

*Д.Л. Колесник*

**МУТАЦИОННОЕ ДАВЛЕНИЕ В ГЕНЕ ЧЕЛОВЕЧЕСКОГО ПРОБЕЛКА  
ТРАНСФОРМИРУЮЩЕГО ФАКТОРА РОСТА АЛЬФА**

*Научный руководитель: ассист. А.А. Акуневиц*

*Кафедра общей химии*

*Белорусский государственный медицинский университет, г. Минск*

*D.L. Kolesnik*

**MUTATIONAL PRESSURE IN THE GENE OF HUMAN  
PROTRANSFORMING GROWTH FACTOR ALPHA**

*Tutor: assistant A.A. Akunevich*

*Department of General Chemistry*

*Belarusian State Medical University, Minsk*

**Резюме.** В статье показано, что высокие значения GC-насыщенности кодирующей последовательности пробелка человеческого трансформирующего фактора роста альфа могут быть следствием действовавшего ранее сильного мутационного GC-давления. Анализ частот возникновения герминативных и соматических мутаций, а также нуклеотидных замен в интронах показал, что ген пробелка TGF- $\alpha$  находится под влиянием мутационного AT-давления.

**Ключевые слова:** трансформирующий фактор роста альфа; мутационное давление; герминативные мутации; соматические мутации.

**Resume.** The article shows that high values of GC-content of human protransforming growth factor alpha coding sequence may be a consequence of the previous strong mutational GC-pressure. An analysis of the germline and somatic mutations rates, as well as nucleotide substitutions in introns, showed that human protransforming growth factor alpha gene is under the influence of mutational AT-pressure.

**Keywords:** transforming growth factor alpha; mutational pressure; germline mutations; somatic mutations.

**Актуальность.** Трансформирующий фактор роста альфа (TGF- $\alpha$ , от англ. transforming growth factor alpha) представляет собой полипептид, состоящий из 50 аминокислотных остатков. Активный TGF- $\alpha$  отщепляется от более крупного интегрального мембранного гликопротеина, связывается с рецептором эпидермального фактора роста (EGFR, от англ. epidermal growth factor receptor) и стимулирует пролиферацию и дифференцировку клеток. Соматические и герминативные мутации TGF- $\alpha$  являются причиной нарушений данных процессов [1]. Повышенная экспрессия TGF- $\alpha$ , а также возникающие в его структуре мутации ассоциированы со многими опухолевыми заболеваниями (например, гепатоцеллюлярной карциномой, раком желудка). Повышенный уровень TGF- $\alpha$  также ассоциирован с предопухолевым состоянием желудка — хроническим гипертрофическим полиаденоматозным гастритом [2]. Анализ нуклеотидного состава кодирующей последовательности пробелка TGF- $\alpha$  и возникающих в ней герминативных и соматических мутаций позволит прояснить генетические механизмы заболеваний, ассоциированных с его экспрессией.

**Цель:** проанализировать нуклеотидный состав кодирующей последовательности человеческого пробелка трансформирующего фактора роста альфа и данные о возникающих в ней герминативных и соматических мутациях.

### **Задачи:**

1. Описать характер использования нуклеотидов в кодирующей последовательности пробелка человеческого трансформирующего фактора роста альфа.
2. Определить тип мутационного давления в кодирующей последовательности пробелка человеческого трансформирующего фактора роста альфа.
3. Систематизировать и провести статистическую обработку данных о герминативных и соматических мутациях, возникающих в кодирующей последовательности и интронах пробелка человеческого трансформирующего фактора роста альфа.

**Материал и методы.** Кодирующие последовательности CDS (от англ. coding DNA sequence) пробелка TGF- $\alpha$  человека и приматов были получены из базы данных Ensembl. Для анализа использовался 201-й транскрипт гена человеческого пробелка TGF- $\alpha$  (ID ENST00000295400.11) [3]. Распределение значений GC-состава и частот использования нуклеотидов в различных сайтах CDS пробелка TGF- $\alpha$  было рассчитано с помощью алгоритма VVK Protective buffer (<http://chemres.bsmu.by>). Для дальнейшего анализа использовались рассчитанные значения общей GC-насыщенности (G+C), GC-насыщенности первых, вторых и третьих положений кодонов (1GC, 2GC и 3GC, соответственно), GC-насыщенности четырёхкратно и двукратно вырожденных сайтов третьих положений кодонов (GC4f и GC2f3p, соответственно).

Для анализа герминативных мутаций в CDS и интронах пробелка TGF- $\alpha$  использовались данные о вариабельности генов, размещённые в базе данных Ensembl: по состоянию на март 2022 года общее число синонимичных мутаций в экзонах пробелка TGF- $\alpha$  составило 59, несинонимичных — 128, нонсенс — 5, общее число мутаций в интронах — 23 352 [3]. Для анализа соматических мутаций в CDS пробелка TGF- $\alpha$  опухолевых клеток использовались данные, размещённые в базе данных COSMIC: по состоянию на май 2022 года общее число синонимичных мутаций в экзонах пробелка TGF- $\alpha$  составило 19, несинонимичных — 42, нонсенс — 2 [4]. Проведён статистический анализ частот возникновения нуклеотидных замен, наблюдаемых при герминативном и соматическом мутагенезе в CDS пробелка TGF- $\alpha$ . Доля каждого из типов нуклеотидных замен рассчитывалась как отношение между количеством сайтов с мутациями, фактически наблюдаемыми в клетках (согласно базам данных Ensembl и COSMIC), и общим количеством сайтов CDS пробелка TGF- $\alpha$ , в которых конкретный тип мутации может произойти. Доли различного типа мутаций сравнивались между собой с помощью *t*-теста для относительных величин.

**Результаты и их обсуждение.** Общая GC-насыщенность CDS пробелка TGF- $\alpha$  составляет 60,0%, а уровень 3GC-насыщенности — 71,4%, при этом для последовательности активного TGF- $\alpha$  характерны минимальные значения GC-насыщенности (таблица 1). Мутационное GC-давление приводит к увеличению 3GC-насыщенности, поскольку многие мутации с АТ на GC в этих положениях кодонов являются синонимичными. GC-насыщенность первых и вторых положений кодонов растёт медленнее, поскольку все мутации во вторых положениях кодонов и большинство мутаций в первых положениях кодонов несинонимичны [5].

**Табл. 1.** Показатели GC-насыщенности CDS пробелка человеческого TGF- $\alpha$  и её частей

Часть CDS TGF- $\alpha$	G+C	1GC	2GC	3GC	GC4f	GC2f3p
Пробелок TGF- $\alpha$	0,600	0,609	0,478	0,714	0,682	0,742
Пре-TGF- $\alpha$	0,658	0,692	0,564	0,718	0,654	0,889
Активный TGF- $\alpha$	0,553	0,600	0,440	0,620	0,591	0,630
Пост-TGF- $\alpha$	0,602	0,569	0,458	0,778	0,757	0,808

Высокие значения GC-насыщенности в CDS пробелка человеческого TGF- $\alpha$  могут быть следствием длительного мутационного GC-давления, действовавшего в этом гене у приматов (таблица 2). Даже если мутационное давление изменило своё направление, паттерн распределения GC-насыщенности ( $2GC < 1GC < 3GC$ ) сохранится во многих поколениях.

**Табл. 2.** Показатели GC-насыщенности CDS пробелка TGF- $\alpha$  приматов

Вид	G+C	1GC	2GC	3GC	GC4f	GC2f3p
<i>Cebus capucinus</i>	0,582	0,578	0,464	0,705	0,711	0,652
<i>Cercocebus atys</i>	0,591	0,581	0,473	0,719	0,709	0,708
<i>Chlorocebus pygerythrus</i>	0,599	0,608	0,475	0,715	0,671	0,746
<i>Colobus angolensis</i>	0,587	0,581	0,467	0,713	0,706	0,697
<i>Hylobates lar</i>	0,595	0,575	0,473	0,737	0,723	0,735
<i>Lemur pusillus</i>	0,601	0,615	0,467	0,720	0,670	0,771
<i>Lemur tarsier</i>	0,586	0,600	0,479	0,679	0,605	0,734
<i>Macaca nemestrina</i>	0,587	0,581	0,473	0,707	0,709	0,677
<i>Macaca sylvanus</i>	0,585	0,581	0,473	0,701	0,686	0,692
<i>Mandrillus leucophaeus</i>	0,589	0,581	0,473	0,713	0,698	0,708
<i>Pan paniscus</i>	0,583	0,587	0,473	0,689	0,679	0,687
<i>Pan troglodytes</i>	0,579	0,603	0,464	0,670	0,602	0,714
<i>Papio anubis</i>	0,564	0,592	0,463	0,637	0,626	0,632
<i>Pongo abelii</i>	0,599	0,587	0,473	0,737	0,729	0,727
<i>Prolemur simus</i>	0,605	0,609	0,466	0,739	0,721	0,758
<i>Propithecus coquereli</i>	0,600	0,606	0,463	0,731	0,721	0,742
<i>Rhinopithecus bieti</i>	0,587	0,581	0,461	0,719	0,714	0,701
<i>Rhinopithecus roxellana</i>	0,585	0,581	0,461	0,713	0,702	0,701
<i>Saimiri boliviensis</i>	0,607	0,609	0,478	0,733	0,706	0,742
<i>Troglodytes gorilla</i>	0,595	0,587	0,473	0,725	0,702	0,731

Чтобы выявить текущее направление мутационного давления, был проведён анализ данных о герминативных и соматических мутациях пробелка человеческого TGF- $\alpha$ . При анализе частот возникновения герминативных мутаций в CDS пробелка TGF- $\alpha$  были получены следующие результаты: несинонимичные замены C>T (0,441), G>A (0,272), A>T (0,129) происходят достоверно чаще ( $P < 0,05$ ), чем несинонимичные замены в обратном направлении T>C (0,077), A>G (0,104), T>A (0,023). Несинонимичные замены C>T, A>T и T>G происходят достоверно чаще ( $P < 0,05$ ) синонимичных, а синонимичные замены G>A и A>G происходят достоверно чаще

( $P < 0,05$ ) несинонимичных. При анализе частот возникновения соматических мутаций были получены схожие результаты: несинонимичные замены  $C > T$  (0,176),  $G > A$  (0,152) происходят достоверно чаще ( $P < 0,05$ ), чем несинонимичные замены в обратном направлении  $T > C$  (0,013),  $A > G$  (0,045). Несинонимичные замены  $G > T$  происходят достоверно чаще ( $P < 0,05$ ) синонимичных. При анализе частот возникновения нуклеотидных замен в интронах пробелка TGF- $\alpha$  были получены следующие результаты: замены  $C > T$  (0,364),  $G > A$  (0,318),  $C > A$  (0,115),  $G > T$  (0,107),  $C > G$  (0,104),  $A > T$  (0,060), происходят достоверно чаще ( $P < 0,05$ ), чем замены в обратном направлении  $T > C$  (0,0503),  $A > G$  (0,210),  $A > C$  (0,069),  $T > G$  (0,015),  $G > C$  (0,093),  $T > A$  (0,012). При этом общая GC-насыщенность всех интронов гена пробелка TGF- $\alpha$  составляет 32,9%. Исходя из результатов анализа фактически возникающих герминативных и соматических мутаций, ген пробелка TGF- $\alpha$  находится под влиянием мутационного АТ-давления.

Стоит отметить, что при анализе герминативных и соматических мутаций не было обнаружено достоверных отличий в направлениях синонимичных замен, а ряд несинонимичных замен происходит достоверно чаще синонимичных (например,  $C > T$ ,  $A > T$  и  $G > T$ ). Подобный характер возникновения нуклеотидных замен может свидетельствовать о направленном положительном отборе, способствующем выживанию клеток с мутированным пробелком TGF- $\alpha$  в опухолевых клетках, тогда как полиморфизм генов пробелка TGF- $\alpha$  может определять развитие хейлосклизиса (расщелины губы и нёба) и тяжесть данного заболевания [5, 6].

#### **Выводы:**

1. Для кодирующей последовательности пробелка TGF- $\alpha$  характерны высокие значения GC-насыщенности, что может быть следствием действовавшего ранее мутационного GC-давления.

2. Анализ частот возникновения герминативных и соматических мутаций, а также нуклеотидных замен в интронах показал, что ген пробелка TGF- $\alpha$  находится под влиянием мутационного АТ-давления.

3. В кодирующей последовательности пробелка TGF- $\alpha$  несинонимичные замены в АТ-направлении ( $C > T$ ,  $A > T$  и  $G > T$ ) происходят достоверно чаще, чем синонимичные замены в данном направлении.

#### **Литература**

1. Transforming growth factor alpha (TGF $\alpha$ ) regulates granulosa cell tumor (GCT) cell proliferation and migration through activation of multiple pathways / C. Wang, X. Lv, C. Jiang et al. // PLoS One. – 2012. – № 11. – P. 1-12.

2. TGFA [Electronic resource] / The Human Protein Atlas. – Stockholm, 2022. – Mode of access: <https://www.proteinatlas.org/ENSG00000163235-TGFA/pathology> (date of access: 15.05.2022).

3. TGFA-201 Transcript: ENST00000295400.11 [Electronic resource] / EMBL-EBI. – Hinxton, 2022. – Mode of access: <http://www.ensembl.org/index.html> (date of access: 15.03.2022).

4. Gene TGFA\_ENST00000295400 [Электронный ресурс] / COSMIC. – Hinxton, 2022. – Режим доступа: <https://cancer.sanger.ac.uk/cosmic> (date of access: 12.05.2022).

5. Хрусталёв, В. В. Биохимические механизмы мутационного давления в методологии вычислительной биологии / В. В. Хрусталёв; под ред. Е. В. Барковского. – Минск: БГМУ, 2010. – 213 с.

6. Association between polymorphism of TGFA Taq I and cleft Lip and/or palate: a meta-analysis / C. Feng, E. Zhang, W. Duan et al. // BMC Oral Health. – 2014. – № 14. – P. 1-11.