

Uma série temporal é qualquer conjunto de observações ordenadas no tempo. As sequências de DNA também podem ser consideradas séries temporais, desde que se faça o uso de alguma transformação numérica. A longa dependência é a presença, em uma série temporal, de significativa correlação ou dependência entre as observações separadas por um longo período de tempo. Modelos que incluem diferenciação fracionária $d \in (0,0;0,5)$ são capazes de representar séries temporais com longa dependência. Entre estes modelos encontram-se os ARFIMA, utilizados neste trabalho. No estudo de séries temporais encontramos diversos artigos que analisam a longa dependência em sequências de DNA (ver Peng et al., 1992; Guharay et al., 2000). Nosso objetivo é verificar a existência ou não de longa dependência em sequências de DNA. Para isto, é preciso utilizar uma função que transforme suas bases nitrogenadas (timina, citosina, guanina e adenina) em sequências numéricas. Aqui utilizamos as regras SW e RY, que atribuem os valores -1, 0 ou 1 a cada uma das bases. Analisamos sete sequências de DNA com o objetivo de verificar a existência ou não de longa dependência, seis delas do *Homo sapiens* e uma da espécie *Leishmania brasiliensis*. Para isto, estimamos o parâmetro de diferenciação d através de diversos métodos de estimação paramétricos e semiparamétricos, em versões clássica e robusta. São eles: GPH (ver Geweke e Porter-Hudak, 1983), R (ver Robinson, 1995), BA (ver janela espectral de Bartlett), DFA (ver “Detrended Fluctuation Analysis” em Peng et al., 1994), W (ver Whittle, 1953), R/S(n) (ver Hurst, 1951) e R/S(q) (ver Lo, 1991). Todas as sequências apresentaram a característica de longa dependência, utilizando qualquer uma das transformações propostas. Em cada uma das sequências a longa dependência é estatisticamente significativa, com 95% de confiança, para todos os estimadores considerados.