inTB - User Manual

Release 1.0

Patricia Soares, Renato Alves

August 12, 2013

CONTENTS

ONE

INTRODUCTION

1.1 What is inTB?

inTB is a data integration platform for molecular and clinical epidemiological analysis of *Mycobacterium tuberculosis*.

inTB innovates not only by its extensive integrative nature and the ability to perform sophisticated analysis but also by providing all this in a ready to use tool. This contrasts with existing alternatives that typically require online access and upload of data to a server via the Internet.

With inTB there are no confidentiality or privacy concerns as your data never leaves your computer.

1.2 What can I do with inTB?

You can use inTB to manage your clinical, socio-demographic and molecular data. inTB has several validation mechanisms while adding data that ensure consistency and prevent human errors.

Once some data is present in the system, you can analyse it. Current analysis include:

- Frequency of infection by a given Species/strain over time

- Reinfection and relapse with same or different Species/strains

- Age distribution of infected patients

- Gender distribution of infected patients

- Co-occurrence of diseases and behavior of infected patients

- Geographic distribution of Species/strains.

- Frequency of treatment abandonment

- Drug resistance over time

- Generate yearly reports of different aspects of tuberculosis progression

- Follow tuberculosis evolution and occurrence of new variants

1.3 Is inTB for the non technically savvy?

inTB is designed with end users in mind. Its target audience are primarily clinicians and medical staff interested in extracting more information from existing data.

All of the provided analysis tools were successfully tested with a group of users mostly unfamiliar with tuberculosis. Most charts/plots were considered intuitive and easy to interpret. For detailed descriptions proceed to the *Usage* section of this manual.

Installation was simplified by providing ready to use packages and detailed installation instructions. Once installed you have all the tools required for the job at the distance of a mouse-click. Refer to the *quick installation instructions* to get started.

1.4 Is it possible to use inTB for infections other than tuberculosis?

inTB was originally designed to be used with tuberculosis, however with minor adjustments it can be adapted for other organisms.

If you have interest in using inTB for other organisms don't hesitate to contact us at http://bioinformatics.igc.gulbenkian.pt.

INSTALLATION

2.1 Alternatives

inTB is available in two formats. A ready to use *VirtualBox image* for *quick installation* and the traditional *zip/tar package* for *server installation*.

2.1.1 Quick Install

In order to use the *VirtualBox image* you must first install *VirtualBox*. Instructions for this can be found under *VirtualBox Setup*.

The easiest way to use inTB is to use the *VirtualBox image*s. If however you prefer to install it on a server for production, please refer to *Server install*.

2.1.2 Server install

inTB can **only** be installed on a Linux/Unix machine. If you want to install it on MacOSX or Windows you will need to use the *Quick Install* approach.

In order to follow these instructions you should already have some knowledge about Linux and the Linux shell. If at any point you feel lost, don't hesitate to contact us at http://bioinformatics.igc.gulbenkian.pt

In order to install inTB on your servers you have two options:

1. Transfer all the content of the *VirtualBox image*: to a local disk. This option replaces anything you have on your disks with the content of the VirtualBox disk. It's easier because all dependencies are already included. Detailed instructions under *Reusing VirtualBox content*.

2. Use the "inTB Package" found on the *Download* section of the inTB website. Use this option if you want to install inTB without replacing the system. You will have to setup all dependencies manually. Detailed instructions under *Installing from source*.

2.2 VirtualBox Setup

*VirtualBox image*s are the simplest way to run your own test version of inTB, but first you need to *Installing VirtualBox*.

2.2.1 Installing VirtualBox

Figure 2.1: Download links on VirtualBox's download page

In order to start you need to download and install the VirtualBox application. You can do it by visiting http://www.virtualbox.org and downloading the **VirtualBox platform package** matching your system.

Figure 2.2: Installation warning - can be safely disregarded

If during installation you are warned (multiple times) about incompatible software, please continue the installation anyway. This warning if frequent in some versions of Windows (in particular Windows XP) and can be safely ignored.

Once VirtualBox is installed (reboot if necessary) you can now proceed to the installation of inTB.

2.2.2 Picking the right format

If you don't want to be bothered, **use the 32bits image**.

If on the other hand you would like to get the best out of your computer while running inTB you should use the 64bits image. However, depending on your computer you might need to confirm if it supports *Hardware virtualization* and if it's active.

Detailed instructions are available for *Windows*, *MacOS* and *Linux*.

Windows

On **Windows** you can use the *LeoMoon CPU-V* application to detect *hardware virtualization*. Simply download the application at http://leomoon.com/shop/application/leomoon-cpu-v/ and execute the provided file.

After a few seconds, you should be presented with one of the following three scenarios:
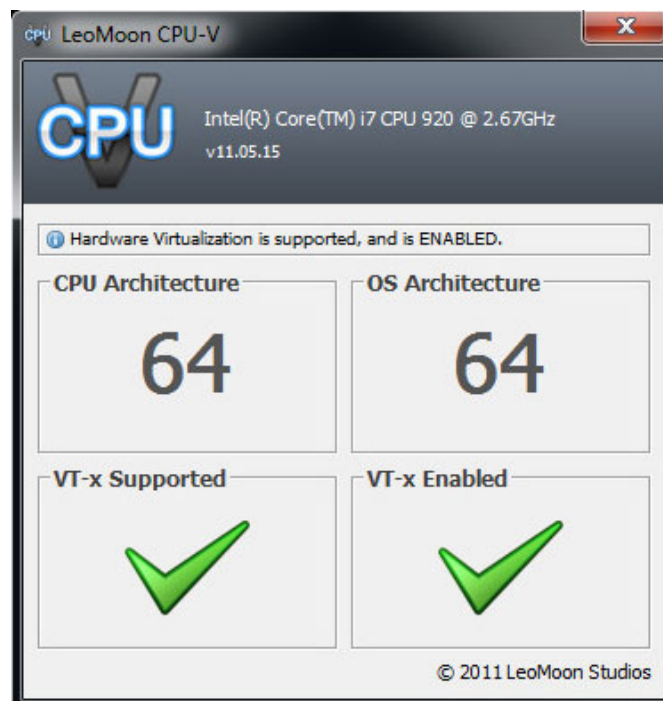

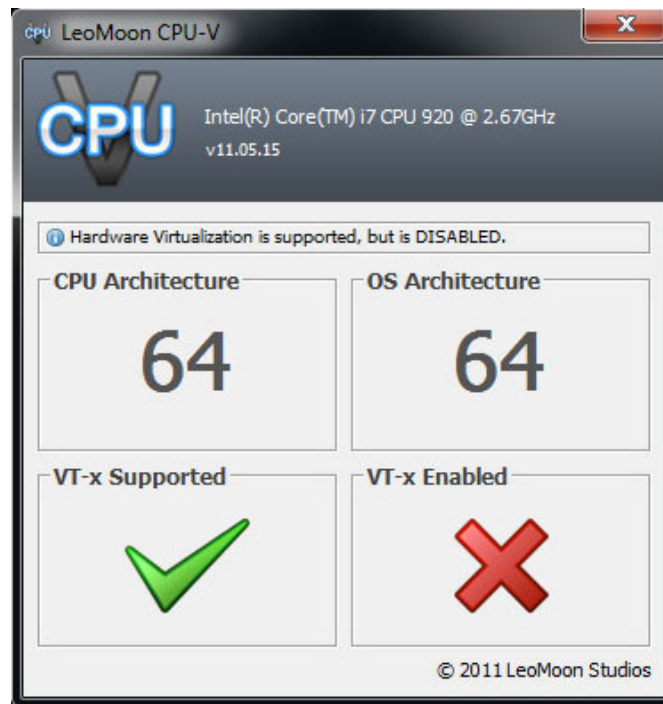Figure 2.3: Virtualization is present, you can use the 64bits image.

Figure 2.4: Virtualization is disabled

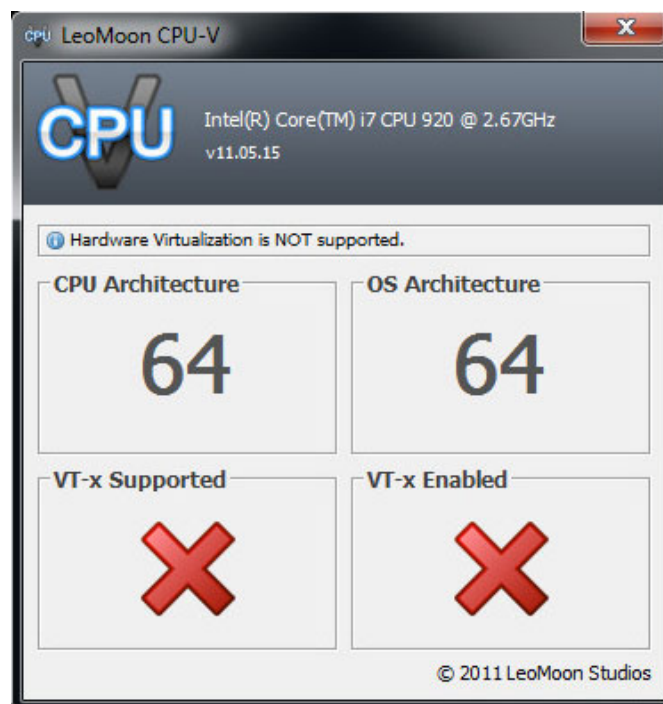Refer to *Disabled Virtualization* under *Advanced Topics* to enable it.

Figure 2.5: Virtualization is not available

Your computer doesn't have *hardware virtualization* - your only option is to use the 32bits image.

MacOS

Intel based MacOS systems have virtualization enabled by default. You should be ok using the 64bits image.

However if after importing and starting the image in VirtualBox you get the error:
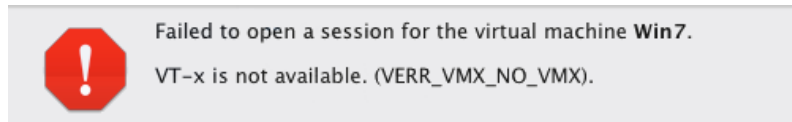

Figure 2.6: Virtualization is not enabled or not available
Try the instructions on http://support.apple.com/kb/TS2744

If the instructions don't work for you, you'll have to use the 32bits image.

Linux

Open a terminal and run the command:

```
egrep '(vmx|svm)' /proc/cpuinfo
```

and if any output including **vmx** or **svm** is generated it means your computer has *hardware virtualization*.

However if after importing and starting the image in VirtualBox you get the error:
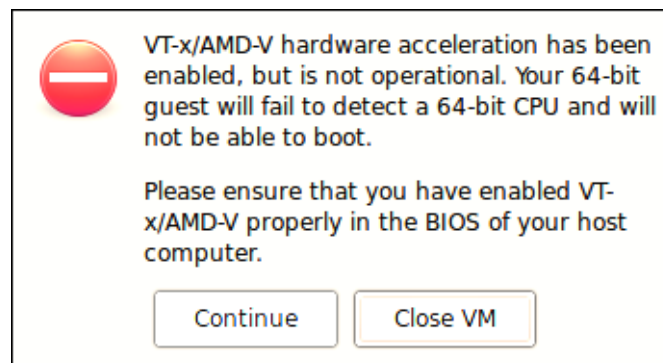

Figure 2.7: Virtualization is not enabled
Refer to *Disabled Virtualization* under *Advanced Topics* to enable it.

2.2.3 Importing inTB

First you should download the correct *VirtualBox image*. Read the section *Picking the right format* if you don't know which to choose.
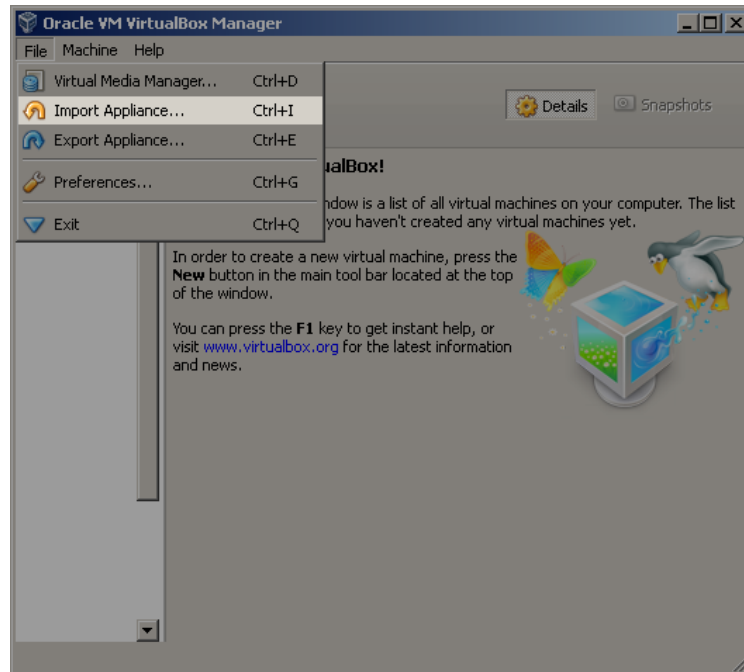

Figure 2.8: How to import the inTB *VirtualBox image*.

Once the download is complete, start VirtualBox, and on the "File" menu select "Import Appliance...".
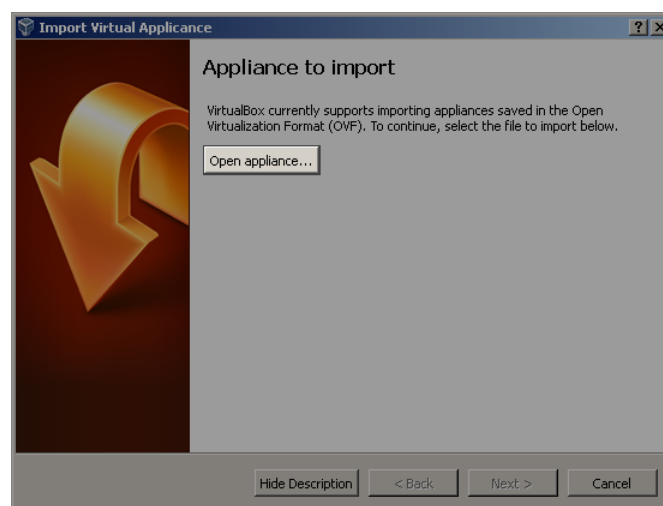

Figure 2.9: The *Import Wizard*

Click on Open appliance and select the *VirtualBox image* you just downloaded, click "Open" followed by "Next".
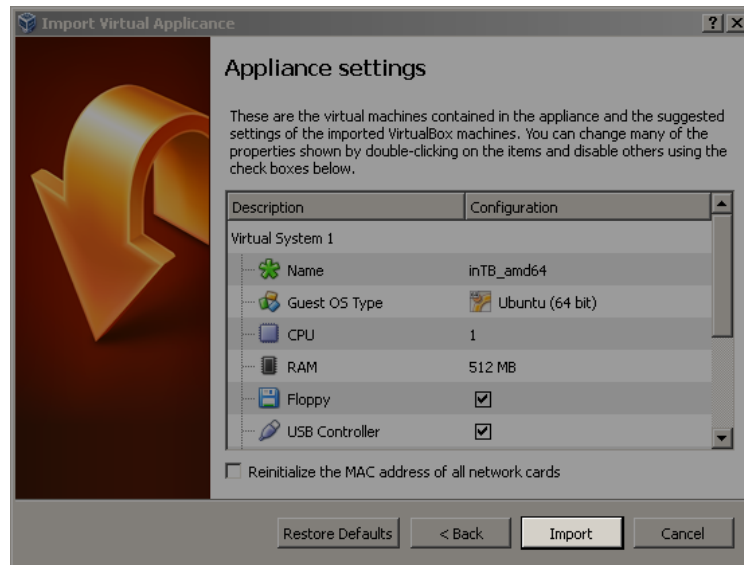
Figure 2.10: inTB's *VirtualBox image* details.
The **Name** should be *inTB_x86* or *inTB_amd64* for the *32* and *64bits* images, respectively.

After a few seconds you will see a summary version of what you are about to import. To continue simple click "Import".

Figure 2.11: Importing in progress.

After a few seconds/minutes you should see a menu entry with the name "inTB_amd64" if you chose the 64bits image or "inTB_x86" if you chose the 32bits image.

inTB is now installed and ready to be used.

2.2.4 Using VirtualBox

On the VirtualBox interface right-click the inTB entry on the list and select "Start".
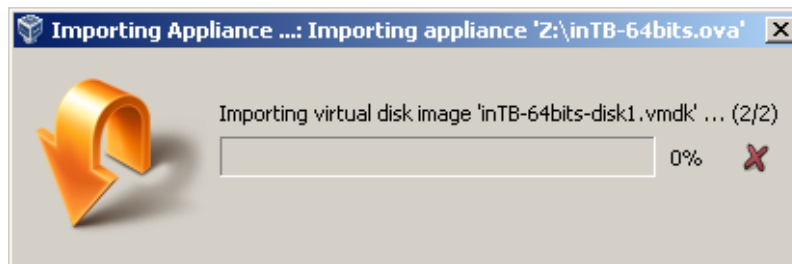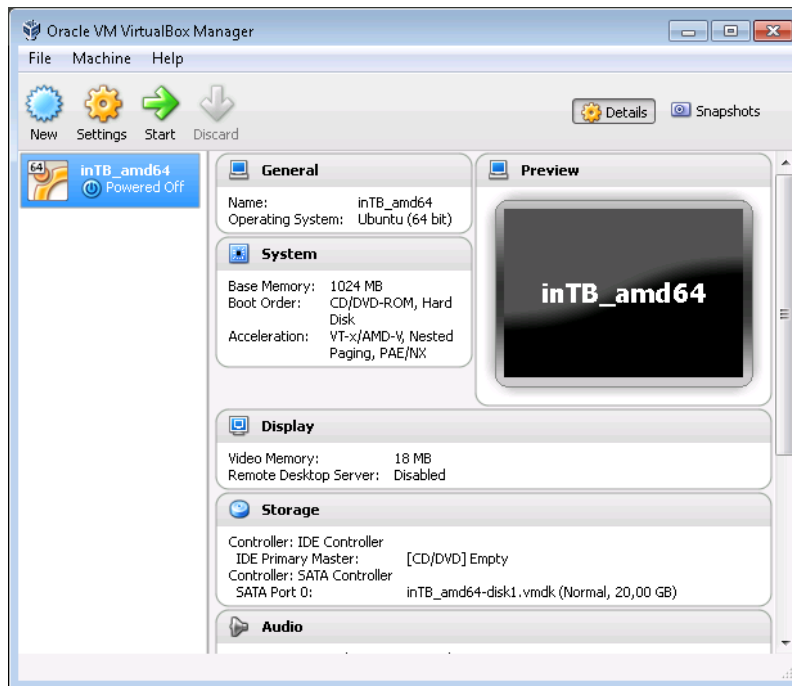

Figure 2.12: Starting inTB

Note: If you picked the 64 bits version and at this point you got an error mentioning VT-x or AMD-V read the *Disabled Virtualization* section under *Advanced Topics*.

If this is the first time you are using VirtualBox you will be presented with a few messages. They explain certain aspects of how VirtualBox works. If you are done with them you can select "Don't show this message again" and press OK.

If your firewall prompts about VirtualBox trying to access the network, please allow it or you may not be able to work with inTB. inTB doesn't require access to external resources but requires that port 8181 is reserved in order to run locally.

2.2.5 Acessing inTB

Now that you have everything ready and VirtualBox running (the black window is visible and it displays login details) you can use inTB.
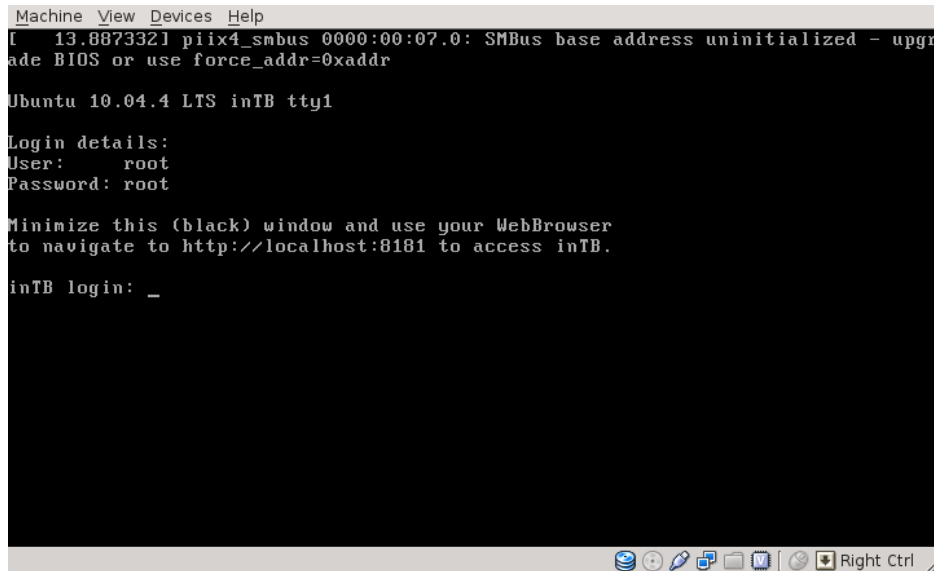
Figure 2.13: inTB ready to be used

To do so, minimize the black window, open a browser (Mozilla Firefox, Internet Explorer, Safari...) and on the address bar type: http://localhost:8181

Figure 2.14: Accessing inTB running locally

If everything worked as expected you should now be using your local installation of inTB.

You can now start adding data via the "INSERT / EDIT" or the "UPLOAD" menu entries. Once some data is loaded, the Browse and Analyze sections will be functional.

2.2.6 Stopping inTB

While working with inTB you should leave the black window minimized. Once you are done you can close it and select "Send the shutdown signal".

Figure 2.15: Stopping inTB

A few messages will be displayed and the black window should close after a few seconds.
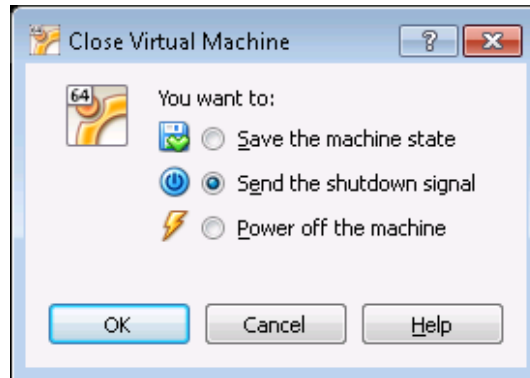

Note: The inTB *VirtualBox image*s are built with conservative limits. If you find that the system crashes or becomes too slow to be usable, read the section *Increasing inTB resources* under *Advanced Topics*.

2.3 Server installation

2.3.1 Reusing VirtualBox content

Requirements:

- 20GB or bigger disk (where inTB will be installed).

- An extra disk or usb drive for storing the raw image (20GB) during the conversion process.

Steps:

1. If you already have VirtualBox installed you can skip to step 3.

2. Download the Oracle VM VirtualBox software from http://www.virtualbox.org or install it via your package repository.

3. Run VirtualBox.

4. On the menu go to File > Import Appliance... , browse to where you downloaded the virtualbox image and select the ".ova" image. Only Use the 32bits.ova image if your server doesn't support 64bits systems. If you use the 32 bits image, everytime 64bits is mentioned from now on, replace it by 32bits.

5. Once importing is finished, open a Linux shell.

6. On the shell type:

```
VBoxManage list hdds | grep inTB-64bits
```

and you should see something like:

```
Location:    /home/user/.VirtualBox/Machines/inTB_amd64/inTB-
  64bits-disk1.vmdk
```

The /home/user... part will be used in the next step.

7. Using the Location from the previous line type:

```
VBoxManage clonehd /your/disk/location/inTB-64bits-disk1.vmdk
  inTB-raw.img --format RAW
```

Once the command finishes there should be a file called inTB-raw.img in your current folder.

8. The aim of this step is to transfer the contents of the inTB-raw.img to the disks of your server.

   Since there is too much variability to cover here on how you can access your server and how its disks are configured, the following instructions assume that you can connect your server disks into your current computer.

   If you have your server disk connected to your current computer run:

```
ls /dev/sd*
```

This should give you something like:

```
/dev/sda   /dev/sda1  ... /dev/sdb  /dev/sdb1
```

The first letter after "sd" identifies the disk and the number identifies the partition number. From the list of all disks you should identify which one corresponds to your server. *Make sure of this or you may accidentally delete all content from the wrong disk.*

If you only have 2 disks connected on your computer (including USB disks and pen drives) you should have "sda" and "sdb". To know which one corresponds to your computer and which one to the server run:

```
mount | grep sd
```

It should print something with a similar structure to this:

```
/dev/sda1 on / type ext3 (rw,errors=remount-ro)
```

In the case above "sda" is my computer's disk which in turn makes "sdb" the server disk I have connected to it.

If you already had a system installed on your server's disk you may try to identify it simply by comparing the number of partitions in each disk or additionally by running:

```
mkdir /mnt/testdisk
sudo mount /dev/sdb1 /mnt/testdisk
(or sdb2, sdb3... depending on your partition scheme)
```

and then checking the content of /mnt/testdisk for something that identifies it as your server's disk.

Before continuing, make sure you run:

```
cd
sudo umount /dev/sdb1
```

Finally, now that you are absolutely sure: that "sdb" (or whatever is your case) is the server's disk, run:

```
dd if=inTB-raw.img of=/dev/sdb
```

9. If all commands were successful so far you should now have inTB installed on the disk as a 20GB partition. All remaining disk space will be left for you to partition to your needs.

10. Additionally, if you plan on using this setup as a public production system, you should follow the steps below - section "inTB in Production".

2.3.2 Installing from source

Requirements:

- 10MB for inTB.

- 200-300MB for dependencies.

- Some extra space for temporary files.

Steps:

1. Download "inTB Package" from the Download section of inTB website.

2. Create the folder /var/django.

3. Unpack the contents of "inTB Package" to /var/django.

4. Install the following python dependencies either using your Linux package manager or manually using easy_install or pip:

```
django - 1.4.0
MySQL-python (aka mysqldb) - 1.2.3
PIL (aka python imaging) - 1.1.7
numpy - 1.4.1
scipy - 0.7.0
networkx - 0.99
pygraphviz - 0.99
chardet - 2.0.1
reportlab - 2.4
simplejson - 2.0.9
biopython - 1.54
pycairo - 1.8.8
matplotlib - 0.99.1.1
```

5. Install the following software:

```
Python - 2.6.5
MySQL server - 5.1.63 (or above)
Apache - 2.2.14 (or above)
Apache's mod_wsgi - 2.8 (or above)
graphviz - 2.20.2
clustalw - 2.0.10 (or above)
phylip - 3.68 (or above)
PhyML - 3.0 (or above)
```

Note: Other versions of the software may work but were not tested

---

6. Go into /var/django/inTB/data/ and load the provided SQL dump into a MySQL database (you can delete the data folder afterwards).

7. Edit the config file located at /var/django/inTB/inTBsite/settings.py and change database connection details accordingly.

---

8. Configure mod_wsgi in Apache following the instructions at https://docs.djangoproject.com/en/1.4/howto/deployment/wsgi/modwsgi/

9. Configure Apache to serve the contents of the media folder located at /var/django/inTB/media/ and adjust MEDIA_URL on settings.py accordingly.

10. If everything went well you should now have inTB running locally.

11. Additionally you should do step 5 of the *inTB in Production* section.
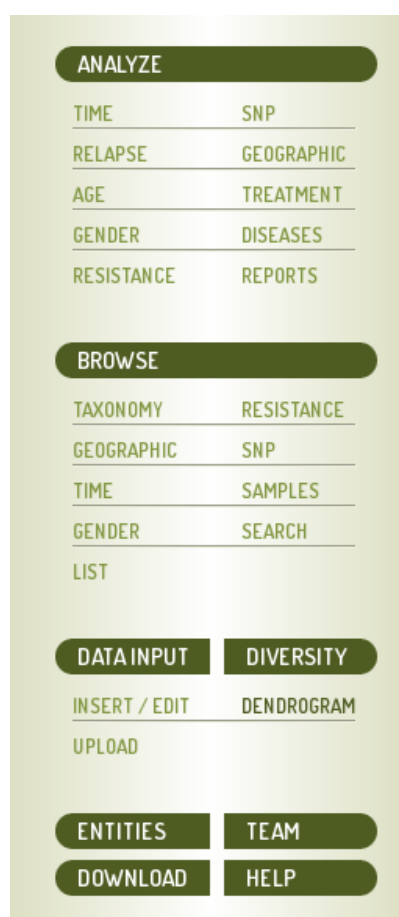
USAGE

3.1 Navigation menu

Figure 3.1: inTB's navigation menu

The menu is organized in 4 sections, *Analyze*, *Browse*, *Data Input* and *Diversity*.

Additionally, 4 buttons exist at the bottom of the menu with institutional (**Entities** and **Team**) and *Download* information and the present *User manual*.

3.1.1 Analyze

Under this section you have all the analysis you can perform in inTB with exception of the phylogenetic analysis which live under the *Diversity* menu section.

- **Time** - the number of infected cases, for all Species/strains, along the years

- **SNP**

    – *Percentage* - visualize SNP variation in collected samples.

    – *Regions* - compare SNP variation, grouping data by region where samples were collected

    – *Resistance* - correlate resistance to antibiotics with presence or absence of snp variants

- **Relapse**

    – *Species* - visualize frequency of reinfection by different species

    – *Strains* - same as above for strains

- **Geographic**

    – *Districts* - compare species variation by district

    – *Counties* - same as above for counties

    – *Citizenship* - proportion of cases by citizenship of affected patients

- **Age** - age distribution of infected patients

- **Treatment**

    – *Previous treatment* - distribution of number of treatments per patient

    – *Interruption* - treatment duration before interruption

- **Gender** - gender distribution of patients

- **Diseases** - proportion of diseases which are associated with tuberculosis infections

- **Resistance** - resistance variation over time

- **Reports** - generate PDF or *CSV* reports summarizing different variables

3.1.2 Browse

Under the browse section you navigate existing data using the different pre-defined filters. Additionally you can perform a custom search with a multitude of personalized filters.

- **Taxonomy** - navigate episodes filtered by species/strain

- **Resistance** - navigate episodes filtered by antibiotics' resistance

- **Geographic**

    – *Districts* - episodes filtered by district

– *Counties* - episodes filtered by county

- **SNP** - explore all SNPs currently in the system and associated information
- **Time** - episodes filtered by year of infection
- **Samples** - display the molecular information of all samples in the system
- **Gender** - episodes filtered by year of infection
- **Search** - perform a custom search with user selected filters
- **List** - manage/compare the episodes that were hand-picked for comparison

3.1.3 Data Input

This section includes the two existing approaches to add data to inTB.

For detailed information on adding data to inTB, please refer to the *Adding data* section of this manual.

- **Insert / Edit** - Insert and update data in the system using a sequential set of forms
- **Upload** - Insert data in the system by uploading a *CSV*

3.1.4 Diversity

In this section you can visualize trees built for each individual genotyping method, as well as one tree (supertree) which tries to incorporate all information from all others methods into a single tree.

Reference species/strains are also marked in the tree which allows automatic classification of all samples in the system.

- **Dendrogram**

  – *SNP* - tree built using SNP information

  – *MIRU-VNTR* - tree built using MIRU information

  – *RFLP* - tree built using RFLP information

  – *Spoligotype* - tree built using Spoligotype information

  – *Supertree* - tree built by merging all results from the other four methods - this method should be analyzed with care and not without comparing with each individual method.

3.1.5 Download

In this section you can download the present *User Manual* and also the *VirtualBox image*s and the packages for advanced installation.

For more information on installation please refer to *Installation*.

3.2 Data access

3.2.1 Summary results

The results page gives an overview of the most relevant information for each patient.

At the top of the page, information about how many episodes match the chosen criteria (both with and without molecular data), as well as instructions on how to interpret the table are provided.

Figure 3.2: Episode statistics

The third column contains the number of episodes each patient had while the color of the number corresponds to the number of samples for each episode. Each episode can have more than one sample. The same column is also a link to the patient file where the user can see all the information about the current episode.

Figure 3.3: Number of infections per patient

The remaining columns contain graphical summaries of the most important information, and more details are available by hovering the cursor over each of these icons. By default, the results table will only show the episodes that have associated molecular data, although a link is provided to show all the episodes.

Figure 3.4: Clinical data - Hovered with mouse

3.2.2 Detailed results

The patient file presents information about the current episode of disease, other episodes, data on follow-up and contact and a report of fields with missing information.

Figure 3.5: Patient file

To visualize the molecular information associated to each episode, the user has to click on the genotyping method name and the table will expand showing the information. SNP polymorphisms are colored for quick identification.

Figure 3.6: Molecular information (in the patient file)

3.3 Adding data

When inserting data, either through the web form or uploading a *CSV* file the user has to select which type of data to insert: clinical, socio-demographic, molecular, resistance, follow-up or contact data. For every type of data a patient identifier will be requested, but it is not mandatory for molecular and resistance data, where only a sample code is required. This enables the asynchronous uploading of data that can later be linked, when further data becomes available.

3.3.1 Symbols used

Adding new options

While adding information, in particular when the database is mostly empty, a lot of the fields will not include any options other than the *no information* representation "`------`".

Figure 3.7: Adding patient information

In order to prevent inserting different variations of the same information, options already on the system are displayed for selection. If the desired option is not yet present among the available ones, one can click on the **green plus (+)** sign in order to add it.

A window will show up prompting the user for new information.

Figure 3.8: Adding new option

Once confirmed, the new option will be automatically selected and will be available in the future.

Selecting dates

Dates can be selected by clicking on the *small blue calendar icon* or by typing in the date in the correct format. If the blue calendar is clicked a widget will be displayed where you can select the date.
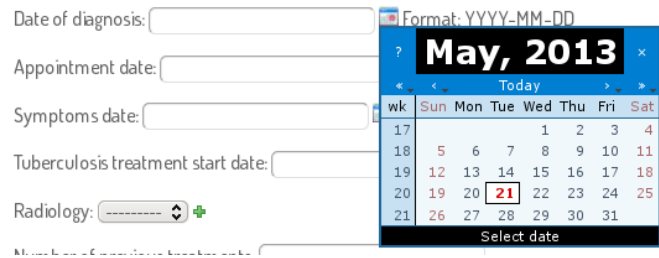
Figure 3.9: Picking a date using the widget

3.3.2 Single entry - Using the forms

Figure 3.10: Inserting/Editing data

Information is divided into **Patient** and **Infection** information. The former includes *clinical* and *socio-demographic* information, as well as *contacts* and *follow-up* records. The latter includes molecular information about any collected samples.

While adding patient data using the forms the *patient identifier* will always be a required field. For each patient inserted into the database, the system automatically creates an episode number. If the patient reappears into the system, a new episode will be associated for this patient, meaning the patient had, at least, two distinct clinical episodes.

When adding infection information, the *patient-ID* can be omitted and instead a *sample identifier* will be required. Not providing the *patient-ID* will store molecular information without being associated to a patient.

In the case that the user is inserting data for an existing patient, with multiple episodes in another block of information, the user must specify the episode corresponding to the new data. Since the episode is a number created by the system, a link is provided to the patient file where all the information about the episodes is presented. This facilitates the linking of different blocks of information. The same reasoning is applied when inserting molecular and/or resistance data.

Additionally one can edit existing data by using the provided *UPDATE* forms. They are identical in all to the *INSERT* counterpart, with the exception that they will be rendered pre-filled with existing information.

Follow-up information

During a *follow-up* new samples may be collected and patient details may have changed since last visit. To simplify the process when adding a *follow-up* information the user is presented with a copy of data inserted previously for the same patient.

If no information change, you can simply confirm all intermediate steps.

3.3.3 Multiple entries - File upload

While *inserting single entries* may be the most frequent use-case, we provide an alternative for bulk insertion of data. Bulk insert can be useful when migrating data from other systems.

Note: Please note that it's not possible to **edit** existing data using this approach.

Similarly to *Single entry - Using the forms* the user has different options for data upload. The only exception is *follow-up* information which should be included in each file and distinguished by specifying the appropriate value under the column **Type visit.**.

After selecting the type of information to be inserted, the user is presented with an upload form and links to **two template files**. One file (*without errors*) provides a reference for how data should be formatted. The other (*with errors*) allows previewing the validation mechanisms present in inTB.

ADVANCED TOPICS

4.1 Disabled Virtualization

If you have a recent computer but VirtualBox says it doesn't have hardware acceleration, check that virtualization is enabled on the BIOS of your computer. Some computers have virtualization disabled by default.

The BIOS can be accessed by pressing a key imediately after powering on your computer. Usually the key that needs to be pressed is one of "Delete", "F2", "F8", "F9" depending on the computer model.

If you can't find any option on the BIOS related with virtualization, please go back to *Picking the right format* and use the 32 bits version.

4.2 Advanced access to database

In addition to the inTB web interface you can access the database directly by navigating to: http://localhost:8181/phpmyadmin

and login with user "root" and password "intb". The database is called "intb".

This will allow you to change content and structure in a way which is not possible via inTB's interface.

Keep in mind that modifying the database directly may affect inTB's output and depending on the extent of the modification, even cause it to fail.

4.3 Increasing inTB resources

Two things may affect significantly the performance of inTB while running inside VirtualBox: memory and processing power, both of which can be increased as long as sufficient resources exist on your computer.

Additionally, if you loaded large sets of data it may be that the system is running out of space. The current space limit is set to 20GB. If this is your case we recommend that you follow the *Server install* instructions to install inTB on a dedicated computer/server.

To increase either memory or processing power, stop inTB by closing the black window. On the VirtualBox Manager window right-click on inTB and select "Settings". On the newly presented window select "System". Here you can increase the amount of memory. Try to prevent going into the orange/red area or your computer's performance will suffer. If you wish to increase processing power, while still on the "System" section, click on the "Processor" tab and increase the number of processors assigned to it.

Note: If you don't have Hardware virtualization support you won't be able to increase the amount of processors.

For best performance in inTB at least 2048MB of Memory and 2 Processors are recommended.

4.4 Changing the system

If you wish to customize the system beyond what the web interface allows you to do, you can access the virtual machine directly by logging in on the VirtualBox window. The credentials are presented on screen. If the screen is completely black, move the mouse cursor over the black screen, left-click it and press any key on the keyboard.

The VirtualBox is running Ubuntu Lucid (10.04). inTB is built with Django 1.4 and all of its files can be found on /var/django/inTB/ .

4.5 inTB in Production

The provided *VirtualBox image*s are meant for testing purposes and limited use. When using them in production some additional steps need to be taken to ensure the system remains secure. These are beyond the scope of this manual are are indicated here simply as bullet points:

1. Change password for user "root".

2. Change password for user "root" on MySQL.

3. Create a MySQL account for a user specific for intb.

4. Configure inTB to use this intb user (/var/django/inTB/inTBsite/settings.py)

5. Change SECRET_KEY on settings.py (same as point 4) according to https://docs.djangoproject.com/en/dev/ref/settings/#secret-key

TROUBLESHOOTING

Table 5.1: Local installation problems

Problem	Solution
The message "Minimize this (black) window..." is displayed in VirtualBox but cannot access inTB after installation	Check that your firewall allows local connections. Try disabling it temporarily. On MacOS run `sudo ipfw flush`.
The message "Kernel panic..." is displayed in VirtualBox	Restart VirtualBox. If it happens again, and you are using the 64bits image, use the 32bits image instead
You can't find the mouse and the keyboard is locked in VirtualBox	Press the Right Ctrl key (on Windows and Linux) or the Left Cmd key (on MacOS) on your keyboard to release the mouse and keyboard.
I want more follow-up options/alternatives	In order to add other follow-up options you need to use the *Advanced access to database* and insert the desired option in the table **entry_db**. Additionally, the column **order** must match the temporal order of different follow-ups.

GLOSSARY

Clinical form Different symptoms which reflect the type of tuberculosis infection (e.g. Pulmonary, Meningitis,...).

Episode Record of infection. One patient may have more than one episode if infection reoccurs by the same or a different species/strain.

H37RV A strain of M. tuberculosis commonly used as reference.

MIRU-VNTR Mycobacterial Interspersed Repetitive Units - Variable Number Tandem Repeat, a method by which you can count the number of times a sequence is repeated in a given set of locus.

RFLP Restriction fragment length polymorphism, a method based on the restriction pattern of a predefined set of enzymes.

Sample Biological sample which is then analyzed with different methods. More than one sample may be collected for a single episode.

SNP Single Nucleotide Polymorphism, identifying single nucleotide variations scattered across the genome.

Spoligotype A method which detects presence or absence of 43 unique spacers between repeats in a specific locus.

VirtualBox An application which allows running different operating systems as if they were independent applications.

VirtualBox image A full operating system including inTB and all required dependencies. Once *installed*, one click and you are ready to go.

User Manual The present document.

zip/tar package A package which include all files you need for deployment, except dependencies - recommended only for advanced users and system administrators.

Hardware virtualization A technology present in most recent processors which allows virtualizing an operating system. In practice this means running an operating system as if it was an application. This technology is also called VT-x on Intel processors and AMD-V or SVM on AMD processors.

CSV A Comma Separated Values (CSV) file. Contains information in a spreadsheet like format which can be easily processed by other tools. This format can easily be generated

from most software that can handle spreadsheets. Examples include OpenOffice (and variants), Microsoft Office and iWorks' Numbers, among several others.

Patient identifier A numerical entry which identifies the patient. Typically this would be the National Health identification ID. This number is never displayed in inTB.

Sample identifier A alpha-numerical entry which identifies a sample collected from a patient. This field has no required format. Typically this would be a sample code generated by genotyping or sequencing platforms.

Follow-up Typically a medical visit after the first evaluation of the clinical status of the patient. A follow-up appointment can be performed before or after treatment has started.

INDEX