

# Pre-course

---

It is worthwhile to prepare your computer before the course. We will make extensive use of R, various packages and other support software.

You will definitely need R installed. It can be downloaded from: <http://cran.r-project.org/>

Course details and links to materials can be found at <http://www-personal.une.edu.au/~cgondro2/Rcourse.htm>

or download the prac data directly from

<http://www-personal.une.edu.au/~cgondro2/Rcourse/practicals.zip> (it's a big download at ~350Mb).

## Must have software

To build pdf files you'll need *LaTeX* installed. On Windows machines a good option is MiKTeX <http://miktex.org/download>. The *MiKTeX Portable* version also works well if you don't want a full setup on your machine (<http://miktex.org/portable/about>). For Macs, MacTeX is a good option <http://mirror.aarnet.edu.au/pub/CTAN/systems/mac/mactex/MacTeX.pkg>.

## Additional software

An R editor to help writing scripts. Try *Tinn-R* (<http://sourceforge.net/projects/tinn-r/>) if you are on Windows, or choose one you like from [http://sciviews.org/\\_rgui/projects/Editors.html](http://sciviews.org/_rgui/projects/Editors.html).

R's inbuilt editor for Macs is not too bad (in Windows it is horrible – might as well use notepad). TextWrangler is also nice (<http://www.barebones.com/products/textwrangler/>) but you'll need to install a syntax highlighter (<http://www.r-bloggers.com/textwrangler-and-r/>).

We'll also need a proper text editor to view large files. For Windows, have a look at *TextPad* (<http://www.textpad.com/download/>) and for really large files the *Large Text File Viewer* (<http://www.swiftgear.com/ltfviewer/features.html>), *EmEditor* (<http://www.emeditor.com/#download>) or *V File Viewer* (<http://www.fileviewer.com/Download.html>).

A database browser to look at the content of SQLite databases is also handy to have installed. *SQL Studio* works well (<http://sqlitestudio.pl/?act=download>) on Windows and Macs, or *SQLite Browser* (<http://sourceforge.net/projects/sqlitebrowser/>) for Windows.

Visual Studio (<http://www.visualstudio.com/en-us/products/visual-studio-express-vs.aspx>) to create a graphical user interface for applications using R as the *engine* under the hood. Note that this is Windows only; on Mac Xcode can be used (<https://developer.apple.com/xcode/>) and GTK+ (<http://www.gtk.org/>) for Linux.

## Packages

We will need the following R packages for the course (and their dependencies):

made4, car, nlme, lme4, RSQLite, gplots, rrBLUP, pedigree, pegas, hierfstat, snpStats, adegnet, StAMPP, snpStats, ape, affy, ShortRead, affyPLM, ABarray, plier, vsn, limma, GenomicFeatures, GenomicAlignments, edgeR, annotate, biomaRt, bovine.db, org.Bt.eg.db, GO.db, topGO, Category, compiler, bigmemory, snowfall, Rcpp, egDII, R2HTML.

To download and install these packages, start R and cut and paste the following in the R command line (note that you will need an internet connection for this):

```
setRepositories(graphics=F, ind=1:6)
packs=c("made4","car","nlme","lme4","RSQLite","gplots","rrBLUP","pedigree","pegas","hierfstat","sn
pStats","adegenet","StAMPP","snpStats","ape","affy","ShortRead","affyPLM","ABarray","plier","vsn"
,"limma","GenomicFeatures","GenomicAlignments","edgeR","annotate","biomaRt","bovine.db","org
.Bt.eg.db","GO.db","topGO","Category","compiler","bigmemory","snowfall","Rcpp","egDII","R2HTM
L")
install.packages(packs)
```

## Additional Analyses Software

<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> A java program to evaluate quality of RNAseq data

<https://java.com/en/> Java runtime (for FastQC)

<http://www.usadellab.org/cms/?page=trimmomatic> Program for cleaning RNAseq data (also requires Java)

<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml> bowtie2 to align the sequence data against a reference genome

Indexes can be downloaded from Illumina:

[http://support.illumina.com/sequencing/sequencing\\_software/igenome.html](http://support.illumina.com/sequencing/sequencing_software/igenome.html)

<http://www.htslib.org/> samtools to convert SAM to BAM, sort and index sequence data

<http://www.clustal.org/> ClustalW for multiple sequence alignment