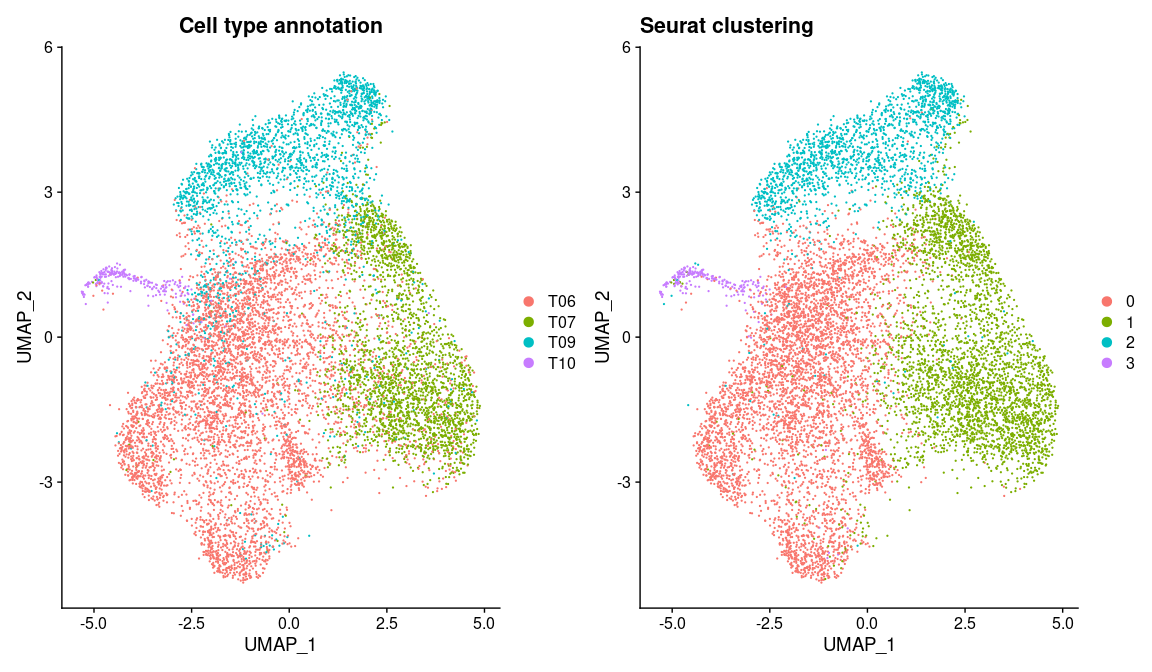
Figure Draft - Three plots version

Xinyi Lin

3/12/2022

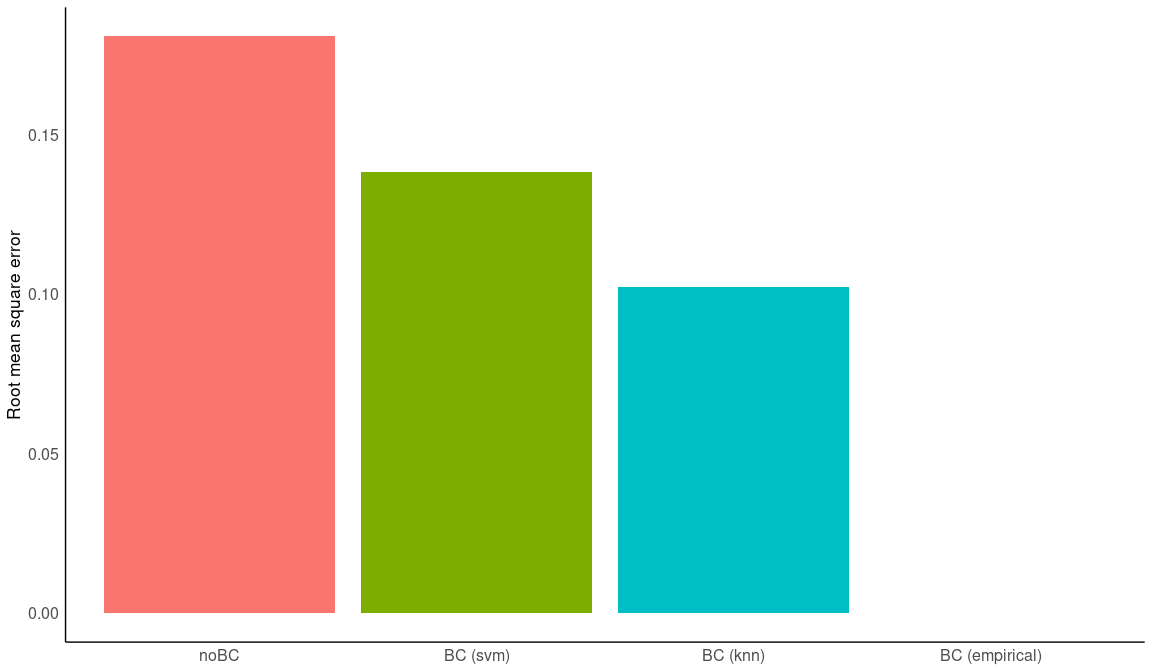
## Bias Correction validation

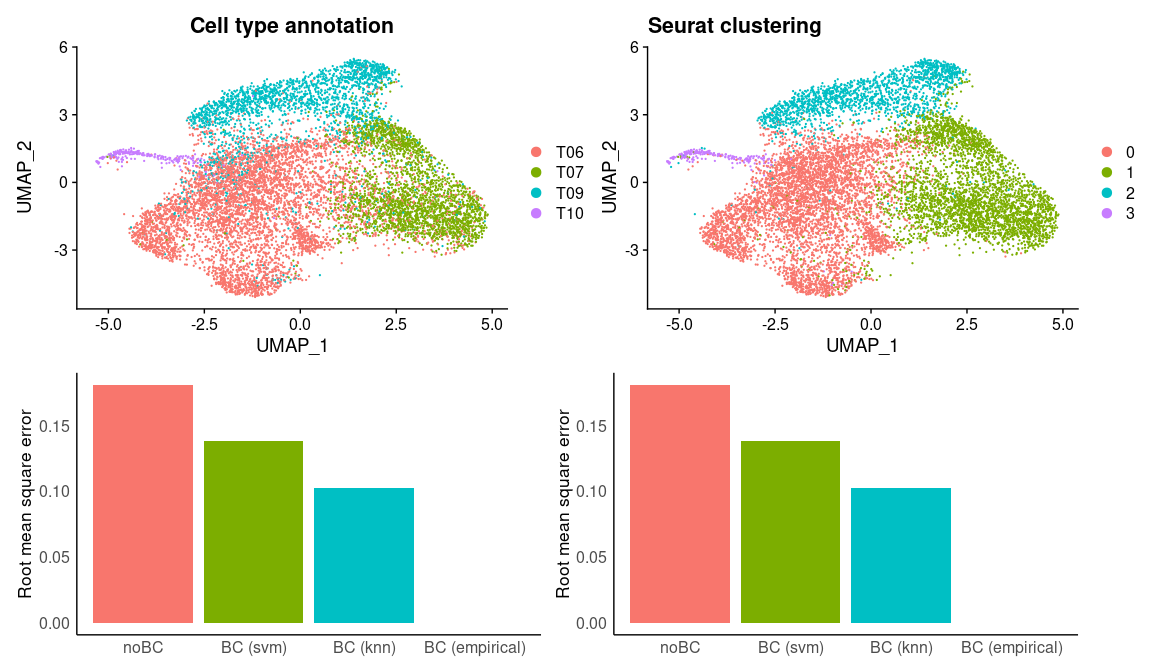
Supplementary figures



## T06 T07 T09 T10  
## true\_proportion 0.5176000 0.2473000 0.2124000 0.022700000  
## cur\_proportion 0.4757000 0.3378000 0.1654000 0.021100000  
## bs\_true\_proportion 0.5175566 0.2473243 0.2124182 0.022700893  
## bs\_knn\_proportion 0.5359492 0.2801164 0.1633053 0.020629093  
## bs\_svm\_proportion 0.5544828 0.2795524 0.1575926 0.008372169

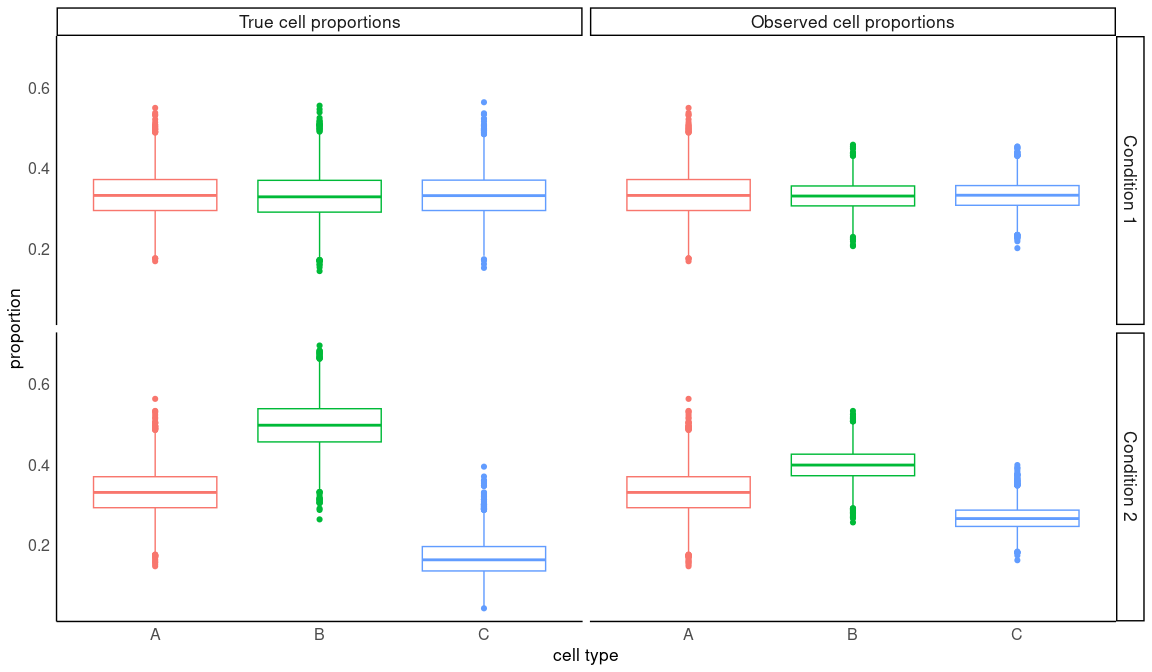
## .  
## rmse\_cur 1.810000e-01  
## rmse\_true\_bs 8.672239e-05  
## rmse\_svm\_bs 1.382705e-01  
## rmse\_knn\_bs 1.023313e-01

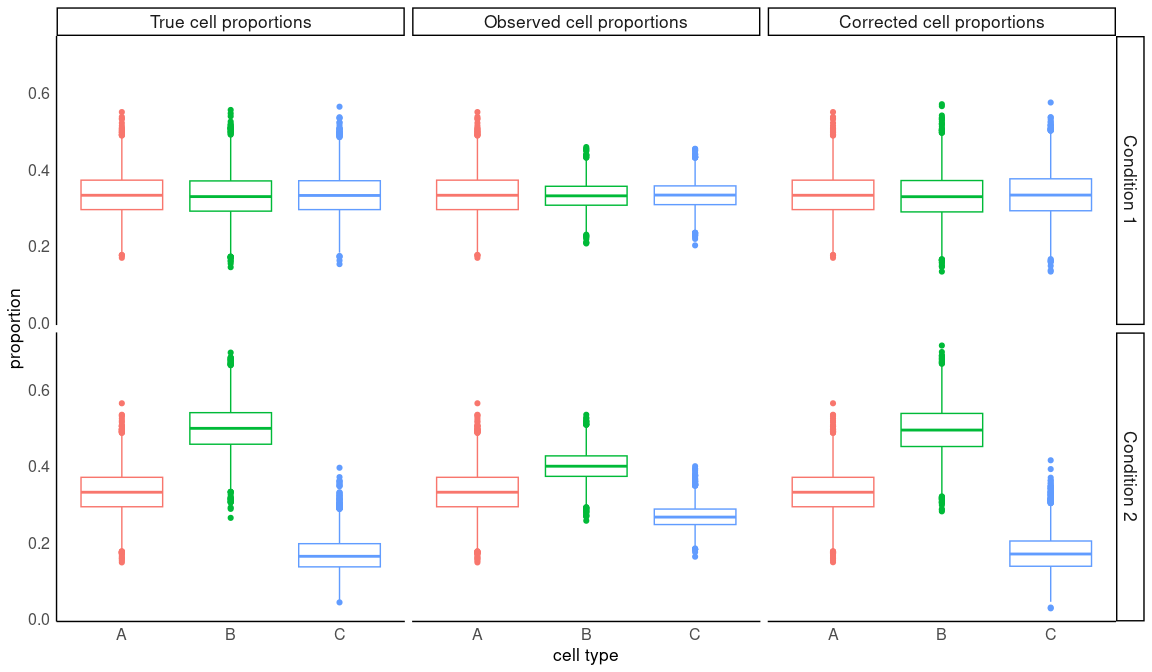




## Saving 12 x 7 in image

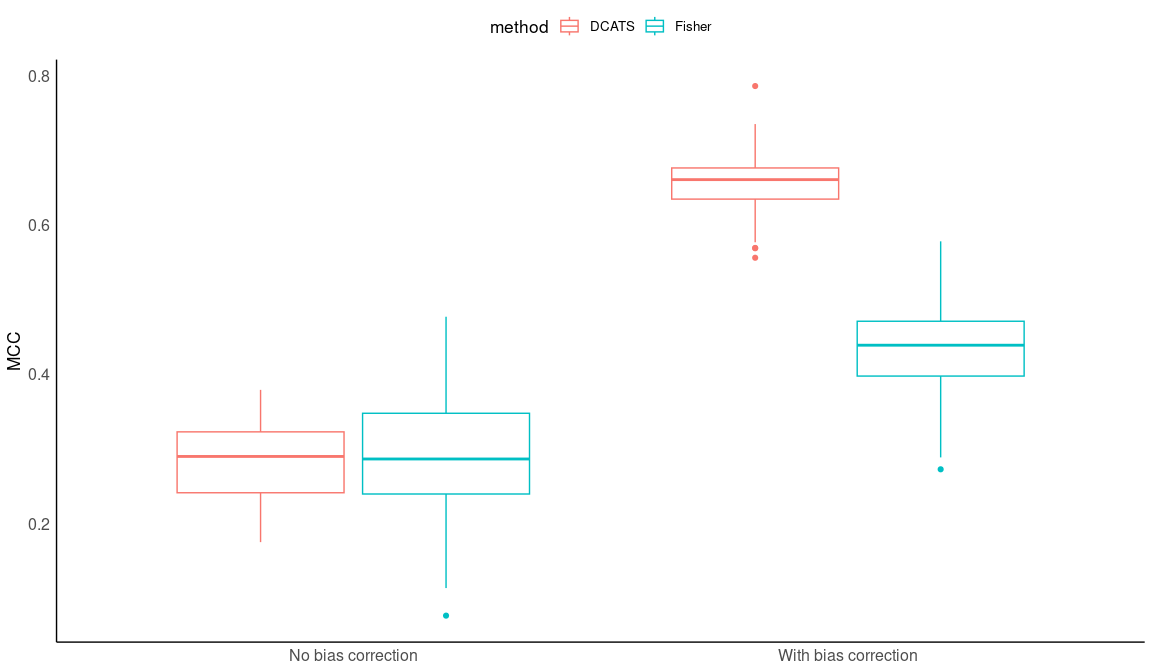
### Figure 1 - B





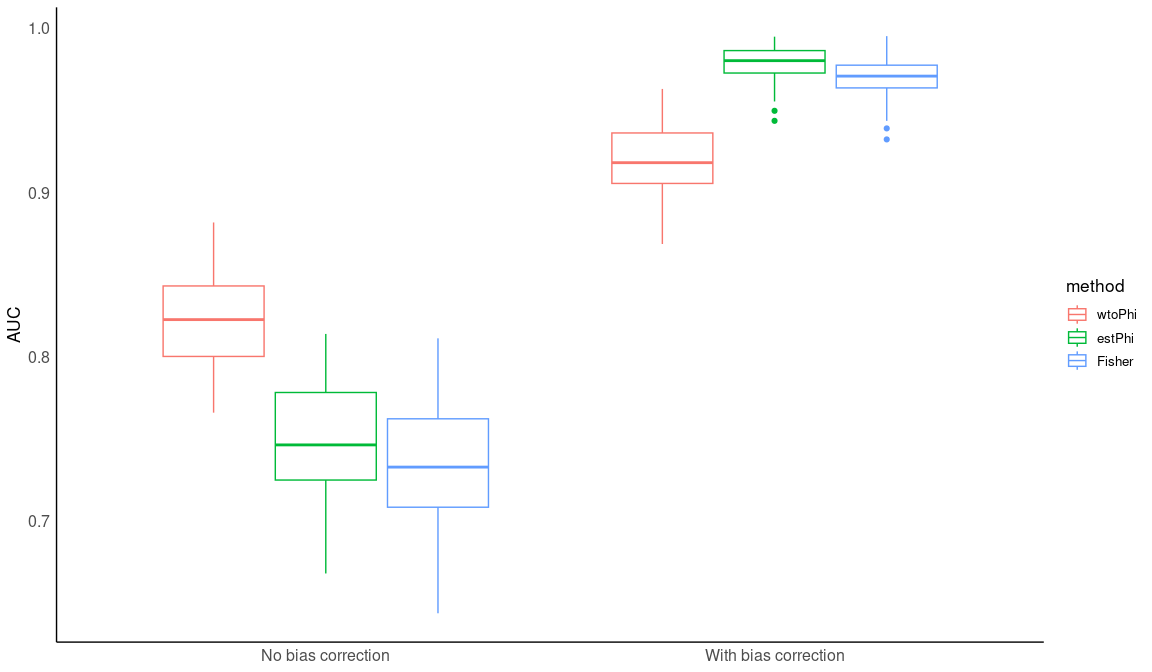
## Saving 12 x 7 in image

### Figure 2 - C

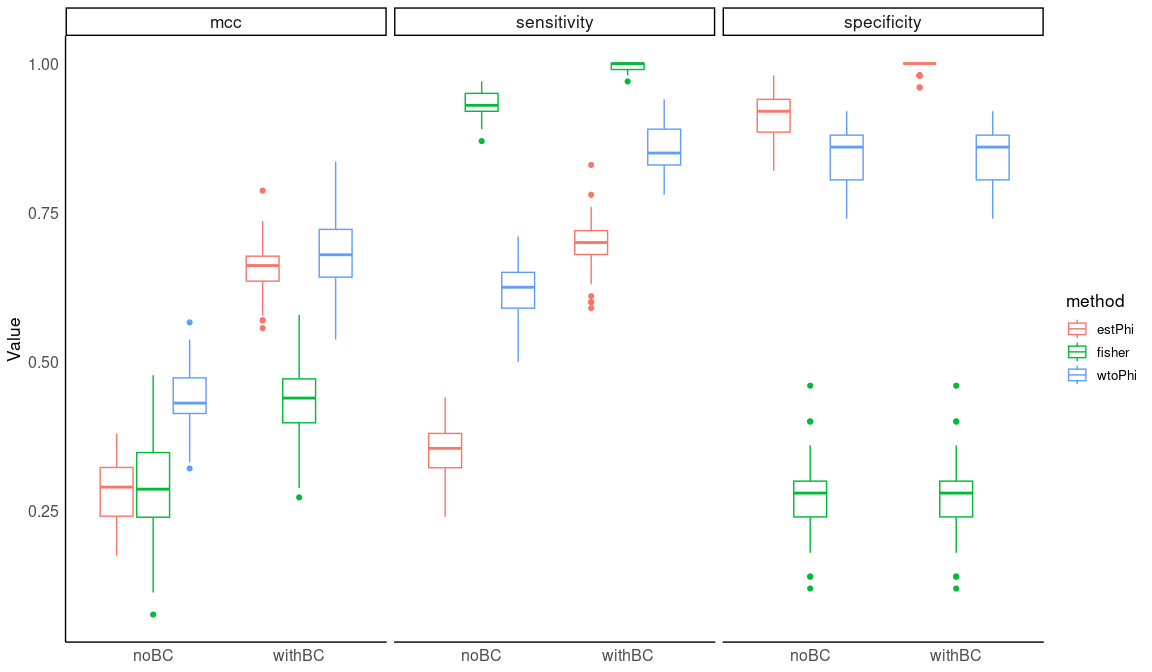


## Saving 12 x 7 in image

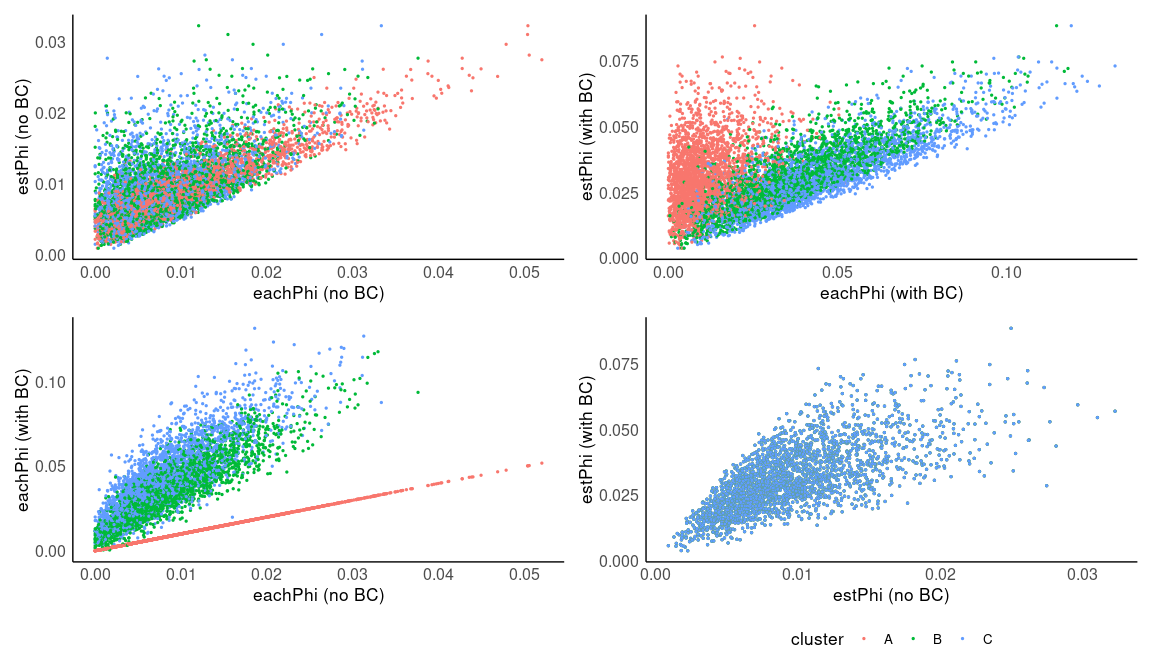
Simulation 1: Supplementary Plots

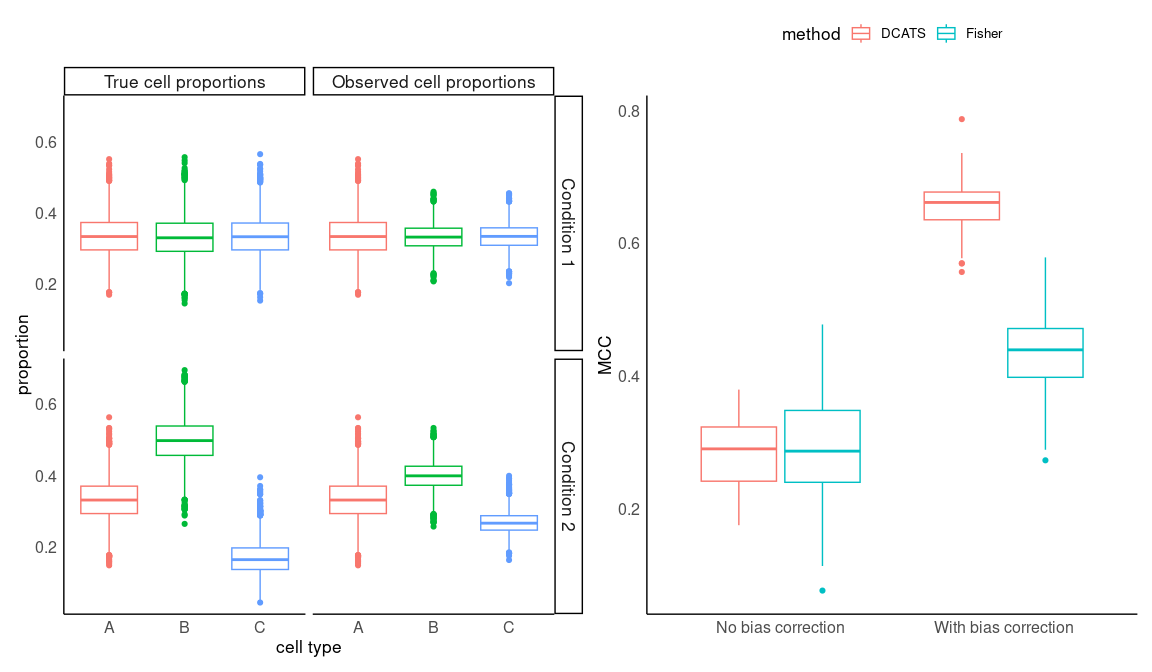


## Saving 12 x 7 in image



## Saving 12 x 7 in image





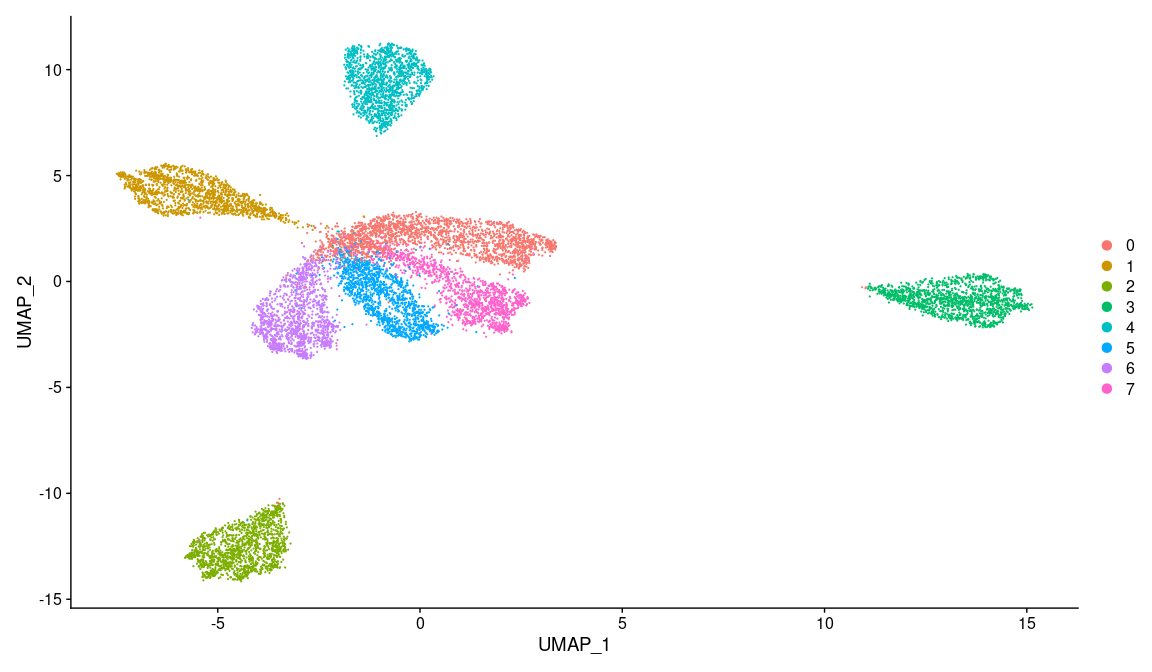
## Figure 2

### Figure 2-A

The Distribution of cell pool

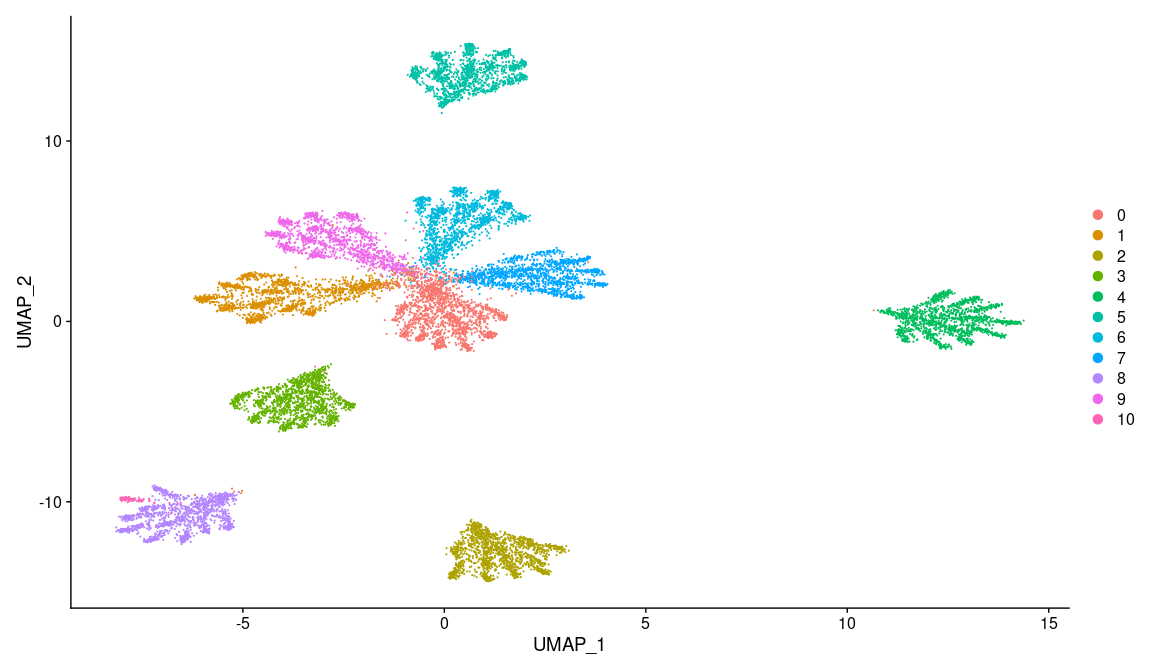
## Centering and scaling data matrix

## PC\_ 1   
## Positive: Gene862, Gene593, Gene473, Gene94, Gene511, Gene508, Gene724, Gene830, Gene623, Gene584   
## Gene662, Gene396, Gene439, Gene214, Gene524, Gene714, Gene285, Gene485, Gene395, Gene19   
## Gene840, Gene191, Gene676, Gene88, Gene418, Gene111, Gene569, Gene713, Gene179, Gene434   
## Negative: Gene754, Gene241, Gene650, Gene117, Gene705, Gene244, Gene665, Gene971, Gene864, Gene443   
## Gene734, Gene540, Gene908, Gene280, Gene378, Gene369, Gene456, Gene136, Gene26, Gene401   
## Gene159, Gene267, Gene594, Gene965, Gene559, Gene174, Gene322, Gene703, Gene148, Gene935   
## PC\_ 2   
## Positive: Gene273, Gene88, Gene327, Gene713, Gene173, Gene981, Gene397, Gene60, Gene349, Gene724   
## Gene356, Gene511, Gene24, Gene150, Gene14, Gene624, Gene373, Gene418, Gene161, Gene819   
## Gene57, Gene126, Gene817, Gene142, Gene214, Gene763, Gene485, Gene191, Gene46, Gene345   
## Negative: Gene171, Gene260, Gene138, Gene36, Gene189, Gene444, Gene68, Gene743, Gene93, Gene145   
## Gene806, Gene883, Gene809, Gene799, Gene726, Gene507, Gene170, Gene672, Gene896, Gene83   
## Gene147, Gene117, Gene53, Gene715, Gene867, Gene184, Gene673, Gene778, Gene125, Gene810   
## PC\_ 3   
## Positive: Gene655, Gene94, Gene171, Gene3, Gene241, Gene479, Gene624, Gene179, Gene833, Gene521   
## Gene883, Gene362, Gene513, Gene862, Gene951, Gene593, Gene830, Gene524, Gene391, Gene714   
## Gene797, Gene402, Gene419, Gene902, Gene600, Gene118, Gene584, Gene569, Gene865, Gene235   
## Negative: Gene754, Gene679, Gene458, Gene155, Gene27, Gene531, Gene220, Gene309, Gene208, Gene731   
## Gene298, Gene300, Gene614, Gene953, Gene782, Gene819, Gene428, Gene510, Gene566, Gene420   
## Gene765, Gene791, Gene730, Gene810, Gene255, Gene565, Gene393, Gene170, Gene68, Gene715   
## PC\_ 4   
## Positive: Gene521, Gene475, Gene525, Gene877, Gene888, Gene397, Gene431, Gene655, Gene651, Gene219   
## Gene85, Gene171, Gene235, Gene575, Gene970, Gene34, Gene860, Gene552, Gene201, Gene402   
## Gene780, Gene236, Gene839, Gene557, Gene301, Gene955, Gene36, Gene736, Gene349, Gene638   
## Negative: Gene214, Gene593, Gene862, Gene623, Gene662, Gene27, Gene569, Gene138, Gene473, Gene99   
## Gene511, Gene666, Gene395, Gene560, Gene676, Gene444, Gene809, Gene743, Gene510, Gene713   
## Gene425, Gene132, Gene398, Gene19, Gene723, Gene798, Gene787, Gene439, Gene94, Gene640   
## PC\_ 5   
## Positive: Gene166, Gene624, Gene179, Gene557, Gene970, Gene552, Gene993, Gene909, Gene663, Gene301   
## Gene623, Gene575, Gene799, Gene586, Gene797, Gene508, Gene548, Gene402, Gene877, Gene754   
## Gene830, Gene521, Gene235, Gene234, Gene213, Gene374, Gene404, Gene633, Gene19, Gene736   
## Negative: Gene14, Gene425, Gene362, Gene619, Gene444, Gene713, Gene439, Gene511, Gene70, Gene241   
## Gene6, Gene546, Gene184, Gene31, Gene946, Gene216, Gene781, Gene169, Gene593, Gene676   
## Gene651, Gene142, Gene775, Gene914, Gene67, Gene273, Gene351, Gene714, Gene640, Gene340



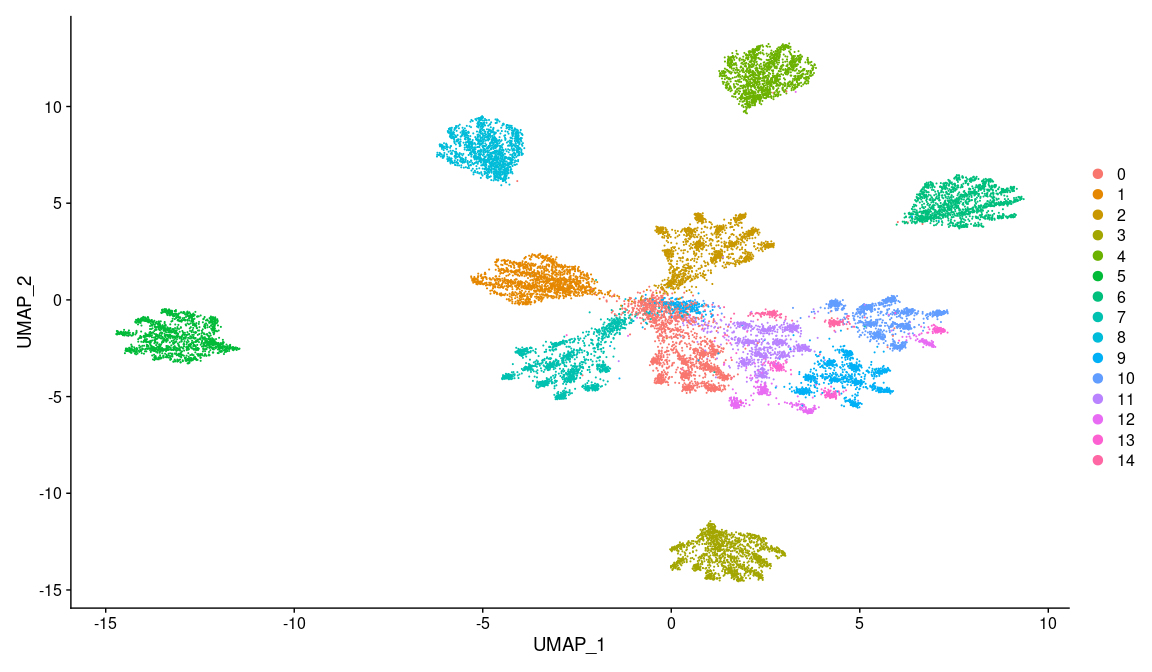
## Centering and scaling data matrix

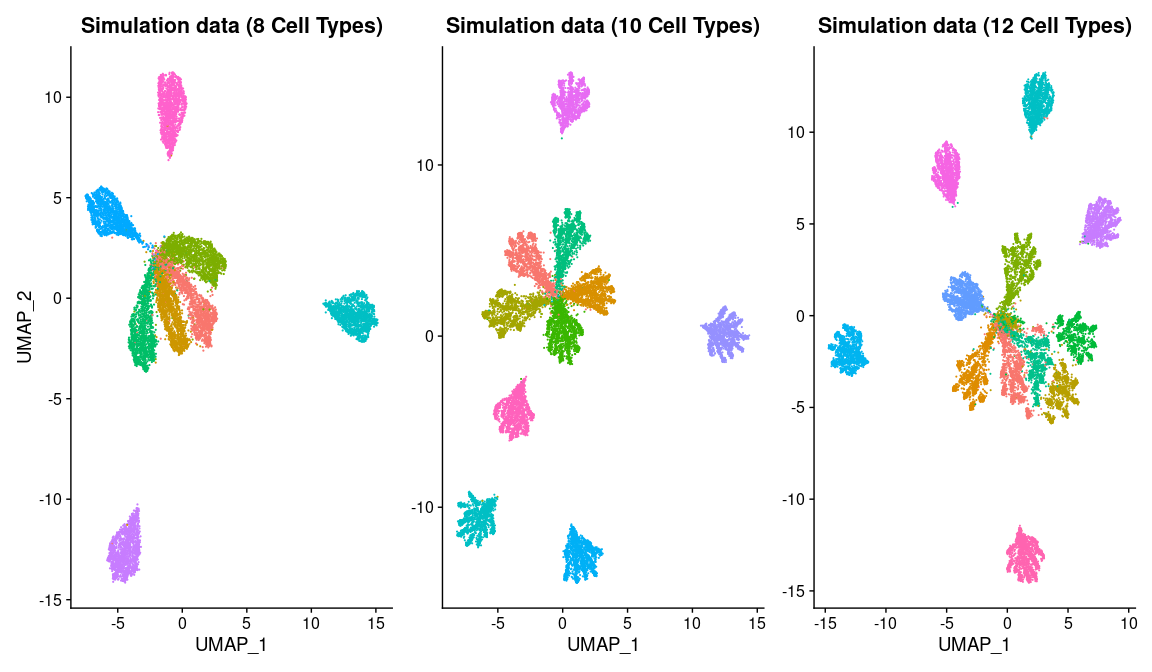
## PC\_ 1   
## Positive: Gene623, Gene412, Gene469, Gene313, Gene433, Gene400, Gene198, Gene589, Gene464, Gene973   
## Gene151, Gene251, Gene654, Gene340, Gene317, Gene56, Gene102, Gene524, Gene392, Gene872   
## Gene344, Gene766, Gene748, Gene980, Gene303, Gene873, Gene811, Gene652, Gene155, Gene485   
## Negative: Gene605, Gene650, Gene746, Gene143, Gene110, Gene153, Gene409, Gene620, Gene714, Gene460   
## Gene755, Gene615, Gene475, Gene370, Gene936, Gene65, Gene950, Gene335, Gene3, Gene921   
## Gene164, Gene328, Gene195, Gene985, Gene821, Gene818, Gene849, Gene793, Gene859, Gene140   
## PC\_ 2   
## Positive: Gene413, Gene418, Gene579, Gene479, Gene373, Gene605, Gene513, Gene190, Gene559, Gene388   
## Gene601, Gene560, Gene660, Gene816, Gene522, Gene24, Gene203, Gene542, Gene746, Gene912   
## Gene566, Gene98, Gene5, Gene438, Gene241, Gene186, Gene914, Gene672, Gene270, Gene322   
## Negative: Gene258, Gene339, Gene735, Gene469, Gene503, Gene384, Gene820, Gene595, Gene833, Gene556   
## Gene675, Gene617, Gene168, Gene259, Gene96, Gene470, Gene491, Gene84, Gene42, Gene589   
## Gene607, Gene488, Gene321, Gene649, Gene313, Gene552, Gene276, Gene608, Gene347, Gene919   
## PC\_ 3   
## Positive: Gene746, Gene820, Gene842, Gene373, Gene215, Gene96, Gene927, Gene800, Gene648, Gene739   
## Gene20, Gene46, Gene773, Gene913, Gene402, Gene74, Gene86, Gene349, Gene672, Gene347   
## Gene42, Gene203, Gene99, Gene689, Gene291, Gene605, Gene844, Gene237, Gene932, Gene14   
## Negative: Gene524, Gene627, Gene544, Gene464, Gene23, Gene510, Gene384, Gene251, Gene623, Gene400   
## Gene155, Gene872, Gene313, Gene781, Gene875, Gene809, Gene38, Gene340, Gene366, Gene56   
## Gene642, Gene595, Gene469, Gene315, Gene344, Gene992, Gene973, Gene490, Gene589, Gene127   
## PC\_ 4   
## Positive: Gene731, Gene807, Gene608, Gene811, Gene385, Gene274, Gene402, Gene513, Gene356, Gene12   
## Gene234, Gene809, Gene158, Gene510, Gene508, Gene315, Gene362, Gene118, Gene616, Gene11   
## Gene958, Gene339, Gene544, Gene443, Gene605, Gene961, Gene266, Gene938, Gene276, Gene384   
## Negative: Gene689, Gene866, Gene67, Gene533, Gene746, Gene219, Gene552, Gene584, Gene564, Gene941   
## Gene267, Gene350, Gene595, Gene695, Gene503, Gene777, Gene148, Gene326, Gene934, Gene854   
## Gene613, Gene59, Gene617, Gene290, Gene883, Gene344, Gene711, Gene488, Gene352, Gene56   
## PC\_ 5   
## Positive: Gene464, Gene158, Gene927, Gene265, Gene64, Gene173, Gene362, Gene230, Gene690, Gene739   
## Gene258, Gene424, Gene412, Gene508, Gene668, Gene902, Gene259, Gene349, Gene672, Gene954   
## Gene664, Gene236, Gene99, Gene938, Gene653, Gene523, Gene597, Gene400, Gene560, Gene695   
## Negative: Gene627, Gene148, Gene42, Gene267, Gene438, Gene644, Gene270, Gene103, Gene912, Gene347   
## Gene777, Gene126, Gene800, Gene957, Gene813, Gene911, Gene131, Gene809, Gene384, Gene839   
## Gene987, Gene59, Gene84, Gene1, Gene276, Gene607, Gene262, Gene434, Gene658, Gene117



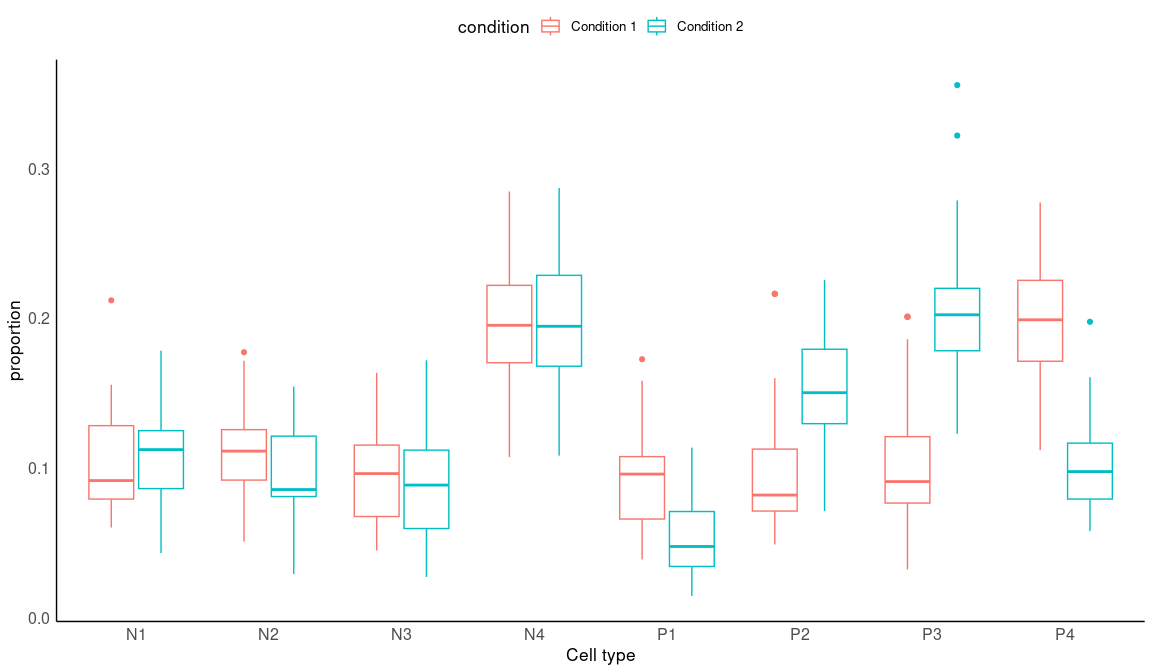
## Centering and scaling data matrix

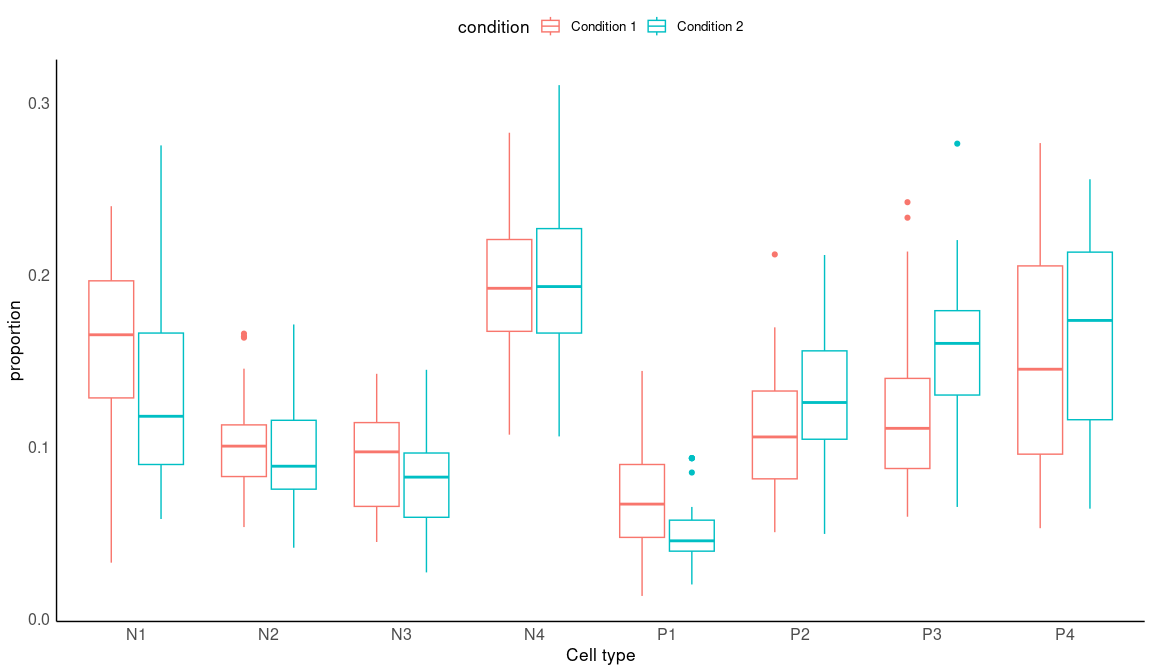
## PC\_ 1   
## Positive: Gene422, Gene948, Gene338, Gene23, Gene233, Gene837, Gene895, Gene131, Gene387, Gene552   
## Gene209, Gene428, Gene740, Gene396, Gene366, Gene835, Gene554, Gene149, Gene269, Gene516   
## Gene488, Gene522, Gene670, Gene571, Gene560, Gene189, Gene65, Gene700, Gene937, Gene866   
## Negative: Gene958, Gene216, Gene810, Gene868, Gene452, Gene870, Gene563, Gene44, Gene820, Gene925   
## Gene329, Gene848, Gene767, Gene972, Gene660, Gene311, Gene377, Gene156, Gene207, Gene323   
## Gene721, Gene507, Gene161, Gene604, Gene953, Gene126, Gene712, Gene673, Gene626, Gene580   
## PC\_ 2   
## Positive: Gene675, Gene736, Gene393, Gene777, Gene255, Gene799, Gene66, Gene925, Gene441, Gene661   
## Gene966, Gene881, Gene665, Gene847, Gene15, Gene580, Gene360, Gene44, Gene180, Gene502   
## Gene600, Gene685, Gene712, Gene208, Gene867, Gene819, Gene26, Gene869, Gene519, Gene323   
## Negative: Gene21, Gene972, Gene197, Gene500, Gene95, Gene56, Gene821, Gene745, Gene725, Gene642   
## Gene127, Gene813, Gene784, Gene362, Gene790, Gene960, Gene43, Gene341, Gene927, Gene911   
## Gene493, Gene126, Gene602, Gene803, Gene588, Gene673, Gene351, Gene576, Gene399, Gene344   
## PC\_ 3   
## Positive: Gene222, Gene56, Gene989, Gene957, Gene639, Gene526, Gene311, Gene143, Gene799, Gene215   
## Gene671, Gene238, Gene74, Gene328, Gene966, Gene815, Gene917, Gene409, Gene1000, Gene932   
## Gene165, Gene860, Gene347, Gene704, Gene717, Gene275, Gene707, Gene721, Gene9, Gene290   
## Negative: Gene339, Gene216, Gene608, Gene870, Gene287, Gene728, Gene803, Gene952, Gene995, Gene436   
## Gene227, Gene745, Gene137, Gene797, Gene360, Gene911, Gene661, Gene810, Gene859, Gene443   
## Gene407, Gene766, Gene993, Gene252, Gene887, Gene505, Gene434, Gene316, Gene566, Gene820   
## PC\_ 4   
## Positive: Gene632, Gene573, Gene874, Gene809, Gene381, Gene762, Gene526, Gene191, Gene804, Gene213   
## Gene790, Gene174, Gene593, Gene3, Gene838, Gene248, Gene783, Gene382, Gene881, Gene991   
## Gene889, Gene813, Gene333, Gene7, Gene328, Gene193, Gene394, Gene710, Gene237, Gene857   
## Negative: Gene215, Gene584, Gene887, Gene284, Gene681, Gene310, Gene409, Gene777, Gene74, Gene350   
## Gene230, Gene239, Gene406, Gene21, Gene956, Gene377, Gene80, Gene252, Gene664, Gene338   
## Gene391, Gene110, Gene534, Gene363, Gene495, Gene900, Gene407, Gene316, Gene810, Gene431   
## PC\_ 5   
## Positive: Gene790, Gene733, Gene1000, Gene948, Gene328, Gene534, Gene495, Gene143, Gene848, Gene953   
## Gene884, Gene56, Gene7, Gene476, Gene275, Gene791, Gene341, Gene717, Gene844, Gene878   
## Gene641, Gene347, Gene839, Gene689, Gene870, Gene206, Gene424, Gene287, Gene116, Gene314   
## Negative: Gene129, Gene582, Gene207, Gene744, Gene246, Gene707, Gene858, Gene137, Gene125, Gene665   
## Gene778, Gene682, Gene350, Gene538, Gene711, Gene905, Gene952, Gene200, Gene960, Gene322   
## Gene580, Gene979, Gene321, Gene21, Gene589, Gene987, Gene838, Gene83, Gene503, Gene851





Draw the proportion of different cell types and ROC curves





### AUC curve

## Setting levels: control = N, case = P

## Setting direction: controls > cases

## Setting levels: control = N, case = P

## Setting direction: controls > cases

## Setting levels: control = N, case = P

## Setting direction: controls > cases

## Setting levels: control = N, case = P

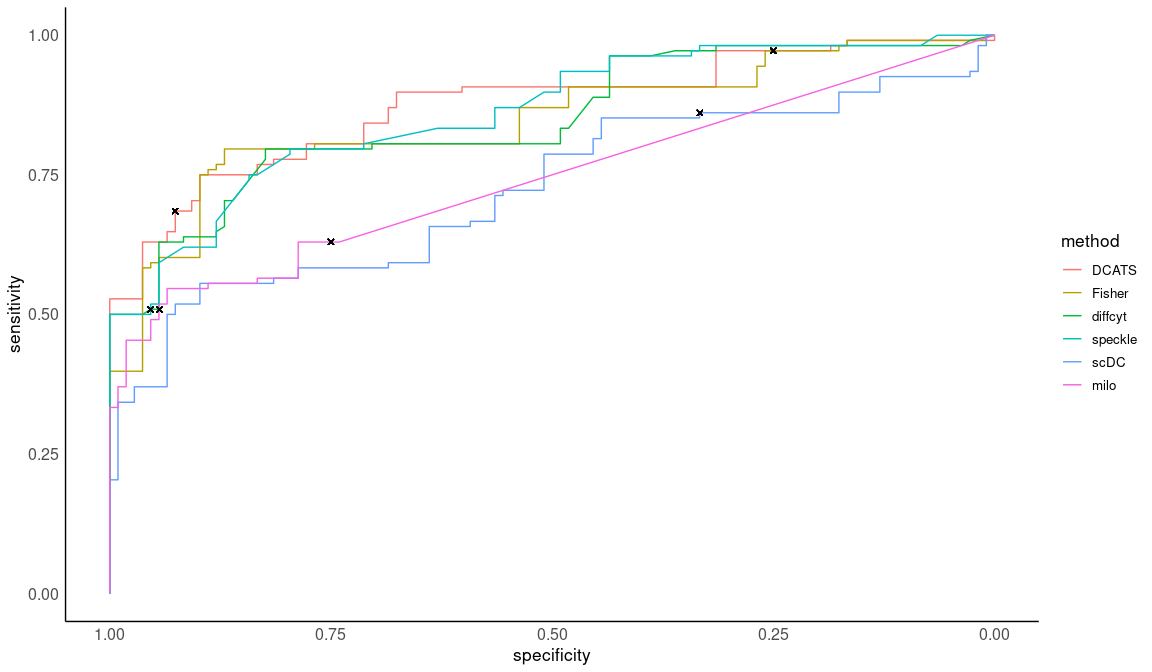
## Setting direction: controls > cases

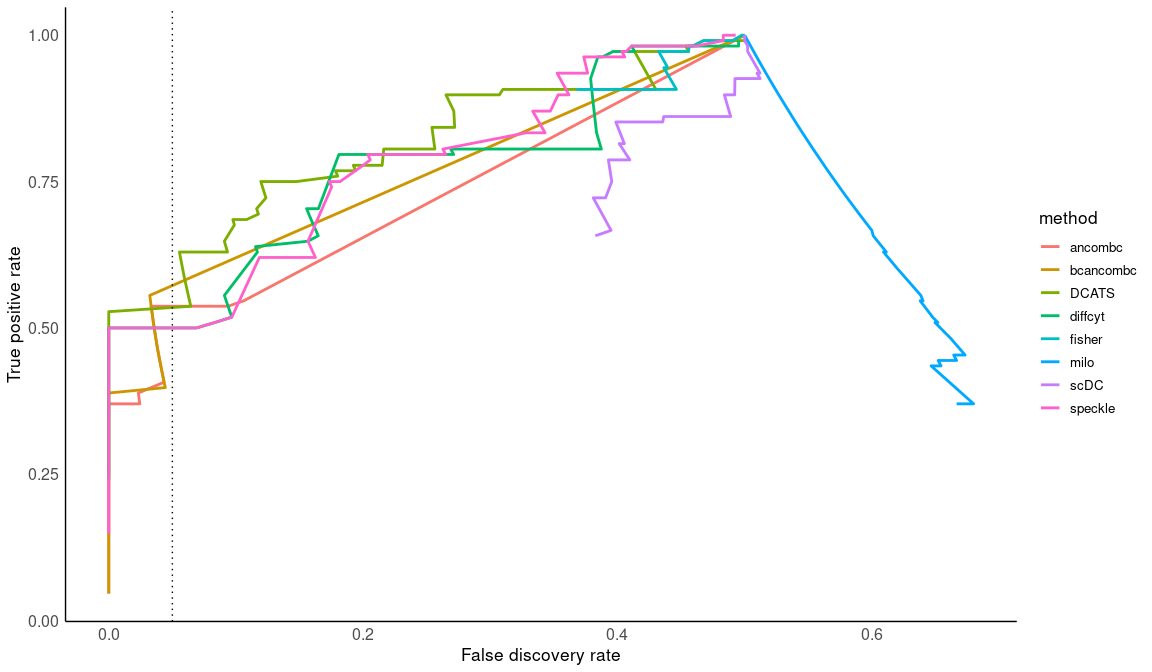
## Setting levels: control = N, case = P

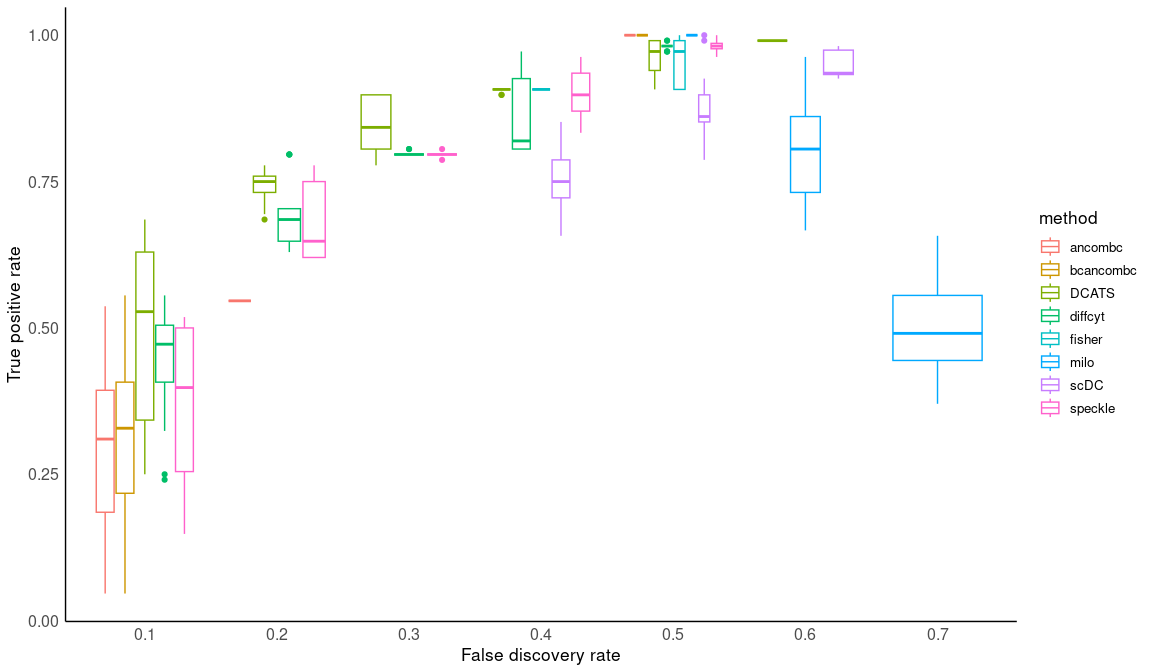
## Setting direction: controls > cases

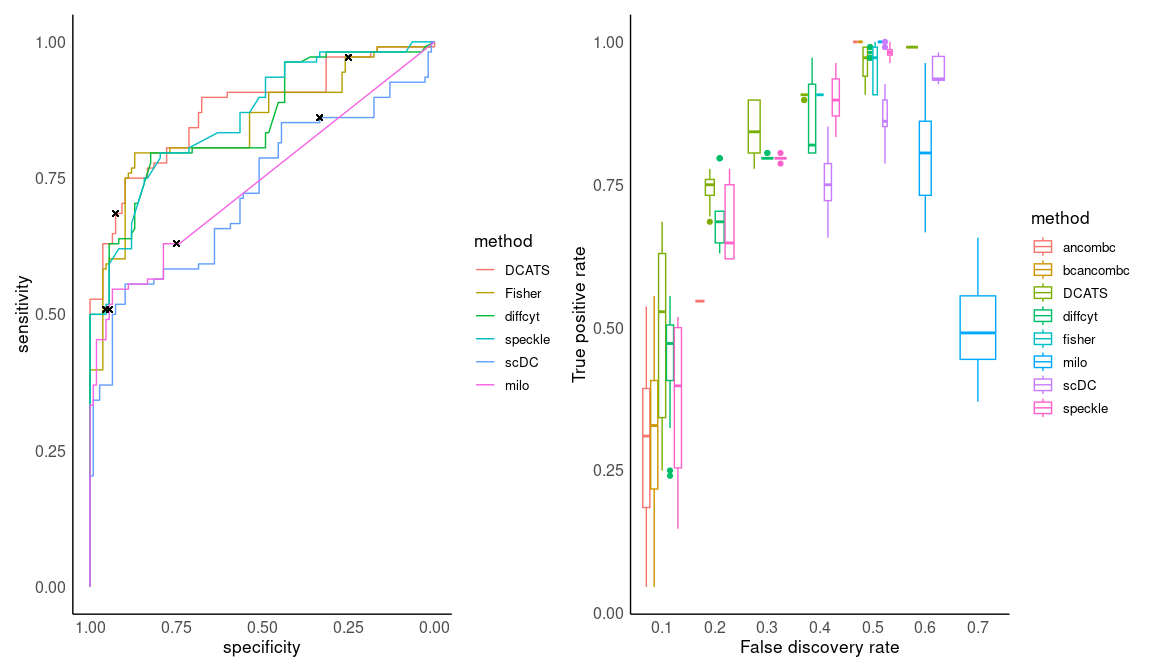
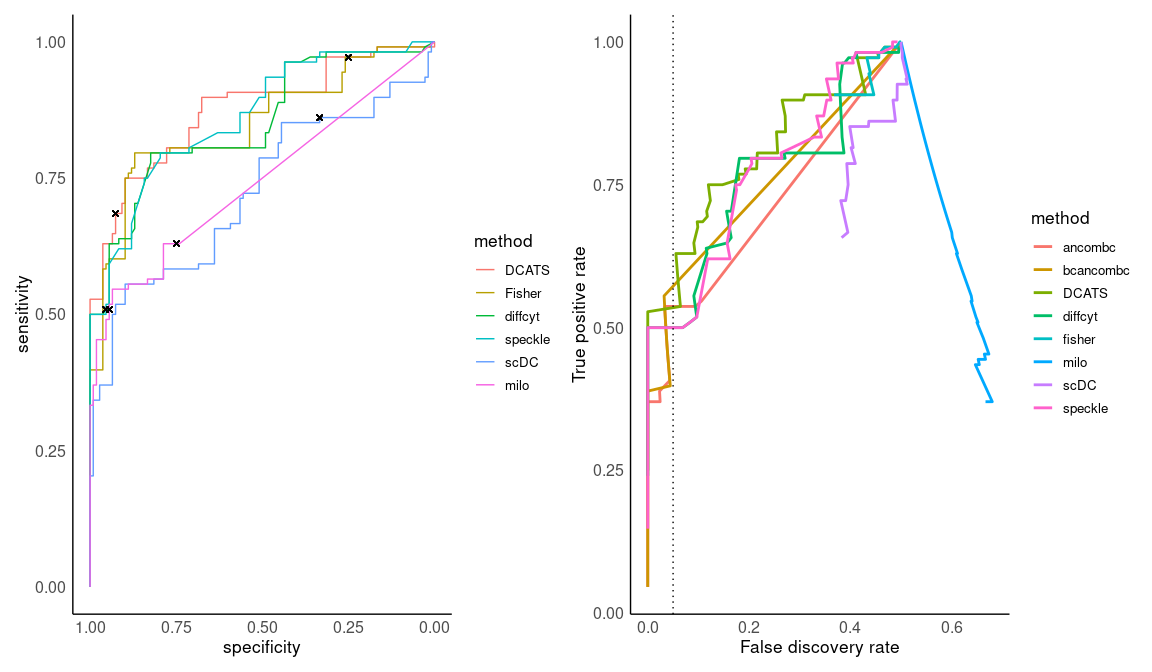
## Setting levels: control = N, case = P

## Setting direction: controls < cases

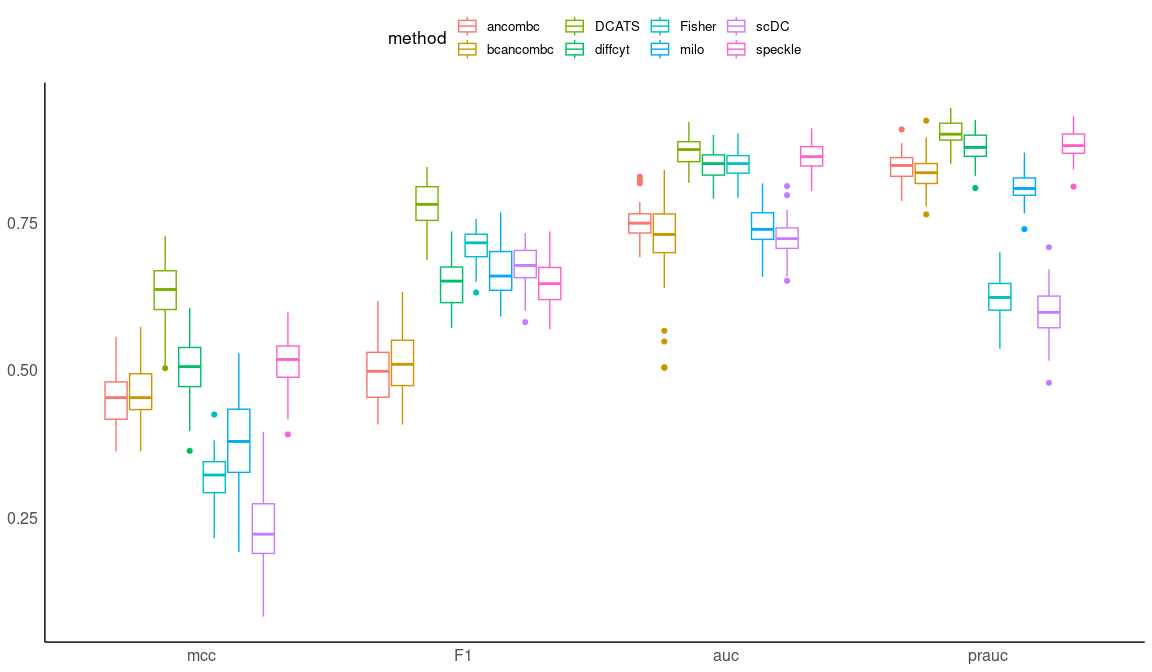








## calculate confidence interval

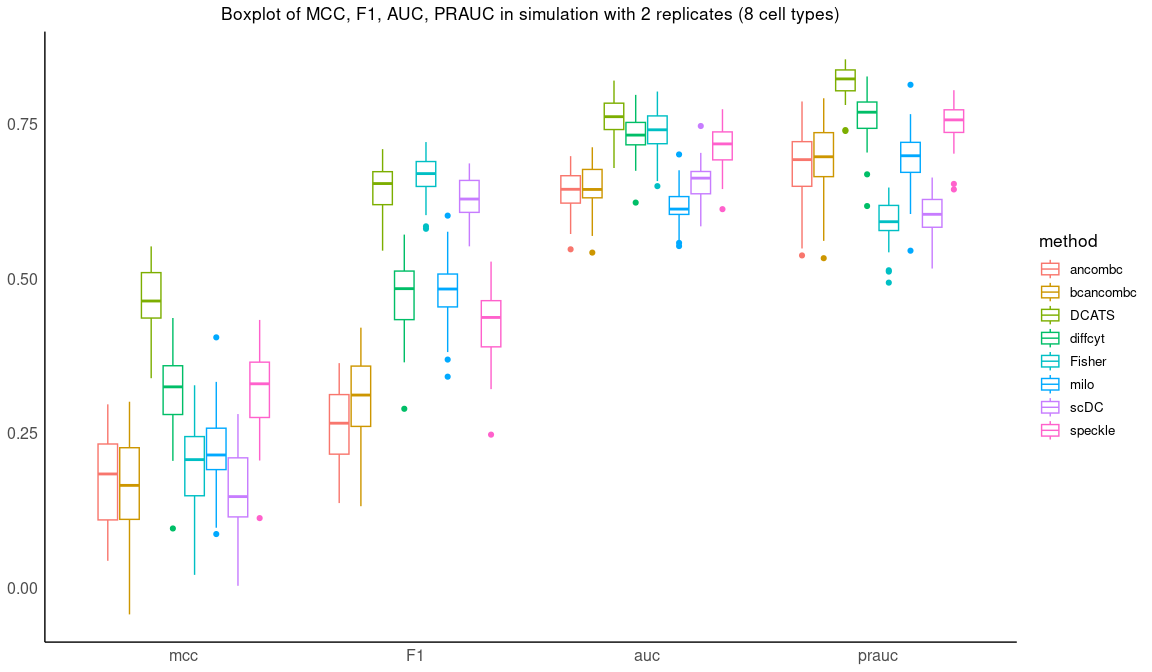


### Figure 2-B

#### Different numbers of replicates

## [1] "/storage/holab/linxy/DCATS/simulation/current/replicates2&2\_K8\_con100\_splatter3000&3000para.RData"  
## [2] "/storage/holab/linxy/DCATS/simulation/current/replicates3&3\_K8\_con100\_splatter3000&3000para.RData"  
## [3] "/storage/holab/linxy/DCATS/simulation/current/replicates4&4\_K8\_con100\_splatter3000&3000para.RData"

### calculate confidence interval



## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

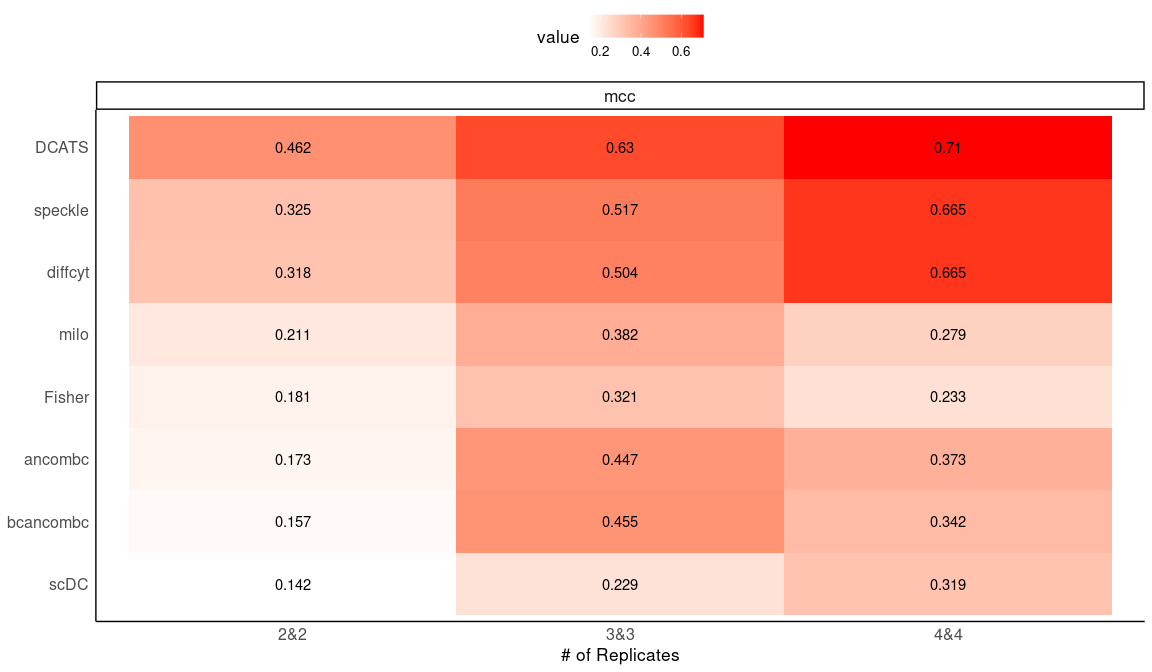
## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

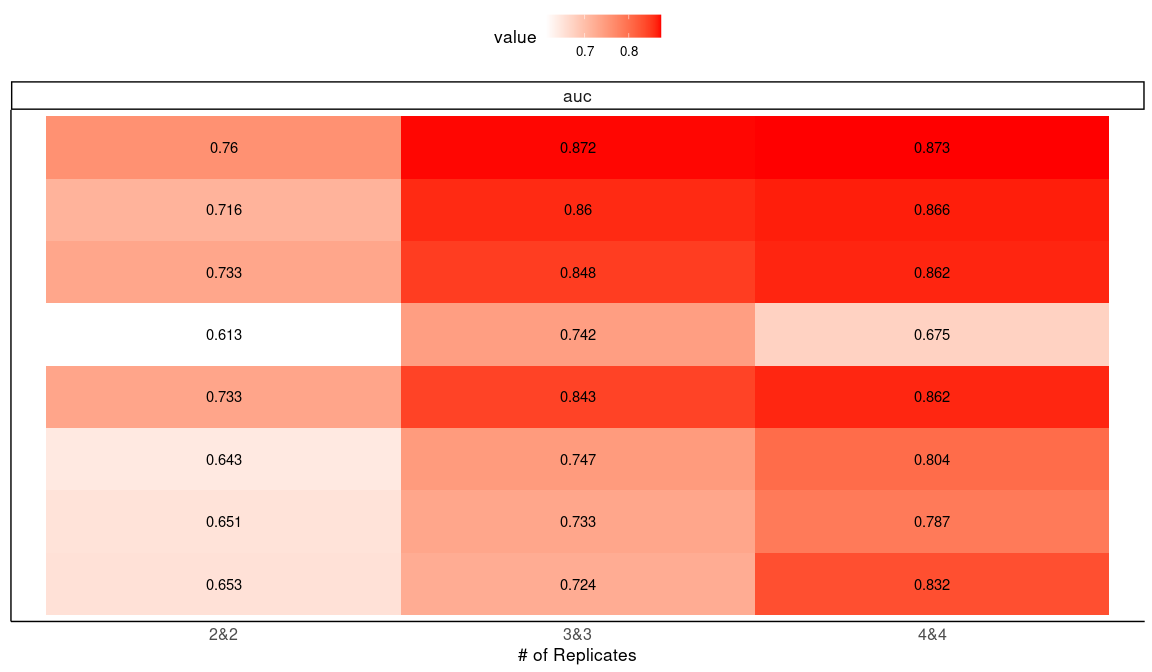
## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

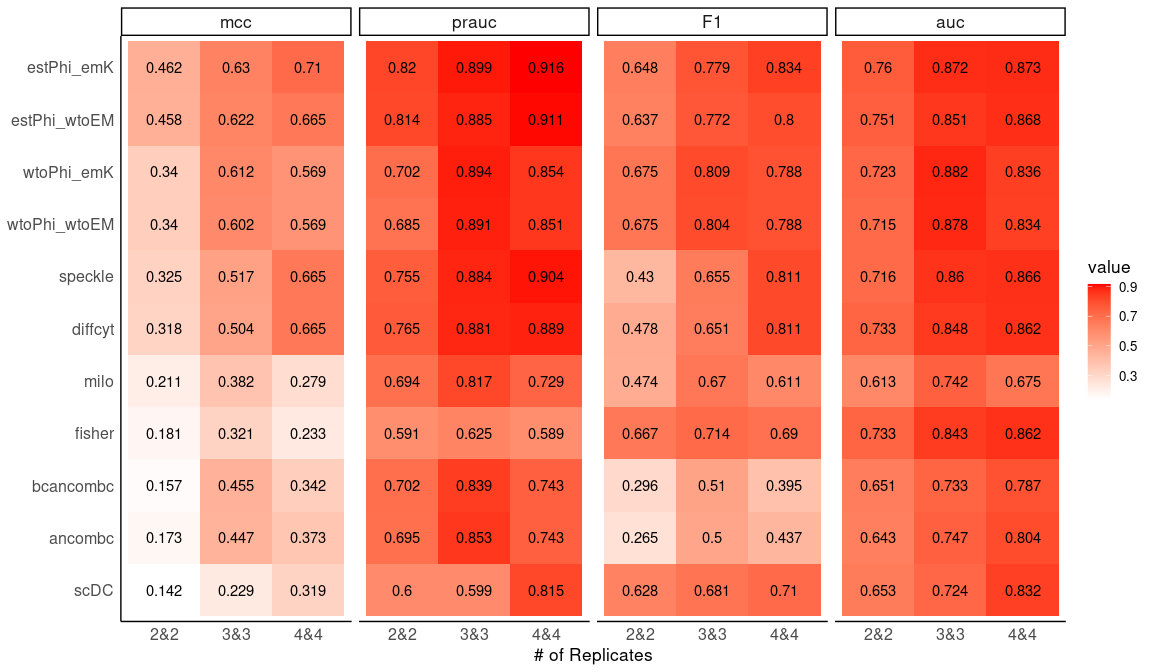
## Warning in if (!is.na(oper[[idx]])) {: the condition has length > 1 and only the  
## first element will be used

## method mcc auc sensitivity specificity F1 replicates  
## 1 estPhi\_emK 0.462 0.760 0.527 0.902 0.648 2&2  
## 2 estPhi\_wtoEM 0.458 0.751 0.509 0.911 0.637 2&2  
## 3 estPhi\_emU 0.451 0.743 0.598 0.839 0.680 2&2  
## 4 fisher 0.181 0.733 0.848 0.304 0.667 2&2  
## 5 diffcyt 0.318 0.733 0.339 0.920 0.478 2&2  
## 6 wtoPhi\_emK 0.340 0.723 0.688 0.652 0.675 2&2  
## 7 speckle 0.325 0.716 0.286 0.955 0.430 2&2  
## 8 wtoPhi\_wtoEM 0.340 0.715 0.688 0.652 0.675 2&2  
## 9 scDC 0.142 0.653 0.732 0.402 0.628 2&2  
## 10 ancombc\_knn 0.157 0.651 0.188 0.920 0.296 2&2  
## 11 ancombc 0.173 0.643 0.161 0.946 0.265 2&2  
## 12 milo 0.211 0.613 0.366 0.821 0.474 2&2  
## 13 wtoPhi\_emK 0.612 0.882 0.824 0.787 0.809 3&3  
## 14 wtoPhi\_wtoEM 0.602 0.878 0.815 0.787 0.804 3&3  
## 15 estPhi\_emK 0.630 0.872 0.685 0.926 0.779 3&3  
## 16 speckle 0.517 0.860 0.509 0.954 0.655 3&3  
## 17 estPhi\_emU 0.622 0.856 0.713 0.898 0.786 3&3  
## 18 estPhi\_wtoEM 0.622 0.851 0.676 0.926 0.772 3&3  
## 19 diffcyt 0.504 0.848 0.509 0.944 0.651 3&3  
## 20 fisher 0.321 0.843 0.972 0.250 0.714 3&3  
## 21 ancombc 0.447 0.747 0.333 1.000 0.500 3&3  
## 22 milo 0.382 0.742 0.630 0.750 0.670 3&3  
## 23 ancombc\_knn 0.455 0.733 0.343 1.000 0.510 3&3  
## 24 scDC 0.229 0.724 0.861 0.333 0.681 3&3  
## 25 estPhi\_emK 0.710 0.873 0.759 0.940 0.834 4&4  
## 26 estPhi\_emU 0.684 0.873 0.793 0.888 0.833 4&4  
## 27 estPhi\_wtoEM 0.665 0.868 0.707 0.940 0.800 4&4  
## 28 speckle 0.665 0.866 0.741 0.914 0.811 4&4  
## 29 diffcyt 0.665 0.862 0.741 0.914 0.811 4&4  
## 30 fisher 0.233 0.862 0.922 0.250 0.690 4&4  
## 31 wtoPhi\_emK 0.569 0.836 0.802 0.767 0.788 4&4  
## 32 wtoPhi\_wtoEM 0.569 0.834 0.802 0.767 0.788 4&4  
## 33 scDC 0.319 0.832 0.888 0.388 0.710 4&4  
## 34 ancombc 0.373 0.804 0.284 0.983 0.437 4&4  
## 35 ancombc\_knn 0.342 0.787 0.250 0.983 0.395 4&4  
## 36 milo 0.279 0.675 0.569 0.707 0.611 4&4





Simulation 2: Supplementary Plots

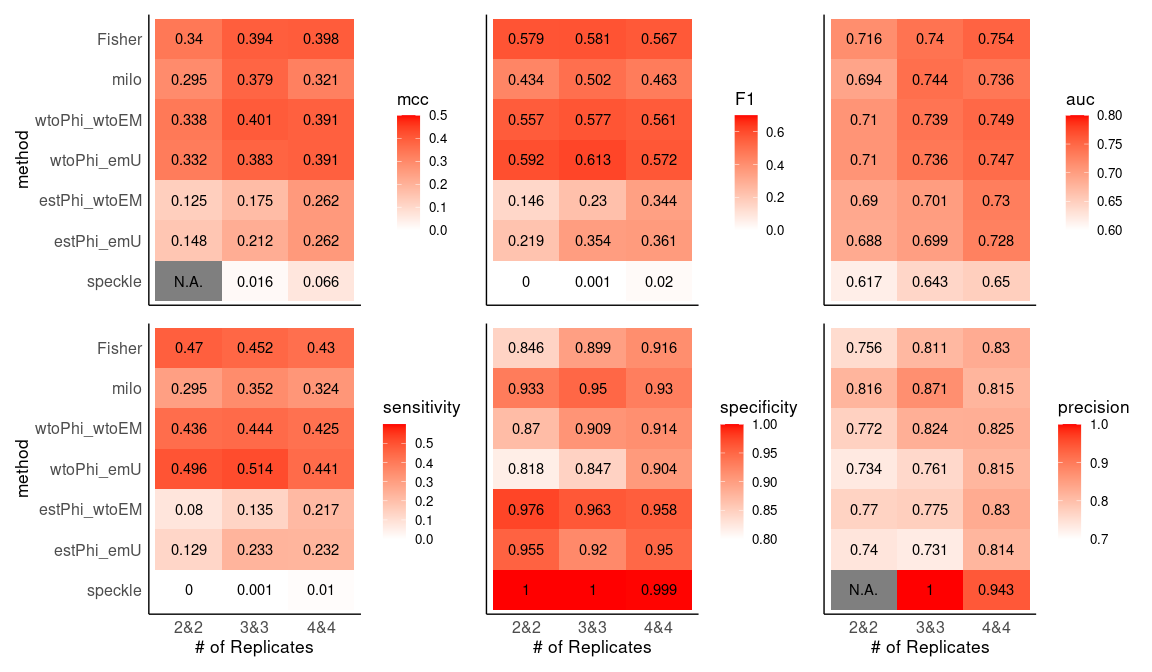


Details of statistics value

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| method | mcc | auc | prauc | sensitivity | specificity | precision | F1 | replicates |
| estPhi\_emK | 0.462 | 0.760 | 0.820 | 0.527 | 0.902 | 0.843 | 0.648 | 2&2 |
| estPhi\_wtoEM | 0.458 | 0.751 | 0.814 | 0.509 | 0.911 | 0.851 | 0.637 | 2&2 |
| wtoPhi\_emK | 0.340 | 0.723 | 0.702 | 0.688 | 0.652 | 0.664 | 0.675 | 2&2 |
| wtoPhi\_wtoEM | 0.340 | 0.715 | 0.685 | 0.688 | 0.652 | 0.664 | 0.675 | 2&2 |
| speckle | 0.325 | 0.716 | 0.755 | 0.286 | 0.955 | 0.865 | 0.430 | 2&2 |
| diffcyt | 0.318 | 0.733 | 0.765 | 0.339 | 0.920 | 0.809 | 0.478 | 2&2 |
| milo | 0.211 | 0.613 | 0.694 | 0.366 | 0.821 | 0.672 | 0.474 | 2&2 |
| fisher | 0.181 | 0.733 | 0.591 | 0.848 | 0.304 | 0.549 | 0.667 | 2&2 |
| ancombc | 0.173 | 0.643 | 0.695 | 0.161 | 0.946 | 0.750 | 0.265 | 2&2 |
| bcancombc | 0.157 | 0.651 | 0.702 | 0.188 | 0.920 | 0.700 | 0.296 | 2&2 |
| scDC | 0.142 | 0.653 | 0.600 | 0.732 | 0.402 | 0.550 | 0.628 | 2&2 |
| estPhi\_emK | 0.630 | 0.872 | 0.899 | 0.685 | 0.926 | 0.902 | 0.779 | 3&3 |
| estPhi\_wtoEM | 0.622 | 0.851 | 0.885 | 0.676 | 0.926 | 0.901 | 0.772 | 3&3 |
| wtoPhi\_emK | 0.612 | 0.882 | 0.894 | 0.824 | 0.787 | 0.795 | 0.809 | 3&3 |
| wtoPhi\_wtoEM | 0.602 | 0.878 | 0.891 | 0.815 | 0.787 | 0.793 | 0.804 | 3&3 |
| speckle | 0.517 | 0.860 | 0.884 | 0.509 | 0.954 | 0.917 | 0.655 | 3&3 |
| diffcyt | 0.504 | 0.848 | 0.881 | 0.509 | 0.944 | 0.902 | 0.651 | 3&3 |
| bcancombc | 0.455 | 0.733 | 0.839 | 0.343 | 1.000 | 1.000 | 0.510 | 3&3 |
| ancombc | 0.447 | 0.747 | 0.853 | 0.333 | 1.000 | 1.000 | 0.500 | 3&3 |
| milo | 0.382 | 0.742 | 0.817 | 0.630 | 0.750 | 0.716 | 0.670 | 3&3 |
| fisher | 0.321 | 0.843 | 0.625 | 0.972 | 0.250 | 0.565 | 0.714 | 3&3 |
| scDC | 0.229 | 0.724 | 0.599 | 0.861 | 0.333 | 0.564 | 0.681 | 3&3 |
| estPhi\_emK | 0.710 | 0.873 | 0.916 | 0.759 | 0.940 | 0.926 | 0.834 | 4&4 |
| estPhi\_wtoEM | 0.665 | 0.868 | 0.911 | 0.707 | 0.940 | 0.921 | 0.800 | 4&4 |
| speckle | 0.665 | 0.866 | 0.904 | 0.741 | 0.914 | 0.896 | 0.811 | 4&4 |
| diffcyt | 0.665 | 0.862 | 0.889 | 0.741 | 0.914 | 0.896 | 0.811 | 4&4 |
| wtoPhi\_emK | 0.569 | 0.836 | 0.854 | 0.802 | 0.767 | 0.775 | 0.788 | 4&4 |
| wtoPhi\_wtoEM | 0.569 | 0.834 | 0.851 | 0.802 | 0.767 | 0.775 | 0.788 | 4&4 |
| ancombc | 0.373 | 0.804 | 0.743 | 0.284 | 0.983 | 0.943 | 0.437 | 4&4 |
| bcancombc | 0.342 | 0.787 | 0.743 | 0.250 | 0.983 | 0.935 | 0.395 | 4&4 |
| scDC | 0.319 | 0.832 | 0.815 | 0.888 | 0.388 | 0.592 | 0.710 | 4&4 |
| milo | 0.279 | 0.675 | 0.729 | 0.569 | 0.707 | 0.660 | 0.611 | 4&4 |
| fisher | 0.233 | 0.862 | 0.589 | 0.922 | 0.250 | 0.552 | 0.690 | 4&4 |

nhoods level

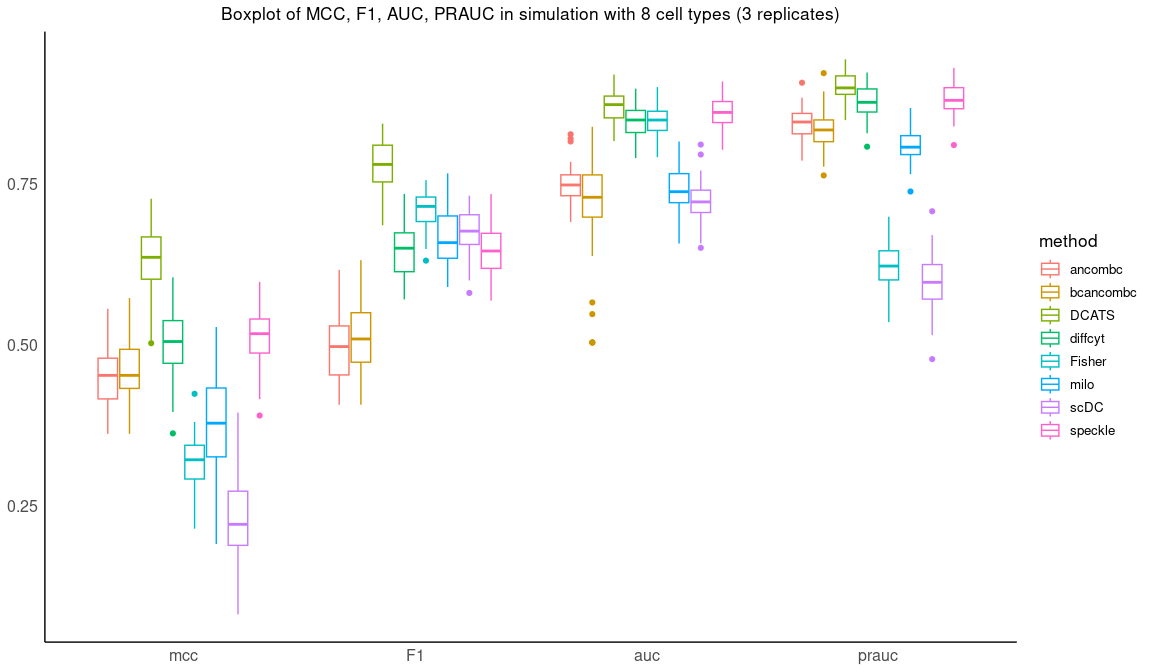
## method mcc auc sensitivity specificity F1 replicates prauc  
## 1 fisher 0.340 0.7157843 0.470 0.846 0.579 2&2 0.716  
## 2 wtoPhi\_emU 0.332 0.7096903 0.496 0.818 0.592 2&2 0.710  
## 3 wtoPhi\_wtoEM 0.338 0.7095942 0.436 0.870 0.557 2&2 0.710  
## 4 milo 0.295 0.6937905 0.295 0.933 0.434 2&2 0.694  
## 5 estPhi\_wtoEM 0.125 0.6895417 0.080 0.976 0.146 2&2 0.690  
## 6 estPhi\_emU 0.148 0.6875149 0.129 0.955 0.219 2&2 0.688  
## 7 speckle NaN 0.6173524 0.000 1.000 0.000 2&2 0.617  
## 8 milo 0.379 0.7439506 0.352 0.950 0.502 3&3 0.744  
## 9 fisher 0.394 0.7403612 0.452 0.899 0.581 3&3 0.740  
## 10 wtoPhi\_wtoEM 0.401 0.7393887 0.444 0.909 0.577 3&3 0.739  
## 11 wtoPhi\_emU 0.383 0.7361322 0.514 0.847 0.613 3&3 0.736  
## 12 estPhi\_wtoEM 0.175 0.7009994 0.135 0.963 0.230 3&3 0.701  
## 13 estPhi\_emU 0.212 0.6994401 0.233 0.920 0.354 3&3 0.699  
## 14 speckle 0.016 0.6428572 0.001 1.000 0.001 3&3 0.643  
## 15 fisher 0.398 0.7535549 0.430 0.916 0.567 4&4 0.754  
## 16 wtoPhi\_wtoEM 0.391 0.7493532 0.425 0.914 0.561 4&4 0.749  
## 17 wtoPhi\_emU 0.391 0.7465912 0.441 0.904 0.572 4&4 0.747  
## 18 milo 0.321 0.7364727 0.324 0.930 0.463 4&4 0.736  
## 19 estPhi\_wtoEM 0.262 0.7304013 0.217 0.958 0.344 4&4 0.730  
## 20 estPhi\_emU 0.262 0.7279494 0.232 0.950 0.361 4&4 0.728  
## 21 speckle 0.066 0.6503482 0.010 0.999 0.020 4&4 0.650



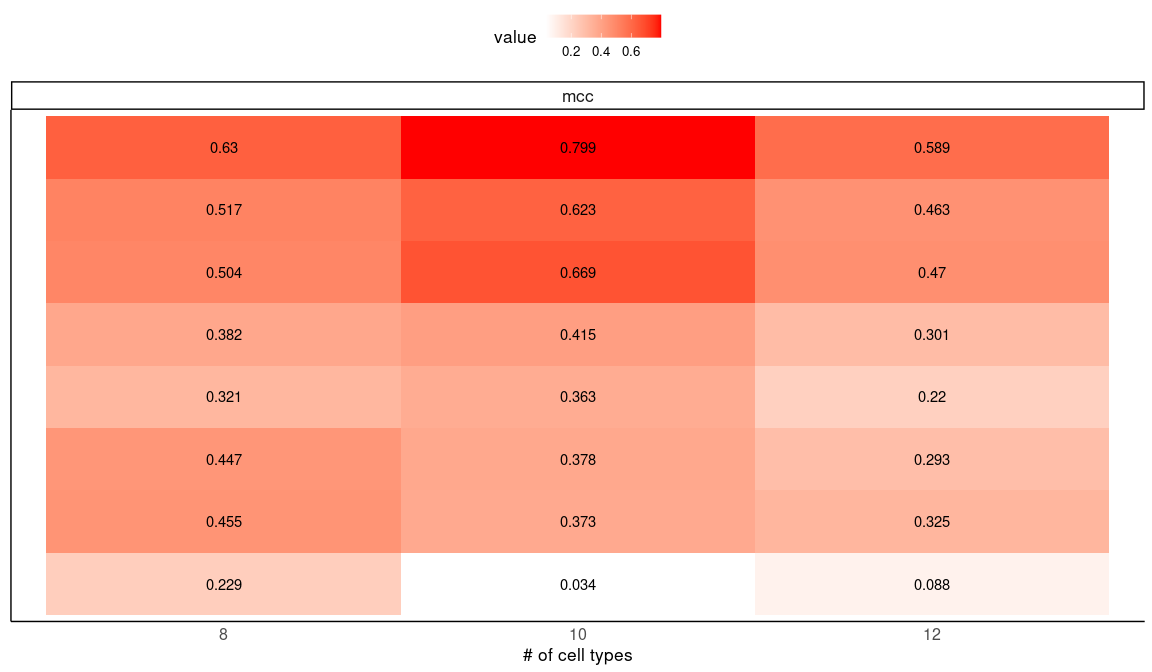
#### Different numbers of clusters

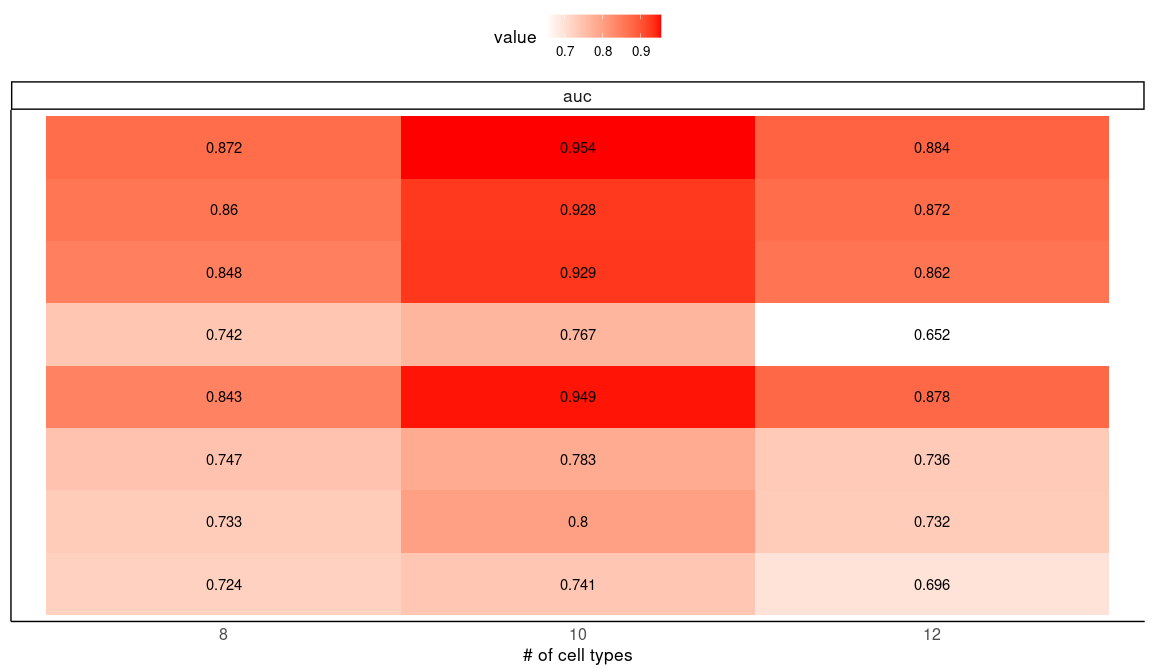
## [1] "/storage/holab/linxy/DCATS/simulation/current/replicates3&3\_K10\_con100\_splatter3000&3000para.RData"  
## [2] "/storage/holab/linxy/DCATS/simulation/current/replicates3&3\_K12\_con100\_splatter3000&3000para.RData"  
## [3] "/storage/holab/linxy/DCATS/simulation/current/replicates3&3\_K8\_con100\_splatter3000&3000para.RData"

### calculate confidence interval

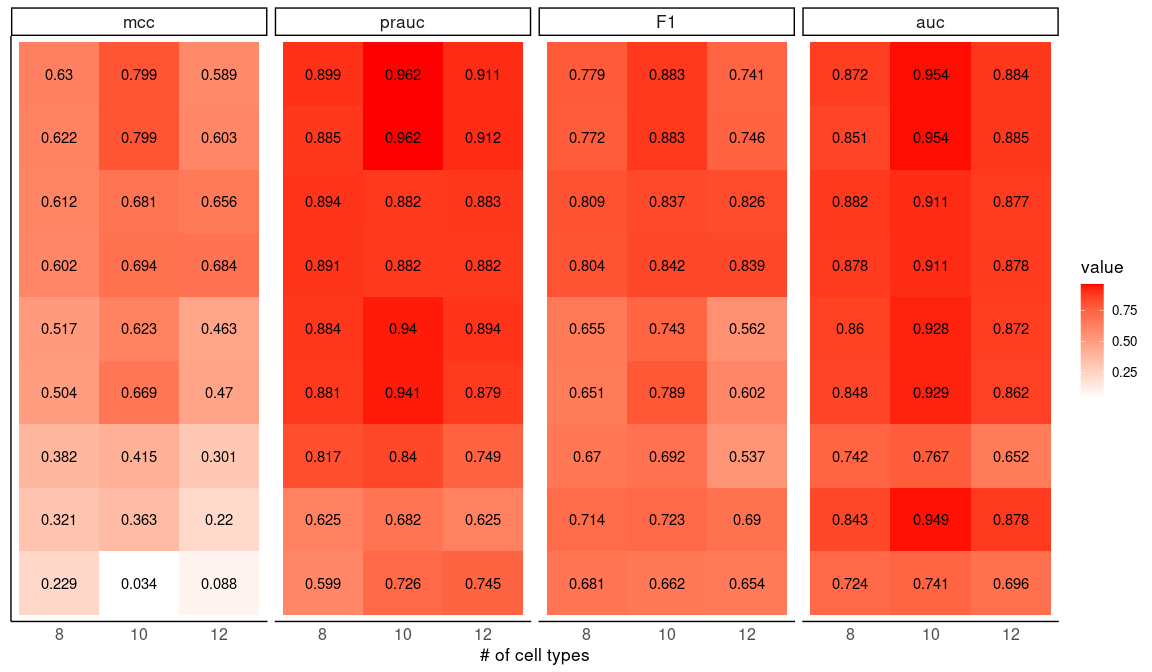


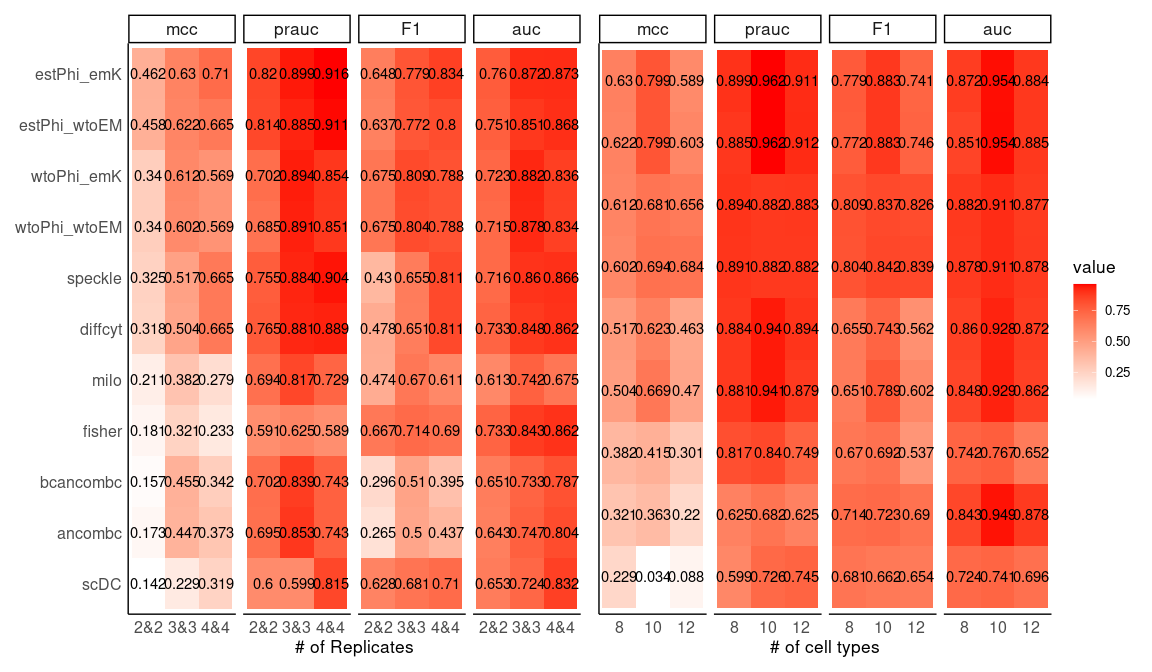
## method mcc auc sensitivity specificity F1 clustersN  
## 1 wtoPhi\_emK 0.612 0.882 0.824 0.787 0.809 8  
## 2 betabin\_null 0.602 0.878 0.815 0.787 0.804 8  
## 3 estPhi\_emK 0.630 0.872 0.685 0.926 0.779 8  
## 4 speckle 0.517 0.860 0.509 0.954 0.655 8  
## 5 estPhi\_null 0.622 0.851 0.676 0.926 0.772 8  
## 6 diffcyt 0.504 0.848 0.509 0.944 0.651 8  
## 7 fisher 0.321 0.843 0.972 0.250 0.714 8  
## 8 milo 0.382 0.742 0.630 0.750 0.670 8  
## 9 scDC 0.229 0.724 0.861 0.333 0.681 8  
## 10 estPhi\_null 0.799 0.954 0.807 0.980 0.883 10  
## 11 estPhi\_emK 0.799 0.954 0.807 0.980 0.883 10  
## 12 fisher 0.363 0.949 1.000 0.233 0.723 10  
## 13 diffcyt 0.669 0.929 0.673 0.967 0.789 10  
## 14 speckle 0.623 0.928 0.607 0.973 0.743 10  
## 15 betabin\_null 0.694 0.911 0.820 0.873 0.842 10  
## 16 wtoPhi\_emK 0.681 0.911 0.820 0.860 0.837 10  
## 17 milo 0.415 0.767 0.660 0.753 0.692 10  
## 18 scDC 0.034 0.741 0.967 0.047 0.662 10  
## 19 estPhi\_null 0.603 0.885 0.628 0.944 0.746 12  
## 20 estPhi\_emK 0.589 0.884 0.628 0.933 0.741 12  
## 21 betabin\_null 0.684 0.878 0.822 0.861 0.839 12  
## 22 fisher 0.220 0.878 0.950 0.194 0.690 12  
## 23 wtoPhi\_emK 0.656 0.877 0.817 0.839 0.826 12  
## 24 speckle 0.463 0.872 0.400 0.978 0.562 12  
## 25 diffcyt 0.470 0.862 0.450 0.956 0.602 12  
## 26 scDC 0.088 0.696 0.889 0.172 0.654 12  
## 27 milo 0.301 0.652 0.422 0.850 0.537 12





Simulation 2: Supplementary Plot



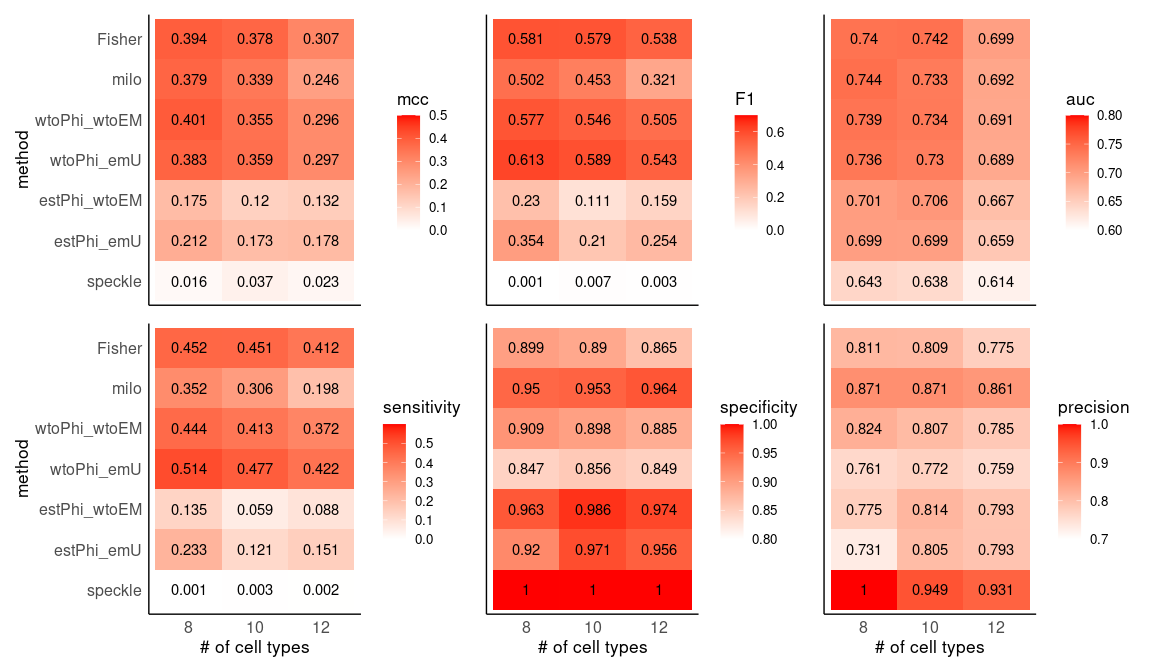


Details of statistics value

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| method | mcc | auc | prauc | sensitivity | specificity | precision | F1 | clustersN |
| estPhi\_emK | 0.630 | 0.872 | 0.899 | 0.685 | 0.926 | 0.902 | 0.779 | 8 |
| estPhi\_wtoEM | 0.622 | 0.851 | 0.885 | 0.676 | 0.926 | 0.901 | 0.772 | 8 |
| wtoPhi\_emK | 0.612 | 0.882 | 0.894 | 0.824 | 0.787 | 0.795 | 0.809 | 8 |
| wtoPhi\_wtoEM | 0.602 | 0.878 | 0.891 | 0.815 | 0.787 | 0.793 | 0.804 | 8 |
| speckle | 0.517 | 0.860 | 0.884 | 0.509 | 0.954 | 0.917 | 0.655 | 8 |
| diffcyt | 0.504 | 0.848 | 0.881 | 0.509 | 0.944 | 0.902 | 0.651 | 8 |
| bcancombc | 0.455 | 0.733 | 0.839 | 0.343 | 1.000 | 1.000 | 0.510 | 8 |
| ancombc | 0.447 | 0.747 | 0.853 | 0.333 | 1.000 | 1.000 | 0.500 | 8 |
| milo | 0.382 | 0.742 | 0.817 | 0.630 | 0.750 | 0.716 | 0.670 | 8 |
| fisher | 0.321 | 0.843 | 0.625 | 0.972 | 0.250 | 0.565 | 0.714 | 8 |
| scDC | 0.229 | 0.724 | 0.599 | 0.861 | 0.333 | 0.564 | 0.681 | 8 |
| estPhi\_wtoEM | 0.799 | 0.954 | 0.962 | 0.807 | 0.980 | 0.976 | 0.883 | 10 |
| estPhi\_emK | 0.799 | 0.954 | 0.962 | 0.807 | 0.980 | 0.976 | 0.883 | 10 |
| wtoPhi\_wtoEM | 0.694 | 0.911 | 0.882 | 0.820 | 0.873 | 0.866 | 0.842 | 10 |
| wtoPhi\_emK | 0.681 | 0.911 | 0.882 | 0.820 | 0.860 | 0.854 | 0.837 | 10 |
| diffcyt | 0.669 | 0.929 | 0.941 | 0.673 | 0.967 | 0.953 | 0.789 | 10 |
| speckle | 0.623 | 0.928 | 0.940 | 0.607 | 0.973 | 0.958 | 0.743 | 10 |
| milo | 0.415 | 0.767 | 0.840 | 0.660 | 0.753 | 0.728 | 0.692 | 10 |
| ancombc | 0.378 | 0.783 | 0.857 | 0.267 | 0.993 | 0.976 | 0.419 | 10 |
| bcancombc | 0.373 | 0.800 | 0.857 | 0.260 | 0.993 | 0.975 | 0.411 | 10 |
| fisher | 0.363 | 0.949 | 0.682 | 1.000 | 0.233 | 0.566 | 0.723 | 10 |
| scDC | 0.034 | 0.741 | 0.726 | 0.967 | 0.047 | 0.503 | 0.662 | 10 |
| wtoPhi\_wtoEM | 0.684 | 0.878 | 0.882 | 0.822 | 0.861 | 0.855 | 0.839 | 12 |
| wtoPhi\_emK | 0.656 | 0.877 | 0.883 | 0.817 | 0.839 | 0.835 | 0.826 | 12 |
| estPhi\_wtoEM | 0.603 | 0.885 | 0.912 | 0.628 | 0.944 | 0.919 | 0.746 | 12 |
| estPhi\_emK | 0.589 | 0.884 | 0.911 | 0.628 | 0.933 | 0.904 | 0.741 | 12 |
| diffcyt | 0.470 | 0.862 | 0.879 | 0.450 | 0.956 | 0.910 | 0.602 | 12 |
| speckle | 0.463 | 0.872 | 0.894 | 0.400 | 0.978 | 0.947 | 0.562 | 12 |
| bcancombc | 0.325 | 0.732 | 0.842 | 0.206 | 0.994 | 0.974 | 0.339 | 12 |
| milo | 0.301 | 0.652 | 0.749 | 0.422 | 0.850 | 0.738 | 0.537 | 12 |
| ancombc | 0.293 | 0.736 | 0.754 | 0.172 | 0.994 | 0.969 | 0.292 | 12 |
| fisher | 0.220 | 0.878 | 0.625 | 0.950 | 0.194 | 0.541 | 0.690 | 12 |
| scDC | 0.088 | 0.696 | 0.745 | 0.889 | 0.172 | 0.518 | 0.654 | 12 |

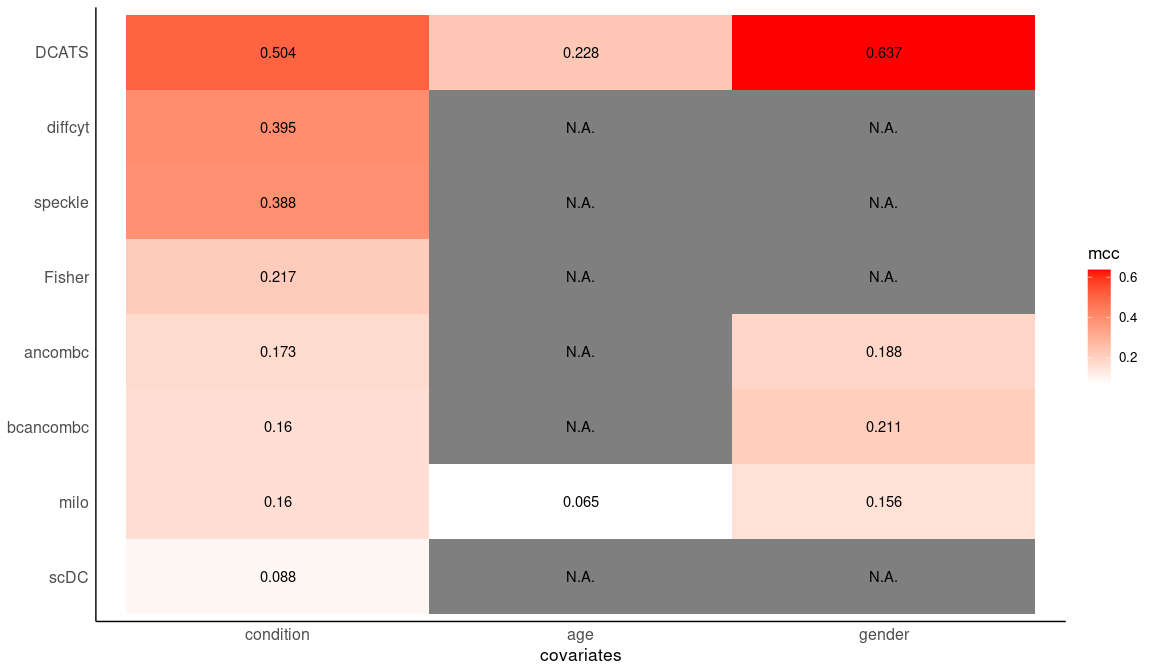
nhoods level

## method mcc auc sensitivity specificity F1 clustersN prauc  
## 1 milo 0.379 0.7439506 0.352 0.950 0.502 8 0.744  
## 2 fisher 0.394 0.7403612 0.452 0.899 0.581 8 0.740  
## 3 wtoPhi\_wtoEM 0.401 0.7393887 0.444 0.909 0.577 8 0.739  
## 4 wtoPhi\_emU 0.383 0.7361322 0.514 0.847 0.613 8 0.736  
## 5 estPhi\_wtoEM 0.175 0.7009994 0.135 0.963 0.230 8 0.701  
## 6 estPhi\_emU 0.212 0.6994401 0.233 0.920 0.354 8 0.699  
## 7 speckle 0.016 0.6428572 0.001 1.000 0.001 8 0.643  
## 8 fisher 0.378 0.7419091 0.451 0.890 0.579 10 0.742  
## 9 wtoPhi\_wtoEM 0.355 0.7336701 0.413 0.898 0.546 10 0.734  
## 10 milo 0.339 0.7330529 0.306 0.953 0.453 10 0.733  
## 11 wtoPhi\_emU 0.359 0.7301010 0.477 0.856 0.589 10 0.730  
## 12 estPhi\_wtoEM 0.120 0.7064403 0.059 0.986 0.111 10 0.706  
## 13 estPhi\_emU 0.173 0.6985041 0.121 0.971 0.210 10 0.699  
## 14 speckle 0.037 0.6383595 0.003 1.000 0.007 10 0.638  
## 15 fisher 0.307 0.6987307 0.412 0.865 0.538 12 0.699  
## 16 milo 0.246 0.6924343 0.198 0.964 0.321 12 0.692  
## 17 wtoPhi\_wtoEM 0.296 0.6910424 0.372 0.885 0.505 12 0.691  
## 18 wtoPhi\_emU 0.297 0.6891288 0.422 0.849 0.543 12 0.689  
## 19 estPhi\_wtoEM 0.132 0.6670580 0.088 0.974 0.159 12 0.667  
## 20 estPhi\_emU 0.178 0.6589271 0.151 0.956 0.254 12 0.659  
## 21 speckle 0.023 0.6143789 0.002 1.000 0.003 12 0.614

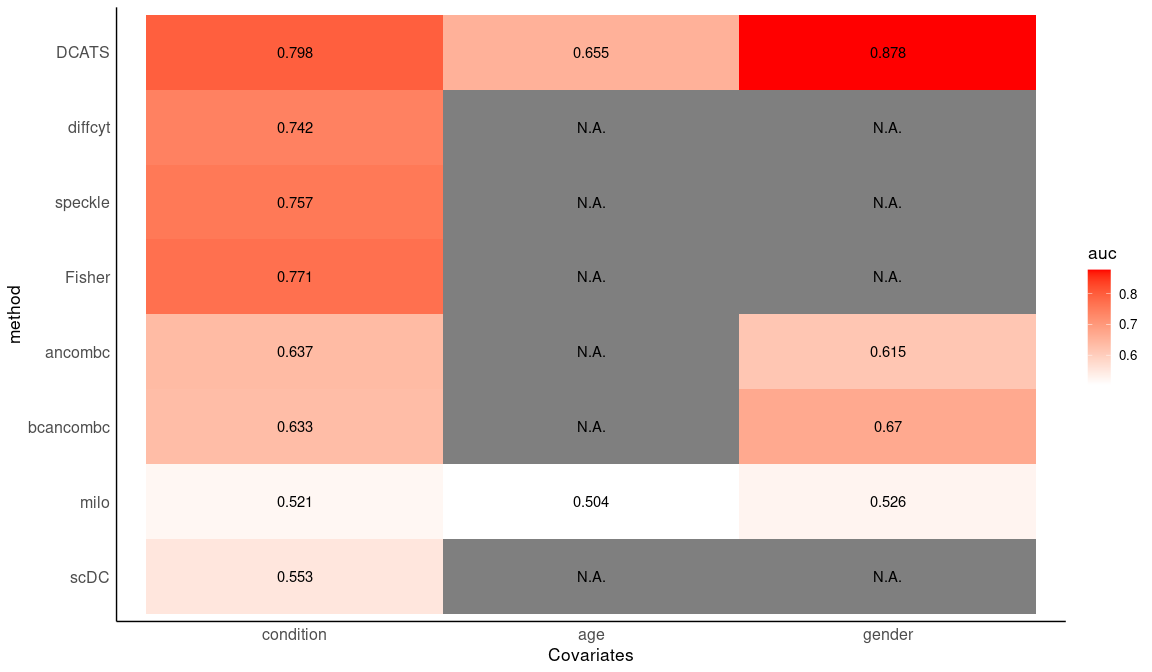


### Figure 2 - C

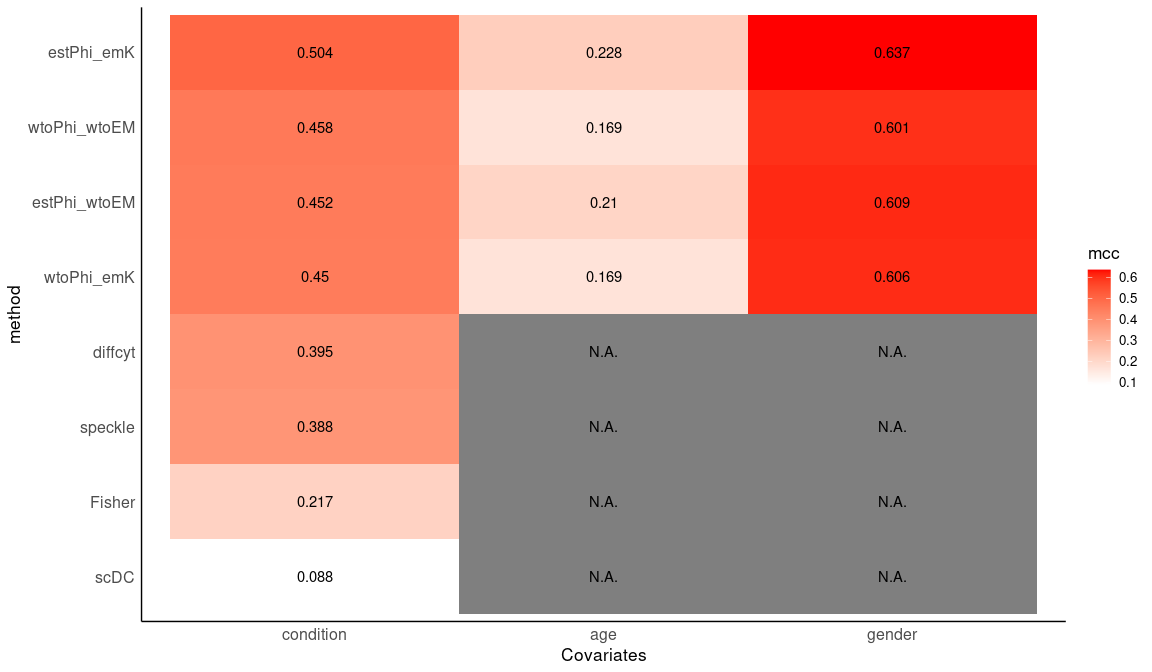
## method mcc auc prauc sensitivity specificity  
## 1 estPhi\_emK 0.50398592 0.7975347 0.8360534 0.600000000 0.8833333  
## 2 estPhi\_emSVM 0.48180779 0.7926736 0.8315759 0.575000000 0.8833333  
## 3 wtoPhi\_emK 0.45000981 0.7883681 0.8191560 0.625000000 0.8166667  
## 4 wtoPhi\_emU 0.44080462 0.7866319 0.8152553 0.625000000 0.8083333  
## 5 betabin\_null 0.45775864 0.7852431 0.8163428 0.633333333 0.8166667  
## 6 wtoPhi\_emSVM 0.45000981 0.7828125 0.8159578 0.625000000 0.8166667  
## 7 wtoPhi\_emT 0.44858910 0.7800694 0.8142453 0.633333333 0.8083333  
## 8 estPhi\_emU 0.46861576 0.7794792 0.8113149 0.625000000 0.8333333  
## 9 estPhi\_null 0.45221368 0.7757639 0.8103624 0.541666667 0.8833333  
## 10 estPhi\_emT 0.44387735 0.7723958 0.7864954 0.608333333 0.8250000  
## 11 fisher 0.21707238 0.7709722 0.5918117 0.883333333 0.2916667  
## 12 speckle 0.38765744 0.7571528 0.7894897 0.333333333 0.9666667  
## 13 diffcyt 0.39498278 0.7419792 0.7861028 0.341666667 0.9666667  
## 14 ancombc\_true 0.16012815 0.6576389 0.7907836 0.050000000 1.0000000  
## 15 ancombc\_svm 0.18569534 0.6411458 0.7751151 0.066666667 1.0000000  
## 16 ancombc 0.17332893 0.6374306 0.7457444 0.058333333 1.0000000  
## 17 ancombc\_knn 0.16012815 0.6333681 0.7671184 0.050000000 1.0000000  
## 18 scDC 0.08838835 0.5531250 0.5428683 0.708333333 0.3750000  
## 19 milo 0.16012815 0.5208333 0.7583333 0.050000000 1.0000000  
## 20 estPhi\_emU 0.26248980 0.6592014 0.6859386 0.300000000 0.9083333  
## 21 estPhi\_emSVM 0.22827719 0.6573264 0.6792429 0.241666667 0.9250000  
## 22 estPhi\_emK 0.22827719 0.6550347 0.6768958 0.241666667 0.9250000  
## 23 wtoPhi\_emU 0.20321418 0.6461806 0.6486288 0.333333333 0.8416667  
## 24 estPhi\_emT 0.19949224 0.6450694 0.6523249 0.275000000 0.8833333  
## 25 wtoPhi\_emK 0.16853174 0.6430208 0.6490306 0.300000000 0.8416667  
## 26 estPhi\_null 0.21004201 0.6422222 0.6646640 0.225000000 0.9250000  
## 27 wtoPhi\_emSVM 0.17732484 0.6413889 0.6472692 0.308333333 0.8416667  
## 28 betabin\_null 0.16853174 0.6364236 0.6430603 0.300000000 0.8416667  
## 29 wtoPhi\_emT 0.15762208 0.6332639 0.6340066 0.300000000 0.8333333  
## 30 milo 0.06468462 0.5041667 0.0000000 0.008333333 1.0000000  
## 31 estPhi\_emK 0.63663932 0.8780367 0.9071110 0.618644068 0.9750000  
## 32 estPhi\_emSVM 0.63663932 0.8749647 0.9046625 0.618644068 0.9750000  
## 33 estPhi\_emU 0.65054678 0.8741525 0.9052171 0.686440678 0.9416667  
## 34 estPhi\_null 0.60925927 0.8698093 0.8984629 0.584745763 0.9750000  
## 35 estPhi\_emT 0.58360346 0.8618291 0.8666093 0.644067797 0.9166667  
## 36 wtoPhi\_emU 0.60847335 0.8586864 0.8879009 0.694915254 0.9000000  
## 37 wtoPhi\_emK 0.60647100 0.8585452 0.8877027 0.703389831 0.8916667  
## 38 wtoPhi\_emSVM 0.60847335 0.8583333 0.8875228 0.694915254 0.9000000  
## 39 betabin\_null 0.60093879 0.8568856 0.8865683 0.686440678 0.9000000  
## 40 wtoPhi\_emT 0.58592163 0.8461158 0.8710085 0.669491525 0.9000000  
## 41 ancombc\_knn 0.21119430 0.6699506 0.7714001 0.084745763 1.0000000  
## 42 ancombc\_true 0.17554669 0.6668785 0.8130453 0.059322034 1.0000000  
## 43 ancombc\_svm 0.19991857 0.6569915 0.7919754 0.076271186 1.0000000  
## 44 ancombc 0.18807484 0.6150071 0.7843676 0.067796610 1.0000000  
## 45 milo 0.15556094 0.5259887 0.6545680 0.084745763 0.9833333  
## precision F1 factor  
## 1 0.8372093 0.69902913 condition  
## 2 0.8313253 0.67980296 condition  
## 3 0.7731959 0.69124424 condition  
## 4 0.7653061 0.68807339 condition  
## 5 0.7755102 0.69724771 condition  
## 6 0.7731959 0.69124424 condition  
## 7 0.7676768 0.69406393 condition  
## 8 0.7894737 0.69767442 condition  
## 9 0.8227848 0.65326633 condition  
## 10 0.7765957 0.68224299 condition  
## 11 0.5549738 0.68167203 condition  
## 12 0.9090909 0.48780488 condition  
## 13 0.9111111 0.49696970 condition  
## 14 1.0000000 0.09523810 condition  
## 15 1.0000000 0.12500000 condition  
## 16 1.0000000 0.11023622 condition  
## 17 1.0000000 0.09523810 condition  
## 18 0.5312500 0.60714286 condition  
## 19 1.0000000 0.09523810 condition  
## 20 0.7659574 0.43113772 age  
## 21 0.7631579 0.36708861 age  
## 22 0.7631579 0.36708861 age  
## 23 0.6779661 0.44692737 age  
## 24 0.7021277 0.39520958 age  
## 25 0.6545455 0.41142857 age  
## 26 0.7500000 0.34615385 age  
## 27 0.6607143 0.42045455 age  
## 28 0.6545455 0.41142857 age  
## 29 0.6428571 0.40909091 age  
## 30 1.0000000 0.01652893 age  
## 31 0.9605263 0.75257732 gender  
## 32 0.9605263 0.75257732 gender  
## 33 0.9204545 0.78640777 gender  
## 34 0.9583333 0.72631579 gender  
## 35 0.8837209 0.74509804 gender  
## 36 0.8723404 0.77358491 gender  
## 37 0.8645833 0.77570093 gender  
## 38 0.8723404 0.77358491 gender  
## 39 0.8709677 0.76777251 gender  
## 40 0.8681319 0.75598086 gender  
## 41 1.0000000 0.15625000 gender  
## 42 1.0000000 0.11200000 gender  
## 43 1.0000000 0.14173228 gender  
## 44 1.0000000 0.12698413 gender  
## 45 0.8333333 0.15384615 gender



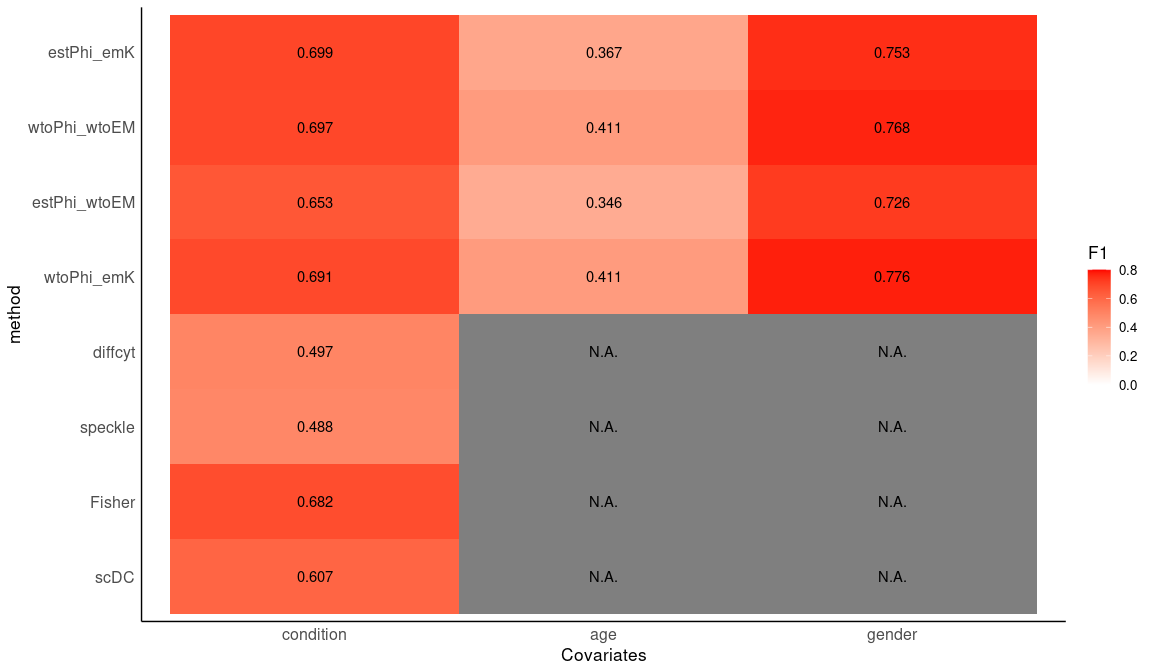
## Saving 12 x 7 in image



## Saving 12 x 7 in image



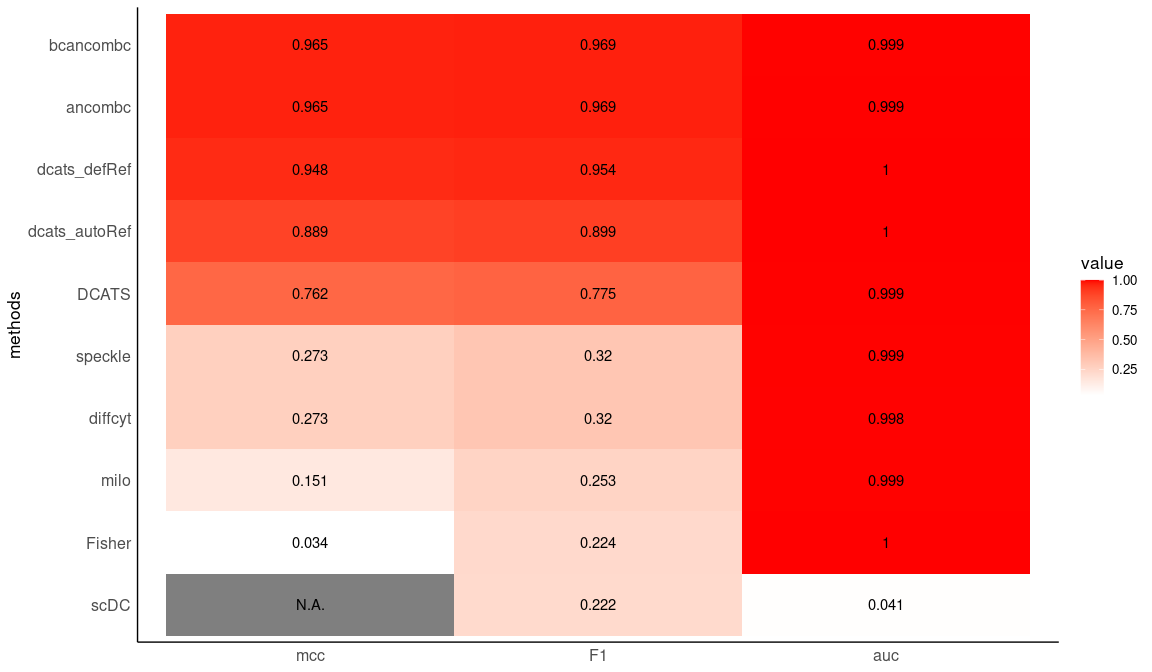
## Saving 12 x 7 in image



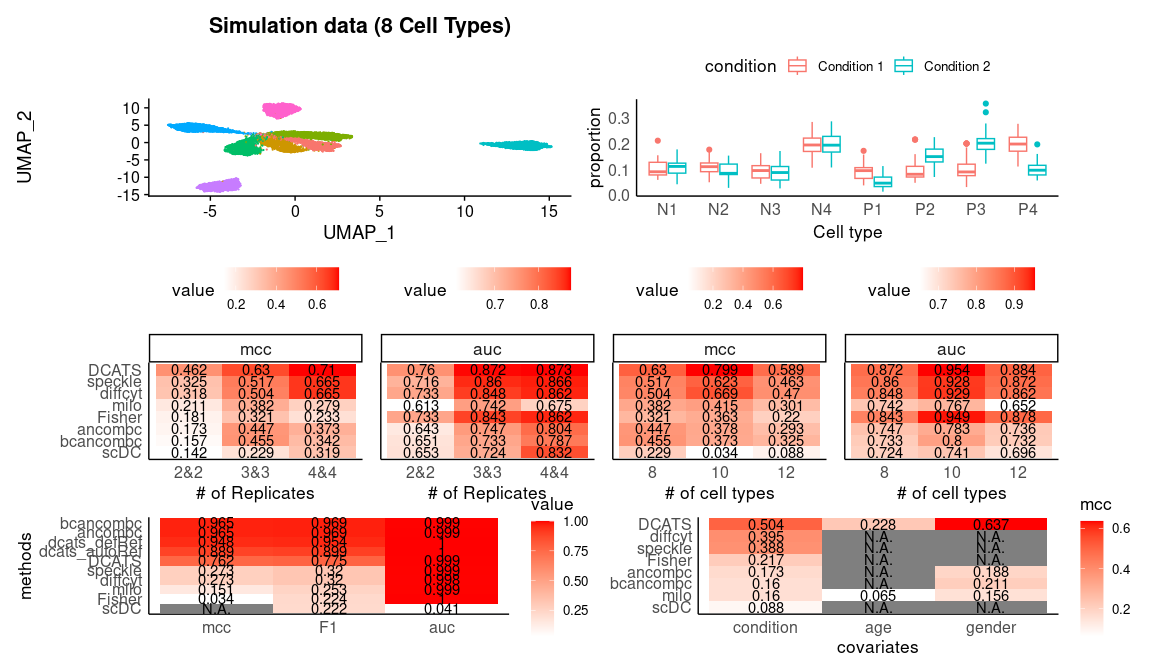
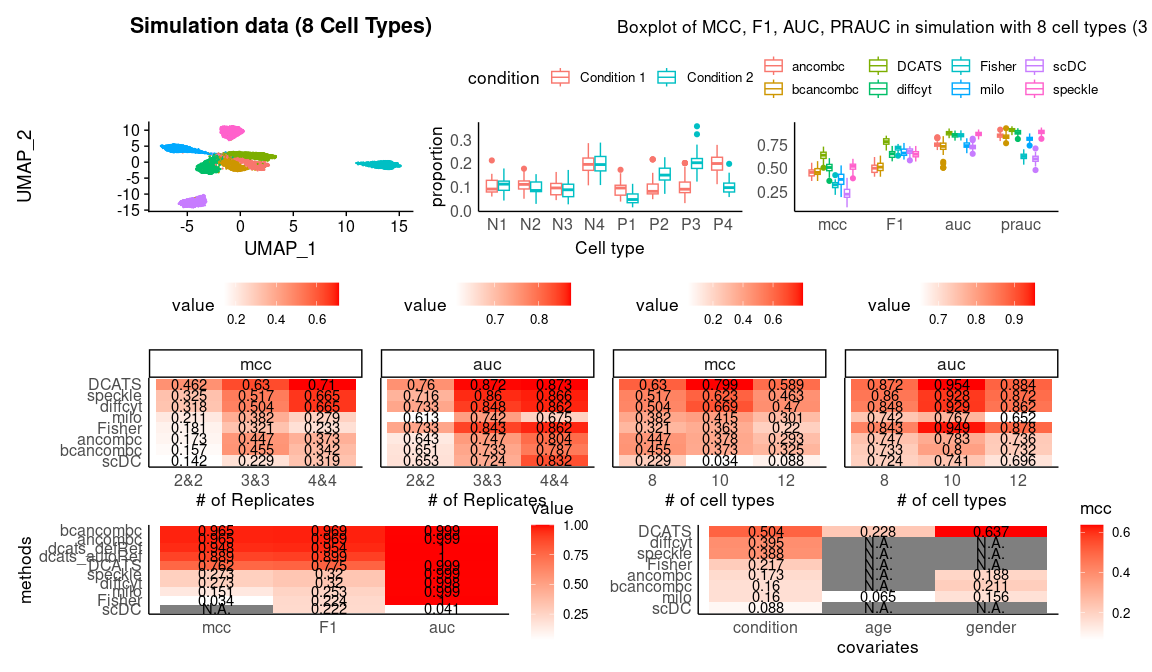
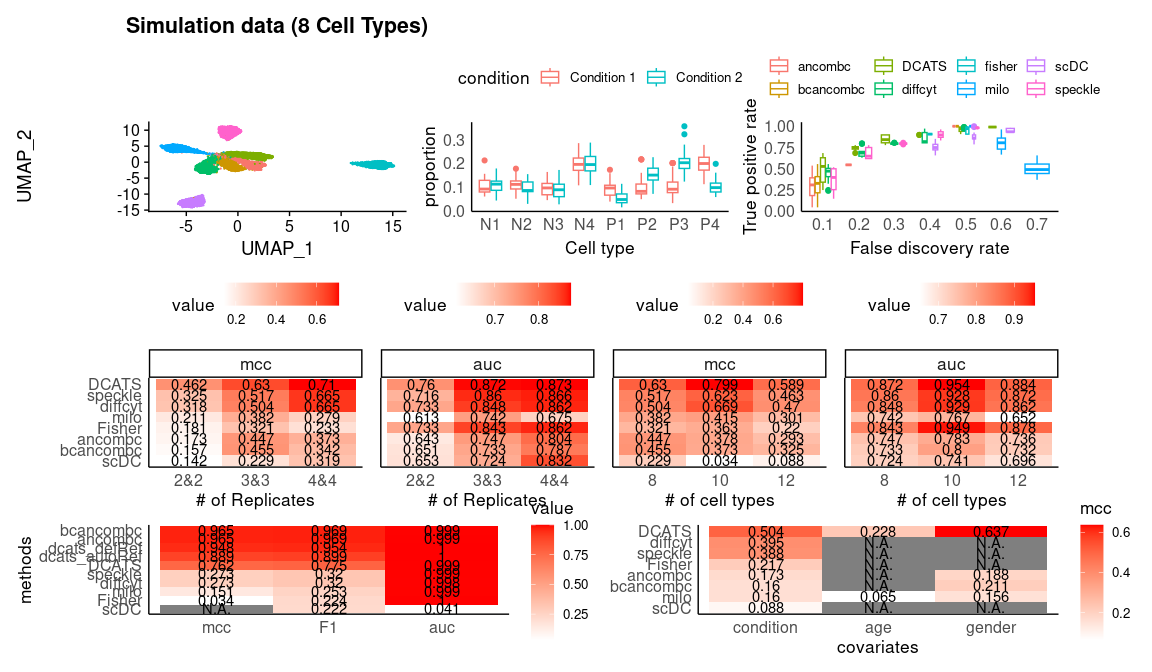
## Saving 12 x 7 in image

### extreme case1

## methods mcc auc prauc sensitivity specificity  
## 1 ancombc 0.96474656 0.99903374 1.0000000 1 0.99078341  
## 2 ancombc\_knn 0.96474656 0.99881076 1.0000000 1 0.99078341  
## 3 ancombc\_svm 0.96474656 0.99903374 1.0000000 1 0.99078341  
## 4 emSVM\_truth3 0.94824030 0.99977702 1.0000000 1 0.98617512  
## 5 emK\_truth3 0.94824030 0.99977702 0.9687500 1 0.98617512  
## 6 ancombc\_true 0.94824030 0.99576334 0.9393939 1 0.98617512  
## 7 emSVM\_truth2 0.91720763 0.99977702 0.9687500 1 0.97695853  
## 8 emK\_truth2 0.91720763 0.99977702 0.9687500 1 0.97695853  
## 9 emK\_org3 0.88852332 1.00000000 1.0000000 1 0.96774194  
## 10 emK\_refK3 0.87496729 0.99992567 1.0000000 1 0.96313364  
## 11 emSVM\_org3 0.87496729 1.00000000 0.9687500 1 0.96313364  
## 12 emSVM\_refSVM3 0.86189161 1.00000000 0.9687500 1 0.95852535  
## 13 estPhi\_null 0.79204116 0.99895942 1.0000000 1 0.93087558  
## 14 estPhi\_emSVM 0.77151675 0.99895942 1.0000000 1 0.92165899  
## 15 estPhi\_emK 0.76169215 0.99925673 1.0000000 1 0.91705069  
## 16 emK\_refK2 0.75213980 1.00000000 1.0000000 1 0.91244240  
## 17 emK\_org2 0.75213980 1.00000000 1.0000000 1 0.91244240  
## 18 emSVM\_refSVM2 0.74284637 1.00000000 0.9687500 1 0.90783410  
## 19 emSVM\_org2 0.74284637 1.00000000 0.9687500 1 0.90783410  
## 20 estPhi\_emT 0.67648143 0.99925673 0.9393939 1 0.87096774  
## 21 estPhi\_emU 0.67648143 0.99940538 1.0000000 1 0.87096774  
## 22 speckle 0.27293966 0.99881076 0.7045455 1 0.39170507  
## 23 diffcyt 0.27293966 0.99843913 0.5636364 1 0.39170507  
## 24 betabin\_null 0.19468147 0.99672960 0.5438596 1 0.23963134  
## 25 wtoPhi\_emU 0.19468147 0.99502007 0.5438596 1 0.23963134  
## 26 wtoPhi\_emSVM 0.18993429 0.99620931 0.5254237 1 0.23041475  
## 27 wtoPhi\_emK 0.18993429 0.99576334 0.5344828 1 0.23041475  
## 28 wtoPhi\_emT 0.18755234 0.99635796 0.5254237 1 0.22580645  
## 29 milo 0.15065500 0.99895942 NA 1 0.15668203  
## 30 fisher 0.03407991 1.00000000 1.0000000 1 0.00921659  
## 31 scDC NaN 0.04140033 0.1260335 1 0.00000000  
## precision F1  
## 1 0.9393939 0.9687500  
## 2 0.9393939 0.9687500  
## 3 0.9393939 0.9687500  
## 4 0.9117647 0.9538462  
## 5 0.9117647 0.9538462  
## 6 0.9117647 0.9538462  
## 7 0.8611111 0.9253731  
## 8 0.8611111 0.9253731  
## 9 0.8157895 0.8985507  
## 10 0.7948718 0.8857143  
## 11 0.7948718 0.8857143  
## 12 0.7750000 0.8732394  
## 13 0.6739130 0.8051948  
## 14 0.6458333 0.7848101  
## 15 0.6326531 0.7750000  
## 16 0.6200000 0.7654321  
## 17 0.6200000 0.7654321  
## 18 0.6078431 0.7560976  
## 19 0.6078431 0.7560976  
## 20 0.5254237 0.6888889  
## 21 0.5254237 0.6888889  
## 22 0.1901840 0.3195876  
## 23 0.1901840 0.3195876  
## 24 0.1581633 0.2731278  
## 25 0.1581633 0.2731278  
## 26 0.1565657 0.2707424  
## 27 0.1565657 0.2707424  
## 28 0.1557789 0.2695652  
## 29 0.1448598 0.2530612  
## 30 0.1260163 0.2238267  
## 31 0.1250000 0.2222222



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| methods | mcc | auc | prauc | sensitivity | specificity | precision | F1 |
| NA | 0.917 | 1.000 | 0.969 | 1 | 0.977 | 0.861 | 0.925 |
| NA | 0.752 | 1.000 | 1.000 | 1 | 0.912 | 0.620 | 0.765 |
| Fisher | 0.034 | 1.000 | 1.000 | 1 | 0.009 | 0.126 | 0.224 |
| ancombc | 0.965 | 0.999 | 1.000 | 1 | 0.991 | 0.939 | 0.969 |
| bcancombc | 0.965 | 0.999 | 1.000 | 1 | 0.991 | 0.939 | 0.969 |
| DCATS | 0.762 | 0.999 | 1.000 | 1 | 0.917 | 0.633 | 0.775 |
| speckle | 0.273 | 0.999 | 0.705 | 1 | 0.392 | 0.190 | 0.320 |
| milo | 0.151 | 0.999 | NA | 1 | 0.157 | 0.145 | 0.253 |
| diffcyt | 0.273 | 0.998 | 0.564 | 1 | 0.392 | 0.190 | 0.320 |
| scDC | NaN | 0.041 | 0.126 | 1 | 0.000 | 0.125 | 0.222 |





|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| methods | mcc | auc | prauc | sensitivity | specificity | precision | F1 |
| emK\_defRef | 0.917 | 1.000 | 0.969 | 1 | 0.977 | 0.861 | 0.925 |
| emK\_autoRef | 0.752 | 1.000 | 1.000 | 1 | 0.912 | 0.620 | 0.765 |
| Fisher | 0.034 | 1.000 | 1.000 | 1 | 0.009 | 0.126 | 0.224 |
| ancombc | 0.965 | 0.999 | 1.000 | 1 | 0.991 | 0.939 | 0.969 |
| bcancombc | 0.965 | 0.999 | 1.000 | 1 | 0.991 | 0.939 | 0.969 |
| estPhi\_emK | 0.762 | 0.999 | 1.000 | 1 | 0.917 | 0.633 | 0.775 |
| speckle | 0.273 | 0.999 | 0.705 | 1 | 0.392 | 0.190 | 0.320 |
| milo | 0.151 | 0.999 | NA | 1 | 0.157 | 0.145 | 0.253 |
| diffcyt | 0.273 | 0.998 | 0.564 | 1 | 0.392 | 0.190 | 0.320 |
| scDC | NaN | 0.041 | 0.126 | 1 | 0.000 | 0.125 | 0.222 |

## Figure 4

### real-world data 1 - Experiment 7

The ‘group’ column started with ‘B’ is the indicators of replicates

## `summarise()` has grouped output by 'batch', 'condition'. You can override  
## using the `.groups` argument.

## `summarise()` has grouped output by 'condition'. You can override using the  
## `.groups` argument.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | truth | betabin\_null | wtoPhi\_emSVM | estPhi\_null | estPhi\_emSVM | fisher | scDC | speckle | milo\_pct | treatment |
| Endocrine | N | 0.130 | 0.130 | 0.288 | 0.273 | 0.000 | 0.360 | 0.714 | 0.000 | Hpoly.Day3 |
| Enterocyte | N | 0.641 | 0.641 | 0.448 | 0.410 | 0.000 | 0.002 | 0.748 | 0.071 | Hpoly.Day3 |
| Enterocyte.Progenitor | N | 0.103 | 0.103 | 0.315 | 0.278 | 0.000 | 0.341 | 0.714 | 0.067 | Hpoly.Day3 |
| Goblet | N | 0.437 | 0.437 | 0.589 | 0.571 | 0.002 | 0.003 | 0.748 | 0.000 | Hpoly.Day3 |
| Stem | N | 0.254 | 0.254 | 0.222 | 0.164 | 0.032 | 0.010 | 0.714 | 0.000 | Hpoly.Day3 |
| TA | N | 0.883 | 0.883 | 0.883 | 0.975 | 0.205 | 0.099 | 0.993 | 0.000 | Hpoly.Day3 |
| TA.Early | N | 0.288 | 0.288 | 0.640 | 0.612 | 0.002 | 0.006 | 0.781 | 0.088 | Hpoly.Day3 |
| Tuft | P | 0.040 | 0.040 | 0.070 | 0.061 | 0.000 | 0.069 | 0.714 | 0.000 | Hpoly.Day3 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | truth | betabin\_null | wtoPhi\_emSVM | estPhi\_null | estPhi\_emSVM | fisher | scDC | speckle | milo\_pct | treatment |
| Endocrine | N | 0.390 | 0.390 | 0.925 | 0.867 | 0.379 | 0.035 | 0.923 | 0.000 | Hpoly.Day10 |
| Enterocyte | P | 0.091 | 0.091 | 0.009 | 0.006 | 0.000 | 0.002 | 0.149 | 0.472 | Hpoly.Day10 |
| Enterocyte.Progenitor | N | 0.132 | 0.132 | 0.285 | 0.233 | 0.000 | 0.103 | 0.349 | 0.162 | Hpoly.Day10 |
| Goblet | P | 0.051 | 0.051 | 0.116 | 0.105 | 0.000 | 0.009 | 0.235 | 0.217 | Hpoly.Day10 |
| Stem | N | 0.824 | 0.824 | 0.865 | 0.830 | 0.279 | 0.406 | 0.923 | 0.096 | Hpoly.Day10 |
| TA | N | 0.852 | 0.852 | 0.836 | 0.592 | 1.000 | 0.572 | 0.923 | 0.647 | Hpoly.Day10 |
| TA.Early | P | 0.015 | 0.015 | 0.132 | 0.083 | 0.000 | 0.224 | 0.235 | 0.266 | Hpoly.Day10 |
| Tuft | P | 0.020 | 0.020 | 0.013 | 0.010 | 0.000 | 0.000 | 0.112 | 0.500 | Hpoly.Day10 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | truth | betabin\_null | wtoPhi\_emSVM | estPhi\_null | estPhi\_emSVM | fisher | scDC | speckle | milo\_pct | treatment |
| Endocrine | N | 0.426 | 0.426 | 0.798 | 0.784 | 0.429 | 0.840 | 0.761 | 0.308 | Salmonella |
| Enterocyte | P | 0.008 | 0.008 | 0.000 | 0.000 | 0.000 | 0.022 | 0.008 | 0.457 | Salmonella |
| Enterocyte.Progenitor | N | 0.146 | 0.146 | 0.561 | 0.495 | 0.000 | 0.357 | 0.647 | 0.019 | Salmonella |
| Goblet | N | 0.681 | 0.681 | 0.821 | 0.802 | 0.558 | 0.253 | 0.864 | 0.100 | Salmonella |
| Stem | P | 0.031 | 0.031 | 0.059 | 0.046 | 0.000 | 0.080 | 0.094 | 0.320 | Salmonella |
| TA | P | 0.026 | 0.026 | 0.192 | 0.318 | 0.000 | 0.801 | 0.216 | 0.538 | Salmonella |
| TA.Early | P | 0.030 | 0.030 | 0.261 | 0.268 | 0.000 | 0.390 | 0.216 | 0.015 | Salmonella |
| Tuft | N | 0.587 | 0.587 | 0.833 | 0.844 | 0.129 | 0.059 | 0.761 | 0.000 | Salmonella |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Tuft | Endocrine | Enterocyte.Progenitor | Stem | Goblet | TA.Early | Enterocyte | TA |
| origin | P | N | N | N | N | N | N | N |
| DCATS | \* | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| fisher | \*\*\* | \*\*\* | \*\*\* | \*\* | \*\*\* | \*\*\* | \*\*\* | n.s. |
| scDC | \* | n.s. | n.s. | \*\*\* | \*\*\* | \*\*\* | \*\*\* | \* |
| speckle | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| milo | N | N | N | N | N | N | N | N |

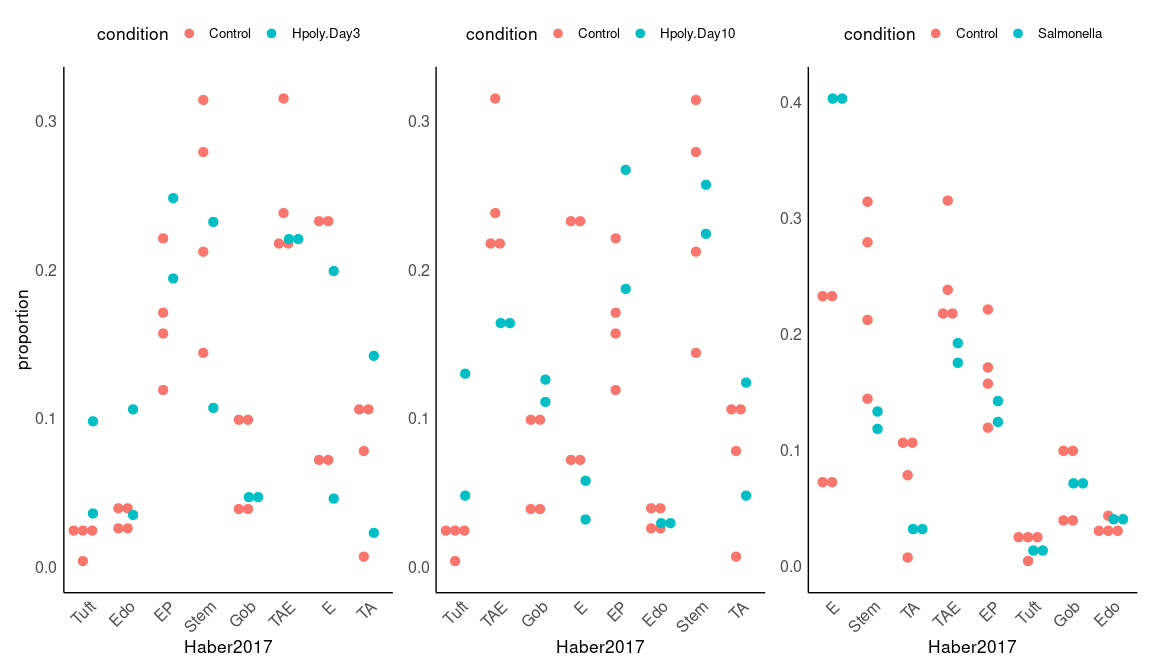
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Tuft | TA.Early | Goblet | Enterocyte | Enterocyte.Progenitor | Endocrine | Stem | TA |
| origin | P | P | P | P | N | N | N | N |
| DCATS | \*\* | \* | n.s. | \*\*\* | n.s. | n.s. | n.s. | n.s. |
| fisher | \*\*\* | \*\*\* | \*\*\* | \*\*\* | \*\*\* | n.s. | n.s. | n.s. |
| scDC | \*\*\* | n.s. | \*\*\* | \*\*\* | n.s. | \*\* | n.s. | n.s. |
| speckle | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| milo | P | P | P | P | N | N | N | P |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Enterocyte | Stem | TA | TA.Early | Enterocyte.Progenitor | Tuft | Goblet | Endocrine |
| origin | P | P | P | P | N | N | N | N |
| DCATS | \*\*\* | \*\* | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| fisher | \*\*\* | \*\*\* | \*\*\* | \*\*\* | \*\*\* | n.s. | n.s. | n.s. |
| scDC | \*\* | \* | n.s. | n.s. | n.s. | \* | n.s. | n.s. |
| speckle | \*\*\* | \* | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| milo | P | P | P | N | N | N | N | P |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Endocrine | Enterocyte | Enterocyte.Progenitor | Goblet | Stem | TA | TA.Early | Tuft |
| Endocrine | 0.9785933 | 0.0006868 | 0.0005495 | 0.0093209 | 0.0102934 | 0.0114537 | 0.0087549 | 0 |
| Enterocyte | 0.0000000 | 0.9800824 | 0.0170330 | 0.0000000 | 0.0000000 | 0.0000000 | 0.0000000 | 0 |
| Enterocyte.Progenitor | 0.0000000 | 0.0185440 | 0.9401099 | 0.0000000 | 0.0005147 | 0.0572687 | 0.0087549 | 0 |
| Goblet | 0.0091743 | 0.0000000 | 0.0000000 | 0.9826897 | 0.0056613 | 0.0017621 | 0.0019455 | 0 |
| Stem | 0.0030581 | 0.0006868 | 0.0000000 | 0.0026631 | 0.8625836 | 0.0933921 | 0.0345331 | 0 |
| TA | 0.0000000 | 0.0000000 | 0.0335165 | 0.0000000 | 0.0761709 | 0.7506608 | 0.0617704 | 0 |
| TA.Early | 0.0000000 | 0.0000000 | 0.0082418 | 0.0039947 | 0.0303654 | 0.0748899 | 0.8793774 | 0 |
| Tuft | 0.0091743 | 0.0000000 | 0.0005495 | 0.0013316 | 0.0144107 | 0.0105727 | 0.0048638 | 1 |

Supplementary Figures: Distribution of each cell type after bias correction

## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.



## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.

## real-world data2

## cell tsne1 tsne2 ind condition cluster clusterRes  
## 1 AAACATACAATGCC-1 -4.277833 -19.294709 107 ctrl 5 CD4 T cells  
## 2 AAACATACATTTCC-1 -27.640373 14.966629 1016 ctrl 9 CD14+ Monocytes  
## 3 AAACATACCAGAAA-1 -27.493646 28.924885 1256 ctrl 9 CD14+ Monocytes  
## 4 AAACATACCAGCTA-1 -28.132584 24.925484 1256 ctrl 9 CD14+ Monocytes  
## 5 AAACATACCATGCA-1 -10.468194 -5.984389 1488 ctrl 3 CD4 T cells  
## 6 AAACATACCTCGCT-1 -24.367997 20.429285 1256 ctrl 9 CD14+ Monocytes  
## multiplets  
## 1 doublet  
## 2 singlet  
## 3 singlet  
## 4 doublet  
## 5 singlet  
## 6 singlet

## `summarise()` has grouped output by 'condition'. You can override using the  
## `.groups` argument.

## # A tibble: 16 × 3  
## # Groups: condition [2]  
## condition clusterRes n  
## <chr> <chr> <int>  
## 1 ctrl B cells 1488  
## 2 ctrl CD14+ Monocytes 3365  
## 3 ctrl CD4 T cells 6005  
## 4 ctrl CD8 T cells 1409  
## 5 ctrl Dendritic cells 227  
## 6 ctrl FCGR3A+ Monocytes 906  
## 7 ctrl Megakaryocytes 166  
## 8 ctrl NK cells 1051  
## 9 stim B cells 1392  
## 10 stim CD14+ Monocytes 3082  
## 11 stim CD4 T cells 6028  
## 12 stim CD8 T cells 1225  
## 13 stim Dendritic cells 245  
## 14 stim FCGR3A+ Monocytes 1008  
## 15 stim Megakaryocytes 180  
## 16 stim NK cells 1279

## # A tibble: 0 × 2  
## # … with 2 variables: cell <chr>, n <int>

## BARCODE RD.TOTL RD.PASS RD.UNIQ N.SNP BEST SNG.1ST  
## 1 AAACATACAATGCC-1 5799 280 269 185 DBL-107-1244-0.500 107  
## 2 AAACATACATTTCC-1 5466 592 501 236 SNG-1016 1016  
## 3 AAACATACCAGAAA-1 4337 330 300 154 SNG-1256 1256  
## 4 AAACATACCAGCTA-1 7120 418 338 179 DBL-1256-1244-0.500 1256  
## 5 AAACATACCATGCA-1 2422 86 76 54 SNG-1488 1488  
## 6 AAACATACCTCGCT-1 5312 561 497 207 SNG-1256 1256  
## SNG.LLK1 SNG.2ND SNG.LLK2 SNG.LLK0 DBL.1ST DBL.2ND ALPHA LLK12  
## 1 -78.5388 1244 -92.3864 -75.9878 107 1244 0.5 -62.8745  
## 2 -61.8525 1256 -193.3940 -135.9317 1016 1256 0.5 -93.0213  
## 3 -50.6213 1015 -100.6965 -67.2587 101 1256 0.5 -64.5022  
## 4 -70.2734 1244 -110.3962 -82.1916 1256 1244 0.5 -67.3787  
## 5 -16.1294 1015 -31.7146 -27.0167 1015 1488 0.5 -18.5646  
## 6 -66.3586 1015 -180.3066 -115.0556 1256 101 0.5 -101.4645  
## LLK1 LLK2 LLK10 LLK20 LLK00 PRB.DBL PRB.SNG1  
## 1 -78.5388 -92.3864 -88.6649 -94.1498 -77.4107 1.00e+00 1  
## 2 -61.8525 -193.3940 -103.2284 -168.3927 -132.6010 4.15e-15 1  
## 3 -107.8190 -50.6213 -111.3319 -64.5022 -72.4244 1.95e-07 1  
## 4 -70.2734 -110.3962 -80.3745 -103.1706 -85.1143 7.21e-01 1  
## 5 -31.7146 -16.1294 -30.2045 -22.5295 -28.2980 1.28e-02 1  
## 6 -66.3586 -190.9367 -101.4645 -194.9529 -124.2197 8.42e-17 1

## BARCODE RD.TOTL RD.PASS RD.UNIQ N.SNP BEST SNG.1ST SNG.LLK1  
## 1 AAACATACCAAGCT-1 2491 239 228 141 SNG-101 101 -31.2822  
## 2 AAACATACCCCTAC-1 3747 282 274 173 SNG-1488 1488 -37.4261  
## 3 AAACATACCCGTAA-1 2844 144 133 108 SNG-1244 1244 -35.0043  
## 4 AAACATACCCTCGT-1 3575 199 189 115 SNG-1488 1488 -31.5489  
## 5 AAACATACGAGGTG-1 2649 170 158 99 SNG-1488 1488 -33.2782  
## 6 AAACATACGCGAAG-1 9874 880 822 345 SNG-101 101 -139.7810  
## SNG.2ND SNG.LLK2 SNG.LLK0 DBL.1ST DBL.2ND ALPHA LLK12 LLK1  
## 1 1015 -81.3393 -59.5871 101 1015 0.5 -47.8026 -31.2822  
## 2 1256 -104.7874 -74.9247 1256 1488 0.5 -53.8991 -104.7874  
## 3 1488 -61.5604 -51.6974 1488 1244 0.5 -36.9775 -61.5604  
## 4 1244 -79.8040 -57.9696 1256 1488 0.5 -42.8116 -83.5802  
## 5 1244 -66.4966 -48.9522 101 1488 0.5 -36.4330 -75.8228  
## 6 1256 -266.4763 -170.1924 101 1488 0.5 -154.9333 -139.7810  
## LLK2 LLK10 LLK20 LLK00 PRB.DBL PRB.SNG1  
## 1 -81.3393 -32.1591 -47.8026 -64.9647 1.07e-08 1  
## 2 -37.4261 -101.7666 -64.5707 -79.2115 1.00e-08 1  
## 3 -35.0043 -72.6479 -50.8297 -52.6835 2.03e-02 1  
## 4 -31.5489 -80.4661 -49.8506 -58.0733 2.38e-06 1  
## 5 -33.2782 -72.2344 -36.4330 -48.3017 8.06e-03 1  
## 6 -273.8806 -136.9040 -154.9333 -175.5614 3.82e-08 1

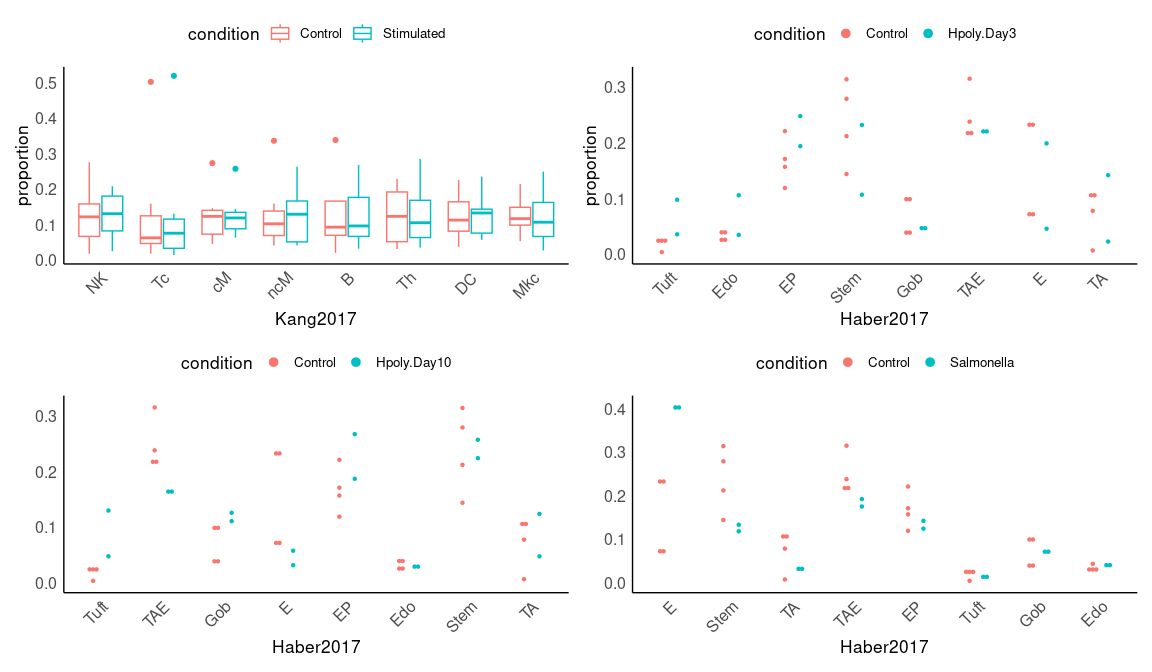
## cell batch  
## 1 AAACATACATTTCC-1 SNG-1016  
## 2 AAACATACCAGAAA-1 SNG-1256  
## 3 AAACATACCATGCA-1 SNG-1488  
## 4 AAACATACCTCGCT-1 SNG-1256  
## 5 AAACATACCTGGTA-1 SNG-1039  
## 6 AAACATACGATGAA-1 SNG-1488

## `summarise()` has grouped output by 'condition', 'clusterRes'. You can override  
## using the `.groups` argument.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | truth | betabin\_null | wtoPhi\_emSVM | estPhi\_null | estPhi\_emSVM | fisher | scDC | speckle | milo\_pct |
| B cells | N | 0.997 | 0.997 | 0.995 | 0.995 | 0.338 | 0.000 | 0.989 | 0.032 |
| CD14+ Monocytes | N | 0.510 | 0.510 | 0.537 | 0.527 | 0.001 | 0.480 | 0.989 | 0.186 |
| CD4 T cells | N | 0.953 | 0.953 | 0.926 | 0.921 | 0.252 | 0.408 | 0.989 | 0.058 |
| CD8 T cells | N | 0.734 | 0.734 | 0.581 | 0.566 | 0.008 | 0.348 | 0.989 | 0.187 |
| Dendritic cells | N | 0.496 | 0.496 | 0.815 | 0.960 | 0.252 | 0.333 | 0.989 | 0.000 |
| FCGR3A+ Monocytes | N | 0.603 | 0.603 | 0.729 | 0.335 | 0.008 | 0.077 | 0.989 | 0.101 |
| Megakaryocytes | N | 0.738 | 0.738 | 0.820 | 0.364 | 0.862 | 0.761 | 0.989 | 0.167 |
| NK cells | N | 0.244 | 0.244 | 0.514 | 0.508 | 0.000 | 0.018 | 0.989 | 0.071 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NK | Tc | cM | ncM | B | Th | DC | Mkc |
| origin | N | N | N | N | N | N | N | N |
| DCATS | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| fisher | \*\*\* | \*\*\* | \*\*\* | \*\*\* | n.s. | n.s. | n.s. | n.s. |
| scDC | \*\* | n.s. | n.s. | \* | \*\*\* | n.s. | n.s. | n.s. |
| speckle | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. | n.s. |
| milo | N | N | N | N | N | N | N | N |

## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.



## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.  
## Bin width defaults to 1/30 of the range of the data. Pick better value with  
## `binwidth`.

## Figure 4

### real-world data 3

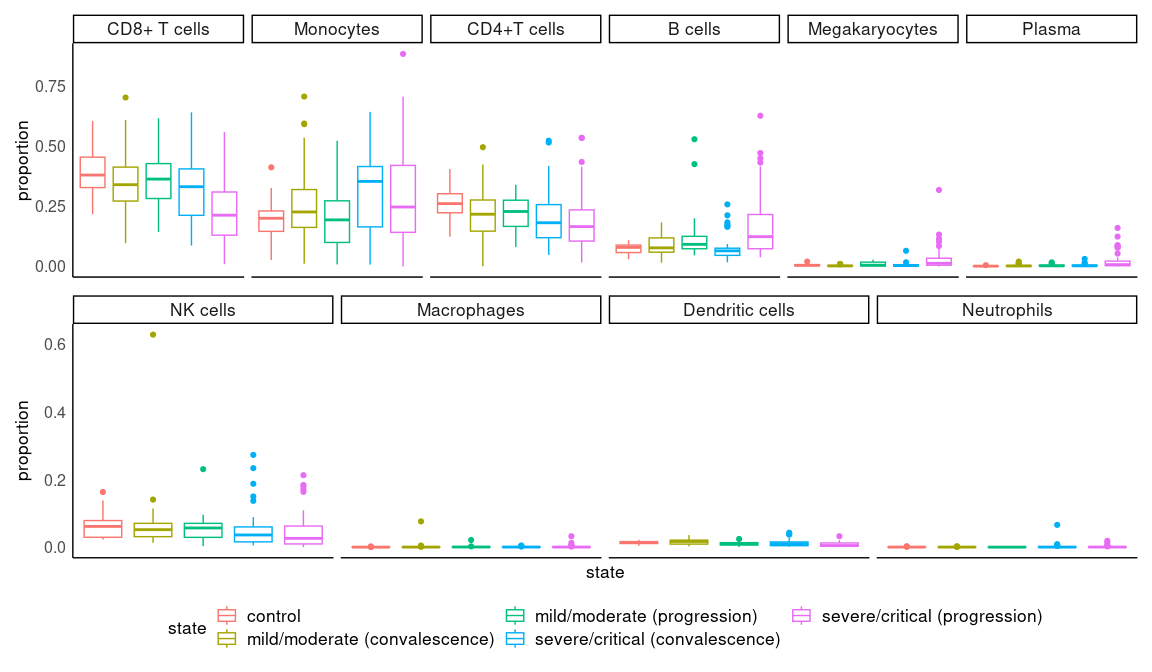
Test between different groups

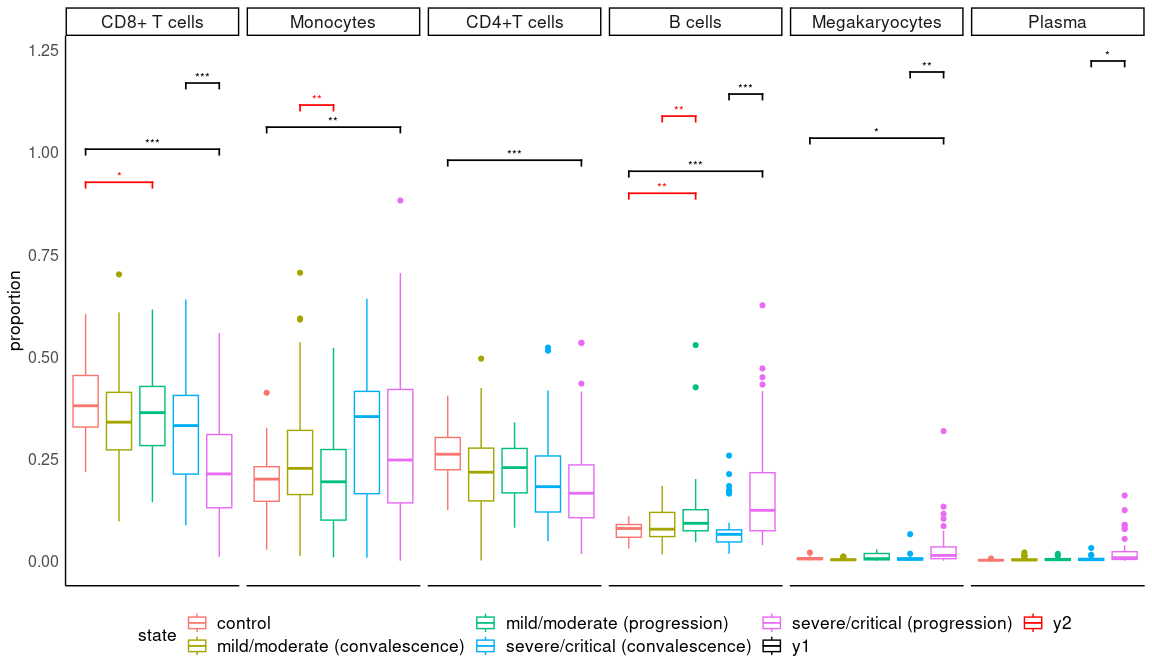
control vs mild/moderate\_progression

control vs severe/critical\_progression

mild/moderate\_convalescence vs mild/moderate\_progression

severe/critical\_convalescence vs severe/critical\_progression





### Real world data 4

## [1] "Empty cell type exists in at least one conidtion; adding replicate & condition specific pseudo count:"

## [1] "Empty cell type exists in at least one conidtion; adding replicate & condition specific pseudo count:"

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | condition | condition | condition | condition |
| Alveolar\_macrophage | 0.0021 | 0.0226 | 0.0849 | 0.1475 |
| B\_cells | 0.0038 | 0.0315 | 0.0068 | 0.0280 |
| Capillary\_endothelial\_cells | 0.0330 | 0.1210 | 0.0141 | 0.0357 |
| Ccl17+/Cd103-/Cd11b-\_dendritic\_cells | 0.6332 | 0.7383 | 0.1099 | 0.1814 |
| Cd103+/Cd11b-\_dendritic\_cells | 0.5535 | 0.7025 | 0.1916 | 0.2634 |
| CD209+/Cd11b+\_dendritic\_cells | 0.4379 | 0.6021 | 0.1819 | 0.2610 |
| Cd4+\_T\_cells | 0.0109 | 0.0514 | 0.0050 | 0.0280 |
| CD8+\_T\_cells | 0.0671 | 0.1533 | 0.0079 | 0.0291 |
| Ciliated\_cells | 0.8631 | 0.8901 | 0.0053 | 0.0280 |
| classical\_monocyte\_(Ly6c2+) | 0.0001 | 0.0020 | 0.0001 | 0.0038 |
| Club\_cells | 0.0660 | 0.1533 | 1.0000 | 1.0000 |
| Eosinophils | 0.0207 | 0.0853 | 0.0121 | 0.0357 |
| Fn1+\_macrophage | 0.3597 | 0.5162 | 0.0341 | 0.0663 |
| Gamma-Delta\_T\_cells | 1.0000 | 1.0000 | 0.1324 | 0.2080 |
| Goblet\_cells | 0.5168 | 0.6821 | 0.0091 | 0.0301 |
| Interstitial\_Fibroblast | 0.0697 | 0.1533 | 0.9195 | 0.9482 |
| Interstitial\_macrophages | 0.0820 | 0.1691 | 0.0035 | 0.0280 |
| Lipofibroblast | 0.3004 | 0.4507 | 0.6183 | 0.6582 |
| low\_quality\_cells | 0.2152 | 0.3738 | 0.0186 | 0.0438 |
| lymphatic\_endothelial\_cells | 0.2415 | 0.3984 | 0.3332 | 0.4229 |
| Megakaryocytes | 0.0019 | 0.0226 | 0.0020 | 0.0280 |
| Mesothelial\_cells | 0.6881 | 0.7383 | 0.0036 | 0.0280 |
| Mki67+\_proliferating\_cells | 0.0556 | 0.1533 | 0.4807 | 0.5665 |
| Natural\_Killer\_cells | 0.6935 | 0.7383 | 0.0209 | 0.0457 |
| Neutrophils | 0.6669 | 0.7383 | 0.0614 | 0.1126 |
| non-classical\_monocyte\_(Ly6c2-) | 0.0051 | 0.0339 | 0.0061 | 0.0280 |
| Plasma\_cells | 0.0090 | 0.0492 | 0.0136 | 0.0357 |
| red\_blood\_cells | 0.2999 | 0.4507 | 0.1496 | 0.2244 |
| Smooth\_muscle\_cells | 0.1801 | 0.3301 | 0.4979 | 0.5666 |
| Type\_2\_pneumocytes | 0.0496 | 0.1533 | 0.2228 | 0.2942 |
| Type1\_pneumocytes | 0.6713 | 0.7383 | 0.4557 | 0.5570 |
| vascular\_endothelial\_cells | 0.0578 | 0.1533 | 0.0222 | 0.0457 |
| Vcam1+\_endothelial\_cells | 0.1540 | 0.2990 | 0.5367 | 0.5904 |

## condition condition condition condition  
## Alveolar\_macrophage TRUE TRUE FALSE FALSE  
## B\_cells TRUE TRUE TRUE TRUE  
## Capillary\_endothelial\_cells TRUE FALSE TRUE TRUE  
## Ccl17+/Cd103-/Cd11b-\_dendritic\_cells FALSE FALSE FALSE FALSE  
## Cd103+/Cd11b-\_dendritic\_cells FALSE FALSE FALSE FALSE  
## CD209+/Cd11b+\_dendritic\_cells FALSE FALSE FALSE FALSE  
## Cd4+\_T\_cells TRUE FALSE TRUE TRUE  
## CD8+\_T\_cells FALSE FALSE TRUE TRUE  
## Ciliated\_cells FALSE FALSE TRUE TRUE  
## classical\_monocyte\_(Ly6c2+) TRUE TRUE TRUE TRUE  
## Club\_cells FALSE FALSE FALSE FALSE  
## Eosinophils TRUE FALSE TRUE TRUE  
## Fn1+\_macrophage FALSE FALSE TRUE FALSE  
## Gamma-Delta\_T\_cells FALSE FALSE FALSE FALSE  
## Goblet\_cells FALSE FALSE TRUE TRUE  
## Interstitial\_Fibroblast FALSE FALSE FALSE FALSE  
## Interstitial\_macrophages FALSE FALSE TRUE TRUE  
## Lipofibroblast FALSE FALSE FALSE FALSE  
## low\_quality\_cells FALSE FALSE TRUE TRUE  
## lymphatic\_endothelial\_cells FALSE FALSE FALSE FALSE  
## Megakaryocytes TRUE TRUE TRUE TRUE  
## Mesothelial\_cells FALSE FALSE TRUE TRUE  
## Mki67+\_proliferating\_cells FALSE FALSE FALSE FALSE  
## Natural\_Killer\_cells FALSE FALSE TRUE TRUE  
## Neutrophils FALSE FALSE FALSE FALSE  
## non-classical\_monocyte\_(Ly6c2-) TRUE TRUE TRUE TRUE  
## Plasma\_cells TRUE TRUE TRUE TRUE  
## red\_blood\_cells FALSE FALSE FALSE FALSE  
## Smooth\_muscle\_cells FALSE FALSE FALSE FALSE  
## Type\_2\_pneumocytes TRUE FALSE FALSE FALSE  
## Type1\_pneumocytes FALSE FALSE FALSE FALSE  
## vascular\_endothelial\_cells FALSE FALSE TRUE TRUE  
## Vcam1+\_endothelial\_cells FALSE FALSE FALSE FALSE