LiftOver test

```
library(GenomicRanges)
library(rtracklayer)
library(liftOver)
library(R.utils)
```

liftover using R

```
# Make a GRanges object from this sample range
gr <- GRanges(seqnames = Rle(c("chr3")), ranges = IRanges(start = 195084601, end = 195195500))
# The chain file ('hg19ToHg38.over.chain') can be found at:
# http://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/
path <- "http://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToHg38.over.chain.gz"
download.file(path, "hg19ToHg38.over.chain.gz")
gunzip("hg19ToHg38.over.chain.gz", skip = TRUE)
```

```
[1] "hg19ToHg38.over.chain"
attr(,"temporary")
[1] FALSE
```

```
chain = import.chain("hg19ToHg38.over.chain")
```

```
gr_lift = liftOver(gr, chain)
gr_lift[[1]]
```

```
GRanges object with 21 ranges and 0 metadata columns:
       seqnames              ranges strand
          <Rle>           <IRanges>  <Rle>
   [1]      chr3 195363872-195458801      *
   [2]      chr3 195458814-195458974      *
   [3]      chr3 195458975-195459914      *
   [4]      chr3 195459917-195461019      *
   [5]      chr3 195461021-195461138      *
   ...       ...                 ...    ...
  [17]      chr3 195473622-195473900      *
  [18]      chr3 195473906-195473939      *
  [19]      chr3 195473941-195474362      *
  [20]      chr3 195474364-195474383      *
  [21]      chr3 195474385-195474797      *
  -------
  seqinfo: 1 sequence from an unspecified genome; no seqlengths
```

This results in 21 ranges for the single given range. Note the start coordinate of the first range and the end coordinate of the last range: 195363872 and 195474797

liftover using the web client:

Navigate to the web based tool @ https://genome.ucsc.edu/cgi-bin/hgLiftOver and submit the following coordinates (same as in above R chunk) in the submission box:

`chr3 195084601 195195500` (OR formatted as `chr3:195084601-195195500` also works as input.)

Select "Human" in both the Original and New Genome dropdown menus, and select "Feb. 2009 (GRCh37/hg19)" for the Original Assembly dropdown and "Dec. 2013 (GRCh38/hg38)" for the New Assembly. Click Submit

The web tool returns a downloadable text file that should read:

`chr3    195363872    195474797` (Or `chr3:195363872-195474797`, if it was submitted in the alternate format)

Note that these coordinates match first start coordinate and the last end coordinate from the 21 ranges produced by liftover in R.

liftOver in a terminal

Instructions are here: http://genometoolbox.blogspot.com/2013/11/install-liftover-locally-on-unix.html but also summarized below.

Run `liftOver` on a Mac by downloading it @ http://hgdownload.cse.ucsc.edu/admin/exe/macOSX.x86_64/ and scrolling down to and downloading the `liftOver` package. Once downloaded move it to your working directory, and modify the permissions to make executable (i.e. `chmod u+x liftOver`). Make an input file with the coordinates, (e.g make a file `gr.bed` containing one line, tab-separated:

```
chr3    195084601    195195500
```

To run liftOver, the usage is:

```
liftOver gr.bed hg19ToHg38.over.chain gr_lift.bed gr_lift.unmapped
```

This returns the same results as the web tool. The file `gr_lift.bed` contains

```
chr3    195363872        195474797
```

```
sessionInfo()
```

```
R version 4.0.3 (2020-10-10)
Platform: x86_64-apple-darwin17.0 (64-bit)
Running under: macOS Big Sur 10.16

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
```

```
 [1] stats4     parallel  stats     graphics  grDevices utils     datasets
 [8] methods    base

other attached packages:
 [1] R.utils_2.10.1
 [2] R.oo_1.24.0
 [3] R.methodsS3_1.8.1
 [4] liftOver_1.14.0
 [5] Homo.sapiens_1.3.1
 [6] TxDb.Hsapiens.UCSC.hg19.knownGene_3.2.2
 [7] org.Hs.eg.db_3.12.0
 [8] GO.db_3.12.1
 [9] OrganismDbi_1.32.0
[10] GenomicFeatures_1.42.1
[11] AnnotationDbi_1.52.0
[12] Biobase_2.50.0
[13] gwascat_2.22.0
[14] rtracklayer_1.50.0
[15] GenomicRanges_1.42.0
[16] GenomeInfoDb_1.26.2
[17] IRanges_2.24.1
[18] S4Vectors_0.28.1
[19] BiocGenerics_0.36.0
[20] knitr_1.31

loaded via a namespace (and not attached):
 [1] MatrixGenerics_1.2.1      httr_1.4.2
 [3] bit64_4.0.5               splines_4.0.3
 [5] assertthat_0.2.1          askpass_1.1
 [7] BiocManager_1.30.10       BiocFileCache_1.14.0
 [9] RBGL_1.66.0               blob_1.2.1
[11] BSgenome_1.58.0           GenomeInfoDbData_1.2.4
[13] Rsamtools_2.6.0           yaml_2.2.1
[15] progress_1.2.2            pillar_1.4.7
[17] RSQLite_2.2.3             lattice_0.20-41
[19] glue_1.4.2                digest_0.6.27
[21] XVector_0.30.0            htmltools_0.5.1.1
[23] Matrix_1.3-2              XML_3.99-0.5
[25] pkgconfig_2.0.3           biomaRt_2.46.2
[27] zlibbioc_1.36.0           purrr_0.3.4
[29] BiocParallel_1.24.1       tibble_3.0.6
[31] openssl_1.4.3             generics_0.1.0
[33] ellipsis_0.3.1            cachem_1.0.1
[35] SummarizedExperiment_1.20.0 survival_3.2-7
[37] magrittr_2.0.1            crayon_1.4.0
[39] memoise_2.0.0             evaluate_0.14
[41] xml2_1.3.2                graph_1.68.0
[43] tools_4.0.3               prettyunits_1.1.1
[45] hms_1.0.0                 formatR_1.7
[47] lifecycle_0.2.0           matrixStats_0.58.0
[49] stringr_1.4.0             DelayedArray_0.16.1
[51] snpStats_1.40.0           Biostrings_2.58.0
[53] compiler_4.0.3            rlang_0.4.10
[55] grid_4.0.3                RCurl_1.98-1.2
```

```
[57] VariantAnnotation_1.36.0    rappdirs_0.3.3
[59] bitops_1.0-6               rmarkdown_2.6
[61] DBI_1.1.1                  curl_4.3
[63] R6_2.5.0                   GenomicAlignments_1.26.0
[65] dplyr_1.0.4                fastmap_1.1.0
[67] bit_4.0.4                  readr_1.4.0
[69] stringi_1.5.3              Rcpp_1.0.6
[71] vctrs_0.3.6               dbplyr_2.1.0
[73] tidyselect_1.1.0          xfun_0.20
```