

# Autophagy: possible origin of essential amino acids

Genshiro Esumi

Department of Pediatric Surgery, Hospital of the University of Occupational and Environmental Health, 1-1, Iseigaoka, Yahatanishi-ku, Kitakyushu-shi, Fukuoka, 807-8556, Japan

## Abstract

Many people know that essential amino acids are amino acids that humans cannot synthesize<sup>1</sup>. However, far fewer know that these components are typical of most heterotrophic eukaryotes<sup>23</sup>. No one has directly explained why organisms have maintained such amino acid requirements<sup>345</sup>. This paper presents a convincing hypothesis that can clearly explain these questions. In an analysis of a food composition table, the first principal component of the amino acid compositions of daily foods was found to separate essential and non-essential amino acids<sup>6</sup>. Regarding all foods from eukaryotes, this finding may indicate that all eukaryotes have common components rich in essential amino acids. Previous studies have indicated that yeast, a eukaryote, can use its cytoplasmic components as a primary source of amino acids by autophagy<sup>7-11</sup>, and all terrestrial organism cells showed very similar amino acid compositions<sup>12-14</sup>. If all eukaryotic cells depend on their cytoplasmic components as primary amino acid sources and if all eukaryotes maintain cytoplasmic components rich in essential amino acids, it would be reasonable for heterotrophic eukaryotes to have lost the need to synthesize amino acids that are stored in large amounts inside their cells. I therefore concluded that the autophagy of cytoplasmic components was a possible origin of essential amino acids.

## Essential amino acids: the mystery of conservation

Adult humans cannot synthesize nine of the 20 amino acids that compose their body proteins<sup>12345</sup>. These have been classified as essential amino acids. This means that these nine amino acids must be obtained from the diet<sup>12</sup>. Such nutritional requirements are not rare in other organisms<sup>23</sup>; in fact, all mammals and vertebrates are known to share almost the same requirements<sup>23</sup>. Surprisingly, not only such animals but also some Protozoa<sup>2</sup>, Amoebozoa<sup>15</sup>, Alveolata<sup>16</sup> and Myxomycota<sup>17</sup> share similar amino acid

requirements. Ultimately, almost all heterotrophic eukaryotes share almost the same amino acid requirements. However, the reason so many organisms lost their ability to synthesize these similar amino acids and how the disability has remained so long are still mysteries<sup>3</sup>.

To date, the most popular theory seems to be that the ancient common ancestor of present eukaryotes lost this ability<sup>4</sup> because of the abundance of essential amino acids in its situation<sup>3</sup>. However, this theory does not explain why such a phenotype has been maintained for such a long evolutionary time.

How did they originate? How have they been conserved?

### **Our foods and essential amino acids**

If our circumstances defined our essential amino acids, the relevant aspects would mainly be what we have eaten, our foods. Thus, I analysed food composition tables. The standard food composition table in Japan 2015 is a set of food composition tables that are publicly accessible on the Japanese government website<sup>618</sup> and report various nutritional components of daily foods and ingredients. The amino acid composition table consists of the total amino acid compositions of 1558 daily foods. Originally, they included 18 food classifications, but for the analysis, I reclassified them into 3 groups: (1) animal foods (n=620), (2) plant foods (n=657), and (3) others (n=281). Principal component analysis was performed on all listed foods, and each food was plotted on a graph. The three groups were separately plotted (Fig. 1).

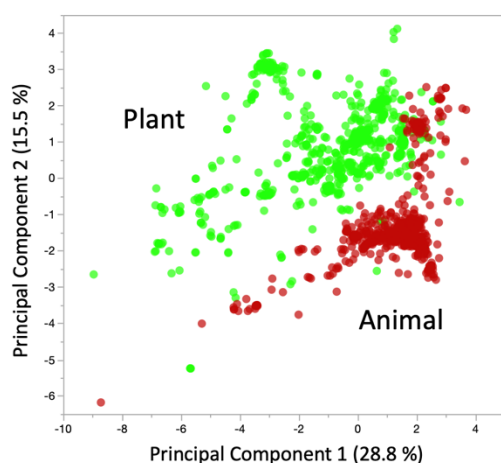


Figure 1. Principal component plot for amino acid composition of listed food

The amino acid composition of each food is plotted by its first and second principal components. Plant foods (up, light green dots) and animal foods (down, dark red dots)

were plotted separately. Other foods plotted were spread across both regions (not shown for clarity).

The first principal component's contribution ratio was 28.8%, and that of the second was 15.5%. The first principal component vector clearly separated essential amino acids from non-essential ones. The only exception was tyrosine (Fig. 2).

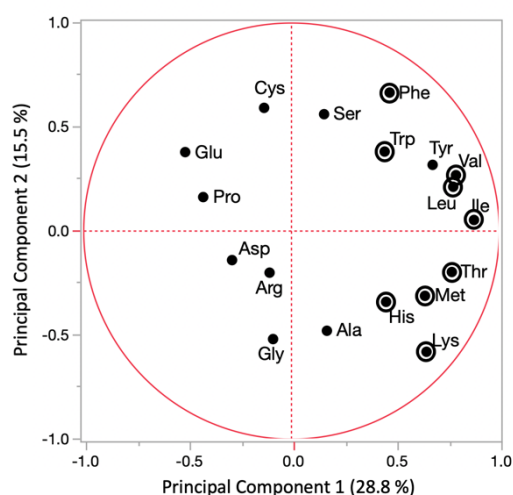


Figure 2. Vector plot of the first and second principal components

The amino acid vectors of the first and second principal components are plotted. Nine human essential amino acids are circled. Essential and non-essential amino acids were separated by the first principal component except for tyrosine (Tyr).

Not only the first principal component vector of all listed foods but also that of each of the three groups independently showed positive correlation to our essential amino acids (Fig. 3).

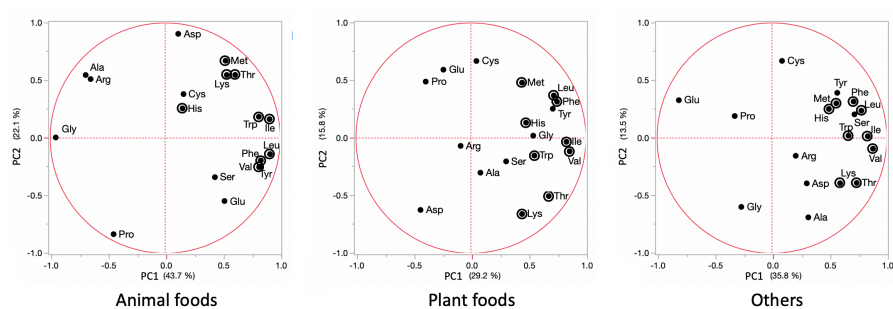


Figure 3. Vector plot of the first and second principal components of each group  
The amino acid vectors of the first and second principal components were plotted for each group. Nine human essential amino acids are circled. Each principal component vector independently showed positive correlations to essential amino acids in all three groups.

Why do our essential amino acids appear in the vectors of the first principal component of our foods? Later, I found the possibility that this could be the origin of essential amino acids.

### **Main component of our body**

In this research, a food composition table was analysed. Because all foods were from eukaryotic organisms, the table is also a database of the amino acid compositions of the bodies of eukaryotic organisms.

Then, here is a question. What is the most predominant component of the eukaryotic organism body? Water constitutes approximately 60% of body mass among mammals and other multicellular animals. The intracellular fluid constitutes ~40% of the body mass<sup>19</sup>. Water also constitutes 70-80% of each cell<sup>20,21</sup>. These results indicate that cells constitute more than half of the body mass of animals. Hypothetically, if there are common essential amino acid-rich components in each eukaryotic cell, then they could be recognized as the first principal component vector. Do eukaryotes truly have such components?

Before this question is considered, some examples of evolution will be presented.

### **Buchnera and aphids**

Aphids are insects that are commonly found on garden plants. It is known that almost all aphids have *Buchnera* inside their cells as a symbiote. *Buchnera* are round bacteria and are a close relative of *Escherichia coli*. While *E. coli* can synthesize all amino acids, *Buchnera* does not synthesize several amino acids<sup>22</sup> that their host aphids synthesize. *Buchnera* provide the host with its essential amino acids and obtain other amino acids from the host<sup>23,24</sup>. In this way, the two organisms form a perfect complementary pair<sup>25</sup> (Table 1).

	Ala	Arg	Asn	Asp	Cys	Glu	Gly	His	Ile	Leu
Host										
<b>Aphid</b>		●						●	●	●
Parasite	▽	▽	▽	▽	▽			▽	▽	▽
<b>Buchnera</b>	●		●	●	●					

	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Gln
Host										
<b>Aphid</b>	●	●	●			●	●		●	
Parasite	▽		▽	▽	▽	▽	▽	▽	▽	▽
<b>Buchnera</b>		●		●	●			●		●

Table 1. Amino acid requirements of aphids and Buchnera<sup>25</sup>

The amino acid requirements of selected organisms are shown. Downward-pointing triangles show complementary pairs of amino acid synthesis abilities.

This example indicates that organisms can dynamically change their ability to synthesize amino acids according to their circumstances.

### Legionella and its hosts

*Legionella pneumophila* is known to be the causative agent of Legionnaires' disease. Because their main hosts are amoebas, these bacteria enter a diverse range of free-living amoeba cells by mimicking their food bacteria and then invade and proliferate inside the amoeba cells<sup>15</sup>. *Legionella* obtains its needed amino acids by hijacking hosts' polyubiquitination-and-proteasome degradation pathway. On the other hand, *Legionella* can parasitize not only amoebas but also some cellular slime moulds and human macrophages<sup>15</sup>. Because these hosts are heterotrophic eukaryotes, it is not surprising that their amino acid requirements have many overlaps. However, it is interesting that we can find some overlaps between the amino acid requirements of *Legionella* and its hosts<sup>15</sup> (Table 2).

Why are these requirements synchronized? No existing theory can directly explain this point.

	Ala	Arg	Asn	Asp	Cys	Glu	Gly	His	Ile	Leu
Parasite										
<b>Legionella</b>	●				●				●	●
Host										
<b>Amoeba</b>	●								●	●
<b>Slime moulds</b>	●							●	●	●
<b>Human</b>								●	●	●

	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Gln
Parasite										
<b>Legionella</b>		●			●	●			●	
Host										
<b>Amoeba</b>		●			●	●			●	
<b>Slime moulds</b>	●	●	●		●	●	●		●	
<b>Human</b>	●	●	●		●	●	●		●	

Table 2. Amino acid requirements of Legionella and its hosts<sup>15</sup>

The amino acid requirements of selected organisms are shown. Squares show overlaps between the requirements of Legionella and its hosts.

### Considerations regarding the loss of synthesis ability

What kind of explanation can address the above questions?

I believe that Buchnera was able to stop synthesizing certain amino acids because these amino acids were abundant inside the host cells. Buchnera therefore lost the need and consequently the ability to synthesize these amino acids.

If the same reasoning is applied to Legionella, the amino acids that Legionella no longer synthesizes must have been abundant inside the host amoeba cells. This indicates that Legionella can use abundant Arg Cys Ile Leu Met Ser Thr and Val in host cells (see Table 2). Many of these are also essential amino acids for humans, which indicates that these amino acids are abundant in the host cytoplasm (see Table 2 again). Thus, this assumption may support our hypothesis that eukaryotes have common essential-amino-acid-rich components inside their cells.

However, how could such a component originate from our essential amino acids? Autophagy can provide this explanation.

### Autophagy

In 1992, Ohsumi *et al* revealed that under nutrient-deficient conditions, budding yeast sequesters its own cytoplasmic components into vacuoles in the form of “autophagic bodies”<sup>7</sup>. By autophagy, cytoplasmic components are then degraded and reused as a nutritional resource<sup>910</sup>. They also revealed that autophagy-defective mutants of yeast

could not survive under nitrogen starvation conditions<sup>8</sup>, and autophagic degradation of the cytoplasmic components is essential for survival during starvation<sup>10</sup>. They also speculated that autophagic degradation might be the primary source of intracellular amino acids<sup>10,11</sup>.

Autophagy is now believed to be common in all eukaryotes<sup>26</sup>. In 1975, Young showed that human daily total body protein synthesis is several-fold larger than protein intake<sup>27</sup>. At that time, how humans obtained such large amounts of amino acids for this protein synthesis was not clearly explained. Now, however, this is explained mainly by autophagy. Thus, autophagic degradation can be assumed to be the primary source of intracellular amino acids.

### **Compositions of cytoplasmic components**

If autophagy is the primary source of amino acids in all eukaryotes, it is necessary to confirm whether the cytoplasmic components contain abundant essential amino acids because these components are the main ones degraded by autophagy during nutritional starvation. However, only fragmented information was available. From 1997 to 2000, Okayasu and Sorimachi *et al.* used high-performance liquid chromatography and showed that cultured cells of all terrestrial organisms have similar amino acid compositions<sup>12–14</sup>. Another genome study showed that prokaryotes tend to keep proteins rich in amino acid residues that demand high synthetic cost inside their cells<sup>28</sup>, and another genome study showed that eukaryotes keep proteins rich in hydrophobic amino acid residues inside their cells<sup>29</sup>. Neither high synthetic cost<sup>28</sup> nor hydrophobicity<sup>30</sup> is completely consistent with our essential amino acids, but both have positive correlations and might have affected the accumulation of essential amino acids in the cytoplasm. With this fragmented information, I therefore assumed that the amino acid compositions of eukaryotic cytoplasmic components have essential-amino-acid-rich compositions and have been conserved. Thus, cytoplasmic components could be the source of the synchronization of *Legionella* and its hosts' amino acid requirements because organisms need not synthesize these amino acids that are abundant in their cytoplasmic components.

### **Hypothesis and principal component vector**

The hypothesis that all eukaryotic cytoplasmic components have essential-amino-acid-rich compositions explains that the essential-amino-acid-oriented vector of the first principal component might reflect cytoplasmic amino acid contents in each food. Because eukaryotes can store particularly valuable amino acids inside their cells, it is

reasonable that they lost the need to synthesize those amino acids. Therefore, I concluded that the autophagic degradation of cytoplasmic components was a possible origin of essential amino acids.

### **Unsolved questions**

The cytoplasmic hypothesis explained many unsolved questions regarding essential amino acids. However, some new or unsolved questions remain. Two of them will be addressed here.

The first question is why the non-essential amino acid tyrosine was plotted among essential amino acids on the vector plot. This might be because tyrosine lies two steps downstream of the essential amino acid phenylalanine on the synthetic pathway and was omitted from the list of essential amino acids because it could be synthesized from another amino acid. However, the ability to synthesize tyrosine can be said to have been lost with the ability to synthesize phenylalanine. Therefore, in our hypothesis, the indication of the vector plot that tyrosine may be abundant in cytoplasmic components suggests that tyrosine might well be reclassified as an essential amino acid.

The second question is arginine requirement. During this research, it was observed that among all heterotrophic eukaryotes, only adult humans seemed to have regained the ability to synthesize arginine (data not shown). The cytoplasmic hypothesis does not explain why arginine requirements are so fundamental and why adult humans are the only exception. Humans might be losing eukaryotic essential amino acids. Perhaps such exceptions and unsolved questions are waiting for another theory to be established.

### **Essential amino acids in evolution**

Eukaryotes have been differentiated from prokaryotes. Upon becoming eukaryotes, they gained the ability of autophagy and began using intracellular proteins as a source of their amino acids. This allowed them to store valuable amino acids inside their cells and thereby achieve survival under starvation.

In the era of unicellular organisms, for prokaryotes, amino acid requirements must have been a factor defined by the surrounding environments. However, for eukaryotes, with the acquisition of autophagy, such amino acid requirements might have become a factor defined by their own cytoplasmic components, the primary source of their amino acid supply. Eukaryotes could survive because after eating other organisms, they were guaranteed the ability to obtain rich essential amino acids from their cytoplasmic components.

Later, over time, unicellular organisms evolved into multicellular organisms. With the



acquisition of diversity, animals obtained protein-rich bodies, and plants obtained bodies of carbohydrates with poor protein composition. For organisms that eat plants, amino acid requirements such as essential amino acids might not be beneficial. That must be why aphids formed a symbiotic pair with Buchnera to regain the ability to synthesize all amino acids and become one of the most abundant pests in the world. However, even for aphids, the ability to synthesize essential amino acids itself could not be regained. What was lost seems not to be easily regained.

Essential amino acids may be a conserved remnant of our unicellular era in which we survived in harsh nutritional environments.

### Competing interest declaration

The author declares no competing interests.

### References

1. ROSE, W. C. Amino acid requirements of man. *Fed. Proc.* **8**, 546–52 (1949).
2. Gietzen, D. W. & Rogers, Q. R. Nutritional homeostasis and indispensable amino acid sensing: A new solution to an old puzzle. *Trends Neurosci.* **29**, 91–99 (2006).
3. Oda, H. Essential Amino Acids and Nonessential Amino Acids in Evolution. *Nippon Eiyo Shokuryo Gakkaishi* **60**, 137–149 (2007).
4. Reeds, P. J. Dispensable and Indispensable Amino Acids for Humans. *J. Nutr.* **130**, 1835S–1840S (2000).
5. Gutiérrez-preciado, B. A. *et al.* An Evolutionary Perspective on Amino Acids. *Nat. Educ.* **3**, 1–6 (2010).
6. Office for Resources, Policy Division Science and Technology Policy Bureau. Ministry of Education, Culture, Sports, Science and Technology, J. STANDARD TABLES OF FOOD COMPOSITION IN JAPAN - 2015 - (Seventh Revised Edition) - Amino Acids -. [https://www.mext.go.jp/a\\_menu/syokuhinseibun/](https://www.mext.go.jp/a_menu/syokuhinseibun/) (2015).
7. Takeshige, K., Baba, M., Tsuboi, S., Noda, T. & Ohsumi, Y. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *J. Cell Biol.* **119**, 301–312 (1992).
8. Tsukada, M. & Ohsumi, Y. Isolation and characterization of autophagy-defective mutants of *Saccharomyces cerevisiae*. *FEBS Lett.* **333**, 169–174 (1993).
9. Baba, M., Takeshige, K., Baba, N. & Ohsumi, Y. Ultrastructural analysis of the autophagic process in yeast: Detection of autophagosomes and their

- characterization. *J. Cell Biol.* **124**, 903–913 (1994).
10. Onodera, J. & Ohsumi, Y. Autophagy is required for maintenance of amino acid levels and protein synthesis under nitrogen starvation. *J. Biol. Chem.* **280**, 31582–31586 (2005).
  11. Scott, S. V. *et al.* Cytoplasm-to-vacuole targeting and autophagy employ the same machinery to deliver proteins to the yeast vacuole. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 12304–12308 (1996).
  12. Okayasu, T., Ikeda, M., Akimoto, K. & Sorimachi, K. The amino acid composition of mammalian and bacterial cells. *Amino Acids* **13**, 379–391 (1997).
  13. Sorimachi, K. Evolutionary changes reflected by the cellular amino acid composition. *Amino Acids* **17**, 207–226 (1999).
  14. Sorimachi, K., Okayasu, T., Akimoto, K. & Niwa, A. Conservation of the basic pattern of cellular amino acid composition during biological evolution in plants. *Amino Acids* **18**, 193–197 (2000).
  15. Price, C. T. D., Richards, A. M., Von Dwingelo, J. E., Samara, H. A. & Abu Kwaik, Y. Amoeba host-Legionella synchronization of amino acid auxotrophy and its role in bacterial adaptation and pathogenic evolution. *Environ. Microbiol.* **16**, 350–358 (2014).
  16. DEWEY, V. C. & KIDDER, G. W. The Influence of Folic Acid, Threonine and Glycine on Serine Synthesis in Tetrahymena. *J. Gen. Microbiol.* **22**, 72–78 (1960).
  17. Franke, J. & Kessin, R. A defined minimal medium for axenic strains of Dictyostelium discoideum. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 2157–2161 (1977).
  18. Watanabe, T. & Kawai, R. Advances in food composition tables in Japan- Standard Tables Of Food Composition in Japan – 2015 – (Seventh Revised Edition)–. *Food Chem.* **238**, 16–21 (2018).
  19. Andersen, O. S. Cellular Electrolyte Metabolism. in *Encyclopedia of Metalloproteins* (eds. Kretsinger, R. H., Uversky, V. N. & Permyakov, E. A.) 580–587 (Springer New York, 2013). doi:10.1007/978-1-4614-1533-6\_223.
  20. Ellis, R. J. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597–604 (2001).
  21. Ellis, R. J. & Minton, A. P. Join the crowd. *Nature* **425**, 27–28 (2003).
  22. Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. & Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Nature* **407**, 81–86 (2000).
  23. Baumann, P. BIOLOGY OF BACTERIOCYTE-ASSOCIATED ENDOSYMBIONTS OF PLANT SAP-SUCKING INSECTS. *Annu. Rev. Microbiol.* **59**, 155–189 (2005).

24. Pál, C. *et al.* Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**, 667–670 (2006).
25. Takahata, N. & Satta, Y. Amino Acids and Genome Modifications. *Nippon Eiyo Shokuryo Gakkaishi* **60**, 131–135 (2007).
26. Mizushima, N. The role of mammalian autophagy in protein metabolism. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **83**, 39–46 (2007).
27. Young, V. R., Steffee, W. P., Pencharz, P. B., Winterer, J. C. & Scrimshaw, N. S. Total human body protein synthesis in relation to protein requirements at various ages. *Nature* **253**, 192–194 (1975).
28. Smith, D. R. & Chapman, M. R. Economical evolution: Microbes reduce the synthetic cost of extracellular proteins. *MBio* **1**, 28–32 (2010).
29. Gaur, R. K. Amino acid frequency distribution among eukaryotic proteins. *I/OAB J.* **5**, 6–11 (2014).
30. Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M. & Hodges, R. S. Relationship of sidechain hydrophobicity and  $\alpha$ -helical propensity on the stability of the single-stranded amphipathic  $\alpha$ -helix. *J. Pept. Sci.* **1**, 319–329 (1995).

## **Methods**

### **Food composition table**

The food composition table can be downloaded as an Excel file from the Japanese government website<sup>6</sup>.

[https://www.mext.go.jp/a\\_menu/syokuhinseibun/1365450.htm](https://www.mext.go.jp/a_menu/syokuhinseibun/1365450.htm)

Among the tables in the file, “Table 1: Amino acids per 100 g edible portion” was used for this research.

### **Reclassification of food groups**

The original table has 18 food groups, as classified by the table creators<sup>18</sup>.

They were reclassified into three groups as follows. (1) Animal foods (n=620): Fish, molluscs and crustaceans, Meat, Eggs, Milk and milk products. (2) Plant foods (n=657): Cereals, Potatoes and starches, Sugars and sweeteners, Pulses, Nuts and seeds, Vegetables, Fruits, Mushrooms. (3) Others (n=281): Algae, Fats and oils, Confectioneries, Beverages, Seasonings and spices, Prepared foods.

### **Data preparation and analyses**

In the table, some data are listed as letters, so they were changed into numerical data before analysis.

There are 19, not 20, amino acids on the list for two reasons. The first is that the table creators measured the amounts of asparagine and glutamine within the amounts of aspartic acid and glutamic acid, respectively. The second is that hydroxy prolines were independently measured and are listed in this table. For the purposes of this research, hydroxy prolines were combined with proline after converting them to weights of the same molar amounts. The molecular weights used for these calculations were taken from the appendix document of the table, which is accessible on the same website.

Note that the listed amounts of each amino acid are the weights of the amino acids, not the weights of the amino acid residues. For the principal component analysis, each amino acid weight was calculated as the amino acid weight ratio by dividing the total weight of these 18 listed amino acids. Then, all 18 amino acid weight ratios were incorporated into the principal component.

### **Principal component analysis**

Calculations of principal components and their vectors were performed by using JMP v.14.3.0 software (SAS Institute Inc.).