

УДК 616-092

## Метод выявления клеточных состояний и определение их количества при эпителиально-мезенхимальных переходах

Ситдикова К.К., Устинова Е.Е., Черкашенко В.Н., Сабурина И.Н.

Федеральное государственное бюджетное научное учреждение «Научно-исследовательский институт общей патологии и патофизиологии»  
125315, Москва, ул. Балтийская, д. 8

В сложных биологических динамических системах состояния системы часто ассоциируются со стабильными режимами поведения (например, аттракторы) или характерными фазами процесса. Биологические системы являются объектами с высокой размерностью, и их состояния характеризуются множеством показателей. Этот факт делает задачу обнаружения таких состояний сложно выполнимой, а рассмотрение низкоразмерных портретов системы часто не позволяет обнаруживать состояния систем. Статьи, посвященные обнаружению и визуализации таких состояний, обычно оперируют дву- или трёхмерными изображениями. Обнаружение таких состояний по данным многомерных временных рядов является ключевой задачей математической биологии и биофизики.

Мы предложили использовать для обнаружения и визуализации состояний искусственную нейронную сеть SOM. В работе анализировался процесс эпителиально-мезенхимального перехода, определяющий клеточную пластичность, играющий огромную роль в клеточной и тканевой биологии и определяющий такие важные патологии, как рак. Нами было показано, что в динамике ЭМП содержится 6 состояний в 16-мерном пространстве. Предложенный метод позволяет выявлять и интерпретировать состояния многомерных биологических систем, что может быть использовано в поиске методов управления этими состояниями.

**Ключевые слова:** эпителиально-мезенхимальный переход; ландшафт Уоддингтона; аттрактор; генная регуляторная сеть; кластер; искусственная нейронная сеть; клеточная пластичность.

**Для цитирования:** Ситдикова К.К., Устинова Е.Е., Черкашенко В.Н., Сабурина И.Н. Метод выявления клеточных состояний и определение их количества при эпителиально-мезенхимальных переходах. *Патогенез*. 2025; 23(4): 79–90.

**DOI:** 10.48612/path/2310-0435.2025.04.79-90

**Для корреспонденции:** Сабурина Ирина Николаевна, e-mail: saburina@mail.ru

**Финансирование:** Обзор написан при поддержке Министерства здравоохранения РФ (Государственное задание № FGUFU-2025-0004).

**Конфликт интересов:** Авторы заявляют об отсутствии конфликта интересов

**Поступила:** 14.11.2025.

## Method for identifying cellular states and determining their quantity during epithelial-mesenchymal transitions

Sitdikova K.K., Ustinova E.E., Cherkashenko V.N., Saburina I.N.

Institute of General Pathology and Pathophysiology  
Baltijskaya St. 8, Moscow 125315, Russian Federation

In complex biological dynamic systems, system states are often associated with stable behavioral modes (e.g., attractors) or characteristic process phases. Biological systems are high-dimensional objects; their states are characterized by a multitude of parameters. This fact makes the task of detecting such states challenging, and considering low-dimensional system portraits often fails to detect system states. Articles devoted to the detection and visualization of such states typically employ two- or three-dimensional images. Detecting such states from multidimensional time series data is a key task in mathematical biology and biophysics. We proposed using the SOM artificial neural network for state detection and visualization. In this work, we analyzed the process of epithelial-mesenchymal transition, which determines cellular plasticity and plays a crucial role in cell and tissue biology, as well as in such important pathologies as cancer. We demonstrated that the EMT dynamics contain six states in a 16-dimensional space. The proposed method allows one to identify and interpret the states of multidimensional biological systems, which can be used in the search for methods of controlling these states.

**Key words:** epithelial-mesenchymal transition; Waddington landscape; attractor; gene regulatory network; cluster; artificial neural network; cellular plasticity.

**For citation:** Sitdikova K.K., Ustinova E.E., Cherkashenko V.N., Saburina I.N. [Method for identifying cellular states and determining their quantity during epithelial-mesenchymal transitions]. *Patogenez [Pathogenesis]*. 2025; 23(4): 79–90 (in Russian).

**DOI:** 10.48612/path/2310-0435.2025.04.79-90

**For correspondence:** Saburina Irina Nikolaevna, e-mail: saburina@mail.ru

**Funding:** The support of a grant from the Ministry of Health of the Russian Federation (State Task #FGUFU-2025-0004)

**Conflict of interest:** The authors declare no conflict of interest.

**Received:** 14.11.2025.

## Основные сокращения

ЭМП — эпителиально-мезенхимальный переход, РС — регуляторная сеть, ГРС (GRN) — генная регуляторная сеть, ИНС (ANN) — искусственная нейронная сеть, СОКК (SOM) — самоорганизующиеся карты Кохоннена, ВР — временные ряды, ОДУ — обыкновенные дифференциальные уравнения.

## Введение

### Общебиологическая значимость процесса ЭМП

Эпителиально-мезенхимальный переход (ЭМП) — важный биологический процесс, который играет ключевую роль как в нормальном развитии, так и в патологии. Этот процесс характеризуется переходом клеток из эпителиального состояния, где они плотно связаны друг с другом, в мезенхимальное состояние, которое позволяет клеткам стать более подвижными и инвазивными [1]. В норме ЭМП обеспечивает стандартное развитие и формирование тканей и органов, однако нарушение регуляции процессов ЭМП может привести к наиболее важной проблеме современной медицины — развитию онкологических заболеваний, и в отдельных случаях эти нарушения усложняют раковую терапию за счёт своего участия в процессах метастазирования и развития лекарственной устойчивости у раковых клеток [1]. В ходе ЭМП клетки не сразу переходят из эпителиального в мезенхимальный тип, а проходят несколько промежуточных состояний. Количество и определение этих промежуточных состояний всё ещё является актуальным вопросом, и характеристика этих промежуточных состояний, вероятно, может являться ключевой как для прогнозирования перехода клетки в опухолевое состояние, так и для онкотерапии в целом [2].

ЭМП влияет на множество клеточных процессов, регулируя миграцию и инвазию клеток в окружающие ткани, клеточную пластичность, изменяя клеточную морфологию, и влияя на экспрессию множества клеточных маркеров, ассоциированных с ЭМП [1]. Это не только важно для нормального клеточного самообновления и развития тканей, но также обеспечивает клеточный ответ на стресс и повреждение. Так, при нарушении целостности ткани переход клеток в мезенхимальное состояние обеспечивает миграцию клеток к месту повреждения, их дифференцировку в нужный тип и адаптацию к новым условиям за счёт пластичности мезенхимальных клеток [1]. Таким образом ЭМП обеспечивает поддержание тканевого гомеостаза, причём этот переход является обратимым, и по окончании процесса восстановления клетки могут подвергнуться мезенхимально-эпителиальному переходу.

ЭМП также играет ключевую роль в процессах старения. Так, повышенная активность ЭМП, вызванная воспалением и окислительным стрессом, приводит к повышенному синтезу внеклеточного матрикса и накоплению фибробластов, что в результате приводит к развитию фи-

броза в стареющих тканях [1]. Это также истощает регенеративный потенциал тканей, что, соответственно, приводит к невозможности поддерживать тканевой гомеостаз и старению.

ЭМП характеризуется изменениями в экспрессии различных клеточных маркеров, что отражает переход клеток от эпителиального к мезенхимальному фенотипу. Основные маркеры, изменяющиеся при ЭМП, включают Е-кадгерин, N-кадгерин, Виментин, транскрипционные факторы (Snail, Twist и Slug),  $\beta$ -катенин, Фибронектин [1]. Эти вещества служат важными маркерами процесса ЭМП и могут быть использованы для оценки агрессивности опухолей и их метастатического потенциала. Помимо указанных маркеров, в регуляции ЭМП также важную роль играют микроРНК [3]. Эти небольшие некодирующие РНК регулируют экспрессию генов на посттранскрипционном уровне, что позволяет им влиять на различные процессы в ЭМП, в том числе на регуляцию транскрипционных факторов (miR-34a, miR-200), поддержание баланса между эпителиальными и мезенхимальными клетками (miR-200), и процессы передачи сигнала (miR-612) [3]. Изменения экспрессии этих белков и микроРНК могут иметь большое прогностическое значение в онкодиагностике, а также для поиска новых методов терапии рака.

Как было уже отмечено выше, ЭМП играет ключевую роль в опухолевой прогрессии и метастазировании рака. Если процесс запуска ЭМП нарушен и клетки начинают самообновляться чаще, то это также повышает вероятность дифференцировки в опухолевую клетку. Более того, снижение клеточной адгезии в процессе ЭМП увеличивает подвижность опухолевых клеток и упрощает их распространение по организму, ремоделирование внеклеточного матрикса в ходе ЭМП ещё больше способствует миграции раковых клеток [1, 4]. В некоторых случаях при онкологии наблюдается развитие внутриопухолевой гетерогенности — клеточного разнообразия между клетками одной опухоли. Наличие подобной гетерогенности является едва ли не главным фактором неэффективности терапии, низкой выживаемости и высокой частоты рецидивов [4]. Связано это, в первую очередь, с тем, что увеличение клеточного разнообразия в опухоли повышает вероятность появления подтипов клеток, устойчивых к различным лекарственным препаратам, а также повышает количество мишеней, что усложняет эффективность таргетной терапии. Также важным является изменение в микроокружении опухоли, что может привести к формированию дополнительных барьеров и усложнить проникновение препаратов к целевым клеткам [4].

Хотя в данном обзоре были кратко отмечены лишь некоторые, наиболее важные участники ЭМП, в этот процесс на самом деле вовлечены тысячи молекул, и полноценная визуализация всех состояний, путей и участников этого процесса невозможна ввиду высокой размерности этих данных. Но возможно рассмотреть проекции и упрощения этих данных для получения более понятной картины. Для подобной визуализации требуется использование

целого ряда математических методов [5]. Так, благодаря методам снижения размерности (РСА, UMAP и т.д.) становится возможным рассмотреть проекции экспрессии тысяч генов, вовлеченных в ЭМП, в более удобном 3D или 2D пространстве. Использование методов кластеризации (SOM, DTW и т.д.) позволяет выделить основные пути и переходы в ЭМП, а построение топографических карт и энергетических ландшафтов позволяет визуализировать различные состояния при ЭМП и вероятности переходов из одного состояния в другое [5].

### Ландшафт Уоддингтона

Концепция ландшафта Уоддингтона была предложена британским биологом Конрадом Уоддингтоном в 1950-х годах как метафора для описания процессов клеточной дифференциации и развития. Эта модель визуализирует клеточные состояния как «долины» и «холмы» в трёхмерном пространстве, где клетки могут перемещаться между различными состояниями, представляя собой динамическую систему, в которой стабильные состояния (аттракторы) соответствуют различным типам клеток. Автор использовал ландшафт как метафору для описания того, как клетки могут «скатываться» в более стабильные состояния, что отражает их дифференциацию в специализированные типы клеток.

Концепция ландшафта Уоддингтона основана на ряде ключевых идей, которые оказали значительное влияние на понимание процессов развития и дифференциации клеток [6]. Вот основные из них:

- **Эпигенетический ландшафт.** Введена метафора эпигенетического ландшафта, где клеточные состояния представлены как «долины» и «холмы». Эта модель иллюстрирует, как клетки могут переходить между различными состояниями, где «долины» представляют стабильные состояния, а «холмы» — барьеры, которые необходимо преодолеть для изменения состояния [7, 8].
- **Канализирование.** Описан процесс канализирования, при котором клеточные судьбы становятся более предсказуемыми и устойчивыми к изменениям окружающей среды. Это означает, что клетки, находясь в определенных условиях, имеют тенденцию развиваться в определенные типы, что упрощает понимание их дифференциации [9, 10].
- **Генетическая ассимиляция.** Проведены эксперименты, показывающие, что фенотипы, вызванные воздействием окружающей среды, могут стать наследственными через искусственный отбор. Это явление, известное как генетическая ассимиляция, подчеркивает, как развитие может влиять на эволюцию [11].
- **Сложные динамические системы.** Уоддингтон рассматривал развитие как стохастическую динамическую систему, где альтернативные траектории могут контролироваться как внешними сигналами, так и внутренними переменными. Это

понимание помогает объяснить, как различные факторы могут влиять на клеточные судьбы и их изменения [12].

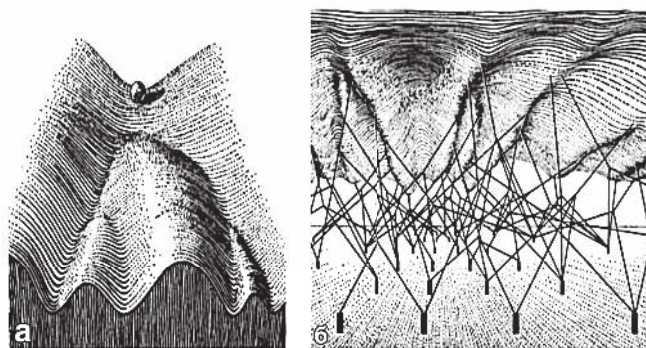
### • Интеграция различных биологических явлений.

Автор стремился создать единую теоретическую систему, которая объединяет все биологические явления, независимо от их сложности. Его подход к использованию образов и метафор помогал объяснять сложные концепции и привлекал внимание к важным вопросам в биологии [13].

Эти идеи Уоддингтона продолжают оказывать влияние на современные исследования в области клеточной биологии, генетики и эволюционной теории, открывая новые горизонты для понимания сложных процессов, происходящих в живых организмах. В частности, концепция ландшафта оказывает глубокое влияние и на методы терапии серьёзных заболеваний человека [14].

Ландшафт Уоддингтона сам по себе является трехмерной моделью, которая иллюстрирует различные клеточные состояния и их переходы на уровне метафоры (рис. 1). На рис. 1-а представлена поверхность, иллюстрирующая ландшафт. По оси, направленной на зрителя из рисунка располагается ось времени, ортогонально к ней — значение того или иного маркера в генной регуляторной сети (ГРС), а по вертикальной оси, направленной из плоскости «время-маркер», располагается такой показатель, как «энергия», требуемая для перехода из состояния (одного «оврага») в состояние (в другой «овраг»). Овраги ландшафта представляют собой состояния (аттракторы), а гребни между ними отражают «потенциальные барьеры», через которые должны перебираться клетки для перехода из состояния в состояние. Таким образом, классическое изображение ландшафта Уоддингтона, по сути, является одномерной иллюстрацией процессов.

Между генотипом и фенотипом существует взаимосвязь, которая реализуется за счет ГРС, топология которых и формирует ландшафт Уоддингтона. Иллюстрация этого утверждения представлена на рис. 1-б, где под поверхностью ландшафта располагаются «гены», влияющие на профиль ландшафта через связи, отображаемые линиями. Их сокращение или удлинение изменяют профиль ландшафта.



**Рис. 1.** Метафора ландшафта Уоддингтона: а — поверхность, иллюстрирующая ландшафт Уоддингтона; б — роль ГРС в изменении профиля ландшафта Уоддингтона [15].



## Многомерность процессов и порождаемые проблемы

Визуализация ландшафта Уоддингтона представляет собой сложную задачу, особенно из-за многомерности данных, связанных с клеточной дифференциацией и эпигенетическими изменениями. Однако реальные биологические данные часто имеют гораздо более высокую размерность. Для регуляторных сетей (РС), описывающих разные биологические процессы, мерность признакового пространства может достигать от нескольких десятков до сотен маркёров, учёт которых необходим для описания анализируемых процессов. Таким образом, основная проблема при попытке практического применения данной модели к реальным данным — высокая размерность РС. Именно многомерность регуляторных сетей не позволяет использовать метафору ландшафта Уоддингтона для анализа практических данных [16].

Увеличение размерности признакового пространства может изменять интерпретацию ландшафта Уоддингтона, уменьшая или увеличивая количество выявленных кластеров (состояний) в данных. Сказанное иллюстрируется **рис. 2** (по [17, 18]). На **рис. 2-а** приведены два представления одного и того же ландшафта; левое представление — в двумерном пространстве, а правое — в трёхмерном пространстве. На первом визуально можно четко выделить три состояния, обозначенные цифрами 1–3, а при переходе к трёхмерному представлению оказывается, что состояний 2.

Типичная иллюстрация ландшафта из научных публикаций всегда использует две переменные (маркёра), что подтверждается **рис. 2-а**.

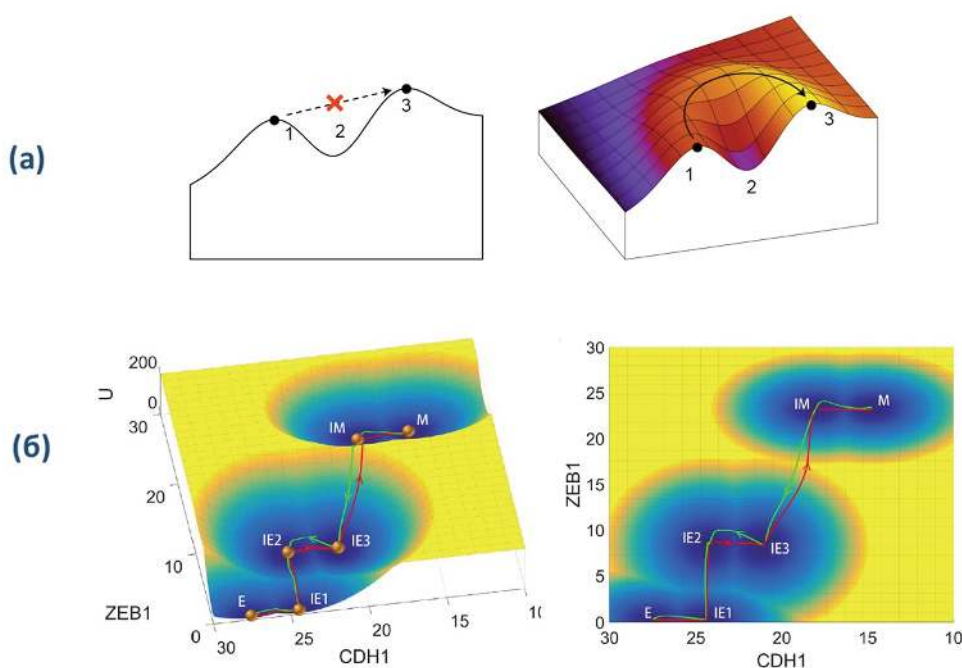
Таким образом, высокая размерность признакового пространства непосредственно влияет на способность выявлять состояния клеток:

- Гетерогенность клеток. Высокая размерность позволяет более точно захватывать гетерогенность клеточных популяций, что важно для понимания различных клеточных состояний и их динамики. Например, в исследованиях, связанных с иммунным ответом, использование методов выбора признаков на основе распределений генов может улучшить идентификацию динамики на уровне популяции.
- Идентификация клеточных типов. Снижение размерности помогает выделить специфические маркёры, связанные с определенными клеточными типами, что критически важно для точной аннотации клеток и понимания их функций в различных биологических контекстах.

Таким образом, мы можем утверждать, что задача визуализации ландшафта Уоддингтона является крайне важной в силу широкой применимости самой концепции ландшафта в биологии и медицине.

### Самоорганизующаяся карта (SOM) Кохонена

Очень мощным инструментом для анализа данных, является искусственная нейронная сеть, называемая самоорганизующейся картой (SOM) Кохонена, предложенная финским исследователем Тейво Кохоненом в 1982 году. Этот метод используется в качестве инструмента кластеризации и визуализации в исследовательском анализе данных. Основная цель SOM — преобразовать сложное многомерное входное пространство в более простое маломерное (обычно двумерное) дискретное выходное пространство, сохра-



**Рис. 2.** Проблема применимости ландшафта Уоддингтона: а — влияние увеличения размерности ландшафта на оценку количества состояний [17]; б — иллюстрация ландшафта Уоддингтона как низкоразмерного изображения поверхности [18].

ная взаимосвязи (т. е. топологию) в данных [19-21]. Процедура SOM имеет ряд преимуществ по сравнению с методами, решающими аналогичные задачи. Во-первых, при её применении нет необходимости в предположениях относительно распределений переменных и не требуется независимость переменных. Во-вторых, она легко реализуется и способна решать нелинейные задачи очень высокой сложности.

Сеть Кохонена (рис. 3, А, левая часть) — это однослойная сеть, каждый нейрон которой соединен со всеми компонентами  $n$ -мерного входного вектора. Входной вектор — это описание одного объекта из множества объектов, подлежащих кластеризации. В качестве нейронов сети Кохонена применяются линейные взвешенные сумматоры. Каждый  $j$ -й нейрон описывается вектором весов  $w_j$ :

$$w_j = (w_{1j}, w_{2j}, w_{3j}, \dots, w_{mj})$$

где  $m$  — число компонентов входных векторов, а входной вектор имеет вид

$$x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{im})$$

В отличие от большинства других нейронных сетей, в SOM слой входных узлов напрямую подключен к выходному слою вычислительных узлов без какого-либо скрытого слоя. Узлы (искусственные нейроны сети) во входном слое обозначают атрибуты (признаки), переменные, содержащиеся во входных данных. Каждый элемент входных данных представлен  $m$ -мерным входным вектором

$$x = (x_1, x_2, x_3, \dots, x_m),$$

элементы которого указывают значения атрибутов конкретного набора данных. Выходной слой, также известный как «слой Кохонена» или «слой SOM», представляет собой низкоразмерную визуализацию данных. Как уже говорилось, узлы в выходном слое организованы в виде двумерной сетки. Обычно топология сети организована в виде прямоугольной или гексагональной сетки в традиционной SOM (рис. 3, А, правая часть). Очевидно, что эти топологии имеют разные свойства. В прямоугольной топологии каждый внутренний узел имеет четырёх соседей, тогда как в гексагональной топологии — шесть. Количество нейронов в слое Кохонена определяет разрешающую способность данной модели и задается эвристикой при помощи формулы

$$K = 5\sqrt{N}.$$

В сетях Кохонена используется обучение без учителя. Для обучения сети применяются механизмы конкуренции. При подаче на вход сети вектора  $x$  побеждает тот нейрон, вектор весов которого в наименьшей степени отличаются от входного вектора. Для нейрона-победителя выполняется соотношение

$$d(x, w_j) = \min_{1 \leq i \leq n} d(x, w_i)$$

где  $n$  — количество нейронов,  $j$  — номер нейрона-победителя,  $d(x, w)$  — расстояние (в смысле выбранной метрики) между векторами  $x$  и  $w$ . Чаще всего в качестве меры

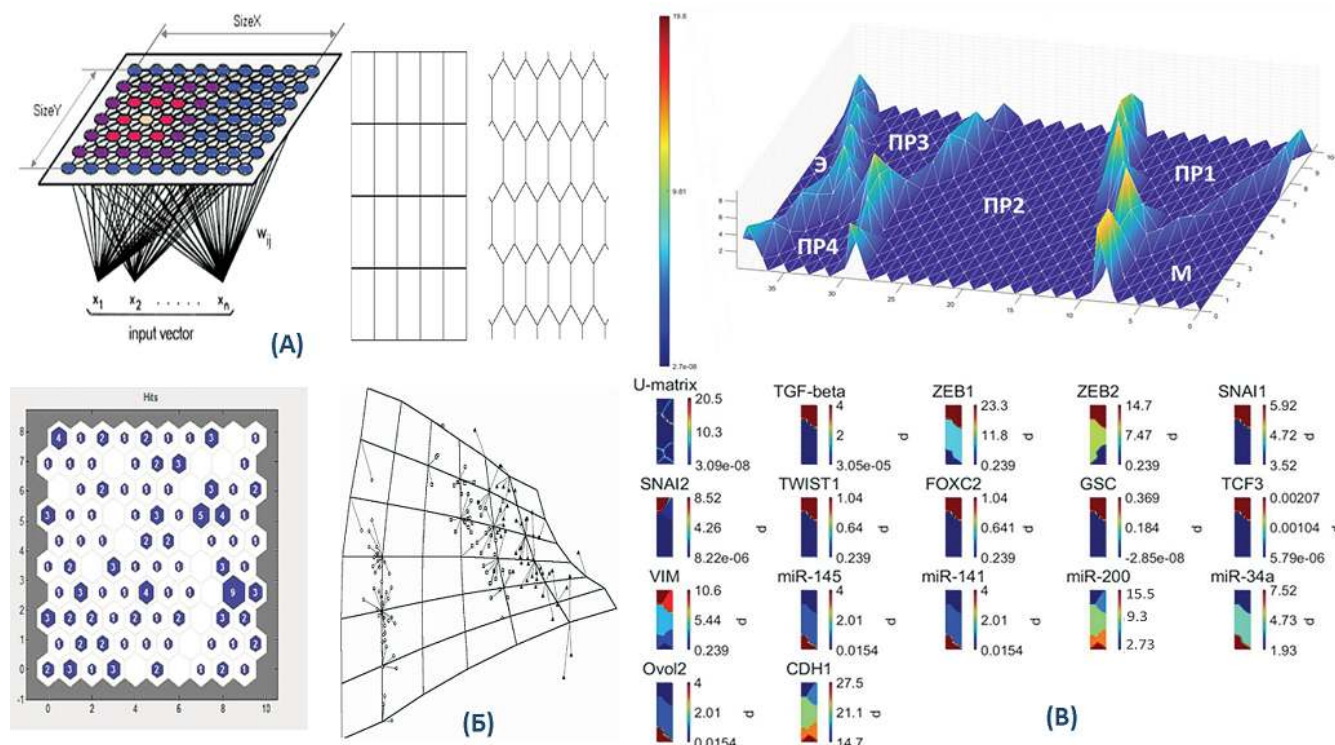


Рис. 3. Метод самоорганизующихся карт Кохонена: А — структура и топология сети Кохонена [22], Б — плотность распределения данных по нейронам сети деформирует поверхность, отображающую данные в пространстве пониженной размерности [21], В — представление результатов в виде двух картинок; верхняя — матрица похожести, выявляющая кластерную структуру, и нижняя — «раскраска» полученной кластерной структуры значениями анализируемых маркеров. Структуры были получены с использованием библиотеки Matlab [23].

расстояния используется евклидова метрика. В процессе обучения таких сетей на входы последовательно подаются данные, в процессе последовательной подачи на вход сети обучающих примеров определяется наиболее схожий нейрон. Этот нейрон объявляется победителем и является центром при подстройке весов у соседних нейронов.

Вокруг нейрона-победителя образуется окружение (neighborhood), или радиус обучения (radius of learning). Радиус обучения определяет, сколько нейронов кроме нейрона-победителя участвуют в обучении (т.е. корректируют свои веса) на данной итерации. Под радиусом в данном случае подразумевается расстояние в пространстве векторов весов нейронов. Т.е., любой нейрон, расстояние от вектора весов которого до вектора весов нейрона-победителя менее радиуса обучения, участвует в коррекции весов на данной итерации. Радиус обучения максимален на первой итерации и уменьшается с увеличением числа итераций таким образом, что в конце обучения корректирует свои веса только нейрон победитель. Веса нейрона-победителя и всех нейронов, лежащих в пределах его радиуса обучения, подвергаются обучению (адаптации) по правилу Кохонена:

$$w_i^{k+1} = w_i^k + \vartheta_i^k [\tilde{x} - w_i^k],$$

где  $\tilde{x}$  — входной вектор,  $k$  — номер цикла обучения,  $\vartheta_i^k$  — коэффициент скорости обучения  $i$ -го нейрона из радиуса обучения в  $k$ -ом цикле обучения. Веса нейронов, находящихся за пределами радиуса обучения, не изменяются.

В процессе обучения исходная сетка нейронов деформируется, стягиваясь в местах плотного расположения данных и растягиваясь в местах отсутствия данных. Места скопления большого количества данных (рис. 3, Б, левая часть) сетка, под «весом» данных сильно деформируется (углубляется), что иллюстрируется рисунком (рис. 3, Б, правая часть). Эти места сетки окрашиваются в синий цвет, характеризующий малые расстояния (похожесть между данными), а в местах отсутствия данных она окрашивается в желто-красную гамму, отмечая области непохожести данных. Таким образом формируется карта, напоминающая географическую карту, на которой синие области маркируют выявленные кластеры данных, а красные — границы, разделяющие выявленные кластеры (рис. 3, В, верхняя часть). На следующем шаге, найденная кластерная структура «раскрашивается» цветами, которые в данном случае характеризуют числовые значения исходных переменных, которые характерны для конкретной точки карты Кохонена, что обеспечивает интерпретируемость полученной кластерной картины (рис. 3, В, нижняя часть).

## Материалы и методы исследования

Данный раздел работы опирается на результаты численных экспериментов с теоретической моделью, описывающей ЭМП. Теоретическая модель ЭМП представлена в виде направленного графа, взятого из работы [17], и приведена на рис. 4.

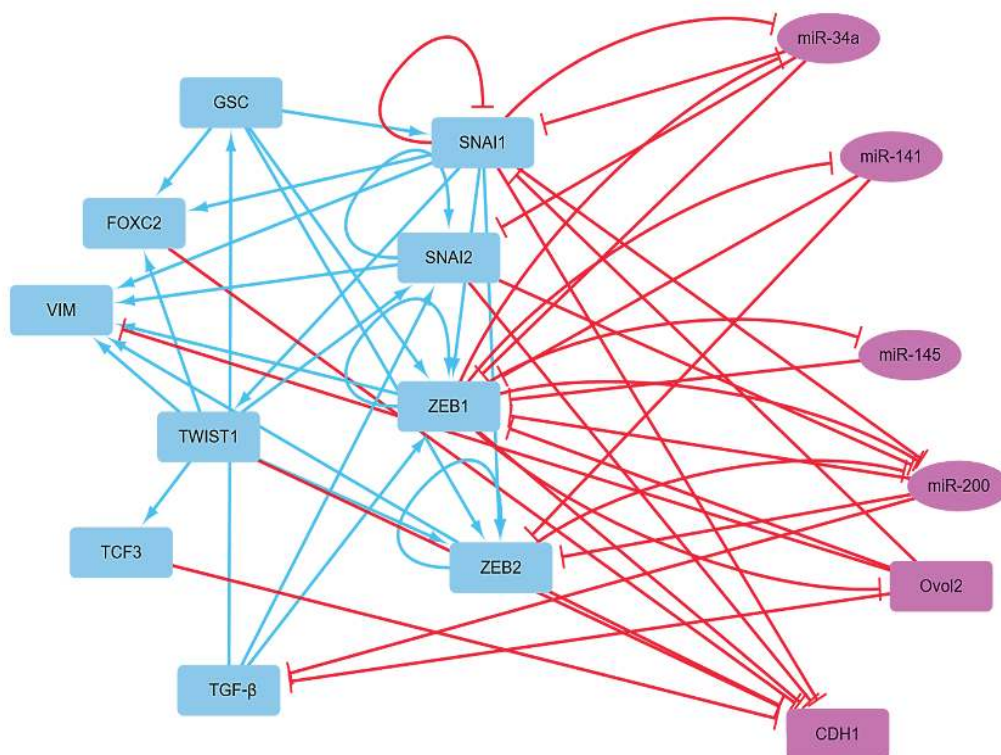


Рис. 4. Схема РС эпителиально-мезенхимального перехода [18]. Розовым цветом обозначены маркеры эпителия, а синим цветом — мезенхимальные маркеры.



Для достижения цели работы нам для анализа были необходимы временные ряды (ВР) экспрессии маркеров, формирующих РС ЭМП. Рассматриваемая РС ЭМП, как видно из **рис. 4**, включает в себя 16 компонентов — генов или микро-РНК. Здесь  $x_i$  ( $i = 1, 2, 3, \dots, 16$ ) означают уровни экспрессии маркеров TGF- $\beta$ , ZEB1, ZEB2, SNAI1, SNAI2, TWIST1, FOXC2, GSC, TCF3, VIM, miR-145, miR-141, miR-200, miR-34a, Ovov2 и CDH1 соответственно. На основе представленной на **рис. 4**, РС в виде направленного взвешенного графа с обратными связями мы составили систему обыкновенных дифференциальных уравнений (ОДУ). Синие ребра графа обозначают активационное влияние, а красные — подавление экспрессии маркеров. Шесть генов розового цвета являются Э-маркерами, а остальные гены, представленные синим цветом — это М-маркеры. Система ОДУ рассматриваемой регуляторной сети в [19], изображенной на **рис. 4**, имеет вид:

$$F_i = \dot{x}_i = \frac{dx_i}{dt} = \sum_{j \in A_i} \frac{a_{ji} x_j^n}{S_{ji}^n + x_j^n} + \sum_{j \in B_i} \frac{b_{ji} s_{ji}^n}{S_{ji}^n + x_j^n} - k x_i \quad (1)$$

где  $F_i$  представляет собой «движущую силу» временной эволюции активности  $i$ -го гена ( $i = 1, 2, 3, \dots, 16$ ).

Параметры, входящие в систему (1), можно проинтерпретировать с биологической точки зрения следующим образом:

- $k_i$  — скорость деградации
- $A_i$  — множество генов, которые активируют  $i$ -й ген в сети
- $B_i$  — множество генов, ингибирующих  $i$ -й ген в сети
- $a_{ji}$  — константа активации  $i$ -го гена со стороны  $j$ -го
- $b_{ji}$  — константа, которая отвечает за ингибирование  $i$ -го гена со стороны  $j$ -го
- $S_{ji}$  — пороговый уровень регулирования от  $x_j$  к  $x_i$
- $n$  — коэффициент Хилла.

Для завершения формирования системы дифференциальных уравнений, эквивалентной направленному графу, описывающему ЭМП нам было необходимо задать число-

вые значения параметров, входящих в уравнение (1) которые представлены в **табл. 1**.

Для описания временной эволюции (ВР) уровней экспрессии каждого из 16 маркёров (генов или микроРНК) мы использовали метод Рунге-Кутты из пакета математического моделирования Matlab. Интегрирование данной системы при разных начальных условиях генерировало множество траекторий (ВР) уровней экспрессии 16 маркёров, которые служили исходными данными для анализа по поиску состояний изучаемой системы. Поиск скрытой в многомерных ВР ЭМП структуры (кластеров) и анализ полученных синтетических ВР, описывающих ЭМП отвечает на вопросы:

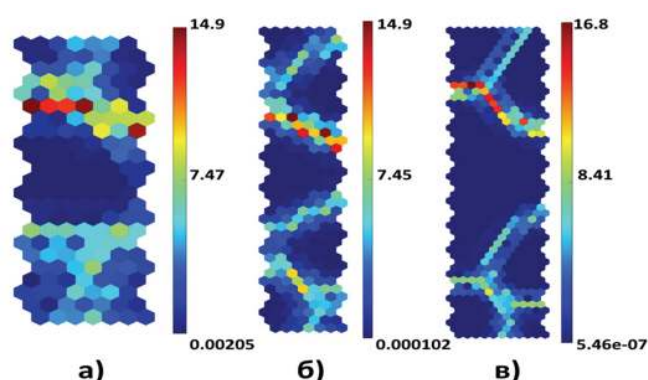
- Различимы ли между собой временные срезы пробных траекторий, запущенных при различных начальных условиях, на карте Кохонена?
- Какие фазовые переменные в наибольшей степени оказывают влияние на формирование выявленной кластерной структуры?
- Какова биологическая интерпретация каждого из кластеров?
- Являются ли выделенные кластеры устойчивыми?
- Какова природа процессов перехода между разными кластерами, за счет которых и формируется ЭМП?

## Результаты исследования

Для решения проблемы поиска состояний рассматриваемой нелинейной многомерной системы мы обратились к описанному выше методу визуализации многомерных данных SOM. Этот метод позволяет увидеть кластерную структуру данных и установить, какие из фазовых переменных участвуют в их формировании. В качестве входных данных для SOM использованы временные срезы (16-мерные вектора) синтетических экспрессионных ВР, полученных интегрированием системы (1), при разных начальных условиях (решение задачи Коши). **Рис. 5** иллюстрирует то, как меняется рельеф карты Кохонена в зависимости от количества пробных траекторий.

Таблица 1.  
Расшифровка смысла параметров системы ОДУ (1) и их числовые значения

$a$	константа активирования	2 мкмоль/час
$\tilde{a}$	константа самоактивации	7,9 мкмоль/час
$b$	константа ингибирования	4 мкмоль/час
$S_a$	пороговое значение для величины $a$	5,8 мкмоль
$S_b$	пороговое значение для величины $b$	0,21 мкмоль
$k$	уровень деградации	1/час
$n$	коэффициента Хилла	4



**Рис. 5.** SOM-карты временного среза пробных траекторий в зависимости от их количества. На рисунках «а-в» представлены срезы для 100, 250 и 1000 траекторий, полученных для различных начальных значений, соответственно (результаты нашего моделирования при разных значениях количества траекторий).

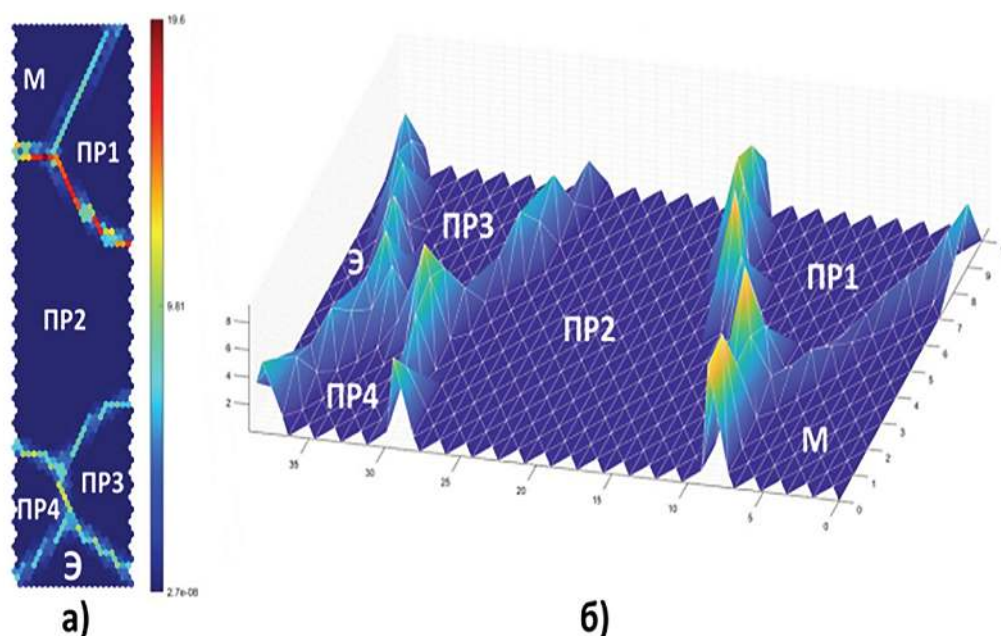
Этот рельеф представляет собой складчатую структуру. Две основные образующие этого рельефа — «моря» и разделяющие их границы, похожие на «горы». Шкала, приведенная рядом с каждой из SOM-карт — это шкала расстояний, подсчитанных по евклидовой норме между входными векторами. Эта шкала отражает степень похожести входных векторов. Так каждый из темно-синих шестиугольников на карте отображает некоторый набор входных векторов (в нашем случае временных срезов пробных траекторий), отличающихся друг от друга минимальным образом. Из этого следует, что каждая из крупных областей темно-синих шестиугольников, похожих на «моря», отображает некоторый класс векторов, похожих друг на друга и не похожих на остальные входные данные. В свою очередь, любой из нейронов карты более светлых цветов (от голубого до красного) изображает разграничение между классами и напоминает «горы» («хребты», разделяющие соседние кластеры (состояния). На **рис. 5** видно, что по мере увеличения числа пробных траекторий от 100 до 1000 SOM-карта контрастируется, а выявляемые на ней кластеры становятся более чёткими с хорошо видимыми границами между соседними состояниями. За счёт увеличения количества пробных траекторий, и, как следствие — похожих векторов внутри одного временного среза, рельеф карты преобразуется и формируются массивные темно-синие «моря» с четкими, разделяющими соседние кластеры границами.

При запуске достаточного количества пробных траекторий системы (1) и построении карт Кохонена нам удалось решить задачу кластеризации временных срезов пробных траекторий и выделить 6 кластеров, представленных на **рис. 6**.

На **рис. 6-б** представлена 3-мерная картина выявленных в ЭМП состояний. По осям  $x$  и  $y$  в плоскости расположены номера нейронов сети SOM, представляющие сложные нелинейные выражения исходных значений анализируемых маркеров, а по вертикальной оси — похожесть анализируемых векторов (временных срезов многомерных траекторий динамики экспрессии маркеров ЭМП). На этом рисунке хорошо видны границы между состояниями. Высота границ говорит о степени различий между временными точками траекторий при переходе из одного состояния в другое. Видно, что самая высокая граница отделяет состояния М и Пр1 от Пр2 и остальных состояний ЭМП. Чуть меньшая по высоте граница отделяет состояние Пр2 от состояний Пр3, Пр4 и Э. Площадь, которую занимает состояние Пр2 самая большая, что указывает на большую вариативность анализируемых маркеров в этом состоянии.

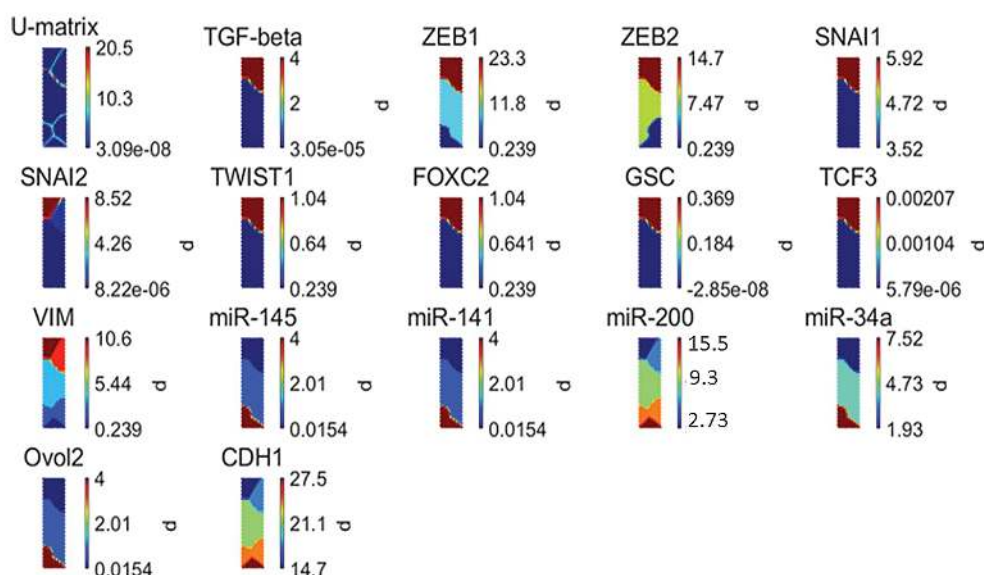
Помимо свойства SOM-карт отображать кластерную структуру данных мы воспользовались ещё одной особенностью этого метода. Чтобы сформулировать биологическую интерпретацию для найденных состояний (кластеров), необходимо установить какие переменные играют определяющую роль в формировании обнаруженной на **рис. 6** кластерной структуры. В нашем случае фазовые переменные нелинейной динамической системы (1) — это гены и микро-РНК регуляторной сети.

На **рис. 7** изображена подробная расшифровка SOM-карты, представленной на предыдущем рисунке. Видно, что SOM-карта имеет кластерную структуру с чёткими границами, и каждая из переменных участвует в её формировании. Цвет на рисунках окрашивает SOM-карту и соответствует уровню экспрессии каждого из 16 анализируемых маркеров, контрастируя состояния динамической системы,



**Рис. 6.** SOM-карты временного среза пробных траекторий в зависимости от их количества: а — кластерная структура в двумерном представлении, б — в трёхмерном представлении. Буквами Э и М обозначены эпителий и мезенхима соответственно. Промежуточные состояния обозначены как Пр1, Пр2, Пр3 и Пр4.





**Рис. 7.** Биологическая интерпретация выявленных состояний для динамической системы (1). Видно, что каждая из 16 переменных (маркеров) в разной степени участвует в формировании каждого из выявленных состояний ЭМП.

отражающих этапность эпителиально-мезенхимального перехода. Красный цвет соответствует высоким уровням экспрессии маркеров, а синий — низким уровням. В силу этого мы можем каждое из выявленных состояний проинтерпретировать с точки зрения биологии. Из **рис. 7** видно, что среди 16 анализируемых маркеров наиболее значимыми для выявленной кластерной структуры и, соответственно, в формировании всех состояний ЭМП являются следующие маркеры: CDH1, miR-200 и VIM. Об этом говорит тот факт, что раскраска SOM карты для этих маркеров полностью повторяет выявленную кластерную структуру, и все выявленные шесть состояний отличаются по уровню экспрессии анализируемых маркеров.

Результаты сопоставления уровней генной экспрессии ЭМП из [21] и значений переменных векторов из выделенных кластеров позволили установить, что самый верхний кластер SOM-карты на **рис. 6-а** соответствует мезенхиме, а нижний — эпителию. Остальные 4 кластера отвечают промежуточным состояниям. Наличие высокой границы между состояниями Пр1 и Пр2 говорит о том, что состояния Пр1 и Пр2 разведены достаточно далеко в пространстве маркеров.

Кратко рассмотрим остальные маркеры ЭМП. Маркеры ZEB1, ZEB2, miR145, miR141, miR34a и Ovof2 хорошо контрастируют состояние М и Пр1 и Э и Пр4. Для первой пары состояний эти маркеры имеют очень высокие значения уровней экспрессии. И, наоборот, для состояний Пр4 и Э значения уровней экспрессии для этих маркеров имеют очень низкие значения. Маркеры Ovof2 и miR141,145 и 34a демонстрируют обратную картину: для состояний М и Пр1 они имеют низкие значения уровней экспрессии, а для состояний Пр4 и Э — высокие значения. Остальные из неназванных маркеров (TGF- $\beta$ , SNAIL1, SNAIL2, TWIST2, FOXC2, GSC, TCF3) хорошо контрастируют со-

стояние М и Пр1 от остальных состояний. Для состояний М и Пр1 уровни их экспрессии очень велики, а для остальных состояний от Пр2 и до состояния Э их уровни экспрессии низки.

На основании полученных результатов и выявленных в многомерных данных ВР кластеров (состояний) мы можем связать между собой понятия состояния в ЭМП и аттрактора исследуемой нелинейной динамической системы. В силу этого, естественным образом встает вопрос об устойчивости аттракторов, увиденных в динамике многомерных образов ЭМП с помощью SOM-карты. Ввиду отсутствия методов классификации многомерных аттракторов и визуализации многомерного фазового пространства, мы провели численный эксперимент с выявленными состояниями (аттракторами). Мы численно исследовали поведение пробных траекторий, запущенных из областей фазового пространства, которые отображаются на SOM-карте как один кластер. Т.е., задавая пробные начальные условия вблизи границ выделенных аттракторов, мы анализировали дальнейшее поведение этих траекторий. В результате этих численных экспериментов было показано, что все выделенные в ЭМП, описываемом системой (1), состояния представляли собой устойчивые аттракторы; пробные траектории, запускаемые из точек начальных условий, близким к границам между выделенными состояниями, заканчивались в центрах выявленных кластеров — притягивались к нейрону, содержащему наиболее похожие друг на друга вектора данных.

Важно заметить два обстоятельства, характеризующие поведение пробных траекторий в нашем численном эксперименте по анализу устойчивости аттракторов:

- Вне зависимости от кластера траектории, изображенные на SOM-карте, не выходят за границы своего кластера.

- Траектории устремляются к некоторой точке внутри своего кластера.

Первый факт подтверждает утверждение о том, что выделенные кластеры являются аттракторами (областями притяжения траекторий в фазовом пространстве динамической системы с **рис. 4** и системы (1). Второе замечание позволяет нам утверждать, что выявленные аттракторы являются устойчивыми.

Отсутствие переходов через границу между кластерами и притяжение траекторий от точек близких к «горам» к точкам, расположенным в центре «моря», указывает на устойчивость каждого из 6 выделенных аттракторов. Устойчивость аттракторов, отвечающих промежуточным и терминальным состояниям клеток, ставит вопрос о механизме перехода клеток из одного состояния в другое.

На основании данных, полученных в нашей работе, мы высказываем гипотезу о регуляторной функции стресса и экспрессионного шума в реализации механизма ЭМП. Для смены клеточного фенотипа в процессе реализации ЭМП клетке, которая находится в центре аттрактора (в самом «глубоком» месте кластера) необходимо получить «импульс», который перебросит её через «потенциальный барьер» в следующее состояние. Клетки, которые находятся на периферии аттрактора, могут получать такой «импульс» за счет экспрессионного шума.

На результатах искусственной численной модели ЭМП мы показали возможность выявления состояний биологических систем по множеству характеризующих эти системы признаков (многомерных паттернов) при помощи предложенного нами инструментария.

## Выводы

Полученные результаты математического моделирования позволяю сделать ряд выводов:

1. Анализ литературы показал, что в настоящее время исследователи при выявлении состояний молекулярных, клеточных и тканевых динамических

систем работают в двумерном пространстве маркеров, и важность этих маркеров в формировании состояний оценивают в низкоразмерных пространствах.

2. Из анализа литературы следует, что многомерность играет огромную роль в процессе выявления состояний, и работа с низкоразмерными рядами может коренным образом изменять структуру (количество кластеров) в многомерных данных, что приводит к неправильной интерпретации.
3. Для выявления состояний динамической системы важного биологического процесса — ЭМП — нами был предложен метод кластеризации многомерных экспрессионных временных рядов при помощи ИНС — самоорганизующейся сети Кохоннена (SOM).
4. Предложенная нами ИНС позволила выявить в синтетических многомерных экспрессионных рядах 6 кластеров: два терминальных — эпителий и мезенхима и 4 промежуточных состояния.
5. Предложенная нами ИНС позволила определить три ведущих маркера, которые в значительной степени формируют полученную картину; этими маркерами являются CDH1, VIM и miR200, а не обычно используемые при анализе ЭМП ZEB1 и ZEB2, SNAIL1 и SNAIL2.
6. Выделенные нами в многомерных экспрессионных временных рядах кластеры интерпретируются как аттракторы динамической системы, устойчивость которых мы оценили в численном эксперименте. Было показано, что все 6 выделенных кластеров являются устойчивыми.
7. Экспериментальный факт существования ЭМП требует наличия механизма этого перехода. Поэтому на основании полученных данных мы высказали гипотезу о регуляторной роли в реализации ЭМП двух процессов — клеточного стресса и экспрессионного шума.

## Список литературы

1. Dongre A., Weinberg R.A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 2019; 20(2): 69–84. DOI: 10.1038/s41580-018-0080-4
2. Ta C.H., Nie Q., Hong T. Controlling stochasticity in epithelial-mesenchymal transition through multiple intermediate cellular states. *Discrete Contin. Dyn. Syst. — Ser. B.* 2016; 21(7): 2275–2291. DOI: 10.3934/dcdsb.2016047
3. Díaz-López A., Moreno-Bueno G., Cano A. Role of microRNA in epithelial to mesenchymal transition and metastasis and clinical perspectives. *Cancer Manag. Res.* 2014; 6: 205–216. DOI: 10.2147/CMAR.S38156
4. Francou A., Anderson K.V. The Epithelial-to-Mesenchymal Transition in Development and Cancer. *Annu. Rev. Cancer Biol.* 2020; 4: 197–220. DOI: 10.1146/annurev-cancerbio-030518-055425
5. Font-Clos F., Zapperi S., La Porta C.A.M. Topography of epithelial-mesenchymal plasticity. *Proc. Natl. Acad. Sci.* 2018; 115(23): 5902–5907. DOI: 10.1073/pnas.1722609115
6. Nicholson D.J., Dupré J. *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press, 2018. 403 p.
7. Wu Y. Transform the Information Landscape Applying Waddington Landscape to Information Theory of Aging and Epigenetic Rejuvenation. *bioRxiv*. 2024. DOI:10.1101/2024.07.20.604423
8. Henikoff S. The epigenetic landscape: An evolving concept. *Front. Epigenet. Epigenom.* 2023; 1: 1176449. DOI: 10.3389/freae.2023.1176449
9. Valcourt J.R., Huang R., Kundu S. Changing the Waddington landscape to control mesendoderm competence. *bioRxiv*. 2020. DOI: https://doi.org/10.1101/867010
10. Nicol-Benoit F., le Goff P., Michel D. Drawing a Waddington landscape to capture dynamic epigenetics. *Biol. Cell.* 2013; 105: 576–584. DOI: 10.1111/boc.201300029
11. Waddington K. Genetic assimilation. *Adv. Genet.* 1961; 10: 257–293. DOI: 10.1016/s0065-2660(08)60119-4

12. Matsushita Y., Kaneko K. Homeorhesis in Waddington's Landscape by Epigenetic Feedback Regulation. *Phys. Rev. Research*. 2020; 2: 023083. DOI: 10.1103/PhysRevResearch.2.023083
13. Gruber P., Jeronimidis G. Has biomimetics arrived in architecture? *Bioinspir. Biomim.* 2012; 7(1): 010201. DOI: 10.1088/1748-3182/7/1/010201
14. Feinberg A.P., Levchenko A. Epigenetics as a mediator of plasticity in cancer. *Science*. 2023; 379 (6632): eaaw3835. DOI: 10.1126/science.aaw3835
15. Fagan M.B. Waddington redux: models and explanation in stem cell and systems biology. *Biol. Philos.* 2012; 27: 179–213. DOI: 10.1007/s10539-011-9294-y
16. Schmidt M., Loeffler-Wirth H., Binder H. Developmental scRNAseq Trajectories in Gene- and Cell-State Space-The Flatworm Example. *Genes (Basel)*. 2020; 11(10): 1214. DOI: 10.3390/genes11101214
17. Greenbury S.F., Louis A.A., Ahnert S.E. The structure of genotype-phenotype maps makes fitness landscapes navigable. *bioRxiv*, 2021. DOI: 10.1101/2021.10.11.463990
18. Lang J., Nie Q., Li C. Landscape and kinetic path quantify critical transitions in epithelial-mesenchymal transition. *Biophys. J.* 2020; 120(20): 4484–4500. DOI: 10.1016/j.bpj.2021.08.043
19. Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps. *Biol. Cybern.* 1982, 43 (1): 59–69. DOI: 10.1007/BF00337288
20. Kohonen T. *Self-Organizing Maps*. Berlin — Heidelberg: Springer, 1997. 501 p.
21. Зиновьев А.Ю. *Визуализация многомерных данных*. Красноярск: Издательство Красноярского государственного технического университета, 2000. 180 с.
22. Valles-Perez I. *Self-Organizing Maps — the Kohonen's algorithm explained*. ivape3. Режим доступа: <https://ivape3.blogs.uv.es/2015/03/15/self-organizing-maps-the-kohonens-algorithm-explained/> Дата обращения: 19.10.2025.
23. Ilarinieminen Ilarinieminen/SOM-Toolbox, 2025. Режим доступа: <https://github.com/ilarinieminen/SOM-Toolbox> Дата обращения: 19.10.2025.

## References

1. Dongre A., Weinberg R.A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell. Biol.* 2019; 20(2): 69–84. DOI: 10.1038/s41580-018-0080-4
2. Ta C.H., Nie Q., Hong T. Controlling stochasticity in epithelial-mesenchymal transition through multiple intermediate cellular states. *Discrete Contin. Dyn. Syst. — Ser. B*. 2016; 21(7): 2275–2291. DOI: 10.3934/dcdsb.2016047
3. Díaz-López A., Moreno-Bueno G., Cano A. Role of microRNA in epithelial to mesenchymal transition and metastasis and clinical perspectives. *Cancer Manag. Res.* 2014; 6: 205–216. DOI: 10.2147/CMAR.S38156
4. Francou A., Anderson K.V. The Epithelial-to-Mesenchymal Transition in Development and Cancer. *Annu. Rev. Cancer Biol.* 2020; 4: 197–220. DOI: 10.1146/annurev-cancerbio-030518-055425
5. Font-Clos F., Zapperi S., La Porta C.A.M. Topography of epithelial-mesenchymal plasticity. *Proc. Natl. Acad. Sci.* 2018; 115(23): 5902–5907. DOI: 10.1073/pnas.1722609115
6. Nicholson D.J., Dupré J. *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press, 2018. 403 p.
7. Wu Y. Transform the Information Landscape Applying Waddington Landscape to Information Theory of Aging and Epigenetic Rejuvenation. *bioRxiv*. 2024. DOI:10.1101/2024.07.20.604423
8. Henikoff S. The epigenetic landscape: An evolving concept. *Front. Epigenet. Epigenom.* 2023; 1: 1176449. DOI: 10.3389/frae.2023.1176449
9. Valcourt J.R., Huang R., Kundu S. Changing the Waddington landscape to control mesendoderm competence. *bioRxiv*. 2020. DOI: <https://doi.org/10.1101/867010>
10. Nicol-Benoit F., le Goff P., Michel D. Drawing a Waddington landscape to capture dynamic epigenetics. *Biol. Cell.* 2013; 105: 576–584. DOI: 10.1111/boc.201300029
11. Waddington K. Genetic assimilation. *Adv. Genet.* 1961; 10: 257–293. DOI: 10.1016/s0065-2660(08)60119-4
12. Matsushita Y., Kaneko K. Homeorhesis in Waddington's Landscape by Epigenetic Feedback Regulation. *Phys. Rev. Research*. 2020; 2: 023083. DOI: 10.1103/PhysRevResearch.2.023083
13. Gruber P., Jeronimidis G. Has biomimetics arrived in architecture? *Bioinspir. Biomim.* 2012; 7(1): 010201. DOI: 10.1088/1748-3182/7/1/010201
14. Feinberg A.P., Levchenko A. Epigenetics as a mediator of plasticity in cancer. *Science*. 2023; 379 (6632): eaaw3835. DOI: 10.1126/science.aaw3835
15. Fagan M.B. Waddington redux: models and explanation in stem cell and systems biology. *Biol. Philos.* 2012; 27: 179–213. DOI: 10.1007/s10539-011-9294-y
16. Schmidt M., Loeffler-Wirth H., Binder H. Developmental scRNAseq Trajectories in Gene- and Cell-State Space-The Flatworm Example. *Genes (Basel)*. 2020; 11(10): 1214. DOI: 10.3390/genes11101214
17. Greenbury S.F., Louis A.A., Ahnert S.E. The structure of genotype-phenotype maps makes fitness landscapes navigable. *bioRxiv*, 2021. DOI: 10.1101/2021.10.11.463990
18. Lang J., Nie Q., Li C. Landscape and kinetic path quantify critical transitions in epithelial-mesenchymal transition. *Biophys. J.* 2020; 120(20): 4484–4500. DOI: 10.1016/j.bpj.2021.08.043
19. Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps. *Biol. Cybern.* 1982, 43 (1): 59–69. DOI: 10.1007/BF00337288
20. Kohonen T. *Self-Organizing Maps*. Berlin — Heidelberg: Springer, 1997. 501 p.
21. Zinoviev A. Yu. *[Visualization of Multidimensional Data]*. Krasnoyarsk: Publishing House of Krasnoyarsk State Technical University, 2000. 180 p.
22. Valles-Perez I. *Self-Organizing Maps — the Kohonen's algorithm explained*. ivape3. Available at: <https://ivape3.blogs.uv.es/2015/03/15/self-organizing-maps-the-kohonens-algorithm-explained/> Retrieved: 19.10.2025.
23. Ilarinieminen Ilarinieminen/SOM-Toolbox, 2025. Available at: <https://github.com/ilarinieminen/SOM-Toolbox> Retrieved: 19.10.2025.

### Сведения об авторах:

**Ситдикова Ксения Камильевна** — младший научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0000-0001-8870-7403>

**Устинова Екатерина Евгеньевна** — младший научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0009-0008-8976-0212>



---

*Черкашенко Владимир Николаевич* — научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0009-0009-0071-471X>

*Сабурина Ирина Николаевна* — доктор биологических наук, заведующая лабораторией клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0000-0003-2014-2535>