Практическое занятие 4

Компьютерная система ANDSystem. Реконструкция и анализ ассоциативных генных сетей. Выявление центральных вершин в генных сетях. Кластеризация генных сетей.

**Цели занятия:** знакомство с системой ANDSystem; овладение методами реконструкции и анализа ассоциативных генных сетей с помощью системы ANDSystem; ознакомление с методами оценки центральностей вершин в графах и кластеризации генных сетей.

В ходе практического занятия нужно выполнить описанные ниже упражнения (этапы 1, 2 и 3) и приготовить отчет, включающий:

(а) фамилию,

(б) результаты выполнения упражнений в виде xlsx файла.

Отчет надо отправить на адрес **tiys@bionet.nsc.ru** в приложении к письму в виде xlsx файла.

# *Этап 1. Знакомство с системой ANDSystem. Реконструкция и анализ ассоциативных генных сетей.*

**Цель этапа:** знакомство с системой ANDSystem и интерфейсом системы ANDVisio; овладение методами реконструкции и анализа ассоциативных генных сетей с помощью системы ANDSystem.

## Шаг 1. Установка программы ANDVisio

1.1.1. Скачайте архив с программой ANDVisio, пройдя по ссылке: <http://www-bionet.sscc.ru/and/cell/service/download/update.html>

1.1.2. Распакуйте архив с программой ANDVisio и запустите программу ANDVisio.exe.

1.1.3. В разделе **File→Connect** введите логин и пароль: student.

## Шаг 2. Реконструкция генной сети, ассоциированной с болезнью Альцгеймера (“Alzheimer's disease”), с помощью ANDSystem

1.2.1. В верхней части окна нажмите на иконку «**Query Wizard**».

1.2.2. Во вкладке «**Organism**» отметьте «**Homo sapiens**».

1.2.3. Выберите вкладку «**Synonyms**». В поле «**Synonym**» введите “**alzheimer\***”. В поле «**Type**» выберите “**Disease**” и нажмите кнопку “**Add**”.

1.2.4. В открывшемся окне выберите “**Alzheimer disease**” и переместите это название в правую часть окна, либо перетащив мышкой, либо нажав на кнопку “>”. Нажмите “**Ok**”.

1.2.5. Отметьте галочкой окошко “**Direct relations only**”

1.2.6. Во вкладке «**Object filter**» отметьте только «**Protein**» и «**Gene**».

1.2.7. Во вкладке «**Interaction type filter**» отметьте все типы взаимодействий («**Check all**»). Нажмите “**Ok**”.

1.2.8. В полученной ассоциативной сети нажмите правой кнопкой мыши на любой объект, нажмите «**Property**». Убедитесь, что таким образом Вы можете посмотреть синонимы объектов и ссылки на другие базы данные.

1.2.9. В полученной ассоциативной сети нажмите правой кнопкой мыши на любой зеленый (чёрный) шарик, обозначающий взаимодействие между объектами. Нажмите «**Property**». Убедитесь, что таким образом Вы можете посмотреть тип связи, участников взаимодействия, предложение, в котором описано данное взаимодействие.

1.2.10. Сохраните полученную сеть в файлы «**Alz.and**», «**Alz.tsv**» в форматах AND XML (.and) и Simple table format (.stsv), нажав на кнопку «**Save as**» в верхней части окна. Ознакомьтесь с форматами выдачи генной сети .and и .stsv. Для этого откройте эти файлы в «**Notepad ++**» и посмотрите их содержимое.

## Шаг 3. Реконструкция сети взаимодействий между генами, ассоциированными с болезнью Альцгеймера по данным OMIM, с помощью ANDSystem и проверка связности сети с помощью FunGeneNet

1.3.1. Зайдите на сайт OMIM (<http://www.omim.org/>).

1.3.2. Введите в поисковую строку запрос «**Alzheimer's disease**» и нажмите кнопку «**Search**».

1.3.3. Сохраните результат поиска в xls файл, нажав на кнопку «**Download As**».

1.3.4. Отсортируйте сохраненную таблицу по колонке «**Entrez Gene ID**» и сохраните идентификаторы из этой колонки в файл «**GeneID\_from\_OMIM.txt**».

1.3.5. Откройте FunGeneNet (<http://www-bionet.sscc.ru/fungenenet/index.php>), выберите в выпадающих списках: **«DATA SOURCE» = «ANDSystem», «ORGANISM» = человек, «INTERACTION TYPE» = «all types» и «DATABASE FOR UPLOAD LIST» = «GeneID».**

1.3.6. Загрузите файл «**GeneID\_from\_OMIM.txt**» и нажмите «**CONTINUE**».

1.3.7. Скопируйте значение **T-test p-value** в xls файл с результатами из OMIM в первую ячейку нового листа «**connectivity**».

1.3.8. Загрузите полученный список «**Entrez Gene ID**» в систему ANDSystem. Для этого откройте новую вкладку нажав на кнопку “**+**”. В новой вкладке нажмите на на иконку “**Query Wizard**”. В окне «**Synonyms**» удалите старый запрос с вкладки «**Synonym**».

1.3.9. Во вкладке “**Database links**” уберите отметку “**Direct relations only**” и выберете уровень «**0**» в поле «**Level**».

1.3.10. Вставьте список «**Entrez Gene ID**» из таблицы OMIM в поле “**Database ID**”. В поле “**Database name**” выберите “**GeneID**”. Нажмите кнопку “**Add**”.

1.3.11. Убедитесь, что во вкладке «**Interaction type filter**» отмечены все типы взаимодействий («**Check all**»). Нажмите “**Ok**”.

1.3.12. Нажмите на кнопку раскладки “**Relayout graph**”.

1.3.13. Сохраните полученную сеть в файлы «**OMIM.and**» «**OMIM.stsv**» в форматах AND XML (.and) и Simple table format (.stsv), нажав на кнопку «**Save as**» в верхней части окна.

## Шаг 4. Проверка связности сети дифференциально экспрессирующихся генов болезни Альцгеймера

1.4.1. Скопируйте список идентификаторов UniGene (или других идентификаторов) для генов, полученных в результате анализа дифференциальной экспрессии на прошлом занятии из файла **«practice3\_ВашаФамилия.xls»** с листа **«Alz\_diff»** в файл «**UniGene\_alz\_differ\_exp.txt**»

1.4.2. Откройте FunGeneNet (<http://www-bionet.sscc.ru/fungenenet/index.php>), выберите в выпадающих списках: **«DATA SOURCE» = «ANDSystem», «ORGANISM» = человек, «INTERACTION TYPE» = «all types» и «DATABASE FOR UPLOAD LIST» = «UniGene».**

1.4.3. Загрузите файл «**UniGene\_alz\_differ\_exp.txt**» и нажмите «**CONTINUE**».

1.4.4. Скопируйте значение **T-test p-value** в xls файл с результатами из OMIM в первую ячейку нового листа «**diff\_exp\_connectivity**».

1.4.5. Сделайте вывод какой сети стоит отдать приоритет в изучении на основании их связности.

## Шаг 5. Знакомство с интерфейсом AndVisio системы ANDSystem

1.5.1. Отметьте в сети, полученной на шаге 3, белки **APP** и **CD33**. Для этого в левой части окна нажмите на **«+»** рядом со словом **«Protein»**. В открывшемся списке найдите белки **APP** и **CD33**. После выделения **APP** удерживайте клавишу **Сtrl** и нажмите левой кнопкой мыши на **CD33**. Белки **CD33** и **APP** будут подсвечены синим цветом в списке и в сети.

1.5.2. В верхней части окна нажмите на кнопку **«Analysis»** и выберите **«Find shortest pathway»**.

1.5.3. В левом меню появятся кратчайшие пути, щёлкните на **«+»** рядом с **«Path 1»** и в раскрывшемся списке щёлкните на **«+»** рядом с **«Relation»** и **«association»** и выделите все взаоимодействия. С помощью **Select→Select Neighbors (Ctrl+N)** выделите все соседние вершины для этих взаимодействий. Если хорошо видите одну из вершин, потяните за неё чтобы вытащить весь путь от APP до CD33, если вершины плохо видно, нажмите **Select→Invert selection (Ctrl+I)** и клавишу **Delete**, после чего перенесите вершины, в то место экрана, которое не было занято графом и нажмите **Edit→Undo (Ctrl+Z)**.

1.5.4. Отметьте в сети белок **APOE**. В верхней части окна нажмите на кнопку **«Analysis»** и выберите **«Find circles»**.

1.5.5. В верхней части окна нажмите на кнопку «**Analysis**» и выберите «**Find fundamental rings**».

1.5.6. В верхней части окна нажмите на кнопку **«Analysis»** и выберите «**Statistics**», нажмите на «**Save data as …»** и сохраните статистику в файл «**OMIM\_S.xlsx**». Ознакомьтесь с выдачей программы.

1.5.7. Ознакомьтесь с возможностями функции **«Layout»** в верхней части окна.

1.5.8. Ознакомьтесь с возможностями функции **«Select»** в верхней части окна.

1.5.9. Ознакомьтесь с возможностями функции **«View»** в верхней части окна.

1.5.10. Сделайте пересечение двух графов «**Alz**» и «**OMIM**». Для этого нажмите в верхней части окна на кнопку **Edit→Intersect Graphs**.

1.5.11. Найдите взаимодействия между белками и генами общими для двух графов «**Alz**» и «**OMIM**». Для этого во вкладке с пересечением двух графов «**Alz**» и **«OMIM»** выделите все объекты (**Ctrl+a**) и нажмите в верхней части окна кнопку «**Expand**». Выберите «**level**» равный **0**. Нажмите «**OK**».

1.5.12. Добавьте в сеть пересечения объект заболевание «**Alzheimer disease**». Для этого нажмите в верхней части окна на кнопку **Edit→Add objects**. В окне **«Synonyms»** введите «**Glaucoma, Open-Angle**». Нажмите «**OK**». Остальные параметры оставьте без изменений. Нажмите «**OK**».

1.5.13. Сделайте объединение двух графов «**Alz**» и «**OMIM**». Для этого нажмите в верхней части окна на кнопку **Edit→Union Graphs**.

1.5.14. Скопируйте граф из пункта 1.5.12. Для этого выделите все объекты (**Ctrl+a**), нажмите в верхней части окна на кнопку **Edit→Copy (Ctrl+C)**. Откройте новую вкладку, нажав на **«+»**. Вставьте скопированные объекты, нажав на **Edit→Paste (Ctrl+V)**.

1.5.15. В скопированном графе отфильтруйте гены и все связи кроме связей типа «**association**». Для этого нажмите в верхней части окна на кнопку «**Filter**». Во вкладке «**Object type filter**» уберите галочку напротив **«Gene»**. Во вкладке **«Interaction type filter»** нажмите «**Uncheck all**» и поставьте галочку напротив «**association**». Нажмите «**OK**».

1.5.16. Уберите несвязанные объекты в сети из пункта 1.5.14. Для этого нажмите в верхней части окна на кнопку **Edit→Hide unconnected (Ctrl+H)**.

**Шаг 6. Поиск путей в генных сетях с помощью функции «Pathway Wizard» в системе ANDSystem.**

1.6.1. Откройте новую вкладку нажав на кнопку «+». В новой вкладке нажмите на иконку «**Pathway Wizard**». В открывшемся окне выберите параметр “**Pathway length**” равный 2.

1.6.2. Нажмите на квадрат с номером 1. В открывшейся форме введите “glaucoma\*”, в поле «**Type**» выберите “**Disease**” и нажмите кнопку “**Add**”. В открывшемся окне выберите “**Glaucoma, Open-Angle**” и переместите это название в правую часть окна, либо перетащив мышкой, либо нажав на кнопку **“>”**. Нажмите “**Ok**”. Нажмите “**Close**”.

1.6.3. Нажмите на квадрат со стрелкой с номером 2. В открывшемся окне во вкладке “**Database filter**” выберите все базы данных (“**Check all**”). Во вкладке “**Interaction type filter**” выберите все взаимодействия (“**Check all**”). Нажмите “**Close**”.

1.6.4. Нажмите на квадрат с номером 3. В открывшейся форме во вкладке “**Organism**” выберите “**Homo sapiens**”. Во вкладке “**Object filter**” поставьте галочку напротив “**Protein”**. Нажмите “**Close**”.

1.6.5. Нажмите на квадрат со стрелкой с номером 4. В открывшемся окне во вкладке “**Database filter**” выберите все базы данных (“**Check all**”). Во вкладке “**Interaction type filter**” выберите все взаимодействия (“**Check all**”). Нажмите “**Close**”.

1.6.6. Нажмите на квадрат с номером 5. В открывшейся форме введите “**obesity**”, в поле «**Type**» выберите “**Phenotype**” и нажмите кнопку “**Add**”. В открывшемся окне выберите “**Obesity**” и переместите это название в правую часть окна, либо перетащив мышкой, либо нажав на кнопку “>”. Нажмите “**Ok**”. Нажмите “**Close**”. Нажмите “**Find pathways**”.

1.6.7. Сохраните полученную сеть в файлы «**G\_O.and**» «**G\_O.stsv**» в форматах AND XML (.and) и Simple table format (.stsv), нажав на кнопку «**Save as**» в верхней части окна.

1.6.8. Самостоятельно найдите регуляторный контур взаимо-активации генов IRF1 и STAT1, приведенный на рисунке 1, с помощью функции “**Pathway Wizard**” системы ANDSystem.





Рисунок 1. Регуляторный контур взаимо-активации генов IRF1 и STAT1.

Для этого задайте в шаблоне длиной 4 следующих участников:

1. Белок IRF1

2. все типы взаимодействий.

3. ген STAT1

4. экспрессия (expression)

5. белок STAT1

6. регуляция экспрессии (expression regulation, expression downregulation, expression upregulation).

7. ген IRF1

8. экспрессия (expression)

9. белок IRF1

1.6.9. Сохраните полученную сеть в файлы «**IRF1\_STAT1.and**» «**IRF1\_STAT1.stsv**» в форматах AND XML (.and) и Simple table format (.stsv), нажав на кнопку «**Save as**» в верхней части окна.

**Шаг 7. Формирование отчета по этапу 1**

1.7.1. Создайте xls файл с названием **practice4\_ВашаФамилия.xls**

1.7.2. Назовите лист «**Alz**» и скопируйте в него таблицу из файла «**Alz.stsv**».

1.7.3. Назовите лист «**OMIM**» и скопируйте в него таблицу из файла «**OMIM.stsv**».

1.7.4. Назовите лист «**OMIM\_S**» и скопируйте в него таблицу из вкладки «**Object statistics**» файла «**OMIM\_S.xlsx**». Отсортируйте таблицу по убыванию показателя центральности «**Betweenness centrality**». Выделите цветом наиболее центральный ген или белок.

1.7.5. Назовите лист «**G\_O**» и скопируйте в него таблицу из файла «**G\_O.stsv**».

1.7.6. Назовите лист «**IRF1\_STAT1**» и скопируйте в него таблицу из файла «**IRF1\_STAT1.stsv**».

# *Этап 2. Поиск кластеров в ассоциативных генных сетях с помощью плагина ClusterViz для CytoScape.*

**Цель этапа:** знакомство с программой для кластеризации генных сетей ClusterViz.

**Шаг 1. Установка плагина ClusterViz для CytoScape.**

2.1.1. Запустите программу CytoScape. В верхней части окна нажмите на кнопку “Apps” → “App Manager”. Введите в поле “Search” название плагина “ClusterViz”. Установите его.f

**Шаг 2. Загрузка сети в CytoScape.**

2.2.1. В xls файле с названием **practice4\_ВашаФамилия.xls** создайте новый лист “**Clust**” и скопируйте рядом из листа «**OMIM**» колонки ObjName1 и ObjName2.

2.2.2. Сохраните эти две колонки в текстовый файл “Clust.txt”

2.2.3. Загрузите сеть из файла “**Clust.txt**” в программу CytoScape через **File → import → network → file**.

2.2.4. В открывшемся окне нажмите на «**ObjName1**» и нажмите на зеленый кружок (“**Source** **Node**”), затем нажмите на «**ObjName2**» и нажмите на красный с белым кружок (“**Target** **Node**”). Нажмите “**Ok**”.

**Шаг 3. Кластеризация сети с помощью плагина ClusterViz для CytoScape.**

2.3.1. Запустите ClusterViz через кнопку **Apps → ClusterViz → Open Viz**.

2.3.2. Выберите алгоритм кластеризации **FAG-EC** или **MCODE** и нажмите «**Analyze current network**».

2.3.3. Экспортируйте полученные кластеры с помощью кнопки “**Export**” в правой части окна в файл “**Clusters.txt”**.

**Шаг 4. Формирование отчета по этапу 2**

2.4.1. Убедитесь, что в xls файл с названием **practice4\_ВашаФамилия.xls** присутствует лист «**Clust**».

2.4.2. Создайте лист «**Clusters**» и скопируйте в него содержимое файла “**Clusters.txt”**.