Comparison of two trees

General information

Both trees are **rooted**. Both trees contain **91 nodes**. Both trees contain **46 leaves**. Both trees have branch lengths. 40 leaves are shared between the two trees; **6 leaves** are present only in the first tree, while **6** are present only in the second tree. These are summarised in the following table:

In both trees	Only in first tree	Only in second tree

Acaryochloris.marina.MBIC11017
Angiopteris.evecta.vascCT
Anthoceros.formosae.strCT
Chaetosphaeridium.globosum.strCT
Coccomya.sp.C169
Coffea.arabica.vascCT
Cyanobium.gracile.PCC6307
Cyanophora.paradoxa.glauCT
Cyanothece.sp.CCY0110
Cyanothece.sp.PCC7424
Cyanothece.sp.PCC7425
Fischerella.sp.JSC11
Floydiella.terrestris.cCT
Gnetum.parvifolium.vascCT
Illicium.oligandrum.vascCT
Lyngbya.majuscula.3L
Mesostigma.viride.strCT
Microcystis.aeruginosa.NIES843
Nephroselmis.olivacea.cCT
Nodosilinea.nodulosa.PCC7104
Nostoc.azollae.0708
Oscillatoria.sp.PCC6506
P.marinus.str.AS9601
P.marinus.str.MIT9303
P.marinus.str.MIT9515
Pedinomonas.minor.cCT
Pinus.contorta.vascCT
Porphyra.purpurea.rCT
Pseudanabaena.sp.PCC7367
Stigeoclonium.helveticum.cCT
Synechococcus.sp.JA33Ab
Synechococcus.sp.PCC7002
Synechococcus.sp.RCC307
Synechococcus.sp.RS9916
Synechococcus.sp.WH7803
Synechococcus.sp.WH8102
Synechococcus.spongiarum.SH4
Synedra.acus.bacCT
Thalassiosira.pseudonana.bacCT
Zygnema.circumcarinatum.strCT

Anabaena.cylindrica.PCC7122
Anabaena.variabilis.ATCC29413
Chlorella.vulgaris.cCT
Fistulifera.sp.JPCC.DA0580.bacCT
Pseudanabaena.biceps.PCC7429
Trichodesmium.erythraeum.IMS101

Amborella.trichopoda.vascCT
Arthrospira.sp.PCC8005
Cryptomeria.japonica.vascCT
Nostoc.punctiforme.PCC73102
P.marinus.subsp.marinus.str.CCMP1375
Pleurocapsa.sp.PCC7327

The following analyses are based on the subset of leaves shared between the two trees.

The trees have the same topology. The edge-length distance between them is **2.22**.

Considering the 79 splits in common between the two trees, the average difference in length between a split from the first tree and the corresponding split from the second tree is **-0.02534** (89% highest-density interval: -0.5762 — 0.2402). **Figure 1** shows the distribution of the difference in length between splits from the first tree and the corresponding splits from the second tree.
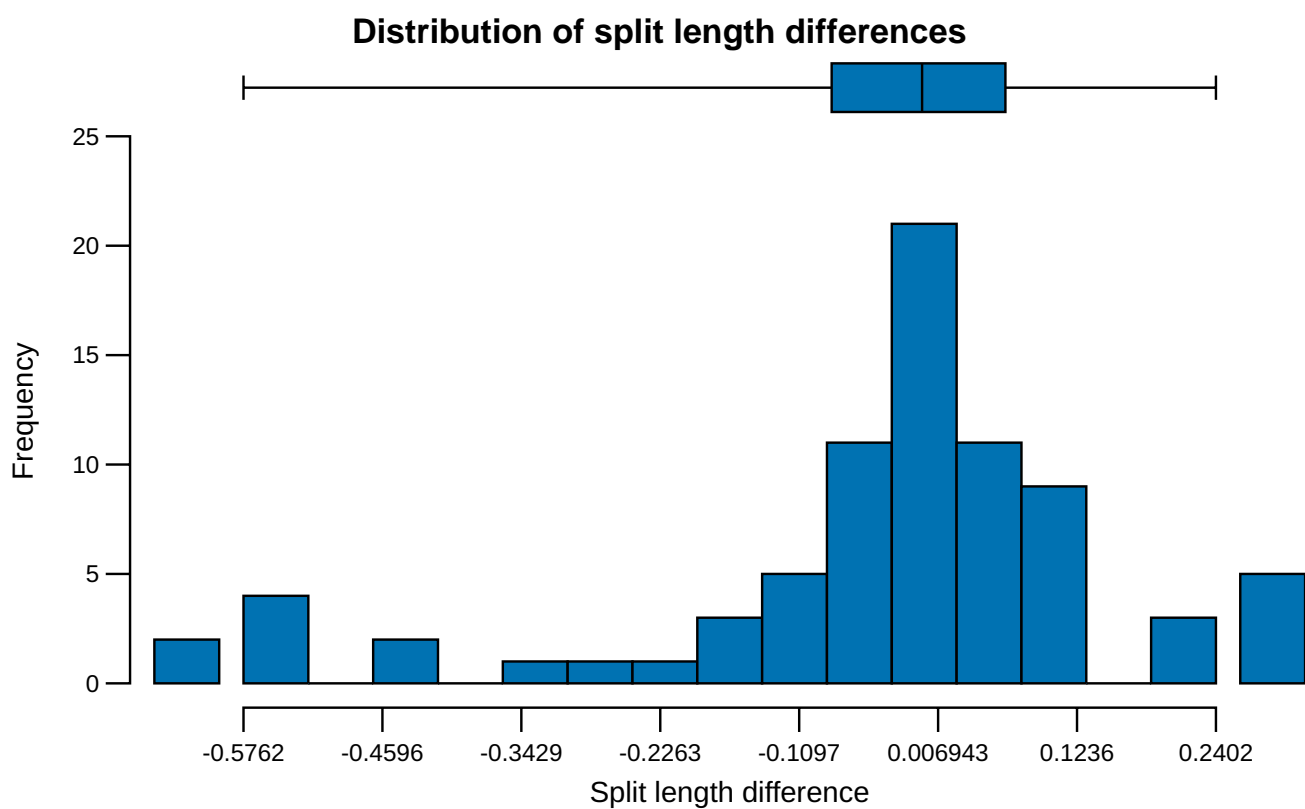
Figure 1. Distribution of split length differences. The histogram shows the distribution of the difference in length between splits from the first tree and the corresponding splits from the second tree. 2 values smaller than -0.5762 are shown in the underflow bin; 5 values greater than 0.2402 are shown in the overflow bin. The box and whisker plot at the top represents the median branch length, the interquartile range, and the 89% HDI.

Tree shape statistics

The **Sackin index** of the trees is **338** (YHK model normalisation: 0.5145, PDA model normalisation: 1.08). According to the Sackin index, the balancing of the trees is not significantly different than expected under the YHK model. The trees are significantly more balanced than expected under the PDA model (p ≈ 0.008). **Figure 2** shows the expected distribution of Sackin index values under the YHK and PDA models.
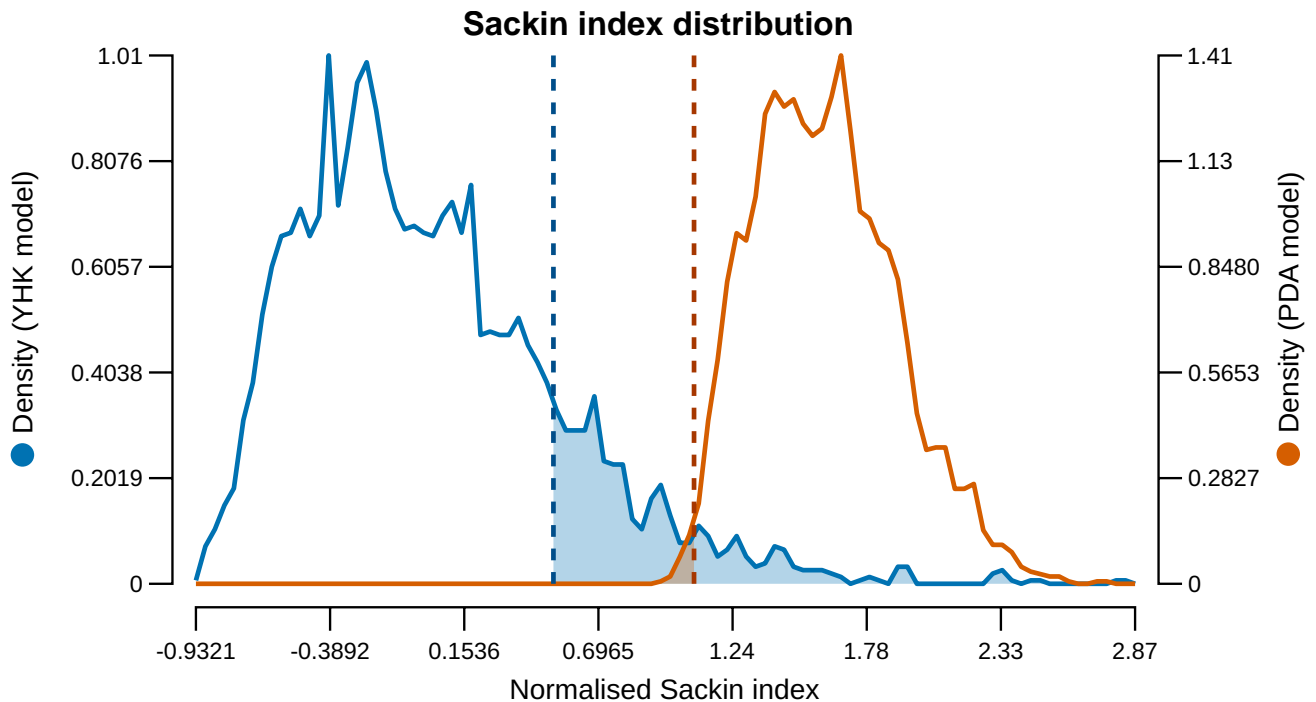
Figure 2. Distribution of Sackin index values. The curve shows the expected distribution of (normalised) Sackin index values under the null hypothesis for a tree containing 40 tips, computed using a Monte Carlo approach (2000 random trees were sampled according to the YHK or PDA model). The dashed line represents the observed value of the Sackin index, normalised according to the YHK or PDA model. The blue plot refers to the YHK model, while the orange plot to the PDA model.

The **Colless index** of the trees is **186** (YHK model normalisation: 1.23, PDA model normalisation: 0.5962). According to the Colless index, the trees are significantly less balanced than expected under the YHK model (p ≈ 0.05). The trees are significantly more balanced than expected under the PDA model (p ≈ 0.04). **Figure 3** shows the expected distribution of Colless index values under the YHK and PDA models.
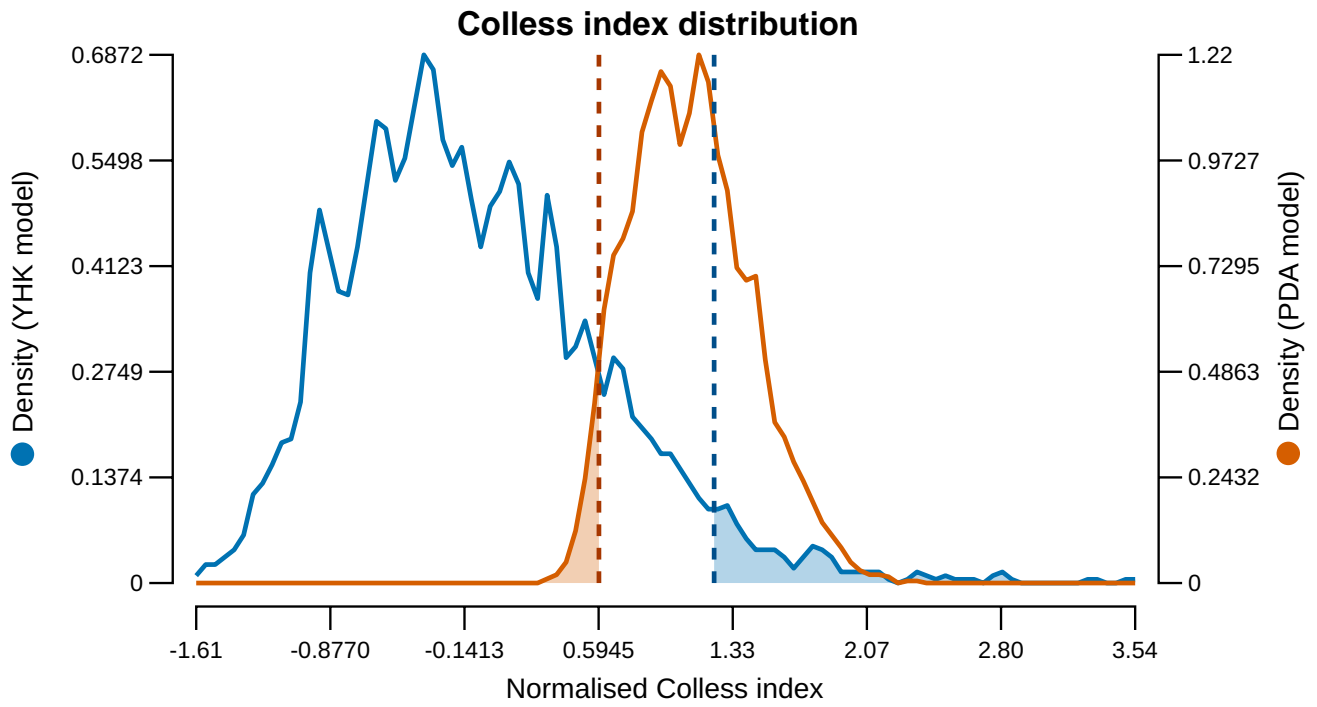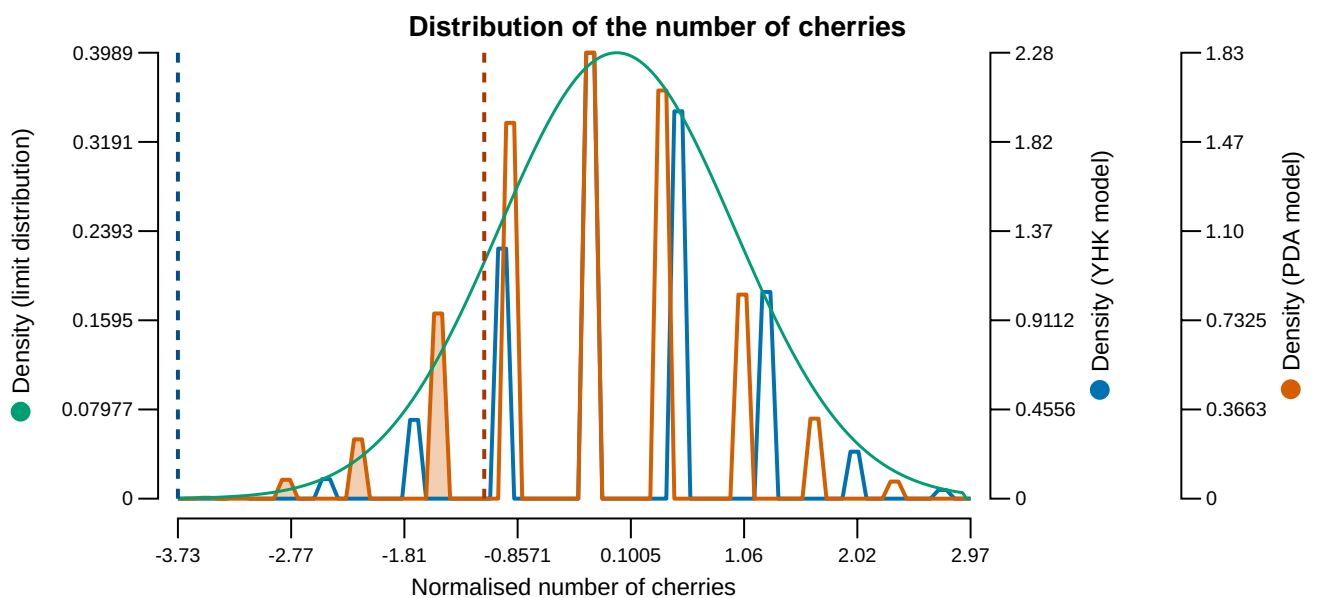
Figure 3. Distribution of Colless index values. The curve shows the expected distribution of (normalised) Colless index values under the null hypothesis for a tree containing 40 tips, computed using a Monte Carlo approach (2000 random trees were sampled according to the YHK or PDA model). The dashed line represents the observed value of the Colless index, normalised according to the YHK or PDA model. The blue plot refers to the YHK model, while the orange plot to the PDA model.

The trees have **10** cherries (YHK model normalisation: -3.73, PDA model normalisation: -1.14). The number of cherries is significantly lower than expected under the YHK model (MC approximation: p ≈ 0, Gaussian approximation: p ≈ 0.00010). The number of cherries does not differ significantly from the number expected under the PDA model. **Figure 4** shows the expected distribution of the number of cherries under the YHK and PDA models.

Figure 4. Distribution of the number of cherries. The curve shows the expected distribution of the (normalised) number of cherries under the null hypothesis for a tree containing 40 tips, computed using a Monte Carlo approach (2000 random trees were sampled according to the YHK or PDA model). The dashed line represents the observed value of the number of cherries, normalised according to the YHK or PDA model. The blue plot refers to the YHK model, while the orange plot to the PDA model. The green curve represents the limit distribution for the (normalised) number of cherries as the number of tips grows to infinity.