Analysis of CNV around RHT for stiching

Ricardo H. Ramirez-Gonzalez

Testing that we can get all the chromosome:

```
regions <- getFullChromosomes(covs_db)
kable(regions)
```

seqnames	start	end	width	strand
chr1A__part1	1	471268514	471268514	*
chr1A__part2	1	122700062	122700062	*
chr1B__part1	1	438710978	438710978	*
chr1B__part2	1	251131730	251131730	*
chr1D__part1	1	452132239	452132239	*
chr1D__part2	1	43247170	43247170	*
chr2A__part1	1	461985541	461985541	*
chr2A__part2	1	318408600	318408600	*
chr2B__part1	1	453189109	453189109	*
chr2B__part2	1	348037928	348037928	*
chr2D__part1	1	462213638	462213638	*
chr2D__part2	1	189609282	189609282	*
chr3A__part1	1	453944834	453944834	*
chr3A__part2	1	296650429	296650429	*
chr3B__part1	1	448149495	448149495	*
chr3B__part2	1	382653875	382653875	*
chr3D__part1	1	476202024	476202024	*
chr3D__part2	1	139247838	139247838	*
chr4A__part1	1	452361832	452361832	*
chr4A__part2	1	291985939	291985939	*
chr4B__part1	1	450934365	450934365	*
chr4B__part2	1	222522897	222522897	*
chr4D__part1	1	450983086	450983086	*
chr4D__part2	1	58852321	58852321	*
chr5A__part1	1	453009176	453009176	*
chr5A__part2	1	256532231	256532231	*
chr5B__part1	1	451214723	451214723	*
chr5B__part2	1	261770196	261770196	*
chr5D__part1	1	451892688	451892688	*
chr5D__part2	1	114152980	114152980	*
chr6A__part1	1	452282082	452282082	*
chr6A__part2	1	165497012	165497012	*
chr6B__part1	1	452040062	452040062	*
chr6B__part2	1	268910784	268910784	*
chr6D__part1	1	450491070	450491070	*
chr6D__part2	1	23010327	23010327	*
chr7A__part1	1	450010157	450010157	*
chr7A__part2	1	286652930	286652930	*

seqnames	start	end	width	strand
chr7B_part1	1	453799701	453799701	*
chr7B_part2	1	296791078	296791078	*
chr7D_part1	1	453789122	453789122	*
chr7D_part2	1	184858467	184858467	*
chrUn	1	480870096	480870096	*

First we want to stich all the coverages for a single group.

seqnames	start	end	width	strand	norm_cov	cnv_level	line	max_gap
chr4D_part1	18174768	18177343	2576	*	1.7185726	2	WATDE0009	20
chr4D_part1	18461838	18465037	3200	*	0.9087390	1	WATDE0009	20
chr4D_part1	18466638	18473037	6400	*	0.2273286	0	WATDE0009	20
chr4D_part1	18686931	18691130	4200	*	0.7435594	1	WATDE0009	20
chr4D_part1	18692931	18698530	5600	*	0.8697436	1	WATDE0009	20
chr4D_part1	18721885	18724284	2400	*	1.0229174	1	WATDE0009	20
chr4D_part1	18781062	18785418	4357	*	1.6233020	2	WATDE0009	20
chr4D_part1	18961637	18962036	400	*	2.6892301	3	WATDE0009	20

seqnames	start	end	width	strand	norm_cov	cnv_level	line	max_gap
chr4D_part1	18173168	18177343	4176	*	1.9962180	2	WATDE0039	20
chr4D_part1	18461838	18465437	3600	*	1.3768191	1	WATDE0039	20
chr4D_part1	18465638	18468037	2400	*	0.0412816	0	WATDE0039	20
chr4D_part1	18468638	18474386	5749	*	0.3035551	0	WATDE0039	20
chr4D_part1	18686931	18690130	3200	*	0.3328953	0	WATDE0039	20
chr4D_part1	18691331	18700102	8772	*	1.0276488	1	WATDE0039	20
chr4D_part1	18721885	18724684	2800	*	1.4079341	1	WATDE0039	20
chr4D_part1	18725285	18728824	3540	*	0.9052041	1	WATDE0039	20
chr4D_part1	18778862	18785418	6557	*	2.0949333	2	WATDE0039	20
chr4D_part1	18961637	18964636	3000	*	0.2378474	0	WATDE0039	20

seqnames	start	end	width	strand	norm_cov	cnv_level	line	max_gap
chr4D_part1	18462238	18465437	3200	*	1.703345	2	WATDE0821	20
chr4D_part1	18465638	18467437	1800	*	1.620880	2	WATDE0821	20
chr4D_part1	18468638	18472037	3400	*	1.712566	2	WATDE0821	20

seqnames	start	end	width	strand	norm_cov	cnv_level	line	max_gap
chr4D_part1	18472438	18472637	200	*	1.809434	2	WATDE0812	20
chr4D_part1	18694331	18697130	2800	*	1.404684	1	WATDE0812	20

```
sessionInfo()
```

```
## R version 4.0.0 (2020-04-24)
## Platform: x86_64-apple-darwin19.4.0 (64-bit)
## Running under: macOS Catalina 10.15.6
##
```

```
## Matrix products: default
## BLAS/LAPACK: /usr/local/Cellar/openblas/0.3.9/lib/libopenblasp-r0.3.9.dylib
##
## locale:
## [1] en_GB/en_GB/en_GB/C/en_GB/en_GB
##
## attached base packages:
##  [1] grid      parallel  stats4    stats     graphics  grDevices utils
##  [8] datasets  methods   base
##
## other attached packages:
##  [1] bio.cnv_0.1.0.000    fields_10.3          maps_3.3.0
##  [4] spam_2.5-1           dotCall64_1.0-0      ggplot2_3.3.2
##  [7] RSQLite_2.2.0        GenomicRanges_1.40.0 GenomeInfoDb_1.24.0
## [10] IRanges_2.22.2       S4Vectors_0.26.1     BiocGenerics_0.34.0
## [13] knitr_1.28           devtools_2.3.0       usethis_1.6.1
##
## loaded via a namespace (and not attached):
##  [1] bitops_1.0-6        matrixStats_0.56.0   fs_1.4.1
##  [4] bit64_0.9-7         rprojroot_1.3-2      tools_4.0.0
##  [7] backports_1.1.8     R6_2.4.1             DBI_1.1.0
## [10] colorspace_1.4-1    withr_2.2.0          tidyselect_1.1.0
## [13] prettyunits_1.1.1   processx_3.4.2       bit_1.1-15.2
## [16] compiler_4.0.0      chron_2.3-55         cli_2.0.2
## [19] quantreg_5.61       SparseM_1.78         desc_1.2.0
## [22] scales_1.1.1        callr_3.4.3          stringr_1.4.0
## [25] digest_0.6.25       rmarkdown_2.3        XVector_0.28.0
## [28] pkgconfig_2.0.3     htmltools_0.4.0      sessioninfo_1.1.1
## [31] highr_0.8           rlang_0.4.7          rstudioapi_0.11
## [34] generics_0.0.2      dplyr_1.0.1          RCurl_1.98-1.2
## [37] magrittr_1.5        GenomeInfoDbData_1.2.3 Matrix_1.2-18
## [40] Rcpp_1.0.4.6        munsell_0.5.0        fansi_0.4.1
## [43] proto_1.0.0         lifecycle_0.2.0      sqldf_0.4-11
## [46] stringi_1.4.6       yaml_2.2.1           zlibbioc_1.34.0
## [49] pkgbuild_1.0.8      plyr_1.8.6           blob_1.2.1
## [52] crayon_1.3.4        lattice_0.20-41      hash_2.2.6.1
## [55] ps_1.3.3            pillar_1.4.4         optparse_1.6.6
## [58] reshape2_1.4.4      pkgload_1.1.0        glue_1.4.1
## [61] evaluate_0.14       remotes_2.1.1        vctrs_0.3.2
## [64] testthat_2.3.2      MatrixModels_0.4-1   gtable_0.3.0
## [67] getopt_1.20.3       purrr_0.3.4          assertthat_0.2.1
## [70] gsubfn_0.7          xfun_0.14            e1071_1.7-3
## [73] class_7.3-17        tibble_3.0.1         conquer_1.0.1
## [76] memoise_1.1.0       ellipsis_0.3.1
```