

РАЗРАБОТКА ИММУНОГЕНЕТИЧЕСКОЙ МОДЕЛИ РИСКА РАЗВИТИЯ ТУБЕРКУЛЕЗА ЛЕГКИХ МЕТОДОМ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ

Беляева С.В.^{1,2}, Сташкевич Д.С.¹, Сулова Т.А.^{1,2}, Бурмистрова А.Л.¹

¹ФГБОУ ВО «Челябинский государственный университет», Челябинск, e-mail: shshvetlana@yandex.ru;

²ГБУЗ «Челябинская областная станция переливания крови», Челябинск

Не вызывает сомнений, что развитие одного из самых распространенных заболеваний в мире – туберкулеза – определяется не только условиями среды, свойствами возбудителя, но и способностью организма хозяина противостоять инфекциям за счет активации собственного иммунитета. Несмотря на многочисленные исследования, в настоящее время нет маркеров, которые однозначно могли бы помочь в диагностике заболевания, и туберкулез остается одной из основных причин смертности в мире. Часто туберкулез диагностируют уже в распространенных формах, что ведет к более длительному и агрессивному лечению. С помощью метода логистической регрессии мы рассмотрели возможность раннего прогнозирования развития тяжелых или легких форм заболевания на основе генетических характеристик больного. В предыдущих исследованиях нами были определены маркеры – гены HLA, цитокины и TLR и их комбинации, по которым исследуемые группы больных туберкулезом различались статистически значимо. В данной работе мы использовали ранее данные для прогнозирования риска развития определенных фенотипов туберкулеза. Анализ полученных логит-моделей свидетельствует о том, что значимое влияние на развитие разных форм заболевания оказывает носительство генов HLA: для очагового туберкулеза – аллеля HLADQB1*06:02-8, а тяжелого фиброзно-кавернозного – HLA B*08 и HLA DRB1*03. Это позволит разработать методы профилактики развития тяжелых форм и подобрать схемы лечения с целью персонализированного подхода.

Ключевые слова: туберкулез легких, полиморфизм генов, HLA, TLR, цитокины, логистическая регрессия.

THE CREATION OF IMMUNOGENETIC MODEL OF THE RISK OF DEVELOPMENT OF PULMONARY TUBERCULOSIS BY THE METHOD OF LOGISTIC REGRESSION

Belyaeva S.V.^{1,2}, Stashkevich D.S.¹, Suslova T.A.^{1,2}, Burmistrova A.L.¹

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

²Chelyabinsk Regional Blood Transfusion Station, Chelyabinsk

There is no doubt that the development of tuberculosis, which is one of the most common diseases in the world, is determined not only by environmental conditions, the properties of the pathogen, but also by the ability of the host organism to resist infections by activating its own immunity. Despite numerous studies, currently there are no markers that could definitely help in the diagnosis of the disease, and tuberculosis remains one of the leading causes of death in the world. Tuberculosis is often diagnosed in more common forms, which leads to longer and more aggressive treatment. Using the logistic regression method, we examined the possibility of early prediction of the development of severe or mild forms of the disease based on the patient's genetic characteristics. In previous studies, we identified markers - genes HLA, cytokines and TLR and their combinations, according to which the studied groups of patients with tuberculosis differed significantly. In this work, we used the data obtained earlier to predict the risk of developing certain phenotypes of tuberculosis. The analysis of the obtained logit models indicates that the carriage of HLA genes has a significant effect on the development of various forms of the disease: for focal tuberculosis – the HLADQB1*06:02-8 allele, and for severe fibro-cavernous – HLA B * 08 and HLA DRB1 * 03. This will allow developing methods for preventing the development of severe forms and choosing treatment regimens with the aim of a personalized approach.

Keywords: pulmonary tuberculosis, gene polymorphism, HLA, TLR, cytokines, logistic regression.

Туберкулез (ТБ) – хроническое инфекционное заболевание, вызываемое *Mycobacterium tuberculosis*, которое приобрело характер пандемии и представляет серьезную угрозу здоровью человека, поскольку является одной из основных причин смертности в мире. Согласно статистике около трети населения земного шара инфицировано микобактериями, но болеют лишь 1/10, что обусловлено наследуемой устойчивостью к

развитию активных форм заболевания [1]. Это определяется способностью организма хозяина противостоять инфекции за счет активации собственного иммунитета, уровень которого в значительной степени зависит от генотипа человека. В настоящее время генетические маркеры уже используются для корректировки лечения и постановки диагноза при многих заболеваниях [2]. Исследования показывают, что туберкулез является генетически детерминированным инфекционным заболеванием и генетический полиморфизм генов хозяина служит одним из механизмов, который приводит к прогрессированию от инфицирования к активным формам [1], а также к формированию многообразия клинических фенотипов туберкулеза легких: от трудно диагностируемых легких форм с небольшими очаговыми гранулемами и без бактериовыделения до тяжелых форм с обширными полостями и высокой бактериальной нагрузкой, заканчивая образованием внелегочных комплексов и распространенных процессов с диссеминацией возбудителя [3]. Известно большое количество потенциальных генетических маркеров, ассоциированных с развитием туберкулеза. Это гены Toll-like рецепторов, цитокинов (IL6, IL1, IFN γ , TNFA, IL10, IL12) и системы HLA (Human Leukocyte Antigens) [4]. Исследования генетических вариантов, продукты которых играют важную роль при туберкулезе, доказывают, что генетические компоненты, ассоциированные с предрасположенностью/устойчивостью к заболеванию, определяются комбинациями нескольких предрасполагающих генов, а не одним геном [4]. Сегодня ясно, что гены и их продукты образуют генетические комплексы, взаимодействие которых и их экспрессия определяются окружением и внутренним состоянием организма, демонстрируя многообразие фенотипических проявлений в зависимости от качества сигналов. На основе создания моделей предрасположенности к развитию определенных форм туберкулеза с помощью метода логистической регрессии возможно проводить анализ ассоциаций полиморфизмов с риском развития тяжелых или легких форм. Это создает предпосылки для выделения среди больных групп повышенного риска и разработки комплекса профилактических мероприятий для конкретного индивидуума с целью повышения устойчивости к прогрессированию инфекции.

Цель исследования: построение моделей логистической регрессии на основе анализа комплекса полиморфизмов генов наследуемого иммунитета (HLA, цитокинов и TLR) для оценки риска развития различных форм туберкулеза легких.

Материал и методы исследования

Выборка больных туберкулезом легких состояла из 77 больных русской популяции, пациентов ГБУЗ «Челябинский областной клинический противотуберкулезный диспансер» из Челябинска и Челябинской области. Больные имели активную форму туберкулеза и

относились к социально адаптированным слоям населения. Средний возраст больных составил $47,38 \pm 1,22$ года. Согласно Российской клинической классификации (Российская клиническая классификация, Приказ МЗ РФ от 21.03.2003 г. № 109) были выделены группы больных с разными по тяжести формами туберкулеза легких: с очаговой, инфильтративной и фиброзно-кавернозной.

Выделение геномной ДНК выполняли из крови на колонках согласно инструкции производителя «Ахуген», США. Генотипирование SNPs генов цитокинов и TLR осуществляли следующим образом: методом ПЦР в гене TLR4(299)*G/A (реактивы ООО НПФ «Литех», Москва), методом ПДРФ в генах TNFA(-308)*G/A и TNFA(-238)*G/A (реактивы ИХБиФМ СО РАН, г. Новосибирск) и IL1 β (+3953)*C/T (реактивы «Синтол», Москва, и «Fermentas», США). Для типирования генов HLA использовали метод ПЦР с сиквенс-специфическими праймерами (реактивы Protrans, Germany). Амплификацию проводили на приборе Veriti («Applied Biosystems», США). Для детекции использовали электрофоретический метод в 3%-ном агарозном геле и 8%-ном ПААГе.

Для выявления лиц с высоким и низким риском развития различных форм туберкулеза легких было выполнено моделирование вероятности наличия определенного фенотипа туберкулеза с помощью метода логистического регрессионного анализа с учетом генотипа обследуемого человека. Построение модели логистической регрессии выполнено с помощью программного пакета PAST V.3.18 [5].

Для оценки качества использовали следующие параметры: чувствительность и специфичность исследуемой модели. Также для моделей проводили ROC-анализ следующим образом: строили ROC-кривую и рассчитывали площадь под ROC-кривой. Для количественной оценки ROC-анализа использовали 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 [6].

Модель логистической регрессии имеет вид:

$$\boxed{\text{Логит}(P) = \text{Константа} + a_1X_1 + a_2X_2 + \dots + a_nX_n} \quad (1)$$

где P – вероятность отнесения пациента к группе риска развития определенного фенотипа туберкулеза,

$\text{Логит}(P) = \ln\left(\frac{P}{1-P}\right)$, a_i – коэффициенты регрессии, X_i – значения вошедших в модель показателей.

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

$$P = \left(\frac{1}{1 + e^{-\text{Логит}(P)}} \right) \quad (2)$$

Если $P < 0,5$ – пациент относится к группе 0 (отсутствие риска), если $P > 0,5$ – к группе 1 (высокий риск).

Результаты исследования и их обсуждение

В рамках предыдущих исследований [7, 8] нами были выявлены потенциальные маркеры: гены HLA, цитокинов и TLR и их комбинации, по которым исследуемые группы больных туберкулезом различались статистически значимо. В данной работе мы применяли полученные ранее данные для прогнозирования риска развития определенных фенотипов туберкулеза. Для этого мы воспользовались аппаратом множественной логистической регрессии.

Логистическая регрессия – это разновидность множественной регрессии, общее назначение которой состоит в анализе связи между несколькими независимыми переменными (называемыми также регрессорами, или предикторами) и зависимой переменной. С помощью логистической регрессии можно оценивать вероятность того, что при носительстве комбинаций определенных генов у больного разовьется определенный фенотип заболевания. В модель были включены все гены, которые по результатам предыдущих исследований имели статистическую значимость для данного фенотипа туберкулеза [7, 8].

На первом этапе нами были рассмотрены потенциальные маркеры, ассоциированные с развитием легкой очаговой формы (HLADQB1*06:02-8, DRB1*15, В*15, TNFA(-308)*G/G, TLR4(299)*G, TLR4(399)*G, TNFA(-308)*G), и принадлежность к женскому полу (данная форма чаще отмечается у женщин).

Результаты анализа для больных с очаговой формой представлены в таблице 1.

Таблица 1

Параметры и характеристика модели логистической регрессии для прогноза развития очагового туберкулеза (n=13)

Показатели	Коэффициент регрессии ± станд. ошибка	Отношение шансов (95%-ный доверит. интервал)	Оценка значимости P
HLADQB1*06:02-8	1,35± 0,62	3,84 (1,14–12,99)	0,0303
HLADRB1*15	1,26± 0,66	3,52 (0,97–12,78)	0,0550
HLAB*15	1,41 ±0,81	4,08 (0,84–19,77)	0,0810
TNFA(-308)*G/G	2,09± 1,08	8,10 (0,98–67,27)	0,0527
Пол (женский)	-1,15±0,64	0,32(0,09–1,12)	0,0740

TLR4(299)*G	1,38± 0,77	3,96(0,88–17,87)	0,0729
TLR4(399)*G	1,25±0,75	3,50 (0,80–15,23)	0,0951
TNFA(-308)*G	12,68±2235,18	4,89 (1,15–20,72)	0,9955
Константа	-2,2336		

Модель полностью: $\chi^2 = 17,74$; $df=8$; $P<0,05$
Чувствительность: 30,77%, специфичность: 98,63%
Диагностическая эффективность: 64,7%
Доля верно классифицированных: 88,37%
Площадь под ROC-кривой: 0,805±0,0762 (95% ДИ: 0,705–0,882).

Анализ полученной логит-модели свидетельствует о том, что значимое влияние на развитие очагового туберкулеза оказывают аллели HLA-DQB1*06:02-8. Диагностическая эффективность такой модели составила 64,7%. Данный ген находится в сильном неравновесном сцеплении с HLA-DRB1*15 и входит в состав расширенного анцестрального гаплотипа AH7.1 (C*07-B*07-DRB1*15:01-DQA1*01:02- DQB1*06:02), который с высокой частотой встречается у европейского населения.

На втором этапе исследования мы рассмотрели потенциальные маркеры, ассоциированные с развитием тяжелой фиброзно-кавернозной формы (HLA-B*08, DRB1*03, IL1 β (+3953)*T, IL1 β (+3953)*T/T), и принадлежность к мужскому полу (данная форма чаще имеет место у мужчин).

Результаты анализа для больных с фиброзно-кавернозной формой представлены в таблице 2.

Таблица 2

Параметры и характеристика модели логистической регрессии для прогноза развития фиброзно-кавернозного туберкулеза легких (n=11)

Показатели	Коэффициент регрессии ± станд. ошибка	Отношение шансов (95%-ный доверит. интервал)	Оценка значимости <i>P</i>
HLA-B*08	1,81±0,70	6,11 (1,54–24,20)	0,0099
HLA-DRB1*03	1,57±0,67	4,80 (1,29–17,88)	0,0194
IL1 β (+3953)*T	0,42±0,749	1,52 (0,35–6,60)	0,5772
IL1 β (+3953)*T/T	1,06±0,874	2,89 (0,52–16,03)	0,2240
Пол (мужской)	1,92±1,126	6,84 (0,75–62,11)	0,0876
Константа	-2,3979		

Модель полностью: $\chi^2 = 12,69$; $df=5$; $P < 0,05$

Чувствительность: 18,18%, специфичность: 98,67%

Диагностическая эффективность: 58,43%

Доля верно классифицированных: 88,37%

Площадь под ROC-кривой: $0,808 \pm 0,0818$ (95% ДИ: 0,709–0,885).

Полученная логит-модель свидетельствует о том, что значимое влияние на развитие фиброзно-кавернозного туберкулеза оказывают аллели HLA B*08 и HLA DRB1*03. Диагностическая эффективность такой модели составила 58,43%. Данные гены входят в состав высококонсервативного анцестрального гаплотипа HLA AN8.1 (A*01-B*08-TNFA(-308)*A-BfS-C4AQ0-C4B1-DRB1*03-DQA1*05:01-DQB1*02:01) – классического иммуногенетического маркера европеоидов, который встречается у четверти населения. Он известен своей «гиперреактивностью» за счет гиперпродукции TNF α . Вероятно, повышенный воспалительный ответ полезен на ранних стадиях заболевания, но в конечном итоге становится разрушительным, особенно если хроническая инфекция развивается у пожилых пациентов [9].

Для инфильтративной формы, как было показано в предыдущих исследованиях, были установлены в качестве кандидатных полиморфизм гена TNFA(-308)*A и его гетерозиготный генотип TNFA(-308)*G/A. Но при использовании метода логистической регрессии ни один из оцененных показателей не проявил себя в качестве значимого предиктора.

Данные результаты свидетельствуют о том, что прогностическим потенциалом для ранней диагностики развития определенной формы туберкулеза обладают только гены системы HLA. Это комплекс генов, расположенных на коротком плече 6-й хромосомы и обеспечивающих регуляцию иммунного ответа [9]. Для очаговой и фиброзно-кавернозной форм были обнаружены значимые предикторы, что дает нам возможность выявить минимальное количество потенциальных биомаркеров, которые можно использовать в качестве кандидатных полиморфизмов для данных фенотипов туберкулеза. Генетические характеристики пациента (гены системы HLA) позволяют определять вероятность развития тяжелой (фиброзно-кавернозной) или легкой (очаговой) формы туберкулеза легких. Пример расчета вероятности развития определенной формы туберкулеза (фиброзно-кавернозной) рассмотрен ниже (табл. 3).

Пример вычисления логита у пациента с туберкулезом, в генотипе которого есть аллели генов HLA B*08 и DRB1*03

Показатель	Вариант ответа	Код	Коэффициент регрессии a_i	Произведение (Код $\times a_i$)	Результат
HLAB*08	имеется	1	1,81	1,81 \times 1	1,81
HLADRB1*03	имеется	1	1,57	1,57 \times 1	1,57
Константа					-2,4849
Логит (P)				Сумма	0,8951

$$\text{Вероятность } P_A = \left(\frac{1}{1 - e^{-0,8951}} \right) = \left(\frac{1}{0,59143326} \right) = 1,691$$

Поскольку P больше 0,5, у данного пациента есть высокий риск развития фиброзно-кавернозной формы туберкулеза.

Заключение. Несмотря на многочисленные исследования, в настоящее время нет маркеров, которые однозначно могли бы помочь в диагностике туберкулеза. Прогнозирование развития различных форм заболевания с помощью метода множественной логистической регрессии позволяет провести предиктивную диагностику до формирования клинических признаков, что важно для повышения эффективности лечения. Использование генетических маркеров и математических методов исследования развития тяжелой или легкой формы заболевания даст возможность выделить среди больных группы высокого риска. Это необходимо для трансформации иммунновоспалительного ответа с гиперпродуктивного на толерантный для снижения деструкции пораженного органа. В настоящее время данный подход является оптимальным, поскольку применение комплекса диагностических мероприятий способствует выбору правильной тактики лечения, а также снижает риск прогрессирования инфекции и развития тяжелых форм туберкулеза.

Список литературы

1. Aravindan P.P. Host genetics and tuberculosis: Theory of genetic polymorphism and tuberculosis. Lung India. 2019. Vol. 36. no.3. P.244-252. DOI: 10.4103/lungindia.lungindia_146_15.

2. Никулина С.Ю., Чернова А.А., Шульман В.А., Верещагина Т.Д., Чернов В.Н. Модель логистической регрессии как дополнительный математический метод прогнозирования нарушений сердечной проводимости в семьях г. Красноярска // Российский кардиологический журнал. 2014. №10 (114). С. 46–52. DOI: 10.15829/1560-4071-2014-10-46-52.
3. Gideon H.P., Flynn J.L. Latent tuberculosis: what the host “sees”? Immunological Reviews. 2011. Vol. 50. Is. 2-3. P. 202-12. DOI: 10.1007/s12026-011-8229-7.
4. Cai L., Li Z., Guan X., Cai K., Wang L., Liu J., Tong Y. The Research Progress of Host Genes and Tuberculosis Susceptibility. Oxidative Medicine and Cellular Longevity. 2019. Vol. 2019. id 9273056. DOI:10.1155/2019/9273056.
5. Hammer O., Harper D.A.T., Ryan P.D. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 2001. Vol. 4. is. 1. art. 4. P. 9 https://palaeo-electronica.org/2001_1/past/issue1_01.htm (дата обращения 20.08.2020).
6. Григорьев С.Г., Лобзин Ю.В., Скрипченко Н.В. Роль и место логистической регрессии и ROC-анализа в решении медицинских диагностических задач // Журнал инфектологии. 2016. 8(4). С.36-45. DOI: 10.22625/2072-6732-2016-8-4-36-45.
7. Беляева С.В. Гены иммунного ответа и их комбинации в качестве предиктивных маркеров потенциального риска развития активного туберкулеза легких и его клинических фенотипов у представителей русской популяции Челябинской области: автореф. дис. ... канд. биол. наук. Челябинск, 2014. 23 с.
8. Бурмистрова А.Л., Беляева С.В. Визуализация генетических комбинаций, участвующих в формировании воспалительного ответа при туберкулезе легких, методом канонического анализа соответствий // Вестник Челябинского государственного университета. 2015. № 21 (376). Биология. Вып. 3. С. 12–16.
9. Candore G, Lio D., Colonna Romano G., Caruso C. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. Autoimmunity Reviews. 2002. Vol. 1. Is. 1–2. P. 29–35. DOI: 10.1016/S1568-9972(01)00004-0.