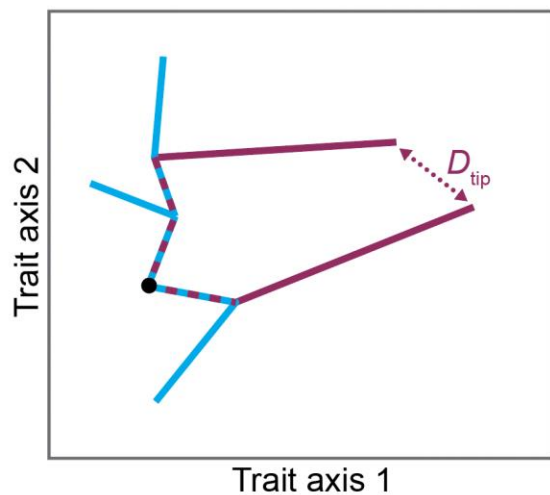


From dolphins and ichthyosaurs to succulent Cactaceae and Euphorbiaceae, phenotypic convergence provides some of the best known and most striking examples of evolution. Although convergence can result from a variety of processes (e.g., shared developmental constraints), its association with adaptation to shared selective environments is perhaps best known. Given the level of interest in phenotypic convergence, there is an understandably large body of research focusing on a wide range of study taxa. But how does one go about testing for phenotypic convergence? How do we go beyond qualitative similarities to quantitative tests? Put another way, how can we be more certain that the similarities we observe between distantly related lineages actually are convergence? Luckily, a wealth of methods for quantitatively testing for convergence have been proposed in recent years. However, recent work has also shown that some of these methods may have potentially serious biases (see Grossnickle et al., 2024 to learn more). The *convevol* package for *R* includes functions for applying two related statistical tests of phenotypic convergence. The popular *C*-measures (Stayton, 2015) and the newer *Ct*-measures, which incorporate several changes to minimize the aforementioned biases. In this tutorial we'll focus on *Ct*-measures.

First, we need some data to work with. Let's suppose we have some hypothetical study group and we're interested in whether a shared ecological trait is associated with phenotypic convergence in some taxa. Functions in *convevol* give us the ability to test such hypotheses, by evaluating whether putatively convergent lineages are closer to each other in morphospace than were their respective ancestors (e.g, see the focal lineages in the figure below). As already mentioned, the *convevol* package includes functions to measure convergence via the *C*-measures of Stayton (2015) and the recently updated *Ct*-measures of Grossnickle et al. (2024). Both the *C*-measures and *Ct*-measures compare the phenotypic distances between tips (D_{tip}) to the maximum phenotypic distance between lineages at any point in their evolutionary history (D_{max} or $D_{max,t}$). While the original *C*-measures allow D_{max} to be measured between any two internal or external nodes, the *Ct*-measures constrain $D_{max,t}$ to be measured between two lineages at a specific point in time. While this means *Ct*-measures can only be computed on time calibrated trees, it also solves some of the undesirable behavior of the original metrics when dealing with taxa which are outlying in morphospace. For more information on the theory behind these methods and their limitations, see Grossnickle et al. (2024). For now, we will walk through a hypothetical example and demonstrate some of the functions in the *convevol* package.



Focal lineages (D_{tip} = distance between tips)
 'Ancestral' lineages
 ● MRCA

To use the *Ct*-measures we first need a time calibrated phylogeny including the taxa of interest. The requirement for trees to be time calibrated is one of the major differences between the *Ct*-measures (Grossnickle et al. 2024) and the *C*-measures (Stayton 2015). The tree of a fictitious group is included in the package for testing purposes.

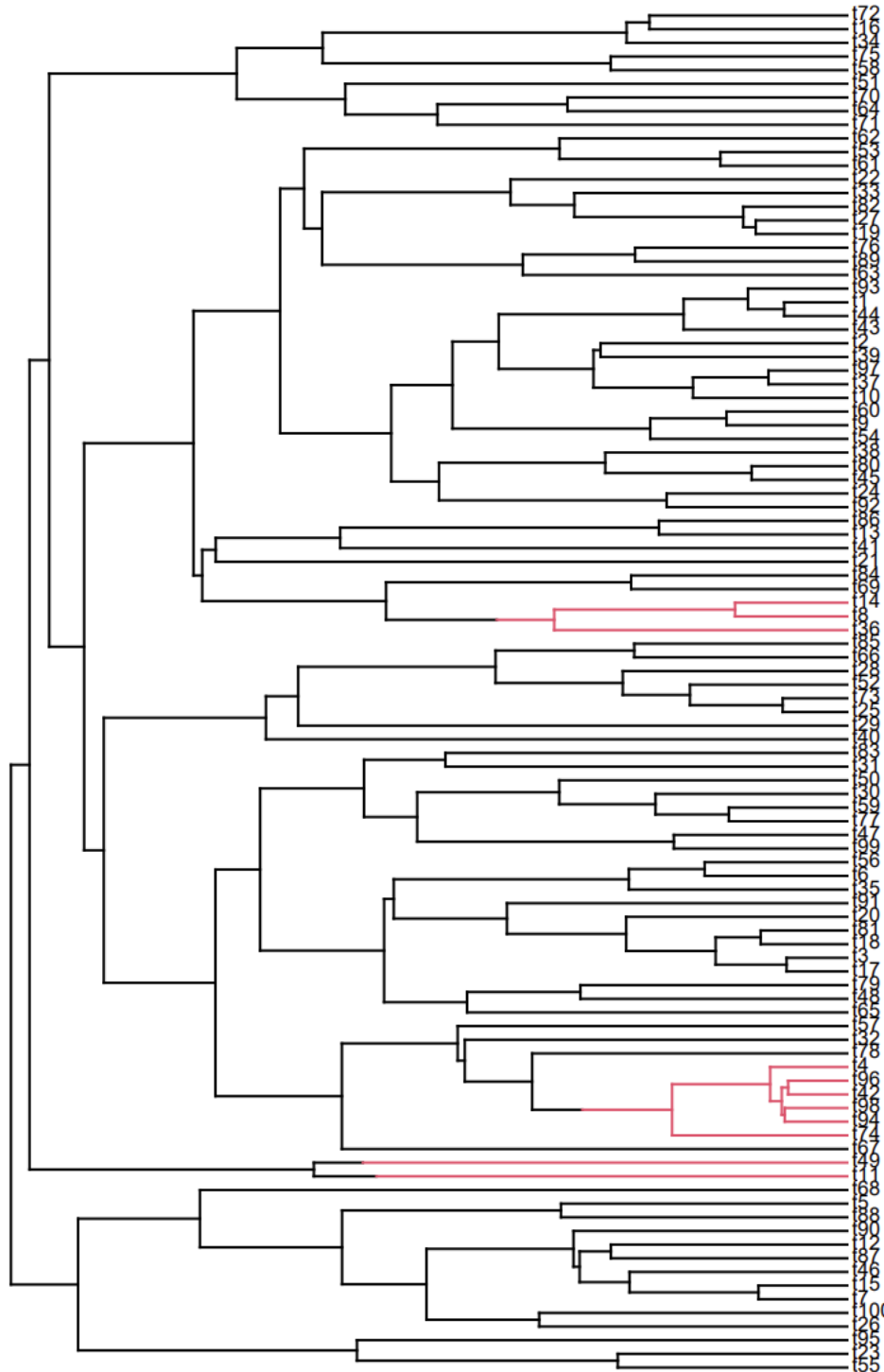
Let's load this in now:

```
library(convevol)
data("phy")
```

For our example, let's pretend that tips 4, 96, 42, 98, 94, 74, 14, 8, 36, 49, and 11 all share some ecological trait potentially associated with morphological convergence.

Let's save a vector with the tip names of these putatively convergent taxa and map it onto the tree:

```
conv.tax <- c("t4","t96","t42","t98","t94","t74","t14","t8","t36","t49","t11")
tax <- rep(0,length(phy$tip))
names(tax) <- phy$tip
tax[conv.tax] <- 1
set.seed(12345)
map <- make.simmap(phy,tax)
plot(map)
```



Here we can see that in our example we have three distinct clades possessing the ecological trait of interest, associated with three (or four?) separate evolutionary origins. We'll come back to this pattern in just a minute because it will be important to how we group our taxa for our analyses. But, first, we'll need some morphological data to test for convergence.

A simulated dataset is also included with the package for testing purposes. The simulated data include six morphological traits. In our focal lineages, four of those traits were simulated as evolving toward a shared adaptive peak (i.e., they're convergent), while the remaining two traits evolve via Brownian motion. (Using a mix of convergent and non-convergent traits may better reflect typical empirical datasets, such as that in Grossnickle et al. 2020.). For the non-focal lineages, we'll leave all six traits to evolve via a standard Brownian motion model (i.e., the lineages should be divergent).

Let's load these data in now (for those interested, these data were simulated using the OUwie package, with $\alpha = 0.1$, $\theta = 20$, $\sigma^2 = 0.1$ for convergent traits/taxa and $\theta = 0$, $\sigma^2 = 0.1$ for all other traits/taxa).

```
data("trait")
```

Now we're ready to start testing the *convevol* functions. Our first step is to think about which taxa we'd like to compare. *Ct*-measures (and *C*-measures) are computed by making pairwise comparisons between individual taxa of interest. However, not all of the possible pairwise comparisons are meaningful in our case because our focal taxa fall in three distinct clades. It is unlikely that we want to test whether two members of one of these clades are morphologically convergent with each other because our hypothesized driver of that morphological convergence (i.e., their shared ecology) was inherited from a common ancestor. Rather, we're much more likely to be interested in whether members of the clade {t14, t8, t36} are morphologically convergent with members of the clade {t4, t96, t42, t98, t94} or {t49, t11}. The "group" argument included in the *Ct*-measures allow us to modify which taxa we compare, and tailor this to our particular hypothesis. We can do this by providing a vector which includes a group name for each of our putatively convergent tips. If this vector is provided, *convevol* will only calculate *Ct*-measures for intergroup comparisons.

Let's define four groups corresponding to each independent origin of our ecology of interest, assuming that the most recent common ancestor of t49 and t11 did not possess this trait (but we'll come back to them). Note that if you're using the group function, you must assign all of your focal taxa to a group, but it is alright for some groups to include a single taxon.

```
group <- rep(1, length(conv.tax))
names(group) <- conv.tax
group[c("t14","t8","t36")] <- 2
group["t49"] <- 3
group["t11"] <- 4
```

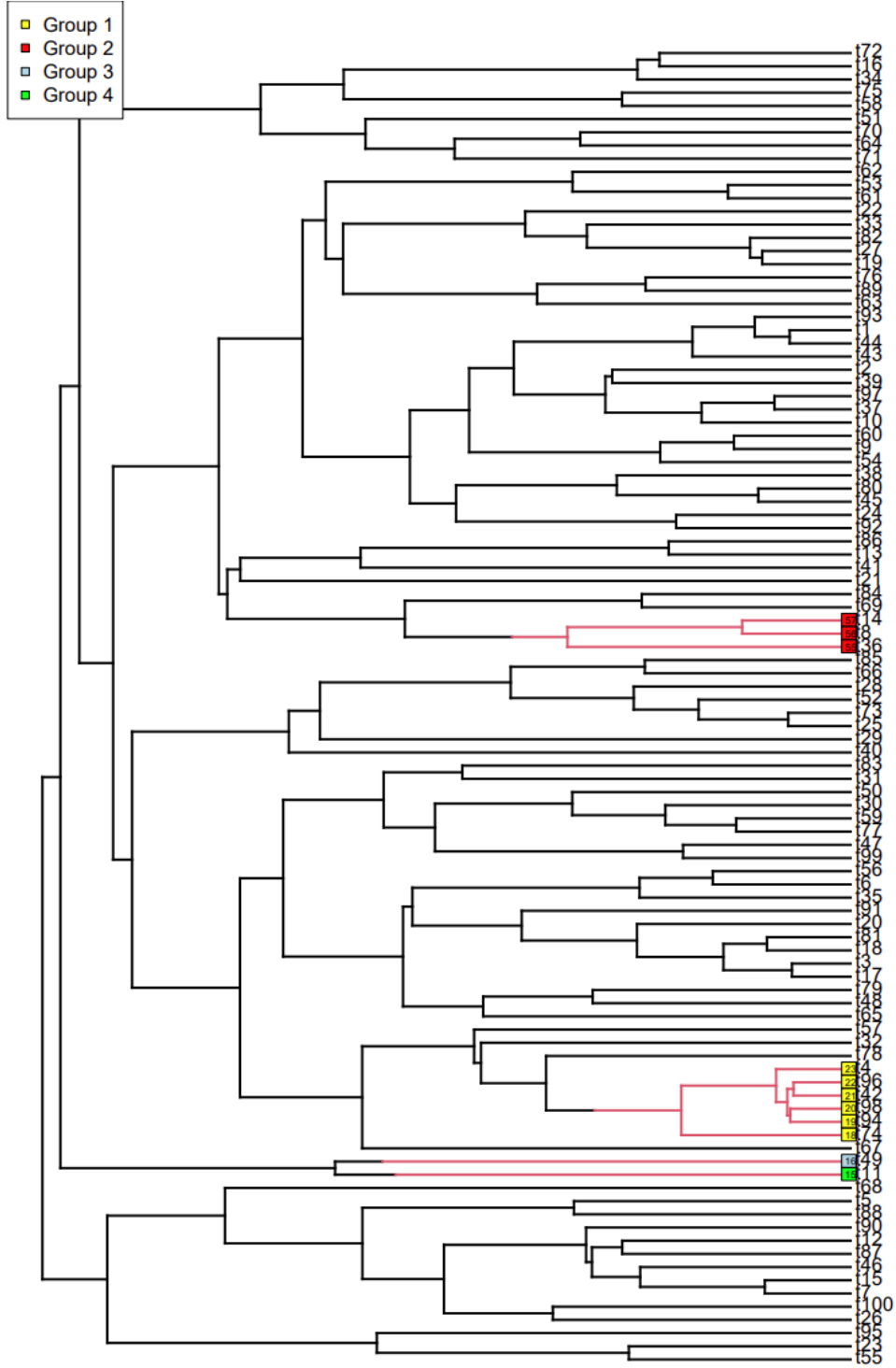
And now let's map our groups onto the tree.

```
col <- group
col[col == 1] <- "yellow"
col[col == 2] <- "red"
col[col == 3] <- "light blue"
col[col == 4] <- "green"
```

```

plot(map)
tiplabels(tip = match(names(group),phy$tip),bg = col,cex = 0.5)
legend("topleft",pch = 22, legend = c("Group 1","Group 2","Group 3","Group 4"), pt.bg =
c("yellow","red","light blue","green"), bg = "white")

```



Before we proceed to our analysis it's a good idea to use *convevol's* *pwCheck* function to ensure we've defined our groups in a way that doesn't cause any issues for the analyses. As already mentioned, *Ct*-measures measure $D_{\max.t}$ (the maximum historical phenotypic distance between lineages) at a specific point in time. They do this by measuring the phenotypic distance between two focal lineages at each internal node along the path between their most recent common ancestor and tips. The largest of these measures becomes $D_{\max.t}$. Occasionally you may find that there are tip pairings without any valid candidates for $D_{\max.t}$. This happens when there are no internal nodes between the tips and two taxa's most recent common ancestor (i.e., sister taxa). In most cases we likely don't want to test for convergence between these taxa anyway, but especially in large trees it is not always easy to know if we have these pairings in our dataset. The *pwCheck* function helps us by going through each pairwise comparison and calculating the number of valid points over which measurements of $D_{\max.t}$ can be made, given our phylogeny, list of putatively convergent taxa, and any potential groupings.

```
check <- pwCheck(phy,conv.tax,group)
check$taxa
```

The output returned shows a summary of the number of candidate $D_{\max.t}$ measurements that can be made for all possible pairwise comparisons. Note that the median value is 13, which means that $D_{\max.t}$ is drawn from 13 candidate measurements between lineages. Note, also, that the minimum is zero. This is problematic and means that at least one of our pairwise comparisons is between two lineages which have no valid time point from which to measure $D_{\max.t}$. The call to the "taxa" element in our *pwCheck* output returns a matrix with the offending taxa. In this case, it looks like we cannot compare t49 and t11 – as can be seen in the phylogeny above, there are no internal nodes between the tips of t49 and t11 and their most recent common ancestor, meaning that there are no points at which candidate $D_{\max.t}$ measurements can be obtained.

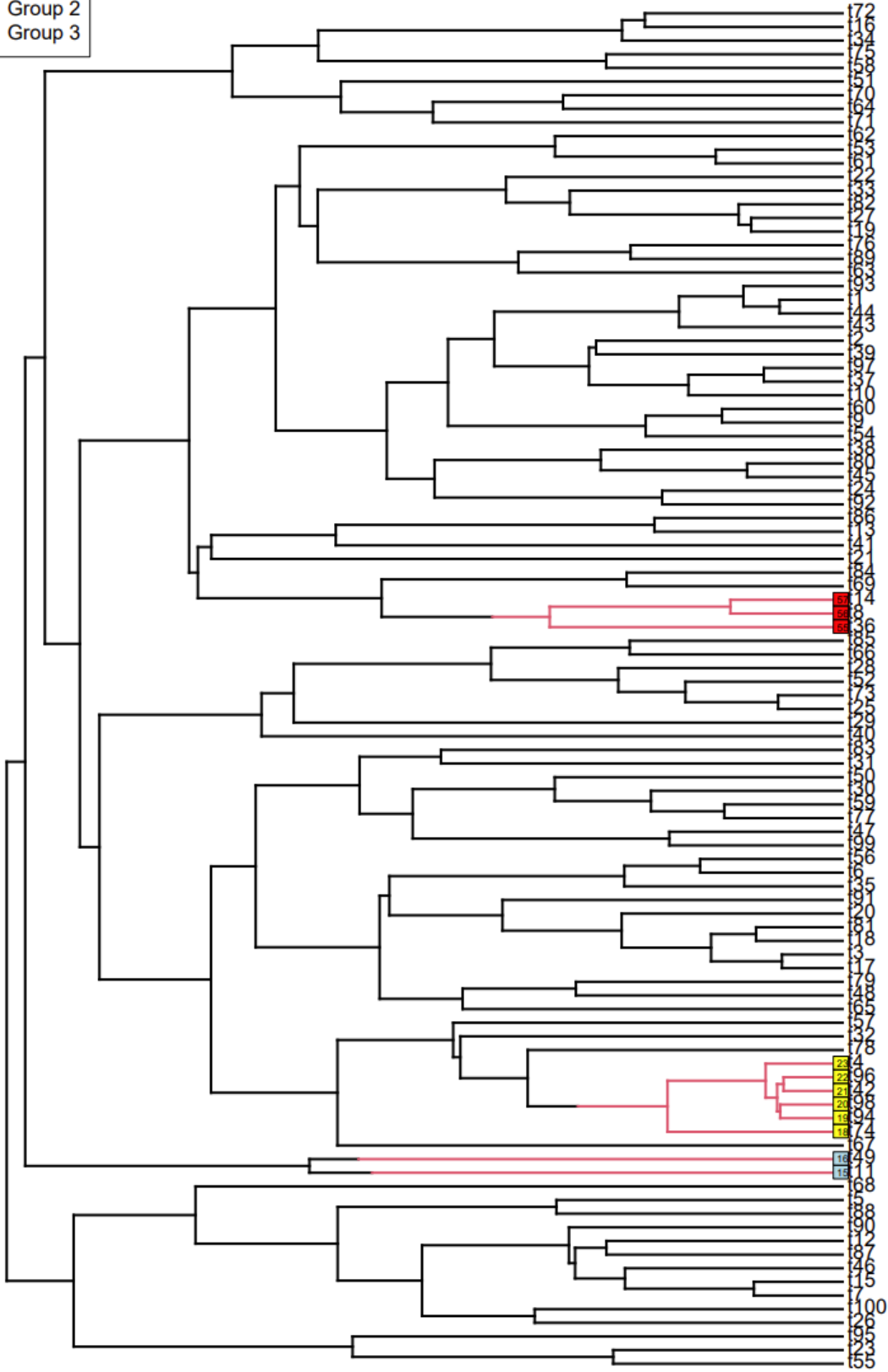
Because of the issue with t49 and t11, if we were to run our *Ct*-measures now, we'd get an error. To avoid this, let's re-format the group object to prevent comparing these two taxa. This can easily be done by placing them in the same group.

```
group["t11"] <- 3
check.ii <- pwCheck(phy,conv.tax,group)

## plot new groups
col.ii <- group
col.ii[col.ii == 1] <- "yellow"
col.ii[col.ii == 2] <- "red"
col.ii[col.ii == 3] <- "light blue"

plot(map)
tiplabels(tip = match(names(group),phy$tip),bg = col.ii,cex = 0.5)
legend("topleft",pch = 22, legend = c("Group 1","Group 2","Group 3"), pt.bg =
c("yellow","red","light blue"), bg = "white")
```

- Group 1
- Group 2
- Group 3



When we run *pwCheck* again, we see that the minimum has now increased to seven. We should be ready to run our *Ct*-measures. We'll start by just calculating the *Ct*-values without a significance test. This is generally advisable to ensure everything runs smoothly. Although computational run times for the *Ct*-measures (and *C*-measures) have been dramatically reduced relative to earlier versions of the package, significance tests can still be time consuming, which makes any troubleshooting more challenging.

Let's use the *calcConvCt* function to compute overall *Ct* values using our grouped taxa and simulated data.

```
Ct <- calcConvCt(phy,trait,conv.tax,group)
```

Now let's go through the different components of the output one by one.

```
Ct$mean          ## overall results, average of all pairwise comparisons
Ct$grp.mean      ## overall, group specific, and group weighted means (i.e.,
                  averaged so unique intergroup comparisons have equal weight)
```

```
> Ct$mean
      Ct1      Ct2      Ct3      Ct4
0.56001969 8.37499714 0.13634313 0.02031296
> Ct$grp.mean## overall, group specific, and group weighted means
      overall      1_2      1_3      2_3 overall.weighted
Ct1 0.56001969 0.55579489 0.61165444 0.46942454 0.54562463
Ct2 8.37499714 7.02221590 11.74533055 5.69267407 8.15340684
Ct3 0.13634313 0.11691997 0.18826599 0.09076687 0.13198428
Ct4 0.02031296 0.01767653 0.02770912 0.01342993 0.01960519
> |
```

The first of these components show the overall average *Ct*1–*Ct*4 values for all possible pairwise comparisons. The second shows group averaged comparisons. You'll notice that the first column matches the results shown by calling *Ct\$mean*, while the next three columns show *Ct* values calculated for each possible intergroup comparison. The final column again provides averaged results, but this time averaged giving equal weight to each group pairing as opposed to each taxon pairing (i.e., so one large group doesn't disproportionately influence results). For more in-depth discussion on how to interpret these values, see Stayton (2015) and Grossnickle et al. (2024). For now, we will focus on *Ct*1, which reflects the proportion of the maximum distance between two lineages ($D_{\max,t}$) that has been closed by subsequent evolution. That our comparisons here are positive supports the hypothesis that the focal taxa are morphologically convergent because the tips lie substantially closer to each other in morphospace than their ancestors did.

Next, let's look at an individual pairwise comparison as an example.

```
Ct$Cmat          ## Ct values for each pairwise comparisons - these are
                  calculated from the meas.path tables (see next lines)
```


Ct\$meas.path[1]

meas.path table showing measurements used to calculate Ct values for this pairwise comparison

```
> Ct$Cmat## Ct values for each pairwise comparisons - these are calculated from the meas.path tables (see next lines)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14]
Ct1 0.60507119 0.49836985 0.57081212 0.57822349 0.6936123 0.66816670 0.55080915 0.59961609 0.56689421 0.65725071 0.65824224 0.55585310 0.56953236 0.56253366
Ct2 7.64479952 6.29667652 7.21195171 11.23977434 13.1827436 8.44198255 6.95922327 7.57587676 11.01955050 12.49165804 8.31659155 7.02295123 7.19578255 10.93478804
Ct3 0.12921463 0.10413259 0.11949916 0.17795865 0.2142101 0.14224915 0.11474257 0.12515002 0.17396814 0.20237908 0.13931324 0.11512777 0.11818650 0.17167969
Ct4 0.01924371 0.01885018 0.01815414 0.02651643 0.0311002 0.02125041 0.01781796 0.01907022 0.02599689 0.02946982 0.02093477 0.01767838 0.01811344 0.02579692
      [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24] [,25] [,26] [,27] [,28]
Ct1 0.67778796 0.56682489 0.49398592 0.51163877 0.53330333 0.63101096 0.59175602 0.47835310 0.52622101 0.54330593 0.65913990 0.59849557 0.46618423 0.55240280
Ct2 12.89197790 7.16157487 5.48287378 6.46432339 10.17221098 11.99294723 7.47656833 6.04877402 6.64856328 10.56103065 12.52756378 7.56171951 5.91529496 6.97935837
Ct3 0.20752390 0.12123327 0.09081066 0.10727281 0.16128809 0.19516503 0.12726250 0.10063985 0.11092568 0.16931646 0.20494391 0.12686890 0.09712024 0.11480995
Ct4 0.03039065 0.01802733 0.01380165 0.01627218 0.02399788 0.02829328 0.01882024 0.01521357 0.01673596 0.02491516 0.02955453 0.01903458 0.01489015 0.01756865
      [,29] [,30] [,31] [,32] [,33] [,34] [,35] [,36]
Ct1 0.56249293 0.48429845 0.36231616 0.48627744 0.41056509 0.48814144 0.48838594 0.58086119
Ct2 10.93399631 13.00575252 4.39377919 5.89704778 4.97889020 5.91965242 5.92261743 7.04405740
Ct3 0.17192218 0.20983668 0.07000928 0.09644893 0.07771714 0.09479641 0.09261635 0.11301309
Ct4 0.02579505 0.03068259 0.01036563 0.01391208 0.01174600 0.01396541 0.01397240 0.01661806
> Ct$meas.path[1]## meas.path table showing measurements used to calculate Ct values for this pairwise comparison
[[1]]
      node path      height node.anc1 node.anc2 node.anc3 node.anc4 node.anc5 node.anc6 int.anc1 int.anc2 int.anc3 int.anc4 int.anc5 int.anc6 anc.diff
X23 23 tip 50.000000 16.310461 16.264286 15.625030 17.295711 -1.27078671 -0.42772005 NA NA NA NA NA NA NA 4.9897460
X126 126 P1 45.395670 15.473236 15.462726 15.496426 16.041950 -0.79260483 -0.27214763 17.501514 17.005466 16.909059 16.142709 0.2310271 3.4587940 4.8443154
X125 125 P1 39.527033 12.569614 13.006561 13.134819 13.313785 -0.64053265 -0.27133348 16.021467 15.660701 15.411480 14.853601 -0.1773503 3.3445436 6.3082674
X124 124 P1 31.161815 5.577642 5.706699 7.174040 7.137010 -0.95151297 0.60638838 12.769762 12.953024 12.265231 12.080975 -0.7474636 2.9405036 6.2634585
X123 123 P1 27.138581 4.010511 3.637584 5.343463 5.369104 -0.96057943 0.26995928 10.412427 10.597823 9.835992 9.798649 -0.9565251 2.5177269 11.5881677
X122 122 P1 26.707596 3.918622 3.494999 5.177346 5.213949 -0.99231777 0.25252041 10.161670 10.347293 9.577586 9.555871 -0.9787637 2.4833921 11.3631399
X121 121 P1 19.786039 3.478890 2.514759 3.713430 3.978510 -1.05615237 -0.62581126 6.529946 7.024328 6.177951 6.378051 -1.2562543 1.7775211 6.8792005
X120 120 P1 12.221165 3.253239 2.197399 2.449922 3.637926 -1.65465845 -0.69729825 3.306719 4.679106 3.838619 4.226646 -1.4143684 0.7386082 3.2490564
X119 119 P1 5.548962 3.192961 3.153614 2.492531 3.703945 -1.38185103 -0.01552789 3.243700 3.662727 2.817595 3.866568 -1.3705439 0.1409671 0.6469166
X118 118 mrca 4.373167 3.317332 3.520558 2.655739 3.859357 -1.36437332 0.04975347 NA NA NA NA NA NA NA 4.9897460
X57 57 tip 50.000000 17.842774 17.685002 17.439154 16.596977 0.53826411 3.24368422 NA NA NA NA NA NA NA 4.9897460
X160 160 P2 43.289113 17.345382 16.694567 16.666532 15.934873 0.09046105 3.55721031 14.430976 14.581081 14.648723 15.062671 -0.7380183 -0.2718554 5.7568213
X159 159 P2 32.479753 13.541458 13.724022 13.060467 12.828117 -0.67902534 2.94616771 6.679226 6.856791 8.113159 8.110160 -0.9025181 0.4681035 12.1315627
X158 158 P2 22.426758 7.655090 7.842989 6.994556 7.129056 -1.20106036 2.14018105 3.643564 2.888741 4.271945 4.449856 -1.0317981 -0.2907093 7.8208431
X157 157 P2 11.435485 2.971958 4.435534 3.595659 4.003203 -1.43078992 0.63070781 3.246141 2.309997 2.454940 3.645700 -1.6225316 -0.6170170 2.7596585
X156 156 P2 10.910771 2.907927 4.311038 3.555681 3.899453 -1.39868285 0.55691521 3.241401 2.385196 2.458290 3.650892 -1.6010777 -0.5634013 2.5263005
```

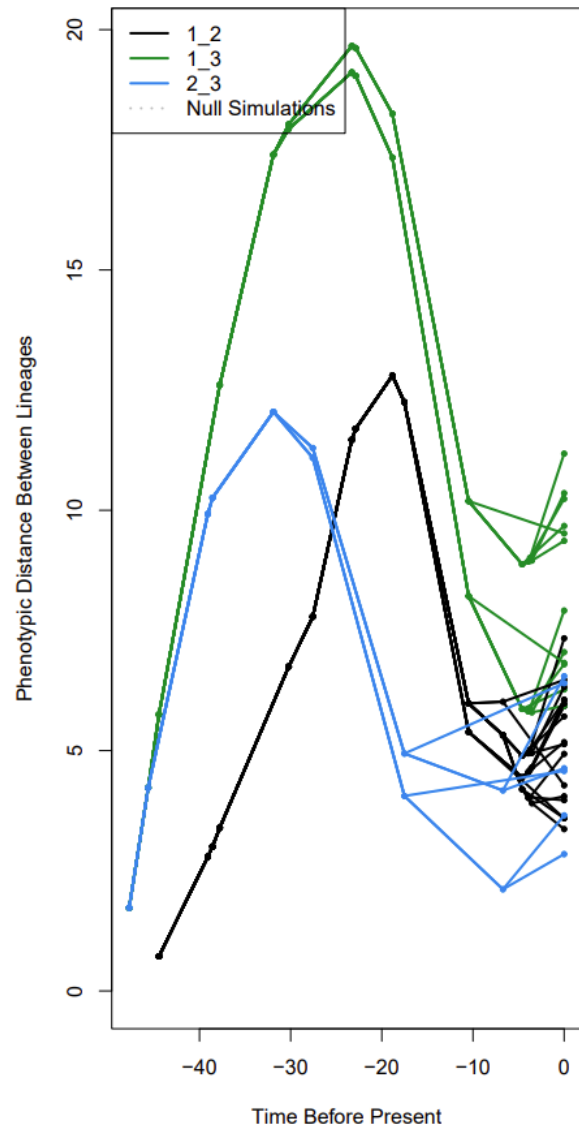
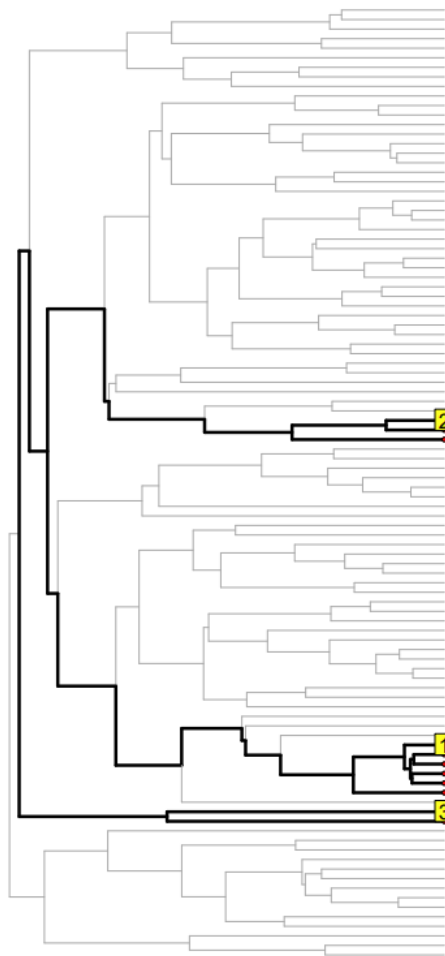


The first of these commands returns the *Ct* values calculated for all 36 pairwise comparisons. The second shows the measurements made between the two lineages that forms the basis of the first of these 36 pairwise comparisons (i.e., the first column of the *Cmat* matrix). For each node along the phylogenetic paths between the two focal tips to their most recent common ancestor, this table shows the reconstructed (or observed) value for each of our six morphological traits (the node.anc1 – node.anc6 columns), the corresponding values reconstructed at the same time point along the opposite path (the int.anc1 – int.anc6 columns), and the euclidean distance between them (anc.diff). $D_{max.t}$ is the largest value in the anc.diff column (excluding tips). $Ct1$ is calculated as the distance between tips (i.e., D_{tip}) divided by $D_{max.t}$, with the resulting value subtracted from one. Thus, in this example $Ct1$ is $1 - (4.98/12.63)$, or 0.61. Note that this value matches the first $Ct1$ value in the *Ct\$Cmat* output.

The last thing we'll do before conducting a significance test is plot the distance between our lineages through time using *convevol's* plotting function.

```
plotCt(Ct,phy,conv.tax,groups = group)
```

The lefthand panel of the output shows our focal lineages on the tree, with group labels, while the righthand panel shows distance through time. The latter is color coded to show different intergroup comparisons. Note the dramatic decrease in phenotypic distance observed for all pairwise comparisons, which is indicative of morphological convergence.



Now let's rerun the same analysis, this time including a significance test. This test compares observed morphological distances between lineages to those expected under Brownian motion, obtained using simulated data. As such we'll need to specify the number of simulations we'd like to conduct. In the interest of time, we'll limit ourselves to 25 simulated datasets, but you'll likely want to run more in empirical cases. This analysis should take only a few minutes to run on most personal computers.

```
set.seed(12345)
system.time(Ct.sig <- convSigCt(phy,trait,conv.tax,group,nsim = 25))
```

Again, let's look at the individual components of the output.

```
Ct.sig$grp.mean
```

Ct.sig\$grp.pvals

```
> Ct.sig$grp.mean
  overall      1_2      1_3      2_3 overall.weighted
Ct1 0.56001969 0.55579489 0.61165444 0.46942454 0.54562463
Ct2 8.37499714 7.02221590 11.74533055 5.69267407 8.15340684
Ct3 0.13634313 0.11691997 0.18826599 0.09076687 0.13198428
Ct4 0.02031296 0.01767653 0.02770912 0.01342993 0.01960519
> Ct.sig$grp.pvals
  overall  1_2 1_3 2_3 overall.weighted
Ct1      0 0.00  0  0              0
Ct2      0 0.04  0  0              0
Ct3      0 0.04  0  0              0
Ct4      0 0.00  0  0              0
> |
```

The grp.mean table here is the same as we saw above, but now we also have the grp.pvals table which indicates whether the values in the grp.mean table are significantly different from the null expectation (calculated from our simulated datasets). Note that the layout of the grp.pvals table is exactly the same as grp.mean. It also appears that all these values are statistically significant. This coupled with the positive *Ct1* values strongly suggests convergent evolution (unsurprising since that's what we simulated these taxa to do)!

Next we can look in a bit more depth at the statistics calculated for our simulated datasets.

Ct.sig\$sim.avg ## average Ct measurements for all pairwise comparisons for each NULL simulation

Ct.sig\$sim.path[1] ## path table for first pairwise comparison (tip 23 and tip 57 from first and second NULL simulated dataset)

Ct.sig\$sim.path[length(Ct.sig\$meas.path)+1]

```

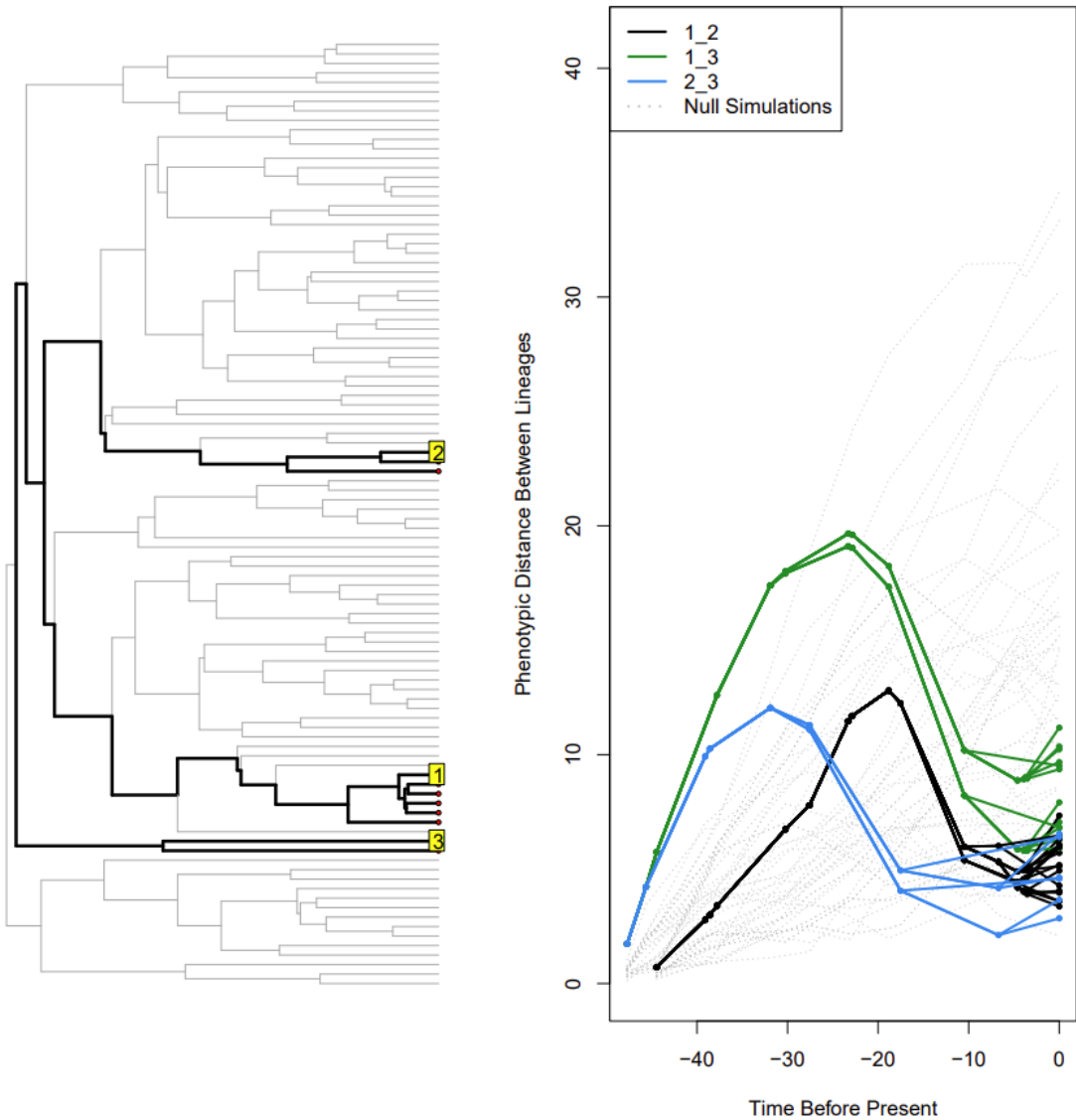
> Ct.sig$sim.avg## average Ct measurements for all pairwise comparisons for each NULL simulation
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]      [,9]      [,10]      [,11]      [,12]      [,13]
Ct1 -0.0280131706 -0.058194381 0.241631803 -0.0618295511 -0.379370914 0.0384645684 -0.074228273 4.634013e-02 -0.048029878 -0.145555937 -0.123079434 -0.091316699 -0.0556609740
Ct2 -0.1679748111 -0.583294398 4.017163815 -0.495364483 -1.849455990 0.1503590614 -0.1054075248 5.820698e-03 -0.882808651 -1.144947508 -1.607633027 -0.921306237 -0.0912092258
Ct3 -0.0096829500 -0.027354332 0.102242571 -0.0184066326 -0.082996803 0.0129219591 -0.031174063 4.689660e-03 -0.0273707202 -0.044648724 -0.052531842 -0.031750915 -0.0099609315
Ct4 -0.0002879078 -0.001175301 0.008660369 -0.0009404045 -0.004512115 0.0003943592 -0.002362684 5.918772e-05 -0.001936204 -0.002617246 -0.00297809 -0.001869079 -0.000193035
      [,14]      [,15]      [,16]      [,17]      [,18]      [,19]      [,20]      [,21]      [,22]      [,23]      [,24]      [,25]
Ct1 -0.048885584 -0.114170795 -0.182900556 -0.182696698 -0.064343530 -0.069669623 0.031539475 -0.062364914 -0.096094872 -1.135782e-02 -0.16552096 0.0105953386
Ct2 -0.459973282 -0.921072682 -1.173803992 -0.983378593 -0.965175640 -1.402238086 0.868369051 -1.020768948 -0.596393273 -9.000075e-02 -1.69601503 0.1901278031
Ct3 -0.020658699 -0.035968918 -0.045567739 -0.045184212 -0.026820534 -0.043300242 0.020866179 -0.030188509 -0.016262989 -7.195224e-03 -0.05747110 0.0070234484
Ct4 -0.001015988 -0.001868015 -0.002519498 -0.002023056 -0.002451426 -0.002608245 0.0001952631 -0.001953552 -0.001201438 -0.00326427 0.0004625535
>
> Ct.sig$sim.path[1]## path table for first pairwise comparison (tip 23 and tip 57 from first and second NULL simulated dataset)
$meas.path1
  node path height node.anc1 node.anc2 node.anc3 node.anc4 node.anc5 node.anc6 int.anc1 int.anc2 int.anc3 int.anc4 int.anc5 int.anc6 anc.diff
X23 23 tip 50.000000 -4.9235987 -4.7258953 -4.546837 -3.0187349 3.0684288 -4.7166419 NA NA NA NA NA NA 10.2192899
X126 126 P1 45.395670 -6.1870316 -5.2699007 -5.546160 -3.7998870 3.1218836 -4.5017613 1.5705269 -1.033847 -1.355272 -0.6701643 1.5575882 -2.0721815 10.6692292
X125 125 P1 39.527033 -5.5586747 -4.4255861 -4.883623 -3.4312915 2.2155584 -4.4885215 0.4434426 -1.691067 -1.777106 -1.1112422 1.5285763 -2.0684007 8.0538818
X124 124 P1 31.161815 -7.3622664 -6.1891851 -5.830069 -5.3234233 1.3029623 -3.7813363 -0.9196970 -2.478180 -2.505462 -2.1117759 1.5895869 -1.8120462 8.9780962
X123 123 P1 27.135851 -7.1739462 -6.0784334 -5.404218 -5.2096230 0.7640389 -3.8673195 -1.5067221 -2.962790 -3.120028 -2.5780568 1.5577066 -1.5276243 7.7505748
X122 122 P1 26.707596 -7.1484412 -6.0945125 -5.384220 -5.2353721 0.7206041 -3.8851718 -1.5691659 -3.014340 -3.185401 -2.6276566 1.5543154 -1.4973694 7.6581613
X121 121 P1 19.786039 -6.6799848 -6.6911456 -5.641558 -5.9358719 0.5401536 -3.0761137 -2.5017077 -3.800025 -3.971327 -3.3266721 1.4596690 -1.1045384 6.3360516
X120 120 P1 12.221165 -6.1485940 -7.3846900 -6.476256 -6.6587826 0.2144153 -2.1721122 -3.3850480 -4.574628 -4.350770 -3.908240 1.2856446 -0.8487203 5.5242914
X119 119 P1 5.548962 -4.3912519 -5.4253471 -5.044712 -4.5137365 0.4661975 -1.3113934 -3.9310705 -4.889599 -4.587150 -4.2024521 0.6402260 -1.1495989 0.9279486
X118 118 mrca 4.373167 -4.0555936 -4.9680237 -4.663420 -4.2880585 0.5134172 -1.12267571 NA NA NA NA NA NA 10.2192899
X57 57 tip 50.000000 2.6677106 -0.3306576 -1.048372 -0.6954226 1.7043647 -1.8899677 NA NA NA NA NA NA 10.2192899
X160 160 P2 43.289113 1.0685473 -1.3555676 -1.495684 -0.6586082 1.4904356 -2.1555472 -5.9614819 -4.966833 -5.308341 -3.6675790 2.7965567 -4.4970088 9.6561485
X159 159 P2 32.479753 -0.7275286 -2.3195379 -2.304278 -1.9591343 1.6000232 -1.9051545 -7.0781109 -5.911330 -5.680957 -5.0253183 1.4467416 -3.8927532 8.8322340
X158 158 P2 22.426758 -2.1933544 -3.5296299 -3.838873 -3.1234567 1.5204168 -1.1949413 -6.8587107 -6.463518 -5.543378 -5.6686166 0.6090178 -3.3847863 6.7366530
X157 157 P2 11.435485 -3.4767908 -4.6550772 -4.390178 -3.9692857 1.2675706 -0.8199125 -5.9416597 -7.153969 -6.307686 -6.4061945 0.2440637 -2.0707590 4.9454518
X156 156 P2 10.910771 -3.3632298 -4.239347 -3.8120739 1.2184936 -0.7977452 -5.8034590 -6.999883 -6.195106 -6.2375041 0.2638643 -2.0030705 4.9098045
>
> Ct.sig$sim.path[length(Ct.sig$meas.path)+1]
$meas.path1
  node path height node.anc1 node.anc2 node.anc3 node.anc4 node.anc5 node.anc6 int.anc1 int.anc2 int.anc3 int.anc4 int.anc5 int.anc6 anc.diff
X23 23 tip 50.000000 -1.2772798 0.8942641 -0.03392312 -1.0247405 2.2467692543 1.6492080 NA NA NA NA NA NA 18.848017
X126 126 P1 45.395670 -0.4500963 0.2846356 0.77294966 -0.8562682 2.3262803650 1.3653047 -6.5111055 -7.60531607 -12.16702430 -6.5791432 1.46114649 -2.7778424 17.807176
X125 125 P1 39.527033 -1.2455230 -0.3245666 0.39889596 -1.4595998 2.3301477346 1.6320219 -5.5955516 -6.63947663 -10.47044739 -5.7810722 1.21503147 -2.4484742 14.612060
X124 124 P1 31.161815 -1.5512992 -0.6756788 0.13870759 -1.7694192 2.3205848354 1.2933522 -4.3750440 -5.42753717 -8.09880671 -4.5364668 0.69581868 -2.0303869 10.944706
X123 123 P1 27.135851 -1.4609231 -0.9907141 0.14631791 -1.4169914 2.4631370200 1.1116951 -4.3421903 -5.24432216 -7.34644007 -4.5163400 0.36591631 -1.9170390 10.280785
X122 122 P1 26.707596 -1.5265667 -1.0854135 0.08949624 -1.4348372 2.4647235288 1.0618753 -4.3386956 -5.22483298 -7.26640836 -4.5141991 0.33587019 -1.5049819 10.097966
X121 121 P1 19.786039 -1.5864852 -1.4248228 -0.27219414 -1.5787022 2.2678273934 0.4846649 -4.0389150 -4.63478635 -5.90261055 -4.0938242 -0.1214479 -1.7483247 8.064773
X120 120 P1 12.221165 -0.9062046 -1.0060036 0.05888804 -0.8238257 1.5398131318 -0.2773799 -3.2802070 -3.50256232 -4.28748980 -2.9608824 -0.46804002 -1.6448131 6.417586
X119 119 P1 5.548962 -1.6423120 -1.6970264 -1.22542952 -1.5210589 0.2655526901 -1.0585089 -2.1495322 -2.31471295 -2.23526895 -2.0024023 0.01590577 -1.1888912 1.403428
X118 118 mrca 4.373167 -1.9605806 -2.1164356 -1.87401036 -1.8583698 0.1300587278 -1.1025242 NA NA NA NA NA NA 10.2192899
X57 57 tip 50.000000 -7.1006936 -8.3722444 -13.38538256 -6.8789686 1.5024099602 -3.0532137 NA NA NA NA NA NA 10.848017
X160 160 P2 43.289113 -6.2413592 -7.2544337 -11.60503018 -6.4419682 1.4422677797 -2.651854 -0.7356160 0.06596141 0.63867889 -1.0728350 2.32766856 1.4610432 16.743275
X159 159 P2 32.479753 -4.3857990 -5.4875144 -8.34510122 -4.5430555 0.7893633299 -2.0674924 -1.5031242 -0.62036117 0.17969863 -1.7206072 2.32209147 1.3467095 11.253597
X158 158 P2 22.426758 -4.3037620 -5.0300191 -6.46641110 -4.4927982 -0.0004565698 -1.7844582 -1.5636251 -1.29533103 -0.13420169 -1.5238147 2.34294739 0.7048827 9.058496
X157 157 P2 11.435485 -3.2014085 -3.3849708 -4.11974514 -2.8321778 -0.5040370205 -1.6340625 -0.9928843 -1.08737435 -0.09234566 -0.9059278 1.38976370 -0.3693611 5.939096
X156 156 P2 10.910771 -3.0112720 -3.2189385 -3.88266400 -2.6592130 -0.5046493807 -1.5827381 -1.0507732 -1.14171769 -0.19334682 -0.9607595 1.28955345 -0.4307905 5.403735

```

The first element here shows Ct values calculated for each simulated dataset (notice that the number of columns is equal to the number of simulations). The next shows the measurements from one simulated dataset for the comparison between “t23” and “t57”. It has exactly the same layout as the meas.path table we looked at earlier, but was computed using different data. The final element shows measurements for the same comparison, but using a different simulated dataset. Note that the number of tables in the “sim.path” component of the output will equal the number of pairwise comparisons multiplied by the number of simulations – in our case 36*25 = 900.

Finally, let’s graph our results again, this time adding in our same traces from our null simulations. Note that the value given for nsim here indicates the number of pairwise comparisons from the null distribution, not the number of simulated datasets (as it indicates above). As such in our example, the maximum value is 900, not 25.

plotCt(Ct.sig,phy,conv.tax,nsim = 50, groups = group)



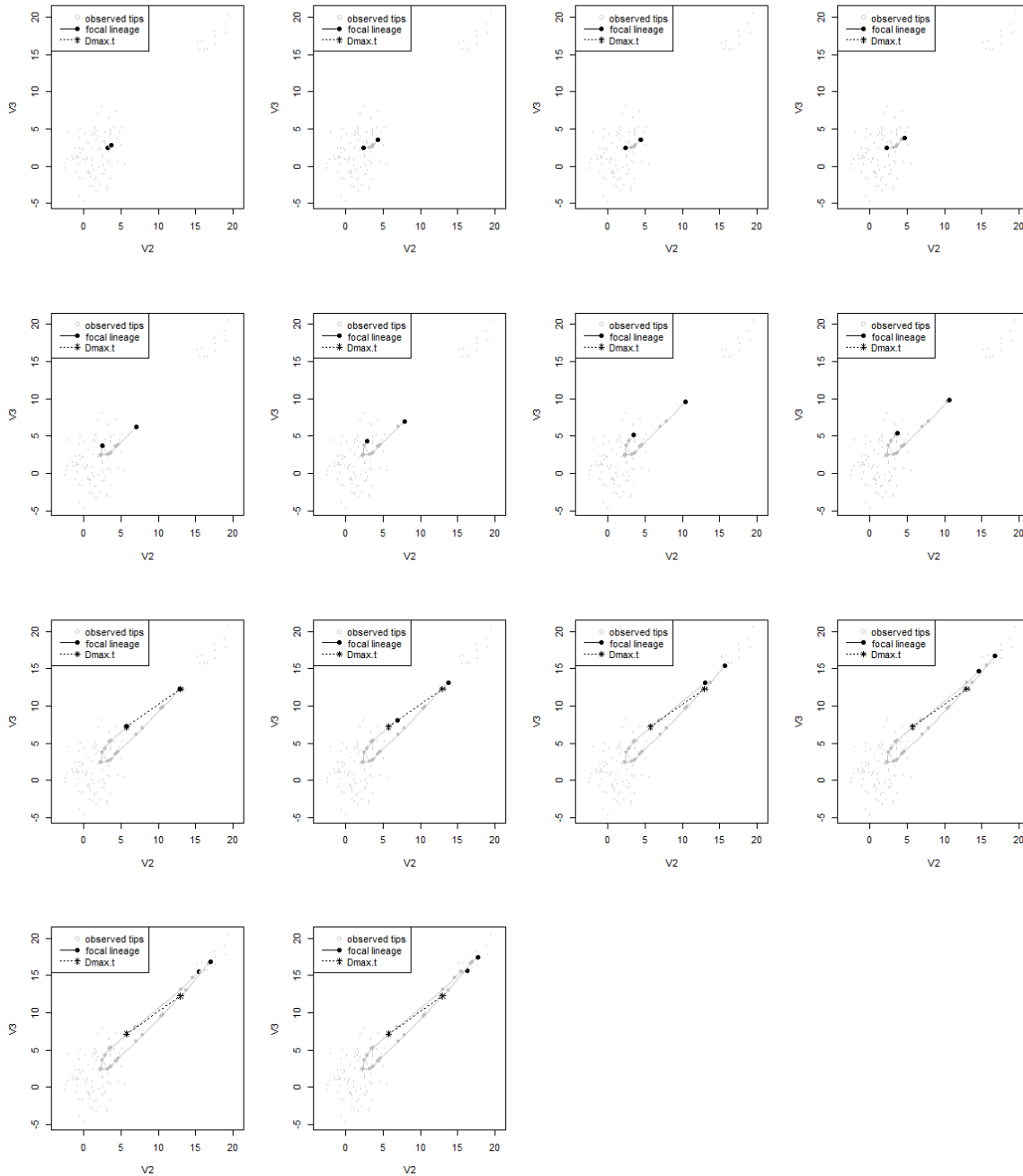
In contrast to our focal lineages, which show a decrease in phenotypic distances between lineages following a peak, null simulations show a general increase in the morphological distance between the same lineages through time.

In addition to the distance through time plots we've been using, *convevol* also provides the option to plot two-dimensional morphospace time series using the *plot2D* function. This allows us to visualize how two lineages move through morphospace over time, much like the example figure from the start of this tutorial. Let's do this now for our results. First, we'll need to choose the taxa we'd like to visualize. In this case let's start off with "t4" and "t14", which come from groups 1 and 2 respectively (it will shortly become clear why we've chosen these two taxa). We'll also need to choose which of our six simulated traits we'd like to visualize. If you're having trouble deciding, note that *plot2D* does allow the user to run and visualize principal components analysis (even if *Ct* values were calculated on untransformed data).

However, for our purposes we can just arbitrarily choose two of the traits we know were simulated to be convergent.

```
tip <- c("t4","t14")  
foc.trt <- c("V2","V3")  
par(mfrow = c(4,4))  
X <- plot2D(Ct, phy, tip, foc.trt, trait, cex = 0.5, leg = TRUE) ## plot morphospace
```

```
## choose species to highlight  
## choose focal traits to plot  
## set parameters to panel all 15 plots
```



It's possible to visualize these timeseries in a variety of ways. By default, each time slice will be output as a single plot, but these can be paneled (as in our example), or the `save = TRUE` option can be used to output each as a separate png file which will be combined into an animated gif (using functions from the *magick* package).

Although this is a handy way to visualize our data, the location of $D_{\max.t}$ is potentially a bit troubling in our particular example. We can clearly see that $D_{\max.t}$ is largely capturing differences between focal lineages that accumulate after one of those lineages starts getting "pulled toward" their shared selective optimum. If we'd prefer $D_{\max.t}$ not to record differences that arise after species begin shifting toward the morphospace region where they putatively are convergent, *convevol* allows users to run more conservative analyses (for more detail, including discussion of the rationale for conservative analyses, see Grossnickle et al., 2024). These analyses fix $D_{\max.t}$ to be measured before the origin of the stem lineage of the oldest group of putatively convergent taxa for each pairwise comparison. Although this resolves potential bias associated with asynchronous movement of lineages to the shared optimum, users should note that 1) *Ct* values will typically be considerably lower when using this option, and 2) many candidate nodes for $D_{\max.t}$ may be ruled out, meaning some taxon comparisons will no longer be possible. As such, it's a good idea for us to go back to the handy *pwCheck* function before we run new analyses. We can run *pwCheck* as before, we just need to indicate that we'd like to run a conservative analysis (i.e., `conservative = TRUE`).

```
check.iii <- pwCheck(phy, conv.tax, group, conservative = TRUE) ## run new pwCheck
check.iii$taxa ## the taxa pairings which are invalid
check.iii$group ## the group pairings which are invalid
```

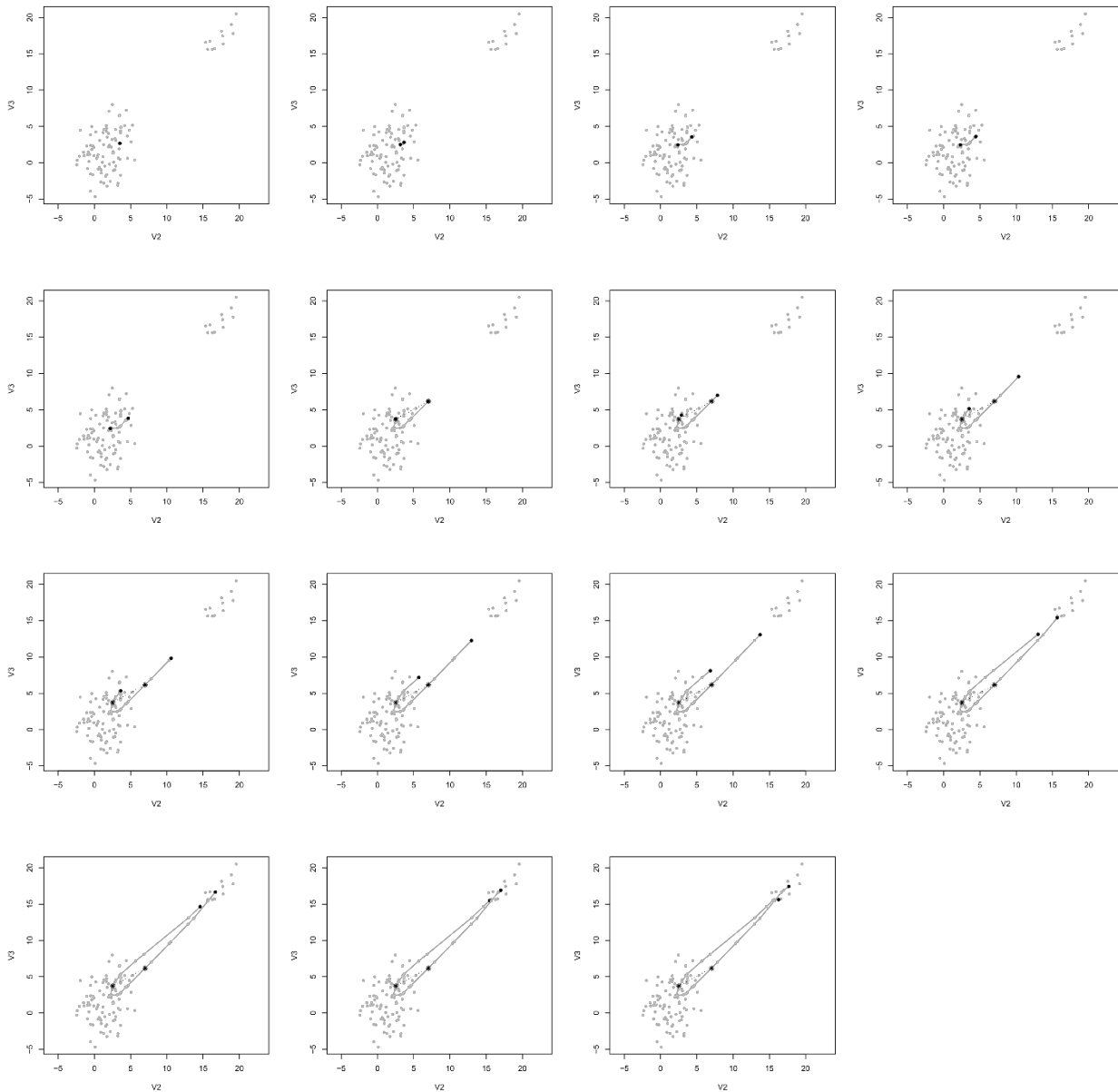
```
> check.iii <- pwCheck(phy, conv.tax, group, conservative = TRUE) ## run new pwCheck
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0    0.0     2.5     2.5    5.0     5.0
> check.iii$taxa ## the taxa pairings which are invalid
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16] [,17] [,18]
[1,] "t4"  "t4"  "t96" "t96" "t42" "t42" "t98" "t98" "t94" "t74" "t14" "t14" "t8"  "t8"  "t36" "t36"
[2,] "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11" "t49" "t11"
> check.iii$group
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16] [,17] [,18]
[1,]  1    1    1    1    1    1    1    1    1    1    1    1    2    2    2    2    2    2
[2,]  3    3    3    3    3    3    3    3    3    3    3    3    3    3    3    3    3    3
> |
```

Unfortunately, not all of our pairwise comparisons are compatible with a conservative analysis. If we look at the returned list of uninformative taxon and group pairings, we can clearly see that all comparisons with species in group 3 fail to pass our check. This is because the parent node of the stem lineage leading to this group is also its most recent common ancestor with species in the other two groups. In this particular case we have a few options, we could split group 3 in two again, and run separate analyses alternating which of the two taxa are included. Or, we can cut our losses and just drop group 3 altogether. For simplicity's sake we'll do the latter.

```
conv.tax <- names(group[group %in% c(1,2)]) ## remove taxa in group 3 from the analysis
group <- group[group %in% c(1,2)]
Ct.cons <- calcConvCt(phy, trait, conv.tax, group, conservative = TRUE)
```

And now we can plot our two-dimensional morphospace again to see if the questionable $D_{\max.t}$ placement has been resolved.

```
par(mfrow = c(4,4))  
Y <- plot2D(Ct.cons, phy, tip, foc.trt, trait, cex = 0.5)
```



Although $D_{\max.t}$ looks a lot shorter than it was before, this looks much better. Furthermore, we're still getting positive $Ct1$ values suggesting convergence, albeit weaker convergence than originally thought. Finally, we can run another significance test to confirm whether our results are still significant.


```
set.seed(12345)
Ct.cons.sig <- convSigCt(phy, trait, conv.tax, group, nsim = 25, conservative = TRUE)
```

```
> set.seed(12345)
> Ct.cons.sig <- convSigCt(phy, trait, conv.tax, group, nsim = 25, conservative = TRUE)
[[1]]
  overall  1_2 overall.weighted
Ct1  0.184 0.184           0.184
Ct2  1.267 1.267           1.267
Ct3  0.021 0.021           0.021
Ct4  0.003 0.003           0.003

[[2]]
  overall  1_2 overall.weighted
Ct1  0.04 0.04           0.04
Ct2  0.04 0.04           0.04
Ct3  0.04 0.04           0.04
Ct4  0.04 0.04           0.04
```

They are, but not as convincingly as before. This really drives home a trait of the *Ct*-measures. They are generally more conservative (particularly in terms of raw *Ct* values) than the original *C*-measures. Given the earlier *C*-measures propensity to mischaracterize divergent lineages as convergent (see Grossnickle et al., 2024 for more details), we think this is a reasonable trade-off. Regardless, in our example, the combination of statistically significant results, positive *Ct* values, and evidence from our visualizations suggest that example taxa show phenotypic convergence.

References

Grossnickle, D. M., et al. (2020). Incomplete convergence of gliding mammal skeletons. *Evolution*, 74, 2662-2680.

Grossnickle, D. M., Brightly, W. H., et al. (2024). Challenges and advances in methods for measuring phenotypic convergence. *Evolution*. (Advanced article, <https://doi.org/10.1093/evolut/qpae081>)

Stayton, C. T. (2015). The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution*, 69, 2140-2153.