**User Manual for the DNARNA package**

**Introduction:**

This guide should be read in conjunction with the associated publication that describes the theory in detail. The program is open source and can be freely distributed. It is intended to update the program as required and we would be pleased to assist with readers' requests and queries. - please contact the authors.

DNARNA is a companion package to the paper (ref to be added). The program was used to carry out the calculations.

This package is tested with R v. 4.1.2. It does not work on versions lower than R v.4.

**Step 1**: **Download files into a local folder:**

Navigate to the folder:

Either:

<https://github.com/peterdgill/DNAVagMucosa>

Or:

https://www.dropbox.com/sh/yp1f21l8bu7327e/AAA-SkcXVR\_Y4V-erR524zBOa?dl=0

and download " DNARNA\_0.1.0.tar.gz" (the source file) and the datafile " Results\_DNA\_rfu.csv". Put the files into the same folder which will be used for the R directory.

**Step 2:** In the R-console, set the working directory using the setwd() command to the folder where the files are loaded e.g. setwd("C:/myfolder"). Alternatively set the folder from File>Change dir and use the browser to point to the folder location

**Step 3:** Load packages: Set the CRAN mirror by navigating to Packages>Set CRAN mirror (fig 1) and choose a connection. Copy/paste the following code into the R console, press return. This step only needs to be carried out once for the version of R that you are using:

# Package names

packages <-c("plyr","poibin","rstanarm","ResourceSelection","DescTools","Rcpp","fitdistrplus","bootstrap")

# Install packages not yet installed

installed\_packages <- packages %in% rownames(installed.packages())

if (any(installed\_packages == FALSE)) {

install.packages(packages[!installed\_packages])

}

# Packages loading

invisible(lapply(packages, library, character.only = TRUE))

**Step 4: Load the DNARNA package:**

In the R-window navigate to Packages>Install packages from local files (fig 1) and navigate to the folder which contains the source file and data file.

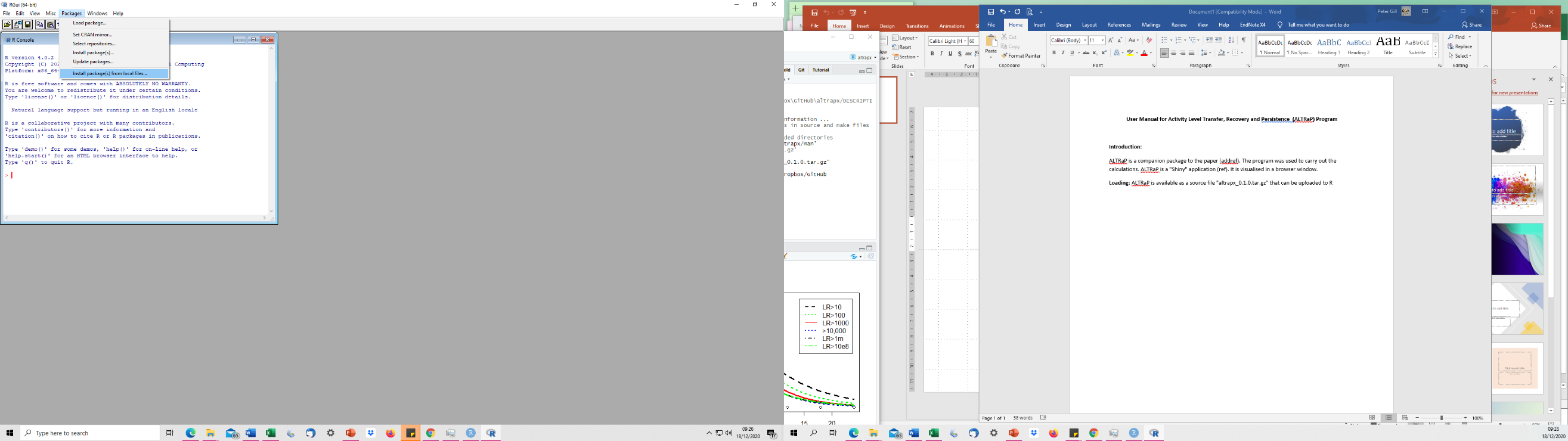


Fig 1: R gui to load packages from local file

Next load the library by typing into the console:

library(DNARNA)

and press enter

**Step 5:**

To run the program, type:

BNprog("Boxershorts","Results\_DNA\_rfu.csv") for boxershorts analysis

BNprog("Penile swabs","Results\_DNA\_rfu.csv") for penile swabs analysis

BNprog("Fingernail swabs","Results\_DNA\_rfu.csv") for fingernail swabs analysis

Note that you can use your own "Results" files but they must be in the same format (.csv) – see section X for details.

**Step 6:**

The results are written to .csv files in your working directory. They can be opened in excel.

The .csv files are labelled according to the Bayesian network nodes D, V and DV as follows:

Dminus.csv ## D-

Dplus.csv ## D+

Vminus.csv ##V-

Vplus.csv ## V+

DplusVplus.csv ## D+V+

DplusVminus.csv ## D+V-

DminusVplus.csv ## D-V+

DminusVminus.csv ## D-V-

The tables generated reproduce the LR results in the paper: additional files provide much more information for the interested user. All LRs are log10 values (fig 1). Timescales are between 0-35h in steps of 5h.

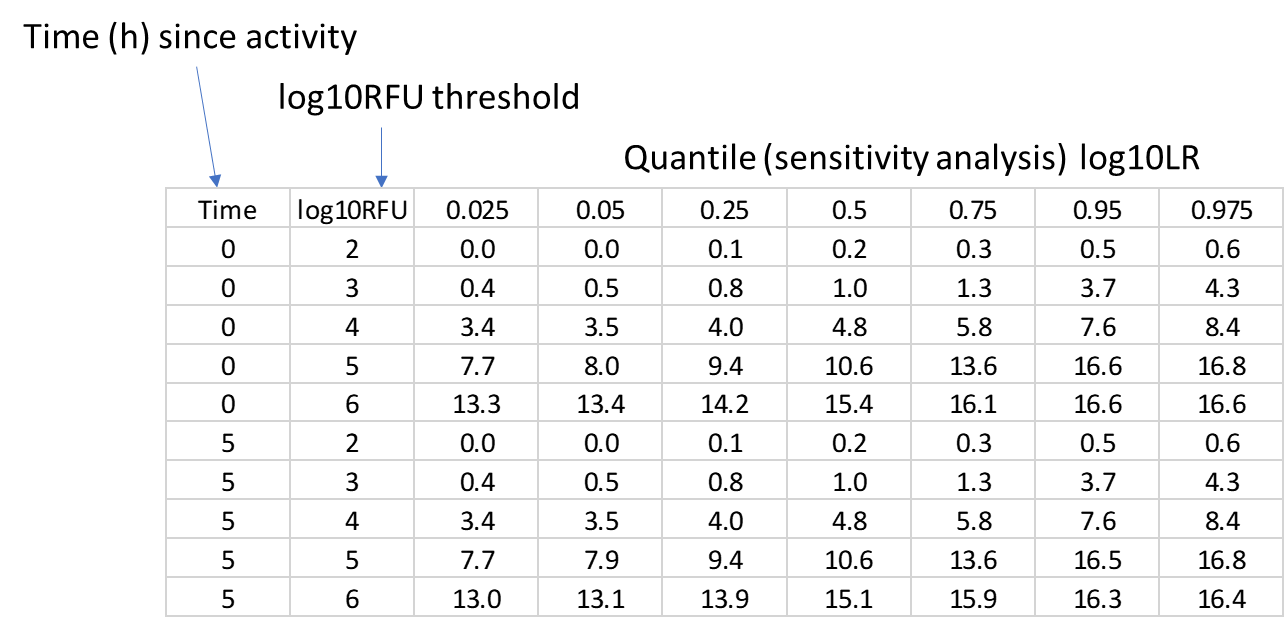


Fig 1: Section of printout of log10LRs from DplusVplus.csv file showing 0 and 5h results only.

**Datafile**

The .csv datafile contains all of the information collected in relation to the samples. Program XX only imports results from "Location", "Time\_point", "Ave\_rfu\_POI". Identifiers cannot be altered as the program will not recognise the data. The datafile is imported into the R-program as a dataframe.

**A list of datafile headings**

Donor: a code identifying the 'suspect' donor

POI: A code identifying the POI ('victim')

Sample: A code to identify the sample

T number: Internal identity number assigned by LIMS system

No\_handwash: Number of handwashes reported by the 'suspect' donor during the course of the experiment

Location: Either Fingernail swabs, penile swabs or Boxershorts

Time\_point: Time between the activity and collection of intimate contact samples (0-36) hours. Boxershorts results are all set to time 0. If the results are from non-intimate contact sampling, they are designated "Background".

DNA\_quant: The DNA quantity measured in ng/ul

Total\_rfu: Sum of all peak heights across all loci, except for amelogenin

Ave\_rfu\_loci: The average per locus sum of peak heights

Input vol: The volume in ul taken from the eluant and forwarded to the PCR reaction

Dil\_fac: The dilution factor applied i.e. the dilution factor of the DNA extract into the PCR reaction

NOC: Number of contributors

MxProp\_D: The mixture proportion of the suspect donor

rfu\_D: the total RFU across loci adjusted by the mixture proportion of the suspect donor

Ave\_rfu\_D: The average RFU per locus adjusted by the mixture proportion of the suspect donor

MxProp\_POI: The mixture proportion of the POI (victim)

rfu\_POI: the total RFU across loci adjusted by the mixture proportion of the POI (victim)

Ave\_rfu\_POI: The average RFU per locus adjusted by the mixture proportion of the POI (victim)

MxProp\_U: The mixture proportion of an unknown donor if present

rfu\_U: the total RFU across loci adjusted by the mixture proportion of the unknown donor

Ave\_rfu\_U: The average RFU per locus adjusted by the mixture proportion of the unknown donor

LR\_mle: The subsource LR

Vaginal\_mucosa\_result: either "Detected" or "Not detected"