3.6 Assembly and analysis of Hsa21 interaction networks

3.6.1 Computational generation of a Hsa21 interaction network

To integrate the set of 108 Hsa21 proteins and 547 direct interactors into a proteome-wide interaction network, iterative searches were performed in the available human PPI data (BIND, HPRD and large-scale human Y2H PPI sets; see Methods) aimed at the identification of interactors of the Hsa21 protein partners. It has to be noted that 30 Hsa21 proteins interact with themselves, thus being able to form dimers or oligomers.

In the first round of search for indirect interactors, it was found that 346 of the 547 direct Hsa21 interactors (proteins in 'shell 1') participate in 630 interactions with other shell 1 proteins (Figure 3-19, yellow sphere). This means that close to 2/3 of the proteins in shell 1 (63%) are interconnected among themselves, a finding that underlines the importance of network analyses for the understanding of the interplay of Hsa21 proteins.

In addition, 437 proteins in shell 1 were found to form 4,779 interactions with 2,542 proteins that form shell 2 of the network (Figure 3-19, green sphere). The connectivity within shell 2 is even higher than in shell 1 (72%, 3,826 PPIs among 1,837 proteins of shell 2), probably due to over-representation of proteins with several binding partners ('hub proteins').

When performing a second round of searching indirect binding partners, 1,420 proteins of shell 2 were found to form 4,354 PPIs with 2,463 different proteins in what is then shell 3 (Figure 3-19, blue sphere). The intra-shell connectivity in shell 3 is only 35% of all shell 3 proteins, involving 709 PPIs among 859 different proteins. Hub proteins seem to have already been assigned to previous shells. The entire Hsa21 PPI network described here is composed of 5,660 proteins establishing a total of 14,982 connections. PPI data forming this network can be accessed online via the '21ppi' database (http://chr21.molgen.mpg.de/21ppi).

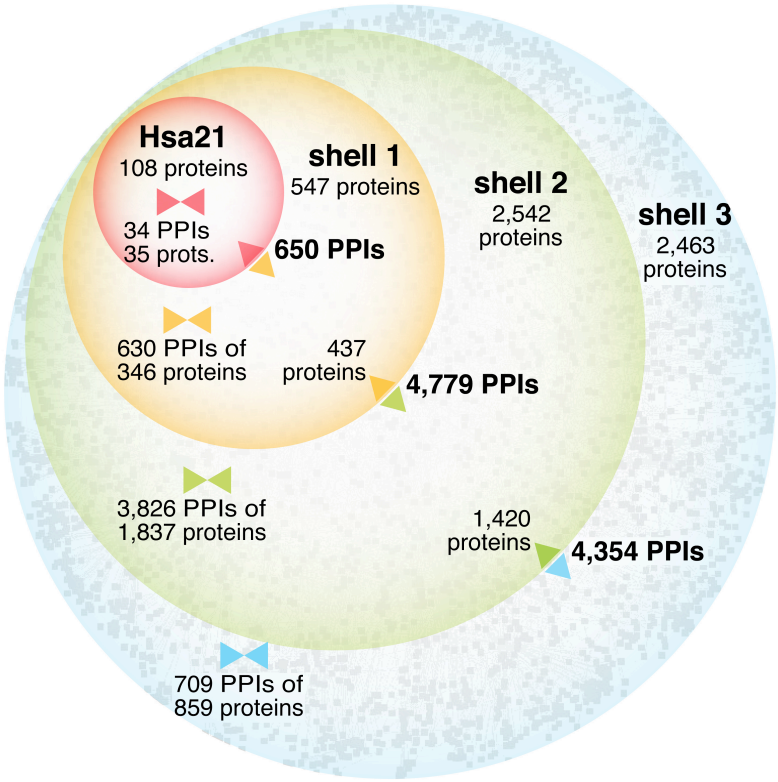

Figure 3-19. Composition of the Hsa21 global interaction network. First, interaction data was combined from the Hsa21 Y2H and interolog screens, from manual literature and automated database searches to obtain a comprehensive list of 684 PPIs for 108 Hsa21 proteins (red sphere). 34 PPIs connect 35 proteins among Hsa21, while the remainder of 650 PPIs connects 105 proteins of Hsa21 with 547 direct interactors (shell 1, orange sphere). Then, a list of all human PPIs retrieved from the HPRD and BIND databases and the MDC and CCSB large-scale Y2H screens was screened electronically for PPI connections of shell 1, leading to the indirect interactors (shell 2, green sphere) and distant interactors (shell 3, blue sphere). Arrows inside shells describe the number of PPIs among proteins in one shell (intra-shell connections). Arrows between different shells denote the numbers of PPIs between them (inter-shell connections). All interaction data, i.e. 14,982 PPIs among 5,660 human proteins, for Hsa21 proteins and their direct, indirect and distant interactors can be accessed using the 21ppi database (http://chr21.molgen.mpg.de/21ppi).

Topology of the Hsa21 interaction network

The topology of the entire Hsa21 protein interaction network was analyzed by determining the frequency distribution of the links (i.e. interactions). The average number of connections per node (i.e. protein) is 2.64, where 758 proteins (13%) have 10 or more connections, while 2,659 proteins (47%) have only one or two connections. As expected for a scale-free and highly clustered interaction network resulting from incomplete sampling (Stumpf *et al.* 2005; Friedel and Zimmer 2006), the linear fit in a log-log plot of the connectivity distribution is shaped by a small fraction of highly connected nodes, following the power law $P(k) \sim k^{-\lambda}$ with P(k) as the frequency that nodes have k connections, and with $\lambda=1.95$. The mean path length between two proteins is 4.65.

Identification of hub proteins

Several Hsa21 proteins exhibit features of cellular hubs, displaying a multitude of PPIs to different proteins. The top five Hsa21 hubs are the amyloid precursor protein APP (60 PPIs), nuclear receptor interacting protein NRIP1 (41), the integrin beta chain ITGB2 (36), runt-related transcription factor RUNX1 (24) and calcium binding protein S100B (22). These proteins are linked to a broad variety of biological processes, ranging from transcriptional regulation, signal transduction and calcium signaling to roles in immune defense, response to oxidative stress and others.

3.6.2 Functional analysis of interaction subnetworks

For the identification of larger functional networks in the Hsa21-centered interaction data, Gene Ontology (GO) associations were used as functional classifiers. About 40% of all human genes are annotated in the GO database concerning molecular function, biological process or cellular component. In contrast to randomly chosen protein pairs sharing only ca. 14% of GO terms, genuinely interacting proteins are enriched for similar GO terms (Lehner and Fraser 2004; Rual *et al.* 2005; Wu *et al.* 2006), a feature that was also observed here in the Hsa21 PubMed, interolog and Y2H data sets (Table 3-10).

Table 3-10. Enrichment of shared GO annotations in the Hsa21 interaction data set

Data set	Molecular function* (GO level 3)	Biological process* (GO level 4)	Cellular component* (GO level 4)
Hsa21 PubMed	57% (39/69)	77% (53/69)	68% (17/25)
Hsa21 interologs	33% (14/42)	55% (22/40)	75% (40/53)
Hsa21+MDC screen^{+}	37% (10/27)	28% (8/29)	50% (8/16)
CCSB Y2H screen^{+}	25%	23%	49%

*Percentages of shared GO terms were calculated as indicated in brackets, namely as the number of pairs with common GO terms divided by the number of pairs where both partners have a GO annotations at the respective GO level.
$^{+}$Shared GO terms were calculated for the 56 combined PPIs derived from the Hsa21 Y2H screen and the MDC screen (Stelzl *et al.* 2005). For comparison, shared GO terms were also calculated for the interaction data derived from the CCSB Y2H screen (Rual *et al.* 2005).

Using GO annotations for analysis of the Hsa21 interaction network, thirteen transcription factors (TFs) encoded by Hsa21 could be linked via 138 PPIs with 119 proteins in shell 1 (Fig. 3-20). 78 of these interactors (66%) are transcription factors themselves, while the remainder (41 proteins, 33%) has a different or no functional

annotation. The TF network analysis reveals a compelling interconnection of eight Hsa21 TFs in a large network via ten other TFs. In this, the hub transcription factor NRIP1 forms 41 direct connections to other proteins, 26 of which also being involved in transcriptional regulation (Figure 3-20, upper part). Moreover, a smaller network connects the three Hsa21 TFs PKNOX1, OLIG1 and OLIG2, and two isolated PPI sets involve the transcription factors HMGN1 and SIM2.

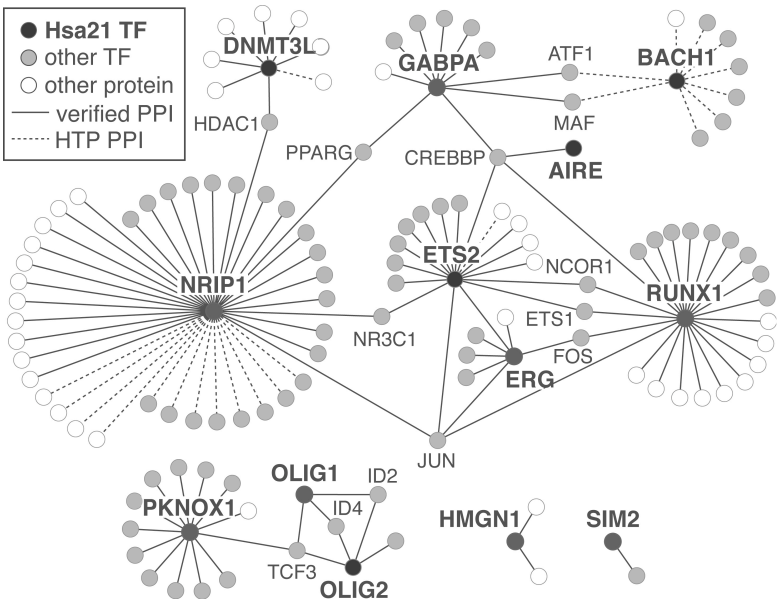

Figure 3-20. The transcription factors (TFs) of Hsa21 are highly interconnected via other transcriptional regulators. Thirteen Hsa21 TFs (black nodes) are connected to 119 proteins by 138 PPIs. 78 of the interactors (66%) are themselves transcription factors (grey nodes), while the remainder (41 proteins, 34%) has a different or no functional annotation (white nodes). The PPIs shown here are derived from sources of different reliability: (straight lines) 83% of the PPIs were verified using individual biochemical assays; (dotted lines) 17% of the PPIs were found in high-throughput (HTP) screens using Y2H assays or binding on microarrays. Four PPI networks for Hsa21 TFs can be derived from the combined interaction data: (Upper part) A large network involves eight Hsa21 TFs that are interconnected via ten other TFs; (Lower part, left) A smaller network connects PKNOX1, OLIG1 and OLIG2; (Lower part, right) Two currently isolated PPI sets involve the Hsa21 transcription factors HMGN1 and SIM2.

GO annotation of proteins, even when available, oftentimes does not reveal details about the role of the proteins in biological networks. Therefore, the next step was to associate functional attributes by inference through combining more specific information of the biological role of proteins in cellular processes.

3.6.3 Prediction of new protein functions from interaction networks

Most proteins are associated to more than one single biological process or signaling pathway, and these can often be identified through analysis of protein-protein interaction data. Based on the novel PPIs identified here by Y2H and verified by other assays, several examples are discussed below for functional association by inference.

For this, all available PPI data was used to integrate new Hsa21 interactions into known biological processes. Of special interest were molecular functions and biological processes shared by several proteins, and special emphasis was put on new three- and four-protein-interaction loop motifs indicative of potential functional modules (Goldberg and Roth 2003; Yeger-Lotem *et al.* 2004). Since many proteins are associated to more than one single molecular function or biological process, this kind of annotation is often not trivial. On basis of five examples, it will be shown that the extended Hsa21 interaction network described above can be used to infer new functional roles for Hsa21 proteins.

Pivotal role of NRIP1 in transcription regulation via acetylation/ deacetylation

NRIP1, the nuclear receptor interacting protein 1, is a hormone-recruited transcriptional cofactor. The regulation of NRIP1 activity by estrogens (Thenot *et al.* 1999) and retinoids (Kerley *et al.* 2001) confers to NRIP1 an important regulatory role in hormone signaling. For example, NRIP1 controls the expression of many genes involved in lipid and carbohydrate metabolism (Parker *et al.* 2006). NRIP1 acts mainly as a repressor of nuclear receptor activity (Augereau *et al.* 2006), but transactivation effects of NRIP1 have also been reported (Castet *et al.* 2006). NRIP1 has the capacity to directly interact with the estrogen receptor ESR1 (Cavailles *et al.* 1995) and twenty other nuclear hormone receptors (Figure 3-21), indicating a role of NRIP1 as hub protein in nuclear hormone signaling. In this, the negative regulatory function of NRIP1 involves the interaction with chromatin remodeling complexes through its recruitment to target genes and direct binding of histone deacetylases, such as HDAC1 (Wei *et al.* 2000).

Here, a novel interaction of NRIP1 was discovered, establishing two new interaction loops associated with transcriptional regulation (Figure 3-21). NRIP1 was found in

the Hsa21 Y2H to bind to HTATIP (HIV-1 Tat interacting protein, alias TIP60), a histone acetyltransferase with multiple activities involved in cellular signaling, cell cycle and checkpoint control. This interaction was confirmed by overexpression of both proteins in COS-1 cells, showing a colocalization in discrete speckles at the nuclear periphery (Figure 3-15C). HTATIP does not only acetylate histones, but also participates in nuclear receptor signaling, where it competes with histone deacetylase HDAC1 in the acetylation/deacetylation of the androgen receptor (Gaughan *et al.* 2002), establishing a new three-protein interaction loop (Figure 3-21). In line with this result, the dual actions of HTATIP and HDAC1 can also involve the proliferative transcription factor STAT3, which regulates gene expression in cell cycle progression (Xiao *et al.* 2003; Yuan *et al.* 2005).

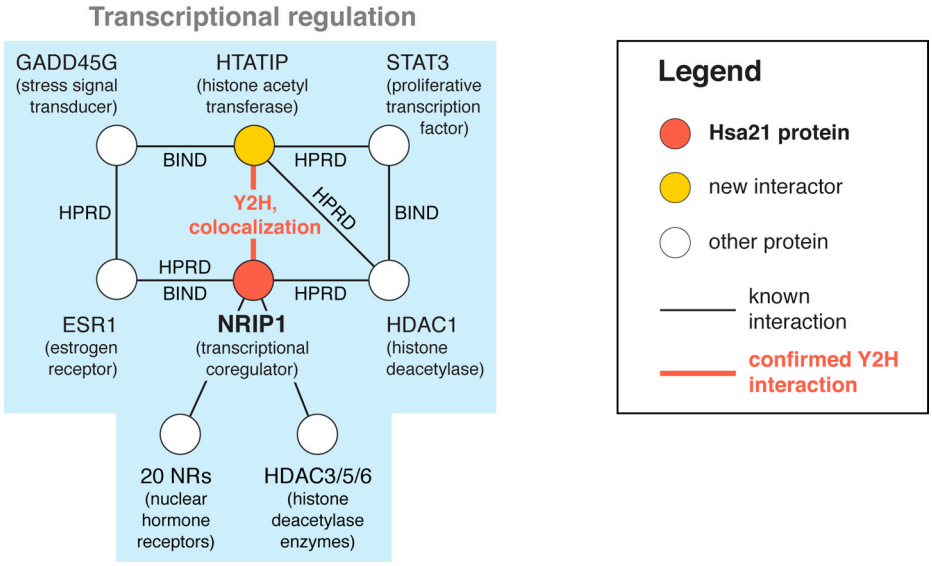

Figure 3-21. Integration of the new interaction identified for NRIP1 into transcriptional regulation. The new Y2H PPI confirmed by cellular colocalization was linked to known interaction data derived from database retrieval (HPRD and BIND). The biological process shared by several proteins is highlighted (blue area), and the respective process (transcription) is indicated. Proteins are named according to HUGO nomenclature, with short descriptions of protein function added in brackets. Nodes and connections are colored as indicated in the legend box. The interaction between NRIP1 and HTATIP, confirmed by nuclear colocalization in COS-1, closes two interaction loops associated with transcriptional regulation.

Moreover, HTATIP can bind to GADD45G (Goehler *et al.* 2004), a coactivator of nuclear hormone receptors implicated in terminal differentiation, growth suppression, and apoptosis. GADD45G, in turn, can interact with the estrogen receptor ESR1 (Yi *et al.* 2000), as does NRIP1, closing a second interaction loop associated with transcriptional regulation (Figure 3-21). Taken together, this data suggests an influence of NRIP1 on both acetylation and deacetylation of histones, nuclear

receptors and possibly other transcription factors related to hormone and stress signaling as well as cell cycle control and apoptosis.

UBE2G2 can be linked to central nervous system development and mental retardation

The ubiquitin-conjugating enzyme UBE2G2 is linked to a number of signaling pathways through its cooperation with CBL (Figure 3-22), a receptor ubiquitin ligase transferring activated ubiquitin from E2 to the substrate. CBL catalyzes receptor protein-tyrosine kinase ubiquitination and terminates signaling transduction by marking active receptors for degradation (Dikic and Schmidt 2007). UBE2G2 can also bind to PARC, a Parkin-like ubiquitin ligase, and to the ubiquitin-protein ligase UBE3A (Huang *et al.* 1999), whose loss of expression is associated to Angelman syndrome, a severe neurogenetic disorder (Kishino *et al.* 1997).

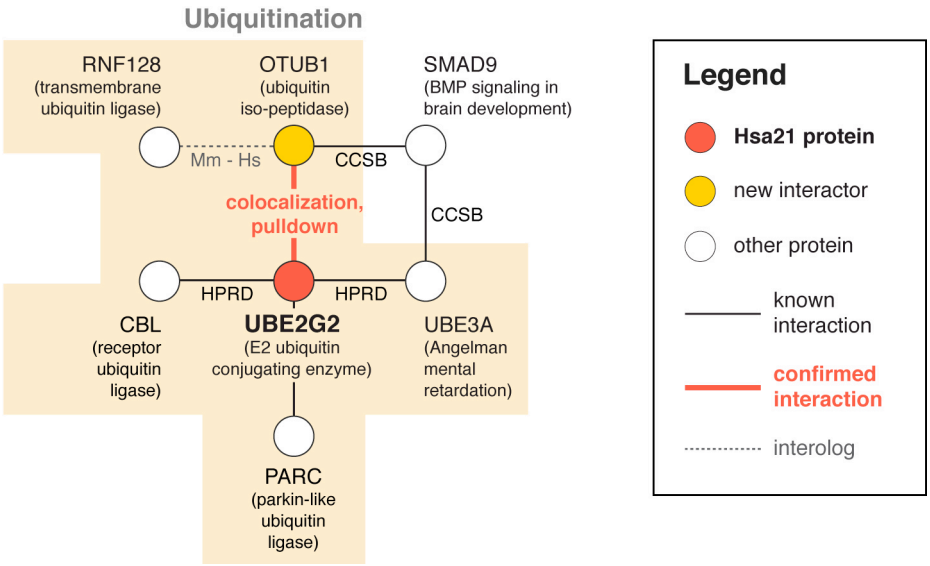

Figure 3-22. Integration of the new interaction identified for UBE2G2 into ubiquitination processes. The new interolog PPI confirmed by cellular colocalization and pull-down assays was linked to known interaction data derived from database retrieval (HPRD), a previous large Y2H screen (CCSB) and an interolog (*Hs: Homo sapiens*; *Mm: Mus musculus*). The biological process shared by several proteins is highlighted (orange area), and the respective process (ubiquitination) is indicated. Proteins are named according to HUGO nomenclature, with short descriptions of protein function added in brackets. Nodes and connections are colored as indicated in the legend box. The interaction between UBE2G2 and OTUB1, confirmed by cytoplasmic colocalization in COS-1 and pull-down experiments, closes one interaction loop associated with ubiquitination.

A novel interaction of UBE2G2 was discovered by Hsa21 interolog testing, establishing a new interaction loop associated with ubiquitination, brain development and mental retardation (Figure 3-22). As described above, it was found that UBE2G2

can bind to the ubiquitin iso-peptidase otubain 1 (OTUB1). The interaction was indicated by cytoplasmic colocalization of both proteins in COS-1 cells (Figure 3-17B) and confirmed by in vitro pull-down assays (Figure 3-17C). OTUB1 has the capacity to cleave poly-ubiquitin chains, but not ubiquitin, from peptides, and may thus function as an editing protease controlling polyubiquitin chain growth on substrates (Balakirev *et al.* 2003). Two OTUB1 isoforms are epistatic regulators of the expression of E3 ubiquitin ligase RNF128 (alias GRAIL), regulating the resultant anergy phenotype in T cells (Soares *et al.* 2004). OTUB1 can also interact with SMAD9, a transcription factor and effector of BMP signaling involved in central nervous system development (Hester *et al.* 2005). As shown in Figure 3-22, the confirmed new interaction between UBE2G2 and OTUB1 establishes an interaction loop, suggesting that UBE2G2 may be involved in regulating SMAD9- and UBE3A-dependent processes in brain development. In this, signaling events could be controlled via recruitment of two different components of the ubiquitination machinery, both ubiquitin ligases as well as editing proteases.

PCP4 is likely to play a role in calcium-dependent regulation of p53 transcriptional activity

The Purkinje cell protein PCP4 (alias PEP-19) is a neuronal polypeptide abundantly expressed in cerebellar Purkinje cells (Cabin *et al.* 1996) and involved in calcium signaling. PCP4 has the ability to interact with calmodulin (CALM1), a ubiquitous EF-hand calcium-binding protein, and to modulate its calcium binding properties (Putkey *et al.* 2003). As shown in Figure 3-23, the Hsa21 Y2H screen resulted in two novel interactors of PCP4. One of them is TP53 (tumor suppressor protein p53), a transcription factor functioning as key survival and apoptosis checkpoint which is found mutated in a large fraction of cancer cases (Chipuk and Green 2006).

In addition, as described above, an interaction between PCP4 and C1orf103 (alias RIF-1) was found. C1orf103 is a novel nuclear receptor corepressor that associates with the nuclear matrix (Li *et al.* 2007). Interestingly, both C1orf103 and TP53 have been independently reported to interact with the transcriptional corepressor NOC2L (alias NIR), an inhibitor of histone acetyltransferases exerting a negative control on the transcriptional activity of TP53 (Hublitz *et al.* 2005; Rual *et al.* 2005). This connection establishes a new interaction loop associated with transcriptional regulation (Figure 3-23). As described above, it was shown that overexpressed PCP4 and TP53 colocalize in COS-1 nuclei in a pattern excluding the nucleoli (Figure 3-

15B), which is reminiscent of that seen for NOC2L and TP53 (Hublitz *et al.* 2005), a finding that further confirms the interaction of PCP4 with TP53.

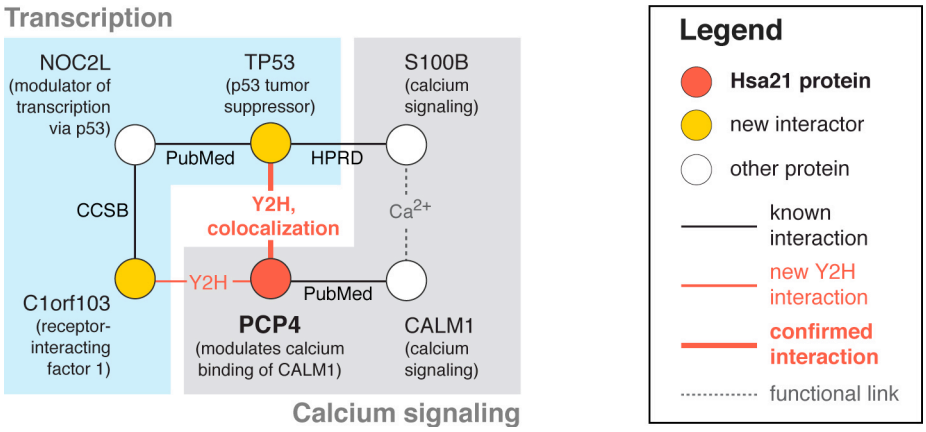

Figure 3-23. Integration of the new interactions identified for PCP4 into calcium signaling and transcriptional regulation processes. The new Y2H PPIs confirmed by cellular colocalization and pull-down assays was linked to known interaction data derived from database retrieval (PubMed, HPRD) and a previous large Y2H screen (CCSB). Biological processes shared by several proteins are highlighted (colored areas), and the respective processes are indicated. Proteins are named according to HUGO nomenclature, with short descriptions of protein function added in brackets. Nodes and connections are colored as indicated in the legend box. The interactions of PCP4 with TP53, confirmed by colocalization in nucleus and cell periphery of COS-1 cells, and with C1orf103 close one interaction loop associated with transcription and a functional loop associated with calcium signaling.

The DNA binding and transcriptional activity of TP53 can also be negatively regulated by another member of the EF-hand superfamily, the calcium-binding S100B protein, which can inhibit TP53 tetramer formation and phosphorylation mediated by PKC (Lin *et al.* 2001). This link establishes a functional loop associated with calcium signaling (Figure 3-23). TP53 has been shown before to bind to cytoplasmic and nuclear F-actin in the presence of free calcium ions (Okorokov *et al.* 2002). It could therefore be possible that PCP4, which can alter the calcium-binding properties of calmodulin by its IQ motif, might regulate the binding of TP53 to actin or its activity in a calcium-dependent manner. However, a preliminary analysis showed no colocalization of PCP4 with rhodamine-phalloidin, a marker for the actin cytoskeleton (data not shown).

Taken together, this data strongly suggests that PCP4 may play a role in the modulation of TP53 transcriptional activity, very likely coupled to the level of the intracellular second messenger calcium, which might be reflected by the colocalization of PCP4 and TP53 not only in the nucleus, but also in the cell periphery (Figure 3-15B).

MCM3AP as a link between DNA replication, activator-dependent transcription and neuronal outgrowth

Minichromosome maintenance (MCM) proteins are essential components of pre-replication complexes, which limit DNA replication to once per cell cycle (Tye 1999). The MCM3 acetylating protein (MCM3AP) encoded on Hsa21 was shown to be a potent natural inhibitor of the initiation of DNA replication and cell cycle progression, mediated through acetylation of MCM3 (Takei *et al.* 2002). As shown in Figure 3-24, another reported interactor of MCM3AP is the glucocorticoid receptor (GR) NR3C1. Interaction of the latter with MCM3AP was shown to enhance GR-regulated cell proliferation (Osman *et al.* 2006), a finding that confers to MCM3AP an interesting dual role as both activator and inhibitor of DNA replication.

In the Hsa21 Y2H screen described above, four novel interactors of MCM3AP were identified. While one novel interaction with a putative transmembrane protein called MGC13379 remains without functional connection for the time being (not shown), another new interaction connects MCM3AP to DNA replication, as supported by the known interaction with MCM3 (Figure 3-24). NUDT5, an antitumor candidate member of the Nudix hydrolase family, hydrolyzes 8-oxodGTP to a monophosphate, thus preventing mutations caused by the misincorporation of 8-oxoguanine into the DNA of replicating human cells (Ishibashi *et al.* 2003). The interaction of MCM3AP and NUDT5 might prevent initiation of DNA replication in a cellular state with high concentrations of oxidized guanine nucleotides.

Another novel interaction (MCM3AP with ENY2) was confirmed by pull-down experiments (Figure 3-16B). The Drosophila enhancer of yellow 2, ortholog of ENY2, has previously been shown to participate in transcriptional activation via direct interaction with TAF(II)40, a TATA binding protein-associated factor which is part of the activator-dependent transcription complex TFIID (Georgieva *et al.* 2001). The TFIID complex, in turn, has been reported to bind NR3C1 (Ford *et al.* 1997), which establishes an interaction loop associated with activator-dependent transcription (Figure 3-24).

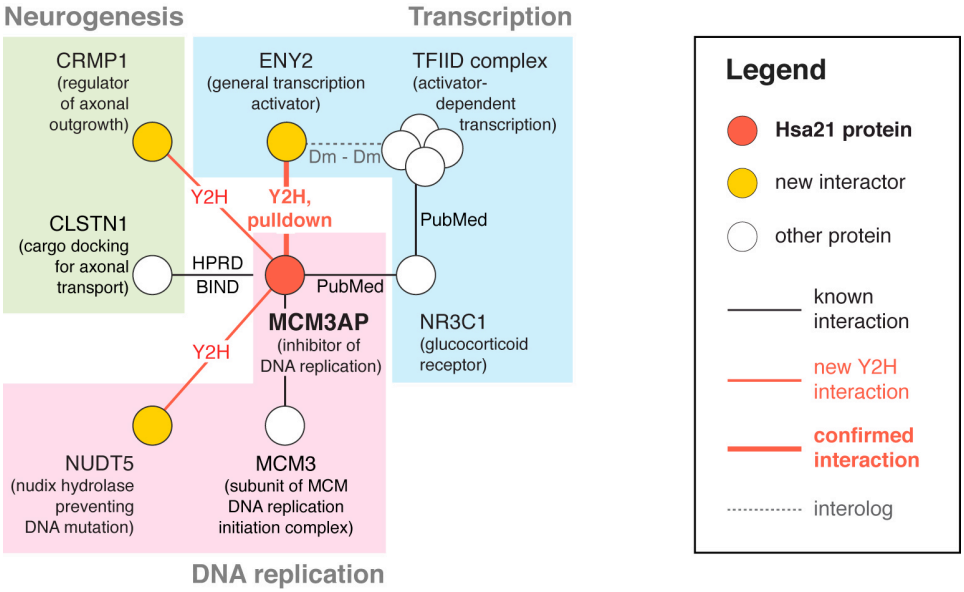

Figure 3-24. Integration of the new interactions identified for MCM3AP into DNA replication, transcription and neurogenesis. The new Y2H PPIs, one of them confirmed by pull-down assays, were linked to known interaction data derived from database retrieval (PubMed, HPRD, BIND) and an interolog from fly (*Dm*: *Drosophila melanogaster*). Biological processes shared by several proteins are highlighted (colored areas), and the respective processes are indicated. Proteins are named according to HUGO nomenclature, with short descriptions of protein function added in brackets. Nodes and connections are colored as indicated in the legend box. The interaction between MCM3AP and ENY2, confirmed by pull-down experiments, closes one interaction loop associated with transcription. Two more Y2H interactors of MCM3AP involved in DNA replication and neurogenesis are supported by known interactors linked to these processes, while one unknown new interactor remains anonymous (not shown).

A fourth novel interaction connects MCM3AP with CRMP1 (collapsin response mediator protein 1), a regulator of axonal outgrowth involved in the mediation of semaphorin-induced actin depolymerisation and growth cone collapse in axons (Goshima *et al.* 1995). This finding is supported by the known interaction of MCM3AP with the axonal protein calsyntenin 1 (CLSTN1), a postsynaptic transmembrane protein involved in organelle transport along axons (Konecna *et al.* 2006). Interestingly, developing CRMP1(-/-) mice show a decrease in granule cell proliferation and apoptosis in external granule cell layers (Charrier *et al.* 2006), features reminiscent of the Down syndrome phenotype (Olson *et al.* 2004).

Taken together, the new interactions of MCM3AP identified here place this Hsa21-encoded protein at a unique position between the initiation of DNA replication, cell proliferation via glucocorticoid signaling, axonal transport and axonal outgrowth coupled to cerebellar development. Based on these findings, future studies might elucidate a potential contribution of MCM3AP to the pathogenesis of cerebellar defects in Down syndrome.

C21orf127 is involved in protein synthesis termination and metabolism of proteins and transcripts

As last example for inferring novel functional connections from the Hsa21 interaction network, those PPIs were extracted that connect the Hsa21-encoded protein C21orf127, via several new interaction loops, to a number of proteins regulating transcription, translational termination, mRNA decay and mRNA transport along microtubules. In this, six new PPIs are involved (Figure 3-25) that were identified by the Hsa21 Y2H screen and interolog analysis described above.

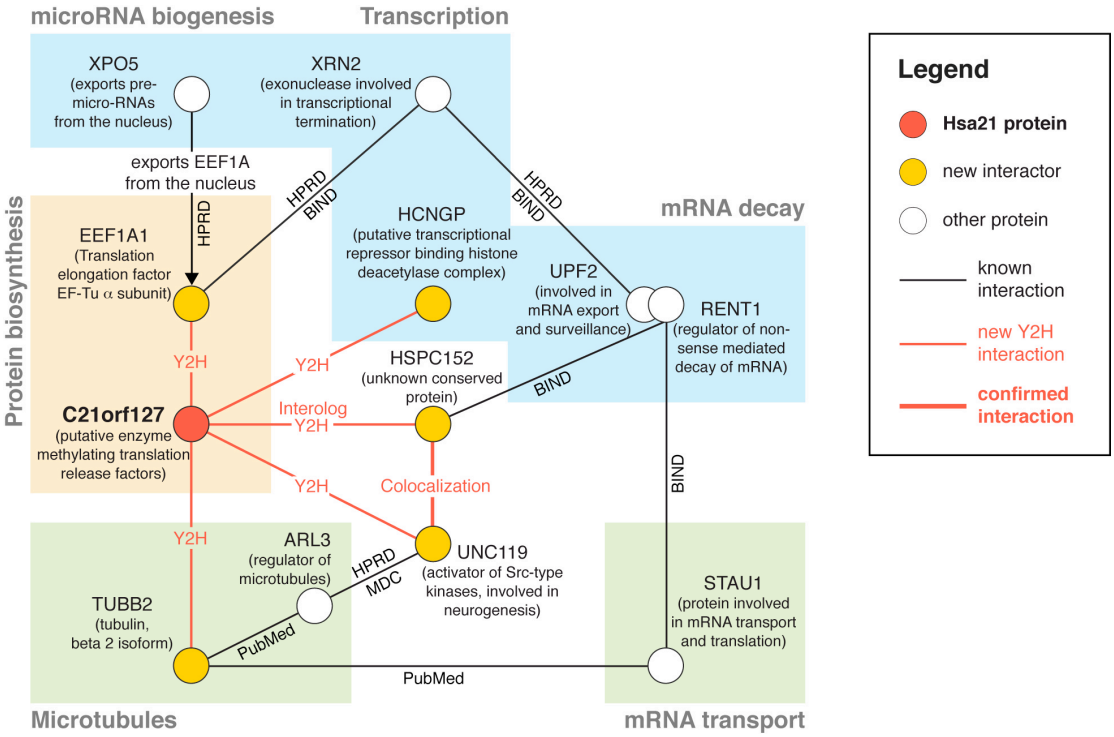

Figure 3-25. Integration of the new interactions identified for C21orf127 into protein biosynthesis, transcription and mRNA transport and metabolism processes. The new PPIs, one of them supported by an interolog and one by nuclear colocalization, were linked to known interaction data derived from database retrieval (PubMed, HPRD, BIND) and a previous large-scale Y2H screen (MDC). Biological processes shared by several proteins are highlighted (colored areas), and the respective processes are indicated. Proteins are named according to HUGO nomenclature, with short descriptions of protein function added in brackets. Nodes and connections are colored as indicated in the legend box. An extended interaction network is established by new Y2H interactors of C21orf127, supported by nuclear colocalization of two interactors in the nucleus of COS-1 cells. Several interaction loops can be observed, placing C21orf127 between proteins regulating translation, transcription, mRNA export from the nucleus, nonsense-mediated mRNA decay and mRNA transport along microtubules.

Although C21orf127 is officially called N6AMT1 for putative N-6 adenine-specific DNA methyltransferase 1, no evidence for an adenine-methyltransferase activity could be detected after recombinant protein expression in the nucleus of C2C12 cells

(Ratel *et al.* 2006). Instead, sequence alignments show that the yeast ortholog of C21orf127 is YDR140w, a methyltransferase whose substrate is the release factor ternary complex associated with termination of protein synthesis (Heurgue-Hamard *et al.* 2005). In line with this, a new interactor of C21orf127 was found that is also involved in protein biosynthesis, namely the eukaryotic translation elongation factor 1 alpha1 subunit (EEF1A1), a part of the elongation factor-1 complex responsible for delivery of aminoacyl tRNAs to the ribosome (Mansilla *et al.* 2002). The EEF1A isoforms are interaction hubs with 76 currently known interactions, and among the most prominent multifunctional proteins (Lamberti *et al.* 2004). Besides their role in protein biosynthesis, they have been implicated in several cellular processes, including embryogenesis, senescence, oncogenic transformation, cell proliferation and organization of the cytoskeleton.

As shown in Figure 3-25 (upper part), C21orf127 is connected via EEF1A1 to exportin 5 (XPO5), which exports pre-microRNAs and also the EEF1A proteins from the nucleus, and to an exoribonuclease (XRN2) involved in transcriptional termination. C21orf127 itself is linked to transcription by two other new interactors, namely the putative transcriptional repressor HCNGP (Li *et al.* 2004a) and, via the unknown conserved protein HSPC152, to the regulator of nonsense transcripts RENT1 (alias UPF1) and the associated UPF2 protein involved in mRNA nuclear export and mRNA surveillance. Interestingly, the latter was shown to bind to XRN2 (Lehner and Sanderson 2004), establishing an interaction loop associated with biogenesis and metabolism of transcripts. A nuclear function of C21orf127 can also be inferred from the new interactions identified here with UNC119 and HSPC152, for which colocalization was observed in the nucleus of COS-1 cells (Figure 3-17A), establishing a triangle-shaped interaction loop (Figure 3-25).

Finally, C21orf127 is linked to mRNA transport and nonsense-mediated mRNA decay (Figure 3-25, lower part). Interestingly, NMD complexes, with RENT1 as component, are also involved in the termination of translation, as suspected for C21orf127. In this, STAU1, the staufen microtubule-binding mRNA transport protein, binds directly to RENT1 and elicits mRNA decay when tethered downstream of a termination codon (Kim *et al.* 2005). Intriguingly, C21orf127 bound to tubulin beta 2 (TUBB2) in the Hsa21 Y2H screen, and thus potentially also to microtubules, as does STAU1. Further support comes from the interaction loop formed by C21orf127 and its Y2H interactor UNC119 (alias HRG4, a functionally conserved and essential neural

protein) via the regulator of microtubules ARL3 (Van Valkenburgh *et al.* 2001) to TUBB2 (Zhou *et al.* 2006).

Taken together, data suggests that the function of C21orf127 alias HEMK2 is at the interface between transcription, translation, RNA processing and RNA transport. However, the temporal and special network dynamic of these interactions are not known, and it is likely that they occur neither at the same time nor in the same cellular context, as shown by different cellular localizations of the components. Notwithstanding, since many of the components described here are highly conserved among many organisms, it will be interesting to find out more about the functional interplay in this potentially ancient interaction network.

3.6.4 Retrieval of pathways with involvement of Hsa21 proteins

On the level of cellular signaling networks, the state of each cell is integrated from the combination of numerous protein interactions. The entire Hsa21 ORF data set (206 ORFs) and the PPI data set (108 Hsa21 proteins with connections to 547 direct interactors) was used for systematic identification of connections between Hsa21 proteins and signal transduction pathways. To find these connections, three major pathway databases were interrogated, namely TransPath, KEGG and STKE (Gough 2002; Kanehisa *et al.* 2006; Krull *et al.* 2006). Of the Hsa21 proteins themselves, only five gene products (IL10RB, IFNAR2, IFNGR2, IFNAR1 and APP), which are all membrane-bound proteins, were directly annotated as parts of signaling cascades, namely the interleukin (IL), interferon (IFN) and c-Jun N-terminal kinase (JNK) pathways. All other pathway associations were established via pathway annotation of direct interactors of Hsa21 proteins.

The TransPath database was found to contain the highest number of pathway connections for Hsa21 proteins (35 Hsa21 proteins connected via 59 interactors to nine pathways), while KEGG and STKE resulted in significantly lower number of connections for these proteins (25 Hsa21 proteins via 36 interactors to six pathways and 16 Hsa21 proteins via 17 interactors to five pathways, respectively).

All annotations were retrieved from the TransPath database using the ArrayAnalyzer tool provided by TransPath, which is designed to find pathway associations (hits) for lists of Entrez GeneIDs and to calculate p-values for the probabilities that the

observed number of hits in each pathway occur by random chance. During this analysis, the focus was not entirely on low p-values, since all pathway connections should be retrieved, but pathways showing an interesting number of connections to Hsa21 proteins were mostly associated with low p-values, depending on the individual pathway sizes.

Figure 3-26 shows a color-coded matrix for the nine TransPath pathways with connection to Hsa21 proteins and their direct interactors. For instance, fifteen Hsa21 proteins are indirectly associated to the epidermal growth factor (EGF) pathway via 21 PPIs to proteins in this pathway. Based on the new PPIs identified in the Hsa21 Y2H screens, six novel connections could be established for five Hsa21 proteins: PPIAL3, C21orf91 and NRIP1 with the androgen receptor (AR) pathway, PCP4 with the p38 MAP kinase pathway, and TMPRSS3, which was previously not associated to any known pathway, can now be linked to both the c-Jun N-terminal kinas (JNK) and p38 MAP kinase pathways.

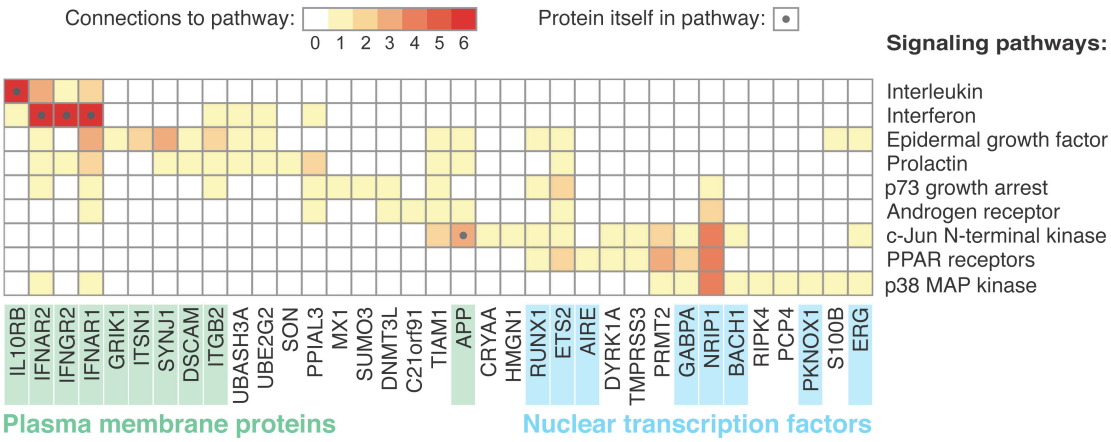

Figure 3-26. Hsa21 proteins and direct interactors are connected to nine major signal transduction pathways. Included in this connection matrix are six new connections arising from the PPIs identified in the Hsa21 Hsa21 Y2H screens, three of which were verified by cellular colocalization or pull-down experiments. The comprehensive Hsa21 PPI data set was used to query cellular pathways annotated in the TransPath database for all Hsa21 proteins and their 547 direct interactors. A connection matrix was established to show the number of links to pathways for each Hsa21 protein in a color code from yellow to red. In total, 35 Hsa21 proteins (gene symbols below connection matrix) are connected to nine major signal transduction pathways (names on right hand side). Five Hsa21 proteins are themselves annotated as components of interferon, interleukin and c-Jun N-terminal kinase signaling cascades (marked with dots in the respective cells). All other connections are established via 136 PPIs of Hsa21 proteins with 59 proteins in pathways. Cell colors indicate the total number of connections for each Hsa21 protein. Proteins were sorted according to their pattern of pathway connections, revealing that membrane proteins (green) are connected to a different set of pathways than nuclear transcription factors (blue).

Using the Hsa21 PPI data set, connections were established between a total of 35 Hsa21 proteins via 136 PPIs to 59 pathway-annotated interactors. All pathway

connection data can be found in Supplemental Table S4 (see http://chr21.molgen.mpg.de/21ppi/S4_pathways.xls).

Hsa21 pathway connection data, as presented here, might enlighten our understanding of protein function and of pathway perturbations due to trisomy 21. Future studies will be directed towards candidate Hsa21 proteins whose overexpression might trigger perturbations in signaling through direct PPIs.