microPopGut: R package for simulating microbial populations in the human colon

Helen Kettle

Overview of microPopGut

MicroPopGut is an R package which simulates/predicts the growth of interacting microbial populations in the human colon based on the solution of a system of ordinary differential equations (ODEs). It models the colon as three compartments (proximal, transverse and distal) with a microbial community based on 10 microbial functional groups (MFGs) in each one. Protein and carbohydrates (resistant starch and non-starch polysaccharide (NSP)) enter the proximal compartment. This inflow can be modelled as a constant rate or with fluctuations to simulate meals (which have passed through the stomach and small intestine). The outflow from the distal compartment can be modelled as a constant flow or bowel movements can be simulated.

To simulate microbial growth we use our previous R package, microPop, which is a generic package for modelling microbial communities. The 10 microbial functional groups used in microPopGut are instrinsic data frames in microPop (see Table 1 in the paper Kettle et al. 2018, in the link below for details of the MFGs). The main difference to the usual use of microPop used here is that we include water as a state variable. Water is injested by the host and absorbed through the colon walls and its volume is not constant. When calculating concentrations we use the current volume of water in each compartment.

the microPop package

For background info, the microPop package is described in the paper:

Kettle H, G Holtrop, P Louis, HJ Flint. 2018. microPop: Modelling microbial populations and communities in R. Methods in Ecology and Evolution, 9(2), p399-409. doi: 10.1111/2041-210X.12873

https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12873

The user specifies the system via a number of input files (csv files that become dataframes) and the function `microPopModel()` will construct and solve the necessary equations (ordinary differential equations) and provide an output containing the solution (e.g. the concentrations of microbes, substrates and metabolites at the required time points) as well as all the settings/parameters involved in the simulation, and plots of the microbes and resources over time.

Also see the webpage:

https://www.bioss.ac.uk/people/helen/microPop/microPop.html

Getting started with microPopGut

Install the package and add the library:

```
install.packages('microPopGut')
```

```
#> Loading required package: usethis
#> i Loading microPopGut
#> Loading required package: microPop
#> Loading required package: deSolve
#> Loading required package: visNetwork
```

```
library(microPopGut)
```

Basic model run

Note that simulating a complex microbial community of 10 MFGs in 3 compartments is fairly slow to run, so the code evaluated in this tutorial only simulates very short time periods.

In this first example we only use two microbial groups and have constant inflow and outflow to and from the colon and we run the simulation for 2 hours.

```
sim.time.h=2 #time to simulate in hours

m.out=microPopGut(
    numDays=sim.time.h/24, #number of days to simulate
    time.step=1/24/60, #time step at which you want output
    transitTime=1.25,  #time taken to move through colon (days)
    microbeNames=c('Bacteroides','ButyrateProducers1'), #vector of MFG names
                  #(these are data frames provided in the microPop package)
    microbeNames.short=c('Bacteroides'='B','ButyrateProducers1'='BP1') #abbreviated names used for plot

)
#> [1] "simulating growth in compartment 1 - please wait!"
#> [1] "using microPopGut"
#> [1] "Set up completed, ODE solver called..."
#> [1] "simulating growth in compartment 2 - please wait!"
#> [1] "using microPopGut"
#> [1] "Set up completed, ODE solver called..."
#> [1] "simulating growth in compartment 3 - please wait!"
#> [1] "using microPopGut"
#> [1] "Set up completed, ODE solver called..."
```

Here we can see several messages telling us which compartment the processor is currently simulating - this is helpful if the model is running over several hours.

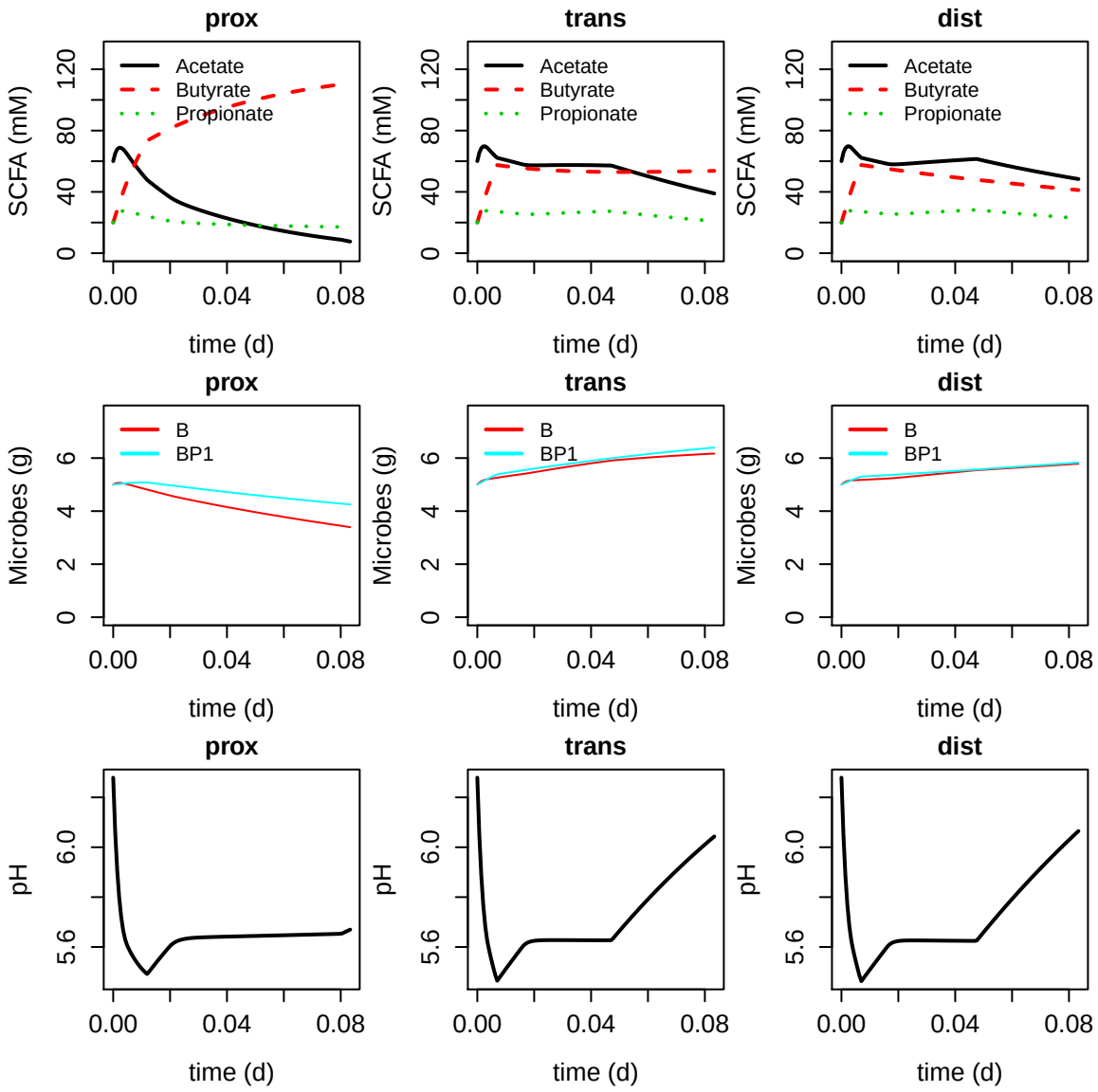
To look at the results of the model there are two inbuilt functions. The first is **verification()** (see below) which gives a summary of the concentrations of the short chain fatty acids (SCFA) in each compartment, which are released as the microbes grow (averaged over the time period **start.av** to **fin.av**), as well as the fraction of the incoming water that reaches the end of the colon (should be about 0.1, i.e. 10%) and the microbial outflow from the colon (should be around 16 g/d). Note, TSCFA is the total SCFA.

```
time=m.out$solution[[1]][,'time']
verification(m.out,start.av=0.8*max(time),fin.av=max(time))
#>                   prox trans  dist
#> TSCFA (mM)      136.1 118.5 117.7
#> Acetate (mM)     10.1  42.8  51.1
#> Butyrate (mM)   108.7  53.4  42.7
```

```
#> Propionate (mM)  17.3  22.3  23.9
#> [1] "Fraction of the incoming water that leaves colon is 0.12"
#> [1] "fecal microbe output rate is 17.04 g/d"
```

Next we use **plotMPG()** to see a summary plot of the SCFA concentration, the mass of each microbial group and the pH in each model compartment.

```
plotMPG(m.out)
```



Changing host diet

A standard western diet is the default setting in the **microPopGut()** but this can be changed via the input arguments, **init** (i.e. the starting mass in each of the 3 model compartments) and **inflow** which is the mass inflow each day of carbohydrate, protein and water. Here you can also specify the division of carbohydrate into resistant starch and NSP using **RS.frac**. In this example we change from 10 g/d of protein to zero and add an extra 10 g/d to the carbohydrate default of 50 g/d. Results are shown in Fig. 1.

```
m.out=microPopGut(
    numDays=1,
    time.step=1/24/60,
    transitTime=1.25,
    microbeNames=c('Bacteroides','ButyrateProducers1'),
    microbeNames.short=c('Bacteroides'='B','ButyrateProducers1'='BP1'),

    #initial mass in each compartment
    init=list(
        C=2, #carbohydrate (g)
        P=0, #protein (g)
        B=10, #biomass (g)
        Acetate=0.3606, #g
        Propionate=0.1482, #g
        Butyrate=0.1762, #g
        W=100), #water (g)

    #inflow from diet
    inflow=list(
        C=60, #carbohydrate (g/d)
        P=0, #protein (g/d)
        W=1100, #water (g/d)
        RS.frac=0.78) #fraction of C that is resistant starch (rest is NSP)

)

time=m.out$solution[[1]][,'time']
verification(m.out,start.av=0.8*max(time),fin.av=max(time))
dev.new()
plotMPG(m.out)
```

Meal composition/fluctuations

The default setting in **microPopGut()** is constant inflow but there is also the option to include a more realistic inflow which aims to represent intermittent host eating. This is specified in the **meals** input list.

The **gamma.mag** option in the **meals** list controls the magnitude of the fluctuations in meal composition and size by controlling the spread of the gamma distribution that the values are drawn from for each substrate.

```
m.out=microPopGut(
    numDays=1,
    time.step=1/60/24,
    transitTime=1.25,
    microbeNames='Bacteroides',
    microbeNames.short=c('Bacteroides'='B'),
    meals=list(
        seed=1,
        fluc.inflow=TRUE,
        fluc.subst.comp=TRUE,
        plotInflow=TRUE,
        saveInflowFig=TRUE,
        breakfast.start=7, #time (24 h clock) to start breakfast
```

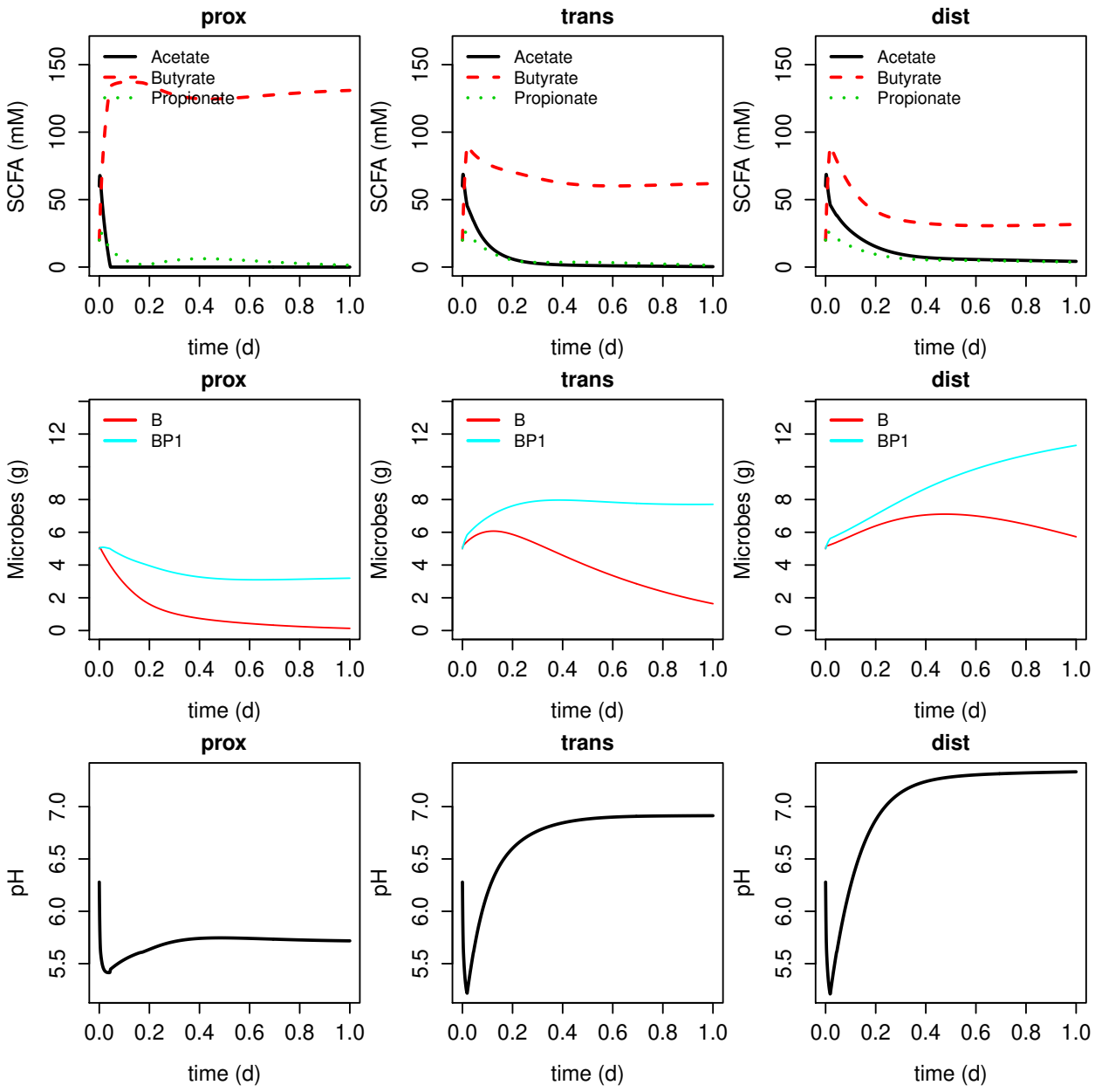


Figure 1: Summary results for a diet with no protein and 2 MFGs

```
        lunch.start=13,     #time (24 h clock) to start lunch
        dinner.start=19,    #time (24 h clock) to start dinner
        meal.duration.h=0.5, #length of time eating one meal in hours
        time.to.reach.colon.h=7, #time take to pass through stomach and small intestine in hours
        gamma.mag=1 #scaling factor to control the fluctuations in meals
    )
)
```

Figs. 2-4 shows meals over 7 days for gamma.mag equal to 0.1, 1 and 1.9 respectively for 3 arbitrary substrates A, B and C, and water. Note that after stochastically generating the composition the output is scaled so that the mean substrate values are maintained over the time period (this information is printed to screen when you run the model). Due to smoothing delays as the meals pass through the small intestine and stomach the final inflow to the colon may have slightly smaller means (this is also printed to screen).

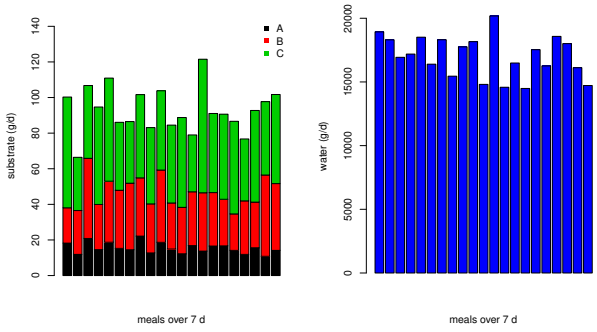


Figure 2: Meal composition for gamma.mag=0.1

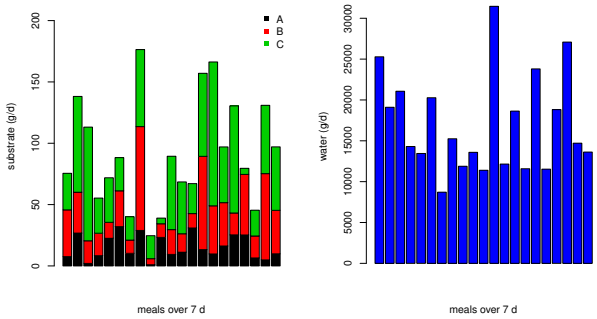


Figure 3: Meal composition for gamma.mag=1

The "meals" are then passed through an ODE model representing the stomach and small intestine - the amount of smoothing caused by this is controlled by the time.to.reach.colon.h item in the meals list. The fig. below shows the input to the colon for gamma.mag=1.9 and time.to.reach.colon.h=7 hours.

Full gut model (10 MFGs with meals for 7 days)

In this example we show the settings for simulating the full model. In this example we also change the pH tolerance and the maximum growth rate on resistant starch of the LactateProducers group. The code below

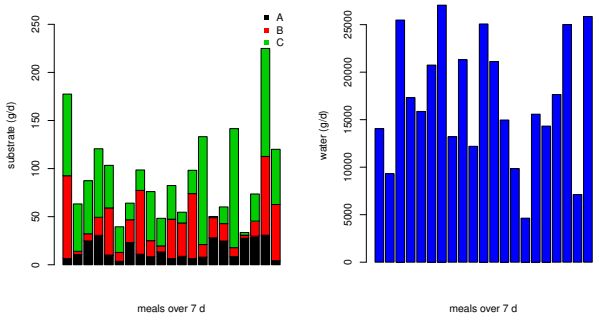


Figure 4: Meal composition for gamma.mag=1.9

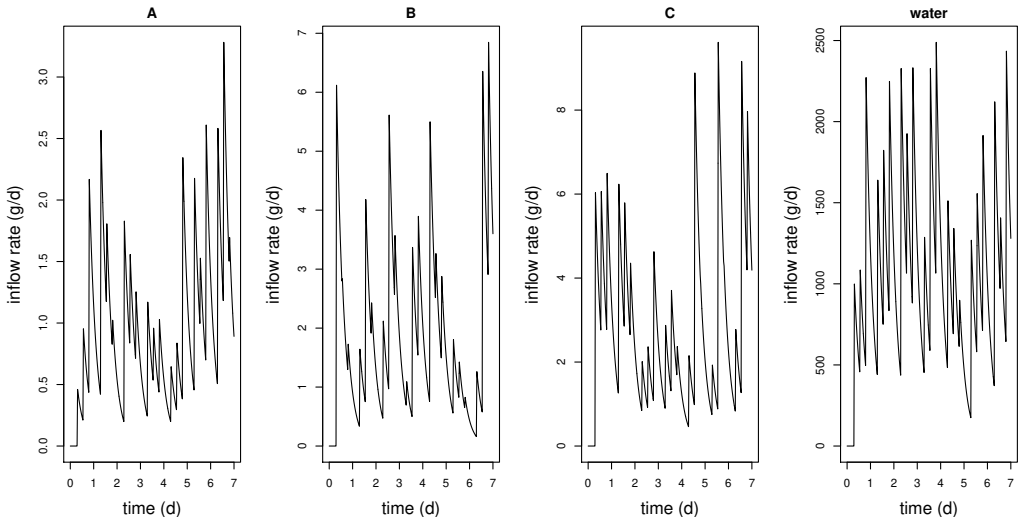


Figure 5: Substrate inflows to colon (after passing through stomach and small intestine model) for gamma.mag=1.9

takes over an hour to run so if you want to try it out it is better to change **num.days** from 5, to 1 or less, or to try fewer groups e.g. microbeNames=microbeNames[1:2].

```
microbeNames = c('Bacteroides','ButyrateProducers1',
                 'ButyrateProducers2','ButyrateProducers3',
                 'LactateProducers','PropionateProducers',
                 'Methanogens','NoButyFibreDeg',
                 'NoButyStarchDeg','Acetogens')

microbeNames.short = c('Bacteroides'='B','ButyrateProducers1'='BP1',
                   'ButyrateProducers2'='BP2','ButyrateProducers3'='BP3',
                   'LactateProducers'='LP','PropionateProducers'='PP',
                   'Methanogens'='M','NoButyFibreDeg'='NBFD',
                   'NoButyStarchDeg'='NBSD','Acetogens'='A')

#Change pH tolerance & Gmax(RS) for lactate producers
#read LPc from batch file
LactateProducers['pHcorners',2:5]=c(4.5,5.25,7.2,7.95)
LactateProducers['maxGrowthRate','RS']=7

m.out=microPopGut(
    numDays=5,
    time.step=1/24/60,
    transitTime=1.25,
    microbeNames=microbeNames,
    microbeNames.short=microbeNames.short,
    meals=list(
        seed=1,
        fluc.inflow=TRUE,
        fluc.subst.comp=TRUE,
        plotInflow=TRUE,
        saveInflowFig=TRUE,
        breakfast.start=7,
        lunch.start=13,
        dinner.start=19,
        meal.duration.h=0.5,
        time.to.reach.colon.h=7,
        gamma.mag=1
    )
)
```

See Fig. 6 for the results of **plotMPG(m.out)**

Bowel Movements

By default the model is set up for zero bowel movements per day (BMpd) i.e there is constant outlow. To turn on bowel movements, set **BMpd** in the **bowel.movements** input to either 1, 2 or 3 (see code below). The **start.BM.time** list gives the start times for 1, 2 or 3 bowel movements per day and the bowel movement duration is set using **BM.duration.h**. In our model we assume that bowel movements only affect the distal part of the colon.

```
  microbeNames = c('Bacteroides','ButyrateProducers1','ButyrateProducers2','ButyrateProducers3','Lactat

microbeNames.short = c('Bacteroides'='B','ButyrateProducers1'='BP1','ButyrateProducers2'='BP2','Butyrat
```

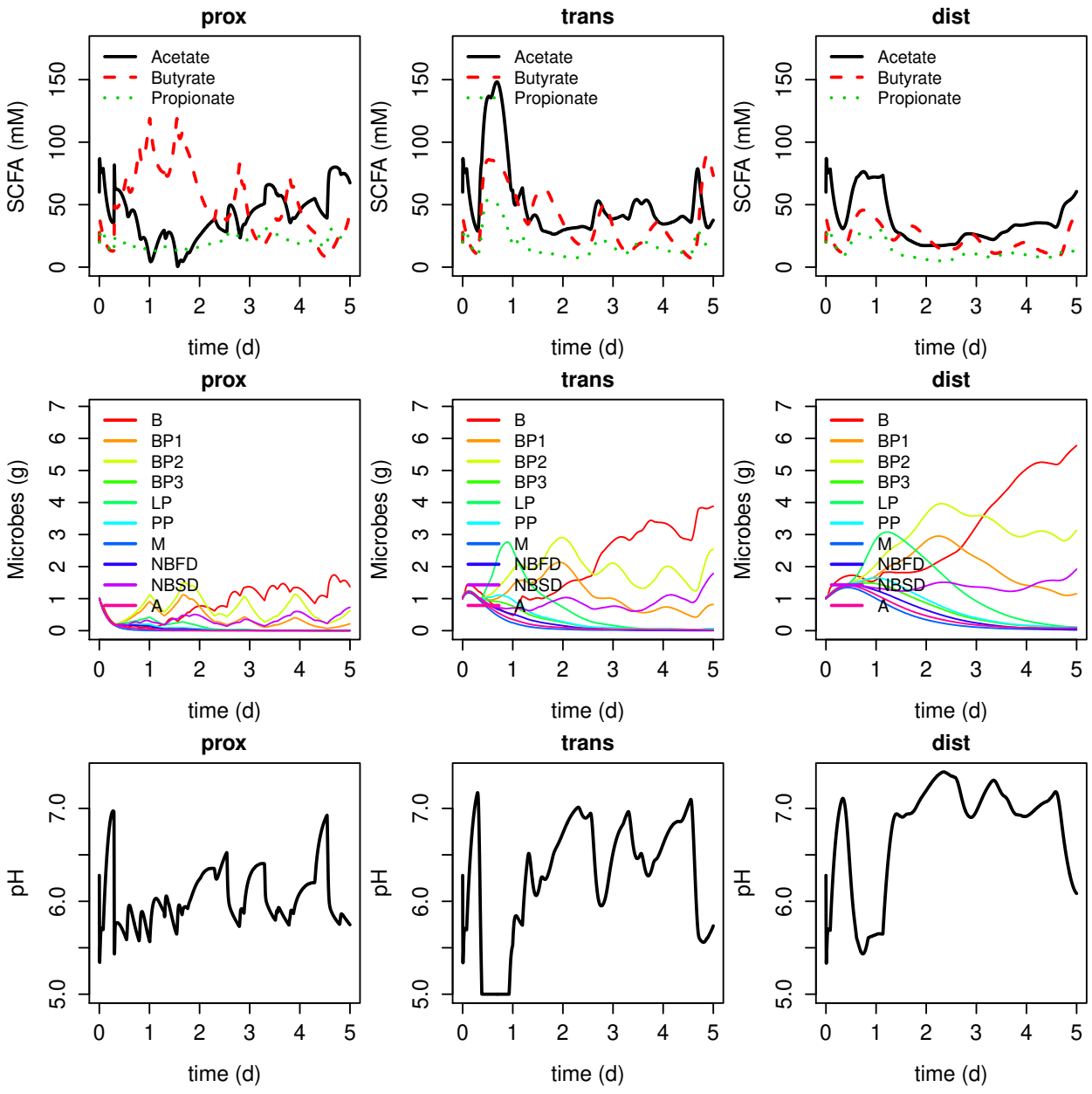


Figure 6: Summary results for 10 MFGs over 5 days

```
sim.time.h=48 #time to simulate in hours

m.out=microPopGut(
    numDays=sim.time.h/24,
    time.step=1/24/60,
    transitTime=1.25,
    microbeNames=microbeNames[1:2],
    microbeNames.short=microbeNames.short[1:2],
    bowel.movements = list(
        BM.duration.h = 15/60,
        frac.distal.emptied = 0.9,
        BMpd = 3,
        start.BM.time = list(7, c(7, 19), c(7, 15, 21)))
)

time=m.out$solution[[1]][,'time']
verification(m.out,start.av=0.8*max(time),fin.av=max(time))
dev.new()
plotMPG(m.out)
```

Multiple strains per group

By default there is only one strain in each MFG but using the functionality of microPop (on which microPopGut is based) we can add multiple strains in each group. In versions of microPop from 1.6 onwards we can have different numbers of strains in each group. The parameters for each strain in a group are drawn randomly from a given range for that particular group. The particulars of this are controlled by the **strainOptions** input argument.

```
m.out=microPopGut(
    numDays=2,
    time.step=1/24/60,
    transitTime=1.25,
    microbeNames=c('Bacteroides','ButyrateProducers1'),
    microbeNames.short=('Bacteroides'='B','ButyrateProducers1'='BP1'),
    numStrains=c('Bacteroides'=3,'ButyrateProducers1'=2),
    strainOptions = list(
        randomParams = c("halfSat", "yield", "maxGrowthRate",
        "pHtrait"),
        seed = 3,
        distribution = "uniform",
        percentTraitRange = 10,
        maxPHshift = 0.1,
        applyTradeOffs = FALSE,
        tradeOffParams = NULL,
        paramsSpecified = FALSE,
        paramDataName = NULL)
)

time=m.out$solution[[1]][,'time']
```

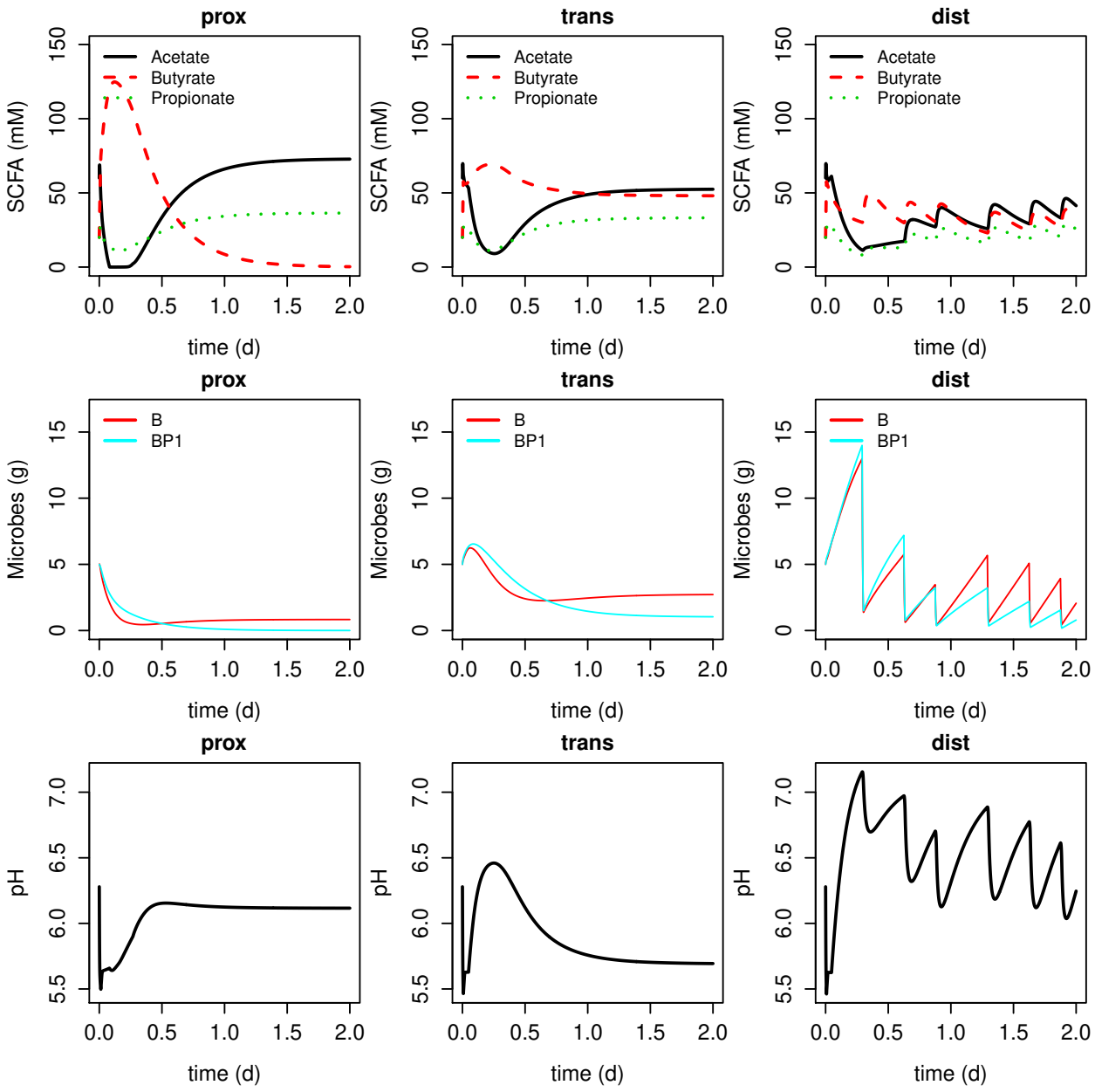


Figure 7: Summary results for 2 MFGs over 2 days with 3 bowel movements per day

```
verification(m.out,start.av=0.8*max(time),fin.av=max(time))
plotMPG(m.out)
```

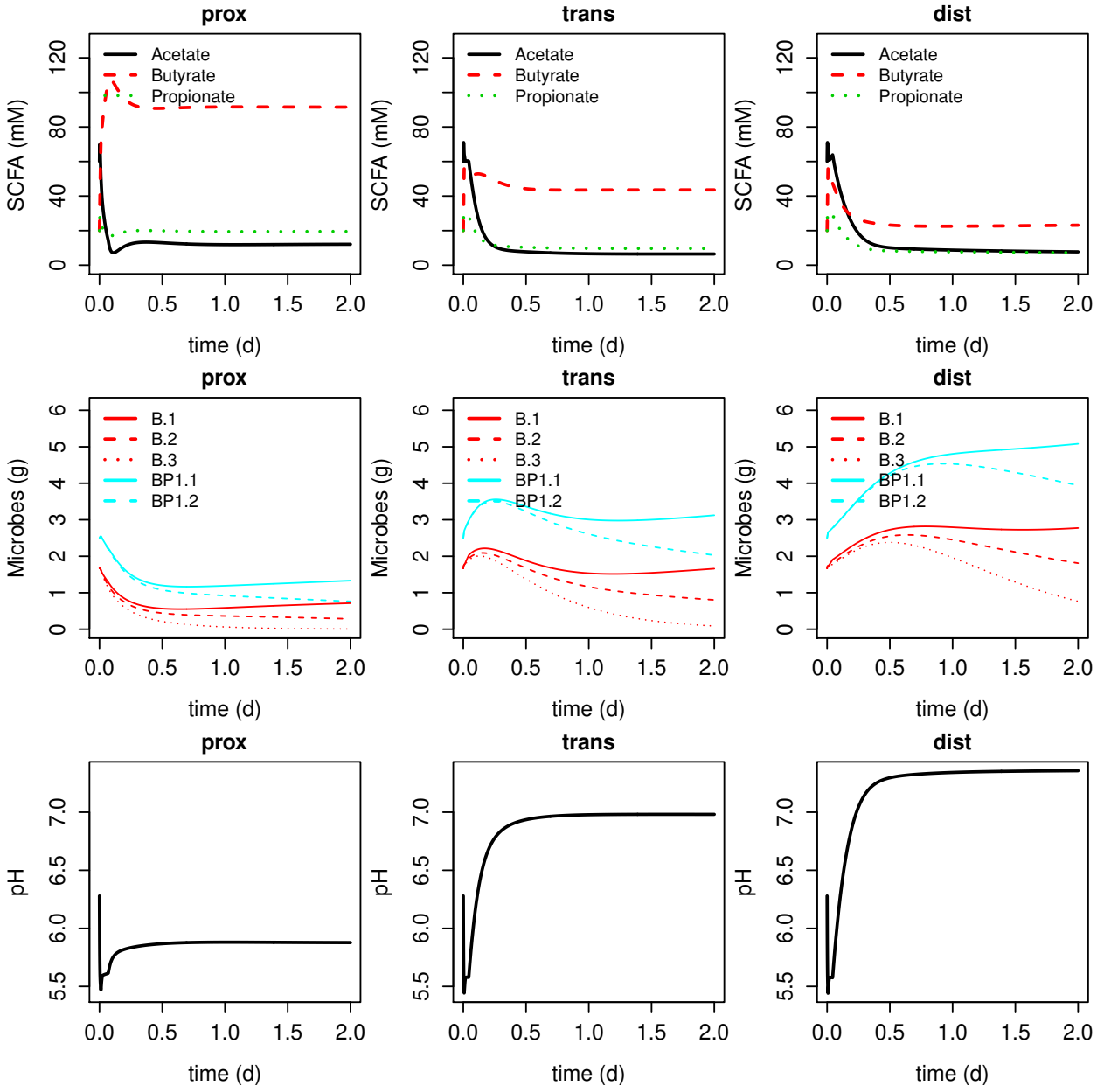


Figure 8: Summary results for 2 MFGs: Bacteroides with 3 strains and ButyrateProducers1 with 2 strains

Troubleshooting

Failure of ODE solver

If you get warnings about time steps then this mean the ODE solver is failing - generally because the problem is too stiff i.e. there are rapid changes in time. Ways to deal with this are to make your time step

(microPopGut input argument) smaller, to make gamma.mag smaller (nearer to zero) which will decrease the size of the meal fluctuations, or to alter the tolerances in the ODE solver (see ode.options list of microPopGut input arguments). If you are running the system with only one group, adding more groups often makes the system more stable as this slows down the growth of each individual group due to resource competition.