



HAL
open science

Scripts for filtering SNPs from RAD sequencing; analysis of Mediterranean octocorals along a depth gradient

D. Aurelle

► **To cite this version:**

D. Aurelle. Scripts for filtering SNPs from RAD sequencing; analysis of Mediterranean octocorals along a depth gradient. 2024. hal-04622544

HAL Id: hal-04622544

<https://hal.science/hal-04622544>

Submitted on 24 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SNPfiltR analyses

Didier AURELLE

2024-05-14

Filter SNPs from RAD sequencing

First check the filtering parameters you used in your assembly program (e.g. ipyrad)

From <https://devonderaad.github.io/SNPfiltR/> If too many data, first hard filtering Here first filtering with vcfTools

Install packages:

```
#Install current release from CRAN install.packages("SNPfiltR")
```

```
#Install current development version directly from GitHub library(devtools) devtools::install_github("DevonDeRaad/SNPfiltR")
```

Set working directory and import files

```
setwd("~/Téléchargements/outfiles_RAD_speciation_reference")
```

```
#read vcf into this R session as vcfR  
vcfR <- read.vcfR("RAD_deepevo_10_1kb.recode.vcf")
```

```
## Scanning file to determine attributes.  
## File attributes:  
## meta lines: 10  
## header_line: 11  
## variant count: 34277  
## column count: 132  
## Meta line 10 read in.  
## All meta lines processed.  
## gt matrix initialized.  
## Character matrix gt created.  
## Character matrix gt rows: 34277  
## Character matrix gt cols: 132  
## skip: 0  
## nrows: 34277  
## row_num: 0  
## Processed variant 1000Processed variant 2000Processed variant 3000Processed variant 4000Processed variant 5000  
## All variants processed
```

```
# check the metadata present in your vcf  
vcfR
```

```
## ***** Object of Class vcfR *****  
## 123 samples  
## 1696 CHROMs  
## 34,277 variants  
## Object size: 41.3 Mb  
## 0 percent missing data  
## *****          *****          *****
```

```
# import popmap file. Two column popmap with the same format as stacks, and the columns must be named '  
popmap_RAD <- read.csv("poplist_initiale_deepevo.txt", header = TRUE, sep = ",")
```

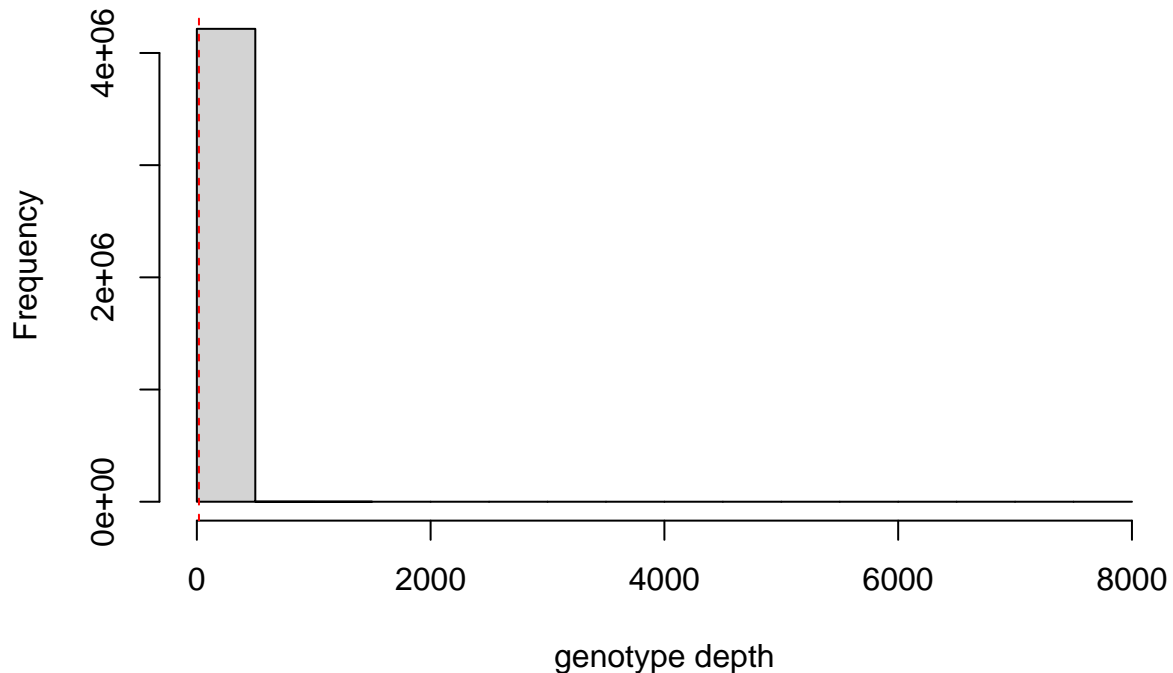
Visualize distributions of depth of sequencing and other statistics

NB: my outfiles from vcftools did not include genotype

```
hard_filter(vcfR=vcfR)
```

```
## no depth cutoff provided, exploratory visualization will be generated.
```

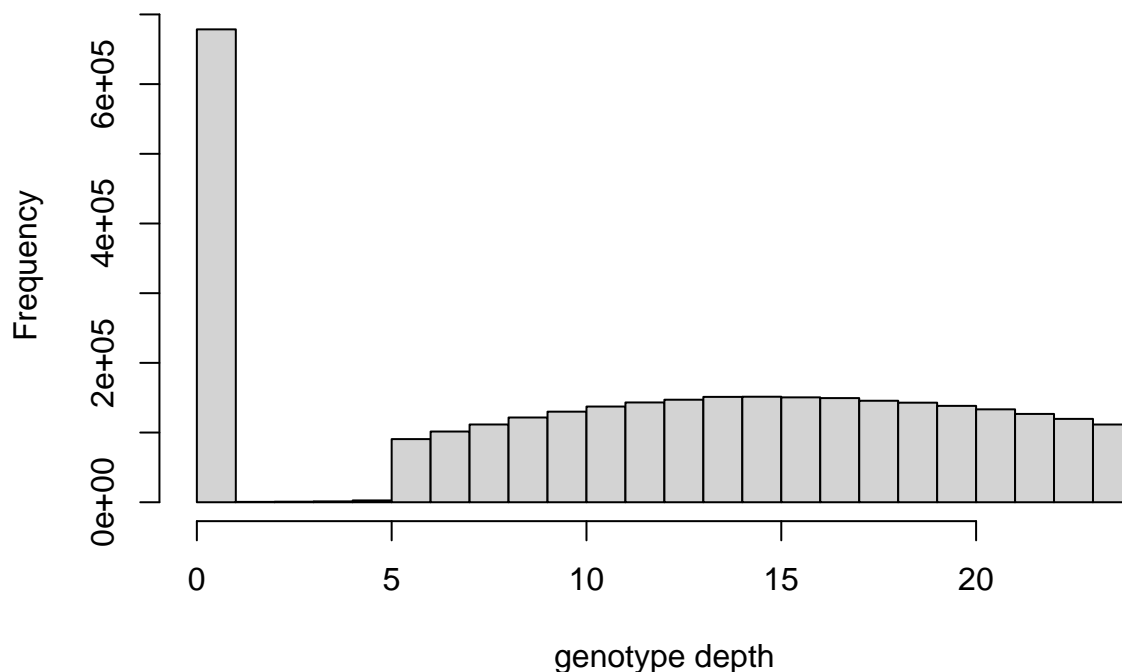
Histogram of dp.matrix



```
## no genotype quality cutoff provided, exploratory visualization will be generated.
```

```
## genotype quality not specified in input vcf
```

Histogram of dp.matrix[dp.matrix < 25]



```
## ***** Object of Class vcfR *****  
## 123 samples  
## 1696 CHROMs  
## 34,277 variants  
## Object size: 41.3 Mb  
## 0 percent missing data  
## *****          *****          *****
```

filter on allele balance (see <https://www.ddocent.com/filtering/>)

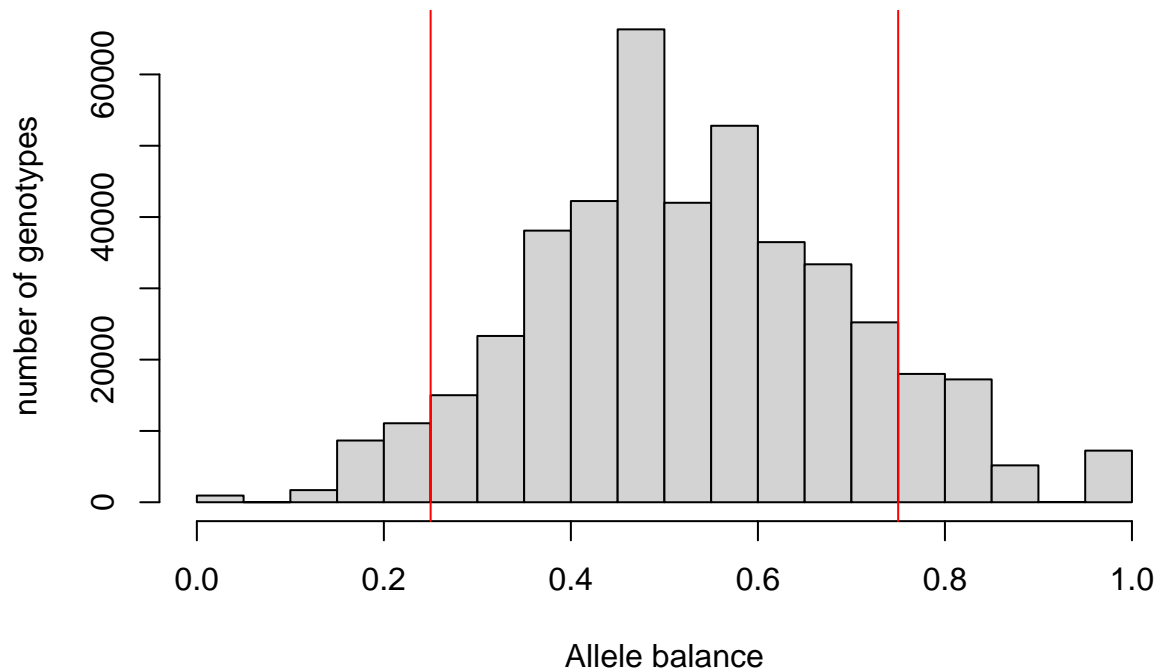
NB: this part can be quite long, that's why it can be useful to do some hard filter before these analyses (see above)

```
vcfR<-filter_allele_balance(vcfR)
```

```
## [1] "Warning, allele depth encoded only as raw CATG counts, must first index out relevant allele depth"  
## [1] "CATG format converted to REF,ALT allele depth"
```

```
## 14.81% of het genotypes (1.89% of all genotypes) fall outside of 0.25 - 0.75 allele balance ratio and
```

allele balance distribution



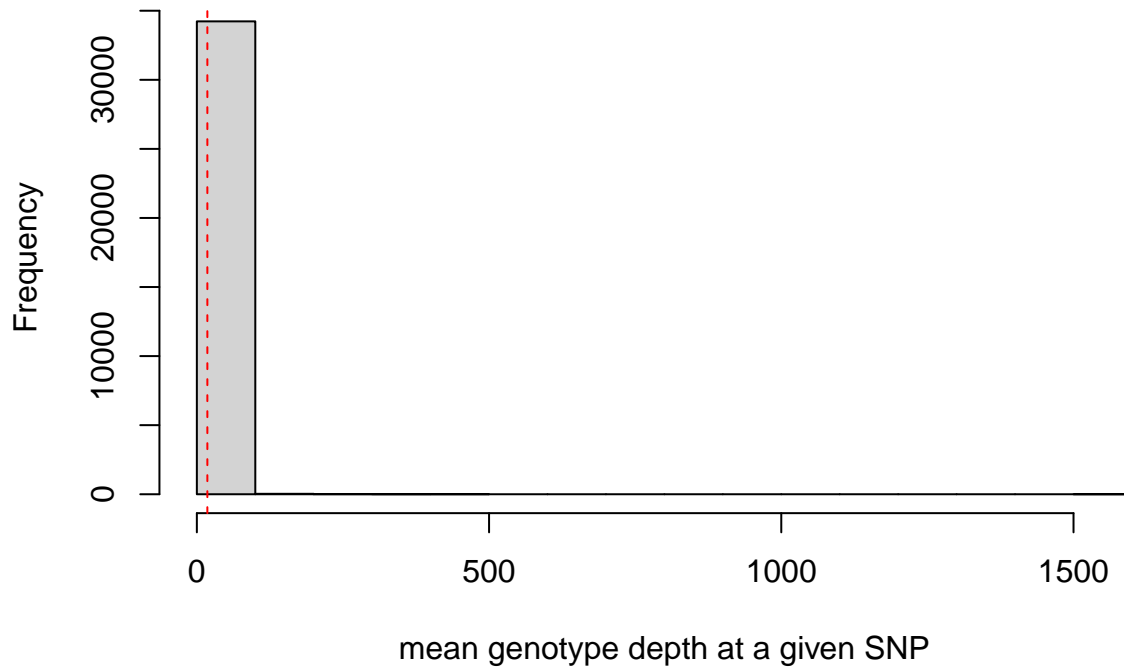
filter on max depth cutoff

```
#visualize and pick appropriate max depth cutoff
```

```
max_depth(vcfR)
```

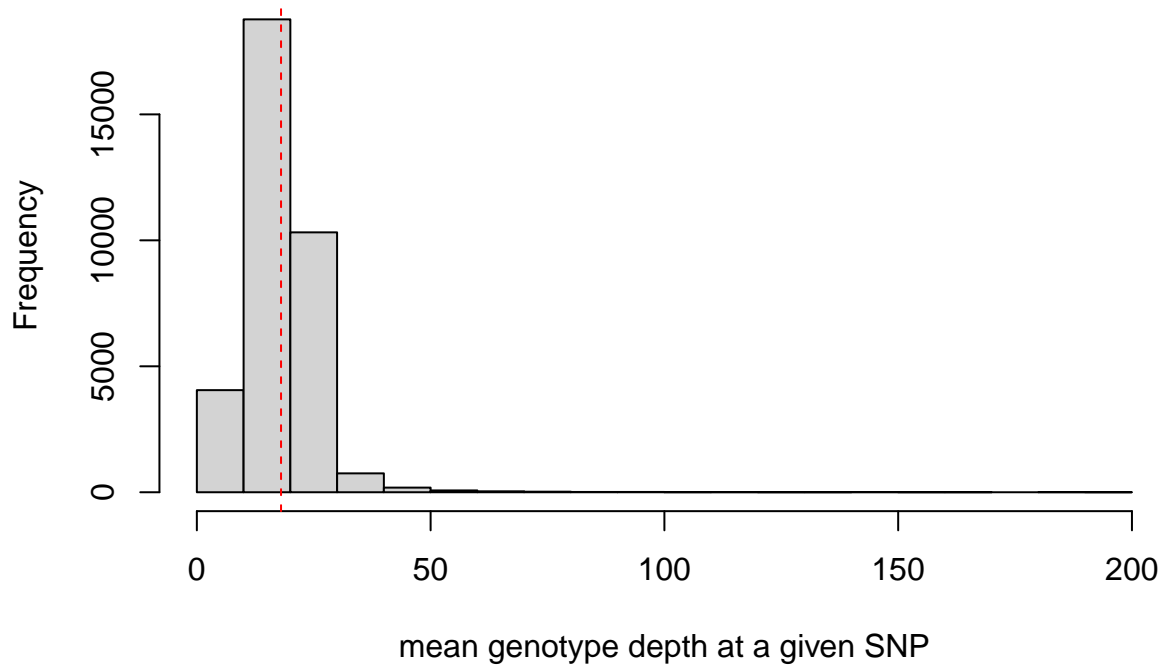
```
## cutoff is not specified, exploratory visualization will be generated.
```

depth of all called SNPs



dashed line indicates a mean depth across all SNPs of 18

depth of SNPs < 200

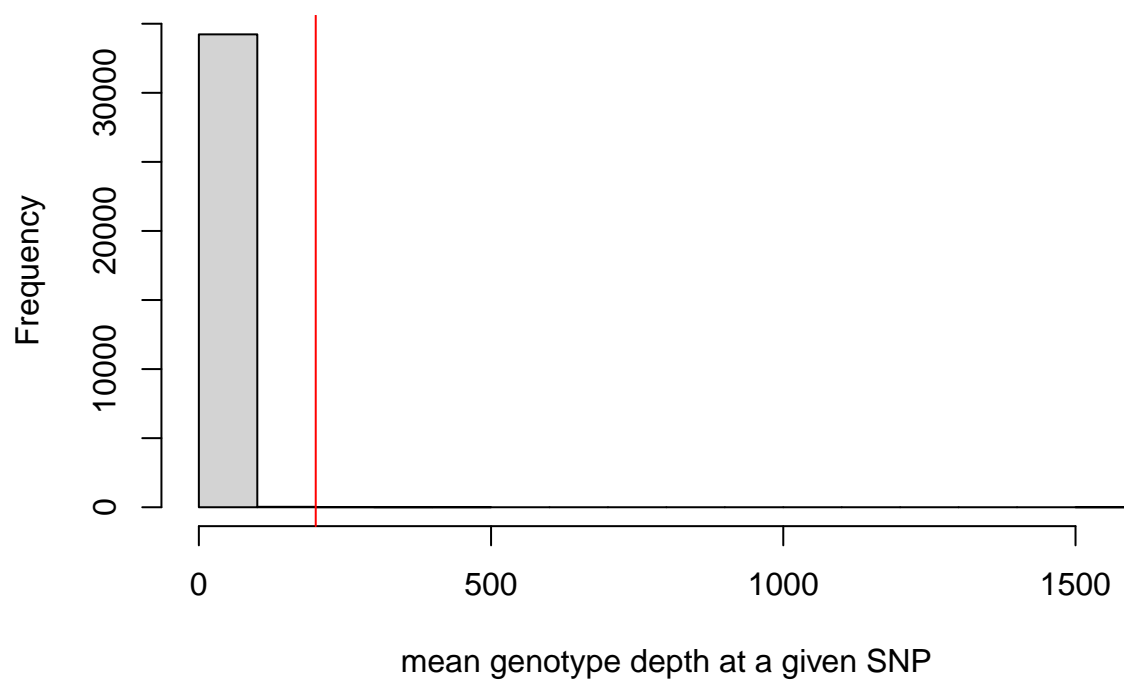


```
#filter vcf by the max depth cutoff you chose  
vcfR<-max_depth(vcfR, maxdepth = 200)
```

```
## maxdepth cutoff is specified, filtered vcfR object will be returned
```

```
## 0.04% of SNPs were above a mean depth of 200 and were removed from the vcf
```

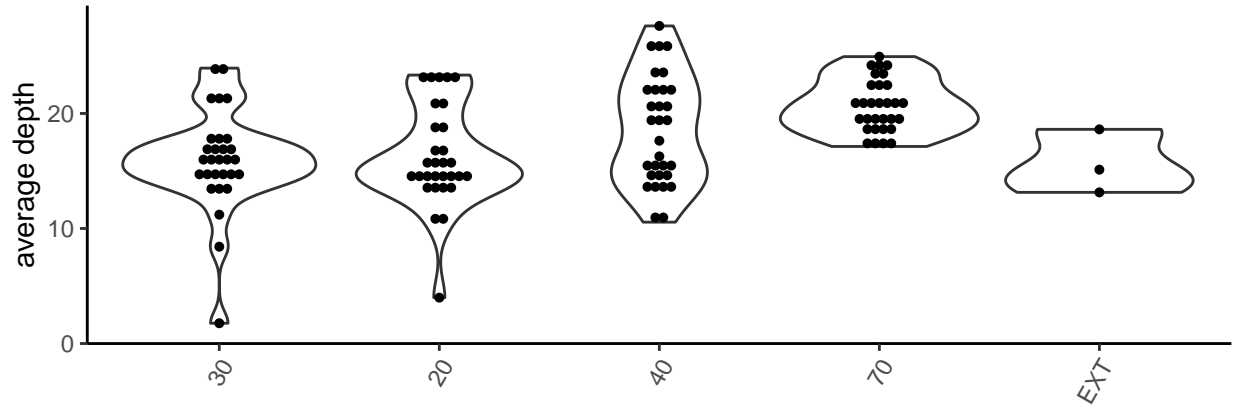
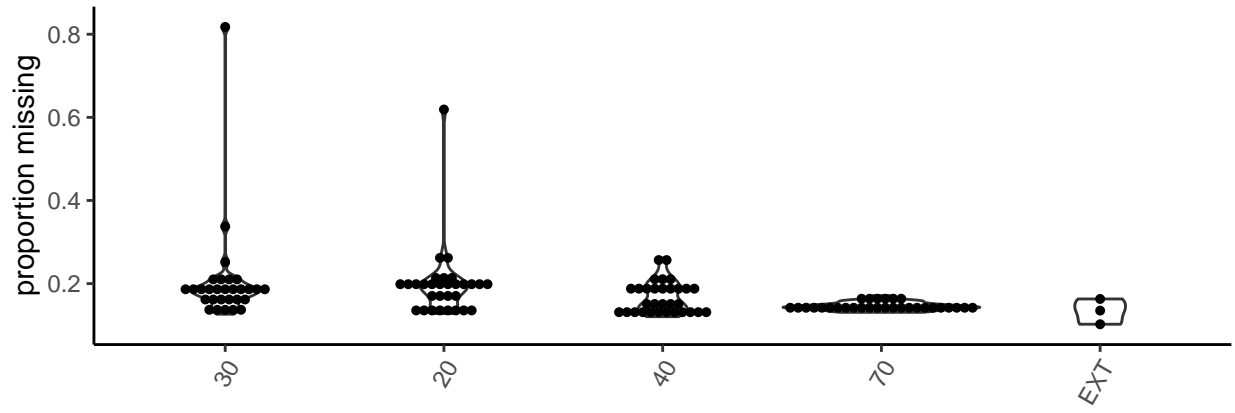
max depth cutoff



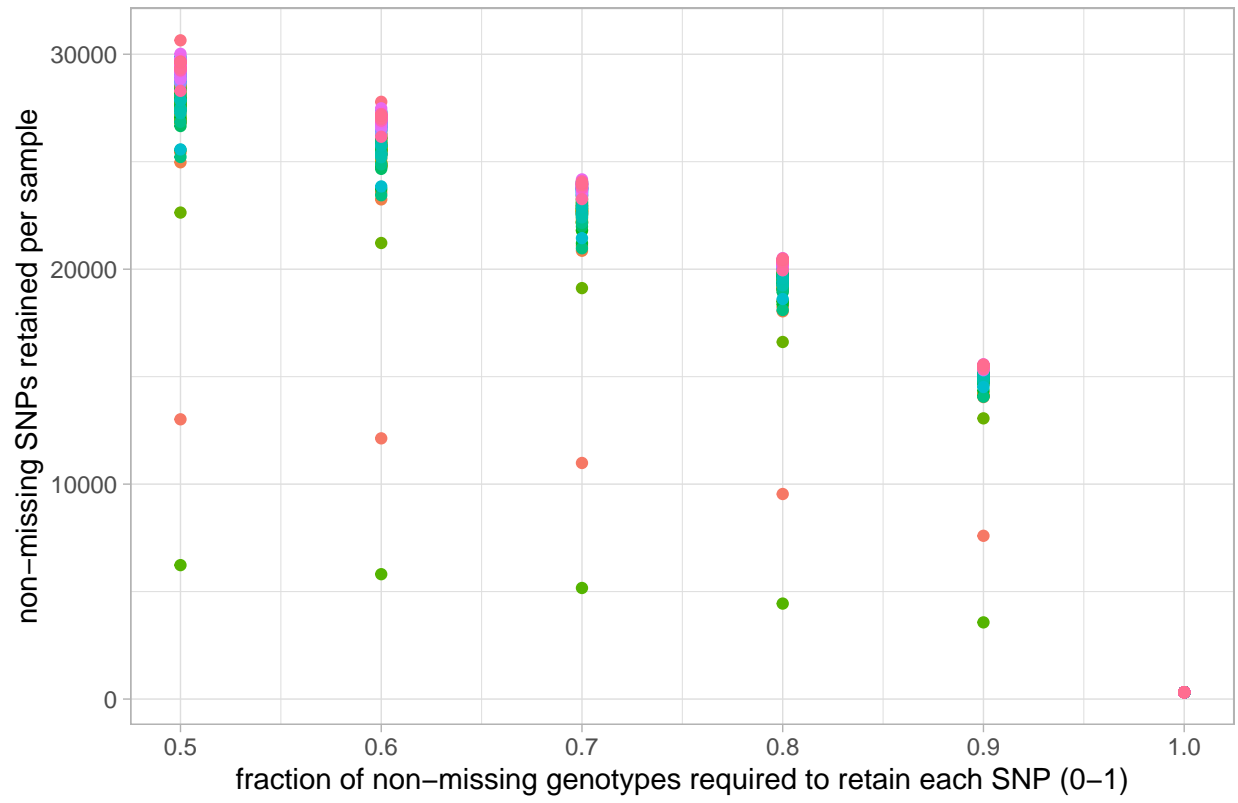
explore the impact of cutoff on missing data per samples

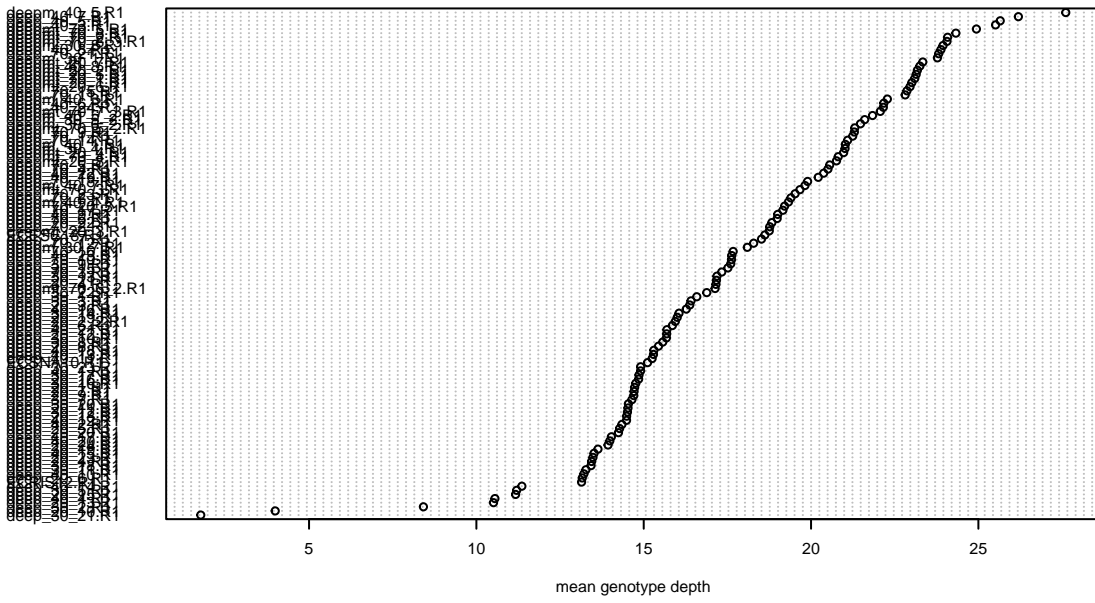
```
#run function to visualize samples  
missing_by_sample(vcfR=vcfR, popmap = popmap_RAD)
```

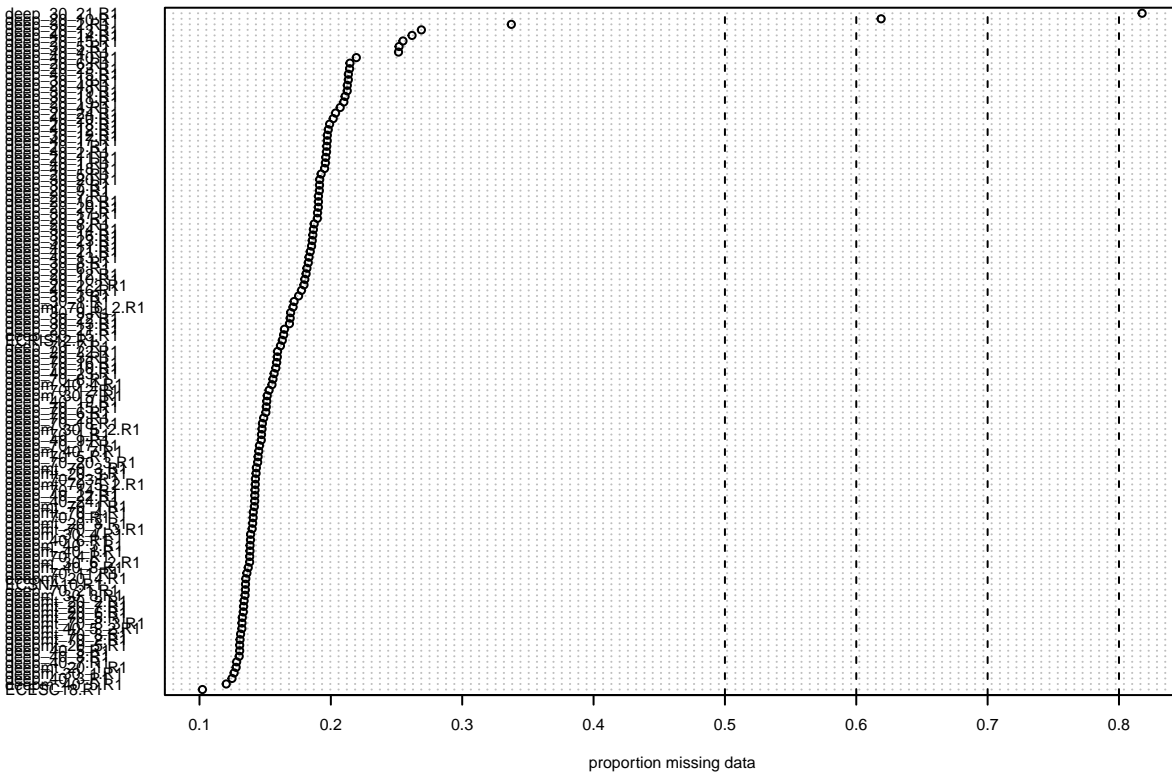
```
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## 'binwidth'.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## 'binwidth'.
```

SNPs retained by filtering scheme







```
## $missing.by.filter
##          indiv  filt snps.retained
## 1      ECESC18.R1  0.5      30647
## 2      ECRIS12.R1  0.5      28574
## 3      ECSNA10.R1  0.5      29530
## 4      deep_20_1.R1 0.5      27430
## 5      deep_20_10.R1 0.5      13015
## 6      deep_20_11.R1 0.5      27448
## 7      deep_20_12.R1 0.5      27962
## 8      deep_20_13.R1 0.5      24974
## 9      deep_20_15.R1 0.5      25428
## 10     deep_20_16.R1 0.5      27641
## 11     deep_20_18.R1 0.5      27345
## 12     deep_20_19.R1 0.5      26978
## 13     deep_20_20.R1 0.5      27650
## 14     deep_20_21.R1 0.5      28518
## 15     deep_20_22.R1 0.5      28700
## 16     deep_20_23.R1 0.5      26828
## 17     deep_20_24.R1 0.5      27190
## 18     deep_20_2_2.R1 0.5      28022
## 19     deep_20_3.R1  0.5      27671
## 20     deep_20_4.R1  0.5      26894
## 21     deep_20_5.R1  0.5      27586
## 22     deep_20_6.R1  0.5      26823
## 23     deep_20_7.R1  0.5      27652
## 24     deep_20_8.R1  0.5      27753
```

## 25	deep_20_9.R1	0.5	27613
## 26	deep_30_1.R1	0.5	28270
## 27	deep_30_10.R1	0.5	27981
## 28	deep_30_11.R1	0.5	26952
## 29	deep_30_12.R1	0.5	27408
## 30	deep_30_13.R1	0.5	28398
## 31	deep_30_14.R1	0.5	27775
## 32	deep_30_15.R1	0.5	28551
## 33	deep_30_16.R1	0.5	27784
## 34	deep_30_17.R1	0.5	27664
## 35	deep_30_18.R1	0.5	26860
## 36	deep_30_19.R1	0.5	26886
## 37	deep_30_2.R1	0.5	22634
## 38	deep_30_20.R1	0.5	27612
## 39	deep_30_21.R1	0.5	6231
## 40	deep_30_22.R1	0.5	28371
## 41	deep_30_23.R1	0.5	27806
## 42	deep_30_3.R1	0.5	28163
## 43	deep_30_4.R1	0.5	27070
## 44	deep_30_5.R1	0.5	25544
## 45	deep_30_6.R1	0.5	27940
## 46	deep_30_7.R1	0.5	27627
## 47	deep_30_8.R1	0.5	27889
## 48	deep_30_9.R1	0.5	28361
## 49	deep_40_1.R1	0.5	27460
## 50	deep_40_10.R1	0.5	26660
## 51	deep_40_11.R1	0.5	27827
## 52	deep_40_12.R1	0.5	27382
## 53	deep_40_13.R1	0.5	27886
## 54	deep_40_14.R1	0.5	25220
## 55	deep_40_15.R1	0.5	26857
## 56	deep_40_16.R1	0.5	28091
## 57	deep_40_17.R1	0.5	27412
## 58	deep_40_18.R1	0.5	27471
## 59	deep_40_19.R1	0.5	28974
## 60	deep_40_2.R1	0.5	27437
## 61	deep_40_20.R1	0.5	27257
## 62	deep_40_21.R1	0.5	27851
## 63	deep_40_22.R1	0.5	29286
## 64	deep_40_23.R1	0.5	28804
## 65	deep_40_24.R1	0.5	29287
## 66	deep_40_3.R1	0.5	29876
## 67	deep_40_4.R1	0.5	25572
## 68	deep_40_5.R1	0.5	29675
## 69	deep_40_6.R1	0.5	29396
## 70	deep_40_7.R1	0.5	29764
## 71	deep_40_8.R1	0.5	29701
## 72	deep_40_9.R1	0.5	29127
## 73	deep_70_1.R1	0.5	29119
## 74	deep_70_10.R1	0.5	28760
## 75	deep_70_11.R1	0.5	29496
## 76	deep_70_12.R1	0.5	28921
## 77	deep_70_14.R1	0.5	29285
## 78	deep_70_15.R1	0.5	29002

## 79	deep_70_16.R1	0.5	28732
## 80	deep_70_17.R1	0.5	29171
## 81	deep_70_18.R1	0.5	29093
## 82	deep_70_2.R1	0.5	29061
## 83	deep_70_20_3.R1	0.5	29227
## 84	deep_70_21.R1	0.5	29537
## 85	deep_70_23.R1	0.5	29288
## 86	deep_70_3.R1	0.5	28701
## 87	deep_70_4.R1	0.5	29423
## 88	deep_70_5.R1	0.5	29206
## 89	deep_70_6.R1	0.5	29014
## 90	deep_70_7.R1	0.5	28627
## 91	deep_70_8.R1	0.5	28826
## 92	deep_70_9.R1	0.5	29341
## 93	deepm_30_1.R1	0.5	29823
## 94	deepm_30_4.R1	0.5	29400
## 95	deepm_30_5_2.R1	0.5	29109
## 96	deepm_30_6_2.R1	0.5	29434
## 97	deepm_30_7.R1	0.5	28971
## 98	deepm_30_8.R1	0.5	29554
## 99	deepm_40_1.R1	0.5	29420
## 100	deepm_40_3.R1	0.5	29414
## 101	deepm_40_4.R1	0.5	28839
## 102	deepm_40_5.R1	0.5	30021
## 103	deepm_40_5_2.R1	0.5	29622
## 104	deepm_40_7.R1	0.5	29197
## 105	deepm_40_8.R1	0.5	29482
## 106	deepmt_20_1.R1	0.5	29773
## 107	deepmt_20_2.R1	0.5	29567
## 108	deepmt_20_3.R1	0.5	29256
## 109	deepmt_20_4.R1	0.5	29526
## 110	deepmt_20_5.R1	0.5	29674
## 111	deepmt_20_6.R1	0.5	29596
## 112	deepmt_20_7.R1	0.5	29596
## 113	deepmt_20_8.R1	0.5	29354
## 114	deepmt_70_1.R1	0.5	29293
## 115	deepmt_70_2.R1	0.5	29663
## 116	deepmt_70_3.R1	0.5	29249
## 117	deepmt_70_4.R1	0.5	29339
## 118	deepmt_70_5.R1	0.5	29627
## 119	deepmt_70_5_2.R1	0.5	29290
## 120	deepmt_70_6_2.R1	0.5	28296
## 121	deepmt_70_7_3.R1	0.5	29357
## 122	deepmt_70_8.R1	0.5	29652
## 123	deepmt_70_8_3.R1	0.5	29614
## 124	ECESC18.R1	0.6	27786
## 125	ECRIS12.R1	0.6	26061
## 126	ECSNA10.R1	0.6	26881
## 127	deep_20_1.R1	0.6	25390
## 128	deep_20_10.R1	0.6	12130
## 129	deep_20_11.R1	0.6	25347
## 130	deep_20_12.R1	0.6	25843
## 131	deep_20_13.R1	0.6	23247
## 132	deep_20_15.R1	0.6	23573

## 133	deep_20_16.R1	0.6	25485
## 134	deep_20_18.R1	0.6	25291
## 135	deep_20_19.R1	0.6	24766
## 136	deep_20_20.R1	0.6	25485
## 137	deep_20_21.R1	0.6	26265
## 138	deep_20_22.R1	0.6	26427
## 139	deep_20_23.R1	0.6	24820
## 140	deep_20_24.R1	0.6	25139
## 141	deep_20_2_2.R1	0.6	25858
## 142	deep_20_3.R1	0.6	25526
## 143	deep_20_4.R1	0.6	24969
## 144	deep_20_5.R1	0.6	25536
## 145	deep_20_6.R1	0.6	24842
## 146	deep_20_7.R1	0.6	25548
## 147	deep_20_8.R1	0.6	25595
## 148	deep_20_9.R1	0.6	25528
## 149	deep_30_1.R1	0.6	26049
## 150	deep_30_10.R1	0.6	25821
## 151	deep_30_11.R1	0.6	24942
## 152	deep_30_12.R1	0.6	25338
## 153	deep_30_13.R1	0.6	26133
## 154	deep_30_14.R1	0.6	25656
## 155	deep_30_15.R1	0.6	26273
## 156	deep_30_16.R1	0.6	25668
## 157	deep_30_17.R1	0.6	25540
## 158	deep_30_18.R1	0.6	24839
## 159	deep_30_19.R1	0.6	24887
## 160	deep_30_2.R1	0.6	21218
## 161	deep_30_20.R1	0.6	25572
## 162	deep_30_21.R1	0.6	5810
## 163	deep_30_22.R1	0.6	26087
## 164	deep_30_23.R1	0.6	25702
## 165	deep_30_3.R1	0.6	25936
## 166	deep_30_4.R1	0.6	24761
## 167	deep_30_5.R1	0.6	23727
## 168	deep_30_6.R1	0.6	25803
## 169	deep_30_7.R1	0.6	25535
## 170	deep_30_8.R1	0.6	25731
## 171	deep_30_9.R1	0.6	26135
## 172	deep_40_1.R1	0.6	25388
## 173	deep_40_10.R1	0.6	24667
## 174	deep_40_11.R1	0.6	25756
## 175	deep_40_12.R1	0.6	25323
## 176	deep_40_13.R1	0.6	25724
## 177	deep_40_14.R1	0.6	23441
## 178	deep_40_15.R1	0.6	24879
## 179	deep_40_16.R1	0.6	25930
## 180	deep_40_17.R1	0.6	25346
## 181	deep_40_18.R1	0.6	25405
## 182	deep_40_19.R1	0.6	26658
## 183	deep_40_2.R1	0.6	25362
## 184	deep_40_20.R1	0.6	25207
## 185	deep_40_21.R1	0.6	25745
## 186	deep_40_22.R1	0.6	26918

## 187	deep_40_23.R1	0.6	26608
## 188	deep_40_24.R1	0.6	26873
## 189	deep_40_3.R1	0.6	27343
## 190	deep_40_4.R1	0.6	23850
## 191	deep_40_5.R1	0.6	27148
## 192	deep_40_6.R1	0.6	26941
## 193	deep_40_7.R1	0.6	27266
## 194	deep_40_8.R1	0.6	27166
## 195	deep_40_9.R1	0.6	26799
## 196	deep_70_1.R1	0.6	26762
## 197	deep_70_10.R1	0.6	26560
## 198	deep_70_11.R1	0.6	27093
## 199	deep_70_12.R1	0.6	26613
## 200	deep_70_14.R1	0.6	26881
## 201	deep_70_15.R1	0.6	26634
## 202	deep_70_16.R1	0.6	26468
## 203	deep_70_17.R1	0.6	26870
## 204	deep_70_18.R1	0.6	26759
## 205	deep_70_2.R1	0.6	26795
## 206	deep_70_20_3.R1	0.6	26898
## 207	deep_70_21.R1	0.6	27026
## 208	deep_70_23.R1	0.6	26882
## 209	deep_70_3.R1	0.6	26485
## 210	deep_70_4.R1	0.6	26940
## 211	deep_70_5.R1	0.6	26836
## 212	deep_70_6.R1	0.6	26727
## 213	deep_70_7.R1	0.6	26319
## 214	deep_70_8.R1	0.6	26537
## 215	deep_70_9.R1	0.6	26955
## 216	deepm_30_1.R1	0.6	27270
## 217	deepm_30_4.R1	0.6	27006
## 218	deepm_30_5_2.R1	0.6	26780
## 219	deepm_30_6_2.R1	0.6	27061
## 220	deepm_30_7.R1	0.6	26674
## 221	deepm_30_8.R1	0.6	27126
## 222	deepm_40_1.R1	0.6	27016
## 223	deepm_40_3.R1	0.6	27061
## 224	deepm_40_4.R1	0.6	26593
## 225	deepm_40_5.R1	0.6	27485
## 226	deepm_40_5_2.R1	0.6	27174
## 227	deepm_40_7.R1	0.6	26834
## 228	deepm_40_8.R1	0.6	27095
## 229	deepmt_20_1.R1	0.6	27282
## 230	deepmt_20_2.R1	0.6	27123
## 231	deepmt_20_3.R1	0.6	26898
## 232	deepmt_20_4.R1	0.6	27115
## 233	deepmt_20_5.R1	0.6	27179
## 234	deepmt_20_6.R1	0.6	27146
## 235	deepmt_20_7.R1	0.6	27127
## 236	deepmt_20_8.R1	0.6	26962
## 237	deepmt_70_1.R1	0.6	26924
## 238	deepmt_70_2.R1	0.6	27216
## 239	deepmt_70_3.R1	0.6	26920
## 240	deepmt_70_4.R1	0.6	27028

## 241	deepmt_70_5.R1	0.6	27213
## 242	deepmt_70_5_2.R1	0.6	26992
## 243	deepmt_70_6_2.R1	0.6	26161
## 244	deepmt_70_7_3.R1	0.6	27009
## 245	deepmt_70_8.R1	0.6	27153
## 246	deepmt_70_8_3.R1	0.6	27141
## 247	ECESC18.R1	0.7	24103
## 248	ECRIS12.R1	0.7	22692
## 249	ECSNA10.R1	0.7	23401
## 250	deep_20_1.R1	0.7	22569
## 251	deep_20_10.R1	0.7	10984
## 252	deep_20_11.R1	0.7	22588
## 253	deep_20_12.R1	0.7	22962
## 254	deep_20_13.R1	0.7	20857
## 255	deep_20_15.R1	0.7	21089
## 256	deep_20_16.R1	0.7	22696
## 257	deep_20_18.R1	0.7	22534
## 258	deep_20_19.R1	0.7	21811
## 259	deep_20_20.R1	0.7	22624
## 260	deep_20_21.R1	0.7	23276
## 261	deep_20_22.R1	0.7	23397
## 262	deep_20_23.R1	0.7	22143
## 263	deep_20_24.R1	0.7	22410
## 264	deep_20_2_2.R1	0.7	22941
## 265	deep_20_3.R1	0.7	22672
## 266	deep_20_4.R1	0.7	22325
## 267	deep_20_5.R1	0.7	22708
## 268	deep_20_6.R1	0.7	22211
## 269	deep_20_7.R1	0.7	22742
## 270	deep_20_8.R1	0.7	22725
## 271	deep_20_9.R1	0.7	22693
## 272	deep_30_1.R1	0.7	23071
## 273	deep_30_10.R1	0.7	22909
## 274	deep_30_11.R1	0.7	22188
## 275	deep_30_12.R1	0.7	22591
## 276	deep_30_13.R1	0.7	23124
## 277	deep_30_14.R1	0.7	22809
## 278	deep_30_15.R1	0.7	23243
## 279	deep_30_16.R1	0.7	22780
## 280	deep_30_17.R1	0.7	22707
## 281	deep_30_18.R1	0.7	22157
## 282	deep_30_19.R1	0.7	22228
## 283	deep_30_2.R1	0.7	19119
## 284	deep_30_20.R1	0.7	22797
## 285	deep_30_21.R1	0.7	5168
## 286	deep_30_22.R1	0.7	23111
## 287	deep_30_23.R1	0.7	22822
## 288	deep_30_3.R1	0.7	22990
## 289	deep_30_4.R1	0.7	21801
## 290	deep_30_5.R1	0.7	21211
## 291	deep_30_6.R1	0.7	22947
## 292	deep_30_7.R1	0.7	22643
## 293	deep_30_8.R1	0.7	22852
## 294	deep_30_9.R1	0.7	23202

## 295	deep_40_1.R1	0.7	22547
## 296	deep_40_10.R1	0.7	21987
## 297	deep_40_11.R1	0.7	22939
## 298	deep_40_12.R1	0.7	22504
## 299	deep_40_13.R1	0.7	22827
## 300	deep_40_14.R1	0.7	20974
## 301	deep_40_15.R1	0.7	22158
## 302	deep_40_16.R1	0.7	23047
## 303	deep_40_17.R1	0.7	22499
## 304	deep_40_18.R1	0.7	22600
## 305	deep_40_19.R1	0.7	23459
## 306	deep_40_2.R1	0.7	22607
## 307	deep_40_20.R1	0.7	22408
## 308	deep_40_21.R1	0.7	22866
## 309	deep_40_22.R1	0.7	23803
## 310	deep_40_23.R1	0.7	23661
## 311	deep_40_24.R1	0.7	23749
## 312	deep_40_3.R1	0.7	24070
## 313	deep_40_4.R1	0.7	21447
## 314	deep_40_5.R1	0.7	23943
## 315	deep_40_6.R1	0.7	23849
## 316	deep_40_7.R1	0.7	24035
## 317	deep_40_8.R1	0.7	23922
## 318	deep_40_9.R1	0.7	23749
## 319	deep_70_1.R1	0.7	23662
## 320	deep_70_10.R1	0.7	23536
## 321	deep_70_11.R1	0.7	23972
## 322	deep_70_12.R1	0.7	23580
## 323	deep_70_14.R1	0.7	23742
## 324	deep_70_15.R1	0.7	23592
## 325	deep_70_16.R1	0.7	23432
## 326	deep_70_17.R1	0.7	23764
## 327	deep_70_18.R1	0.7	23714
## 328	deep_70_2.R1	0.7	23716
## 329	deep_70_20_3.R1	0.7	23792
## 330	deep_70_21.R1	0.7	23821
## 331	deep_70_23.R1	0.7	23767
## 332	deep_70_3.R1	0.7	23491
## 333	deep_70_4.R1	0.7	23756
## 334	deep_70_5.R1	0.7	23717
## 335	deep_70_6.R1	0.7	23659
## 336	deep_70_7.R1	0.7	23310
## 337	deep_70_8.R1	0.7	23464
## 338	deep_70_9.R1	0.7	23829
## 339	deepm_30_1.R1	0.7	23973
## 340	deepm_30_4.R1	0.7	23848
## 341	deepm_30_5_2.R1	0.7	23685
## 342	deepm_30_6_2.R1	0.7	23905
## 343	deepm_30_7.R1	0.7	23651
## 344	deepm_30_8.R1	0.7	23959
## 345	deepm_40_1.R1	0.7	23825
## 346	deepm_40_3.R1	0.7	23927
## 347	deepm_40_4.R1	0.7	23569
## 348	deepm_40_5.R1	0.7	24177

## 349	deepm_40_5_2.R1	0.7	23983
## 350	deepm_40_7.R1	0.7	23824
## 351	deepm_40_8.R1	0.7	23945
## 352	deepmt_20_1.R1	0.7	24027
## 353	deepmt_20_2.R1	0.7	23954
## 354	deepmt_20_3.R1	0.7	23751
## 355	deepmt_20_4.R1	0.7	23948
## 356	deepmt_20_5.R1	0.7	24017
## 357	deepmt_20_6.R1	0.7	24063
## 358	deepmt_20_7.R1	0.7	23929
## 359	deepmt_20_8.R1	0.7	23768
## 360	deepmt_70_1.R1	0.7	23799
## 361	deepmt_70_2.R1	0.7	24009
## 362	deepmt_70_3.R1	0.7	23832
## 363	deepmt_70_4.R1	0.7	23935
## 364	deepmt_70_5.R1	0.7	24087
## 365	deepmt_70_5_2.R1	0.7	23904
## 366	deepmt_70_6_2.R1	0.7	23262
## 367	deepmt_70_7_3.R1	0.7	23909
## 368	deepmt_70_8.R1	0.7	23981
## 369	deepmt_70_8_3.R1	0.7	24018
## 370	ECESC18.R1	0.8	20214
## 371	ECRIS12.R1	0.8	19182
## 372	ECSNA10.R1	0.8	19760
## 373	deep_20_1.R1	0.8	19348
## 374	deep_20_10.R1	0.8	9540
## 375	deep_20_11.R1	0.8	19448
## 376	deep_20_12.R1	0.8	19631
## 377	deep_20_13.R1	0.8	18040
## 378	deep_20_15.R1	0.8	18243
## 379	deep_20_16.R1	0.8	19404
## 380	deep_20_18.R1	0.8	19385
## 381	deep_20_19.R1	0.8	18509
## 382	deep_20_20.R1	0.8	19422
## 383	deep_20_21.R1	0.8	19889
## 384	deep_20_22.R1	0.8	19979
## 385	deep_20_23.R1	0.8	19046
## 386	deep_20_24.R1	0.8	19251
## 387	deep_20_2_2.R1	0.8	19641
## 388	deep_20_3.R1	0.8	19381
## 389	deep_20_4.R1	0.8	19257
## 390	deep_20_5.R1	0.8	19450
## 391	deep_20_6.R1	0.8	19104
## 392	deep_20_7.R1	0.8	19508
## 393	deep_20_8.R1	0.8	19481
## 394	deep_20_9.R1	0.8	19425
## 395	deep_30_1.R1	0.8	19734
## 396	deep_30_10.R1	0.8	19628
## 397	deep_30_11.R1	0.8	19091
## 398	deep_30_12.R1	0.8	19389
## 399	deep_30_13.R1	0.8	19779
## 400	deep_30_14.R1	0.8	19596
## 401	deep_30_15.R1	0.8	19812
## 402	deep_30_16.R1	0.8	19539

## 403	deep_30_17.R1	0.8	19480
## 404	deep_30_18.R1	0.8	19048
## 405	deep_30_19.R1	0.8	19099
## 406	deep_30_2.R1	0.8	16613
## 407	deep_30_20.R1	0.8	19533
## 408	deep_30_21.R1	0.8	4442
## 409	deep_30_22.R1	0.8	19749
## 410	deep_30_23.R1	0.8	19590
## 411	deep_30_3.R1	0.8	19696
## 412	deep_30_4.R1	0.8	18501
## 413	deep_30_5.R1	0.8	18370
## 414	deep_30_6.R1	0.8	19660
## 415	deep_30_7.R1	0.8	19376
## 416	deep_30_8.R1	0.8	19608
## 417	deep_30_9.R1	0.8	19820
## 418	deep_40_1.R1	0.8	19367
## 419	deep_40_10.R1	0.8	18933
## 420	deep_40_11.R1	0.8	19718
## 421	deep_40_12.R1	0.8	19346
## 422	deep_40_13.R1	0.8	19611
## 423	deep_40_14.R1	0.8	18099
## 424	deep_40_15.R1	0.8	19060
## 425	deep_40_16.R1	0.8	19743
## 426	deep_40_17.R1	0.8	19247
## 427	deep_40_18.R1	0.8	19419
## 428	deep_40_19.R1	0.8	19987
## 429	deep_40_2.R1	0.8	19441
## 430	deep_40_20.R1	0.8	19248
## 431	deep_40_21.R1	0.8	19574
## 432	deep_40_22.R1	0.8	20284
## 433	deep_40_23.R1	0.8	20244
## 434	deep_40_24.R1	0.8	20305
## 435	deep_40_3.R1	0.8	20440
## 436	deep_40_4.R1	0.8	18611
## 437	deep_40_5.R1	0.8	20378
## 438	deep_40_6.R1	0.8	20337
## 439	deep_40_7.R1	0.8	20392
## 440	deep_40_8.R1	0.8	20347
## 441	deep_40_9.R1	0.8	20232
## 442	deep_70_1.R1	0.8	20170
## 443	deep_70_10.R1	0.8	20135
## 444	deep_70_11.R1	0.8	20376
## 445	deep_70_12.R1	0.8	20128
## 446	deep_70_14.R1	0.8	20238
## 447	deep_70_15.R1	0.8	20182
## 448	deep_70_16.R1	0.8	20044
## 449	deep_70_17.R1	0.8	20269
## 450	deep_70_18.R1	0.8	20216
## 451	deep_70_2.R1	0.8	20207
## 452	deep_70_20_3.R1	0.8	20293
## 453	deep_70_21.R1	0.8	20268
## 454	deep_70_23.R1	0.8	20265
## 455	deep_70_3.R1	0.8	20109
## 456	deep_70_4.R1	0.8	20225

## 457	deep_70_5.R1	0.8	20223
## 458	deep_70_6.R1	0.8	20148
## 459	deep_70_7.R1	0.8	19953
## 460	deep_70_8.R1	0.8	20078
## 461	deep_70_9.R1	0.8	20264
## 462	deepm_30_1.R1	0.8	20334
## 463	deepm_30_4.R1	0.8	20281
## 464	deepm_30_5_2.R1	0.8	20180
## 465	deepm_30_6_2.R1	0.8	20318
## 466	deepm_30_7.R1	0.8	20209
## 467	deepm_30_8.R1	0.8	20378
## 468	deepm_40_1.R1	0.8	20306
## 469	deepm_40_3.R1	0.8	20371
## 470	deepm_40_4.R1	0.8	20114
## 471	deepm_40_5.R1	0.8	20512
## 472	deepm_40_5_2.R1	0.8	20406
## 473	deepm_40_7.R1	0.8	20314
## 474	deepm_40_8.R1	0.8	20371
## 475	deepmt_20_1.R1	0.8	20410
## 476	deepmt_20_2.R1	0.8	20358
## 477	deepmt_20_3.R1	0.8	20241
## 478	deepmt_20_4.R1	0.8	20353
## 479	deepmt_20_5.R1	0.8	20345
## 480	deepmt_20_6.R1	0.8	20414
## 481	deepmt_20_7.R1	0.8	20357
## 482	deepmt_20_8.R1	0.8	20245
## 483	deepmt_70_1.R1	0.8	20301
## 484	deepmt_70_2.R1	0.8	20431
## 485	deepmt_70_3.R1	0.8	20302
## 486	deepmt_70_4.R1	0.8	20404
## 487	deepmt_70_5.R1	0.8	20500
## 488	deepmt_70_5_2.R1	0.8	20386
## 489	deepmt_70_6_2.R1	0.8	19955
## 490	deepmt_70_7_3.R1	0.8	20345
## 491	deepmt_70_8.R1	0.8	20437
## 492	deepmt_70_8_3.R1	0.8	20462
## 493	ECESC18.R1	0.9	15283
## 494	ECRIS12.R1	0.9	14648
## 495	ECSNA10.R1	0.9	14980
## 496	deep_20_1.R1	0.9	14852
## 497	deep_20_10.R1	0.9	7596
## 498	deep_20_11.R1	0.9	15006
## 499	deep_20_12.R1	0.9	15096
## 500	deep_20_13.R1	0.9	14050
## 501	deep_20_15.R1	0.9	14229
## 502	deep_20_16.R1	0.9	14944
## 503	deep_20_18.R1	0.9	14910
## 504	deep_20_19.R1	0.9	14117
## 505	deep_20_20.R1	0.9	14968
## 506	deep_20_21.R1	0.9	15221
## 507	deep_20_22.R1	0.9	15252
## 508	deep_20_23.R1	0.9	14784
## 509	deep_20_24.R1	0.9	14868
## 510	deep_20_2_2.R1	0.9	15072

## 511	deep_20_3.R1	0.9	14882
## 512	deep_20_4.R1	0.9	14852
## 513	deep_20_5.R1	0.9	14944
## 514	deep_20_6.R1	0.9	14754
## 515	deep_20_7.R1	0.9	15032
## 516	deep_20_8.R1	0.9	14965
## 517	deep_20_9.R1	0.9	14961
## 518	deep_30_1.R1	0.9	15141
## 519	deep_30_10.R1	0.9	15090
## 520	deep_30_11.R1	0.9	14765
## 521	deep_30_12.R1	0.9	14951
## 522	deep_30_13.R1	0.9	15169
## 523	deep_30_14.R1	0.9	15062
## 524	deep_30_15.R1	0.9	15175
## 525	deep_30_16.R1	0.9	15007
## 526	deep_30_17.R1	0.9	14998
## 527	deep_30_18.R1	0.9	14766
## 528	deep_30_19.R1	0.9	14756
## 529	deep_30_2.R1	0.9	13060
## 530	deep_30_20.R1	0.9	14986
## 531	deep_30_21.R1	0.9	3568
## 532	deep_30_22.R1	0.9	15128
## 533	deep_30_23.R1	0.9	15067
## 534	deep_30_3.R1	0.9	15126
## 535	deep_30_4.R1	0.9	14096
## 536	deep_30_5.R1	0.9	14353
## 537	deep_30_6.R1	0.9	15104
## 538	deep_30_7.R1	0.9	14896
## 539	deep_30_8.R1	0.9	15045
## 540	deep_30_9.R1	0.9	15194
## 541	deep_40_1.R1	0.9	14935
## 542	deep_40_10.R1	0.9	14646
## 543	deep_40_11.R1	0.9	15157
## 544	deep_40_12.R1	0.9	14894
## 545	deep_40_13.R1	0.9	15067
## 546	deep_40_14.R1	0.9	14067
## 547	deep_40_15.R1	0.9	14750
## 548	deep_40_16.R1	0.9	15138
## 549	deep_40_17.R1	0.9	14820
## 550	deep_40_18.R1	0.9	14979
## 551	deep_40_19.R1	0.9	15300
## 552	deep_40_2.R1	0.9	15007
## 553	deep_40_20.R1	0.9	14876
## 554	deep_40_21.R1	0.9	15043
## 555	deep_40_22.R1	0.9	15471
## 556	deep_40_23.R1	0.9	15442
## 557	deep_40_24.R1	0.9	15478
## 558	deep_40_3.R1	0.9	15486
## 559	deep_40_4.R1	0.9	14518
## 560	deep_40_5.R1	0.9	15515
## 561	deep_40_6.R1	0.9	15482
## 562	deep_40_7.R1	0.9	15507
## 563	deep_40_8.R1	0.9	15467
## 564	deep_40_9.R1	0.9	15436

## 565	deep_70_1.R1	0.9	15399
## 566	deep_70_10.R1	0.9	15392
## 567	deep_70_11.R1	0.9	15508
## 568	deep_70_12.R1	0.9	15407
## 569	deep_70_14.R1	0.9	15446
## 570	deep_70_15.R1	0.9	15482
## 571	deep_70_16.R1	0.9	15365
## 572	deep_70_17.R1	0.9	15448
## 573	deep_70_18.R1	0.9	15449
## 574	deep_70_2.R1	0.9	15409
## 575	deep_70_20_3.R1	0.9	15442
## 576	deep_70_21.R1	0.9	15422
## 577	deep_70_23.R1	0.9	15442
## 578	deep_70_3.R1	0.9	15359
## 579	deep_70_4.R1	0.9	15398
## 580	deep_70_5.R1	0.9	15408
## 581	deep_70_6.R1	0.9	15410
## 582	deep_70_7.R1	0.9	15299
## 583	deep_70_8.R1	0.9	15395
## 584	deep_70_9.R1	0.9	15455
## 585	deepm_30_1.R1	0.9	15485
## 586	deepm_30_4.R1	0.9	15452
## 587	deepm_30_5_2.R1	0.9	15415
## 588	deepm_30_6_2.R1	0.9	15491
## 589	deepm_30_7.R1	0.9	15414
## 590	deepm_30_8.R1	0.9	15512
## 591	deepm_40_1.R1	0.9	15455
## 592	deepm_40_3.R1	0.9	15508
## 593	deepm_40_4.R1	0.9	15334
## 594	deepm_40_5.R1	0.9	15573
## 595	deepm_40_5_2.R1	0.9	15517
## 596	deepm_40_7.R1	0.9	15436
## 597	deepm_40_8.R1	0.9	15501
## 598	deepmt_20_1.R1	0.9	15504
## 599	deepmt_20_2.R1	0.9	15500
## 600	deepmt_20_3.R1	0.9	15446
## 601	deepmt_20_4.R1	0.9	15476
## 602	deepmt_20_5.R1	0.9	15487
## 603	deepmt_20_6.R1	0.9	15500
## 604	deepmt_20_7.R1	0.9	15471
## 605	deepmt_20_8.R1	0.9	15442
## 606	deepmt_70_1.R1	0.9	15487
## 607	deepmt_70_2.R1	0.9	15519
## 608	deepmt_70_3.R1	0.9	15443
## 609	deepmt_70_4.R1	0.9	15490
## 610	deepmt_70_5.R1	0.9	15566
## 611	deepmt_70_5_2.R1	0.9	15519
## 612	deepmt_70_6_2.R1	0.9	15345
## 613	deepmt_70_7_3.R1	0.9	15499
## 614	deepmt_70_8.R1	0.9	15521
## 615	deepmt_70_8_3.R1	0.9	15538
## 616	ECESC18.R1	1.0	315
## 617	ECRIS12.R1	1.0	315
## 618	ECSNA10.R1	1.0	315

## 619	deep_20_1.R1	1.0	315
## 620	deep_20_10.R1	1.0	315
## 621	deep_20_11.R1	1.0	315
## 622	deep_20_12.R1	1.0	315
## 623	deep_20_13.R1	1.0	315
## 624	deep_20_15.R1	1.0	315
## 625	deep_20_16.R1	1.0	315
## 626	deep_20_18.R1	1.0	315
## 627	deep_20_19.R1	1.0	315
## 628	deep_20_20.R1	1.0	315
## 629	deep_20_21.R1	1.0	315
## 630	deep_20_22.R1	1.0	315
## 631	deep_20_23.R1	1.0	315
## 632	deep_20_24.R1	1.0	315
## 633	deep_20_2_2.R1	1.0	315
## 634	deep_20_3.R1	1.0	315
## 635	deep_20_4.R1	1.0	315
## 636	deep_20_5.R1	1.0	315
## 637	deep_20_6.R1	1.0	315
## 638	deep_20_7.R1	1.0	315
## 639	deep_20_8.R1	1.0	315
## 640	deep_20_9.R1	1.0	315
## 641	deep_30_1.R1	1.0	315
## 642	deep_30_10.R1	1.0	315
## 643	deep_30_11.R1	1.0	315
## 644	deep_30_12.R1	1.0	315
## 645	deep_30_13.R1	1.0	315
## 646	deep_30_14.R1	1.0	315
## 647	deep_30_15.R1	1.0	315
## 648	deep_30_16.R1	1.0	315
## 649	deep_30_17.R1	1.0	315
## 650	deep_30_18.R1	1.0	315
## 651	deep_30_19.R1	1.0	315
## 652	deep_30_2.R1	1.0	315
## 653	deep_30_20.R1	1.0	315
## 654	deep_30_21.R1	1.0	315
## 655	deep_30_22.R1	1.0	315
## 656	deep_30_23.R1	1.0	315
## 657	deep_30_3.R1	1.0	315
## 658	deep_30_4.R1	1.0	315
## 659	deep_30_5.R1	1.0	315
## 660	deep_30_6.R1	1.0	315
## 661	deep_30_7.R1	1.0	315
## 662	deep_30_8.R1	1.0	315
## 663	deep_30_9.R1	1.0	315
## 664	deep_40_1.R1	1.0	315
## 665	deep_40_10.R1	1.0	315
## 666	deep_40_11.R1	1.0	315
## 667	deep_40_12.R1	1.0	315
## 668	deep_40_13.R1	1.0	315
## 669	deep_40_14.R1	1.0	315
## 670	deep_40_15.R1	1.0	315
## 671	deep_40_16.R1	1.0	315
## 672	deep_40_17.R1	1.0	315

## 673	deep_40_18.R1	1.0	315
## 674	deep_40_19.R1	1.0	315
## 675	deep_40_2.R1	1.0	315
## 676	deep_40_20.R1	1.0	315
## 677	deep_40_21.R1	1.0	315
## 678	deep_40_22.R1	1.0	315
## 679	deep_40_23.R1	1.0	315
## 680	deep_40_24.R1	1.0	315
## 681	deep_40_3.R1	1.0	315
## 682	deep_40_4.R1	1.0	315
## 683	deep_40_5.R1	1.0	315
## 684	deep_40_6.R1	1.0	315
## 685	deep_40_7.R1	1.0	315
## 686	deep_40_8.R1	1.0	315
## 687	deep_40_9.R1	1.0	315
## 688	deep_70_1.R1	1.0	315
## 689	deep_70_10.R1	1.0	315
## 690	deep_70_11.R1	1.0	315
## 691	deep_70_12.R1	1.0	315
## 692	deep_70_14.R1	1.0	315
## 693	deep_70_15.R1	1.0	315
## 694	deep_70_16.R1	1.0	315
## 695	deep_70_17.R1	1.0	315
## 696	deep_70_18.R1	1.0	315
## 697	deep_70_2.R1	1.0	315
## 698	deep_70_20_3.R1	1.0	315
## 699	deep_70_21.R1	1.0	315
## 700	deep_70_23.R1	1.0	315
## 701	deep_70_3.R1	1.0	315
## 702	deep_70_4.R1	1.0	315
## 703	deep_70_5.R1	1.0	315
## 704	deep_70_6.R1	1.0	315
## 705	deep_70_7.R1	1.0	315
## 706	deep_70_8.R1	1.0	315
## 707	deep_70_9.R1	1.0	315
## 708	deepm_30_1.R1	1.0	315
## 709	deepm_30_4.R1	1.0	315
## 710	deepm_30_5_2.R1	1.0	315
## 711	deepm_30_6_2.R1	1.0	315
## 712	deepm_30_7.R1	1.0	315
## 713	deepm_30_8.R1	1.0	315
## 714	deepm_40_1.R1	1.0	315
## 715	deepm_40_3.R1	1.0	315
## 716	deepm_40_4.R1	1.0	315
## 717	deepm_40_5.R1	1.0	315
## 718	deepm_40_5_2.R1	1.0	315
## 719	deepm_40_7.R1	1.0	315
## 720	deepm_40_8.R1	1.0	315
## 721	deepmt_20_1.R1	1.0	315
## 722	deepmt_20_2.R1	1.0	315
## 723	deepmt_20_3.R1	1.0	315
## 724	deepmt_20_4.R1	1.0	315
## 725	deepmt_20_5.R1	1.0	315
## 726	deepmt_20_6.R1	1.0	315

```

## 727 deepmt_20_7.R1 1.0 315
## 728 deepmt_20_8.R1 1.0 315
## 729 deepmt_70_1.R1 1.0 315
## 730 deepmt_70_2.R1 1.0 315
## 731 deepmt_70_3.R1 1.0 315
## 732 deepmt_70_4.R1 1.0 315
## 733 deepmt_70_5.R1 1.0 315
## 734 deepmt_70_5_2.R1 1.0 315
## 735 deepmt_70_6_2.R1 1.0 315
## 736 deepmt_70_7_3.R1 1.0 315
## 737 deepmt_70_8.R1 1.0 315
## 738 deepmt_70_8_3.R1 1.0 315
##
## $unfiltered.stats
##
##          samples prop.missing mean.depth
## ECESC18.R1      ECESC18.R1  0.1022386 18.619915
## ECRIS12.R1      ECRIS12.R1  0.1630330 13.142337
## ECSNA10.R1      ECSNA10.R1  0.1350144 15.115100
## deep_20_1.R1    deep_20_1.R1  0.1968888 14.730064
## deep_20_10.R1   deep_20_10.R1 0.6189184  3.988593
## deep_20_11.R1   deep_20_11.R1 0.1961591 14.530832
## deep_20_12.R1   deep_20_12.R1 0.1811575 15.686253
## deep_20_13.R1   deep_20_13.R1 0.2689198 10.512544
## deep_20_15.R1   deep_20_15.R1 0.2549105 11.172086
## deep_20_16.R1   deep_20_16.R1 0.1905554 14.848636
## deep_20_18.R1   deep_20_18.R1 0.1990777 14.488990
## deep_20_19.R1   deep_20_19.R1 0.2099057 16.001862
## deep_20_20.R1   deep_20_20.R1 0.1904679 14.245194
## deep_20_21.R1   deep_20_21.R1 0.1646674 17.173556
## deep_20_22.R1   deep_20_22.R1 0.1594431 18.829667
## deep_20_23.R1   deep_20_23.R1 0.2142836 13.483600
## deep_20_24.R1   deep_20_24.R1 0.2036016 13.934000
## deep_20_2_2.R1  deep_20_2_2.R1 0.1795523 15.944435
## deep_20_3.R1    deep_20_3.R1  0.1897382 16.377228
## deep_20_4.R1    deep_20_4.R1  0.2125325 13.442716
## deep_20_5.R1    deep_20_5.R1  0.1925984 14.279933
## deep_20_6.R1    deep_20_6.R1  0.2145463 13.164219
## deep_20_7.R1    deep_20_7.R1  0.1905554 14.711809
## deep_20_8.R1    deep_20_8.R1  0.1873158 15.441999
## deep_20_9.R1    deep_20_9.R1  0.1912559 14.702691
## deep_30_1.R1    deep_30_1.R1  0.1720223 16.585843
## deep_30_10.R1   deep_30_10.R1 0.1801944 15.677921
## deep_30_11.R1   deep_30_11.R1 0.2108397 13.273052
## deep_30_12.R1   deep_30_12.R1 0.1973266 14.513831
## deep_30_13.R1   deep_30_13.R1 0.1684908 17.188704
## deep_30_14.R1   deep_30_14.R1 0.1867029 16.055105
## deep_30_15.R1   deep_30_15.R1 0.1640837 17.516309
## deep_30_16.R1   deep_30_16.R1 0.1862943 14.764620
## deep_30_17.R1   deep_30_17.R1 0.1898549 14.859899
## deep_30_18.R1   deep_30_18.R1 0.2132913 13.426009
## deep_30_19.R1   deep_30_19.R1 0.2124741 13.633811
## deep_30_2.R1    deep_30_2.R1  0.3373902  8.417407
## deep_30_20.R1   deep_30_20.R1 0.1915185 14.542535
## deep_30_21.R1   deep_30_21.R1 0.8175875  1.764372

```

## deep_30_22.R1	deep_30_22.R1	0.1690453	16.883330
## deep_30_23.R1	deep_30_23.R1	0.1858273	14.908116
## deep_30_3.R1	deep_30_3.R1	0.1754371	16.413966
## deep_30_4.R1	deep_30_4.R1	0.2071039	17.158068
## deep_30_5.R1	deep_30_5.R1	0.2519336	11.203383
## deep_30_6.R1	deep_30_6.R1	0.1817996	15.856176
## deep_30_7.R1	deep_30_7.R1	0.1913726	14.646741
## deep_30_8.R1	deep_30_8.R1	0.1828211	15.568957
## deep_30_9.R1	deep_30_9.R1	0.1693956	17.600755
## deep_40_1.R1	deep_40_1.R1	0.1958381	14.906771
## deep_40_10.R1	deep_40_10.R1	0.2193912	13.213930
## deep_40_11.R1	deep_40_11.R1	0.1853311	15.251477
## deep_40_12.R1	deep_40_12.R1	0.1980270	14.487302
## deep_40_13.R1	deep_40_13.R1	0.1834048	15.291642
## deep_40_14.R1	deep_40_14.R1	0.2618276	11.359450
## deep_40_15.R1	deep_40_15.R1	0.2133497	13.516004
## deep_40_16.R1	deep_40_16.R1	0.1776260	16.271937
## deep_40_17.R1	deep_40_17.R1	0.1972390	14.034587
## deep_40_18.R1	deep_40_18.R1	0.1952252	15.303398
## deep_40_19.R1	deep_40_19.R1	0.1512127	20.214252
## deep_40_2.R1	deep_40_2.R1	0.1967137	14.348698
## deep_40_20.R1	deep_40_20.R1	0.2017920	13.988015
## deep_40_21.R1	deep_40_21.R1	0.1843096	15.692939
## deep_40_22.R1	deep_40_22.R1	0.1421358	20.377697
## deep_40_23.R1	deep_40_23.R1	0.1568164	17.630002
## deep_40_24.R1	deep_40_24.R1	0.1420483	22.160352
## deep_40_3.R1	deep_40_3.R1	0.1247702	25.513172
## deep_40_4.R1	deep_40_4.R1	0.2514958	10.553441
## deep_40_5.R1	deep_40_5.R1	0.1306074	25.652930
## deep_40_6.R1	deep_40_6.R1	0.1388086	22.175219
## deep_40_7.R1	deep_40_7.R1	0.1283017	26.194980
## deep_40_8.R1	deep_40_8.R1	0.1302279	23.881614
## deep_40_9.R1	deep_40_9.R1	0.1470391	19.005049
## deep_70_1.R1	deep_70_1.R1	0.1472142	21.245287
## deep_70_10.R1	deep_70_10.R1	0.1581298	17.622184
## deep_70_11.R1	deep_70_11.R1	0.1360943	22.813078
## deep_70_12.R1	deep_70_12.R1	0.1529930	18.284934
## deep_70_14.R1	deep_70_14.R1	0.1422818	21.093758
## deep_70_15.R1	deep_70_15.R1	0.1508624	22.866616
## deep_70_16.R1	deep_70_16.R1	0.1587427	17.672261
## deep_70_17.R1	deep_70_17.R1	0.1456382	19.159233
## deep_70_18.R1	deep_70_18.R1	0.1478855	19.897645
## deep_70_2.R1	deep_70_2.R1	0.1488778	18.766358
## deep_70_20_3.R1	deep_70_20_3.R1	0.1443248	19.209913
## deep_70_21.R1	deep_70_21.R1	0.1349269	23.814780
## deep_70_23.R1	deep_70_23.R1	0.1425736	19.523316
## deep_70_3.R1	deep_70_3.R1	0.1592388	17.331718
## deep_70_4.R1	deep_70_4.R1	0.1382541	20.505699
## deep_70_5.R1	deep_70_5.R1	0.1446166	20.557763
## deep_70_6.R1	deep_70_6.R1	0.1505414	18.987970
## deep_70_7.R1	deep_70_7.R1	0.1616321	18.519343
## deep_70_8.R1	deep_70_8.R1	0.1559408	19.398010
## deep_70_9.R1	deep_70_9.R1	0.1408516	21.294151
## deepm_30_1.R1	deepm_30_1.R1	0.1263754	23.775851

```

## deepm_30_4.R1      deepm_30_4.R1      0.1388962  21.017106
## deepm_30_5_2.R1   deepm_30_5_2.R1   0.1476228  21.603034
## deepm_30_6_2.R1   deepm_30_6_2.R1   0.1381957  21.478204
## deepm_30_7.R1     deepm_30_7.R1     0.1515629  18.095480
## deepm_30_8.R1     deepm_30_8.R1     0.1346934  23.944078
## deepm_40_1.R1     deepm_40_1.R1     0.1384584  21.024482
## deepm_40_3.R1     deepm_40_3.R1     0.1384876  22.280236
## deepm_40_4.R1     deepm_40_4.R1     0.1551820  19.326214
## deepm_40_5.R1     deepm_40_5.R1     0.1203631  27.610872
## deepm_40_5_2.R1   deepm_40_5_2.R1   0.1322710  21.837522
## deepm_40_7.R1     deepm_40_7.R1     0.1451420  19.818093
## deepm_40_8.R1     deepm_40_8.R1     0.1371158  23.251982
## deepmt_20_1.R1    deepmt_20_1.R1    0.1278931  23.012878
## deepmt_20_2.R1    deepmt_20_2.R1    0.1339346  23.178117
## deepmt_20_3.R1    deepmt_20_3.R1    0.1430698  18.752067
## deepmt_20_4.R1    deepmt_20_4.R1    0.1352771  20.974666
## deepmt_20_5.R1    deepmt_20_5.R1    0.1306074  23.142301
## deepmt_20_6.R1    deepmt_20_6.R1    0.1333508  22.957874
## deepmt_20_7.R1    deepmt_20_7.R1    0.1335843  23.338622
## deepmt_20_8.R1    deepmt_20_8.R1    0.1405014  20.761000
## deepmt_70_1.R1    deepmt_70_1.R1    0.1417564  23.098894
## deepmt_70_2.R1    deepmt_70_2.R1    0.1308700  24.080525
## deepmt_70_3.R1    deepmt_70_3.R1    0.1433033  19.667828
## deepmt_70_4.R1    deepmt_70_4.R1    0.1409392  20.811675
## deepmt_70_5.R1    deepmt_70_5.R1    0.1325920  24.941168
## deepmt_70_5_2.R1  deepmt_70_5_2.R1  0.1424569  21.306727
## deepmt_70_6_2.R1  deepmt_70_6_2.R1  0.1712343  17.132599
## deepmt_70_7_3.R1  deepmt_70_7_3.R1  0.1400928  22.069504
## deepmt_70_8.R1    deepmt_70_8.R1    0.1312786  24.327418
## deepmt_70_8_3.R1  deepmt_70_8_3.R1  0.1323585  24.061694

```

drop samples according to the retained cutoff

```

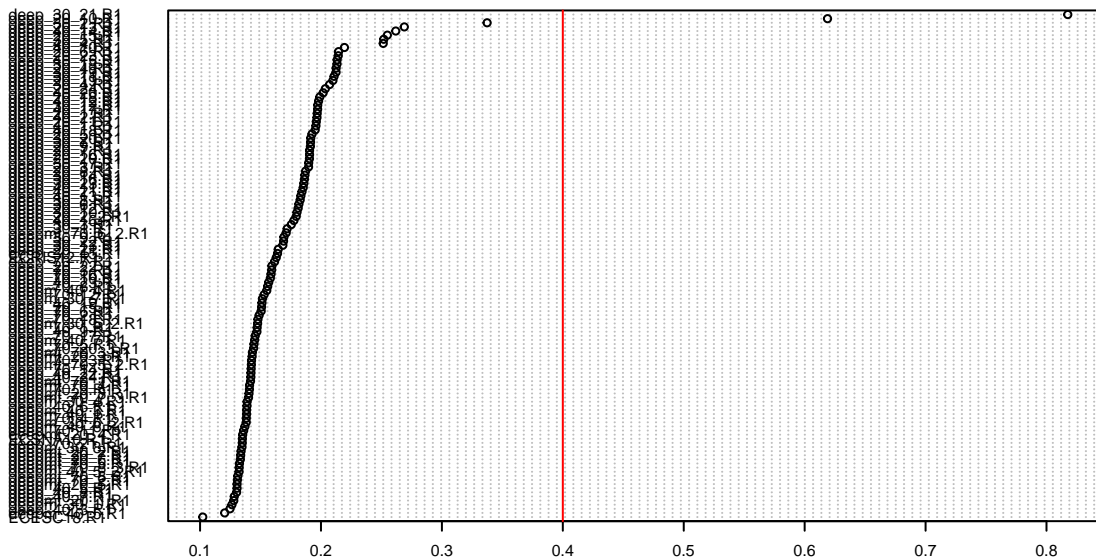
#run function to drop samples above the threshold we want from the vcf
vcfR<-missing_by_sample(vcfR=vcfR, cutoff = .4)

```

```

## 2 samples are above a 0.4 missing data cutoff, and were removed from VCF

```



remove individuals from popmap and remove invariant sites generated by filtering

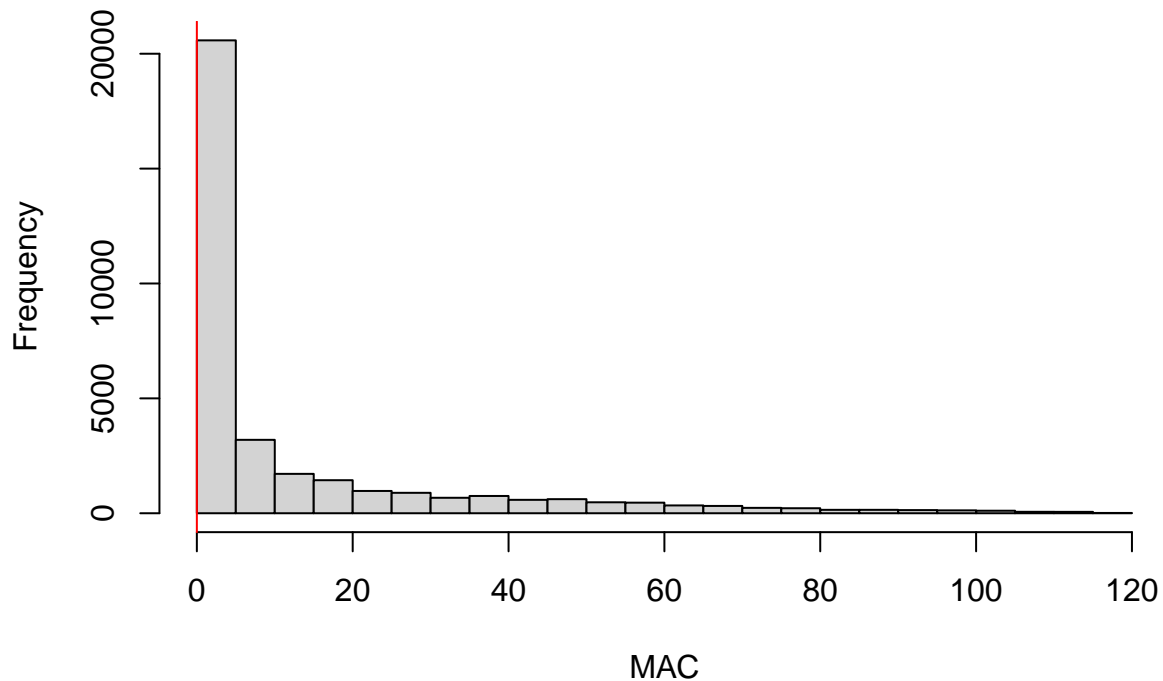
```
#subset popmap to only include retained individuals
popmap_RAD <-popmap_RAD[popmap_RAD$id %in% colnames(vcfR@gt),]

#remove invariant sites generated by dropping individuals
vcfR<-min_mac(vcfR, min.mac = 1)
```

```
## Warning in class(gt.matrix) <- "numeric": NAs introduits lors de la conversion
## automatique
```

```
## 29.63% of SNPs fell below a minor allele count of 1 and were removed from the VCF
```

folded SFS



No filter on minor allele count here, as the datafile can be used for model based inferences.

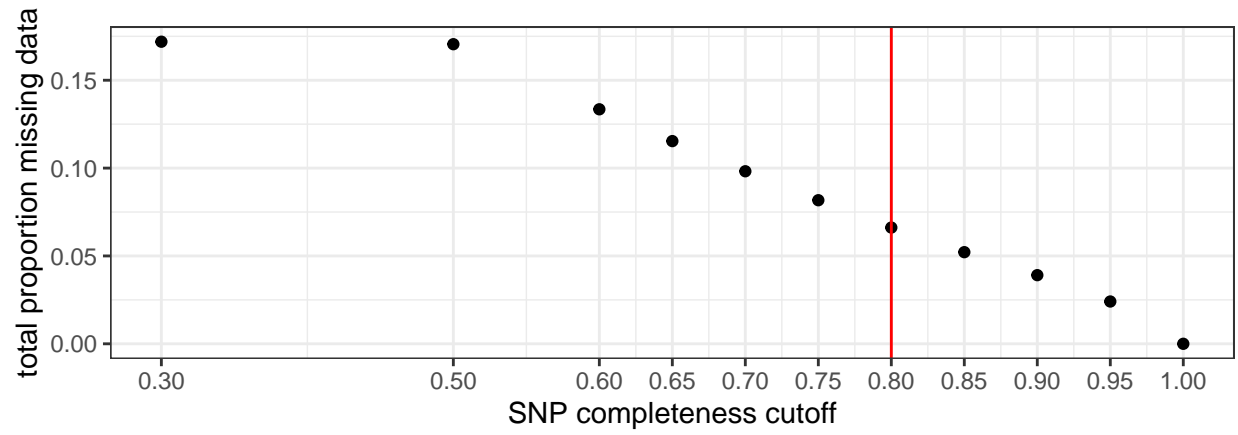
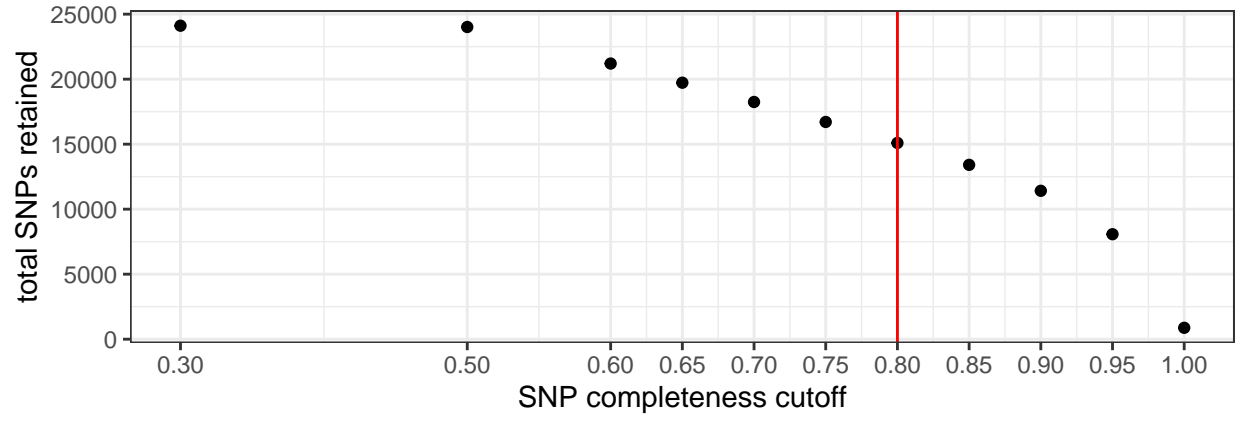
assess if missing data do not create spurious clustering

```
#verify that missing data is not driving clustering patterns among the retained samples  
miss<-assess_missing_data_pca(vcfR=vcfR, popmap = popmap_RAD, thresholds = .8, clustering = FALSE)
```

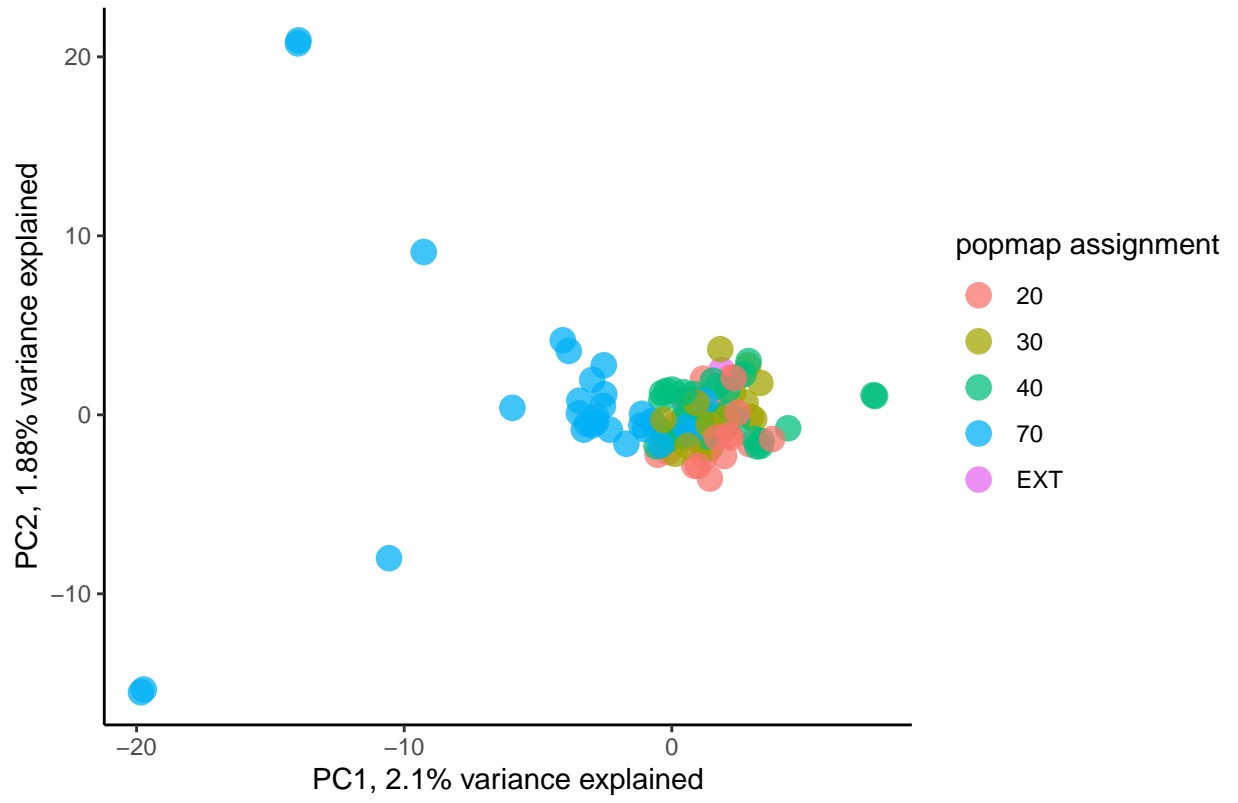
```
## cutoff is specified, filtered vcfR object will be returned
```

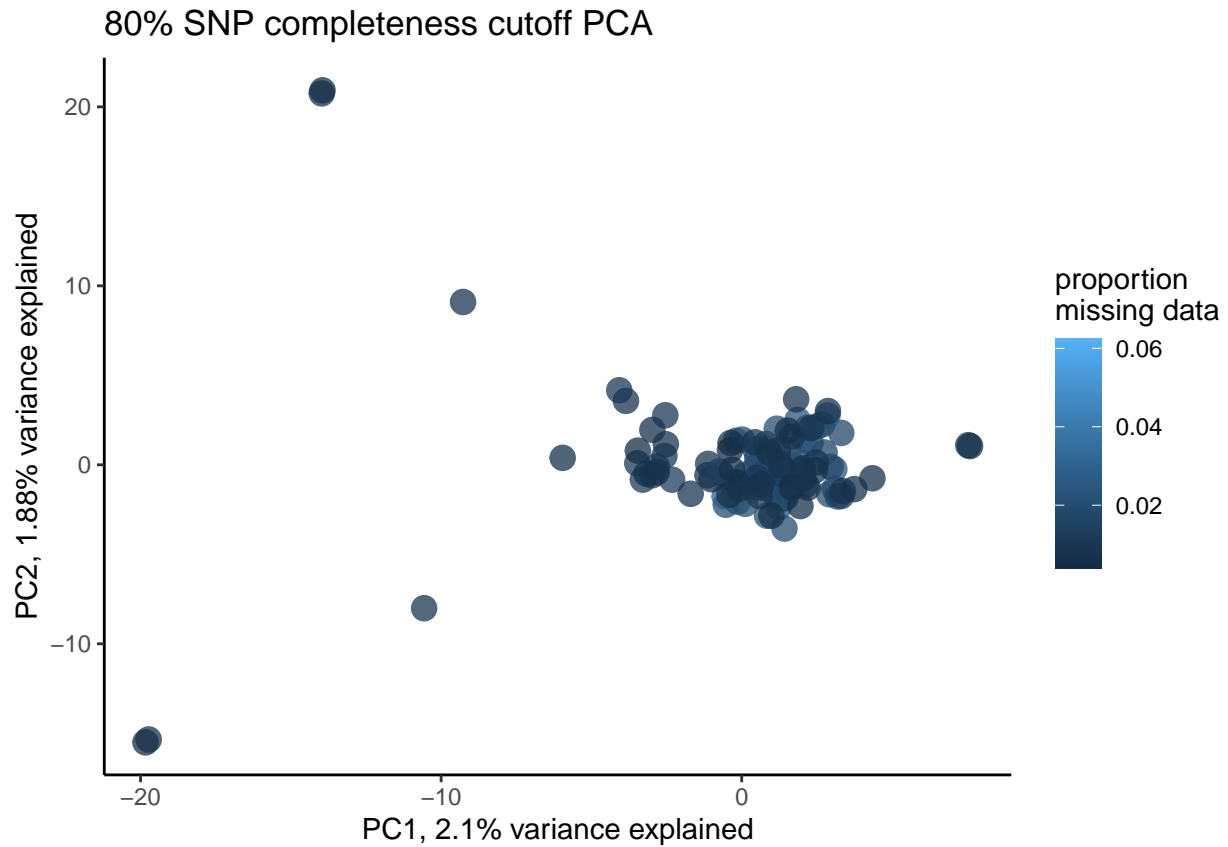
```
## 37.39% of SNPs fell below a completeness cutoff of 0.8 and were removed from the VCF
```

```
## Le chargement a nécessité le package : adegenet
```



80% SNP completeness cutoff PCA





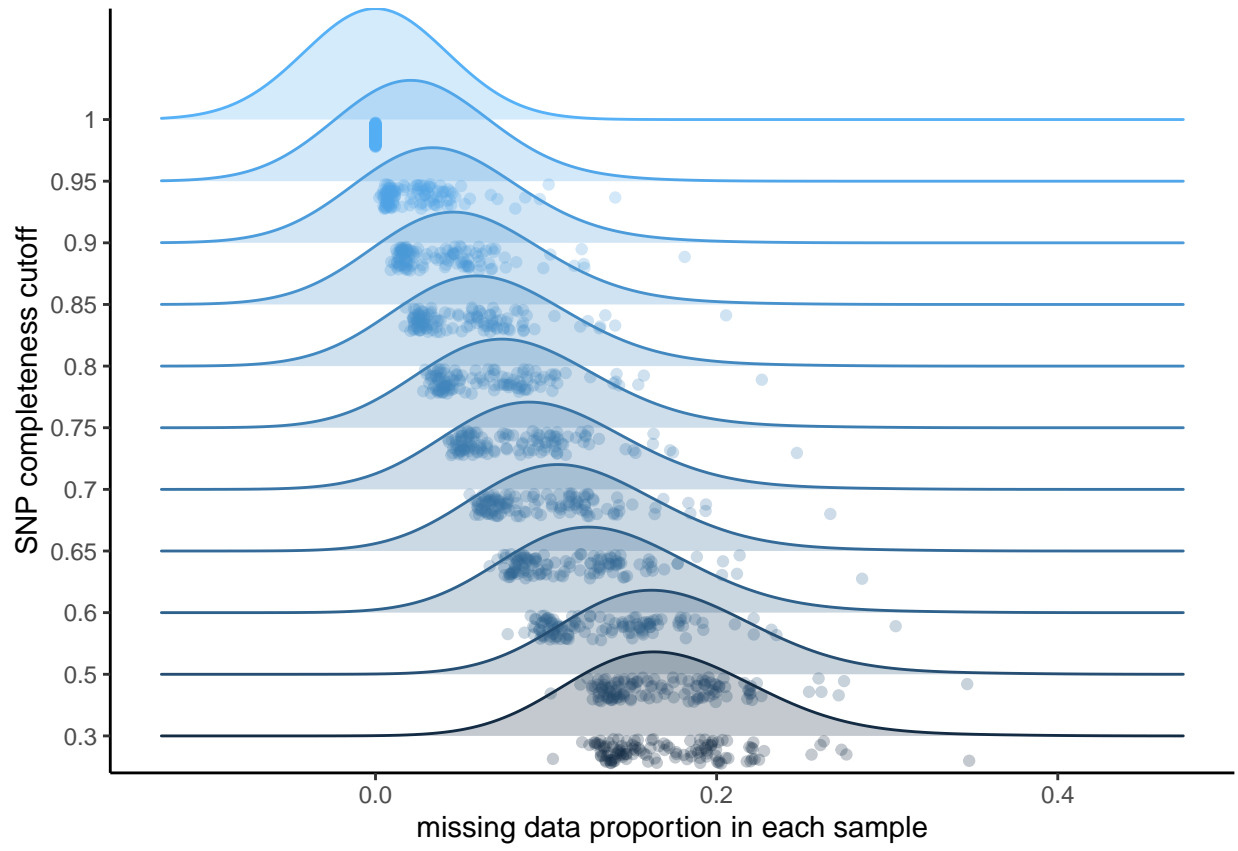
check the impact of various cutoffs on the missingness of each sample

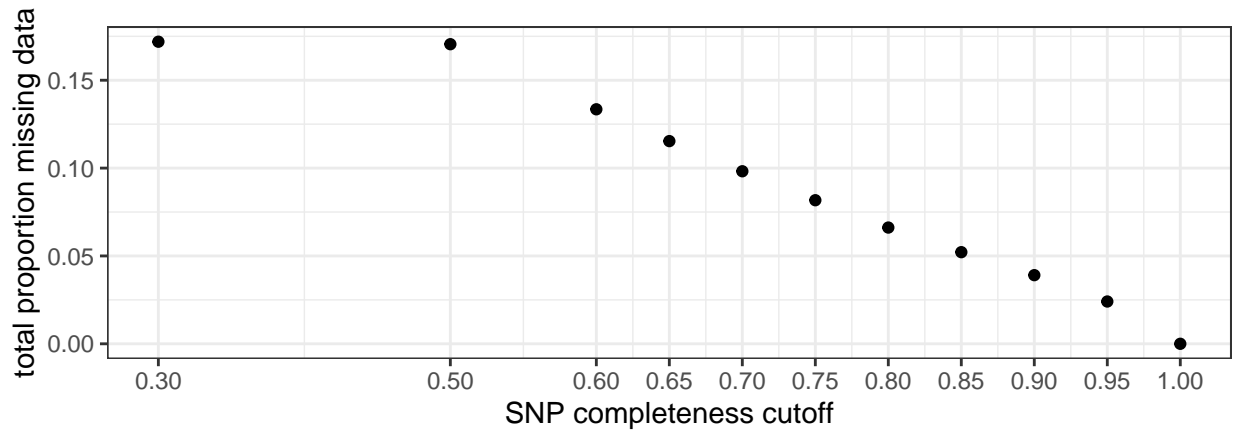
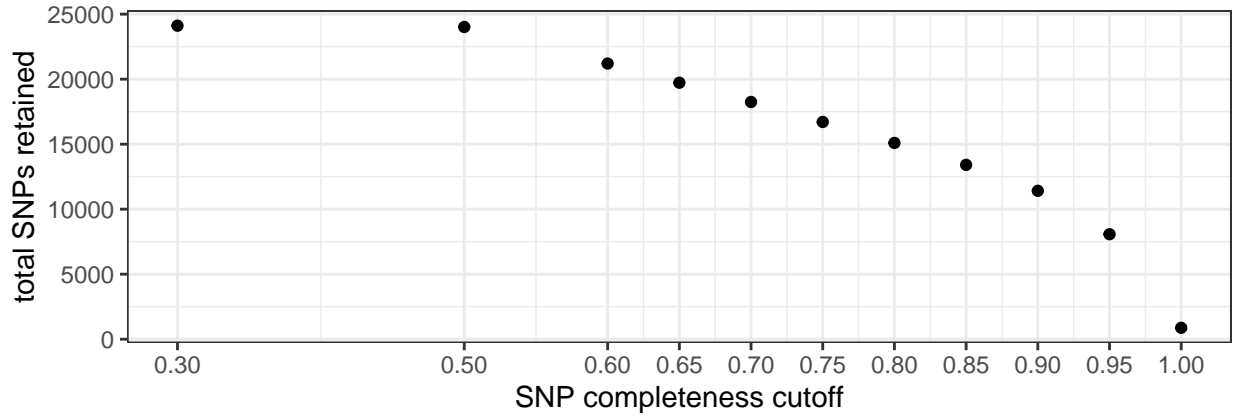
```
missing_by_snp(vcfR)
```

```
## cutoff is not specified, exploratory visualizations will be generated
```

```
## Warning in ggridges::geom_density_ridges(jittered_points = TRUE, position =  
## "raincloud", : Ignoring unknown parameters: 'size'
```

```
## Picking joint bandwidth of 0.0418
```





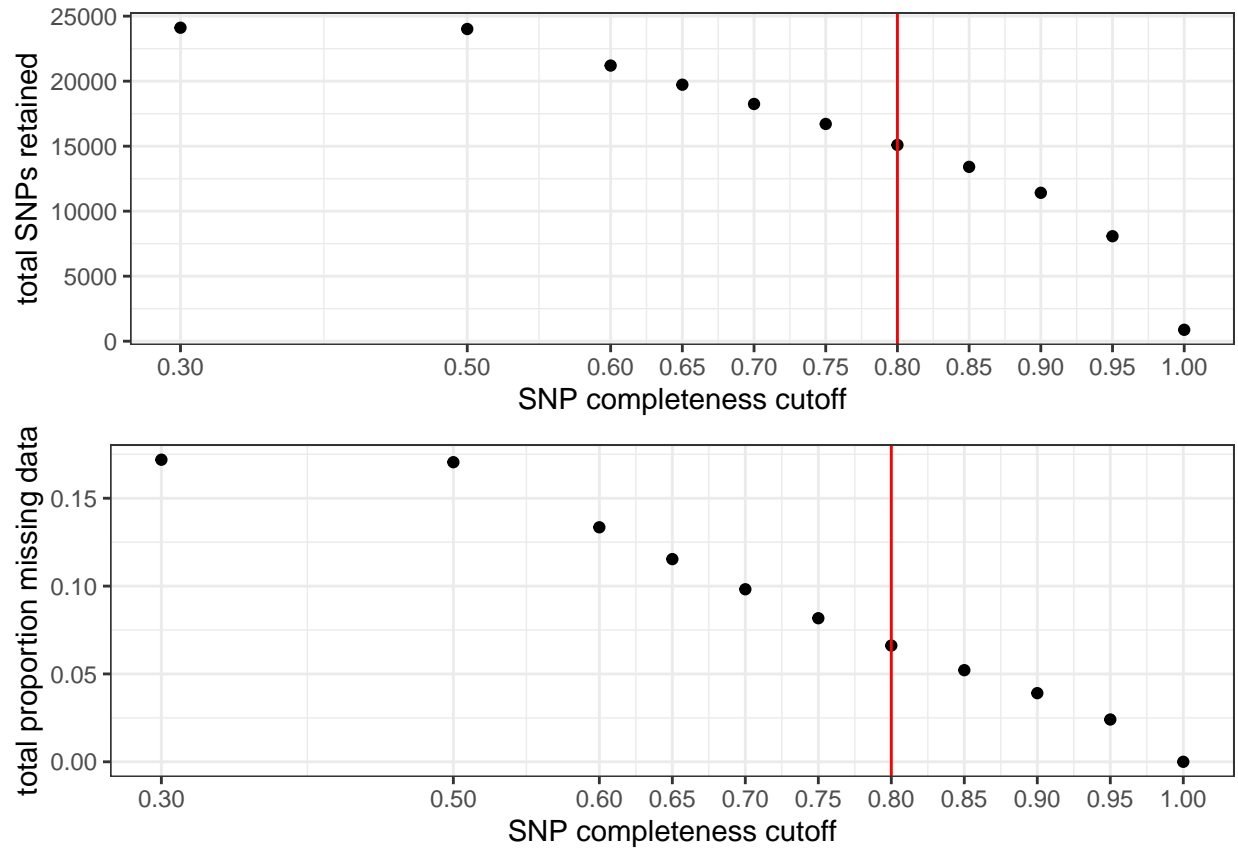
```
##      filt missingness snps.retained
## 1  0.30  0.17192832      24111
## 2  0.50  0.17052046      24012
## 3  0.60  0.13346673      21201
## 4  0.65  0.11535938      19728
## 5  0.70  0.09819463      18244
## 6  0.75  0.08171915      16708
## 7  0.80  0.06615184      15095
## 8  0.85  0.05215012      13409
## 9  0.90  0.03906883      11413
## 10 0.95  0.02406980       8075
## 11 1.00  0.00000000       877
```

choose a threshold and filter vcf

```
vcfR<-missing_by_snp(vcfR, cutoff = .80)
```

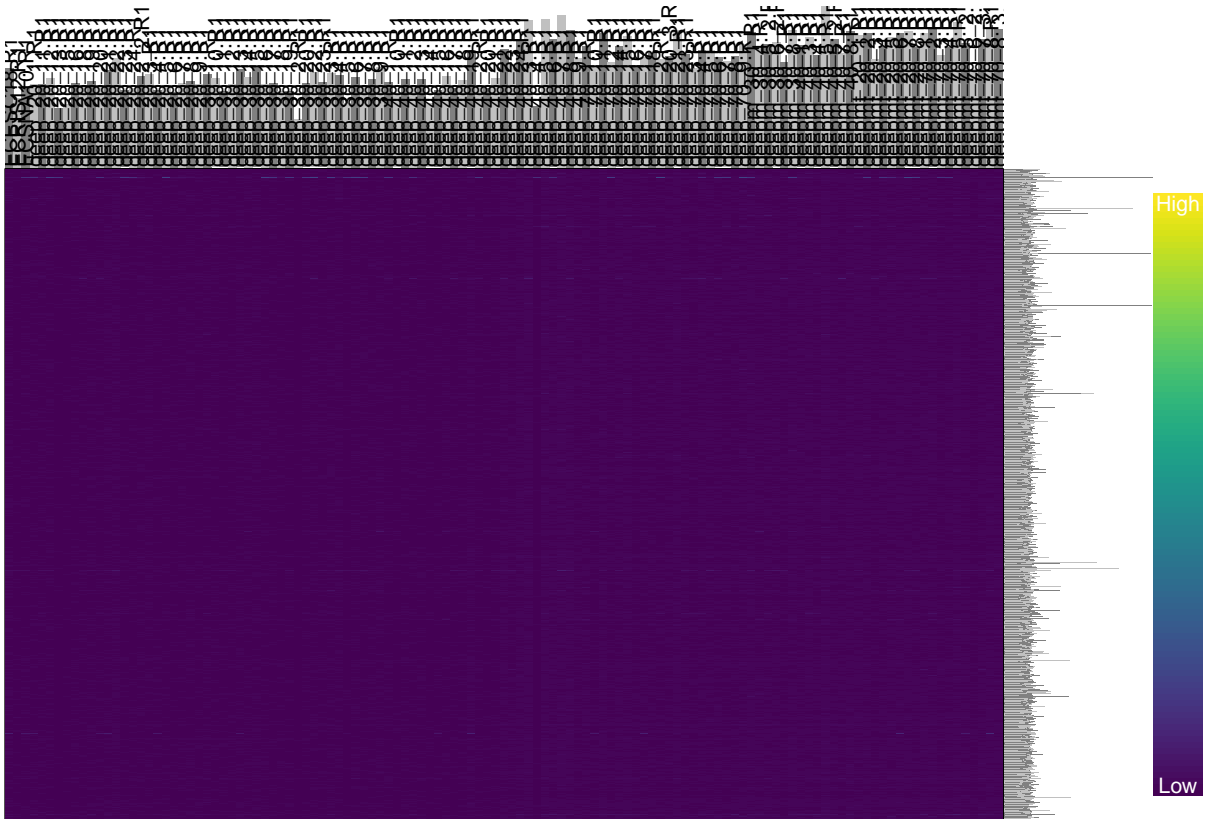
```
## cutoff is specified, filtered vcfR object will be returned
```

```
## 37.39% of SNPs fell below a completeness cutoff of 0.8 and were removed from the VCF
```



check variation of depth among samples

```
#plot depth per snp and per sample
dp <- extract.gt(vcfR, element = "DP", as.numeric=TRUE)
heatmap.bp(dp, rlabels = FALSE)
```



check the obtained vcf

```
vcfR
```

```
## ***** Object of Class vcfR *****
## 121 samples
## 1192 CHROMs
## 15,095 variants
## Object size: 18.2 Mb
## 1.465 percent missing data
## *****          *****          *****
```

write vcf for further analyses

```
#write out vcf with all SNPs
vcfR::write.vcf(vcfR, "RAD_deepevo_SNPfiltr.vcf.gz")
```