

O conceito de informação em biologia

Romeu Cardoso Guimarães

Como citar: GUIMARÃES, R. C. O conceito de informação em biologia. *In:* GONZALES, M. E. Q.; BROENS, M. C. (org.). **Encontro com as Ciências Cognitivas**. Marília: Unesp Marília Publicações, 1998. 2 v. p. 69-94. DOI: <https://doi.org/10.36311/1998.978-85-86738-03-4.p69-94>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

O CONCEITO DE INFORMAÇÃO EM BIOLOGIA

Romeu Cardoso GUIMARÃES¹

Introdução

O termo informação tem sido muito usado em biologia, especialmente na genética molecular (Alberts e col., 1994; Watson e col., 1988; para economia de citações, consultar estas para os detalhes referentes à maioria dos conceitos e dados de bioquímica e biologia celular, deste ponto em diante). Refere-se, comumente, à propriedade do DNA e dos genes de conter a informação ontogenética (a 'programação' própria de cada espécie para o desenvolvimento dos organismos), em óbvia analogia aos programas de computadores. Mais para o caso humano, tornou-se comum a analogia da programação genética com a memória ROM e da programação aprendida com a memória RAM.

O termo memória (Guimarães, 1997b) refere-se ao caráter evolutivo e histórico da programação genética, indicando que estão gravados, nos genes, adequações funcionais às condições de vida das gerações anteriores da espécie (informação filogenética) que voltam a se manifestar nas gerações presentes. [Fica implícita a 'expectativa' de que as condições ambientais também não sejam muito diferentes das anteriores; caso contrário, haverão dificuldades adaptativas].

Diz-se que os genes são estoques de informação. O termo 'programação' genética é analogia discutível, porque muitas são as variações na expressão dos genes em decorrência de interações diversas com as condições ambientais (Pereira e col., 1996); é, também, limitada desde que a hereditariedade biológica é de células inteiras, através do contínuo reprodutivo e a genética é somente uma parte mais estável do sistema.

¹ Departamento de Biologia Geral – Instituto de Ciências Biológicas – UFMG - 31270-901 - Belo Horizonte, MG. E-mail: romeucg@mono.icb.ufmg.br - Homepage: <http://www.icb.ufmg.br/~romeucg>

Metáforas?

Notam-se grandes dificuldades na definição dos termos acima, quando extrapolados da linguagem humana, essencialmente teleológica e tecnológica, para a natureza biológica. É interessante investigar se são somente metáforas úteis, ou se adquirem sentidos específicos na biologia e, em assim sendo, quais são esses.

Vida

São dificuldades semelhantes às da definição de vida. Neste caso, parece satisfatória a caracterização de vida como ‘um conjunto de propriedades dos seres vivos’ (Guimarães, 1994b). Os seres vivos conhecidos são objetos concretos, as células, das bactérias aos organismos multicelulares. Suas propriedades são objetos abstratos, descrições de comportamentos. Os biólogos podem não obter consenso quanto à listagem mínima, necessária e suficiente, para caracterizar os seres vivos, dada sua complexidade e o grande número de propriedades específicas. Qualquer recorte em sua caracterização pode não satisfazer a todos. Talvez a solução seja fazer listagem por ordem decrescente de importância, que seja aproximadamente consensual. Pode-se começar pelas propriedades de metabolismo e reprodução, acrescentar evolução e outras, com limite de demarcação difícil. Pode-se destacar a autonomia funcional e a compartimentação (enclausuramento) frente ao ambiente externo que, no entanto, são sempre parciais.

Ao se exigir o metabolismo, exclui-se os vírus. Esses passam a ser considerados como partes ou pedaços de células, vagantes e em circulação entre as suas células hospedeiras. Como gerar um conceito de vida referente a objetos artificiais, hipotéticos, com ‘vida mínima’?

Objeto abstrato

O termo informação é, assim como vida, um objeto conceitual, abstrato. Refere-se a propriedades interativas de objetos que lhes permitem participar de conjuntos, esses sendo chamados de sistemas. Surge, então, o problema de definir os mínimos de informação e de sistema. A propriedade interativa mínima

pode ser um quantum de energia (Stonier, 1990) e o sistema mínimo pode ter somente dois elementos em interação.

Chega-se, assim, a limiares quase triviais, esvaziando o significado e até dificultando o emprego dos termos em situações mais complexas. Faz-se necessário classificar e introduzir qualificações. A classificação 'minimalista' de Stonier, identificando informação com as formas organizadas de energia (excluindo a térmica), parece-me inadequada à biologia; o calor pode ser percebido pelos seres vivos como informação útil e adaptativa.

Causalidade

Os movimentos e transformações em sistemas são causados por forças ou formas de energia, atuando sobre objetos ou sobre outras forças. Em algumas situações, prefere-se falar em causalidade informacional, em contraposição à energética (Pereira e Gonzales, 1995; Debrun, 1996a), referindo-se às transformações resultantes, que são menos previsíveis, como que influenciadas em grau maior pelo objeto receptor da força empregada, que parece poder 'interpretar' a força que incidiu sobre ele e reagir com certa liberdade.

Assim, a distinção introduzida, entre causalidade energética e informacional, decorre mais de propriedades do receptor do que da fonte ou do sinal. A primeira seria reservada a sistemas toscos e simples, com maior previsibilidade de respostas ou transformações, quando o agente (energético) praticamente 'dita uma ordem'. A segunda seria aplicada a sistemas complexos e com alto grau de conectividade interna, o que resulta em maiores possibilidades regulatórias e modulatórias, de modo que as respostas não são unívocas. Esse é o caso dos sistemas biológicos, ecológicos e sociais.

Na biologia, não é comum a ocorrência de grandes diferenças de potenciais energéticos; as barreiras energéticas são vencidas por transformações graduais, em passos sucessivos menores nas cadeias metabólicas, que são amplamente interconectadas em redes, de modo que os resultados fisiológicos são mais adequados às explicações informacionais do que às simplesmente energéticas.

Estruturas e modulações

Há, por certo, diferenças de intensidade da força na atuação de certos elementos dentro do sistema biológico. O componente genético tende a ser mais forte, porque apresenta maior estabilidade ao longo de gerações e a ele corresponde a produção de estruturas (principalmente RNA, proteínas e outras derivadas destas). O componente ambiental tem atuação mais regulatória, favorecendo ou dificultando, modulando a expressão dos genes, portanto, tendendo a ser mais fraco. No entanto, a complexidade de quase todos os comportamentos e interações gênicas é tão grande que possibilita a ocorrência, em geral, de modulações fortes (Guimarães, 1992, 1994a). Assim, pode-se dizer que mesmo a função genética seria informacional.

Fontes, sinais e receptores

Parece não ser frutífera a tentativa de distinção classificatória (entre sistemas energéticos e informacionais) baseada nos tipos de fontes produtoras de sinais ou nos tipos de sinais. Esses podem ser tão simples como um objeto físico (lâmpada, sol) emissor de quanta energéticos (fótons) ou tão complexos como um ser humano emissor de ondas de choque (sons), e ambos podem se envolver em resultados ou transformações dos mais variados tipos. Fótons podem simplesmente aquecer objetos fisicamente, sem maiores consequências, ou produzir matéria biológica, pela fotossíntese. Sons podem produzir movimentos físicos ou resultar na comunicação humana.

É evidente que o receptor se torna o principal elemento definidor do resultado da interação e a classificação dos tipos de sistemas deve basear-se, principalmente, nas propriedades dos receptores. Para estudar o conceito de informação em biologia, é necessário definir com clareza os processos envolvidos no sistema biológico.

O sistema biológico é uma estrutura hipercíclica

A estrutura básica e o modo de operação interna dos seres vivos estão esquematizados na **Figura 1**, como 3 ciclos funcionais superpostos, um dentro do outro, em arranjo do tipo 'casca de cebola' (Guimarães, 1996). Essa apresentação

é um modo de configurar hiperciclos, séries de ciclos interconectados e encadeados (ver Eigen e Schuster 1979; Eigen 1993).

Componentes e funções

Os elementos fundamentais são dois tipos de biopolímeros, genes e proteínas. As funções básicas são:

- a) metabolismo, que fornece os monômeros (nucleotídeos 'N' e aminoácidos 'aa'), a serem polimerizados, respectivamente, nos genes e proteínas;
- b) replicação dos genes e sua transcrição, sempre em RNA, copiando os genes, como polímeros-modelos. A função de gene pode ser cumprida, indistintamente, por DNA ou RNA, e esses são considerados quase 'sinônimos', molecularmente, sendo chamados genericamente de ácidos nucleicos (AN ou NA). Na fisiologia celular, DNA é gene e RNAs são intermediários (de vários tipos, como: mensageiro 'm' e de transferência 't' dos aminoácidos, na forma de 't-aa') que exercem funções diversas na síntese de proteínas.
- c) tradução dos genes em proteínas, polimerizando aminoácidos segundo a sequência-modelo do RNA mensageiro. Entende-se tradução como 'mudança de alfabeto', de uma sequência de N-N-N-N... para aa-aa-aa-...; a troca é, mecanicamente, pequena mas de grandes consequências funcionais.

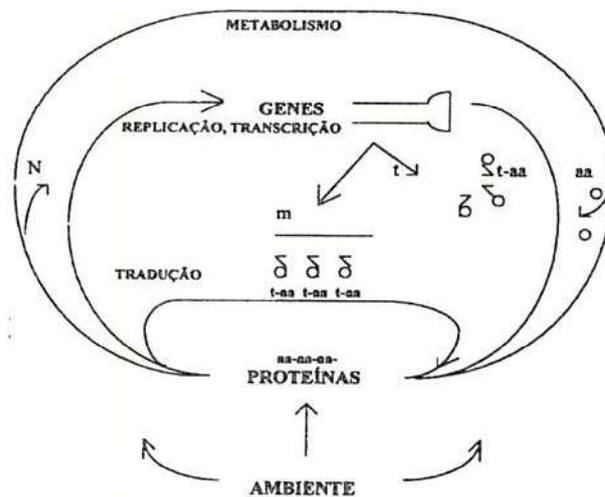


Figura 1. Esquema da estrutura hipercíclica básica da célula.

Consulte '*Componentes e funções*' a '*Moléculas vivas?*' para detalhes.

Genes e cabideiros

As funções dos genes podem ser descritas por analogia com as de cabideiros de guarda-roupas ou de vitrines de lojas, onde se penduram objetos que têm significados (informação) para os respectivos sistemas de usuários, os habitantes da casa ou os vendedores e compradores; ou de pontes- ou talhas-rolantes de indústrias de montagens, que transportam objetos segundo planejamento detalhado de seqüência temporal e de tipos, adequados ao sistema de produção (Figura 2).

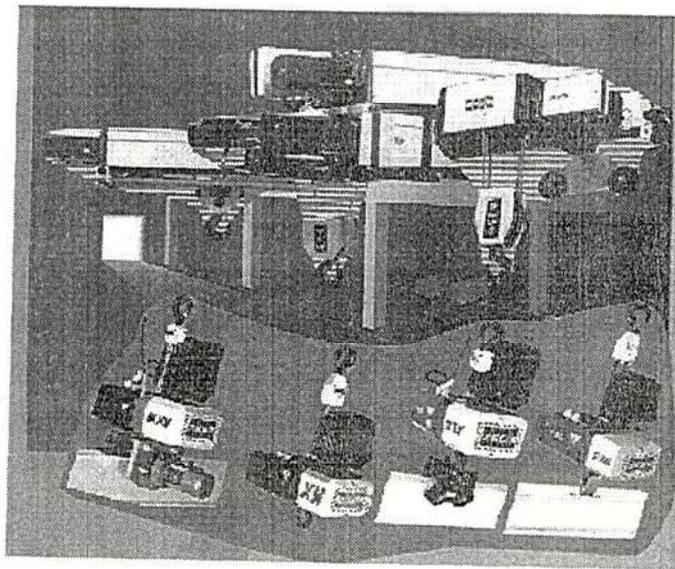


Figura 2. Modelo de cabideiros para a função informacional dos biopolímeros.

A descrição molecular diz que os ácidos nucleicos são fios poliméricos onde as seqüências de monômeros (nucleotídeos) são os cabides ou ganchos. Seus radicais específicos (átomos expostos das bases) são encadeados em seqüência ordenada pelo sistema de replicação e tradução, adequadamente ao cumprimento de todas as funções dos organismos, através da dinâmica hipercíclica (Guimarães, 1997c).

Duas seqüências de cabides são mostradas, como análogas às seqüências dos monômeros dos biopolímeros. Há algumas diferenças entre os cabides, mas todos são adequados à montagem de cadeias, pelas suas estruturas que formam uma 'espinha dorsal', amarradas e formando os trilhos (seqüência superior) dos

cabideiros. Os ganchos correspondem aos radicais dos monômeros (bases dos nucleotídeos ou cadeias laterais dos aminoácidos) e estão expostos, neste caso, para o espaço entre as duas cadeias, sugerindo interações internas à conformação espacial dos polímeros, como nas duplas fitas de DNA ou RNA. A cadeia inferior ainda não está polimerizada, os monômeros estão somente alinhados consecutivamente. Do mesmo modo como os sistemas de usuários dos cabides podem pendurar objetos nos ganchos, de maneira adequada à sua utilização (neste caso, planejada pelos usuários), o sistema hipercíclico celular organizou, evolutivamente, as sequências dos monômeros, de modo que sua replicação seja fácil e fiel e sua tradução produza proteínas adequadas à realização das funções biológicas, de metabolismo, replicação e tradução.

Enzimas e ribozimas

As funções são realizadas, principalmente, pelas proteínas. São enzimas (catalisadores) que (Figura 1) produzem os monômeros, realizam as transformações metabólicas e energéticas, e polimerizam os monômeros, copiando os genes, na replicação e na transcrição, e traduzindo-os em proteínas; em todas as polimerizações, seguem-se as sequências dos genes e RNAs.

Há RNAs que podem desempenhar atividades catalíticas, mas essas são quantitativamente minoritárias. É possível que essas atividades (de ribozimas; Soares, 1997) tenham desempenhado papel maior em épocas precoces da evolução das células e protocélulas, depois tendo sido transferidas para as proteínas.

Moléculas vivas?

O sistema bimolecular (genes e proteínas) provoca questionamentos de difícil solução, quanto ao modo de origem. Há os genocêntricos, propondo início pelo RNA, valendo-se de suas propriedades potencialmente auto-replicadoras e potencialmente auto-tradutoras (ver Soares, 1997). Um RNA poderia se ligar a aminoácido, tornando-se um RNA de transferência ou transportador (Figura 1, t e t-aa) e se alinhar sobre outro RNA (que se torna um RNA mensageiro (Figura 1, m),

de acordo com as regras de complementaridade de bases; assim, os aminoácidos estariam alinhados e aproximados (Figura 1, tradução), podendo se polimerizar em proteínas (Figura 1, aa-aa-aa ...).

O RNA é chamado pelos genecêntricos de uma 'molécula viva'. As maiores dificuldades com essa proposta estão na obtenção primária do RNA e de seus monômeros, moléculas já muito complexas. Essa abordagem é 'de cima para baixo', como construir uma pirâmide começando pelo ápice (Guimarães, 1995). Por isso, muitos sugerem que tais eventos foram tardios na evolução das células.

Sistemas proteico-metabólicos

A outra corrente é dos sistêmicos, começando pelas proteínas (Guimarães, 1995). Essas são de fácil obtenção pela química orgânica pré-biótica. Podem se agregar em glóbulos, também facilmente, com interessantes propriedades protometabólicas, até reprodutivas. Note-se a distinção entre replicação, cópia de modelo molecular, e reprodução, divisão de objetos complexos em 'filhos'.

Esta é uma abordagem 'de baixo para cima', começando pela base da pirâmide. No entanto, falta-lhe um mecanismo de estabilidade replicativa. Diz-se que as proteínas têm potenciais somente hetero-sintéticos (de fazer outras coisas) mas não os autosintéticos (de fazer a si mesmas, como poderiam fazer os RNA auto-replicadores). Sugere-se que o RNA teria surgido no interior de agregados proteico-metabólicos.

Coevolução

Pode-se, ainda, propor que ambos processos ocorreram, cada um em evolução paralela, eventualmente a confluir; uma origem dupla para o sistema biológico (Dyson, 1985). O processo de junção poderia ser descrito tanto como simbiose quanto como parasitismo (de tipo viral) do sistema metabólico pelo sistema replicador. As reatividades integradoras e vinculadoras (afinidades constitutivas) entre ácidos nucleicos e proteínas são múltiplas.

Ambas hipóteses de origem 'por um dos componentes' são possíveis:

proteínas (hipótese metabólico-sistêmica) poderiam polimerizar RNA ou estes (hipótese genecêntrica) poderiam ter origem independente; as ribozimas poderiam ser autoreplicativas ou até atuar como ativadoras de aminoácidos (como sintetases de t-aa). Pode ser igualmente razoável a hipótese de origem dupla, por coevolução e posterior integração. Talvez as primeiras posturas sejam tendenciosidades da ciência e cultura ocidentais, de preferir inícios singulares e monofatoriais. No momento, é mais razoável dizer que a vida surgiu no momento em que o vínculo associativo e funcional nucleoproteico (de genes com proteínas) se estabeleceu.

Cristais

As argilas e cristais poderiam ser facilitadoras de polimerizações, tanto de ácidos nucleicos como de proteínas, com a vantagem de lhes oferecer unidade primordial, por origem comum. As polimerizações podem seguir ordenação pelas superfícies dos cristais. Estes são estruturas físicas estáveis, providenciando repetitividade às sínteses (Guimarães, 1997a).

Alguns cristais (como sulfeto ferroso, pirita), podem ainda funcionar como 'células fotoelétricas', captando energia solar e a transferindo para moléculas orgânicas, providenciando um tipo de protometabolismo (Danchin, 1990).

Quando ácidos nucleicos se tornam genes

Ácidos nucleicos não são genes, *a priori*. Tornam-se genes quando adquirem a função de modelos para síntese de proteínas ou algum outro tipo de função dentro do sistema biológico (Pardini e Guimarães, 1992). Antes disso, podem ser somente moléculas replicadoras, ainda que auto-replicadoras. Esta situação, meramente replicativa, é análoga à dos cristais, portanto trivial com respeito à biologia, não justificando sua conotação de 'molécula viva'.

Geração de informação biológica

O processo de um ácido nucleico tornar-se gene é o do sistema

hipercíclico. Todos os componentes incluídos na Figura 1 são necessários. O processo envolve dinâmica mutacional e populacional e é o mesmo da evolução do sistema.

Os genes sofrem mutações, na replicação (com o DNA e as enzimas celulares sofisticadas atuais, da ordem de 10^{-9} por base) e na transcrição (da ordem de 10^{-6} ; essa maior frequência decorre da falta de processos enzimáticos de reparo, que atuam em manter a estrutura regular da dupla fita helicoidal do DNA). Em sistemas virais, como os da gripe e da AIDS, com genes de RNA e polimerases mais erráticas, as taxas de mutação são maiores, da ordem de 10^{-4} (Eigen, 1993; Holmquist e Filipiski, 1994). A tradução, nas células atuais, é tão errática como estes sistemas virais de baixa fidelidade.

Assim, o sistema envolve *populações* de variantes dos genes e das proteínas. Dentre as variantes, algumas serão funcionalmente eficientes, capazes de realizar a ciclização. O processo de fixar o sistema de correspondências de letras, das trinças de nucleotídeos (codons) com os aminoácidos (o código genético), já foi modelado (Wills, 1993) probabilisticamente, através de flutuações nessas ciclizações interativas.

Dependurando correspondências

Poder-se-ia, em situação limite, iniciar com qualquer sequência nucleotídica. Desde que algumas proteínas produzidas sejam catalizadoras de metabolismo, replicação e tradução, o processo evolutivo se inicia, com fixação das correspondências de letras (o código genético) e geração de sequências (genes) mutuamente adequadas às sínteses de ambas. É possível que o metabolismo seja, inicialmente, externo (ambiental), posteriormente endogenizado.

Códigos genéticos

Parece adequado o uso do termo código para descrever o sistema de correspondências entre as trinças de bases dos genes e os aminoácidos das proteínas. Foi gerado e fixado pelo processo hipercíclico celular, evolutivamente. Configura um tipo padrão (canônico) para quase todos os tipos de organismos ('universal') mas apresenta variações pontuais, como linguagens particulares de certos

tipos de organismos.

Parece razoável, também, a analogia de ácido nucleico (molécula que pode ser sintetizada com qualquer sequência, por máquinas laboratoriais) com um cristal ou um disco magnético ou ótico virgem, somente formatado mas não gravado, e a de gene com um disco gravado pelo sistema hipercíclico celular. As analogias dos cabideiros e pontes-rolantes são dessa mesma ordem. A analogia entre a memória genética e a ROM, e entre a cultural e a RAM, requerem as ressalvas decorrentes da complexidade dos sistemas de leitura e expressão dos genes, que diferem fundamentalmente dos programas de computadores, principalmente pela plasticidade e historicidade biológicas.

Biodiversidade e auto-organização

O estabelecimento dos ciclos, em vinculação 'fechada', envolve somente forças positivas, de auto-alimentação e reforço retrógrado. Os elementos menos eficientes simplesmente são deixados de lado, não participando da dinâmica evolutiva. Assim, não cabe a denominação de seleção natural ao processo. É previsível que o sistema possa gerar variedade de combinações de tipos de elementos (biodiversidade): tipos diversos de genes e de proteínas, em ciclização, poderiam ser acomodados e conviver com eficiência similar, dentro dos mesmos limites de condições metabólicas. Não é consequência inevitável a competição eliminatória, podem ocorrer até 'ecossistemas' de hiperciclos interdependentes, uns tornando-se partes do ambiente de outros, e arranjos mais complexos (**Figura 3**).



Figura 3. Sistemas de hiperciclos, como indivíduos, formam populações de variantes, que podem conviver como em 'comunidades ecológicas'.

Auto-organização primária

Parece ser justificada a consideração de ácidos nucleicos e proteínas como elementos 'realmente distintos', e sua vinculação hipercíclica e 'urobórica' como um evento de auto-organização primária (Debrun, 1996c). É óbvio que não poderiam ser 'absolutamente distintos', o que impediria suas interações e compatibilidades mútuas.

Informação e complexidade

A organização cíclica é multifatorial e complexa, indicando que o processo pode ser qualificado como gerado por e gerador de informação. As interações moleculares envolvem, singularmente, bits ou quanta energéticos, mas a descrição do conjunto se torna extremamente difícil nesses termos puramente energéticos e o termo informação se torna útil. O processo energético atuante sobre um passo se relaciona com todos os outros, em graus diversos e por vias amplamente ramificadas.

Vias eferente e aferente da informação

A informação biológica pode ser descrita como distribuída entre os elementos do conjunto, por adequação mútua. A sequência de bases nos genes é entendida como depositária de grande parte da informação, quando nela confluem atribuições múltiplas, correspondentes à estrutura de todas as proteínas. O processo cíclico resulta em bidirecionalidade na transferência de informação.

Na *via eferente* dos genes para as proteínas, o processo é de correspondência direta e linear de sequências, por transcrição e tradução, de modo mecanicista. Na *via aferente* das proteínas para os genes, o processo é funcional, de modo informacional, através das diversas eficiências catalíticas de um grande número de enzimas, em conjunto. O percurso evolutivo é de sofisticar o conjunto, com obtenção de maior fidelidade de replicação, transcrição e tradução (*via eferente*), e confiabilidade reprodutiva do sistema (*via aferente*).

A *via eferente* pode ser caracterizada como envolvendo *informação*

em sentido fraco, desde que decorre de mecanismos energéticos simples, da replicação à tradução dos genes, em via de descrição detalhada pela bioquímica. Pelo contrário, a via aferente envolve *informação em sentido forte*. O desempenho interativo interno ao organismo e deste com o ambiente, desemboca integrada e globalmente, afunilada e concentradamente, nos eventos reprodutivos singulares e momentâneos. Parece projeto talvez inalcançável tentar descrever com detalhes esses eventos, ou quantificar a informação envolvida no processo.

A genética populacional tenta medir, em termos de ‘valor adaptativo’ darwiniano, o quanto determinados genótipos influenciam a eficiência reprodutiva de amostras de populações de indivíduos portadores desses genótipos, em relação a amostras de outras populações de indivíduos que não os portam. Tais medidas são sempre relativas às amostras selecionadas e aos contextos interativos delas, individual e ambiental.

A via aferente é, também, ‘*redatora e editora*’ da informação hereditária, no percurso evolutivo. Variantes (mutações genéticas ou fisiológicas) que resultam em melhor desempenho funcional e reprodutivo, nos contextos interativos orgânico e populacional, frente àquele contexto ambiental, serão porção maior das gerações sucessivas.

Estratégias de confiabilidade

No entanto, o sistema não evoluiu em perfeição ou fidelidade absolutas. Não se conhecem os processos envolvidos na gênese de um aparente ‘equilíbrio’, com um ótimo de fidelidade, sem a estática do perfeito (certamente inatingível) e sem as catástrofes decorrentes de erros abundantes. A utilização de genes segmentados em pequenos módulos (chamados de exons) e de proteínas ‘primárias’ (as cadeias polipeptídicas) de tamanho pequeno foi parte da estratégia encontrada.

As estruturas celulares proteicas grandes são construídas por agregação de proteínas ‘primárias’ menores. Proteínas grandes têm alta probabilidade de conter defeitos internos e suas populações menor fração das funcionalmente

adequadas. As doenças genéticas humanas mais frequentes incidem sobre genes muito grandes.

Informação 'top-down' e 'bottom-up'

A Teoria da Informação, desenvolvida para a tecnologia da comunicação e codificação (ver Shannon e Weaver, 1949, em Silva, 1996), poderia ter aplicabilidade direta no estudo de partes da via eferente, 'de cima para baixo', das estruturas genéticas para as proteicas (Atlan, 1992). Os erros na transmissão da informação dos genes para as proteínas são, também, usados como explicação para os processos de envelhecimento e morte celular. É aparente que a aplicação dos procedimentos shannonianos para a via aferente, 'de baixo para cima', de retorno da 'informação sistêmica' para o processo de geração, estocagem e editoração da 'informação genética' (das funções para as estruturas), será mais complexo.

Mutações markovianas

A sofisticação da tecnologia do DNA está implementando o entendimento do processo mutacional. Até esse momento, prevalecia a idéia de que as mutações seriam essencialmente aleatórias, mas já se começa a acumular evidências explicativas de porque algumas sequências tendem a mudar com maior frequência ('pontos quentes') que outras e em quais direções (Holmquist e Filipiski, 1994). Parece que a explicação por aleatoriedade permanecerá importante e majoritária, mas não será a explicação total. Mecanismos markovianos deverão ocupar papel de destaque, entre outros.

Informação nas interações com o ambiente

O organismo interage com o ambiente através dos comportamentos; no nível molecular, principalmente através das proteínas e do metabolismo. Seu desempenho fisiológico e reprodutivo pode ser afetado, identificando-se pressões seletivas externas e a introdução de *processos darwinianos*.

Destaca-se que o processo de seleção natural por fatores ambientais envolve principalmente forças negativas, atuantes na eliminação de variantes menos adequadas àquele ambiente. Os efeitos positivos, favorecedores das variantes restantes, seriam somente corolários, secundários à seleção negativa. Os organismos que não foram prejudicados naquelas interações estariam somente liberados para desenvolver suas próprias potencialidades internas. A seleção, a cada momento, seria isotrópica com relação ao processo evolutivo subsequente, não lhe indicando uma direção.

Forças positivas e negativas

Pelo contrário, o processo de 'fechamento' dos ciclos funcionais biológicos envolve principalmente forças de realimentação positiva, o que justifica o modelo da auto-organização. Ressalte-se, no entanto, que forças negativas também podem participar da 'economia interna' biológica, principalmente como reguladoras e moduladoras dos ajustes funcionais. Alguns passos do metabolismo podem gerar produtos tóxicos ou inibidores de outras reações, exaurir substratos (por bloqueio, sequestro ou transformação e destruição de produtos) que seriam necessários para outros passos, indicando analogias com a seleção darwiniana (predação, competição etc).

Fica, assim, difícil e arbitrário tentar demarcar os limites dos processos de auto-organização ou darwinianos simplesmente através da caracterização do sentido, positivo ou negativo, das forças atuantes.

Forças isotrópicas e anisotrópicas

Estudos moleculares contribuem, também, para demonstrar que uma direcionalidade evolutiva ou organizacional pode ser obtida mesmo na ausência da identificação de forças intrinsecamente direcionais. No ambiente aquoso, as interações moleculares direcionais predominantes são as polares, hidrofílicas e eletromagnéticas. Diz-se que as moléculas que não interagem com a água (apolares, hidrofóbicas e hidroapáticas) simplesmente são 'empurradas' para longe, para nichos sem água, sem participação intrínseca dos elementos apolares. O processo de

afastamento dos não interativos pode ser visto como adirecional e isotrópico, com respeito à participação dos componentes apolares, sendo somente secundário à atuação das forças polares direcionais, que permanecem exógenas.

Por outro lado, pode-se dizer que, apesar das ‘forças hidrofóbicas’ não existirem como tais (alguns defendem sua realidade, talvez como ‘ressonâncias’ estabelecidas entre os elementos apolares), os objetos hidrofóbicos podem contribuir, significativamente, para a ordenação estrutural do sistema, no sentido em que irão se acomodar uns aos outros, em ajustamento espacial de alta precisão. Por exemplo, a estrutura regular das hélices de DNA e RNA depende fundamentalmente do empilhamento planar das bases, altamente hidrofóbicas, formando um ambiente interno hidrofóbico entre as duas fitas pareadas. A energia térmica necessária para romper o empilhamento (que possibilitará a entrada de água ou outros elementos polares naquele ambiente, entre as bases) pode ser até superior à energia envolvida nas pontes de hidrogênio, que promovem o pareamento das bases.

Parcimônia energética

Dentre as forças envolvidas na configuração da complexidade dos seres vivos, o princípio da economia de recursos (um tipo de parcimônia) não é generalizadamente vigente. Os organismos podem ser descritos como máquinas de alta eficiência em aproveitamento energético, mas também apresentam aspectos de muito esbanjamento e desperdício.

Distinguem-se dois tipos de percursos evolutivos, como ‘estratégias de permanência’. O princípio fundamental ou constitutivo é de crescer e reproduzir, tanto quanto e sempre que possível. Esse processo pode produzir excessos e desperdícios: grandes números de gametas, sementes e filhos, muitos deles inexoravelmente impossibilitados de permanecer (não somente por defeitos internos ou falta de garantia de condições ambientais para tal, mas também pelo simples excesso - p. ex., o número de gametas masculinos é muito maior que dos femininos, em muitas espécies); muitas duplicações gênicas com amplificação e redundância, às vezes sem funções detectáveis; etc.

O princípio da parcimônia energética prevaleceu nas bactérias, que

seguiram percurso de alta rapidez reprodutiva (ciclos de até cerca de 20 min), o que exige o mínimo de redundância interna e organismos mais simples. Este princípio não vigora nos eucariotos, das leveduras, algas unicelulares e protozoários até os fungos multicelulares, plantas e animais. Seu percurso foi de menor taxa reprodutiva (mínimo de 12 h), genomas grandes e altamente redundantes, com mais tempo disponível e maior exploração das potencialidades regulatórias. Nas bactérias, produz-se grande número de indivíduos variantes, em populações clonais, mantendo a simplicidade dos organismos; nos eucariotos, as espécies acumulam grande número de variantes internas em seus genomas e organismos mais complexos. A seleção natural, nestes, não é muito eficiente em eliminar ‘excessos ou desperdícios’, que podem ser desde ‘neutros’ até benéficos, quanto ao aspecto de aproveitamento regulatório.

Redes neurais combinatórias

A biologia já se apressa em incorporar os avanços desenvolvidos pelas redes neurais computacionais (Gonzales, 1996). Observa-se que os processos regulatórios bioquímicos nos organismos complexos envolvem sinais (como hormônios, neurotransmissores, organizadores morfogênicos, fatores de crescimento, etc) pouco numerosos e, cada um, de efeitos relativamente ambíguos. Esses atuam sobre grande variedade de células-alvo (receptores) e produzem efeitos diversos (conforme o tipo de receptor e conforme sua circunstância ou contexto espacial e temporal).

As especificidades de respostas são obtidas por: (a) controle eficiente e rápido das concentrações dos efetores (intensidade de cada sinal), que são tipicamente de meia-vida curta (rapidamente degradados, instáveis) e produzidos ou liberados em baixa concentração; (b) receptores de alta afinidade pelos sinais; (c) como nas redes neurais artificiais, atuação combinatória de efetores múltiplos sobre cada alvo, cada combinação (em tipos, momentos e intensidades dos sinais) adquirindo especificidade e univocidade de resposta, cada sinal participando de um contexto único.

Formalização shannoniana

Dadas essas características, parece difícil o projeto visando uma formalização shannoniana aplicada à bioquímica. Há subsistemas que poderiam se assemelhar aos shannonianos, com fonte - canal - veículo (sinal) - receptor, como em glândulas endócrinas - circulação em líquidos - hormônios - células-alvo, mas se assemelham mais às redes neurais. Ainda mais, na maioria das situações biológicas, os contatos moleculares são diretos, em 'corpo a corpo', cada elemento interagindo com outro simultaneamente e mutuamente como fonte, sinal e receptor.

Informação restrita no DNA

É interessante relacionar a quantidade de informação estimada como manifestada por um biopolímero com suas potencialidades de assumir diferentes conformações espaciais, desde que são essas as formas funcionalmente ativas. A sequência linear dos polímeros é somente indicativa das conformações atingíveis.

O DNA (**Figura 4**) é caso especial de baixos potenciais, porque suas reatividades (além de serem menores que as do RNA, por não possuir uma hidroxila no açúcar; fator que deve ter sido importante na sua destinação como gene) estão majoritariamente bloqueadas na dupla fita, por complementaridade completa das bases. As proteínas que com ele interagem têm que desenvolver propriedades fortes de penetrar no interior da dupla fita para interagir com as bases, geralmente rompendo os pareamentos destas. Os RNA têm algumas características mais relacionadas com as das proteínas (**Figura 4**).

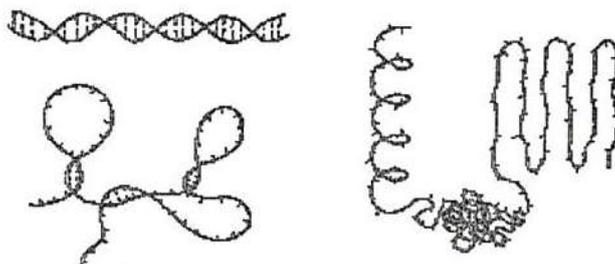


Figura 4. Esquema da estrutura dos biopolímeros.

A estrutura do DNA (acima, à esquerda) é composta por uma dupla fita (hélice) completa, em que as bases estão inteiramente (mas dinamicamente) ocupadas, cada uma com sua complementar, em percurso anti-paralelo. A estrutura do RNA (abaixo, à esquerda) corresponde a uma das fitas do DNA, onde só é possível complementaridade parcial e segmentar interna, com formação de estrutura espacial, tridimensional complexa (hélices e alças). Vários segmentos têm radicais expostos para o ambiente, com reatividade facilitada. Nas proteínas (à direita), os tipos de estruturas básicas são 4: hélices-alfa, ângulos, novelos e fitas-beta, estas podendo ser paralelas ou antiparalelas. A variabilidade e maleabilidade das conformações espaciais segue ordem crescente DNA < RNA < proteínas. As sequências dos radicais dedicados a interações internas definem as conformações espaciais fundamentais e as sequências expostas para fora definem as reatividades interativas com outros elementos, com o ambiente e modulações conformacionais.

Conformações de RNA e proteínas

O número de conformações calculadas para um RNA pequeno (de cerca de 100 bases), é da ordem de 10^6 , enquanto que o de uma proteína de igual tamanho seria de 5×10^{47} . A diferença decorre, principalmente, do tipo de alfabeto que compõe os polímeros, sendo 4 letras no RNA e 20 na proteína. No caso dos RNA mensageiros, o diferencial quanto à cadeia linear é menor, porque usam-se 3 nucleotídeos por aminoácido.

A diferença principal reside, portanto, nas regras de estruturação espacial dos dois polímeros. As possibilidades conformacionais são limitadas para o RNA, porque têm que obedecer as regras do pareamento de bases, de somente 2 tipos, enquanto as proteínas são muito mais plásticas, porque as reatividades dos aminoácidos são mais diversificadas. Isto explica porque estas são adequadas para cumprir a grande diversidade de funções metabólicas, inclusive as de copiar os modelos na replicação e na tradução, e porque os RNAs, ainda que tenham potenciais catalíticos, não os cumprem com grande diversidade.

Informação e as vias biossintéticas

Essas diferenças informacionais (conformacionais dos biopolímeros) estão relacionadas com a complexidade dos sistemas de sua síntese. Os sistemas replicativo e transcritivo são mais simples, envolvendo menor número de proteínas (e seus genes correspondentes, da ordem de dezenas) do que os sistemas de tradução (envolvendo pelo menos uma centena de proteínas e RNAs). A regulação da síntese de proteínas específicas para cada função envolve, também, transcrição específica dos RNAs correspondentes, sendo a complexidade distribuída nos dois passos, mas permanecendo maior na tradução. Na via aferente, das funções orgânicas para a reprodução, estão envolvidos todos os genes e suas interações com o ambiente, em conjunto (**Figura 5**).

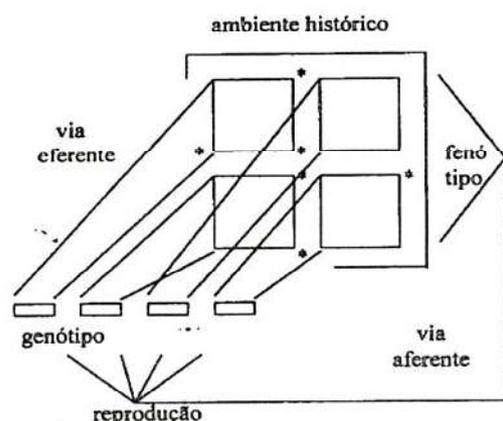


Figura 5a. Os percursos ontogenético e evolutivo dos sistemas biológicos.

Fig. 5a. As vias dos processos informacionais.

A expressão do conjunto genotípico, e suas variações, produz (via eferente, dos genes para as proteínas), espaços fenotípicos interativos (informação) progressivamente ampliados: dos genes para as proteínas, nas redes de interações mútuas entre as proteínas e nas interações do conjunto com o ambiente histórico. Este conjunto retorna (via aferente, do fenótipo para os genes) informação, em forma conjunta e integrada, nos eventos reprodutivos, que será distribuída entre os

genes e suas variantes. A evolução é do conjunto fenotípico (do organismo), que inclui os caracteres genéticos, no contínuo reprodutivo. No desenvolvimento ontogenético, o retorno promove modulação da expressão dos genes. Na evolução através de gerações, a via aferente testa o desempenho das variantes e permite maior crescimento das eficientes, confirmando e atestando a qualidade das boas alternativas.

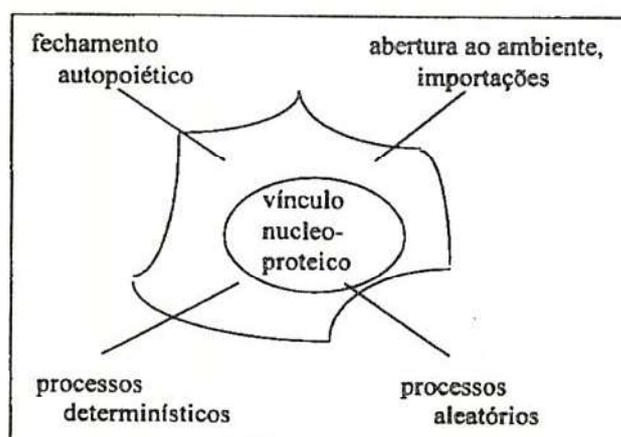


Figura 5b. Processos envolvidos na ontogênese e evolução.

A célula se origina pelo estabelecimento do vínculo nucleoproteico. Os processos aleatórios são exemplificados pela difusão e pelos encontros moleculares brownianos; a difusão é limitada e o funcionamento é garantido pelo enclausuramento parcial do sistema. As altas afinidades químicas, como as especificidades enzimáticas, replicativas e metabólicas, exemplificam os processos determinísticos. O sistema é semi-aberto ao ambiente, do qual capta energia e matéria prima, às vezes importando componentes complexos, como em transferência horizontal de genes, simbioses e parasitismos, e em interações que resultam nas adaptações. O sistema é, também, semi-fechado sobre si mesmo, auto-produzindo-se no contínuo reprodutivo evolucionário, com ampla interação de processos de auto-organização e de seleção darwiniana. Adaptado de Debrun (1996b).

Codificações superpostas na 'linguagem molecular'

A mesma sequência de um biopolímero pode conter várias codificações superpostas, reforçando a sua caracterização como sistêmica e informacional, quanto a origem e funções. Sua leitura ou interpretação, pelo sistema biológico, pode ser feita de maneiras diversas, conforme o contexto e a circunstância de momento.

DNA tem sinais para interação com proteínas arquiteturais do cromossoma, com proteínas reguladoras da replicação e da transcrição, e outros que serão manifestos somente nas suas cópias de RNA e daí por diante na cadeia de expressão gênica. RNA tem sinais para adquirir conformação em hélices e alças, ou estruturas de ordem superior, para interação com proteínas reguladoras de sua estabilidade e de suas múltiplas funções, ou para interação com outros RNAs. Proteínas têm sinais para aquisição de suas estruturas conformacionais internas, para interação com outras proteínas ou com RNA ou DNA, para reação com substratos e efetores de regulação. Muitos sinais são superpostos, condensados no mesmo segmento de sequência, a serem utilizados em momentos e situações diversas. Isto resulta em maior complexidade estrutural 'linguística' dos textos biológicos, do que dos textos humanos (Popov et al, 1996).

Decomposição da informação

Desde que cada elemento, genes ou proteínas, só faz sentido em relação ao sistema hipercíclicamente interativo, não parece ser cabível uma tentativa de quantificação da informação nos elementos, isoladamente, extraídos do contexto funcional, para posterior integração da informação total. Seria possível decompor a informação circulante no interior de um sistema e distribuí-la entre suas partes? Talvez faça mais sentido computar o potencial informacional genérico dos tipos de biopolímeros, conforme indicado acima, ou de algum subsistema, desde que seja fechado e isolado, como na 'biologia em tubos de ensaio', assim obtendo compreensão de partes.

Significados 'a posteriori'

Informação é uma força, uma propriedade relacional, conexional, funcional, interativa de um elemento, que o permite fazer parte de um conjunto funcional. Pode ser caracterizada em vários níveis de complexidade, em sentidos fracos e fortes. Por sua atuação, o sistema sofre modificações funcionais, como consequências e respostas. O significado da informação reside, principalmente, no sistema receptor. Dizer que um sinal contém informação ou instrução é descrição de experiência de observações, *a posteriori*, feitas pelo observador do sistema, ou raciocínio teleológico, referente ao planejador do sistema.

Agradecimentos: CNPQ, FAPEMIG

Referências Bibliográficas

- ALBERTS, B., BRAY, D., LEWIS, J., RAFF, M., ROBERTS, K., WATSON, J. D. *Molecular biology of the cell*. New York: Garland, 1994.
- ATLAN, H. Entre o cristal e a fumaça. Rio de Janeiro: Zahar, 1992. Tradução de: Entre le cristal et la fumée.
- DANCHIN, A. *Une aurore de pierres aux origines de la vie*. Paris: Seuil, 1990.
- DEBRUN, M. Comunicação pessoal, 1996a.
- _____. A idéia de auto-organização. In: DEBRUN M, GONZALES, M. E. Q., PESSOA JÚNIOR, O. (Org.) *Auto-organização*. Campinas: UNICAMP, 1996b. p. 3-23. (Coleção CLE, n. 18).
- _____. A dinâmica de auto-organização primária. In: DEBRUN, M, GONZALES, M. E. Q., PESSOA JÚNIOR, O. (Org.) *Auto-organização*. Campinas: UNICAMP, 1996c. p. 25-59. (Coleção CLE, n. 18).
- DYSON, F. *Origins of life*. Cambridge: Cambridge Univ. Press, 1985.
- EIGEN, M. The origin of genetic information. *Gene* n. 135, p. 37-47, 1993.
- EIGEN, M., SCHUSTER, P. *The hypercycle*. Springer: Berlin, 1979.
- GONZALES, M. E. Q. Ação, causalidade e ruído nas redes neurais auto-organizadas. In: GONZALES M. E. Q., PESSOA JÚNIOR, O. (Org.). *Auto-organização*. Campinas: UNICAMP, 1996, p 3-23. (Coleção CLE, n. 18)
- GUIMARÃES R. C. A base celular dos ciclos de memória. In: ENCONTRO BRASILEIRO INTERNACIONAL DE CIÊNCIAS COGNITIVAS, 1996, Campus de Goytacazes. Org. AMB SOARES. UENF/FENORTE/FAPERJ, P. 17-24, 1996.

- _____. Systemic approaches in genetics. *Ciência e Cultura*, v. 44, P. 314-19, 1992.
- _____. Genética pós-moderna. *Revista Ciências Biomédicas*, v. 14, p. 65-81, 1994a.
- _____. *Procuram-se ilusões*. São Paulo: Paulicéia. 1994b.
- _____. Linguistics of biomolecules and the protein-first hypothesis for the origins of cells. In: PONNAMPERUMA, C, CHELA-FLORES, J., KLUWER, DORDRECHT. *Chemical evolution: structure and model of the first cell*, p. 193-9, 1995. (também *J. Biol. Phys.* 20, 1994).
- _____. Uma estrutura básica da célula. In: MARI H, DOMINGUES I, PINTO J. (Org.). *Estruturalismo - memória e repercussões*. Rio de Janeiro: Diadorim/UFMG, 1996, p. 267-76.
- _____. Pesquisa sobre origem da vida em fase de sondagens. *Ciência Hoje* (Rio de Janeiro), v. 22, n. 127, p. 8-121, 1997a.
- _____. Non-coding sequences original, genes derived: hydrophobic and dinucleotide type specificities of the aminoacyl-tRNA synthetase classes. *Abstr. Int'l Soc. Mol. Evol. Symp.* 'Junk DNA: the role and the evolution of non-coding sequences', Guanacaste, Costa Rica, p 6, 1997c.
- HOLMQUIST, G. P., FILIPSKI, J. Organization of mutations along the genome: a prime determinant of genome evolution. *Trends Evol. Ecol.* 9, p. 65-9, 1994.
- PARDINI, M. I. M. C., GUIMARÃES, R. C. A systemic concept of the gene. *Revista Brasileira de Genética*, v. 15, p. 713-21, 1992.
- PEREIRA JÚNIOR, A., GONZALES, M. E. Q. Informação, organização e linguagem. In: ÉVORA, F. (Org.) *Espaço e tempo*. Campinas: UNICAMP, 1995. p 255-90. (Coleção CLE, n. 15).
- PEREIRA JÚNIOR, A., GUIMARÃES, R. C., CHAVES JR, J. Auto-organização na biologia: nível ontogenético. In: DEBRUN, M., GONZALES, M. E. Q., PESSOA JR, O. (Org.) *Auto-organização*. Campinas: UNICAMP, 1996. p. 239-69. (Coleção CLE, n. 18)
- POPOV, O., SEGAL, D.M., TRIFONOV, E. N. Linguistic complexity of protein sequences as compared to texts of human languages. *BioSystems*, v. 38, p. 65-74, 1996.
- SILVA, J. J. Informação e auto-organização. In: DEBRUN, M., GONZALES, M. E. Q., PESSOA JÚNIOR, O. (Org.) *Auto-organização*. Campinas: UNICAMP, 1996. p 103-28. (Coleção CLE, n. 18).

- SOARES, T. A. Introdução ao estudo das ríbozimas. In: Monografias 5. Ribeirão Preto: Soc. Brasil. Genét., 1997, p 57-156.
- STONIER, T. *Information and the internal structure of the universe*. Londres: Springer, 1990.
- WATSON, J. D., HOPKINS, N. H., ROBERTS, J. W., STEITZ, J. A., WEINER, A. M. *Molecular biology of the gene*. Cummings: Menlo Park, 1988.
- WILLS, P. R. Self-organization of genetic coding. *J. Theor. Biol.* v. 162, p. 267-87, 1993.