

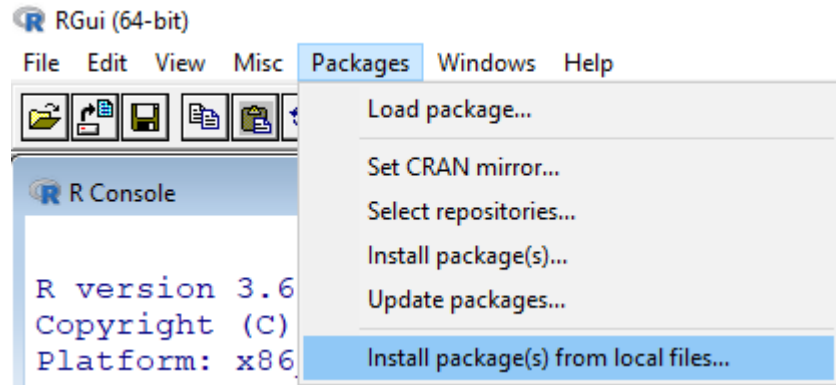
seq2lus

A R-package to convert UAS sequences to
Longest Uninterrupted Stretch (LUS)

By Øyvind Bleka and Rebecca Just

Installation

- Requires R (at least v3.0.0)
- Install required R-packages in R:
 - `install.packages(c("gWidgetstcltk","readxl"))`
 - Select the “seq2lus” zip-file when clicking
 - “Install package(s) from local files...”

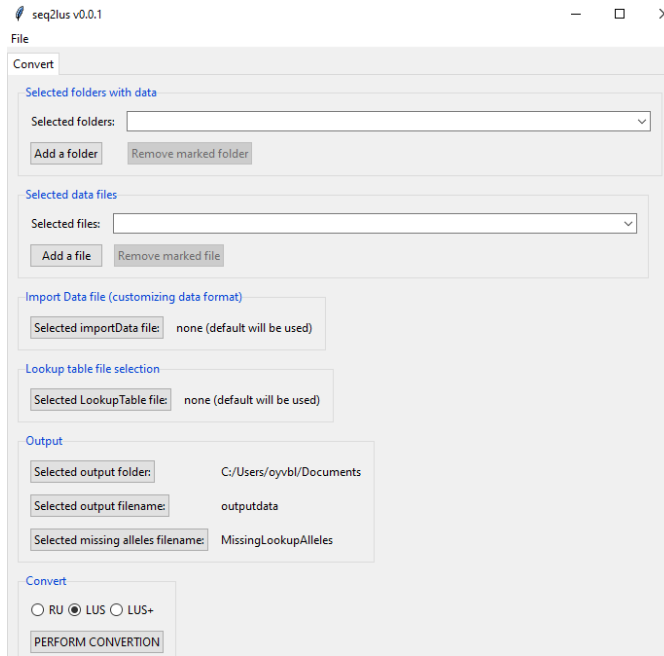


Installation (from github)

- Requires R (at least v3.0.0)
- Install required R-packages in R:
install.packages(c("gWidgetstcltk", "readxl", "devtools"))
- Install directly from github:
library(devtools)
install_github("oyvble/seq2lus")

Opening the GUI

- After installation, run following in commando:
library(seq2lus);cgui()



File

Convert

Selected folders with data

Selected folders:

Add a folder

Remove marked folder

Selected data files

Selected files:

Add a file

Remove marked file

Import Data file (customizing data format)

Selected importData file: none (default will be used)

Set back to default

Lookup table file selection

Selected LookupTable file: none (default will be used)

Set back to default

Output

Selected output folder: C:/Users/oyvbl/Documents

Selected output filename: outputdata

Selected missing alleles filename: MissingLookupAlleles

Convert

 RU LUS LUS+

PERFORM CONVERSION

Select folder(s) with datafiles

Select specific datafile(s)

Allows for data to be in another format than (data customization):
[SampleName Marker Allele1,.....]

The excel lookup-file used for conversion (can optionally be selected)

Select where data is saved after conversion (all profiles are stored in one file)

Filename for storing non-converted alleles

Select allele type to convert to

The converter

- Based on using the “convert” R-function
- A specific allele type is selected (RU, LUS, LUS+)
 - 1) Takes sequences from the imported data
 - 2) Look up sequences in LookupTable file
 - An excel file created by Rebecca Just
 - 3) Extract corresponding alleles
 - Adding up coverages if several sequences give same allele type
 - 4) Store converted data in a single file
 - 5) Sequences not found in LookupTable are stored in “missingfile”

Important notes

- Only either evidence profiles or reference profiles can be converted at the same time:
 - This is because of different data formats
 - Evids can have many alleles (with coverage/height columns)
 - Refs only have two allele columns
- Settings will be remembered once selected
- Settings can be stored as a project file

Default data import

- The evid-data returned by the function `importData` must be in one of the two specific formats:
 - “Wide-format”:
[SampleName Marker Allele1 Allele2...Height1 Height2...]
 - “Long-format”:
[SampleName Marker Allele Height]
- “Covarage” instead of “Height” is accepted as name.

Customized data import

- It's possible to create your own importData function and store it in an R-script which is selected at "Select importData file"
 - The function MUST return one of the formats from the previous slide.
 - An example file to read from XL-spreadsheet is found in the installation folder:
~\R-installation\seq2lus\importData_UAS.R

References

1. R.S. Just, J.A. Irwin, Use of the LUS in sequence allele designations to facilitate probabilistic genotyping of NGS-based STR typing results, *Forensic. Sci. Int.: Genet.* 34 (2018) 197-205.
2. R.S. Just, J. Le, J.A. Irwin, LUS+: Extension of the LUS designator concept to differentiate most sequence alleles for 27 STR loci, *Forensic Sci Int.: Reports*, in press.
3. K.B. Gettings, L.A. Borsuk, D. Ballard, M. Bodner, B. Budowle, L. Devesse, et al., STRSeq: A catalog of sequence diversity at human identification Short Tandem Repeat loci, *Forensic. Sci. Int.: Genet.* 31 (2017) 111-117.