immunoClust - Automated Pipeline for Population Detection in Flow Cytometry

Till Sörensen*

October 14, 2015

Contents

*till-antoni.soerensen@charite.de

1 Licensing

Under the Artistic License, you are free to use and redistribute this software. However, we ask you to cite the following paper if you use this software for publication.

> Sörensen, T., Baumgart, S., Durek, P., Grützkau, A. and Häupl, T.
> immunoClust - an automated analysis pipeline for the identification of immunophenotypic signatures in high-dimensional cytometric datasets.
> *Cytometry A* (accepted).

2 Overview

immunoClust presents an automated analysis pipeline for uncompensated fluorescence and mass cytometry data and consists of two parts. First, cell events of each sample are grouped into individual clusters (cell-clustering). Subsequently, a classification algorithm assorts these cell event clusters into populations comparable between different samples (meta-clustering). The clustering of cell events is designed for datasets with large event counts in high dimensions as a global unsupervised method, sensitive to identify rare cell types even when next to large populations. Both parts use model-based clustering with an iterative Expectation Maximization (EM) algorithm and the Integrated Classification Likelihood (ICL) to obtain the clusters.

The cell-clustering process fits a mixture model with t-distributions. Within the clustering process a optimisation of the $asinh$-transformation for the fluorescence parameters is included.

The meta-clustering fits a Gaussian mixture model for the meta-clusters, where adjusted Bhattacharyya-Coefficients give the probability measures between cell- and meta-clusters.

Several plotting routines are available visualising the results of the cell- and meta-clustering process. Additional helper-routines to extract population features are provided.

3 Getting started

The installation on *immunoClust* is normally done within the Bioconductor.

The core functions of *immunoClust* are implemented in C/C++ for optimal utilization of system resources and depend on the GNU Scientific Library (GSL) and Basic Linear Subprogram (BLAS). When installing *immunoClust* form source using Rtools be aware to adjust the GSL library and include pathes in src/Makevars.in or src/Makevars.win (on Windows systems) repectively to the correct installation directory of the GSL-library on the system.

immunoClust relies on the *flowFrame* structure imported from the *flowCore*-package for accessing the measured cell events from a flow cytometer device.

4 Example Illustrating the immunoClust Pipeline

The functionality of the immunoClust pipeline is demonstrated on a dataset of blood cell samples of defined composition that were depleted of particular cell subsets by magnetic cell sorting. Whole blood leukocytes taken from three healthy individuals, which were experimentally modified by the depletion of one particular cell type per sample, including granulocytes (using CD15-MACS-beads), monocytes (using CD14-MACS-beads), T lymphocytes (CD3-MACS-beads), T helper lymphocytes (using CD4-MACS-beads) and B lymphocytes (using CD19-MACS-beads).

The example datasets contain reduced (10.000 cell-events) of the first Flow Cytometry (FC) sample in `dat.fcs` and the *immunoClust* cell-clustering results of all 5 reduced FC samples for the first donor in `dat.exp`. The full sized dataset is published and available under http://flowrepository.org/id/FR-FCM-ZZWB.

4.1 Cell Event Clustering

```
> library(immunoClust)
```

The cell-clustering is performed by the `cell.process` function for each FC sample separately. Its major input are the measured cell-events in a *flowFrame*-object imported from the *flowCore*-package.

```
> data(dat.fcs)
> dat.fcs
```

```
flowFrame object '2d36b4cf-da0f-4b8d-9a4c-fc7e4f5fccc8'
with 10000 cells and 7 observables:
               name desc  range   minRange maxRange
$P2           FSC-A <NA> 262144    0.00000   262143
$P5           SSC-A <NA> 262144 -111.00000   262143
$P8          FITC-A CD14 262144 -111.00000   262143
$P9            PE-A CD19 262144 -111.00000   262143
$P12          APC-A CD15 262144 -111.00000   262143
$P13      APC-Cy7-A  CD4 262144 -111.00000   262143
$P14 Pacific Blue-A  CD3 262144  -98.93999   262143
171 keywords are stored in the 'description' slot
```

In the `parameters` argument the parameters (named as observables in the `flowFrame`) used for cell-clustering are specified. When omitted all determined parameters are used.

```
> pars=c("FSC-A","SSC-A","FITC-A","PE-A","APC-A","APC-Cy7-A","Pacific Blue-A")
> res.fcs <- cell.process(dat.fcs, parameters=pars)
```

The `summary` method for an *immunoClust*-object gives an overview of the clustering results.

```
> summary(res.fcs)
```

```
** Experiment Information **
Experiment name: immunoClust Experiment
Data Filename:   fcs/12443.fcs
Parameters:   FSC-A SSC-A FITC-A PE-A APC-A APC-Cy7-A Pacific Blue-A
Description:  NA NA CD14 CD19 CD15 CD4 CD3

** Data Information **
Number of observations: 10000
Number of parameters:   7
Removed from above:    318 (3.18%)
Removed from below:    0 (0%)

** Transformation Information **
htrans-A:   0.000000 0.000000 0.009777 0.008345 0.006325 0.010357 0.024461
htrans-B:   0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000
htrans-decade:   -1

** Clustering Summary **
Number of clusters: 13
Cluster     Proportion  Observations
       1      0.038458           377
       2      0.054362           522
       3      0.040146           389
       4      0.013472           123
       5      0.024707           237
       6      0.013874           138
```

```
       7      0.007784            79
       8      0.114307          1108
       9      0.028758           283
      10      0.005154            50
      11      0.014736           140
      12      0.007232            69
      13      0.637007          6167

    Min.      0.005154            50
    Max.      0.637007          6167

** Information Criteria **
Log likelihood: -257037.9 -257215.1 -175889.7
BIC: -257037.9
ICL: -257215.1
```

With the bias argument of the cell.process function the number of clusters in the final model is controlled.

```
> res2 <- cell.process(dat.fcs, bias=0.25)
> summary(res2)

** Experiment Information **
Experiment name: immunoClust Experiment
Data Filename:   fcs/12443.fcs
Parameters:   FSC-A SSC-A FITC-A PE-A APC-A APC-Cy7-A Pacific Blue-A
Description:  NA NA CD14 CD19 CD15 CD4 CD3

** Data Information **
Number of observations: 10000
Number of parameters:   7
Removed from above:    318 (3.18%)
Removed from below:    0 (0%)

** Transformation Information **
htrans-A:   0.000000 0.000000 0.009412 0.006860 0.007320 0.009057 0.024744
htrans-B:   0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000
htrans-decade:   -1

** Clustering Summary **
Number of clusters: 15
Cluster     Proportion  Observations
       1      0.054924           523
       2      0.021434           212
       3      0.007793            72
       4      0.010996           105
       5      0.035356           341
       6      0.007845            77
       7      0.037925           377
       8      0.014764           147
       9      0.094634           923
      10      0.029426           285
      11      0.005168            50
      12      0.011802           114
      13      0.022892           215
      14      0.007213            69
      15      0.637828          6172
```

```
    Min.      0.005168            50
    Max.      0.637828          6172

** Information Criteria **
Log likelihood: -255718.1 -255988.2 -174333.2
BIC: -255718.1
ICL: -255988.2
```

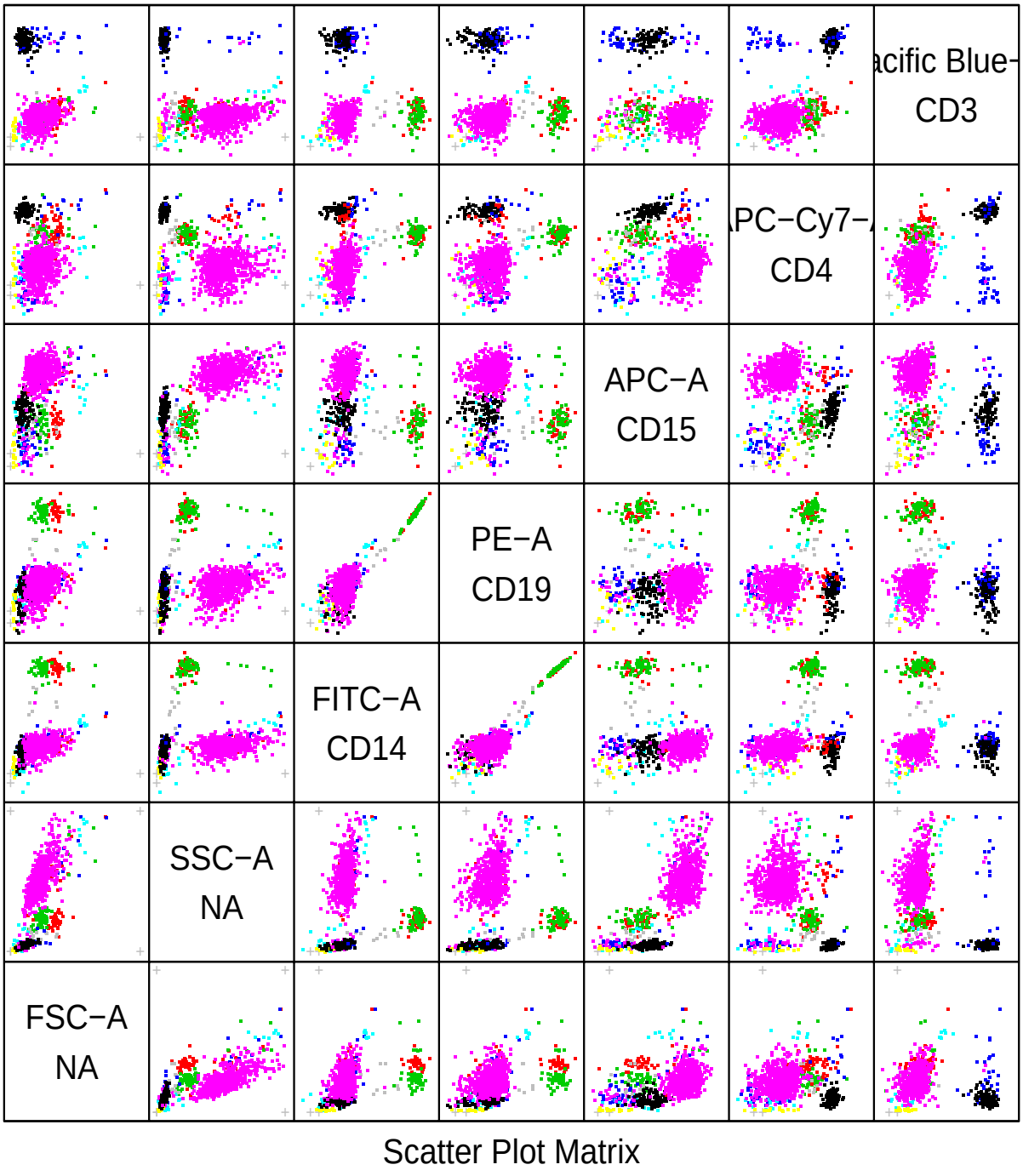
An ICL-bias of 0.3 is reasonable for fluorescence cytometry data based on our experiences, whereas the number of clusters increase dramatically when a `bias` below 0.2 is applied. A principal strategy for the ICL-bias in the whole pipeline is the use of a moderately small `bias` (0.2 - 0.3) for cell-clustering and to optimise the `bias` on meta-clustering level to retrieve the common populations across all samples.

For plotting the clustering results on cell event level, the optimised $asinh$-transformation has to be applied to the raw FC data first.

```
> dat.transformed <- trans.ApplyToData(res.fcs, dat.fcs)
```

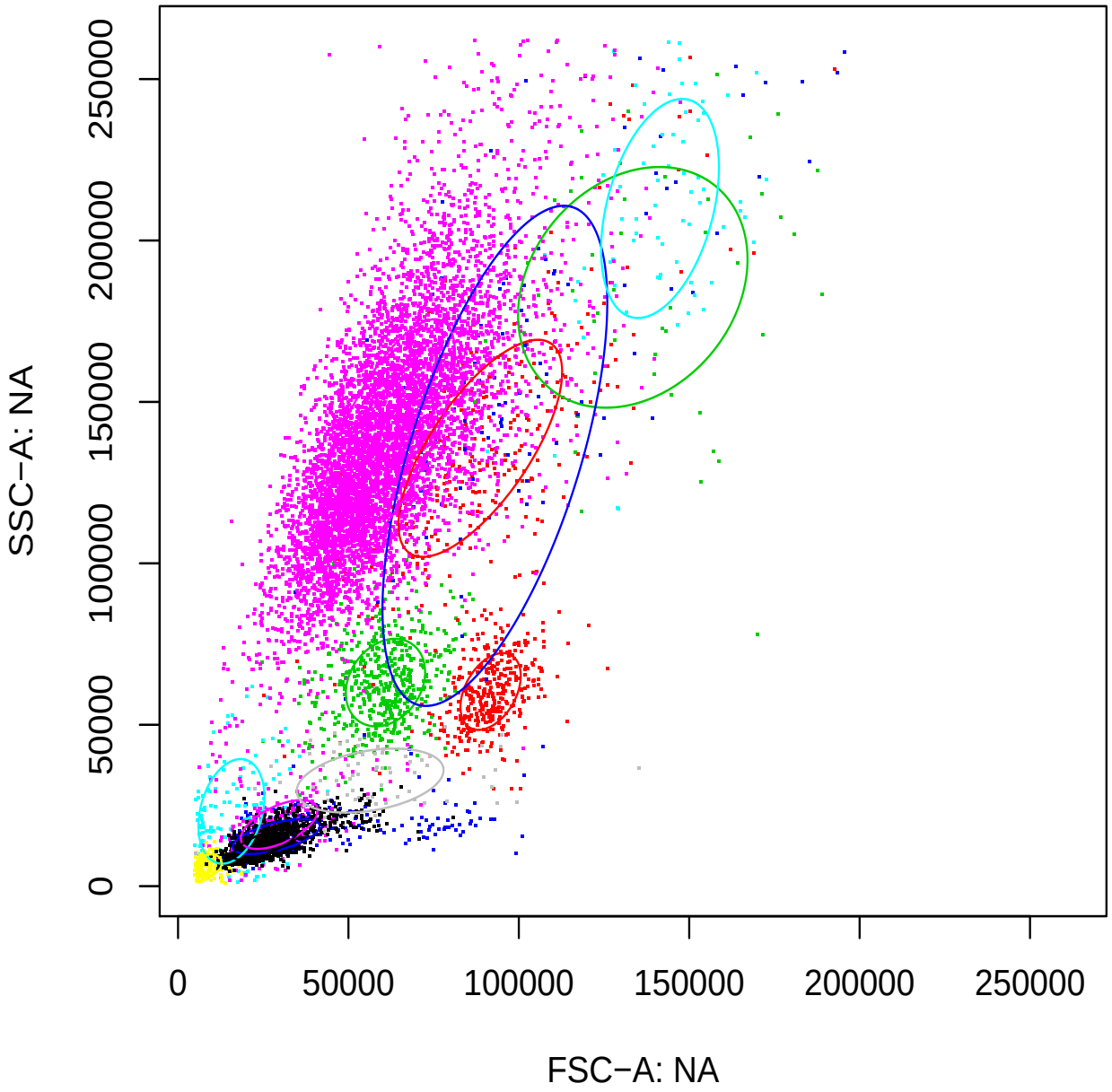
A scatter plot matrix of all used parameters for clustering is obtained by the `splom` method.

```
> splom(res.fcs, dat.transformed, N=1000)
```



For a scatter plot of 2 particular parameters the `plot` method can be used, where parameters of interest are specified in the `subset` argument.

```
> plot(res.fcs, data=dat.transformed, subset=c(1,2))
```



4.2 Meta Clustering

For meta-clustering the cell-clustering results of all FC samples obtained by the `cell.process` function are collected in a `vector` of *immunoClust*-objects and processed by the `meta.process` function.

```
> data(dat.exp)
> meta<-meta.process(dat.exp, meta.bias=0.3)
```

```
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -417035
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -361615
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -356427
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -320472
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -293923
The EM (0) with 6 clusters required 5 iterations, has tolerance 0 and loglike -271096
The EM (0) with 7 clusters required 5 iterations, has tolerance 0 and loglike -265420
The EM (0) with 8 clusters required 5 iterations, has tolerance 0 and loglike -265503
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -361615
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -32433.8
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -31253.1
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -25897.2
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -25897.2
The EM (0) with 4 clusters required 6 iterations, has tolerance 0 and loglike -23543.7
The EM (0) with 5 clusters required 6 iterations, has tolerance 0 and loglike -21419.7
The EM (0) with 6 clusters required 6 iterations, has tolerance 0 and loglike -19893.8
The EM (0) with 7 clusters required 6 iterations, has tolerance 0 and loglike -19773.5
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -329191
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -292848
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -266657
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -243826
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -238149
The EM (0) with 6 clusters required 5 iterations, has tolerance 0 and loglike -238229
The EM (0) with 7 clusters required 5 iterations, has tolerance 0 and loglike -238337
The EM (0) with 8 clusters required 6 iterations, has tolerance 0 and loglike -213524
The EM (0) with 8 clusters required 6 iterations, has tolerance 0 and loglike -268185
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -12148.4
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -10720.9
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -10600.1
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -9441.51
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -8316.22
The EM (0) with 5 clusters required 6 iterations, has tolerance 0 and loglike -8316.22
The EM (0) with 5 clusters required 7 iterations, has tolerance 0 and loglike -8316.22
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -974.599
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -960.715
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -11438.9
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -9714.12
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -8788.02
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -8253.46
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -50670.5
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -47727.5
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -33665.7
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -32811
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -31377.3
The EM (0) with 6 clusters required 5 iterations, has tolerance 0 and loglike -30842.3
The EM (0) with 7 clusters required 5 iterations, has tolerance 0 and loglike -26960.4
The EM (0) with 8 clusters required 5 iterations, has tolerance 0 and loglike -24576.1
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -182391
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -179117
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -179037
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -151027
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -150809
The EM (0) with 6 clusters required 5 iterations, has tolerance 0 and loglike -147446
```

```
The EM (0) with 7 clusters required 5 iterations, has tolerance 0 and loglike -147431
The EM (0) with 8 clusters required 5 iterations, has tolerance 0 and loglike -146420
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -3602.57
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -3625.07
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -3634.4
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -3292.66
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -3291.78
The EM (0) with 6 clusters required 5 iterations, has tolerance 0 and loglike -3213.84
The EM (0) with 7 clusters required 5 iterations, has tolerance 0 and loglike -3054.99
The EM (0) with 8 clusters required 5 iterations, has tolerance 0 and loglike -3077.89
The EM (0) with 18 clusters required 10 iterations, has tolerance 0 and loglike -214875
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -9149.49
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 7 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 8 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 2 clusters required 7 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 2 clusters required 8 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -21428.5
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -17236.9
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -15723.6
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -15784.3
The EM (0) with 4 clusters required 6 iterations, has tolerance 0 and loglike -15184
The EM (0) with 4 clusters required 7 iterations, has tolerance 0 and loglike -15184
The EM (0) with 4 clusters required 8 iterations, has tolerance 0 and loglike -15184
The EM (0) with 4 clusters required 9 iterations, has tolerance 0 and loglike -15184
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -127281
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -123894
```

```
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -108800
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -96963.2
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 6 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 7 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1874.54
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -1633.7
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -1593.52
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -1548.82
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1704.94
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -1732.18
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -749.672
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -764.914
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -704.838
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -249.155
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -267.748
The EM (0) with 20 clusters required 10 iterations, has tolerance 0 and loglike -210630
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -9149.49
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 7 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 8 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 2 clusters required 7 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 2 clusters required 8 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -127281
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -123894
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -108800
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -96963.2
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 6 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 7 iterations, has tolerance 0 and loglike -89054.3
```

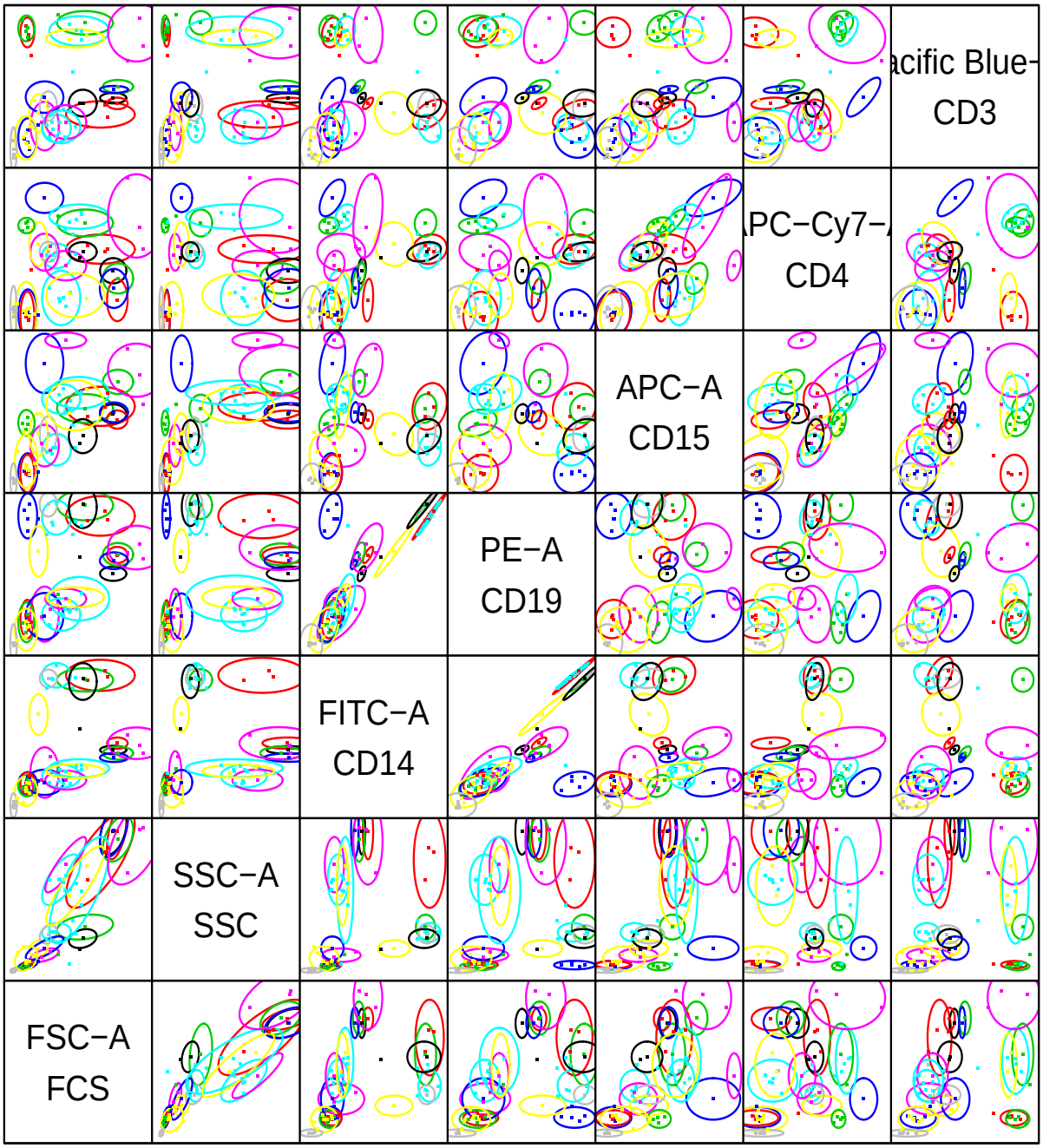
```
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -13329.7
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -13329.7
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -11445.7
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 7 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 8 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 9 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 10 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -730.193
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -679.75
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -902.043
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -834.63
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1704.94
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -1732.18
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -749.672
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -764.914
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -704.838
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -249.155
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -267.748
The EM (0) with 22 clusters required 10 iterations, has tolerance 0 and loglike -210637
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -1068.35
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -8420.49
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -10316.3
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -9149.49
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 7 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 3 clusters required 8 iterations, has tolerance 0 and loglike -7955.91
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -1048.57
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -3727.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -9338.05
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -21864.8
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 2 clusters required 7 iterations, has tolerance 0 and loglike -17696.4
```

```
The EM (0) with 2 clusters required 8 iterations, has tolerance 0 and loglike -17696.4
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -127281
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -123894
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -108800
The EM (0) with 4 clusters required 5 iterations, has tolerance 0 and loglike -96963.2
The EM (0) with 5 clusters required 5 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 6 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 5 clusters required 7 iterations, has tolerance 0 and loglike -89054.3
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -13329.7
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -13329.7
The EM (0) with 2 clusters required 6 iterations, has tolerance 0 and loglike -11445.7
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 7 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 8 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 9 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 3 clusters required 10 iterations, has tolerance 0 and loglike -10806.8
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 6 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 7 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 8 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 9 iterations, has tolerance 0 and loglike -5023.39
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -1704.94
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -1732.18
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 3 clusters required 6 iterations, has tolerance 0 and loglike -1742.32
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -749.672
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -764.914
The EM (0) with 3 clusters required 5 iterations, has tolerance 0 and loglike -704.838
The EM (0) with 1 clusters required 5 iterations, has tolerance 0 and loglike -249.155
The EM (0) with 2 clusters required 5 iterations, has tolerance 0 and loglike -267.748
The EM (0) with 22 clusters required 10 iterations, has tolerance 0 and loglike -210637
The EM (0) with 6 clusters required 4 iterations, has tolerance 0 and loglike -217770
```

The obtained `list`-object contains the meta-clustering result in `$res.clusters`, and the used cell-clusters information in `$dat.clusters`. Additionally a meta-clustering using only the scatter parameter is performed within the `meta.process` function with results in `$res.scatter` and `$dat.scatter`. In a preliminary state of development an automated hierarchical gating on the meta-cluster is performed with results in `$gating`.

A scatter plot matrix of the meta-clustering is again obtained by the `splom` method.

```
> splom(meta$res.clusters, meta$dat.clusters$M, ellipse=TRUE)
```

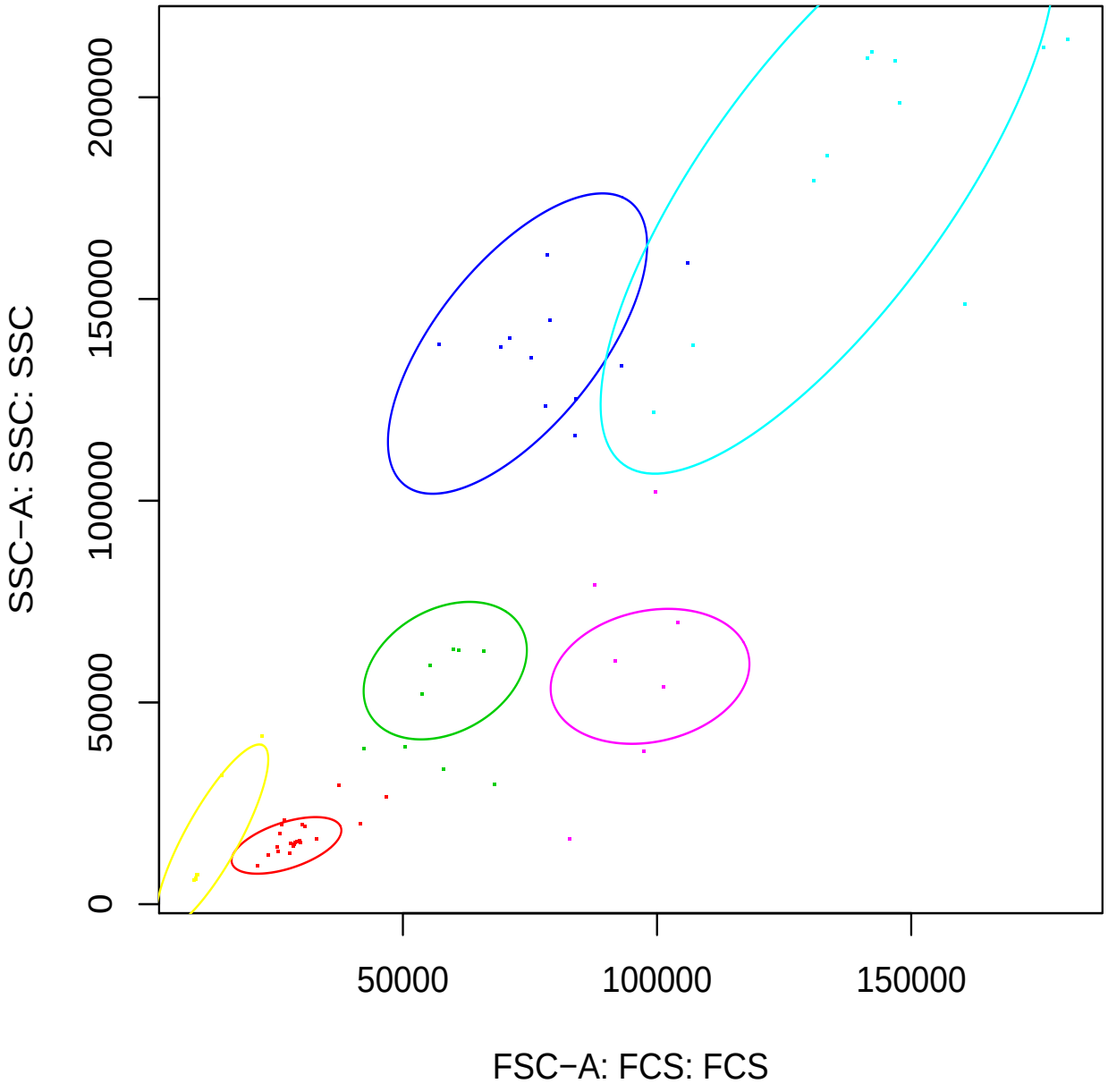


Scatter Plot Matrix

In these scatter plots each cell-cluster is marked by a point of its centre. With the `ellipse=TRUE` argument the meta-clusters are outlined by ellipses of the 90% quantile.

A scatter plot of the scatter parameter distribution of the cell-clusters is obtained by the `plot` method using `$res.scatter`.

```
> plot(meta$res.scatter, data=meta$dat.scatter$M)
```



The scatter distribution is divided into 6 major regions `P1 - P6` with default colors red, green, blue, cyan, magenta and yellow.

The event numbers of each meta-cluster and each sample are extracted in a numeric matrix by the `meta.numEvents` function.

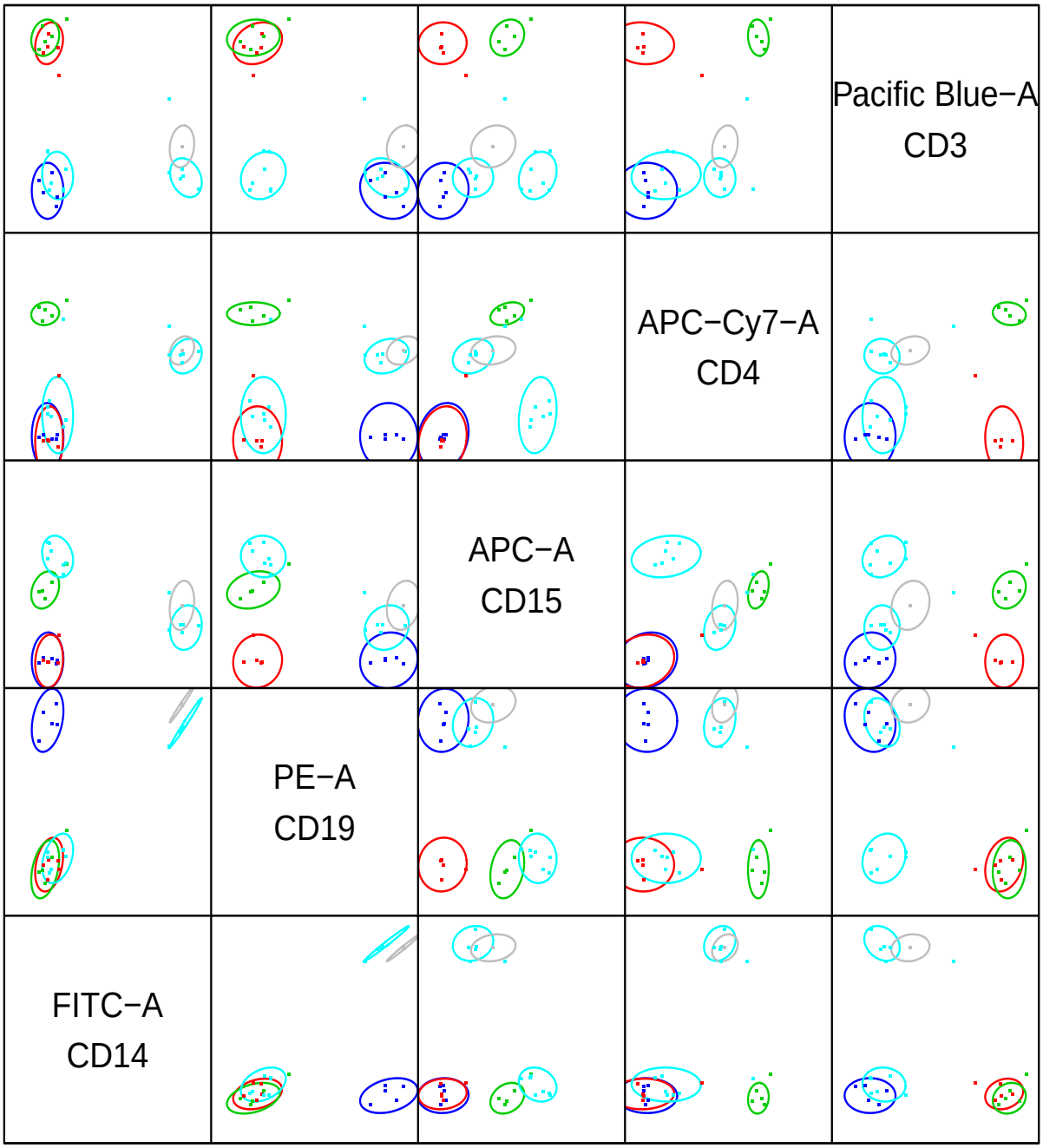
```
> tbl <- meta.numEvents(meta, out.all=FALSE)
> tbl[,1:5]
```

	12543	12546	12549	12552	12555
P1_CD3neg_CD19neg_CD4neg.14.yellow	344	695	780	527	400
P1_CD3neg_CD19neg_CD4pos.5.magenta	71	145	0	0	0
P1_CD3neg_CD19pos.3.blue	0	926	452	331	325
P1_CD3pos_CD4neg.9.red	389	1079	574	433	46
P1_CD3pos_CD4pos.10.green3	1107	3425	1585	0	0
P2_CD14neg.11.blue	0	220	0	0	0
P2_CD14pos_CD3neg.4.cyan	898	0	0	761	950
P2_CD14pos_CD3pos.7.gray	0	1447	0	0	0
P2_CD14dim.6.yellow	0	173	0	0	0
P3_CD3neg_CD15neg.12.cyan	6459	0	5717	7280	7417
P3_CD3neg_CD15pos.13.magenta	0	24	0	0	0
P3_CD3pos_CD4neg.22.yellow	0	0	40	0	0
P3_CD3pos_CD4pos.20.cyan	143	8	199	0	0
P4_CD14neg_CD3neg_CD15neg_CD19neg.16.black	70	0	0	0	0
P4_CD14neg_CD3neg_CD15neg_CD19pos.17.red	0	0	0	95	0
P4_CD14neg_CD3neg_CD15neg_CD19pos.19.blue	0	0	132	0	0
P4_CD14neg_CD3neg_CD15pos.18.green3	0	77	0	0	0
P4_CD14neg_CD3pos.21.magenta	0	103	10	0	0
P4_CD14neg_CD3pos.21.magenta_CD15pos.21.magenta	0	103	10	0	0
P4_CD14pos.1.red	50	0	0	62	94
P5_CD3neg.8.black	0	923	0	0	0
P5_CD3pos.2.green3	0	102	0	0	0
P6.15.gray	151	495	247	247	278

Each row denotes a meta-cluster and each column a data sample used for meta-clustering. The row names give the automated generated gating name, the meta-cluster index and the default color used in the plot routines for each meta-cluster. With an argument `out.all=TRUE` additionally the event numbers in each gating hierarchy level are extracted. In the last columns additionally the meta-cluster centre values in each parameter are given, which helps to identify the meta-clusters. Further export functions retrieve relative cell event frequencies and sample meta-cluster centre values in a particular parameter.

Picking the meta-clusters of the five commonly found population, with respect to the technical depletion the scatter plot matrix reduces to

```
> splom(meta$res.clusters, meta$dat.clusters$M,
+ include=c(3,9,10,4,7,12), subset=3:7, ellipse=TRUE)
```



The whole analysis is performed on uncompensated FC data, thus the high CD19 values on the CD14-population is explained by spillover of FITC into PE. The variation of the CD3 expression in the CD14-population of sample 12546 is caused artificially due to depletion of the granulocytes, which constitute about 60% - 75% of the cells in the other samples.

5 Session Info

The documentation and example output was compiled and obtained on the system:

```
> toLatex(sessionInfo())
```

- R version 3.2.2 (2015-08-14), x86_64-pc-linux-gnu
- Locale: LC_CTYPE=en_US.UTF-8, LC_NUMERIC=C, LC_TIME=en_US.UTF-8, LC_COLLATE=C, LC_MONETARY=en_US.UTF-8, LC_MESSAGES=en_US.UTF-8, LC_PAPER=en_US.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=en_US.UTF-8, LC_IDENTIFICATION=C
- Base packages: base, datasets, grDevices, graphics, grid, methods, stats, utils
- Other packages: flowCore 1.36.0, immunoClust 1.2.0, lattice 0.20-33
- Loaded via a namespace (and not attached): Biobase 2.30.0, BiocGenerics 0.16.0, BiocStyle 1.8.0, DEoptimR 1.0-3, Rcpp 0.12.1, cluster 2.0.3, corpcor 1.6.8, graph 1.48.0, mvtnorm 1.0-3, parallel 3.2.2, pcaPP 1.9-60, robustbase 0.92-5, rrcov 1.3-8, stats4 3.2.2, tools 3.2.2