

```

% read in data

%data will be used from 4 young mice and 4 old mice

% reading in data for young mouse 1
genetableY1 = readtable('YoungM1.txt');
genelist = genetableY1(:,1); %list containing gene names
genelist = table2array(genelist);
genetableY1(:,1) = []; % processing of txt file
genetableY1 = table2array(genetableY1);

% reading in data for young mouse 2
genetableY2 = readtable('YoungM2.txt');
genetableY2(:,1) = []; % processing of txt file
genetableY2 = table2array(genetableY2);

% reading in data for young mouse 3
genetableY3 = readtable('YoungM3.txt');
genetableY3(:,1) = []; % processing of txt file
genetableY3 = table2array(genetableY3);

% reading in data for young mouse 4
genetableY4 = readtable('YoungM4.txt');
genetableY4(:,1) = []; % processing of txt file
genetableY4 = table2array(genetableY4);

% combine data to create 1 young mouse dataset
genetableY = cat(2, genetableY1, genetableY2, genetableY3, genetableY4);

%preparing data for old mice samples

%mouse 1
genetable01 = readmatrix('OldM1.txt');
genetable01(:,1) = [];

% reading in data for young mouse 2
genetable02 = readmatrix('OldM2.txt');
genetable02(:,1) = [];

% reading in data for young mouse 3
genetable03 = readmatrix('OldM3.txt');
genetable03(:,1) = [];

% reading in data for young mouse 4
genetable04 = readmatrix('OldM4.txt');
genetable04(:,1) = [];

%combining data from each sample
genetable0 = cat(2, genetable01, genetable02, genetable03, genetable04);

%matrix of both Young and Old samples
genetableBoth = cat(2, genetableY, genetable0);

```

```

% Use cell-type marker genes defined in paper to identify cell types

%oligodendrocyte precursor (OPC)
OPCgene = find(strcmp(genelist, 'Pdgfra'));
OPCindY = find(genetableY(OPCgene,:) > 1);
OPCind0 = find(genetable0(OPCgene,:) > 1);

OPCindY1 = find(genetableY1(OPCgene,:) > 1);
OPCind01 = find(genetable01(OPCgene,:) > 1);

OPCindY2 = find(genetableY2(OPCgene,:) > 1);
OPCind02 = find(genetable02(OPCgene,:) > 1);

OPCindY3 = find(genetableY3(OPCgene,:) > 1);

```

```

OPCind03 = find(genetable03(OPCgene,:)>1);

OPCindY4 = find(genetableY4(OPCgene,:)>1);
OPCind04 = find(genetable04(OPCgene,:)>1);

%oligodendrocyte (OLG)
OLGgene = find(strcmp(genelist,'Cldn11'));
OLGindY = find(genetableY(OLGgene,:)>1);
OLGind0 = find(genetable0(OLGgene,:)>1);

%astrocytes (ASC)
ASCgene = find(strcmp(genelist,'Gja1'));
ASCindY = find(genetableY(ASCgene,:)>1.75);
ASCind0 = find(genetable0(ASCgene,:)>1.75);

%mature neurons (mNEUR)
mNEURgene = find(strcmp(genelist,'Sytn1'));
mNEURindY = find(genetableY(mNEURgene,:)>1);
mNEURind0 = find(genetable0(mNEURgene,:)>1);

%endothelial cells (EC)
ECgene = find(strcmp(genelist,'Cldn5'));
ECindY = find(genetableY(ECgene,:)>1.25);
ECind0 = find(genetable0(ECgene,:)>1.25);

ECindY1 = find(genetableY1(ECgene,:)>1);
ECind01 = find(genetable01(ECgene,:)>1);

ECindY2 = find(genetableY2(ECgene,:)>1);
ECind02 = find(genetable02(ECgene,:)>1);

ECindY3 = find(genetableY3(ECgene,:)>1);
ECind03 = find(genetable03(ECgene,:)>1);

ECindY4 = find(genetableY4(ECgene,:)>1);
ECind04 = find(genetable04(ECgene,:)>1);

%microglia (MG)
MGgene = find(strcmp(genelist,'Tmem119'));
MGindY = find(genetableY(MGgene,:)>1);
MGind0 = find(genetable0(MGgene,:)>1);
MGY = genetableY(:,MGindY);

```

Unrecognized function or variable 'genetable'.
 Index in position 2 exceeds array bounds (must not exceed 1).

```

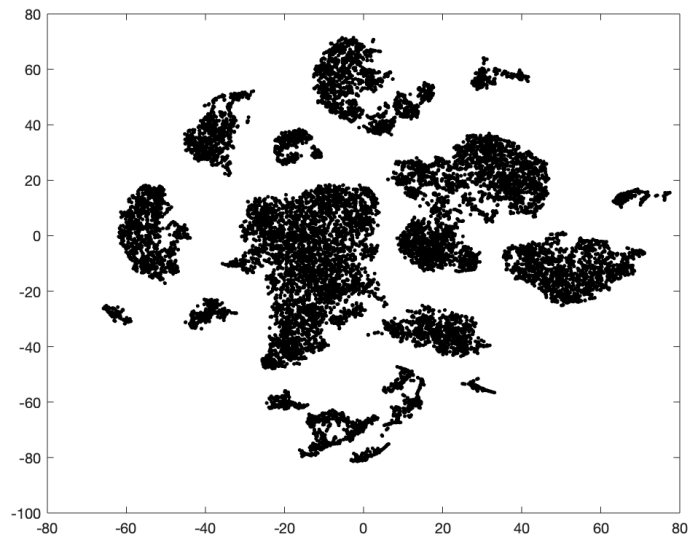
tsne = tsne(genetableBoth', "Algorithm","barneshut", "NumPCAComponents",20);
%tsneOPC01 = tsne(genetable01(:,OPCind01)', "Algorithm","barneshut", "NumPCAComponents",50);

```

```

plot(tsne(:,1),tsne(:,2),'k.')

```



```

hold on
plot(tsne(OPCind0,1),tsne(OPCind0,2), 'r.')
plot(tsne(OPCindY,1),tsne(OPCindY,2), 'r.')

plot(tsne(MGind0,1),tsne(MGind0,2), 'b.')
plot(tsne(MGindY,1),tsne(MGindY,2), 'b.')

plot(tsne(OLGind0,1),tsne(OLGind0,2), 'g.')
plot(tsne(OLGindY,1),tsne(OLGindY,2), 'g.')

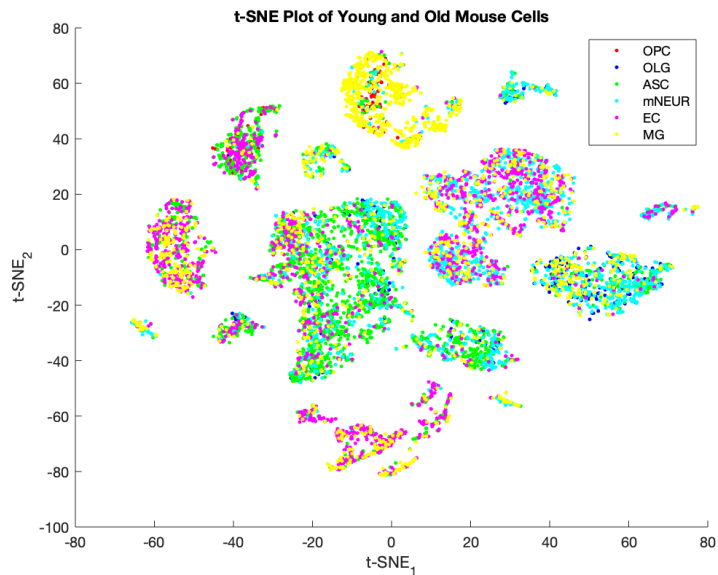
plot(tsne(ASCind0,1),tsne(ASCind0,2), 'c.')
plot(tsne(ASCindY,1),tsne(ASCindY,2), 'c.')

plot(tsne(mNEURind0,1),tsne(mNEURind0,2), 'm.')
plot(tsne(mNEURindY,1),tsne(mNEURindY,2), 'm.')

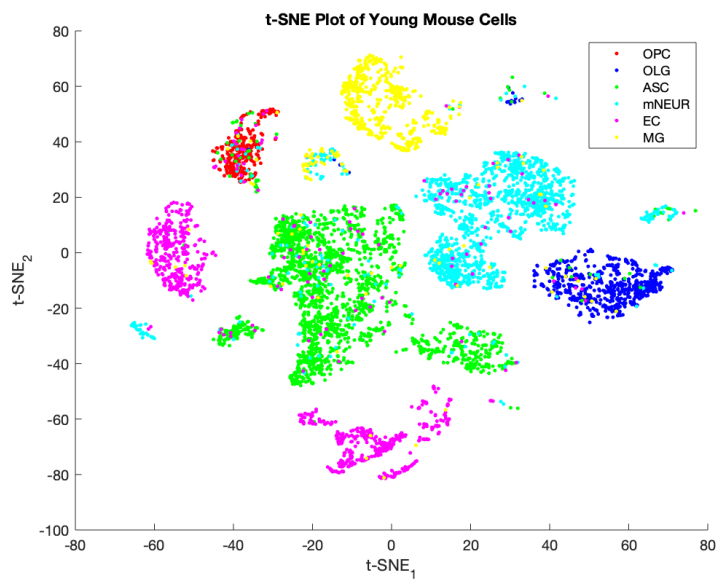
plot(tsne(ECind0,1),tsne(ECind0,2), 'y.')
plot(tsne(ECindY,1),tsne(ECindY,2), 'y.')

ylabel('t-SNE_2')
xlabel('t-SNE_1')
legend('OPC', 'OLG', 'ASC', 'mNEUR', 'EC', 'MG')
title('t-SNE Plot of Young and Old Mouse Cells')
hold off

```



```
hold on
plot(tsne(OPCindY,1),tsne(OPCindY,2),'r.')
plot(tsne(MGindY,1),tsne(MGindY,2),'b.')
plot(tsne(OLGindY,1),tsne(OLGindY,2),'g.')
plot(tsne(ASCindY,1),tsne(ASCindY,2),'c.')
plot(tsne(mNEURindY,1),tsne(mNEURindY,2),'m.')
plot(tsne(ECindY,1),tsne(ECindY,2),'y.')
hold off
ylabel('t-SNE2')
xlabel('t-SNE1')
legend('OPC', 'OLG', 'ASC', 'mNEUR', 'EC', 'MG')
title('t-SNE Plot of Young Mouse Cells')
```



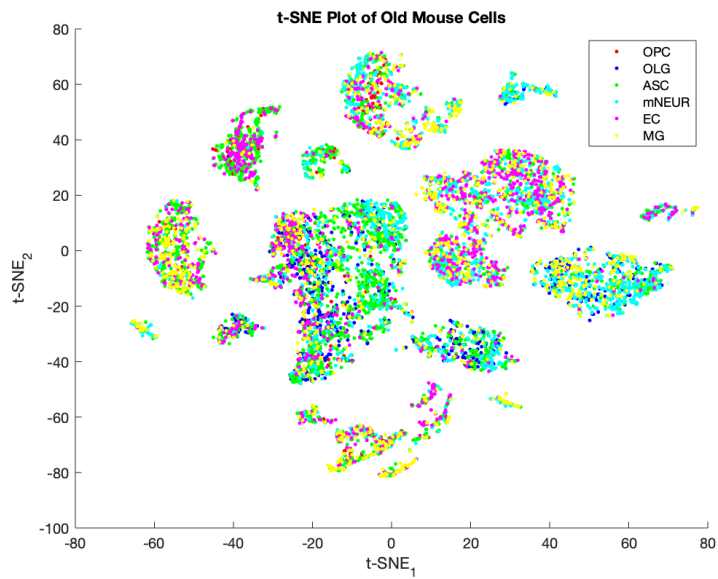
```
hold on
plot(tsne(OPCind0,1),tsne(OPCind0,2), 'r.')
plot(tsne(MGind0,1),tsne(MGind0,2),'b.')
plot(tsne(OLGind0,1),tsne(OLGind0,2),'g.')
plot(tsne(ASCind0,1),tsne(ASCind0,2),'c.')
plot(tsne(mNEURind0,1),tsne(mNEURind0,2),'m.')
plot(tsne(ECind0,1),tsne(ECind0,2),'y.')
hold off
ylabel('t-SNE2')
xlabel('t-SNE1')

```

```

legend('OPC', 'OLG', 'ASC', 'mNEUR', 'EC', 'MG')
title('t-SNE Plot of Old Mouse Cells')

```



```

% analyze cellular composition across young and old groups

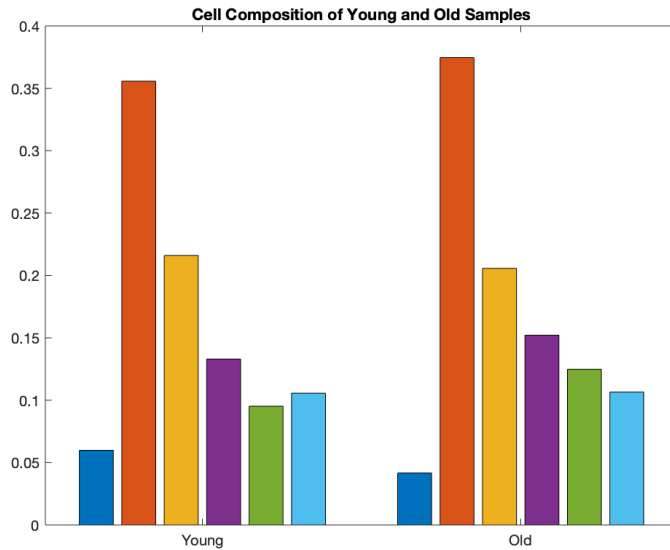
% first aggregate data/calculate cellular compositions
younghist = [length(OPCindY), length(OLGindY), length(ASCindY), length(mNEURindY), length(ECindY), length(MGindY)];
oldhist = [length(OPCind0), length(OLGind0), length(ASCind0), length(mNEURind0), length(ECind0), length(MGind0)];

compY1 = [];
for i=1:6
    new = younghist(i)./7166; %fix this
    compY1 = [compY1, new];
end

comp01 = [];
for i=1:6
    new = oldhist(i)./9471;
    comp01 = [comp01, new];
end

% plot in bar graph
bar([compY1; comp01])
set(gca, 'xticklabel', ["Young"; "Old"])
%legend('OPC', 'OLG', 'ASC', 'mNEUR', 'EC', 'MG')
title('Cell Composition of Young and Old Samples')

```



```
% investigate statistical significance using Wilcoxon rank sum test
```

```
OPCY = [length(OPCindY1)/1370, length(OPCindY2)/1666, length(OPCindY3)/1851, length(OPCindY4)/2279];
OPCO = [length(OPCind01)/1823, length(OPCind02)/1489, length(OPCind03)/3984, length(OPCind04)/2175];
```

```
ECY = [length(ECindY1)/1370, length(ECindY2)/1666, length(ECindY3)/1851, length(ECindY4)/2279];
ECO = [length(ECind01)/1823, length(ECind02)/1489, length(ECind03)/3984, length(ECind04)/2175];
```

```
[p,h,stats] = ranksum(OPCY, OPCO)
```

```
p = 0.0286
```

```
h = logical
```

```
1
```

```
stats = struct with fields:
```

```
ranksum: 26
```

```
[p,h,stats] = ranksum(ECY, ECO)
```

```
p = 0.1143
```

```
h = logical
```

```
0
```

```
stats = struct with fields:
```

```
ranksum: 12
```

```
%Identifying age-related genes
```

```
%MG analysis
```

```
%calculate average gene expression for each cell type
```

```
MGYavg = mean(genetableY(:,MGindY));
```

```
MGOavg = mean(genetable0(:,MGind0));
```

```
%corrcoef(MGYavg, MGOavg)
```

```
%restricting to highly expressed genes
```

```
highMGO = MGOavg > .5; % genes must be expressed above this level
```

```
highMGY = MGYavg > .5;
```

```
highExpMGY = MGYavg(highMGY); %genes we found
```

```
highExpMGO = MGOavg(highMGY); %these are the expression values
```

```
MGhighGenes = genelist(highMGY) %these are the names
```

```
MGhighGenes = 931@1 cell array
```

```
'Tram1'
'Rpl7'
'Tceb1'
'Ogfrl1'
'Ptpn18'
'Cox5b'
'Rpl31'
'Sf3b1'
'Hspe1'
'Ndufb3'
```

```
% identifying restricted genes w/ fold change above 10%
MGfoldchange = highExpMG0./highExpMGY;
MGindstouse = MGfoldchange > 1.1 | MGfoldchange < .909 & MGfoldchange ~= 0;

MGYdatatoUse = highExpMGY(MGindstouse)%gene expression values
```

```
MGYdatatoUse = 10398
    0.6359    2.0590    0.5194    1.1571    2.0144    0.6038    0.7449    0.5081    0.9900    0.5136    0.5274
```

```
MG0datatoUse = highExpMG0(MGindstouse)
```

```
MG0datatoUse = 10398
    0.7531    2.5027    0.5825    1.2764    2.3442    0.6961    1.0489    0.5605    1.3790    0.6023    0.6403
```

```
MGhighGenes = MGhighGenes(MGindstouse) %gene names
```

```
MGhighGenes = 39801 cell array
'Tram1'
'Rpl7'
'Tceb1'
'Ptpn18'
'Rpl31'
'Sf3b1'
'Hspe1'
'Ndufb3'
'Eef1b2'
'Slc11a1'
```

```
MGfoldtoUse = MG0datatoUse./MGYdatatoUse
```

```
MGfoldtoUse = 10398
    1.1844    1.2155    1.1214    1.1031    1.1637    1.1529    1.4080    1.1030    1.3929    1.1727    1.2141
```

```
% sort gene expression values
sorted = sort(MGfoldtoUse, "descend")
```

```
sorted = 10398
    1.6723    1.5801    1.5779    1.4301    1.4204    1.4201    1.4184    1.4080    1.4016    1.3929    1.3575
```

```
inds = [];
for i = 1:length(MGfoldtoUse) %finding names of associated genes
    new = find(MGfoldtoUse==sorted(1,i));
    inds = [inds, new];
end
orderedgenes = MGhighGenes(inds)
```

```
orderedgenes = 39801 cell array
Rows 389:398
'Ly11'
'Jund'
'Bin2'
'Basp1'
'Tm6sf1'
'Klf2'
'Dusp1'
'Cirbp'
'Ppp2ca'
'Pld1'
```

```
highnames = orderedgenes(1:15)
```

```
highnames = 1501 cell array
'Rpl6'
'Klk8'
'PISD'
'Rpl12'
'Tmcc3'
'H2-K1'
'Lyz2'
'Hspe1'
'Herpud1'
'Eef1b2'
```

```
lownames = orderedgenes(length(orderedgenes)-14:length(orderedgenes))
```

```
lownames = 1501 cell array
'Ndrp2'
'Fam212a'
'Cebpb'
'Rpl13-ps3'
'Mt3'
'Lyl1'
'Jund'
'Bin2'
'Basp1'
'Tm6sf1'
```

```
names = cat(1,highnames,lownames)
```

```
names = 3001 cell array
'Rpl6'
'Klk8'
'PISD'
'Rpl12'
'Tmcc3'
'H2-K1'
'Lyz2'
'Hspe1'
'Herpud1'
'Eef1b2'
```

```
% make a heatmap of top 15 highest and lowest expressed genes
heatmapHi = sorted(1:15);
heatmapLow = sorted(length(sorted)-14:length(sorted));
heatmptotal = cat(2,heatmapHi,heatmapLow);
heatmptotal = log10(heatmptotal)
```

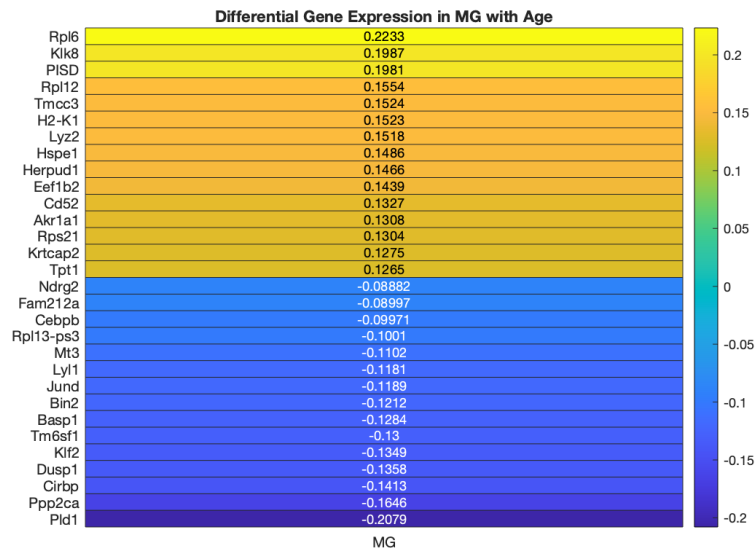
```
heatmptotal = 1030
    0.2233    0.1987    0.1981    0.1554    0.1524    0.1523    0.1518    0.1486    0.1466    0.1439    0.1327
```

```
h = heatmap(heatmptotal')
```

```
h =
    HeatmapChart with properties:
        XData: {'1'}
        YData: {3001 cell}
        ColorData: [3001 double]
```

```
Show all properties
```

```
title("Differential Gene Expression in MG with Age");
h.XDisplayLabels = {'MG'};
h.YDisplayLabels = names;
h.ColorScaling = 'scaled';
h.Colormap = parula
```

```
h =
HeatmapChart (Differential Gene Expression in MG with Age) with properties:
```

```
    XData: {'1'}
    YData: {3001 cell}
    ColorData: [3001 double]
```

Show all properties

```
find(orderedgenes == "Csflr")
```

```
ans = 356
```

```
find(genelist == "Ctss")
```

```
ans = 2950
```

```
%dataToUse_MG = MGY1avg(indstouse)
```

```
%Analysis of age-related gene changes across cell-types
```

```
%finding most up/downregulated genes across all cell types
```

```
%calculate average gene expression of each cell type
```

```
OPCYavg = mean(genetableY(:,OPCindY));
OPC0avg = mean(genetable0(:,OPCind0));
```

```
OLGYavg = mean(genetableY(:,OLGindY));
OLG0avg = mean(genetable0(:,OLGind0));
```

```
ASCYavg = mean(genetableY(:,ASCindY));
ASC0avg = mean(genetable0(:,ASCind0));
```

```
mNEURYavg = mean(genetableY(:,mNEURindY));
mNEURO0avg = mean(genetable0(:,mNEURind0));
```

```
ECYavg = mean(genetableY(:,ECindY));
EC0avg = mean(genetable0(:,ECind0));
```

```
%average gene expression for all cells
```

```
Yavg = [OPCYavg;OLGYavg;ASCYavg;mNEURYavg;ECYavg;MGYavg];
Oavg = [OPC0avg;OLG0avg;ASC0avg;mNEURO0avg;EC0avg;MG0avg];
meanYavg = mean(Yavg);
meanOavg = mean(Oavg);
```

```
%restricting to highly expressed genes
```

```

highind0 = mean0avg > .5; % genes must be expressed above this level
highindY = meanYavg > .5;
highExpY = meanYavg(highindY); %genes we found
highExp0 = mean0avg(highindY); %these are the expression values
highGenesAll = genelist(highindY) %these are the names

```

```

highGenesAll = 100201 cell array
'Rpl7'
'Tceb1'
'Ptp4a1'
'Cox5b'
'Rpl31'
'Pantr1'
'Sf3b1'
'Hspe1'
'Ndufb3'
'Sumo1'

```

```
TotalFoldChange = highExp0./highExpY %this one doesn't
```

```

TotalFoldChange = 101002
    1.1429    1.0198    1.0206    0.9855    1.1245    0.8769    1.0961    1.1953    1.0674    0.9514    1.1888

```

```
allsorted = sort(TotalFoldChange, 'ascend');
```

```

lowest1 = find(TotalFoldChange==allsorted(1,1));
lowest2 = find(TotalFoldChange==allsorted(1,2));
lowest3 = find(TotalFoldChange==allsorted(1,3));
highest1 = find(TotalFoldChange==allsorted(1,length(allsorted)-1));
highest2 = find(TotalFoldChange==allsorted(1,length(allsorted)-2));
highest3 = find(TotalFoldChange==allsorted(1,length(allsorted)-3));
highGenesAll(lowest1)

```

```

ans = 101 cell array
{'Basp1'}

```

```
highGenesAll(lowest2)
```

```

ans = 101 cell array
{'Rpl13-ps3'}

```

```
highGenesAll(lowest3)
```

```

ans = 101 cell array
{'Ppp2ca'}

```

```
highGenesAll(highest1)
```

```

ans = 101 cell array
{'Rpl6'}

```

```
highGenesAll(highest2)
```

```

ans = 101 cell array
{'Rpl12'}

```

```
highGenesAll(highest3)
```

```

ans = 101 cell array
{'Ptges3'}

```

```
% genes to validate
```

```

% upregulated from paper
B2mind = find(genelist=="B2m");
Rpl6ind = find(genelist=="Rpl6");
Malat1ind = find(genelist=="Malat1");

```

```

%downregulated from paper
mtND1ind=find(genelist=="mt-Nd1");

```

```

Aldocind=find(genelist=="Aldoc");
Sepw1ind=find(genelist=="Sepw1");

%traditional aging marker
Ctssind = find(genelist=="Ctss");

%newly discovered marker in paper
Apoc1ind = find(genelist=="Apoc1");
Calyind = find(genelist=="Caly");

%bidirectional genes
Cd9ind = find(genelist=="Cd9");
Rpl37aind = find(genelist=="Rpl37a");

%genes we found most up/downregulated
Basp1ind = find(genelist=="Basp1");
Rpl12ind = find(genelist=="Rpl12");

```

```

highGenesAll = 100201 cell array

```

```

'Rpl7'
'Tceb1'
'Ptp4a1'
'Cox5b'
'Rpl31'
'Pantr1'
'Sf3b1'
'Hspe1'
'Ndufb3'
'Sumo1'

```

```

TotalFoldChange = 101002

```

```

1.1429 1.0198 1.0206 0.9855 1.1245 0.8769 1.0961 1.1953 1.0674 0.9514 1.1888

```

```

ans = 101 cell array

```

```

{'Basp1'}

```

```

ans = 101 cell array

```

```

{'Rpl13-ps3'}

```

```

ans = 101 cell array

```

```

{'Ppp2ca'}

```

```

ans = 101 cell array

```

```

{'Rpl6'}

```

```

ans = 101 cell array

```

```

{'Rpl12'}

```

```

ans = 101 cell array

```

```

{'Ptges3'}

```

```

allfoldchange = 0avg./Yavg;
%find genes of interest
B2m = allfoldchange(:,B2mind);
Rpl6 = allfoldchange(:,Rpl6ind);
Malat1 = allfoldchange(:,Malat1ind);
mtND1 = allfoldchange(:,mtND1ind);
Aldoc = allfoldchange(:,Aldocind);
Sepw1 = allfoldchange(:,Sepw1ind);
Ctss = allfoldchange(:,Ctssind);
Caly = allfoldchange(:,Calyind);
Cd9 = allfoldchange(:,Cd9ind);
Rpl37a = allfoldchange(:,Rpl37aind);
Basp1 = allfoldchange(:,Basp1ind);
Rpl12 = allfoldchange(:,Rpl12ind);

total = [B2m,Rpl6,Rpl12, Malat1, Ctss,Caly, mtND1,Aldoc,Sepw1, Cd9, Rpl37a,Basp1];
total = log10(total)

```

```

total = 6012

```

```
0.0739 0.1007 0.1107 0.0236 0.1321 0.0316 -0.0101 0.0326 -0.0303 -0.0510 -0.0181
0.1627 0.1642 0.1178 0.0029 0.2893 0.0173 -0.0093 -0.0557 -0.0196 -0.0342 0.0230
0.1053 0.1642 0.1062 0.0140 0.1890 0.1728 -0.0177 -0.0301 -0.0313 -0.0636 0.0204
0.0534 0.1729 0.1235 0.0176 0.3916 -0.0032 -0.0220 -0.0572 -0.0178 -0.0731 0.0785
-0.0137 0.0984 0.0753 0.0136 0.0760 0.0747 -0.0071 -0.0565 -0.0457 0.1340 0.0098
0.0303 0.2233 0.1554 0.0051 -0.0129 -0.0829 -0.0218 -0.0379 -0.0272 0.0257 0.0353
```

```
h2 = heatmap(total')
```

```
h2 =
```

```
HeatmapChart with properties:
```

```
  XData: {601 cell}
  YData: {1201 cell}
  ColorData: [1206 double]
```

```
Show all properties
```

```
title("Validation of Selected Aging Signatures");
h2.XDisplayLabels = {'OPC','OLG','ASC','mNEUR','EC','MG'};
h2.YDisplayLabels = {'B2m','Rpl6','Rpl12','Malat1','Ctss','Caly','mt-Nd1','Aldoc','Sepw1','Cd9','Rpl37a','Basp1'};
h2.ColorScaling = 'scaled';
h2.Colormap = parula
```



```
h2 =
```

```
HeatmapChart (Validation of Selected Aging Signatures) with properties:
```

```
  XData: {601 cell}
  YData: {1201 cell}
  ColorData: [1206 double]
```

```
Show all properties
```

```
genetableY1(high_indY1)
```

```
ans = 15969701
1.3403
1.3403
1.8931
1.3403
1.3403
1.3403
1.8931
1.8931
1.3403
1.3403
1.3403
```

```
genetable01(high_ind01)
```

```
ans = 60540601
    1.9902
        0
        0
        0
    1.9902
    1.9902
        0
        0
    1.9902
        0
```

```
MG_high = cat(2, genetableY1(high_indY1),genetable01(high_ind01))
```

Error using cat
Dimensions of arrays being concatenated are not consistent.

```
MG_Y1vs01 = tsne(MG_high, "Algorithm","barneshut","NumPCAComponents",50)
```

```
sce=SingleCellExperiment(genetableY1,genelist)
```

```
sce =
SingleCellExperiment with properties:
                X: [1469901370 double]
                g: [1469901 string]
                s: [137003 double]
                c: [137001 double]
    c_cell_cycle_tx: []
    c_cell_type_tx: []
    c_cluster_id: []
    c_batch_id: []
    c_cell_id: [137001 double]
    list_cell_attributes: {}
    list_gene_attributes: {}
    struct_cell_embeddings: [101 struct]
    struct_cell_clusterings: [101 struct]
    table_attributes: [000 table]
    NumCells: 1370
```

```
sc_scatter(sce);
```

Error using uipushtool
First argument must be a valid parent, such as a Figure or Panel object.

Error in sc_scatter_sce/i_addbutton (line 86)
pt3 = uipushtool(barhandle, 'Separator', septag);

Error in sc_scatter_sce (line 130)
i_addbutton(2,0,@call_scscatter,"IMG00107.GIF","New SC_SCATTER")

Error in sc_scatter (line 252)
sc_scatter_sce(X);

```
remove = [];
counter = 0;
for column = 1:1371
    for row = 1:14699
        if genetable(row, column) > 0
            counter = counter + 1;
        end
    end
    if counter < 250
        remove = [remove, column];
    end
    counter = 0;
end
```

```
processed = genetable;
for i = 1:323
    processed(:, remove(1,i)) = [];
end
```

