SoupX

Kimberly Olney, PhD

04/14/2024

Contents

SoupX to remove ambinet RNA

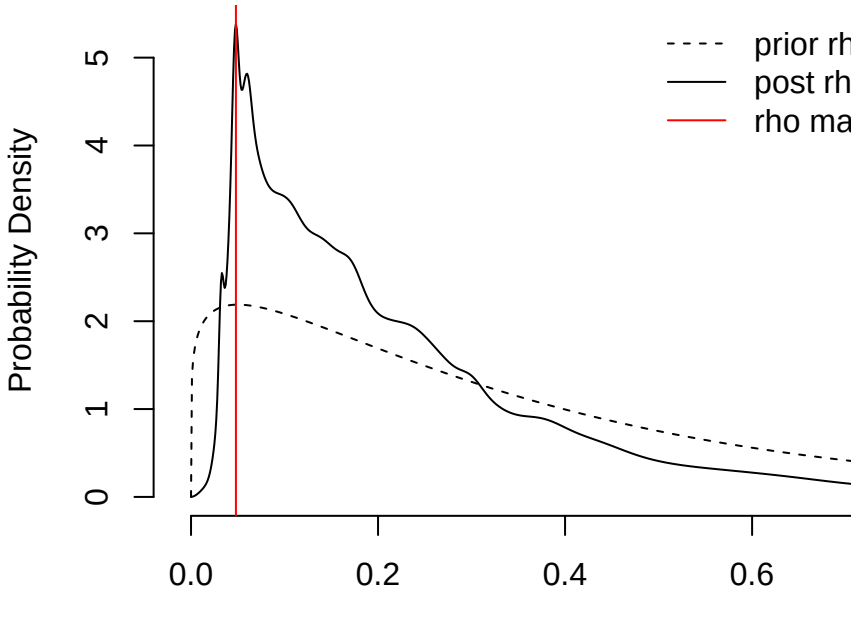
The vignette
Estimate the composition of this soup, i.e. what fraction of UMIs are derived from the soup in each droplet and produce a corrected count table with the soup based expression removed.

The method to do this consists of three parts:

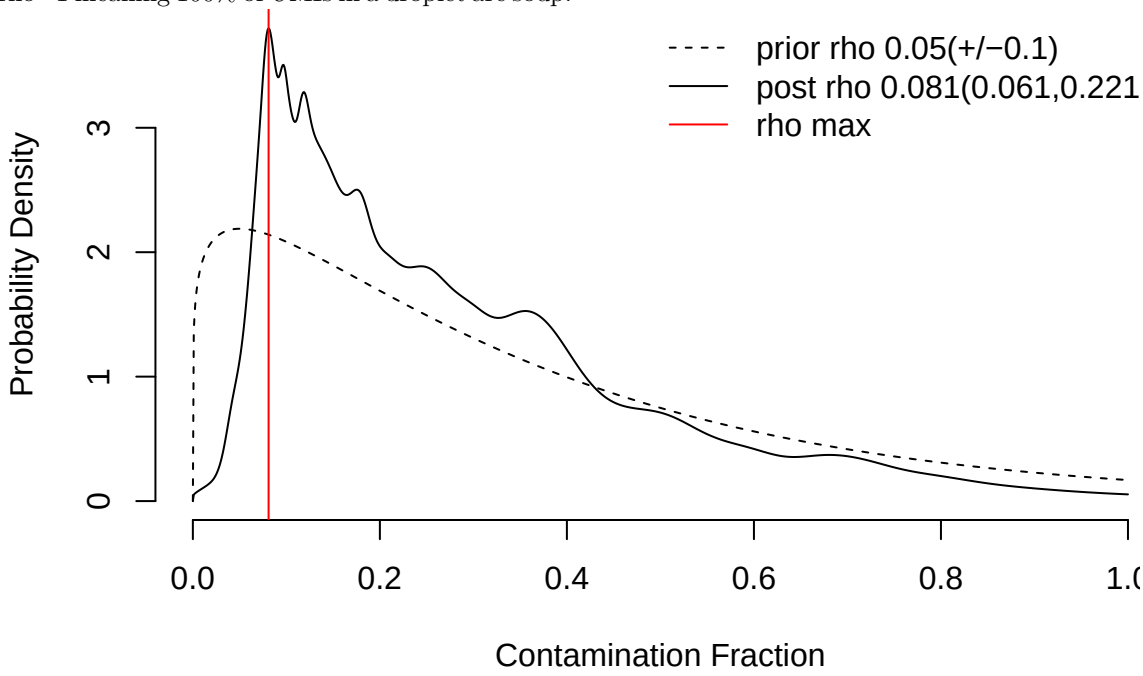
1. Calculate the profile of the soup.
2. Estimate the cell specific contamination fraction.
3. Infer a corrected expression matrix.

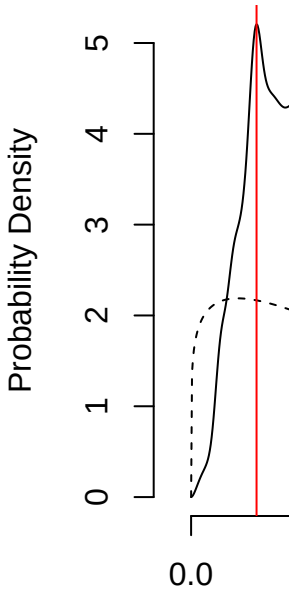
```
knitr::opts_chunk$set(root.dir = ".", echo=FALSE, message=FALSE, warning=FALSE, background=TRUE)
```

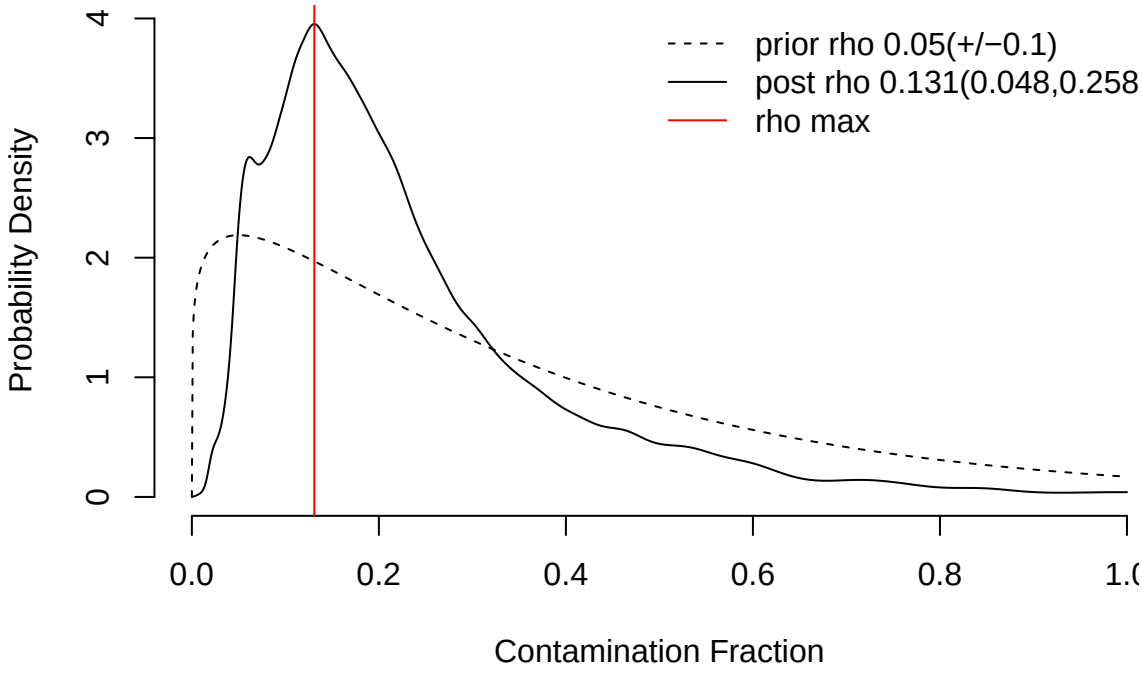
The contamination fraction is parametrised as rho in the code, with rho=0 meaning no contamination and

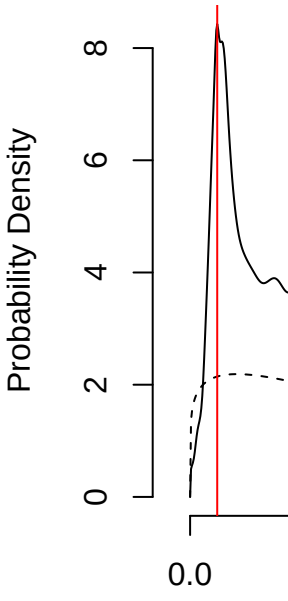


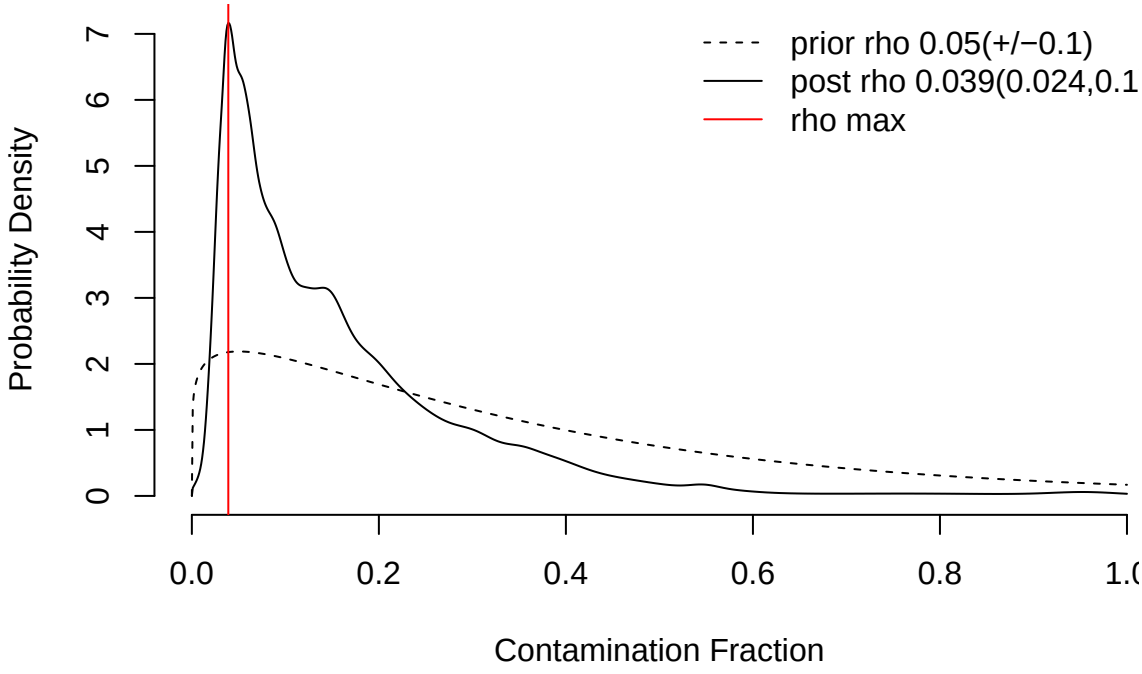
rho=1 meaning 100% of UMIs in a droplet are soup.

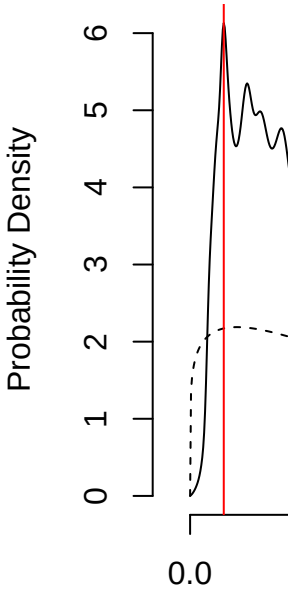


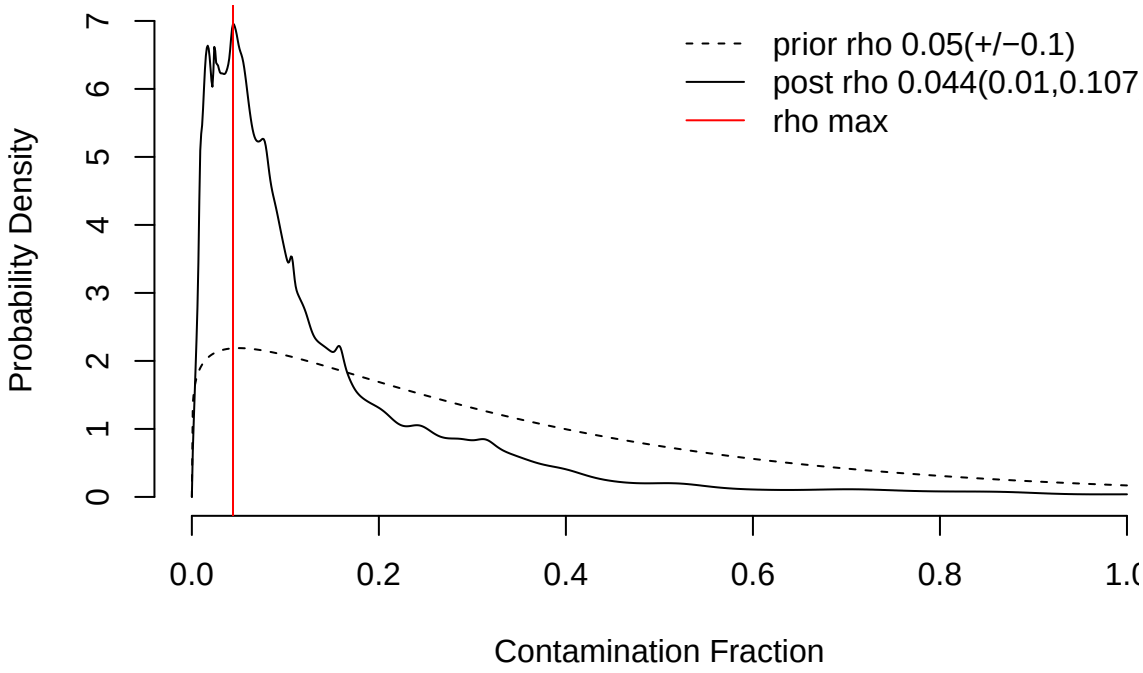


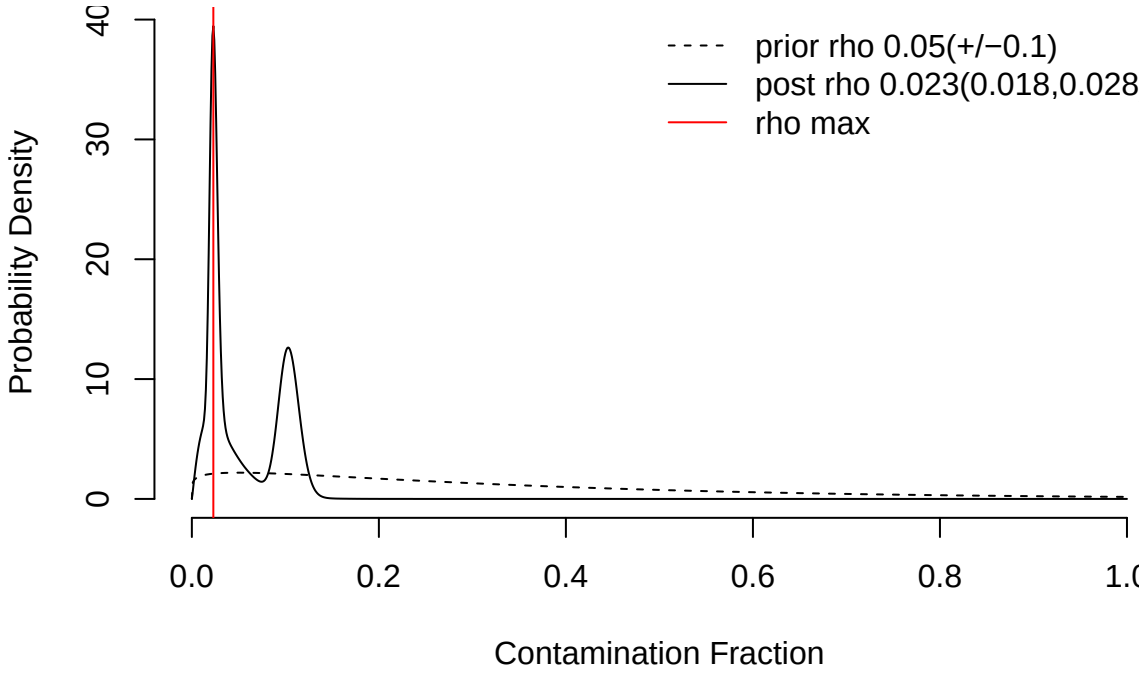

prior rho 0.05(+/−0.1)
post rho 0.131(0.048,0.258
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4

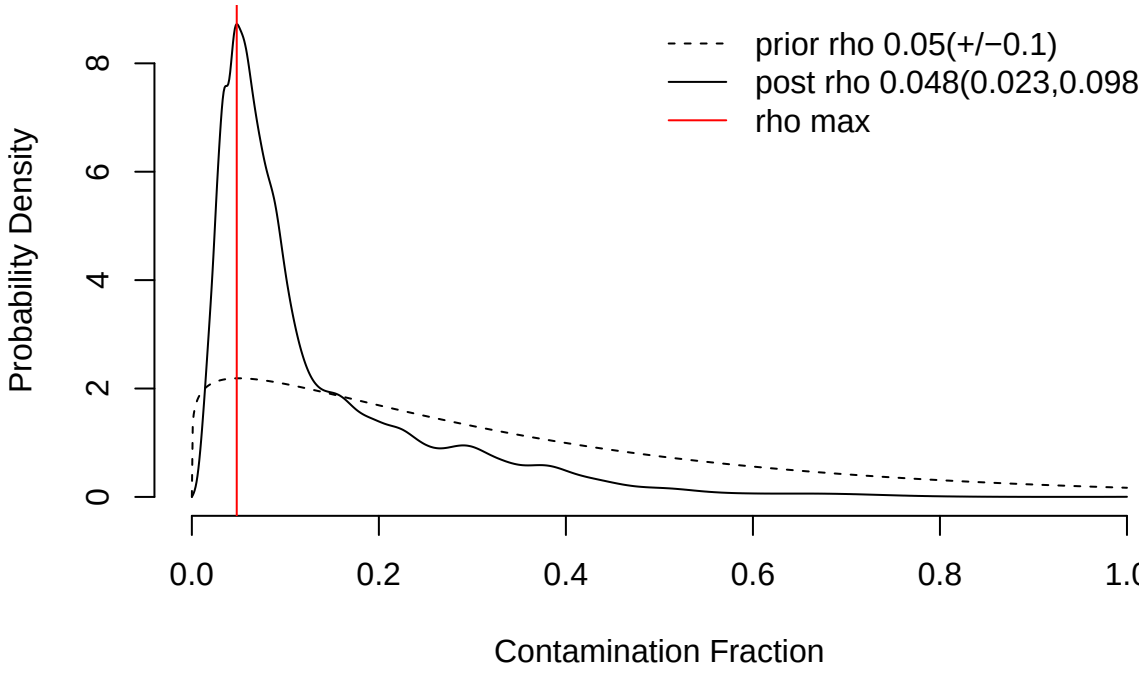

Probability Density
0.0
0
2
4
6
8

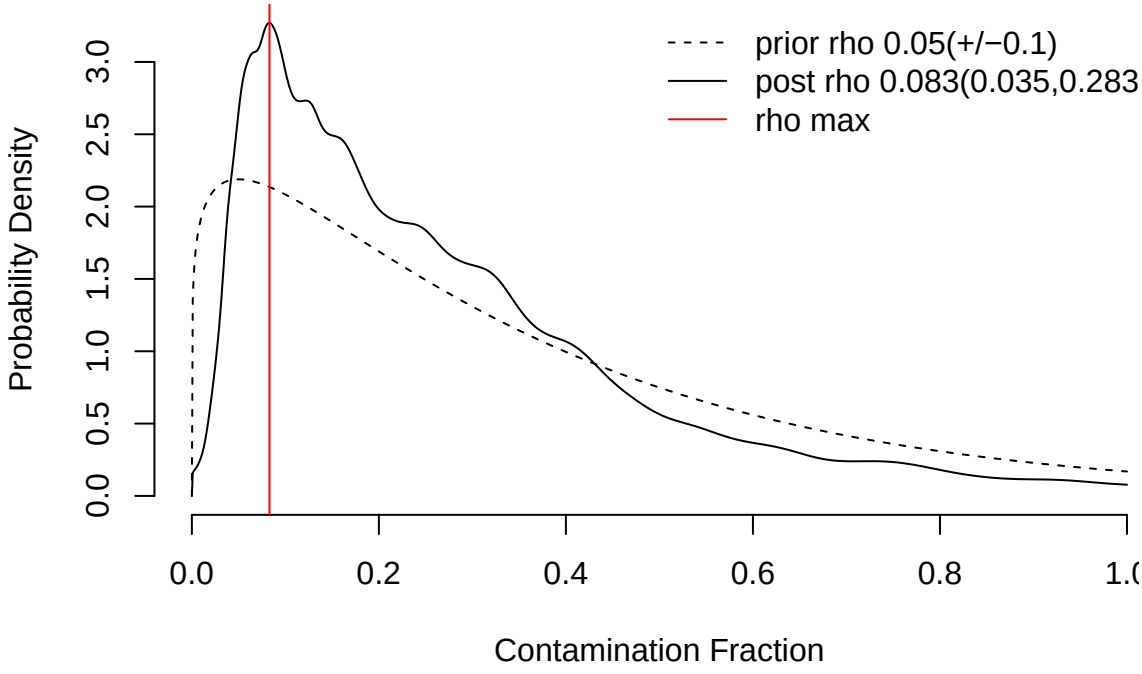

prior rho 0.05(+/−0.1)
post rho 0.039(0.024,0.1
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
5
6
7

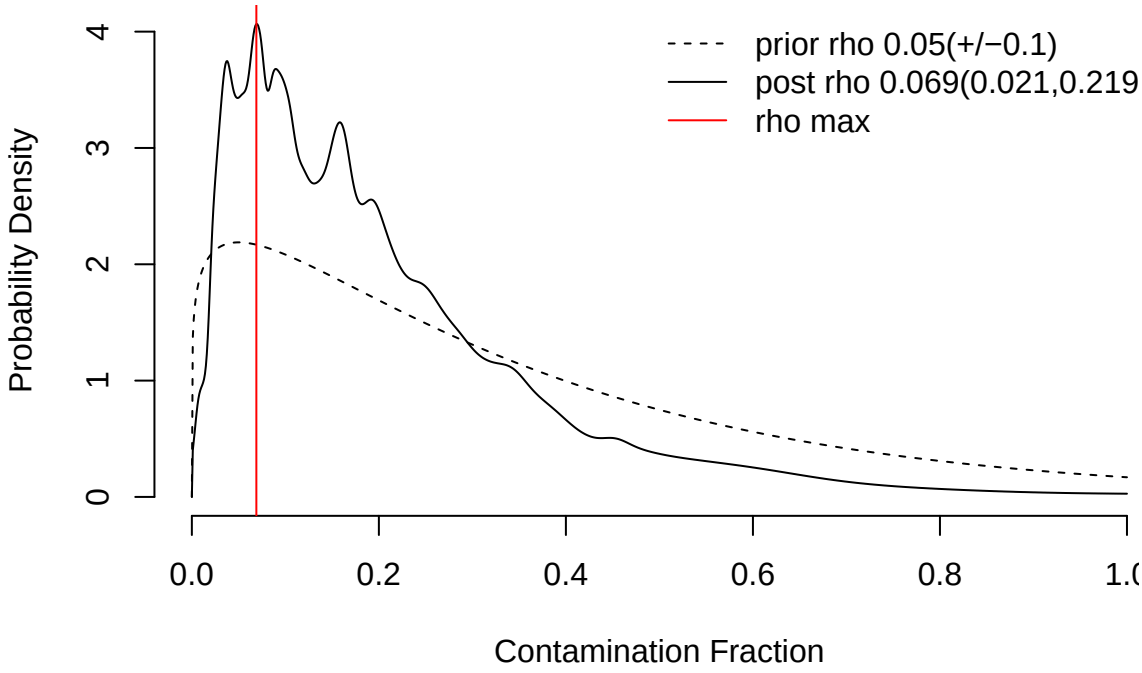

Probability Density
0.0
0
1
2
3
4
5
6

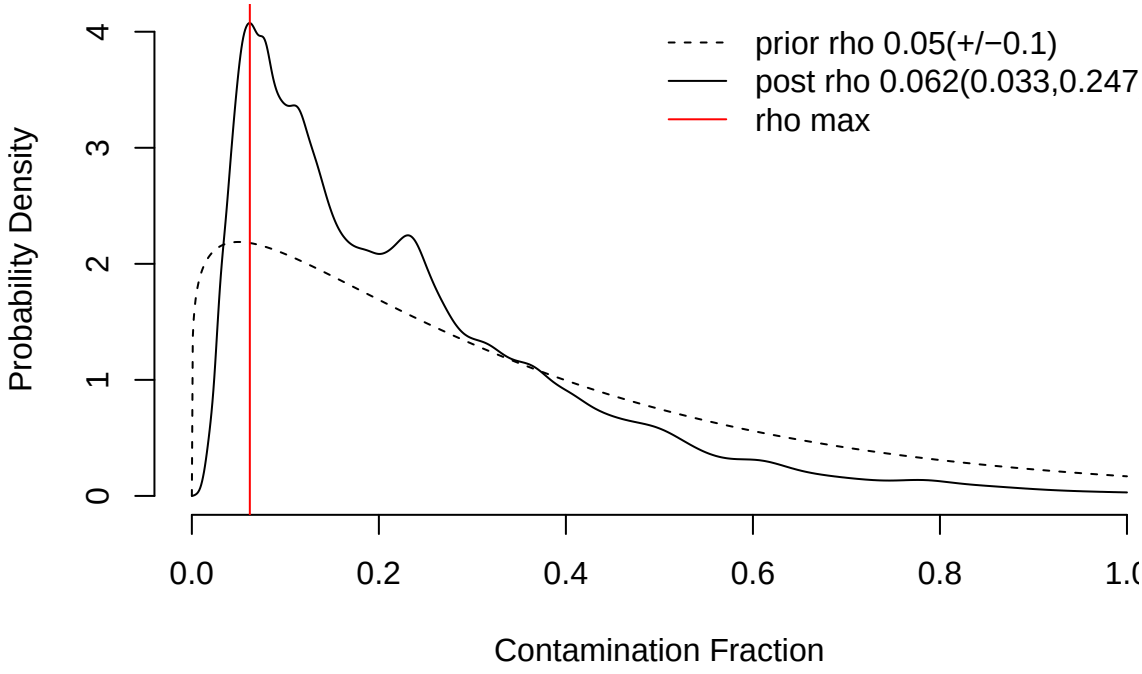

prior rho 0.05(+/−0.1)
post rho 0.044(0.01,0.107
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8

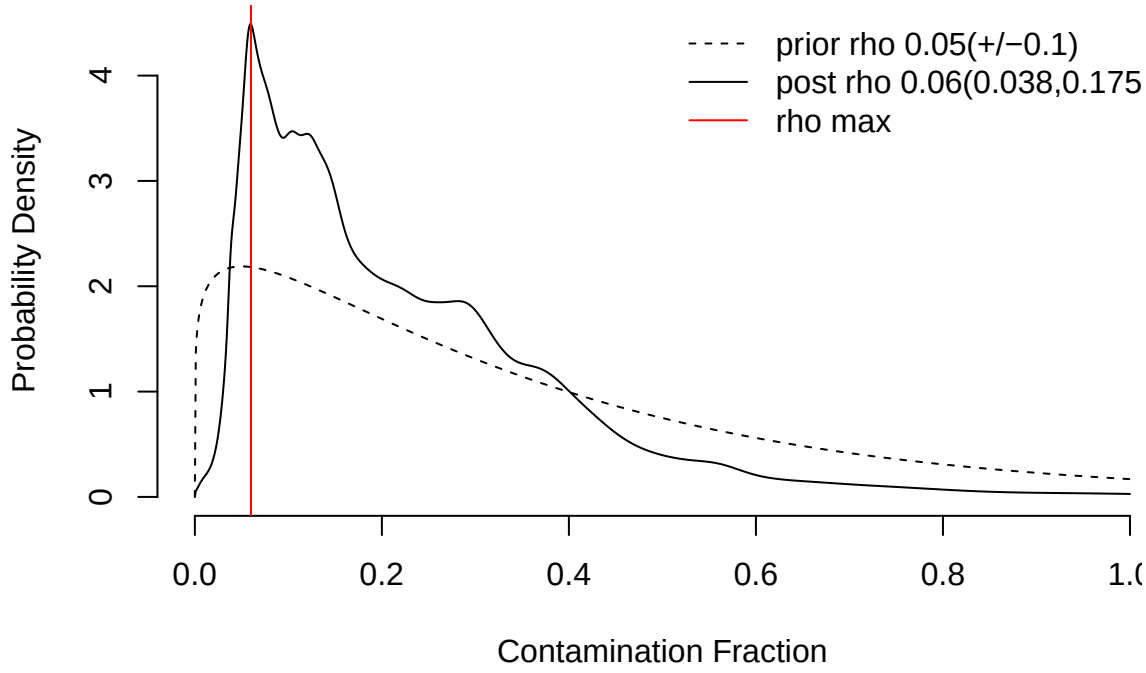

prior rho 0.05(+/−0.1)
post rho 0.023(0.018,0.028
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8


prior rho 0.05(+/−0.1)
post rho 0.048(0.023,0.098
rho max
Probability Density
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
Contamination Fraction


prior rho 0.05(+/−0.1)
post rho 0.083(0.035,0.283
rho max
Probability Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.0
0.2
0.4
0.6
0.8
Contamination Fraction


prior rho 0.05(+/−0.1)
post rho 0.069(0.021,0.219
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4


prior rho 0.05(+/−0.1)
post rho 0.062(0.033,0.247
rho max
Probability Density
Contamination Fraction
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4




```
##    sampleID rhoEst
## 1      P_69  0.048
## 2       B_7  0.081
## 3      P_65  0.070
## 4       B_6  0.131
## 5      P_52  0.029
## 6      P_62  0.039
## 7      P_67  0.036
## 8       J_3  0.044
## 9       J_4  0.079
## 10     P_60  0.023
## 11     P_71  0.033
## 12     B_10  0.048
## 13      B_8  0.018
## 14      B_9  0.083
## 15     P_68  0.032
## 16     P_54  0.069
## 17     P_53  0.036
## 18     P_64  0.062
## 19      J_2  0.034
## 20      J_1  0.060
```