Assessing Viable Culinary Alternatives to Salmon With Evolutionary Relationships

Owen Liu[1], Sai Chung[1]

The evolutionary relationships among various aquatic species were examined to seek viable alternatives to salmon in the culinary industry. The protein samples of five different aquatic species (A, B, C, D, E), along with a Precision Plus Kaleidoscope Protein Ladder and actin myosin standard were run through SDS-PAGE. The bands produced on the gel indicated that Sample C had a consistently low percentage similarity (mean = 25.44%, STDEV = 8.58%) with the other samples, while Sample D had the percentage greatest similarities with the other samples (mean: 40.31%, STDEV = 23.00%). Computational analysis was performed on the α-actin amino acid sequences of *S. salar, T. albacares*, *O. vulgaris*, *H. hippoglossus*, and *O. massambicus*, suggesting the a link between salmon and tilapia species. The efficacy of a particular species in place of salmon was not supported and further studies should be conducted to localize certain proteins.

INTRODUCTION

Salmon is a delicacy in many types of cuisine around the world, but many questions have been raised in terms of its harvestable sustainability. Industrial salmon farms are highly susceptible to disease and parasites, while many species of wild salmon are endangered. (1) (2)

To assess the suitability of different types of aquatic organisms as a substitute for salmon, tracing evolutionary relationships could lend valuable insight into the selection process. Evolution is the cumulative change of heritable traits within biological populations. (3) Ecological niches pertain to the interrelationships a species has between the biotic and abiotic factors of an ecosystem. (4) Species with more similar ecological niches would share features such as environmental conditions and food sources required to survive. Speciation, the divergence of a population into groups that are unable to reproduce and create fertile offspring, is heavily driven by the differentiation of ecological niches. (5) As a result, species with closer evolutionary relationships may have more similar ecological niches and would be sustained in a similar way.

Cladograms are character state change models used to trace evolutionary histories. (6) Constructing a cladogram may provide a visual representation of the species most closely related to salmon. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a technique that separates proteins by molecular weight, allowing for the biological relatedness between different organisms to be analyzed. (7) This study aimed to identify whether examination of the evolutionary relationships between various aquatic species with SDS-PAGE could be used to evaluate the efficacy of potential culinary alternatives to salmon.

MATERIALS AND METHODS

250 µL of Laemmli sample buffer was pipetted into each of the five microcentrifuge tubes. 1.7 ± 0.5 g of protein flesh from each sample species (A, B, C , D , E) was transferred individually into each microcentrifuge tube. The microcentrifuge tubes were capped, agitated, and left to incubate for five minutes at room temperature. The buffer, exclusively, from each microcentrifuge tube was poured into another five separate screw cap tubes. The five screw cap tubes along with a tube containing the actin and myosin standard were placed into a 95 °C water bath for five minutes. 10 µL of buffer from the tubes including the actin and myosin standard, as well as 5 µL of Precision Plus Kaleidoscope Protein Ladder was pipetted into individual lanes of a Mini-PROTEAN® TGX™Precast Gel that was contained in a vertical electrophoresis chamber filled with 1x TGS buffer. The gel ran on a 300 V power supply for 18 minutes and was then removed from the chamber and transferred to a staining tray. The gel was rinsed and stained with Bio-Safe Coomassie stain for one hour.

[1]Biotech SHSM, A.Y. Jackson Science Department

RESULTS

The results of the Mini-PROTEAN® TGX™Precast Gel after electrophoresis from three different groups with the same samples were noted. In each gel, blue bands were observed to have travelled away from the anode of the electrophoresis chamber at varying distances, thicknesses, and quantities.

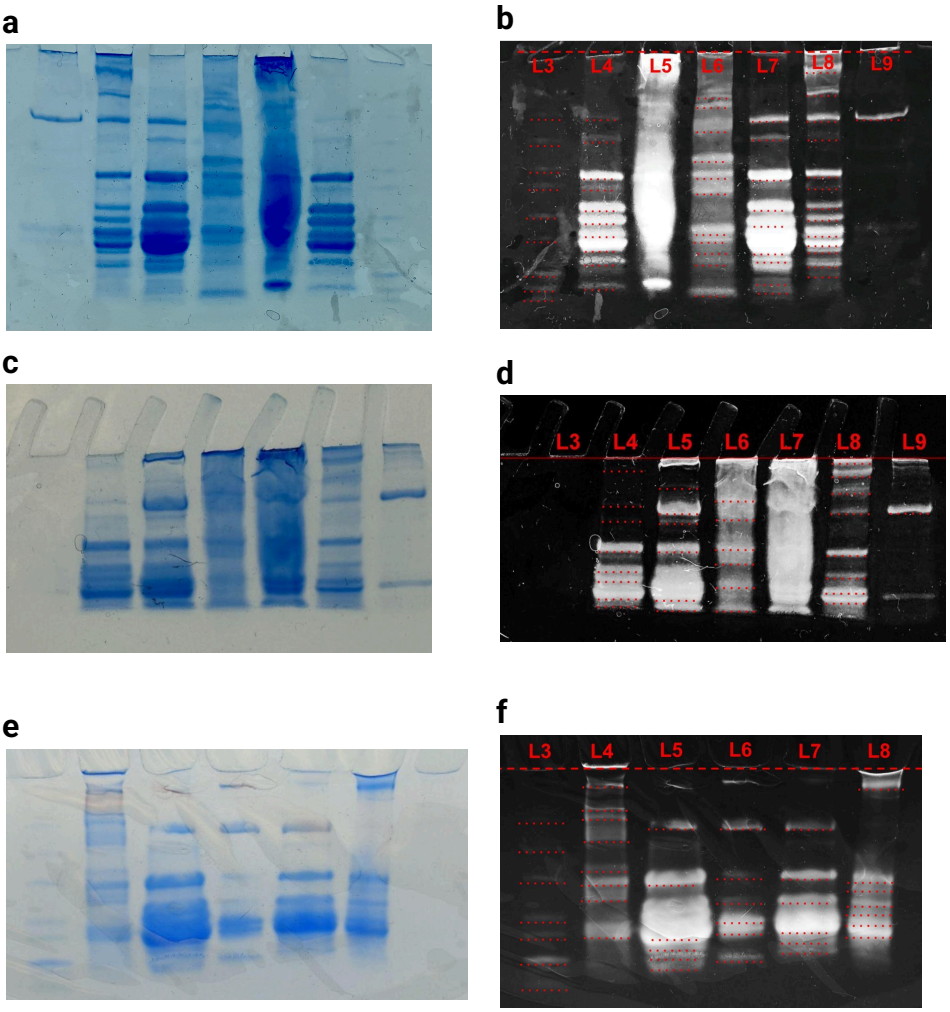

Figure 1 | Resulting Mini-PROTEAN® TGX™Precast Gels with protein samples following electrophoresis and staining. a, Original photo of Group 3 gel. L3 contains the Precision Plus Kaleidoscope Protein Ladder, L4 contains Sample A, L5 contains Sample B, L6 contains Sample C, L7 contains Sample D, L8 contains Sample E, and L9 contains the actin myosin standard. **b,** Edited photo of Group 3 gel, with annotated bands. Each lane produced some type of visible staining. Bands in L5 were indistinguishable and unmarked. **c,** Original photo of group 6 gel. Lane designations same as the Group 6 gel. **d,** Edited photo of Group 6 gel, with annotated bands. Bands in L7 were indistinguishable and unmarked, and no bands were observed in L3. **e,** Original photo of Group 4 gel. L3 contains the Precision Plus Kaleidoscope Protein Ladder, L4 contains Sample C, L5 contains Sample A, L6 contains Sample B, L7 contains Sample D, and L8 contains Sample E. **f,** Edited photo of the Group 4 gel. No bands were observed from the actin myosin standard lane, so L9 was omitted. All other lanes produced distinguishable bands.

Some lanes within the gels did not produce any visible bands. Particularly, the Precision Plus Kaleidoscope Protein Ladder in the Group 6 gel and the actin myosin standard in the Group 4 gel lacked any observable staining. Lanes with insufficient contrast to outline the edges of different bands (L5 in Group 3 gel, L7 in Group 6 gel) were also omitted from further analysis. A notably thick band was observed in L5 (Sample B) of both the Group 6 and the Group 4 gels.

The retention factor (R_f) of a band is the ratio between the distance traveled by the band of concern and the distance traveled by the furthest migrated band within the gel. To standardize the measured results between the different gels, the R_f of each band was calculated. In both the Group 3 and Group 4 gel, the furthest traveled band was observed in L3, the Precision Plus Kaleidoscope Protein Ladder, while the furthest traveled band in Group 6 was observed in L6. Since the Group 6 gel lacked visible bands from the ladder lane, the retention factors were then corrected by corresponding the most prominent band produced by the actin myosin standard to the same band observed in the Group 3 gel. The mean R_f of a band corresponding to a specific lane was tabulated. If the presence of a band in a lane was not supported by at least two separate gels, the R_f for that band was discarded from further analysis.

a

$$R_f = \frac{\text{distance traveled}}{\text{furthest traveled}}$$

b

$$\mu = \frac{\Sigma R_f}{N}$$

c

$$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}}$$

$$2SE_x = 2 \cdot \frac{\sigma_x}{\sqrt{N}}$$

Figure 2 | Formulae used to calculate to standardize the data between each gel through finding retention factors. a, Formula to calculate retention factor of an individual band. Distances were measured between annotated red dashed lines from Fig 1b, d, f. **b,** Formula to calculate mean of retention factors for a shared band between different gels. If the retention factor of a band in one sample of a gel was not similar to the retention factors of any other bands in corresponding samples from other gels, the mean was not calculated and omitted from Table 1. **c,** Formulae to calculate the standard deviation and standard mean of error for each mean set of retention factors.

Since the bands from the Precision Plus Protein Kaleidoscope Standard only displayed on the Group 3 and the Group 4 gel, the data from those gels were used to predict the molecular weight of a band using R_f values. Taking the average retention factors from these lanes, 6 data points were associated with the log of a known molecular weight value and used to construct a standard curve. Bands that showed up with retention factors exclusive to a single gel were omitted.

Sample B produced 4 resulting bands, the fewest among the different samples. Both Sample A and E produced 9 resulting bands, the most among the different samples. All samples produced a band with a mean R_f between 0.509 - 0.521. All samples except Sample C produced a band with a mean R_f of approximately 0.280.

Table 1 | Mean retention factors of bands according to samples from all gels, with 95% confidence intervals.

Mean $R_f(\mu)$	$2SE_x$
Sample A	
0.262	0.026
0.339	0.052
0.537	0.030
0.628	0.027
0.698	0.010
0.742	0.005
0.819	0.030
0.866	0.017
0.939	0.032
Sample B	
0.276	0.002
0.486	0.038
0.707	0.002
0.760	0.004
Sample C	
0.190	0.005
0.224	0.014
0.322	0.018
0.456	0.016
0.524	0.015
0.608	0.042
0.769	0.007
Sample D	
0.282	0.002
0.512	0.009
0.627	0.022
0.728	0.052
0.808	0.012
0.846	0.001
Sample E	
0.089	0.006
0.175	0.005
0.279	0.004
0.505	0.022
0.571	0.029
0.632	0.001
0.675	0.005
0.728	0.014
0.777	0.005

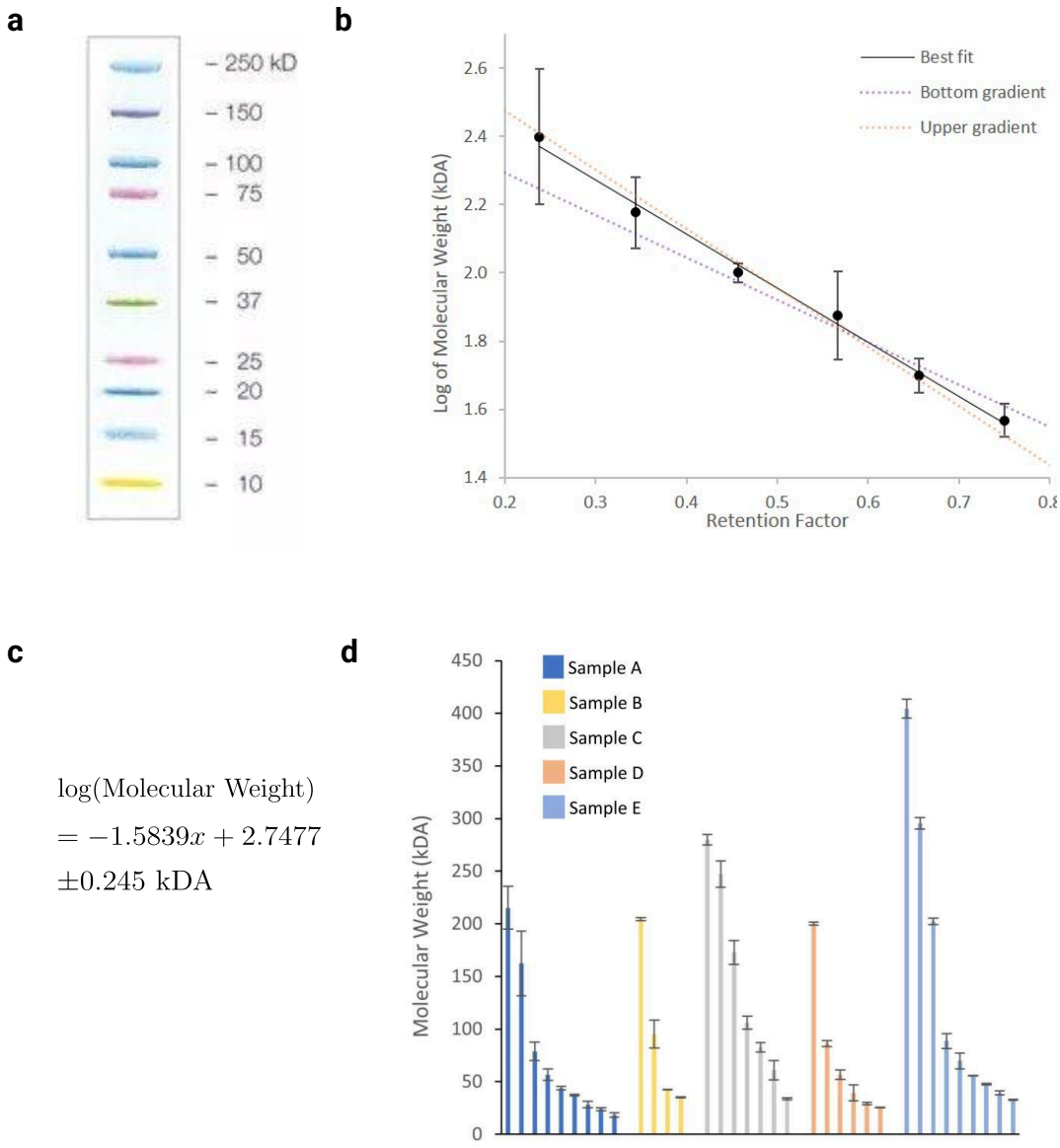

Figure 3 | Calculating molecular weights of bands produced by running samples through SDS-PAGE. a, Known molecular weight values from Precision Plus Protein Kaleidoscope Standard. First six were associated with bands produced in the Group 3 and the Group 4 gel. **b,** Mean retention factor values for bands in ladder lanes plotted against the log of corresponding known molecular weight values. The best fit line, bottom gradient (shallowest possible gradient that crosses through all plotted error bars) and the upper gradient (steepest possible gradient that crosses through all plotted error bars) were included. Each line had a negative, linear gradient. **c,** Equation for the standard curve had a negative, linear trend, with a small gradient uncertainty. **d,** Bar chart of estimated molecular weights for bands from each sample using mean retention factors, with 95% confidence intervals. Each sample had a range of different molecular weights, with Sample E having bands with the greatest molecular weights.

A strong, negative linear correlation was identified from the best fit line of the data points. The molecular weights of the mean R_f values from Table 1 were estimated with the curve, yielding bands with a range of 386.11 kDA (18.21 kDA - 404.33 kDA). The estimations had an average molecular weight of 105.7 kDA, while the standard deviation was 95.26 kDA, indicating a significant variation of protein weights.

Table 2 | Number of shared bands between each pair of sample species.

	A	B	C	D	E
A	9	3	2	5	4
B		4	2	3	2
C			7	1	2
D				6	4
E					9

Sample C only shared up to 2 bands with each other sample species. The greatest number of bands shared between a pair of sample species was 5 bands, observed with Sample A and Sample D,

a

$$\%\ \text{similarity} = \frac{\text{number of shared bands}}{\text{average number of bands}} \cdot 100$$

b

$$\%\ \text{similarity}_{A,B} = \frac{3}{\frac{1}{2}(9+4)} \cdot 100\% = 46.15\%$$

Figure 4 | Determining % similarity between different species with molecular weights of each band. General formula for % similarity between two different fish samples. **b,** Sample calculation for percentage similarity between Sample A and Sample B.

Table 3 | Percentage similarities calculated from shared bands between sample species.

	A	B	C	D	E
A	100%	46.15%	25.00%	66.67%	44.44%
B		100%	36.36%	60.00%	30.77%
C			100%	15.38%	25.00%
D				100%	53.33%
E					100%

The average % similarity Sample C had with each other sample was only 25.44%, while the average % similarity calculated between the selection of any two samples was 40.31%. In addition, the % similarities of Sample C had the least variation compared to the other samples, having a standard deviation of 8.58%.

Sample D had the greatest average % similarity with the other samples at 48.85%. Sample D shared a 66.67% similarity with Sample A, the greatest between any other pair of samples. It also had a 15.38% similarity with Sample C, the lowest % similarity between any pair of samples. In turn, the % similarities of Sample D had a standard deviation of 23.00%, indicating that it had the greatest variation among all other samples.

Running the samples through SDS-PAGE produced bands that traveled different distances across the gel. With the bands produced by the protein ladder of known molecular weights, a standard curve was constructed to estimate the molecular weights of the other bands. The estimated molecular weights of the proteins within each sample were used to determine that a notable similarity was found between Sample A and Sample D, and that Sample C was drastically distinct from the other samples.

DISCUSSION

Cladogram

The results highlighted a significant distinction in the traveled distances of the bands produced by Sample C, suggesting that Sample C was the most evolutionary distant species from the tested samples. In turn, Sample C was placed on the cladogram as an outgroup.

Since Sample A and Sample D shared the highest % similarity among any other pair of samples, they were placed in a subgroup. Sample E was placed further away from C than Sample B because the % similarity between E and C (25.00%) was lower than between B and C (36.36%).

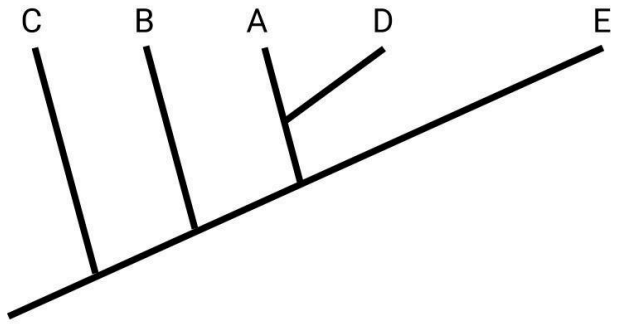

Figure 5 | Hypothesized relationships of sample species based on % similarities of resulting bands. Sample C placed as an outgroup such that it is depicted as the most unrelated among the sample species, while Sample A and D represented to be most related to each other.

Identifying Samples

Actin is a highly abundant and conserved protein in eukaryotic cells. (8) Amino acid sequences of the α-actin protein from *S. salar* (Atlantic salmon) (9), *T. albacares* (yellowfin tuna) (10), *O. vulgaris* (common Octopus) (11), *H. hippoglossus* (Atlantic halibut) (12), and *O. mossambicus* (Mozambique tilapia) (13) were accessed from the NCBI Protein Database with the Biopython package. (14)

Levenshtein distance is a measurement that represents the similarity between two strings. (15) With the Levenshtein Python C extension module, the Levenshtein distance algorithm was used to compute the percentage identity between the amino acid sequences in the form of a FASTA file from each pair of reference species. (16)

Table 4 | Percentage Identity of α-actin amino acid sequences from various aquatic species.

	Salmon	Tuna	Octopus	Halibut	Tilapia
Salmon	100%	91.40%	93.32%	91.40%	99.21%
Tuna		100%	85.88%	100%	91.68%
Octopus			100%	85.88%	92.53%
Halibut				100%	91.68%
Tilapia					100%

The α-actin amino acid sequence from the yellowfin tuna and Atlantic halibut were 100% identical, suggesting that they were the most related among the analyzed species. The percent identity between the α-actin amino acid sequence of salmon and tilapia was also high: 99.21%. Octopus did not have a particularly high percentage identity with a single other species, so it was considered an outgroup from the rest of the species.

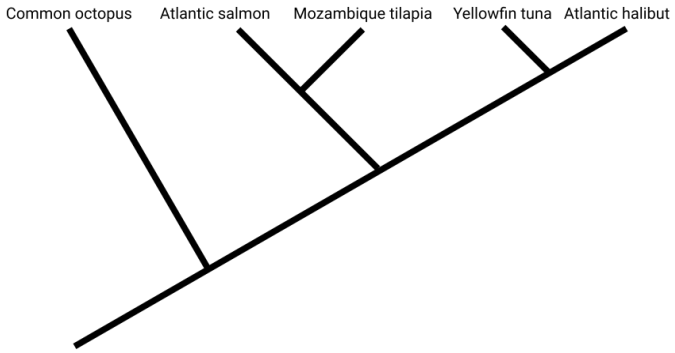

Figure 6 | Hypothesized relationships of five reference species based on percentage identity of α-actin amino acid sequence. Atlantic salmon and Mozambique tilapia were placed under the same subgroup, and Yellowfin Tuna and Atlantic halibut were placed under the same subgroup. Common octopus considered an outgroup from the rest of the species.

Both cladograms represented by Figure 5 and Figure 6 had one distinct outgroup, suggesting that the source of Sample C may also be from an octopus.

No clear matches were found between the other species of each cladogram. The largest discrepancy between the cladograms is that one subgroup was prevalent among the sample species, while two subgroups were prevalent among the five known reference species.

A possible assignment is that Sample B is derived from salmon because it had the greatest percentage identity to the outgroup (36.36% in Table 3, 93.32% in Table 4) than the other species. Accordingly, Sample D may have been from tilapia because it shared the greatest percentage identity to salmon (99.21%) while Sample D was the most similar to Sample B (60.00%). However, no reasonable distinction could be made between Sample A and Sample E because tuna and halibut had 100% percentage identity.

Salmon and Tilapia

Salmon and tilapia can adapt to live in both freshwater and saltwater environments. (17)(18) Majority of both are farmed by commercial aquaculture farms. (19)(20) Both also have significantly lower mercury levels than the other reference fish species.

Table 5 | FDA Published Mercury Levels in Commercial Fish and Shellfish. (1990 - 2012) (21)

Species	Mercury concentration (ppm)
Tilapia	0.013
Salmon (canned)	0.014
Salmon (fresh/frozen)	0.022
Halibut	0.241
Tuna (fresh/frozen yellowfin)	0.354

Mercury poisoning acquired from seafood has been associated with many adverse health effects, especially toward pregnant women. (22)(23)

In terms of key differences, salmon grows larger than tilapia. The largest species of tilapia, the Nile tilapia (*Oreochromis niloticus*) grows up to 4.3 kg while the largest species of salmon, the chinook salmon (*Oncorhynchus tshawytscha*) grows up to 61.4 kg. (24)(25) Salmon are found along the shores of the North Atlantic and Pacific oceans (26). Tilapia is native to Africa, but has been introduced to some U.S. States to control aquatic plant growth.

Uncertainties and Limitations

Although a link was established between salmon and tilapia, the cladogram of the reference species could not effectively support a correlation with the cladogram of the sample species. Some lanes in the gels did not produce expected bands, suggesting an error during the preparation of these samples. A possible explanation for the absence of bands in the Group 6 lane containing the Precision Plus Kaleidoscope Protein Ladder (Fig. 1d) may have been that the samples used were expired; expired protein standards may result in missing bands due to protein degradation. (27) The oversmearing in various lanes including L5 of the Group 3 gel (Fig. 1b) and L7 of the Group 6 gel (Fig. 1d) may have been caused by the overloading of wells. (28)

For increased distinction of bands within each lane, the gels could have been run for longer. In addition, the procedure was limited to protein separation by molecular weight. However, SDS-PAGE alone would be unable to separate different proteins of the same molecular weight. (29) An enhancement to the analysis may have involved detection of specific proteins with certain types of primary and secondary antibodies to trigger immunofluorescent signals. (30)

Published 19 June 2023.

1. Salmon virus originally from the Atlantic, spread to B.C. wild salmon from farms. *UBC Science - Faculty of Science at the University of British Columbia* https://science.ubc.ca/news/salmon-virus-originally-atlantic-spread-bc-wild-salmon-farms (2021).
2. Where are salmon most endangered? | U.S. Geological Survey. https://www.usgs.gov/faqs/where-are-salmon-most-endangered#:~:text=Certain%20populations%20of%20sockeye%20salmon,River%20may%20already%20be%20extinct. (2013).
3. Evolution by Natural Selection | manoa.hawaii.edu/ExploringOurFluidEarth. https://manoa.hawaii.edu/exploringourfluidearth/biological/what-alive/evolution-natural-selection#:~:text=Evolution%20is%20the%20change%20in,phenomenon%20known%20as%20natural%20selection.
4. Polechová, J. & Storch, D. Ecological Niche. in *Elsevier eBooks* 72–80 (2019). doi:10.1016/b978-0-12-409548-9.11113-3.
5. Spor, A. *et al.* Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of Saccharomyces cerevisiae. *BMC Evolutionary Biology* **9**, 296 (2009).
6. Mickevich, M. F. & Weller, S. C. EVOLUTIONARY CHARACTER ANALYSIS: TRACING CHARACTER CHANGE ON A CLADOGRAM. *Cladistics* **6**, 137–170 (1990).
7. Nowakowski, A. B., Wobig, W. J. & Petering, D. H. Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions. *Metallomics* **6**, 1068–1078 (2014).
8. Dominguez, R. & Holmes, K. C. Actin Structure and Function. *Annual Review of Biophysics* **40**, 169–186 (2011).
9. actin, alpha, cardiac muscle 2 [Salmo salar] - Protein - NCBI. https://www.ncbi.nlm.nih.gov/protein/XP_014000428.1?report=fasta.
10. actin alpha 1, skeletal muscle a isoform X2 [Thunnus albacares] - Protein - NCBI. https://www.ncbi.nlm.nih.gov/protein/XP_044197324.1?report=fasta.
11. actin [Octopus vulgaris] - Protein - NCBI. https://www.ncbi.nlm.nih.gov/protein/ACM77788.1?report=fasta.
12. actin, alpha cardiac muscle 1a isoform X2 [Hippoglossus hippoglossus] - Protein - NCBI. https://www.ncbi.nlm.nih.gov/protein/XP_034435820.1?report=fasta.
13. alpha-actin [Oreochromis mossambicus] - Protein - NCBI. https://www.ncbi.nlm.nih.gov/protein/BAA90689.1?report=genpept.
14. Documentation · Biopython. https://biopython.org/wiki/Documentation.
15. Cuelogic Insights. The Levenshtein Algorithm. *Cuelogic An LTI Company* https://www.cuelogic.com/blog/the-levenshtein-algorithm (2021).
16. python-Levenshtein. *PyPI* https://pypi.org/project/python-Levenshtein/ (2023).
17. Schiewe, M. H. Salmon. in *Elsevier eBooks* 522–531 (2013). doi:10.1016/b978-0-12-384719-5.00293-8.
18. Cruz-Suarez, L. E., Marie, D. R., Salazar, M. T. & El-Sayed, A. Tilapia culture in salt water: environmental requirements, nutritional implications and economic potentials. *ResearchGate* (2006).
19. Maine Seafood Guide - Atlantic Salmon - Maine Sea Grant - University of Maine. *Maine Sea Grant* https://seagrant.umaine.edu/maine-seafood-guide/salmon/ (2023).
20. Tilapia farming guide - Understanding the five needs of tilapia. *Lakeway Tilapia* https://lakewaytilapia.com/How_To_Raise_Tilapia.php.
21. NutritionU.S. Food And Drug Administration (2022).
22. Stratakis, N. *et al.* Association of Fish Consumption and Mercury Exposure During Pregnancy With Metabolic Health and Inflammatory Biomarkers in Children. *JAMA Network Open* **3**, e201007 (2020).
23. Bose-O'Reilly, S., McCarty, K. M., Steckling, N. & Lettmeier, B. Mercury Exposure and Children's Health. *Current Problems in Pediatric and Adolescent Health Care* **40**, 186–215 (2010).
24. Oreochromis niloticus summary page. *FishBase* https://fishbase.mnhn.fr/summary/Oreochromis-niloticus.html.
25. Oncorhynchus tshawytscha summary page. *FishBase* https://www.fishbase.se/summary/SpeciesSummary.php?ID=244&AT=king+salmon.
26. Salmon. *The Nature Conservancy*.
27. Protein Standards and Ladders Support — Troubleshooting | Thermo Fisher Scientific - IE. https://www.thermofisher.com/ca/en/home/technical-resources/technical-reference-library/protein-electrophoresis-western-blotting-support-center/protein-standards-ladders-support/protein-standards-ladders-support-troubleshooting.html.
28. Smeared protein gels. https://www.ruf.rice.edu/~bioslabs/studies/sds-page/gelgoofs/smear.html.
29. SDS-PAGE. https://bio.davidson.edu/genomics/method/SDSPAGE/SDSPAGE.html.
30. Quintero-Ronderos, P. Analysis of proteins and antibodies. *Autoimmunity - NCBI Bookshelf* https://www.ncbi.nlm.nih.gov/books/NBK459443/#:~:text=Immunofluorescence%20and%20immunohistochemistry%20are%20methods,done%20directly%20on%20microscopy%20slides. (2013).

Acknowledgements The authors thank the Group 3 researchers lead by Susie Choi and Group 4 researchers lead by Benjamin Guan for providing images of their gel after SDS-PAGE, the Biotech SHSM department for supplying the equipment and materials needed to conduct this study, and Isabella Wong for printing this publication.