

Recent Advances in Predicting Protein Nuclear Export Signals

Szu-Chin Fu

Sequence Analysis Team, CBRC, AIST

The therapeutic potential of the CRM1-mediated nuclear export pathway and disease relevance of its cargo proteins has gained recognition in recent years. CRM1 cargo proteins are often recognized by a classical nuclear export signal (leucine-rich NES), but this signal is notoriously difficult to predict from sequence alone. We propose a new predictor, NESsential, which uses predicted structural disorder and other features to predict NES's from amino acid sequence. We demonstrate that NESsential can identify promising NES-containing candidate proteins (albeit at low coverage), but other methods cannot. Meanwhile, only about one third of known CRM1 cargo proteins are accessible in a single database since the last compilation in 2003. We therefore develop a new database, ValidNESs, making a threefold increase in database entries for public use. Finally, we integrate a web interface to prediction by NESsential into ValidNESs, enabling valuable hints to be gained by *in silico* prediction