Mapping OMIM mutated residues on PDB protein structures

Peluso D, Via A, Ausiello G, Helmer-Citterich M

Centro di Bioinformatica Molecolare, Department of Biology,
University of Tor Vergata, Rome

Motivation

The OMIM database (1) is a collection of hereditary point mutations associated to diseases in Homo sapiens. Such mutations have been recently mapped onto the Swiss-Prot sequences (2). Despite the increasing number of protein structures stored in the PDB database, a correspondence between OMIM mutated residues and PDB residues has not yet been established. This type of information could help in providing insight into the molecular mechanisms that cause hereditary diseases. In this work, we performed an accurate mapping of OMIM mutations onto 3D protein structures. The results of the entire procedure will be used to further annotate amino acids in the pdbFun (3) database and will be available through the resource web site (http://pdbfun.uniroma2.it/).

Methods

The Swiss-Prot sequence numbers of OMIM missense mutations, provided by Martin and coworkers (2), have been transferred to PDB structures via the seq2struct resource (4), which establishes reliable links between Uniprot sequences and PDB or SCOP structures. Swiss-Prot sequences have been aligned to the sequences extracted from the ATOM coordinates of the PDB files, by retaining only sequence-structure pairs displaying a sequence identity greater than 90%. A supplementary BLAST local alignment, based on more stringent thresholds, was performed in a short region including the mutated amino acid on the Swiss-Prot sequence, in order to obtain a punctual residue mapping.

Results

In a non-redundant PDB set of structures, we found 3376 mutations associated to 1308 mutation sites, which are linked to 178 OMIM entries. As soon as the mapping is complete, we intend to use the OMIM mapping on protein structures to explore and further analyze cases where the effects of a mutation onto a protein structure (might) account for the protein misfunction and, hence, for the associated disease (5). In the future, we want to apply the mapping procedure and the subsequent structural analysis also to SNPs (6).

Contact email: daniele@cbm.bio.uniroma2.it

References

1. Hamosh,A., Scott,AF., Amberger,J., Valle,D., McKusick,VA. (2000). Mendelian Inheritance in Man (OMIM). Hum Mutat.;15(1):57-61.
2. Andrew C. R. Martin. (2005). Mapping OMIM mutations to SwissProt . Bioinformatics, Application Note, Vol. 00 no.00 pages 1-2
3. Ausiello,G., Zanzoni,A., Peluso,D., Via,A., Helmer-Citterich,M. (2005). pdbFun: mass selection and fast comparison of annotated PDB residues. Nucleic Acids Res.;33(Web Server issue):W133-7.
4. Via,A., Zanzoni,A., Helmer-Citterich,M. (2005). Seq2Struct: a resource for establishing sequence-structure links. Bioinformatics;21(4):551-3. Epub 2004 Sep 28.
5. Yue,P., Li,Z. and Moult,J. (2005). Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease. J. Mol. Biol. 353,495-473.
6. Wang,Z. and Moult, J. (2001). SNPs, Protein Structure, and Disease. Human Mutation 17:263-270.