

# Phylogenetic placement visualization with treeio-ggtree method

## 0.1 Parse jplace file and visualize the placement results

```
# Required packages
library(ggplot2)
library(treeio)
library(ggtree)
library(colorspace)
library(aplot)

# Processing data
tree1 <- read.jplace(file = "../exampledata/Drosophila/Drosophila.jplace")

#Mapping nplace to color and line size of branches
#Set color range from 0.5 to 5 by limits =c(0.5, 5)
#Set line size range by scale_size_continuous()
p1 <- ggtree(tree1,
             layout = "rect",
             aes(
               color=nplace,
               size=nplace
             )
             ) +
  # scale_color_viridis_c(option="plasma",limits =c(0.5, 5)) +
  geom_tiplab(size=4) +
  scale_color_continuous_sequential(palette = "Sunset",limits =c(0.5, 1)) +
  scale_size_continuous(range=c(0.1, 1.5)) +
  xlim(0,0.2)

#Using gradient color and line size by continuous parameter(including color,size,all)
p2 <- ggtree(tree1,
             layout = "rect",
             aes(color=nplace,size=nplace),
             continuous = "all",
             nsplit = 400
             ) +
  # scale_color_viridis_c(option="plasma",limits =c(0.5, 5)) +
  geom_tiplab(size=4) +
  scale_color_continuous_sequential(palette = "Sunset",limits =c(0.5, 1)) +
  scale_size_continuous(range=c(0.1, 1.5)) +
  xlim(0,0.2)

#Annotated the placements on the branches
#Here we showed nplace>20
#Using ggrepel package
p3 <- p1 + #ggrepel::geom_label_repel(
  geom_label(
    data=td_filter(nplace>0),
    aes(x=branch,label=nplace),
    label.size=0.25,size=4
  )
)

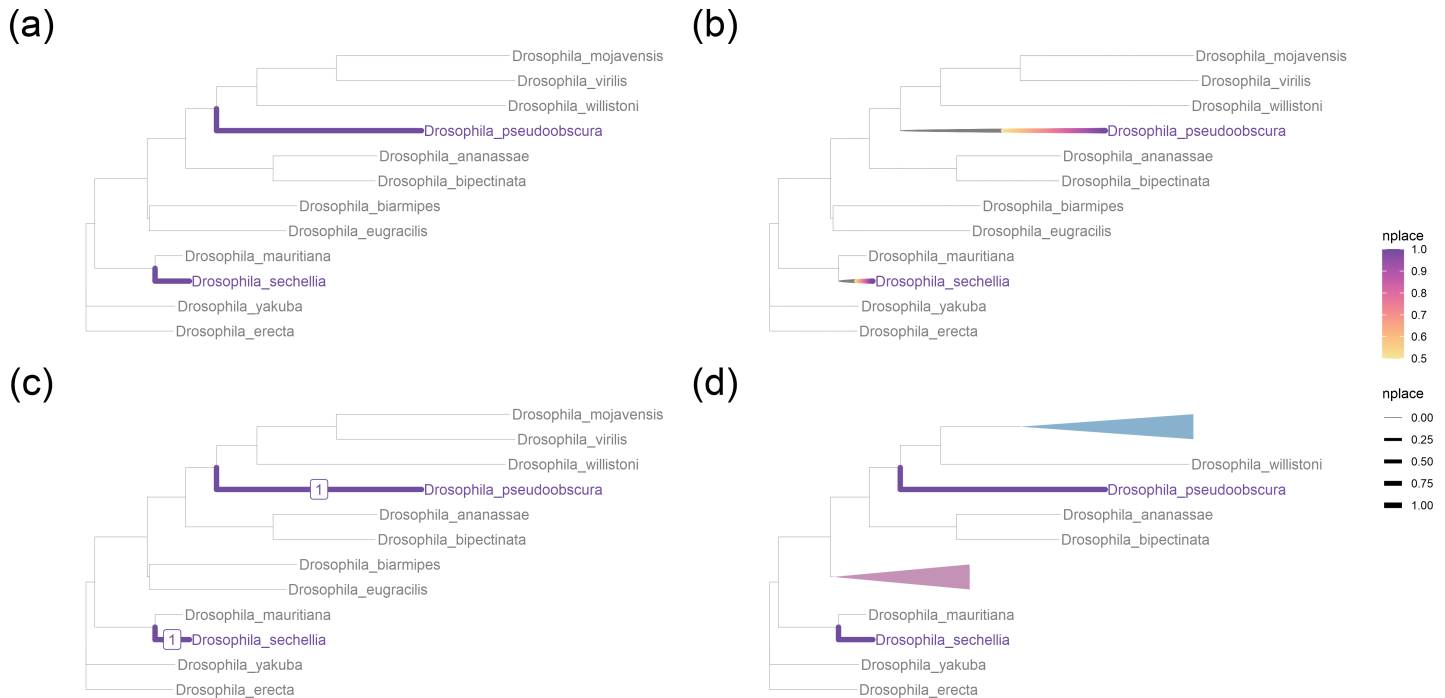
#Collapse clades without placements
#To visualize the node number
# ggtree(tree1,layout = "rect") + geom_text2(aes(x=branch,label=node))
```

```

p4 <- p1 %>% collapse(
  node=20,
  'min',
  fill='#11659a',
  alpha=.5,
  clade_name="collapsed clade 1"
) %>% collapse(
  node=16,
  'min',
  fill='#8b2671',
  alpha=.5,
  clade_name="collapsed clade 2"
)

plot_list(p1, p2,p3,p4,
  tag_levels = list(c('(a)', '(b)', '(c)', '(d)'), '1'),
  ncol =2,
  guides = 'collect', tag_size = 30)

```



## 0.2 Visualizing other placement features

```

# Required packages
library(treeio)
library(ggtree)
library(ggplot2)
library(tidytree)
library(colorspace)
library(dplyr)

# Load example datasets
tree1 <- read.jplace(file = "../exampledata/TaraOceans/Tara_Oceans_OTUs.jplace")

```

```

p1 <- ggtree(tree1,
  layout = "rect",
  branch.length = "none",
  aes(
    color=nplace,
    size=nplace
  )
) +
  geom_tiplab(size=2,offset=0.02) +
  scale_color_continuous_sequential(palette = "Sunset",limits =c(1, 7)) +
  scale_size_continuous(range=c(0.2, 1.6)) +
  xlim(0,22)

# Extract placement informations
pla <- get.placements(tree1)

# Get the max like_weight_ratio of each node
pla2 <- group_by(pla, .data$node) %>%
  filter(.data$like_weight_ratio == max(.data$like_weight_ratio))

# Link additional placement data onto reference tree
p2 <- p1 %<+% pla +
  geom_label(
    aes(x=branch,label=like_weight_ratio),
    size=2
  )

# Annotate Group Labels
p3 <- p2 + geom_strip(1,6,label = "Ohter Eukaryota", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#B0171F",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(7,39,label = "Holomycota", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#EE6A50",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(40,48,label = "Metazoa", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#FFC125",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(49,55,label = "Filasterea", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#4DAF4A",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(56,65,label = "Dermocystida", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#9ACD32",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(66,87,label = "Ichthyophonida", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#87CEFA",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(88,90,label = "Corallochytreia", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#7B68EE",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(91,104,label = "Craspedida", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#D15FEE",offset.text = 0.2,
  barsize = 5,extend = 1) +
  geom_strip(105,129,label = "Acanthoecida", align = TRUE, alpha=.8,
  fontsize=5,offset = 4, color = "#984EA3",offset.text = 0.2,
  barsize = 5,extend = 1)

# Collapse nonimportant clades
p4 <- p3 %>% collapse(node=131,

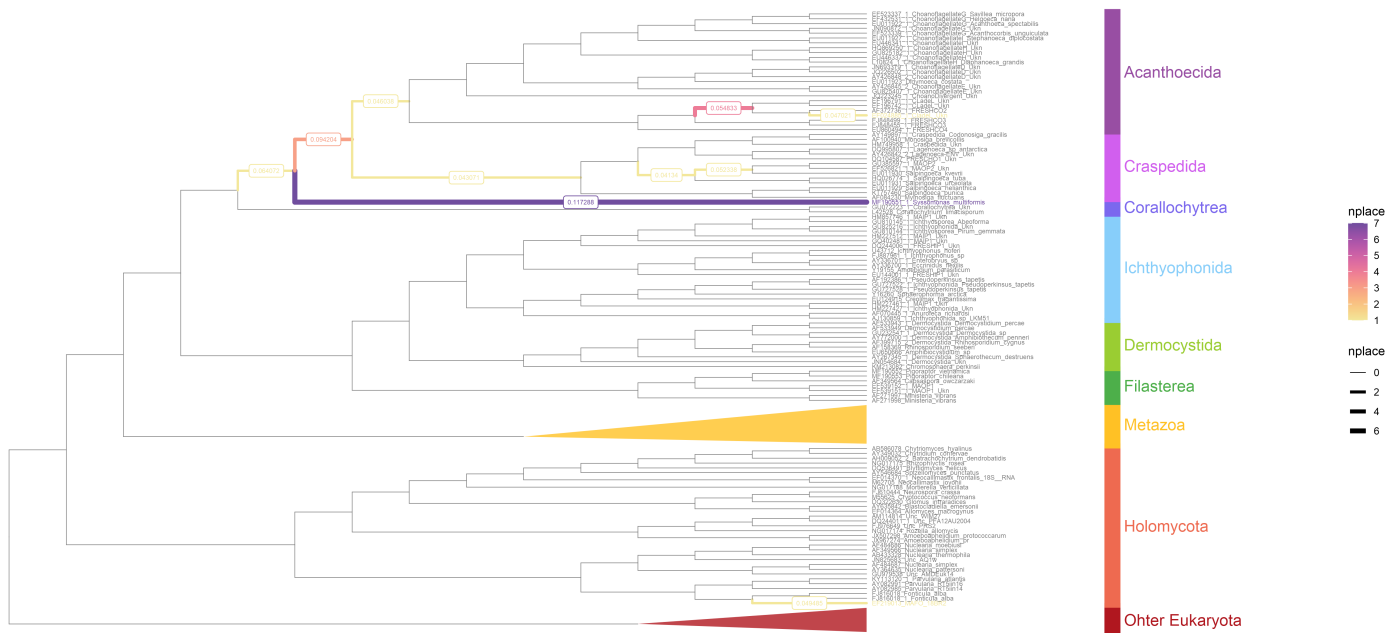
```

```

    'min',
    fill='#B0171F',
    alpha=.8,
    clade_name="Ohter Eukaryota") %>%
collapse(node=170,
  'min',
  fill='#FFC125',
  alpha=.8,
  clade_name="Metazoa")

```

p4



### 0.3 Extract subtree of interest from large reference tree

```

# Required packages
library(treeio)
library(ggtree)
library(ggtreeExtra)
library(ggplot2)
library(tidytree)
library(colorspace)
library(patchwork)
library(aplot)

tree1 <- read.jplace(file = "../exampledata/subtree/pplacer_Amt_subtree.jplace")

p1 <- ggtree(tree1,
  layout = "circular",
  aes(
    color=nplace,
    size=nplace
  )
) +
  scale_color_continuous_sequential(palette = "Sunset", limits = c(0.5, 1)) +
  scale_size_continuous(range=c(0.1, 1.5)) +
  geom_highlight(node = 659, fill="#8b2671", alpha=0.3)

```

```
Amt_tiplabel <- read.csv(file = "../exampledata/subtree/Amt_tiplabel.csv",header = TRUE)
head(Amt_tiplabel)
```

```
##   node      tiplabel      Group
## 1     1 Asterochloris Trebouxiphyceae
## 2     2 Asterochloris Trebouxiphyceae
## 3     3      Coccomyxa Trebouxiphyceae
## 4     4 Stichococcus Trebouxiphyceae
## 5     5      Galdieria   Cyanidiales
## 6     6      Galdieria   Cyanidiales
```

```
taxon_groups <- c("Trebouxiphyceae","Cyanidiales","Non-Cyanidiales",
                  "Chlamydomonadales","Tetraselmis ssp.,"Pyramimonas ssp.",
                  "Nephroselmis pyriformis CCMP717","Bathycoccus ssp.",
                  "Ostreococcus","Mantoniella ssp.,"Micromonas ssp.",
                  "Embryophyta")
```

```
Amt_tiplabel$Group <- factor(Amt_tiplabel$Group, levels=taxon_groups)
cols <- c("#FFC125","#87CEFA","#7B68EE","#191970","#800080",
          "#9ACD32","#D15FEE","#FFC0CB","#EE6A50","#8DEEEE",
          "#006400","#800000","#B0171F")
```

```
p2 <- p1%<+% Amt_tiplabel +
  geom_fruit(
    geom = geom_tile,
    mapping = aes(fill = Group),
    width = 0.2,
    offset = 0.01
  ) +
  scale_fill_manual(
    name = "",
    values = cols,
    guide = guide_legend(
      keywidth = 1,
      keyheight = 1,
      order = 3
    ),
    na.translate = FALSE
  )
```

```
# Extracting subtree from the whole tree
```

```
tree3 <- tree_subset(tree1, node=659, levels_back=0)
```

```
tree4 <- as_tibble(tree3)
```

```
tree4$label2 <- c("Ost_RCC809","Ost_mediterraneus_RCC2593","Ost_mediterraneus_RCC1107","Ost_mediterraneus_RCC1600")
```

```
tree3@phylo$tip.label <- tree4[match(tree3@phylo$tip.label,tree4$label),]$label2
```

```
p3 <- ggtree(tree3,
  layout = "rect",
  aes(
    color=nplace,
    size=nplace
  )
) +
  scale_color_continuous_sequential(palette = "Sunset",limits =c(0.5, 1)) +
  scale_size_continuous(range=c(0.1, 1.5)) +
  geom_tiplab(size=3,offset=0.02) +
  geom_tippoint(size=3) +
  geom_cladelab(node=5, label="saltern1",geom='label',offset=0.2,fontsize =3) +
  geom_cladelab(node=6, label="saltern2",geom='label',offset=0.15,fontsize =3) +
```

```

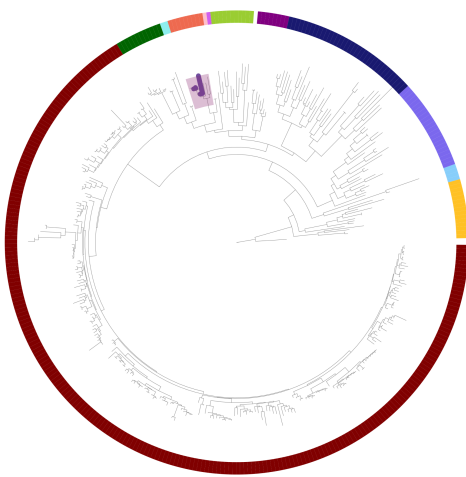
ggtitle("Ostreococcus subtree") +
# theme_tree2() +
xlim(0,0.6)

layout <- c(
  area(t = 0, l = 0, b = 4, r = 10),
  area(t = 1, l = 10, b = 3, r = 16)
)

plot_list(p2, p3,
  tag_levels = list(c('(a)', '(b)'), '1'),
  ncol = 2,
  guides = 'collect', tag_size = 20, design = layout)

```

(a)



(b) *Ostreococcus* subtree

