

Podstawowe reguły odczytywania informacji genetycznej zawartej w DNA współczesnych organizmów zostały odkryte ponad 50 lat temu i nazwane standardowym kodem genetycznym. Znamy więc wszystkie przypisania aminokwasów do trójek nukleotydów tworzących tzw. kodony i dzięki temu jesteśmy w stanie przetłumaczyć kodujący fragment sekwencji DNA na sekwencję aminokwasów, z których tworzone jest białko. Do dziś naukowcy nie są jednak zgodni co do tego, jak wyglądał ten mechanizm u początku powstawania życia na Ziemi, w jaki sposób i pod wpływem jakich czynników ewoluował oraz dlaczego jest uniwersalny, czyli taki sam dla niemal wszystkich współczesnych organizmów. Wielu badaczy próbowało znaleźć odpowiedzi na te pytania, a największą popularność zdobyły hipotezy: stereochemiczna, koewolucyjna i adaptatywna.

Pierwsza z nich postuluje występowanie pewnych strukturalnych i fizykochemicznych relacji i oddziaływań między fragmentami kwasów nukleinowych (DNA, RNA) i aminokwasami. Do tej pory takie oddziaływania zostały wykryte jedynie w przypadku pięciu z dwudziestu aminokwasów, zatem nie można uzasadnić obecnej struktury kodu genetycznego tylko tą hipotezą.

Hipoteza koewolucyjna zakłada, że kodony w pierwotnej i prostszej wersji kodu genetycznego kodowały tylko kilka aminokwasów oraz, że wraz z ewolucją biochemicznej organizacji komórki nowo syntetyzowane aminokwasy przejmowały kodony używane wcześniej do kodowania aminokwasów będącymi prekursorami do tych nowych w szlakach biosyntetycznych. Ta teoria daje dość prawdopodobny obraz ewolucji kodu genetycznego, ale w dalszym ciągu nie tłumaczy powstania konkretnych par kodon-aminokwas, obecnych w standardowym kodzie genetycznym.

Trzecia hipoteza postuluje, że struktura naszego kodu genetycznego jest optymalna ze względu na odporność na mutacje i błędy na etapie translacji, prowadzące do podmiany aminokwasu w nowo syntetyzowanym białku na inny niż ten oryginalnie kodowany w DNA. Nietrudno zauważyć, że w istocie większość aminokwasów kodowanych jest przez czwórki lub pary kodonów, różniące się tylko nukleotydem na trzecim miejscu w kodonie. Co ciekawe, aminokwasy o podobnych właściwościach fizykochemicznych mają często ten sam nukleotyd na drugiej pozycji w kodonie, więc nawet zmiana pozycji pierwszej nie spowoduje przyłączenia aminokwasu o radykalnie innych cechach i tym samym zaburzenia funkcjonalności syntetyzowanego białka.

Aby sprawdzić zasadność hipotezy adaptatywnej, wielu badaczy przeprowadza obliczenia porównujące odporność standardowego kodu genetycznego z najlepiej zoptymalizowanymi po tym względem teoretycznymi alternatywami. Znalezienie najbardziej odpornych na mutacje kodów i wiarygodna ocena odporności standardowego kodu genetycznego stanowi jednak wyzwanie, ponieważ wiele czynników wpływa na prawdopodobieństwo zmiany kodowanego aminokwasu w białku i powinny one zostać wzięte pod uwagę w modelu. Celem tego projektu jest właśnie stworzenie modelu uwzględniającego wiele parametrów przy ocenie jakości kodu. Zbadamy wpływ wielu czynników na optymalność struktury standardowego kodu genetycznego: presji mutacyjnej (czyli prawdopodobieństw mutacji pojedynczych nukleotydów), używalności kodonów (czyli częstość występowania poszczególnych kodonów w genach), typu mutacji (mutacje punktowe, insercje i delecje), oraz różnych właściwości aminokwasów. Za pomocą tzw. algorytmu genetycznego przeszukamy przestrzeń możliwych teoretycznych alternatyw w celu znalezienia kodów zarówno najlepiej, jak i najgorzej zoptymalizowanych w omawianym sensie i porównamy ich cechy z własnościami kodu standardowego. Da nam to bardziej dokładny obraz tego, jak dobry jest nasz kod na tle innych, niż dotychczasowe porównania jedynie z kodami najlepszymi i losowo wybranymi.

Informacje uzyskiwane za pomocą tego typu badań są bardzo istotne dla naukowców zajmujących się modyfikowaniem kodu genetycznego w żywych organizmach i projektowaniu sztucznych układów biologicznych. Tego typu badania dynamicznie się rozwijają oraz dopracowywane są techniki przypisywania aminokwasów do kodonów i ich wprowadzania do organizmów. Wiedza o tym, które przypisania kodonów są korzystne dla organizmu, a które mogłyby być zmienione dla poprawy danej właściwości, jest bardzo istotna. Dzięki tym technikom będzie można produkować peptydy i białka z nowymi aminokwasami o pożądanym właściwościach oraz konstruować organizmy o nowych cechach.