HBP D18/D19 2021 Bryophyte Sorting Subsample Experiment

Courtney L. Meier

March 28, 2022

Goal

To determine the efficacy of subsampling the BRY growth form across various proportions of HBP samples in all D18/D19 sites in order to reduce long sort times associated with bryophytes from current-year clipped biomass. Subsampling is only evaluated for the BRY functional group; all other functional groups are sorted from the full clip strip.

Experimental Setup and Analyses

- Subsample the BRY growth from from as many `clipIDs` as possible, in both Distributed and Tower plots.
- For each clipID harvested in the field, test procedure by creating subsamples with the following percentages of the total freshMass:
  - 10%
  - 15%
  - 25%
  - 50%
  - 100% (sum of all subsamples)
- When subsampling is employed, calculate dryMass as follows: `dM = fM * (ssDM/ssFM)`, where:
  - dM = estimated dryMass of BRY in the clipID
  - fM = total freshMass in the clipID
  - $ssDM$ = subsampleDryMass of BRY in the subsample
  - $ssFM$ = fresh mass of all biomass in the subsample (BRY + other clipped material)
- Compare dryMass results calculated via subsampling with dryMass obtained with no subsampling, and use linear mixed-effects models to analyze results.

Procedure

1. Perform clip harvest in the field as normal, and bring clipped biomass back to the laboratory in cold storage as normal.
2. Thoroughly mix biomass from each clipID to homogenize as thoroughly as possible.
   a. For large amounts of biomass, and when there is more than one bag of biomass for a given clipID, use a large bag, box, tray or equivalent vessel to mix the biomass.
3. For each clipID, weigh and record to 0.01 g:
   a. **`freshMass`** = total fresh mass in the clipID (all clipped material)

4. Based on the **freshMass**, calculate the desired subsample fresh masses for testing. For example, assuming **freshMass** = 100 g, the target subsample fresh masses are:
    a. 10% subsample –> 10 g
    b. 15% subsample –> 15 g
    c. 25% subsample –> 25 g
    d. 50% subsample –> 50 g
5. Label a coin envelope for each subsample above with the information below.
    a. **subsampleTest**: 10%, 15%, 25% or 50%.
    b. **clipID**
    c. **collectDate**
6. Weigh each subsample created above (all clipped material), and record the information below.
    a. **subsampleTest**: as above
    b. **clipID**
    c. **collectDate**
    d. **subsampleFreshMass**: To the nearest 0.01 g; for subsamples < 0.5 g total mass, weigh to the nearest 0.0001 g
7. Sort BRY from other clipped material for each subsample, and place sorted, current-year BRY biomass into the corresponding labeled coin envelope.
    a. Sorted OSD may be discarded at this point.
8. Dry subsamples until dry; minimum of 48 h @ 65 °C, track drying progress as normal.
9. Remove dry samples from the oven one at a time, and immediately weigh and record:
    a. **subsampleDryMass**: To the nearest 0.01 g; for masses < 0.5 g, weigh to the nearest 0.0001 g.

Analyses and Graphs

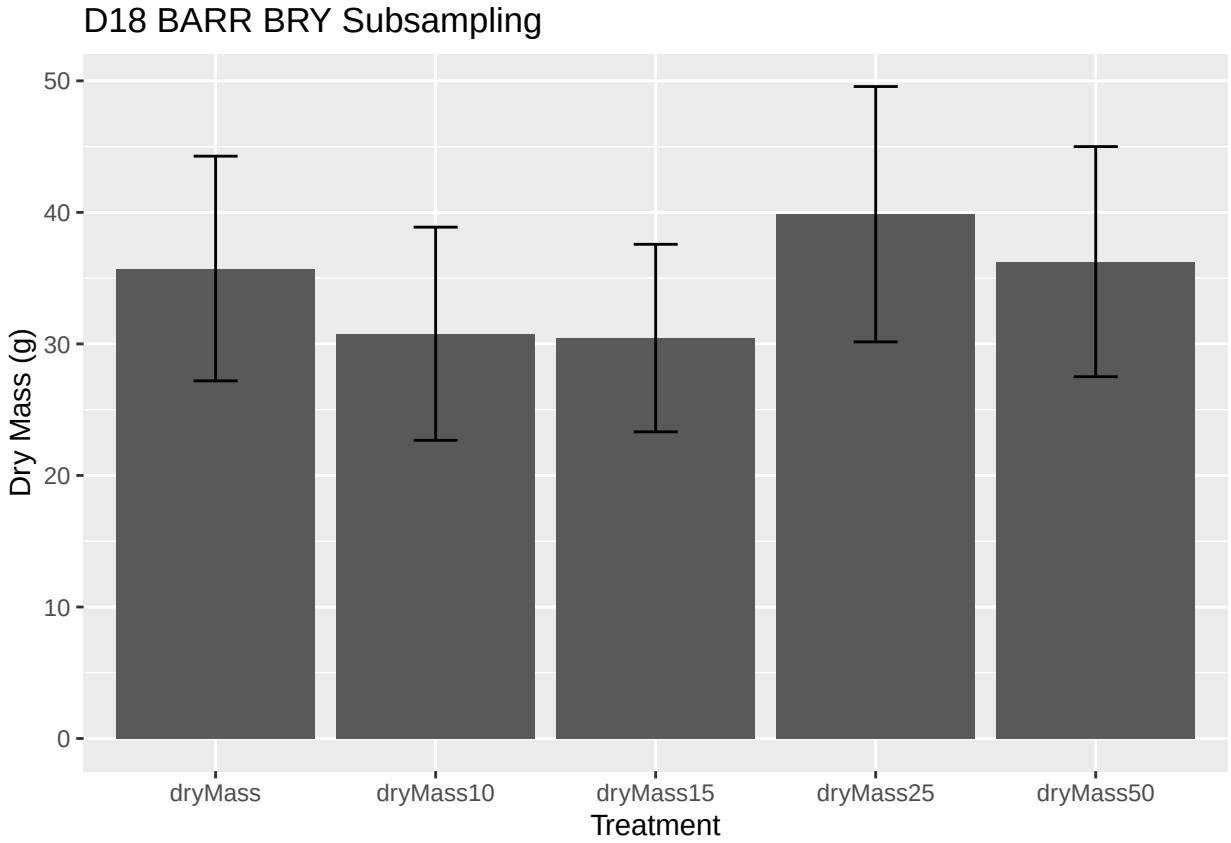
Using formula from above, `dryMass` values are calculated for BRY from `freshMass` value from whole clip-strip, as well as subsample fresh and dry mass values.

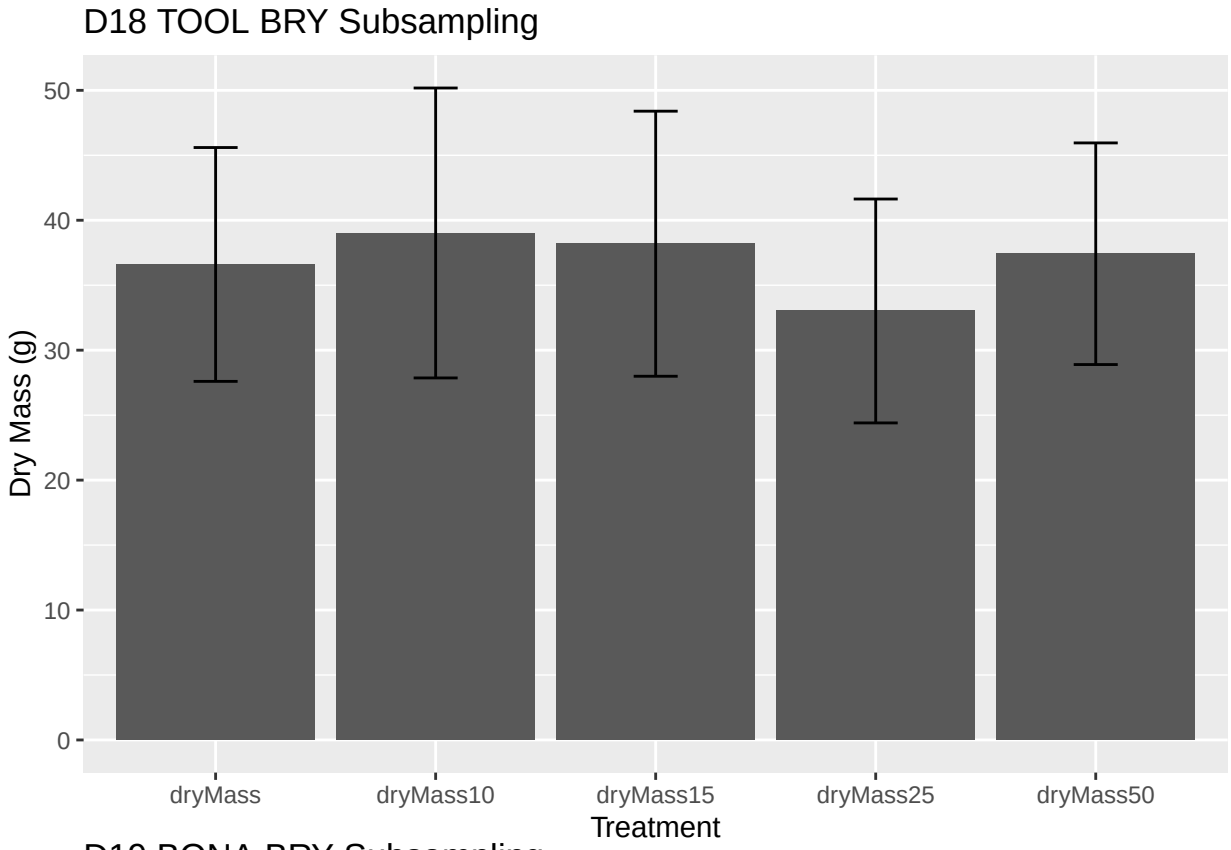
Note: The 15% subsample group has the greatest uncertainty due to the fact it was consistently sorted last and therefore lost the most water.

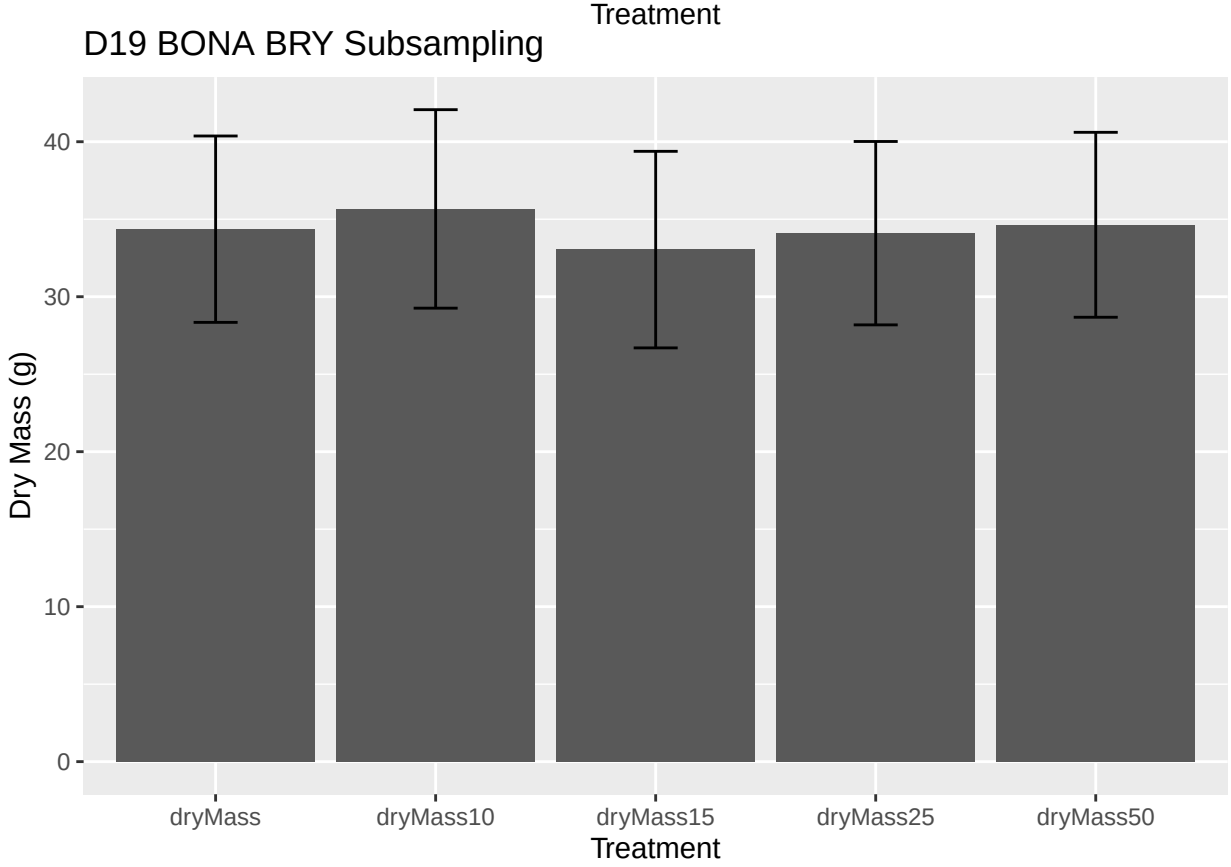
```
# Read in data collected by D18/D19
if(file.exists("/Users/cmeier")){
  inputPath <- "~/Documents/gitRepositories/neon-plant-sampling/hbpSampling/subsampleSortExperiments/"
}

if(file.exists("/Users/Pajaro")){
  inputPath <- "~/Documents/workDocuments/gitRepositories/neon-plant-sampling/hbpSampling/subsampleSortE
}

df <- openxlsx::read.xlsx(
  xlsxFile = paste0(inputPath, "hbp_d1819_sortExperiment2021_data.xlsx"),
  sheet = "R_input",
  startRow = 1,
  colNames = TRUE
)


#   Calculate total observed dryMass for each Clip Strip, and estimated
#   dryMass from subsampling
df <- df %>%
  dplyr::mutate(
    dryMass10 = round(freshMass*(ssDM10/ssFM10), digits = 2),
    .before = dryMass
  ) %>%
  dplyr::mutate(
    dryMass15 = round(freshMass*(ssDM15/ssFM15), digits = 2),
    .before = dryMass
  ) %>%
  dplyr::mutate(
    dryMass25 = round(freshMass*(ssDM25/ssFM25), digits = 2),
    .before = dryMass
  ) %>%
  dplyr::mutate(
    dryMass50 = round(freshMass*(ssDM50/ssFM50), digits = 2),
    .before = dryMass
  ) %>%
  dplyr::arrange(domainID, siteID, clipID)


##  Re-arrange data for lmer models and ggplot2
#   Create long data frame for lmer models
longDF <- df %>%
  dplyr::select(
    domainID,
    siteID,
    clipID,
    collectDate,
```

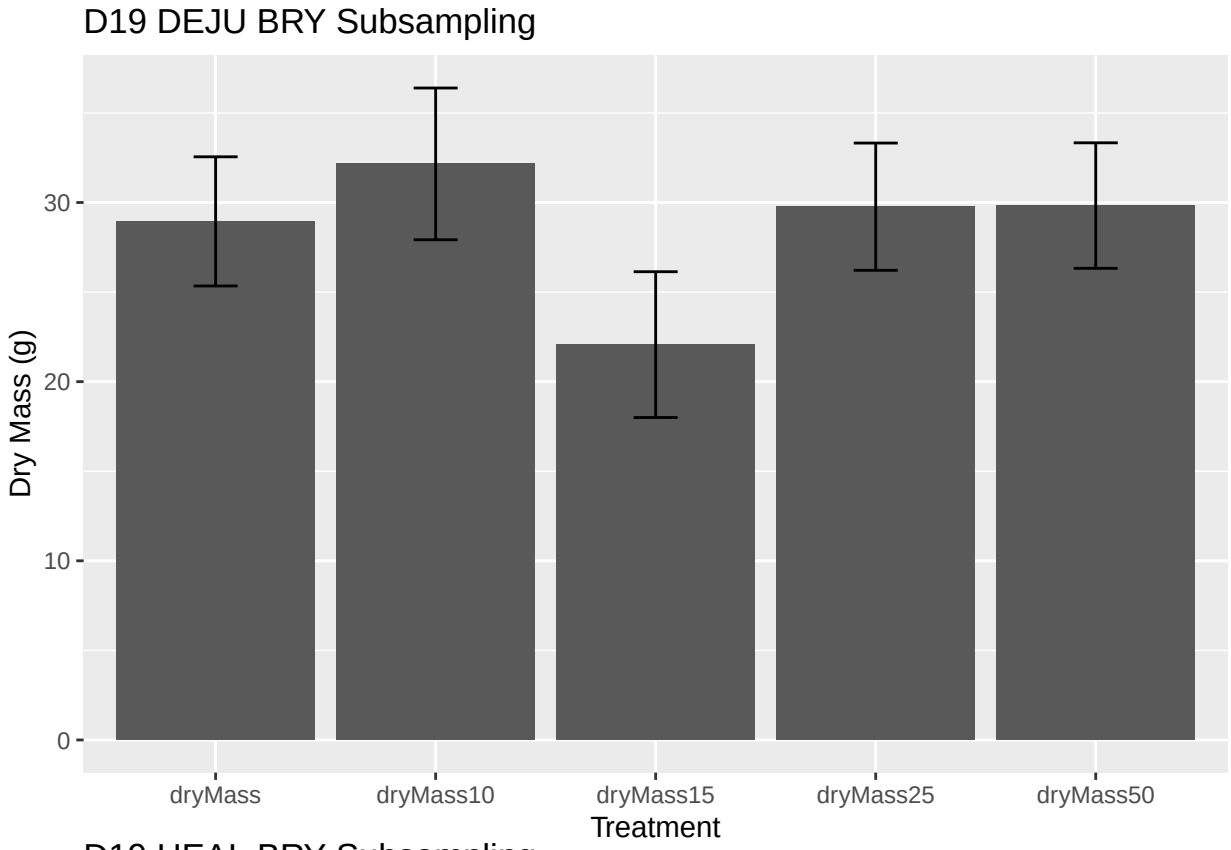
```
    dryMass10,
    dryMass15,
    dryMass25,
    dryMass50,
    dryMass
  ) %>%
  tidyr::pivot_longer(
    cols = tidyr::starts_with("dry"),
    names_to = "treatment",
    values_to = "estimatedDryMass"
  ) %>%
  dplyr::arrange(domainID, siteID, clipID)

#   Create summary data frame for ggplot2
summaryDF <- longDF %>%
  dplyr::group_by(domainID, siteID, treatment) %>%
  dplyr::summarise(
    count = n(),
    mean = mean(estimatedDryMass),
    sd = sd(estimatedDryMass),
    se = sd/sqrt(count)
  )
```

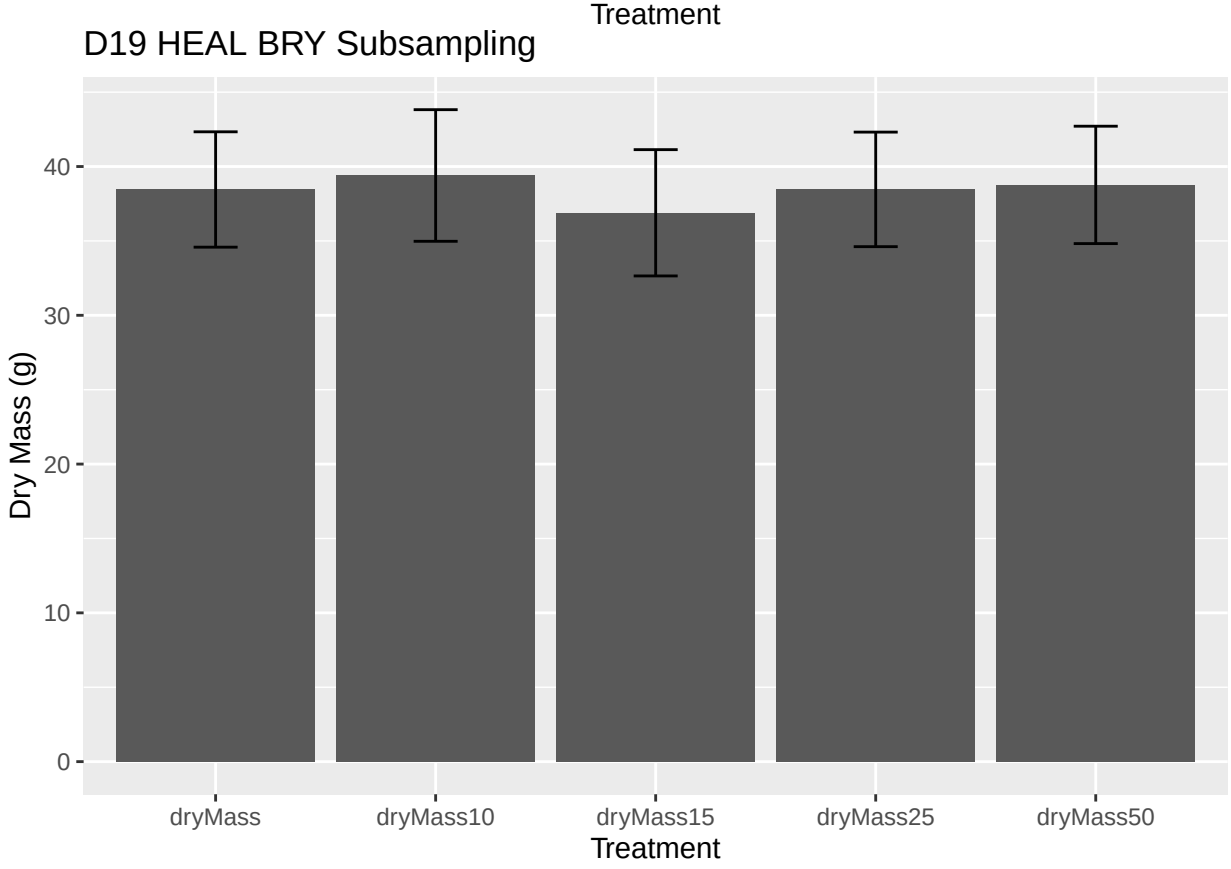
For each site, we graph the true observed BRY dryMass alongside the estimated dryMass calculated via subsampling.




D18 TOOL BRY Subsampling
Dry Mass (g)
50
40
30
20
10
0
dryMass
dryMass10
dryMass15
dryMass25
dryMass50
Treatment


D19 BONA BRY Subsampling
Dry Mass (g)
40
30
20
10
0
dryMass
dryMass10
dryMass15
dryMass25
dryMass50
Treatment


D19 DEJU BRY Subsampling
Dry Mass (g)
0
10
20
30
dryMass
dryMass10
dryMass15
dryMass25
dryMass50
Treatment


D19 HEAL BRY Subsampling
Dry Mass (g)
0
10
20
30
40
dryMass
dryMass10
dryMass15
dryMass25
dryMass50
Treatment

Results: Mixed-Effects model analysis

- Models are run separately for each site, as the aim is to understand site-specific effects of `treatment` on `estimatedDryMass`.
- Models use **clipID** as a random effect; benefit of random effect is it can account for variation at level of clip-strip unrelated to treatment - e.g., variation due to moisture on day clip-strip was harvested, effect of moss abundance or species composition at particular location, etc.
- The version of the `lmer()` function in the `lmerTest` package is used because the assumptions for this model allow standard anova output to be assessed.

```
### BARR subsampling model
#   Create model with random intercept by clipID
barrMod <- lmerTest::lmer(
  formula = estimatedDryMass ~ treatment + (1|clipID),
  data = longDF %>% dplyr::filter(siteID == "BARR"),
  REML = FALSE
)

#   Assess model outputs
summary(barrMod)
```

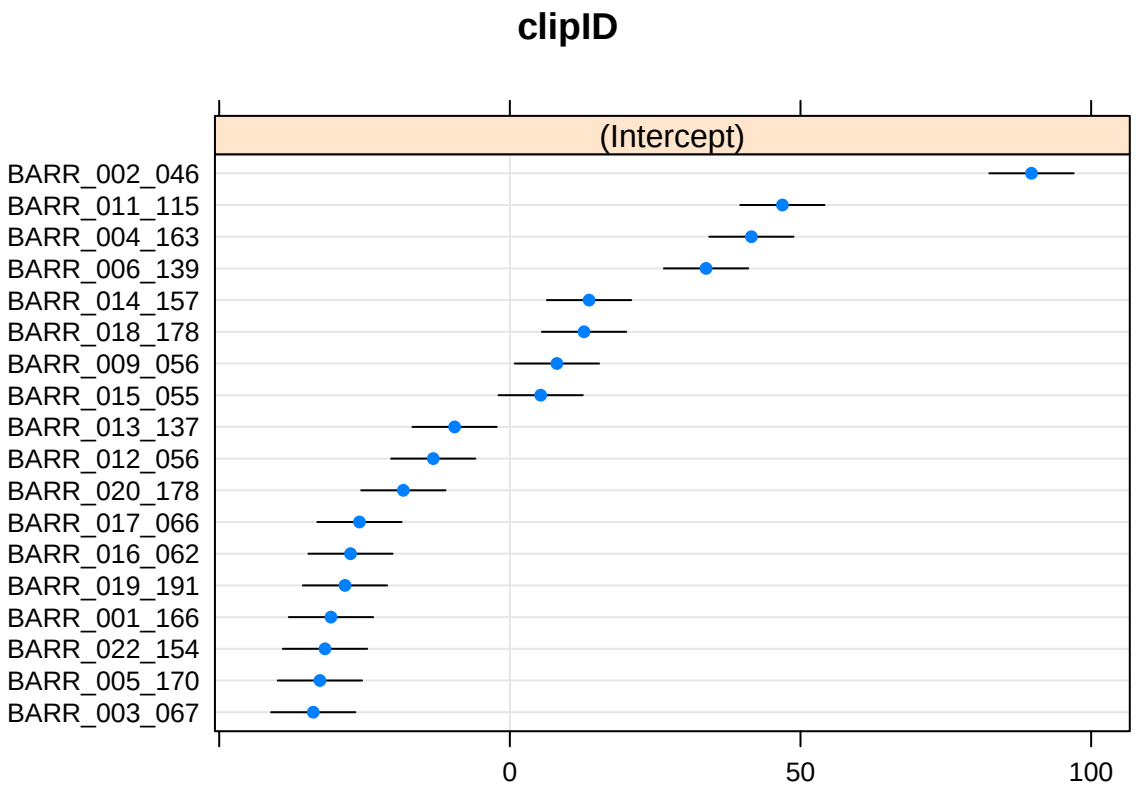
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estimatedDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(siteID == "BARR")
##
##      AIC      BIC   logLik deviance df.resid
##    730.6    748.1   -358.3    716.6       83
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9352 -0.3454 -0.0274  0.3990  3.7469
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 1155.50  33.993
##  Residual               69.19   8.318
## Number of obs: 90, groups:  clipID, 18
##
## Fixed effects:
##                    Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)         35.7317     8.2485 19.7333   4.332 0.000333 ***
## treatmentdryMass10  -4.9578     2.7727 72.0000  -1.788 0.077968 .
## treatmentdryMass15  -5.2878     2.7727 72.0000  -1.907 0.060495 .
## treatmentdryMass25   4.1300     2.7727 72.0000   1.490 0.140711
## treatmentdryMass50   0.5222     2.7727 72.0000   0.188 0.851134
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.168
## trtmntdrM15 -0.168  0.500
## trtmntdrM25 -0.168  0.500  0.500
## trtmntdrM50 -0.168  0.500  0.500  0.500
```

```
anova(barrMod)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value   Pr(>F)
## treatment   1145  286.26     4    72  4.1373 0.004502 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lattice::dotplot(ranef(barrMod, condVar = TRUE))
```

```
## $clipID
```



```r
### TOOL subsampling model
#   Create TOOL linear mixed-effects model to assess subsampling treatment
toolMod <- lmerTest::lmer(
  formula = estimatedDryMass ~ treatment + (1|clipID),
  data = longDF %>% dplyr::filter(siteID == "TOOL"),
  REML = FALSE
)

#   Assess model outputs
summary(toolMod)
```

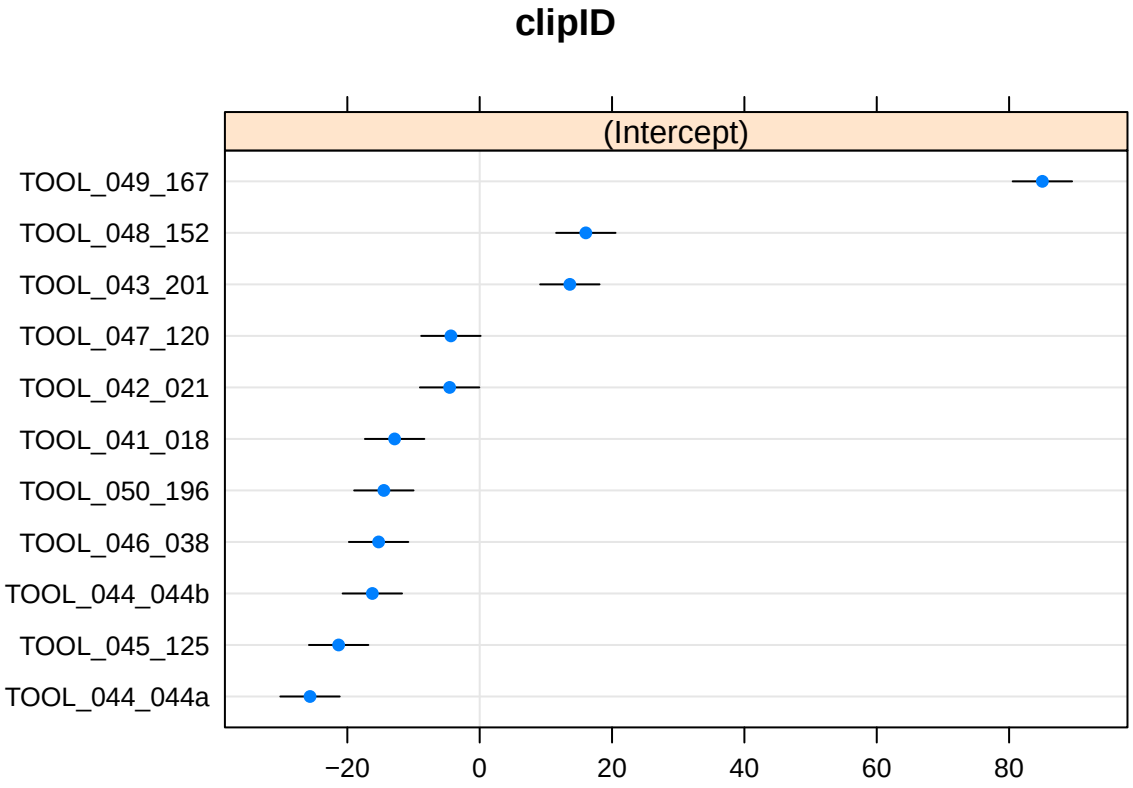
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estimatedDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(siteID == "TOOL")
##
##      AIC      BIC   logLik deviance df.resid
##    406.5    420.5   -196.2    392.5       48
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.2630 -0.4229  0.0324  0.3333  3.3273
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 886.59   29.776
##  Residual              26.36    5.134
## Number of obs: 55, groups:  clipID, 11
##
## Fixed effects:
##                    Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)         36.5982     9.1102 11.5248   4.017  0.00185 **
## treatmentdryMass10   2.4227     2.1893 44.0000   1.107  0.27447
## treatmentdryMass15   1.5964     2.1893 44.0000   0.729  0.46977
## treatmentdryMass25  -3.5800     2.1893 44.0000  -1.635  0.10914
## treatmentdryMass50   0.8255     2.1893 44.0000   0.377  0.70796
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.120
## trtmntdrM15 -0.120  0.500
## trtmntdrM25 -0.120  0.500  0.500
## trtmntdrM50 -0.120  0.500  0.500  0.500
```

```r
anova(toolMod)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value  Pr(>F)
## treatment 237.56  59.389     4    44  2.2528 0.07863 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lattice::dotplot(ranef(toolMod, condVar = TRUE))
```

```
## $clipID
```



```r
### BONA subsampling model
#   Create BONA linear mixed-effects model to assess subsampling treatment
bonaMod <- lmerTest::lmer(
  formula = estimatedDryMass ~ treatment + (1|clipID),
  data = longDF %>% dplyr::filter(siteID == "BONA"),
  REML = FALSE
)

#   Assess model outputs
summary(bonaMod)
```

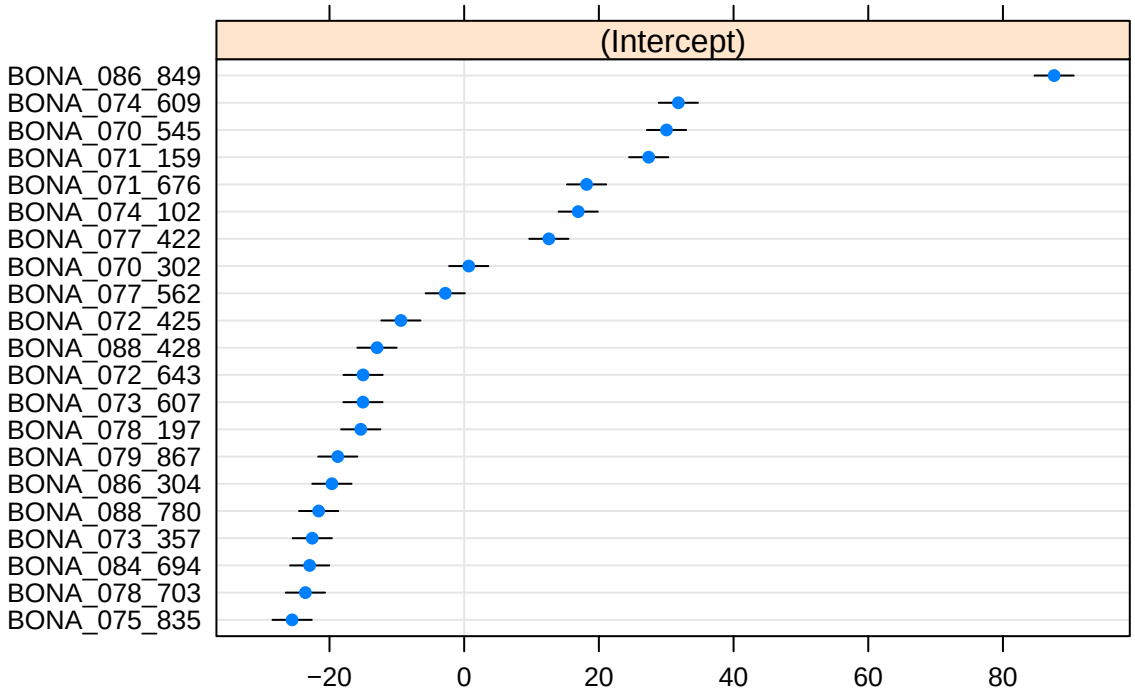
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estimatedDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(siteID == "BONA")
##
##      AIC      BIC   logLik deviance df.resid
##    686.6    705.1   -336.3    672.6       98
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3328 -0.3997 -0.0437  0.2640  2.7606
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 740.93   27.220
##  Residual              11.07    3.327
## Number of obs: 105, groups:  clipID, 21
##
## Fixed effects:
##                    Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)         34.3576     5.9841 21.5028   5.741 9.74e-06 ***
## treatmentdryMass10   1.3076     1.0268 84.0000   1.273    0.206
## treatmentdryMass15  -1.3186     1.0268 84.0000  -1.284    0.203
## treatmentdryMass25  -0.2581     1.0268 84.0000  -0.251    0.802
## treatmentdryMass50   0.2843     1.0268 84.0000   0.277    0.783
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.086
## trtmntdrM15 -0.086  0.500
## trtmntdrM25 -0.086  0.500  0.500
## trtmntdrM50 -0.086  0.500  0.500  0.500
```

```r
anova(bonaMod)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## treatment 75.514  18.878     4    84  1.7052 0.1565
```

```r
lattice::dotplot(ranef(bonaMod, condVar = TRUE))
```

```
## $clipID
```


clipID
(Intercept)
BONA_086_849
BONA_074_609
BONA_070_545
BONA_071_159
BONA_071_676
BONA_074_102
BONA_077_422
BONA_070_302
BONA_077_562
BONA_072_425
BONA_088_428
BONA_072_643
BONA_073_607
BONA_078_197
BONA_079_867
BONA_086_304
BONA_088_780
BONA_073_357
BONA_084_694
BONA_078_703
BONA_075_835
−20
0
20
40
60
80

```r
### DEJU subsampling model
#   Create DEJU linear mixed-effects model to assess subsampling treatment
dejuMod <- lmerTest::lmer(
  formula = estimatedDryMass ~ treatment + (1|clipID),
  data = longDF %>% dplyr::filter(siteID == "DEJU"),
  REML = FALSE
)

#   Assess model outputs
summary(dejuMod)
```

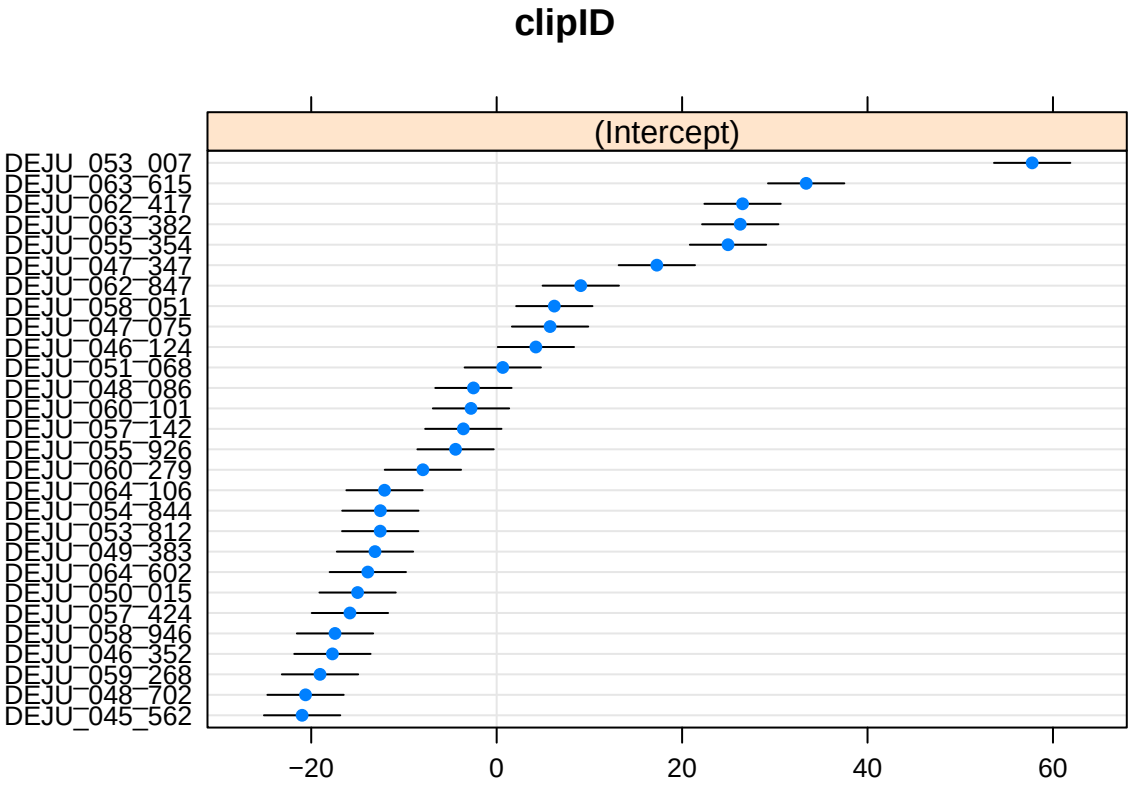
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estimatedDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(siteID == "DEJU")
##
##      AIC      BIC   logLik deviance df.resid
##    970.0    990.6   -478.0    956.0      133
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6154 -0.4695 -0.0427  0.3524  3.1870
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 368.52   19.197
##  Residual              22.31    4.724
## Number of obs: 140, groups:  clipID, 28
##
## Fixed effects:
##                     Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)          28.9461     3.7361  30.7269   7.748 1.03e-08 ***
## treatmentdryMass10    3.2100     1.2625 112.0000   2.543   0.0124 *
## treatmentdryMass15   -6.8789     1.2625 112.0000  -5.449 3.05e-07 ***
## treatmentdryMass25    0.8214     1.2625 112.0000   0.651   0.5166
## treatmentdryMass50    0.8825     1.2625 112.0000   0.699   0.4860
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.169
## trtmntdrM15 -0.169  0.500
## trtmntdrM25 -0.169  0.500  0.500
## trtmntdrM50 -0.169  0.500  0.500  0.500
```

```r
anova(dejuMod)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
## treatment 1632.5  408.14     4   112   18.29 1.373e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lattice::dotplot(ranef(dejuMod, condVar = TRUE))
```

```
## $clipID
```



```r
### HEAL subsampling model
#   Create HEAL linear mixed-effects model to assess subsampling treatment
healMod <- lmerTest::lmer(
  formula = estimatedDryMass ~ treatment + (1|clipID),
  data = longDF %>% dplyr::filter(siteID == "HEAL"),
  REML = FALSE
)

#   Assess model outputs
summary(healMod)
```

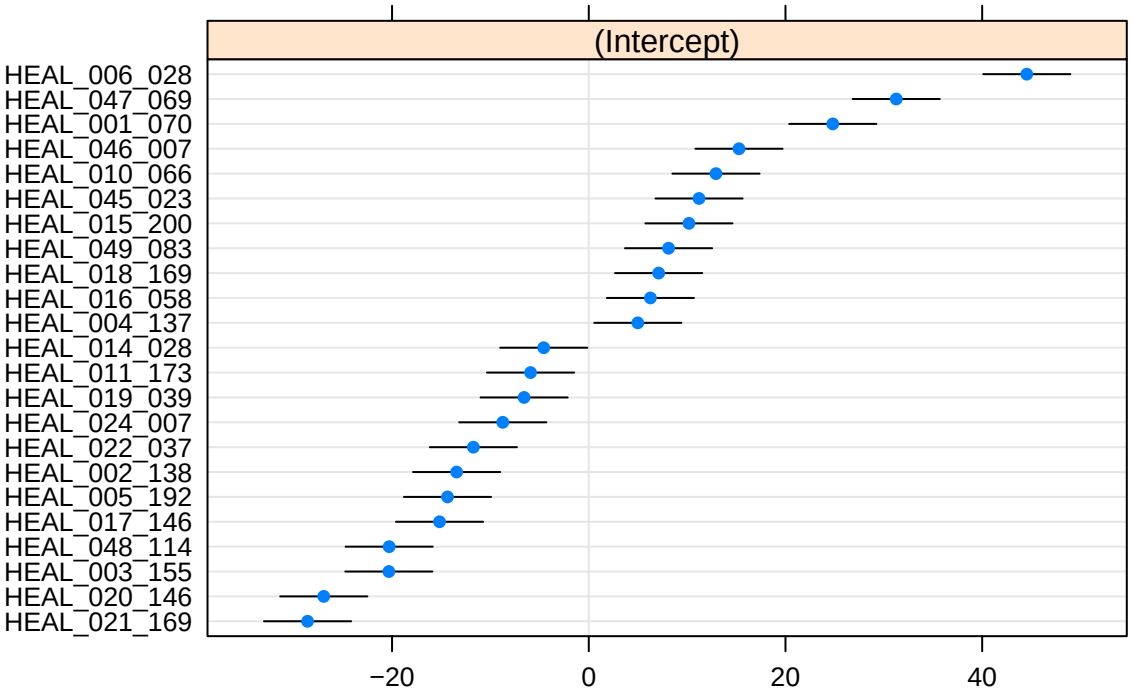
```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estimatedDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(siteID == "HEAL")
##
##      AIC      BIC   logLik deviance df.resid
##    811.3    830.5   -398.6    797.3      108
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4102 -0.3412 -0.0339  0.3460  4.5668
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 339.26   18.419
##  Residual              25.95    5.094
## Number of obs: 115, groups:  clipID, 23
##
## Fixed effects:
##                     Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)        38.458261   3.984799 25.831956   9.651 4.71e-10 ***
## treatmentdryMass10  0.940870   1.502034 92.000000   0.626    0.533
## treatmentdryMass15 -1.567826   1.502034 92.000000  -1.044    0.299
## treatmentdryMass25  0.006087   1.502034 92.000000   0.004    0.997
## treatmentdryMass50  0.307826   1.502034 92.000000   0.205    0.838
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.188
## trtmntdrM15 -0.188  0.500
## trtmntdrM25 -0.188  0.500  0.500
## trtmntdrM50 -0.188  0.500  0.500  0.500
```

```r
anova(healMod)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## treatment 78.626  19.656     4    92  0.7576 0.5555
```

```r
lattice::dotplot(ranef(healMod, condVar = TRUE))
```

```
## $clipID
```


clipID
(Intercept)
HEAL_006_028
HEAL_047_069
HEAL_001_070
HEAL_046_007
HEAL_010_066
HEAL_045_023
HEAL_015_200
HEAL_049_083
HEAL_018_169
HEAL_016_058
HEAL_004_137
HEAL_014_028
HEAL_011_173
HEAL_019_039
HEAL_024_007
HEAL_022_037
HEAL_002_138
HEAL_005_192
HEAL_017_146
HEAL_048_114
HEAL_003_155
HEAL_020_146
HEAL_021_169
−20
0
20
40

Summary

1. BARR results
    - 10% subsampling 14% lower than control, marginally significant ($p = 0.078$)
    - 15% subsampling 15% lower than control, marginally significant ($p = 0.060$)
    - 25% subsampling 12% higher than control ($p = 0.14$)
    - 50% subsampling virtually identical to control ($p = 0.85$)
2. TOOL results
    - 10% subsampling 6.6% higher than control ($p = 0.27$)
    - 15% subsampling 4.4% higher than control ($p = 0.47$)
    - 25% subsampling about 10% lower than control, marginally significant ($p = 0.11$)
    - 50% subsampling less than 3% higher than control ($p = 0.71$)
3. BONA results
    - 10% subsampling 3.8% higher than control ($p = 0.21$)
    - 15% subsampling 3.8% lower than control ($p = 0.20$)
    - 25% subsampling $< 1\%$ lower than control ($p = 0.80$)
    - 50% subsampling $< 1\%$ higher than control ($p = 0.78$)
4. DEJU results
    - 10% subsampling 11% higher than control ($p < 0.05$)
    - 15% subsampling 24% lower than control ($p < 0.0001$)
    - 25% subsampling 2.8% higher than control ($p = 0.52$)
    - 50% subsampling 3% higher than control ($p = 0.49$)
5. HEAL results
    - 10% subsampling 2.4% higher than control ($p = 0.53$)
    - 15% subsampling 4.2% lower than control ($p = 0.30$)
    - 25% subsampling virtually identical to control ($p = 0.997$)
    - 50% subsampling virtually identical to control ($p = 0.84$)

Summary of recommendations for BRY subsampling by site:

- BARR: 50%
- TOOL: 50% (conservative due to variable results at lower subsampling levels)
- BONA: 10%
- DEJU: 25%
- HEAL: 10%

Summary of recommendations for BRY subsampling by domain:

- D18 (BARR, TOOL): 50%
- D19 (BONA, DEJU, HEAL): 25%

D18/19 leads indicate it will be easier to train and execute BRY subsampling correctly if a single subsampling target is used per domain.