Chicken Eggs Are a Practical and Common Exome-Matched Diet for Multicellular Eukaryotic Organisms

Esumi, Genshiro

Department of Pediatric Surgery, Hospital of the University of Occupational and Environmental Health, Kitakyushu, Japan

Abstract

A 2017 study reported that a diet reflecting the average exomic amino acid composition of a fruit fly—referred to as its "exome-matched diet"—maximized both its lifespan and reproductive output. Building on this insight, a species-specific exome-matched diet was proposed as a candidate for an organism's optimal amino acid composition.

In this analysis, I used publicly available Reference Proteomes data for 81 different species to calculate each species' exome-matched diet amino acid composition. I then compared these compositions to approximately 2,000 food items listed in the official Japanese Standard Tables of Food Composition to determine which foods best matched each species' exome profile. While there was substantial variability among most prokaryotic and certain unicellular eukaryotic organisms, my findings revealed that, with a few exceptions, whole chicken eggs or egg-based products most closely approximated the calculated exome-matched amino acid composition for a broad range of eukaryotic species, especially multicellular organisms. Consequently, for these organisms, whole chicken eggs appear to serve as a practical exome-matched diet.

Although the exome-matched diet does not automatically define an absolutely optimal amino acid composition for nutrition, the finding that the exome-matched diets of multicellular eukaryotes are essentially represented by chicken eggs aligns well with established nutritional principles. This consistency, paradoxically, further suggests the validity of the exome-matched diet concept.

Keywords: Exome-matched diet, Chicken eggs, Amino acid composition, Food composition, Nutrition

Email: esumi@clnc.uoeh-u.ac.jp

Background

Proteins in living organisms are composed of up to twenty types of amino acids, and heterotrophic organisms—including humans—depend on external sources for many of these amino acids. Furthermore, because it is well known that the amino acid profile of a given food source can significantly influence such organism's nutritional status, the questions "What do we eat?" and "Which amino acids does this food contain?" have long been central in nutritional science. However, considering how to establish an optimal amino acid composition involves analyzing a potentially vast number of combinations—up to twenty different amino acids—making it a highly complex task and posing a longstanding challenge in determining an optimal dietary amino acid composition.

Meanwhile, a 2017 study reported that a diet reflecting the average exomic amino acid composition of *Drosophila melanogaster*—referred to as the "exome-matched diet"—maximized both lifespan and reproductive output in fruit flies [1]. Building on this insight, the exome-matched diet was proposed as a candidate for an optimal amino acid composition for nutrition.

If the exome-matched diet concept is indeed valid, it naturally raises the questions: "What would such a diet look like for humans, and which foods would match it?" and "What about other organisms—how would their exome-matched diets align with actual food items?" Investigating these questions requires two key data resources: (1) exome information from various species to compute each exome-matched diet composition, and (2) a comprehensive database of amino acid compositions for everyday foods. For the former, I used publicly available "reference proteome" datasets from multiple species, and for the latter, I took advantage of the comprehensive Japanese food composition table, which contains data for approximately 2,000 distinct foods. I therefore calculated exome-matched amino acid compositions for a variety of organisms and compared them against the amino acid profiles of numerous foods in the database.

In this study, by leveraging this approach, I aimed to identify, for each organism, what its practical exome-matched diet would be.

Subjects and Methods

Reference Proteomes

For the present analysis of various organisms' exomes, I used the "reference proteomes" dataset published by EMBL-EBI [2]. This dataset spans the three domains of life (Archaea, Bacteria, and Eukaryotes) and includes amino acid sequences from 1,547,370 proteins across a total of 81 different species. In this study, I analyzed each protein sequence by counting the number of each amino acid residue, then divided by the total number of residues in that protein to obtain its amino acid composition (which sums to 1). By averaging these compositions for all proteins within a species, I derived the exome-matched diet amino acid composition for each of the 81 species. This procedure was applied to every species listed, regardless of its actual feeding capabilities (e.g., plants and other organisms that may not consume external protein sources).

Japanese Standard Tables of Food Composition

Next, I used the Standard Tables of Food Composition published by the Japanese Ministry of Education, Culture, Sports, Science and Technology, the latest version of which provides amino acid data for 1,954 foods (including both raw ingredients and processed products) [3]. Each food is annotated with a measured weight for every amino acid. To convert these weights into molar compositions, I divided each amino acid's weight by its respective molecular weight, as indicated in the tables' supporting documentation.

Because of certain measurement and analytical constraints on amino acids, aspartic acid (Asp) and asparagine (Asn) are typically reported together, and glutamic acid (Glu) and glutamine (Gln) likewise. Consequently, the dataset ultimately contained 18 amino acids per food. By dividing each individual amino acid's moles by the sum of all measured amino acids, I obtained an 18-component molar composition for every food item (so that each total is 1).

Finding the Food with the Minimum Distance

To determine which food item's amino acid composition best matches each organism's exome, I first aligned the 20 amino acids computed from the exome data with the 18 amino acids available in the food database. In other words, I merged Asn with Asp and Gln with Glu to produce an 18-amino-acid composition for each organism's exome. I then calculated the distance between each organism's exome-matched composition and each of the 1,954 foods using the angular distance metric:

$$d(x, y) = \arccos\left(\frac{x \cdot y}{\|x\| \|y\|}\right).$$

For each species, I identified the single food item that yielded the minimum distance. Additionally, because the Japanese food composition table was originally presented in Japanese, I translated the identified food names into English where appropriate.

Data Processing

All data processing and table creation in this study were performed using Microsoft Excel (Microsoft Corp., Redmond, WA, USA), and all graphs were generated with JMP 18 (SAS Institute Inc., Cary, NC, USA).

Results

Overview of the Studied Species and Protein Counts

Table 1 lists the 81 species included in this analysis, along with their IDs, taxonomic domain, cell organization type, and the number of exons/proteins in each reference proteome dataset [2].

Average Amino Acid Composition for Each Exome

Table 2 presents the average amino acid composition calculated for each of the 81 listed species, following the methods described in the Subjects and Methods section. The values in this table represent the exome-matched diet amino acid compositions as originally described in the 2017 paper [3].

Closest Food Items to Each Exome-Matched Composition

Table 3 lists, for each organism, the food item that exhibited the minimum distance to that organism's exome-matched amino acid composition, along with the corresponding distance value. In many species, raw whole chicken eggs emerged as the closest match. I therefore also calculated the distance from each organism's exome-matched composition to the egg composition and the difference between these distances. If this difference was sufficiently small, I deemed raw whole chicken eggs a practical approximation of that organism's exome-matched diet.

Reordered Species by Relative Distance to Chicken Eggs

To facilitate more intuitive inspection, I reordered the species in **Table 4** according to the descending difference between their distance to raw whole chicken eggs and their distance to their best-matching food. This reordering highlights how closely each species' exome composition aligns with raw whole chicken eggs relative to its top-ranked food item. As a result, for most species—particularly multicellular eukaryotic organisms—raw whole chicken eggs could be considered their practical exome-matched diet among the listed foods.

Additional Analyses

I produced several figures to further clarify these findings. **Figure 1** compares the distributions of amino acid compositions, by species and by amino acid, for the exomes of all 81 organisms. These distributions range from amino acids—such as Met, His, and Leu—that appear similarly distributed across species, to others—like Ala, Ile, and Lys—whose distributions vary considerably among different organisms.

Because the amino acid composition distance between chicken eggs and the average exome of most multicellular eukaryotes was particularly small, I prepared **Figure 2** to illustrate the amino acid composition distributions of these multicellular eukaryotes, while marking the chicken egg composition with a vertical black line. The results indicate that the exomes of these organisms cluster closely together, with the black line (i.e., the chicken egg composition) positioned near the center of their distributions.

Finally, in **Figure 3**, I focused on three specific species: *Drosophila melanogaster* (the fruit fly) from the 2017 study, *Gallus gallus* (the chicken species from which the eggs originate), and *Homo sapiens* (humans). For each species, I plotted its exomic amino acid distribution and indicated both the mean amino acid composition of that species with a dashed vertical line and the chicken egg composition with a solid black line. As observed in **Figure 2**, these three distributions overlap substantially, and their mean values are quite similar. The dashed lines lie somewhat closer together than they do to the solid black line, suggesting that these organisms' exomes resemble one another even more closely than they resemble the chicken egg composition itself.

Discussion

In 1838, G.J. Mulder identified a fundamental class of molecular components shared by both animals and plants, naming them "proteins" [4]. Over the following century, it gradually became clear that these proteins are assemblies of amino acids. In 1935, the twentieth and final amino acid, threonine, was discovered by W.C. Rose, who subsequently demonstrated that rats could gain weight on a nutrient source containing only amino acids, with no intact protein [5]. This finding showed that amino acid nutrition could replace protein nutrition. Rose further reported that removing certain amino acids from the diet induced a negative nitrogen balance in rats, thereby establishing these as "essential amino acids" [6]. As a result, the importance of amino acid nutrition became widely recognized.

Nevertheless, determining which amino acids—and in what quantities—should be ingested has remained a major challenge. There are as many as twenty amino acids to consider, and several can be omitted without disrupting nitrogen balance. In other words, organisms exhibit some degree of resilience to low-amino-acid diets, making the analysis of optimal amino acid composition extremely difficult. Consequently, identifying which amino acid profile is truly optimal has remained a central problem in clinical nutrition.

In the midst of this longstanding complexity, a 2017 study introduced the concept of an "exome-matched diet," wherein the average amino acid composition of an organism's exomic proteins was proposed to represent a potentially optimal nutritional profile—at least in the case of *Drosophila melanogaster* [1]. Before becoming aware of that study, I had demonstrated that, for a variety of organisms, the distribution of amino acid compositions in the exome closely approximates a single-peaked binomial distribution for each species [7]. I further suggested that this bell-shaped distribution arises because the amino acid compositions of protein-coding genes in the exome are constrained by the organism's proteome synthesis resources [7]. In a binomial distribution, the peak of this bell shape corresponds to the average of the distribution. Consequently, it seemed plausible that the observation of bell-shaped amino acid composition distributions in various exomes, and the notion that the average of these exomic compositions (i.e., the exome-matched diet) might be optimal for that exome, could be two sides of the same coin. This possibility prompted me to undertake the present study.

In the present work, by comparing each organism's average exomic amino acid composition with data from Japan's official food composition table, I found that—particularly among multicellular eukaryotes—chicken eggs emerged as the closest match to each exome-matched diet. For example, in **Table 4**, *Oryza sativa* (a multicellular eukaryote) appears relatively low in the list, suggesting at first glance that its amino acid composition might diverge considerably from that of chicken eggs. However, the difference between the distance to its best-matching food and the distance to chicken eggs is only about 0.05—a relatively small gap within the distribution of distances—leading me to conclude that, in practical terms, chicken eggs still approximate the exome-matched composition for *O. sativa*. Moreover, any species ranked above *O. sativa* in **Table 4**, including mostly eukaryotes but also some bacteria, would similarly have its exome-matched diet approximated by chicken eggs. Consequently, for the majority of these 81 species, chicken eggs appear to serve as a practical approximation of their exome-matched diet.

However, in this study, I initially reduced the 20 amino acids to 18 by merging Asp with Asn and Glu with Gln. To evaluate how this consolidation might affect the results, I investigated the balance of Asp versus Asn and Glu versus Gln across different species. It is conceivable that species vary considerably in their usage ratios of Asp/Asn and Glu/Gln, which could alter the rank order of distances observed in this analysis.

Therefore, as an additional step, I introduced two indices (Asp–Asn skew and Glu–Gln skew) defined as follows:

$$\text{Asp-Asn skew} = \frac{\text{Asp} - \text{Asn}}{\text{Asp} + \text{Asn}}, \qquad \text{Glu-Gln skew} = \frac{\text{Glu} - \text{Gln}}{\text{Glu} + \text{Gln}}.$$

I calculated these indices for the proteins of each species and plotted their distributions in **Figure 4**. As shown, both Asp–Asn skew and Glu–Gln skew exhibit variability among Archaea, Bacteria, and unicellular eukaryotes, whereas the distributions in multicellular eukaryotes, including chickens (and thus chicken eggs), are notably uniform. This suggests that the relative usage of Asp versus Asn and Glu versus Gln is effectively constant within the proteomes of multicellular eukaryotes. Consequently, analyzing 18 amino acids (merging Asp with Asn and Glu with Gln) is still valid for comparisons involving multicellular eukaryotes. Based on these findings, I conclude that the scope of this study's conclusions is most appropriately restricted to multicellular eukaryotes.

Naturally, a chicken egg itself is the resource used to construct the entire body of a chick—that is, a young chicken. Accordingly, the amino acid composition of a chicken egg can be viewed as having been evolutionarily optimized to provide the raw materials for the chick's entire proteome. In this study, I found that the exome-matched diet, computed as the average amino acid composition of each species' exome, was quite similar to that of the chicken egg. Because binomial distributions can be determined solely by their averages, this interpretation supports my hypothesis that the bell-shaped, binomial-like distributions of exomic amino acid compositions in many species arise because these exomes are constrained by the amino acid composition of the resource they rely on—in this case, chicken eggs.

As shown in **Figure 2** and **Figure 3**, multicellular eukaryotes have amino acid compositions that closely resemble one another. In a prior report, I used the same set of 81 species to compute pairwise angular distances among each species' average exomic amino acid composition, then constructed a phylogenetic tree through clustering analysis [8]. That analysis revealed that the exomic amino acid compositions of so-called "animals"—including chickens and humans—are particularly similar to one another. Therefore, if chicken eggs represent a proteome-matched diet for chickens (i.e., chicks), it is plausible that in animals more broadly, chicken eggs are similarly optimized as an amino acid resource for whole-body protein synthesis.

Based on these findings, I conclude that for multicellular eukaryotes, the actually optimized proteome-matched diet can be approximated in practice by chicken eggs, and that the concept of the exome-matched diet serves as one way to estimate this composition. However, from the standpoint of clinical nutrition, does an optimally balanced amino acid profile for humans indeed correspond to chicken eggs? In an earlier study, I examined published data on the amino acid composition of a fetal pig's entire body and discovered that this composition could be approximated by combining the pig's average exomic amino acid composition with the amino acid composition of type I collagen [9]. Given that the pig's exome distribution is nearly the same as that of humans, and that their collagen genes are largely homologous—and considering that the pig's exome composition can also be approximated by chicken eggs—it is reasonable to hypothesize that in multicellular eukaryotes (such as humans and pigs) whose extracellular matrix contains large amounts of collagen, the truly optimized proteome-matched diet might be represented by a combination of chicken eggs and collagen (essentially gelatin), in contrast to fruit flies whose extracellular matrix is predominantly composed of a non-proteinaceous substance called cuticle.

Although the oral intake of collagen remains a subject of debate, the results presented here—encompassing whole-body amino acid composition, exomic amino acid composition, and food amino acid composition—suggest that collagen supplementation could, in principle, be beneficial from both an exome-based and proteome-based nutritional perspective. While the exome-matched diet concept does not automatically provide the exact amino acid composition for optimal nutrition, it does offer a valuable starting point for considering what optimal nutrition might look like.

Throughout this study, I have demonstrated that, in practical terms, the exome-matched diet for animals coincides with chicken eggs. Eggs have long been considered a reference for optimal nutrition in clinical settings, and this finding aligns well with that notion [10]. Consequently, the very plausibility of this result paradoxically reinforces the validity of the exome-matched diet approach.

Conclusion

By comparing the amino acid compositions of various organisms' exomes with those of different foods, I found that, for multicellular eukaryotes, the exome-matched diet can be practically represented by chicken eggs. While the candidate for optimal nutrition suggested by the exome-matched diet concept (i.e., chicken eggs) does not automatically provide the exact amino acid composition for truly optimal nutrition, it nonetheless offers a valuable starting point for considering what desirable and appropriate nutrition might look like.

Reference

1. **Piper, M. D. W., Soultoukis, G. A., Blanc, E., Mesaros, A., Herbert, S. L., Juricic, P., He, X., Atanassov, I., Salmonowicz, H., Yang, M., Simpson, S. J., Ribeiro, C., & Partridge, L.** (2017). Matching dietary amino acid balance to the in silico-translated exome optimizes growth and reproduction without cost to lifespan. *Cell Metabolism*, 25(5), 1206–1221. https://doi.org/10.1016/j.cmet.2017.04.020

2. **EMBL-EBI.** (2024). Reference Proteomes (Release 2024_02) [Database]. Retrieved January 7, 2025, from https://www.ebi.ac.uk/reference_proteomes/

3. **Ministry of Education, Culture, Sports, Science and Technology [MEXT].** (2020). 日本食品標準成分表2020版（八訂） アミノ酸成分表編 第2章第1表 [Standard tables of food composition in Japan 2020 (Eighth revised edition), amino acid composition table, Chapter 2, Table 1]. Retrieved January 15, 2025, from https://www.mext.go.jp/a_menu/syokuhinseibun/index.htm

4. **Mulder, G. J.** (1838). Sur la composition de quelques substances animales. *Bulletin des Sciences Physiques et Naturelles en Néerlande*: 104. https://archive.org/details/bulletindesscien00leyd/page/104/mode/2up

5. **McCoy, R. H., Meyer, C. E., & Rose, W. C.** (1935). FEEDING EXPERIMENTS WITH MIXTURES OF HIGHLY PURIFIED AMINO ACIDS. *Journal of Biological Chemistry* (Vol. 112, Issue 1, pp. 283–302). https://doi.org/10.1016/s0021-9258(18)74986-7

6. **ROSE W. C.** (1957). The amino acid requirements of adult man. *Nutrition abstracts and reviews,* 27(3), 631–647. http://www.ncbi.nlm.nih.gov/pubmed/18146542

7. **Esumi, G.** (2023). The distributions of amino acid compositions of proteins in an organism's proteome uniformly approximate binomial distributions. *Jxiv*. https://doi.org/10.51094/jxiv.408

8. **Esumi, G.** (2025). Construction of a Phylogenetic Tree Based on the Average Amino Acid Composition of Exomes. *Jxiv*. https://doi.org/10.51094/jxiv.1044

9. **Esumi, G.** (2023). The Total Body Amino Acid Composition of an Animal Could Be Explained by a Mixture of the Average Composition of Its Proteome as an Intracellular Composition Estimate and the Type I Collagen Composition as an Extracellular Composition Estimate. *Jxiv*. https://doi.org/10.51094/jxiv.424

10. **Marmorshtein, S. I., Trahktenberg, A. K., & Bidiak, I. V.** (1965). Protein requirements: Reports of a joint FAO-WHO expert group. *World Health Organization Technical Report Series,* (301), 1–71. https://pubmed.ncbi.nlm.nih.gov/14280833/

No.	Scientific Name	Organism ID	Domain	Cell Organization	Protein Count
1	*Halobacterium salinarum* (strain ATCC 700922 / JCM 11081 / NRC-1) (Halobacterium halobium)	64091	archaea	unicellular	2427
2	*Thermococcus kodakarensis* (strain ATCC BAA-918 / JCM 12380 / KOD1) (Pyrococcus kodakaraensis (strain KOD1))	69014	archaea	unicellular	2301
3	*Methanosarcina acetivorans* (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A)	188937	archaea	unicellular	4468
4	*Methanocaldococcus jannaschii* (strain ATCC 43067 / DSM 2661 / JAL-1 / JCM 10045 / NBRC 100440) (Methanococcus jannaschii)	243232	archaea	unicellular	1787
5	*Saccharolobus solfataricus* (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2) (Sulfolobus solfataricus)	273057	archaea	unicellular	2937
6	*Korarchaeum cryptofilum* (strain OPF8)	374847	archaea	unicellular	1602
7	*Nitrosopumilus maritimus* (strain SCM1)	436308	archaea	unicellular	1795
8	*Mycobacterium tuberculosis* (strain ATCC 25618 / H37Rv)	83332	bacteria	unicellular	3999
9	*Escherichia coli* (strain K12)	83333	bacteria	unicellular	4416
10	*Helicobacter pylori* (strain ATCC 700392 / 26695) (Campylobacter pylori)	85962	bacteria	unicellular	1554
11	*Streptomyces coelicolor* (strain ATCC BAA-471 / A3(2) / M145)	100226	bacteria	unicellular	8039
12	*Neisseria meningitidis* serogroup B (strain MC58)	122586	bacteria	unicellular	2001
13	*Leptospira interrogans* serogroup Icterohaemorrhagiae serovar Lai (strain 56601)	189518	bacteria	unicellular	3676
14	*Fusobacterium nucleatum* subsp. nucleatum (strain ATCC 25586 / DSM 15643 / BCRC 10681 / CIP 101130 / JCM 8532 / KCTC 2640 / LMG 13131 / VPI 4355)	190304	bacteria	unicellular	2046
15	*Pseudomonas aeruginosa* (strain ATCC 15692 / DSM 22644 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1)	208964	bacteria	unicellular	5564
16	*Bacillus subtilis* (strain 168)	224308	bacteria	unicellular	4267
17	*Aquifex aeolicus* (strain VF5)	224324	bacteria	unicellular	1553
18	*Bradyrhizobium diazoefficiens* (strain JCM 10833 / BCRC 13528 / IAM 13628 / NBRC 14792 / USDA 110)	224911	bacteria	unicellular	8253
19	*Bacteroides thetaiotaomicron* (strain ATCC 29148 / DSM 2079 / JCM 5827 / CCUG 10774 / NCTC 10582 / VPI-5482 / E50)	226186	bacteria	unicellular	4782
20	*Rhodopirellula baltica* (strain DSM 10527 / NCIMB 13988 / SH1)	243090	bacteria	unicellular	7271
21	*Deinococcus radiodurans* (strain ATCC 13939 / DSM 20539 / JCM 16871 / CCUG 27074 / LMG 4051 / NBRC 15346 / NCIMB 9279 / VKM B-1422 / R1)	243230	bacteria	unicellular	3084
22	*Geobacter sulfurreducens* (strain ATCC 51573 / DSM 12127 / PCA)	243231	bacteria	unicellular	3402
23	*Mycoplasma genitalium* (strain ATCC 33530 / DSM 19775 / NCTC 10195 / G37) (Mycoplasmoides genitalium)	243273	bacteria	unicellular	483
24	*Thermotoga maritima* (strain ATCC 43589 / DSM 3109 / JCM 10099 / NBRC 100826 / MSB8)	243274	bacteria	unicellular	1852
25	*Gloeobacter violaceus* (strain ATCC 29082 / PCC 7421)	251221	bacteria	unicellular	4406
26	*Chlamydia trachomatis* (strain D/UW-3/Cx)	272561	bacteria	unicellular	895
27	*Thermodesulfovibrio yellowstonii* (strain ATCC 51303 / DSM 11347 / YP87)	289376	bacteria	unicellular	1982
28	*Chloroflexus aurantiacus* (strain ATCC 29366 / DSM 635 / J-10-fl)	324602	bacteria	unicellular	3850
29	*Dictyoglomus turgidum* (strain DSM 6724 / Z-1310)	515635	bacteria	unicellular	1743
30	*Synechocystis sp.* (strain PCC 6803 / Kazusa)	1111708	bacteria	unicellular	3508
31	*Chlamydomonas reinhardtii* (Chlamydomonas smithii)	3055	eukaryota	unicellular	18832
32	*Physcomitrium patens* (Spreading-leaved earth moss) (Physcomitrella patens)	3218	eukaryota	**multicellular**	47782
33	*Arabidopsis thaliana* (Mouse-ear cress)	3702	eukaryota	**multicellular**	41596
34	*Zea mays* (Maize)	4577	eukaryota	**multicellular**	63281
35	*Leishmania major*	5664	eukaryota	unicellular	8038
36	*Paramecium tetraurelia*	5888	eukaryota	unicellular	39461
37	*Caenorhabditis elegans*	6239	eukaryota	**multicellular**	28553
38	*Helobdella robusta* (Californian leech)	6412	eukaryota	**multicellular**	23328
39	*Ixodes scapularis* (Black-legged tick) (Deer tick)	6945	eukaryota	**multicellular**	20496
40	*Tribolium castaneum* (Red flour beetle)	7070	eukaryota	**multicellular**	18505
41	*Anopheles gambiae* (African malaria mosquito)	7165	eukaryota	**multicellular**	14411
42	*Drosophila melanogaster* (Fruit fly)	7227	eukaryota	**multicellular**	23539
43	*Ciona intestinalis* (Transparent sea squirt) (Ascidia intestinalis)	7719	eukaryota	**multicellular**	17311
44	*Branchiostoma floridae* (Florida lancelet) (Amphioxus)	7739	eukaryota	**multicellular**	38648
45	*Lepisosteus oculatus* (Spotted gar)	7918	eukaryota	**multicellular**	22463
46	*Danio rerio* (Zebrafish) (Brachydanio rerio)	7955	eukaryota	**multicellular**	46840
47	*Oryzias latipes* (Japanese rice fish) (Japanese killifish)	8090	eukaryota	**multicellular**	36138
48	*Xenopus laevis* (African clawed frog)	8355	eukaryota	**multicellular**	61769
49	*Xenopus tropicalis* (Western clawed frog) (Silurana tropicalis)	8364	eukaryota	**multicellular**	37693
50	*Gallus gallus* (Chicken)	9031	eukaryota	**multicellular**	43968
51	*Macaca mulatta* (Rhesus macaque)	9544	eukaryota	**multicellular**	44416
52	*Gorilla gorilla gorilla* (Western lowland gorilla)	9595	eukaryota	**multicellular**	44726
53	*Pan troglodytes* (Chimpanzee)	9598	eukaryota	**multicellular**	48794
54	*Homo sapiens* (Human)	9606	eukaryota	**multicellular**	104573
55	*Canis lupus familiaris* (Dog) (Canis familiaris)	9615	eukaryota	**multicellular**	43672
56	*Bos taurus* (Bovine)	9913	eukaryota	**multicellular**	37871
57	*Mus musculus* (Mouse)	10090	eukaryota	**multicellular**	63289
58	*Rattus norvegicus* (Rat)	10116	eukaryota	**multicellular**	49582
59	*Monodelphis domestica* (Gray short-tailed opossum)	13616	eukaryota	**multicellular**	36221
60	*Thalassiosira pseudonana* (Marine diatom) (Cyclotella nana)	35128	eukaryota	unicellular	11612
61	*Daphnia magna*	35525	eukaryota	**multicellular**	26600
62	*Plasmodium falciparum* (isolate 3D7)	36329	eukaryota	unicellular	5369
63	*Oryza sativa* subsp. japonica (Rice)	39947	eukaryota	**multicellular**	49224
64	*Dictyostelium discoideum* (Social amoeba)	44689	eukaryota	unicellular	12746
65	*Nematostella vectensis* (Starlet sea anemone)	45351	eukaryota	**multicellular**	24445
66	*Monosiga brevicollis* (Choanoflagellate)	81824	eukaryota	unicellular	9156
67	*Phytophthora ramorum* (Sudden oak death agent)	164328	eukaryota	unicellular	15349
68	*Giardia intestinalis* (strain ATCC 50803 / WB clone C6) (Giardia lamblia)	184922	eukaryota	unicellular	4900
69	*Cryptococcus neoformans* var. neoformans serotype D (strain JEC21 / ATCC MYA-565) (Filobasidiella neoformans)	214684	eukaryota	unicellular	6746
70	*Candida albicans* (strain SC5314 / ATCC MYA-2876) (Yeast)	237561	eukaryota	unicellular	6037
71	*Ustilago maydis* (strain 521 / FGSC 9021) (Corn smut fungus)	237631	eukaryota	unicellular	6805
72	*Yarrowia lipolytica* (strain CLIB 122 / E 150) (Yeast) (Candida lipolytica)	284591	eukaryota	unicellular	6454
73	*Schizosaccharomyces pombe* (strain 972 / ATCC 24843) (Fission yeast)	284812	eukaryota	unicellular	5132
74	*Phaeosphaeria nodorum* (strain SN15 / ATCC MYA-4574 / FGSC 10173) (Glume blotch fungus) (Parastagonospora nodorum)	321614	eukaryota	unicellular	15998
75	*Aspergillus fumigatus* (strain ATCC MYA-4609 / CBS 101355 / FGSC A1100 / Af293) (Neosartorya fumigata)	330879	eukaryota	unicellular	9648
76	*Neurospora crassa* (strain ATCC 24698 / 74-OR23-1A / CBS 708.71 / DSM 1257 / FGSC 987)	367110	eukaryota	unicellular	10266
77	*Trichomonas vaginalis* (strain ATCC PRA-98 / G3)	412133	eukaryota	unicellular	50190
78	*Puccinia graminis* f. sp. *tritici* (strain CRL 75-36-700-3 / race SCCL) (Black stem rust fungus)	418459	eukaryota	unicellular	15808
79	*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c) (Baker's yeast)	559292	eukaryota	unicellular	6091
80	*Sclerotinia sclerotiorum* (strain ATCC 18683 / 1980 / Ss-1) (White mold) (Whetzelinia sclerotiorum)	665079	eukaryota	unicellular	14445
81	*Batrachochytrium dendrobatidis* (strain JAM81 / FGSC 10211) (Frog chytrid fungus)	684364	eukaryota	unicellular	8610
				Total	1547370

Table 1. Overview of the Studied Species and Protein Counts

Table 1 lists the 81 species included in this analysis, along with their organism IDs, taxonomic domain, cell organization type, and the number of exons/proteins in each reference proteome dataset. The domain and protein count columns are color-coded according to domain classification, and multicellular organisms are displayed in bold to improve readability.

No.	Scientific Name	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
1	*Halobacterium salinarum*	0.12407	0.00889	0.0888	0.06979	0.03058	0.08201	0.02265	0.03687	0.01821	0.08469	0.0191	0.02153	0.0464	0.02691	0.06561	0.05464	0.06794	0.0949	0.01126	0.02514
2	*Thermococcus kodakarensis*	0.07304	0.00605	0.04609	0.09242	0.04258	0.07421	0.01543	0.07038	0.0718	0.10734	0.02497	0.02987	0.04216	0.01729	0.06113	0.04855	0.04367	0.08396	0.01264	0.03645
3	*Methanosarcina acetivorans*	0.06629	0.01453	0.04976	0.07941	0.04731	0.06994	0.01726	0.07562	0.07015	0.09589	0.02711	0.04246	0.03945	0.02558	0.04721	0.06736	0.05082	0.068	0.01016	0.03571
4	*Methanocaldococcus jannaschii*	0.05447	0.01407	0.05329	0.08678	0.04249	0.06184	0.01392	0.1071	0.10671	0.09512	0.02504	0.05102	0.0327	0.01459	0.03919	0.04432	0.03922	0.06867	0.00701	0.04246
5	*Saccharolobus solfataricus*	0.05343	0.00756	0.04685	0.07053	0.0438	0.06169	0.01298	0.09485	0.08221	0.10399	0.02327	0.04882	0.03589	0.02106	0.04918	0.06635	0.04584	0.07477	0.00995	0.04701
6	*Korarchaeum cryptofilum*	0.06846	0.00938	0.04855	0.08472	0.03745	0.07669	0.01348	0.08305	0.05775	0.10988	0.02782	0.02675	0.04342	0.01481	0.07305	0.07177	0.03527	0.07344	0.01108	0.03318
7	*Nitrosopumilus maritimus*	0.05963	0.01131	0.05955	0.07541	0.04374	0.06223	0.01819	0.08494	0.09008	0.08403	0.02842	0.04776	0.03691	0.03234	0.03552	0.06945	0.05444	0.06666	0.00898	0.03041
8	*Mycobacterium tuberculosis*	0.13255	0.01007	0.05865	0.04819	0.02836	0.09087	0.02317	0.04236	0.0218	0.09669	0.02099	0.02323	0.05782	0.03129	0.07804	0.05515	0.05916	0.08598	0.01504	0.02057
9	*Escherichia coli*	0.09305	0.01318	0.05002	0.0576	0.03944	0.07025	0.02335	0.0617	0.04742	0.10621	0.03054	0.03902	0.04268	0.04376	0.05607	0.05831	0.0534	0.07084	0.01514	0.028
10	*Helicobacter pylori*	0.06687	0.01225	0.04646	0.06903	0.05481	0.05555	0.02189	0.07188	0.09247	0.11448	0.02528	0.05466	0.03218	0.03653	0.0363	0.06709	0.0417	0.05719	0.00745	0.03593
11	*Streptomyces coelicolor*	0.13599	0.00884	0.06061	0.05768	0.02622	0.09479	0.0239	0.02903	0.02075	0.10077	0.0185	0.01645	0.06188	0.02648	0.08583	0.05005	0.06105	0.08621	0.01524	0.01972
12	*Neisseria meningitidis*	0.09655	0.01324	0.05037	0.06095	0.04411	0.07361	0.02219	0.05997	0.05986	0.09822	0.02674	0.04048	0.04248	0.03915	0.05665	0.05563	0.05087	0.06652	0.01217	0.03026
13	*Leptospira interrogans*	0.05031	0.01006	0.04632	0.07204	0.0582	0.06183	0.01719	0.08154	0.08268	0.10395	0.02057	0.05026	0.03772	0.03269	0.04483	0.07716	0.0494	0.05664	0.01145	0.03517
14	*Fusobacterium nucleatum*	0.05234	0.00858	0.0527	0.0799	0.05173	0.05921	0.01188	0.10135	0.1049	0.09525	0.02538	0.06217	0.02491	0.02151	0.03207	0.05801	0.04602	0.06088	0.00634	0.04486
15	*Pseudomonas aeruginosa*	0.11635	0.0114	0.05248	0.06156	0.03557	0.08272	0.02222	0.04153	0.02988	0.12462	0.02175	0.02523	0.05084	0.04232	0.07702	0.05463	0.04113	0.06885	0.01504	0.02486
16	*Bacillus subtilis*	0.07337	0.00897	0.05087	0.07343	0.0458	0.06647	0.02285	0.07454	0.07458	0.09641	0.03031	0.04034	0.03454	0.03849	0.04144	0.06179	0.05295	0.0672	0.01024	0.03541
17	*Aquifex aeolicus*	0.05813	0.00893	0.04224	0.09674	0.05124	0.06659	0.01555	0.07351	0.09662	0.10561	0.0205	0.0356	0.04013	0.02066	0.0495	0.04767	0.04139	0.07947	0.00937	0.04055
18	*Bradyrhizobium diazoefficiens*	0.12369	0.01047	0.05365	0.05283	0.03696	0.08185	0.02124	0.05213	0.0374	0.09763	0.02633	0.02647	0.05292	0.0316	0.07496	0.05847	0.05299	0.07343	0.01345	0.02154
19	*Bacteroides thetaiotaomicron*	0.06771	0.0139	0.05323	0.06664	0.04671	0.06457	0.01886	0.07206	0.06924	0.0925	0.02863	0.04971	0.03669	0.03432	0.04655	0.06205	0.05549	0.06357	0.01271	0.04485
20	*Rhodopirellula baltica*	0.08859	0.01789	0.05437	0.05502	0.03876	0.07142	0.02544	0.04779	0.03579	0.09128	0.02706	0.03306	0.05542	0.03938	0.07983	0.0791	0.05724	0.06811	0.01593	0.01851
21	*Deinococcus radiodurans*	0.12104	0.00765	0.0497	0.05762	0.03138	0.09042	0.02157	0.03185	0.02746	0.11759	0.02005	0.02303	0.06085	0.04118	0.07576	0.05194	0.05744	0.07643	0.01467	0.02237
22	*Geobacter sulfurreducens*	0.10093	0.01443	0.05256	0.06639	0.03912	0.08291	0.02089	0.05658	0.04475	0.10136	0.02614	0.02788	0.0467	0.0273	0.07124	0.05314	0.05342	0.0785	0.0102	0.02556
23	*Mycoplasma genitalium*	0.05619	0.00984	0.04591	0.05393	0.06123	0.04645	0.01697	0.08423	0.10079	0.10697	0.01714	0.07179	0.02895	0.04575	0.03396	0.06381	0.05272	0.06213	0.00941	0.03183
24	*Thermotoga maritima*	0.05742	0.00847	0.04814	0.08921	0.05199	0.06882	0.01575	0.07238	0.07894	0.10056	0.02553	0.03494	0.03891	0.01979	0.05662	0.05611	0.04436	0.0866	0.01079	0.03467
25	*Gloeobacter violaceus*	0.10994	0.01139	0.04863	0.06323	0.03664	0.08016	0.01917	0.04493	0.0308	0.11385	0.0197	0.02606	0.05579	0.04251	0.0758	0.0543	0.05052	0.07519	0.01542	0.02599
26	*Chlamydia trachomatis*	0.07537	0.01766	0.04345	0.0657	0.04805	0.06224	0.02262	0.06605	0.06105	0.11303	0.02206	0.03348	0.04284	0.04167	0.05089	0.07965	0.04938	0.06615	0.00928	0.0294
27	*Thermodesulfovibrio yellowstonii*	0.05953	0.01214	0.04616	0.08053	0.0499	0.06168	0.01614	0.10021	0.09635	0.0966	0.02422	0.04223	0.0373	0.02707	0.04214	0.05656	0.04469	0.06233	0.00845	0.03577
28	*Chloroflexus aurantiacus*	0.11165	0.00913	0.04999	0.05489	0.03395	0.07508	0.02203	0.06084	0.01942	0.11375	0.0219	0.0257	0.05778	0.04189	0.07782	0.04814	0.05758	0.07602	0.01582	0.02663
29	*Dictyoglomus turgidum*	0.05175	0.00781	0.04687	0.08251	0.05036	0.06613	0.01364	0.09922	0.08789	0.10536	0.02077	0.04564	0.03887	0.0212	0.04543	0.05539	0.04123	0.06691	0.01081	0.04221
30	*Synechocystis sp.*	0.08285	0.01119	0.0493	0.06173	0.04029	0.07021	0.01904	0.0628	0.04545	0.1143	0.02225	0.03915	0.05043	0.05466	0.05253	0.05863	0.05362	0.06619	0.01606	0.02931
31	*Chlamydomonas reinhardtii*	0.15477	0.01565	0.04349	0.0509	0.02383	0.10551	0.0216	0.02349	0.03227	0.08804	0.02077	0.02148	0.07044	0.04449	0.065	0.07292	0.04813	0.0649	0.01302	0.01929
32	*Physcomitrium patens*	0.07484	0.02155	0.04843	0.06009	0.04002	0.06676	0.02582	0.04853	0.05496	0.09671	0.02654	0.03897	0.0497	0.03905	0.06068	0.08663	0.05225	0.06816	0.01383	0.02647
33	*Arabidopsis thaliana*	0.06273	0.01994	0.05248	0.0656	0.04369	0.0642	0.02281	0.05348	0.06493	0.09354	0.02598	0.0433	0.04848	0.03445	0.05462	0.09068	0.05136	0.06676	0.01238	0.02861
34	*Zea mays*	0.09753	0.01959	0.05225	0.05851	0.03508	0.07529	0.02553	0.04017	0.04821	0.0923	0.02409	0.03175	0.0599	0.03554	0.0682	0.08228	0.04772	0.06834	0.01287	0.02485
35	*Leishmania major*	0.11223	0.01983	0.04876	0.0607	0.0325	0.06396	0.02691	0.03322	0.03921	0.09125	0.02512	0.02826	0.05521	0.04032	0.07125	0.08224	0.05839	0.07266	0.01129	0.0267
36	*Paramecium tetraurelia*	0.03532	0.01676	0.04883	0.06691	0.04863	0.03576	0.01871	0.08293	0.09275	0.09577	0.02274	0.06816	0.03047	0.09121	0.03745	0.069	0.04489	0.04457	0.00769	0.04145
37	*Caenorhabditis elegans*	0.064	0.0217	0.05155	0.06312	0.04829	0.05487	0.02308	0.06134	0.06279	0.08519	0.02846	0.04835	0.0496	0.04117	0.05239	0.08043	0.05844	0.06179	0.01113	0.03232
38	*Helobdella robusta*	0.05147	0.02567	0.05646	0.05683	0.04471	0.04953	0.02556	0.06147	0.07335	0.08879	0.02498	0.07006	0.03854	0.04047	0.04827	0.08343	0.05604	0.06032	0.01101	0.03304
39	*Ixodes scapularis*	0.08029	0.02489	0.04887	0.05652	0.04033	0.06947	0.02683	0.03855	0.04984	0.09581	0.02322	0.03298	0.05936	0.03782	0.07051	0.07833	0.05572	0.07014	0.01288	0.02766
40	*Tribolium castaneum*	0.06016	0.0219	0.05065	0.0638	0.04406	0.05509	0.02475	0.05852	0.06969	0.09195	0.02324	0.0507	0.05071	0.03975	0.0555	0.07415	0.05709	0.06388	0.01131	0.0331
41	*Anopheles gambiae*	0.07664	0.02044	0.05163	0.06165	0.03768	0.06547	0.02669	0.04995	0.05456	0.08998	0.02429	0.04352	0.05282	0.04528	0.0567	0.07537	0.05846	0.06528	0.01046	0.03311
42	*Drosophila melanogaster*	0.07466	0.01998	0.05093	0.0619	0.03684	0.0634	0.02629	0.05018	0.05675	0.09008	0.02489	0.04678	0.05417	0.05009	0.05564	0.08075	0.05563	0.05963	0.01036	0.03105
43	*Ciona intestinalis*	0.05675	0.02469	0.05016	0.05876	0.04303	0.05446	0.02561	0.05727	0.06568	0.08966	0.02772	0.05207	0.04579	0.04095	0.05217	0.0817	0.06162	0.06652	0.01237	0.03304
44	*Branchiostoma floridae*	0.0698	0.02301	0.05797	0.06595	0.03545	0.06886	0.02527	0.04221	0.05654	0.084	0.02618	0.03857	0.05582	0.04528	0.05799	0.07488	0.06323	0.06654	0.01258	0.02989
45	*Lepisosteus oculatus*	0.0674	0.02433	0.04895	0.06803	0.03908	0.06396	0.02496	0.04619	0.05868	0.09678	0.023	0.0374	0.05473	0.04547	0.05754	0.08232	0.05429	0.06445	0.01257	0.02987
46	*Danio rerio*	0.06315	0.02343	0.05175	0.06919	0.03743	0.06012	0.02709	0.04646	0.06054	0.09262	0.02514	0.03983	0.05454	0.04697	0.05588	0.08788	0.05704	0.06193	0.01124	0.02777
47	*Oryzias latipes*	0.06564	0.02437	0.04994	0.06538	0.0399	0.0628	0.02691	0.04478	0.05759	0.09657	0.02436	0.03807	0.05582	0.04544	0.05741	0.08627	0.05484	0.06393	0.01244	0.02755
48	*Xenopus laevis*	0.0607	0.02291	0.05061	0.0704	0.03798	0.05895	0.02587	0.05151	0.06417	0.09248	0.02458	0.04323	0.05475	0.04659	0.05256	0.08566	0.05624	0.05961	0.0114	0.0298
49	*Xenopus tropicalis*	0.06238	0.02337	0.04966	0.06975	0.03797	0.06065	0.02556	0.05127	0.06322	0.09356	0.02402	0.04213	0.0554	0.04635	0.05229	0.08451	0.05626	0.05986	0.01158	0.03019
50	*Gallus gallus*	0.07187	0.02402	0.04757	0.06957	0.03727	0.06569	0.0254	0.04561	0.05998	0.09608	0.02312	0.03725	0.05805	0.0453	0.05749	0.08078	0.05228	0.06145	0.01284	0.02839
51	*Macaca mulatta*	0.07029	0.02363	0.04586	0.06801	0.03871	0.06623	0.02679	0.04407	0.0577	0.10017	0.02191	0.03524	0.06247	0.04662	0.05799	0.08248	0.05252	0.05923	0.01343	0.02665
52	*Gorilla gorilla gorilla*	0.07005	0.02395	0.04662	0.06869	0.03786	0.06607	0.02597	0.0443	0.05895	0.10004	0.02366	0.03552	0.06075	0.04644	0.05765	0.08017	0.05231	0.06023	0.01341	0.02736
53	*Pan troglodytes*	0.07068	0.02395	0.04636	0.06835	0.03773	0.06657	0.02598	0.04408	0.05832	0.09981	0.02347	0.03538	0.06142	0.04635	0.05804	0.08043	0.0522	0.0601	0.01343	0.02734
54	*Homo sapiens*	0.07089	0.02315	0.04663	0.06878	0.03692	0.0669	0.02553	0.04258	0.05684	0.10069	0.0249	0.03473	0.06186	0.04699	0.05886	0.08145	0.05215	0.05991	0.0138	0.02644
55	*Canis lupus familiaris*	0.07254	0.02311	0.04653	0.06768	0.03749	0.06831	0.02539	0.04417	0.05761	0.09954	0.02268	0.0351	0.06315	0.04556	0.0594	0.07978	0.0516	0.06017	0.0129	0.0273
56	*Bos taurus*	0.0713	0.02382	0.04636	0.06718	0.03858	0.0669	0.02538	0.0446	0.05717	0.10151	0.02267	0.03483	0.06163	0.04533	0.05833	0.08009	0.05221	0.06139	0.01314	0.0276
57	*Mus musculus*	0.06952	0.024	0.04673	0.06759	0.03784	0.06468	0.02568	0.04322	0.05658	0.10134	0.02516	0.03479	0.06035	0.04689	0.05773	0.08368	0.05318	0.06072	0.0131	0.02724
58	*Rattus norvegicus*	0.06832	0.02372	0.04726	0.06758	0.03902	0.06446	0.02572	0.0453	0.05891	0.10025	0.02371	0.03573	0.059	0.04595	0.05676	0.08176	0.05387	0.06188	0.01267	0.02814
59	*Monodelphis domestica*	0.06337	0.02324	0.04617	0.06856	0.04076	0.06326	0.02571	0.05124	0.06268	0.09972	0.02359	0.03937	0.05748	0.04605	0.05437	0.08115	0.05338	0.05803	0.01277	0.02911
60	*Thalassiosira pseudonana*	0.07733	0.01676	0.06188	0.06541	0.03605	0.07067	0.02319	0.04822	0.05527	0.08523	0.02611	0.04426	0.04645	0.03681	0.05342	0.08862	0.05928	0.06554	0.01169	0.02779
61	*Daphnia magna*	0.06545	0.02211	0.04769	0.05763	0.04517	0.05589	0.02655	0.05446	0.05946	0.09247	0.02491	0.04584	0.05232	0.04182	0.05925	0.08426	0.05903	0.06343	0.01281	0.02944
62	*Plasmodium falciparum*	0.02533	0.01834	0.05968	0.06987	0.04696	0.03194	0.02273	0.09324	0.11662	0.08264	0.02265	0.12125	0.02226	0.02818	0.02996	0.06374	0.04238	0.04154	0.00589	0.05479
63	*Oryza sativa*	0.10101	0.02119	0.04967	0.05574	0.03353	0.07968	0.0267	0.03922	0.0423	0.09099	0.02443	0.02969	0.06261	0.03226	0.07658	0.08252	0.04755	0.0678	0.01409	0.02246
64	*Dictyostelium discoideum*	0.03294	0.01611	0.05093	0.0582	0.04944	0.04721	0.01802	0.08516	0.08054	0.08467	0.01823	0.10532	0.03813	0.04601	0.02889	0.09229	0.05858	0.04532	0.00769	0.03632
65	*Nematostella vectensis*	0.06266	0.0278	0.0502	0.05619	0.04028	0.06259	0.02953	0.05458	0.0609	0.09165	0.0242	0.04234	0.05043	0.03775	0.05596	0.07574	0.06042	0.06676	0.01377	0.03624
66	*Monosiga brevicollis*	0.1039	0.01909	0.05702	0.05765	0.03427	0.0609	0.02881	0.03788	0.039	0.10162	0.0233	0.0333	0.05456	0.0458	0.06438	0.07597	0.05964	0.06657	0.01199	0.02435
67	*Phytophthora ramorum*	0.09438	0.01622	0.05824	0.06615	0.03737	0.06446	0.0231	0.03814	0.05226	0.09103	0.02554	0.03314	0.04708	0.04151	0.06074	0.07889	0.05956	0.07304	0.01248	0.02666
68	*Giardia intestinalis*	0.08114	0.02487	0.05416	0.06029	0.03465	0.05496	0.02393	0.0567	0.05261	0.102	0.02345	0.03905	0.04728	0.03946	0.05605	0.08587	0.06189	0.06132	0.00718	0.03311
69	*Cryptococcus neoformans*	0.08439	0.01117	0.05258	0.06512	0.03483	0.07231	0.02338	0.04871	0.05388	0.08853	0.02211	0.03394	0.0654	0.03767	0.05888	0.08756	0.05785	0.06111	0.01411	0.02647
70	*Candida albicans*	0.05197	0.01137	0.05762	0.06384	0.04413	0.05221	0.02096	0.07141	0.07562	0.09152	0.01838	0.06506	0.04436	0.04456	0.03835	0.08693	0.06096	0.0548	0.00984	0.0361
71	*Ustilago maydis*	0.10327	0.01213	0.05634	0.05243	0.03353	0.06561	0.0263	0.04315	0.04756	0.08919	0.02011	0.0343	0.05665	0.04384	0.06263	0.09638	0.06023	0.05991	0.01242	0.02402
72	*Yarrowia lipolytica*	0.07856	0.01291	0.05863	0.06502	0.03924	0.06433	0.02381	0.04992	0.06197	0.08679	0.02347	0.04042	0.05265	0.04181	0.05065	0.07893	0.06037	0.06737	0.01231	0.03084
73	*Schizosaccharomyces pombe*	0.06338	0.0154	0.05183	0.06532	0.04602	0.05104	0.02269	0.06145	0.06787	0.09658	0.02207	0.05059	0.04659	0.03825	0.05073	0.08978	0.05405	0.06082	0.01123	0.03431
74	*Phaeosphaeria nodorum*	0.08965	0.01512	0.0546	0.05947	0.03661	0.06767	0.02579	0.04783	0.05028	0.08506	0.02565	0.03634	0.05971	0.0398	0.06176	0.07982	0.0613	0.06069	0.01536	0.0275
75	*Aspergillus fumigatus*	0.08656	0.01378	0.0541	0.06038	0.03735	0.06902	0.02433	0.0496	0.04711	0.0905	0.0221	0.03559	0.05949	0.03955	0.06283	0.08212	0.05931	0.06276	0.01493	0.02861
76	*Neurospora crassa*	0.08676	0.01307	0.05429	0.06272	0.03435	0.07338	0.02471	0.04442	0.05181	0.0828	0.02244	0.03684	0.06379	0.04134	0.06171	0.08237	0.06223	0.06061	0.01412	0.02625
77	*Trichomonas vaginalis*	0.05199	0.01552	0.05795	0.07639	0.04988	0.04147	0.01959	0.07515	0.08908	0.0833	0.02467	0.06123	0.04428	0.04052	0.04005	0.07025	0.05467	0.05306	0.0088	0.04214
78	*Puccinia graminis*	0.07212	0.01573	0.0524	0.05697	0.0361	0.06061	0.0281	0.05123	0.05661	0.09193	0.02135	0.04427	0.06631	0.04628	0.05736	0.09364	0.05969	0.05225	0.01279	0.02424
79	*Saccharomyces cerevisiae*	0.05717	0.01413	0.0555	0.06335	0.04529	0.0515	0.02196	0.06456	0.07492	0.09471	0.0224	0.05807	0.04343	0.03951	0.04671	0.0861	0.05799	0.05745	0.01078	0.03444
80	*Sclerotinia sclerotiorum*	0.07255	0.01488	0.05248	0.06552	0.03796	0.06927	0.02306	0.05681	0.05839	0.0845	0.02715	0.04194	0.05239	0.0367	0.05929	0.08422	0.05947	0.0593	0.01489	0.02923
81	*Batrachochytrium dendrobatidis*	0.06999	0.01564	0.0559	0.05387	0.03911	0.05439	0.02688	0.06248	0.0596	0.09164	0.02551	0.04485	0.04863	0.04543	0.04799	0.09136	0.06376	0.06177	0.01104	0.03017

Table 2. Average Amino Acid Composition for Each Exome

Table 2 presents the average amino acid composition calculated for each of the 81 listed species, following the methods described in the Subjects and Methods section. Each amino acid is displayed using its three-letter abbreviation, and these values represent the exome-matched diet amino acid compositions as originally described in the 2017 paper.

No.	Scientific Name	Domain	Cell Organization	Food Name with Minimum Distance (Original)	Food Name (Translated)	Minimum Distance	Distance to Chicken egg	Difference in Distance
1	*Halobacterium salinarum*	archaea	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (rough wood ear mushroom), raw	0.17901	0.33652	0.15751
2	*Thermococcus kodakarensis*	archaea	unicellular	＜鳥肉類＞　がちょう　フォアグラ　ゆで	Goose foie gras, boiled	0.19505	0.26320	0.06814
3	*Methanosarcina acetivorans*	archaea	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.15668	0.18091	0.02422
4	*Methanocaldococcus jannaschii*	archaea	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.29952	0.35872	0.05920
5	*Saccharolobus solfataricus*	archaea	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.24545	0.28177	0.03632
6	*Korarchaeum cryptofilum*	archaea	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.24776	0.26271	0.01495
7	*Nitrosopumilus maritimus*	archaea	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.20435	0.22397	0.01961
8	*Mycobacterium tuberculosis*	bacteria	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (wood ear mushroom), raw	0.21200	0.38494	0.17295
9	*Escherichia coli*	bacteria	unicellular	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.16533	0.20999	0.04466
10	*Helicobacter pylori*	bacteria	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.22286	0.24350	0.02064
11	*Streptomyces coelicolor*	bacteria	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (rough wood ear mushroom), raw	0.24709	0.42026	0.17318
12	*Neisseria meningitidis*	bacteria	unicellular	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.13717	0.20343	0.06626
13	*Leptospira interrogans*	bacteria	unicellular	あずき　あん　さらしあん　（乾燥あん）	Azuki bean paste, “sarashi-an” (dried, strained)	0.22042	0.23642	0.01600
14	*Fusobacterium nucleatum*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.30023	0.33912	0.03889
15	*Pseudomonas aeruginosa*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.24326	0.33598	0.09273
16	*Bacillus subtilis*	bacteria	unicellular	＜魚類＞　（いわし類）　缶詰　アンチョビ	Anchovies, canned	0.14853	0.19155	0.04302
17	*Aquifex aeolicus*	bacteria	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.22422	0.29108	0.06686
18	*Bradyrhizobium diazoefficiens*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.19091	0.31555	0.12465
19	*Bacteroides thetaiotaomicron*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.15332	0.19003	0.03671
20	*Rhodopirellula baltica*	bacteria	unicellular	（きくらげ類）　しろきくらげ　ゆで	White wood ear mushroom, boiled	0.19690	0.21847	0.02157
21	*Deinococcus radiodurans*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.23913	0.36567	0.12653
22	*Geobacter sulfurreducens*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.16630	0.26802	0.10172
23	*Mycoplasma genitalium*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.26663	0.30055	0.03392
24	*Thermotoga maritima*	bacteria	unicellular	＜魚類＞　（いわし類）　缶詰　アンチョビ	Anchovies, canned	0.21831	0.24777	0.02946
25	*Gloeobacter violaceus*	bacteria	unicellular	ひじき　ほしひじき　ステンレス釜　ゆで	Hijiki seaweed (dried), boiled in a stainless steel pot	0.21696	0.30981	0.09285
26	*Chlamydia trachomatis*	bacteria	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.15976	0.17156	0.01179
27	*Thermodesulfovibrio yellowstonii*	bacteria	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.25777	0.29709	0.03932
28	*Chloroflexus aurantiacus*	bacteria	unicellular	ひじき　ほしひじき　ステンレス釜　ゆで	Hijiki seaweed (dried), boiled in a stainless steel pot	0.24502	0.34906	0.10403
29	*Dictyoglomus turgidum*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.26892	0.30949	0.04057
30	*Synechocystis sp.*	bacteria	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.15967	0.21706	0.05738
31	*Chlamydomonas reinhardtii*	eukaryota	unicellular	ふのり　素干し	Funori seaweed, sun-dried	0.19346	0.41326	0.21980
32	*Physcomitrium patens*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11573	0.12156	0.00582
33	*Arabidopsis thaliana*	eukaryota	**multicellular**	うずら卵　全卵　生	Quail egg, whole, raw	0.10327	0.10879	0.00552
34	*Zea mays*	eukaryota	**multicellular**	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.16682	0.19403	0.02721
35	*Leishmania major*	eukaryota	unicellular	なめこ　カットなめこ　生	Nameko mushroom (cut), raw	0.19080	0.22590	0.03510
36	*Paramecium tetraurelia*	eukaryota	unicellular	＜和生菓子・和半生菓子類＞　まんじゅう　くずまんじゅう　こしあん入り	Kuzu manju (with strained red bean paste filling)	0.25026	0.33125	0.08100
37	*Caenorhabditis elegans*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11068	0.11342	0.00274
38	*Helobdella robusta*	eukaryota	**multicellular**	らいまめ　全粒　ゆで	Lima beans, whole, boiled	0.17143	0.19213	0.02070
39	*Ixodes scapularis*	eukaryota	**multicellular**	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.17354	0.18360	0.01006
40	*Tribolium castaneum*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.12557	0.13754	0.01197
41	*Anopheles gambiae*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.10830	0.11186	0.00356
42	*Drosophila melanogaster*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.09526	0.09835	0.00309
43	*Ciona intestinalis*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.10877	0.12900	0.02023
44	*Branchiostoma floridae*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.12305	0.12752	0.00448
45	*Lepisosteus oculatus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.11719	0.11719	0.00000
46	*Danio rerio*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10583	0.10583	0.00000
47	*Oryzias latipes*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.11735	0.11735	0.00000
48	*Xenopus laevis*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10934	0.10934	0.00000
49	*Xenopus tropicalis*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10978	0.10978	0.00000
50	*Gallus gallus*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.12249	0.12446	0.00197
51	*Macaca mulatta*	eukaryota	**multicellular**	鶏卵　全卵　ポーチドエッグ	Poached egg (whole chicken egg)	0.14976	0.15037	0.00061
52	*Gorilla gorilla gorilla*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.14360	0.14516	0.00156
53	*Pan troglodytes*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.14555	0.14707	0.00151
54	*Homo sapiens*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.15052	0.15190	0.00138
55	*Canis lupus familiaris*	eukaryota	**multicellular**	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.14544	0.15293	0.00750
56	*Bos taurus*	eukaryota	**multicellular**	鶏卵　全卵　ポーチドエッグ	Poached egg (whole chicken egg)	0.14985	0.15090	0.00104
57	*Mus musculus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.14382	0.14382	0.00000
58	*Rattus norvegicus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.13612	0.13612	0.00000
59	*Monodelphis domestica*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.13631	0.13631	0.00000
60	*Thalassiosira pseudonana*	eukaryota	unicellular	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.10134	0.10272	0.00138
61	*Daphnia magna*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11428	0.11945	0.00518
62	*Plasmodium falciparum*	eukaryota	unicellular	＜スナック類＞　ポテトチップス　成形ポテトチップス	Snack: Potato chips (processed/formed type)	0.37345	0.47958	0.10613
63	*Oryza sativa*	eukaryota	**multicellular**	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.19315	0.24398	0.05083
64	*Dictyostelium discoideum*	eukaryota	unicellular	らいまめ　全粒　ゆで	Lima beans, whole, boiled	0.24019	0.33630	0.09611
65	*Nematostella vectensis*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11243	0.14331	0.03088
66	*Monosiga brevicollis*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.15799	0.19945	0.04146
67	*Phytophthora ramorum*	eukaryota	unicellular	鶏卵　全卵　いり	Chicken egg, whole, scrambled	0.12735	0.13394	0.00658
68	*Giardia intestinalis*	eukaryota	unicellular	うずら卵　全卵　生	Quail egg, whole, raw	0.11722	0.12489	0.00767
69	*Cryptococcus neoformans*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.14473	0.15295	0.00822
70	*Candida albicans*	eukaryota	unicellular	らいまめ　全粒　乾	Lima beans, whole, dried	0.16923	0.20023	0.03100
71	*Ustilago maydis*	eukaryota	unicellular	（きくらげ類）　きくらげ　乾	Wood ear mushroom, dried	0.16378	0.17618	0.01240
72	*Yarrowia lipolytica*	eukaryota	unicellular	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.09449	0.09830	0.00381
73	*Schizosaccharomyces pombe*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.10983	0.12324	0.01341
74	*Phaeosphaeria nodorum*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.13406	0.15251	0.01845
75	*Aspergillus fumigatus*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.14143	0.15494	0.01351
76	*Neurospora crassa*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.13414	0.15911	0.02497
77	*Trichomonas vaginalis*	eukaryota	unicellular	あずき　あん　さらしあん　（乾燥あん）	Azuki bean paste, “sarashi-an” (dried, strained)	0.21071	0.23685	0.02615
78	*Puccinia graminis*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.12502	0.14652	0.02150
79	*Saccharomyces cerevisiae*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.14833	0.16198	0.01365
80	*Sclerotinia sclerotiorum*	eukaryota	unicellular	鶏卵　全卵　生	Chicken egg, whole, raw	0.12031	0.12031	0.00000
81	*Batrachochytrium dendrobatidis*	eukaryota	unicellular	うずら卵　全卵　生	Quail egg, whole, raw	0.10212	0.11412	0.01200

Table 3. Closest Food Items to Each Exome-Matched Composition

Table 3 lists, for each organism, the food item that exhibited the minimum distance to that organism’s exome-matched amino acid composition, along with the corresponding distance value. In many species, raw whole chicken eggs emerged as the closest match. I therefore also calculated the distance from each organism’s exome-matched composition to the egg composition and the difference between these distances. If this difference was sufficiently small, I deemed raw whole chicken eggs a practical approximation of that organism’s exome-matched diet.

No.	Scientific Name	Domain	Cell Organization	Food Name with Minimum Distance (Original)	Food Name (Translated)	Minimum Distance	Distance to Chicken egg	Difference in Distance
46	*Danio rerio*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10583	0.10583	0.00000
48	*Xenopus laevis*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10934	0.10934	0.00000
49	*Xenopus tropicalis*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.10978	0.10978	0.00000
45	*Lepisosteus oculatus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.11719	0.11719	0.00000
47	*Oryzias latipes*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.11735	0.11735	0.00000
80	*Sclerotinia sclerotiorum*	eukaryota	unicellular	鶏卵　全卵　生	Chicken egg, whole, raw	0.12031	0.12031	0.00000
58	*Rattus norvegicus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.13612	0.13612	0.00000
59	*Monodelphis domestica*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.13631	0.13631	0.00000
57	*Mus musculus*	eukaryota	**multicellular**	鶏卵　全卵　生	Chicken egg, whole, raw	0.14382	0.14382	0.00000
51	*Macaca mulatta*	eukaryota	**multicellular**	鶏卵　全卵　ポーチドエッグ	Poached egg (whole chicken egg)	0.14976	0.15037	0.00061
56	*Bos taurus*	eukaryota	**multicellular**	鶏卵　全卵　ポーチドエッグ	Poached egg (whole chicken egg)	0.14985	0.15090	0.00104
54	*Homo sapiens*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.15052	0.15190	0.00138
60	*Thalassiosira pseudonana*	eukaryota	unicellular	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.10134	0.10272	0.00138
53	*Pan troglodytes*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.14555	0.14707	0.00151
52	*Gorilla gorilla gorilla*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.14360	0.14516	0.00156
50	*Gallus gallus*	eukaryota	**multicellular**	鶏卵　全卵　加糖全卵	Chicken egg, whole (sweetened)	0.12249	0.12446	0.00197
37	*Caenorhabditis elegans*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11068	0.11342	0.00274
42	*Drosophila melanogaster*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.09526	0.09835	0.00309
41	*Anopheles gambiae*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.10830	0.11186	0.00356
72	*Yarrowia lipolytica*	eukaryota	unicellular	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.09449	0.09830	0.00381
44	*Branchiostoma floridae*	eukaryota	**multicellular**	鶏卵　全卵　水煮缶詰	Chicken egg, whole, canned in water (boiled)	0.12305	0.12752	0.00448
61	*Daphnia magna*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11428	0.11945	0.00518
33	*Arabidopsis thaliana*	eukaryota	**multicellular**	うずら卵　全卵　生	Quail egg, whole, raw	0.10327	0.10879	0.00552
32	*Physcomitrium patens*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11573	0.12156	0.00582
67	*Phytophthora ramorum*	eukaryota	unicellular	鶏卵　全卵　いり	Chicken egg, whole, scrambled	0.12735	0.13394	0.00658
55	*Canis lupus familiaris*	eukaryota	**multicellular**	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.14544	0.15293	0.00750
68	*Giardia intestinalis*	eukaryota	unicellular	うずら卵　全卵　生	Quail egg, whole, raw	0.11722	0.12489	0.00767
69	*Cryptococcus neoformans*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.14473	0.15295	0.00822
39	*Ixodes scapularis*	eukaryota	**multicellular**	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.17354	0.18360	0.01006
26	*Chlamydia trachomatis*	bacteria	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.15976	0.17156	0.01179
40	*Tribolium castaneum*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.12557	0.13754	0.01197
81	*Batrachochytrium dendrobatidis*	eukaryota	unicellular	うずら卵　全卵　生	Quail egg, whole, raw	0.10212	0.11412	0.01200
71	*Ustilago maydis*	eukaryota	unicellular	（きくらげ類）　きくらげ　乾	Wood ear mushroom, dried	0.16378	0.17618	0.01240
73	*Schizosaccharomyces pombe*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.10983	0.12324	0.01341
75	*Aspergillus fumigatus*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.14143	0.15494	0.01351
79	*Saccharomyces cerevisiae*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.14833	0.16198	0.01365
6	*Korarchaeum cryptofilum*	archaea	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.24776	0.26271	0.01495
13	*Leptospira interrogans*	bacteria	unicellular	あずき　あん　さらしあん　（乾燥あん）	Azuki bean paste, “sarashi-an” (dried, strained)	0.22042	0.23642	0.01600
74	*Phaeosphaeria nodorum*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.13406	0.15251	0.01845
7	*Nitrosopumilus maritimus*	archaea	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.20435	0.22397	0.01961
43	*Ciona intestinalis*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.10877	0.12900	0.02023
10	*Helicobacter pylori*	bacteria	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.22286	0.24350	0.02064
38	*Helobdella robusta*	eukaryota	**multicellular**	らいまめ　全粒　ゆで	Lima beans, whole, boiled	0.17143	0.19213	0.02070
78	*Puccinia graminis*	eukaryota	unicellular	鶏卵　卵黄　乾燥卵黄	Chicken egg, yolk, dried	0.12502	0.14652	0.02150
20	*Rhodopirellula baltica*	bacteria	unicellular	（きくらげ類）　しろきくらげ　ゆで	White wood ear mushroom, boiled	0.19690	0.21847	0.02157
3	*Methanosarcina acetivorans*	archaea	unicellular	うずら卵　水煮缶詰	Quail egg, canned in water	0.15668	0.18091	0.02422
76	*Neurospora crassa*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.13414	0.15911	0.02497
77	*Trichomonas vaginalis*	eukaryota	unicellular	あずき　あん　さらしあん　（乾燥あん）	Azuki bean paste, “sarashi-an” (dried, strained)	0.21071	0.23685	0.02615
34	*Zea mays*	eukaryota	**multicellular**	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.16682	0.19403	0.02721
24	*Thermotoga maritima*	bacteria	unicellular	＜魚類＞　（いわし類）　缶詰　アンチョビ	Anchovies, canned	0.21831	0.24777	0.02946
65	*Nematostella vectensis*	eukaryota	**multicellular**	うずら卵　水煮缶詰	Quail egg, canned in water	0.11243	0.14331	0.03088
70	*Candida albicans*	eukaryota	unicellular	らいまめ　全粒　乾	Lima beans, whole, dried	0.16923	0.20023	0.03100
23	*Mycoplasma genitalium*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.26663	0.30055	0.03392
35	*Leishmania major*	eukaryota	unicellular	なめこ　カットなめこ　生	Nameko mushroom (cut), raw	0.19080	0.22590	0.03510
5	*Saccharolobus solfataricus*	archaea	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.24545	0.28177	0.03632
19	*Bacteroides thetaiotaomicron*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.15332	0.19003	0.03671
14	*Fusobacterium nucleatum*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.30023	0.33912	0.03889
27	*Thermodesulfovibrio yellowstonii*	bacteria	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.25777	0.29709	0.03932
29	*Dictyoglomus turgidum*	bacteria	unicellular	はやとうり　果実　白色種　生	Chayote (white variety), raw	0.26892	0.30949	0.04057
66	*Monosiga brevicollis*	eukaryota	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.15799	0.19945	0.04146
16	*Bacillus subtilis*	bacteria	unicellular	＜魚類＞　（いわし類）　缶詰　アンチョビ	Anchovies, canned	0.14853	0.19155	0.04302
9	*Escherichia coli*	bacteria	unicellular	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.16533	0.20999	0.04466
63	*Oryza sativa*	eukaryota	**multicellular**	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.19315	0.24398	0.05083
30	*Synechocystis sp.*	bacteria	unicellular	＜魚類＞　にしん　かずのこ　塩蔵　水戻し	Herring roe (“kazunoko”), salted, soaked (desalinated)	0.15967	0.21706	0.05738
4	*Methanocaldococcus jannaschii*	archaea	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.29952	0.35872	0.05920
12	*Neisseria meningitidis*	bacteria	unicellular	＜鳥肉類＞　にわとり　［副品目］　肝臓　生	Chicken liver (raw)	0.13717	0.20343	0.06626
17	*Aquifex aeolicus*	bacteria	unicellular	＜魚類＞　かます　焼き	Barracuda, grilled	0.22422	0.29108	0.06686
2	*Thermococcus kodakarensis*	archaea	unicellular	＜鳥肉類＞　がちょう　フォアグラ　ゆで	Goose foie gras, boiled	0.19505	0.26320	0.06814
36	*Paramecium tetraurelia*	eukaryota	unicellular	＜和生菓子・和半生菓子類＞　まんじゅう　くずまんじゅう　こしあん入り	Kuzu manju (with strained red bean paste filling)	0.25026	0.33125	0.08100
15	*Pseudomonas aeruginosa*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.24326	0.33598	0.09273
25	*Gloeobacter violaceus*	bacteria	unicellular	ひじき　ほしひじき　ステンレス釜　ゆで	Hijiki seaweed (dried), boiled in a stainless steel pot	0.21696	0.30981	0.09285
64	*Dictyostelium discoideum*	eukaryota	unicellular	らいまめ　全粒　ゆで	Lima beans, whole, boiled	0.24019	0.33630	0.09611
22	*Geobacter sulfurreducens*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　ゆで	Wakame seaweed (parboiled & salted), desalinated, boiled	0.16630	0.26802	0.10172
28	*Chloroflexus aurantiacus*	bacteria	unicellular	ひじき　ほしひじき　ステンレス釜　ゆで	Hijiki seaweed (dried), boiled in a stainless steel pot	0.24502	0.34906	0.10403
62	*Plasmodium falciparum*	eukaryota	unicellular	＜スナック類＞　ポテトチップス　成形ポテトチップス	Snack: Potato chips (processed/formed type)	0.37345	0.47958	0.10613
18	*Bradyrhizobium diazoefficiens*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.19091	0.31555	0.12465
21	*Deinococcus radiodurans*	bacteria	unicellular	わかめ　湯通し塩蔵わかめ　塩抜き　生	Wakame seaweed (parboiled & salted), desalinated, raw	0.23913	0.36567	0.12653
1	*Halobacterium salinarum*	archaea	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (rough wood ear mushroom), raw	0.17901	0.33652	0.15751
8	*Mycobacterium tuberculosis*	bacteria	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (wood ear mushroom), raw	0.21200	0.38494	0.17295
11	*Streptomyces coelicolor*	bacteria	unicellular	（きくらげ類）　あらげきくらげ　生	Arage kikurage (rough wood ear mushroom), raw	0.24709	0.42026	0.17318
31	*Chlamydomonas reinhardtii*	eukaryota	unicellular	ふのり　素干し	Funori seaweed, sun-dried	0.19346	0.41326	0.21980

Table 4. Reordered Species by Relative Distance to Chicken Eggs

To facilitate more intuitive inspection, I reordered the species in Table 4 according to the descending difference between their distance to raw whole chicken eggs and their distance to their best-matching food. This reordering highlights how close each species' exome composition is to raw whole chicken eggs relative to its top-ranked food item. As a result, for most species—particularly multicellular eukaryotic organisms—raw whole chicken eggs could be considered their practical exome-matched diet among the listed foods.

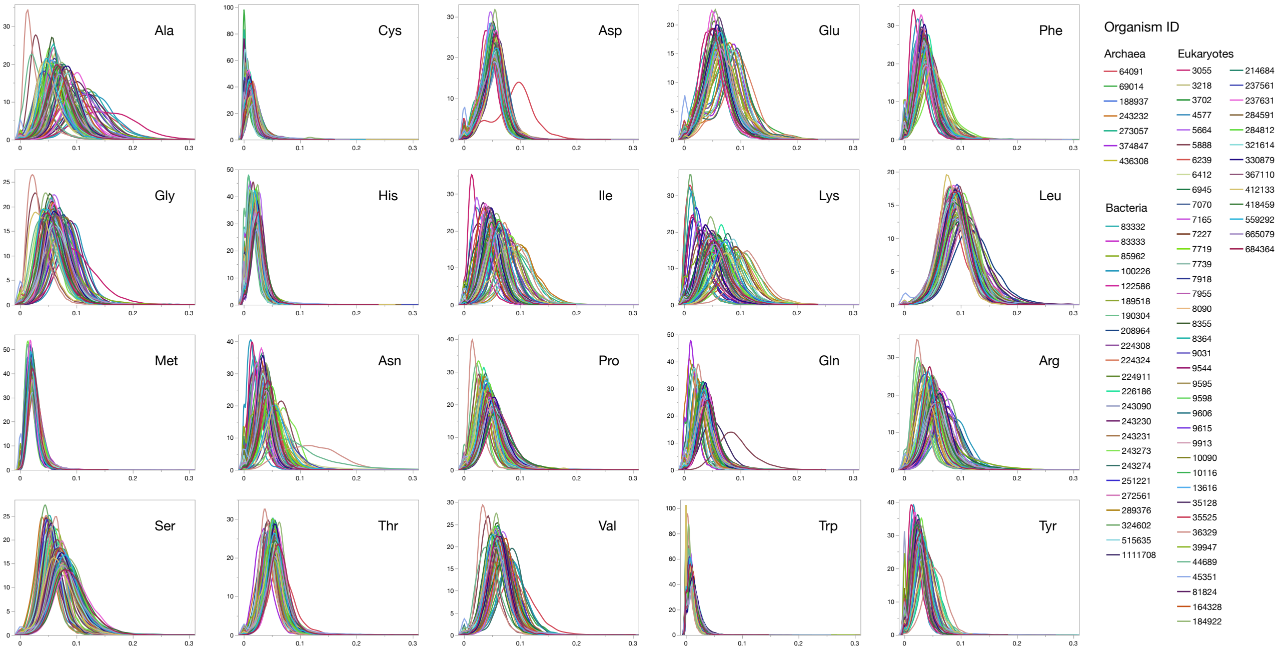

Figure 1. Distributions of Amino Acid Residue Compositions in the Exomic Proteins of 81 Species from the Three Domains of Life

Figure 1 compares the distributions of amino acid compositions, by species and by amino acid, for all 81 organisms included in this study. These distributions range from amino acids—such as Met, His, and Leu—that appear similarly distributed across species, to others—like Ala, Ile, and Lys—whose distributions show considerable variation among different organisms.

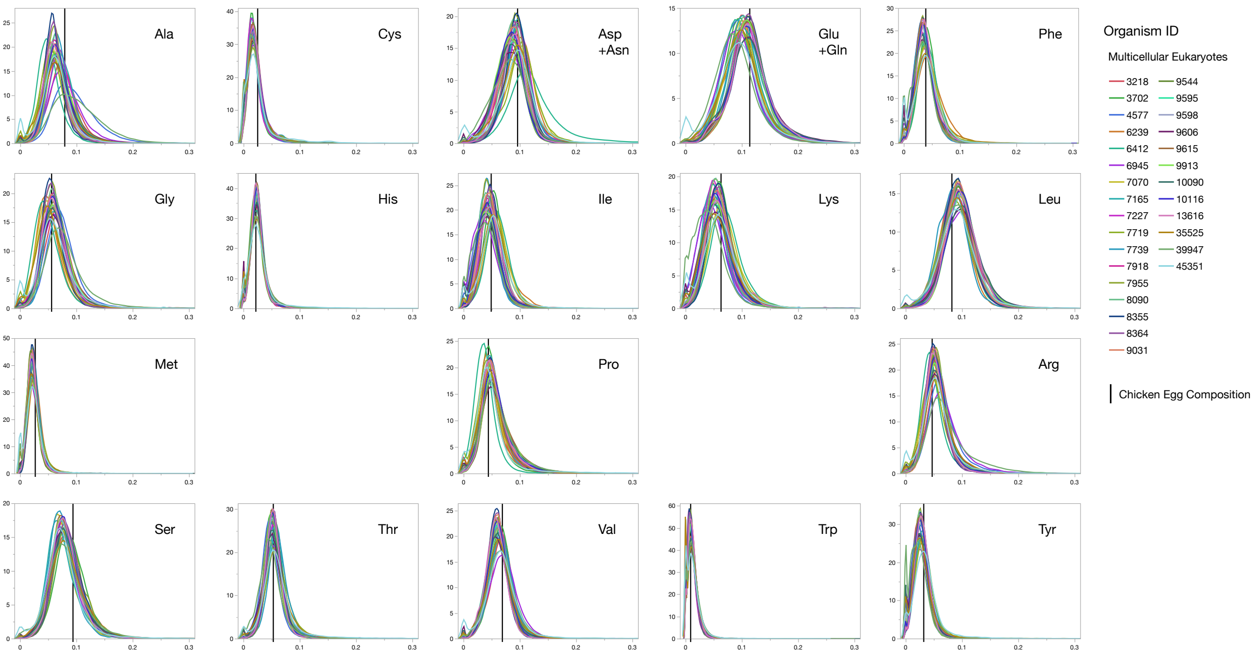

Figure 2. Distributions of Amino Acid Residue Compositions in the Exomic Proteins of Multicellular Eukaryotes and the Chicken Egg Composition

Because the amino acid composition distance between chicken eggs and the average exome of most multicellular eukaryotes was particularly small, Figure 2 illustrates the amino acid composition distributions of these eukaryotes while marking the chicken egg composition with a vertical black line. The results indicate that these exomes cluster closely together, with the black line (i.e., the chicken egg composition) positioned near the center of their distributions.

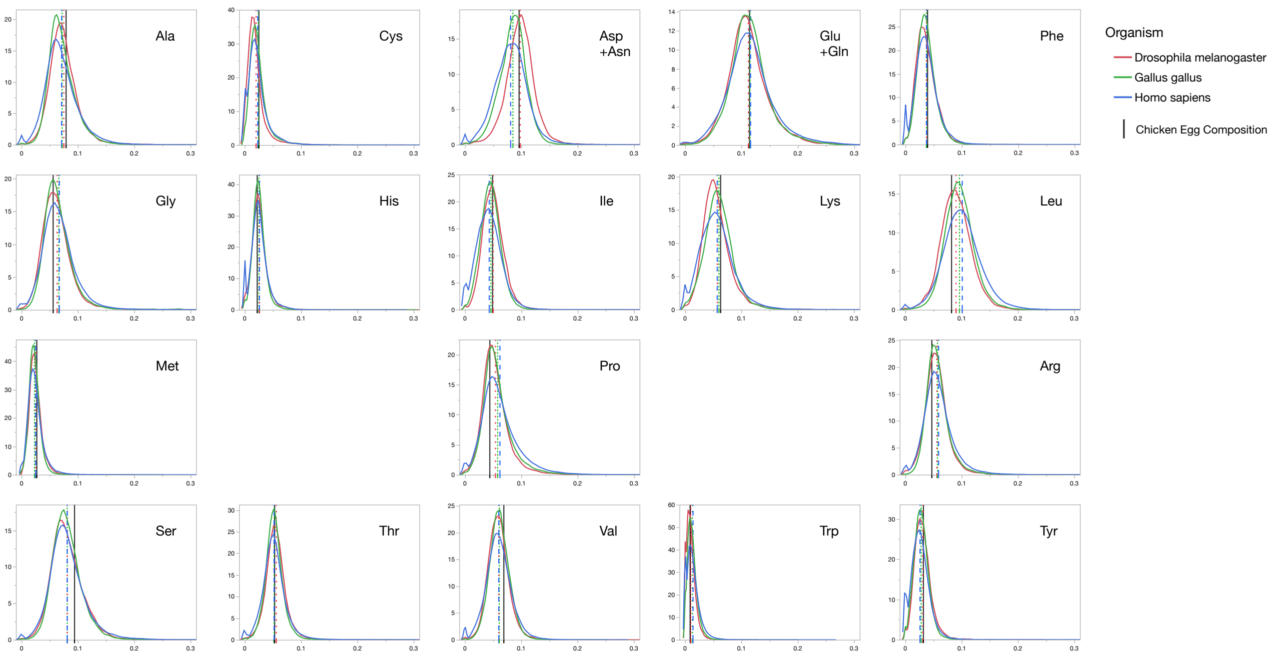

Figure 3. Distributions and Averages of Amino Acid Residue Compositions in the Exomic Proteins of Three Species and the Chicken Egg Composition

Figure 3 focuses on three specific species: *Drosophila melanogaster* (the fruit fly), *Gallus gallus* (the chicken), and *Homo sapiens* (humans). Each species' exomic amino acid distribution is plotted, with dashed vertical lines representing the mean composition of that species and a solid black line marking the chicken egg composition. As observed in Figure 2, these distributions overlap considerably, and their means are very similar. However, the dashed lines lie somewhat closer together than they do to the solid black line, suggesting that these exomes resemble each other more closely than they resemble the chicken egg composition itself.

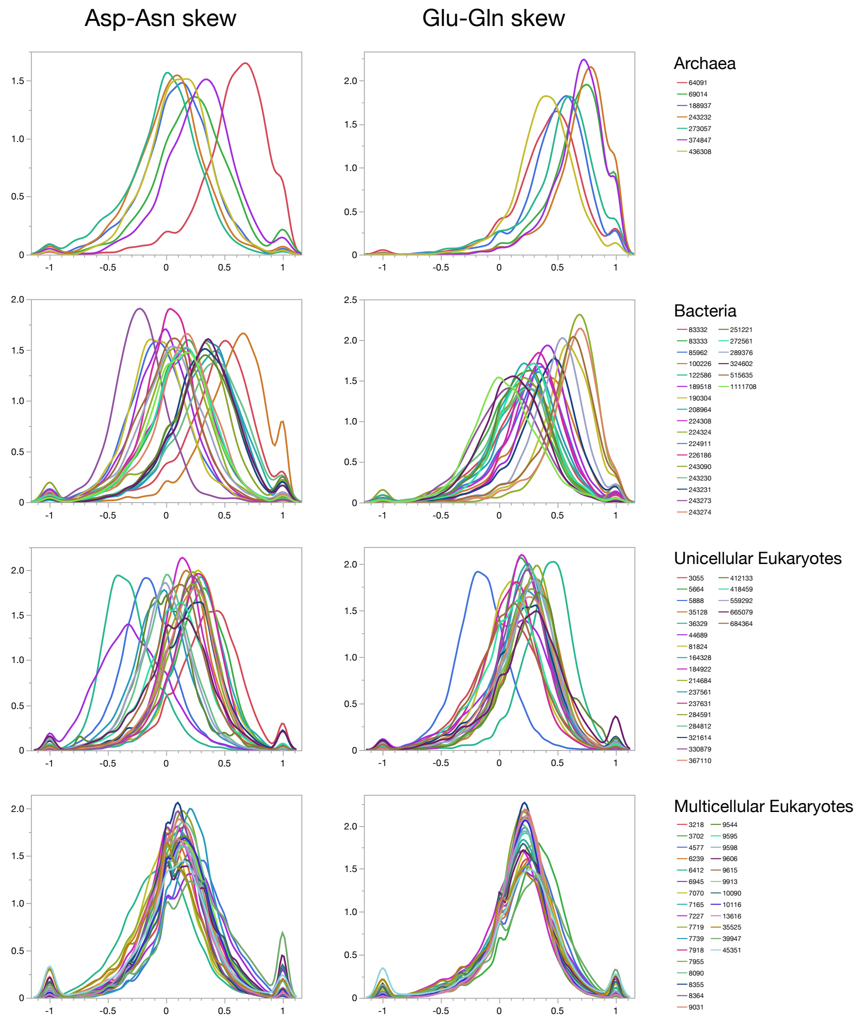

Figure 4. Distributions of Asp–Asn and Glu–Gln Skew across Four Organism Groups

Figure 4 shows the distributions of two skew indices—Asp–Asn (left) and Glu–Gln (right)—for Archaea, Bacteria, unicellular eukaryotes, and multicellular eukaryotes. Each panel plots the skew index for individual proteins within each domain or group. Whereas Archaea, Bacteria, and unicellular eukaryotes display considerable variation in both skew values, multicellular eukaryotes exhibit notably uniform distributions, suggesting that the relative usage of Asp versus Asn and Glu versus Gln is effectively constant in their proteomes.