VCF_convert

Jinliang Yang

2/2/2017

VCF AGPv3 => AGPv4

setid

```
bcftools annotate --set-id +'%CHROM\_%POS\_%REF\_%FIRST_ALT' merged_flt_ad_c3.vcf.gz -r 3:140088574-160
```

get BED file

```
bcftools query chr3_agpv3_140-160M.vcf.gz -f '%CHROM\t%POS\t%POS\t%ID\n' > chr3_agpv3_140-160M.bed
```

Convert using Assembly Converter

site
```
d <- read.table("~/dbcenter/HapMap/HapMap3/chr3_agpv3_140-160M.bed", header=FALSE)
d$V2 <- d$V2 - 1
d$V2 <- format(d$V2, scientific = FALSE)
d$V3 <- format(d$V3, scientific = FALSE)
write.table(d, "~/dbcenter/HapMap/HapMap3/chr3_agpv3_140-160M.bed", sep="\t", row.names=FALSE, quote=FA
```

```
d <- read.table("~/Documents/Github/Misc/chr3_agpv3_140-160M.bed", header=T)
d$snpid <- 1:nrow(d)
d$V2 <- format(d$V2, scientific = FALSE)
d$V3 <- format(d$V3, scientific = FALSE)
write.table(d[1:350000, -4], "~/Documents/Github/Misc/chr3_agpv3_140-160M_1.bed", sep="\t", row.names=F
write.table(d[350001:nrow(d), -4], "~/Documents/Github/Misc/chr3_agpv3_140-160M_2.bed", sep="\t", row.na
```

```
out1 <- read.table("~/Documents/Github/Misc/largedata/output_chr3_agpv3_140-160M_1.bed")
out2 <- read.table("~/Documents/Github/Misc/largedata/output_chr3_agpv3_140-160M_2.bed")
out <- rbind(out1, out2)
names(out)[1:3] <- c("newchr", "news", "newe")

outd <- merge(d, out, by.x="snpid", by.y="V4")
```

BCFtools Query

```
bcftools query -f '%CHROM\t%POS\t%REF\t%ALT[\t%IUPACGT]\n'  file.vcf.gz > bisnp.txt
```