

```
(* © Orly Alter 2022, All Rights Reserved *)
```

```
(* BME 6770: Genomic Signal Processing *)
```

```
(* The Pseudoinverse Projection of Measured Data *)
```

```
(* Initialize *)
```

```
Clear["Global`*"]
```

```
(* The SVD of Cell Cycle mRNA Expression *)
```

```
(* Read Cell Cycle Expression Data *)
```

```
a = 1;
```

```
b = 7;
```

```
matrix =
```

```
  Import["https://alterlab.org/network_decomposition/data/Cell_Cycle_Expression.txt", "Table"];
```

```
header = matrix[[1]];
```

```
matrix = Drop[matrix, {1}];
```

```
{genes, arrays} = Dimensions[matrix] - {a, b};
```

```
matrix =
```

```
  Sort[Transpose[Join[{Table[Count[matrix[[c]], "Null"], {c, 1, genes + a}], Transpose[matrix]]]]];
```

```
pos = Position[matrix[All, 1], 1][[1, 1]];
```

```
matrix = Join[{header}, Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes + a}], {1}]]]]];
```

```
{genes, arrays} = Dimensions[matrix] - {a, b}
```

```
{2493, 18}
```

```
arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}], {1, b}]]];
```

```
matrix = Drop[matrix, {1}];
```

```
matrix = Transpose[matrix];
```

```
genenames = Drop[matrix, {b + 1, b + arrays}];
```

```
matrix = Drop[matrix, {1, b}];
```

```
matrix = Transpose[matrix];
```

```
Clear[a, b];
```

```
(* Compute the SVD *)
```

```
{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
```

```
eigenexpressions = Diagonal[eigenexpressions];
```

```
eigengenes = Transpose[eigengenes];
```

```
rank = arrays - Count[eigenexpressions, 0.]
```

```
fractions = eigenexpressions^2 / Sum[eigenexpressions[[a]]^2, {a, 1, rank}];
```

```
entropy = -N[Sum[fractions[[a]] * Log[fractions[[a]], {a, 1, rank}] / Log[rank]];
```

```
entropy = N[Round[100 * entropy] / 100]
```

```
18
```

```
0.16
```

```
(* 1. What can we learn from the number of rank of freedom about the data? *)
```

```
(* 2. What can we learn from the entropy about the data? *)
```

(* Create Fractions Bar Chart Display *)

```
gridx = Table[a, {a, 0.25, 0.75, 0.25}];
framex = Table[{gridx[[a]], Rotate[gridx[[a]], Pi / 2]}, {a, 1, 3}];
gridx = Table[{gridx[[a]], RGBColor[0, 0, 0]}, {a, 1, 3}];
framey = Table[{a + 1, rank - a}, {a, 0, rank - 1}];
labelx = ColumnForm[
  {"(b) Eigenexpression Fraction", StringJoin["d = ", ToString[entropy]]},
  Center];
g2 = Show[BarChart[
  Table[fractions[[rank - a]], {a, 0, rank - 1}],
  BarOrigin → Left,
  PlotRange → {{0.0001, 0.9999}, {0.5, rank + 0.5}},
  AspectRatio → 1,
  ChartStyle → Red],
  Axes → False,
  Frame → True,
  FrameTicks → {None, framey, framex, None},
  FrameLabel → {None, None, labelx, None},
  GridLines → {gridx, None},
  BaseStyle → {FontFamily → "Courier"}];
```

(* Create Eigengenes Raster Display *)

```
contrast = 3.5;
displaying = Table[
  If[contrast * eigengenes[[i, j]] > 0,
    If[contrast * eigengenes[[i, j]] < 1, {contrast * eigengenes[[i, j]], 0}, {1, 0}],
    If[contrast * eigengenes[[i, j]] > -1, {0, -contrast * eigengenes[[i, j]]}, {0, 1}]],
  {i, 1, rank}, {j, 1, arrays}];
framex = Table[{a - 0.5, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = Table[{a + 1 - 0.5, rank - a}, {a, 0, rank - 1}];
labely = "Eigengenes";
labelx = "(a) Arrays";
g1 = Show[
  Graphics[
    Raster[
      Table[
        {displaying[[i, j, 1]], displaying[[i, j, 2]], 0},
        {i, rank, 1, -1}, {j, 1, arrays}]]],
  AspectRatio → 1,
  Frame → True,
  FrameTicks → {None, framey, framex, None},
  FrameLabel → {None, labely, labelx, None},
  BaseStyle → {FontFamily → "Courier"}];
```

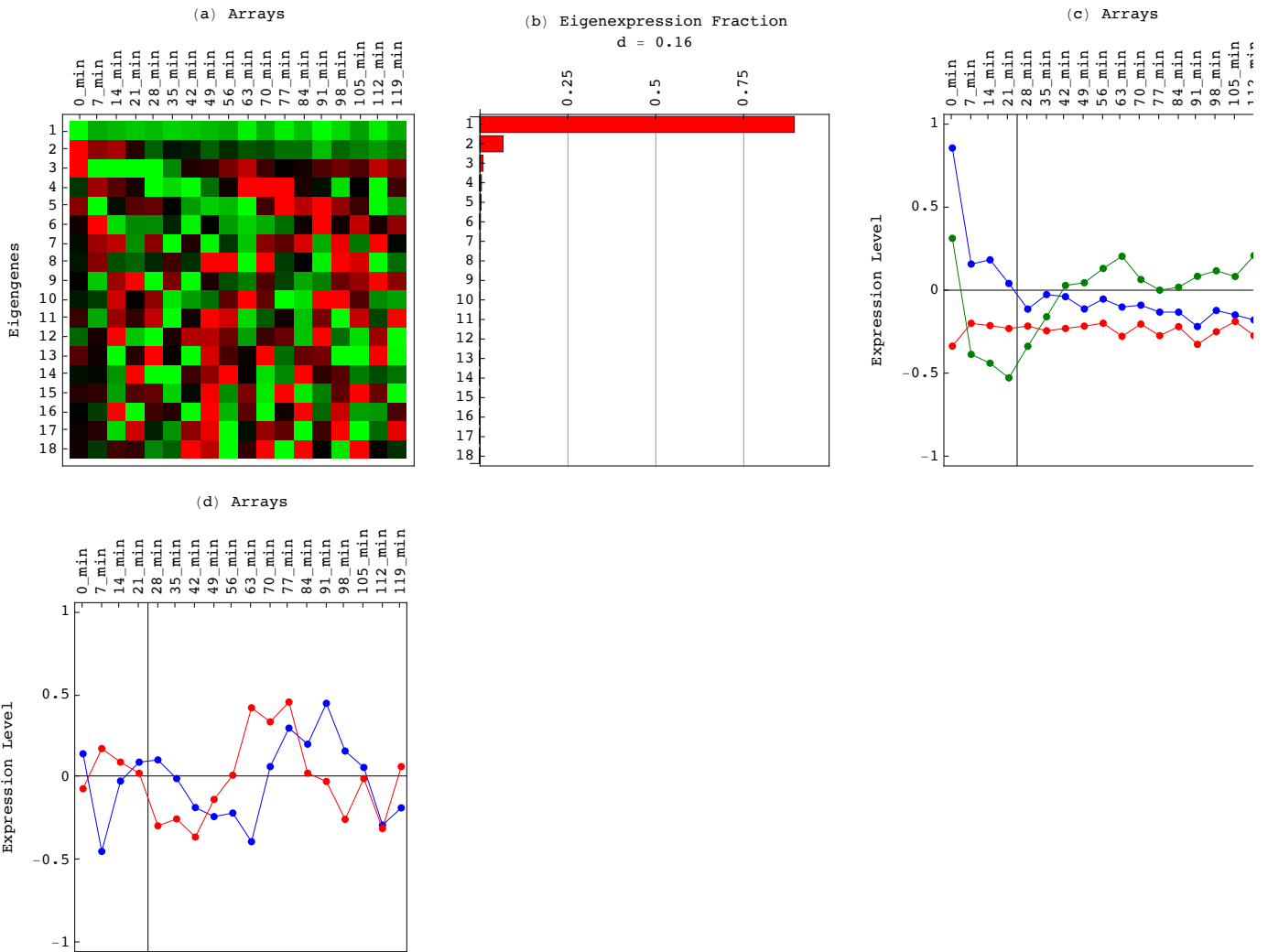
(* Create Selected Eigengenes Graph Display *)

```
p = Table[0, {n, 1, 5}];
color =
  {RGBColor[1, 0, 0], RGBColor[0, 0, 1], RGBColor[0, 0.5, 0], , RGBColor[1, 0.5, 0], RGBColor[1, 0, 0]};
labelx = "(c) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n, a]]}, {a, 1, arrays}],
  points = Table[Point[coordinates[[a]]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
    Frame → True,
    FrameLabel → {None, labely, labelx, None},
    GridLines → {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}}},
    FrameTicks → {None, framey, framex, None},
    BaseStyle → {FontFamily → "Courier"}],
  p[[n]] = Show[g,
    AspectRatio → 1.05,
    PlotRange → {-1.05, 1.05},
    DisplayFunction → Identity]],
  {n, 1, 3}]
g3 = Show[{p[[3]], p[[2]], p[[1]]}];

p = Table[0, {n, 1, 2}];
color = {RGBColor[1, 0, 0], RGBColor[0, 0, 1]};
labelx = "(d) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n + 3, a]]}, {a, 1, arrays}],
  points = Table[Point[coordinates[[a]]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
    Frame → True,
    FrameLabel → {None, labely, labelx, None},
    GridLines → {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}}},
    FrameTicks → {None, framey, framex, None},
    BaseStyle → {FontFamily → "Courier"}],
  p[[n]] = Show[g,
    AspectRatio → 1.05,
    PlotRange → {-1.05, 1.05},
    DisplayFunction → Identity]],
  {n, 1, 2}]
g4 = Show[{p[[2]], p[[1]]}];
```

(* Display the SVD *)

```
Show[GraphicsGrid[{{g1, g2, g3}}, Spacings -> -20],
ImageSize -> 750]
Show[GraphicsGrid[{{g4}},
ImageSize -> 250]
```



(* 3. What do the top three eigengenes look like? *)

(* 4. What biological processes might they describe? *)

(* Compute P-Value of Enrichment of Annotations Assuming Hypergeometric Distribution *)

```
headerx = {{
ColumnForm[{" ", " ", "Eigenarray"}, Left],
ColumnForm[{"Most Likely", "Parallel", "Association"}, Left],
ColumnForm[{"P-Value of", "Parallel", "Association"}, Left],
ColumnForm[{"Most Likely", "Antiparallel", "Association"}, Left],
ColumnForm[{"P-Value of", "Antiparallel", "Association"}, Left]}}
```

(* Cell Cycle Annotations *)

```

annotations = genenames[[6]];
stages = {"M/G1", "G1", "S", "S/G2", "G2/M", "None"};

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}}]]
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

```

{53, 117, 32, 63, 91, 2137}

```

Do[{
  pattern = Transpose[Sort[
    TransposeJoin[{Transpose[eigenarrays][[c]], {annotations}}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
    probability = Table[{
      Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
        Binomial[genes, most]], {b, table[[a, 3]], most}],
      stages[[a]],
      {a, 1, Dimensions[stages][[1]]},
      parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
      parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]}],
      table = Table[{
        stages[[a]],
        numbers[[a]],
        Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
        {a, 1, Dimensions[stages][[1]]},
        probability = Table[{
          Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
            Binomial[genes, most]], {b, table[[a, 3]], most}],
          stages[[a]],
          {a, 1, Dimensions[stages][[1]]},
          antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
          antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]}],
      {c, 1, Dimensions[counter][[1]]}
    }
  ]
}

```

(* Display the Enrichments *)

```

table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]

```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	M/G1	4.5×10^{-5}	G2/M	1.8×10^{-7}
2	M/G1	3.6×10^{-9}	G2/M	6.2×10^{-9}
3	G2/M	5.2×10^{-14}	G1	1.9×10^{-34}
4	M/G1	2.2×10^{-33}	S	4.3×10^{-19}
5	G1	5.9×10^{-57}	G2/M	$6. \times 10^{-19}$

(* Pheromone Response Annotations *)

```
annotations = genenames[[7]];
stages = {"Up", "Down", "None"};
```

```
most = 150;
```

```
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]}, {a, 1, Dimensions[stages][[1]]}]]
```

```
counter = Table[{a}, {a, 1, 5}];
```

```
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

```
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

```
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

```
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

```
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

```
{71, 82, 2340}
```

```
Do[{
```

```
  pattern = Transpose[Sort[
```

```
    TransposeJoin[{Transpose[eigenarrays][[c]], {annotations}], OrderedQ[{{#2}, {#1}}] &][[2]],
```

```
  table = Table[{
```

```
    stages[[a],
```

```
    numbers[[a],
```

```
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
```

```
    {a, 1, Dimensions[stages][[1]]},
```

```
  probability = Table[{
```

```
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
```

```
      Binomial[genes, most]], {b, table[[a, 3], most}],
```

```
    stages[[a]],
```

```
    {a, 1, Dimensions[stages][[1]]},
```

```
  parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
```

```
  parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
```

```
  table = Table[{
```

```
    stages[[a],
```

```
    numbers[[a],
```

```
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
```

```
    {a, 1, Dimensions[stages][[1]]},
```

```
  probability = Table[{
```

```
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
```

```
      Binomial[genes, most]], {b, table[[a, 3], most}],
```

```
    stages[[a]],
```

```
    {a, 1, Dimensions[stages][[1]]},
```

```
  antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
```

```
  antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
```

```
{c, 1, Dimensions[counter][[1]]}]
```

(* Display the Enrichments *)

```
table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]
```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	Up	2.5×10^{-1}	Up	4.1×10^{-14}
2	Up	7.9×10^{-37}	Down	4.1×10^{-33}
3	Up	4.5×10^{-7}	Up	5.4×10^{-5}
4	Up	5.4×10^{-5}	Down	7.1×10^{-10}
5	Down	1.4×10^{-11}	Up	8.4×10^{-4}

(* 5. What annotations do the top four eigenarrays associate with? *)

(* 6. Are the eigenarray associations consistent with the eigene patterns? *)

(* The SVD Reconstruction of Cell Cycle mRNA Expression *)

(* Compute the SVD Reconstruction by Removing the First Eigene and Eigenarray *)

```
{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
eigenexpressions[[1]] = 0;
matrix = Dot[eigenarrays, Dot[DiagonalMatrix[eigenexpressions], eigengenes]];
```

(* Examine the SVD-Reconstructed Data by Using the SVD *)

```
{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
rank = arrays - Count[eigenexpressions, 0.]
fractions = eigenexpressions^2 / Sum[eigenexpressions[[a]]^2, {a, 1, rank}];
entropy = -N[Sum[fractions[[a]] * Log[fractions[[a]]], {a, 1, rank}] / Log[rank];
entropy = N[Round[100 * entropy] / 100]
```

17

0.5

```
(* Create Fractions Bar Chart Display *)
```

```
gridx = Table[a, {a, 0.25, 0.75, 0.25}];  
framex = Table[{gridx[[a]], Rotate[gridx[[a]], Pi / 2]}, {a, 1, 3}];  
gridx = Table[{gridx[[a]], RGBColor[0, 0, 0]}, {a, 1, 3}];  
framey = Table[{a + 1, rank - a}, {a, 0, rank - 1}];  
labelx = ColumnForm[  
  {"(b) Eigenexpression Fraction", StringJoin["d = ", ToString[entropy]]},  
  Center];  
g2 = Show[BarChart[  
  Table[fractions[[rank - a]], {a, 0, rank - 1}],  
  BarOrigin → Left,  
  PlotRange → {{0.0001, 0.9999}, {0.5, rank + 0.5}},  
  AspectRatio → 1,  
  ChartStyle → Red],  
  Axes → False,  
  Frame → True,  
  FrameTicks → {None, framey, framex, None},  
  FrameLabel → {None, None, labelx, None},  
  GridLines → {gridx, None},  
  BaseStyle → {FontFamily → "Courier"}];
```

```
(* Create Eigengenes Raster Display *)
```

```
contrast = 3.5;  
displaying = Table[  
  If[contrast * eigengenes[[i, j]] > 0,  
    If[contrast * eigengenes[[i, j]] < 1, {contrast * eigengenes[[i, j]], 0}, {1, 0}],  
    If[contrast * eigengenes[[i, j]] > -1, {0, -contrast * eigengenes[[i, j]]}, {0, 1}]],  
  {i, 1, rank}, {j, 1, arrays}];  
framex = Table[{a - 0.5, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];  
framey = Table[{a + 1 - 0.5, rank - a}, {a, 0, rank - 1}];  
labely = "Eigengenes";  
labelx = "(a) Arrays";  
g1 = Show[  
  Graphics[  
    Raster[  
      Table[  
        {displaying[[i, j, 1]], displaying[[i, j, 2]], 0},  
        {i, rank, 1, -1}, {j, 1, arrays}]]],  
    AspectRatio → 1,  
    Frame → True,  
    FrameTicks → {None, framey, framex, None},  
    FrameLabel → {None, labely, labelx, None},  
    BaseStyle → {FontFamily → "Courier"}];
```



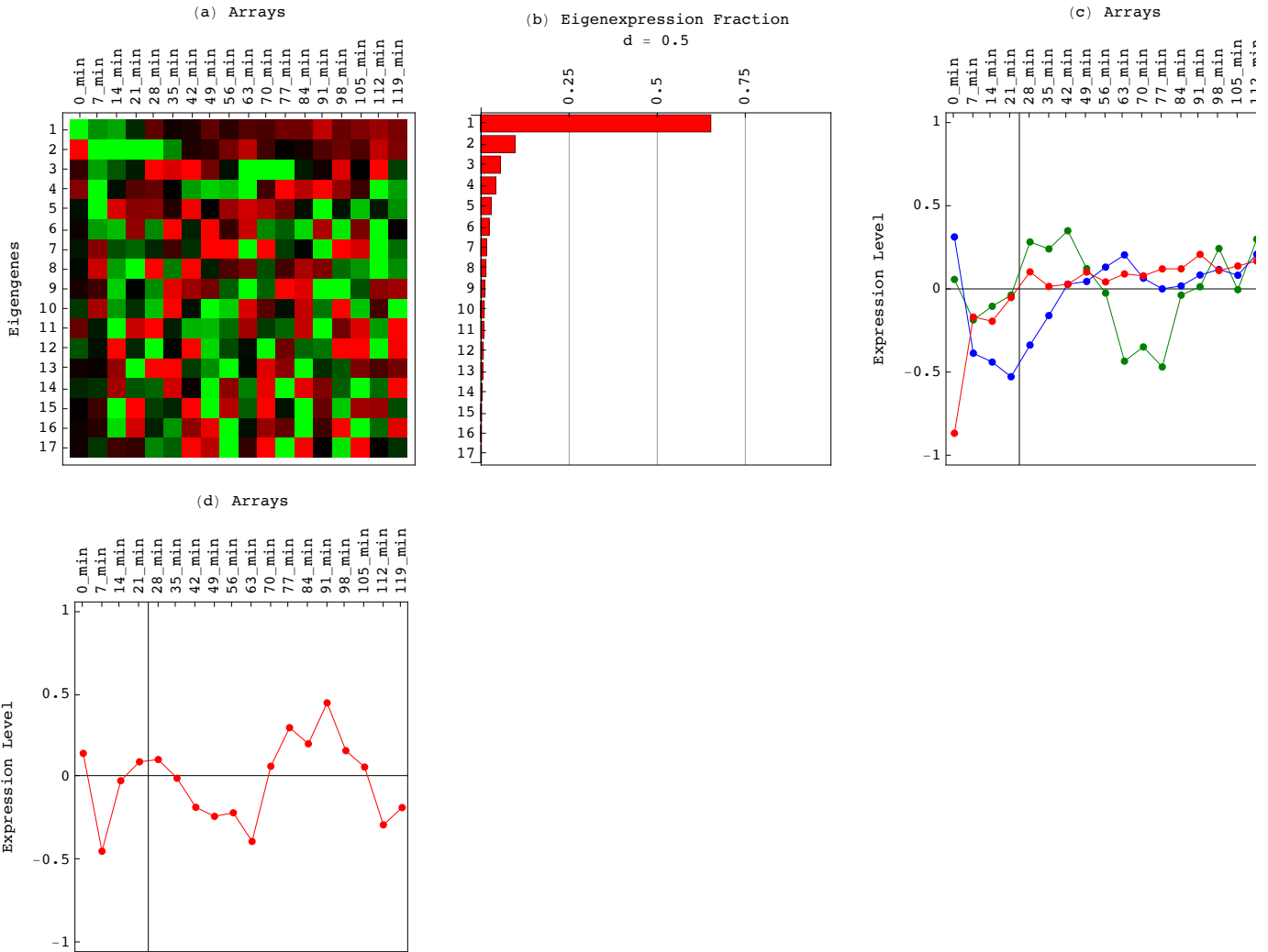
```
(* Create Selected Eigengenes Graph Display *)
```

```
p = Table[0, {n, 1, 3}];
color = {RGBColor[1, 0, 0], RGBColor[0, 0, 1], RGBColor[0, 0.5, 0]};
labelx = "(c) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n, a]], {a, 1, arrays}},
  points = Table[Point[coordinates[[a]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
    Frame → True,
    FrameLabel → {None, labely, labelx, None},
    GridLines → {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}}},
    FrameTicks → {None, framey, framex, None},
    BaseStyle → {FontFamily → "Courier"}],
  p[[n]] = Show[g,
    AspectRatio → 1.05,
    PlotRange → {-1.05, 1.05},
    DisplayFunction → Identity]],
{n, 1, 3}]
g3 = Show[{p[[3]], p[[2]], p[[1]]};
```

```
p = Table[0, {n, 1, 2}];
color = {RGBColor[1, 0, 0], RGBColor[0, 0, 1]};
labelx = "(d) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n + 3, a]], {a, 1, arrays}},
  points = Table[Point[coordinates[[a]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
    Frame → True,
    FrameLabel → {None, labely, labelx, None},
    GridLines → {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}}},
    FrameTicks → {None, framey, framex, None},
    BaseStyle → {FontFamily → "Courier"}],
  p[[n]] = Show[g,
    AspectRatio → 1.05,
    PlotRange → {-1.05, 1.05},
    DisplayFunction → Identity]],
{n, 1, 1}]
g4 = Show[{p[[1]]};
```

(* Display the SVD *)

```
Show[GraphicsGrid[{{g1, g2, g3}}, Spacings -> -20],
ImageSize -> 750]
Show[GraphicsGrid[{{g4}},
ImageSize -> 250]
```



(* 7. What do the top three eigengenes look like after reconstruction? *)

(* 8. Which eigengenes in the original data do they correlate with? *)

(* 9. What biological processes might they describe? *)

(* Compute P-Value of Enrichment of Annotations Assuming Hypergeometric Distribution *)

```
headerx = {{
ColumnForm[{" ", " ", "Eigenarray"}, Left],
ColumnForm[{"Most Likely", "Parallel", "Association"}, Left],
ColumnForm[{"P-Value of", "Parallel", "Association"}, Left],
ColumnForm[{"Most Likely", "Antiparallel", "Association"}, Left],
ColumnForm[{"P-Value of", "Antiparallel", "Association"}, Left]}};
```

(* Cell Cycle Annotations *)

```
annotations = genenames[[6]];
stages = {"M/G1", "G1", "S", "S/G2", "G2/M", "None"};

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}}];
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

{53, 117, 32, 63, 91, 2137}

Do[{
  pattern = Transpose[Sort[
    TransposeJoin[{Transpose[eigenarrays][[c]], {annotations}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
  {c, 1, Dimensions[counter][[1]]}]
```

(* Display the Enrichments *)

```
table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]
```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	G2/M	6.2×10^{-9}	M/G1	3.6×10^{-9}
2	G2/M	5.2×10^{-14}	G1	1.9×10^{-34}
3	S	4.3×10^{-19}	M/G1	2.2×10^{-33}
4	G1	5.9×10^{-57}	G2/M	$6. \times 10^{-19}$
5	G1	3.5×10^{-15}	M/G1	9.5×10^{-4}

(* Pheromone Response Annotations *)

```
annotations = genenames[[7]];
stages = {"Up", "Down", "None"};

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}]
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
```

{71, 82, 2340}

```
Do[{
  pattern = Transpose[Sort[
    Transpose[Join[{Transpose[eigenarrays][[c]], {annotations}}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a],
    numbers[[a],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2], most - b] /
      Binomial[genes, most]], {b, table[[a, 3], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  parallelannotation[c] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  parallelprobability[c] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  table = Table[{
    stages[[a],
    numbers[[a],
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2], most - b] /
      Binomial[genes, most]], {b, table[[a, 3], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  antiannotation[c] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  antiprobability[c] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  {c, 1, Dimensions[counter][[1]]}
```

(* Display the Enrichments *)

```
table = Join[headerx, Transpose[Join[Transpose[counter],  
  Transpose[parallelannotation],  
  Transpose[parallelprobability],  
  Transpose[antiannotation],  
  Transpose[antiprobability]]]];  
TableForm[table]
```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	Down	4.1×10^{-33}	Up	7.9×10^{-37}
2	Up	4.5×10^{-7}	Up	5.4×10^{-5}
3	Down	7.1×10^{-10}	Up	5.4×10^{-5}
4	Down	1.4×10^{-11}	Up	8.4×10^{-4}
5	Down	1.4×10^{-11}	Up	4.5×10^{-7}

(* 10. What annotations do the top four eigenarrays associate with after reconstruction? *)

(* 11. Which eigenarrays in the original data do they correlate with? *)

(* 12. Are the eigenarray associations consistent with the eigengene patterns? *)

(* The Pseudoinverse Projection of the Cell Cycle mRNA Expression Data onto DNA-Binding of Cell Cycle Transcription Factors *)

(* Read Cell Cycle Binding Basis *)

```
a = 1;  
b = 7;
```

```
matrix = Import["https://alterlab.org/network_decomposition/data/Cell_Cycle_Binding.txt", "Table"];  
{genes, arrays} = Dimensions[matrix] - {a, b}
```

```
{2120, 12}
```

```
arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}]], {1, b}]];  
matrix = Drop[matrix, {1}];  
matrix = Sort[Transpose[Join[  
  {Table[If[Position[genenames, matrix[[c, 1]] == {}, 1, 0], {c, 1, genes}], Transpose[matrix]]]];  
pos = Position[matrix[[All, 1]], 1][[1, 1]];  
matrix = Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes}]], {1}]]];  
{genes, arrays} = Dimensions[matrix] - {0, b}
```

```
{960, 12}
```

```
matrix = Transpose[matrix];  
genenames = Drop[matrix, {b + 1, b + arrays}];  
matrix = Drop[matrix, {1, b}];  
matrix = Transpose[matrix];  
Clear[a, b];
```

```

(* Compute the SVD Reconstruction by Removing the First Eigengene and Eigenarray *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
eigenexpressions[[1]] = 0;
matrix = Dot[eigenarrays, Dot[DiagonalMatrix[eigenexpressions], eigengenes]];

(* Save as Basis *)

basis = matrix;
basisarrays = arrays;

(* Read Cell Cycle Expression Data *)

a = 1;
b = 7;

matrix =
  Import["https://alterlab.org/network_decomposition/data/Cell_Cycle_Expression.txt", "Table"];
{genes, arrays} = Dimensions[matrix] - {a, b}

{4153, 18}

arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}]], {1, b}]];
matrix = Drop[matrix, {1}];
matrix = Sort[Transpose[Join[
  {Table[If[Position[genenames, matrix[[c, 1]]] == {}, 1, 0], {c, 1, genes}], Transpose[matrix]]]];
pos = Position[matrix[All, 1], 1][[1, 1]];
matrix = Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes}]], {1}]]];
{genes, arrays} = Dimensions[matrix] - {0, b}

{960, 18}

matrix = Transpose[matrix];
genenames = Drop[matrix, {b + 1, b + arrays}];
matrix = Drop[matrix, {1, b}];
matrix = Transpose[matrix];
Clear[a, b];

(* Compute the SVD Reconstruction by Removing the First Eigengene and Eigenarray *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
eigenexpressions[[1]] = 0;
matrix = Dot[eigenarrays, Dot[DiagonalMatrix[eigenexpressions], eigengenes]];

(* Compute the Pseudoinverse Projection of the Data onto the Basis *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[basis, basisarrays - 1];
matrix = Dot[Dot[eigenarrays, Transpose[eigenarrays], matrix]];
{genes, arrays} = Dimensions[matrix]

{960, 18}

```

```
(* Examine the Pseudoinverse-Projected Data by Using the SVD *)
```

```
{eigenarrays, eigenexpressions, eigengenex} = SingularValueDecomposition[matrix, arrays];  
eigenexpressions = Diagonal[eigenexpressions];  
eigengenes = Transpose[eigengenes];  
rank = arrays - Count[eigenexpressions, 0.]  
fractions = eigenexpressions^2 / Sum[eigenexpressions[[a]]^2, {a, 1, rank}];  
entropy = -N[Sum[fractions[[a]] * Log[fractions[[a]]], {a, 1, rank}] / Log[rank];  
entropy = N[Round[100 * entropy] / 100]
```

```
11
```

```
0.56
```

```
(* 13. How did the number of rank of freedom and entropy change? *)
```

```
(* 14. What can we learn from these changes about the pseudoinverse-projected data? *)
```

```
(* Create Fractions Bar Chart Display *)
```

```
gridx = Table[a, {a, 0.25, 0.75, 0.25}];  
framex = Table[{gridx[[a]], Rotate[gridx[[a]], Pi / 2]}, {a, 1, 3}];  
gridx = Table[{gridx[[a]], RGBColor[0, 0, 0]}, {a, 1, 3}];  
framey = Table[{a + 1, rank - a}, {a, 0, rank - 1}];  
labelx = ColumnForm[  
  {"(b) Eigenexpression Fraction", StringJoin["d = ", ToString[entropy]]},  
  Center];  
g2 = Show[BarChart[  
  Table[fractions[[rank - a]], {a, 0, rank - 1}],  
  BarOrigin -> Left,  
  PlotRange -> {{0.0001, 0.9999}, {0.5, rank + 0.5}},  
  AspectRatio -> 1,  
  ChartStyle -> Red],  
  Axes -> False,  
  Frame -> True,  
  FrameTicks -> {None, framey, framex, None},  
  FrameLabel -> {None, None, labelx, None},  
  GridLines -> {gridx, None},  
  BaseStyle -> {FontFamily -> "Courier"}];
```

(* Create Eigengenes Raster Display *)

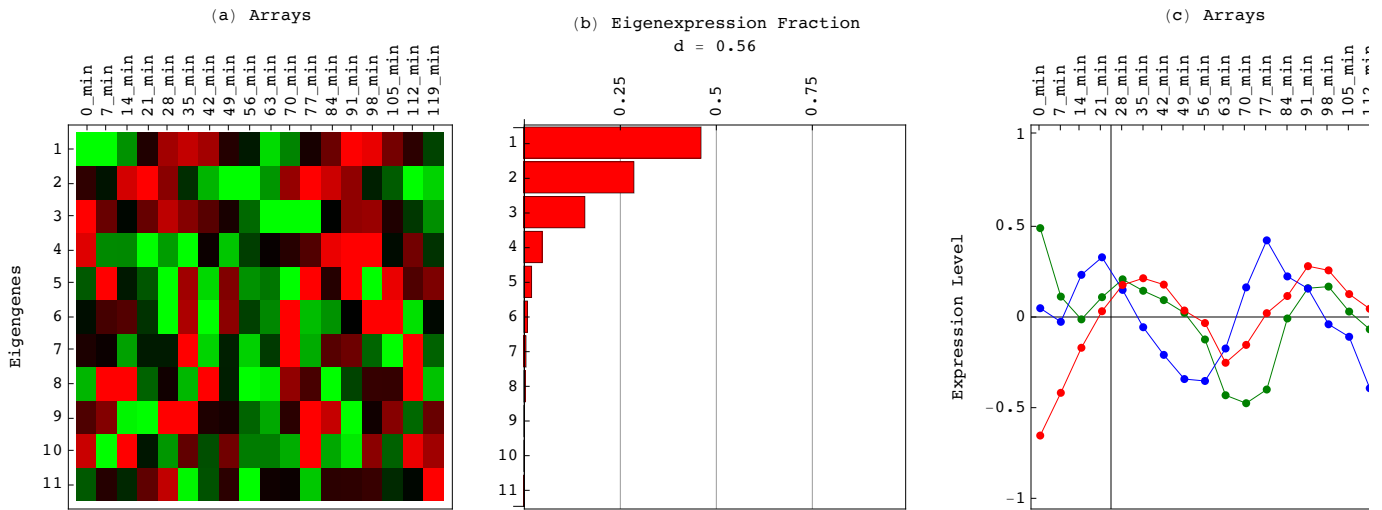
```
contrast = 3.5;
displaying = Table[
  If[contrast * eigengenes[[i, j]] > 0,
    If[contrast * eigengenes[[i, j]] < 1, {contrast * eigengenes[[i, j]], 0}, {1, 0}],
    If[contrast * eigengenes[[i, j]] > -1, {0, -contrast * eigengenes[[i, j]]}, {0, 1}]],
  {i, 1, rank}, {j, 1, arrays}];
framex = Table[{a - 0.5, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = Table[{a + 1 - 0.5, rank - a}, {a, 0, rank - 1}];
labely = "Eigengenes";
labelx = "(a) Arrays";
g1 = Show[
  Graphics[
    Raster[
      Table[
        {displaying[[i, j, 1]], displaying[[i, j, 2]], 0},
        {i, rank, 1, -1}, {j, 1, arrays}]]],
  AspectRatio -> 1,
  Frame -> True,
  FrameTicks -> {None, framey, framex, None},
  FrameLabel -> {None, labely, labelx, None},
  BaseStyle -> {FontFamily -> "Courier"}];
```

(* Create Selected Eigengenes Graph Display *)

```
p = Table[0, {n, 1, 5}];
color =
  {RGBColor[1, 0, 0], RGBColor[0, 0, 1], RGBColor[0, 0.5, 0], , RGBColor[1, 0.5, 0], RGBColor[1, 0, 0]};
labelx = "(c) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n, a]], {a, 1, arrays}},
  points = Table[Point[coordinates[[a]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
  Frame -> True,
  FrameLabel -> {None, labely, labelx, None},
  GridLines -> {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}},
  FrameTicks -> {None, framey, framex, None},
  BaseStyle -> {FontFamily -> "Courier"}],
  p[[n]] = Show[g,
  AspectRatio -> 1.05,
  PlotRange -> {-1.05, 1.05},
  DisplayFunction -> Identity}],
{n, 1, 3}]
g3 = Show[{p[[3]], p[[2]], p[[1]]};
```


(* Display the SVD *)

Show[GraphicsGrid[{{g1, g2, g3}}, Spacings -> -20], ImageSize -> 750]



(* 15. What do the top three eigengenes look like after the pseudoinverse projection? *)

(* 16. Which eigengenes in the original data do they correlate with? *)

(* 17. What biological processes might they describe? *)

(* Compute P-Value of Enrichment of Annotations Assuming Hypergeometric Distribution *)

```
headerx = {{
  ColumnForm[{" ", " ", "Eigenarray"}, Left],
  ColumnForm[{"Most Likely", "Parallel", "Association"}, Left],
  ColumnForm[{"P-Value of", "Parallel", "Association"}, Left],
  ColumnForm[{"Most Likely", "Antiparallel", "Association"}, Left],
  ColumnForm[{"P-Value of", "Antiparallel", "Association"}, Left]}};
```

(* Cell Cycle Annotations *)

```
annotations = genenames[[6]];
stages = {"M/G1", "G1", "S", "S/G2", "G2/M", "None"};

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}}];
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

{38, 66, 17, 28, 35, 776}
```

```

Do[{
  pattern = Transpose[Sort [
    Transpose[Join[{Transpose[eigenarrays][[c]], {annotations}}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]}],
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]}],
  parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]}],
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]}],
  antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  {c, 1, Dimensions[counter][[1]]}}

```

(* Display the Enrichments *)

```

table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]

```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	G1	2.3×10^{-7}	M/G1	2.4×10^{-6}
2	G1	$1. \times 10^{-8}$	G2/M	7.9×10^{-5}
3	S	4.8×10^{-5}	M/G1	5.8×10^{-5}
4	G1	8.5×10^{-3}	G1	1.6×10^{-4}
5	S	1.1×10^{-1}	G1	1.3×10^{-1}

(* Pheromone Response Annotations *)

```

annotations = genenames[[7]];
stages = {"Up", "Down", "None"};

```

```

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}]]
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

{35, 48, 877}

```

```

Do[{
  pattern = Transpose[Sort[
    Transpose[Join[{Transpose[eigenarrays][[c]], {annotations}}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3], most]},
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]},
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3], most]},
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]},
  antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]},
  {c, 1, Dimensions[counter][[1]]}

```

(* Display the Enrichments *)

```

table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]

```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	Down	2.6×10^{-11}	Up	4.4×10^{-3}
2	Down	3.2×10^{-5}	None	2.2×10^{-1}
3	Down	1.1×10^{-2}	None	7.3×10^{-2}
4	Down	1.1×10^{-2}	Down	5.7×10^{-2}
5	Down	1.1×10^{-1}	Down	2.6×10^{-2}

```
(* 18. What annotations do the top four eigenarrays associate with after the pseudoinverse projection? *)
```

```
(* 19. Which eigenarrays in the original data do they correlate with? *)
```

```
(* 20. Are the eigenarray associations consistent with the eigengene patterns? *)
```

```
(* The Pseudoinverse Projection of the Cell Cycle mRNA Expression Data onto DNA-Binding of Development Transcription Factors *)
```

```
(* Read Cell Cycle Expression Data *)
```

```
a = 1;
b = 7;

matrix =
  Import["https://alterlab.org/network_decomposition/data/Cell_Cycle_Expression.txt", "Table"];
header = matrix[[1]];
matrix = Drop[matrix, {1}];
{genes, arrays} = Dimensions[matrix] - {a, b};
matrix =
  Sort[Transpose[Join[{Table[Count[matrix[[c]], "Null"], {c, 1, genes + a}], Transpose[matrix]]]];
pos = Position[matrix[All, 1], 1][[1, 1]];
matrix = Join[{header}, Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes + a}], {1}]]]];
{genes, arrays} = Dimensions[matrix] - {a, b}

{2493, 18}

arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}], {1, b}]];
matrix = Drop[matrix, {1}];

matrix = Transpose[matrix];
genenames = Drop[matrix, {b + 1, b + arrays}];
matrix = Drop[matrix, {1, b}];
matrix = Transpose[matrix];
Clear[a, b];
```

```
(* Read Development Binding Basis *)
```

```
a = 1;
b = 7;

matrix = Import["https://alterlab.org/network_decomposition/data/Develop_Binding.txt", "Table"];
{genes, arrays} = Dimensions[matrix] - {a, b}

{2476, 12}

arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}], {1, b}]];
matrix = Drop[matrix, {1}];
matrix = Sort[Transpose[Join[
  {Table[If[Position[genenames, matrix[[c, 1]] == {}, 1, 0], {c, 1, genes}], Transpose[matrix]]]];
pos = Position[matrix[All, 1], 1][[1, 1]];
matrix = Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes}], {1}]]]];
{genes, arrays} = Dimensions[matrix] - {0, b}

{1116, 12}
```

```

matrix = Transpose[matrix];
genenames = Drop[matrix, {b + 1, b + arrays}];
matrix = Drop[matrix, {1, b}];
matrix = Transpose[matrix];
Clear[a, b];

(* Compute the SVD Reconstruction by Removing the First Eigengene and Eigenarray *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
eigenexpressions[[1]] = 0;
matrix = Dot[eigenarrays, Dot[DiagonalMatrix[eigenexpressions], eigengenes]];

(* Save as Basis *)

basis = matrix;
basisarrays = arrays;

(* Read Cell_Cycle_Expression_3.txt *)

a = 1;
b = 7;

matrix =
  Import["https://alterlab.org/network_decomposition/data/Cell_Cycle_Expression.txt", "Table"];
{genes, arrays} = Dimensions[matrix] - {a, b}

{4153, 18}

arraynames = Transpose[Drop[Transpose[Drop[matrix, {a + 1, a + genes}]], {1, b}]];
matrix = Drop[matrix, {1}];
matrix = Sort[Transpose[Join[
  {Table[If[Position[genenames, matrix[[c, 1]] == {}, 1, 0], {c, 1, genes}], Transpose[matrix]]]];
pos = Position[matrix[All, 1], 1][[1, 1]];
matrix = Sort[Transpose[Drop[Transpose[Drop[matrix, {pos, genes}]], {1}]]];
{genes, arrays} = Dimensions[matrix] - {0, b}

{1116, 18}

matrix = Transpose[matrix];
genenames = Drop[matrix, {b + 1, b + arrays}];
matrix = Drop[matrix, {1, b}];
matrix = Transpose[matrix];
Clear[a, b];

(* Compute the SVD Reconstruction by Removing the First Eigengene and Eigenarray *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[matrix, arrays];
eigenexpressions = Diagonal[eigenexpressions];
eigengenes = Transpose[eigengenes];
eigenexpressions[[1]] = 0;
matrix = Dot[eigenarrays, Dot[DiagonalMatrix[eigenexpressions], eigengenes]];

(* Compute the Pseudoinverse Projection of the Data onto the Basis *)

{eigenarrays, eigenexpressions, eigengenes} = SingularValueDecomposition[basis, basisarrays - 1];
matrix = Dot[Dot[eigenarrays, Transpose[eigenarrays], matrix]];
{genes, arrays} = Dimensions[matrix]

{1116, 18}

```

```
(* Examine the Pseudoinverse-Projected Data by Using the SVD *)
```

```
{eigenarrays, eigenexpressions, eigengenex} = SingularValueDecomposition[matrix, arrays];  
eigenexpressions = Diagonal[eigenexpressions];  
eigengenes = Transpose[eigengenes];  
rank = arrays - Count[eigenexpressions, 0.]  
fractions = eigenexpressions^2 / Sum[eigenexpressions[[a]]^2, {a, 1, rank}];  
entropy = -N[Sum[fractions[[a]] * Log[fractions[[a]]], {a, 1, rank}] / Log[rank];  
entropy = N[Round[100 * entropy] / 100]
```

```
11
```

```
0.27
```

```
(* 21. How did the number of rank of freedom and entropy change? *)
```

```
(* 22. What can we learn from these changes about the pseudoinverse-projected data? *)
```

```
(* Create Fractions Bar Chart Display *)
```

```
gridx = Table[a, {a, 0.25, 0.75, 0.25}];  
framex = Table[{gridx[[a]], Rotate[gridx[[a]], Pi / 2]}, {a, 1, 3}];  
gridx = Table[{gridx[[a]], RGBColor[0, 0, 0]}, {a, 1, 3}];  
framey = Table[{a + 1, rank - a}, {a, 0, rank - 1}];  
labelx = ColumnForm[  
  {"(b) Eigenexpression Fraction", StringJoin["d = ", ToString[entropy]]},  
  Center];  
g2 = Show[BarChart[  
  Table[fractions[[rank - a]], {a, 0, rank - 1}],  
  BarOrigin -> Left,  
  PlotRange -> {{0.0001, 0.9999}, {0.5, rank + 0.5}},  
  AspectRatio -> 1,  
  ChartStyle -> Red],  
  Axes -> False,  
  Frame -> True,  
  FrameTicks -> {None, framey, framex, None},  
  FrameLabel -> {None, None, labelx, None},  
  GridLines -> {gridx, None},  
  BaseStyle -> {FontFamily -> "Courier"}];
```

(* Create Eigengenes Raster Display *)

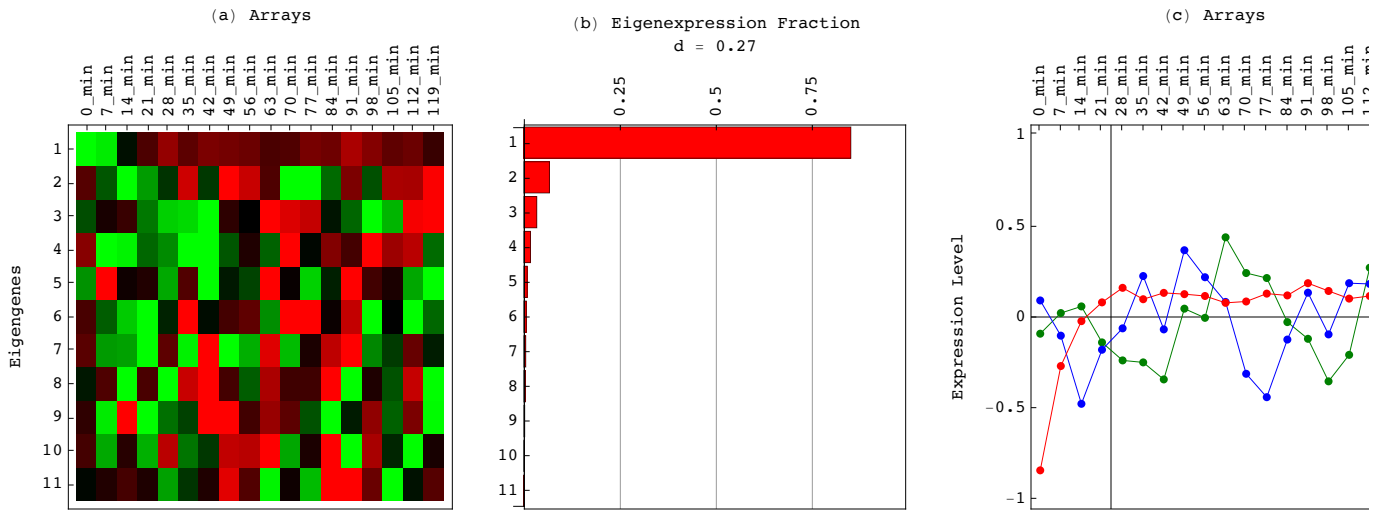
```
contrast = 3.5;
displaying = Table[
  If[contrast * eigengenes[[i, j]] > 0,
    If[contrast * eigengenes[[i, j]] < 1, {contrast * eigengenes[[i, j]], 0}, {1, 0}],
    If[contrast * eigengenes[[i, j]] > -1, {0, -contrast * eigengenes[[i, j]]}, {0, 1}]],
  {i, 1, rank}, {j, 1, arrays}];
framex = Table[{a - 0.5, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = Table[{a + 1 - 0.5, rank - a}, {a, 0, rank - 1}];
labely = "Eigengenes";
labelx = "(a) Arrays";
g1 = Show[
  Graphics[
    Raster[
      Table[
        {displaying[[i, j, 1]], displaying[[i, j, 2]], 0},
        {i, rank, 1, -1}, {j, 1, arrays}]]],
  AspectRatio → 1,
  Frame → True,
  FrameTicks → {None, framey, framex, None},
  FrameLabel → {None, labely, labelx, None},
  BaseStyle → {FontFamily → "Courier"}];
```

(* Create Selected Eigengenes Graph Display *)

```
p = Table[0, {n, 1, 5}];
color =
  {RGBColor[1, 0, 0], RGBColor[0, 0, 1], RGBColor[0, 0.5, 0], , RGBColor[1, 0.5, 0], RGBColor[1, 0, 0]};
labelx = "(c) Arrays";
labely = "Expression Level";
framex = Table[{a - 1, Rotate[arraynames[[1, a]], Pi / 2]}, {a, 1, arrays}];
framey = {-1, -0.5, 0, 0.5, 1};
Do[{
  coordinates = Table[{a - 1, eigengenes[[n, a]], {a, 1, arrays}},
  points = Table[Point[coordinates[[a]], {a, 1, arrays}],
  line = Line[coordinates],
  g = Show[
    {Graphics[{color[[n]], PointSize[0.022], points}],
    Graphics[{color[[n]], line}]},
  Frame → True,
  FrameLabel → {None, labely, labelx, None},
  GridLines → {{{3.5, RGBColor[0, 0, 0]}}, {{0, RGBColor[0, 0, 0]}},
  FrameTicks → {None, framey, framex, None},
  BaseStyle → {FontFamily → "Courier"}],
  p[[n]] = Show[g,
  AspectRatio → 1.05,
  PlotRange → {-1.05, 1.05},
  DisplayFunction → Identity]},
{n, 1, 3}]
g3 = Show[{p[[3]], p[[2]], p[[1]]};
```

(* Display the SVD *)

Show[GraphicsGrid[{{g1, g2, g3}}, Spacings -> -20], ImageSize -> 750]



(* 23. What do the top three eigengenes look like after the pseudoinverse projection? *)

(* 24. Which of the eigengenes are significant in terms of the fraction of the information they capture in the data? *)

(* 25. Which eigengenes in the original data do they correlate with? *)

(* 26. What biological processes might they describe? *)

(* Compute P-Value of Enrichment of Annotations Assuming Hypergeometric Distribution *)

```
headerx = {{
  ColumnForm[{" ", " ", "Eigenarray"}, Left],
  ColumnForm[{"Most Likely", "Parallel", "Association"}, Left],
  ColumnForm[{"P-Value of", "Parallel", "Association"}, Left],
  ColumnForm[{"Most Likely", "Antiparallel", "Association"}, Left],
  ColumnForm[{"P-Value of", "Antiparallel", "Association"}, Left]}};
```

(* Cell Cycle Annotations *)

```
annotations = genenames[[6]];
stages = {"M/G1", "G1", "S", "S/G2", "G2/M", "None"};

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}}];
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

{37, 71, 18, 30, 50, 910}
```



```

Do[{
  pattern = Transpose[Sort [
    Transpose[Join[{Transpose[eigenarrays][[c]], {annotations}}], OrderedQ[{{#2}, {#1}}] &][[2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]}],
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]}],
  parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  table = Table[{
    stages[[a]],
    numbers[[a]],
    Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
    {a, 1, Dimensions[stages][[1]]}],
  probability = Table[{
    Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
      Binomial[genes, most]], {b, table[[a, 3]], most}],
    stages[[a]],
    {a, 1, Dimensions[stages][[1]]}],
  antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
  antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
  {c, 1, Dimensions[counter][[1]]}}

```

(* Display the Enrichments *)

```

table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]

```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	G2/M	2.2×10^{-1}	M/G1	6.7×10^{-3}
2	None	1.8×10^{-3}	M/G1	1.4×10^{-7}
3	M/G1	6.7×10^{-3}	S	8.2×10^{-2}
4	G1	1.4×10^{-1}	None	1.2×10^{-1}
5	G2/M	1.2×10^{-1}	G2/M	3.5×10^{-1}

(* Pheromone Response Annotations *)

```

annotations = genenames[[7]];
stages = {"Up", "Down", "None"};

```

```

most = 150;
numbers = Flatten[Table[{Count[Flatten[annotations], stages[[a]]], {a, 1, Dimensions[stages][[1]]}]]
counter = Table[{a}, {a, 1, 5}];
probability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
parallelprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiannotation = Table[{0}, {a, 1, Dimensions[counter][[1]]}];
antiprobability = Table[{0}, {a, 1, Dimensions[counter][[1]]}];

{44, 48, 1024}

```

```

Do[{
  pattern = Transpose[Sort[
    Transpose[Join[{Transpose[eigenarrays][[c]], {annotations}], OrderedQ[{{#2}, {#1}}] &][[2]],
    table = Table[{
      stages[[a]],
      numbers[[a]],
      Count[Flatten[Drop[pattern, {most + 1, genes}]], stages[[a]]],
      {a, 1, Dimensions[stages][[1]]},
    probability = Table[{
      Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
        Binomial[genes, most]], {b, table[[a, 3], most}],
      stages[[a]],
      {a, 1, Dimensions[stages][[1]]},
    parallelannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
    parallelprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
    table = Table[{
      stages[[a]],
      numbers[[a]],
      Count[Flatten[Drop[pattern, {1, genes - most}]], stages[[a]]],
      {a, 1, Dimensions[stages][[1]]},
    probability = Table[{
      Sum[N[Binomial[table[[a, 2]], b] * Binomial[genes - table[[a, 2]], most - b] /
        Binomial[genes, most]], {b, table[[a, 3], most}],
      stages[[a]],
      {a, 1, Dimensions[stages][[1]]},
    antiannotation[[c]] = {Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 2]],
    antiprobability[[c]] = {ScientificForm[Sort[probability, OrderedQ[{{#1}, {#2}}] &][[1, 1], 2]],
    {c, 1, Dimensions[counter][[1]]}
}

```

(* Display the Enrichments *)

```

table = Join[headerx, Transpose[Join[Transpose[counter],
  Transpose[parallelannotation],
  Transpose[parallelprobability],
  Transpose[antiannotation],
  Transpose[antiprobability]]]];
TableForm[table]

```

Eigenarray	Most Likely Parallel Association	P-Value of Parallel Association	Most Likely Antiparallel Association	P-Value of Antiparallel Association
1	Down	3.1×10^{-1}	Up	3.5×10^{-6}
2	None	1.8×10^{-1}	Down	7.8×10^{-3}
3	Up	1.2×10^{-1}	Down	1.8×10^{-1}
4	None	$4. \times 10^{-1}$	Down	4.7×10^{-1}
5	None	1.8×10^{-1}	None	$4. \times 10^{-1}$

(* 27. What annotations do the top four eigenarrays associate with after the pseudoinverse projection? *)