Journal of Molecular Biology Research; Vol. 5, No. 1; 2015
ISSN 1925-430X E-ISSN 1925-4318
Published by Canadian Center of Science and Education

Identifying Significant Biological Markers in Klotho Gene Variants Across Wide Ranging Taxonomy

Tommy Rodriguez[1]

[1] Pangaea Biosciences, Department of Research & Development, Miami, FL, USA

Correspondence: Tommy Rodriguez, Pangaea Biosciences, [redacted]
USA. Tel: 1-305-389-4156. E-mail: trodriguez@pangaeabio.com

Received: November 4, 2014 Accepted: December 7, 2014 Online Published: January 9, 2015

doi:10.5539/jmbr.v5n1p11 URL: http://dx.doi.org/10.5539/jmbr.v5n1p11

Abstract

Biological aging is marked by progressively degenerative physiological change that causes damage to tissues and organs. Errors in biopolymers accumulate over time; mitochondrial dysfunction, telomere attrition, and wider genomic instability lead to an altered state of intercellular communication. In this investigation, my focus will be aimed at examining and identifying specifically critical biomarkers in genetic variants of KLOTHO (a transmembrane protein involved in the genetic regulation of age-related disease) among organisms with varied life spans that range across wide taxonomical rankings. Here, I investigate the correlation between lower and higher frequency α-amino acid compositions in Klotho protein factors within a grouped methodology; as to also include several demonstrative techniques in comparative sequence analysis for inferring relatedness in evolutionary context.

Keywords: Klotho, KL gene, protein, protein evolution, comparative sequence analysis, multiple sequence alignment, Kalign, phylogenetc analysis, amino acid composition analysis

1. Introduction

Promising new research in life extension has led to remarkable advancements in our basic understanding of the molecular mechanisms associated with aging. Within the last two decades, much of this research has focused on identifying potential genomic candidates for longevity, while seeking to explain how these individual genes can affect the biological process of aging. One of the proteins in particular (KLOTHO) has been the subject of several recently published papers. A handful of studies now suggest that genetic variants of KLOTHO [encoded by the KL gene] are associated with human aging and tumor suppression, and trials on model organisms involving KLOTHO variants has shown improvement in cognition and deceleration in age-related development (Dubal et al., 2014). In one such case, Dubal et al. (2014) demonstrated that systemic overexpression of Klotho variants in transgenic mice enhanced cognition and increased longevity by an average ratio of ±25 percent; whereas Klotho-deficient mice manifested a syndrome resembling accelerated human aging and displayed extensive and accelerated arteriosclerosis (Dubal et al., 2014).

For reasons that are not yet fully understood, Klotho-associated mechanisms, "change cellular calcium homeostasis, by both increasing the expression and activity of TRPV5 and decreasing that of TRPC6 (Kurosu et al., 2005)." Moreover, altered mineral-ion homeostasis could be a cause of premature aging-like phenotypes (Kurosu et al., 2005). In order to gain a more comprehensive understanding of the underlying functions in molecular components of KLOTHO, we should begin by examining KL-derivative patterns across a wide evolutionary spectrum, without limiting ourselves to one individual taxa or another. Because the lifespan of organisms vary widely among species, a comparative approach could help us identify a set of unique signatures in the molecular variation patterns of KLOTHO. In turn, this may help provide meaningful reference toward a fully systematic investigation. Throughout the course of this paper, I seek to address those areas thoroughly.

It should then be noted: This paper does not presume on a solution to longevity; nor does it seek to draw parallels between the many collective biological features that allow organisms to persist in semi-optimal states. Instead, my focus will be aimed at identifying and examining specifically critical biomarkers, or sequence motifs, in genetic variants of KLOTHO among organisms with varied life spans that range across wide taxonomical rankings. Here, I investigate the correlation between lower and higher frequency α-amino acid compositions in Klotho protein factors within a grouped methodology; as to also include several demonstrative techniques in comparative sequence analysis for inferring relatedness in evolutionary context.

2. Methods

2.1 Sequence Selection and Group Categorization

Biological aging is marked by progressively degenerative physiological change that causes [irreversible] damage to tissues and organs. Errors in biopolymers accumulate over time; mitochondrial dysfunction, telomere attrition, and wider genomic instability lead to an altered state of intercellular communication. Within a broader spectrum, degrees of longevity vary widely among organisms. Only a tiny percentage of species exhibit characteristics that is counter-intuitive to universal mechanisms of biological aging. *Turritopsis dohrnii* and *Polycelis feline* are two such examples of eukaryotic organisms that appear to exhibit, "limitless telomere regenerative capacity fueled by a population of highly proliferative adult stem cells (Tan et al., 2012)." Several other species of marine Crustaceans share similar traits that prolong biological aging.

Individual-scale longevity is largely determined by inheritance; genes can explain up to 35 percent (American Federation for Aging Research, 2012). As noted earlier, recent studies indicate that Klotho gene variants may be potential candidates for the genetic regulation of age-related disease. For this purpose, I selected the KL gene as a prime candidate for analysis and evaluation. Furthermore, using homologous variants of Klotho protein sequences as a primary source for comparative studies, I grouped and categorized six distinct genomic datasets of KLOTHO [in FASTA format] by the following criteria:

i. Group *type a* { *Heterocephalus glaber*; *Chelonia mydas*; *Danio rerio*; } – consisting of organisms that are recognized as having favorable biological attributes in one or more of the following areas: life cycle duration, tumor-suppression, cancer-resistance, immunity, tissue regeneration.

ii. Group *type b* { *Homo sapiens*; *Mus musculus*; *Mesocricetus auratus*; } – consisting of organisms that reside within a moderate range of natural life expectancy due to age-related decline, disease, and predispositions.

The six protein sequences used for curating the experimental dataset were obtained via NCBI Protein database. The accession references for each individual sequence are detailed in the sections below.

2.1.1 Regarding group *type a*

Two of three *type a* organisms exhibit extraordinarily prolonged life cycles. *Heterocephalus glaber* are highly resistant to cancer, and can maintain a youthful vascular function and cellular oxidant-antioxidant phenotype relatively longer and are better protected against aging-induced oxidative stress than shorter-living rodents (Buffenstein, 2009; Csiszar et al., 2007; Pérez et al., 2009). *Danio rerio* have shown the ability to regenerate their fins, skin, and heart (Jopling, 2010; Sun et al., 2009). Jopling et al. (2010) demonstrated that [zebrafish] can fully regenerate its heart after amputation of up to 20 percent of the ventricle. Although very little is known about the average lifespan of green sea turtles, *Chelonia mydas* reach sexual maturity between 20 to 50 years of age (United States Fish and Wildlife Service, 2005).

2.2 Arriving at Reliable Models for Comparative Analysis

2.2.1 Performing Pairwise Sequence Alignment

Following group categorization, I deployed UGENE's pairwise sequence alignment (PSA) tool in order to identify similarity probabilities among my six protein sequences [length: 1,037]. Initially, I produced five pairwise sequence alignments by pairing each Klotho protein sequence against *Homo sapiens* [*type b*] and then grouping the results in numerical sequence [see below]. In this operation, Hirschberg algorithm was selected for optimal matching. A reliable algorithmic selection, Hirschberg increases efficiency by maximizing the sum of pairwise scores with quasi-gap penalties (Chao & Zhang, 2008). The Hirschberg algorithm can be derived from the Needleman-Wunsch algorithm by observing the following (Hirschberg, 1975):

if $(Z,W) = NW$[score] (X,Y) is the optimal alignment of (X,Y), and $X = X^l + X^r$ is an arbitrary partition of X, there exists a partition $Y^l + Y^r$ of Y such that NW[score] $(X,Y) = NW$[score] $(X^l,Y^l) + NW$[score] (X^r,Y^r)

When calculating the highest number of consistent patterns in a local alignment, gap penalty scores are often disregarded (Lassmann & Sonnhammer, 2005). However, because I am interested in obtaining a global alignment rather than calculating common subsequences, I applied a gap open penalty score of ten; including a gap extension penalty of two. The resulting PSA provided a first glimpse into similarity probabilities. As noted below, my numerically grouped PSA outputs infer high degree homology between three of the six sequences and two of six sequences; now categorized as group *type a* and group *type b*. PSA validates sequence placement among *type a* organisms and *type b* organisms.

Table 1. Similarity probabilities among *type a* and *type b* (PSA)

Group	Identification	Similarity appx.	FASTA description
type *b*	Homo sapiens	100%	>gi|2618596|dbj|BAA23382.1
type *b*	Mus musculus	85%	>gi|2618594|dbj|BAA23381.1
type *b*	Mesocricetus auratus	85%	>gi|524966649|ref|XP_005083219.1
type *a*	Heterocephalus glaber	54%	>gi|351702474|gb|EHB05393.1
type *a*	Chelonia mydas	52%	>gi|465963328|gb|EMP30051.1
type *a*	Danio rerio	50%	>gi|528492619|ref|XP_690797.4

2.2.2 Performing Multiple Sequence Alignment

A series of multiple sequence alignment (MSA) operations were then performed in preparation for phylogenetic analysis. In contrast to pairwise alignment, multiple sequence alignment can reveal subtle similarities among *groups* of proteins (Chao & Zhang, 2008). Here, I selected Kalign for multiple sequence alignment; an accurate and fast MSA algorithm (Lassmann & Sonnhammer, 2005). Kalign is an extension of Wu-Manber approximate pattern-matching algorithm, which is based on Levenshtein distances. This strategy enables Kalign to estimate sequence distances faster and more accurately than other popular iterative methods. Comparisons done by Lassmann and Sonnhammer (2005) show that Kalign is about 10 times faster than ClustalW and, depending on the alignment size, up to 50 times faster than other iterative methods; Kalign also delivers better overall resolution (Lassmann & Sonnhammer, 2005).

Kalign is noted for producing optimal execution times, and this procedure would require minimal computational resources. First, I initiated UGENE's multiple sequence alignment tool by importing and processing the six protein sequences in FASTA format. Due to parameter setting sensitivity in protein data types, Kalign for MSA gap penalty scores were modified slightly during successive intervals until an optimal global alignment was obtained. Each interval resulted in a 3,092 base-pair alignment, followed by a phylogenetic diagram. These operations were then repeated in successive fashion, upon conducting protein to nucleotide conversions.

2.3 Reverse KLOTHO Protein-DNA Translation for Phylogenetic Reconstruction

Generally speaking, protein sequences are intolerant of change in evolutionary context. Over the span of evolutionary time, protein sequences undergo selective constraints for protein function and protein structure, and these are conserved over much longer periods than individual codons (Martin & Palumbi, 1993). The most direct evolutionary changes to protein occur at the amino and carboxyl termini in the form of domain insertions, repetitions, and deletions (Marsh & Teichmann, 2010). Likewise, we must also consider the possibility that convergent evolution can occur to produce apparent similarity between proteins that are evolutionarily unrelated, but perform similar functions and have similar structures (Bastien, 2008). Thus, multiple substitutions at a single DNA base more accurately reflect mutational history (Martin & Palumbi, 1993).

The challenges in utilizing protein sequences to infer divergence events led me to consider a secondary option to protein-protein comparison in the context of phylogenetic reconstruction. Perhaps it would be necessary to cross-check the original diagram(s) produced by protein-protein comparison with a nucleotide derivative. Using my custom-based API translator (Protein to DNA Bio Translator), I reverse translated each sequence from its original data type, namely protein, into a workable [and theoretical] dataset made exclusively of KL-derivative nucleotide sequences (DNA).

Following conversion, the six DNA sequences were imported to UGENE's bioinformatics software, and upon file import and MSA execution, an alternate radial phylogenetic diagram was generated that would allow me to cross-confirm the original(s) obtained by protein-protein comparison. I implemented PHYLIP neighbor-joining method coupled with distance matrix model F84 on the 3,092 base-pair alignment; this procedure would also require additional bootstrapping compilers to help evaluate the strength of the nodes. Lastly, it is worth noting that the resulting phylogenetic tree(s) do not assume an evolutionary clock; it is in effect an unrooted tree.

The PHYLIP neighbor-joining method is capable of generating highly probable diagrams in scenarios involving low degrees of variance, regardless of dataset size. An accurate and statically consistent polynomial-time algorithm, PHYLIP neighbor-joining does not assume that all lineages evolve at the same rate (as proteins evolve at different rates), and it constructs a tree by successive clustering of lineages, setting branch lengths as the lineages join [where a set of n taxa requires $n - 3$ iterations; each step is repeated by $(n - 1) \times (n - 1)$] (Felsenstein, 1981). For illustration purposes, the following formulas demonstrate a standard neighbor-joining Q-matrix algorithm:

$$Q(i,j) = (n-2)\, d(i,j) - \sum \{n, k = 1\}\, d(i,k) - \sum \{n, k = 1\}\, d(j,k) \quad (1)$$

Pair to node (distances):

$$(f,u) = \tfrac{1}{2}\, d(f,g) + \tfrac{1}{2}(n-2)\, [\, \sum \{n, k = 1\}\, d(f,k) - \sum \{n, k = 1\}\, d(g,k)\,] \quad (2)$$

Taxa to node (distances):

$$d(u,k) = \tfrac{1}{2}\, [\, d(f,k) + d(g,k) - d(f,g)\,] \quad (3)$$

*2.4 Analyzing α-*Amino Acid *Compositions*

Given its mass and length, the molecular weight measurements of protein structures are fundamentally important to its biochemical characterization and function. Moreover, the α-amino acid composition of each protein structure may contribute to the overall quality of a protein. If certain α-amino acids are optimal for protein structure, natural selection should have acted over evolutionary time to increase the frequency of these α-amino acids (Anthis et al., 2013). As noted by Mannakee and Gutenkunst (2012), catalytic domains in protein evolve faster, while non-catalytic domains in protein evolve more slowly. This may also suggest that networks typically evolve under stabilizing selection (Mannakee & Gutenkunst, 2012).

In addition to a phylogenetic reconstruction, this investigation puts a strong emphasis on discerning meaningful patterns, or sequence motifs, in lower and higher frequency α-amino acid compositions. By comparing the results of composition frequency percentages among group *type a* and group *type b*, I hope to identify any potentially significant molecular markers in evolutionary conserved chemical properties of Klotho that may have increased or decreased over evolutionary time, within a particular group and across a wider evolutionary spectrum. This procedure incorporates a few methods and techniques, such as α-amino acid residue calculations for determining molecular weight and frequency percentage.

A number of highly efficient web applications are well suited for α-amino acid composition analysis. In this phase of my investigation, I use Composition/Molecular Weight Calculation tool (University of Delaware, 2014) for obtaining the sum ratio of α-amino acid residues and approximate residue charge, and Protein Calculator (Anthis et al., 2013) to determine atomic compositions. I then apply the resulting figures and datasets in comparative analysis in order to evaluate the overall α-amino acid distributions.

The sum ratios in α-amino acid counts vary slightly due to irregularities in protein sequence length. This would also explain discrepancies in PSA similarity approximations. Amino acid counts are quantified by the total number of residues in an individual protein sequence. These figures are later arranged and disbursed as to determine frequency percentage within a protein sequence arrangement. The total estimate in molecular weight per arrangement is also deciphered by accounting for the total sum of individual residue weights. Lastly, my molecular weight formula integrates the following two variables: (x) the individual sum residue weight(s); (y) subtracted by the molecular weight of H_2O (18.01528) (University of Delaware, 2014).

Molecular weight = sum of individual residues weights - water molecular weight ´ (number of residues - 1)

a) Homo sapiens ($C_{5337}H_{8024}N_{1410}O_{1452}S_{30}$): 116131.78
b) Mesocricetus auratus ($C_{5357}H_{8070}N_{1434}O_{1449}S_{24}$): 116514.11
c) Mus musculus ($C_{5338}H_{8043}N_{1431}O_{1454}S_{28}$): 116424.91
d) Danio rerio ($C_{5178}H_{7752}N_{1380}O_{1437}S_{30}$): 113287.91
e) Chelonia mydas ($C_{3896}H_{5817}N_{979}O_{1045}S_{23}$): 83826.69
f) Heterocephalus glaber ($C_{3528}H_{5347}N_{919}O_{973}S_{21}$): 76876.45

3. Results

3.1 Reconstructing a Phylogeny Based on Klotho Gene Variants

The resulting base-pair alignments yielded nearly identical outputs. At first glance, *type a* organisms and *type b* organisms are each placed together within their respective groups – with four mammalian candidates on one end of the diagram(s), and the remaining non-mammalian candidates to the other [see Figure 1]. When cross-checked with morphology, the corresponding nodes reside in correct order of taxonomy on both diagrams. On one end, three of six organisms belonging to *order rodentia* coincide in proper placement according to scientific classification; while my fourth *mammalian* species (Homo sapiens) falls within a fairly close proximity; followed by the remaining candidates.

The phylogenies illustrated in this study depict gene families within paralogy regions consistent with the early evolution of vertebrates. Because similar variations of KL-derivative proteins occur in all species of *Chordata* – ranging from fish, reptiles to mammals – we may safely infer that, in fact, Klotho family proteins belong to ancient

paralogy lineages. While having the same basic molecular functions, Klotho protein variants may have undergone only tiny degrees of modification throughout evolutionary history (originating at the nucleotide level). As the next section will demonstrate, these tiny variations are further observed among my candidate sequences. Alas, the current MSA models help further verify group categorization, and the extent of evolutionary relatedness among group *type a* and group *type b* candidates, including node length and probable time scales in decimal format, are detailed below.

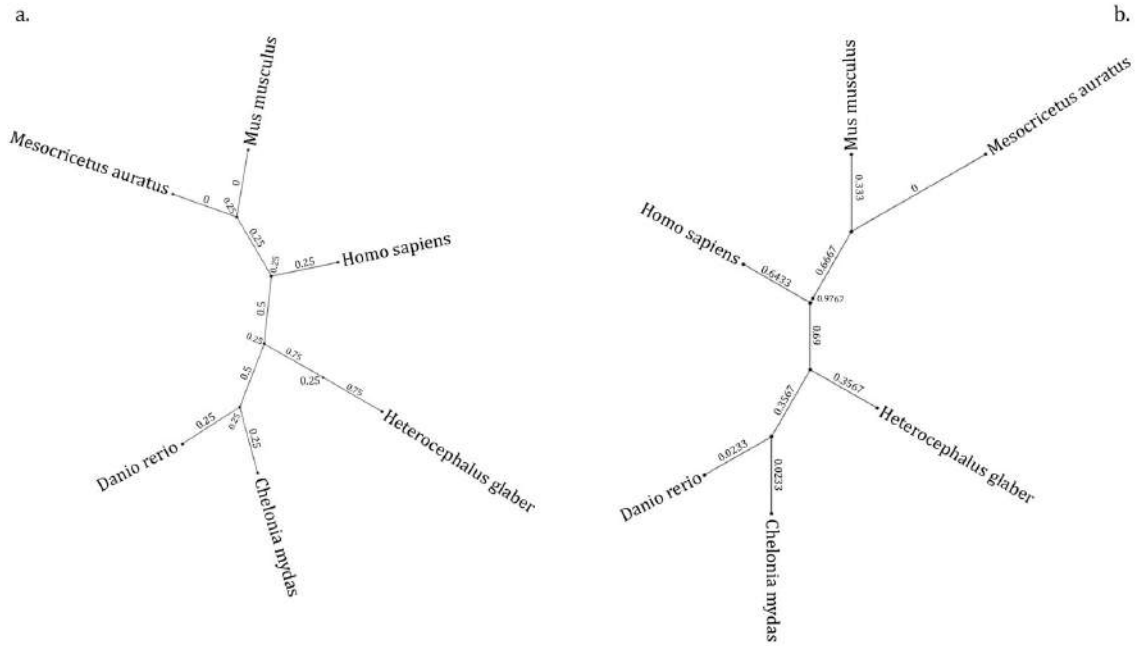


Figure 1. Radial phylogenetic trees. a) Protein-protein output (Klotho). b) Reverse protein-DNA output (Klotho)

3.2 Minimal Degrees of Variation in Evolutionary Conserved Chemical Properties of Klotho

The chemical properties of the α-amino acids of proteins determine the biological activity of the protein, and it conveys a vast array of chemical versatility (University of Arizona, 2014). Evolution itself exerts pressure to preserve α-amino acid residues that bear invaluable functional roles. According to Dokholyan et al. (2002), proteins may retain significantly critical chemical properties (or molecular signatures) that contribute, in lower or higher frequency values, to protein stability via their interaction with other α-amino acid residues. To this end, we might expect to observe a large percentage of evolutionary conserved α-amino acids across entire protein families, with only minimal degrees of variation occurring among wide ranging taxonomical groups.

Indeed, in the context of evolutionary conservation, only three of twenty α-amino acids in this study exceed disparity ratios that average more than 0.093 percent. Comparative analyses of α-amino acid compositions highlight only three strongly distinct molecular signatures indicative of each group. As Figure 2 and Figure 3 illustrate, a correlation between elevated levels of L-Lysine coupled with low quantities of L-Alanine and L-Arginine are featured in group *type a* [as compared to group *type b*]; where the average disparity ratio is ±1.8 percent (L-Lysine), ±1.6 percent (L-Alanine), and ±1.34 percent (L-Arginine). A max disparity ratio of >3.1 percent occurs in L-Lysine molecules [*type a*]. Increased levels of L-Lysine are significant in terms of optimal structure and fold, and suggest that natural selection may have acted to increase the frequency of this molecule.

It is worthwhile to note that L-Alanine and L-Arginine are manufactured internally via biosynthetic pathways; whereas L-Lysine must be ingested as lysine or lysine-containing proteins (plants synthesize L-Lysine from aspartic acid) (Beals, 1999). While the implications of dietary L-Lysine are inconsequential to a study involving techniques in comparative sequence analysis, various other studies suggest that increased dietary L-Lysine plays a role in calcium absorption; a function that is also inadvertently linked to KLOTHO (Tan et al., 2012). As noted in the sections above, "Klotho stimulates calcium reabsorption in the distal convoluted tubule by deglycosylating and stabilizing the epithelial calcium channel TRPV5 on the surface of cellular membrane (Robinson, 2012)." This, and proficiencies in tissue regeneration and other immunity-related health benefits are also linked to increased levels of dietary L-Lysine (Falco, 1995).

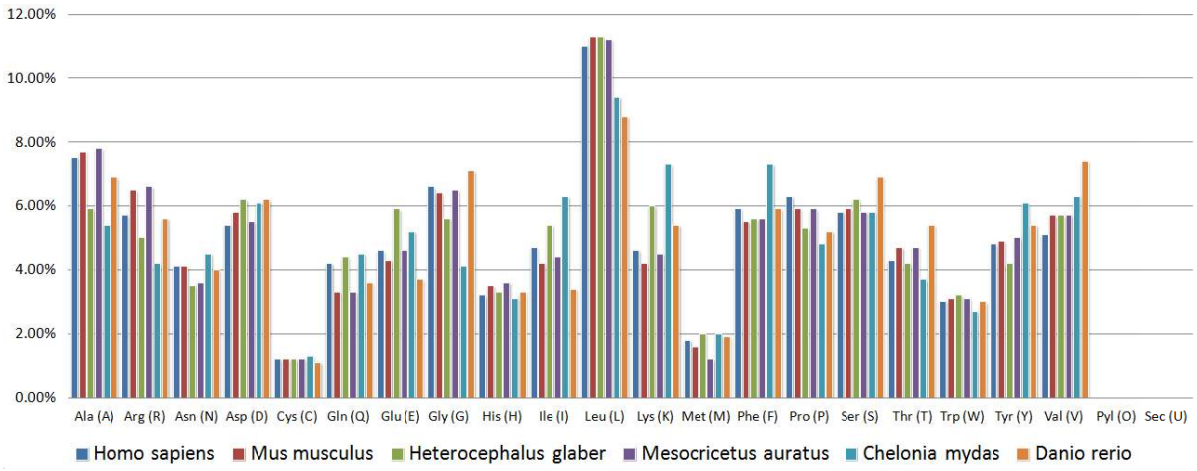


Figure 2. Amino acid composition analysis (six sequences)

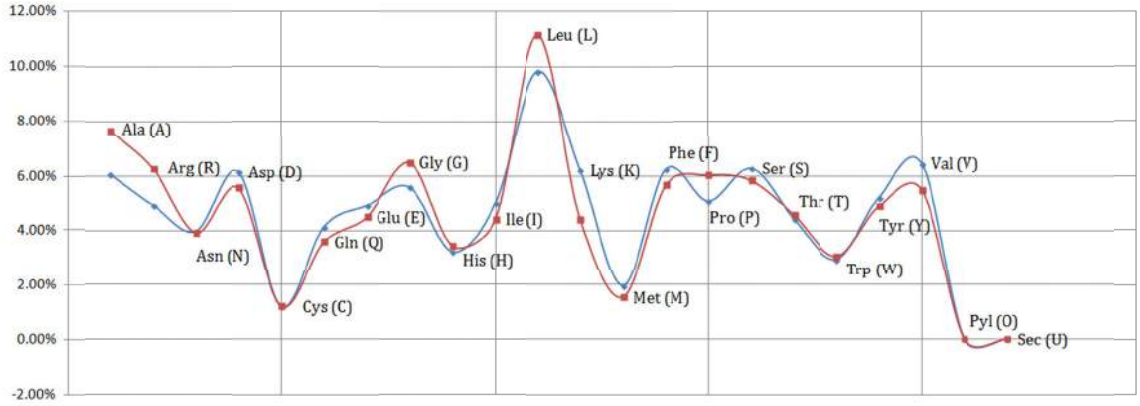


Figure 3. Amino acid composition analysis (*type a* highlighted in blue; *type b* highlighted in red)

4. Discussion

Proteins are essential building blocks of living cells; and life itself can be viewed as resulting substantially from the chemical activity of proteins (Lupas, 2014). Systematic approaches directed toward the evolution of proteins suggest that protein interactions were possibly the first forms of self-reproducing life on earth. As such, the chemical properties of each protein have been developed over billions of years. The α-amino acid composition of proteins contributes to the overall quality of a protein. If certain α-amino acids are optimal for protein structure, natural selection should have acted over evolutionary time to increase the frequency of those α-amino acids (Anthis et al., 2013).

As previously noted, the chemical properties of proteins are also conserved over much longer periods than individual codons; and frequency patterns in Klotho gene variants across wide ranging taxonomy confirm it. Yet, conflicting distribution ratios among L-Lysine, L-Alanine, and L-Arginine [in group *type a* versus group *type b*] is indicative of natural selection acting upon frequency distribution. Could this have consequences in terms of protein function that contribute toward Klotho-associated anti-aging mechanisms?

In fully examining the data presented in this study, it is tempting for me to draw premature [and definitive] conclusions based upon the preliminary appearance of a modest correlation; but correlation does not necessarily imply causation. Instead of over-speculating on functional cause-and-effect and probable implications thereof, I would simply stress the primary scope of this paper is not to presume on solutions, but rather, on associations and identification. From the perspective of comparative sequence analysis, I have outlined a set of significantly distinct biomarkers in molecular variation patterns of KLOTHO across wide ranging taxonomy; and, in the broadest sense, I demonstrated homology and relatedness based on protein variants of Klotho that is consistent with the early evolution of vertebrates. There is still considerable work to do to improve our understanding of age-related

mechanisms and the protein functions that play a decisive role. And while a great deal about protein evolution remains unresolved, much can be learned about protein evolution by trying to reconstruct ancient paralogy lineages via comparative analysis of modern proteins. With this in mind, I believe that my results may help provide meaningful reference toward a fully systematic investigation.

References

United States Fish and Wildlife Service. (2005). *Green sea turtle (Chelonia mydas)*. North Florida Field Office.

American Federation for Aging Research. (2006). *Theories of aging -- Why and how do we age?*

Florida State University. (2007). Chemists Kill Cancer Cells With Light-activated Molecules. *ScienceDaily*. Retrieved from http://www.sciencedaily.com/releases/2007/08/070808132019.htm

Minnesota State University. (2012). Molecular Weight Estimations and SDS PAGE.

University of Arizona. (2014). *The Chemistry of Amino Acids*.

Agnisola, C., & Femiano, S. (2006). Studies on routine metabolism in adult zebrafish, Danio rerio. *Acta Physiol., 187*.

Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical Science,* 364-376.

Andreollo, N. A., Santos, E. F. D., Araújo, M. R., & Lopes, L. R. (2012). Rat's age versus human's age: what is the relationship? ABCD. *Arquivos Brasileiros de Cirurgia Digestiva (São Paulo), 25*(1), 49-51. http://dx.doi.org/10.1590/S0102-67202012000100011

Anthis, N. J., & Clore, G. M. (2013). Sequence-specific determination of protein and peptide concentrations by asorbance at 205 nm. *Protein Science, 22*, 851-858. http://dx.doi.org/10.1002/pro.2253

Bastien, O, Aude, J. C., Roy, S., & Marechal, E. (2004). Fundamentals of massive automatic pairwise alignments of protein sequences: theoretical significance of Z-value statistics. *Bioinformatics, 20*(4), 534–537. http://dx.doi.org/10.1093/bioinformatics/btg440

Beals, M., Gross, L., & S. Harrell, S. (1999). *Amino Acid Frequency*. University of Tennessee. Retrieved from http://www.tiem.utk.edu/~gross/bioed/webmodules/α-amino acid.htm

Bouchard, C., Tremblay, A., Nadeau, A., Despres, J. P., Theriault, G., Boulay, M. R., ... & Fournier, G. (1989). Genetic effect in resting and exercise metabolic rates. *Metabolism, 38*(4), 364-370. http://dx.doi.org/10.1016/0026-0495(89)90126-1

Bryan, J. K. (1980). Aspartate family and branched-chain amino acids. In B. J. Miflin (Ed.), *The Biochemistry of Plants* (Vol. 5, pp. 403-452). New York: Academic Press.

Buffenstein, R. (2008). Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. *Journal of Comparative Physiology B, 178*(4), 439-445. http://dx.doi.org/10.1007/s00360-007-0237-5

Chao, K. M., & Zhang, L. (2008). *Sequence comparison: theory and methods* (Vol. 7). Springer.

Chen, C., Sander, J. E., & Dale, N. M. (2003). The effect of dietary lysine deficiency on the immune response to Newcastle disease vaccination in chickens. *Avian diseases, 47*(4), 1346-1351. http://dx.doi.org/10.1637/7008

Chen, L., Bush, D. R. (1997). LHT1, a lysine- and histidine-specific amino acid transporter in Arabidopsis. *Plant Physiol., 115*, 1127-1134. http://dx.doi.org/10.1104/pp.115.3.1127

Composition/Molecular Weight Calculation tool. (ND). University of Delaware. Retrieved from http://pir.georgetown.edu/pirwww/search/comp_mw.shtml

Csiszar, A., Labinskyy, N., Orosz, Z., Xiangmin, Z., Buffenstein, R., & Ungvari, Z. (2007). Vascular aging in the longest-living rodent, the naked mole rat. *American Journal of Physiology-Heart and Circulatory Physiology, 293*(2), H919-H927. http://dx.doi.org/10.1152/ajpheart.01287.2006

DALY, T. J. M., WILLIAMS, L. A., & BUFFENSTEIN, R. (1997). Catecholaminergic innervation of interscapular brown adipose tissue in the naked mole‐rat (Heterocephalus glaber). *Journal of anatomy, 190*(3), 321-326. http://dx.doi.org/10.1046/j.1469-7580.1997.19030321.x.

Dena, B. (2014). Life Extension Factor Klotho Enhances Cognition, Cell Reports, http://dx.doi.org/10.1016/j.celrep.2014.03.076.

Dokholyan, N. V., Mirny, L. A., & Shakhnovich, E. I. (2002). Understanding conserved amino acids in proteins. *Physica A: Statistical Mechanics and its Applications, 314*(1), 600-606. http://dx.doi.org/10.1016/S0378-4371(02)01079-8

Enstipp, M. R., Ciccione, S., Gineste, B., Milbergue, M., Ballorain, K., Ropert-Coudert, Y., ... & Georges, J. Y. (2011). Energy expenditure of freely swimming adult green turtles (Chelonia mydas) and its link with body acceleration. *The Journal of experimental biology, 214*(23), 4010-4020. http://dx.doi.org/10.1242/jeb.062943

Falco, S. C., Guida, T., Locke, M., Mauvais, J., Sanders, C., Ward, R. T., & Webber, P. (1995). Transgenic canola and soybean seeds with increased lysine. *Biotechnology (N.Y.), 13*, 577-582. http://dx.doi.org/10.1038/nbt0695-577

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution, 17*(6), 368–376. http://dx.doi.org/10.1007/BF01734359.

Galili, G. (1995). Regulation of lysine and threonine synthesis. *Plant Cell, 7*, 899-906. http://dx.doi.org/10.1105/tpc.7.7.899

Gibbs, R. A., Weinstock, G. M., Metzker, M. L., Muzny, D. M., Sodergren, E. J., Scherer, S., ... & Gill, R. (2004). Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature, 428*(6982), 493-521. http://dx.doi.org/10.1038/nature02426

Griffith, R. S., Norins, A. L., & Kagan, C. (1978). A multicentered study of lysine therapy in herpes simplex infection. *Dermatology, 156*(5), 257-267. http://dx.doi.org/10.1159/000250926

Hirschberg, D. S. (1975). A linear space algorithm for computing maximal common subsequences. *Communications of the ACM, 18*(6), 341–343. http://dx.doi.org/10.1145/360825.360861.

IUPAC, I. (1984). Nomenclature and symbolism for amino acids and peptides. *Eur J Biochem, 138*, 9-37.

Jopling, C., Sleep, E., Raya, M., Martí, M., Raya, A., & Belmonte, J. C. I. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature, 464*(7288), 606-609. http://dx.doi.org/10.1038/nature08899

Jumpertz, R., Hanson, R. L., Sievers, M. L., Bennett, P. H., Nelson, R. G., & Krakoff, J. (2011). Higher energy expenditure in humans predicts natural mortality. *The Journal of Clinical Endocrinology & Metabolism, 96*(6), E972-E976. http://dx.doi.org/10.1210/jc.2010-2944

Kim, B. (2008). Thyroid hormone as a determinant of energy expenditure and the basal metabolic rate. *Thyroid, 18*(2), 141-144. http://dx.doi.org/10.1089/thy.2007.0266

Kurosu, H., Yamamoto, M., Clark, J. D., Pastor, J. V., Nandi, A., Gurnani, P., … Kuro-o, M. (2005). Suppression of aging in mice by the hormone Klotho. *Science, 309*(5742), 1829-1833. http://dx.doi.org/10.1126/science.1112766

Lassmann, T., & Sonnhammer, E. L. (2005). Kalign—An Accurate and Fast Multiple Sequence Alignment Algorithm. *BMC Bioinformatics, 6*, 298. http://dx.doi.org/10.1186/1471-2105-6-298

Little, A. G., & Seebacher, F. (2014). Thyroid hormone regulates cardiac performance during cold acclimation in zebrafish (Danio rerio). The Journal of experimental biology, 217(5), 718-725. http://dx.doi.org/10.1242/jeb.096602

Mannakee, B., & Gutenkunst, R. (2012). *Influence of Protein Network Dynamics on Protein Evolution*. University of Arizona.

Marsh, J. A., & Teichmann, S. A. (2010). How do proteins gain new domains. *Genome Biol, 11*(7), 126. http://dx.doi.org/10.1186/gb-2010-11-7-126

Martin, A. P., & Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences, 90*(9), 4087-4091. http://dx.doi.org/10.1073/pnas.90.9.4087

Mueller, P., & Diamond, J. (2001). Metabolic rate and environmental productivity: well-provisioned animals evolved to run and idle fast. *Proceedings of the National Academy of Sciences, 98*(22), 12550-12554. http://dx.doi.org/10.1073/pnas.221456698

NHLBI Communications Office. (2004). *Scientists Compare Rat Genome With Human, Mouse -- Analysis Yields New Insights into Medical Model, Evolutionary Process*. National Institutes of Health.

Pérez, V. I., Buffenstein, R., Masamsetti, V., Leonard, S., Salmon, A. B., Mele, J., ... & Chaudhuri, A. (2009). Protein stability and resistance to oxidative stress are determinants of longevity in the longest-living rodent, the naked mole-rat. *Proceedings of the National Academy of Sciences, 106*(9), 3059-3064. http://dx.doi.org/10.1073/pnas.0809620106

Robinson, M. (2012). *In response to: Which is more informative: a phylogenetic tree based on alignment of*

protein α-amino acid sequences or one based on the corresponding DNA sequences? Los Alamos National Laboratory. Retrieved from Research Gate.

Rozing, M. P., Westendorp, R. G., de Craen, A. J., Frölich, M., Heijmans, B. T., Beekman, M., ... & van Heemst, D. (2009). Low serum free triiodothyronine levels mark familial longevity: the Leiden Longevity Study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, glp200.

Speakman, J. R., Selman, C., McLaren, J. S., & Harper, E. J. (2002). Living fast, dying when? The link between aging and energetics. *The Journal of nutrition, 132*(6), 1583S-1597S.

Sun, P., Zhang, Y., Yu, F., Parks, E., Lyman, A., Wu, Q., ... & Hsiai, T. K. (2009). Micro-electrocardiograms to study post-ventricular amputation of zebrafish heart. *Annals of biomedical engineering, 37*(5), 890-901. http://dx.doi.org/10.1007/s10439-009-9668-3

Tan, T. C., Rahman, R., Jaber-Hijazi, F., Felix, D. A., Chen, C., Louis, E. J., & Aboobaker, A. (2012). Telomere maintenance and telomerase activity are differentially regulated in asexual and sexual worms. *Proceedings of the National Academy of Sciences, 109*(11), 4209-4214. http://dx.doi.org/10.1073/pnas.1118885109

Torres, P. U., Prie, D., Molina-Bletry, V., Beck, L., Silve, C., & Friedlander, G. (2007). Klotho: an antiaging protein involved in mineral and vitamin D metabolism. *Kidney international, 71*(8), 730-737. http://dx.doi.org/10.1038/sj.ki.5002163

Vogt, G., Etzold, T., & Argos, P. (1995). An assessment of amino acid exchange matrices in aligning protein sequences: the twilight zone revisited. *Journal of molecular biology, 249*(4), 816-831. http://dx.doi.org/10.1006/jmbi.1995.0340

Vogt, G., Etzold, T., & Argos, P. (1995). An assessment of amino acid exchange matrices in aligning protein sequences: the twilight zone revisited. *Journal of molecular biology, 249*(4), 816-831. http://dx.doi.org/10.1006/jmbi.1995.0340

Warwicker, J., Charonis, S., & Curtis, R. A. (2013). Lysine and Arginine Content of Proteins: Computational Analysis Suggests a New Tool for Solubility Design. *Molecular pharmaceutics, 11*(1), 294-303. http://dx.doi.org/10.1021/mp4004749

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).