

Udział białka p53 w długowieczności

Autor: Aleksandra Gomółka

Na podstawie Bartas M, Brázda V, Volná A, Červeň J, Pečinka P, Zawacka-Pankau JE. The Changes in the p53 Protein across the Animal Kingdom Point to Its Involvement in Longevity.

Wstęp

P53, zwane niegdyś „strażnikiem genomu” to białko o aktywności czynnika transkrypcyjnego, które wiążąc się z DNA w obszarze promotorowym moduluje ekspresję wielu genów. Białko p53 jest zaangażowane w regulację różnych procesów komórkowych, a w szczególności aktywację mechanizmów naprawy DNA lub indukcję apoptozy w odpowiedzi na uszkodzenia. Dzięki swym właściwościom uznawane jest za kluczowy supresor guza. Mutacje w genie kodującym białko p53 są obserwowane w dużym odsetku nowotworów i korelują z niekorzystnym rokowaniem. Uznaje się, że poprzez funkcję „tłumienia” nowotworów, p53 odgrywa rolę regulatora długowieczności u różnych organizmów. Ostatnie badania sugerują, że p53 może regulować starzenie się i długowieczność, niezależnie od funkcji supresora nowotworowego¹.

Coraz więcej dowodów wskazuje na to, że aktywność p53 może odgrywać kluczową rolę w procesie starzenia, jednak niewiele wiadomo na temat molekularnych czynników przedłużonej długości życia w królestwie zwierząt, w tym ludzi. Z tego względu, założeniem przeprowadzonej analizy było zbadanie związku między długowiecznością (maksymalną długością życia) a sekwencjami białka p53. W tym celu porównano sekwencje białka p53 długowiecznych zwierząt z sekwencjami pochodzącymi od organizmów o krótkiej długości życia. Dokonano, także porównania sekwencji tego białka z sekwencjami pochodzącymi od dwóch organizmów, które zostały uznane za najciekawsze w kontekście długowieczności i wysokiej witalności (*Heterocephalus glaber* oraz *Balaena mysticetus*). Wybrane sekwencje poddano analizie przewidywania struktury białek, a także zwrócono uwagę na ich pochodzenie i zmiany ewolucyjne.

¹ Feng Z, Lin M, Wu R. The Regulation of Aging and Longevity: A New and Complex Role of p53. Genes Cancer. 2011 Apr;2(4):443-52. doi: 10.1177/1947601911410223. PMID: 21779512; PMCID: PMC3135645.

Materiały i metody

Dane dotyczące długowieczności i maksymalnej długości życia uzyskano z bazy danych AnAge (ang. The Animal Ageing and Longevity Database)^{2,3}. Wybrane sekwencje pobrano z bazy danych NCBI protein oraz UniProt. Przeszukania dokonano za pomocą kwerend, określających naukową nazwę szukanego organizmu, a także nazwę białka (cellular tumor antygen p53 lub p53). W przypadku organizmów, których homologii p53 nie były obecne w ww. bazach danych, sekwencje pobrano z zasobów zgromadzonych w ramach badań, przeprowadzonych przez Bartas i wsp⁴. W znacznej większości sekwencje pochodziły z genu TP53, jednak w analizie uwzględniono, również inne geny. Przy braku dostępności pełnej sekwencji białka wykorzystano izoformy p53, które także mogą odgrywać istotną rolę w modulowaniu długowieczności. Ostatecznie w analizie użyto 136 sekwencji, pochodzących zarówno od organizmów uznanych za długowieczne, jak i o krótkim czasie życia. Dalszą część badania przeprowadzono z wykorzystaniem programu MEGA, który umożliwia wykonanie pełnej analizy filogenetycznej. Wyrównania miejsc homologicznych sekwencji dokonano za pomocą algorytmu MUSCLE (ustawienia domyślne), który jest preferowaną i uznawaną za najbardziej dokładną metodę dopasowania^{5,6}. Długość sekwencji po wyrównaniu wynosiła ok. 1000 pz. Następnie do skorygowania wielokrotnych podstawień aminokwasowych, występujących podczas ewolucji sekwencji, wybrano odpowiedni model substytucji. W tym celu wykorzystano wbudowaną funkcję programu MEGA, która wybiera i dostosowuje model do analizowanych danych – Find Best DNA/Protein Models. Ostatecznie, analizę przeprowadzono za pomocą modelu substytucji - JTT+G. Analizę najbardziej wiarygodnej wartości parametru G dla wybranych sekwencji przeprowadzono z wykorzystaniem funkcji – Estimate Gamma Parameter for Site Rates programu MEGA i uzyskano wynik równy ok. 0.68. Drzewo filogenetyczne zostało skonstruowane za pomocą metody Neighbor Joining (Jones-Taylor-Thornton model, Gamma Distribution: 0.68, Pairwise deletion, Number of threads: 3). Parametr *pairwise deletion* pozwala na usunięcie występujących w sekwencjach luk lub brakujących danych, powstałych w wyniku dopasowania. Decyzja o usunięciu luki jest podejmowana na podstawie porównań parami. Do oszacowania wiarygodności drzewa wykorzystano metodę bootstrap z liczbą iteracji/powtórzeń równą 100. Następnie wyznaczono parametr OMD, określający średnią odległość między dwiema sekwencjami (ang. overall mean distance). Dokonano tego za pomocą funkcji programu MEGA – Compute Overall Mean Distance i uzyskano wynik ok. 0.86.

² <https://genomics.senescence.info/species/index.html>

³ DE MAGALHÃES, J.P. and COSTA, J. (2009), A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, 22: 1770-1774

⁴ Bartas M, Brázda V, Volná A, Červeň J, Pečinka P, Zawacka-Pankau JE. The Changes in the p53 Protein across the Animal Kingdom Point to Its Involvement in Longevity. *Int J Mol Sci*. 2021 Aug 7.

⁵ Barry G. Hall, Building Phylogenetic Trees from Molecular Data with MEGA, *Molecular Biology and Evolution*, Volume 30, Issue 5, May 2013, Pages 1229–1235

⁶ Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113 (2004).

W dalszej kolejności przeprowadzono predykcję struktury drugorzędowej białka dla trzech wybranych gatunków (*Heterocephalus glaber*, *Balaena mysticetus* oraz *Homo Sapiens*). Do odtworzenia sekwencji ancestralnej wykorzystano funkcję Infer Ancestral Sequences, dostępną w programie MEGA. Predykcję struktur białkowych wykonano za pomocą programu AlphaFold⁷ oraz SWISS-MODEL⁸. Jako uzupełnienie analizy przygotowano wykresy przedstawiające długość życia różnych gatunków, a także cechy fenotypowe (wygląd) przedstawicieli każdej z grup.

Wyniki

Początkowym etapem analizy była weryfikacja, czy u wybranych gatunków o zwiększonej długości życia obserwujemy występowanie zmian w sekwencji p53. To podejście pozwoliło na wstępną weryfikację i ocenę zasadności prowadzonych badań. Szczególną uwagę zwrócono na wieloryba grenlandzkiego (*Balaena mysticetus*). *Balaena mysticetus*⁹ jest najdłużej żyjącym zwierzęciem w wybranym do analizy zbiorze danych. Szacuje się, że wieloryb grenlandzki żyje ponad 200 lat, co czyni go najdłużej żyjącym ssakiem. Zwierzęta te pozostają wolne od chorób znacznie dłużej niż ludzie, co w szczególności w kontekście raka, sugeruje posiadanie mechanizmów przeciwnowotworowych. W związku z tym poznanie cech zapewniających długowieczność temu gatunkowi dostarczyłoby kluczowych zasobów i wiedzy do poprawy ludzkiego zdrowia i życia.

Do analizy wykorzystano 9 gatunków z rzędu Cetacea (walenie)(Figura 1). *Balaena mysticetus* wykazuje znacznie dłuższy czas życia, a analiza sekwencji p53 potwierdziła występowanie substytucji leucyny w regionie bogatym w prolinę¹⁰ (Figura 2). Inne zmiany nie były unikalne dla tego gatunku. Jednak, aby stwierdzić czy taka modyfikacja może wpłynąć na aktywność białka, należałoby przeprowadzić bardziej szczegółowe badania funkcjonalne. Wspomniana substytucja leucyny odpowiada reszcie aminokwasowej nr 77 w ludzkim (*Homo sapiens*) wariantcie tego białka (Figura 3).

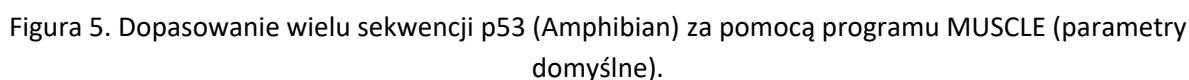
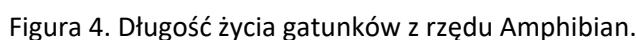
⁷ <https://alphafold.ebi.ac.uk/>

⁸ <https://swissmodel.expasy.org/interactive>

⁹ <http://www.bowhead-whale.org/>

¹⁰ <http://www.e-biotechnologia.pl/artykuly/bialko-p53>

Do analizy wykorzystano 6 gatunków z rzędu Amphibian (płazy)(Figura 4). Analiza sekwencji p53 pochodzącej od *Proteus anguinus* wykazała występowanie insercji dwóch aminokwasów – seryny i argininy w centralnej domenie¹¹, wiążącej się z DNA (Figura 5). Modyfikacja odpowiada insercji po aminokwasie L188 w ludzkim p53 (Figura 6). Dodatkowo, wykazano występowanie innych, unikalnych modyfikacji takich jak: L21F, P26H, D358_N359insH. Ponownie, aby stwierdzić czy takie zmiany mogą wpłynąć na aktywność białka, należałoby przeprowadzić bardziej szczegółowe badania.



5

[illegible]

Figura 6. Dopasowanie sekwencji p53 (*Proteus anguinus* vs *Homo sapiens*) za pomocą programu MUSCLE (parametry domyślne).

U gatunków z rzędu Chiroptera, również wykazano szczególne zmiany w sekwencji p53 dla najdłużej żyjących organizmów *Myotis lucifugus* oraz *Myotis brandtii* (Figura 7). Blisko spokrewnione nietoperze objawiały szczególny układ z insercją siedmiu reszt aminokwasowych w centralnej domenie wiążącej DNA¹² (po aminokwasie nr 295 w ludzkiej sekwencji p53)(Figura 8). Zaobserwowano, również występowanie licznych, unikalnych dla tych dwóch gatunków substytucji pojedynczych aminokwasów oraz polimorfizmów typu indel (np. L76_C77insG, D64_L66delV). Porównując sekwencję p53 pochodzącą od *Myotis lucifugus* oraz *Myotis brandtii* z sekwencją *Homo sapiens* wykazano podobną substytucję 7 aminokwasów (P300_P301insPGSPKQP), również znajdującą się w regionie wiążącym (Figura 9). Możemy sądzić, że występujące zmiany mają wpływ na funkcjonowanie białka, jednak aby to potwierdzić konieczne jest przeprowadzenie bardziej szczegółowych badań.

¹² <http://www.e-biotechnologia.pl/artykuly/bialko-p53>

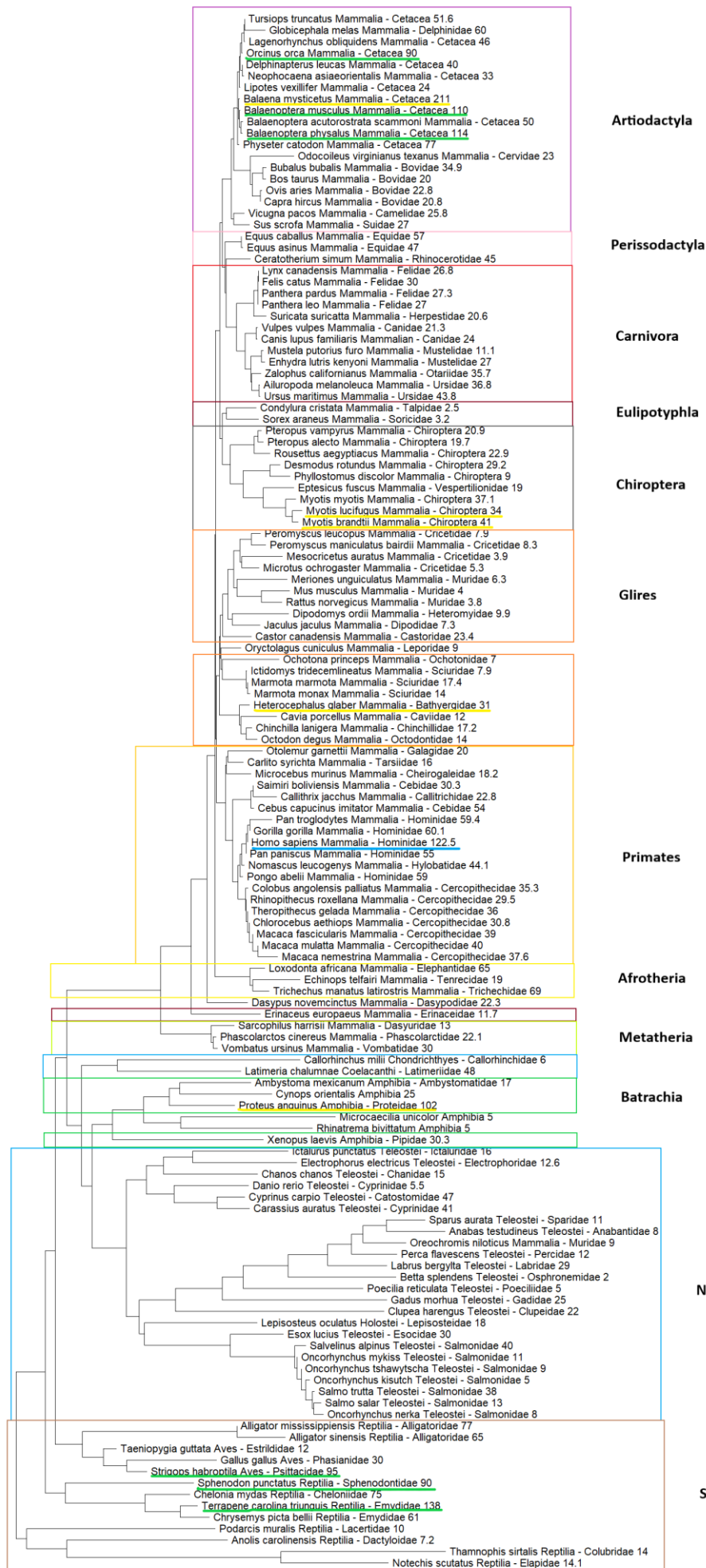
Dokonano porównania ludzkiej sekwencji p53 z sekwencją pochodzącą od *Heterocephalus glaber* w celu identyfikacji zmian mających możliwy wpływ na zwiększoną odporność i przedłużenie czasu życia (Figura 10).

[illegible]

Analiza wybranych, długowiecznych gatunków doprowadziła do wniosku, że sekwencja aminokwasowa białka p53 może być związana z długością życia. Dlatego kolejnym etapem było przeprowadzenie analizy filogenetycznej, pozwalającej na korelację sekwencji aminokwasowej p53 z długością życia dla większego zbioru danych.

¹³ <http://www.naked-mole-rat.org/>

8



Zestaw danych z sekwencjami p53 został podzielony na kilka grup, na podstawie relacji filogenetycznych. W znacznej większości sekwencje grupują się zgodnie z taksonomią, ale nie wszystkie. Wskazuje to na możliwe występowanie kilku równoległych procesów ewolucyjnych, prowadzących do zmodyfikowanej aktywności p53. Topologia drzewa nie pozwala jednak na jednoznaczne określenie związku sekwencji białka z długością życia. Nie zaobserwowano wyraźnego oddzielenia gatunków uznanych za długowieczne od tych o krótkim czasie życia. Układ drzewa przedstawia jedynie ewolucyjne zależności pomiędzy gatunkami. Należy pamiętać, jednak że otrzymane drzewo filogenetyczne jest oszacowaniem rzeczywistych relacji między taksonami, a ich przodkami. Mimo bliskiego pokrewieństwa niektórych gatunków wykazano znaczne różnice w długości życia, co stanowi podstawę do przeprowadzenia dalszych badań nad korelacją maksymalnej długości życia ze zmianami w sekwencji p53.

Do odtworzenia sekwencji ancestralnej wybrano grupę zewnętrzną (ang. outgroup), składającą się z gatunków wchodzących w skład najbardziej odległego taksonu (*Thamnophis sirtalis*, *Notechis scutatus*, *Anolis carolinensis*, *Podarcis muralis*). Poniżej przedstawiono wybrane sekwencje ancestralne (z najmniejszą liczbą przerw), które zostały zaproponowane przez program MEGA.

Loxodonta africana Mammalia - Elephantidae 65:

MFSINSTLAALVCRTSPPQNPGLSLRSLLFHSLASPLPTGKLLALTCHGDCPALCQKPRGGCWDWEFPFPCAHTGAKSFQL

Rhinatrema bivittatum Amphibia 5:

DVMGVPTDRFRAPDKNDSCAGGLESLWTAMDLYPEDADTHPFGDLFPCCLETVFSEEGHCSTPELSLACCCPEGHPPALAM

W następnym kroku przeprowadzono predykcję struktury białka p53 dla *Heterocephalus glaber*, *Balaena mysticetus* oraz *Homo Sapiens* w celu identyfikacji zmian mających możliwy wpływ na funkcjonowanie i aktywność produktu białkowego. Dokonano tego za pomocą programu AlphaFold, dostępnego online.

AlphaFold to system sztucznej inteligencji, który przewiduje trójwymiarową strukturę białka na podstawie jego sekwencji aminokwasowej. AlphaFold przyznaje punktację ufności na resztę (pLDDT), mieszczącą się między 0 a 100. Regiony poniżej 50 pLDDT mogą być uznawane za nieustrukturyzowane. Program pozwala, również na określenie przewidywanego błędu predykcji struktury (Figura 12, Figura 14, Figura 16). Odcień zieleni na wykresach PAE (ang. Predicted Aligned Error) wskazuje na oczekiwany błąd odległości (w jednostce Å). Kolor w punkcie (x, y) odpowiada oczekiwanemu błędowi odległości w pozycji reszty x, gdy przewidywanie i prawdziwa struktura są wyrównane z resztą y. Ciemnozielony kolor określa niski błąd, natomiast jasnozielony kolor wysoki błąd.

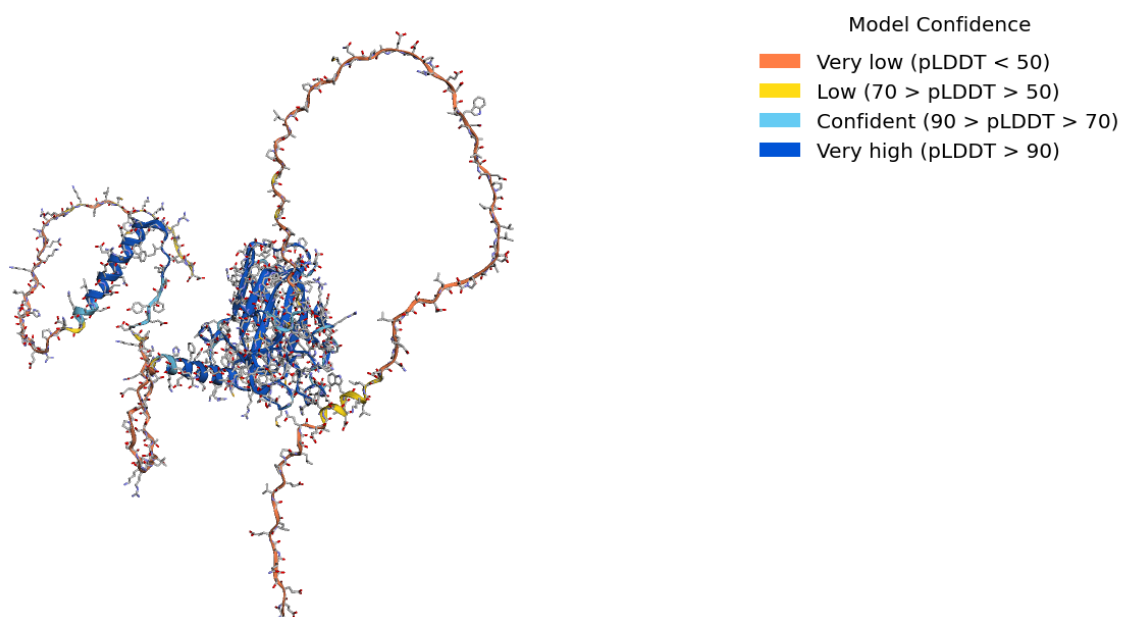


Figura 11. Przewidywana struktura p53 dla *Balaena mysticetus*.

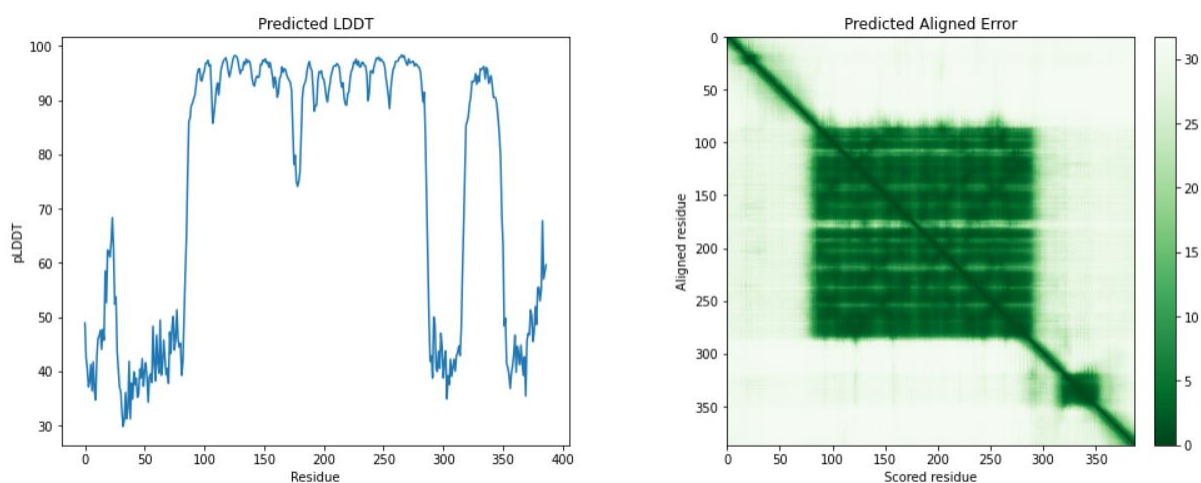


Figura 12. Przewidywany błąd wyrównania struktury p53 dla *Balaena mysticetus*.

*Parametr LDDT (pLDDT) najlepiej nadaje się do określania pewności wewnątrz domeny, podczas gdy przewidywany błąd wyrównania (PAE) najlepiej nadaje się do określania pewności między domenami lub między łańcuchami.

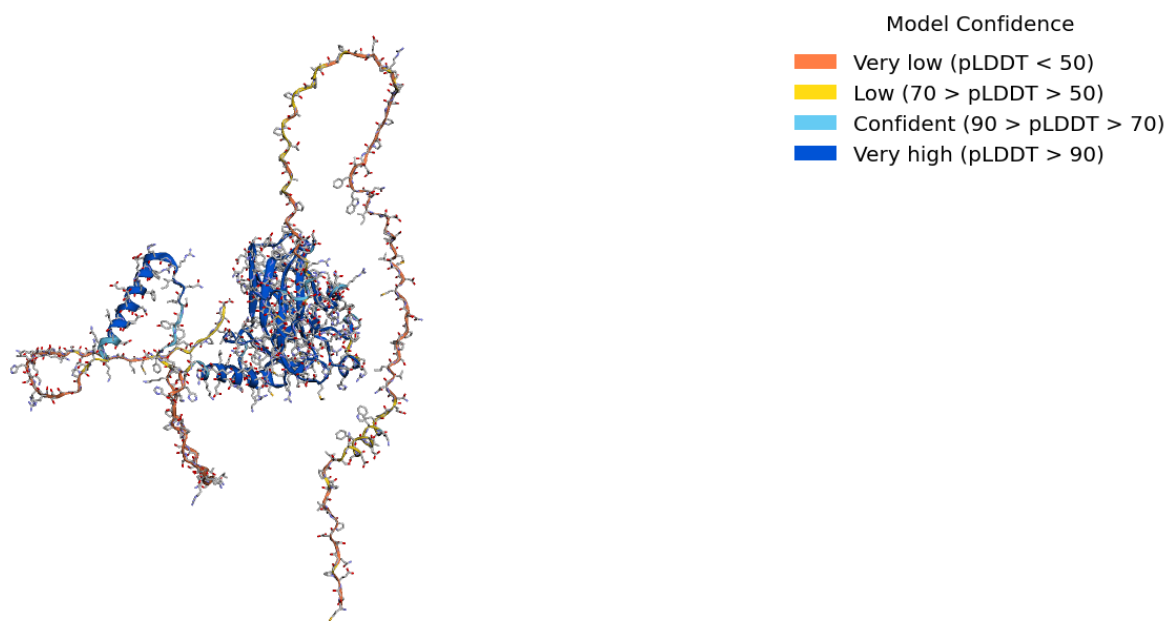


Figura 13. Przewidywana struktura p53 dla *Homo sapiens*.

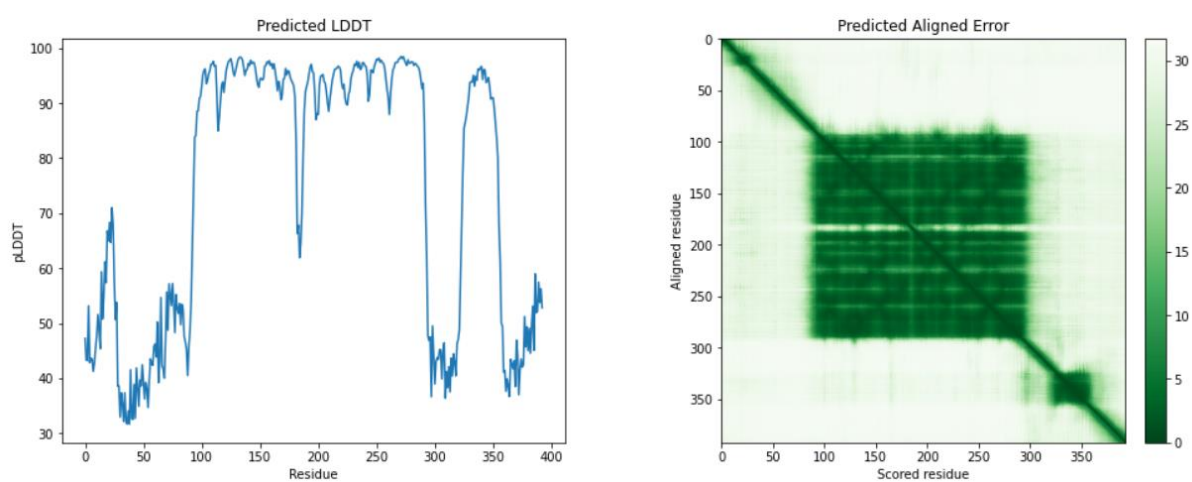


Figura 14. Przewidywany błąd wyrównania struktury p53 dla *Homo sapiens*.

*Parametr LDDT (pLDDT) najlepiej nadaje się do określania pewności wewnątrz domeny, podczas gdy przewidywany błąd wyrównania (PAE) najlepiej nadaje się do określania pewności między domenami lub między łańcuchami.

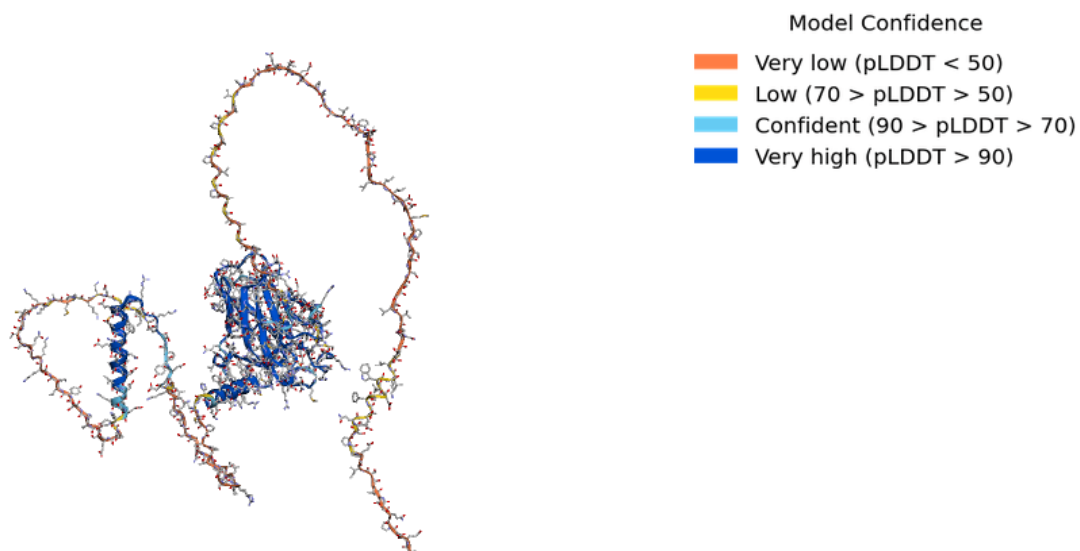


Figura 15. Przewidywana struktura p53 dla *Heterocephalus glaber*.

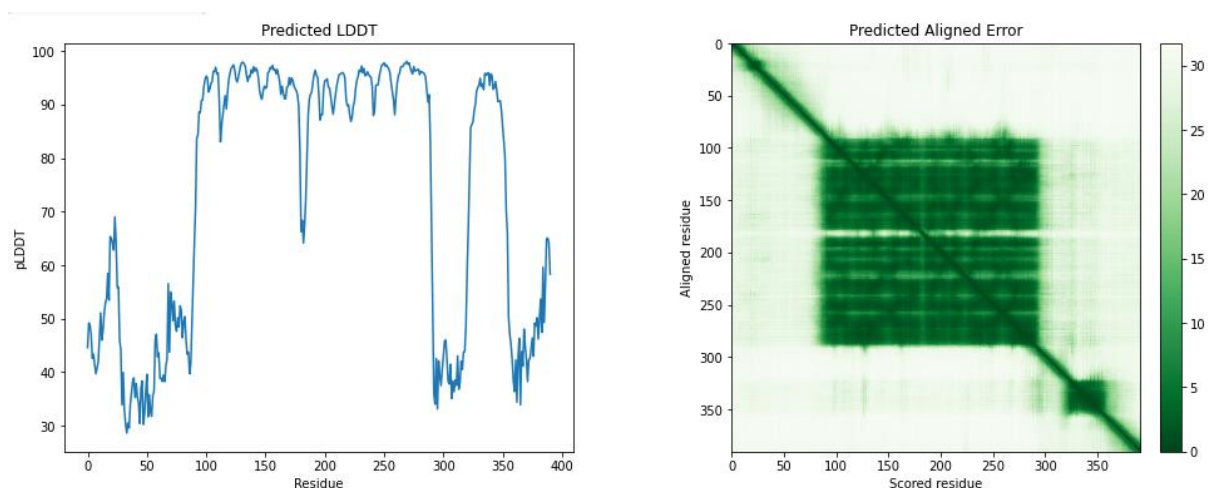


Figura 16. Przewidywany błąd wyrównania struktury p53 dla *Heterocephalus glaber*.

*Parametr LDDT (pLDDT) najlepiej nadaje się do określania pewności wewnątrz domeny, podczas gdy przewidywany błąd wyrównania (PAE) najlepiej nadaje się do określania pewności między domenami lub między łańcuchami.

Przewidywane struktury p53 dla ww. gatunków wykazują względne podobieństwo. Szczególne podobieństwo obserwujemy w rejonach o największej wiarygodności (kolor niebieski), co potwierdza konserwatywność domeny rdzeniowej p53. Największe zmiany zaobserwowano w regionach uznanych za najmniej wiarygodne.

Wykorzystując serwer SWISS-MODEL¹⁵ dokonano modelowania struktur 3D tych białek (Figura 17). Modelowanie homologiczne, zwane również modelowaniem porównawczym białka, polega na konstrukcji modelu białka na podstawie jego sekwencji aminokwasowej i eksperymentalnej struktury powiązanego białka homologicznego. Takie podejście pozwala na zdobycie informacji m.in. o występowaniu konserwatywnych reszt, które mogą stabilizować fałdowanie, uczestniczyć w wiązaniu małych cząsteczek i ligandów lub asocjować z innymi białkami¹⁶. Ponownie obserwujemy wysokie podobieństwo w konserwatywnych regionach rdzeniowych, odpowiedzialnych za wiązanie p53 do DNA. Należy pamiętać, jednak że uzyskane struktury nie odzwierciedlają rzeczywistego wyglądu białek, są jedynie predykcją możliwej reprezentacji.

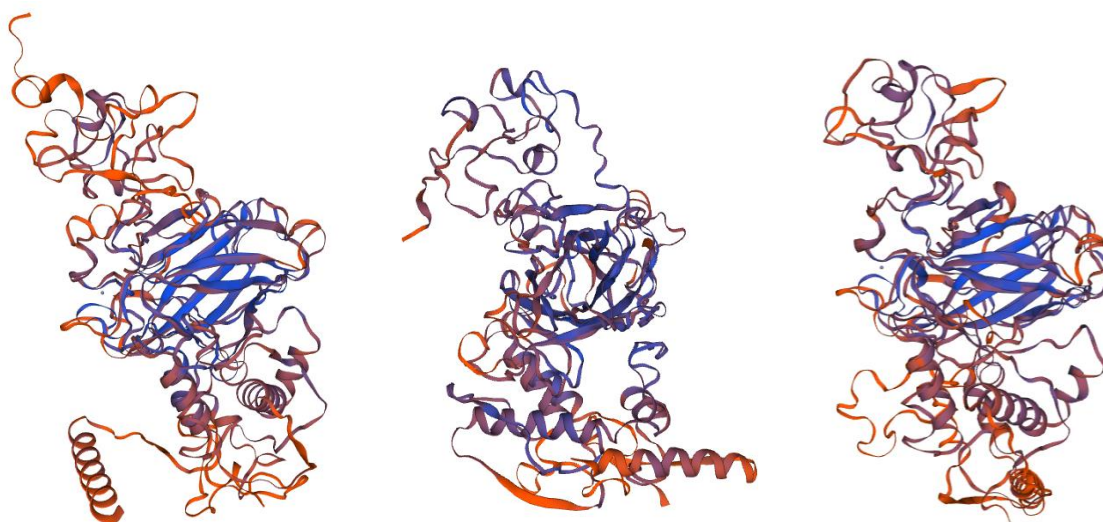


Figura 17. Przewidywana struktura p53, kolejno dla *Balaena mysticetus* (86.01% seq. identity), *Homo sapiens* (100% seq. identity), *Heterocephalus glaber* (85.84% seq. identity).

Wnioski i dyskusja

Białko p53 odgrywa istotną rolę w kontroli cyklu komórkowego, chroni przed stresem komórkowym, a także wpływa na stabilność genomu. Zmieniona funkcjonalność zmutowanego białka p53 przyczynia się do proliferacji nieprawidłowych komórek i w konsekwencji powstawania transformacji nowotworowych. Jednak rola zmian sekwencji aminokwasowej p53 w królestwie zwierząt nie była systematycznie badana. Przeprowadzona analiza ujawniła unikalne zmiany sekwencji tego białka dla kilku długowiecznych gatunków, w tym *Myotis brandtii*, *Myotis lucifugus*, *Balaena mysticetus*, *Heterocephalus glaber* i *Proteus anguinus*, które nie były obecne u ich bliskich krewnych, wykazujących krótką żywotnością. Jednak, aby wnioskować o korelacji określonych odmian białka p53 (ze zmienioną aktywnością, bądź funkcją) z długością życia, należałoby przeprowadzić dalszą

¹⁵ <https://swissmodel.expasy.org/interactive>

¹⁶ <https://en.wikipedia.org/wiki/Swiss-model>

eksplorację sekwencji p53 u gatunków wykazujących długowieczność, a także dokonać statystycznej oceny występujących modyfikacji.

Warto, również zwrócić uwagę na ograniczenia wykonanej analizy, mające wpływ na ostateczne wnioskowanie o wpływie p53 na długowieczność wybranych gatunków. Świadomość istnienia możliwych zmian i ulepszeń pozwala na przeprowadzenie dokładniejszej analizy, by z większą pewnością odpowiedzieć na postawione pytanie badawcze.

Przyrównanie sekwencji za pomocą algorytmu MUSCLE pozwoliło na zidentyfikowanie regionów podobieństwa między badanymi sekwencjami. W sekwencjach zostały wprowadzone luki, które reprezentują historyczne insercje i delecje, a ich celem jest wyrównanie miejsc homologicznych w tej samej kolumnie. Przeprowadzone dopasowanie jest jedynie oszacowaniem pozycji insercji i delecji. Sposób wykonanego przyrównania, wybór odpowiedniego algorytmu dopasowania, a także metoda traktowania przerw (ang. gaps) ma wpływ na jakość drzewa filogenetycznego. Dlatego istotne jest dokonanie oceny jakości wyrównania.

W kolejnym kroku dostosowano model substytucji do analizowanych danych. Wybrany model JTT zakłada, że wszystkie pozycje (w naszym przypadku aminokwasy) ewoluują w tym samym tempie, a tempo ewolucji jest stałe w czasie. Model Juke-Cantora ma jednak pewne ograniczenia. Niektóre rodzaje zmian mogą występować z większą częstością, w porównaniu do innych. Aby uwzględnić występowanie takich różnic należałoby wykorzystać inne modele podstawień aminokwasowych np. model Kimury.

Uzyskany parametr OMD (0.86) wskazuje, że sekwencje w wybranym do analizy zbiorze danych są zróżnicowane. Mniejsza wartość parametru OMD wskazuje na znaczne podobieństwo sekwencji. Parametr ten można wykorzystać do porównania różnorodności genetycznej w obrębie populacji lub między różnymi populacjami, ale również do oceny dokładności rekonstrukcji filogenetycznych. Jeśli ogólna średnia odległość między sekwencjami jest duża, może to wskazywać na mniej dokładne rekonstrukcje, ponieważ sekwencje bardziej się od siebie różnią. Należy zauważyć jednak, że średnią odległość należy interpretować w kontekście konkretnego analizowanego zbioru danych, a także stawianych pytań badawczych. Przy przyrównaniu sekwencji z różnych gatunków możemy oczekiwać większych średnich odległości.

Do konstrukcji drzewa filogenetycznego wykorzystano całe, dostępne sekwencje p53. Jednak ze względu na niskie podobieństwo między domenami N-końcowymi i C-końcowymi p53 u różnych gatunków, a także wykazane mutacje w domenie wiążącej DNA, należałoby ograniczyć analizę do konserwatywnej domeny rdzeniowej. Takie podejście pozwoliłoby na odtworzenie bardziej prawdopodobnych relacji ewolucyjnych między gatunkami.

W normalnych komórkach poziom białka p53 jest niski. Uszkodzenia DNA i inne sygnały stresowe mogą powodować wzrost poziomu tego białka (w celu zatrzymania proliferacji komórek, naprawy DNA, apoptozy). Stężenie komórkowe p53 musi być ściśle regulowane. Wywołany przewlekłym stresem, wysoki poziom p53 może przyspieszać proces starzenia, co zostało zaobserwowane w przypadku niektórych gatunków o krótkiej długości życia ¹⁷. Długowieczne zwierzęta mogły, więc rozwinąć „ulepszone”, ale mniej aktywne białka p53, które nadal mogą przyczyniać się do naprawy uszkodzeń DNA i apoptozy. Prawdopodobne jest także, że zmiany aminokwasów w sekwencji p53 u długowiecznych gatunków umożliwiały interakcję tego białka z innymi białkami lub kwasami nukleinowymi w celu indukowania programów ekspresji genów, różnych od tych występujących u gatunków o normalnej długości życia. Niewątpliwe jest, również występowanie innych mechanizmów ochrony przed chorobami, a także starzeniem się komórek, które nie są bezpośrednio związane z aktywnością p53. We wnioskowaniu o długowieczności należałoby, także uwzględnić adaptacyjne zmiany, pozwalające organizmom na ucieczkę przed maszyną starzenia. W związku z tym, istnieje potrzeba dalszych badań nad mechanizmem zmian w sekwencji p53, aby określić ich wpływ na opóźnienie starzenia i długowieczność.

¹⁷ Bartas M, Brázda V, Volná A, Červeň J, Pečinka P, Zawacka-Pankau JE. The Changes in the p53 Protein across the Animal Kingdom Point to Its Involvement in Longevity. *Int J Mol Sci*. 2021 Aug 7.