Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 8, n. 1, 2021.

Trabalho apresentado no XL CNMAC, Evento Virtual - Co-organizado pela Universidade do Mato Grosso do Sul (UFMS).

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Análise das propriedades físico-químicas dos aminoácidos por meio das distâncias de Hamming associadas ao rotulamento C do código genético

Roberta Siqueira Fernandes[1]
Discente do Programa de Pós-Graduação em Estatística Aplicada e Biometria, UNIFAL, Alfenas-MG
Anderson José de Oliveira[2]
Professor Adjunto do curso Matemática-Licenciatura, Instituto de Ciências Exatas, UNIFAL, Alfenas-MG

Resumo. Estudos realizados a partir da segunda metade do século XX, tornaram a natureza do código genético conhecida. Desse modo, a modelagem algébrica envolvendo o código genético tem sido alvo de diversas pesquisas, tendo como principais objetivos a identificação das propriedades, características e implicações do modelo a ser estudado. Por meio do mapeamento das bases nitrogenadas adenina, citosina, guanina, timina/uracila, $\{A, C, G, T/U\}$ com o anel $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, é possível obter 24 permutações, as quais podem ser divididas em 3 rotulamentos (A, B e C), de acordo com as características geométricas dos mesmos. O objetivo do presente trabalho é apresentar uma análise das diferenças e semelhanças físico-químicas dos aminoácidos, através da caracterização biológica acerca da construção do Diagrama de Hasse e dos cálculos das distâncias de Hamming entre os códons, relacionados a uma permutação do rotulamento C, os quais poderão ser utilizados no processo de análise de fenômenos mutacionais.

Palavras-chave. Rotulamento C, Diagrama de Hasse, Modelagem, Álgebra, Código Genético

1 Introdução

Há tempos que diversos pesquisadores buscam representar o que mundo biológico realiza, a fim de identificar suas características, propriedades, funções e implicações do modelo estudado. Além disso, a modelagem matemática associada ao código genético é uma área de pesquisa que está em constante expansão e aprimoramento de técnicas, pois vários autores procuram estruturar e modelar o código genético, com o intuito de validar suas hipóteses.

Os códons do código genético são formados por uma trinca de bases nitrogenadas com a possibilidade de 64 combinações distintas que são consideradas por [8] em seu estudo. Nesse trabalho, é apresentado um modelo por meio da construção da tabela do código genético. A construção é feita utilizando uma bijeção do alfabeto biológico com o alfabeto matemático $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{00, 01, 10, 11\}$. Desta forma, o seguinte mapeamento é proposto: $G - 00$, $U - 10$, $A - 01$ e $C - 11$, onde são definidas operações no conjunto das quatro bases do DNA, onde G representa guanina, U uracila, A adenina e C citosina. Vale ressaltar que o procedimento realizado por [8] leva em consideração a importância biológica das bases nitrogenadas, fundamentada pela ocorrência de erros encontrada nos códons, analisada na Biologia.

Segundo [9], além do código genético representar uma extensão do alfabeto de quatro letras das bases do DNA e RNA e estabelecer o emparelhamento entre as bases, devido as ligações de

[1]robertaf_mat@hotmail.com.
[2]anderson.oliveira@unifal-mg.edu.br.

hidrogênio que ocorre entre elas, foi observada uma associação entre códons com U na posição da segunda base e a hidrofobia dos aminoácidos. Já os códons com A na segunda posição da base codificam aminoácidos hidrofílicos ou polares. Acredita-se que a ordem dos códons deva refletir nas propriedades físico-químicas dos aminoácidos. Partindo da ordem das quatro bases do DNA na rede booleana, [9] em seu estudo propõem a utilização de uma álgebra de Lie do código genético sobre o corpo de Galois das quatro bases do DNA, permitindo compreender melhor a lógica subjacente ao código genético. Concluiu-se então, que a atribuição dos códons e as propriedades físico-químicas dos aminoácidos não estão conectadas ao acaso, ou seja, a origem do código genético não foi aleatória. Além disso, o modelo proposto por [9] pode ajudar na compreensão de eventos mutacionais nos processos de evolução molecular.

Uma outra proposta é a de [5], na qual é apresenta a modelagem algébrica do código genético, com o objetivo de identificar suas propriedades, características e implicações do modelo. É apresentada uma representação do código genético por meio de estruturas algébricas, que buscam formas de explicar fenômenos biológicos. São apresentadas as estruturas dos diagramas de Hasse, reticulados booleanos, estruturas de grupos, corpos, anéis e extensões de Galois. Os reticulados booleanos, bem como sua representação através do diagrama de Hasse, podem ser ferramentas eficazes na análise de algumas propriedades associadas ao código genético.

Em [3], é apresentada a existência de códigos corretores de erros e protocolos de comunicação em sequências de DNA, usando para isso estruturas matemáticas e em [6], é apresentado um modelo de comunicação biológico de importação de proteínas mitocondriais, que se baseia em um sistema de comunicação padrão, tendo como objetivo identificar estruturas matemáticas associadas às sequências de DNA. Através de uma analogia do alfabeto que está relacionado com o conjunto de nucleotídeos $N = \{A, C, G, T/U\}$ (A = Adenina, C = Citosina, G = Guanina e T/U - Timina/Uracila) e o alfabeto 4-ário na estrutura de anel, denotado por $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ e sabendo que é desconhecido o mapeamento entre $N \leftrightarrow \mathbb{Z}_4$, a sequência de DNA será rotulada conforme as 24 permutações entre $N \leftrightarrow \mathbb{Z}_4$. Pode-se organizar esse mapeamento em três conjuntos, denominados rotulamentos A, B e C, contendo 8 permutações cada, de acordo com a caracterização geométrica de cada um, que produzem um diferente nível de não-linearidade para as sequências. Além disso, a diferença de cada rotulamento é a associação de complementaridade dos nucleotídeos.

Por meio deste trabalho, tem-se como propósito a apresentação da construção do Diagrama de Hasse para a permutação 2031 do rotulamento C, a caracterização biológica acerca dessa construção e uma análise das diferenças e semelhanças físico-químicas dos aminoácidos acerca dos cálculos das distâncias de Hamming entre os códons, os quais poderão ser utilizados no processo de análise de fenômenos mutacionais.

2 Revisão de Conceitos

Nesta Seção serão apresentados os principais conceitos teóricos utilizados neste trabalho. As referências utilizadas foram [1], [2] e [4].

2.1 Elementos de Biologia

A célula é a menor unidade morfofisiológica de um ser vivo. Ela foi descoberta por Robert Hooke (1635-1703), no ano de 1665, ao examinar um pedaço de cortiça em seu microscópio. Os nucleotídeos são unidades de moléculas de um ácido nucleico que estão presentes nas células e que estão presentes no metabolismo como transportador de energia. Em todas as células vivas se encontram os ácidos nucleicos, que têm a função de conter, transmitir e traduzir informações genéticas dos seres vivos. Todos os organismos vivos apresentam ácidos nucleicos na forma de ácido desoxirribonucleico (DNA) e ácido ribonucleico (RNA).

DOI: 10.5540/03.2021.008.01.0363

O DNA é um polímero, ou seja, uma longa cadeia de nucleotídeos. Sua função é armazenar informação genética dos seres vivos. Conforme [1], o DNA é formado por uma dupla hélice com giro para a direita e parece uma escada de caracol molecular. Cada nucleotídeo do DNA é constituído por um grupo fosfato, uma pentose (desoxirribose) e uma dentre quatro bases nitrogenadas: adenina (A), guanina (G), citosina (C) e timina (T), que estão interligadas por pontes de hidrogênio. Já o RNA é um polímero de fita simples e possui uma pentose (ribose) em seus nucleotídeos em vez da desoxirribose encontrada no DNA. O RNA possui a uracila (U) no lugar da timina (T) em sua composição. O DNA e o RNA são os responsáveis pela síntese proteica. Para que esse processo ocorra, é necessário que a célula transcreva e traduza a informação do código genético.

As proteínas apresentam-se sob inúmeras formas e tamanhos e são as macromoléculas de maior abundância nos seres vivos. Elas estão presentes em todas as células e desempenham funções vitais dos seres vivos. As proteínas são polímeros constituídos basicamente de aminoácidos.

Cada molécula de DNA se diferencia pela sequência de bases nitrogenadas que elas apresentam, e essa sequência é que vai formar o código genético, ou seja, essa sequência de bases do DNA se relaciona com a sequência correspondente de 20 aminoácidos para a formação das proteínas. Então, essa correspondência permite identificar o aminoácido específico de cada códon. Esses códons são formados por uma trinca de bases nitrogenadas com 64 combinações possíveis. As bases são formadas pela adenina, citosina, guanina e timina/uracila, que são representadas pelas letras A, C, G e T/U respectivamente, e representam o alfabeto do DNA.

2.2 Elementos de Álgebra

Definição 2.1. *Um conjunto não vazio G e uma operação $(x, y) \mapsto x * y$ sobre G é chamado grupo se essa operação estiver sujeita às seguintes propriedades: Associativa: $(a * b) * c = a * (b * c), \forall\ a, b, c \in G$; Existência de Elemento Neutro: $\exists$ um elemento $e \in G$ tal que $a * e = e * a = a$, $\forall\ a \in G$; Elemento Simétrico: $\forall\ a \in G$, $\exists$ um elemento $a^{'} \in G$ tal que $a * a^{'} = a^{'} * a = e$.*

Se o grupo obedecer a propriedade da comutatividade, ou seja, se $a * b = b * a, \forall\ a, b \in G$, então esse grupo recebe o nome de *grupo comutativo ou abeliano.*

Definição 2.2. *Um conjunto não vazio A e um par de operações sobre A, respectivamente uma adição $(x, y) \mapsto x+y$ e uma multiplicação $(x, y) \mapsto xy$ (ou $x.y$), é chamado de anel se estiver sujeito as seguintes propriedades - Para a soma: $(A, +)$ é um grupo abeliano; Associativa: $(a + b) + c = a + (b + c)$, $\forall\ a, b, c \in A$; Comutativa: $a + b = b + a$, $\forall\ a, b \in A$; Existência de Elemento Neutro: $\exists$ um elemento $0_A \in A$ tal que $a + 0_A = 0_A + a = a, \forall\ a \in A$; Existência de Elemento Oposto: $\forall\ a \in A, \exists -a \in A$, tal que $a+(-a) = (-a)+a = 0_A$. Para a multiplicação: Associativa: $a.(b.c) = (a.b).c$, $\forall\ a, b, c \in A$; Distributiva (em relação à adição, à direita e à esquerda): $a.(b + c) = a.b + a.c$ e $(b + c).a = b.a + c.a$, $\forall\ a, b, c \in A$.*

Definição 2.3. *Quando um conjunto possui um número finito de elementos, a relação de ordem possui uma representação gráfica adequada para as suas propriedades. Essa representação é denominada* ***"diagrama de Hasse"*** *ou* ***"diagrama de linha"****.*

Definição 2.4. *O peso de Hamming de um vetor v, cuja notação é $\omega(v)$, é definido como sendo o número de elementos não nulos em v. Para um vetor binário, o peso de Hamming é igual ao número de dígitos "1" contidos em v.*

Definição 2.5. *A distância de Hamming entre dois vetores códigos, v e x, cuja notação é $d(v, x)$, é definido como sendo o número de posições em que os dígitos dos dois vetores que são diferentes entre si. Para o caso binário, a distância de Hamming pode ser determinada facilmente pela propriedade de adição módulo-2, pois ela é igual ao número de dígitos "1" contidos no vetor resultante da operação $v \oplus x$.*

$$d(v, x) = \omega(v \oplus x). \tag{1}$$

DOI: 10.5540/03.2021.008.01.0363

3 Desenvolvimento

O diagrama de Hasse é uma forma organizada de apresentar os códons do código genético, com o objetivo de analisar as propriedades dos aminoácidos e classificar os códons de acordo com a sua hidropacidade e outras características. Para a construção utilizou-se a permutação 2031 do rotulamento C. Uma observação é que a atribuição $\{0, 1, 2, 3\}$ de $\mathbb{Z}_4$ é feita em relação a ordem $\{A, C, G, U\}$ em N, ou seja, as bases são estabelecidas na seguinte ordem: adenina (A), citosina (C), guanina (G) e por fim, uracila (U). Utilizando a permutação escolhida, foi estabelecida a seguinte associação, levando em consideração a ordem das bases apresentadas anteriormente e a associação com elementos do conjunto $\mathbb{Z}_2 \times \mathbb{Z}_2$: $A - 11$, $C - 00$, $G - 01$ e $U - 10$.

O diagrama de Hasse (Figura 1) é composto por 64 códons, dispostos em 7 linhas. Cada linha é apresentada da seguinte forma: Na **1ª linha** tem-se o códon que apresenta seu elemento máximo, atribuído por 11. Na **7ª linha** tem-se o códon que apresenta elemento mínimo, atribuído por 00. A **2ª linha** é constituída por seis códons organizados da esquerda para a direita. Biologicamente, como se a leitura fosse no sentido $5' - 3'$. A **6ª linha** é constituída no sentido oposto à **2ª linha** ($3' - 5'$), ou seja, da direita para a esquerda, respeitando a complementariedade algébrica: $00 - 11$ e $01 - 10$. As bases nitrogenadas referentes a essa complementaridade algébrica são: A - C e G - U, ou vice-versa. A **3ª linha** e **5ª linha** são constituídas por quinze códons cada uma, e segue o mesmo procedimento das linhas 2 e 6. A **4ª linha** é constituída por vinte códons e a escrita dos códons é realizada respeitando a complementaridade algébrica seguindo das laterais para o centro.

Cada códon disposto no diagrama de Hasse apresenta sua representação vetorial. O número de códons existente em cada linha está relacionado com o número de zeros (0) ou uns (1) que cada códon possui. Além disso, no diagrama de Hasse, a relação de ordem é feita tomando como base um determinado número de arestas. Cada aresta determina a distância de Hamming entre códons.

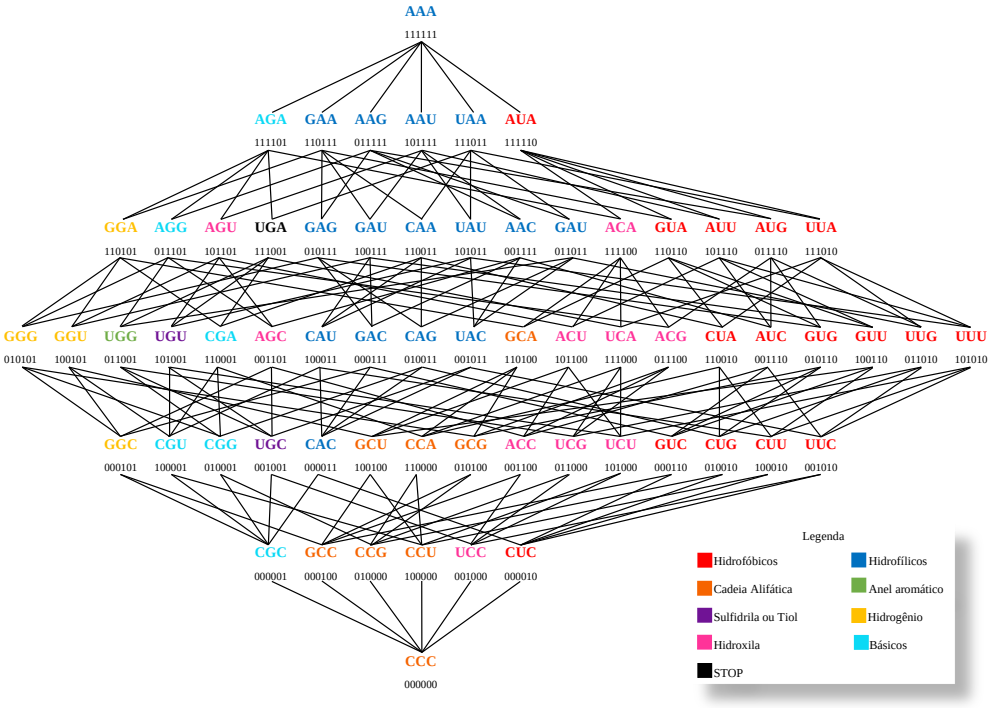

Figura 1: Diagrama de Hasse do Rotulamendo C para a permutação 2031.

DOI: 10.5540/03.2021.008.01.0363

Analisando a complementaridade das bases nitrogenadas do Rotulamento C, nota-se a união das bases adenina (A) com citosina (C) e guanina (G) com uracila (U). Considere como exemplo a cadeia CUC, UCC, CCU, CCG, GCC, CGC tem como imagem a anti-cadeia AUA, UAA, AAU, AAG, GAA, AGA, seguindo a complementaridade algébrica dos códons. Na lateral direita estão os códons hidrofóbicos, aqueles que não possuem "afinidade" com a água (estão representados em vermelho) e no centro superior no diagrama, estão os códons hidrofílicos, aqueles que possuem "afinidade"com a água (estão representados em azul). Diferente dos códons hidrofóbicos não observa-se uma regularidade na distribuição dos códons hidrofílicos (Figura 1).

O códon UGG (anel aromático), ficou distante do códon UGA (STOP), e ele se encontra na cadeia lateral direita. Os códons com característica cadeia alifática estão no centro inferior do diagrama e os códons com característica hidroxila, no centro do diagrama. Na lateral direita, estão os códons com característica hidrogênio, básicos, sulfidrila ou tiol e alguns códons com caraterística hidroxila.

Conforme a proposta de [7], serão apresentadas a seguir as médias das distâncias de Hamming entre os pares de aminoácidos, como as distâncias médias entre os seus respectivos códons (Tabela 1). Foram calculadas as distâncias de Hamming entre cada um dos códons que codificam determinado aminoácido, em seguida, foi realizada a média entre essas distâncias.

A distância de Hamming entre os códons é o número de posições (base) em que os códons se diferem, por meio da associação estabelecida do conjunto $\mathbb{Z}_2 \times \mathbb{Z}_2$. Pode-se notar que os valores das médias das distâncias de Hamming maiores ou iguais a 4 aparecem em negrito, pois são considerados valores altos para essas distâncias, de acordo com [7].

Tabela 1: Médias das distâncias de Hamming entre pares de aminoácidos do Rotulamento C.

*	G	W	C	R	S	V	L	F	M	I	E	D	Y	K	N	Q	H	A	T	P	STOP
G	1	3	3	2	3,33	3	**4,33**	**5**	**4**	**4**	2	2	**4**	3	3	3	3	2	3	3	3,67
W	3	0	1,5	1,83	2,17	**5**	3,5	3,5	3	**4,33**	3,5	**4,5**	2,5	2,5	3,5	2,5	3,5	**4**	3	3	1,33
C	3	1,5	0,5	2,17	1,83	**5**	3,83	2,5	**4,5**	3,83	**4,5**	3,5	1,5	3,5	2,5	3,5	2,5	**4**	3	3	2,17
R	2	1,83	2,17	1,83	2,83	**4**	3,67	**4,17**	**4,17**	**4,39**	2,83	3,17	3,17	3,17	3,5	2,5	2,83	3	3,33	2,67	2,5
S	3,33	2,17	1,83	2,83	1,83	2,67	3,39	2,5	3,17	2,94	**4,5**	**4,17**	2,83	3,5	3,17	**4,17**	3,83	3	2	2,67	2,83
V	3	**5**	**5**	**4**	2,67	1	2,33	3	2	2	2	2	**4**	3	3	3	3	2	3	3	**4,33**
L	**4,33**	3,5	3,83	3,67	3,39	2,33	1,39	1,83	2,5	2,67	3,17	3,5	2,83	3,5	3,83	2,17	2,5	3,33	3,67	2,38	2,83
F	**5**	3,5	2,5	**4,17**	2,5	3	1,83	0,5	2,5	1,83	**4,5**	3,5	1,5	3,5	2,5	3,5	2,5	**4**	3	3	2,83
M	**4**	3	**4,5**	**4,17**	3,17	2	2,5	2,5	0	1,33	2,5	3,5	3,5	1,5	2,5	3,5	3,5	3	2,25	**4**	3
I	**4**	**4,33**	3,83	**4,39**	2,94	2	2,67	1,83	1,33	0,89	3,17	2,83	2,83	2,17	1,83	**4,17**	3,83	3	2	3,92	3,44
E	2	3,5	**4,5**	2,83	**4,5**	2	3,17	**4,5**	2,5	3,17	0,5	1,5	3,5	1,5	2,5	1,5	2,5	3	**4**	**4**	2,67
D	2	**4,5**	3,5	3,17	**4,17**	2	3,5	3,5	3,5	2,83	1,5	0,5	2,5	2,5	1,5	2,5	1,5	3	**4**	**4**	3,83
Y	**4**	2,5	1,5	3,17	2,83	**4**	2,83	1,5	3,5	2,83	3,5	2,5	0,5	2,5	1,5	2,5	1,5	**5**	**4**	**4**	1,83
K	3	2,5	3,5	3,17	3,5	3	3,5	3,5	1,5	2,17	1,5	2,5	2,5	0,5	1,5	2,5	3,5	**4**	3	**5**	1,83
N	3	3,5	2,5	3,5	3,17	3	3,83	2,5	2,5	1,83	2,5	1,5	1,5	1,5	0,5	3,5	2,5	**4**	3	**5**	2,83
Q	3	2,5	3,5	2,5	**4,17**	3	2,17	3,5	3,5	**4,17**	1,5	2,5	2,5	2,5	3,5	0,5	1,5	**4**	**5**	3	1,83
H	3	3,5	2,5	2,83	3,83	3	2,5	2,5	3,5	3,83	2,5	1,5	1,5	3,5	2,5	2,5	0,5	**4**	**5**	3	2,83
A	2	**4**	**4**	3	3	2	3,33	**4**	3	3	3	3	**5**	**4**	**4**	**4**	**4**	1	2	2	**4,67**
T	3	3	3	3,33	2	3	3,67	3	2,25	2	**4**	**4**	**4**	3	3	3	**5**	2	1	3	3,67
P	3	3	3	2,67	2,67	3	2,38	3	**4**	3,92	**4**	**4**	**4**	**5**	**5**	**5**	3	2	3	1	3,67
STOP	3,67	1,33	2,17	2,5	2,83	**4,33**	2,83	2,83	3	3,44	2,67	3,83	1,83	1,83	2,83	2,83	2,83	**4,67**	3,67	3,67	0,89

*Representação dos aminoácidos por um único símbolo. A - Alanina, C - Cisteína, D - Ác. Aspártico, E - Ác. Glutâmico, F - Fenilalanina, G - Glicina, H - Histidina, I - Isoleucina, K - Lisina, L - Leucina, M - Metionina, N - Asparagina, P - Prolina, Q - Glutamina, R - Arginina, S - Serina, T - Treonina, V - Valina, W - Triptofano, Y - Tirosina.

Segundo [7], quando se obtém valores maiores nas médias das distâncias de Hamming entre pares de códons, é sabido que elas são as mais perigosas, pois pode-se alterar as propriedades dos aminoácidos e as funções biológicas das proteínas, acarretando até riscos de grandes mutações genéticas. Pode-se notar que aminoácidos com grandes diferenças em suas propriedades apresentaram altos valores de distâncias de Hamming.

4 Análise dos Resultados

Com base nas características de cada aminoácido codificado pelos seus respectivos códons, serão analisadas algumas diferenças e semelhanças biológicas acerca dos cálculos obtidos na Tabela 1.

O códon UGG codifica o aminoácido triptofano (W) e os códons GUU, GUC, GUA e GUG codificam o aminoácido valina (V). A distância média entre o códon UGG e os códons GUU, GUC, GUA e GUG é igual a 5. O triptofano (W) é um anel aromático e a valina (V) é um aminoácido hidrofóbico. Os códons AUU, AUC e AUA codificam o aminoácido isoleucina (I), que é hidrofóbico. A distância média entre esses códons e o códon do aminoácido triptofano é igual a 4,33. Os códons GAU e GAC codificam o aminoácido ácido aspártico (D), e sua distância média com o códon que codifica o aminoácido triptofano (W) é igual 4,5. O ácido aspártico (D) é hidrofílico, sendo este diferente do triptofano. Os códons GCU, GCC, GCA e GCG codificam o aminoácido alanina (A), sendo uma cadeia alifática. A sua distância média com o códon do aminoácido triptofano é igual a 4. Já os códons UGU e UGC codificam o aminoácido cisteína (C), que são códons com características sulfidrila ou tiol. A distância média com o códon de aminoácido triptofano é igual a 1,5. Pode-se observar que no diagrama da Figura 1 esses códons estão próximos do códon do aminoácido triptofano (W), concluindo assim, que eles possuem características em comum.

Os códons CGU, CGC, CGA, CGG, AGA e AGG codificam o aminoácido arginina (R), que são códons classificados como básicos e os códons GUU, GUC, GUA e GUG codificam o aminoácido valina (V), que possui característica hidrofóbica. Pode-se notar que a distância média entre os códons desses aminoácidos é igual a 4. Os códons UUU e UUC codificam o aminoácido fenilalanina (F). Como é um aminoácido com característica hidrofóbica, pode-se notar que a sua distância média entre o aminoácido arginina (R) é igual a 4,17. O códon AUG codifica o aminoácido metionina (M), sendo este hidrofóbico. A distância média dos códons do aminoácido arginina (R) entre os códons do aminoácido metionina (M) é igual a 4,17. Já os códons GGU, GGC, GGA e GGG codificam o aminoácido glicina (G), que são códons que tem hidrogênio como característica. A distância média com o códon do aminoácido arginina (R) é igual a 2. Pode-se notar que tanto os códons que codificam o aminoácido arginina (R) quanto os códons que codificam o aminoácido glicina (G) estão próximos um do outro, no diagrama (Figura 1), concluindo assim que eles possuem características em comum.

Os códons ACU, ACC, ACA e ACG codificam o aminoácido treonina (T), que é classificado como hidroxila. A distância média com o códon do aminoácido serina (S) é igual a 2. Como os aminoácidos serina (S) e treonina (T) são classificados como hidroxila, possuindo as mesmas características, faz sentido o fato delas estarem tão próximas uma da outra no diagrama (Figura 1). Assim, justifica o valor da média entre eles ser menor.

Pode-se observar os outros casos na Tabela 1 com o diagrama (Figura 1) construído. Além disso, se forem efetuadas as análises de todos os casos, nos quais, as distâncias médias entre as distâncias de Hamming foram maiores ou iguais a 4, as suas distâncias estão refletidas no Diagrama de Hasse, com base nas diferenças biológicas entre os códons.

5 Considerações Finais

Diante do exposto, percebe-se uma interessante caracterização algébrica associada ao mapeamento do código genético, permitindo uma análise biológica acerca da construção do diagrama de Hasse para a permutação 2031 do rotulamento C, de acordo com as características dos aminoácidos. Além disso, pode-se observar com a Tabela 1, que o diagrama de Hasse reflete de forma significativa as características físico-químicas dos aminoácidos baseados nos cálculos das distâncias de Hamming, de forma a poder utilizar esse resultados em estudos de fenômenos mutacionais.

DOI: 10.5540/03.2021.008.01.0363

Agradecimentos

Agradecemos à Universidade Federal de Alfenas - UNIFAL-MG, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes e a Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG.

Referências

[1] Alberts, B. e et al. *Biologia Molecular da Célula.* 5ª. ed. Porto Alegre: ArtMed, 2010.

[2] Domingues, H. H. e Iezzi, G. *Álgebra Moderna.* 4º. ed. São Paulo: Atual, 2003.

[3] Faria, L. C. B. D. e Palazzo Júnior, R. Existências de códigos corretores de erros e protocolos de comunicação em sequências de DNA. Tese de Doutorado. Unicamp, 2011.

[4] Menezes, P. B. *Matemática Discreta:* para computação e informática. 2º. ed. Porto Alegre: Sagra Luzzatto, 2005.

[5] Oliveira, A. J. e Palazzo Júnior, R. Análise Algébrica dos Rotulamentos Associados ao Mapeamento do Código Genético. Dissertação de Mestrado. Unicamp, 2012.

[6] Rocha, A. S. L. e Palazzo Júnior, R. Modelo de sistema de comunicação digital para o mecanismo de importação de proteínas mitocondriais através de códigos corretores de erros. Tese de Doutorado. Unicamp, 2010.

[7] Sánchez, R., Morgado, E. and Grau, R. The genetic code boolean lattice. *MATCH Commun. Math.Comp.Chem*, v. 52, n. 52, p. 29-46, 2004. ISSN: 0340-6253.

[8] Sánchez, R., Morgado, E. and Grau, R. Gene algebra from a genetic code algebraic structure. *Journal of Mathematical Biology*, v. 51, n. 4, p. 431-457, 2005. DOI: 10.1007/s00285-005-0332-8.

[9] Sánchez, R.; Grau, R. and Morgado, E. A novel Lie algebra of the genetic code over the Galois field of four DNA bases, *Mathematical Biosciences*, v. 202, p. 156–174, 2006. DOI: 10.1016/j.mbs.2006.03.017.

