

Literature curation of protein interactions

Michael Cusick

Nowadays every project in interactome mapping, large or small, requires a substantial contribution of functional annotation and quality assessment from curation. An empirical framework for assessing quality parameters of high-throughput binary interactome mapping finds that high-throughput datasets can be as good as, or better, than literature-curated interactome datasets, for multiple species (yeast, human, worm). In developing reference datasets necessary for this empirical framework, we evaluated existing curation of protein interaction experiments reported in the literature, finding that curation can often, but not always, be error-prone and possibly of lower quality than commonly assumed. This problem is not due to curator ineptitude, but to the simple reality that extraction of accurate information from long free-text documents is much harder than often appreciated. There are initiatives and projects underway to improve the quality of literature curation of protein interactions. Several stories of how curation improves understanding of interactome networks will be reported.