```
In [1]: gut_samples = [
        ['chic.1', "mgm4440283.3"],
        ['chic.2', "mgm4440284.3"],
        ['cow.1', "mgm4441679.3"],
        ['cow.2', "mgm4441680.3"],
        ['mus.1', "mgm4440463.3"],
        ['mus.2', "mgm4440464.3"],
        ['fish.1', "mgm4441695.3"],
        ['fish.2', "mgm4441696.3"]]

        gut_ids   = slice_column(gut_samples, 1)
        gut_names = slice_column(gut_samples, 0)

        print gut_ids
        print gut_names
```

```
['mgm4440283.3', 'mgm4440284.3', 'mgm4441679.3', 'mgm4441680.3', 'mgm4440463.3',
'mgm4440464.3', 'mgm4441695.3', 'mgm4441696.3']
['chic.1', 'chic.2', 'cow.1', 'cow.2', 'mus.1', 'mus.2', 'fish.1', 'fish.2']
```

```
In [2]: gut_data = analysis.Analysis(gut_ids, 'metagenome', annotation='function', source='Sub
        print gut_data.ids()
```

```
[u'mgm4440283.3', u'mgm4440284.3', u'mgm4440463.3', u'mgm4440464.3', u'mgm4441679.3',
u'mgm4441695.3', u'mgm4441696.3']
```

```
In [3]: gut_data.annotations()[:10]
```

```
Out[3]: [u'(GlcNAc)2_Catabolic_Operon',
        u'16S_rRNA_modification_within_P_site_of_ribosome',
        u'2-Ketogluconate_Utilization',
        u'2-methylcitrate_to_2-methylaconitate_metabolism_cluster',
        u'2-phosphoglycolate_salvage',
        u'4-Hydroxyphenylacetic_acid_catabolic_pathway',
        u'5-FCL-like_protein',
        u'ABC-type_iron_transport_system',
        u'ABC_transporter_[iron.B12.siderophore.hemin]',
        u'ABC_transporter_alkylphosphonate_(TC_3.A.1.9.1)']
```

```
In [4]: print "python sparse matrix"
        print gut_data.matrix[:10]
        print "r dense matrix"
        print gut_data.Rmatrix
        print "r dense normalized matrix"
        print gut_data.NRmatrix
```

```
python sparse matrix
[[0, 0, 6], [0, 1, 7], [0, 2, 8], [0, 3, 6], [0, 4, 33], [0, 5, 20], [1, 0, 224],
[1, 1, 136], [1, 2, 17], [1, 3, 14]]
r dense matrix
         [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
   [1,]    6    7    8    6   33   20    0    0
   [2,]  224  136   17   14   85   47   21   27
   [3,]    0    0    0    0    0    0   44   41
   [4,]    1    0    0    0    0    0    0    0
   [5,]   22   13    4    0   10    6   13    6
   [6,]   18    6    0    0    2    0    4    4
   [7,]  663  423   43   31  317  212  152  124
   [8,]    0    0    0    0    1    1    2    3
   [9,]    1    0    0    0    1    0    4    2
  [10,]   20   23    1    0    3    8   47   44
  [11,]   83   70    3    2   67   59  114  202
  [12,]   33   41    1    4   57   38   32   29
  [13,]    1    0    0    0    0    0    0    0
  [14,]  111  124    6   11  177  113   19   32
  [15,]    1    1    0    0    0    0    7    1
  [16,]    1    1    0    0    0    2    3    0
  [17,]   27   26    3    8   14    4    6    0
  [18,]   85   64    2    2   19    8   18   16
  [19,]    4    5    0    0    1    4    0    0
  [20,]  161  110   15    8  107   93   19   20
  [21,]    0    1    0    0    1    2    8    0
  [22,]  159  146   21   16   90   75   23   24
  [23,]   31   17    5    1   14   12    0    5
  [24,]  144   85    6    4   50   32   25   26
  [25,]  302  187   25   16  165  138   79   79
  [26,]  208  110   16   13  128  104  101   82
  [27,]   89   52    9    1   17   18   10    8
  [28,]    0    1    0    0    0    0    0    0
  [29,]    0    0    1    0    2    1    0    0
  [30,]    1    0    2    0    2    6    2    2
  [31,]   54   44    9    3   30   28    7    1
  [32,]   30    5    2    0    3    6    0    0
  [33,]   26   17    1    0   11    4    0    3
  [34,]  125   54    8    3   26   19    1    0
  [35,]  153  117   10   17  138   95   64   35
  [36,]  147  102   23    8  109  101   48   41
  [37,]    0    0    0    0    0    1    0    0
  [38,]  171   82    3    4   13   13  100  168
  [39,]    2    2    2    2    1    3   33   62
  [40,]   23   11    0    0    3    4   13   11
  [41,]   45   10    2    0   11    8   58   75
  [42,]   48   31    5    1   21   14   51   34
  [43,]    7    8    2    3   16    6    7    9
  [44,]    4    4    0    0    0    0    0    3
  [45,]    0    0    0    0    0    0    0    0
  [46,]    0    0    0    0    1    2    0    0
  [47,]  140   59   10    1   41   35   35   18
  [48,]  371  264   28   24  314  199  157   99
  [49,]    2    1    1    0   20    4    0    0
```

In [5]:
```
apco = gut_data.get_pco(method='bray-curtis', normalize=1)
print apco
```
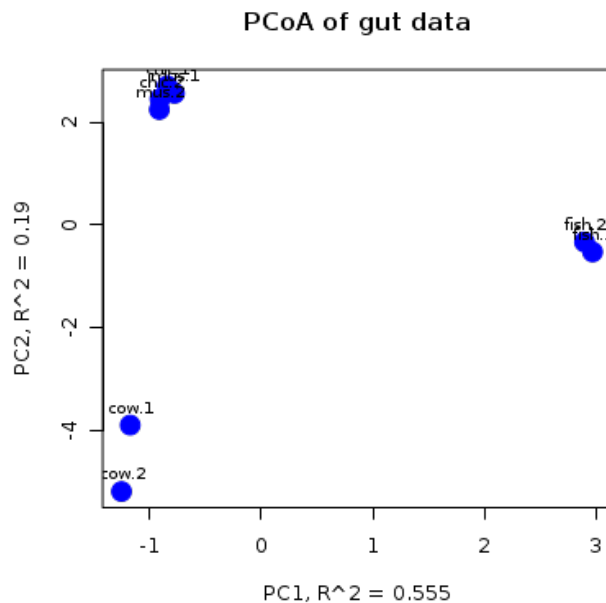
```
$values
 PCO1  PCO2  PCO3  PCO4  PCO5  PCO6  PCO7  PCO8
0.555 0.190 0.086 0.067 0.063 0.022 0.017 0.000

$vectors
        [,1]   [,2]  [,3] [,4]   [,5]    [,6]    [,7] [,8]
[1,] -0.84   2.70 -6.66 -2.3 -0.076   7.18 -15.40 0.35
[2,] -0.91   2.45 -4.21 -3.3 -0.169  -7.99  18.48 0.35
[3,] -1.17  -3.90 -3.27  9.5  2.646  -2.48   0.47 0.35
[4,] -1.25  -5.19  2.48 -8.1 -2.015   1.55  -1.44 0.35
[5,] -0.78   2.56  6.10  1.3 -0.113 -14.41 -10.03 0.35
[6,] -0.91   2.24  5.63  3.0 -0.466  15.79   7.40 0.35
[7,]  2.97  -0.53 -0.84  2.0 -9.708  -0.56   0.26 0.35
[8,]  2.89  -0.34  0.77 -2.1  9.901   0.93   0.26 0.35

$distances
          1     2     3     4     5     6     7
2 0.056
3 0.123 0.120
4 0.139 0.130 0.104
5 0.095 0.084 0.127 0.133
6 0.094 0.086 0.120 0.130 0.058
7 0.185 0.185 0.200 0.207 0.183 0.186
8 0.183 0.182 0.199 0.205 0.179 0.184 0.099
```

In [6]:
```
afile = gut_data.plot_pco(labels=gut_names, title='PCoA of gut data')
Image(filename=afile)
```

Out[6]:

In [7]:
```
afile = gut_data.plot_heatmap(labels=gut_names, title='PCoA of gut data')
Image(filename=afile)
```

```
Loading required package: gplots
Loading required package: gtools

Attaching package: 'gtools'

The following object(s) are masked from 'package:matR':

    permutations

Loading required package: gdata
gdata: read.xls support for 'XLS' (Excel 97-2004) files ENABLED.

gdata: read.xls support for 'XLSX' (Excel 2007+) files ENABLED.

Attaching package: 'gdata'

The following object(s) are masked from 'package:stats':

    nobs

The following object(s) are masked from 'package:utils':

    object.size

Loading required package: caTools
Loading required package: bitops
Loading required package: grid
Loading required package: KernSmooth
KernSmooth 2.23 loaded
Copyright M. P. Wand 1997-2009
Loading required package: MASS

Attaching package: 'gplots'

The following object(s) are masked from 'package:stats':

    lowess

Loading required package: matlab

Attaching package: 'matlab'

The following object(s) are masked from 'package:stats':

    reshape

The following object(s) are masked from 'package:utils':

    find, fix

The following object(s) are masked from 'package:base':

    sum
```

Out[7]:

PCoAofgutdata::ward_clustering

In [7]: