

ГАПЛОТИПЫ TNFA-HLA II КЛАССА КАК ПРОГНОСТИЧЕСКИЕ МАРКЕРЫ РЕВМАТОИДНОГО АРТРИТА

Сташкевич Д.С.¹, Хромова Е.Б.¹, Девальд И.В.^{1,2}, Ходус Е.А.³, Филиппова Ю.Ю.¹,
Бурмистрова А.Л.¹

¹Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет», Челябинск, e-mail: stashkevich_dary@mail.ru;

²Федеральное государственное бюджетное образовательное учреждение высшего образования «Южно-Уральский государственный медицинский университет», Челябинск;

³Клиника профессора Кинзерского, Челябинск

Ревматоидный артрит (РА) – это хроническое воспалительное заболевание соединительной ткани, характеризующееся прогрессирующим поражением суставов и специфическими системными нарушениями, с доказанной генетической предрасположенностью и вовлечением факторов окружающей среды. В статье приведен анализ ассоциации расширенных гаплотипов, образованных однонуклеотидными полиморфизмами промоторной области гена TNFA и генными семействами локусов HLA-DRB1, HLA-DQA1, HLA-DQB1, с предрасположенностью к ревматоидному артриту в популяции русских Челябинской области. Расчет частот гаплотипов и параметров сцепления проводился с помощью программы Arlequin ver. 3.5. Достоверность различий оценивалась с использованием критериев χ^2 , OR с расчетом 95%-ного доверительного интервала. С применением метода логистической регрессии получена модель на основе распределения двухлокусных гаплотипов с относительно высокой диагностической эффективностью 82,33%, позволяющая определить предрасположенность к ревматоидному артриту у русских Челябинской области. Полученная логит-модель свидетельствует о том, что значимое влияние на формирование предрасположенности к ревматоидному артриту у русских Челябинской области имеет двухлокусный гаплотип TNFA -1031*t- HLA DRB1*04, содержащий функционально значимый аллельный вариант TNFA -1031*t и генную специфичность HLA DRB1*04, являющуюся главным фактором восприимчивости к РА.

Ключевые слова: гаплотипы TNFA-HLA II класса, ревматоидный артрит, неравновесное сцепление, предрасположенность, модель логистической регрессии.

CLASS II TNFA-HLA HAPLOTYPES AS PREDICTIVE MARKERS OF RHEUMATOID ARTHRITIS

Stashkevich D.S.¹, Khromova E.B.¹, Devald I.V.^{1,2}, Khodus E.A.³, Filippova Y.Y.¹,
Burmistrova A.L.¹

¹Chelyabinsk State University, Chelyabinsk, e-mail: stashkevich_dary@mail.ru;

²South Ural State Medical University, Chelyabinsk;

³Clinic of Professor Kinzersky, Chelyabinsk

Rheumatoid arthritis (RA) is a chronic inflammatory disease of the connective tissue characterized by progressive joint damage and specific systemic disorders, with a proven genetic predisposition and involvement of environmental factors. The article analyzes the association of extended haplotypes formed by single nucleotide polymorphisms of the TNFA gene promoter region and gene families of loci HLA-DRB1, HLA-DQA1, HLA-DQB1, with a predisposition to rheumatoid arthritis in the Russian population of the Chelyabinsk region. The haplotype frequencies and linkage parameters were calculated using the Arlequin ver. 3.5. The significance of differences was assessed using the χ^2 , OR criteria with the calculation of a 95% confidence interval. Using the logistic regression method, a model was obtained based on the distribution of two-locus haplotypes with a relatively high diagnostic efficiency of 82.33%, which makes it possible to determine the predisposition to rheumatoid arthritis in Russians of the Chelyabinsk region. The resulting logit model indicates that the two-locus haplotype TNFA-1031 * t-HLA DRB1 * 04, which contains a functionally significant allelic variant of TNFA-1031 * t and the gene specificity of HLA DRB1, has a significant effect on the formation of a predisposition to rheumatoid arthritis in Russians of the Chelyabinsk region * 04 which is the main factor of susceptibility to RA.

Keywords: haplotypes TNFA-HLA II class, rheumatoid arthritis, linkage disequilibrium, predisposition logistic regression model.

Одним из актуальных направлений современного здравоохранения является развитие персонафицированной медицины. Основным подходом в случае изучения мультифакторных патологий служит предсказание на основе геномных данных вероятности возникновения того или иного заболевания и его диагностики с учетом индивидуальных показателей пациента, в том числе биомаркеров различной молекулярной природы. Среди таких заболеваний в области ревматологии лидирующее место занимает ревматоидный артрит (РА) – это хроническое воспалительное заболевание соединительной ткани, характеризующееся прогрессирующим поражением суставов и специфическими системными нарушениями. На сегодняшний день число пациентов с аутоиммунным артритом составляет почти 1% населения мира, что создает актуальную проблему для систем здравоохранения. Очевидной особенностью РА является клинический полиморфизм, представленный широким разнообразием симптомов, клинических форм и темпов прогрессирования. РА считается многофакторным заболеванием с доказанной генетической предрасположенностью и вовлечением факторов окружающей среды [1].

Данные *genome-wide association studie* (GWAS) свидетельствуют, что основные гены предрасположенности к РА принадлежат локусу главного комплекса гистосовместимости (МНС). Этот локус, охватывающий 4 Мбр, находится в области 6р21.3 и содержит около 250 генов, кодирующих антигенные белки для распознавания Т-клетками. Считается, что до 60% возможных полиморфизмов (генетических вариантов), ответственных за восприимчивость к РА, происходят из локуса МНС [2]. Исторически первыми кандидатами были гены и аллели локуса HLA DRB1, позже, с развитием методов молекулярной биологии, были выявлены маркерные аллели в локусах DQA1, DQB1. Анализ литературных данных показывает, что в различных популяциях мира установлены предрасполагающие аллели, кодирующие позиции ассоциированных с РА аминокислот, которые расположены в середине эпитоп-связывающих карманов молекул HLA. Таким образом, замены могут напрямую влиять на антигенную специфичность и приводить к связыванию белка с остатками цитруллина как основного патогенетического фактора развития аутоиммунитета. Так, в наших работах определены ассоциативные (HLA DRB1*01, *03, *04) и протективные гены (DRB1*11, *13) в русской популяции Челябинской области. Другая группа геномных маркеров предрасположенности к РА – это гены цитокинов и их рецепторов как ключевых компонентов сигнальных путей и ко-стимулирующих факторов. Возможные механизмы, которые реализуются через различные SNP генов цитокинов, согласно литературным данным, – это влияние на воспалительные реакции и дифференцировку клеток [3, 4]. У больных ревматоидным артритом русской популяции Челябинской области нами установлено, что генотип T/C – 1031 TNF α – маркер предрасположенности к ревматоидному артриту у женщин, наличие

системных проявлений ассоциировано с гетерозиготным генотипом G/A – 308 TNF α [5]. Поскольку гены HLA I, II и III класса локализируются на одной и той же хромосоме, вероятнее всего полиморфизмы TNFA находятся в неравновесных сцеплениях с локусами –A, –B, –B1, –DQA1, –DQB1 системы HLA и формируют определенные сочетания – гаплотипы. В группе больных РА русской популяции определен ряд гаплотипов TNFA-HLA DRB1, TNFA-HLA DQA1, TNFA-HLA DQB1 [6], частоты которых были повышены, и по критерию отношения шансов они могут рассматриваться как потенциальные биомаркеры предрасположенности к РА. В продолжение исследования новой задачей является поиск комплекса связанных факторов, ассоциированных с развитием ревматоидного артрита, из уже установленных для данной популяции частот гаплотипов с использованием метода логистической регрессии. Цель: построение модели логистической регрессии на основе гаплотипов TNFA-HLA DRB1, TNFA-HLA DQA1, TNFA-HLA DQB1, позволяющей выявить наличие/отсутствие ассоциации с предрасположенностью к ревматоидному артриту у русских Челябинской области.

Материалы и методы исследования. Контингент обследуемых лиц. В исследование включены больные с верифицированным диагнозом «ревматоидный артрит» (98 человек, из них: женщин – 81, мужчин – 17, средний возраст – 51,7 \pm 1,3 года). Группа сравнения (132 человека: мужчин – 73, женщин – 59, средний возраст – 34,6 \pm 0,75 года) составлялась на основе популяционной выборки, сформированной случайным образом на основе ДНК-банка потенциальных доноров костного мозга ГБУЗ «Челябинская областная станция переливания крови», проживающих в г. Челябинске и Челябинской области. Популяционная принадлежность определялась по данным генеалогического анамнеза до третьего поколения (согласно рекомендациям 8-го Международного Симпозиума в 1980 г., Лос-Анджелес, США) [5].

Молекулярно-генетические методы определения генетического полиморфизма промоторной области TNFA HLA II класса подробно представлены в нашей статье [6].

В исследовании использовались следующие методы и подходы, принятые для анализа частотных данных. Оценка неравновесного сцепления (D' , частоты гаплотипов, достоверность сцепления) проводилась при помощи EM-алгоритма в программе Arlequin ver 5. [7]

При анализе достоверности различий между группами больных РА и условно здоровых лиц использовались стандартные для иммуногенетики критерий χ^2 , критерий Фишера для одной степени свободы, критерий отношения шансов с расчетом 95%-ного доверительного интервала. Различия считались достоверными при $p \leq 0,05$.

Для построения прогностической модели исхода применялся метод логистической регрессии, выполненный с помощью программного обеспечения MedCalc Version 20 [8]. В анализ были включены гаплотипы, частоты которых были достоверно повышены в группе больных, по критерию отношения шансов их носительство повышает вероятность развития РА.

Для оценки качества исследуемой модели использовали такие параметры, как: диагностическая эффективность, чувствительность и специфичность. Модель логистической регрессии имеет вид:

$$\boxed{\text{Логит}(P) = \text{Константа} + a_1X_1 + a_2X_2 + \dots + a_nX_n}$$

где P – вероятность отнесения пациента к группе заболевания,

$$\text{Логит}(P) = \ln\left(\frac{P}{1-P}\right);$$

a_i – коэффициенты регрессии (равны \ln ОШ (отношения шансов) для каждой из независимых переменных);

X_i – значения вошедших в модель показателей.

Зная коэффициенты регрессии, можно рассчитать $\text{Логит}(P)$ и далее – само значение P :

$$\boxed{P = \left(\frac{1}{1 + e^{-\text{Логит}(P)}}\right)}$$

Если $p < 0,5$, обследуемого следует причислить к группе 0 (отсутствие риска), если $p > 0,5$ – к группе 1 (наличие риска).

Для определения конкретного значения показателя, который вносит большой вклад в значимость модели, использовали ROC-анализ и рассчитывали площадь под ROC-кривой. Для количественной оценки ROC-анализа применяли показатель AUC (area under curve, численное значение клинической значимости диагностического теста). Показатель площади под ROC-кривой выше 0,7 свидетельствует о высоком качестве диагностического теста [9].

Результаты исследования и их обсуждение. Первый этап исследования заключался в оценке достоверных изменений в частотах встречаемости гаплотипов TNFA-HLADRB1, TNFA-HLADQA1, TNFA-HLADQB1 в группах больных ревматоидным артритом и сравнения. Показано наличие неравновесного сцепления и между однонуклеотидными полиморфизмами в точках -1031 Т/С, -863 С/А, 308G/А, -238 G/А и генными семействами в локусах HLADRB1, DQA1, DQB1. При сравнении двух групп (больных и условно здоровых лиц) установлены гаплотипы, частоты которых повышены в группе больных, по критерию OR их носительство повышает вероятность развития РА. Таких гаплотипов получилось 11,

для последующего включения в модель все они закодированы с использованием букв английского алфавита (табл. 1).

Таблица 1

Названия гаплотипов, включенных в модель

гаплотип	OR, 95% CI	Обозначение для модели
TNF -1031*c-DQA1*05:01	11,7 3,4–40,6	a
TNF -1031*t-DQB1*03:02	7,038 3,32–14,92	b
TNF -1031*c-DQB1*02:01	6,18 1,74– 21,99	c
TNF -1031*c-HLADRB1*03	5,39 1,76–16,49	d
TNF -1031*t-HLADRB1*04	7,0 3,9–12,6	e
TNF-238*g - HLADRB1*04	6,69 3,7–12,01	f
TNF -308*g-DQA1*03:01	4,38 2,68–7,15	g
TNF -308*g-DQB1*03:02	3,2 1,79 – 5,74	h
TNF -308*g-DRB1*04	7,32 4,05–13,25	i
TNF -863*a DRB1*03	17,2 2,2–133,1	j
TNF -863*a-DQA1*05:01	28,23 3,75–212,80	k

На втором этапе на основе всех значимых гаплотипов построена модель с высокой статистической значимостью ($\chi^2_{mi}=131,97$, $df=12$, $p<0,0001$), относительно высокой диагностической эффективностью – 82,33%, чувствительностью 72,53% и специфичностью 89,52%, что позволяет определить высокий риск формирования РА. Получившаяся модель логистической регрессии развития РА в зависимости от носительства двухлокусных гаплотипов TNFA- HLA II класса и пола больных имеет вид:

$\text{logit}(p) = -2,88 - 39,68*a + 23,49*b - 1,2*c + 40,0*d + 2,38*e - 21,11*f + 1,98*g - 23,39*h + 20,04*I - 40,71*j + 63,05*k + 2,16*l$. Характеристики модели представлены в таблице 2.

Таблица 2

Параметры и характеристика модели логистической регрессии развития РА в зависимости от носительства гаплотипов и пола

Показатели	Коэффициент регрессии	Стандартная ошибка	P	OR	95% CI
TNFA -1031*c – HLA DQA1*05:01	-39,67952	29481,55	0,99	5,8533E-18	
TNFA -1031*t – HLA DQB1*03:02	23,49184	13001,68	0,99	15,9E+009	
TNFA -1031*c – HLA DQB1*02:01	-1,20699	2,31	0,60	0,29	0,003 – 27,48
TNFA -1031*c – HLA DRB1*03	40,00378	29481,55	0,99	236E+015	
TNFA -1031*t- HLA DRB1*04	2,37912	0,98	0,016	10,79	1,57 – 74,23
TNFA -238*g – HLA DRB1*04	-21,11167	20130,99	0,99	6,7814E-10	
TNFA -308*g – HLA DQA1*03:01	1,97768	1,08	0,067	7,23	0,87 - 59,84
TNFA -308*g – HLA DQB1_03_02	-23,39209	13001,68	0,99	6,9334E-11	
TNFA -308*g – HLA DRB1*04	20,03676	20130,99	0,99	503E+006	
TNFA -863*a- HLA DRB1*03	-40,71439	30914,18	0,99	2,0795E-18	
TNFA -863*a- HLA DQA1*05:01	63,05149	32847,32	0,99	2,415E+27	
Пол	2,16433	0,49	<0,0001	8,71	3,27 - 23,19
Константа	-2,88				
Чувствительность	Модель полностью: $\chi^2_{ml}=131,97$, $df=12$, $p<0,0001$ Чувствительность – 72,53%, специфичность – 89,52% Диагностическая эффективность – 82,33% Площадь под ROC-кривой – 0,893±0,022 (95% CI: 0,84–0,93)				

Модель была проверена на 16 людях, отобранных случайным образом из исследуемых групп (8 больных РА, 8 условно здоровых лиц). Проверка показала, что только 62,5% больных РА совпали с данной моделью, что оказалось ниже полученной чувствительности модели. Однако процент правильно диагностированных случаев среди условно здоровых лиц составил 87,5%, что соответствует заявленной в модели специфичности. Несмотря на полученные результаты и высокую значимость, применение полной модели в клинической практике пока затруднительно, вероятно, требуются дальнейшие исследования с увеличением исследуемых групп. Однако было установлено, что наибольшей статистической значимостью обладали три показателя: пол и два гаплотипа: TNFA -1031*t- HLA DRB1*04 и TNFA -308*g – HLA DQA1*03:01. Интересно, что разные аллельные варианты TNFA

включают в себя HLA DRB1*04 и DQA1*03:01, для которых показана предрасполагающая роль в развитии РА в составе самостоятельного устойчивого гаплотипического сочетания HLA DRB1*04-DQA1*03:01.

Кроме гаплотипов, значимым фактором, формирующим предрасположенность к РА, оказался пол больных ($p < 0,0001$). Полученные нами результаты о значимости пола при формировании модели предрасположенности к ревматоидному артриту подтверждают литературные данные о повышенной частоте встречаемости ревматоидного артрита у женщин, среди заболевших РА соотношение женщины : мужчины составляет от 3:1 до 4:1 [2, 3].

Нами также проведен ROC-анализ для каждого из трех статистически значимых показателей модели. Установлено, что гаплотип TNFA -1031*t- HLA DRB1*04 с чувствительностью 79,7% и специфичностью 74,0% (площадь под ROC-кривой 0,731 (95% CI: 0,67–0,79)) позволяет выявить предрасположенность к ревматоидному артриту. Полученная логит-модель свидетельствует о том, что значимое влияние на формирование предрасположенности к ревматоидному артриту у русских Челябинской области оказывает двухлокусный гаплотип TNFA -1031*t- HLA DRB1*04, содержащий функционально значимый аллельный вариант TNFA -1031*t и генную специфичность HLA DRB1*04, являющуюся главным фактором восприимчивости к РА. Модель показала, что в развитие данной аутоиммунной патологии вносит вклад комплекс генетических факторов, не ограничивающихся локусами HLA, и позволила выявить межгенные связи. Использование в практике различных вариантов анализа данных, в частности построение модели логистической регрессии, позволяет находить новые зависимости, интересные взаимосвязи и направить ход исследования в перспективное русло. Комплексное выявление максимально возможных предикторов даст возможность клиницисту использовать эти данные для прогноза развития и течения РА.

Заключение. Таким образом, нами протестирован один из вариантов построения логистической регрессионной модели риска развития ревматоидного артрита, на основании которой идентифицированы главные генетические факторы и их комбинации в локусах генов TNFA и HLA II класса. Полученные результаты могут служить предпосылкой для оценки риска развития ревматоидного артрита и включены в разработку комплекса профилактических мероприятий, что и является одним из подходов персонифицированной медицины. Одним из сложным моментов разработки таких моделей служит то, что развитие данного заболевания является результатом комплексного воздействия факторов и наличие генетической предрасположенности не является однозначным предиктором. Прогнозирование на основе модели логистической регрессии позволяет провести анализ всех

доступных для анализа факторов и более точно осуществить ранний этап первичной профилактики для данной патологии.

Исследование выполнено при финансовой поддержке РФФИ и Челябинской области в рамках научного проекта № 20-415-740013.

Список литературы

1. Karami J., Aslani S., Jamshidi A., Garshasbi M., Mahmoudi M. Genetic implications in the pathogenesis of rheumatoid arthritis; an updated review. *Gene*. 2019. vol. 702. P. 8–16.
2. Messemaker T.C., Huizinga T.W., Kurreeman F. Immunogenetics of rheumatoid arthritis: Understanding functional implications. *Journal of Autoimmunity*. 2015. vol. 64. P. 74–81.
3. Terao C., Raychaudhuri S., Gregersen P.K. Recent Advances in Defining the Genetic Basis of Rheumatoid Arthritis. *Annual review of genomics and human genetics*. 2016. vol. 31. no. 17. P. 273-301.
4. Deane K.D., Demoruelle M.K., Kelmenson L.B., Kuhn K.A., Norris J.M., Holers V.M. Genetic and environmental risk factors for rheumatoid arthritis. *Best Practice & Research: Clinical Rheumatology*. 2017. vol. 31. no.1. P. 3-18.
5. Сташкевич Д.С. Аллели и генотипы генов основных цитокинов и их межгенные и внутригенные связи в ассоциации с ревматоидным артритом у русской и башкирской популяций: автореф. дис. ... канд. биол. наук. Челябинск, 2009. 22 с.
6. Сташкевич Д.С., Хромова Е.Б., Девальд И.В., Ходус Е.А., Бурмистрова А.Л. Расширенные гаплотипы на основе редких однонуклеотидных полиморфизмов TNFA и HLA DRB1 в ассоциации с ревматоидным артритом// *Российский иммунологический журнал*. 2021. Т. 24. № 4. С. 555-562.
7. Excoffier L. Lischer H.E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 2010. vol. 10. no. 3. P. 564-567.
8. MedCalc: User-friendly statistical software. [Электронный ресурс]. URL: <https://www.medcalc.org/>(дата обращения: 18.12.2021).
9. Филиппова Ю.Ю. Характеристика паттернов информативных лабораторных показателей развития системного воспалительно-инфекционного процесса у больных с тяжелыми термическими травмами: автореф. дис. ... канд. биол. наук. Челябинск, 2013. 22 с.