



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA
CURSO DE GRADUAÇÃO EM FÍSICA

DIÊGO TAVARES CANAFISTULA

DINÂMICA DE REDES DE REGULAÇÃO GENÉTICA: ESTUDO DO UNSAT-FFF

FORTALEZA

2023

DIÊGO TAVARES CANAFISTULA

DINÂMICA DE REDES DE REGULAÇÃO GENÉTICA: ESTUDO DO UNSAT-FFF

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Física.

Orientador: Prof. Dr. Saulo Davi Soares e Reis.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- C219d Canafistula, Diêgo Tavares.
Dinâmica de redes de regulação genética : estudo do unsat-fff / Diêgo Tavares Canafistula. – 2023.
43 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências,
Curso de Física, Fortaleza, 2023.
Orientação: Prof. Dr. Saulo Davi Soares e Reis.
1. Redes de regulação. 2. Genética. 3. Cognição celular. I. Título.

CDD 530

DIÊGO TAVARES CANAFISTULA

DINÂMICA DE REDES DE REGULAÇÃO GENÉTICA: ESTUDO DO UNSAT-FFF

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Física.

Aprovada em: 06/06/2023.

BANCA EXAMINADORA

Prof. Dr. Saulo Davi Soares e Reis (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Humberto de Andrade Carmona
Universidade Federal do Ceará (UFC)

Prof. Dr. Diego Rabelo da Costa
Universidade Federal do Ceará (UFC)

Aos meus avós. Aos meus pais, familiares e
amigos

AGRADECIMENTOS

Agradeço primeiramente à minha família. Minha avó Fátima e meus pais Edileuza e Firmino. Agradeço a todos os meus amigos e professores que me ajudaram nessa longa caminhada desse curso.

Ao Prof. Dr. Saulo Davi Soares e Reis, pela excelente orientação. Aos professores participantes da banca examinadora Prof. Dr. Humberto de Andrade Carmona e Prof. Dr. Diego Rabelo da Costa pelo tempo, pelas valiosas colaborações e sugestões recebidas.

Agradecer a três amigos que foram de extrema importância na hora de escrever esse trabalho, Ana Thaís, André Albuquerque e Cristian Vicson. Agradecer também a toda a coordenação do curso pelo suporte e a paciência que tiveram comigo, principalmente ao Prof Dr. José Ramos e a Rebeca Matos Freire.

“A confiança não vem de fora, a confiança vem de dentro e não há força mais poderosa no mundo que quando você confia em si mesmo.” (VOJVODA, 2022)

RESUMO

A física é usada para o estudo da natureza e de tudo que há nela, sendo a parte mais complexa os seres vivos. Os seres vivos são estudados com mais especificidade na biologia, porém, os outros campos das ciências, podem usar diversas ferramentas desenvolvidas a partir da observação da dinâmica de seres não-vivos, para entender o funcionamento da dinâmica dos seres vivos. Entender essa dinâmica é importante porque nossa vida é conduzida por ela, e qualquer dano antinatural que podemos vir a sofrer, pode ser previsto e impedido dado o nível de conhecimento que temos do todo. Esse trabalho, a partir de uma revisão bibliográfica, procura demonstrar que a dinâmica da menor parte funcional de um ser vivo, está estritamente correlacionada com a dinâmica aparente do ser como um todo. Estudando a cognição das células e como elas reagem a mudanças no ambiente, por meio das redes de transcrição e análise do UNSAT-FFF, podemos entender da onde vem alguns comportamentos e doenças que acabam tendo algum efeito negativo na vida do ser humano.

Palavras-chave: dinâmica transcricional; cognição celular; genética

ABSTRACT

Physics is used to study nature and everything in it, the most complex part being living beings. Living beings are studied more specifically in biology, however, other fields of science can use several tools developed from the observation of the dynamics of non-living beings, to understand the functioning of the dynamics of living beings. Understanding this dynamic is important because our life is driven by it, and any unnatural damage we may suffer can be predicted and prevented given the level of knowledge we have of the whole. This work, based on a bibliographic review, seeks to demonstrate that the dynamics of the smallest functional part of a living being is correlated with the apparent dynamics of the being as a whole. By studying cells and how they react to changes in the environment, through transcription networks, we can understand where some behaviors and diseases come from that end up having a negative effect on human life.

Keywords: transcriptional dynamics; cellular cognition; genetics.

LISTA DE FIGURAS

Figura 1 – Fotografia de uma célula obtida da pele humana e crescida em cultura foi fotografada com um microscópio optico utilizando lentes de contraste de interferência	16
Figura 2 – Figura de representação da fita de DNA e do DNA de dupla helice	17
Figura 3 – Figura de representação da transcrição e tradução de novas proteínas que são codificadas pelo DNA. e do DNA de dupla helice	18
Figura 4 – Figura de representação de um fator de transcrição ativador.	19
Figura 5 – Figura de representação de um fator de transcrição repressor.	19
Figura 6 – Figura do modelo de estudo de genes.	21
Figura 7 – Gráfico da função de hil e função degrau.	23
Figura 8 – Gráfico da função de hil e função degrau.	24
Figura 9 – Gráfico da função de decaimento de Y.	26
Figura 10 – Gráfico da função de nível de expressão de Y.	27
Figura 11 – Figura com tipos de FFL, coerentes e incoerentes.	29
Figura 12 – Figura com o diagrama de um FFL funcionando em lógica AND.	30
Figura 13 – Gráfico que mostra a função da concentração de Y e Z em relação ao tempo.	31
Figura 14 – Figura da representação em grafos de um FFL e um FFF.	33
Figura 15 – Uma rede de transcrição genética da bactéria Escherichia coli.	34
Figura 16 – Figura que mostra o isomorfismo em genes diferentes.	35
Figura 17 – Figura que mostra a representação em grafo do UNSAT-FFF.	36
Figura 18 – Gráfico que mostra a função da concentração de ψ em relação ao tempo.	39

LISTA DE ABREVIATURAS E SIGLAS

AND	Porta lógica que precisa de dois ativadores para um resultado positivo na saída
AR	Autorregulação
C1-FFL	FFL coerente do tipo 1
DNA	Ácido desoxirribonucleico
EDD	Equação diferencial com delay
EDO	Equação diferencial ordinária
FFL	Feed Forward Loop
miRNAs	Micro RNA's
mRNA	Ácido ribonucleico mensageiro
OR	Porta lógica que precisa de um ativador para um resultado positivo na saída
RNA	Ácido ribonucleico
RNAp	Ácido ribonucleico polimerase

LISTA DE SÍMBOLOS

X	Representação de um gene
X^*	Forma ativa do gene
β	Máxima atividade do promotor de um gene
θ	Função degrau
α	Coefficiente de degradação
K	Coefficiente de ativação
Y_{st}	Estado estável de concentração
T_{ON}	Tempo de atraso
S_x	Sinal de ativação
ψ	Concentração do gene Y reescalada
ζ	Concentração do gene Z reescalada
δ	Atividade dos promotores reescalada
f_0	Função inicial de concentração
τ	Intervalo de tempo de delay

SUMÁRIO

1	INTRODUÇÃO	12
2	CÉLULA	15
2.1	A célula e seu funcionamento	16
3	REDES DE TRANSCRIÇÃO E SUA DINÂMICA	21
3.1	Diagrama de regulação	22
3.2	Dinâmica de uma regulação simples	25
3.3	Network motifs	27
4	MOTIFS COMO BASE PARA A BIOLOGIA COMPUTACIONAL	33
4.1	Feed Foward Fiber e a simetria de fibração	33
4.2	A dinâmica do UNSAT-FFF	36
5	CONCLUSÕES E TRABALHOS FUTUROS	41
	REFERÊNCIAS	42

1 INTRODUÇÃO

Quando estudamos física no seu estado mais fundamental, tratamos da dinâmica de interações entre partículas da maneira mais simples que pudermos. O trabalho de um físico é estudar a natureza da maneira mais acurada possível e tentar não só definir o que está acontecendo no momento, mas também tentar prever o que pode vir a acontecer posteriormente. Nas ciências biológicas o foco do estudo são os seres vivos, não podemos entrar na eterna discussão do que é a definição de vida, mas uma boa definição que podemos usar é: fábricas químicas interessantes e organizadas de forma complexa que recebem substâncias de sua vizinhança e as utilizam como matérias-primas para gerar cópias de si próprias, (ALBERTS *et al.*, 2010).

As células foram observadas primeiramente pelo inglês Robert Hooke, (HOOKE, 1968), que a partir de observações feitas em um microscópio rudimentar, conseguiu analisar uma fina camada de cortiça e notar que ela era formada por cavidades poliédricas, que foram chamadas de células. Após a observação de Hooke, as células foram estudadas por outros cientistas, entre eles, Anton van Leeuwenhoek, que foi o primeiro a estudar células vivas e observou diferentes seres unicelulares, (LEEUWENHOEK, 1966).

A descoberta da célula é de extrema importância mas o que de fato colocou a célula no patamar de unidade básica da vida foi a teoria celular primeiramente formulada por dois cientistas, Theodor Schwann e Matthias Schleiden; Matthias escreveu pela primeira vez que “todas as plantas são compostas por células”, (SCHLEIDEN, 1838)) essa foi uma das bases para o desenvolvimento da teoria celular que ficou quase completa com o trabalho de Theodor, no qual ele completou a teoria com seu estudo sobre o crescimento de animais e plantas (SCHWANN, 1838), esses dois trabalhos deram a base necessária para a consolidação da teoria celular. A teoria celular se baseia em alguns postulados sobre as células, (SCHWANN, 1838):

- Todos os seres vivos são formados por uma ou mais células;
- A célula é a unidade funcional e estrutural de todos os seres vivos ;
- Todas as células provêm de células pré-existentes, através da divisão celular;

Com o avanço da tecnologia e invenção do microscópio eletrônico a teoria celular moderna adicionou mais três postulados nos dias de hoje (ALBERTS *et al.*, 2017):

- As células contêm as informações de hereditariedade Ácido desoxirribonucleico (DNA), que são transmitidas de célula a célula;
- Todas as células têm basicamente a mesma composição química;

-Todo o fluxo de energia da vida (metabolismo e bioquímica) ocorre dentro das células.

Com a consolidação da teoria celular, conseguimos ter um ponto de partida para um estudo com mais precisão sobre a vida e a natureza. Considerando a célula como a unidade básica da vida, podemos afirmar que qualquer problema no funcionamento de seres vivos podem ser definidos como algum “mau funcionamento” de algum organismo celular. Os organismos são compostos por uma enormidade de diferentes células, porém, como foi visto nos postulados, elas tem uma mesma composição química e funcionam de uma forma semelhante, por mais que tenham diferentes finalidades. “Toda a biologia é um contraponto entre os dois temas: a admirável variedade em particularidades individuais e a admirável constância nos mecanismos fundamentais” (ALBERTS *et al.*, 2010).

Para exemplificar a grande variedade das células podemos comparar duas conhecidas, o *lactobacillus*, uma bactéria que podemos achar em um pedaço de queijo e o ovo de rã que também é uma só célula. O primeiro tem o tamanho da ordem de micrômetro e o segundo tem o diâmetro na ordem de um milímetro, isso nos mostra a grande diferença de tamanho entre duas células distintas, com finalidades completamente diferentes mas que funcionam de uma forma similar, e é essa similaridade de funcionamento que nos permite generalizar os processos internos de interação que ocorrem em uma célula, para todos os tipos conhecidos, (ALBERTS *et al.*, 2017).

Os seres vivos são acometidos por diversas doenças, o primeiro passo para lidar com elas é entender como a máquina viva funciona e buscar em qual processo dela ocorreu o mau funcionamento. O corpo humano é um agregado de células que possui em torno de 10^{13} células na sua composição (ALBERTS *et al.*, 2010), formando assim um sistema extremamente complexo, e antes de fazermos qualquer afirmação sobre o funcionamento do mesmo, vamos procurar entender o funcionamento de uma célula só, que por sua vez continua sendo um sistema complexo pois contém uma grande gama de moléculas no seu interior interagindo entre si e com o ambiente para o desenvolvimento, sobrevivência e evolução da célula. Assim como na física precisamos primeiro entender como o sistema funciona para poder fazer previsões e prevenções, o estudo do funcionamento da unidade básica da vida nos dá a oportunidade de descobrir e prevenir a grande maioria dos problemas que aparecem na vida dos seres vivos. 16

Quando viramos nossa atenção para as interações que ocorrem dentro da célula, vemos que o problema é muito mais complexo do que estamos acostumados, para os físicos

sistemas com mais de dois corpos interagindo não possuem solução analítica e uma previsão acurada está fora de questão, então como conseguiremos estudar as interações de um sistema com milhares de moléculas e ainda tentar prever algo? A resposta é mais simples do que podemos imaginar, nós vamos focar nas questões qualitativas em vez das quantitativas. As moléculas que vamos observar e buscar entender, são chamadas de fatores de transcrição (ALON, 2019), eles serão uma das nossas ferramentas de estudo daqui pra frente. Mais importante que saber a posição e momento de cada um fator de transcrição, é saber o nível de expressão daquele fator e como ele reage na presença e ausência de outros fatores, que juntos formam uma rede. Essa rede é o nosso objetivo final de estudo a rede de transcrição.

Nesse trabalho nosso objetivo é fazer uma ponte entre a biologia e a física e usar os modelos já existentes de equações diferenciais e sistemas complexos para entender a dinâmica de comunicação nas redes de transcrição genéticas e a partir de uma análise matemática qualitativa, tentar explicar e prever fenômenos que acontecem em sistemas biológicos.

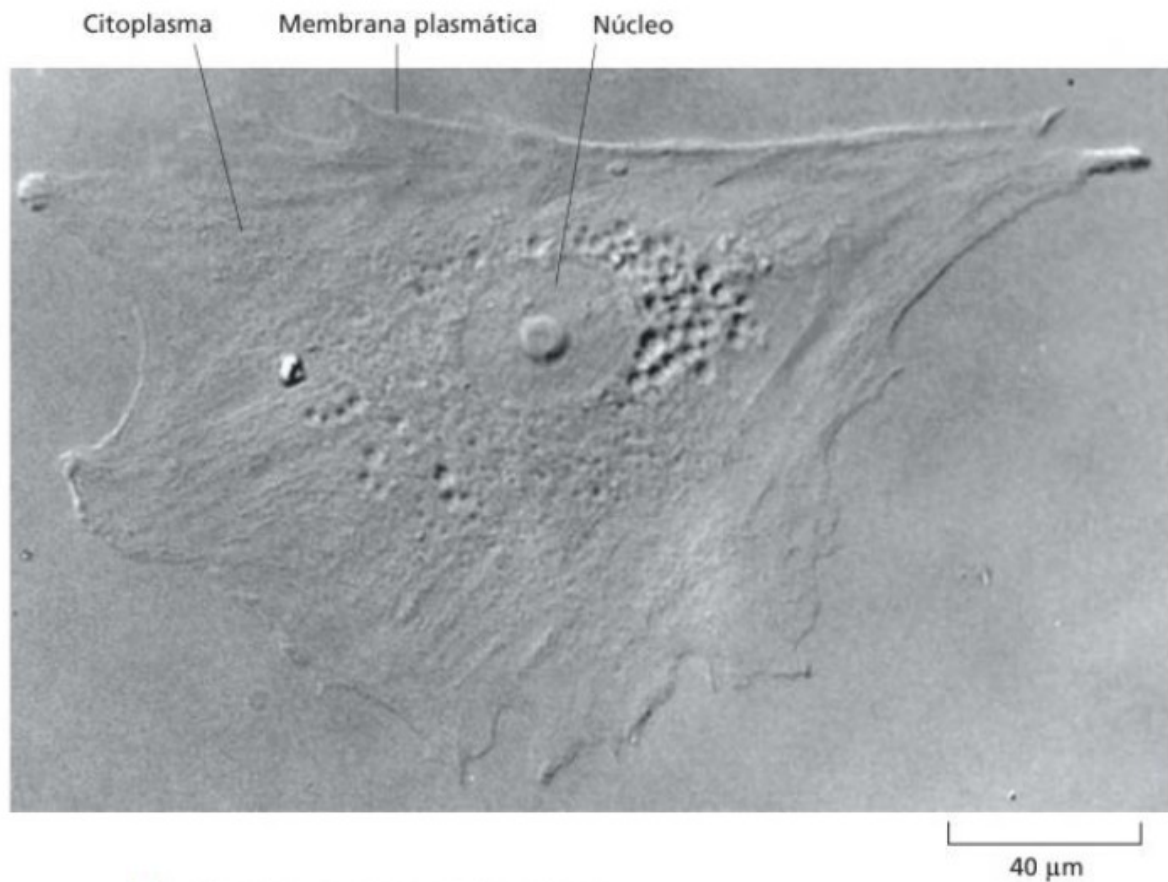
2 CÉLULA

Para iniciar o estudo das redes de transcrição começamos procurando entender o que é uma célula e como ela funciona. A célula é a menor unidade funcional da matéria viva. A célula é delimitada por uma barreira chamada de membrana plasmática que envolve o citoplasma, que é o material composto por uma substância transparente contendo uma mistura de objetos heterogêneos, onde podemos observar na figura 1. As células são classificadas em eucariontes e procariontes, a diferença entre elas está na ausência do núcleo nas células procariontes e em uma estrutura mais complexa nas eucariontes. (ALBERTS *et al.*, 2017).

Dentro da membrana plasmática temos diversos dispositivos que possuem cada um, uma determinada função. A proteína é a principal máquina da formação da célula, ela determina o comportamento e aparência do corpo celular e serve como suporte estrutural, catalisadores químicos e motores moleculares, elas são compostas por aminoácidos, que é o tijolo básico da formação da proteína, e cada organismo utiliza o mesmo grupo de 20 aminoácidos para sintetizar suas proteínas. O núcleo é onde as informações das células são armazenadas e o ribossomo, um grande complexo macromolecular composto de 80 a 90 proteínas individuais e moléculas de ácido ribonucleico (RNA), é o local onde as proteínas são fabricadas (ALBERTS *et al.*, 2017). Um exemplo de célula é mostrado na figura 1.

Cada proteína é uma máquina que realiza uma tarefa específica com precisão, a célula está constantemente monitorando o ambiente em que está imersa, e para cada situação ocorrida precisa de um diferente tipo e uma diferente quantidade de proteína, (ALON, 2019)

Figura 1 – Fotografia de uma célula obtida da pele humana e crescida em cultura foi fotografada com um microscópio óptico utilizando lentes de contraste de interferência



Fonte: (ALBERTS *et al.*, 2017) Cortesia de Casey Cunningham para o livro ALBERTS, Bruce et al. Fundamentos da Biologia Celular. 4. ed. Porto Alegre: Artmed Editora, 2017.

2.1 A célula e seu funcionamento

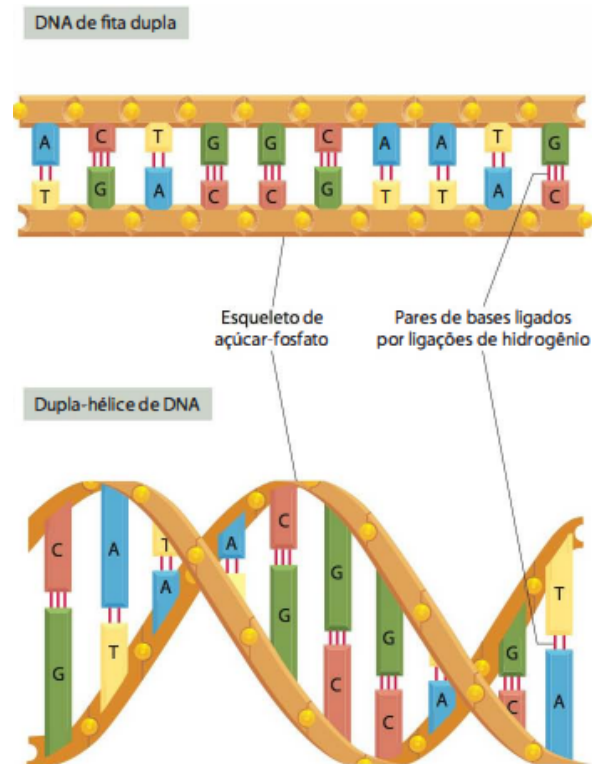
O ambiente de uma célula é complexo e pode ter diferentes sinais, tais como, temperatura, pressão, sinais biológicos de outras células, nutrientes, etc. Os sinais internos da célula também são importantes para o funcionamento do sistema, elas usam os sinais obtidos para produzir as proteínas necessárias para agir em cada diferente tipo de situação, seja no ambiente interno ou externo. Para representar os estados do ambiente as células usam proteínas especiais chamadas de fatores de transcrição, (ALON, 2019).

Essas proteínas podem se ligar ao ácido desoxirribonucleico (DNA), que é uma molécula presente no núcleo da célula encarregada de guardar toda as informações necessárias para manutenção, desenvolvimento e funcionamento de todo os seres vivos. A molécula de DNA é formada por uma fita dupla em formato de hélice. As fitas possuem um esqueleto de fosfato e açúcar e são ligadas por meio de bases ligada por ligações de hidrogênio, (ALBERTS et al.,

2010).

Para guardar a informação que vai ser passada adiante as bases: (A) adenina, (G) guanina, (C) citosina e (T) timina, seguem uma lei de ligação que faz com que a partir de uma fita, pode ser feita uma duplicação do DNA pré-existente. A lei que as bases seguem é: A liga-se com T, e C liga-se com G, como podemos ver na figura 2.

Figura 2 – Figura de representação da fita de DNA e do DNA de dupla hélice



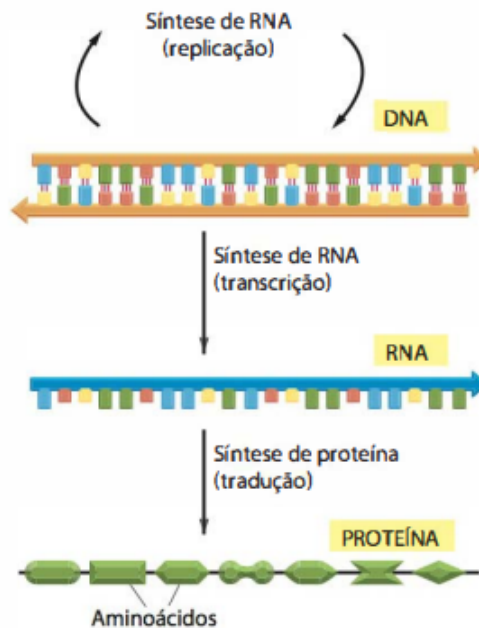
Fonte: (ALBERTS *et al.*, 2010) A figura mostra a forma com que o DNA se liga e a lei que as bases de ligação segue.

Nas células procarióticas o DNA também está presente, porém, não está envolvido em um envelope nuclear como na eucariótica. Por ser uma molécula muito grande, o DNA é subdividido em diversos segmentos menores e cada um desses segmentos possui a informação para a sintetização de uma proteína, eles são os genes, (ALBERTS *et al.*, 2010). Existe uma classe de proteínas que fica encarregada de trazer a informação necessária do ambiente para dar início ao processo de síntese de novas proteínas, essa classe é chamada de fator de transcrição, ele tem a capacidade de se conectar a um gene e controlar a taxa na qual aquele gene será lido e posteriormente, transcrito, para a proteína resultante poder agir no ambiente, (ALON, 2019).

Como temos proteínas como fatores de transcrição e essas proteínas são codificadas

pelos genes que por sua vez são reguladas por outros fatores de transcrição e assim sucessivamente, vemos que esse conjunto de interações formam uma rede de interação complexa, e essa rede pode descrever toda a regulação de interação de transcrição numa célula, chamamos essa rede de Redes de transcrição, (ALON, 2019). A rede de transcrição descreve a interação entre os fatores de transcrição e os genes, no gene conseguimos ter a fórmula necessária para produzir uma determinada proteína, essa proteína será produzida resumidamente em dois passos: transcrição e tradução como vemos na figura 3.

Figura 3 – Figura de representação da transcrição e tradução de novas proteínas que são codificadas pelo DNA. e do DNA de dupla helice



Fonte: (ALBERTS *et al.*, 2010) A figura mostra os principais passos que ocorrem no processo de construção de uma proteína.

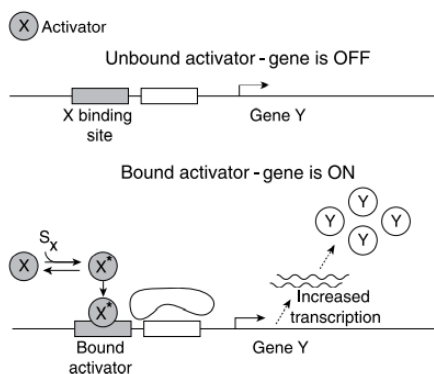
A transcrição começa quando uma molécula chamada, RNA polimerase Ácido ribonucleico polimerase (RNAP) se liga ao DNA por meio do fator de transcrição e começa a copiar o gene em uma molécula descartável chamada de RNA mensageiro Ácido ribonucleico mensageiro (mRNA), (ALON, 2019).

O esqueleto do RNA possui uma grande semelhança com o DNA, porém há uma mudança em uma das bases, em vez de timina (T) o RNA terá a uracila (U). Quando o processo de transcrição termina o mRNA então é transladado para o ribossomo onde vai ser feita a produção

da proteína necessária. Existe uma região do DNA que precede o gene e define a taxa na qual o mRNA é produzido, essa região é chamada de promotor.

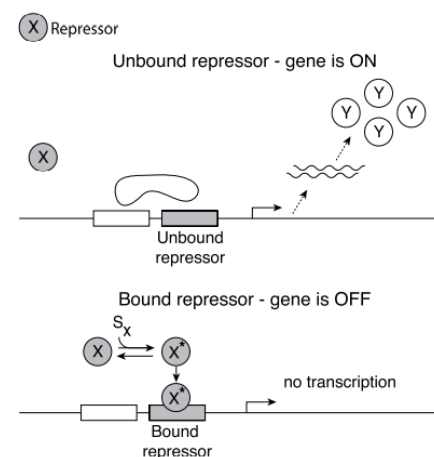
O RNAP atua em todo tipo de gene, e o que vai modular a expressão desses genes são justamente os fatores de transcrição. Os fatores de transcrição podem funcionar de duas formas, ativadores ou repressores, podendo aumentar ou diminuir a taxa transcrição a forma mais intuitiva de se representar a ativação e repressão de fatores de transcrição pode ser vista nas figuras 4 e 5.

Figura 4 – Figura de representação de um fator de transcrição ativador.



Fonte: (ALON, 2019) A figura mostra os principais passos que ocorrem no processo de ativação na transcrição de uma nova proteína.

Figura 5 – Figura de representação de um fator de transcrição repressor.



Fonte: (ALON, 2019) A figura mostra os principais passos que ocorrem no processo de repressão na transcrição de uma nova proteína.

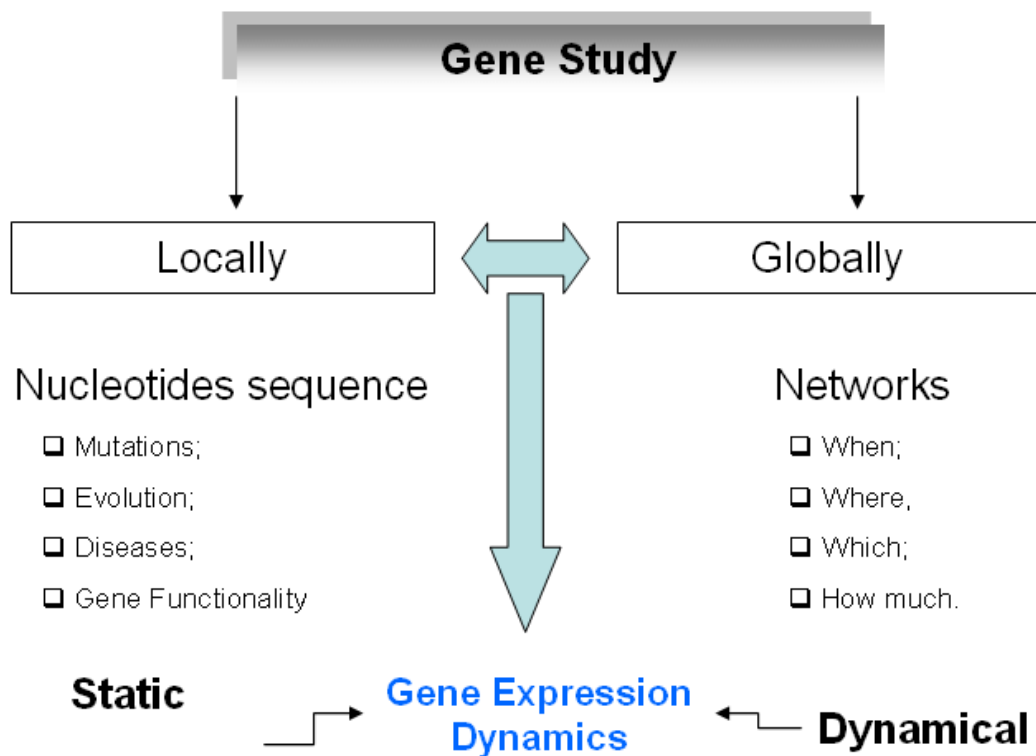
Redes de transcrição são projetadas com uma forte separação de escala de tempo: os

sinais de entrada mudam os fatores de transcrição em uma escala de tempo menor que o segundo, por sua vez para a ligação dos fatores de transcrição como DNA duram alguns segundos, a transcrição e translação do gene leva minutos e a acumulação da proteína obtida pode levar horas, (ALON, 2019).

3 REDES DE TRANSCRIÇÃO E SUA DINÂMICA

O estudo da genética na biologia pode ser feito por dois pontos de vista; antagonistas, porém, complementares: estático e dinâmico. Essas formas de estudar a genética são vistas em Pires (2012). A forma estática busca entender as mutações que um gene pode sofrer, a evolução e as doenças que podem ser derivadas de algum erro no código genético, ela pode nos dizer qual proteína pode ser sintetizada e qual é fator de transcrição que ativa aquele determinado gene, mas falha quando procuramos saber o quanto, como e onde o gene será ativado, essas perguntas são respondidas pela dinâmica e redes. A divisão entre estático e dinâmico vemos com clareza na figura 6.

Figura 6 – Figura do modelo de estudo de genes.



Fonte: (PIRES, 2012) A figura mostra as diferentes formas de se estudar a genética e as principais perguntas relacionadas a cada uma delas.

Para iniciar o estudo da dinâmica, primeiro temos que entender as nossas limitações tecnológicas; a dificuldade de medir o nível real de concentração de proteína é grande, por causa do tamanho da célula e da necessidade de destruir ela para poder “contar” a quantidade de proteínas sintetizadas. As observações dos níveis de expressão feitas até agora, são exclusivamente de mRNA e os dados catalogados são feitos por algumas técnicas bem avançadas, tais como,

real time polymerase chain reaction (PCR), (PFAFFL, 2001), Green Fluorescent Protein (GFP), (ZASLAVER *et al.*, 2006), RNA-Seq, (WANG *et al.*, 2009) e a já consolidada micro-arrays, (BROWN; BOTSTEIN, 1999).

É importante salientar logo no começo que os dados que vamos usar são especificamente dados de concentração de mRNA, e que para um melhor entendimento vamos considerar que o que acontece com o mRNA também vai acontecer com as proteínas sintetizadas, mesmo sabendo que cada um vai ter particularidades na sua dinâmica. Essa aproximação que fazemos com a taxa de transcrição de mRNA com a taxa de sintetização de proteína já foi mostrada que não funciona na grande maioria das vezes por causa da existência dos microRNA's Micro RNA's (miRNAs), que são RNA's que se ligam aos mRNA's e impedem a tradução da proteína, (LAU *et al.*, 2001).

3.1 Diagrama de regulação

Os fatores de transcrição são postos em diagramas com setas que indicam se estão trabalhando como ativadores ou repressores, essas setas não carregam somente o sinal mas também carrega a força de interação que é definida por uma função de entrada (ALON, 2019). Em estudos sobre a regulação da síntese de enzimas indutivas, a enzima é normalmente produzida a uma baixa taxa basal de produção. A partir do momento que se aumenta a concentração do indutor (ativador) até um limite de concentração, a síntese da enzima sobe rapidamente até uma taxa limite de produção, (YAGIL; YAGIL, 1971).

Primeiramente para entender o diagrama de setas vamos pegar um exemplo simples de regulação, Considerando a taxa de produção de uma proteína Y controlada por um fator de transcrição X, quando X regula Y, representado por $X \rightarrow Y$ o número de moléculas de proteína Y produzida por unidade de tempo é uma função da concentração de X na forma de X^* , que é X após ser ativada pelos sinais do ambiente que chamaremos de S_x . Logo, teremos nossa função de Y como:

$$Y = f(X^*). \quad (3.1)$$

A função da Equação (3.2) pode ter duas formas, uma monotonicamente crescente no caso do fator ser um ativador e uma monotonicamente decrescente para o caso do fator ser um repressor, temos uma função extremamente útil que pode representar várias funções de entrada

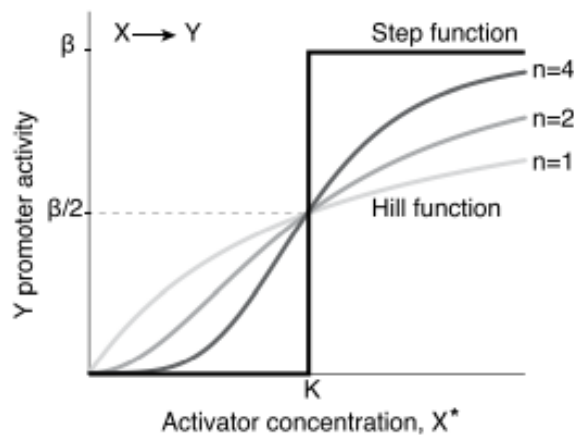
de gene, a função de HILL, e ela tem a seguinte forma:

$$f(X^*) = \beta \frac{(X^{*n})}{(K^n + X^{*n})}. \quad (3.2)$$

$$f(X^*) = \beta \frac{(K^n)}{(K^n + X^{*n})}. \quad (3.3)$$

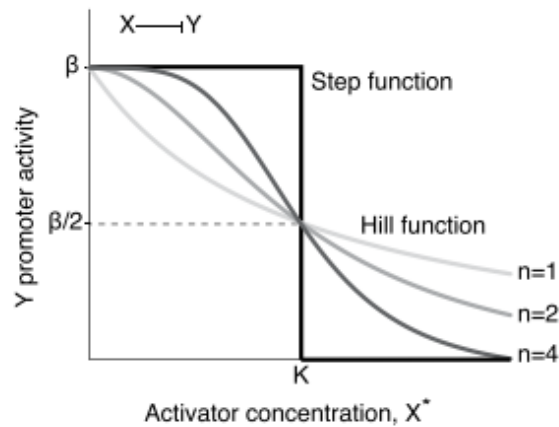
As Equações (3.2) e (3.3), representam a função de Hill para um gene ativador e repressor, respectivamente. A função de Hill tem três parâmetros, K , β e n . O parâmetro K é chamado de coeficiente de ativação e possui unidade de concentração, é basicamente uma barreira de concentração que X precisa ultrapassar para que ative a expressão. O parâmetro β , é a máxima atividade do promotor, onde vemos que ela pode ser alcançada quando temos altas concentrações do ativador, (YAGIL; YAGIL, 1971). O terceiro e último parâmetro é o coeficiente de Hill, n , ele define a inclinação da função, que diz que quanto maior o n , a função se aproxima da função degrau como vemos nas figuras 7 e 8.

Figura 7 – Gráfico da função de hill e função degrau.



Fonte: (ALON, 2019) O gráfico mostra a função de hill e a função degrau para um ativador, juntamente com a forma da função para diferentes tipos de n .

Figura 8 – Gráfico da função de hil e função degrau.



Fonte: (ALON, 2019) O gráfico mostra a função de hill e a função degrau para um repressor, juntamente com a forma da função para diferentes tipos de n .

Na figura 7 e na figura 8 conseguimos visualizar como a função de Hill se aproxima da função degrau com o aumento de n, essa é uma forma importante de conseguir visualizar e facilitar o entendimento da função de entrada, que mais adiante será considerada como uma função degrau para um melhor entendimento. Para um fator de transcrição repressor o símbolo será $X \dashv Y$, como pode ser visto na figura 8, outro fato a se notar é que a função de Hill possui um limite, ela não diverge, por mais que a concentração de X^* seja alta. Isso se dá ao fato de que o ativador e o promotor se ligam em uma relação de 1 para 1, não importando o quão alto seja a concentração de X^* .

Por simplicidade matemática, vamos substituir a função de Hill por uma função degrau, que tem o mesmo comportamento da função de Hill. Como os estados que estamos procurando são os estados LIGADO e DESLIGADO podemos usar como função de entrada as seguintes funções:

$$f(X^*) = \beta \theta(X^* > K). \quad (3.4)$$

$$f(X^*) = \beta \theta(X^* < K). \quad (3.5)$$

As Equações (3.4) e (3.5) representa a função de entrada para um ativador e um repressor, respectivamente, onde o estado ligado temos $f(X^*) = \beta$ e no desligado $f(X^*)=0$, chamaremos esse artifício de aproximação lógica da função de entrada.

3.2 Dinâmica de uma regulação simples

Vamos agora focar no estudo da dinâmica das redes de transcrição, a melhor maneira de iniciar esse assunto é focando na regulação simples, como já vimos na sessão anterior de um fator de transcrição X que regula uma proteína Y , $X \rightarrow Y$. O processo começa com um sinal exterior S_x , que por sua vez transforma nosso X em sua forma ativa X^* e esse rapidamente se liga ao promotor do gene Y . O gene Y começa a ser transcrito e após isso é transladado para a formação da proteína Y , aumentando assim, a concentração da mesma na célula a uma taxa β .

Esse processo de produção de Y é contrabalanceado por dois processos, degradação da proteína e diluição. O primeiro processo é causado pela destruição da proteína por outras proteínas dentro da célula e o segundo é causado pela diluição da mesma decorrente do aumento do volume da célula. Nossa taxa total de remoção de Y será terá a forma:

$$\alpha = \alpha_{dil} + \alpha_{deg}. \quad (3.6)$$

Sabendo os processos que ocorrem dentro da célula, podemos equacionar a variação da concentração da proteína em relação ao tempo, pegando a taxa de produção β e subtraindo pela taxa de remoção α , assim temos:

$$\frac{dY}{dt} = \beta - \alpha Y. \quad (3.7)$$

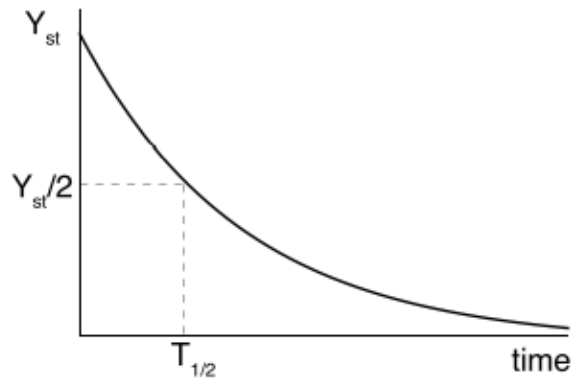
Vamos alcançar um ponto onde Y se manterá em um estado estável, chamaremos esse Y de Y_{st} , para achar esse valor usamos otimização de função e zeramos o lado esquerdo da Equação (3.7) e achamos o seguinte resultado:

$$Y_{st} = \frac{\beta}{\alpha}. \quad (3.8)$$

Esse resultado nos mostra que o valor da concentração do estado estável é diretamente proporcional a β e inversamente proporcional a α , o que faz total sentido. Como já temos o valor do estado estável que Y alcança, agora podemos dar continuidade no estudo da dinâmica fazendo uma pergunta: O que aconteceria se fosse desligada a produção de Y ? Partindo do Y_{st} que é o valor de estabilidade e o maior que Y conseguirá alcançar, por enquanto, resolvemos (3.7) fazendo $\beta = 0$, e a função que obtemos tem a forma matemática mostrada na Equação (3.9) e o gráfico vemos na figura 9:

$$Y(t) = Y_{st}e^{-(\alpha t)}. \quad (3.9)$$

Figura 9 – Gráfico da função de decaimento de Y.

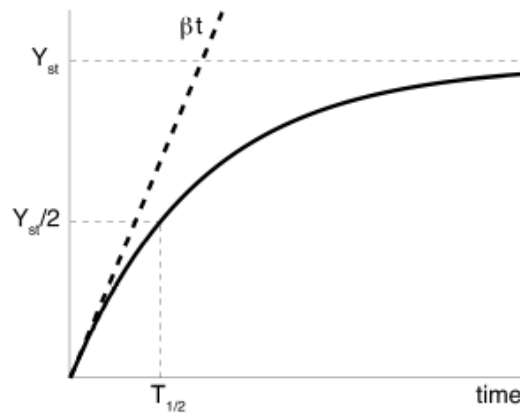


Fonte: (ALON, 2019) O gráfico mostra o que acontece com a função de expressão de Y quando a transcrição é desligada a partir do estado estável .

Seguindo uma lógica semelhante, podemos perguntar: como seria $Y(t)$ se ele fosse ligado a partir do $Y=0$ por um S_x forte o suficiente? Acharemos a resposta da mesma forma que na questão anterior, porém agora o β não é zero e precisamos resolver (3.7). A resposta pode ser descoberta por separação de variáveis e a constante que aparece pode ser definida usando a condição de contorno que $[t=0 / Y=0]$. A função $Y(t)$ que achamos e o gráfico dos níveis de expressão fica:

$$Y(t) = Y_{st}(1 - e^{-(\alpha t)}). \quad (3.10)$$

Figura 10 – Gráfico da função de nível de expressão de Y.



Fonte: (ALON, 2019) O gráfico mostra o que acontece com a função de expressão de Y quando ligamos a ativação a partir do zero .

Podemos notar no gráfico de $Y(t)$ que a concentração Y aumenta gradualmente convergindo para o valor Y_{st} , o sistema se comporta dessa maneira porque enquanto Y aumenta e ganha expressão o termo $-\alpha Y$ começa a ganhar força e diminuir a velocidade com que Y se aproxima de Y_{st} .

3.3 Network motifs

Com o que estudamos agora conseguimos entender como funciona o processo de regulação simples de uma rede de transcrição, agora vamos dar um passo maior e procurar entender como funciona uma rede de transcrição com um número muito maior de interações. As redes de transcrição são extremamente complexas, possui várias setas e interações distintas, porém, existem certos padrões que servem como blocos de construção dessas redes e podemos começar um estudo da dinâmica dessas redes, estudando a dinâmica desses padrões.

Os padrões que procuramos são chamados de network motifs, esses padrões são definidos por estatística, eles aparecem mais frequentemente na vida real do que em uma rede randômica com o mesmo número de fatores de transcrição e interações. A evolução parece ter favorecido os mesmo motifs em sistemas diferentes de diferentes organismos, sugerindo que eles são selecionados com base em suas funções biológicas, inclusive, outras redes biológicas também apresentam os network motifs, como por exemplo as redes neuronais, que em alguns são semelhantes aos motifs encontrados nas redes de transcrição, (ALON, 2007).

O padrão mais simples que podemos ter é a Autorregulação Autorregulação (AR),

essa motif é basicamente quando um fator de transcrição ativa ou reprime ele mesmo, na representação por diagrama é uma seta que sai do fator e volta para ele. Sabendo o que define um motif, vamos agora ver por meio da estatística o motivo dele ser tão importante no nosso estudo. Vamos construir uma rede completamente randômica e para isso, imaginamos um ponto de uma rede que possui um total de N pontos, esse ponto pode escolher qualquer um dos outros $N-1$ pontos, inclusive ele mesmo, para poder se ligar e ativar (reprimir), portanto, a probabilidade dele escolher ele mesmo para tal feito é:

$$P_{self} = \frac{1}{N}. \quad (3.11)$$

Quando colocamos uma quantidade A de setas (interações), de maneira randômica na nossa rede, a média de setas autorreguladoras é definida como A vezes a probabilidade da autorregulação. Tendo nosso modelo formado agora podemos comparar com uma rede real de mesma quantidade de setas e pontos, ou seja, de mesmo A e N .

Muitos estudos relacionam a arquitetura da rede genética e a sua dinâmica de comportamento, a dinâmica vai depender intrinsecamente da forma com que os genes, fatores de transcrição e interação estão distribuídos. Um parâmetro importante que vai definir isso é a conectividade, que basicamente é a razão entre a quantidade de interações e a quantidade de genes, esse número vai nos dizer o quão complexa uma rede pode ser, ele não é o bastante para definir por completo o comportamento que a rede que estamos estudando vai ter, porém, nos proporciona insights sobre a estrutura de todo o sistema regulatório, (THIEFFRY *et al.*, 1998).

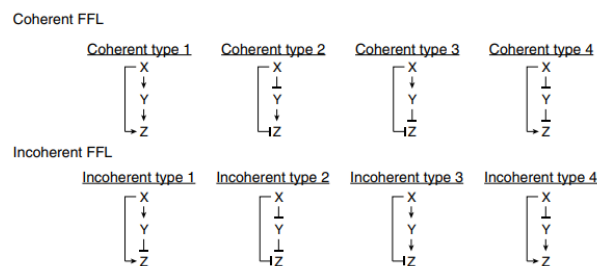
Para esclarecer a recorrência da motif AR, usamos um exemplo de um sistema com 96 genes reguladores da bactéria *Escherichia coli*, no qual 55 desses genes estão envolvidos em regulação interna e possui uma quantidade observada de interações; $A = 63$ e $N = 55$, (THIEFFRY *et al.*, 1998), estatisticamente a autorregulação teria que aparecer:

$$\langle N_{self} \rangle = \frac{A}{N} \approx 1,14. \quad (3.12)$$

A autorregulação apareceria apenas uma vez na nossa rede randômica, porém, na rede real ela aparece 45 vezes, isso é uma diferença bem grande e nos mostra que os motifs são, de fato, um objeto de grande importância no estudo das redes de transcrição, pois, para eles aparecerem muitas vezes indo contra a lógica estatística, o sistema deve ter algum ganho com a presença deles.

Sabemos a definição de motivos e vimos o exemplo mais simples que temos, agora vamos analisar um caso bem conhecido do meio, o Feed Forward Loop Feed Forward Loop (FFL). O FFL é um tipo de Regulação de gene que consiste em um fator X que regula um fator Y que por sua vez ambos regulam um fator Z. Sabendo que a regulação pode ser ativadora ou repressora, conseguimos montar 8 tipos diferentes de FFL, dentro desses 8 tipos conseguimos separar em dois grupos distintos, os coerentes e incoerentes, como veremos na figura 11, (MANGAN *et al.*, 2003).

Figura 11 – Figura com tipos de FFL, coerentes e incoerentes.

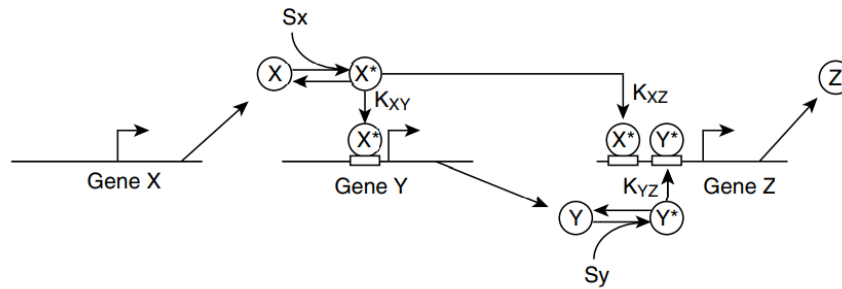


Fonte: (ALON, 2019) A figura mostra os oito possíveis arranjos do feed forward loop e separa em dois grupos, os coerentes e incoerentes.

Para definir os grupos vamos considerar o fator ativador como o sinal positivo, e um repressor como o sinal negativo, a partir disso comparamos o produto dos sinais que chegam no fator Z de forma direta $X \rightarrow Z$ e de forma indireta $X \rightarrow Y \rightarrow Z$, se nos dois caminhos os sinais forem iguais o grupo que estamos observando é o grupo coerente, se os sinais forem diferentes temos o grupo incoerente. A figura 7 mostra as oito formas que o FFL pode ter e seu devido grupo.

A dinâmica do FFL é mais complexa do que as outras que já foram citadas aqui, precisamos investigar como os reguladores X e Y são integrados ao gene Z e como é formada a função de entrada do mesmo. Para facilitar nosso entendimento vamos pegar o FFL coerente do tipo 1, a ativação do gene Z só vai acontecer se o gene X e Y forem ativados simultaneamente, ou seja, nosso FFL é um sistema lógico do tipo Porta lógica que precisa de dois ativadores para um resultado positivo na saída (AND) como na figura 12, outro tipo lógico que pode ser trabalhado é o tipo Porta lógica que precisa de um ativador para um resultado positivo na saída (OR), nesse caso o nosso FFL poderia começar a transcrição de Z somente com um dos dois ativadores, X ou Y.

Figura 12 – Figura com o diagrama de um FFL funcionando em lógica AND.



Fonte: (ALON, 2019) A figura mostra de uma forma acurada os processos de um FFL sendo usado em uma lógica AND.

Na figura 12 podemos ver que o FFL possui o gene Y dependendo diretamente do X e o gene Z dependendo diretamente de X e de Y, que dependem dos fatores externos de ativação S_x e S_y , se um dos fatores não estiver presente a transcrição de Z não é iniciada. Para o início da transcrição existem as mesmas barreiras que vimos em exemplos anteriores, por exemplo, K_{xy} é a concentração necessária que X precisa alcançar para que comece a ativar o gene Y, a mesma lógica vale para K_{xz} e K_{yz} . O gene Y começa a aumentar sua concentração quando X^* excede a barreira K_{xy} e a taxa de produção de Y é β_y , o mesmo vale para Z que por sua vez depende tanto de K_{yz} quanto de K_{xz} , assim temos a taxa de produção vezes a nossa conhecida função degrau:

$$Y = \beta_y \theta(X^* > K_{xy}). \quad (3.13)$$

$$Z = \beta_z \theta(X^* > K_{xz}) \theta(Y^* > K_{yz}). \quad (3.14)$$

A equação da dinâmica agora vai depender de dois termos, o de produção e o de destruição e ficam:

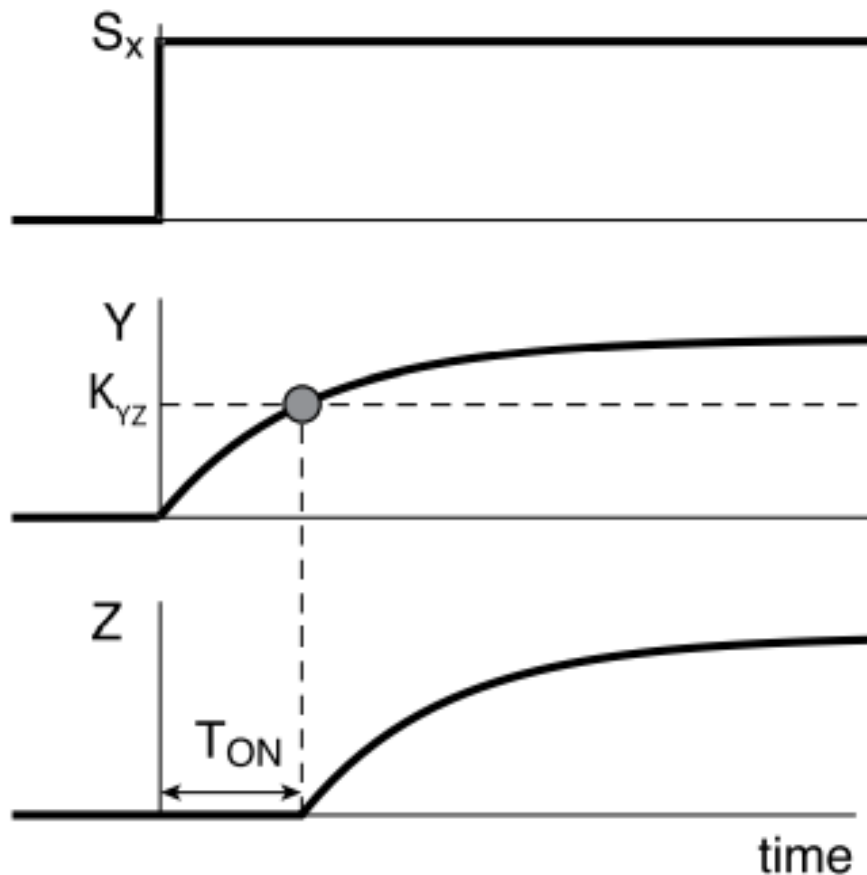
$$\frac{dY}{dt} = \beta_y \theta(X^* > K_{xz}) - \alpha_y Y. \quad (3.15)$$

$$\frac{dY}{dt} = \beta_y \theta(X^* > K_{yz}) - \alpha_y Y. \quad (3.16)$$

Com as Equações (3.15) e (3.16) agora podemos fazer uma análise qualitativa do FFL coerente do tipo 1 (C1-FFL). Para começar a análise vamos considerar que o S_y está sempre

ativo ($Y=Y^*$) e que o S_x é ligado instantaneamente e possui só dois estados, ON e OFF. Após S_x entrar no estado ON, o fator Y começa a ser transcrito e ele precisa chegar a um nível de concentração que ultrapasse a barreira K_{yz} para começar a transcrever Z, isso nos dá um atraso na transcrição, com será observado na figura 13.

Figura 13 – Gráfico que mostra a função da concentração de Y e Z em relação ao tempo.



Fonte: (ALON, 2019) O gráfico mostra o delay presente na transcrição de Z e o valor de Y que Z começa a ser transcrito.

No gráfico da figura 13 conseguimos ver a concentração de Y em relação ao tempo e como ele converge assintoticamente para o estado estável, como foi mostrado nas Equações (3.10) e (3.8). O tempo de atraso, T_{ON} , pode ser encontrado usando a Equação (3.10) e fica:

$$Y^*(T_{ON}) = Y_{st}(1 - e^{-\alpha_y T_{ON}}) = K_{yz}. \quad (3.17)$$

A Equação (3.17) pode ser resolvida algebricamente isolando o T_{ON} e então temos:

$$T_{ON} = \frac{1}{\alpha_y} \ln\left(\frac{1}{1 - \frac{K_{yz}}{Y_{st}}}\right). \quad (3.18)$$

Partindo da Equação (3.18) vemos que esse atraso depende de alguns parâmetros bioquímicos da proteína Y, que é a taxa de degradação de Y e a razão da barreira de ativação com o estado estável de Y. Outra observação que podemos fazer com essa Equação (3.18) é que em um limite onde $\frac{K_{yz}}{Y_{st}}$ tende a 1, essa equação diverge, isso se dá ao fato de que o fator de barreira não pode ser maior que o estado estável que é o “máximo” que nossa proteína consegue alcançar de concentração. Continuando na questão desse atraso que a transcrição de Z tem ao iniciar a partir de um estado OFF, o inverso não acontece, ou seja, se temos um sistema onde Z está sendo transcrito e subitamente S_x ou S_y são desligados, o nível de concentração de Z começa a cair sem nenhum atraso, Basta observar o diagrama exposto na figura 12 e analisar a Equação (3.16) que a função degrau zera e ficamos somente com o fator de degradação alfa.

O estudo da dinâmica desse FFL nos mostrou algo bem interessante e que pode ser de grande valia, o atraso sensível ao sinal que ele possui. O atraso sensível ao sinal é mostrado com mais profundidade e com genes observados da E. coli, (MANGAN; ALON, 2003). Nós observamos que o C1-FFL não reage de forma imediata na transcrição de Z com a ativação de S_x , na realidade para que comece o processo de transcrição, o sinal (S_x) tem que ser maior que o tempo de atraso, T_{ON} , e para encerrar o processo de transcrição o atraso é nulo, ou seja, após o sinal externo ser desligado, a transcrição para. Esse é um excelente mecanismo de defesa que uma célula pode ter contra algum estímulo do ambiente que seja passageiro e que não necessite de uma reação súbita e, as vezes, danosa para a própria célula. Uma deficiência nesse sistema pode levar a grandes problemas na célula, observar e prevenir erros na hora de iniciar ou parar o processo de transcrição de alguma proteína, pode ser uma peça chave para entender algumas doenças que são afetadas por fatores genéticos, tais como, as doenças autoimunes, (BL, 2001).

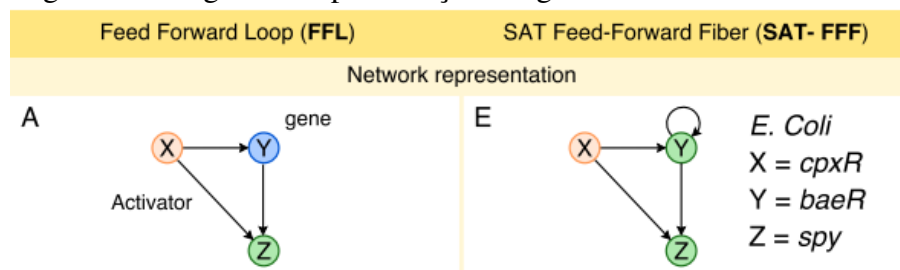
4 MOTIFS COMO BASE PARA A BIOLOGIA COMPUTACIONAL

O Corpo humano é sistema de alta complexidade que até agora não foi completamente desvendado. Relacionar o funcionamento do corpo humano com o funcionamento de computadores não é algo fora do comum, pois ambos são sistemas complexos que são programados para realizar variadas tarefas. Os computadores precisam de vários pequenos sistemas de informação para funcionar, como por exemplo, um relógio. Esses pequenos sistemas também são encontrados dentro dos corpos vivos, como veremos adiante.

4.1 Feed Forward Fiber e a simetria de fibração

Até o momento, discutimos desde a formação básica da célula até a dinâmica que ela possui na reação aos estímulos do ambiente, sabemos da complexidade do corpo celular e a necessidade que ele tem de guardar e transmitir informações, ele funciona de forma bastante semelhante a um computador e para funcionar dessa forma ele precisa das ferramentas necessárias para realizar esse trabalho. O FFL e a AR não são as únicas motifs que achamos 32 nas células, elas foram elucidadas nesse trabalho para facilitar o entendimento de outra motif que também aparece com bastante frequência nas redes de transcrição, a FEED FOWARD FIBER (FFF), (LEIFER *et al.*, 2020). Essa motif que frequentemente aparece nas redes é basicamente um FFL com uma AR, positiva ou negativa, no gene Y como podemos ver na figura 14.

Figura 14 – Figura da representação em grafos de um FFL e um FFF.

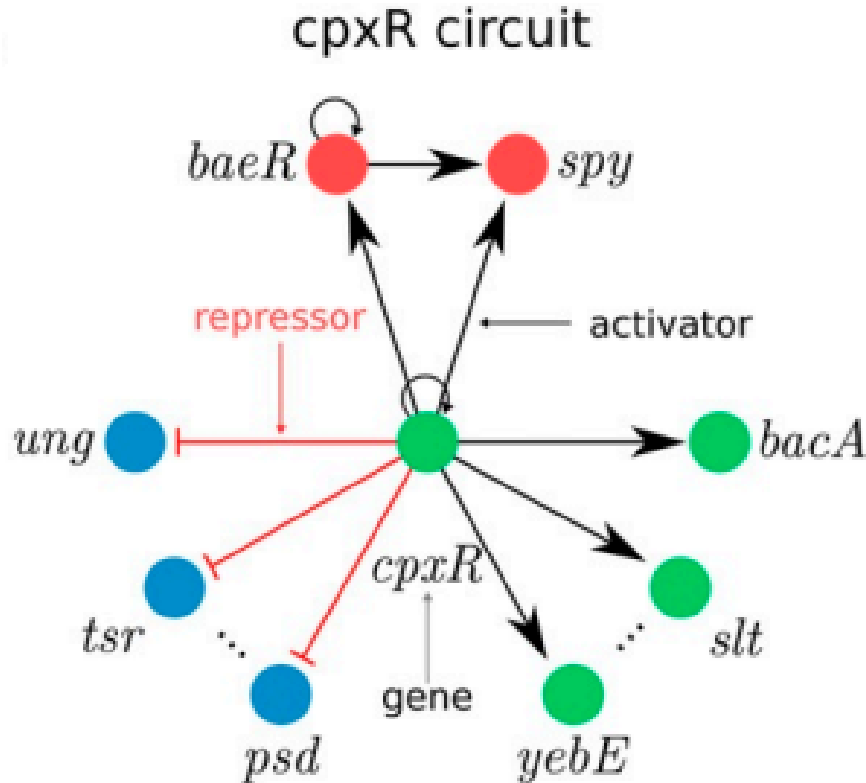


Fonte: (ALON, 2019) A figura mostra os grafos de um FFL e um SAT-FFF e um exemplo real da bactéria *E.colli*

A FFF tem uma característica singular que a FFL não tem, em um limite onde o tempo $\rightarrow \infty$ os níveis de expressão tendem a se igualarem, chamamos isso de sincronização. O FFL além de não sincronizar ele não tem nenhum tipo de oscilação nos níveis de expressão o que é coerente com o uso dele para o atraso sensível de sinal. De acordo com Morone, Leifer e Makse (2020), o FFF sincroniza por causa de uma característica na geometria da árvore de

entrada de cada gene, um exemplo podemos ver na figura 15.

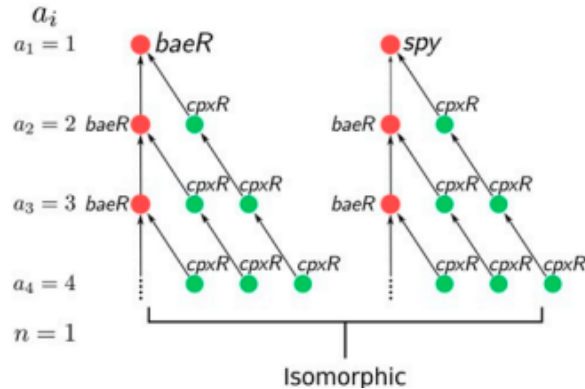
Figura 15 – Uma rede de transcrição genética da bactéria *Escherichia coli*.



Fonte: (MORONE *et al.*, 2020) A figura um exemplo real de uma rede de transcrição da bactéria *E.colli* pode ser vista com suas AR's e formando um FFF

A árvore de entrada de cada gene é formada colocando no topo o gene desejado e abaixo dele todos os outros genes que o regulam, e suas devidas regulações, a partir daí fazemos o mesmo para cada gene dessa segunda camada, e assim por diante. Quando fazemos isso para um segundo gene da mesma rede de regulação e esse segundo gene possui a mesma árvore de entrada do primeiro gene, acontece o que chamamos de isomorfismo e temos a característica geométrica necessária para que os níveis de expressão do gene 1 e gene 2 se sincronizem com o passar do tempo, essa característica se chama simetria de fibração, (MORONE *et al.*, 2020). O exemplo de um isomorfismo veremos na figura 16.

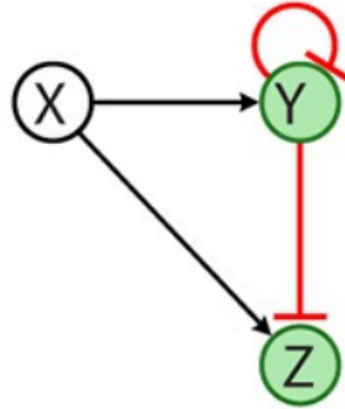
Figura 16 – Figura que mostra o isomorfismo em genes diferentes.



Fonte: (MORONE *et al.*, 2020) A figura um exemplo real de uma rede de transcrição da bactéria *E.colli* pode ser visto como se forma o isomorfismo e a simetria que temos na árvore de entrada.

Possuindo a ferramenta da sincronização, agora podemos nos aprofundar mais no estudo das FFF, vamos discutir agora dois tipos de FFF, SAT-FFF e o UNSAT-FFF. O SAT-FFF é o FFF que foi mostrado na figura 15, ele possui uma AR positiva e uma regulação ativadora de $Y \rightarrow Z$, os dois genes vão sincronizar e tendem a ter os níveis de expressão convergindo para um ponto fixo. O UNSAT-FFF, mostrado na figura 17 é formado Com a mesma disposição de genes que o SAT possui, porém a autorregulação no gene Y e a regulação de $Y \rightarrow Z$ são repressoras, essa formatação faz com que a solução dos níveis de expressão não só convirjam, mas também, oscilem, (ATKINSON *et al.*, 2003). Esse estado de oscilação que o UNSAT-FFF alcança é basicamente uma forma de um fenômeno que ocorre em sistemas complexos, conhecido como frustration, ele acontece quando um sistema não consegue chegar a uma configuração fixa e está sempre sob uma tensão que o faz oscilar indefinidamente. Uma boa explicação desse fenômeno pode ser encontrada nos estudos de vidros de spin, (ANDERSON, 1978).

Figura 17 – Figura que mostra a representação em grafo do UNSAT-FFF.



Fonte: (MORONE *et al.*, 2020) A figura mostra a representação em grafo do UNSAT-FFF com a AR negativa no gene Y e uma regulação repressora de Y para Z.

4.2 A dinâmica do UNSAT-FFF

Na literatura, o estudo da dinâmica do UNSAT-FFF mostra que os níveis de expressão dos genes envolvidos oscilam, (LEIFER *et al.*, 2020). Nessa análise que vamos fazer, descobriremos a causa das oscilações. O primeiro passo é montar as equações de expressão de acordo com o arranjo da motif escolhida, nesse caso o UNSAT-FFF. Diferente de uma autorregulação ativadora, a lógica da função degrau que vamos usar para montar nossa equação de expressão é a da Equação (3.5), ou seja, para começar a reprimir a transcrição do gene a concentração tem que ultrapassar a barreira K , assim fica:

$$\frac{dY}{dt} = \beta_x \beta_y \theta(x_t - k_x) \theta(k_y - y_t) - \alpha y_t. \quad (4.1)$$

$$\frac{dZ}{dt} = \beta_x \beta_z \theta(x_t - k_x) \theta(k_z - z_t) - \alpha z_t. \quad (4.2)$$

Partindo da premissa que S_x sempre está ativo e X^* é sempre maior que K_x , eliminamos a função degrau de X que é sempre 1 e ficamos com uma equação mais enxuta e podemos reescalonar algumas variáveis que ficam na forma:

$$\psi(t) = \frac{y(t)}{k_y}; \zeta(t) = \frac{z(t)}{k_y}; \frac{\beta_x \beta_y}{k_y} = \delta. \quad (4.3)$$

Reescalando as Equações (4.1) e (4.2) com os parâmetros mostrados em (4.3), e usando o artifício de que X é constante e maior que k_x , ficamos com:

$$\frac{d\psi}{dt} = \delta\theta(1 - \psi(t)) - \alpha\psi(t). \quad (4.4)$$

$$\frac{d\zeta}{dt} = \delta\theta(1 - \psi(t)) - \alpha\zeta(t). \quad (4.5)$$

Partindo do princípio de que Z e Y possuem simetria de fibração, logo, eles vão sincronizar e só vamos precisar resolver uma equação, ficaremos com a (4.4). A grande questão a se notar é que uma equação diferencial ordinária não pode ter soluções oscilatórias, então a pergunta que fica é: como vamos achar oscilações na solução das equações de concentração? A resposta pra essa pergunta está no atraso que a presença de uma AR repressora causa na rede de transcrição e também no atraso que esses “circuitos biológicos” sofrem na transição de processos, tais como, transcrição, translação, etc. com esse atraso a nossa equação diferencial ordinária Equação diferencial ordinária (EDO) se transforma numa equação diferencial com delay Equação diferencial com delay (EDD) então a Equação (4.4) fica da forma:

$$\frac{d\psi}{dt} = -\alpha\psi(t) + \delta\theta(1 - \psi(t - \tau)). \quad (4.6)$$

Onde o τ representa esse delay na equação. Quando temos uma EDD ela se torna um problema de infinitas dimensões, pois, não temos condições iniciais do tipo que temos um ponto e o valor da função naquele ponto, nos teremos uma função para um pequeno intervalo, transformando assim, nosso problema em um problema de infinitas dimensões. Essa EDD será resolvida de uma forma chamada, método dos passos, basicamente a função que procuramos será “desenhada” com a soma de várias pequenas funções definidas em pequenos e contíguos intervalos no domínio, ou seja, uma função $\{f_0(t)\}$ possui um domínio $[-\tau, 0]$, que está adjacente a função $\{f_1(t)\}$ que possui o domínio $[0, \tau]$ e assim sucessivamente. Para aplicar esse método primeiro definiremos uma função inicial que chamaremos de $\{f_0(t)\}$ e que está contida no intervalo $[-\tau, 0]$, essa função inicial é o valor inicial que queremos que comece nosso $\psi(t)$, dada

nossas condições iniciais agora vamos definir o próximo intervalo e observar como fica nossa função, em $[0, \tau]$ nossa Equação (4.6) fica:

$$\frac{d\psi}{dt} = -\alpha\psi(t) + \delta\theta(1 - f_0(t - \tau)). \quad (4.7)$$

Multiplicando ambos os lados de (4.7) por $e^{\alpha t}$ e reorganizando a equação com a regra do produto, teremos:

$$e^{\alpha t} \frac{d(\psi e^{\alpha t})}{dt} = \delta\theta(1 - f_0(t - \tau))e^{\alpha t}. \quad (4.8)$$

Integramos os dois lados da Equação (4.8) com os limites de integração variando de 0 a t:

$$\psi e^{\alpha t} - \psi(0) = \delta \int_0^t e^{\alpha t'} \theta(1 - f_0(t' - \tau)) dt'. \quad (4.9)$$

Passamos o $\psi(0)$ para a direita e dividimos a equação por $e^{\alpha t}$ e assim temos:

$$\psi(t) = \psi(0)e^{-\alpha t} + \delta \int_0^t e^{\alpha(t'-t)} \theta(1 - f_0(t' - \tau)) dt'. \quad (4.10)$$

Com a equação 4.10 armada vamos partir para análise dos parâmetros, primeiro consideramos ψ contínua, então no valor inicial $\psi(t = 0) = f_0(t = 0)$, que é onde a função procurada começa, o segundo passo é notar que $\psi(t)$ no intervalo em que procuramos $[0, \tau]$ é justamente o $f_1(t)$. A equação 4.10 com esses ajustes toma a forma:

$$f_1(t) = f_0(0)e^{-\alpha t} + \delta \int_0^t e^{\alpha(t'-t)} \theta(1 - f_0(t' - \tau)) dt'. \quad (4.11)$$

Observamos com a equação (4.11) que a forma da função seguinte tem relação com a forma da função anterior a ela, e a partir disso conseguimos achar a fórmula geral para achar a solução de ψ a partir do intervalo generalizado $[k\tau, (k+1)\tau]$ assumindo que a solução do intervalo anterior $[(k-1)\tau, k\tau]$ é $f_{k-1}(t)$. Fazendo essas alterações nossa Equação (4.6) fica:

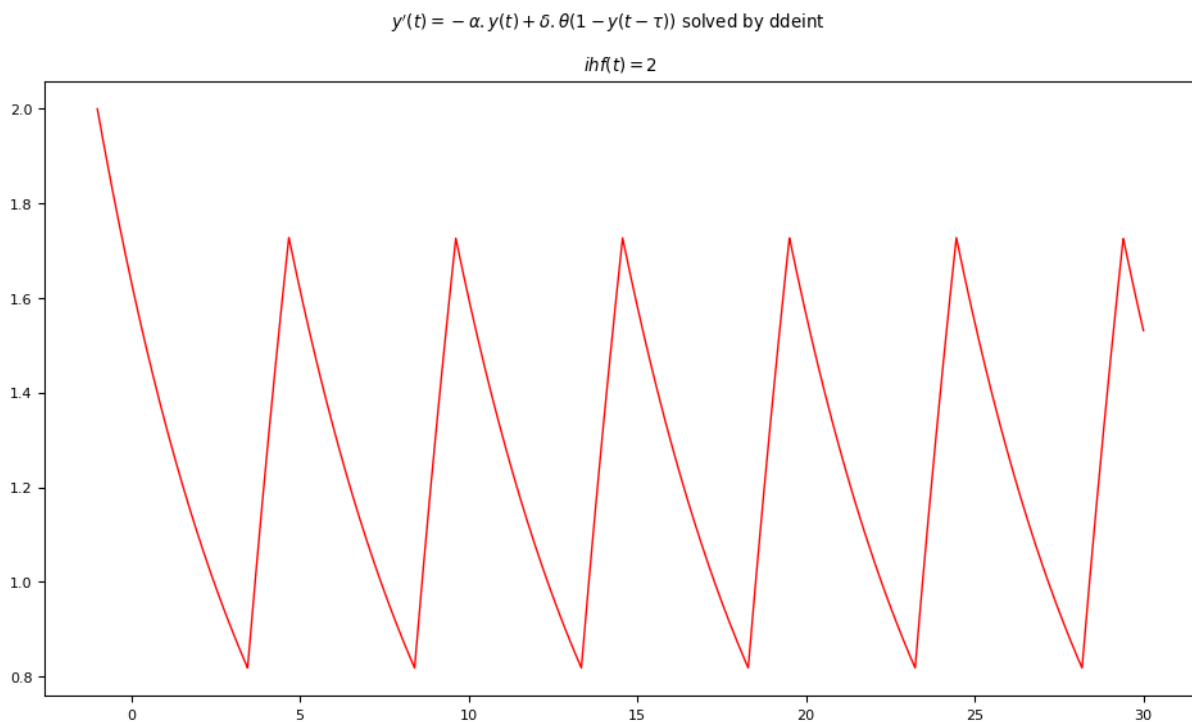
$$\frac{d\psi}{dt} = -\alpha\psi(t) + \delta\theta(1 - f_{k-1}(t - \tau)). \quad (4.12)$$

O mesmo procedimento que aplicamos de (4.6) até (4.11) aplicaremos em (4.12) e assim temos:

$$\psi(t) = \psi(k\tau)e^{\alpha(k\tau-t)} + \delta \int_0^t e^{\alpha(t'-t)} \theta(1 - f_{k-1}(t' - \tau)) dt'. \quad (4.13)$$

Finalmente com a equação 4.13 conseguimos achar a solução de qualquer EDD de mesma forma e para cada intervalo somamos todas as funções e chegamos na função final procurada. A oscilação da função vai aparecer quando definirmos um intervalo de t e fixamos os valores de f_0 , α , δ , τ e calculamos numericamente as funções para cada intervalo. Para a visualização da função de expressão necessitamos de um software que calcule numericamente os valores a partir de parâmetros iniciais. Somente com a equação (4.6), conseguimos achar as oscilações se colocarmos ela para ser resolvida numericamente em algum programa. Usando a linguagem de programação Python e definindo as variáveis, achamos as seguintes oscilações dos níveis de expressão que são demonstrados na figura 18.

Figura 18 – Gráfico que mostra a função da concentração de ψ em relação ao tempo.



Fonte: produzida pelo autor. O gráfico mostra as oscilações que o UNSAT-FFF possui para tempo grande e variáveis definidas.

O eixo Y do gráfico na figura 18, representa os níveis de expressão de ψ , enquanto o eixo X representa o tempo decorrido. Os parâmetros foram escolhidos de tal forma que fossem iguais aos usados no material suplementar do artigo, LEIFER et al., 2020, página 14,

confirmando a validade do código usado nesse trabalho. Os parâmetros foram $f_0 = 2$, $\alpha = 0.2$, $\delta = 1$ e $\tau = 1$. Sendo que f_0 é o nível de expressão inicial e o restante dos parâmetros são frutos do reescalonamento feito na Equação (4.3). O que vimos até então, foi um ramo do estudo da genética, onde estudamos os níveis de expressão de genes com foco em redes (redes de transcrição), e as equações dinâmicas que definem o nível de expressão das proteínas e fatores de transcrição envolvidos nas redes. Essa abordagem, por mais que seja uma abordagem bastante simplificada não só pelas networks motifs, mas também pela relação 1 pra 1 que fazemos entre proteínas e mRNA's, consegue jogar uma luz em um sistema extremamente complexo que é o sistema de interações intracelulares.

5 CONCLUSÕES E TRABALHOS FUTUROS

Vimos que a autorregulação repressora transforma nosso sistema de regulação em basicamente um relógio sincronizado, que é uma peça fundamental em qualquer tipo de dinâmica computacional. A associação de uma rede de transcrição com um computador é de grande importância, pois, muitas funcionalidades do corpo humano são produzidas em ciclos definidos e tem funcionamento periódico, que podem está completamente ligados a esse fenômeno que vimos nas redes de transcrição. Uma das conhecidas funcionalidades é o ciclo circadiano, que basicamente regula as funções biológicas de diversos seres vivos que se repete regularmente com período de aproximadamente 24 horas, entre essas funções podemos destacar a produção de dois hormônios que são estimuladas em duas horas distintas do dia, a melatonina que é um hormônio que regula e estimula nosso sono e começa sua produção bem próximo da hora de dormir e o cortisol que é o hormônio mais conhecido como o hormônio do estresse, e sua maior produção se concentra na parte da manhã ajudando assim a manter nosso corpo com suas perfeitas funcionalidades atuando os dois de forma cíclica.

Com essa análise que fizemos de redes de transcrição, conseguimos mostrar que os corpos celulares funcionam como um computador, guardando informação e reagindo a estímulos para produção de algum objetivo. Os exemplos que ficaram mais claros foram o mecanismo de defesa que o FFL pode proporcionar por meio do seu atraso sensível ao sinal, a sincronização de níveis de expressão dos genes que o FFF produz, e a oscilação periódica do UNSAT-FFF, transformando a rede de transcrição em um relógio, que é uma das bases necessárias para o funcionamento de um computador.

A tecnologia ainda limita um estudo mais caprichado das interações intracelulares, muitas simplificações foram feitas para se ter uma ferramenta matemática que demonstre as interações, sabemos que existem diversos fatores de transcrição interagindo dentro de uma rede de forma que não conseguimos mensurar e por na equação, porém, o estudo dos motifs nos dá uma base sólida para entender as principais características das redes de transcrição. Com o aparecimento dos miRNA's percebemos que a tecnologia precisa ir de alguma forma para o patamar de observar os níveis de expressão das próprias proteínas e não somente dos mRNA's. Com um aparato tecnológico mais refinado, não só para observação, mas também para tratamento de dados, podemos cada vez mais nos aproximar de um pleno conhecimento das interações que ditam o rumo da nossa vida.

REFERÊNCIAS

- ALBERTS, B.; BRAY, D.; HOPKIN, K.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Fundamentos da Biologia Celular-4**. [S. l.]: Artmed Editora, 2017.
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; MORGAN, D.; RAFF, M.; ROBERTS, K.; WALTER, P.; WILSON, J.; HUNT, T. **Biologia molecular da célula**. [S. l.]: Artmed Editora, 2010.
- ALON, U. Network motifs: theory and experimental approaches. **Nature Reviews Genetics**, Nature Publishing Group UK London, v. 8, n. 6, p. 450–461, 2007.
- ALON, U. **An introduction to systems biology: design principles of biological circuits**. [S. l.]: CRC press, 2019.
- ANDERSON, P. The concept of frustration in spin glasses. **Journal of the Less Common Metals**, Elsevier, v. 62, p. 291–294, 1978.
- ATKINSON, M. R.; SAVAGEAU, M. A.; MYERS, J. T.; NINFA, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in escherichia coli. **Cell**, Elsevier, v. 113, n. 5, p. 597–607, 2003.
- BL, M. P. K. J. K. Autoimmune disease: why and where it occurs. **Nat Med**, v. 7, n. 8, p. 899–905, 2001.
- BROWN, P. O.; BOTSTEIN, D. Exploring the new world of the genome with dna microarrays. **Nature genetics**, Nature Publishing Group, v. 21, n. 1, p. 33–37, 1999.
- HOOKE, R. **Micrographia:...** [S. l.]: Holzer, 1968.
- LAU, N. C.; LIM, L. P.; WEINSTEIN, E. G.; BARTEL, D. P. An abundant class of tiny rnas with probable regulatory roles in caenorhabditis elegans. **Science**, American Association for the Advancement of Science, v. 294, n. 5543, p. 858–862, 2001.
- LEEUWENHOEK, A. V. **Arcana naturae detecta ab Antonio van Leeuwenhoek**. [S. l.]: apud Henricum a Krooneveld, 1966.
- LEIFER, I.; MORONE, F.; REIS, S. D.; JR, J. S. A.; SIGMAN, M.; MAKSE, H. A. Circuits with broken fibration symmetries perform core logic computations in biological networks. **PLoS computational biology**, Public Library of Science San Francisco, CA USA, v. 16, n. 6, p. e1007776, 2020.
- MANGAN, S.; ALON, U. Structure and function of the feed-forward loop network motif. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 100, n. 21, p. 11980–11985, 2003.
- MANGAN, S.; ZASLAVER, A.; ALON, U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. **Journal of molecular biology**, Elsevier, v. 334, n. 2, p. 197–204, 2003.
- MORONE, F.; LEIFER, I.; MAKSE, H. A. Fibration symmetries uncover the building blocks of biological networks. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 117, n. 15, p. 8306–8314, 2020.

PFAFFL, M. A new mathematical model for relative quantification in real-time rt-pcr nucleic acids res 29: e45. **Find this article online**, 2001.

PIRES, J. G. On the applicability of computational intelligence in transcription network modelling. **Faculty of Applied Physics and Mathematics. Gdansk University of Technology, Gdansk, Poland, 21th September, 2012.**

SCHLEIDEN, M. J. **Contributions to our Knowledge of Phytogenesis.** [*S. l.: s. n.*], 1838.

SCHWANN, T. Mikroskopische untersuchungen über die uebereinstimmung in der struktur und dem wachsthum der thiere und pflanzen. (**No Title**), 1838.

THIEFFRY, D.; HUERTA, A. M.; PÉREZ-RUEDA, E.; COLLADO-VIDES, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in escherichia coli. **Bioessays**, Wiley Online Library, v. 20, n. 5, p. 433–440, 1998.

WANG, Z.; GERSTEIN, M.; SNYDER, M. Rna-seq: a revolutionary tool for transcriptomics. **Nature reviews genetics**, Nature Publishing Group UK London, v. 10, n. 1, p. 57–63, 2009.

YAGIL, G.; YAGIL, E. On the relation between effector concentration and the rate of induced enzyme synthesis. **Biophysical journal**, Elsevier, v. 11, n. 1, p. 11–27, 1971.

ZASLAVER, A.; BREN, A.; RONEN, M.; ITZKOVITZ, S.; KIKOIN, I.; SHAVIT, S.; LIEBERMEISTER, W.; SURETTE, M. G.; ALON, U. A comprehensive library of fluorescent transcriptional reporters for escherichia coli. **Nature methods**, Nature Publishing Group US New York, v. 3, n. 8, p. 623–628, 2006.