

論文内容の要旨

論文題目 ロイシンリッチ核外移行シグナルに関する研究：データセットの構築及び新規特徴量解析による予測法の改良 (Better understanding and recognition of leucine-rich nuclear export signals: expanded dataset, novel feature analysis, and the development of an improved prediction method)

氏名 傅 思縉 (Fu, Szu-Chin)

Protein sub-cellular localization is an important feature and has been commonly used to support many functional hypotheses. The leucine-rich nuclear export signal (NES) is an important sub-cellular targeting signal, which is involved in processes such as signal transduction and cell cycle regulation. Although 15 years has passed since its discovery, limited structural information and high sequence diversity have hampered understanding of this signal. A consensus sequence was proposed based on early examples, but later evidence demonstrated its low sensitivity (~37%). To raise the sensitivity, a more general consensus sequence has been widely used at a cost of greatly increased spurious matches. Despite continued interest amongst molecular biologists in the function and regulation of NES-containing proteins, further bioinformatic characterization of this import signal remains at a standstill. Indeed, most of the recently discovered NES sites have been identified by the consensus sequence despite its unsatisfactory trade-off. On the other hand, the NetNES server provides the only computational method currently available. Although these two methods have been widely used to attempt to find the correct NES position within potential NES-containing proteins, their performance has not yet been evaluated on the basic task of discriminating NES-containing proteins from other proteins. To better characterize the NES, we propose a new approach, NESsential, not only capable of finding the correct position of many NES's at the site level, but potential NES-containing proteins at the protein level. We also collected 70 NES-containing proteins recently discovered to update the dataset to approximately two-fold larger than NESbase, the largest previously available dataset.