

MBatch 01 Installing on OSX  
Tod Casasent and Andre Schultz  
2022-09-01-1300

## Introduction

These instructions are aimed at people very familiar with OS X and R administration, with administrator access to their environment. If using the versions discussed here, it may be possible to install MBatch with minimal administration knowledge.

NOTE: If these instructions are out of date, please see the Linux install instructions for the newest version: [https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/blob/master/docs/MBatch/MBatch\\_01\\_InstallLinux.pdf](https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/blob/master/docs/MBatch/MBatch_01_InstallLinux.pdf)

## Target Operating System and Installation

These instructions were tested on OSX 10.9.x. These instructions with appropriate modifications should work as a basis for installing MBatch on other systems.

Perform the installs in the order given in this document. Some steps are dependent on previous steps.

## Install R 4x

MBatch is tested and used with R 4.x on other systems. For OSX, we tested with 10.9.x, but 4.x is now available for that. Use the instructions on a local CRAN mirror (<https://cran.r-project.org/mirrors.html>) under "Download R for (Mac) OS X" (as shown in the screenshot).

### Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Then use the "R-4.2.1.pkg" link under "Latest release".

### Latest release:

[R-4.2.1.pkg](#) (notarized and signed)

SHA1-hash: f83a6c96cedd19193255f94cb01381a273073a3a  
(ca. 90MB) for Intel Macs

**R 4.2.1** binary for macOS 10.13 (**High Sierra**) and higher, **Intel 64-bit** build, notarized package.

Contains R 4.2.1 framework, R.app GUI 1.79 in 64-bit for Intel Macs, Tcl/Tk 8.6 libraries and Texinfo 6.7. The latter two components are optional and can be omitted when choosing "custom install", they are only needed if you want to use the terminal package or build package documentation from sources.

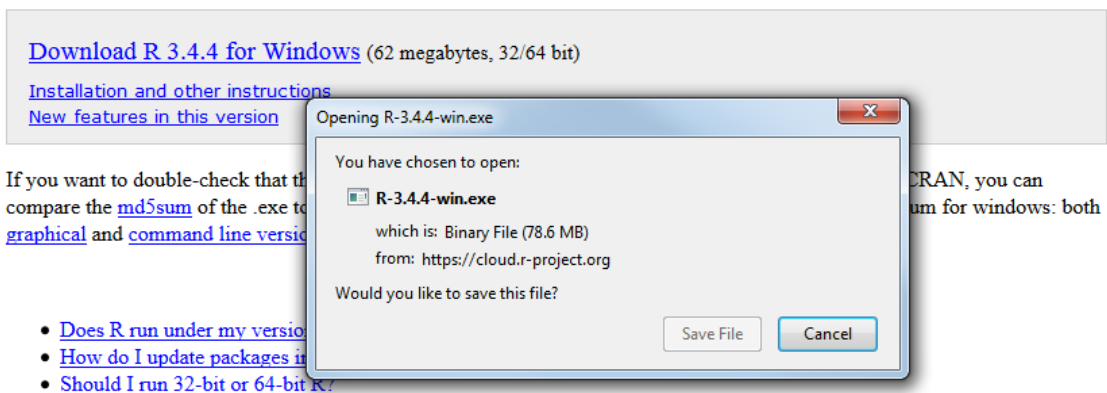
Note: the use of X11 (including `tcltk`) requires [XQuartz](#) to be installed (version 2.7.11 or later) since it is no longer part of macOS. Always re-install XQuartz when upgrading macOS to a new major version.

This release supports Intel Macs, but it is also known to work using Rosetta2 on ARM-based Macs. For native Apple silicon arm64 binary see below.

**Important:** this release uses Xcode 12.4 and GNU Fortran 8.2. If you wish to build packages from sources, you may need to download GNU Fortran 8.2 - see the [GNU Fortran directory](#).

Download the R install program and follow the instructions provided under "Installation and other instructions". (Image is from 3.x for Windows, but 4.x will work similarly.)

### R-3.4.4 for Windows (32/64 bit)



Follow the installation instructions from R-Project.

## Setting Up a Personal R Library

If you have not installed any packages in R, you will need to create a personal package for R libraries. Install a package, as shown below, and create a personal library directory. This will appear similar to the below.

```
> install.packages("devtools", repos = "http://cran.revolutionanalytics.com")
Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)
Warning in install.packages("devtools") :
'lib = "/usr/local/lib/R/site-library"' is not writable
Would you like to use a personal library instead? (y/n) y
Would you like to create a personal library
~/R/x86_64-pc-linux-gnu-library/3.4
to install packages into? (y/n) y
trying URL 'https://cran.revolutionanalytics.com/src/contrib/devtools_1.13.3.tar.gz'
Content type 'application/octet-stream' length 486402 bytes (475 KB)
=====
downloaded 475 KB
```

## Installing Required Packages

MBatch has some required packages.

### Other Required Packages for NGCHM

Within R, run the following installs.

```
# start R (--vanilla says don't load any environments or save any)
R --vanilla

# install packages from CRAN (this installs to site-library so it is available for
all users)

# update to preferred CRAN mirror if desired

# this is at the R command
install.packages("httr", dependencies = TRUE, repos = "http://cran.r-
project.org")
install.packages("magick", dependencies = TRUE, repos = "http://cran.r-
project.org")
install.packages("gert", dependencies = TRUE, repos = "http://cran.r-
project.org")
install.packages("htmlwidgets", dependencies = FALSE, repos = "http://cran.r-
project.org")
```

```

install.packages("shiny", dependencies = TRUE, repos = "http://cran.r-
project.org")

install.packages("usethis", dependencies = TRUE, repos = "http://cran.r-
project.org")

install.packages("covr", dependencies = TRUE, repos = "http://cran.r-
project.org")

install.packages("rversions", dependencies = TRUE, repos = "http://cran.r-
project.org")

install.packages("devtools", dependencies = TRUE, repos = "http://cran.r-
project.org")

library(devtools)

# remotes

devtools::install_github('MD-Anderson-Bioinformatics/tsvio')

devtools::install_github('MD-Anderson-Bioinformatics/NGCHMSupportFiles',
ref='main')

devtools::install_github('MD-Anderson-Bioinformatics/NGCHM-R')

# quit R

# this is at the R command

q()

```

## Other Required Packages

Within R, run the following installs.

```

# start R (--vanilla says don't load any environments or save any)

R --vanilla

message("#*#* reticulate used by MBatch")

install.packages("reticulate", dependencies=TRUE, repos = "http://cran.r-
project.org")

message("#*#* mclust used by oompaBase, ClassDiscovery, and PreProcess (in
install 1)")

install.packages("mclust", dependencies=TRUE, repos = "http://cran.r-
project.org")

message("#*#* Biobase used by oompaBase, ClassDiscovery, and PreProcess
(in install 1)")

install.packages("BiocManager", dependencies=TRUE, repos = "http://cran.r-
project.org")

```

```

BiocManager::install(c("Biobase"), update=FALSE)
message("### oompaBase, ClassDiscovery, and PreProcess")
# used by PCA+ and other clustering algorithms in MBatch
# uses ClassDiscovery functions distanceMatrix and SamplePCA
install.packages(c("oompaBase", "ClassDiscovery", "PreProcess"), dependen-
cies=TRUE, repos=c("http://cran.r-project.org", "http://silicovore.com/OOMPA/"))
message("### Cairo used in MBatch for images")
install.packages("Cairo", dependencies=TRUE, repos = "http://cran.r-
project.org")
message("### BiasedUrn needed for epiR")
install.packages("BiasedUrn", dependencies=TRUE, repos = "http://cran.r-
project.org")
message("### pander needed for epiR")
install.packages("pander", dependencies=TRUE, repos = "http://cran.r-
project.org")
message("### sf needed for epiR")
install.packages("sf", dependencies=TRUE, repos = "http://cran.r-project.org")
message("### lubridate needed for epiR")
install.packages("lubridate", dependencies=TRUE, repos = "http://cran.r-
project.org")
message("### epiR used in correlation density plot")
#install.packages("epiR", dependencies=TRUE, repos = "http://cran.r-
project.org")
#Install version 2.0.40, anything newer fails
install.packages("https://cran.r-project.org/src/contrib/Archive/epiR/epiR_2.0.40.tar.gz")
message("### squash used for dendromat in Supervised Clustering and Hier-
archical Clustering")
install.packages("squash", dependencies=TRUE, repos = "http://cran.r-
project.org")
# used in DAPIR
message("### httr")
install.packages("httr", dependencies=TRUE, repos = "http://cran.r-
project.org")
message("### jsonlite")

```

```

install.packages("jsonlite", dependencies=TRUE, repos = "http://cran.r-
project.org")

# uses dunn.test in MBatchUtils
message("### dunn.test")

install.packages("dunn.test", dependencies=TRUE, repos = "http://cran.r-
project.org")

# used for UMAP
message("### uwot")

install.packages("uwot", dependencies=TRUE, repos = "http://cran.r-
project.org")

message("### devtools")

install.packages("devtools", dependencies=TRUE, repos = "http://cran.r-
project.org")

library(devtools)

message("### jlmelville/vizier")

devtools::install_github("jlmelville/vizier")

# quit R
# this is at the R command
q()

```

## Install MBatch Package

MBatch can be installed from GitHub using the following commands. (The devtools package was installed earlier.)

```

# At Linux command line (this install is for your own user, su to root to install
for everyone)

# start R (--vanilla says don't load any environments or save any)

R --vanilla

# Within R

library(devtools)

devtools::install_github("MD-Anderson-Bioinformatics/BatchEffectsPackage/apps/MBatch")

# quit R
# this is at the R command

```

q()

Note: As a side note, the "tests" directory installed via devtools is not by default included in install.packages or R CMD INSTALL. If installing MBatch via a different method, to include tests via the command line, use "--install-tests".

## Populating Data for Tests

The files located at [https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/tree/master/data/testing\\_static/MATRIX\\_DATA](https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/tree/master/data/testing_static/MATRIX_DATA) go in the ~/Downloads/MATRIX\_DATA folder.

The files located at [https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/tree/master/data/testing\\_static/COMPARE](https://github.com/MD-Anderson-Bioinformatics/BatchEffectsPackage/tree/master/data/testing_static/COMPARE) go in the ~/Downloads/COMPARE folder.

Create a directory named ~/Downloads/output.

## Running the Tests

Finally, run this code to perform the tests. This takes an hour or so to run. Machines with slower processors, slower hard drives, or minimal memory may take longer or be unable to complete the tests.

### # Within R

```
library(MBatch)

# Set these environment variable to override file locations if needed
#
Sys.setenv(MBATCH_TEST_OUTPUT="~/Downloads/output")
Sys.setenv(MBATCH_TEST_INPUT="~/Downloads/MATRIX_DATA")
Sys.setenv(MBATCH_TEST_COMPARE="~/Downloads/COMPARE")
baseDir <- file.path(system.file(package = "MBatch"), "tests")
message(baseDir)
testFiles <- list.files(path=baseDir)
print(testFiles)
results <- c()
for(myFile in testFiles)
{
  message("*****")
  message("*****")
}
```

```

message("**** ", file.path(baseDir, myFile))
message("*****")
message("*****")
test <- source(file.path(baseDir, myFile))
if (isTRUE(test$value))
{
  results <- c(results, paste("Test succeeded for ", myFile, sep=""))
}
else
{
  results <- c(results, paste("Test failed for ", myFile, sep=""))
}
}
print(results)

```

The last part of the output will print whether or not the tests finished properly.



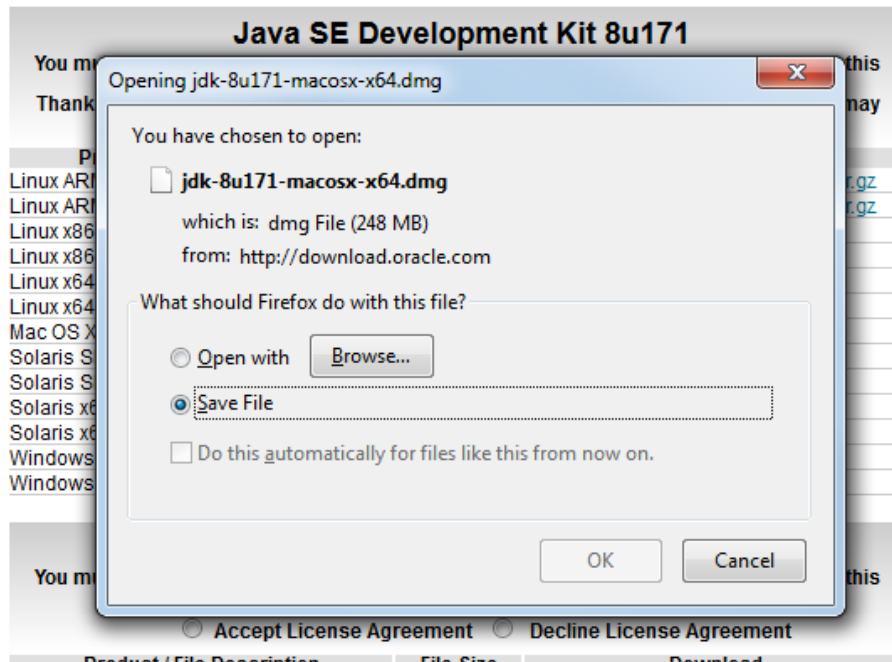
```

2018 05 22 09:59:23.947 DEBUG BSTW5MHDXM2 mbatchStandardLegend before java
2018 05 22 09:59:23.978 DEBUG BSTW5MHDXM2 mbatchStandardLegend after java
No comparable output--no error means 'OK' for now.
> print(results)
[1] "Test succeeded for AN_Adjusted.R"
[2] "Test succeeded for AN_Unadjusted.R"
[3] "Test succeeded for Boxplot_AllSamplesData_Structures.R"
[4] "Test succeeded for Boxplot_AllSamplesRLE_Structures.R"
[5] "Test succeeded for Boxplot_Group_Structures.R"
[6] "Test succeeded for CDP_Files.R"
[7] "Test succeeded for CDP_Plot.R"
[8] "Test succeeded for CDP_Structures.R"
[9] "Test succeeded for EB_withNonParametricPriors.R"
[10] "Test succeeded for EB_withParametricPriors.R"
[11] "Test succeeded for EBNPlus_CombineBatches.R"
[12] "Test succeeded for EBNPlus_Correction_Files.R"
[13] "Test succeeded for EBNPlus_Correction_Structures.R"
[14] "Test succeeded for HierarchicalClustering_Structures.R"
[15] "Test succeeded for MP_ByBatch.R"
[16] "Test succeeded for MP_Overall.R"
[17] "Test succeeded for PCA_DualBatch_Structures.R"
[18] "Test succeeded for PCA_Regular_Structures.R"
[19] "Test succeeded for RBN_Pseudoreplicates.R"
[20] "Test succeeded for RBN_Replicates.R"
[21] "Test succeeded for SupervisedClustering_Batches_Structures.R"
[22] "Test succeeded for SupervisedClustering_Pairs_Structures.R"
> |

```

## JDK - no longer needed, docs retained for historical purposes

For OS X, we use the OS X x64 installation instructions and installer from <http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>. Download and install the JDK8 as per Oracle's instructions.



Follow the installation instructions from Oracle.

## rJava Setup for R

For OS X, rJava is installed from the R command line using the command shown below. The other lines demonstrate a test to make sure the install finished properly.

```
# install rJava from CRAN (this installs to site-library so it is available for all
users)

# update to preferred CRAN mirror if desired

# this is at the R command

install.packages("rJava", dependencies=TRUE, type="source", repos =
"http://cran.revolutionanalytics.com")

# check that the Java version returns properly

# should print something like: [1] "1.8.0_144"

# this is at the R command

library(rJava)

.jinit()

J("java.lang.System")$getProperty("java.version")
```