Structure organization of the innnexin family: integration of computational methods and molecular data - (session: Comparative Genomics and Molecular Evolution)

Chiusano Maria Luisa, Potenza Nicoletta, Del Gaudio Rosanna, Giuseppina Maria Rosaria Russo, Di Giaimo Rossella, Mondola Tiziana and Geraci Giuseppe

Dip. Genetica, Biologia Generale e Molecolare, Napoli

Innexins are a family of membrane proteins involved in the formation of gap junctions in invertebrates. They have been found to participate in several aspects of cell differentiation and in the production of embryonic structures through the formation of specific intercellular channels. These proteins appear to be ubiquitous because, from the discovery in D. melanogaster (Lipshitz and Kankel, 1985) and in C. elegans (Starich et al., 1993), it has been shown that they are present in the mollusc Clione limacina and in the flatworm Girardia tigrina (Panchin et al., 2000), in the annelid Hirudo medicinalis (Alexopoulos et al., 2000) and in the polychaete annelid worm Chaetopterus variopedatus (Potenza et al., 2002). Moreover, several genes encode proteins of this family in each species (Curtin et al., 1999). As an example, the genome of D. melanogaster encodes at least ten different proteins (Stebbings et al., 2002), while C. elegans encodes twentyfive different innexins. Moreover, these proteins are not interchangeable in their function (Curtin et al., 2002).
We present here a computational analysis of this family of proteins. Genes belonging to innexin family were collected from Genbank and their structures were aligned with the corresponding protein sequences and with the information derived from predictive methods to localize transmembrane regions. A phylogenetic analysis of all known innexins was performed on the protein sequences resulting in a tree where insects and other invertebrate innexins are in distinctive clusters when compared to the nematode. While the comparative analysis of the proteins shows similarity at the level of the structural organization and in the clustering of data from the same species, there is a high heterogeneity at amino acid level and at the level of the gene structures. These differences are evident even for genes that are in close contiguity on the same chromosome. The computational methodology reported, consisting in the comparison between different types of structural data, reveals unexpected features also when applied in the study of the organization of multigene families (Chiusano, 1999; Chiusano, 2000).

1) Lipshitz HD, Kankel DR (1985) Specificity of gene action during central nervous system development in Drosophila melanogaster: analysis of the lethal (1) optic ganglion reduced locus. Dev Biol 108(1): 56-77
2) Starich TA, Herman RK, Shaw JE (1993) Molecular and genetic analysis of unc-7, a Caenorhabditis elegans gene required for coordinated locomotion. Genetics 133: 527-541
3) Alexopoulos H, Dykes IM, Bacon JP, Davies JA (2000) Novel innexins in snails and leeches. Eur J Neurosci 12 (Suppll 11): 15
4) Potenza N, del Gaudio R, Rivieccio L, Russo GMR, Geraci G (2002) Cloning and molecular characterization of the first innexin of the phylum Annelida. Expression of the gene during development. J Mol Evol 54: 312-321
5) Curtin KD, Zhang Z, Wyman RJ (1999). Drosophila has several genes for gap junction proteins. Gene 232(2):191-201.
6) Stebbings LA, Todman MG, Phillips R, Greer CE, Tam J, Phelan P, Jacobs K, Bacon JP, Davies JA (2002) Gap junctions in Drosophila: developmental expression of the entire innexin gene family. Mechanisms of Development 113: 197-205
7) Curtin KD, Zhang Z, Wyman RJ, (2002). Gap junction proteins are not interchangeable in development of neural function in the Drosophila visual system. J Cell Sci 115:3379-88
8) Chiusano M.L., D,Onofrio G., Alvarez-Valin F., Jabbari K., Colonna G., Bernardi G. (1999). Correlations of Nucleotide Substitution Rates and Base Composition of Mammalian Coding Sequences with Protein Structure. Gene 238, 23-31.
9) Chiusano M.L., Alvarez-Valin F., Di Giulio M., D'Onofrio G., Ammirato G., Colonna G.and Bernardi G. (2000). Second codon position of genes and the secondary structures of proteins: relationships and implications for the origin of the genetic code. Gene 261, 63-69