

Table 2 Developmental dataset

Expression densities per gene per time-point as obtained from Liscovitch and Chechik (2013).

Table 3 Developmental dataset protein screens

Each gene was matched to its MGI symbol and Ensembl gene identifier. From these, protein sequences were obtained - where multiple isoforms were found, the longest sequence was used for further analyses. Fasta format sequences were screened through PSORT II for cellular localization (PSORT II prediction and probability: gol, Golgi; cyt, cytoplasmic; csk, cytoskeletal; end, endoplasmic reticulum; ext, extracellular, including cell wall; mit, mitochondrial; nuc, nuclear; per, peroxisomal; pla, plasma membrane; ves, vesicles of secretory system). PSORT II uses two-fold k -nearest neighbor (k -NN) algorithm with $k_1 = 9$ and $k_2 = 23$ (Horton and Nakai 1997). Screens for number of predicted transmembrane domains (TMs) with PSORT II were also performed. SignalP detected signal peptides (sigp) on proteins. Big-PI predictor tested different potential GPI-modification sites and returned the GPI best score, GPI profile score, GPI profile independent score and GPI quality (Eisenhaber, Bork et al. 1998, Eisenhaber, Bork et al. 1999, Sunyaev, Eisenhaber et al. 1999, Eisenhaber, Bork et al. 2000). GPI quality of P or S indicated a potential GPI-modification site (P: predicted; S: second predicted site, more than one predicted site). TMHMM screened for the presence (Transmembrane Domain) and number of transmembrane domains (PredHel)(Sonnhammer, von Heijne et al. 1998, Krogh, Larsson et al. 2001). Columns AE, AF and AG show SMART ID, PFAM ID and InterPro short description, respectively. Genes were screened for additional GO terms (see text).