Short Communication

Molecular conservation of the mammalian leptin protein

J.E. Gabriel[1] and K.C.F. Lidani[2]

[1]Colegiado de Ciências Biológicas, Centro de Ciências Agrárias,
Universidade Federal do Vale do São Francisco, Petrolina, PE, Brasil
[2]Departamento de Patologia Médica, Hospital de Clínicas,
Universidade Federal do Paraná, Curitiba, PR, Brasil

Corresponding author: J.E. Gabriel
E-mail: jane.gabriel@univasf.edu.br

ABSTRACT. In this study, we comparatively assessed multiple sequences of the leptin protein from different animal species to establish new insights into conservation degree of biological sequences and evolutionary biology among mammals using computational biology tools. First, amino acid sequences of the leptin protein from *Homo sapiens* (human, P41159), *Sus scrofa* (wild pig, Q29406), *Felis catus* (domestic cat, Q29406), *Rattus norvegicus* (rat, P50596), and *Mus musculus* (mouse, P41160) were randomly searched in the high-quality annotated and non-redundant protein sequence database UniProtKB/Swiss-Prot. A dendogram showing the evolutionary relationships among specimens was constructed from the sequences of interest using the Mega 6.0 software with the neighbor-joining method. The resulting tree presenting the evolutionary relationships among specimens inferred from amino acid sequences of the leptin protein in mammals demonstrated 2 main branches: 1 cluster including the rat and mouse species (0.02) and a second cluster containing both

wild pig and domestic cat species grouped in a sub-branch (0.04 and 0.06, respectively), linking them to the human sequence (0.08). These findings were reinforced by comparing estimates of evolutionary divergence among leptin sequences analyzed. Based on comparative analyses of multiple sequence alignments in the present study, there was a stronger conservation degree of the leptin protein in evolutionarily close species and several conservative changes along the sequences of interest, revealing information regarding the evolutionary biology among mammals.

Key words: Bioinformatics; Conservation degree; Evolutionary biology; Leptin protein; Mammals

INTRODUCTION

Leptin is an adipose tissue-derived circulating protein hormone with multiple biological roles associated with regulation of energy homeostasis and reproductive functions in vertebrates (Zhang et al., 1994; Considine and Caro, 1997; Houseknecht et al., 1998). Several studies have described that increases in leptin signal levels act on the central nervous system to inhibit food intake and/or regulate energy expenditure in order to maintain a constant adipose mass (Friedman and Halaas, 1998). Although the leptin hormone is predominantly expressed by adipocytes, smaller amounts of this peptide are also secreted by cells in the epithelium of the stomach and in the placenta (Bartha et al., 2005). Leptin receptors are highly expressed in areas of the hypothalamus, which is known to be important for regulating body weight, as well as in T lymphocytes and vascular endothelial cells. Furthermore, Zhang et al. (1994) found that mutations in the leptin gene are associated with a myriad of hormonal and metabolic alterations.

The leptin gene and protein have been investigated for their structural and functional aspects and in evolutionary biology studies, indicating phylogenetic relationships among distinct species. Doyon et al. (2001) cloned and sequenced partial nucleotide sequences of leptin from 6 mammalian species and verified that these sequences in chicken and turkey were clustered with mouse and rat leptins, suggesting a convergent or parallel evolution to explain their high similarity. Recently, Yu et al. (2011) examined the evolution of mammalian *leptin* genes in response to selective pressures from life in an aquatic environment, describing positive selection on the leptin gene in toothed whales of the Cetacea and in earless seals of the Pinnipedia.

Over the past several decades, studies using computational biology tools have focused efforts to improve the understanding of living systems through computation approaches (Baxevanis and Ouellette, 2005; Schatz et al., 2010). The biological fields have progressed tremendously in the ability to process and interpret large volumes of biological data using computational biology analyses, advancing molecular and structural characterization of bioactive molecules of interest (Schatz et al., 2010; Gabriel et al., 2013). The purpose of the present study was to comparatively assess multiple sequences of the leptin protein in distinct animal species to establish new insights into the conservation degree of biological sequences and evolutionary biology among mammals using computational biology tools.

MATERIAL AND METHODS

The amino acid sequences of the leptin protein from *Homo sapiens* (human, P41159), *Sus scrofa* (wild pig, Q29406), *Felis catus* (domestic cat, Q29406), *Rattus norvegicus* (rat, P50596), and *Mus musculus* (mouse, P41160) were randomly selected from the high-quality annotated and non-redundant protein sequence database UniProtKB/Swiss-Prot. This database allows for the functional annotation of multiple protein sequences and accurate classification of whole proteins into a network structure reflecting their evolutionary relationships (Boutet et al., 2007). Next, leptin protein sequences were subjected to progressive alignments using the constraint-based multiple protein alignment tool COBALT to annotate multiple protein sequence from conserved domain database, protein motif, and local information of the sequence similarity, including additional data to improve the quality score and significance of the matches, such as expected value (e-value) (Papadopoulos and Agarwala, 2007). E-values from multiple alignment analyses represent the probability of the alignment occurring by chance and are calculated based on the quality of alignment (the score) and the size of the database. Moreover, statistical evaluative parameters concerning identity and similarity values were automatically validated in these comparative analyses. Finally, a dendogram demonstrating the evolutionary relationships among specimens was constructed from sequences of interest using the Mega 6.0 software (Tamura et al., 2013) using the neighbor-joining (NJ) method (Saitou and Nei, 1987) and computed using the Kimura 2-parameter method (Kimura, 1980). The experimental assays involved 5 amino acid sequences in a total of 167 positions in the final dataset. The reliability of different phylogenetic groupings was evaluated using a boot-strap value of 1000 replicates. Divergence times for all branching points in the topology were calculated using the RelTime method (Tamura et al., 2012) using the branch lengths contained in the inferred tree. Bars around each node represent 95% confidence intervals, which were computed using the method described by Tamura et al. (2013).

RESULTS

Phylogenetic groupings were established to investigate the overall evolutionary relationships among specimens inferred from amino acid sequences of the leptin protein in mammals, as shown in Figure 1. The NJ tree showed 2 main branches obtained from multiple alignments of amino acid sequences of the mammalian leptin protein: 1 cluster with the rat and mouse species grouped in a branch (0.02) and a second cluster containing both wild pig and domestic cat species grouped in a sub-branch (0.04 and 0.06, respectively), which was linked to the human sequence (0.08) (Figure 1). Such findings were reinforced by comparing the estimates concerning evolutionary divergence inferred from amino acid sequences of the leptin protein. This demonstrated that these sequences were the most differentiated between human and mouse species (0.213), whereas the sequences of the leptin protein were the least differentiated in human and domestic cat (0.037). Comparative assessments of the human leptin between wild pig and rat species revealed an estimate of evolutionary divergence ranging of 0.162-0.191, respectively, indicating that the leptin protein among these species showed intermediary evolutionary differentiation compared to other sequences examined.
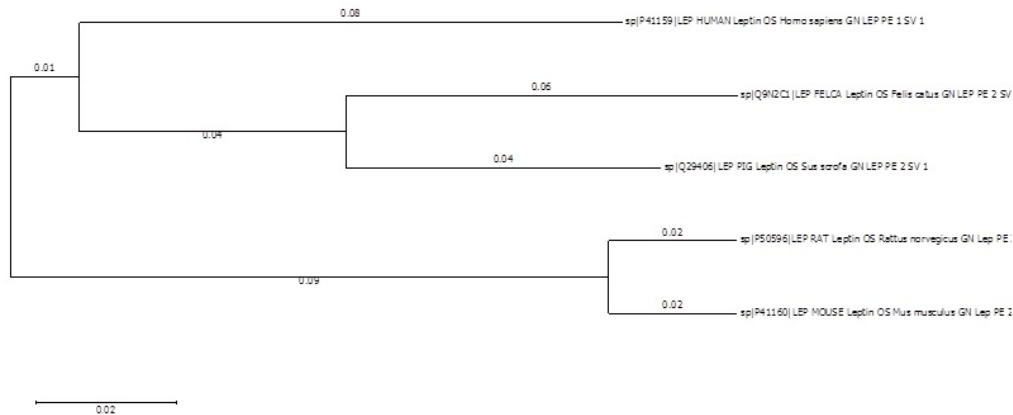
Figure 1. Evolutionary relationships among specimens inferred from amino acid sequences of the leptin protein in different mammals.

Comparative alignment analyses of multiple sequences of the leptin protein revealed an identity value of 71.257% in Swiss-Prot analyses, with 119 residues detected as identical positions in a total content of 167 amino acids (data not shown). Moreover, 30 amino acid residues (approximately 17.96%) were detected as similar positions along the 5 sequences analyzed. In the present study, all comparative assessments generated e-values of 0.003, suggesting that the data obtained in these alignments is significant. Interestingly, comparative alignment of leptin sequences performed exclusively between rat and mouse species revealed a higher identity value corresponding to 96.407%, identifying 6 and 161 amino acid residues in similar and identical positions, respectively (data not shown).

DISCUSSION

The leptin hormone has received increasing attention because of its importance in energy homeostasis and reproductive regulation in vertebrates. A strong conservation degree of multiple sequences of the leptin protein has been observed in distinct mammals, particularly in evolutionarily close species (Figure 1). Sequences of the leptin protein showed high identity between rat and mouse in our study (Figure 1), which agreed with the results obtained by Funahashi et al. (1995) who demonstrated a high degree of conservation of the leptin protein between rat and mouse (96% at the protein level). To gain insight into the structure-function relationships and origin of leptin, Ravishankar Ram et al. (2007) analyzed the amino acid sequences of leptin in 23 species by comparing the frequency specific amino acids, their secondary structure, and sequence homology. They demonstrated extensive conservation within the leptin sequences in all species, particularly the human leptin, which shares a very high degree of homology with gorilla, chimpanzee, and orangutan, indicating a common function of leptin in these related species, and possibly reflecting the closeness of their evolution.

As shown in Figure 1, evolutionary relationships among mammals inferred from amino acid sequences of the leptin protein showed that cat and human species were the least differentiated. Interestingly, Wienberg et al. (1997) revealed widespread conservation of genome organization between the 2 mammalian orders and confirmed 90% of the homologous

genes mapped to both cat and human species by fluorescence *in situ* hybridization analysis. Approximately 30 syntenic segments were previously identified, and the number of translocations was estimated to be in the order of 1 new translocation per 10 million years in the phylogenetic lines, leading to human and cat species (Wienberg et al., 1997).

As described by Denver et al. (2011), while vertebrate leptins show large divergence in their primary amino acid sequence, they form similar tertiary structures and may have similar potencies when tested *in vitro* on heterologous leptin receptors. Amino acid residues (approximately 17.96) were identified to be in similar positions in the leptin protein sequences analyzed. It is well-established that amino acid substitutions with identical chemical properties signal the presence of conservative changes throughout sequences of interest, indicating that an amino acid may be substituted for another with similar chemical properties or for an amino acid with different chemical characteristics (Baxevanis and Ouellette, 2005). Moreover, the relevant evaluative parameters regarding the significance of the matches and their conserved domains were observed in the constraint-based alignments, with an e-value of 0.003. This indicates that the probability of alignments occurring by chance is low (Baxevanis and Ouellette, 2005).

In conclusion, we provided insight into the evolutionary biology of mammals based on the accentuated degree of conservation and the presence of conservative changes along mammalian leptin proteins by comparing multiple sequence alignments of biological molecules.

ACKNOWLEDGMENTS

REFERENCES

Bartha T, Sayed-Ahmed A and Rudas P (2005). Expression of leptin and its receptors in various tissues of ruminants. *Domest. Anim. Endocrinol.* 29: 193-202.

Baxevanis AD and Ouellette BFF (2005). Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. John Wiley and Sons Inc, New Jersey, 560.

Boutet E, Lieberherr D, Tognolli M, Schneider M, et al. (2007). UniProtKB/Swiss-Prot. *Methods Mol. Biol.* 406: 89-112.

Considine RV and Caro JF (1997). Leptin and the regulation of body weight. *Int. J. Biochem. Cell Biol.* 29: 1255-1272.

Denver RJ, Bonett RM and Boorse GC (2011). Evolution of leptin structure and function. *Neuroendocrinology* 94: 21-38.

Doyon C, Drouin G, Trudeau VL and Moon TW (2001). Molecular evolution of leptin. *Gen. Comp. Endocrinol.* 124: 188-198.

Friedman JM and Halaas JL (1998). Leptin and the regulation of body weight in mammals. *Nature* 395: 763-770.

Funahashi T, Shimomura I, Hiraoka H, Arai T, et al. (1995). Enhanced expression of rat obese (ob) gene in adipose tissues of ventromedial hypothalamus (VMH)-lesioned rats. *Biochem. Biophys Res. Commun.* 211: 469-475.

Gabriel JE, Figueiredo DD and Farias RP (2013). Revealing highly conserved regions in the E6 protein among distinct human papillomavirus types using comparative analysis of multiple sequence alignments. *Braz. J. Biol.* 73: 449-450.

Houseknecht KL, Baile CA, Matteri RL and Spurlock ME (1998). The biology of leptin: a review. *J. Anim. Sci.* 76: 1405-1420.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.

Papadopoulos JS and Agarwala R (2007). COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 23: 1073-1079.

Ravishankar Ram M, Beena G, Ragunathan P and Malathi R (2007). Analysis of structure, function, and evolutionary origin of the ob gene product-leptin. *J. Biomol. Struct. Dyn.* 25: 183-188.

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Schatz MC, Delcher AL and Salzberg SL (2010). Assembly of large genomes using second-generation sequencing. *Genome Res.* 20: 1165-1173.

Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, et al. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* 109: 19333-19338.

Tamura K, Stecher G, Peterson D, Filipski A, et al. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.

Wienberg J, Stanyon R, Nash WG, O'Brien PC, et al. (1997). Conservation of human vs. feline genome organization revealed by reciprocal chromosome painting. *Cytogenet Cell Genet.* 77: 211-217.

Yu L, Jin W, Zhang X, Wang D, et al. (2011). Evidence for positive selection on the leptin gene in Cetacea and Pinnipedia. *PLoS One* 6: e26579.

Zhang Y, Proenca R, Maffei M, Barone M, et al. (1994). Positional cloning of the mouse obese gene and its human homologue. *Nature* 372: 425-432.