SUMMARY

Understanding the complex interplay between proteins in physiological contexts poses a great challenge, since a striking number of human proteins still remain without solid functional annotation. To gain insight into the function of proteins encoded by human chromosome 21 (Hsa21) and to identify regulatory networks potentially relevant to Down syndrome (DS), 206 annotated open reading frames (ORFs) on Hsa21 were used as basis for amplification and cloning of ORFs and for collection of available full-length cDNA clones. This work resulted in the largest library of Hsa21 ORF clones available to date, covering 81% of the 206 annotated ORFs. A high success rate was reached in production of recombinant proteins (80%), ensuring appropriate downstream functional genomics analyses.

For high-throughput determination of subcellular protein localizations, 89 constructs encoding N-terminally tagged proteins were introduced into HEK293T cells on transfected cell arrays. Localization information could be obtained for 52 out of the 89 Hsa21 proteins (58%). Of these, 28 subcellular localizations were described for the first time (published in Hu *et al.* 2006). All Hsa21 localization information has been entered into the public Gene Ontology (GO) database.

For the identification of protein-protein interactions (PPIs), 53 constructs encoding non-autoactivating baits were screened against a prey set of 5,640 human proteins using a high-throughput mating array-based yeast two-hybrid (Y2H) system. Altogether, this mating array identified 56 new interactions for 24 different Hsa21 proteins. Twenty-three of the new interactions could be verified in an independent cotransformation Y2H set-up (56% of 41 tested). Cellular colocalization experiments in COS-1 cells and pull-down experiments confirmed five of six tested protein pairs (83%).

All previously known protein interaction data for Hsa21 proteins were retrieved from available public sources. This resulted in a set of 684 new and known PPIs for 108 Hsa21 proteins, representing the largest Hsa21-centered interaction network to date. A proteome-wide interaction network was then generated by computational searches for indirect protein interactors in available data sets. Analysis of transcription factors showed a compelling inter-connection of thirteen Hsa21-encoded transcription factors. Detailed functional analysis of sub-networks involving NRIP1, UBE2G2, PCP4, MCM3AP and C21orf127 revealed novel functional properties of these Hsa21-encoded proteins. 35 Hsa21 proteins were found connected to nine signaling pathways, some of which have been previously implicated in the pathogenesis of DS, while others appear in this context for the first time.

Further expansion of the Hsa21 interaction network will support the identification of new protein functions and regulatory networks. The data presented here shows that a Hsa21 interaction network is a useful resource for elucidating the biological contexts of new PPIs, providing an essential tool for the systems biology of Down syndrome.