

## Особенности функциональной аннотации генов предрасположенности к ревматоидному артриту при использовании Cytoscape

Часовских Н.Ю.

*Сибирский государственный медицинский университет (СибГМУ)  
Россия, 634050, г. Томск, Московский тракт, 2*

### РЕЗЮМЕ

**Цель.** Оценить функциональную аннотацию генов, ассоциированных с ревматоидным артритом, при разных параметрах инструмента ClueGO Cytoscape.

**Материалы и методы.** Гены предрасположенности к ревматоидному артриту были извлечены из публичной базы данных GWAS (каталог ассоциаций однонуклеотидных полиморфизмов с заболеваниями). Генная онтология (Gene Ontology, GO) – функциональная аннотация генов была произведена с помощью алгоритма, реализованного в Cytoscape ClueGO.

**Результаты.** Рассмотрены особенности выполнения функциональной аннотации с помощью плагина ClueGO при разных условиях формирования функциональных групп. В зависимости от исходных параметров, задаваемых в плагине, группировка терминов согласно генной онтологии осуществляется с различной долей обобщения. Меньшее минимальное число генов в группе позволяет сформировать большее число групп, что дает возможность получить более детальную функциональную характеристику.

**Заключение.** Полученные варианты результатов функциональной аннотации могут быть востребованы для дальнейших исследований генетических механизмов ревматоидного артрита.

**Ключевые слова:** функциональная аннотация генов, ревматоидный артрит, GWAS, Cytoscape.

**Конфликт интересов.** Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Источник финансирования.** Автор заявляет об отсутствии финансирования при проведении исследования.

**Для цитирования:** Часовских Н.Ю. Особенности функциональной аннотации генов предрасположенности к ревматоидному артриту при использовании Cytoscape. *Бюллетень сибирской медицины.* 2020; 19 (3): 101–104. <https://doi.org/10.20538/1682-0363-2020-3-101-104>.

---

## Features of functional annotation of rheumatoid arthritis susceptibility genes by Cytoscape

Chasovskikh N.Yu.

*Siberian State Medical University  
2, Moscow Trakt, Tomsk, 634050, Russian Federation*

### ABSTRACT

**Aim.** To evaluate the functional annotation of genes associated with rheumatoid arthritis with different parameters of the ClueGO Cytoscape tool.

**Materials and methods.** Genes of susceptibility to rheumatoid arthritis were extracted from publicly available database GWAS (catalog of associations of single nucleotide polymorphisms with diseases). The Gene Ontology (GO), the functional annotation of genes, was performed using Cytoscape ClueGO. The features of the functional annotation using the plugin ClueGO Cytoscape were analyzed.

---

✉ Часовских Наталья Юрьевна, e-mail: [natacch@mail.ru](mailto:natacch@mail.ru).

**Results.** Depending on the initial parameters specified in the plugin, the grouping of terms according to the gene ontology was carried out with a different degree of generalization. A smaller minimum number of genes in a group allows to form a larger number of groups, which makes it possible to obtain more detailed functional characteristics.

**Conclusion.** The results obtained with different grouping options can be useful for further studies of genetic mechanisms of rheumatoid arthritis.

**Key words:** functional annotation of genes, rheumatoid arthritis, GWAS, Cytoscape.

**Conflict of interest.** The author declares no obvious or potential conflict of interest related to publication of this manuscript.

**Source of financing.** The author states that there is no funding for the study.

**For citation:** Chasovskikh N.Yu. Features of functional annotation of rheumatoid arthritis susceptibility genes by Cytoscape. *Bulletin of Siberian Medicine*. 2020; 19 (3): 101–104. <https://doi.org/10.20538/1682-0363-2020-3-101-104>.

## ВВЕДЕНИЕ

Нарастание объемов геномных и протеомных исследований привело к необходимости аккумуляции их результатов в специализированных базах данных, в том числе на общедоступных онлайн-ресурсах. Огромную роль в систематизации, описании таких данных играет разработанная единая универсальная иерархичная система терминологии – генная онтология (Gene Ontology, GO) [1]. Она позволяет охарактеризовать данные по следующим разделам: биологические процессы, молекулярные функции, клеточные компоненты [2], в соответствии с которыми производится аннотация генов (белков). Чтобы возможности классификации GO были применимы к конкретному набору данных, существуют специализированные инструменты, особое место среди них занимает ClueGO Cytoscape [3], так как позволяет работать одновременно с несколькими списками данных и с сетями.

Возможность описания результатов исследований в терминах генной онтологии играет важную роль в осуществлении функциональной характеристики тысяч генов (например, при микрочипировании). Это позволяет оценивать возможный вклад изучаемых генов в реализацию физиологического ответа или этиопатогенез заболеваний, сгруппировав их соответствующим образом. Особое значение подобный подход имеет для изучения генетических факторов мультифакторных (комплексных) патологий [4], в частности ревматоидного артрита, относительно которого подобных исследований (основанных на результатах полногеномного поиска) не проводилось.

Цель данной работы – оценить функциональную аннотацию генов, ассоциированных с ревматоидным артритом, при разных параметрах инструмента ClueGO Cytoscape.

## МАТЕРИАЛЫ И МЕТОДЫ

Гены предрасположенности к ревматоидному артриту были получены с помощью анализа данных информационного ресурса, содержащего результаты полногеномных исследований GWAS [5] – ассоциаций между однонуклеотидными полиморфизмами и заболеваниями.

Для проведения функционального анализа в плагине ClueGO Cytoscape version 3.2.1 были выбраны типы онтологий, соответствующие сервису генной онтологии Gene Ontology. ClueGO позволяет обновлять данную информацию, поскольку информационный ресурс Gene Ontology является постоянно пополняемым [6].

Оценка функционального сходства генов осуществлялась с помощью гипергеометрического теста, при этом определялась принадлежность генов к конкретным функциям в терминах генной онтологии. Дополнительно плагин позволяет корректировать минимальное число (%) генов, используемых для формирования групп – по умолчанию 3 и 4 соответственно. Также на положительной шкале от 0 до 1 устанавливался уровень каппа-статистики, отражающей функциональные связи между генами (0,4). В случае тестирования большого числа гипотез ClueGO позволяет применять коррекцию для  $P$  (вероятности ошибки первого рода) методами Бонферрони и Беньямини – Хохберга [7]. При этом в таблице результатов представлены оба значения  $P$  для каждого термина.

Функциональные группы создавались путем итеративного слияния первоначально определенных групп на основе заранее установленного порогового значения каппа. При выборе «ведущего» термина в каждой группе программа предлагает ориентироваться (опционально) на их статистическую значимость, число или процент генов.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

С использованием информационного ресурса GWAS был сформирован список из 94 генов, ассоциированных с ревматоидным артритом: *ACOXL*, *AIRE*, *ALS2CR12*, *ANAPC4*, *ANO8*, *ANXA3*, *APOM*, *ARHGEF3*, *ARID5B*, *BAG6*, *BLK*, *BTNL2*, *C1orf159*, *C2*, *C5orf30*, *CBFB*, *CCL21*, *CCR6*, *CD226*, *CD247*, *CD40*, *CDK5RAP2*, *CDK6*, *CLYBL*, *CMKLR1*, *CTTNBP2*, *DDA1*, *DPP4*, *EOMES*, *FADS2*, *FAM107A*, *FAM98B*, *GATA3*, *GATSL3*, *GCHI*, *GRM5*, *GUCY1B2*, *HNRNPA1*, *IFI16*, *IL2RA*, *IL2RB*, *IL6R*, *IRF5*, *JAZF1*, *KIAA1109*, *KIF5A*, *LLGL1*, *MECP2*, *MMEL1*, *MOV10L1*, *MTF1*, *MTG1*, *MTM1*, *NFKBIE*, *OS9*, *PDE2A*, *PDGFA*, *PHF19*, *PHTF1*, *PLCL2*, *PLD4*, *PPIL4*, *PPM1D*, *PPM1E*, *PRKCH*, *PTPN2*, *PTPN22*, *RABEP1*, *RAD51B*, *RASGRP1*, *REL*, *RPP14*, *RTKN2*, *SFTPD*, *SPTBN2*, *STAG1*, *STAT4*, *SUOX*, *SYNGR1*, *TEC*, *TIMMDC1*, *TNFAIP3*, *TNPO3*, *TRHDE*, *TYK2*, *UBASH3A*, *UBE2F*, *VTCN1*, *WDFY4*, *WNT10B*, *YDJC*, *ZBTB12*, *ZNF133*, *ZNF774*.

Для классификации генов были использованы следующие типы онтологий: GO\_ImmuneSystemProcess (процесс иммунной системы), GO\_MolecularFunction (молекулярная функция), GO\_CellularComponent (клеточный компонент), GO\_BiologicalProcess (биологический процесс). Минимальный уровень GO Level равнялся 3, максимальный – 8.

Проведенный далее функциональный анализ генов (при заданном значении минимального числа генов в группах 2,  $p < 0,05$ ) выявил восемь групп генов в соответствии с терминами генной онтологии:

- 1) регуляция продукции интерлейкина (IL)-2 (включает 22 функции);
- 2) сигнальный путь IL-2 (15 функций);
- 3) антиген рецептор-опосредованный сигнальный путь (12 функций);
- 4) продукция IL-12 (пять функций);
- 5) позитивная регуляция G2/M перехода митотического цикла клеток (две функции);
- 6) регуляция нейрональной синаптической пластичности (две функции);
- 7) позитивная регуляция цитотоксичности, связанная с натуральными киллерами (две функции);
- 8) регуляция процессов респираторной системы (две функции).

Кроме того, было выявлено шесть функций, не подвергшихся объединению:

- сигнальный путь IL-6;
- хемотаксис дендритных клеток;
- нейромускульный контроль положения тела;
- негативная регуляция врожденного иммунного ответа;
- ответ на мурамилдипептид;
- регуляция активации тромбоцитов.

Для того чтобы определить наиболее информативные с точки зрения биологической интерпретации результаты, был проведен функциональный анализ генов, ассоциированных с ревматоидным артритом при другом уровне минимального числа генов в группах (три гена),  $p < 0,05$ .

В результате группировки генов было получено пять групп:

- 1) сигнальный путь IL-2 (включает 16 функций);
- 2) рецептор антиген-опосредованный сигнальный путь (12 функций);
- 3) негативная регуляция активации клеток (11 функций);
- 4) продукция IL-12 (три функции);
- 5) регуляция нейрональной синаптической пластичности (две функции).

Четыре функции были представлены отдельными терминами (без объединения):

- негативная регуляция врожденного иммунного ответа;
- дыхательный газовый обмен;
- регуляция метилирования гистонов;
- сигнальный путь рецептора тромбоцитарного фактора роста.

Полученные результаты свидетельствуют, что в зависимости от исходных параметров, задаваемых в плагине ClueGO Cytoscape, группировка терминов генной онтологии, соотнесенных с генами, осуществляется с различной долей обобщения. Так, в первом случае (при минимальном числе генов в группе 2) было сформировано большее число групп по сравнению с таковым во втором варианте анализа (при минимуме генов в группах 3), что позволило получить более детальную функциональную характеристику. При этом для некоторых функций, выявленных про обоих вариантах исследования, число генов в группах было меньше в первом случае по сравнению со вторым. Так, для продукции IL-12 соответствующее число генов равнялось пяти (*CD40*, *CMKLR1*, *IRF5*, *NFKB1*, *REL*) и шести (*CD40*, *CMKLR1*, *IRF5*, *NFKB1*, *REL*, *TRAF3*). Для сигнального пути IL-2 аналогичные показатели составили 14 (*BPI*, *BTNL2*, *CCL21*, *CCR6*, *CMKLR1*, *GATA3*, *IL2RA*, *IL2RB*, *IL6R*, *NFKB1*, *PDGFA*, *PTGIS*, *PTPN2*, *PTPN22*) и 19 (*BPI*, *BTNL2*, *CCL21*, *CCR6*, *CMKLR1*, *GATA3*, *IL2RA*, *IL2RB*, *IL6R*, *NFKB1*, *PDGFA*, *PTGIS*, *PTPN2*, *PTPN22*, *RORA*, *SFTPD*, *TNFAIP3*, *UBASH3A*, *VTCN1*) генов соответственно. При этом для функций антигена рецептор-опосредованного сигнального пути и регуляции нейрональной синаптической пластичности число генов при обоих вариантах исследования оставалось одинаковым.

При детализированной функциональной аннотации в первом варианте исследования выявлена

представленность генов предрасположенности к ревматоидному артриту в следующих функциях регуляции иммунного ответа: сигнальном пути IL-6, который является ключевым цитокином, ответственным за аутоиммунное воспаление [8, 9]; регуляции хемотаксиса дендритных клеток; ответе на мурамилдипептид (элемент клеточной стенки бактерий, активирующий системы как врожденного, так и приобретенного иммунитета). Также определена принадлежность генов функциям позитивной регуляции G2/M перехода митотического цикла клеток, нейромускульного контроля положения тела.

Укрупненные функциональные группы, полученные во втором случае, отражают общую картину, характерную для предыдущего результата: выявлено участие генов в сигнальном пути IL-2 и IL-12. Известно, что IL-2 важен для процесса развития иммунного ответа, так как осуществляет стимулирование клеток-киллеров [10]. IL-12 обладает выраженными провоспалительными свойствами, повышает активность натуральных киллеров, дендритных клеток, по совокупности действий осуществляя связь между врожденным и приобретенным иммунитетом [10–12]. Помимо этого, функциональная группа регуляции нейрональной синаптической пластичности свидетельствует о возможном влиянии генов на процесс нейрональной обработки синаптического сигнала.

## ЗАКЛЮЧЕНИЕ

Полученные результаты свидетельствуют, что при ревматоидном артрите гены предрасположенности влияют не только на реализацию иммунного ответа, опосредованную сигналингом провоспалительных цитокинов (IL-2, -6, -12) и регуляцией иммунцитов, а также на функции нервной системы, в частности обработку синаптического сигнала, нейромускульный контроль положения тела.

Для исследования возможных механизмов заболеваний или физиологических процессов могут быть важны детали относительно участия отдельных сигнальных путей и клеточных реакций. Для этого целесообразно изменять минимальное число генов, объединяемых в функциональную группу в сторону уменьшения (по сравнению со значением, установленным в плагине по умолчанию). Вместе с тем укрупненные группы функциональных характеристик

генов демонстрируют большую наглядность при выявления общих тенденций биологических процессов.

## ЛИТЕРАТУРА

1. Blake A., Christie K.R., Dolan M.E. Gene Ontology Consortium: going forward. *Nucleic Acids Research*. 2015; 43 (1): 1049–1056. DOI: 10.1093/nar/gku1179.
2. Ashburner M., Ball C., Blake J., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M., Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 2000; 25 (1): 25–29. DOI: 10.1093/nar/gku1179.
3. Bindea G., Mlecnik B., Hackl H., Charoentong P., Tosolini M., Kirilovsky A., Fridman W.H., Pagès F., Trajanoski Z., Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009; 25 (8): 1091–1093. DOI: 10.1093/bioinformatics/btp101.
4. Hindorf L.A., Sethupathy P., Junkins H.A., Ramos E.M., Mehta J.P., Collins F.S., Manolio T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *PNAS*. 2009; 106 (23): 9362–9367. DOI: 10.1073/pnas.0903103106.
5. Welter D., MacArthur J., Morales J., Burdett T., Hall P., Junkins H., Klemm A., Flicek P., Manolio T., Hindorf L., Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*. 2014; 42 (1): 1001–1006. DOI: 10.1093/nar/gkt1229.
6. A catalog of published genome-wide association studies. URL: <http://www.genome.gov/gwastudies/>
7. Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B*. 1995; 57 (1): 289–300. DOI: 10.2307/2346101.
8. Удачкина Е.В., Новикова Д.С., Попкова Т.В., Насонов Е.Л. Роль интерлейкина 6 в развитии атеросклероза при ревматоидном артрите. *Современная ревматология*. 2013; 7 (3): 25–32.
9. Tanaka T., Narazaki M., Kishimoto T. IL-6 in inflammation, immunity, and disease. *Cold Spring Harb. Perspect. Biol.* 2014; 6 (10): a016295. DOI: 10.1101/cshperspect.a016295.
10. Ross S.H., Cantrell D.A. Signaling and function of interleukin-2 in T lymphocytes. *Annu. Rev. Immunol.* 2018; (36): 411–433. DOI: 10.1146/annurev-immunol-042617-053352.
11. Новиков А.А., Александрова Е.Н., Диатроптова М.А., Насонов Е.Л. Роль цитокинов в патогенезе ревматоидного артрита. *Научно-практическая ревматология*. 2010; 48 (2): 71–82.
12. Vignali D.A., Kuchroo V.K. IL-12 family cytokines: immunological playmakers. *Nat. Immunol.* 2012; 13 (8): 722–728. DOI: 10.1038/ni.2366.

## Сведения об авторе

**Часовских Наталия Юрьевна**, д-р мед. наук, профессор, кафедра медицинской и биологической кибернетики, СибГМУ, г. Томск. ORCID 0000-0001-6077-0347.

(✉) **Часовских Наталия Юрьевна**, e-mail: [natacch@mail.ru](mailto:natacch@mail.ru).

Поступила в редакцию 01.04.2020

Подписана в печать 16.06.2020