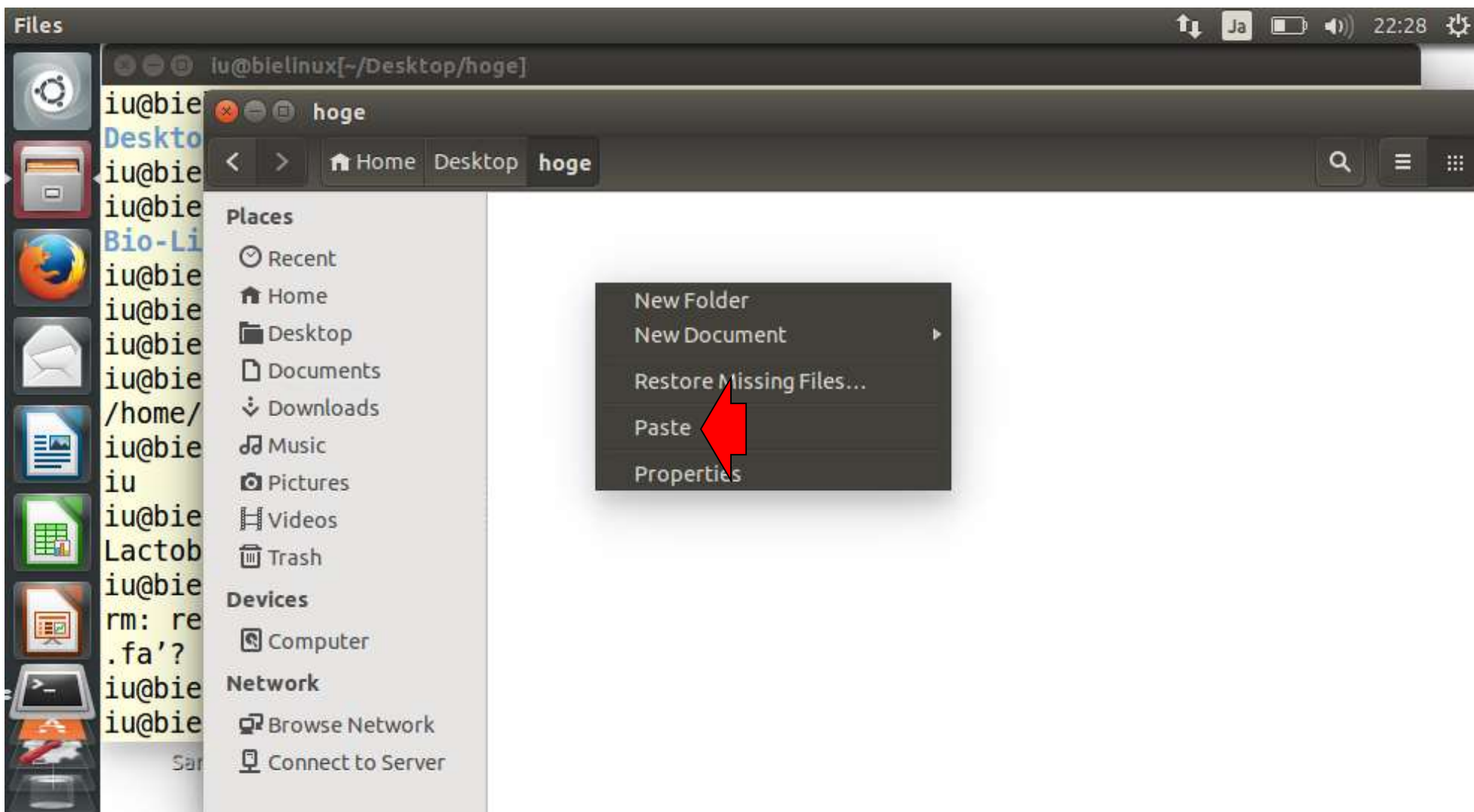
W10-5

やり方2。ゲストOSのデスクトップ上にあるmac_shareフォルダから見られるようにしたFASTAファイルをコピーして…。



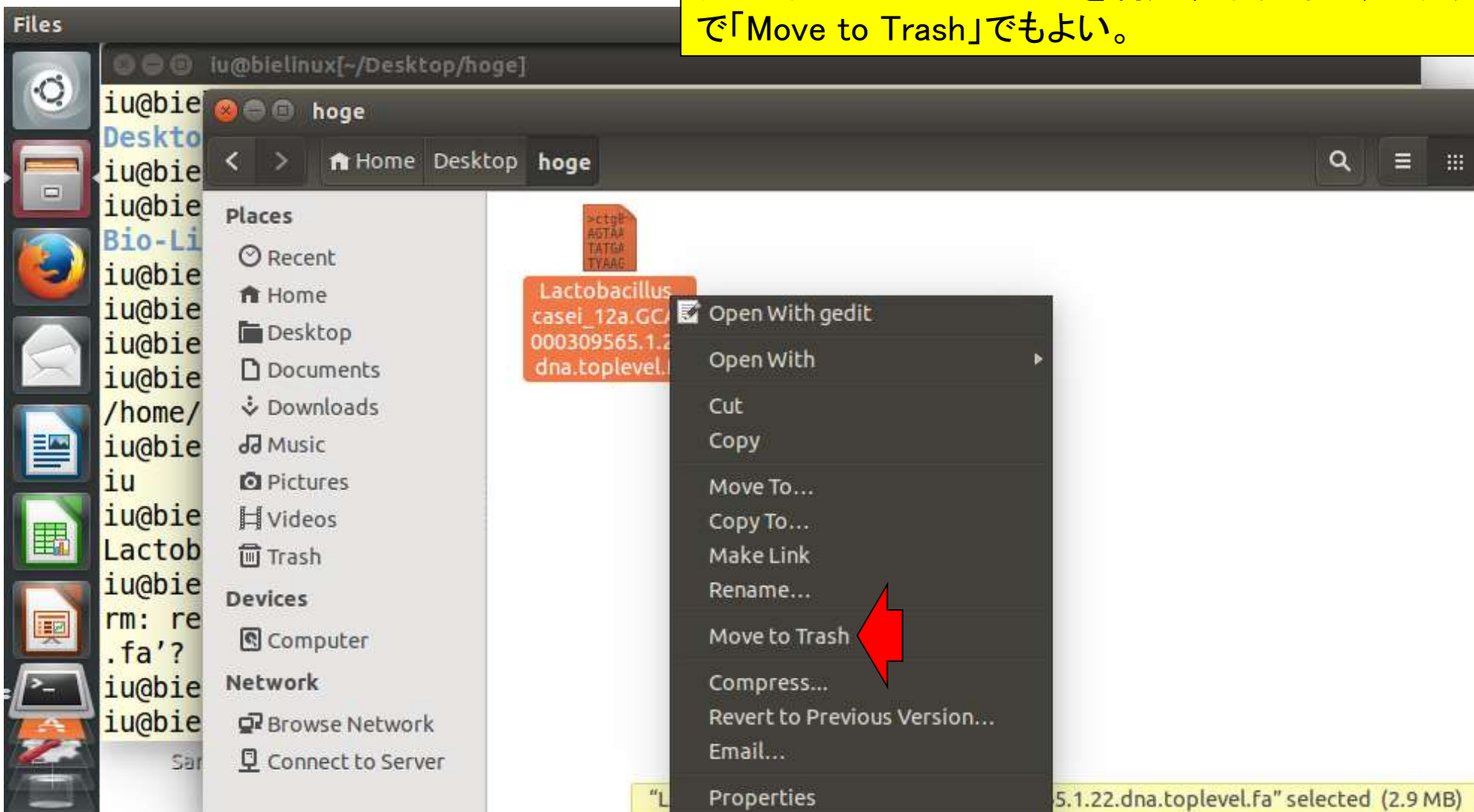
W10-5

やり方2。ゲストOSのデスクトップ上にあるhogeフォルダ上でペースト。



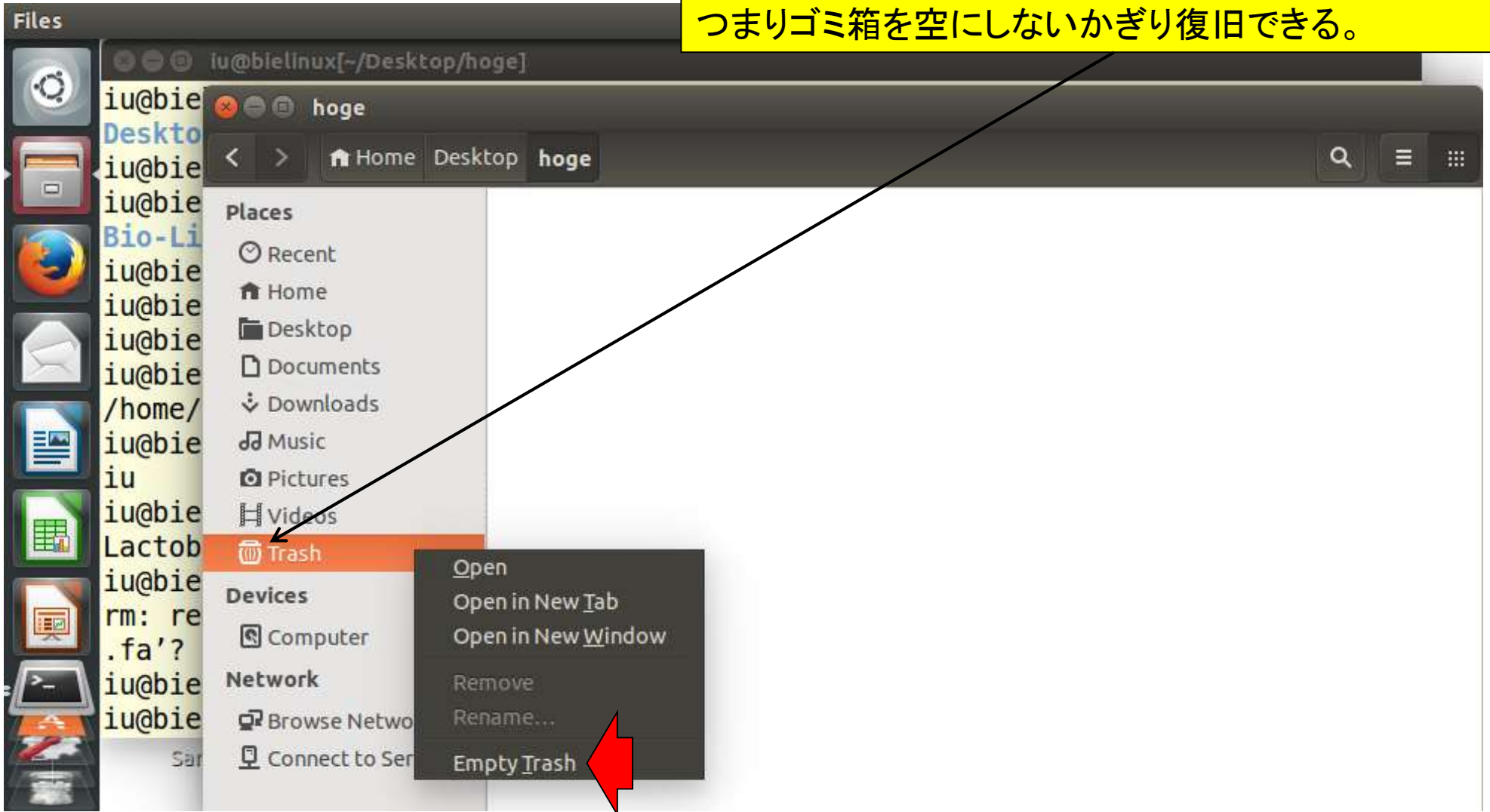
W10-5

やり方2で作成したゲストOSのデスクトップ上にあるhogeフォルダに作成されたFASTAファイルを消すのは、ターミナル上でrmコマンドを利用する以外に、右クリックで「Move to Trash」でもよい。



W10-5

FASTAファイル移動後のごみ箱はMacと同様膨れる。
rmコマンドと違って、この作業自体はゴミ箱への移動なので、ごみ箱アイコン上で右クリックして「Empty Trash」、つまりゴミ箱を空にしないかぎり復旧できる。



W11-1

ホストOS上での通常のダウンロード作業は、目的のgzip圧縮ファイル上で右クリックし「対象をファイルに保存」だが、それだとURL情報を取得できない。

データ取得とチェックサム:

- Ensembl Genomes: Flicek et al., Nucleic Acids Res., 2014
 - Lactobacillus casei 12A (Taxonomy ID: 1051650): Download DNA sequence - Lactobacillus casei 12a.GCA_000309565.1.22.dna.toplevel.fa.gz
 - Ensembl Bacteria - Lactobacillus casei 12A - Download DNA sequence - Lactobacillus casei 12a.GCA_000309565.1.22.dna.toplevel.fa.gz
- Toplevelの概念を説明しているページ
- Release
- Ensembl
- Lactoba
- Ensembl
- L
- E
- Ensembl
- S
- S

FTP ディレクトリ /pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/ / ftp.ensemblgenomes.org

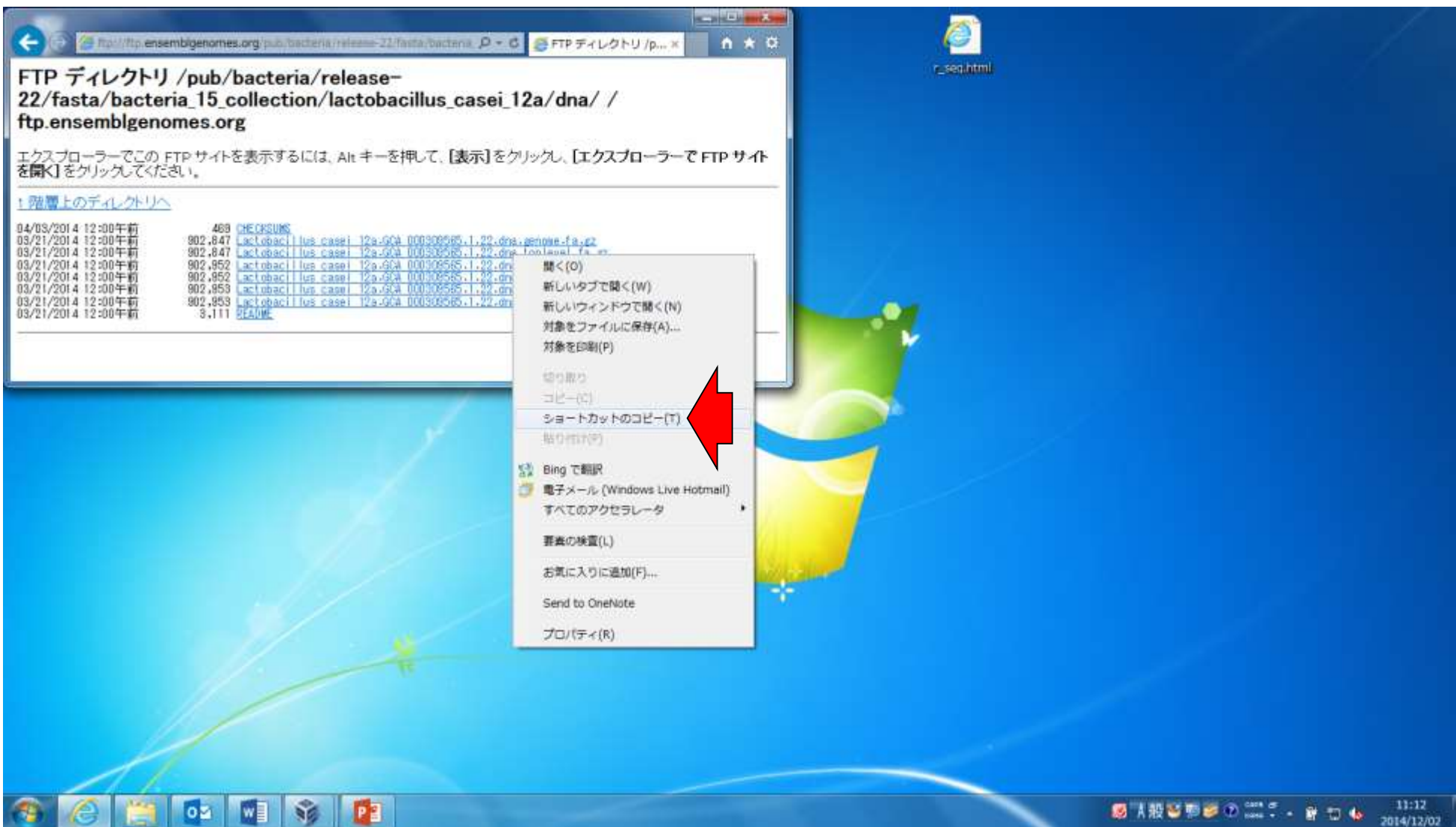
エクスプローラーでこの FTP サイトを表示するには、Alt キーを押して、[表示] をクリックし、[エクスプローラーで FTP サイトを開く] をクリックしてください。

1 階層上のディレクトリへ

04/03/2014 12:00午前	469	CHECKSUMS
03/21/2014 12:00午前	902,847	Lactobacillus casei 12a.GCA_000309565.1.22.dna.genome.fa.gz
03/21/2014 12:00午前	902,847	Lactobacillus casei 12a.GCA_000309565.1.22.dna.toplevel.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus casei 12a.GCA_000309565.1.22.dna.rm.genome.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus casei 12a.GCA_000309565.1.22.dna.rm.toplevel.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus casei 12a.GCA_000309565.1.22.dna.sm.genome.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus casei 12a.GCA_000309565.1.22.dna.sm.toplevel.fa.gz
03/21/2014 12:00午前	3,111	README

W11-1

ホストOSのInternet Explorerの場合は、右クリックの「ショートカットのコピー」でURL情報を取得できる。



W11-1

ゲストOSのFirefoxの場合は、右クリックの「Copy Link Location」でURL情報を取得できる。ここでは、その後の説明も兼ねて1つ上の階層の「Download DNA sequence」のページから行う。

The screenshot shows a virtual machine environment with a Windows 7 desktop. A Firefox browser window is open, displaying the index page for the FTP directory: ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/. The page lists several files with columns for Name, Size, and Last Modified. A right-click context menu is open over the file listing, with a red arrow pointing to the 'Copy Link Location' option. The context menu options include: Open Link in New Tab, Open Link in New Window, Open Link in New Private Window, Bookmark This Link, Save Link As..., Copy Link Location, Search Google for "File...", and Inspect Element (Q). The desktop taskbar shows various application icons, and the system tray displays the date and time as 11:15 on 2014/12/02.

Name	Size	Last Modified
CHECKSUMS	1 KB	2014年04月03日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.f...		
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.f...		
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.f...		
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.toplevel.f...		
README		

W11-1

ファイルを保存したいゲストOSのhogeフォルダ上で「wget -c」と打ち込んでから、右クリックでPaste。スペースの場所に注意。

The screenshot shows a Linux desktop environment with a terminal window open. The terminal displays the following commands and output:

```
iu@bielinux[hoge] pwd [11:19午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls /home/iu/Desktop/hoge [11:19午前]
/bacteriu@bielinux[hoge] wget -c [11:19午前]
```

The terminal output shows a file browser view with a list of files and directories. A context menu is open over the terminal, with the 'Paste' option highlighted by a red arrow. The context menu options are:

- Open Terminal
- Open Tab
- Close Window
- Copy
- Paste
- Profiles
- ✓ Show Menubar

W11-1

ペースト直後の状態。リターンキーを押すとダウンロードが始まる。

The screenshot shows a terminal window with the following content:

```
iu@bielinux[hoge] pwd [11:19午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls /bacteria [11:19午前]
iu@bielinux[hoge] wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22
/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_
12a.GCA_000309565.1.22.dna.toplevel.fa.gz
```

Name	Size	Last Modified
CHECKSUMS	1 KB	2014年04月03日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz	652 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz	652 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.fa.gz	652 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.fa.gz	652 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.fa.gz	652 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.toplevel.fa.gz	652 KB	2014年03月21日 00時00分00秒
README	4 KB	2014年03月21日 00時00分00秒

W11-1

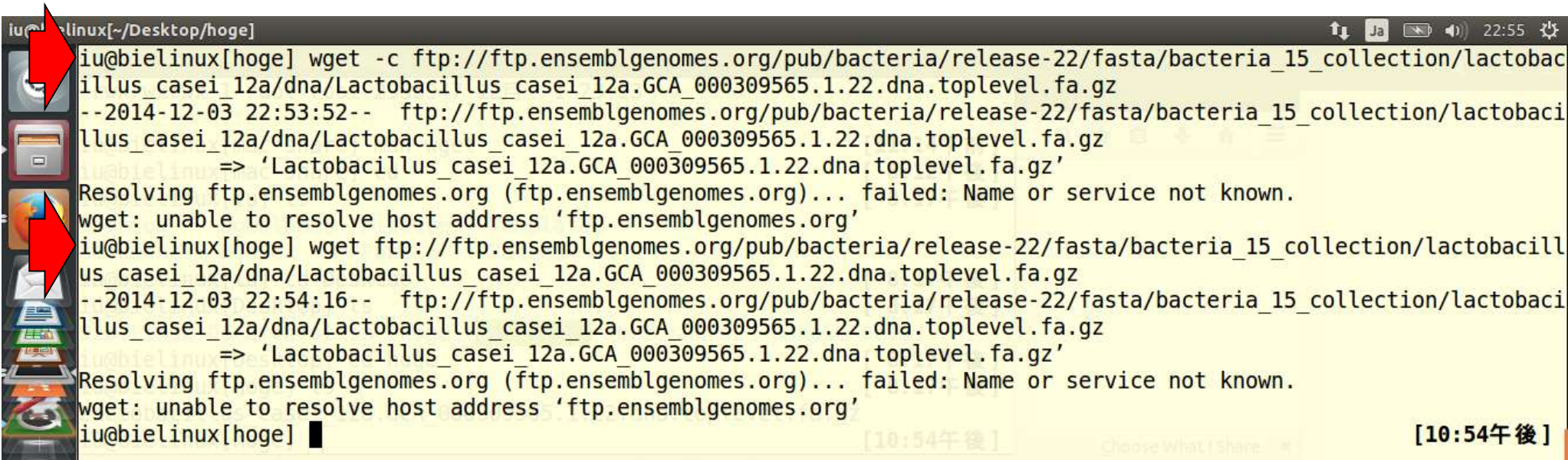
無事ダウンロードが終わると通常のコマンド入力待ち状態となる。ここでは、lsして確認したところまでを示す。

```
Terminal
Index of ftp://ftp.en...
ftp://ftp.ensem... iu@bielinux[~/Desktop/hoge]
Resolving ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)... 193.62.197.94
Connecting to ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)|193.62.197.94|:21... connected.
Logging in as anonymous ... Logged in!
==> SYST ... done. ==> PWD ... done.
^ Up to h ==> TYPE I ... done. ==> CWD (1) /pub/bacteria/release-22/fasta/bacteria_15_c
ollection/lactobacillus_casei_12a/dna ... done.
Nan ==> SIZE Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz ... 902
847
==> PASV ... done. ==> RETR Lactobacillus_casei_12a.GCA_000309565.1.22.dna.
toplevel.fa.gz ... done.
Length: 902847 (882K) (unauthoritative)
100%[=====>] 902,847 215KB/s in 4.1s
2014-12-02 11:42:40 (215 KB/s) - 'Lactobacillus_casei_12a.GCA_000309565.1.22.d
na.toplevel.fa.gz' saved [902847]

iu@bielinux[hoge] ls [11:42午前]
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] [11:43午前]
```

W11-1

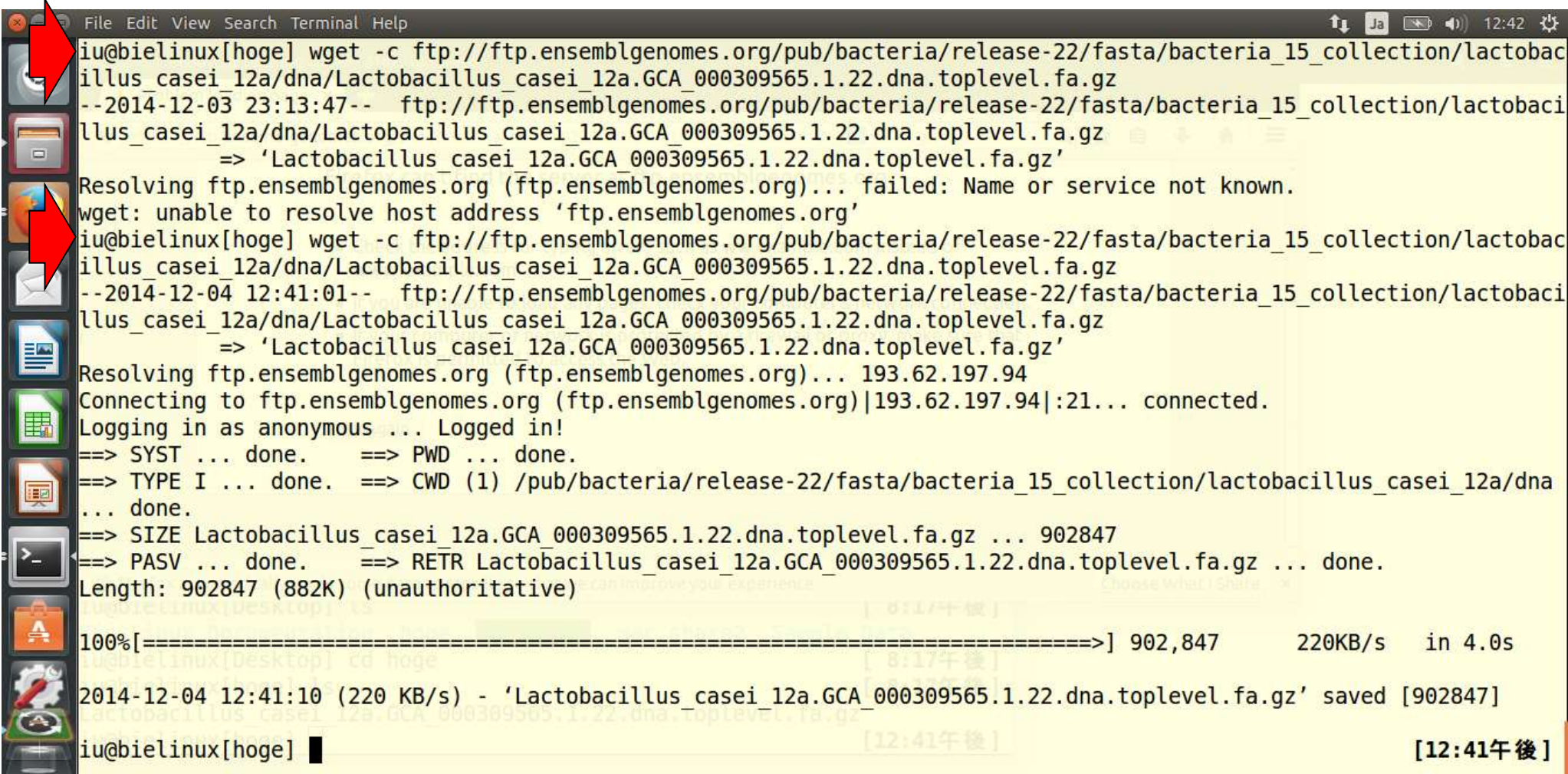
Tips。ときどきダウンロードに失敗した(failed)というメッセージに遭遇することがある。ホストOSのウェブ経由だとうまくいくが、このような場合はゲストOSのウェブもなぜか利用不可能になっている。理由は不明であるが、数十分、あるいは数時間後にリトライするとうまくいく場合が多い。



```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
--2014-12-03 22:53:52-- ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
=> 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz'
Resolving ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)... failed: Name or service not known.
wget: unable to resolve host address 'ftp.ensemblgenomes.org'
iu@bielinux[hoge] wget ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
--2014-12-03 22:54:16-- ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
=> 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz'
Resolving ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)... failed: Name or service not known.
wget: unable to resolve host address 'ftp.ensemblgenomes.org'
iu@bielinux[hoge] █
```

[10:54午後] Choose what to share

W11-1

A terminal window showing two wget commands. The first command fails with a DNS error. The second command succeeds after a delay. Two red arrows point to the first and second command lines.

```
iu@bielinux[hoge] wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
--2014-12-03 23:13:47-- ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
=> 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz'
Resolving ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)... failed: Name or service not known.
wget: unable to resolve host address 'ftp.ensemblgenomes.org'

iu@bielinux[hoge] wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
--2014-12-04 12:41:01-- ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
=> 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz'
Resolving ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)... 193.62.197.94
Connecting to ftp.ensemblgenomes.org (ftp.ensemblgenomes.org)|193.62.197.94|:21... connected.
Logging in as anonymous ... Logged in!
==> SYST ... done.      ==> PWD ... done.
==> TYPE I ... done.    ==> CWD (1) /pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna
... done.
==> SIZE Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz ... 902847
==> PASV ... done.     ==> RETR Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz ... done.
Length: 902847 (882K) (unauthoritative)

100%[=====] 902,847      220KB/s   in 4.0s

2014-12-04 12:41:10 (220 KB/s) - 'Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz' saved [902847]

iu@bielinux[hoge] █
```

W11-2

ファイルサイズは、ホストOS上で眺めると902,847、ゲストOSだと882 KBと微妙に異なっているように見えるが、実は表示のさせ方の違いだけのようである。

The image shows a host OS window displaying an FTP directory listing. The listing includes a 'CHECKSUMS' file with a size of 469 and a 'README' file with a size of 3,111. Other files are listed with sizes of 902,847. A yellow callout box points to these sizes, stating that the difference between the host OS view (902,847) and the guest OS view (882 KB) is due to display differences.

The guest OS window shows a browser view of the same directory listing. The table below summarizes the data shown in the browser view:

Name	Size	Last Modified
CHECKSUMS	1 KB	2014年04月03日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
README	4 KB	2014年03月21日 00時00分00秒

W11-2

実際にダウンロードしたファイルを「ls -la」で眺めると902847となり、ホストOS上のIEで表示されている数値(902,847)と同じ。

The screenshot shows a desktop environment with a web browser and a terminal window. The browser window displays the FTP directory listing for `ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/`. The listing shows several files, with the first one having a size of 902,847. The terminal window shows the command `ls -la` being executed, resulting in a file listing where the size `902847` is highlighted with a red arrow pointing to the same value in the browser's listing.

FTP ディレクトリ /pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/ // ftp.ensemblgenomes.org

1階層上のディレクトリ

日時	サイズ	名前
04/03/2014 12:00午前	469	CHECKSUM
03/21/2014 12:00午前	902,847	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz
03/21/2014 12:00午前	902,847	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.rm.genome.fa.gz
03/21/2014 12:00午前	902,952	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.rm.toplevel.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.sm.genome.fa.gz
03/21/2014 12:00午前	902,953	Lactobacillus_casei_12a.GCA_000309565.1.22.dna.sm.toplevel.fa.gz
03/21/2014 12:00午前	3,111	README

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -la
total 892
drwxrwxr-x 2 iu iu 4096 12月 2 11:42 .
drwx----- 6 iu iu 4096 12月 1 23:11 ..
-rw-rw-r-- 1 iu iu 902847 12月 2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge]
```

W11-2

実際にダウンロードしたファイルを「ls -lh」で眺めると882Kとなり、ゲストOS上のFirefoxで表示されている数値(882K)と同じ。

Index of ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/

Name	Size	Last Modified
CHECKSUMS	1 KB	2014年04月03日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
README		

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd [ 5:09午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [ 5:10午後 ]
total 884K
-rw-rw-r-- 1 iu iu 882K 12月 2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] [ 5:10午後 ]
```

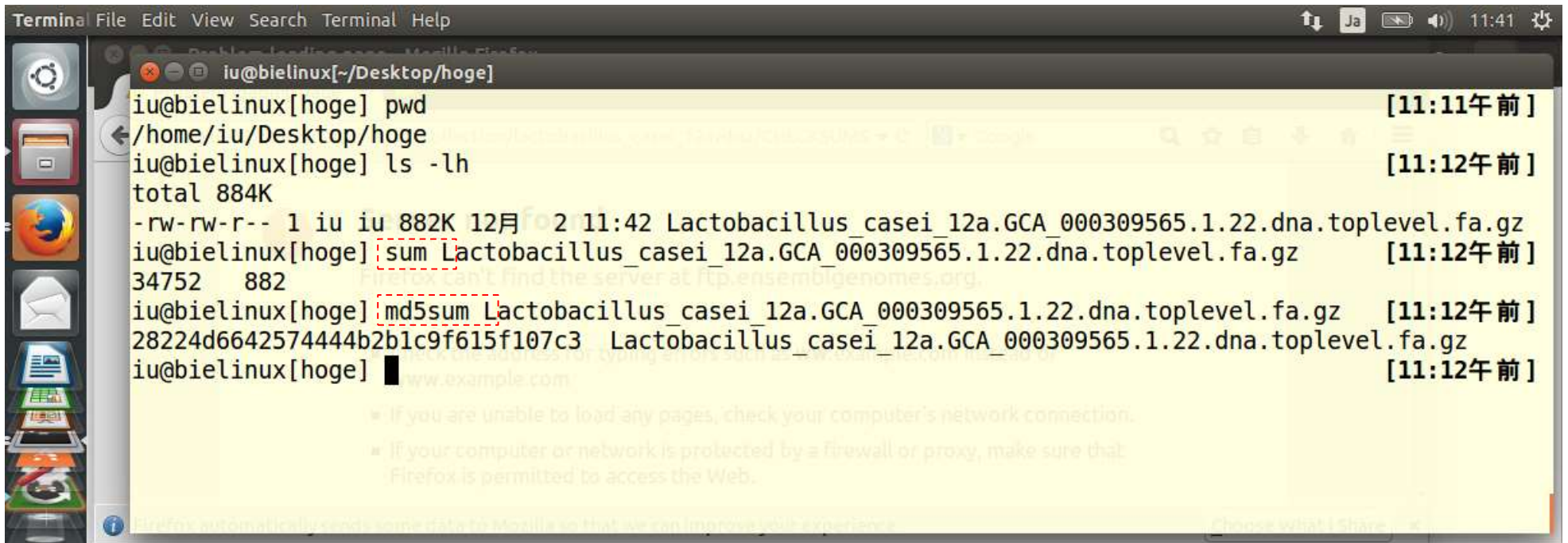
W12

The screenshot shows a web browser window displaying the FTP directory listing for `ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/`. The listing includes a file named `CHECKSUMS` with a size of 1 KB, last modified on 2014年04月03日 00時00分00秒. A red arrow points to this file. Below the browser window, a terminal window shows the same directory listing, with another red arrow pointing to the `CHECKSUMS` file.

Name	Size	Last Modified
CHECKSUMS	1 KB	2014年04月03日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.fa.gz	882 KB	2014年03月21日 00時00分00秒
Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.toplevel.fa.gz	882 KB	2014年03月21日 00時00分00秒
README	4 KB	2014年03月21日 00時00分00秒

W12

sumコマンドとmd5sumコマンド実行結果。赤枠部分のみ打ち込んでタブ補完するテクニックを有効活用すべし。



The image shows a terminal window with the following content:

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd [11:11午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [11:12午前]
total 884K
-rw-rw-r-- 1 iu iu 882K 12月 2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [11:12午前]
34752 882
iu@bielinux[hoge] md5sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [11:12午前]
28224d6642574444b2b1c9f615f107c3 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] [11:12午前]
```

The words 'sum' and 'md5sum' in the terminal are highlighted with red dashed boxes. The terminal also shows a Firefox error message: 'Firefox can't find the server at ftp.ensemblgenomes.org.'

W12

データ提供元(Ensembl)のウェブページではsumコマンド実行結果の「チェックサム」が与えられていることがわかる。同じ値(34752 882)となっていれば、ダウンロードしたファイルの中身はデータ提供元と同じと判断してよい。

```
34752 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.genome.fa.gz
34752 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
14283 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.genome.fa.gz
14283 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna_rm.toplevel.fa.gz
20112 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.genome.fa.gz
20112 882 Lactobacillus_casei_12a.GCA_000309565.1.22.dna_sm.toplevel.fa.gz
12836 4 README
```

```
Terminal
lu@bielinux[~/Desktop/hoge]
lu@bielinux[hoge] pwd [11:11午前]
/home/lu/Desktop/hoge
lu@bielinux[hoge] ls -lh [11:12午前]
total 884K
-rw-rw-r-- 1 lu lu 882K 12月 2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
lu@bielinux[hoge] sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [11:12午前]
34752 882
lu@bielinux[hoge] md5sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [11:12午前]
28224d6642574444b2b1c9f615f107c3 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
lu@bielinux[hoge] █ [11:12午前]
```

W12

Tips。Linuxコマンドもタブ補完可能。「MD5チェックサム」を実行するコマンドは「md…なんだっけ?!」という場合は、「md」まで打ち込んでからTabキーを押す。すると、mdから始まるLinuxコマンドの候補がリストアップされるので、それを眺めながら残りの「md5sum」を打ち込めばよい。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh
total 884K
-rw-rw-r-- 1 iu iu 882K 12月  2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [ 8:13午後 ]
34752 882
iu@bielinux[hoge] md5sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [ 8:13午後 ]
28224d6642574444b2b1c9f615f107e3 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] md [ 8:13午後 ]
md5pass      md5sum.textutils  mdel          mdir          mdust
md5sum       mdatopbm         mdeltree      mdu
```


W13-1

①gzip圧縮ファイルの解凍、②gzip圧縮ファイルの作成、および③チェックサム値の確認。ダウンロード直後のチェックサム値やファイルサイズとは異なるが、基本的に気にしなくてもよい。

```
iu@bielinux[hoge] ls -lh
total 884K
-rw-rw-r-- 1 iu iu 882K 12月  2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] gzip -d Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] ls -lh
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月  2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] gzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] ls -lh
total 856K
-rw-rw-r-- 1 iu iu 853K 12月  2 11:42 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] sum Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
32543 853
iu@bielinux[hoge]
```

W13-2

①gzip圧縮ファイルの解凍は「gzip -d」以外に「gunzip」コマンドでも可能。②bzip2コマンドを用いたbzip2圧縮（拡張子は.bz2）、③bunzip2コマンドを用いたbzip2圧縮ファイルの解凍。bunzip2の代わりに「bzip2 -d」を用いるのでもよい。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh
total 884K
-rw-rw-r-- 1 iu iu 882K 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] ls -lh
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] bzip2 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] ls -lh
total 824K
-rw-rw-r-- 1 iu iu 822K 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.bz2
iu@bielinux[hoge] bunzip2 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.bz2
iu@bielinux[hoge] ls -lh
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge]
```

[12:46午後]

[12:46午後]

[12:48午後]

[12:48午後]

[12:56午後]

[12:56午後]

[12:57午後]

[12:57午後]

圧縮率はbzip2のほうがgzipに比べて高いことがわかる。

W13-3

```
iu@bielinux[hoge] pwd [12:41午後]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [12:46午後]
total 884K
-rw-rw-r-- 1 iu iu 882K 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz [12:46午後]
iu@bielinux[hoge] ls -lh [12:48午後]
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] bzip2 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa [12:48午後]
iu@bielinux[hoge] ls -lh [12:56午後]
total 824K
-rw-rw-r-- 1 iu iu 822K 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.bz2
iu@bielinux[hoge] bunzip2 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.bz2 [12:56午後]
iu@bielinux[hoge] ls -lh [12:57午後]
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月  4 12:41 Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] [12:57午後]
```



W14-1

①mvコマンドでファイル名変更、②grepコマンドで”>”を含む行の数を表示、③headコマンドでgenome.faの最初の10行を表示。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls
Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa
iu@bielinux[hoge] mv Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa genome.fa
iu@bielinux[hoge] ls
genome.fa
iu@bielinux[hoge] grep -c ">" genome.fa
28
iu@bielinux[hoge] head genome.fa
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCCCTCAAGCACCCCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
iu@bielinux[hoge] █
```

[9:54午後]

W14-2

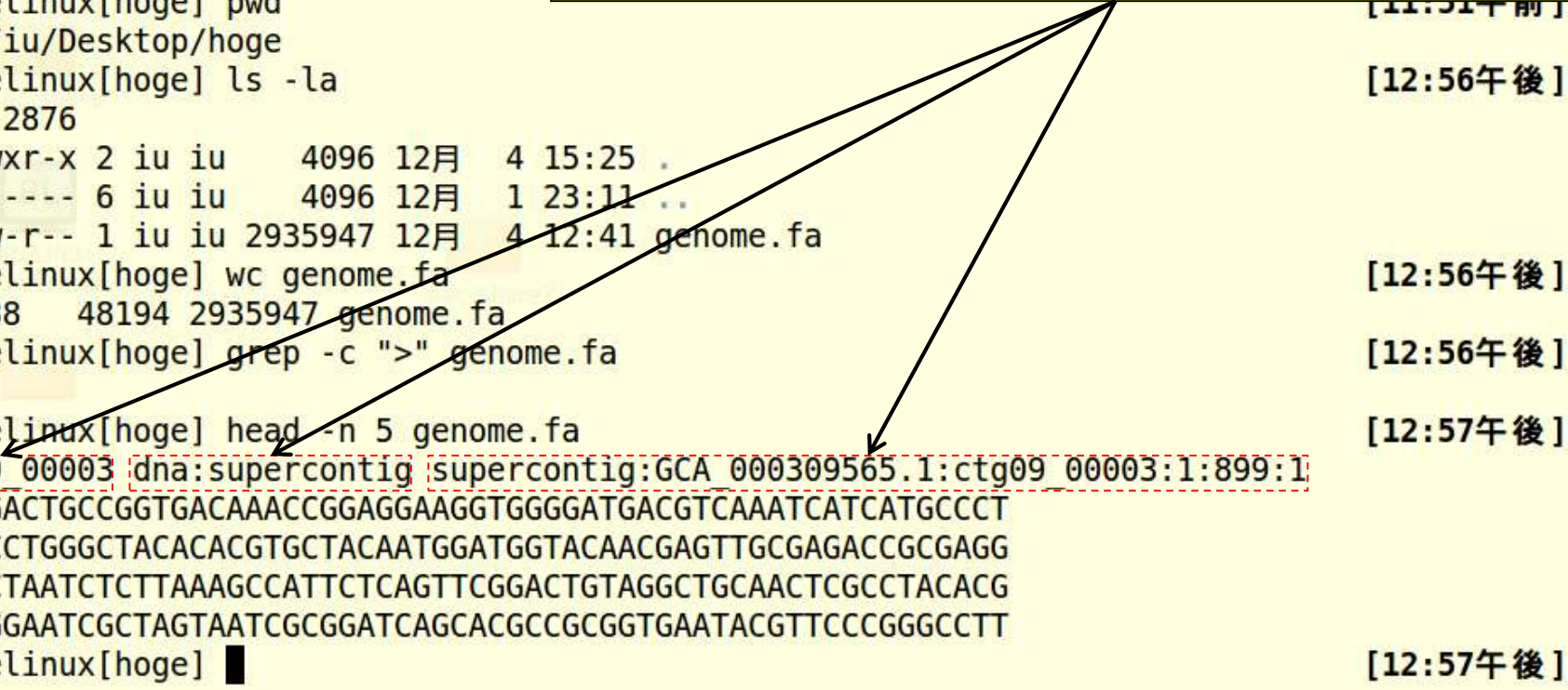
- ① headコマンドでgenome.faの最初の4行分を表示、
- ② tailコマンドでgenome.faの最後の7行分を表示。

```
iu@bielinux[hoge] pwd [11:24午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls [11:24午前]
genome.fa
iu@bielinux[hoge] head -n 4 genome.fa [11:24午前]
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
iu@bielinux[hoge] tail -n 7 genome.fa [11:24午前]
ACCTACTCAGTGTCCATGCTATGCTTGTAGCGTCGAAACCAAAGCAAGCGAGGTTATG
AGTAAATGTCCCAATACGATCCTACACTGTCCGTCCTTGAATACCAGACCATAATATCA
AAGTAGCCTTTGTTTCGTCATGAATATCGCGGCAACGGGGTACGTCGCCGCCAGTATCATG
TGATTGATGCCGAGCTGACTTACCGGTTAACCCGGTGCCCACTGTGTGGCTTTGAGGCCT
TGCACCCTAACGGGTTTTATACGGCCCACGTGCGCGTCCTCAACGGGGTTGAAATACCGA
CAGTCATTGACTTGCACAAGCAACGATGGCGCTGTCATAACTGTTACCACACAGTCAGTG
CCAAGACGCCACTCGTGCAAC
iu@bielinux[hoge] [11:24午前]
```

W14-3

wcコマンドで「行数、単語数、ファイルサイズ」を表示。行数は48138、単語数は48194、ファイルサイズは2935947である。塩基配列部分は実質的に1行1単語となるので、行数と単語数の若干の違いはdescription行のところが出る。

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -la
total 2876
drwxrwxr-x 2 iu iu 4096 12月 4 15:25 .
drwx----- 6 iu iu 4096 12月 1 23:11 ..
-rw-rw-r-- 1 iu iu 2935947 12月 4 12:41 genome.fa
iu@bielinux[hoge] wc genome.fa
48138 48194 2935947 genome.fa
iu@bielinux[hoge] grep -c ">" genome.fa
28
iu@bielinux[hoge] head -n 5 genome.fa
>ctg09 00003 dna:supercontig supercontig:GCA 000309565.1:ctg09 00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATGCCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
iu@bielinux[hoge]
```



W14-3

Description部分は1行あたり3単語。*L. casei* 12Aゲノムは全部で28コンティグ。このファイルの行数は48138なので、予想される単語数は $48138 + 28 * 2 = 48194$ となり、wcコマンド実行結果と同じになる。このようにファイル全体を眺めなくてもwc, grep, headコマンドの組合せでdescription行の記述形式が同じであることを予想して確認もできる。

iu@bielinux[~/Desktop/hoge]

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
```

```
iu@bielinux[hoge] ls -la
total 2876
```

```
drwxrwxr-x 2 iu iu 4096 12月 4 15:25 .
drwx----- 6 iu iu 4096 12月 1 23:11 ..
-rw-rw-r-- 1 iu iu 2935947 12月 4 12:41 genome.fa
```

```
iu@bielinux[hoge] wc genome.fa
```

```
48138 48194 2935947 genome.fa
```

[12:56午後]

```
iu@bielinux[hoge] grep -c ">" genome.fa
28
```

[12:56午後]

```
iu@bielinux[hoge] head -n 5 genome.fa
```

[12:57午後]

```
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
```

```
iu@bielinux[hoge] █
```

[12:57午後]

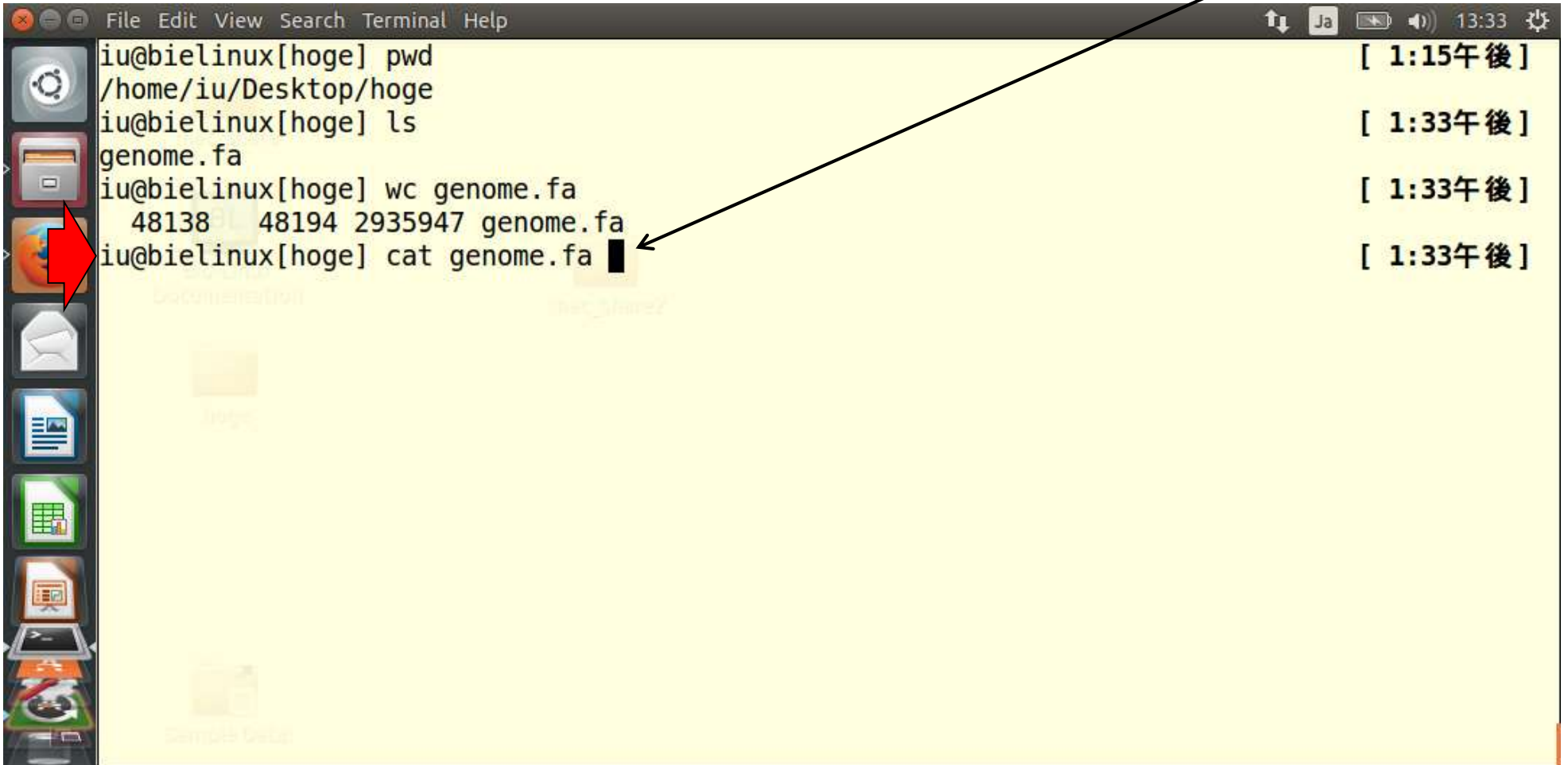
W14-3

grepコマンド実行時に">"を含む行数をカウントするのが「-c」オプションだが、さらに「-v」オプションを追加することで">"を含まない行数をカウントすることができる。

```
iu@bielinux[hoge] pwd [11:51午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -la [12:56午後]
total 2876
drwxrwxr-x 2 iu iu 4096 12月 4 15:25 .
drwx----- 6 iu iu 4096 12月 1 23:11 ..
-rw-rw-r-- 1 iu iu 2935947 12月 4 12:41 genome.fa
iu@bielinux[hoge] wc genome.fa [12:56午後]
48138 48194 2935947 genome.fa
iu@bielinux[hoge] grep -c ">" genome.fa [12:56午後]
28
iu@bielinux[hoge] head -n 5 genome.fa [12:57午後]
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
iu@bielinux[hoge] grep -c -v ">" genome.fa [12:57午後]
48110
iu@bielinux[hoge] [ 1:15午後]
```


W14-4

catコマンド実行前。リターンキーを押すと一気にファイルの全内容が表示される。



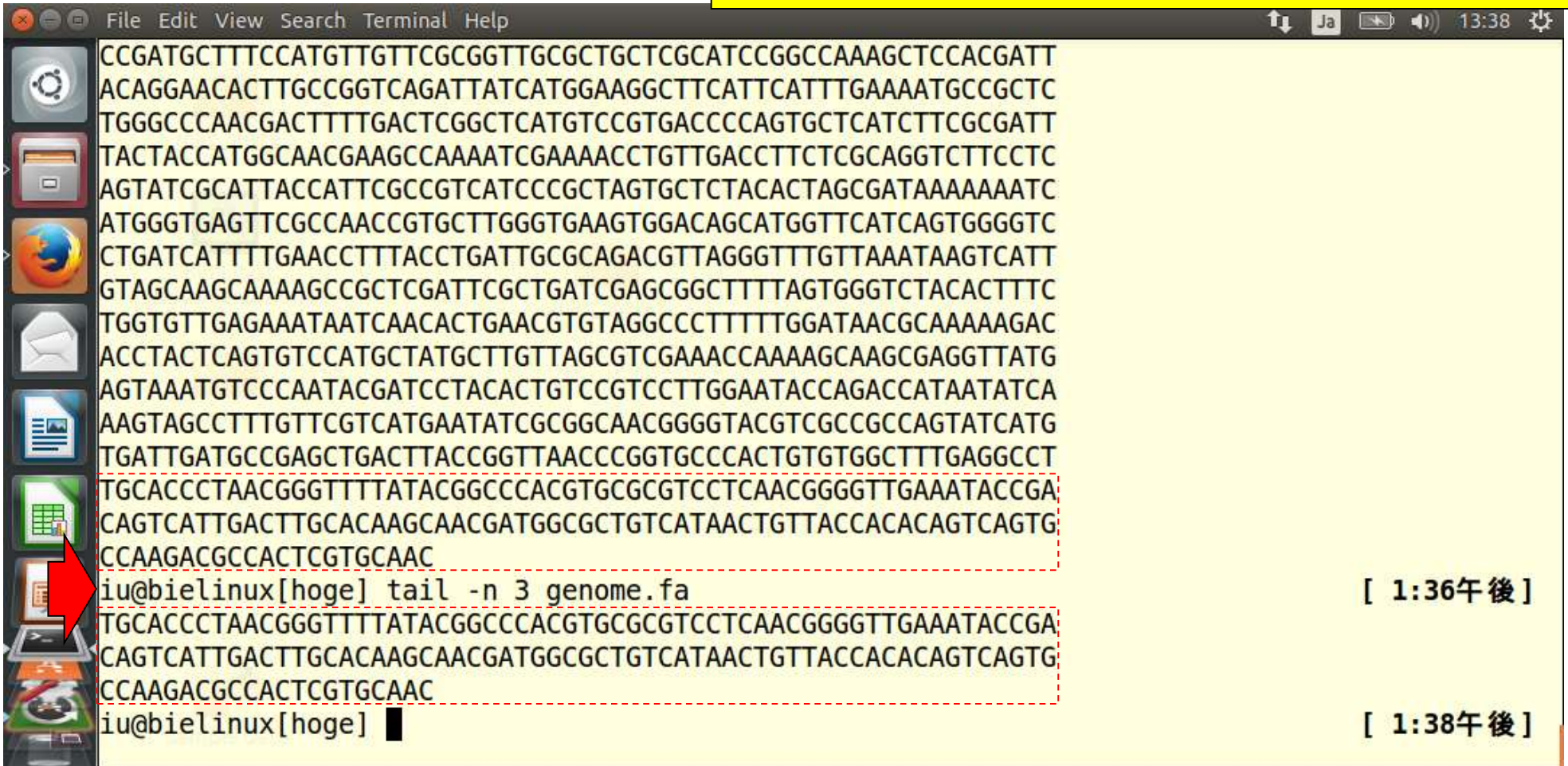
The image shows a terminal window with a yellow background. The terminal output is as follows:

```
iu@bielinux[hoge] pwd [ 1:15午後 ]  
/home/iu/Desktop/hoge  
iu@bielinux[hoge] ls [ 1:33午後 ]  
genome.fa  
iu@bielinux[hoge] wc genome.fa [ 1:33午後 ]  
 48138  48194 2935947 genome.fa  
iu@bielinux[hoge] cat genome.fa [ 1:33午後 ]
```

A red arrow points to the terminal icon in the sidebar, and another red arrow points to the cursor at the end of the `cat genome.fa` command. The terminal window title bar includes 'File Edit View Search Terminal Help', 'Ja', and the time '13:33'.

W14-4

catコマンド実行後。著者の環境では5秒程度で表示が完了する。このファイルは5万行弱なので、概ね1万行につき1秒程度表示に時間がかかっていることがわかる。



```
CCGATGCTTTCCATGTTGTTGCGCGGTTGCGCTGCTCGCATCCGGCCAAAGCTCCACGATT
ACAGGAACACTTGCCGGTCAGATTATCATGGAAGGCTTCATTCATTTGAAAATGCCGCTC
TGGGCCCAACGACTTTTGA CTGCGCTCATGTCCGTGACCCAGTGCTCATCTTCGCGATT
TACTACCATGGCAACGAAGCCAAAATCGAAAACCTGTTGACCTTCTCGCAGGTCTTCCTC
AGTATCGCATTACCATTGCGCGTCATCCCGCTAGTGCTCTACACTAGCGATAAAAAAATC
ATGGGTGAGTTCGCCAACCGTGCTTGGGTGAAGTGGACAGCATGGTTCATCAGTGGGGTC
CTGATCATTTTGAACCTTTACCTGATTGCGCAGACGTTAGGGTTTGTTAAATAAGTCATT
GTAGCAAGCAAAGCCGCTCGATTGCTGATCGAGCGGCTTTTAGTGGGTCTACACTTTC
TGGTGTTGAGAAATAATCAACACTGAACGTGTAGGCCCTTTTTGGATAACGCAAAAAGAC
ACCTACTCAGTGTCCATGCTATGCTTGTAGCGTCGAAACCAAAGCAAGCGAGGTTATG
AGTAAATGTCCCAATACGATCCTACACTGTCCGTCCTTGAATACCAGACCATAATATCA
AAGTAGCCTTTGTTGTCATGAATATCGCGGCAACGGGGTACGTCGCCGCCAGTATCATG
TGATTGATGCCGAGCTGACTTACCGGTTAACCCGGTGCCCACTGTGTGGCTTTGAGGCCT
TGCACCCTAACGGGTTTTATACGGCCCACGTGCGCGTCCTCAACGGGGTTGAAATACCGA
CAGTCATTGACTTGCACAAGCAACGATGGCGCTGTCATAACTGTTACCACACAGTCAGTG
CCAAGACGCCACTCGTGCAAC
iu@bielinux[hoge] tail -n 3 genome.fa [ 1:36午後 ]
TGCACCCTAACGGGTTTTATACGGCCCACGTGCGCGTCCTCAACGGGGTTGAAATACCGA
CAGTCATTGACTTGCACAAGCAACGATGGCGCTGTCATAACTGTTACCACACAGTCAGTG
CCAAGACGCCACTCGTGCAAC
iu@bielinux[hoge] [ 1:38午後 ]
```

W14-4

Tips。catコマンド実行後、全内容表示を途中で止めた
い場合には、「CTRLキー + Cキー」を押せばよい。

```
iu@bielinux[~/Desktop/hoge]
GTTCAAGTCCTCGATAATAGATGAACTCAAATGGCATTGCTTTCCATCGTCTCTTGATT
CAGTCGATTAAATGGCCGCTGAAAAGCGAACACAATCACCAAATAACAAGGCACAGCAAT
TAAAGACACCAAGAATAAGGTGGTGTGTTTGAATGCCTAACACGATCCCTAAAATTAAGAC
AATCCAAATATCAAGAAACAAGGTCATAATCGTGTTTGCCAGCGCATCAATAATCTTATT
AGCATCGGAAAACCGCGACACAATTTCCCCAGTCCGACGCGTGGAAAAGAAGGACATCGG
TAATTCATAAAGGTGTCTGACGTAACCCAAAATGACATCAATCGACAAACGTTGTCCTAA
AACAGCCATCAAAAAATTTGCCCATAGGTTAAAACCGACTGCAGCACATAAGCGAGGAT
GAGACCACTGGCAATCATTCCCAAAGTTGTGCGCATCCCGTTTGAATATAAGTATCAAT
AATCGACTGAAGAAAATAAGACCCGACAATACTAATCAACGTGATCAACACAGCGGCGAC
AATAATATTGGCAATGATTCCTTTTTGTGCGCTAGCATCGGCACAAACGCAAAGAGACT
GCCTTTGTCTCTTTGCGTGGTTGATATTGCGGTTCTGGCGCAATAAAAATAGCGACCCC
CGACCACTCCTTCGCAAAGCGCTCTTTGGGAAGCCAGGTCATCGCTACTGATGAATCAGG
ATCGCCAATCAAGATATGATCCTTTGTGCGCTTAAACACCACATAATAATGAAGCAACTC
GCCACGCTTCAAACATGGACAAGAAACGGATAAGGAACATCATCAAGATCAAACAGCGT
CATATCCGCCTTGACAGCCTTGGTTTCAAATGAAGCGCCTGAGCAGCTCGGACAATCCC
TAACGCTGTTGTTCTTCTGATCCGTCTTCGCCAACTGTGCGCAGATGCGCTAGCGAATA
ATCCGAGCCATTAATTGCAGCAGCATATTAAGCGCCGCAACCCCGCAATCGCGCTCATC
CACCTGAGGCGTGTAGTAACGCTTGTAATACTGTGCTTGGCCGAGCTTTAGGTTGTATTT
GATCATGGGCATGCCTCCTGGTTGGTGTGGATTTTTCTAAATAAGTGATACCTTGAAC
^C%
iu@bielinux[hoge] █
```

[1:45午後]

W14-5

「more genome.fa」実行直後。スペースキーを押すと、次ページに移動。

```
File Edit View Search Terminal Help
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTCGCATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
--More-- (0%)
```

W14-5

スペースキーを1回だけ押して次ページに移動した状態。スペースキーを押し続けると%のところの数値がどんどんあがっていき、100%までいくと最後まで閲覧した状態になって、コマンド入力待ち状態(コマンドプロンプト)となる。途中でやめたい場合は、「qキー」を押すとよい。



```
File Edit View Search Terminal Help
ATGGTTCGCGCAGTTCAAATGCGCAAACCTATTCGCACA
GTGACAGCTTCTGGAGTATTGCCAGCAAGTATGGCATT
ACAACGGAAAATCAATCTACAGCCTGATCTATCCGGGCGAAAGCCTGTATATTAGGTAAC
AACAGCCCTCTGCTCGCTAACGCGGGTGGAGGACTTTTTTAGAACCAAGTAAATTCCT
GTGGATAAGTGGCTTTGTGAAGATTTGCTAGGAACATTGATACAACGGTGTTTTTGTAA
AACCAGATTAGTACCGATTTACGGTTAAAACGTTGTTATATCAGGCATACTATGCCATTT
CACGCTAACTTATCCACAGGGGTTGACATCTAGAATGATATAATATTCATAGAAAGGAGA
GGGAACACATGGATGTCATGAAAATTGCAAATTGTTTTCGCGTCAGGAATTATTCTGATA
TGCAAGTTGATGAGAGTATAGATGAGCTCACTCAGCTAAAGCTGATGAAGCTACTGTATT
TTTCCCAAGGTGTTTTCTCGCCGCATATCATAAGAGATTATTTGATGATCCCATTCTTG
CGTGGAATTATGGTCCGGTGGTCAAACAAGTACATGATAAGTACGTAGGAAAACGTGGAA
TTGTCAATGACAATGTGCCAATTACGAATGCTGACTTGAGCGACTATGAAAATATAAATT
CCGATGAAGAGGCGTCAGCAGTTCTTGAAGCTGTTGATCGTGCCTATGGTTCAATGTCTT
CAAGCGATTTGGTCAACATTACGCACTCGCAAACACCGTGGAAAACGACTCCGCAGAGCG
GCGAAATTAGTACTGATAAAATAGAGAAGTACTTTAAGGATCACGTTGTGACATATGGCT
AAGAAAAAGAAAATTTATATTTCTCAGCCTAGTGTTACAGATCCGGGCGTTTTTTGTCT
AACTCAGTTCGGACAGAGCGGCTCACGTTTAACTTTTCGTTTCTCACAGCTGACAAAAAG
TACTCGATAGCCGGCTTGTCACACGAAGACAGTAAAATCCACAAGAAAATTCTGGAACGA
ATAGAGGCATTGTGAGAGAAGAGATAAAACTCTAGTATTTATTGACGGCAAGGAGTCTGGG
CTGGAGACTTTACCTCAAGAAGAAGTTCATTTTCGATTAATAGCGAGTATATCGATAGT
--More-- (0%)
```

「qキー」を押すと、コマンド入力待ち状態に戻る。

W14-5

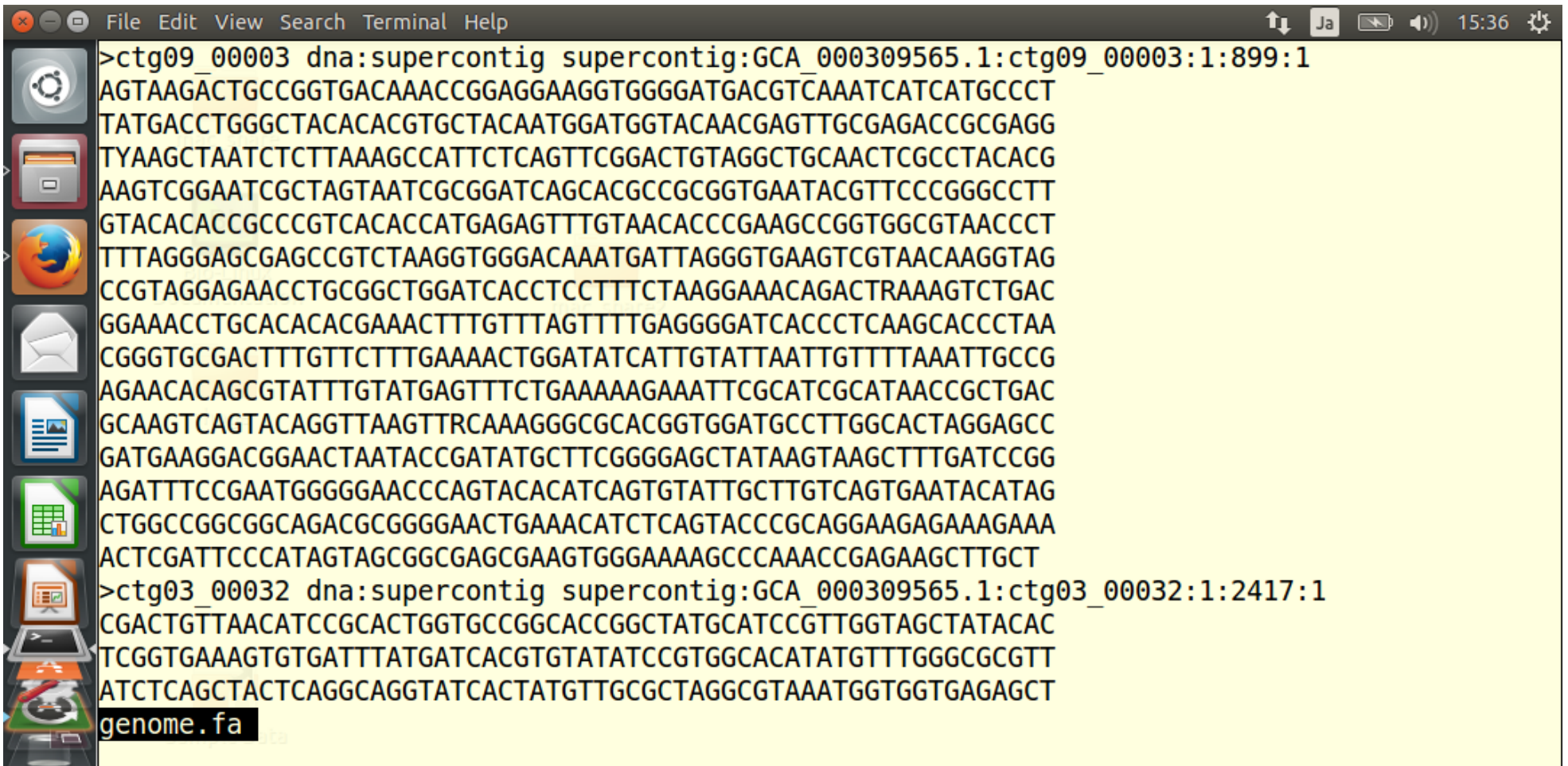


```
ATGGTTCGCGCAGTTCAAATGCGCAAACCTATTCGCACACGTAACACAGTCCGCTCTG
GTGACAGCTTCTGGAGTATTGCCAGCAAGTATGGCATTAGCATGTACACGCTGGCTGCCA
ACAACGGAAAATCAATCTACAGCCTGATCTATCCGGGCGAAAGCCTGTATATTAGGTAAC
AAACAGCCCTCTGCTCGCTAACGCGGGTGGAGGACTTTTTTAGAACCAAGTAAATTCCT
GTGGATAAGTGGCTTTGTGAAGATTTGCTAGGAACATTGATACAACGGTGTGTTTGTAA
AACCAGATTAGTACCGATTTACGGTTAAAACGTTGTTATATCAGGCATACTATGCCATTT
CACGCTAACTTATCCACAGGGGTTGACATCTAGAATGATATAATATTCATAGAAAGGAGA
GGGAACACATGGATGTCATGAAAATTGCAAATTGTTTTCGCGTCAGGAATTATTCTGATA
TGCAAGTTGATGAGAGTATAGATGAGCTCACTCAGCTAAAGCTGATGAAGCTACTGTATT
TTTCCCAAGGTGTTTTCTCGCCGCATATCATAAGAGATTATTTGATGATCCCATTCTTG
CGTGGAATTATGGTCCGGTGGTCAAACAAGTACATGATAAGTACGTAGGAAAACGTGGAA
TTGTCAATGACAATGTGCCAATTACGAATGCTGACTTGAGCGACTATGAAAATATAAATT
CCGATGAAGAGGCGTCAGCAGTTCTTGAAGCTGTTGATCGTGCCTATGGTTCAATGTCTT
CAAGCGATTTGGTCAACATTACGCACTCGCAAACACCGTGGAAAACGACTCCGCAGAGCG
GCGAAATTAGTACTGATAAAATAGAGAAGTACTTTAAGGATCACGTTGTGACATATGGCT
AAGAAAAAGAAAATTTATATTTCTCAGCCTAGTGTTACAGATCCGGGCGTTTTTTGTCT
AACTCAGTTCGGACAGAGCGGCTCACGTTTAACTTTTCGTTTCTCACAGCTGACAAAAAG
TACTCGATAGCCGGCTTGTCACACGAGACAGTAAAATCCACAAGAAAATTCTGGAACGA
ATAGAGGCATTGTGAGAAGAGATAAACTCTAGTATTTATTGACGGCAAGGAGTCTGGG
CTGGAGACTTTACCTCAAGAAGAAGTTCATTTTCGATTAATAGCGAGTATATCGATAGT
iu@bielinux[hoge] █
```

[3:31午後]

「less genome.fa」実行直後。スペースキーを押すと、次ページに移動。

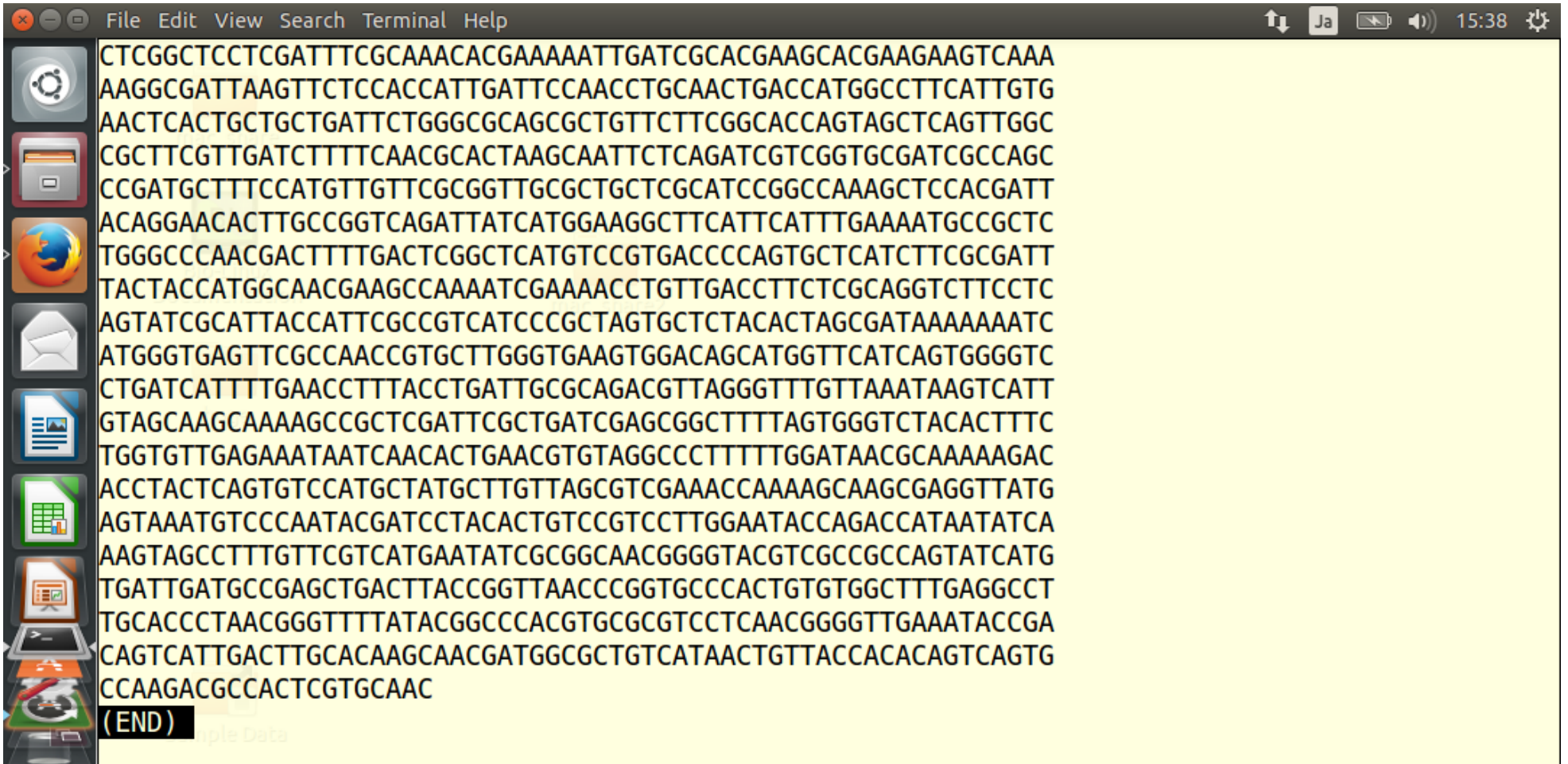
W14-6-1



```
File Edit View Search Terminal Help
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTGCGATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
genome . fa
```

「Gキー」を押して、最終ページに移動した状態。
「gキー」を押すと、最初のページに移動する。

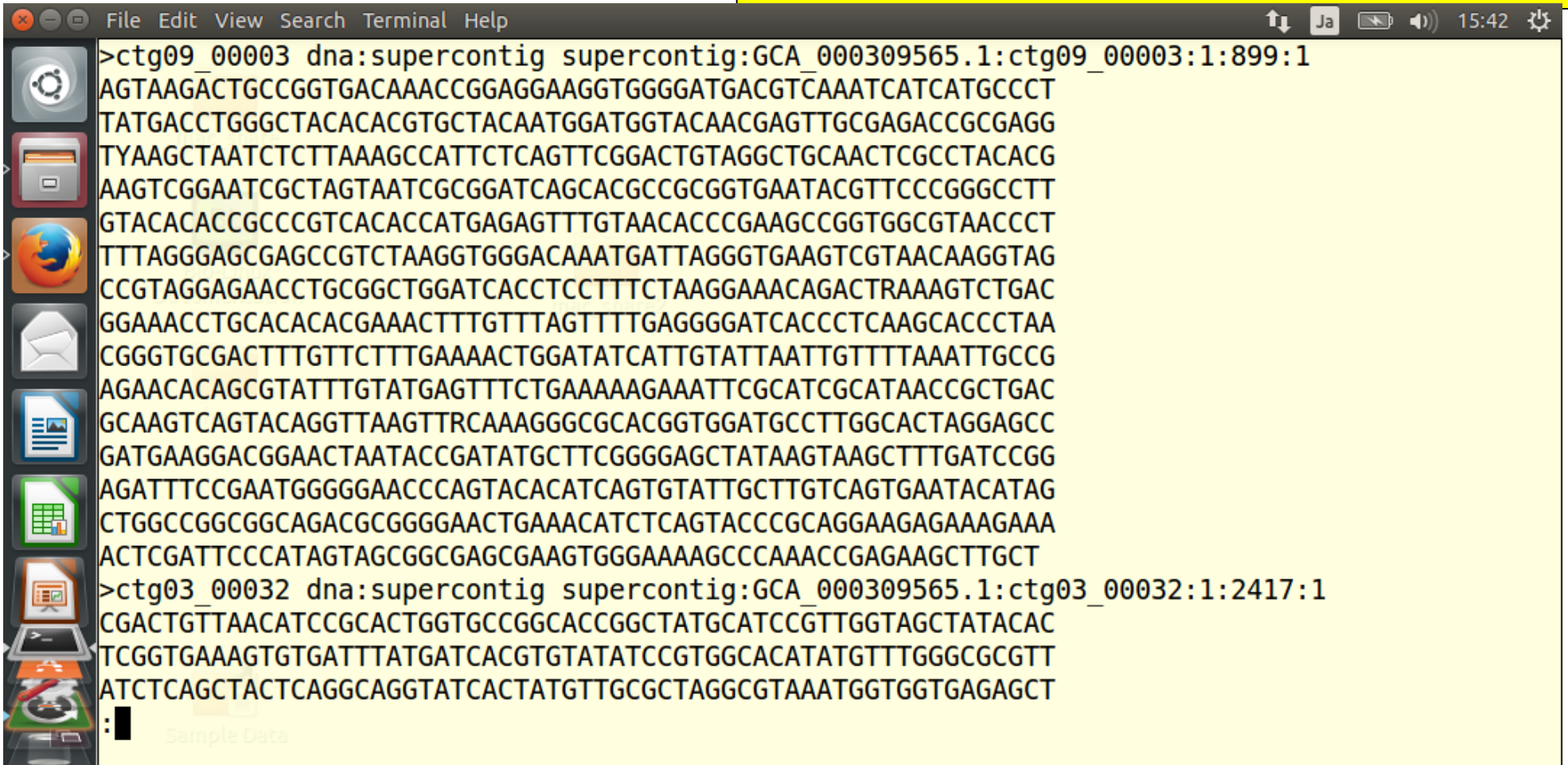
W14-6-1



CTCGGCTCCTCGATTTGCAAACACGAAAAATTGATCGCACGAAGCACGAAGAAGTCAA
AAGGCGATTAAGTTCTCCACCATTGATTCCAACCTGCAACTGACCATGGCCTTCATTGTG
AACTCACTGCTGCTGATTCTGGGCGCAGCGCTGTTCTTCGGCACCAGTAGCTCAGTTGGC
CGTTTCGTTGATCTTTTCAACGCACTAAGCAATTCTCAGATCGTCGGTGCGATCGCCAGC
CCGATGCTTTCCATGTTGTTGCGGGTTGCGCTGCTCGCATCCGGCCAAAGCTCCACGATT
ACAGGAACACTTGCCGGTCAGATTATCATGGAAGGCTTCATTCATTTGAAAATGCCGCTC
TGGGCCCAACGACTTTTGA CTGCGCTCATGTCCGTGACCCAGTGCTCATCTTCGCGATT
TACTACCATGGCAACGAAGCCAAAATCGAAAACCTGTTGACCTTCTCGCAGGTCTTCCTC
AGTATCGCATTACCATTGCGCGTCATCCCGCTAGTGCTCTACACTAGCGATAAAAAAATC
ATGGGTGAGTTCGCCAACCGTGCTTGGGTGAAGTGGACAGCATGGTTCATCAGTGGGGTC
CTGATCATTTTGAACCTTTACCTGATTGCGCAGACGTTAGGGTTTGTAAATAAGTCATT
GTAGCAAGCAAAAGCCGCTCGATTGCTGATCGAGCGGCTTTTAGTGGGTCTACACTTTC
TGGTGTGAGAAATAATCAACACTGAACGTGTAGGCCCTTTTGGATAACGCAAAAAGAC
ACCTACTCAGTGTCCATGCTATGCTTGTAGCGTCGAAACCAAAGCAAGCGAGGTTATG
AGTAAATGTCCCAATACGATCCTACACTGTCCGTCCTTGAATACCAGACCATAATATCA
AAGTAGCCTTTGTTTCGTCATGAATATCGCGGCAACGGGGTACGTCGCCGCCAGTATCATG
TGATTGATGCCGAGCTGACTTACCGGTTAACCCGGTGCCCACTGTGTGGCTTTGAGGCCT
TGCACCCTAACGGGTTTTATACGGCCCACGTGCGCGTCTCAACGGGGTTGAAATACCGA
CAGTCATTGACTTGACACAAGCAACGATGGCGCTGTCATAACTGTTACCACACAGTCAGTG
CCAAGACGCCACTCGTGCAAC
(END)

W14-6-1

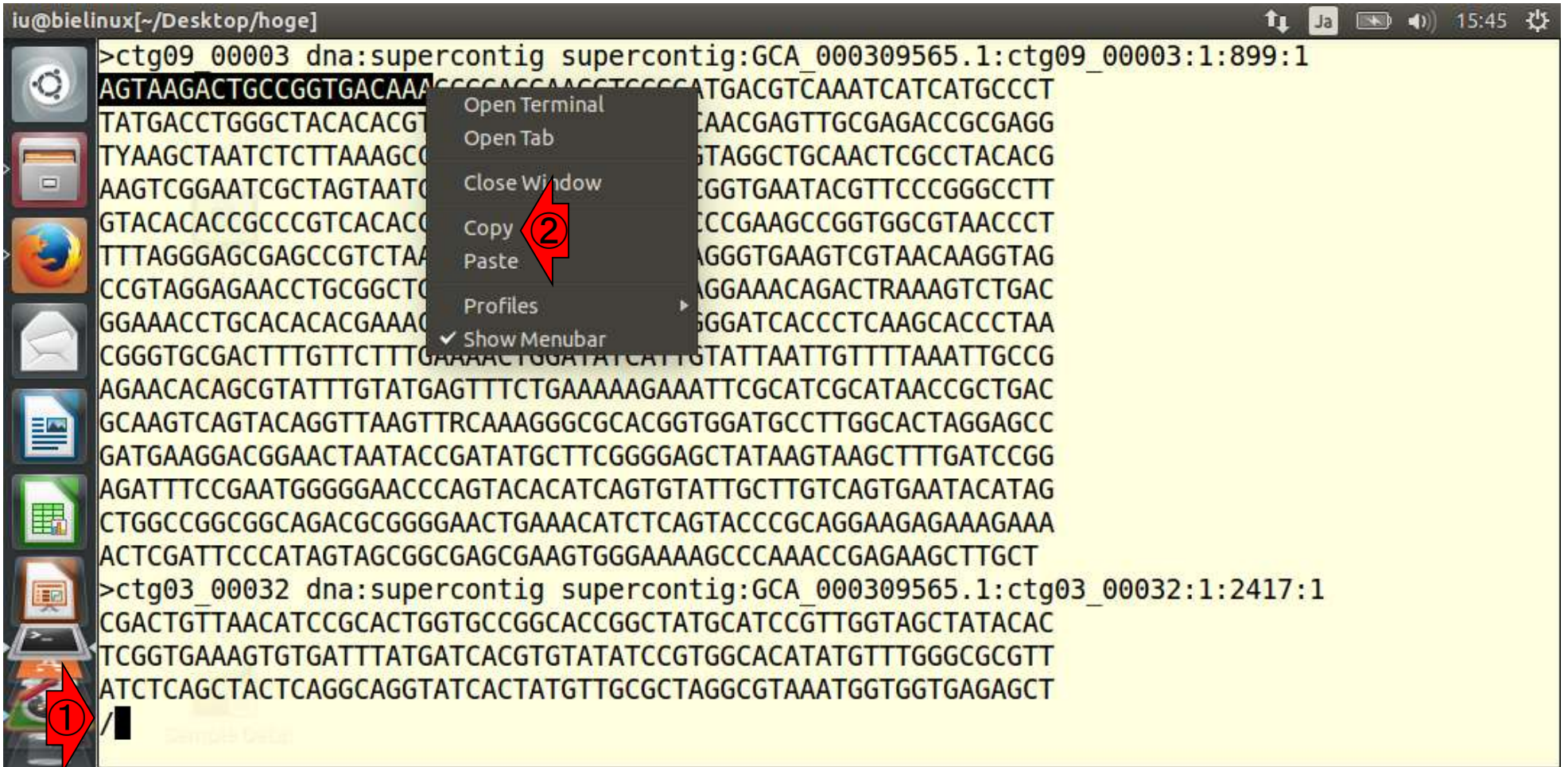
「gキー」を押して、最初のページに移動した状態。
moreコマンドと同様に、「qキー」を押すとコマンド入力
待ち状態に戻る(がここではやらずに次のスライドへ)。



```
File Edit View Search Terminal Help
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTGCGATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
:
Sample Data
```

W14-6-2

文字列検索。①スラッシュ(/)を打ち込んでから、そこにあることがわかっている文字列を②コピー&ペースト。

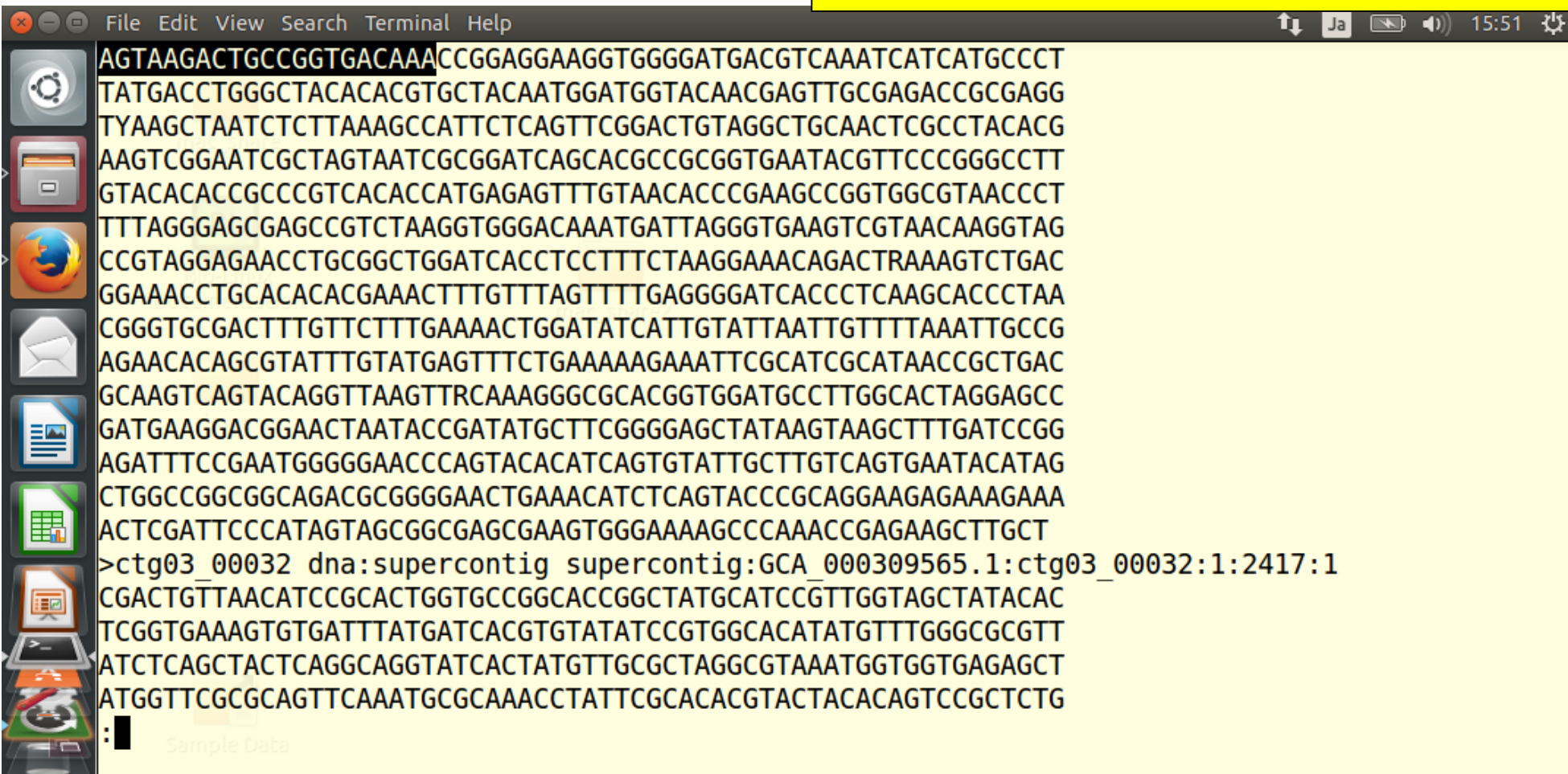


```
iu@bielinux[~/Desktop/hoge]
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACGGGACCAAGCTGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTCAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACGGATCACCCCTCAAGCACCCCTAA
CGGGTGCGACTTTGTTCTTTGAAACCTGGATATCATTGATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTCGCATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAACTAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTCAAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
/
```

The image shows a terminal window with a search for a sequence containing a slash. A context menu is open over the search results, with the 'Copy' option selected. A red arrow points to the 'Copy' option, and another red arrow points to the search input field.

W14-6-2

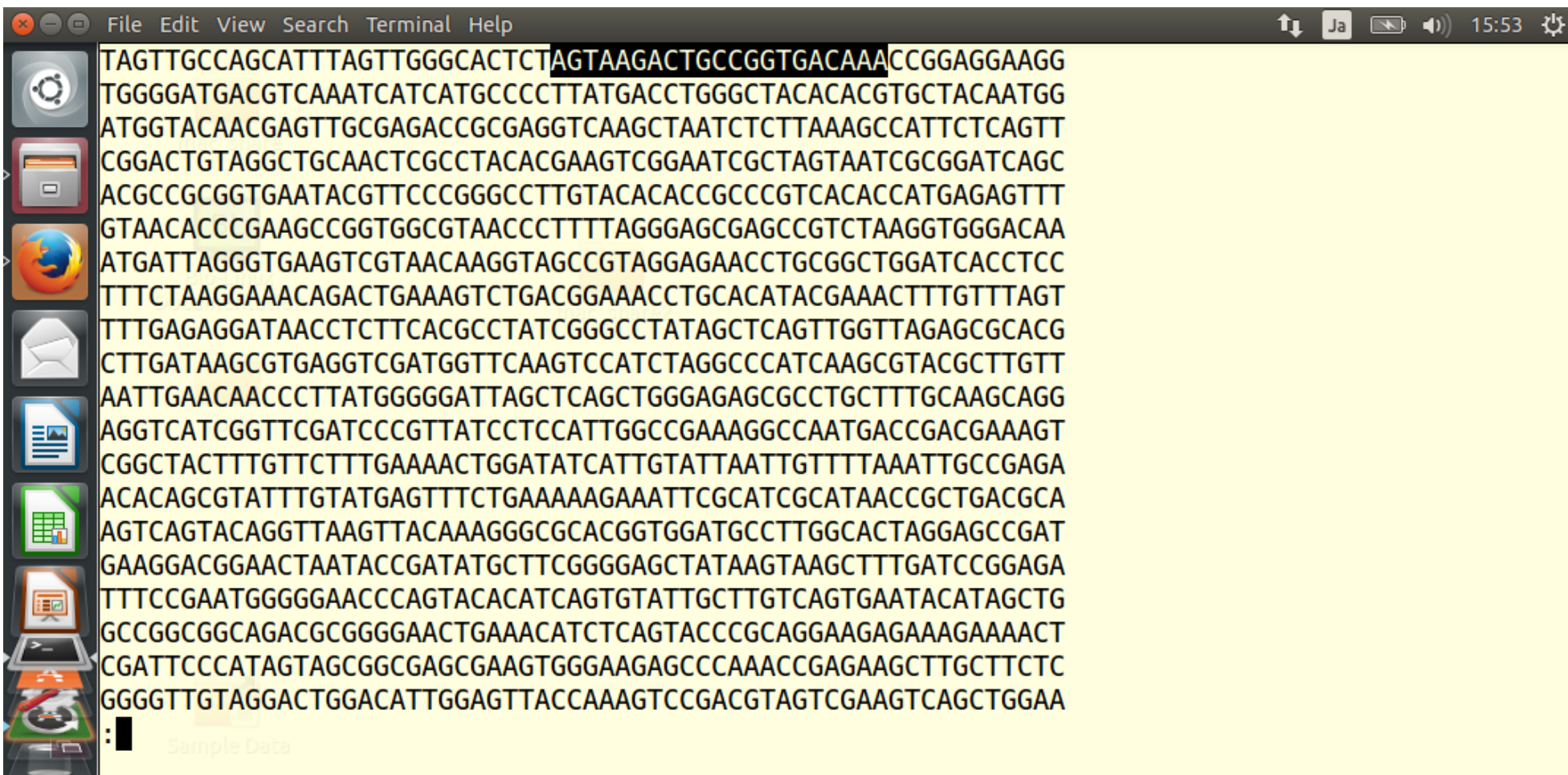
確かにそこが最初にハイライトされるべきところがそうになっていることがわかる。他の箇所にも同じ文字列があれば「nキー」を押すとそこに飛ぶ。



```
AGTAAGACTGCCGGTGACAAA|CCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTCGCATCGCATAACCGCTGAC
GCAAGTCAGTACAGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAACTAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTCAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
ATGGTTCGCGCAGTTCAAATGCGCAAACCTATTCGCACACGTAACACAGTCCGCTCTG
:█
```

Sample Data

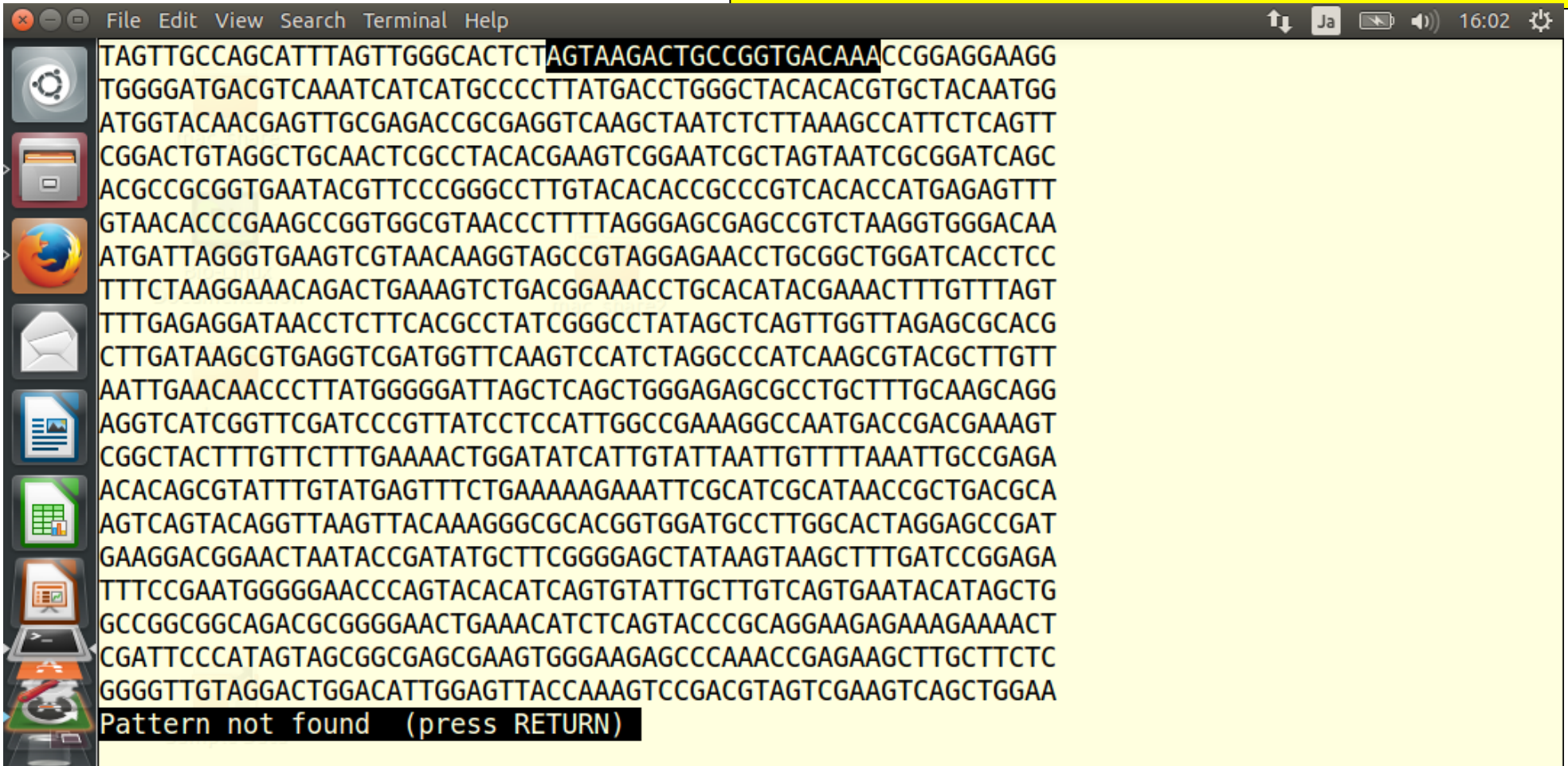
W14-6-2



```
File Edit View Search Terminal Help
TAGTTGCCAGCATTTAGTTGGGCACTCTAGTAAGACTGCCGGTGACAAA CCGGAGGAAGG
TGGGGATGACGTCAAATCATCATGCCCCCTTATGACCTGGGCTACACACGTGCTACAATGG
ATGGTACAACGAGTTGCGAGACCGCGAGGTCAAGCTAATCTCTTAAAGCCATTCTCAGTT
CGGACTGTAGGCTGCAACTCGCCTACACGAAGTCGGAATCGCTAGTAATCGCGGATCAGC
ACGCCGCGGTGAATACGTTCCCGGGCCTGTACACACCGCCCGTCACACCATGAGAGTTT
GTAACACCCGAAGCCGGTGGCGTAACCCTTTAGGGAGCGAGCCGTCTAAGGTGGGACAA
ATGATTAGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCTGGATCACCTCC
TTTCTAAGGAAACAGACTGAAAGTCTGACGGAAACCTGCACATACGAAACTTTGTTTAGT
TTTGAGAGGATAACCTCTTCACGCCTATCGGGCCTATAGCTCAGTTGGTTAGAGCGCACG
CTTGATAAGCGTGAGGTCGATGGTTC AAGTCCATCTAGGCCATCAAGCGTACGCTTGTT
AATTGAACAACCCTTATGGGGGATTAGCTCAGCTGGGAGAGCGCCTGCTTTGCAAGCAGG
AGGTCATCGGTTGATCCCGTTATCCTCCATTGGCCGAAAGGCCAATGACCGACGAAAGT
CGGCTACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCGAGA
ACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTGCGATCGCATAACCGCTGACGCA
AGTCAGTACAGGTTAAGTTACAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCCGAT
GAAGGACGGAAC TAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGGAGA
TTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTCAGTGAATACATAGCTG
GCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAAAC
CGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAGAGCCCAAACCGAGAAGCTTGCTTCTC
GGGTTGTAGGACTGGACATTGGAGTTACCAAAGTCCGACGTAGTCGAAGTCAGCTGGAA
:
Sample Data
```

W14-6-2

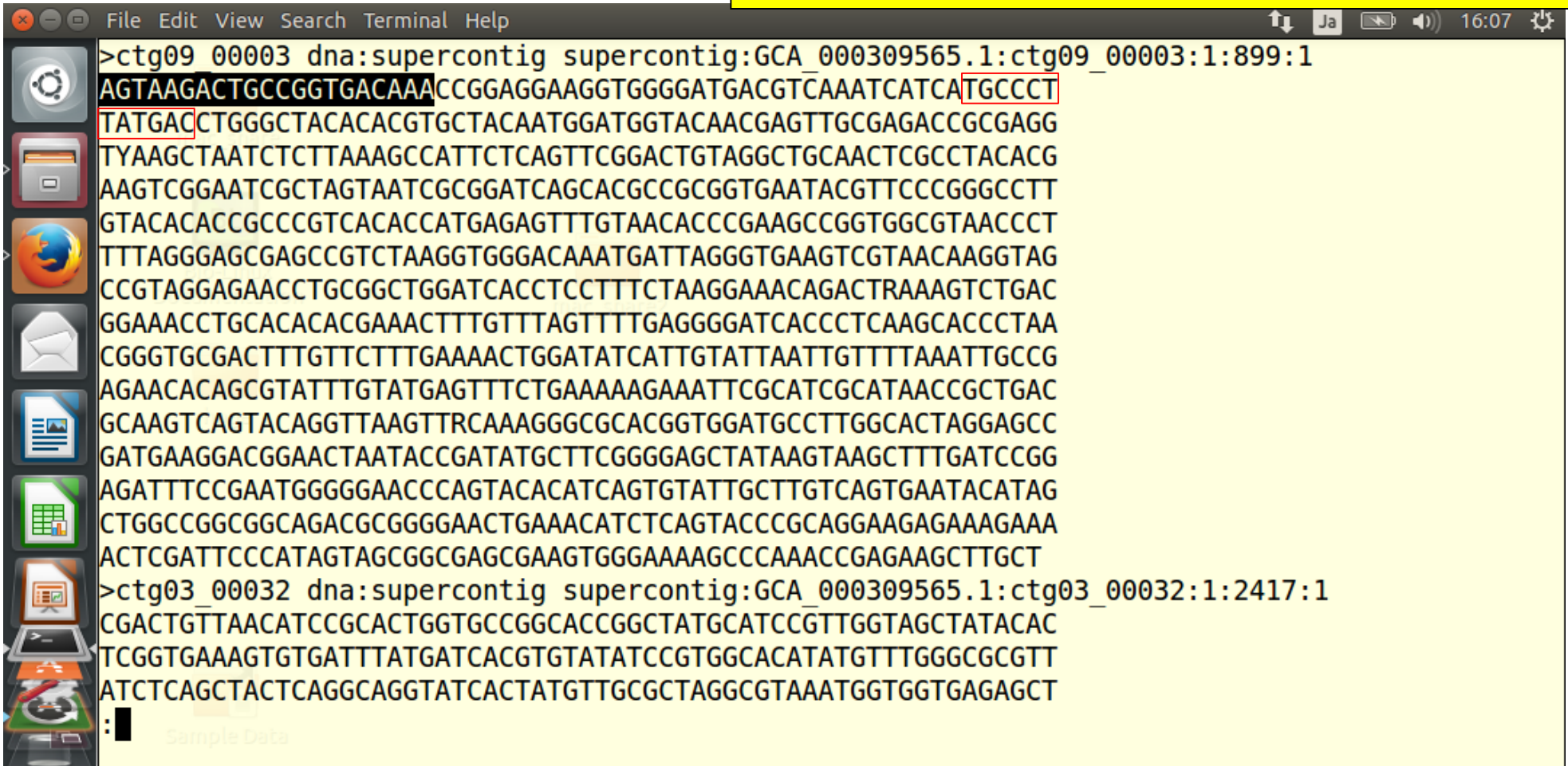
もう1回「nキー」を押すと、もう一致箇所はないというメッセージが出る。いわれるがままにリターンキーを押すと、lessコマンドの中の命令実行待ち状態(:)になる。



```
File Edit View Search Terminal Help
TAGTTGCCAGCATTTAGTTGGGCACTCTAGTAAGACTGCCGGTGACAAAACCGGAGGAAGG
TGGGGATGACGTCAAATCATCATGCCCCCTTATGACCTGGGCTACACACGTGCTACAATGG
ATGGTACAACGAGTTGCGAGACCGCGAGGTCAAGCTAATCTCTTAAAGCCATTCTCAGTT
CGGACTGTAGGCTGCAACTCGCCTACACGAAGTCGGAATCGCTAGTAATCGCGGATCAGC
ACGCCGCGGTGAATACGTTCCCGGGCCTGTACACACCGCCCGTCACACCATGAGAGTTT
GTAACACCCGAAGCCGGTGGCGTAACCCTTTTAGGGAGCGAGCCGTCTAAGGTGGGACAA
ATGATTAGGGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCTGGATCACCTCC
TTTCTAAGGAAACAGACTGAAAGTCTGACGGAAACCTGCACATACGAAACTTTGTTTAGT
TTTGAGAGGATAACCTCTTCACGCCTATCGGGCCTATAGCTCAGTTGGTTAGAGCGCACG
CTTGATAAGCGTGAGGTCGATGGTTC AAGTCCATCTAGGCCATCAAGCGTACGCTTGT
AATTGAACAACCCTTATGGGGGATTAGCTCAGCTGGGAGAGCGCCTGCTTTGCAAGCAGG
AGGTCATCGGTTTCGATCCCGTTATCCTCCATTGGCCGAAAGGCCAATGACCGACGAAAGT
CGGCTACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCGAGA
ACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTCGCATCGCATAACCGCTGACGCA
AGTCAGTACAGGTTAAGTTACAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCCGAT
GAAGGACGGAAC TAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGGAGA
TTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGTCAGTGAATACATAGCTG
GCCGGCGGCAGACGCGGGGAAC TGAACATCTCAGTACCCGCAGGAAGAGAAAAGAAAAC
CGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAGAGCCCAAACCGAGAAGCTTGCTTCTC
GGGTTGTAGGACTGGACATTGGAGTTACCAAAGTCCGACGTAGTCGAAGTCAGCTGGAA
Pattern not found (press RETURN)
```

W14-6-3

lessコマンドの中の命令実行待ち状態(:)で、「gキー」を押して先頭ページに移動した状態。次は、行をまたがる文字列検索がうまくいかない例を紹介。



```
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAA_CCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGAC CTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTGCGATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCGAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
:
Sample Data
```

W14-6-3

```
File Edit View Search Terminal Help 16:20
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAA CCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGAC CTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTGCGATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
/TGCCCTTATGAC
```

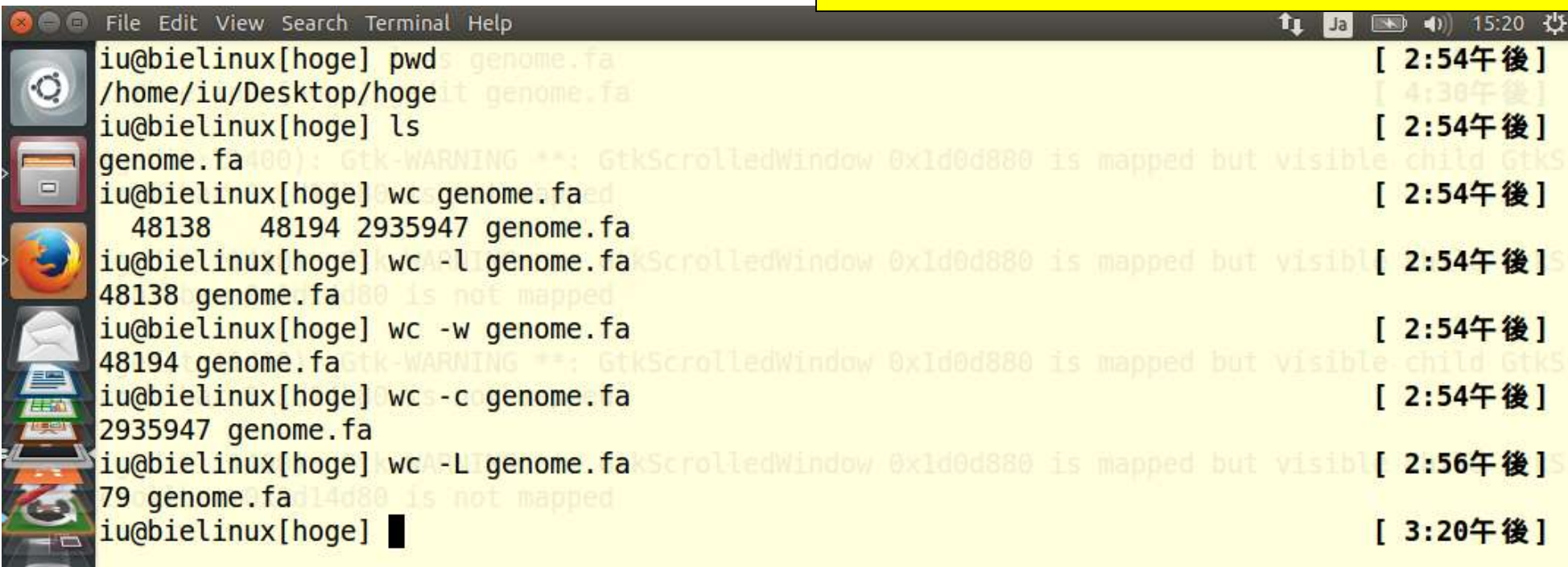
W14-6-3

確かに行をまたがる文字列検索がうまくいかないことがわかります。「リターンキー」を押した後に、「qキー」を押すと、コマンド入力待ち状態に戻る。

```
File Edit View Search Terminal Help
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTGCGATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAATAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCGAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGGCGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTGGGCGCGTT
ATCTCAGCTACTCAGGCAGGTATCACTATGTTGCGCTAGGCGTAAATGGTGGTGAGAGCT
Pattern not found (press RETURN)
```


W15-1

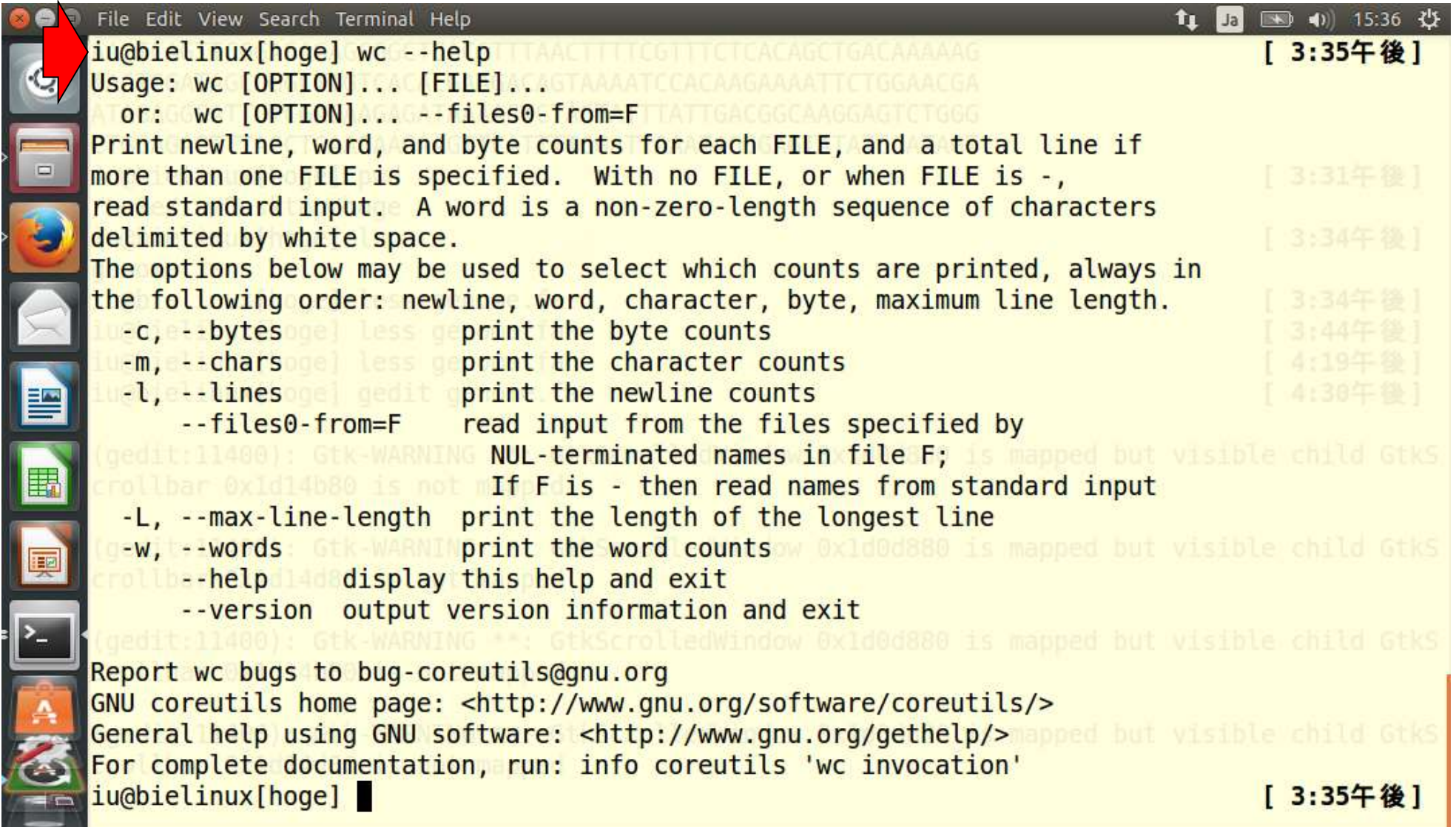
オプションを駆使することで様々な結果を得ることができる。「-L」オプションは、ファイル中で最も長い行の(スペースなどを含む)文字数を返す。



```
iu@bielinux[hoge] pwd /home/iu/Desktop/hoge [ 2:54午後 ]
iu@bielinux[hoge] ls genome.fa [ 2:54午後 ]
iu@bielinux[hoge] wc genome.fa [ 2:54午後 ]
48138 48194 2935947 genome.fa
iu@bielinux[hoge] wc -l genome.fa [ 2:54午後 ]
48138 genome.fa
iu@bielinux[hoge] wc -w genome.fa [ 2:54午後 ]
48194 genome.fa
iu@bielinux[hoge] wc -c genome.fa [ 2:54午後 ]
2935947 genome.fa
iu@bielinux[hoge] wc -L genome.fa [ 2:56午後 ]
79 genome.fa
iu@bielinux[hoge] [ 3:20午後 ]
```

W15-2

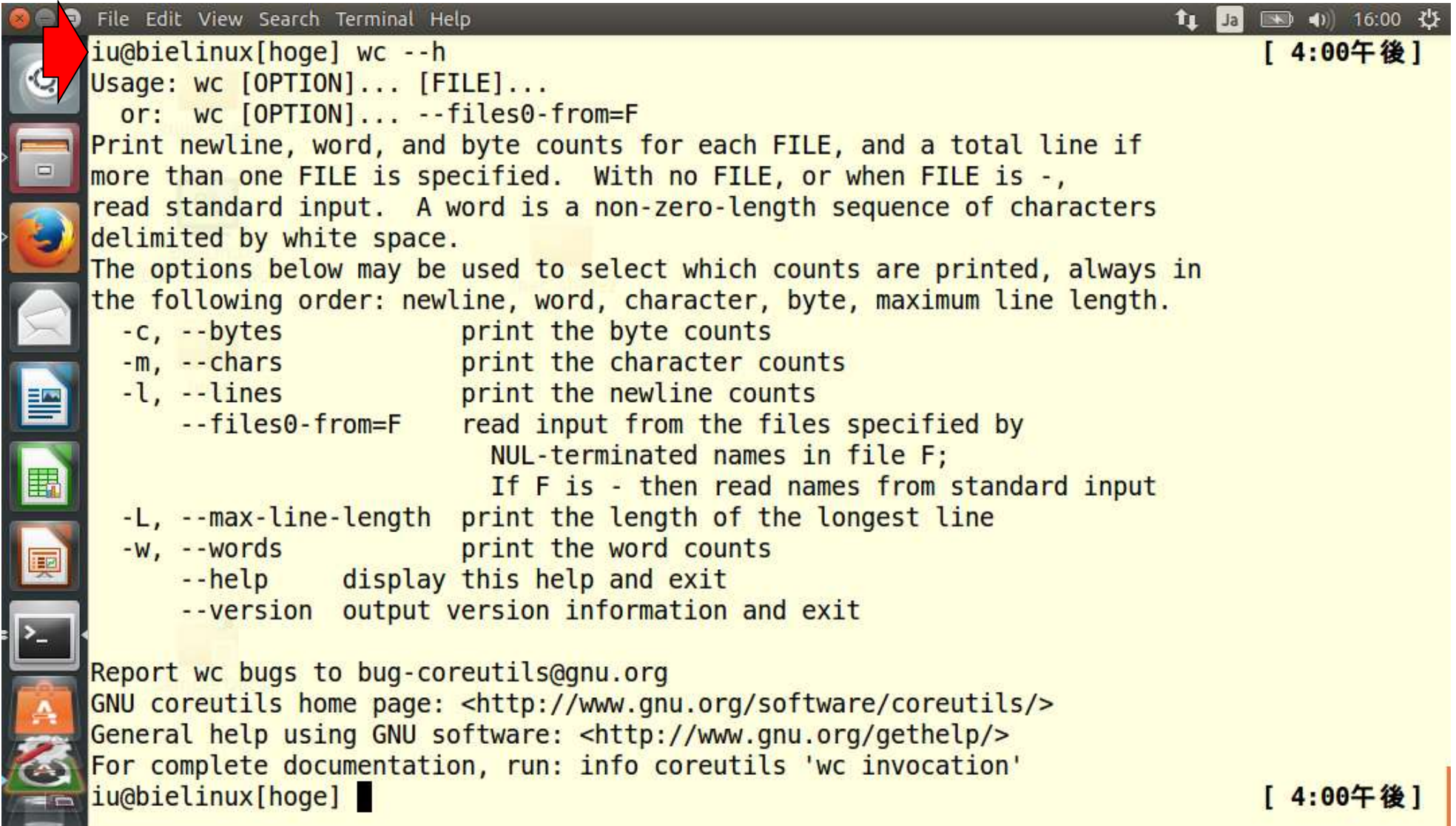
「wc --help」で、利用可能なオプションやその説明を見ることができる。



```
iu@bielinux[hoge] wc --help
Usage: wc [OPTION]... [FILE]...
  or: wc [OPTION]... --files0-from=F
Print newline, word, and byte counts for each FILE, and a total line if
more than one FILE is specified.  With no FILE, or when FILE is -,
read standard input.  A word is a non-zero-length sequence of characters
delimited by white space.
The options below may be used to select which counts are printed, always in
the following order: newline, word, character, byte, maximum line length.
  -c, --bytes          print the byte counts
  -m, --chars          print the character counts
  -l, --lines          print the newline counts
      --files0-from=F  read input from the files specified by
                        F; if F is - then read names from standard input
  -L, --max-line-length print the length of the longest line
  -w, --words          print the word counts
      --help          display this help and exit
      --version       output version information and exit

Report wc bugs to bug-coreutils@gnu.org
GNU coreutils home page: <http://www.gnu.org/software/coreutils/>
General help using GNU software: <http://www.gnu.org/gethelp/>
For complete documentation, run: info coreutils 'wc invocation'
iu@bielinux[hoge]
```

W15-2



```
iu@bielinux[hoge] wc --h
Usage: wc [OPTION]... [FILE]...
  or: wc [OPTION]... --files0-from=F
Print newline, word, and byte counts for each FILE, and a total line if
more than one FILE is specified.  With no FILE, or when FILE is -,
read standard input.  A word is a non-zero-length sequence of characters
delimited by white space.
The options below may be used to select which counts are printed, always in
the following order: newline, word, character, byte, maximum line length.
  -c, --bytes          print the byte counts
  -m, --chars          print the character counts
  -l, --lines          print the newline counts
  --files0-from=F      read input from the files specified by
                        NUL-terminated names in file F;
                        If F is - then read names from standard input
  -L, --max-line-length print the length of the longest line
  -w, --words          print the word counts
  --help              display this help and exit
  --version            output version information and exit

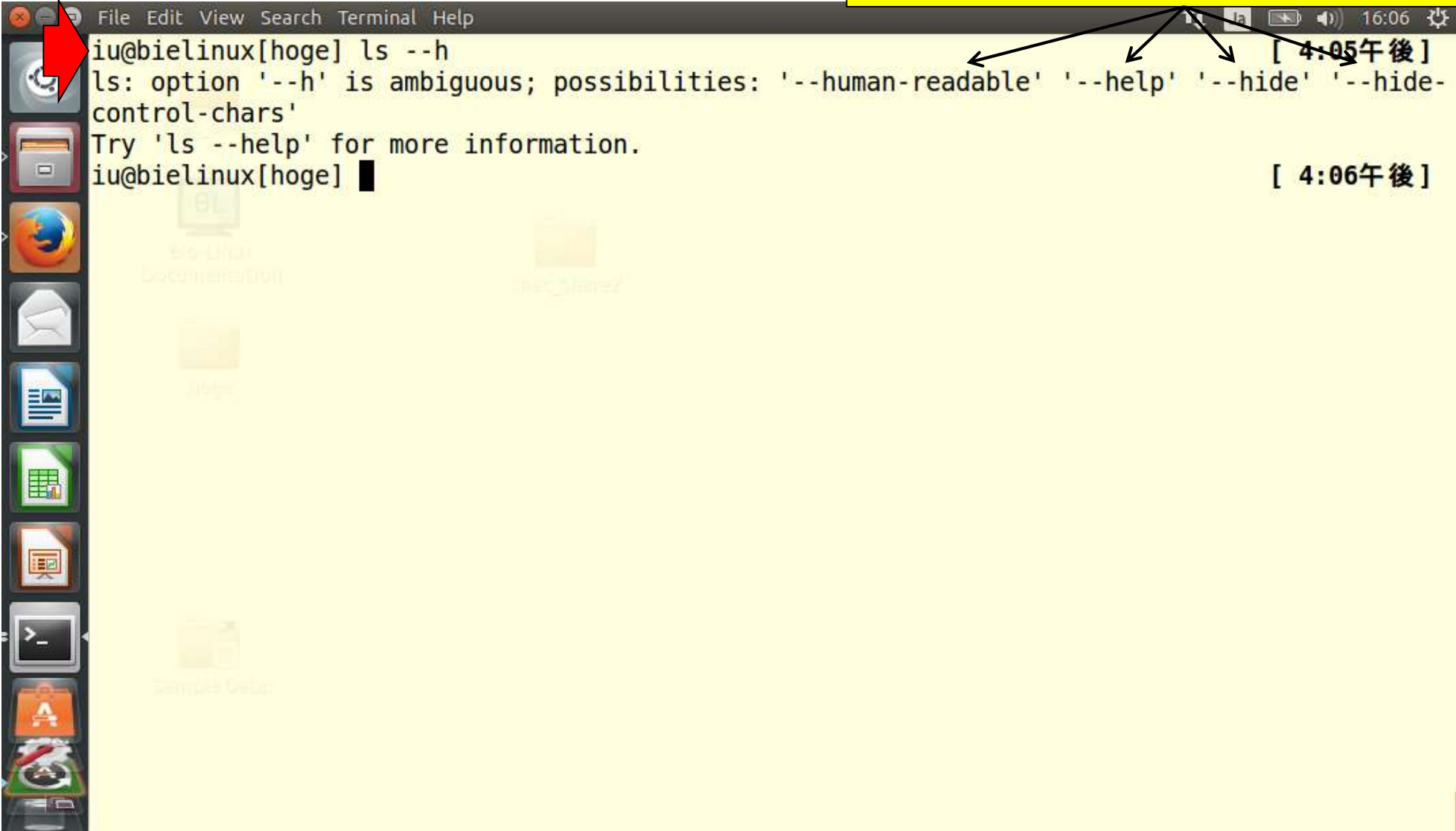
Report wc bugs to bug-coreutils@gnu.org
GNU coreutils home page: <http://www.gnu.org/software/coreutils/>
General help using GNU software: <http://www.gnu.org/gethelp/>
For complete documentation, run: info coreutils 'wc invocation'
iu@bielinux[hoge] █
```

[4:00午後]

[4:00午後]

W15-2

lsコマンドは「--h」はだめ。理由は「--h」から始まる複数のオプションが存在するので、--hだけだと曖昧(ambiguous)だと叱られる。



```
iu@bielinux[hoge] ls --h
ls: option '--h' is ambiguous; possibilities: '--human-readable' '--help' '--hide' '--hide-control-chars'
Try 'ls --help' for more information.
iu@bielinux[hoge] █
```

[4:05午後]

[4:06午後]

W15-3

「man wc」実行直後の状態。詳細なマニュアルを見ることができる。使い方はmoreやlessコマンドと同じ。マニュアル表示画面からの脱出は「qキー」。

```
WC(1) User Commands WC(1)
NAME
wc - print newline, word, and byte counts for each file
SYNOPSIS
wc [OPTION]... [FILE]...
wc [OPTION]... --files0-from=F
DESCRIPTION
Print newline, word, and byte counts for each FILE, and a total line if more than
one FILE is specified. With no FILE, or when FILE is -, read standard input. A
word is a non-zero-length sequence of characters delimited by white space. The
options below may be used to select which counts are printed, always in the fol-
lowing order: newline, word, character, byte, maximum line length.

-c, --bytes
    print the byte counts

-m, --chars
    print the character counts

-l, --lines
    print the newline counts

Manual page wc(1) line 1 (press h for help or q to quit)
```



```
C:\Users\kadota\Desktop\hoge>
```

図5. コマンドプロンプトでの作業例。① cd コマンドを用いた作業ディレクトリの変更、② rename コマンドを用いたファイル名の変更、③ find コマンドを用いた「>」を含む行数のカウント。

ノ酸配列)を1文字表記で記述したものである。FASTQ形式と同様、NGS解析分野においても比較的良好に用いられるファイル形式である。grepは、文字列検索コマンドである。FASTA形式ファイルであるgenome.faを読み込んで、「>」から始まるdescription行を抽出し、出力を行数表示にする-cオプションを与えているため、結果が28(行)と返される。行数表示の-cオプションをつけなかった場合、つまり「grep ">" genome.fa」を実行すると、description情報がそのまま表示される[ウェブ資料21]。grepは応用範囲の極めて広い非常に強力なコマンドである。例えば、アノテーションファイルから特定の染色体名を含む行のみを抽出したり、grep出力結果をパイプでつないで他のコマンドで処理する作業が実際に行われる。ま

findコマンドは存在し、(文字列検索ではなく)ファイルやディレクトリの検索機能をもつ。NGS解析は、Winのコマンドプロンプト環境ではなくLinux環境で行うのが一般的である。バイオインフォ業界では、文字列検索することを「grepする」といい、通常「findする」はファイルやディレクトリ検索のことを指す。つまり、一般にコマンド名で会話が成立するのはLinuxコマンドのみである。

Bio-Linuxの導入

連載第2回は、WinのコマンドプロンプトおよびMacのターミナルという通常利用PC環境を用いて、GUI環境とコマンドライン環境の見栄えの違い、およびWinと

http://www.iu.a.u-tokyo.ac.jp/~kadota/JSLAB_2_kadota.pdf

W16-1

PDFファイル中のコマンドをコピーして実行しようとしてもうまくいかない場合がある。これは連載第2回のPDFファイル。反転部分をコピー。



図5. コマンドプロンプトでの作業例。① cd コマンドを用いた作業ディレクトリの変更、② rename コマンドを用いたファイル名の変更、③ find コマンドを用いた">"を含む行数のカウント。

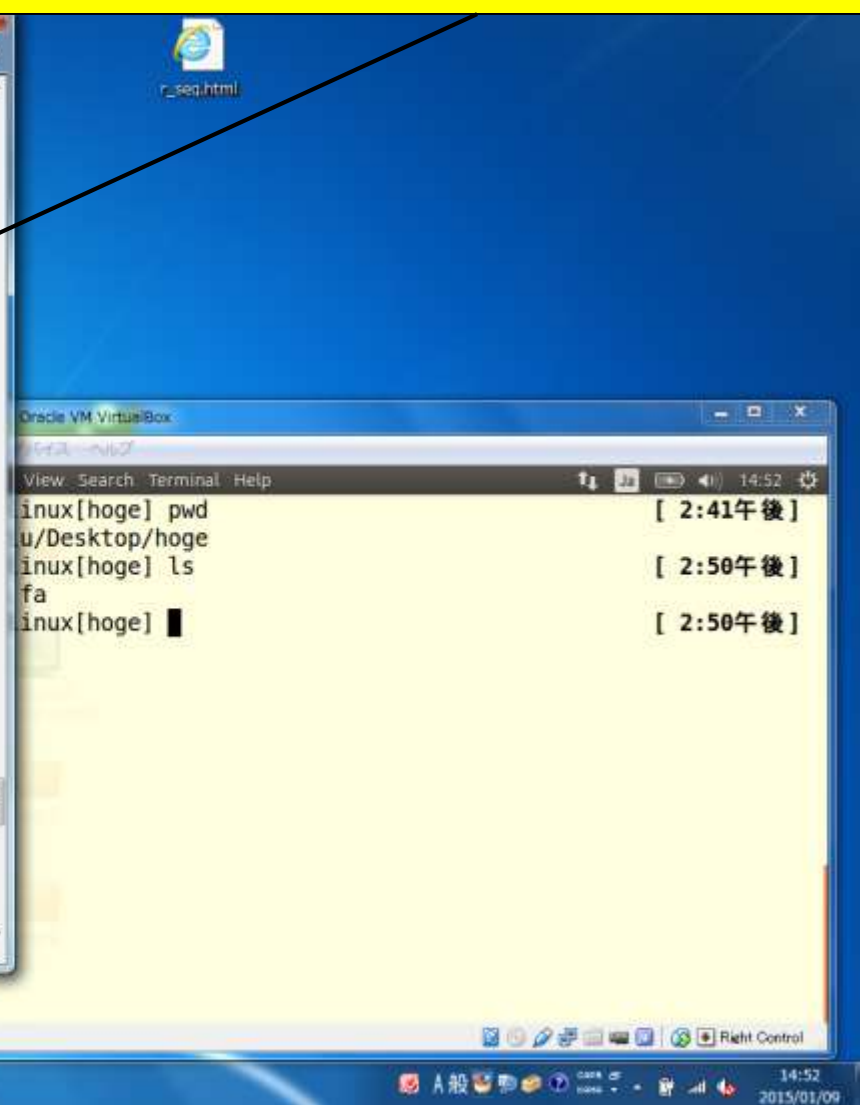
れるファイル形式である。grep は、文字列検索コマンドである。FASTA 形式ファイルである genome.fa を読み込んで、">" から始まる description 行を抽出し、出力を行数表示にする -c オプションを与えているため、結果が 28 (行) と返される。行数表示の -c オプションをつけなかった場合、つまり `grep ">" genome.fa` を実行すると、description 情報がそのまま表示される。

grep は応用範囲の極めて広い非常に強力なコマンドである。例えば、アノテーションファイルから特定の染色体名を含む行のみを抽出したり、grep 出力結果をパイプでつないで他のコマンドで処理する作業が実際に行われる。また最近では、RNA-seq データからのキメラ転写物の同定に用いられるなど、grep コマンドの NGS 解析での有効性も報告されている⁷⁾。もちろんここで示した程度のことであれば、Win のコマンドプロンプト上でも実現可能である(図5: ウェブ資料 22-24)。例えば、Linux の grep に相当する Win のコマンドは find である。しかし、Linux にも find コマンドは存在し、(文字列検索ではなく) ファイルやディレクトリの検索機能をもつ。NGS 解析は、Win のコマンドプロンプト環境ではなく Linux 環境で行うのが一般的である。バイオインフォ業界では、文字列検索する

ことを「grep する」といい、通常「find する」はファイルやディレクトリ検索のことを指す。つまり、一般にコマンド名で会話が成立するのは Linux コマンドのみである。

Bio-Linux の導入

連載第2回は、Win のコマンドプロンプトおよび Mac のターミナルという通常利用 PC 環境を用いて、GUI 環境とコマンドライン環境の見栄えの違い、および Win と Linux の基本コマンドを紹介した。Mac は Linux コマンドをすぐに使えるため、Linux 入門としてはおすすめである。しかし連載第1回でも述べたように、仮想ソフトをインストールして、様々な NGS 解析用プログラムが一通り組み込まれた Bio-Linux⁸⁾ を導入すれば、Win と Mac というホスト OS の違いによらず同じ Linux 環境で解析が可能である。2014年7月末にリリースされた Bio-Linux 8 が最新版である(2014年10月15日調べ)。連載第3回は、Win 用と Mac 用それぞれのインストール手順を掲載しているので、是非自力での Linux 環境構築にチャレンジしてほしい。



W16-1

①右クリックでペースト。手打ち
はうまくいってしまうからダメ！

```
C:\Users\kadota\Desktop\hoge>rename Lactobacillus_casei_12a_GCA_000309565.1.22.d
na_forlevel.fa genome.fa
C:\Users\kadota\Desktop\hoge>find /c >" genome.fa
----- GENME.FA: 28
C:\Users\kadota\Desktop\hoge>
```

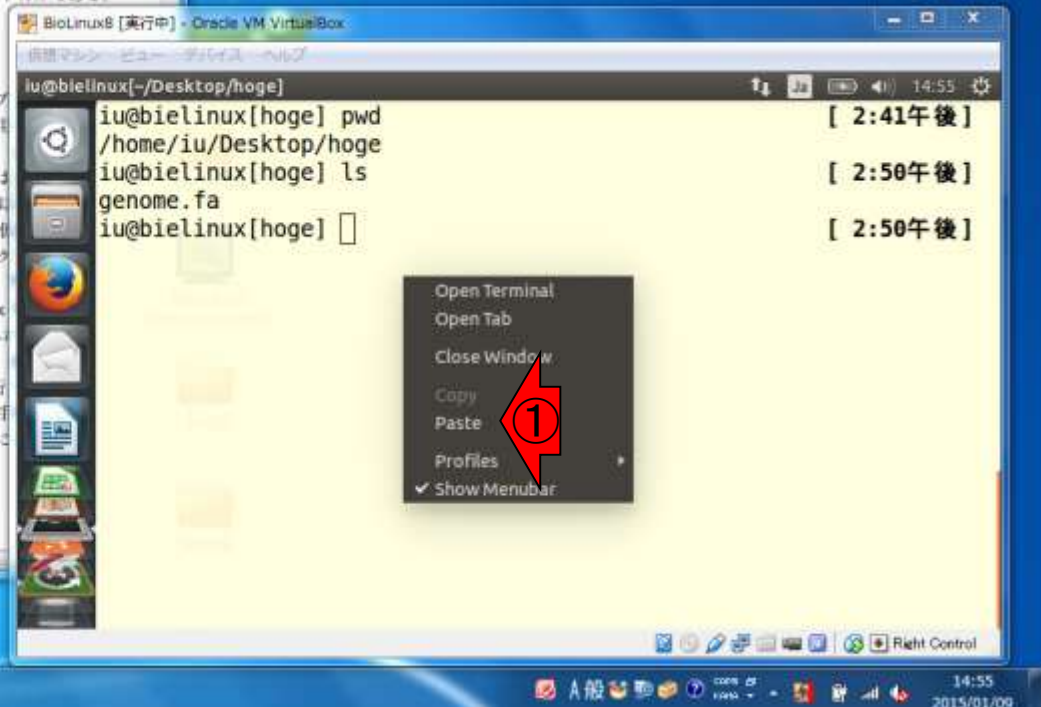
図5. コマンドプロンプトでの作業例。① cd コマンドを用いた作業ディレクトリの変更。② rename コマンドを用いたファイル名の変更。③ find コマンドを用いた">"を含む行数のカウント。

れるファイル形式である。grep は、文字列検索コマンドである。FASTA 形式ファイルである genome.fa を読み込んで、">" から始まる description 行を抽出し、出力を行数表示にする -c オプションを与えているため、結果が 28 (行) と返される。行数表示の -c オプションをつけなかった場合、つまり `grep >" genome.fa` を実行すると、description 情報がそのまま表示される [ウェブ資料 21]。grep は応用範囲の極めて広い非常に強力なコマンドである。例えば、アノテーションファイルから特定の染色体名を含む行のみを抽出したり、grep 出力結果をパイプでつないで他のコマンドで処理する作業が実際に行われる。また最近では、RNA-seq データからのキメラ転写物の同定に用いられるなど、grep コマンドの NGS 解析での有効性も報告されている²⁾。もちろんここで示した程度のことであれば、Win のコマンドプロンプト上でも実現可能である (図 5: ウェブ資料 22-24)。例えば、Linux の grep に相当する Win のコマンドは find である。しかし、Linux にも find コマンドは存在し、(文字列検索ではなく) ファイルやディレクトリの検索機能をもつ。NGS 解析は、Win のコマンドプロンプト環境ではなく Linux 環境で行うのが一般的である。バイオインフォ業界では、文字列検索する

ことを「grep する」といい、通常「find する」はファイルやディレクトリ検索のことを指す。つまり、一般にコマンド名で会話が成立するのは Linux コマンドのみである。

Bio-Linux の導入

連載第 2 回は、Win のコマンドプロンプトのターミナルという通常利用 PC 環境とコマンドライン環境の見栄えの違い、Linux の基本コマンドを紹介した。Mac は下ですぐに使えるため、Linux 入門としては、インストールして、様々な NGS 解析用プログラムが組み込まれた Bio-Linux[®] を導入すれば、いうホスト OS の違いによらず同じ Linux 可能である。2014 年 7 月末にリリースされた最新版である (2014 年 10 月 15 日調べ)。Bio-Linux 8 (ゲスト OS) 環境での解説を行う Win 用と Mac 用それぞれのインストール手順があるので、是非自力での Linux 環境構築してほしい。



W16-1

①ペーストしても、「第2回のW21」と同じ結果にならず、何も表示されない。また、「”」というファイルが作成されてコマンドライン上での削除は初心者は難しい。②ターミナル画面上で通常のダブルクォーテーションで消そうとしても「dquote>」という変なプロンプトが出てしまう。リターンキーを押しても、「qキー」を押してもだめだが、「CTRLキー + Cキー」で復旧できる。どうにかして生成されたファイルと同じ記号のものを入力すれば、一応コマンドライン上で消すことはできる。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls
genome.fa
iu@bielinux[hoge] grep ">" genome.fa
iu@bielinux[hoge] ls
" genome.fa
iu@bielinux[hoge] rm "
dquote>
dquote> q
dquote>
iu@bielinux[hoge] rm "
rm: remove regular empty file ""'? y
iu@bielinux[hoge] ls
genome.fa
iu@bielinux[hoge] █
```

[9:20午後]

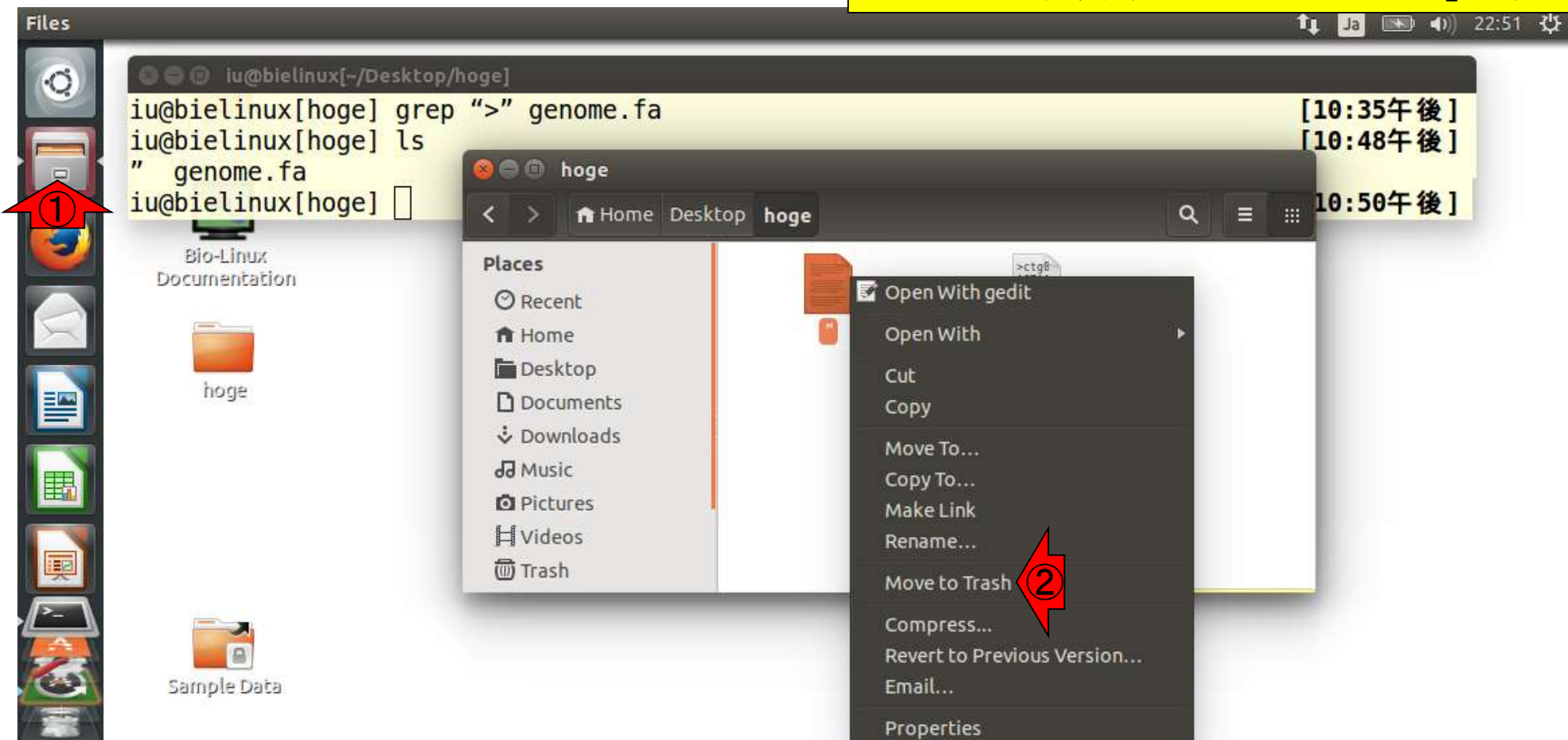
[10:32午後]

[10:35午後]

[10:35午後]

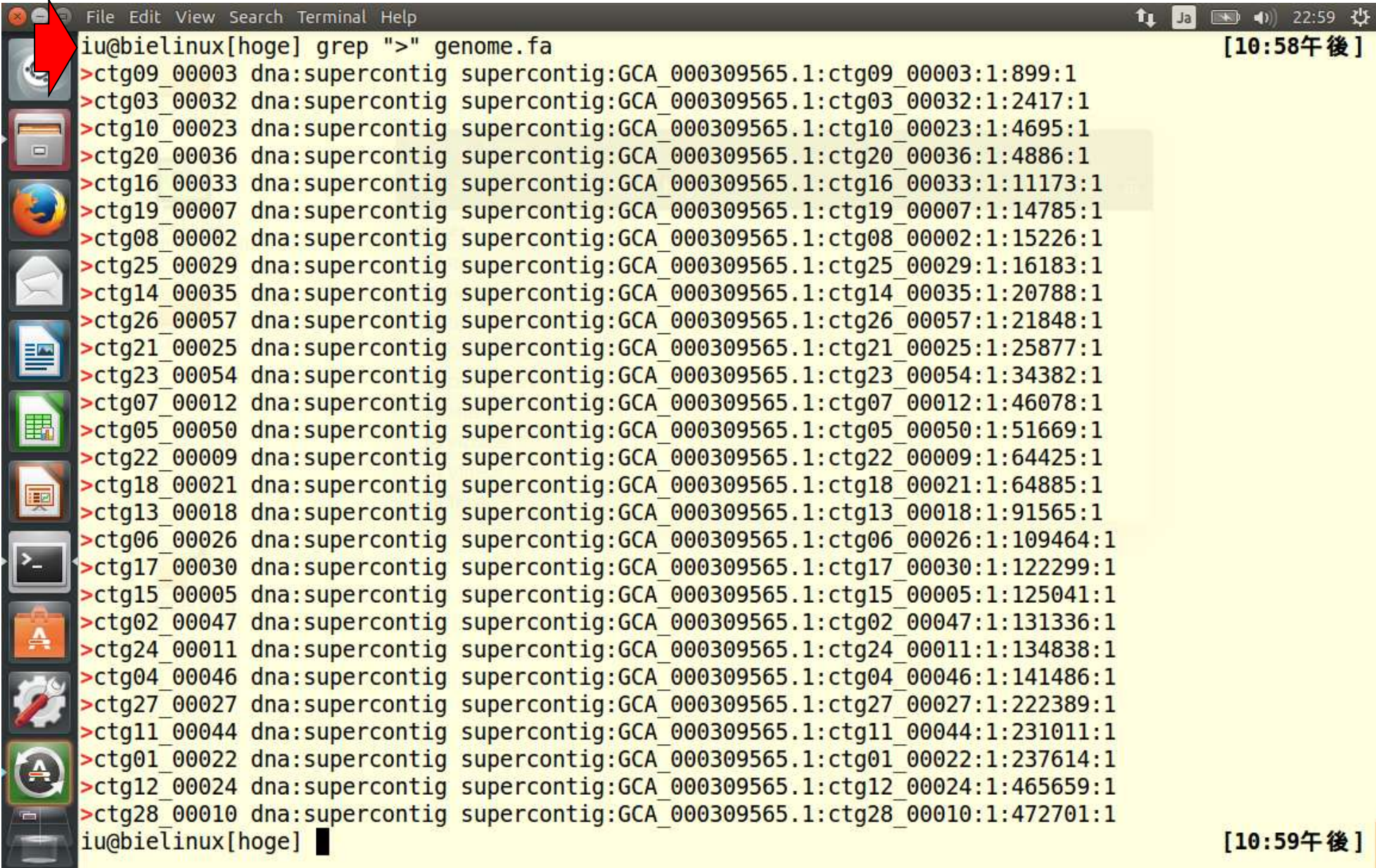
W16-1

別の削除手段。①引き出しみたいなのアイコンからhogeフォルダに移動し、②該当ファイルのアイコン上で右クリックの「Move to Trash」でもよい。



W16-2

Linux上での"正しい"ダブルクォーテーションマークは、開始記号と終了記号が同じほう。



```
iu@bielinux[hoge] grep ">" genome.fa
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
>ctg20_00036 dna:supercontig supercontig:GCA_000309565.1:ctg20_00036:1:4886:1
>ctg16_00033 dna:supercontig supercontig:GCA_000309565.1:ctg16_00033:1:11173:1
>ctg19_00007 dna:supercontig supercontig:GCA_000309565.1:ctg19_00007:1:14785:1
>ctg08_00002 dna:supercontig supercontig:GCA_000309565.1:ctg08_00002:1:15226:1
>ctg25_00029 dna:supercontig supercontig:GCA_000309565.1:ctg25_00029:1:16183:1
>ctg14_00035 dna:supercontig supercontig:GCA_000309565.1:ctg14_00035:1:20788:1
>ctg26_00057 dna:supercontig supercontig:GCA_000309565.1:ctg26_00057:1:21848:1
>ctg21_00025 dna:supercontig supercontig:GCA_000309565.1:ctg21_00025:1:25877:1
>ctg23_00054 dna:supercontig supercontig:GCA_000309565.1:ctg23_00054:1:34382:1
>ctg07_00012 dna:supercontig supercontig:GCA_000309565.1:ctg07_00012:1:46078:1
>ctg05_00050 dna:supercontig supercontig:GCA_000309565.1:ctg05_00050:1:51669:1
>ctg22_00009 dna:supercontig supercontig:GCA_000309565.1:ctg22_00009:1:64425:1
>ctg18_00021 dna:supercontig supercontig:GCA_000309565.1:ctg18_00021:1:64885:1
>ctg13_00018 dna:supercontig supercontig:GCA_000309565.1:ctg13_00018:1:91565:1
>ctg06_00026 dna:supercontig supercontig:GCA_000309565.1:ctg06_00026:1:109464:1
>ctg17_00030 dna:supercontig supercontig:GCA_000309565.1:ctg17_00030:1:122299:1
>ctg15_00005 dna:supercontig supercontig:GCA_000309565.1:ctg15_00005:1:125041:1
>ctg02_00047 dna:supercontig supercontig:GCA_000309565.1:ctg02_00047:1:131336:1
>ctg24_00011 dna:supercontig supercontig:GCA_000309565.1:ctg24_00011:1:134838:1
>ctg04_00046 dna:supercontig supercontig:GCA_000309565.1:ctg04_00046:1:141486:1
>ctg27_00027 dna:supercontig supercontig:GCA_000309565.1:ctg27_00027:1:222389:1
>ctg11_00044 dna:supercontig supercontig:GCA_000309565.1:ctg11_00044:1:231011:1
>ctg01_00022 dna:supercontig supercontig:GCA_000309565.1:ctg01_00022:1:237614:1
>ctg12_00024 dna:supercontig supercontig:GCA_000309565.1:ctg12_00024:1:465659:1
>ctg28_00010 dna:supercontig supercontig:GCA_000309565.1:ctg28_00010:1:472701:1
iu@bielinux[hoge]
```

[10:58午後]

[10:59午後]

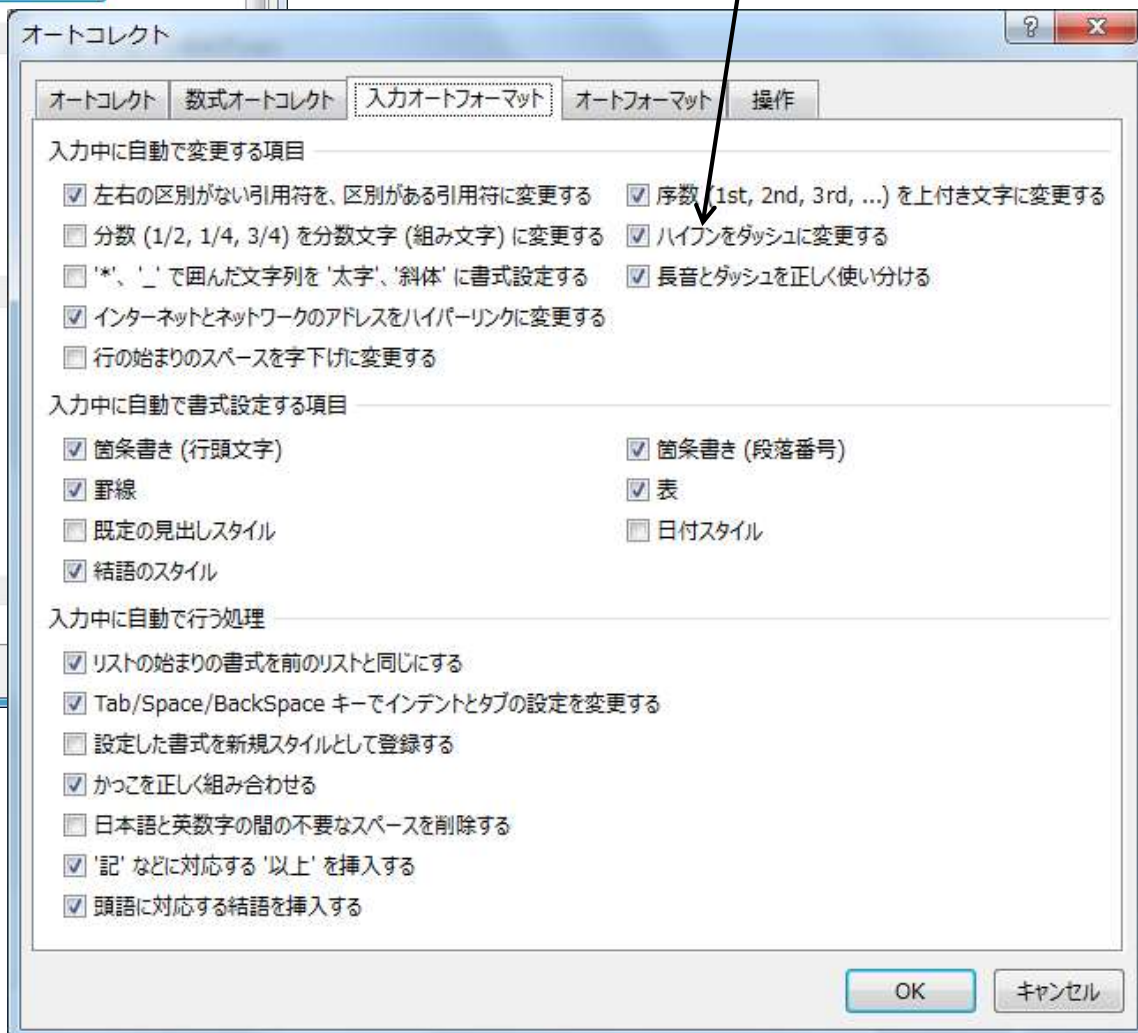
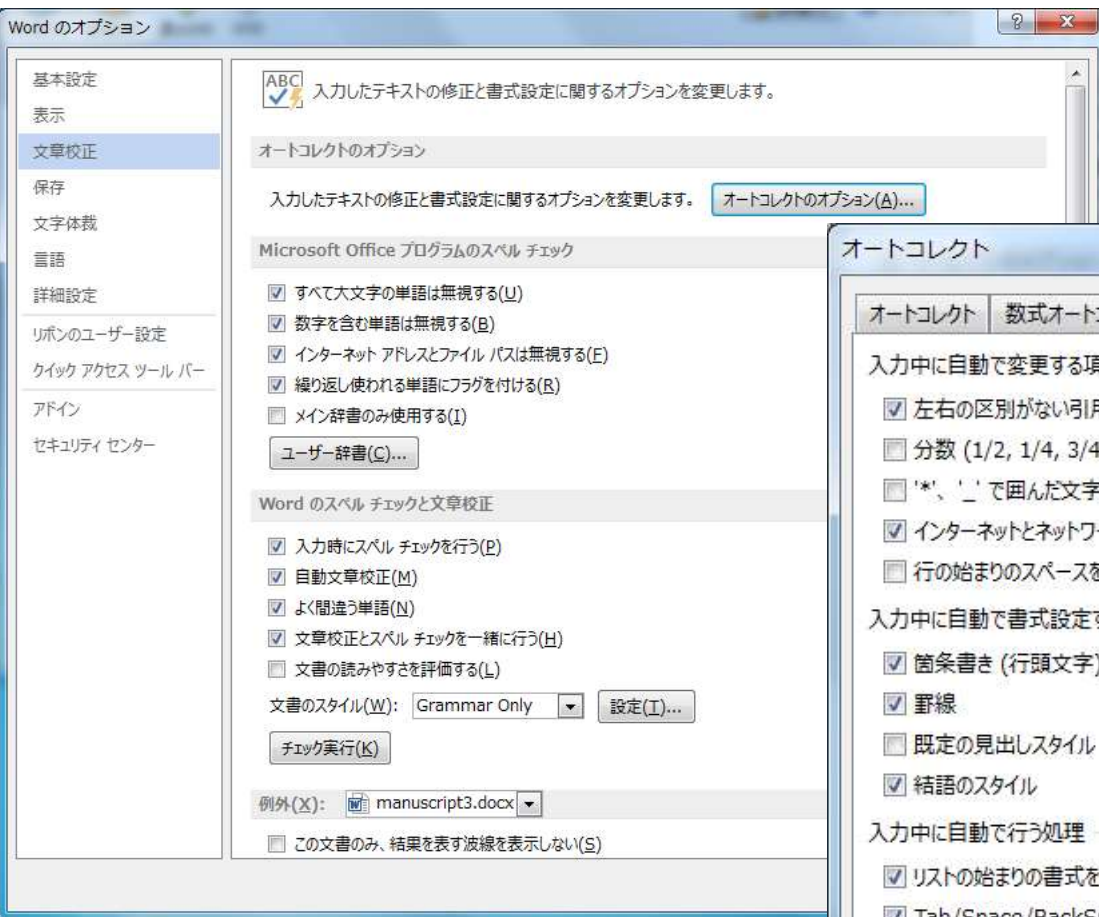
W16-3

原因既知状態で眺めると①「-(ハイフン)」と②「-(ダッシュ)」の違いは明白。Bio-Linuxで用いられているフォントは③いち、と④える、の区別が付きやすい。

```
iu@bielinux[hoge] pwd [ 9:14午前 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls [ 9:14午前 ]
genome.fa
① iu@bielinux[hoge] grep -c ">" genome.fa [ 9:14午前 ]
28
iu@bielinux[hoge] grep -c ">" genome.fa ② [ 9:14午前 ]
grep: >: No such file or directory
③ iu@bielinux[hoge] l [ 9:14午前 ]
zsh: command not found: l
iu@bielinux[hoge] l ④ [10:07午前 ]
zsh: command not found: l
iu@bielinux[hoge] [10:07午前 ]
```

W16-4

著者のデフォルト環境では、Microsoft Word の「オプション」-「オートコレクトのオプション」-「入力オートフォーマット」で「ハイフンをダッシュに変更する」のチェックが入っている。



W17-1

Bio-Linux上で利用可能なテキストエディタは、vi, emacs, geditなど。「vi genome.fa」、「emacs genome.fa」、「gedit genome.fa」などでそれぞれのエディタを起動できる。

The screenshot displays a Linux desktop environment. On the left is a vertical dock with icons for applications like a file manager, browser, and terminal. The main workspace contains two windows:

- Terminal Window:** The title bar reads "iu@bielinux[~/Desktop/hoge]". The command history shows:

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls
genome.fa
iu@bielinux[hoge] gedit genome.fa
```
- gedit Window:** The title bar reads "genome.fa (~/Desktop/hoge) - gedit". The window shows the content of the file genome.fa, which contains DNA sequence data. The visible text is:

```
>ctg09_00003 dna:supercontig
supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAAACTTTGTTTAGTTTTGAGGGGATCACCCCTCAAGCACCTAA
CGGGTGCGACTTTGTTCTTTGAAAACCTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAAGAAATTTCGCATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAAC TAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGCGCAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig
supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTTGGGCGCGTT
```

The gedit window also shows a menu bar with "Open", "Save", "Undo", and other options, and a status bar at the bottom indicating "Plain Text", "Tab Width: 8", "Ln 4, Col 51", and "INS".

W17-1

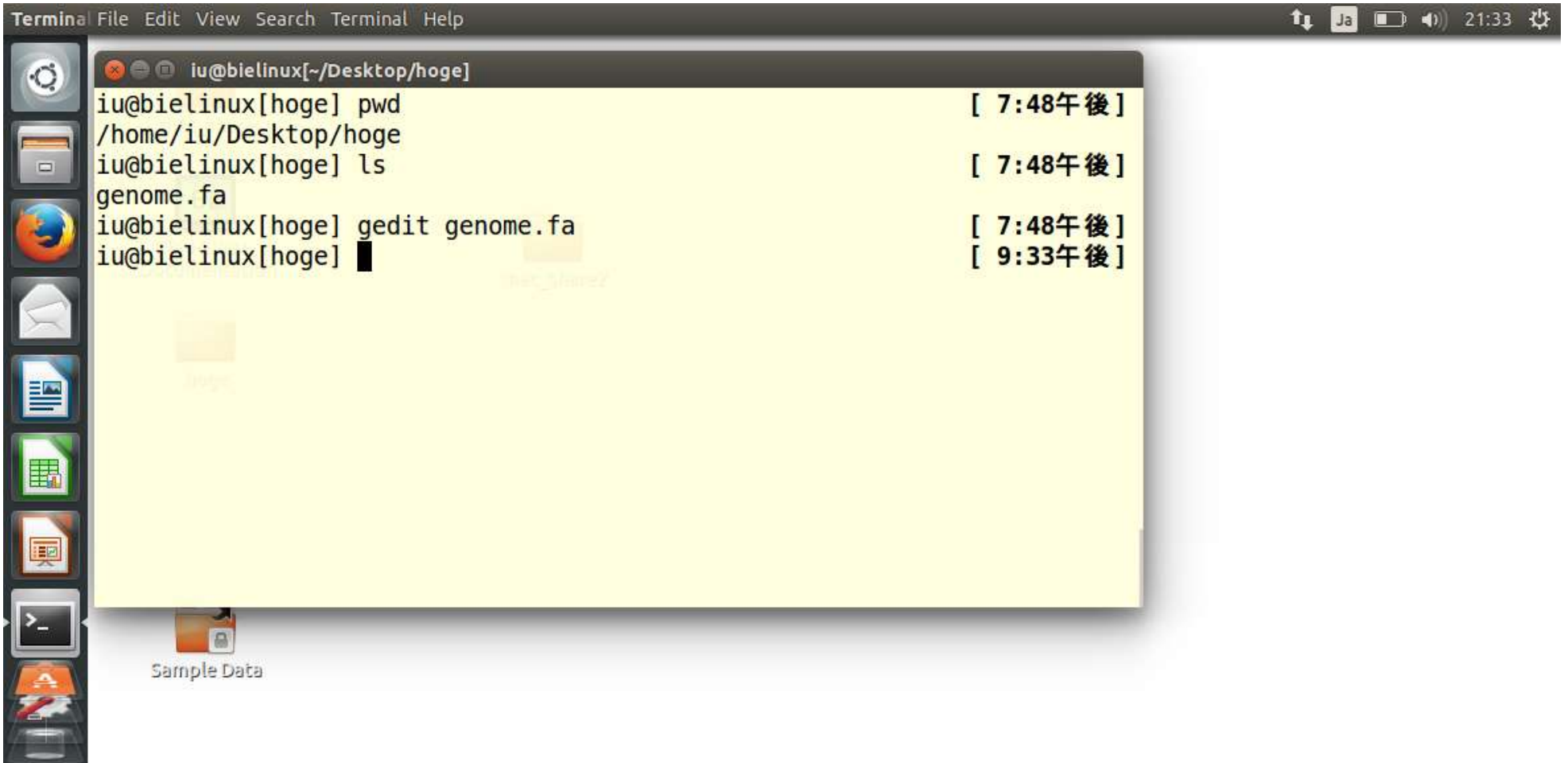
gedit起動中は、geditというコマンド実行中と同義である。そのため、このターミナル上では何か新たなコマンドを打ち込むことはできない。赤矢印の×ボタンを押すとコマンド入力待ち状態になる。

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls
genome.fa
iu@bielinux[hoge] gedit genome.fa
```

```
genome.fa (~/Desktop/hoge) - gedit
genome.fa x
>ctg09_00003 dna:supercontig
supercontig:GCA_000309565.1:ctg09_00003:1:899:1
AGTAAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAACGAGTTGCCGAGACCGCGAGG
TYAAGCTAATCTCTTAAAGCCATTCTCAGTTCGGACTGTAGGCTGCAACTCGCCTACACG
AAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTT
GTACACACCGCCCGTCACACCATGAGAGTTTGTAAACCCGAAGCCGGTGGCGTAACCCT
TTTAGGGAGCGAGCCGTCTAAGGTGGGACAAATGATTAGGGTGAAGTCGTAACAAGGTAG
CCGTAGGAGAACCCTGCGGCTGGATCACCTCCTTTCTAAGGAAACAGACTRAAAGTCTGAC
GGAAACCTGCACACACGAACTTTGTTTAGTTTTGAGGGGATCACCTCAAGCACCTTAA
CGGGTGGCAGCTTTGTTCTTTGAAAAGTGGATATCATTGTATTAATTGTTTTAAATTGCCG
AGAACACAGCGTATTTGTATGAGTTTCTGAAAAGAAATTTCGCATCGCATAACCGCTGAC
GCAAGTCAGTACAGGTTAAGTTRCAAAGGGCGCACGGTGGATGCCTTGGCACTAGGAGCC
GATGAAGGACGGAAC TAATACCGATATGCTTCGGGGAGCTATAAGTAAGCTTTGATCCGG
AGATTTCCGAATGGGGGAACCCAGTACACATCAGTGTATTGCTTGTGAGTGAATACATAG
CTGGCCGGCGGCAGACGCGGGGAACTGAAACATCTCAGTACCCGCAGGAAGAGAAAGAAA
ACTCGATTCCCATAGTAGCGCGGAGCGAAGTGGGAAAAGCCCAAACCGAGAAGCTTGCT
>ctg03_00032 dna:supercontig
supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
CGACTGTTAACATCCGCACTGGTGCCGGCACCGGCTATGCATCCGTTGGTAGCTATACAC
TCGGTGAAAGTGTGATTTATGATCACGTGTATATCCGTGGCACATATGTTTTGGGCGCGTT
```

Plain Text ▾ Tab Width: 8 ▾ Ln 4, Col 51 INS

W17-1



W17-2

The screenshot shows a Linux desktop environment with a terminal window and a Firefox browser window. The terminal window is titled 'iu@bielinux[~/Desktop/hoge]' and displays the following commands and outputs:

```

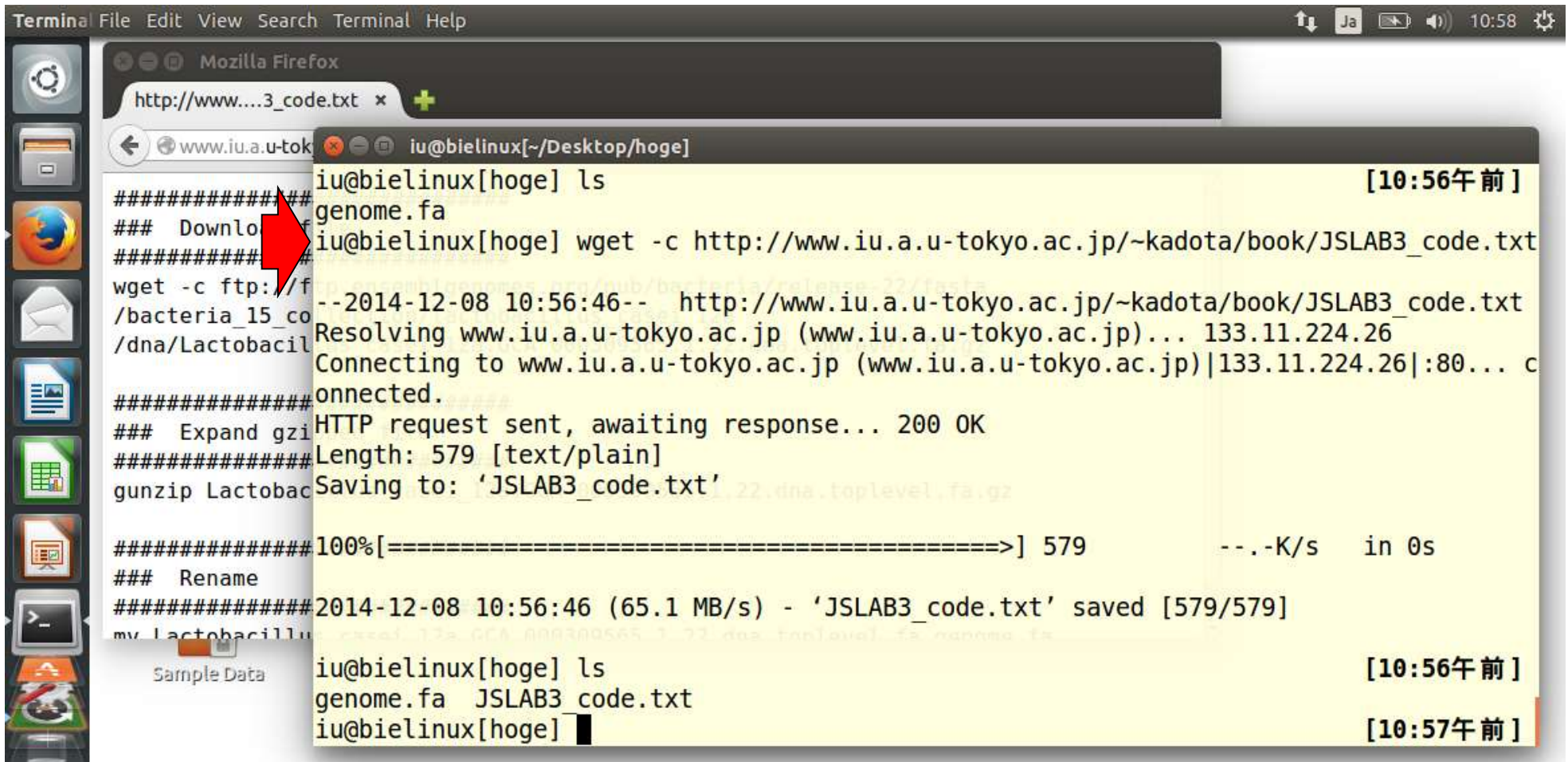
iu@bielinux[hoge] pwd [ 9:33午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls [10:40午前 ]
genome.fa
iu@bielinux[hoge] gedit genome.fa [10:40午前 ]
iu@bielinux[hoge] wget -c http://www.iu.a.u-tokyo.ac.jp/~kadota/book/JSL
AB3_code.txt
Lactobacillus_casei_12a_GCA_000309565.1.22.dna.toplevel.fa.gz
Lactobacillus_casei_12a_GCA_000309565.1.22.dna.toplevel.fa.gz
iu@bielinux[hoge] mv Lactobacillus_casei_12a_GCA_000309565.1.22.dna.toplevel.fa genome.fa
Sample Data

```

The Firefox browser window shows the URL `http://www.iu.a.u-tokyo.ac.jp/~kadota/book/JSLAB3_code.txt` in the address bar. The browser's content area displays the text from the terminal window, which is partially obscured by the terminal window's title bar and content.

(見栄えの関係上若干変更しているが…) wget
コマンドでファイルをダウンロードしてもよい。

W17-2



```
Terminal File Edit View Search Terminal Help
Mozilla Firefox
http://www...3_code.txt x +
www.iu.a.u-tokyo.ac.jp/~kadota/book/JSLAB3_code.txt
iu@bielinux[~/Desktop/hoge]
##### iu@bielinux[hoge] ls [10:56午前]
##### genome.fa
### Download genome.fa
##### iu@bielinux[hoge] wget -c http://www.iu.a.u-tokyo.ac.jp/~kadota/book/JSLAB3_code.txt
#####
wget -c ftp://ftp.iu.a.u-tokyo.ac.jp/pub/ftp.iu.a.u-tokyo.ac.jp/pub/dna/Lactobacillus_22.dna.toplevel.fa.gz
--2014-12-08 10:56:46-- http://www.iu.a.u-tokyo.ac.jp/~kadota/book/JSLAB3_code.txt
Resolving www.iu.a.u-tokyo.ac.jp (www.iu.a.u-tokyo.ac.jp)... 133.11.224.26
Connecting to www.iu.a.u-tokyo.ac.jp (www.iu.a.u-tokyo.ac.jp)|133.11.224.26|:80... connected.
HTTP request sent, awaiting response... 200 OK
### Expand gzip
##### Length: 579 [text/plain]
gunzip Lactobacillus_22.dna.toplevel.fa.gz
##### Saving to: 'JSLAB3_code.txt'
##### 100%[=====] 579 ---K/s in 0s
### Rename
##### 2014-12-08 10:56:46 (65.1 MB/s) - 'JSLAB3_code.txt' saved [579/579]
#####
Sample Data iu@bielinux[hoge] ls [10:56午前]
genome.fa JSLAB3_code.txt
iu@bielinux[hoge] [10:57午前]
```

(記述内容量が少ないことが分かっているので) moreコマンドで眺めているところ。

W17-2

The screenshot shows a terminal window with the following content:

```
##### genome.fa JSLAB3_code.txt
#####
### Download file #####
#####
### Download file #####
#####
wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
### Expand gzipped file #####
#####
gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
### Rename #####
#####
mv Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa genome.fa
#####
iu@bielinux[hoge]
#####
```

A red arrow points to the command `iu@bielinux[hoge] more JSLAB3_code.txt`. The terminal also shows a timestamp `[10:57午前]` and `[11:02午前]`.

W17-2

別のやり方を示すために一旦削除。rmコマンド実行時に、「-f」オプションをつけると「本当に消していいか？」という警告メッセージが表示されない。

```
Terminal
Mozilla Firefox
http://www....3_code.txt x +
www.iu.a.u-tok
iu@bielinux[~/Desktop/hoge]
#####
### Download f
#####
wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/release-22/fasta/bacteria_15_colle
ction/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.top
level.fa.gz
#####
#####
gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
### Expand gzi
#####
gunzip Lactobac
#####
### Rename
#####
mv Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa genome.fa
#####
### Rename
iu@bielinux[hoge] ls [11:02午前]
#####
### genome.fa JSLAB3_code.txt
iu@bielinux[hoge] rm -f JSLAB3_code.txt [11:05午前]
#####
mv Lactobac
iu@bielinux[hoge] ls [11:06午前]
#####
genome.fa
iu@bielinux[hoge] [11:06午前]
```

W17-3

別のやり方。geditを起動して、ウェブ上の情報をコピーし、任意のファイル名で保存するのもよい。

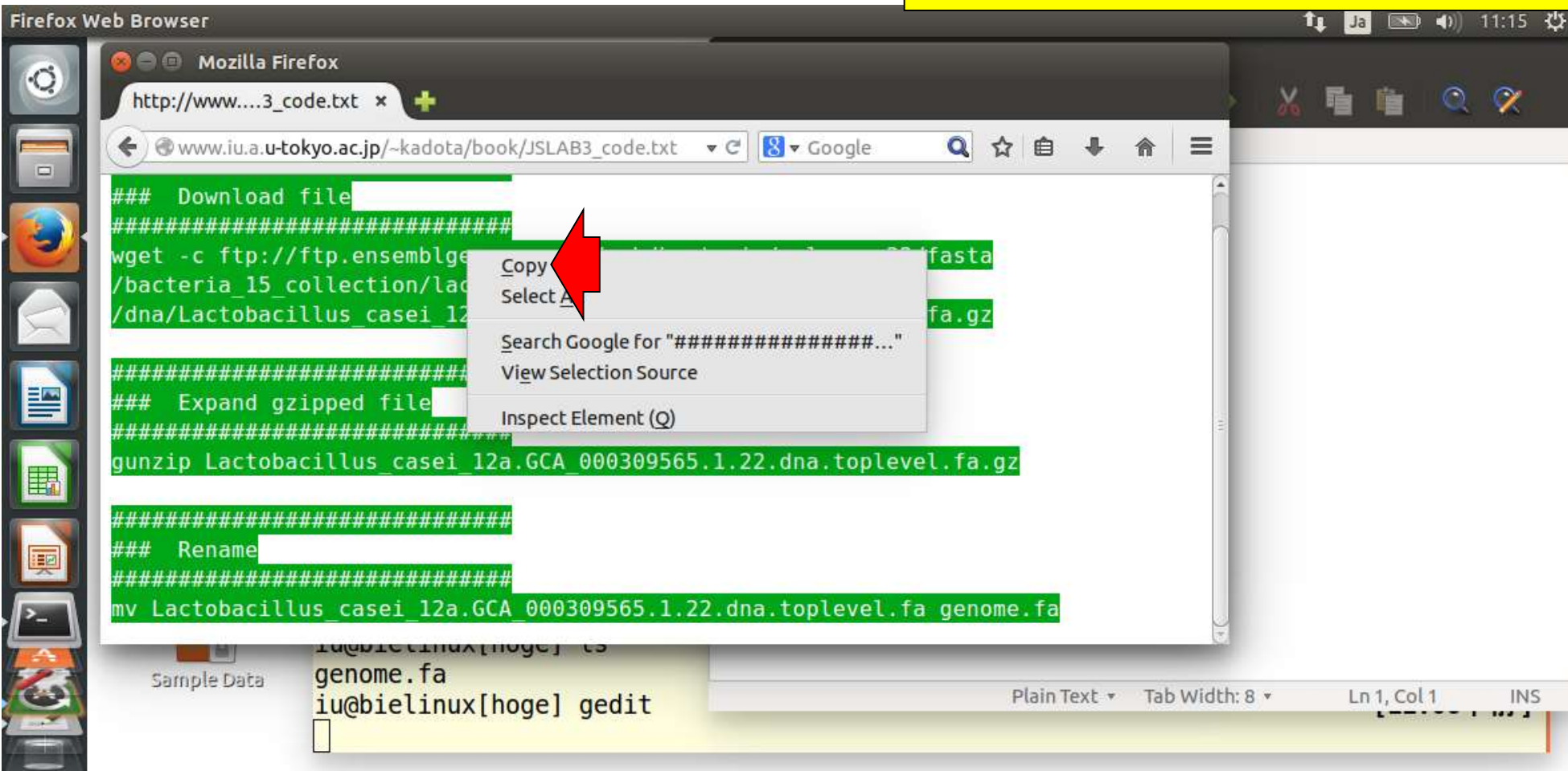
The screenshot shows a Linux desktop environment with a terminal window and a gedit text editor window. The terminal window displays a terminal session with the following commands and output:

```
#####  
#####  
### Download f  
#####  
wget -c ftp://f  
/bacteria_15_co  
/dna/Lactobacil  
#####  
#####  
### Expand gzi  
#####  
gunzip Lactobac  
#####  
### Rename  
#####  
mv Lactobacillus_casei_12a.  
#####  
iu@bielinux[hoge] ls  
### Rename  
genome.fa JSLAB3_code.txt  
#####  
iu@bielinux[hoge] rm -f JSL  
#####  
iu@bielinux[hoge] ls  
mv Lactobacillus  
genome.fa  
iu@bielinux[hoge] gedit
```

A red arrow points from the text "Sample Data" in the terminal window to the terminal prompt. The gedit window is titled "Untitled Document 1 - gedit" and is currently empty.

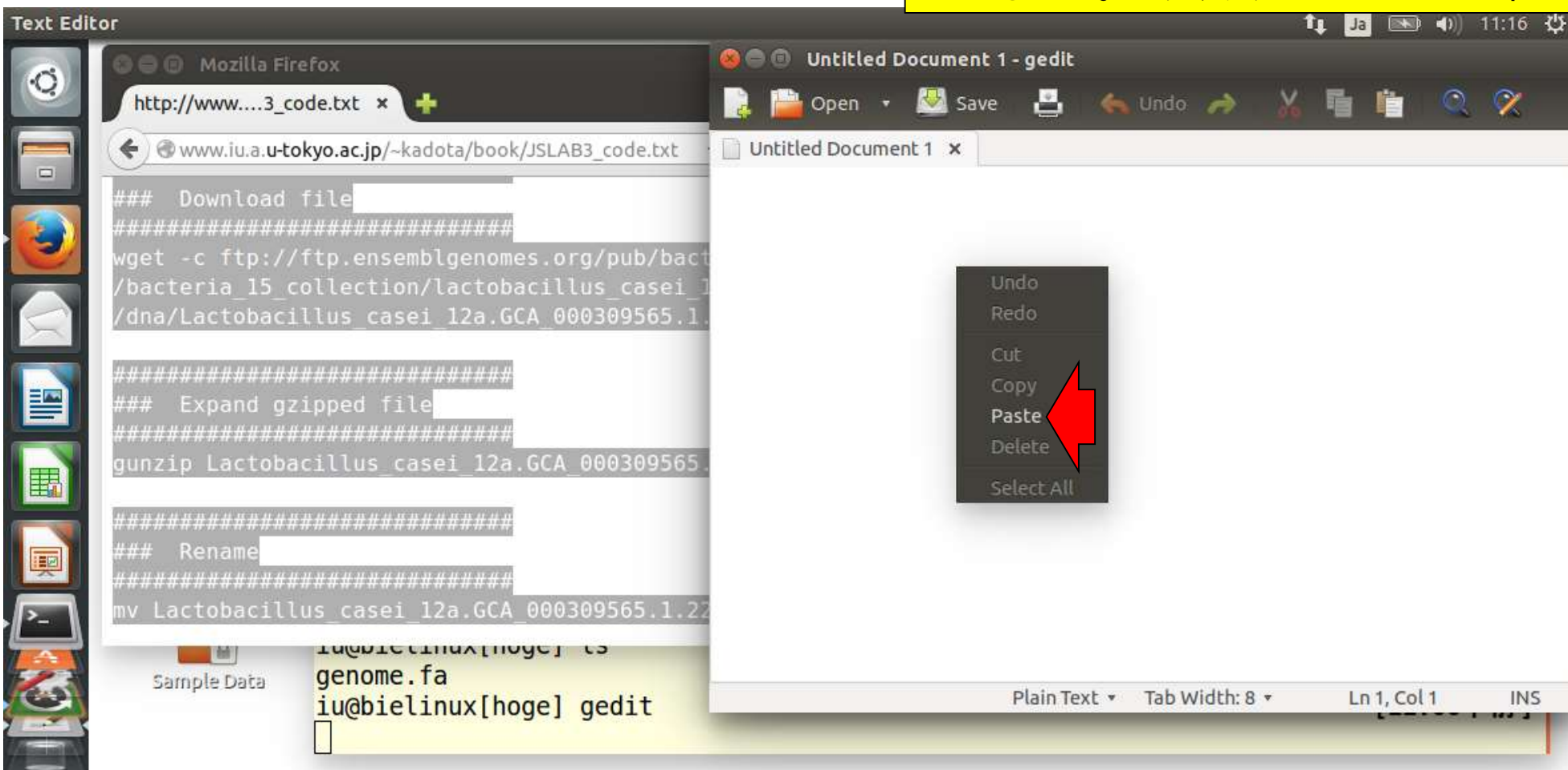
W17-3

別のやり方。geditを起動して、ウェブ上の情報をコピーし、任意のファイル名で保存するのもよい。右クリックのコピーが基本です。



W17-3

別のやり方。geditを起動して、ウェブ上の情報をコピーし、任意のファイル名で保存するのでもよい。右クリックのコピーが基本です。



W17-3

別のやり方。geditを起動して、ウェブ上の情報をコピーし、任意のファイル名で保存するのでもよい。右クリックのコピペが基本です。

The screenshot displays a Linux desktop environment with a terminal window and a gedit editor window. The terminal window shows the following commands and output:

```
### Download file [redacted]
#####
wget -c ftp://ftp.ensemblgenomes.org/pub/bacteria/bacteria_15_collection/lactobacillus_casei_12a/dna/Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
### Expand gzipped file [redacted]
#####
gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
### Rename [redacted]
#####
mv Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa genome.fa
```

The gedit editor window shows the same content being copied from the terminal. A red arrow points to the 'Save' button in the gedit window. The terminal window also shows the prompt `iu@bielinux[hoge] gedit`.

W17-3

デフォルトの保存先は、geditを起動した「カレントディレクトリ」になるようですね。ここではファイル名をJSLAB3_code.txtとしています。

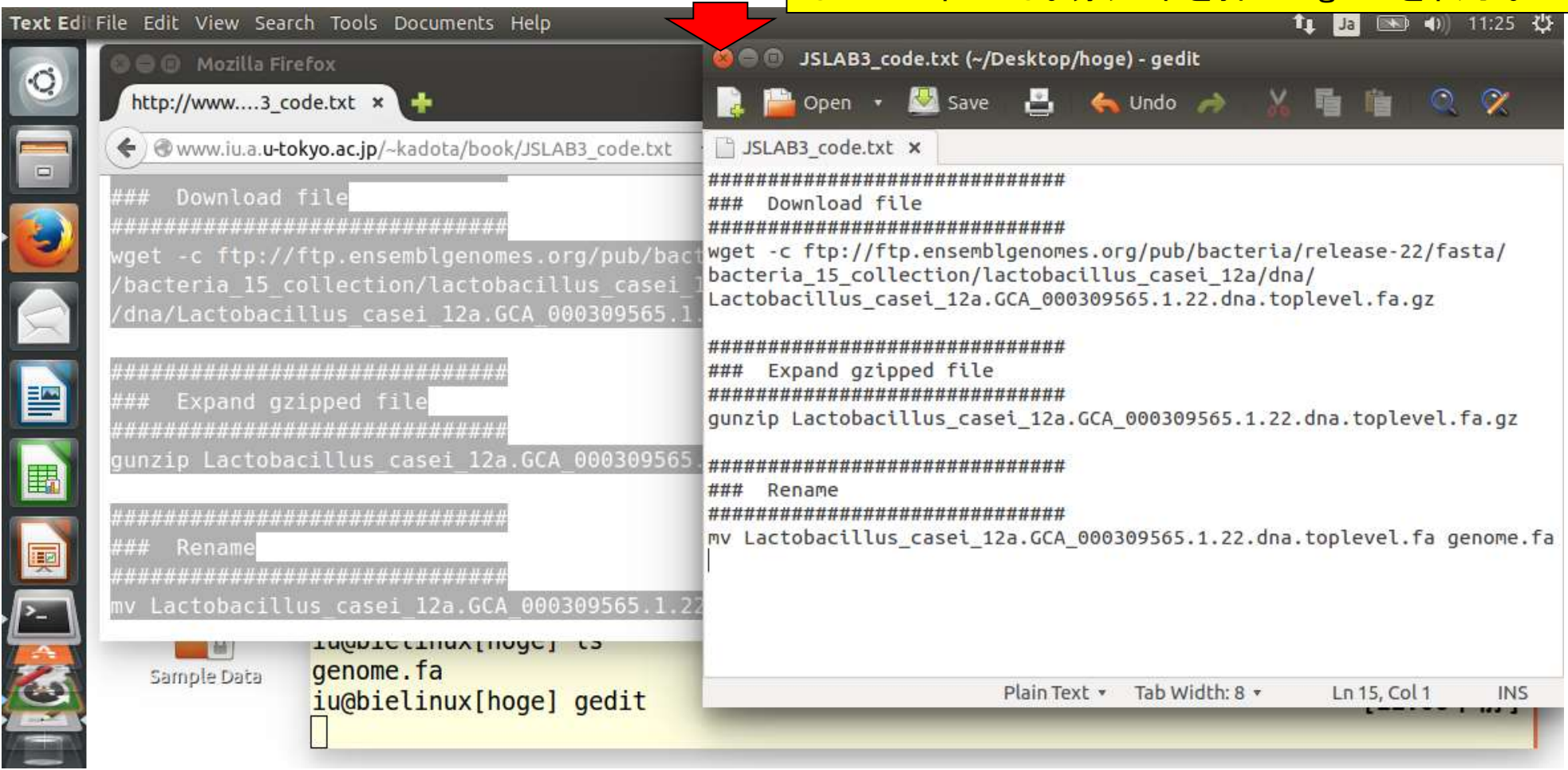
The screenshot shows a Linux desktop environment with a terminal window and a 'Save As' dialog box. The terminal window displays the following commands and output:

```
### Download file
#####
wget -c ftp://ftp.ensemblgenomes.o
/bacteria_15_collection/lactobacil
/dna/Lactobacillus casei 12a.GCA 0
#####
### Expand gzipped file
#####
gunzip Lactobacillus_casei 12a.GCA
#####
### Rename
#####
mv Lactobacillus_casei 12a.GCA 000
```

The 'Save As' dialog box shows the filename 'JSLAB3_code.txt' and the current directory 'iu Desktop hoge'. The 'Places' list includes Recent, Home, Desktop, Documents, Downloads, Music, Pictures, Videos, Trash, and hoge. The file list shows a file named 'genome.fa' with a size of 2.9 MB and a modified date of '木曜日'. The 'Character Encoding' is set to 'Current Locale (UTF-8)' and the 'Line Ending' is set to 'Unix/Linux'. A red arrow points to the 'Save' button.

W17-3

Saveボタンを押した後の状態。Linux環境とはいえ、通常のテキストエディタと同じような感覚で取り扱えることがわかる。赤矢印を押してgeditを終了。



W17-3

ターミナル画面上でlsすると、確かにgedit上で作成したJSLAB3_code.txtが存在することがわかる。

```
Terminal File Edit View Search Terminal Help
Mozilla Firefox
http://www....3_code.txt x +
www.iu.a.u-tok iu@bielinux[~/Desktop/hoge]
### Download f#####
#####
### Expand gzipped file
#####
gunzip Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa.gz
#####
#####
### Expand gzi### Rename
#####
gunzip Lactobacillmv Lactobacillus_casei_12a.GCA_000309565.1.22.dna.toplevel.fa genome.fa
iu@bielinux[hoge] ls [11:02午前]
genome.fa JSLAB3_code.txt
iu@bielinux[hoge] rm -f JSLAB3_code.txt [11:05午前]
iu@bielinux[hoge] ls [11:06午前]
mv Lactobacillugenome.fa
iu@bielinux[hoge] gedit [11:06午前]
iu@bielinux[hoge] ls [11:30午前]
genome.fa JSLAB3_code.txt
iu@bielinux[hoge] [11:30午前]
```

W18-1

リダイレクトで実行結果をファイルに保存。出力ファイル名で指定したheaders.txtが確かに生成されていることがわかる。

```
iu@bielinux[~/Desktop/hoge]
iu@bielinux[hoge] pwd [10:55午前]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls [10:55午前]
genome.fa JSLAB3_code.txt
iu@bielinux[hoge] grep ">" genome.fa > headers.txt [10:55午前]
iu@bielinux[hoge] ls [10:55午前]
genome.fa headers.txt JSLAB3_code.txt
iu@bielinux[hoge] head -n 15 headers.txt [10:55午前]
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
>ctg20_00036 dna:supercontig supercontig:GCA_000309565.1:ctg20_00036:1:4886:1
>ctg16_00033 dna:supercontig supercontig:GCA_000309565.1:ctg16_00033:1:11173:1
>ctg19_00007 dna:supercontig supercontig:GCA_000309565.1:ctg19_00007:1:14785:1
>ctg08_00002 dna:supercontig supercontig:GCA_000309565.1:ctg08_00002:1:15226:1
>ctg25_00029 dna:supercontig supercontig:GCA_000309565.1:ctg25_00029:1:16183:1
>ctg14_00035 dna:supercontig supercontig:GCA_000309565.1:ctg14_00035:1:20788:1
>ctg26_00057 dna:supercontig supercontig:GCA_000309565.1:ctg26_00057:1:21848:1
>ctg21_00025 dna:supercontig supercontig:GCA_000309565.1:ctg21_00025:1:25877:1
>ctg23_00054 dna:supercontig supercontig:GCA_000309565.1:ctg23_00054:1:34382:1
>ctg07_00012 dna:supercontig supercontig:GCA_000309565.1:ctg07_00012:1:46078:1
>ctg05_00050 dna:supercontig supercontig:GCA_000309565.1:ctg05_00050:1:51669:1
>ctg22_00009 dna:supercontig supercontig:GCA_000309565.1:ctg22_00009:1:64425:1
iu@bielinux[hoge] [10:55午前]
```

W18-1

「grep ">" genome.fa > headers.txt」に引き続いて「wc -l headers.txt」とすることでコンティグ数を調べることができる。

```
File Edit View Search Terminal Help
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
>ctg20_00036 dna:supercontig supercontig:GCA_000309565.1:ctg20_00036:1:4886:1
>ctg16_00033 dna:supercontig supercontig:GCA_000309565.1:ctg16_00033:1:11173:1
>ctg19_00007 dna:supercontig supercontig:GCA_000309565.1:ctg19_00007:1:14785:1
>ctg08_00002 dna:supercontig supercontig:GCA_000309565.1:ctg08_00002:1:15226:1
>ctg25_00029 dna:supercontig supercontig:GCA_000309565.1:ctg25_00029:1:16183:1
>ctg14_00035 dna:supercontig supercontig:GCA_000309565.1:ctg14_00035:1:20788:1
>ctg26_00057 dna:supercontig supercontig:GCA_000309565.1:ctg26_00057:1:21848:1
>ctg21_00025 dna:supercontig supercontig:GCA_000309565.1:ctg21_00025:1:25877:1
>ctg23_00054 dna:supercontig supercontig:GCA_000309565.1:ctg23_00054:1:34382:1
>ctg07_00012 dna:supercontig supercontig:GCA_000309565.1:ctg07_00012:1:46078:1
>ctg05_00050 dna:supercontig supercontig:GCA_000309565.1:ctg05_00050:1:51669:1
>ctg22_00009 dna:supercontig supercontig:GCA_000309565.1:ctg22_00009:1:64425:1
iu@bielinux[hoge] wc -l headers.txt [10:55午前]
28 headers.txt
iu@bielinux[hoge] wc headers.txt [11:19午前]
 28  84 2218 headers.txt
iu@bielinux[hoge] grep -c ">" genome.fa [11:20午前]
28
iu@bielinux[hoge] grep -c ">" genome.fa > tmp.txt [11:22午前]
iu@bielinux[hoge] more tmp.txt [11:23午前]
28
iu@bielinux[hoge] [11:23午前]
```

W18-2

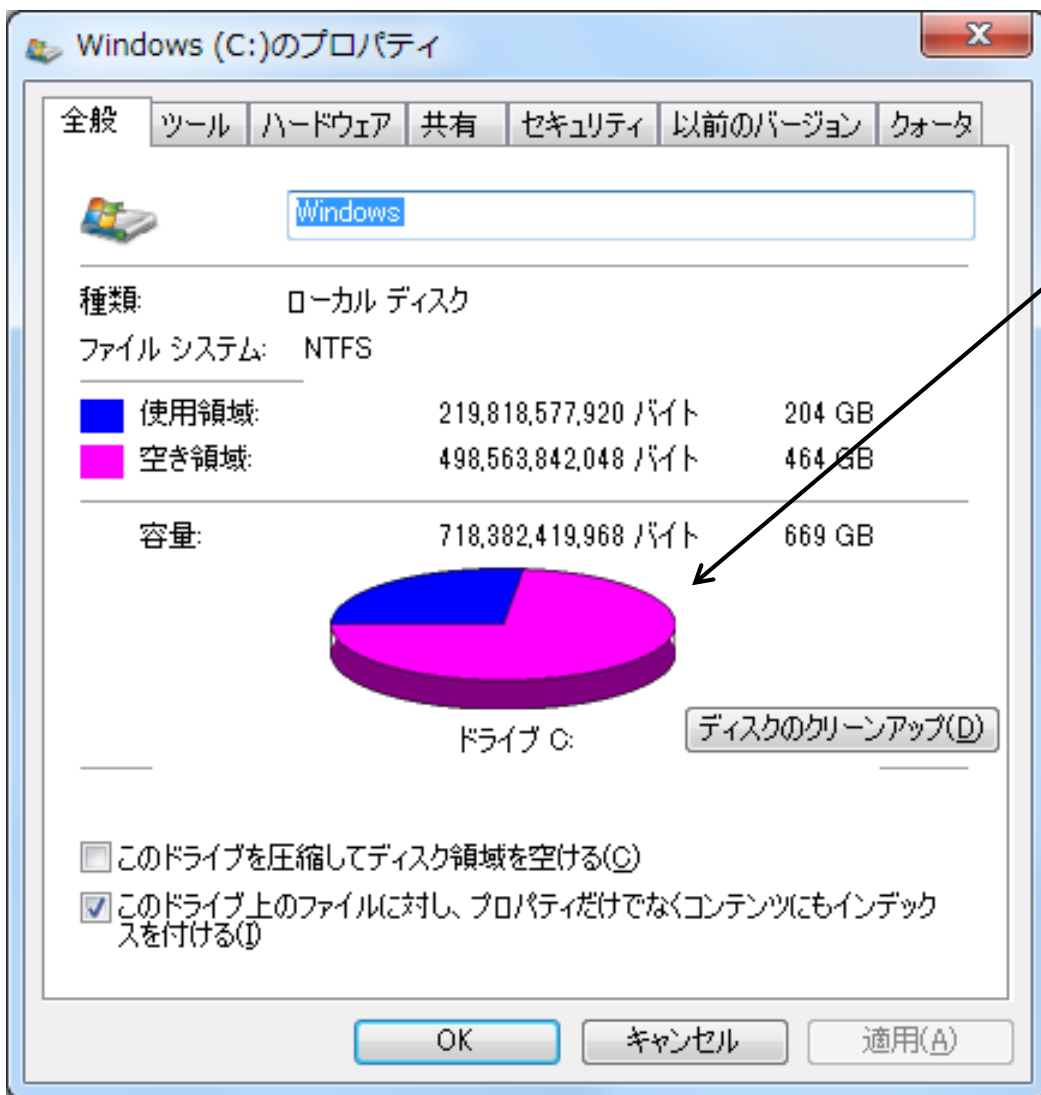
どのディレクトリ上でもいいので「df -h」を実行。Sizeのところが146Gとなっているのは、仮想マシン設定時に割り当てを150GBとしたから。Use%のところが全体の何%を利用したかを表している。ここでは7%となっているが、90%を超えない程度に適宜不要なファイルを削除したほうがいいだろう。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G   8.8G  130G   7% /
none            4.0K    0  4.0K   0% /sys/fs/cgroup
udev            991M   4.0K  991M   1% /dev
tmpfs           201M   920K  200M   1% /run
none            5.0M    0  5.0M   0% /run/lock
none           1002M  156K 1001M   1% /run/shm
none            100M   64K  100M   1% /run/user
iu@bielinux[hoge] cd
iu@bielinux[iu] pwd
/home/iu/Desktop
iu@bielinux[iu] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G   8.8G  130G   7% /
none            4.0K    0  4.0K   0% /sys/fs/cgroup
udev            991M   4.0K  991M   1% /dev
tmpfs           201M   920K  200M   1% /run
none            5.0M    0  5.0M   0% /run/lock
none           1002M  156K 1001M   1% /run/shm
none            100M   64K  100M   1% /run/user
iu@bielinux[iu]
```

[2:37午後]
[2:33午後]
[2:33午後]
[2:33午後]
[2:37午後]
[2:37午後]
[2:37午後]
[2:37午後]

W18-2

「df -h」結果で眺めるUse%は、ホストOS (Windows)上で使用領域と空き領域の割合を眺めているのと同じようなもの。



W18-3

カレントディレクトリ(/home/iu/Desktop/hoge)のディスク使用量(2.9 MB)は、lsコマンド実行結果からも一応分かる。が、ディレクトリ内の全容量を調べるのはduコマンドを利用するのが一般的。

```
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月 4 12:41 genome.fa
-rw-rw-r-- 1 iu iu 2.2K 12月 9 10:55 headers.txt
-rw-rw-r-- 1 iu iu 579 12月 7 21:08 JSLAB3_code.txt
-rw-rw-r-- 1 iu iu 3 12月 9 11:23 tmp.txt
iu@bielinux[hoge] du -sh
2.9M .
iu@bielinux[hoge] cd ..
iu@bielinux[Desktop] pwd
/home/iu/Desktop
iu@bielinux[Desktop] ls -lh
total 12K
drwxr-xr-x 2 iu iu 4.0K 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4.0K 12月 9 11:23 hoge
drwxrwxrwx 1 root root 0 12月 1 23:13 mac_share
drwxrwxr-x 2 iu iu 4.0K 11月 29 14:32 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledatal/
iu@bielinux[Desktop] du -sh
6.4M .
iu@bielinux[Desktop]
```


W18-3

黒矢印の「どっと(.)」はカレントディレクトリという意味。つまり、以下の全ファイルの容量は2.9Mですよという出力結果。

```
iu@bielinux[hoge] pwd [ 3:22午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [ 3:22午後 ]
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月 4 12:41 genome.fa
-rw-rw-r-- 1 iu iu 2.2K 12月 9 10:55 headers.txt
-rw-rw-r-- 1 iu iu 579 12月 7 21:08 JSLAB3_code.txt
-rw-rw-r-- 1 iu iu 3 12月 9 11:23 tmp.txt
iu@bielinux[hoge] du -sh [ 3:22午後 ]
2.9M .
iu@bielinux[hoge] cd .. [ 3:23午後 ]
iu@bielinux[Desktop] pwd [ 3:23午後 ]
/home/iu/Desktop
iu@bielinux[Desktop] ls -lh [ 3:23午後 ]
total 12K
drwxr-xr-x 2 iu iu 4.0K 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4.0K 12月 9 11:23 hoge
drwxrwxrwx 1 root root 0 12月 1 23:13 mac_share
drwxrwxr-x 2 iu iu 4.0K 11月 29 14:32 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledatal/
iu@bielinux[Desktop] du -sh [ 3:23午後 ]
6.4M .
iu@bielinux[Desktop] [ 3:24午後 ]
```

W18-3

「cd ..」は1つ上の階層のディレクトリに移動するときのお約束的な指令。

```
iu@bielinux[hoge] pwd [ 3:22午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [ 3:22午後 ]
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月 4 12:41 genome.fa
-rw-rw-r-- 1 iu iu 2.2K 12月 9 10:55 headers.txt
-rw-rw-r-- 1 iu iu 579 12月 7 21:08 JSLAB3_code.txt
-rw-rw-r-- 1 iu iu 3 12月 9 11:23 tmp.txt
iu@bielinux[hoge] du -sh [ 3:22午後 ]
2.9M .
iu@bielinux[hoge] cd .. [ 3:23午後 ]
iu@bielinux[Desktop] pwd [ 3:23午後 ]
/home/iu/Desktop
iu@bielinux[Desktop] ls -lh [ 3:23午後 ]
total 12K
drwxr-xr-x 2 iu iu 4.0K 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4.0K 12月 9 11:23 hoge
drwxrwxrwx 1 root root 0 12月 1 23:13 mac_share
drwxrwxr-x 2 iu iu 4.0K 11月 29 14:32 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledatal/
iu@bielinux[Desktop] du -sh [ 3:23午後 ]
6.4M .
iu@bielinux[Desktop] [ 3:24午後 ]
```

W18-3

カレントディレクトリが/home/iu/Desktopの状態、2つのコマンド(lsとdu)の実行結果を比較するとduコマンドを使う意義が分かる。

```
iu@bielinux[hoge] pwd [ 3:22午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls -lh [ 3:22午後 ]
total 2.9M
-rw-rw-r-- 1 iu iu 2.8M 12月 4 12:41 genome.fa
-rw-rw-r-- 1 iu iu 2.2K 12月 9 10:55 headers.txt
-rw-rw-r-- 1 iu iu 579 12月 7 21:08 JSLAB3_code.txt
-rw-rw-r-- 1 iu iu 3 12月 9 11:23 tmp.txt
iu@bielinux[hoge] du -sh [ 3:22午後 ]
2.9M .
iu@bielinux[hoge] cd .. [ 3:23午後 ]
iu@bielinux[Desktop] pwd [ 3:23午後 ]
/home/iu/Desktop
iu@bielinux[Desktop] ls -lh [ 3:23午後 ]
total 12K
drwxr-xr-x 2 iu iu 4.0K 11月 10 16:03 Bio-Linux Documentation
drwxrwxr-x 2 iu iu 4.0K 12月 9 11:23 hoge
drwxrwxrwx 1 root root 0 12月 1 23:13 mac_share
drwxrwxr-x 2 iu iu 4.0K 11月 29 14:32 mac_share2
lrwxrwxrwx 1 iu iu 29 11月 10 16:03 Sample Data -> /usr/local/bioinf/sampledatal/
iu@bielinux[Desktop] du -sh [ 3:23午後 ]
6.4M .
iu@bielinux[Desktop] [ 3:24午後 ]
```

W18-3

①cdでカレントディレクトリをホームディレクトリとして
②「du -sh」を実行した結果。ユーザiuは、現在のところトータルで88 MBディスクを使用していることがわかる。見えている数値はヒトによって若干異なる。

```
iu@bielinux[Desktop] pwd
/home/iu/Desktop
iu@bielinux[Desktop] ls
Bio-Linux Documentation hoge mac_share mac_share2 Sample Data
iu@bielinux[Desktop] cd
iu@bielinux[iu] pwd
/home/iu
iu@bielinux[iu] ls
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] du -sh
88M .
iu@bielinux[iu] du -sh Desktop
6.4M Desktop
iu@bielinux[iu] du -sh Documents
27M Documents
iu@bielinux[iu] du -sh Downloads
1.8M Downloads
iu@bielinux[iu]
```

W18-4

ディレクトリ削除例。rmコマンド実行時に-rをつけると「ディレクトリ内のファイルも削除」となる。fオプションも同時につけることによって、ファイル数分の警告メッセージへの対処を回避している。ごみ箱行きではなく消滅なので実行時は十分に注意したほうがよい。

```
iu@bielinux[Desktop] pwd
/home/iu/Desktop
iu@bielinux[Desktop] ls
Bio-Linux Documentation hoge mac_share mac_share2 Sample Data
iu@bielinux[Desktop] du -sh
6.4M .
iu@bielinux[Desktop] du -sh mac_share2
4.0K mac_share2
iu@bielinux[Desktop] ls mac_share2
iu@bielinux[Desktop] rm -rf mac_share2
iu@bielinux[Desktop] ls
Bio-Linux Documentation hoge mac_share Sample Data
iu@bielinux[Desktop] du -sh
6.4M .
iu@bielinux[Desktop]
```

[4:20午後]
[4:21午後]
[4:21午後]
[4:21午後]
[4:21午後]
[4:21午後]
[4:22午後]
[4:22午後]
[4:22午後]

W19-1

パイプ利用例1。「grepとwc」以外にも「grepとhead」などいろいろ組み合わせることができる。

```
iu@bielinux[hoge] pwd [ 5:22午後 ]
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls [ 5:24午後 ]
genome.fa headers.txt JSLAB3_code.txt tmp.txt
iu@bielinux[hoge] grep ">" genome.fa | wc [ 5:24午後 ]
      28      84     2218
iu@bielinux[hoge] grep ">" genome.fa | wc -l [ 5:24午後 ]
28
iu@bielinux[hoge] grep ">" genome.fa | head -n 3 [ 5:25午後 ]
>ctg09_00003 dna:supercontig supercontig:GCA_000309565.1:ctg09_00003:1:899:1
>ctg03_00032 dna:supercontig supercontig:GCA_000309565.1:ctg03_00032:1:2417:1
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
iu@bielinux[hoge] grep ">" genome.fa | tail -n 2 [ 5:25午後 ]
>ctg12_00024 dna:supercontig supercontig:GCA_000309565.1:ctg12_00024:1:465659:1
>ctg28_00010 dna:supercontig supercontig:GCA_000309565.1:ctg28_00010:1:472701:1
iu@bielinux[hoge] [ 5:26午後 ]
```

W19-2

パイプ利用例2。パイプ(|)の両側のスペースはあってもなくてもいいようだ。grepの結果をgrepすることもできる。右側の2回目のgrepでは“_0002”を含む行を出力している。

```
File Edit View Search Terminal Help
genome.fa headers.txt JSLAB3_code.txt tmp.txt
iu@bielinux[hoge]
iu@bielinux[hoge] pwd
/home/iu/Desktop/hoge
iu@bielinux[hoge] ls
genome.fa headers.txt JSLAB3 code.txt tmp.txt
iu@bielinux[hoge] grep ">" genome.fa | wc -l
28
iu@bielinux[hoge] grep ">" genome.fa |wc -l
28
iu@bielinux[hoge] grep ">" genome.fa|wc -l
28
iu@bielinux[hoge] grep ">" genome.fa | grep "_0002"
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
>ctg25_00029 dna:supercontig supercontig:GCA_000309565.1:ctg25_00029:1:16183:1
>ctg21_00025 dna:supercontig supercontig:GCA_000309565.1:ctg21_00025:1:25877:1
>ctg18_00021 dna:supercontig supercontig:GCA_000309565.1:ctg18_00021:1:64885:1
>ctg06_00026 dna:supercontig supercontig:GCA_000309565.1:ctg06_00026:1:109464:1
>ctg27_00027 dna:supercontig supercontig:GCA_000309565.1:ctg27_00027:1:222389:1
>ctg01_00022 dna:supercontig supercontig:GCA_000309565.1:ctg01_00022:1:237614:1
>ctg12_00024 dna:supercontig supercontig:GCA_000309565.1:ctg12_00024:1:465659:1
iu@bielinux[hoge]
```

[6:00午後]
[6:00午後]
[6:00午後]
[6:01午後]
[6:02午後]
[6:02午後]
[6:02午後]
[6:05午後]

W19-3

パイプ利用例3。①3つ以上のコマンド同士をパイプで連結することも可能。②headとtailを連結することで、コマンド単体では概観しづらい箇所も眺めることができる。

```
iu@bielinux[hoge] grep ">" genome.fa | grep "_0002" [ 6:02午後 ]
>ctg10_00023 dna:supercontig supercontig:GCA_000309565.1:ctg10_00023:1:4695:1
>ctg25_00029 dna:supercontig supercontig:GCA_000309565.1:ctg25_00029:1:16183:1
>ctg21_00025 dna:supercontig supercontig:GCA_000309565.1:ctg21_00025:1:25877:1
>ctg18_00021 dna:supercontig supercontig:GCA_000309565.1:ctg18_00021:1:64885:1
>ctg06_00026 dna:supercontig supercontig:GCA_000309565.1:ctg06_00026:1:109464:1
>ctg27_00027 dna:supercontig supercontig:GCA_000309565.1:ctg27_00027:1:222389:1
>ctg01_00022 dna:supercontig supercontig:GCA_000309565.1:ctg01_00022:1:237614:1
>ctg12_00024 dna:supercontig supercontig:GCA_000309565.1:ctg12_00024:1:465659:1
iu@bielinux[hoge] grep ">" genome.fa | grep -c "_0002" [ 6:05午後 ]
8
iu@bielinux[hoge] grep ">" genome.fa | grep "_0002" | wc [ 6:14午後 ]
      8      24     635
iu@bielinux[hoge] grep ">" genome.fa | grep "_0002" | wc -l [ 6:14午後 ]
8
iu@bielinux[hoge] head -n 1010 genome.fa | tail -n 10 [ 6:15午後 ]
GTTTTGTATGCTTTGACCCAAATTCACCGCTGACAGCAACAGTGCAAAGACTACCATCAC
CCGCAACAAAACCTTTATGTGCGGATCTTACCAGTTTCAATAAACGATCTTCACCCATTTA
TTTCACCTCCGCGTTTTTGGCGCGAAACTCAACTTTAAATAAATCAGATAGCGATAACGG
CAAGCTCTCATTAGCAGCAGGGTTCTTCGGCGGCAATCGCGGCCTTGACTTCAGGTGTGAA
ATGCGGGAAAACCGCTTCGATCACGCGTCCACGAATGGTCAGAATCCGACCTTGTTCCCT
GATCACATCAGGAATCTCCTTGTTTTGAAAATAACTGCAGCTTGATCGCCTGCTGTCT
GACGTCCTCTAGCGTCAAGTCTGACTGATCCGACCGGCTTTAACGAACAAAACGCGATC
TGACAAGCCGTCTAATTCGACTAGATTATGTGAGGCGATGAGAAAGGTTCTTTACCATC
AGCAACAGCATCTACCACAAATTGACAATCCGCTCGCGAATCAGCAAATCCAACCCATC
AAATGGTTCATCCAGAAAAACATACGGTGCCCGCTCGCCAGCGCCAACGCCACCATCAC
iu@bielinux[hoge] [ 6:17午後 ]
```


W20-1

NCBI Resources How To Sign in to NCBI

SRA SRA Lactobacillus casei Search Help

SRA

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

Getting Started

- [Understanding and Using SRA](#)
- [How to Submit](#)
- [Login to Submit](#)
- [Download Guide](#)

Tools and Software

- [Download SRA Toolkit](#)
- [SRA Toolkit Documentation](#)
- [SRA-BLAST](#)
- [SRA Run Browser](#)
- [SRA Run Selector](#)

Related Resources

- [dbGaP Home](#)
- [Trace Archive Home](#)
- [BioSample](#)
- [GenBank Home](#)

W20-1

31ヒット。ERXやDRX単位での出力であることがわかる。ERPやDRP単位ではないのでわかりづらいが、タイトルの文字列から「赤枠の2つは同じデータセットに含まれるのだろう」という程度の予想はつく。

NCBI Resources How To Sign in to NCBI

SRA SRA Lactobacillus casei Search

Display Settings: Summary, 20 per page Send to: Filters: Manage Filters

Results: 1 to 20 of 31

- [Population Genomics of Lactobacillus casei](#)
1. 1 ILLUMINA (Illumina HiSeq 2000) run: 5.5M spots, 1.1G bases, 739.1Mb downloads
Accession: ERX358287
- [Population Genomics of Lactobacillus casei](#)
2. 1 ILLUMINA (Illumina HiSeq 2000) run: 5.8M spots, 1.2G bases, 761.4Mb downloads
Accession: ERX358286
- [Whole Genome Sequencing of Lactobacillus casei](#)
3. 1 ILLUMINA (Illumina HiSeq 2000) run: 2.5M spots, 507.9M bases, 263Mb downloads
Accession: ERX529187
- [Whole genome analysis of Lactobacillus casei JCM 1134](#)
4. 1 ION_TORRENT (Ion Torrent PGM) run: 535,510 spots, 161.7M bases, 388.8Mb downloads
Accession: DRX002585

Database	Access		all
	public	controlled	
BioSample	23		23
BioProject	6		6
dbGaP			
GEO Datasets			

W20-2

The screenshot shows a web browser window at the URL <http://www.ebi.ac.uk/ena>. A cookie notice is displayed at the top, with a "Dismiss this notice" button. The ENA logo and "European Nucleotide Archive" text are visible on the left. A search bar contains the text "Lactobacillus casei" with a clear button (x) and a "Search" button. Below the search bar, examples are listed: "Examples: BN000065, histone". A red arrow points to the "Search" button, and the word "Sequence" is written below it. The main navigation menu includes "Home", "Search & Browse", "Submit & Update", "Software", "About ENA", and "Support". The "Popular" section lists several categories: "Submit and update", "Sequence submissions", "Genome assembly submissions", "Submitting environmental sequences", "Citing ENA data", and "Rest URLs for data retrieval".

W20-2

The screenshot shows a web browser window with the URL `http://www.ebi.ac.uk/ena/data/search?query=Lactobacillus+casei`. The page displays a search results page for the European Nucleotide Archive (ENA). The search term is `Lactobacillus casei`. The results are categorized into 'Assembly' and 'Sequence'. Under 'Assembly', it shows 'Assembly (28 results found)' with a link to 'View all 28 results'. Under 'Sequence', it shows 'GCA_000309565 Comparative genomics of Lactobacillus casei'. A red arrow points to the scrollbar on the right side of the browser window, indicating that the search results are located at the bottom of the page.



W20-2

DRP, ERP, SRPのように、3文字目がPになっているIDをクリックすると全体像が分かりやすいのでよい。DRP000852を眺めると、黒矢印部分のDRXやDRRはDRPの下階層のIDであることがわかる。

The screenshot shows a web browser window with the URL <http://www.ebi.ac.uk/ena/data/search?query=Lactobacillus+casei>. The search results are categorized into several sections:

- Experiment (10 results found)**:
 - DRX002585 Ion Torrent PGM sequencing; Whole genome analysis of Lactobacillus casei JCM 1134
 - View all 10 results
- Run (6 results found)**:
 - DRR003254 Ion Torrent PGM sequencing; Whole genome analysis of Lactobacillus casei JCM 1134
 - View all 6 results
- Study (10 results found)**:
 - DRP000852 Genome sequencing of Lactobacillus casei JCM 1134
 - View all 10 results
- Study (Sequence) (55 results found)**:
 - PRJNA165565 Lactobacillus casei UW4 Genome sequencing
 - View all 55 results
- Sample (21 results found)**

Annotations in the image include a red arrow pointing to the DRP000852 entry and two black arrows pointing from the yellow text box to the DRX002585 and DRR003254 entries.

W20-2

DRP, ERP, SRPのように、3文字目がPになっているIDをクリックすると全体像が分かりやすいのでよい。DRP000852を眺めると、黒矢印部分のDRXやDRRはDRPの下階層のIDであることがわかる。

Navigation Read Files

Download files

View: TEXT Download: TEXT

Select columns

Showing results 1 - 1 of 1 results

Study accession	Secondary study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	Fastq files (ftp)	Fastq files (galaxy)
PRJDB638	DRP000852	SAMD00015724	DRS002689	DRX002585	DRR003254	1423732	Lactobacillus casei DSM 20011 = JCM 1134	Ion Torrent PGM	SINGLE	File 1	File 1

For Aspera download, please [download and install Aspera Connect](#)

W20-3

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

DDBJ
DNA Data Bank of Japan

Login & Submit | Databases ▾ | English | Contact

Google™ カスタム検索

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

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

2 検索 DRAsearch

DRA Handbook

DDBJ slideshare

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

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

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

DRAにて「Lactobacillus casei」のデータを検索。ヒット数は2。

W20-3

The screenshot shows the DRASearch web interface. The search criteria are: Organism: Lactobacillus casei, StudyType: (empty), CenterName: (empty), Platform: (empty). The search results show 2 studies. A red arrow points to the 'Search' button.

Search Results (2 studies)

#	STUDY	SUBMISSION	STUDY_TITLE	STUDY_TYPE	ORGANISM	BASES	SU
					Atopobium minutum		
					Atopobium rimae		
					Carnobacterium divergens		
					Carnobacterium maltaromaticum		
					Fructobacillus fructosus		
					Kandleria vitulina		
					Lactobacillus		
					Lactobacillus acetotolerans		
					Lactobacillus acidifarinae		

W20-4

OrganismではなくKeywordで検索するとヒット数は激増。
→検索する場所に気を付けよう。

Accession :

Organism : StudyType :

CenterName : Platform :

Keyword :

Show records Sort by

Search Results (1152 records) / 58 Page

Filtered by
document type:sample(614) study(256) experiment(172) submission(81) run(29)
organism:Lactobacillus fermentum(61) Lactobacillus gasseri(60) Lactobacillus rhamnosus(55) Lactobacillus sp.(48)
Lactobacillus salivarius(28) Lactobacillus casei(18)

#	META_FILE	ACCESSION	STUDY	STUDY_TITLE	STUDY_TYPE	ORGANISM	BASES	SUBMITTED	CENTER_NAME
1	ERA148114.study.xml <_TITLE>Draft complete genome sequence of <i>Lactobacillus casei</i> W56</STUDY_TITLE> <STUDY_TYPE existing	ERP007769	ERP007769	Draft complete genome sequence of Lactobacillus casei W56	Whole Genome Sequencing		0		University of Vienna
2	DRA000821.sample.xml 00015724</PRIMARY_ID> </IDENTIFIERS> <TITLE>Whole genome analysis of <i>Lactobacillus casei</i>	DRS002689	DRP000852			Lactobacillus casei DSM 20011 = JCM 1134	161.7M	2012-10-24	
3	SRA092480.study.xml > <DESCRIPTOR> <STUDY_TITLE>Strain <i>Lactobacillus casei</i> TSDC10.1-1.2 (species <i>Lactobacillus casei</i>) was isolated from	SRP026717							

W20-4

Organism列のところが「Lactobacillus casei 12A」となっているのを発見。

Result ID	Study/Sample ID	Accession	Project	Analysis Type	Organism	Size	Date	Source
7	DRA000956.study.xml	DRP000994	DRP000994		Lactobacillus paracasei	74.1M	2013-03-22	
8	DRA000955.study.xml	DRP000993	DRP000993		Lactobacillus paracasei subsp. paracasei ATCC 25302	83.1M	2013-03-22	
9	SRA061483.study.xml	SRP017156	SRP017156	Lactobacillus casei 12A Project	Lactobacillus casei 12A	53.9G	2012-11-12	JGI
10	SRA061478.study.xml	SRP017154	SRP017154	Lactobacillus casei A2-362 Project	Lactobacillus casei A2-362	48.3G	2012-11-12	JGI
11	ERA297098.sample.xml	ERS429397			Lactobacillus casei			
12	ERA135494.sample.xml	ERS143995			Lactobacillus casei			

W21-1

DRAにて、Organism欄で「Lactobacillus casei 12A」を検索。ヒット数は1。

The screenshot shows the DRASearch web interface. The search criteria are: Organism: Lactobacillus casei 12A, StudyType: (empty), CenterName: (empty), Platform: (empty), Keyword: (empty). The search results show 1 study. The search button is highlighted with a red arrow labeled '2'. The first result is highlighted with a red arrow labeled '3'.

Accession :

Organism : StudyType :

CenterName : Platform :

Keyword :

Show records Sort by

Search Results (1 studies) / 1 Page

#	STUDY	SUBMISSION	STUDY_TITLE	STUDY_TYPE	ORGANISM	BASES	SUBMITTED	CENTER_NAME
1	SRP017156	SRA061483	Lactobacillus casei 12A Project	Transcriptome Analysis	Lactobacillus casei 12A	53.9G	2012-11-12	JGI

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

W21-1

実際にダウンロードするのは「SRRから始まるID」のファイル。DRAは、FASTQ形式のbzip2圧縮ファイルとsra形式ファイルの両方を提供している。SRXから始まるIDを眺めることで、用いた実験機器やpaired or single-endなどの情報が分かる。

The screenshot shows the DRASearch interface for study SRP017156. The browser address bar shows the URL: <http://trace.ddbj.nig.ac.jp/DRASearch/study?acc=SRP017156>. The page title is "DRASearch". Navigation links include "Send Feedback", "Search Home", and "DRA Home".

SRP017156

Study Detail	
Title	Lactobacillus casei 12A Project
Study Type	Transcriptome Analysis
Abstract	none provided
Description	none provided
Center Name	JGI

Navigation

- Submission [SRA061483](#) FTP
- Experiment [SRX204226](#) FASTQ SI
- Experiment [SRX204227](#) FASTQ SI

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

W21-1

SRX204226の実験デザインで得られたデータ(Run)は、
①SRR616268であり、②pairedデータであることがわかる。
FASTQ形式でダウンロードする際に、*_1.fastq.bz2と
*_2.fastq.bz2の2つに分かれている事実からも認識可能。
公共データをダウンロードする際には③のところの記述
内容などを眺めて、RNA鎖の方向性を考慮(stranded)し
ているかどうかをちゃんと認識しておく必要がある。

DRASearch

SRX204226 FASTQ SRA

Experiment Detail	
Title	Lactobacillus casei 12A library
Design Description	Illumina Stranded Std PE (amplified)
Organism	Lactobacillus casei 12A

Library Description	
Name	COHG
Strategy	EST
Source	TRANSCRIPTOMIC
Selection	cDNA
Layout	PAIRED
Orientation	
Nominal Length	300
Nominal Sdev	
Construction Protocol	

Platform	
Platform	ILLUMINA
Instrument Model	Illumina HiSeq 2000

Processing	

Navigation	
Submission	SRA061483 FTP
Study	SRP017156
Sample	SRS375150
Run	SRR616268 FASTQ SRA

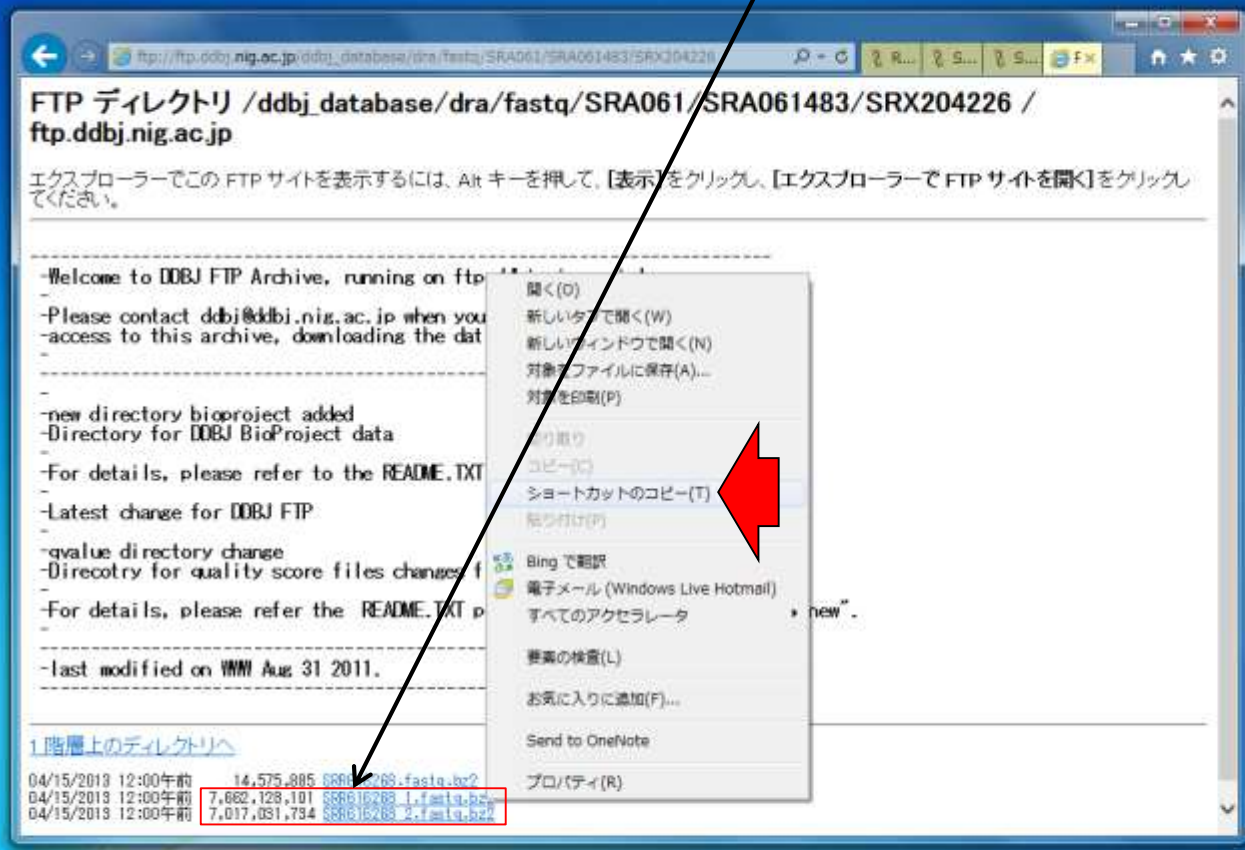
①

②

③

W21-2

ダウンロードしたいFASTQ形式のbzip2圧縮ファイルはこの2つ。それぞれのファイル上で、右クリックの「ショートカットのコピー」でURL情報を取得可能。



W21-3

ENAにて「Lactobacillus casei 12A」
のデータを検索。

The screenshot shows the ENA website interface. At the top, there is a navigation bar with 'EMBL-EBI' on the left and 'Services', 'Research', 'Training', and 'About us' on the right. Below this is a teal banner with the ENA logo and the text 'European Nucleotide Archive'. A search bar is located on the right side of the banner, containing the text 'Lactobacillus casei 12A'. Below the search bar, there are links for 'Advanced' and 'Sequence'. A red arrow points to the 'Search' button. Below the banner is a navigation bar with 'Home', 'Search & Browse', 'Submit & Update', 'Software', 'About ENA', and 'Support'. The main content area is divided into two columns. The left column has a heading 'European Nucleotide Archive' and a paragraph describing the archive. The right column has a heading 'Popular' and a list of links. At the bottom, there is a 'Text Search' box and a 'Latest ENA News' section.

EMBL-EBI

Services Research Training About us

EMBL-EBI ENA European Nucleotide Archive

Lactobacillus casei 12A

Examples: BN000065, histone

Search

Advanced Sequence

Home Search & Browse Submit & Update Software About ENA Support

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. [More about ENA](#)

Access to ENA data is provided though the browser, through search tools, large scale file download and through the API.

Text Search

Examples: BN000065, histone

Popular

- Submit and update
- Sequence submissions
- Genome assembly submissions
- Submitting environmental sequences
- Citing ENA data
- Rest URLs for data retrieval
- Rest URLs to search ENA

Latest ENA News

9 Dec 2014: [ENA Release 122](#)
Release 122 of ENA's

W21-3

「Lactobacillus casei 12A」検索結果。ページ下部に移動するとDRAで見たものと同じIDを発見できる。

The screenshot shows a web browser window with the URL <http://www.ebi.ac.uk/ena/data/search?query=Lactobacillus+casei+12A>. The page header includes the EMBL-ENA logo and navigation links for Services, Research, Training, and About us. A search bar contains the query "Lactobacillus casei 12A" with a "Search" button and links for "Advanced" and "Sequence" search options. Below the search bar is a navigation menu with "Home", "Search & Browse", "Submit & Update", "About ENA", and "Support". A blue banner below the menu contains a message: "Please subscribe to ena-announce mailing list here: listserv.ebi.ac.uk/mailman/listin... to receive alerts about ENA services." The main content area displays "Search results for *Lactobacillus casei* 12A" and lists results in four categories: Assembly (2 results found), Sequence (Release) (9 results found), Coding (Release) (2,683), and Non-coding (Release) (74). The first result under Assembly is "GCA_000309565 Comparative genomics of Lactobacillus casei" with a "View all 2 results" link. The first result under Sequence (Release) is "CP006690 Lactobacillus casei 12A, complete genome." with a "View all 9 results" link. A red arrow points to the bottom of the page, indicating where to scroll to find DRA results.

W21-3

SRP, SRA, SRX, SRRなど沢山のIDが存在するが、SRPというのが一番の大元に近いID。

The screenshot shows a web browser window with the URL <http://www.ebi.ac.uk/ena/data/search?query=Lactobacillus+casei+12A>. The search results are categorized into several sections:

- Non-coding (Release) (74 results found)**
CP006690.1:1654165..1654238:tRNA Lactobacillus casei 12A tRNA-Arg
View all 74 results
- Experiment (2 results found)**
SRX204226 Illumina HiSeq 2000 paired end sequencing; Lactobacillus casei 12A library
View all 2 results
- Run (2 results found)**
SRR616268 Illumina HiSeq 2000 paired end sequencing; Lactobacillus casei 12A library
View all 2 results
- Study (1 results found)**
SRP017156 Lactobacillus casei 12A Project
View all 1 results
- Study (Sequence) (2 results found)**
PRJNA217366 Lactobacillus casei 12A Genome sequencing
View all 2 results
- Sample (1 results found)**
SRS375150 Lactobacillus casei 12A
View all 1 results
- Taxon (1 results found)**

A red arrow points to the 'Study (1 results found)' section, highlighting the SRP017156 entry.

W21-3

The screenshot shows a web browser window with the URL <http://www.ebi.ac.uk/ena/data/view/SRP017156>. The page features a dark header with the EMBL-EBI logo and navigation links for Services, Research, Training, and About us. Below the header is a teal banner with the ENA logo and a search bar. A red arrow points to the scrollbar on the right side of the browser window.

Cookies on EMBL-EBI website
This website uses cookies to store a small amount of information on your computer, as part of the functioning of the site. Cookies used for the operation of the site have already been set. To find out more about the cookies we use and how to delete them, see our [Cookie and Privacy statements](#).
[Dismiss this notice](#)

EMBL-EBI [Services](#) [Research](#) [Training](#) [About us](#)

ENA
European Nucleotide Archive

Search [Advanced](#) [Sequence](#)

Examples: [BN000085](#), [histone](#)

[Home](#) [Search & Browse](#) [Submit & Update](#) [About ENA](#) [Support](#)

[Please subscribe to ena-announce mailing list here: \[listserv.ebi.ac.uk/mailman/listin...\]\(mailto:listserv.ebi.ac.uk/mailman/listin...\) to receive alerts about ENA services.](#)

Study: SRP017156
Lactobacillus casei 12A Project

[Send Feedback](#) [Download: XML](#)

View: [XML](#)

Submitting Centre	Study Type	Read Count	Base Count
JGI	Transcriptome Analysis	269,298,092	53,859,618,400

Secondary accession(s)
[PRJNA217366](#)

Broker Name

W21-3

SRP, SRA, SRX, SRRなど様々な種類のIDの対応関係を俯瞰できるのがENAの長所。

The screenshot shows the ENA data view page for SRP017156. The page includes sections for Description, Abstract, and Navigation. Below the navigation, there are options for Download files, View (TEXT), and Download (TEXT). A table shows the results, with a red box highlighting the first two rows. The table has columns for Study accession, Secondary study accession, Sample accession, Secondary sample accession, Experiment accession, Run accession, Tax ID, Scientific name, Instrument model, Library layout, Fastq files (ftp), and Fastq files (galaxy).

Study accession	Secondary study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	Fastq files (ftp)	Fastq files (galaxy)
PRJNA217366	SRP017156	SAMN01813467	SRS375150	SRX204226	SRR616268	1051650	Lactobacillus casei 12A	Illumina HiSeq 2000	PAIRED	File 1 File 2	File 1 File 2
PRJNA217366	SRP017156	SAMN01813467	SRS375150	SRX204227	SRR616269	1051650	Lactobacillus casei 12A	Illumina HiSeq 2000	PAIRED	File 1 File 2	File 1 File 2

For Aspera download, please download and install Aspera Connect

W21-3

ENAは、FASTQ形式のgzip圧縮ファイルとして提供している。

ENA data view page for SRP017156. The page displays a table of sequencing results. A red arrow points to the 'File 1' and 'File 2' links in the first row. A black arrow points from the text box above to the 'File 1' link. At the bottom, a text box contains the FTP URL for the first file.

Study accession	Secondary study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	Fastq files (ftp)	Fastq files (galaxy)
PRJNA217366	SRP017156	SAMN01813467	SRS375150	SRX204226	SRR616268	1051650	Lactobacillus casei 12A	Illumina HiSeq 2000	PAIRED	File 1 File 2	1 2
PRJNA217366	SRP017156	SAMN01813467	SRS375150	SRX204227	SRR616269	1051650	Lactobacillus casei 12A	Illumina HiSeq 2000	PAIRED	File 1 File 2	File 1 File 2

For Aspera download, please download and install Aspera Connect

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR616/SRR616268/SRR616268_1.fastq.gz

W22-1

①srp017156ディレクトリの作成。ここでは②
/home/iu/Documents/srp017156で作業を行う。

```
iu@bielinux[hoge] cd [ 5:54午後 ]
iu@bielinux[iu] pwd [ 5:55午後 ]
/home/iu
iu@bielinux[iu] ls [ 5:55午後 ]
Desktop Documents Downloads Music Pictures Public Templates Videos
iu@bielinux[iu] cd Documents [ 5:55午後 ]
iu@bielinux[Documents] ls [ 5:55午後 ]
VirtualBox Dropped Files
iu@bielinux[Documents] pwd [ 5:55午後 ]
/home/iu/Documents
① iu@bielinux[Documents] mkdir srp017156 [ 5:55午後 ]
iu@bielinux[Documents] ls -la [ 5:55午後 ]
total 16
drwxr-xr-x 4 iu iu 4096 12月 10 17:55 .
drwxr-xr-x 18 iu iu 4096 12月 10 15:37 ..
drwxrwxr-x 2 iu iu 4096 12月 10 17:55 srp017156
drwx----- 12 iu iu 4096 12月 1 23:13 VirtualBox Dropped Files
iu@bielinux[Documents] cd srp017156 [ 5:55午後 ]
iu@bielinux[srp017156] pwd [ 5:55午後 ]
/home/iu/Documents/srp017156
② iu@bielinux[srp017156] █ [ 5:55午後 ]
```

W22-2

①wgetコマンドを利用してSRR616268_1.fastq.bz2のダウンロードを実行。②DRAはチェックサム情報はないが、ファイルサイズが同じことから正しくダウンロードできたと判断。

```
iu@bielinux[srp017156] wget -c ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_1.fastq.bz2
--2014-12-11 14:25:36-- ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_1.fastq.bz2
=> 'SRR616268_1.fastq.bz2'
Resolving ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)... 133.39.224.12
Connecting to ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)|133.39.224.12|:21... connected.
Logging in as anonymous ... Logged in!
==> SYST ... done.      ==> PWD ... done.
==> TYPE I ... done.    ==> CWD (1) /ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226 ... done.
==> SIZE SRR616268_1.fastq.bz2 ... 7662128101
==> PASV ... done.      ==> RETR SRR616268_1.fastq.bz2 ... done.
Length: 7662128101 (7.1G) (unauthoritative)

100%[=====] 7,662,128,101 3.70MB/s  in 25m 53s
2014-12-11 14:51:30 (4.70 MB/s) - 'SRR616268_1.fastq.bz2' saved [7662128101]

iu@bielinux[srp017156] ls -la
total 7482560
drwxrwxr-x 2 iu iu      4096 12月 11 14:25 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 7662128101 12月 11 14:51 SRR616268_1.fastq.bz2
iu@bielinux[srp017156] █
```

04/15/2013 12:00午前	14,575,885	SRR616268.fastq.bz2
04/15/2013 12:00午前	7,662,128,101	SRR616268_1.fastq.bz2
04/15/2013 12:00午前	7,017,031,734	SRR616268_2.fastq.bz2

W22-2

7.7GBのダウンロードに約26分かかっていることがわかる。
これは東大有線LAN環境での結果であり、ヒトそれぞれ。

```
iu@bielinux[srp017156] wget -c ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_1.fastq.bz2
--2014-12-11 14:25:36-- ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_1.fastq.bz2
=> 'SRR616268_1.fastq.bz2'
Resolving ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)... 133.39.224.12
Connecting to ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)|133.39.224.12|:21... connected.
Logging in as anonymous ... Logged in!
==> SYST ... done.      ==> PWD ... done.
==> TYPE I ... done.    ==> CWD (1) /ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226 ... done.
==> SIZE SRR616268_1.fastq.bz2 ... 7662128101
==> PASV ... done.     ==> RETR SRR616268_1.fastq.bz2 ... done.
Length: 7662128101 (7.1G) (unauthoritative)

100%[=====>] 7,662,128,101 3.70MB/s  in 25m 53s

2014-12-11 14:51:30 (4.70 MB/s) - 'SRR616268_1.fastq.bz2' saved [7662128101]

iu@bielinux[srp017156] ls -la
total 7482560
drwxrwxr-x 2 iu iu      4096 12月 11 14:25 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 7662128101 12月 11 14:51 SRR616268_1.fastq.bz2
iu@bielinux[srp017156] █
```

04/15/2013 12:00午前	14,575,885	SRR616268.fastq.bz2
04/15/2013 12:00午前	7,662,128,101	SRR616268_1.fastq.bz2
04/15/2013 12:00午前	7,017,031,734	SRR616268_2.fastq.bz2

W22-2

wgetの場合はダウンロードに要した時間が表示されるが、他の一般的なコマンドは実行時間は表示されない。実行時間を計測したい場合は、コマンドの左側にtimeをつける。

```
iu@bielinux[srp017156] time wget -c ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_2.fastq.bz2
--2014-12-12 14:38:34-- ftp://ftp.ddbj.nig.ac.jp/ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226/SRR616268_2.fastq.bz2
=> 'SRR616268_2.fastq.bz2'
Resolving ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)... 133.39.224.12
Connecting to ftp.ddbj.nig.ac.jp (ftp.ddbj.nig.ac.jp)|133.39.224.12|:21... connected.
Logging in as anonymous ... Logged in!
==> SYST ... done.      ==> PWD ... done.
==> TYPE I ... done.    ==> CWD (1) /ddbj_database/dra/fastq/SRA061/SRA061483/SRX204226 ... done.
==> SIZE SRR616268_2.fastq.bz2 ... 7017031734
==> PASV ... done.     ==> RETR SRR616268_2.fastq.bz2 ... done.
Length: 7017031734 (6.5G) (unauthoritative)

100%[=====>] 7,017,031,734 7.58MB/s  in 21m 20s

2014-12-12 14:59:55 (5.23 MB/s) - 'SRR616268_2.fastq.bz2' saved [7017031734]

wget -c 14.35s user 131.02s system 11% cpu 21:21.16 total
iu@bielinux[srp017156] ls -la
total 14335136
drwxrwxr-x 2 iu iu      4096 12月 12 14:38 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 7662128101 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 7017031734 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156]
```

[2:59午後]

[3:11午後]

W22-3

②bzip2圧縮ファイルの解凍を行っている。約90分かかっている。①と③で解凍前後に「df -h」を実行し、全体のディスク使用量の推移もチェックしている。

```
iu@bielinux[srp017156] ls -la
total 14335136
drwxrwxr-x 2 iu iu      4096 12月 12 14:38 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 7662128101 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 7017031734 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G  23G  116G  17% /
none            4.0K    0  4.0K   0% /sys/fs/cgroup
udev            991M  4.0K  991M   1% /dev
tmpfs           201M  920K  200M   1% /run
none            5.0M    0  5.0M   0% /run/lock
none           1002M  156K  1001M   1% /run/shm
none            100M   64K  100M   1% /run/user
iu@bielinux[srp017156] bzip2 -d SRR616268_1.fastq.bz2
iu@bielinux[srp017156] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G  57G   82G  41% /
none            4.0K    0  4.0K   0% /sys/fs/cgroup
udev            991M  4.0K  991M   1% /dev
tmpfs           201M  920K  200M   1% /run
none            5.0M    0  5.0M   0% /run/lock
none           1002M  156K  1001M   1% /run/shm
none            100M   64K  100M   1% /run/user
iu@bielinux[srp017156]
```

[2:59午後]

[3:11午後]

[3:24午後]

[4:55午後]

[5:03午後]

W22-3

たった1つのファイルを解凍しただけで、一気にディスク使用量が23GBから57GBに激増していることが分かる。

```
iu@bielinux[srp017156] ls -la
total 14335136
drwxrwxr-x 2 iu iu      4096 12月 12 14:38 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 7662128101 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 7017031734 12月 12 14:59 SRR616268_2.fastq.bz2

iu@bielinux[srp017156] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G  23G  116G  17% /
none            4.0K   0  4.0K   0% /sys/fs/cgroup
udev            991M  4.0K  991M   1% /dev
tmpfs           201M  920K  200M   1% /run
none            5.0M   0  5.0M   0% /run/lock
none           1002M  156K  1001M   1% /run/shm
none            100M   64K  100M   1% /run/user

iu@bielinux[srp017156] bzip2 -d SRR616268_1.fastq.bz2
iu@bielinux[srp017156] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G  57G  82G  41% /
none            4.0K   0  4.0K   0% /sys/fs/cgroup
udev            991M  4.0K  991M   1% /dev
tmpfs           201M  920K  200M   1% /run
none            5.0M   0  5.0M   0% /run/lock
none           1002M  156K  1001M   1% /run/shm
none            100M   64K  100M   1% /run/user

iu@bielinux[srp017156]
```

[2:59午後]

[3:11午後]

[3:24午後]

[4:55午後]

[5:03午後]

W22-3

解凍後のファイルサイズは7GB程度から40GB以上に膨れ上がる。

```
iu@bielinux[srp017156] bzip2 -d SRR616268_1.fastq.bz2 [ 3:24午後 ]
iu@bielinux[srp017156] df -h [ 4:55午後 ]
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G   57G   82G   41% /
none            4.0K    0  4.0K    0% /sys/fs/cgroup
udev            991M   4.0K  991M    1% /dev
tmpfs           201M  920K  200M    1% /run
none            5.0M    0   5.0M    0% /run/lock
none           1002M  156K 1001M    1% /run/shm
none            100M   64K  100M    1% /run/user
iu@bielinux[srp017156] ls -la [ 5:03午後 ]
total 49811440
drwxrwxr-x 2 iu iu      4096 12月 12 16:55 .
drwxr-xr-x 4 iu iu      4096 12月 10 17:55 ..
-rw-rw-r-- 1 iu iu 43989863596 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu  7017031734 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] ls -lh [ 5:14午後 ]
total 48G
-rw-rw-r-- 1 iu iu  41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu  6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] [ 5:30午後 ]
```

W22-4

②2つめのbzip2圧縮ファイルを解凍。①と③の解凍前後で6.6GBから38GBに激増していることがわかる。
④全体のディスク容量も146GBのうち、使用容量が87GB、未使用が52GBとなっていることがわかる。

```
iu@bielinux[srp017156] ls -lh
total 48G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] bunzip2 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] ls -lh
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] df -h
df: 'mac_share': No such file or directory
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       146G   87G   52G   63% /
none            4.0K    0 4.0K   0% /sys/fs/cgroup
udev            991M   4.0K 991M   1% /dev
tmpfs           201M  916K 200M   1% /run
none            5.0M    0 5.0M   0% /run/lock
none            1002M  152K 1001M   1% /run/shm
none            100M   52K  100M   1% /run/user
iu@bielinux[srp017156]
```

[1:32午後]
[1:32午後]
[2:08午後]
[4:07午後]
[4:08午後]

W22-5

bzip2の圧縮パフォーマンス。①41GBが7.2GB、②38GBが6.6GB。処理時間は、それぞれ94分と75分。

```
iu@bielinux[srp017156] pwd [ 5:22午後 ]
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -lh [ 5:23午後 ]
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] time bzip2 SRR616268_1.fastq [ 5:23午後 ]
bzip2 SRR616268_1.fastq 4192.25s user 344.21s system 80% cpu 1:34:10.20 total
iu@bielinux[srp017156] time bzip2 SRR616268_2.fastq [ 6:57午後 ]
bzip2 SRR616268_2.fastq 3569.87s user 294.51s system 86% cpu 1:14:40.80 total
iu@bielinux[srp017156] ls -lh [ 8:12午後 ]
total 14G
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] [ 8:12午後 ]
```

W22-5

bzip2のパフォーマンス。7.2GBの解凍に37分。6.6GBの解凍に34分。

```
iu@bielinux[srp017156] ls -lh [ 1:29午前 ]
total 14G
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] [ 1:29午前 ]
iu@bielinux[srp017156] time bunzip2 SRR616268_1.fastq.bz2 [ 1:29午前 ]
bunzip2 SRR616268_1.fastq.bz2 1395.33s user 145.83s system 68% cpu 37
:14.82 total
iu@bielinux[srp017156] time bunzip2 SRR616268_2.fastq.bz2 [ 2:06午前 ]
bunzip2 SRR616268_2.fastq.bz2 1246.13s user 131.68s system 67% cpu 33
:49.93 total
iu@bielinux[srp017156] ls -lh [ 2:40午前 ]
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] [ 2:40午前 ]
```

W22-6

gzipの圧縮パフォーマンス。41GBが9.5GB、38GBが8.8GB。処理時間は、それぞれ110分と103分。一般に、gzipの圧縮率はbzip2に比べて劣るので合理的。しかし実行時間も遅かったのは想定外。たまたまかもしれない。

```
iu@bielinux[srp017156] ls -lh [ 2:40午前 ]
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] [ 2:40午前 ]
iu@bielinux[srp017156] time gzip SRR616268_1.fastq [ 2:40午前 ]
gzip SRR616268_1.fastq 5290.42s user 429.16s system 86% cpu 1:49:56.3
3 total
iu@bielinux[srp017156] time gzip SRR616268_2.fastq [ 4:30午前 ]
gzip SRR616268_2.fastq 4809.24s user 393.25s system 83% cpu 1:43:22.2
9 total
iu@bielinux[srp017156] ls -lh [ 6:13午前 ]
total 19G
-rw-rw-r-- 1 iu iu 9.5G 12月 11 14:51 SRR616268_1.fastq.gz
-rw-rw-r-- 1 iu iu 8.8G 12月 12 14:59 SRR616268_2.fastq.gz
iu@bielinux[srp017156] [ 6:13午前 ]
```

W22-6

gzipのパフォーマンス。9.5GBの解凍に24分。8.8GBの解凍に17分。解凍の処理速度は、bzip2よりも明らかに速い。

```
iu@bielinux[srp017156] ls -lh
total 19G
-rw-rw-r-- 1 iu iu 9.5G 12月 11 14:51 SRR616268_1.fastq.gz
-rw-rw-r-- 1 iu iu 8.8G 12月 12 14:59 SRR616268_2.fastq.gz
iu@bielinux[srp017156] time gunzip SRR616268_1.fastq.gz
gunzip SRR616268_1.fastq.gz 311.53s user 63.26s system 26% cpu 23:39.92 total
iu@bielinux[srp017156] time gunzip SRR616268_2.fastq.gz
gunzip SRR616268_2.fastq.gz 287.09s user 57.63s system 33% cpu 17:23.14 total
iu@bielinux[srp017156] ls -lh
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156]
```


W23-1

ペアエンドリードなので生データの行数は同じはず。①と②でそれぞれ独立にwcコマンドで行数を調べている。2ファイルともに539,023,984行という結果を得て安心。数が違っていると何かがおかしいので、ちゃんとダウンロードできているのかななどを様々な角度から検証すべし。

```
iu@bielinux[~/Documents/srp017156] [ 7:00午後 ]
iu@bielinux[srp017156] date [ 7:00午後 ]
2014年 12月 13日 土曜日 19:00:44 JST
iu@bielinux[srp017156] wc SRR616268_1.fastq [ 7:00午後 ]
539023984 ← 1078047968 43989863596 SRR616268_1.fastq
iu@bielinux[srp017156] wc SRR616268_2.fastq [ 7:13午後 ]
539023984 ← 1078047968 39947183716 SRR616268_2.fastq
iu@bielinux[srp017156] wc SRR616268_1.fastq SRR616268_2.fastq [ 7:24午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] wc *.fastq [ 7:50午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] time wc *.fastq [ 8:15午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
wc *.fastq 1076.83s user 115.98s system 79% cpu 24:58.14 total
iu@bielinux[srp017156] date [ 8:40午後 ]
2014年 12月 13日 土曜日 20:40:08 JST
iu@bielinux[srp017156] [ 8:41午後 ]
```

W23-1

①wcコマンドのあとに複数のファイル名を与えるやり方。
②ファイル名の最後が「.fastq」で終わるファイル全てに対してwcコマンドを実行せよという命令。この場合、ペアエンドの2つのファイルしかないので実質的には①と同じ結果が得られるが、カレントディレクトリ中に10個の.fastqで終わるファイルが存在するときは10個分実行される。

```
iu@bielinux[~/Documents/srp017156]
iu@bielinux[srp017156] date
2014年 12月 13日 土曜日 19:00:44 JST
iu@bielinux[srp017156] wc SRR616268_1.fastq [ 7:00午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
iu@bielinux[srp017156] wc SRR616268_2.fastq [ 7:13午後 ]
539023984 1078047968 39947183716 SRR616268_2.fastq
① iu@bielinux[srp017156] wc SRR616268_1.fastq SRR616268_2.fastq [ 7:24午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
② iu@bielinux[srp017156] wc *.fastq [ 7:50午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] time wc *.fastq [ 8:15午後 ]
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
wc *.fastq 1076.83s user 115.98s system 79% cpu 24:58.14 total
iu@bielinux[srp017156] date [ 8:40午後 ]
2014年 12月 13日 土曜日 20:40:08 JST
iu@bielinux[srp017156] [ 8:41午後 ]
```

W23-2

約5.4億行のファイルだとwc実行にも一苦労。赤枠部分を眺めるだけでも2つのファイルを実行するのに25分程度かかっていることがわかるが、timeコマンドで実行時間を計測するのが王道。

```
iu@bielinux[~/Documents/srp017156]
iu@bielinux[srp017156] date
2014年 12月 13日 土曜日 19:00:44 JST
iu@bielinux[srp017156] wc SRR616268_1.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
iu@bielinux[srp017156] wc SRR616268_2.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
iu@bielinux[srp017156] wc SRR616268_1.fastq SRR616268_2.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] wc *.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] time wc *.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
wc *.fastq 1076.83s user 115.98s system 79% cpu 24:58.14 total
iu@bielinux[srp017156] date
2014年 12月 13日 土曜日 20:40:08 JST
iu@bielinux[srp017156]
```

[7:00午後]
[7:00午後]
[7:13午後]
[7:24午後]
[7:50午後]
[8:15午後]
[8:40午後]
[8:41午後]

W23-2

NGSデータ解析は、このように各コマンドの実行が十数分から数時間にも及ぶ。そのため、一つのコマンド実行が終わるのを待ち構えて、すぐに次のコマンドを打ち込んでいる...わけではない。ここでは、ホストOS上でEmEditorのようなテキストエディタに一連のスクリプトを書きこんでおき、ターミナル上でコピペで実行している。尚、コピペは「CTRL + CおよびV」ではなく右クリックでやるべし。

```
iu@bielinux[~/Documents/srp017156]
iu@bielinux[srp017156] date
2014年 12月 13日 土曜日 19:00:44
iu@bielinux[srp017156] wc SRR616268_1.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
iu@bielinux[srp017156] wc SRR616268_2.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
iu@bielinux[srp017156] wc SRR616268_1.fastq SRR616268_2.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] wc *.fastq
539023984 1078047968 43989863596 SRR616268_1.fastq
539023984 1078047968 39947183716 SRR616268_2.fastq
1078047968 2156095936 83937047312 total
iu@bielinux[srp017156] time wc *.fastq
wc *.fastq 1076.83s user 115.98s system
iu@bielinux[srp017156] date
2014年 12月 13日 土曜日 20:40:08 JST
iu@bielinux[srp017156]
```

```
無題 * - EmEditor
ファイル(E) 編集(E) 検索(S) 表示(V) ツール(I) ウ-
無題 * 無題 * x
date↓
wc SRR616268_1.fastq↓
wc SRR616268_2.fastq↓
wc SRR616268_1.fastq SRR616268_2.fastq↓
wc *.fastq↓
time wc *.fastq↓
date↓
Text 8行, 1桁 日本語 (シフト JIS)
```

[8:41午後]

もちろん、ゲストOS上で新しいターミナルを開き、geditでスクリプトを作成してコピペでもよい。

W23-2

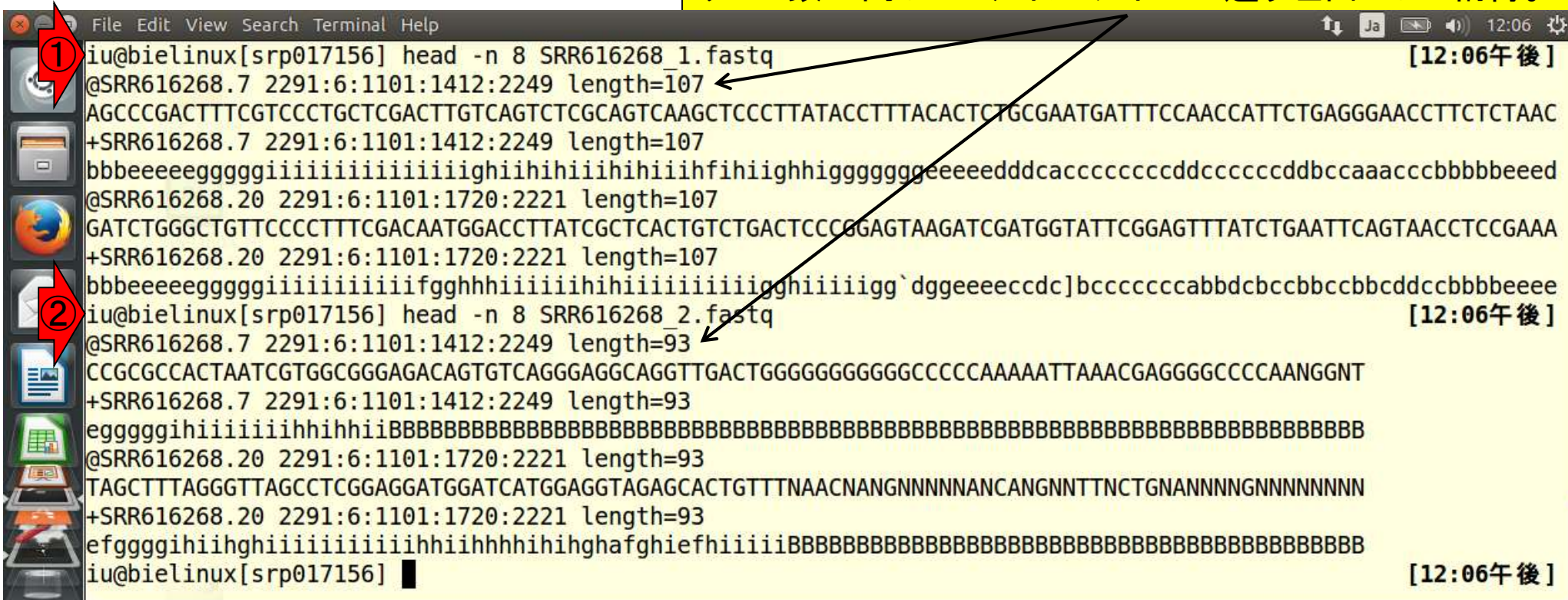
The screenshot displays a Linux desktop environment. On the left is a vertical dock with icons for applications like a terminal, file manager, Firefox, mail, LibreOffice, and a presentation viewer. The main area contains several overlapping windows:

- A terminal window titled `iu@bielinux[~/Documents/srp017156]` showing the execution of `date`, `wc SRR616268_1.fastq`, `wc SRR616268_2.fastq`, and `time wc *.fastq`. The output includes timestamps (e.g., 2014年 12月 13日 土曜日 19:00:44 JST) and file statistics.
- A second terminal window titled `iu@bielinux[~]` is partially visible behind the first.
- A `gedit` window titled `*Untitled Document 1 - gedit` is in the foreground, showing a copy of the terminal output. A context menu is open over the text, with the `Copy` option highlighted by a red arrow.

The system tray at the top right shows the language set to Japanese (Ja), system volume, and the time 21:27.

W24-1

headコマンドで最初の8行分を表示している。リード長は①107塩基と②93塩基で違っていることがわかる。リード数と同じでファイルサイズが違う理由として納得。



```
iu@bielinux[srp017156] head -n 8 SRR616268_1.fastq
@SRR616268.7 2291:6:1101:1412:2249 length=107
AGCCCGACTTTCGTCCCTGCTCGACTTGTCAGTCTCGCAGTCAAGCTCCCTTATACCTTTACTCTGCGAATGATTTCCAACCATTCTGAGGGAACCTTCTCTAAC
+SRR616268.7 2291:6:1101:1412:2249 length=107
bbbeeeeeggggiiiiiiiiiiiiiiiiighiihihihihihihifhiighhiggggggeeeedddcacccccccddccccddbccaaaccbbbbbbeed
@SRR616268.20 2291:6:1101:1720:2221 length=107
GATCTGGGCTGTTCCCTTTCGACAATGGACCTTATCGCTCACTGTCTGACTCCCGAGTAAGATCGATGGTATTCGGAGTTTATCTGAATTCAGTAACCTCCGAAA
+SRR616268.20 2291:6:1101:1720:2221 length=107
bbbeeeeeggggiiiiiiiiiiiiifgghhiiiiiihihihihihihihiigghiiiiigg`dggeeeecdc]ccccccabbdbcbbccbbccddccbbbbeeee
iu@bielinux[srp017156] head -n 8 SRR616268_2.fastq
@SRR616268.7 2291:6:1101:1412:2249 length=93
CCGCGCCACTAATCGTGGCGGGAGACAGTGTGAGGGAGGCAGGTTGACTGGGGGGGGGGCCCCCAAAAATTAACGAGGGGCCCAANGNT
+SRR616268.7 2291:6:1101:1412:2249 length=93
eggggghiiiiiiihhihihiBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
@SRR616268.20 2291:6:1101:1720:2221 length=93
TAGCTTTAGGGTTAGCCTCGGAGGATGGATCATGGAGGTAGAGCACTGTTNAACNANGNNNNNANCANGNNTTCTGNANNNGNNNNNNNN
+SRR616268.20 2291:6:1101:1720:2221 length=93
efgggghiihghiiiiiiiiiiiiihhihhhhiihghafghiefhiiiiBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
iu@bielinux[srp017156]
```

W24-2

tailコマンドで最後の8行分を表示している。黒矢印部分がリードのシリアル番号のようなもの。wc実行結果から134,755,996リードとなっているのになぜ一番最後のリードが135073834となっているのか疑問に思っただけ。

```
iu@bielinux[srp017156] tail -n 8 SRR616268.135073833_1.fastq
@SRR616268.135073833 2291:6:2308:21326:200775 length=107
GCCCCGGTATATTTTCGGTGTGTGCGGCTCGACTTGTGGGGTTTTATGCACTGTTTTAATGGTGGCTGGTTTTGTGGGATGTTTTTTGTTGTCTGTGTTTAGTGAG
+SRR616268.135073833 2291:6:2308:21326:200775 length=107
P\^cJ\`c0[QQ`ddKQbQ[HHP cBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBPJS\cce
@SRR616268.135073834 2291:6:2308:21288:200826 length=107
CGCGTGTGGTGGTGGGTTGTTTTGTGTGGGGTGGATTTTTTTGGTGTGTAGTGGGTTGGGTTATTGGGTATTTTGTGGATCTTTGTTTTTTTGTTTAGGGTG
+SRR616268.135073834 2291:6:2308:21288:200826 length=107
P\\caH0`QSQQR`[PHPHP00^H0^BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBJJPS\\a
iu@bielinux[srp017156] tail -n 8 SRR616268.135073833_2.fastq
@SRR616268.135073833 2291:6:2308:21326:200775 length=93
GGATAAGGTTCTTTGTTGAAAGGGGTACAGTCAAGATCACCAGTTAAGGTCCTAAATTTATGCTAAGTGGAAAAGGATGTGGCGTTGCACAG
+SRR616268.135073833 2291:6:2308:21326:200775 length=93
JQORQJS]bJJJR]bgR[QQ]Q[dgHHPY^OPIY^eghhhhhhghhhfghhiihiifgbghbgfedd`bgeccce`__b`bca_W^bcb^
@SRR616268.135073834 2291:6:2308:21288:200826 length=93
GTTGTGTTGGGATTTGGTGCGGGTGGAGGGTGGGTGTGTTTGGGTTTGGTTGGGGTGGTTGGGTGTGCTTTGGTGGGGAGTTTTTCGGGTT
+SRR616268.135073834 2291:6:2308:21288:200826 length=93
cOQ`HQR`eaBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
iu@bielinux[srp017156] █
```

[12:09午後]

[12:09午後]

DRAは、生データのsraファイルからFASTQファイルを生成する際、クオリティの低いリードなどを除去しているため。
<http://trace.ddbj.nig.ac.jp/dra/faq.html#read-number-fastq>

The screenshot shows a web browser window with the URL `http://trace.ddbj.nig.ac.jp/dra/faq.html#read-number-fastq`. The page content is as follows:

DRA で公開されている fastq のリード数が生データのそれよりも少ないのは何故でしょうか?

DRA では NCBI SRA Toolkit に含まれている `fastq-dump` を使い、以下のオプションで生データである SRA ファイルから fastq ファイルを作成しています。

```
fastq-dump -M 25 -E --skip-technical --split-3 -W <SRA file>
```

- `-M 25`: 25 塩基以上の配列のみを含める。デフォルトは 25。
- `-E`: リードの開始、もしくは終わりに 10 以上の N が存在しない
- `--skip-technical`: technical read を除き biological read のみを出力
- `--split-3`: ペアリードで最初と二番目の biological read をそれぞれ *_1.fastq と *_2.fastq として出力する。一つしか biological read が存在しない場合、*.fastq として出力する。
- `-W`: 指定されていた場合、left と right を clip する

上記の出力条件でリードがフィルタリング、トリミングされるため、一般的に fastq のリード数は SRA ファイルのそれよりも少なくなっています。フィルタリング、トリミングされていない fastq ファイルを得るには以下のコマンドで fastq を生成します。

```
fastq-dump -M 1 --split-3 <SRA file>
```

作成日: 2013年10月8日; 最終更新日: 2014年6月6日

W24-3

DRAは、生データのsraファイルからFASTQファイルを生
成する際、クオリティの低いリードなどを除去しているため。

The image shows two overlapping screenshots of the DRASearch website. The top screenshot displays the study page for SRP017156, and the bottom screenshot displays the experiment page for SRX204226. Red arrows with numbers 1 and 2 point to specific links in the navigation sections.

Top Screenshot: Study Detail

Study Detail	
Title	Lactobacillus casei 12A Project
Study Type	Transcriptome Analysis
Abstract	none provided
Description	
Center	

Navigation

- Submission [SRA061483](#) FTP
- Experiment [SRX204226](#) **1** [FASTQ](#) [SRA](#)
- [SRX204227](#) [FASTQ](#) [SRA](#)

Bottom Screenshot: Experiment Detail

Experiment Detail	
Title	Lactobacillus casei 12A library
Design Description	Illumina Stranded Std PE (amplified)
Organism	Lactobacillus casei 12A

Library Description

Name	COHG
Strategy	EST
Source	TRANSCRIPTOMIC

Navigation

- Submission [SRA061483](#) FTP
- Study [SRP017156](#)
- Sample [SRS375150](#)
- Run [SRR616268](#) **2** [FASTQ](#) [SRA](#)

W24-3

シリアル番号が7のやつはクオリティが高そう。FASTQファイルをheadコマンドで眺めたときに、最初のリードのシリアル番号が7であったのは極めて妥当！結論として、DRAで取り扱うFASTQファイルは生リードよりも若干少なくなるが、事実上気にしなくて良い。

```
http://trace.ddbj.nig.ac.jp/DRASearch/run?acc=SRR616268
>SRR616268.7
AGCCCGACTTTTCGTCCCTGCTCGACTTGTTCAGTCTCGCAGTCAAGCTCCCTTATACCTTTACACTCTGCGAATGATTTCC
AACCAATTCTGAGGGAACCTTCTCTAACCCGCGCCAATAATCGTGGCGGGAGACAGTGTTCAGGGAGGCAGGTTGACTGGGG
GGGGGGGCCCCAAAATTAACGAGGGGCCCAANGGNT
>SRR616268.8
CGCCGTACTCAGGATCCTGG
NNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN
>SRR616268.9
TCCCTTTTCGACAATGGACC
NNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN
>SRR616268.10
GGACTAGGCCAAACAAGAGC
NNNNNNNNNNNTGNNNNNNNN
NNNNNNNNNNNNNNNNNNNN
Copyright©DNA Data Bank of J
```

```
File Edit View Search Terminal Help
iu@bielinux[srp017156] pwd [12:10午前]
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -lh [12:10午前]
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] head -n 8 SRR616268_1.fastq [12:10午前]
@SRR616268.7 2291:6:1101:1412:2249 length=107
AGCCCGACTTTTCGTCCCTGCTCGACTTGTTCAGTCTCGCAGTCAAGCTCCCTTATACCTTTACACTCTGCG
AATGATTTCCAACCATTCTGAGGGAACCTTCTCTAAC
+SRR616268.7 2291:6:1101:1412:2249 length=107
bbbeeeeeggggiiiiiighiihihihihihihfihiighhigggggggeeeedddc
acccccccddccccccdbccaaaccbbbbbbeed
@SRR616268.20 2291:6:1101:1720:2221 length=107
GATCTGGGCTGTTCCCTTTTCGACAATGGACCTTATCGCTCACTGTCTGACTCCCGGAGTAAGATCGATG
GTATTCGGAGTTTATCTGAATTCAGTAACCTCCGAAA
+SRR616268.20 2291:6:1101:1720:2221 length=107
bbbeeeeeggggiiiiiifgghhiiiiihiiiiiiiiiiiigghiiiiigg`dggeeeecdc]
ccccccabbdbccbbccbbccddccbbbbeeee
iu@bielinux[srp017156] [12:10午前]
```

W24-3

シリアル番号が7のやつはクオリティが高そう。FASTQファイルをheadコマンドで眺めたときに、最初のリードのシリアル番号が7であったのは極めて妥当！結論として、DRAで取り扱うFASTQファイルは生リードよりも若干少なくなるが、事実上気にしなくて良い。

```
>SRR616268.7
AGCCCGACTTTTCGTCCCTGCTCGACTTGTGTCAGTCTCGCAGTCAAGCTCCCTTATACCTTTACACTCTGCGAATGATTTCC
AACCAATTCGAGGGAACCTTCTCTAACCCGCGCCACTAATCGTGGCGGGAGACAGTGTTCAGGGAGGCAGGTTGACTGGGG
GGGGGGGCCCCCAAAAATTAACGAGGGGCCCAANGNT

>SRR616268.8
CGCCGTACTCAGGATCCTGG
NNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN

>SRR616268.9
TCCCTTTTCGACAATGGACC
NNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN

>SRR616268.10
GGACTAGGCCAAACAAGAGC
NNNNNNNNNNNTGNNNNNNNN
NNNNNNNNNNNNNNNNNNNN
```

```
iu@bielinux[srp017156] pwd [12:23午前]
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -lh [12:23午前]
total 79G
-rw-rw-r-- 1 iu iu 41G 12月 11 14:51 SRR616268_1.fastq
-rw-rw-r-- 1 iu iu 38G 12月 12 14:59 SRR616268_2.fastq
iu@bielinux[srp017156] head -n 8 SRR616268 2.fastq [12:23午前]
@SRR616268.7 2291:6:1101:1412:2249 length=93
CCGCGCCACTAATCGTGGCGGGAGACAGTGTTCAGGGAGGCAGGTTGACTGGGGGGGGGGGCCCCCAAAA
TTAAACGAGGGGCCCAANGNT
+SRR616268.7 2291:6:1101:1412:2249 length=93
eggggghiiiiiihhihhiBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
BBBBBBBBBBBBBBBBBBBBBBBBBBBB
+SRR616268.20 2291:6:1101:1720:2221 length=93
TAGCTTTAGGGTTAGCCTCGGAGGATGGATCATGGAGGTAGAGCACTGTTTNAACNANGNNNNNANCANG
NNTTCTGNANNNGNNNNNNNN
+SRR616268.20 2291:6:1101:1720:2221 length=93
efgggghihghiiiiiiiiihihhihhihghafghiefhiiiiBBBBBBBBBBBBBBBBBBBB
BBBBBBBBBBBBBBBBBBBBBBBB
iu@bielinux[srp017156] [12:23午前]
```

W25-1

①bzip2とwcをパイプで連結した場合は、コマンドごとに処理時間が出力されるものの、同じ時間が表示されている。つまり、合算したトータルの36分52秒という時間なのだろう。

```
iu@bielinux[srp017156] pwd
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -lh
total 14G
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 | wc
539023984 1078047968 43989863596
bzip2 -d -c SRR616268_1.fastq.bz2 1316.60s user 119.43s system 64% cpu 36:52.07 total
wc 559.21s user 54.61s system 27% cpu 36:52.07 total
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 > hoge_1.fastq
bzip2 -d -c SRR616268_1.fastq.bz2 > hoge_1.fastq 1380.80s user 144.15s system
57% cpu 44:02.97 total
iu@bielinux[srp017156] time wc hoge_1.fastq
539023984 1078047968 43989863596 hoge_1.fastq
wc hoge_1.fastq 562.12s user 59.07s system 75% cpu 13:44.40 total
iu@bielinux[srp017156] ls -lh
total 55G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156]
```



[11:42午後]

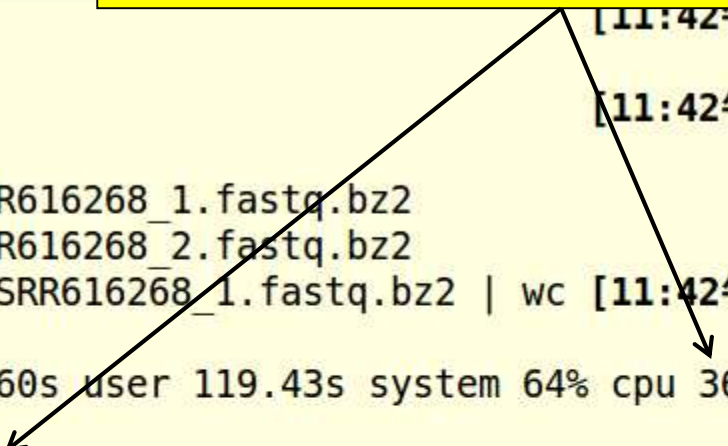
[11:42午後]

[11:42午後]

[1:03午前]

[1:17午前]

[9:41午前]



W25-1

①で-cオプションを入れているので、*.fastq.bz2ファイルは残る。それゆえ、②でもう一度同じ*.fastq.bz2ファイルに対してbzip2コマンドを実行可能。②と③で①と同じことをやっているが、①のパイプで連結する処理時間(約37分)のほうが、中間ファイルhoge_1.fastqを作成しない分だけ早いのだろう。

```
iu@bielinux[srp017156] pwd
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -lh
total 14G
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 | wc [11:42午後]
539023984 1078047968 43989863596
bzip2 -d -c SRR616268_1.fastq.bz2 1316.60s user 119.43s system 64% cpu 36:52.07 total
wc 559.21s user 54.61s system 27% cpu 36:52.07 total
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 > hoge_1.fastq
bzip2 -d -c SRR616268_1.fastq.bz2 > hoge_1.fastq 1380.80s user 144.15s system
57% cpu 44:02.97 total
iu@bielinux[srp017156] time wc hoge_1.fastq [ 1:03午前]
539023984 1078047968 43989863596 hoge_1.fastq
wc hoge_1.fastq 562.12s user 59.07s system 75% cpu 13:44.40 total
iu@bielinux[srp017156] ls -lh [ 1:17午前]
total 55G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] [ 9:41午前]
```

W25-2

①と②は、パイプで連結して最初の400万行分をリダイレクトでファイル保存しているので、十数秒で処理が終わっていることがわかる。

```
iu@bielinux[srp017156] ls -lh
total 55G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
① iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 | head -n 4000000 > subset_1.fastq
bzip2 -d -c SRR616268_1.fastq.bz2 11.17s user 0.97s system 84% cpu 14.294 total
head -n 4000000 > subset_1.fastq 0.33s user 0.46s system 5% cpu 14.293 total
② iu@bielinux[srp017156] time bzip2 -d -c SRR616268_2.fastq.bz2 | head -n 4000000 > subset_2.fastq
bzip2 -d -c SRR616268_2.fastq.bz2 9.70s user 0.87s system 83% cpu 12.623 total
head -n 4000000 > subset_2.fastq 0.30s user 0.42s system 5% cpu 12.622 total
iu@bielinux[srp017156] head -n 4000000 hoge_1.fastq > hoge_subset_1.fastq
iu@bielinux[srp017156] ls -lh
total 56G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 306M 12月 16 10:50 hoge_subset_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
-rw-rw-r-- 1 iu iu 306M 12月 16 10:50 subset_1.fastq
-rw-rw-r-- 1 iu iu 278M 12月 16 10:50 subset_2.fastq
iu@bielinux[srp017156] diff hoge_subset_1.fastq subset_1.fastq
iu@bielinux[srp017156]
```

[10:50午前]

[10:50午前]

[10:50午前]

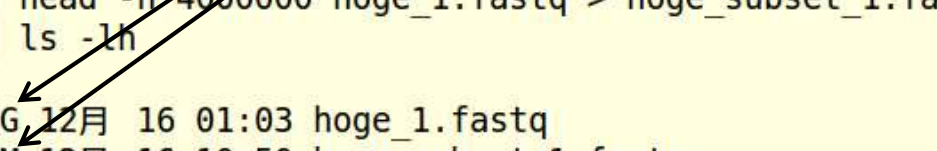
[10:50午前]

[10:53午前]

W25-2

③一旦41GBのhoge_1.fastqという解凍後のファイルを作成したのち、最初の400万行分をリダイレクトでファイル保存したのが306MBのhoge_subset_1.fastq。

```
iu@bielinux[srp017156] ls -lh
total 55G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_1.fastq.bz2 | head -n 4000000 > subset_1.fastq
bzip2 -d -c SRR616268_1.fastq.bz2 11.17s user 0.97s system 84% cpu 14.294 total
head -n 4000000 > subset_1.fastq 0.33s user 0.46s system 5% cpu 14.293 total
iu@bielinux[srp017156] time bzip2 -d -c SRR616268_2.fastq.bz2 | head -n 4000000 > subset_2.fastq
bzip2 -d -c SRR616268_2.fastq.bz2 9.70s user 0.87s system 83% cpu 12.623 total
head -n 4000000 > subset_2.fastq 0.30s user 0.42s system 5% cpu 12.622 total
iu@bielinux[srp017156] head -n 4000000 hoge_1.fastq > hoge_subset_1.fastq
iu@bielinux[srp017156] ls -lh
total 56G
-rw-rw-r-- 1 iu iu 41G 12月 16 01:03 hoge_1.fastq
-rw-rw-r-- 1 iu iu 306M 12月 16 10:50 hoge_subset_1.fastq
-rw-rw-r-- 1 iu iu 7.2G 12月 11 14:51 SRR616268_1.fastq.bz2
-rw-rw-r-- 1 iu iu 6.6G 12月 12 14:59 SRR616268_2.fastq.bz2
-rw-rw-r-- 1 iu iu 306M 12月 16 10:50 subset_1.fastq
-rw-rw-r-- 1 iu iu 278M 12月 16 10:50 subset_2.fastq
iu@bielinux[srp017156] diff hoge_subset_1.fastq subset_1.fastq
iu@bielinux[srp017156]
```



[10:50午前]

[10:50午前]

[10:50午前]

[10:50午前]

[10:53午前]

W25-2

```
iu@bielinux[~/Documents/srp017156] [ 6:46午後 ]
iu@bielinux[srp017156] pwd [ 6:46午後 ]
/home/iu/Documents/srp017156
iu@bielinux[srp017156] ls -l SRR616268_1.fastq.bz2 [ 6:46午後 ]
-rw-rw-r-- 1 iu iu 7662128101 12月 11 2014 SRR616268_1.fastq.bz2
① iu@bielinux[srp017156] bzip2 -d -c SRR616268_1.fastq.bz2 | head -n 4000000
| gzip > subset_1.fastq.gz [ 6:48午後 ]
iu@bielinux[srp017156] ls -l subset_1.fastq.gz [ 6:48午後 ]
-rw-rw-r-- 1 iu iu 74906555 7月 1 18:48 subset_1.fastq.gz
iu@bielinux[srp017156] ls -lh subset_1.fastq.gz [ 6:51午後 ]
-rw-rw-r-- 1 iu iu 72M 7月 1 18:48 subset_1.fastq.gz
iu@bielinux[srp017156] [ 6:52午後 ]
```