## ggtree\_example

## Susan Gogolski

## 2025-04-02

```
library(ggplot2)
library(viridis)
## Loading required package: viridisLite
if (!requireNamespace("BiocManager", quietly = TRUE)){
  install.packages("BiocManager")
}
if (!requireNamespace("ggtree", quietly = TRUE)){
 BiocManager::install("ggtree")
}
library(ggtree)
## ggtree v3.14.0 Learn more at https://yulab-smu.top/contribution-tree-data/
##
## Please cite:
##
## Guangchuang Yu, Tommy Tsan-Yuk Lam, Huachen Zhu, Yi Guan. Two methods
## for mapping and visualizing associated data on phylogeny using ggtree.
## Molecular Biology and Evolution. 2018, 35(12):3041-3043.
## doi:10.1093/molbev/msy194
tree = ggtree::read.tree(file = "tree.nwk")
tree
## Phylogenetic tree with 29 tips and 27 internal nodes.
##
## Tip labels:
    MutS2a_AT5G54090, S_moellendorffii_421325, C_paradoxa_Contig9550, Cyanobacterium_aponinum_WP_09943
##
## Unrooted; includes branch length(s).
ggplot(tree) + geom_tree() +
  geom_tiplab() +
 theme_tree()
```

```
MutS2b_AT1G6507
O_sativa_LOC_Os04g58410
S_moellendorffii_141299
                                                            S_moellendorffii_141299
P_patens_3c17_13784V3.1.p
P_patens_3c17_13780V3.1.p
—M_polymorpha_0028s0103.1.p
—M_endlicherianum_ME000114S10951
—C_reinhardtii_01.g014800.t1.1
—C_braunii_GBG73512
—O_lucimarinus_37894
P_umbilicalis_0346s0006.1.p
pupureum_contin_2084_1
                                                        P_pupureum_contig_2084.1
                                                                                     -Aquifex_aeolicus_AAC07247
                                                                                                         ·Borreliella_burg
                                     Bacillus_subtilis_subsp_CAB14818.1

——Synechocystis_sp.BAA17670.1

— Cyanobacterium_aponinum_WP_099434570.1

C_paradoxa_Contig9550

— C_reinhardtii_01.g040350.t1.1
                           O_lucimarinus_38770
O_lucimarinus_38770
C_braunii_GBG66194
C_braunii_GBG83953
                         M_endlicherianum_ME000591S08501
                  S_moellendorffii_421325
P_patens_3c5_20780V3.1.p
        M_polymorpha_0023s0121.1.p
O_sativa_LOC_Os10g36530.1
-MutS2a_AT5G54090
ggplot(tree) +
   geom_tree(color="firebrick", size=0.5, linetype="dotted", layout="roundrect") +
   geom_tiplab(size=3) +
   theme_tree()
                                                                                      · · · · MutS2b_AT1G6507
                                                                           .... O_sativa_LOC_Os04g58410.1
                                                                          ··· S_moellendorffii_141299
                                                                    P_patens_3c17_13784V3.1.p
                                                                   P_patens_3c17_13780V3.1.p
                                                              ... M_polymorpha_0028s0103.1.p
                                                          .....M_endlicherianum_ME000114S10951
......C_reinhardtii_01.g014800.t1.1
                                                          ····· C_braunii_GBG73512
                                                         .....O_lucimarinus_37894
                                                     ···· P_umbilicalis_0346s0006.1.p
                                                        P_pupureum_contig_2084.1
                                                          _pupureum_comig_zoo4.1
_____Aquifex_aeolicus_AAC07247.1
_____Helic
_____Borreliella_burgdo
                                                     ···· Bacillus_subtilis_subsp_CAB14818.1
                                              · · Synechocystis_sp.BAA17670.1
                                           ··· Cyanobacterium_aponinum_WP_099434570.1
                                     C paradoxa Contig9550
                                    ··· C reinhardtii 01.q040350.t1.1
                                ·····O lucimarinus 38770
                           ····· C_braunii_GBG66194
                           C_braunii_GBG83953
                         M_endlicherianum_ME000591S08501
                        ··S moellendorffii 421325
                P_patens_3c5_20780V3.1.p
            ······ M_polymorpha_0023s0121.1.p
   ···· O_sativa_LOC_Os10g36530.1
 ·····MutS2a_AT5G54090
```

```
tree_rooted = ape::root(tree, outgroup = "Helicobacter_pylori_AAD07685.1", resolve.root=TRUE)
ggplot(tree_rooted) +
    geom_tree() +
    geom_tiplab() +
    theme_tree()
                                                                                                                                                                                     Mut
                                                                                                                                                    O_sativa_
S_moellendorffii_14
                                                                                                         S_moellendorffiii_14
P_patens_3c17_13784V3.
P_patens_3c17_13780V3.
M_polymorpha_0028s0103.1.
M_endlicherianum_ME00011
C_reinhardtii_01.g0
C_braunii_GBG73512
O_lucimarinus_37894
P_umbilicalis_0346s0006.1.p
P_pupureum_contig_2084.1
P_patens_3c5_20780V3.1.p
M_polymorpha_0023s0121.1.p
—___O_sativa_LOC_Os10g
MutS2a_AT5G54090
—_S_moellendorffii_421325
—_M_endlicherianum_ME000591S0850
—_C_braunii_GBG66194
C_braunii_GBG83953
—_C_reinhardtii_01.g040350.t1.1
O_lucimarinus_38770
                                                                                                                                              O_lucimarinus_38770
                                                        C_paradoxa_Contig9550
Synechocystis_sp.BAA17670.1
Cyanobacterium_aponinum_WP_099434570.1
Bacillus_subtilis_subsp_CAB14818.1
                                                                                                                          ·Borreliella_burgdorferi_AAC664
                                                    Aquifex_aeolicus_AAC07247.1
                                                                                                                            -Helicobacter_pylori_AAD0768!
ggplot(tree_rooted) +
```

geom\_tree() +
geom\_tiplab() +
theme\_tree() +
hexpand(0.25)

```
MutS2b_AT1G6507

— S_moellendorffii_141299

— P_patens_3c17_13784V3.1.p

— P_patens_3c17_13784V3.1.p

— M_polymorpha_0028s0103.1.p

— M_endlicherianum_ME000114S10951

— C_reinhardtii_01.g014800.t1.1

— C_braunii_GBG73512

— O_lucimarinus_37894

— P_umbilicalis_0346s0006.1.p

— P_putens_3c5_20780V3.1.p

— M_polymorpha_0023s0121.1.p

— M_polymorpha_0023s0121.1.p

— M_polymorpha_0023s0121.1.p

— O_sativa_LOC_Os10g36530.1

— MutS2a_AT5G54090

— S_moellendorffii_421325

— M_endlicherianum_ME000591S08501

— C_braunii_GBG66194

— C_braunii_GBG66194

— C_braunii_GBG683953

— C_reinhardtii_01.g040350.t1.1

— O_lucimarinus_38770

— C_paradoxa_Contig9550

— Synechocystis_sp.BAA17670.1

— Cyanobacterium_aponinum_WP_099434570.1

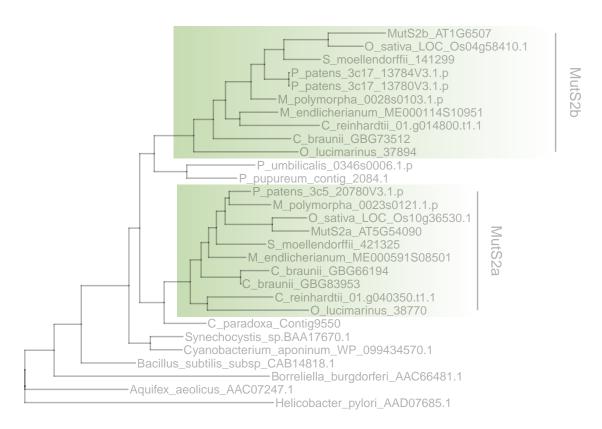
— Bacillus_subtilis_subsp_CAB14818.1

— Borreliella_burgdorferi_AAC66481.1

— Aquifex_aeolicus_AAC07247.1

— Helicobacter_pylori_AAD07685.1
```

```
ggplot(tree_rooted) +
  geom_tree() +
  geom_tiplab(size=3) +
  theme_tree() +
  hexpand(0.25) +
  geom_hilight(node=36, fill="chartreuse4", alpha=.3, type = "gradient", to.bottom=TRUE, extend=1) +
  geom_hilight(node=44, fill="chartreuse4", alpha=.3, type = "gradient", to.bottom=TRUE, extend=1) +
  geom_cladelab(node=36, label = "MutS2a", angle=270, hjust = "center", offset=1, offset.text=0.1) +
  geom_cladelab(node=44, label = "MutS2b", angle=270, hjust = "center", offset=1, offset.text=0.1)
```



```
gene_data = read.delim("gene_data.txt")
gene_data[1:5,]
```

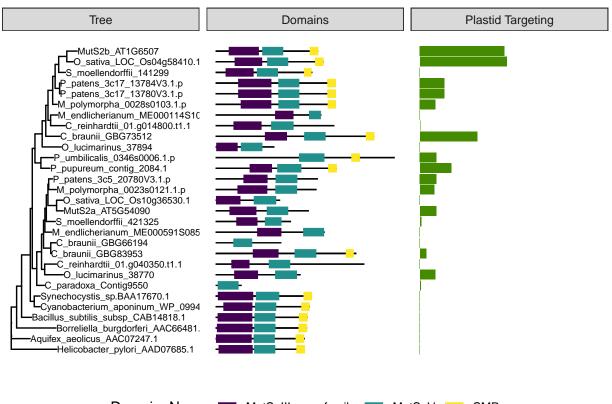
```
##
                  Newick_label Domain_Start Domain_End
                                                                  Domain_Name
## 1
              MutS2a_AT5G54090
                                          118
                                                     348 MutS_III superfamily
## 2
              MutS2a_AT5G54090
                                          436
                                                     619
                                                                        MutS_V
## 3 O_sativa_LOC_Os10g36530.1
                                            3
                                                     212 MutS_III superfamily
## 4 O_sativa_LOC_Os10g36530.1
                                          325
                                                     517
                                                                        MutS_V
## 5
            C_braunii_GBG66194
                                          151
                                                     312
                                                                        MutS_V
##
     Seq_Length Plastid_Targeting
## 1
            796
                              0.28
## 2
            796
                              0.28
## 3
            550
                              0.00
## 4
            550
                              0.00
            561
                              0.00
#store the tree
p1 = ggplot(tree_rooted, aes(x,y)) +
  geom tree() +
  geom_tiplab(size=2.5) +
  hexpand(1.5) +
  theme_tree()
#add a facet with lines corresponding to the length of each protein
p2 = facet_plot(p1, panel = "Domains", data = gene_data, geom = geom_segment, mapping = aes(x=0, xend=S
```

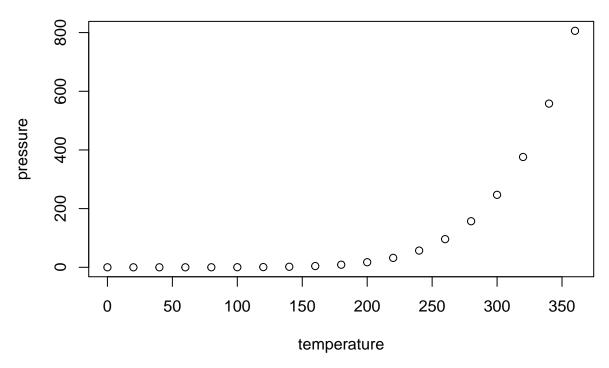
p3 = facet\_plot(p2, panel = "Domains", data = gene\_data, geom=geom\_segment, aes(x=Domain\_Start, xend=Domain\_start, xend=Domain\_

#add the domains on top of those lines

```
theme(legend.position = "bottom") +
scale_color_viridis(discrete = TRUE)

#add another facet with a bar graph showing the probability of plastid targeting
p4 = facet_plot(p3, panel = "Plastid Targeting", data = gene_data, geom=geom_col, aes(x=Plastid_Targeting)
theme(legend.position = "bottom")
p4
```





Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.