

beastscriptr demo

Richel Bilderbeek

2016-04-19

Richel Bilderbeek

2016-04-19

Introduction

The purpose of `beastscriptr` is to create a valid `BEAST2` XML input file from its function argument. In this way, a scientific pipeline using `BEAST2` can be fully scripted, instead of using `BEAUti2` its GUI.

Limitations

`BEAUti2` contains multiple tree priors. This package currently only supports:

- These tree priors:
 - the constant-rate birth-death model
 - the constant-population coalescent model
- DNA data

Demonstration

The library is added first:

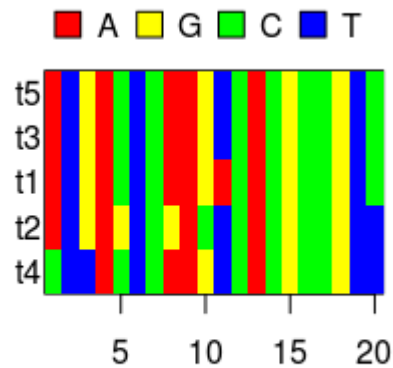
```
library(beastscriptr)
```

A BEAST2 XML input file needs an alignment (as BEAST2 infers phylogenies and parameters on DNA sequences). in this demonstration, a random alignment is created:

```
n_taxa <- 5
dna_sequence_length_nucleotides <- 20
mutation_rate <- 0.1

alignment <- create_random_alignment(
  n_taxa = n_taxa,
  sequence_length = dna_sequence_length_nucleotides,
  rate = mutation_rate
)

image(alignment)
```



This random alignment must be saved to a FASTA file:

```
fasta_filename <- "demo.fas"

phangorn::write.phyDat(
  x = alignment,
  file = fasta_filename,
  format = "fasta"
)
if (file.exists(fasta_filename)) {
```

```

    print("File created")
  } else {
    print("Error: file not created")
  }

## [1] "File created"

```

We also need to define some parameters:

```

# The MCMC chain length BEAST2 will have to use
mcmc_chainlength <- 10000000

# The tree prior BEAST2 will have to use, can be 'birth_death' or 'coalescent_constant_popu
tree_prior <- "birth_death"
# tree_prior <- "coalescent_constant_population"

# The date in YYYYMMDD format.
# This is important, as BEAST2 and BEAUti2 can change formats
date_str <- "20151022"

# The name of the file you intend to let BEAST2 run
output_xml_filename <- "demo.xml"

```

Now we can create our XML file:

```

beast_scripttr(
  input_fasta_filename = fasta_filename,
  mcmc_chainlength = mcmc_chainlength,
  tree_prior = tree_prior,
  date_str = date_str,
  output_xml_filename = output_xml_filename
)

if (file.exists(output_xml_filename)) {
  print("File created")
} else {
  print("Error: file not created")
}

## [1] "File created"

```

The file indeed is a BEAST2 input file:

```

print(readLines(output_xml_filename))

```

```

## [1] "<?xml version=\"1.0\" encoding=\"UTF-8\" standalone=\"no\"?><beast beautitemplate=
## [2] ""
## [3] ""
## [4] "    <data"
## [5] "id=\"demo\""
## [6] "name=\"alignment\">"
## [7] "                <sequence id=\"seq_t4\" taxon=\"t4\" totalcount=\"4\" value=\\
## [8] "                <sequence id=\"seq_t2\" taxon=\"t2\" totalcount=\"4\" value=\\
## [9] "                <sequence id=\"seq_t1\" taxon=\"t1\" totalcount=\"4\" value=\\
## [10] "                <sequence id=\"seq_t3\" taxon=\"t3\" totalcount=\"4\" value=\\
## [11] "                <sequence id=\"seq_t5\" taxon=\"t5\" totalcount=\"4\" value=\\
## [12] "                </data>"
## [13] ""
## [14] ""
## [15] "    "
## [16] ""
## [17] ""
## [18] "    "
## [19] ""
## [20] ""
## [21] "    "
## [22] "<map name=\"Uniform\">beast.math.distributions.Uniform</map>"
## [23] "<map name=\"Exponential\">beast.math.distributions.Exponential</map>"
## [24] "<map name=\"LogNormal\">beast.math.distributions.LogNormalDistributionModel</map>"
## [25] "<map name=\"Normal\">beast.math.distributions.Normal</map>"
## [26] "<map name=\"Beta\">beast.math.distributions.Beta</map>"
## [27] "<map name=\"Gamma\">beast.math.distributions.Gamma</map>"
## [28] "<map name=\"LaplaceDistribution\">beast.math.distributions.LaplaceDistribution</ma
## [29] "<map name=\"prior\">beast.math.distributions.Prior</map>"
## [30] "<map name=\"InverseGamma\">beast.math.distributions.InverseGamma</map>"
## [31] "<map name=\"OneOnX\">beast.math.distributions.OneOnX</map>"
## [32] ""
## [33] ""
## [34] "<run id=\"mcmc\" spec=\"MCMC\" chainLength=\"10000000\">"
## [35] "    <state id=\"state\" storeEvery=\"5000\">"
## [36] "        <tree id=\"Tree.t:demo\" name=\"stateNode\">"
## [37] "            <taxonset id=\"TaxonSet.demo\" spec=\"TaxonSet\">"
## [38] "                <alignment idref=\"demo\"/>"
## [39] "            </taxonset>"
## [40] "        </tree>"
## [41] "        <parameter id=\"birthRate2.t:demo\" lower=\"0.0\" name=\"stateNode\" upper
## [42] "        <parameter id=\"relativeDeathRate2.t:demo\" lower=\"0.0\" name=\"stateNode
## [43] "    </state>"
## [44] ""
## [45] "    <init id=\"RandomTree.t:demo\" spec=\"beast.evolution.tree.RandomTree\" estima
## [46] "        <populationModel id=\"ConstantPopulation0.t:demo\" spec=\"ConstantPopulat

```

```

## [47] "          <parameter id=\"randomPopSize.t:demo\" name=\"popSize\">1.0</parameter>
## [48] "        </populationModel>"
## [49] "      </init>"
## [50] ""
## [51] "    <distribution id=\"posterior\" spec=\"util.CompoundDistribution\">"
## [52] "      <distribution id=\"prior\" spec=\"util.CompoundDistribution\">"
## [53] "        <distribution id=\"BirthDeath.t:demo\" spec=\"beast.evolution.speciat
## [54] "        <prior id=\"BirthRatePrior.t:demo\" name=\"distribution\" x=\"@birthRa
## [55] "          <Uniform id=\"Uniform.0\" name=\"distr\" upper=\"1000.0\"/>"
## [56] "        </prior>"
## [57] "        <prior id=\"DeathRatePrior.t:demo\" name=\"distribution\" x=\"@relativ
## [58] "          <Uniform id=\"Uniform.01\" name=\"distr\"/>"
## [59] "        </prior>"
## [60] "      </distribution>"
## [61] "    <distribution id=\"likelihood\" spec=\"util.CompoundDistribution\">"
## [62] "      <distribution id=\"treeLikelihood.demo\" spec=\"TreeLikelihood\" data=
## [63] "        <siteModel id=\"SiteModel.s:demo\" spec=\"SiteModel\">"
## [64] "          <parameter id=\"mutationRate.s:demo\" estimate=\"false\" name=
## [65] "          <parameter id=\"gammaShape.s:demo\" estimate=\"false\" name=
## [66] "          <parameter id=\"proportionInvariant.s:demo\" estimate=\"false\"
## [67] "          <substModel id=\"JC69.s:demo\" spec=\"JukesCantor\"/>"
## [68] "        </siteModel>"
## [69] "        <branchRateModel id=\"StrictClock.c:demo\" spec=\"beast.evolution
## [70] "          <parameter id=\"clockRate.c:demo\" estimate=\"false\" name=
## [71] "        </branchRateModel>"
## [72] "      </distribution>"
## [73] "    </distribution>"
## [74] "  </distribution>"
## [75] ""
## [76] "  <operator id=\"treeScaler.t:demo\" spec=\"ScaleOperator\" scaleFactor=\"0.5\"
## [77] ""
## [78] "  <operator id=\"treeRootScaler.t:demo\" spec=\"ScaleOperator\" rootOnly=\"true
## [79] ""
## [80] "  <operator id=\"UniformOperator.t:demo\" spec=\"Uniform\" tree=\"@Tree.t:demo
## [81] ""
## [82] "  <operator id=\"SubtreeSlide.t:demo\" spec=\"SubtreeSlide\" tree=\"@Tree.t:demo
## [83] ""
## [84] "  <operator id=\"narrow.t:demo\" spec=\"Exchange\" tree=\"@Tree.t:demo\" weight=
## [85] ""
## [86] "  <operator id=\"wide.t:demo\" spec=\"Exchange\" isNarrow=\"false\" tree=\"@Tree
## [87] ""
## [88] "  <operator id=\"WilsonBalding.t:demo\" spec=\"WilsonBalding\" tree=\"@Tree.t:de
## [89] ""
## [90] "  <operator id=\"BirthRateScaler.t:demo\" spec=\"ScaleOperator\" parameter=\"@b
## [91] ""
## [92] "  <operator id=\"DeathRateScaler.t:demo\" spec=\"ScaleOperator\" parameter=\"@re

```

```

## [93] ""
## [94] "      <logger id=\"tracelog\" fileName=\"demo.log\" logEvery=\"1000\" model=\"@posterior\">
## [95] "          <log idref=\"posterior\"/>
## [96] "          <log idref=\"likelihood\"/>
## [97] "          <log idref=\"prior\"/>
## [98] "          <log idref=\"treeLikelihood.demo\"/>
## [99] "          <log id=\"TreeHeight.t:demo\" spec=\"beast.evolution.tree.TreeHeightLogger\">
## [100] "              <log idref=\"BirthDeath.t:demo\"/>
## [101] "              <log idref=\"birthRate2.t:demo\"/>
## [102] "              <log idref=\"relativeDeathRate2.t:demo\"/>
## [103] "          </logger>
## [104] ""
## [105] "      <logger id=\"screenlog\" logEvery=\"1000\">
## [106] "          <log idref=\"posterior\"/>
## [107] "          <log id=\"ESS.0\" spec=\"util.ESS\" arg=\"@posterior\"/>
## [108] ""
## [109] "          <log idref=\"likelihood\"/>
## [110] "          <log idref=\"prior\"/>
## [111] "      </logger>
## [112] ""
## [113] "      <logger id=\"treelog.t:demo\" fileName=\"$(tree).trees\" logEvery=\"1000\" model=\"@posterior\">
## [114] "          <log id=\"TreeWithMetaDataLogger.t:demo\" spec=\"beast.evolution.tree.TreeWithMetaDataLogger\">
## [115] "      </logger>
## [116] ""
## [117] "</run>"
## [118] ""
## [119] "</beast>"

```

This XML input file should be readable by BEAST2.

All left to do is cleaning up the files created by this demonstration.

```
file.remove(fasta_filename)
```

```
## [1] TRUE
```

```
file.remove(output_xml_filename)
```

```
## [1] TRUE
```