**FOR SIZE SELECTION RANGE OF 495-605 BP**

width(rfsq)

[1] 46131496

ratio

[1] 25.30427

# Multiply final number of loci returned by this number to get estimated amount from whole genome

I used a scaffold-level Downy Woodpecker genome as reference genome sequence – genomes are relatively conserved in size in birds, and used this because it was same order, though this species has especially high GC and TE content and high TE quantity, despite similar size to other avian genomes, and Downy is only genome-sequence Piciformes species (see Manthey2018)

rfsq <- ref.DNAseq("GCA\_000699005.1\_ASM69900v1\_genomic.fna", subselect.contigs = TRUE, prop.contigs = 0.05)

GC(s2c(rfsq))

[1] 0.458941

#Restriction Enzyme 1

#XbaI

#Restriction Enzyme 2

#EcoRI – HF

rfsq.dig <- insilico.digest(rfsq, cs\_5p1, cs\_3p1, cs\_5p2, cs\_3p2, verbose=TRUE)

Number of restriction sites for the first enzyme: 8371

Number of restriction sites for the second enzyme: 8781

Number of type AB and BA fragments:8334

Number of type AA fragments:4189

Number of type BB fragments:4566

rfsq.selected <- adapt.select(rfsq.dig, type="AB+BA", cs\_5p1, cs\_3p1, cs\_5p2, cs\_3p2)

length(rfsq.selected)

[1] 8334

# wide size selection (200-270):  
wid.rfsq <- size.select(rfsq.selected, min.size = 200, max.size = 270, graph=TRUE, verbose=TRUE)

# narrow size selection (210-260):  
nar.rfsq <- size.select(rfsq.selected,, min.size = 210, max.size = 260, graph=TRUE, verbose=TRUE)

#the resulting fragment characteristics can be further examined: boxplot(list(width(rfsq.selected), width(wid.rfsq), width(nar.rfsq)), names=c("All fragments", "Wide size selection", "Narrow size selection"), ylab="Locus size (bp)")

We expect, in entire Downy Woodpecker genome (same order), around 7,262 loci.

**FOR SIZE SELECTION RANGE OF 480-600 BP**

rfsq <- ref.DNAseq("GCA\_000699005.1\_ASM69900v1\_genomic.fna", subselect.contigs = TRUE, prop.contigs = 0.05)

# length of the reference sequence:

width(rfsq)

[1] 50558283

#ratio for the cross-multiplication of the number of fragments and loci at the genomes scale: [1] 22.23019

# computing GC content:

require(seqinr)

[1] 0.4600547

iTru Adapter Total length – 140 bp, meaning size selection range should be 340 – 460

Number of restriction sites for the first enzyme: 9487

Number of restriction sites for the second enzyme: 10074

Number of type AB and BA fragments:9469

Number of type AA fragments:4733

Number of type BB fragments:5287

> length(rfsq.selected)

[1] 9469

wid.rfsq <- size.select(rfsq.selected, min.size = 340, max.size = 460, graph=TRUE, verbose=TRUE)

439 fragments between 340 and 460 bp

> nar.rfsq <- size.select(rfsq.selected, min.size = 340, max.size = 460, graph=TRUE, verbose=TRUE)

439 fragments between 340 and 460 bp

9,759 loci are anticipated