

PROTEIN INTERACTION MAPPING:
USE OF OSPREY TO MAP SURVIVAL OF MOTOR NEURON
PROTEIN INTERACTIONS

by

Margaret Barnhart

An Internship presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Advisor: Dr. Ron Nieman

Technical report number: 04-19

ARIZONA STATE UNIVERSITY

December 2004

ABSTRACT

Spinal Muscular Atrophy is one of the leading genetic causes of death in infants. In humans, the disease state is characterized by homozygous deletion of the telomeric copy of the survival of motor neuron gene (*SMN1*). The centromeric copy, *SMN2*, rescues lethality by producing a small amount of full-length SMN protein as its minor product. The *SMN* gene was first characterized in 1995, and research efforts to describe the molecular mechanisms of SMN protein in the cell have since revealed a highly complex set of functions and interactions for SMN. The large amount of protein-protein interaction data collected for SMN exceeds the limitations imposed by current methods of interaction visualization. Osprey allows a network representation of protein-protein interactions and has been used to describe the recorded sets of interactions of SMN. This method of interaction visualization allows relationships to be drawn between the functions of SMN and analogous proteins, clustering of interactions based on level of interaction or function, and ultimately, the derivation of clues to the critical function of SMN.

TABLE OF CONTENTS

1. Introduction

1.1 Spinal muscular atrophy

1.1.1 [Clinical overview](#)

1.1.2 [Genetic basis of spinal muscular atrophy](#)

1.1.3 [Review of SMN protein functions](#)

1.1.3.1 [Pre-mRNA splicing](#)

1.1.3.2 [Transcription](#)

1.1.3.3 [Ribosome production](#)

1.1.3.4 [Axonal function](#)

1.1.4 [Large-scale visualization of protein interactions](#)

2. Osprey: a network visualization system

2.1 [Main features of Osprey](#)

2.2 Using Osprey to visualize SMN interactions

2.2.1 [Creation of input files](#)

2.2.1.1 [Types of acceptable input files](#)

2.2.1.2 [SMN input files](#)

2.2.1.3 [Superimposition of SMN:SMN files and SMN:GRID files](#)

2.2.1.4 [Gene and edge information](#)

2.2.1.5 [Color indices](#)

2.2.1.6 [Network and connectivity filters](#)

2.2.1.7 [Layouts](#)

3. [Contrasting methods of Visualization systems](#)
 - 3.1 “Cartoon” images
 - 3.2 Fluorescence and live cell imaging
 - 3.3 Osprey: advantages and disadvantages
4. Discussion
5. [Appendix A: bibliography for Osprey interaction datasets](#)
6. [Appendix B: Interaction data files](#)

1. Introduction

1.1 Spinal Muscular Atrophy

1.1.1 Clinical overview

Spinal muscular atrophy (SMA) is an autosomal recessive neurodegenerative disorder characterized by the degeneration of α -motor neurons in the anterior horn of the spinal cord and brainstem, leading to muscular atrophy and paralysis (Melki 1997).

Spinal muscular atrophy is the second most common genetic cause of infant mortality with a recorded frequency of 1 in 6,000 to 1 in 10,000 births (Pearn 1978). SMA exhibits a broad phenotypic range which has been subdivided into three categories. Type I SMA, previously described as Werdnig-Hoffman disease, is the most severe form (Pearn 1980). Infants with Type I exhibit onset of symptoms at birth or before 6 months and generally do not live beyond 2 years. Children born with the intermediate, Type II SMA are unable to stand or walk without aid. Type III SMA, previously described as Kugelberg-Welander disease, is characterized by later onset of nerve degeneration and slower disease progression (Pearn 1980).

1.1.2 The genetic basis of spinal muscular atrophy

The spinal muscular atrophy determining gene, termed survival of motor neurons gene (*SMN*) was characterized in 1995 through genetic analysis and physical mapping (Lefebvre, Burglen et al. 1995). The *SMN* gene is duplicated as an inverted repeat on chromosome 5 at 5q13 in humans. The telomeric copy, *SMN1*, produces the full-length transcript coding for the SMN protein. The centromeric copy, *SMN2*, is highly homologous to *SMN1*, with a limited number (9-14) of translationally-silent differences

in base-pairs (REF). *SMN2* produces a transcript which is mainly alternatively spliced resulting in the expression of a less stable, nonfunctional truncated SMN protein lacking the region coded for by exon 7. A minor product of *SMN2* is full-length, functional SMN protein. A single silent C → T nucleotide transition is the critical difference between the *SMN1* and *SMN2* genes, and is responsible for alternative splicing of *SMN2* (Lorson, Hahnen et al. 1999).

In patients with SMA, telomeric *SMN1* is homozygously mutated or deleted, resulting in non-functional SMN protein from *SMN1* (McAndrew, Parsons et al. 1997). *SMN2* cannot fully compensate for mutation or deletion of the *SMN1* gene since the *SMN2* gene product is mainly alternatively spliced into a less stable and nonfunctional truncated form. Hence, levels of full-length SMN protein in cells are severely reduced in patients with SMA. Fortunately, the small amount of full-length transcript produced by the *SMN2* gene is palliative, and a correlation between number of copies of *SMN2* and disease severity has been observed, with more copies resulting in a less severe SMA phenotype (Coover, Le et al. 1997; Lefebvre, Burlet et al. 1997; Lorson, Hahnen et al. 1999).

The specific role of SMN protein in motor neuron death is not fully understood. *SMN* has been shown to be an essential gene for several organisms, including human, mouse, chicken, *C. elegans*, and yeast (Paushkin, Charroux et al. 2000). The SMN protein is ubiquitously expressed in all tissues of metazoan organisms. In humans and all mammalian cells, SMN is found in both the cytoplasm and nucleus, where it is highly concentrated in nuclear structures termed Gemini or coiled bodies, or gems (Liu and Dreyfuss 1996). SMN concentrations vary widely in tissues as well, with the highest

concentrations occurring in brain and spinal tissue (Broccolini, Engel et al. 1999; Williams, Vinnakota et al. 1999).

1.1.3 Review of SMN protein functions

The SMN protein has been implicated as a major factor in several essential cell functions. These functions include pre-mRNA splicing, transcription, ribosome production, and axon-specific RNP transport. As an essential factor in these cell processes, SMN may prove to have critical functions in the cytoplasm, nucleoplasm, and axon termini.

1.1.3.1 Pre-mRNA splicing

In the late 1990's, the first conceptual picture of SMN function at the molecular level was formed. SnRNPs (for "small nuclear ribonucleoproteins") are RNA-protein complexes which are assembled in the cytoplasm and imported into the nucleus for subsequent assembly into the larger, macromolecular spliceosomes, which catalyze pre-mRNA splicing (Fischer, Liu et al. 1997). Generally, uridine-rich small nuclear RNAs (UsnRNAs) are transcribed by RNA Polymerase II, bound by the nuclear cap-binding complex and Phax (phosphorylated adaptor for protein export) and exported to the cytoplasm. In the cytoplasm the UsnRNAs then bind to the SMN complex which consists of SMN, SmB/B', SmD1, SmD3 as well as several other proteins to be described later in the text. The UsnRNAs complex with the Sm proteins, forming premature UsnRNPs. This premature complex is bound by snurportin, which interacts with importin β to transport the UsnRNPs back into the nucleus (Narayanan, Ospina et al. 2002).

1.1.3.2 Transcription

In 2001, Dreyfuss et al. showed a direct interaction between RNA helicase A and SMN (Pellizzoni, Charroux et al. 2001). RNA helicase A is a DEAH box RNA helicase which binds to RNA polymerase II (Pol II). Pol II assembles into a large complex prior to binding to gene promoter elements, and in this study SMN was proposed to have a possible function in the assembly of the Pol II complex. This idea was supported by the observation that SMN mutant proteins led to the reorganization of the Pol II ‘transcriptome’ complexes and inhibition of transcription *in vivo*. Furthermore, SMN has been found to interact directly with RNA helicase dp103 (Campbell, Hunter et al. 2000), and associate with the transcription corepressor mSin3A (Zou, Barahmand-Pour et al. 2004). Taken together these results strongly support an important role for SMN in transcriptional regulation.

1.1.3.3 Ribosome production

In the nucleolus, snoRNPs (small nucleolar RNPs) function in the cleavage and modification of ribosomal RNA (rRNA). SMN has been shown to interact with GAR1 (Pellizzoni, Baccon et al. 2001), a core component of the box C/D snoRNP, and fibrillarin, a core component of the box H/ACA snoRNP. From these results, Jones et al propose a parallel function to snRNP assembly in snoRNP assembly and thus suggest a critical role for SMN in ribosome production (Jones, Gorzynski et al. 2001).

1.1.3.4 Axon-specific function

SMN has been observed through immunocytochemical analyses to localize in neuronal processes (Battaglia, Princivalle et al. 1997), suggesting an axon-specific function. In SMA, a motor neuron axonopathy is observed and so understanding why full-length SMN is important to neurite axons, and motor neuron axons in particular, is at the core of the research field. In 2003, Zhang et al. directly observed SMN granules moving bidirectionally along the axons using live cell imaging of fluorescently labeled full-length SMN protein (Zhang, Pan et al. 2003). This most recent glimpse into SMN cellular function has yet to be refined, but supports implications of an axon-specific function.

1.1.4 Large Scale-visualization of protein interactions

In less than 10 years, SMA researchers have characterized the SMN gene and proposed four highly plausible or proven functions of the full-length SMN protein at the cellular and molecular levels. As each picture of functionality becomes more refined and increasingly complex and protein-protein interaction datasets grow rapidly, a simple need becomes apparent. Techniques for protein interaction visualization on the proteome scale must be developed in parallel with the massive datasets research is accumulating. In SMA research, as with other disease-related research, there are even more urgent reasons for such tools than those of experimental efficiency and depth of understanding of the proteome as whole.

Surprisingly, few network visualization/manipulation applications exist, and typically have limited data manipulation options. The existing protein interaction database tools are Osprey and BIND (Bader and Hogue 2000; Breitkreutz, Stark et al. 2003). A third, more general network visualization program called Pajek allows users to view data graphics but is not tailored to protein-protein interactions (Batagelj V 1998). Osprey is an application freely available for non-profit use created in the Samuel Lunenfeld research institute at Mt. Sinai Hospital in Toronto, Canada. From a separate lab at the Samuel Lunenfeld research institute, an additional interaction database was created, the Biomolecular Interaction Network Database (BIND). While BIND currently has roughly twice as many recorded interactions as Osprey, it has fewer visualization options. For this project, Osprey was chosen for its ease of data input, portability, and flexible graphical formatting which will hopefully lead to insights into the diverse functions of SMN.

2. Osprey: a network visualization system

2.1 Main features of Osprey

Osprey can be used on all platforms that support the latest Java Plugin, including Windows (95, 98, NT, ME, 2000, XP), Red Hat Linux (versions 7.3 or higher) and Mac OS X (version 10.2 or higher). Osprey was developed using the Sun Microsystems Java Standard Development kit (version 1.4.0_02), and can be used either as a stand-alone application or a viewer for online interaction databases such as the GRID (Breitkreutz, Stark et al. 2003). Imported files may be Osprey (.osp) or text files (.txt, see section 2.2.1.2), and files may be exported in a variety of formats, including JPEG, PNG, SVG, and matrix files (see section 2.2.1.4). An Osprey output file shows a network of protein-

protein interactions, where nodes represent proteins and edges (connecting the nodes) represent interactions (Figure 1).

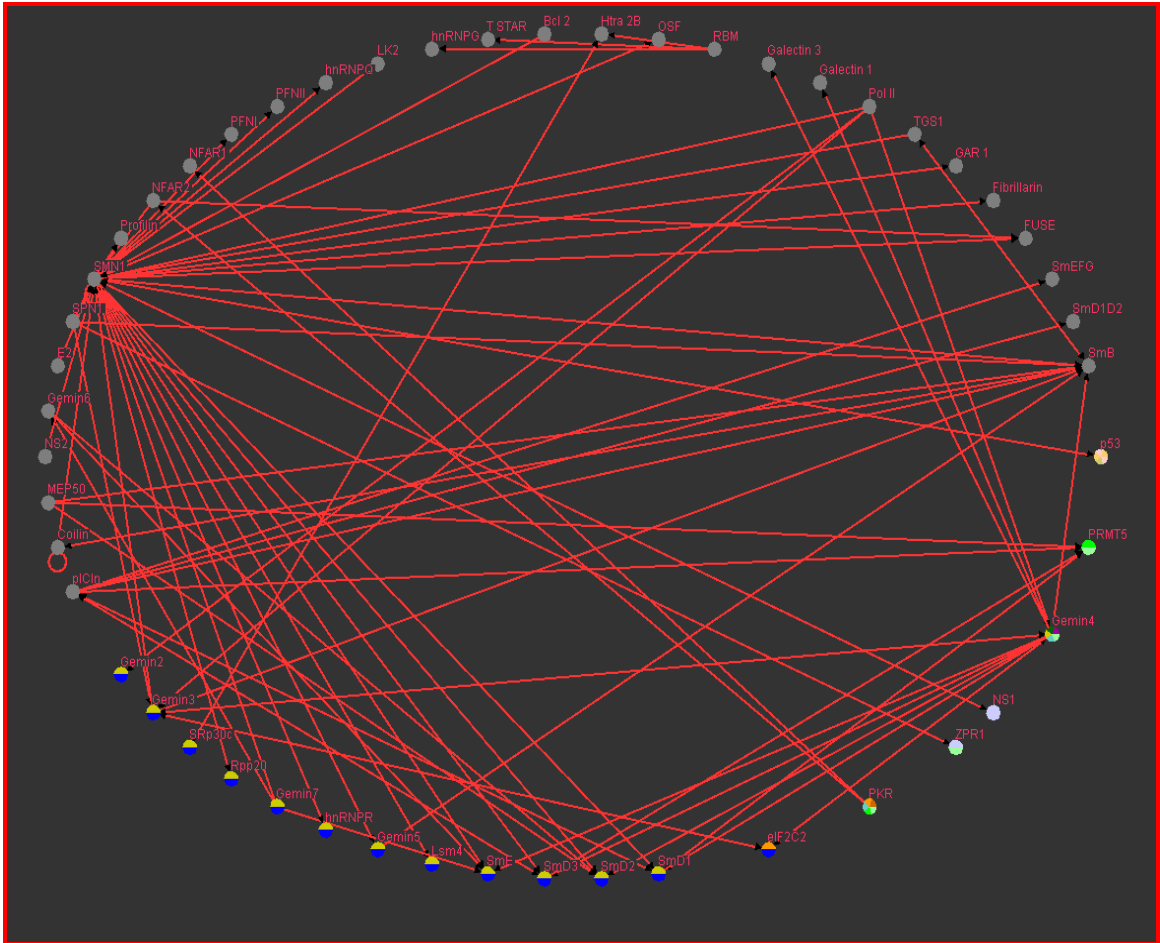


Figure 1: Example of an Osprey output file

Osprey can interactively download online protein interaction datasets from the GRID (General Repository for Interaction Datasets). The most recent release of Osprey (vers. 1.2.0) allows access to the Worm GRID (*C. elegans*), Fly GRID (*D. melanogaster*), Yeast GRID (*S. cerevisiae*), and Human GRID. When Osprey is initiated, the user is

prompted to select a GRID from which datasets will be downloaded. This selection can be permanently set, in this case, to the human GRID. When the user interface is loaded, files may be imported from the human GRID, from user-defined files, or both. One particularly useful aspect of Osprey is its ability to superimpose multiple datasets, either user:user, user:GRID, or GRID:GRID datasets (Figure 2). User:user dataset superimposition is useful when, as in the case of recorded SMN interactions, there are several different experimental methods used by several independent research groups. User:GRID superimposition will give more information to SMN research as more protein interactions are added to the Human GRID.

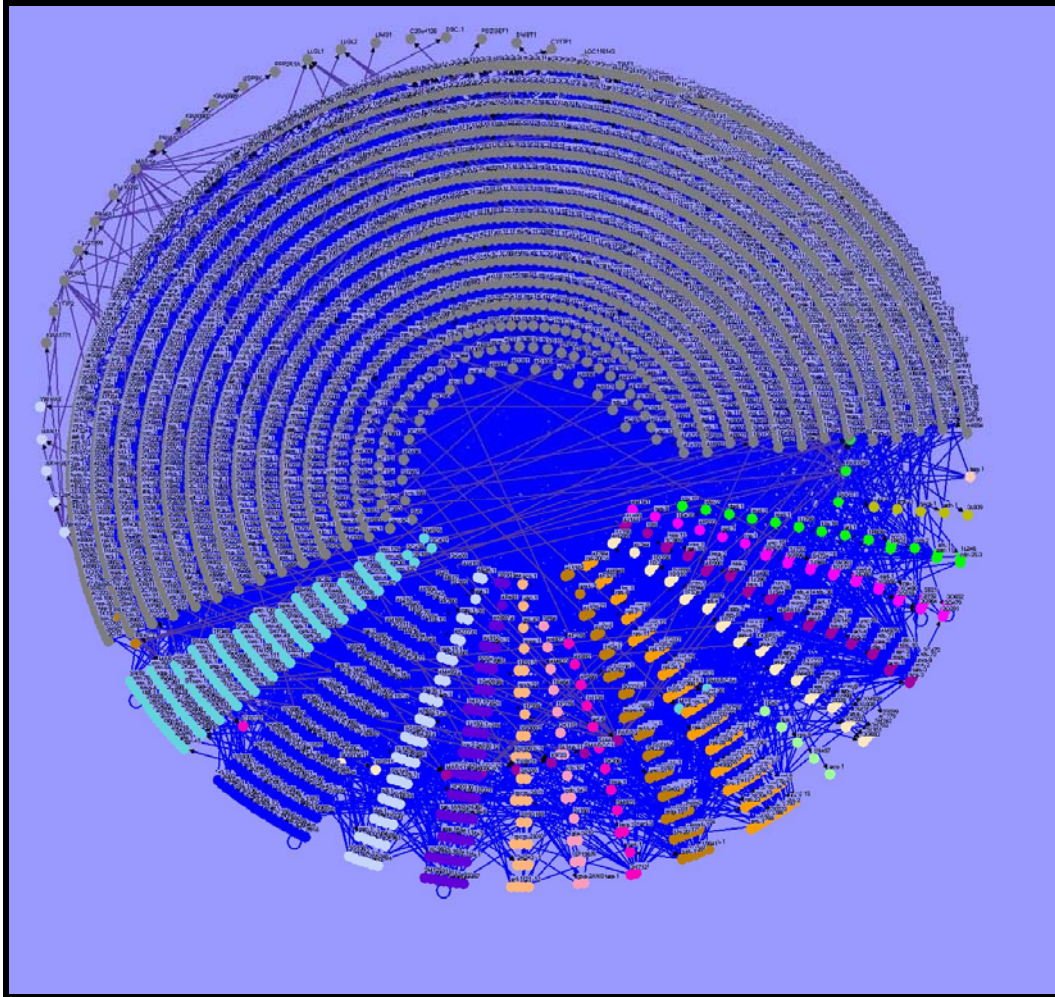


Figure 2: Superimposition of complete worm and human GRID datasets

2.2 Using Osprey to visualize SMN interactions

2.2.1 Creation of input files

2.2.1.1 Types of acceptable input files

In addition to accepting downloaded GRID datasets, Osprey accepts user-created input files. These files are simply tab delimited text files with each column representing a specific set of information. In custom file variation #1, the names of the genes coding

for proteins which interact are placed into two columns, where by convention the ‘Gene A’ in the first column interacts with ‘Gene B’ in the second column (Figure 3).

```
GeneA      GeneB  (*Entries on each line are single tab delimited)
SMN1      FBP
Htra 2B   hnRNPG
Htra 2B   RBM
HnRNPR    SMN1
HnRNPQ    SMN1
Gemin5    SMN1
Gemin5    Gemin2
Gemin5    Gemin3
Gemin5    Gemin4
SMN1      Gemin5
```

Figure 3: custom file variation #1

Custom file variation #2 permits the user to create two additional columns, GeneA screen name and GeneB screen name (Figure 4). Using this input file type Osprey will display gene ontology information about GeneA and GeneB according to the gene names in the first two columns, but the network will display the nodes with the specified alternate names. Custom file variation #3 allows the user to add information into three additional columns: “Experimental system,” “Source,” and “PubMedID” (Figure 5). Osprey will accept any string input into these columns, but given a correct PubMed ID will display a button which links directly to the paper(s) referenced on PubMed.

```
GeneA      GeneB      GeneA Screen Name  GeneB Screen Name
SMN1      FBP        A                  B
Htra 2B   hnRNPG    C                  D
Htra 2B   RBM       C                  E
HnRNPR    SMN1      F                  A
HnRNPQ    SMN1      G                  A
Gemin5    SMN1      H                  A
Gemin5    Gemin2    H                  I
```

Gemin5	Gemin3	H	J
Gemin5	Gemin4	H	K
SMN1	Gemin5	A	H

Figure 4: Custom file variation #2

The “Experimental System” column is intended to denote the type of experiment used to prove the interaction, for example, yeast two-hybrid or affinity precipitation. The “Source” column will list the authors of the paper next to the button which links to PubMed (Figure 6).

GeneA	GeneB	Experimental System	Source	PubMedID
SMN1	FBP	CoIP HEK293	Williams et al	10734235
Htra 2B	hnRNPG	CoIP HEK293	Hofmann et al	12165565
Htra 2B	RBM	CoIP HEK293	Hofmann et al	12165565
HnRNPR	SMN1	CoIP HEK293	Rossoll et al	11773003
HnRNPO	SMN1	CoIP HEK293	Rossoll et al	11773003
Gemin5	SMN1	CoIP HEK293	Gubitz et al	11714716
Gemin5	Gemin2	CoIP HEK293	Gubitz et al	11714716
Gemin5	Gemin3	CoIP HEK293	Gubitz et al	11714716
Gemin5	Gemin4	CoIP HEK293	Gubitz et al	11714716
SMN1	Gemin5	CoIP HEK293	Gubitz et al	11714716

Figure 5: Custom File variation #3

The last type of custom file accepted by Osprey is file variation #4 with two additional columns, 'GeneA screen name' and 'GeneB screen name,' which serve the same purpose as in custom file variation #2.

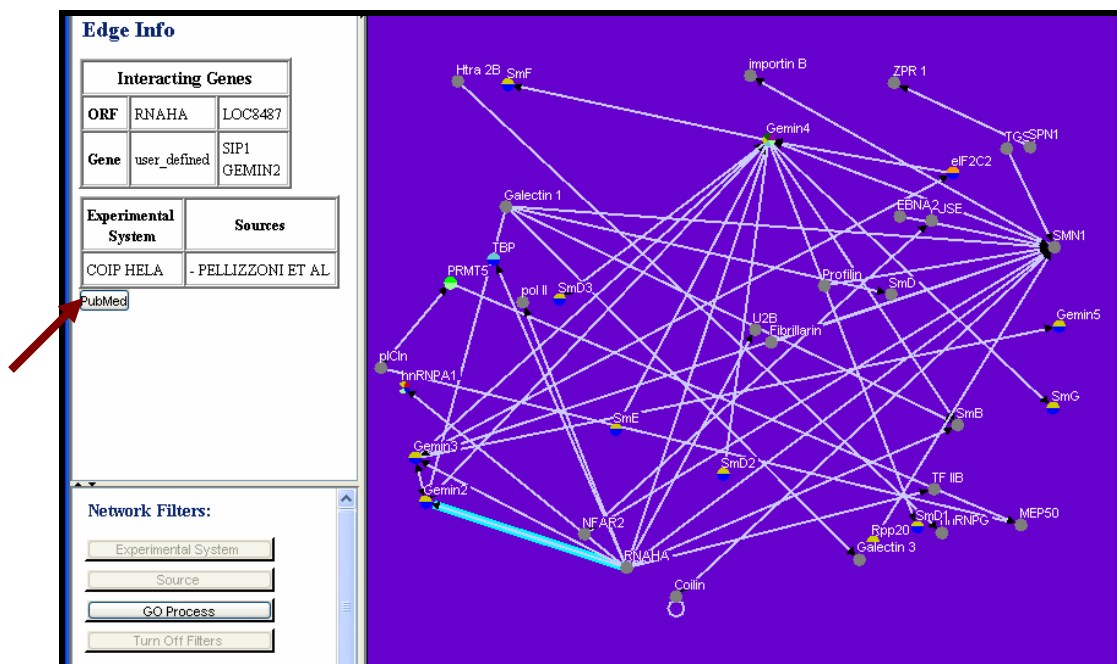


Figure 6: Link to PubMed reference

2.2.1.2 SMN input files

For the SMN input files, custom file type #3 was chosen as it allowed the most comprehensive data input. Custom file type #3 does not allow entry of alternate screen

names for genes, which requires that a dominant choice be made on data entry. This limitation can be reduced by incorporating alternate gene names in an overlay file or by use of a name conversion table when performing searches. However, custom file type #3 results in a single point representation of all proteins in the interaction network, and minimizes ambiguity in complex interactions.

Interactions were determined in a multi-step process. First, a search of NCBI/PubMed for papers containing the terms “SMN”, “SMA.”, “survival motor neuron protein” and “spinal muscular atrophy” was performed, and the results were loaded into the bibliography database tool ProCite (Paquette-Murphy 1996). The abstracts of papers from 1995-2003 were then manually read for information suggesting the presence of primary SMN interactions and in some cases secondary and tertiary interactions, that is, proteins which interact with SMN’s primary interaction partners but not SMN itself. This would preferably have been an automated step, but initial attempts at finding papers based on terms such as “binding,” “interaction,” “association,” etc. alone missed many citations that were known to include reported SMN interactions. Citations to manually selected papers were placed in a group in ProCite and obtained either through the ASU library, personal resources, or PubMed. At this phase there were approximately 200 papers which may have held SMN protein-protein binding data. These papers were examined and relevant data was entered into tab-delimited text files for input into Osprey. A complete bibliography of the files used for the Osprey input files is listed in [Appendix A](#), although some are cross-referenced in this text and placed in the initial bibliography as well.

A critical distinguishing factor in the reports of SMN interactions with regards to this project was the experimental system used to prove the interaction. For example, some of the earlier papers only gave data from yeast two-hybrid experiments (Sato, Eguchi et al. 2000; Kurihara, Menea et al. 2001), which, from a conservative standpoint, should not be considered definitive proof of interaction. A main drawback of the yeast two-hybrid method is the potential to identify protein interacting partners that may never occur *in vivo*. The modification of proteins to create bait and prey molecules can compromise the structure function of proteins giving an inaccurate binding profile. Also, post-translational modifications that occur in humans may not always occur identically in proteins produced by yeast, and this can have a severe impact on binding that would otherwise regularly occur. More recent SMN papers tend to rigorously define interactions using additional methods such as *in vitro* binding assays, *in vivo* coimmunoprecipitation experiments, and immunofluorescence studies (Baccon, Pellizzoni et al. 2002; Gubitz, Mourelatos et al. 2002). Often, separate papers will use different methods to prove the same interaction.

To reconcile the variations in experimental evidence of binding, the SMN-Osprey network was generated by creation of several files, each displaying all SMN protein interactions defined by a single experimental binding system. The files were therefore separated as follows:

- a. all SMN interactions defined using yeast two-hybrid experiments
- b. all SMN interactions defined using Histidine-tag affinity pull down assays
- c. all SMN interactions defined using GST affinity pull down assays
- d. all SMN interactions defined using Maltose binding protein pull down assays

- e. all SMN interactions defined using Biomolecular Interaction Analysis (BIA)
- f. all SMN interactions defined using *in vitro* translated protein affinity binding
- g. all SMN interactions defined through CoIP experiments in 293T cells
- h. all SMN interactions defined through CoIP experiments in A431 cells
- i. all SMN interactions defined through CoIP experiments in COS cells
- j. all SMN interactions defined through CoIP experiments in HEK 293 cells
- k. all SMN interactions defined through CoIP experiments in HeLa cells
- l. all SMN interactions defined through CoIP experiments in MDCK cells
- m. all SMN interactions defined through CoIP experiments in mouse brain cells
- n. all SMN interactions defined through CoIP experiments in mouse spinal cord cells
- o. all SMN interactions defined through CoIP experiments in mouse A92L cells
- p. all SMN interactions defined through CoIP experiments in Raji lymphocytes
- q. all SMN interactions defined through CoIP experiments in Schwann cells
- r. all SMN interactions defined through CoIP experiments in U2OS cells

2.2.1.3 Superimposition of SMN:SMN and SMN:GRID files

Two additional files were created, the first was a combination of files a – f, which constituted the dataset determined through *in vitro* studies, and the second was a combination of files g - r, which constituted the dataset determined through *in vivo* studies. As all of these files can be simultaneously superimposed on one another, the last two files provide a more general, conglomerated view of interactions without requiring superimposition of each file (a-r) individually. The efficiency of this method of

visualization is exemplified in Figure 7, where information from roughly 60 papers recording primary, secondary, and tertiary interactions of SMN is concentrated in a single file which contains detailed information about each protein involved in the interaction, the experimental method used to determine the interaction, and links to sources. The resulting file is also extremely flexible in that nodes and vertices may be filtered by author, experimental method and function, and may be graphically manipulated to suit the user's interests.

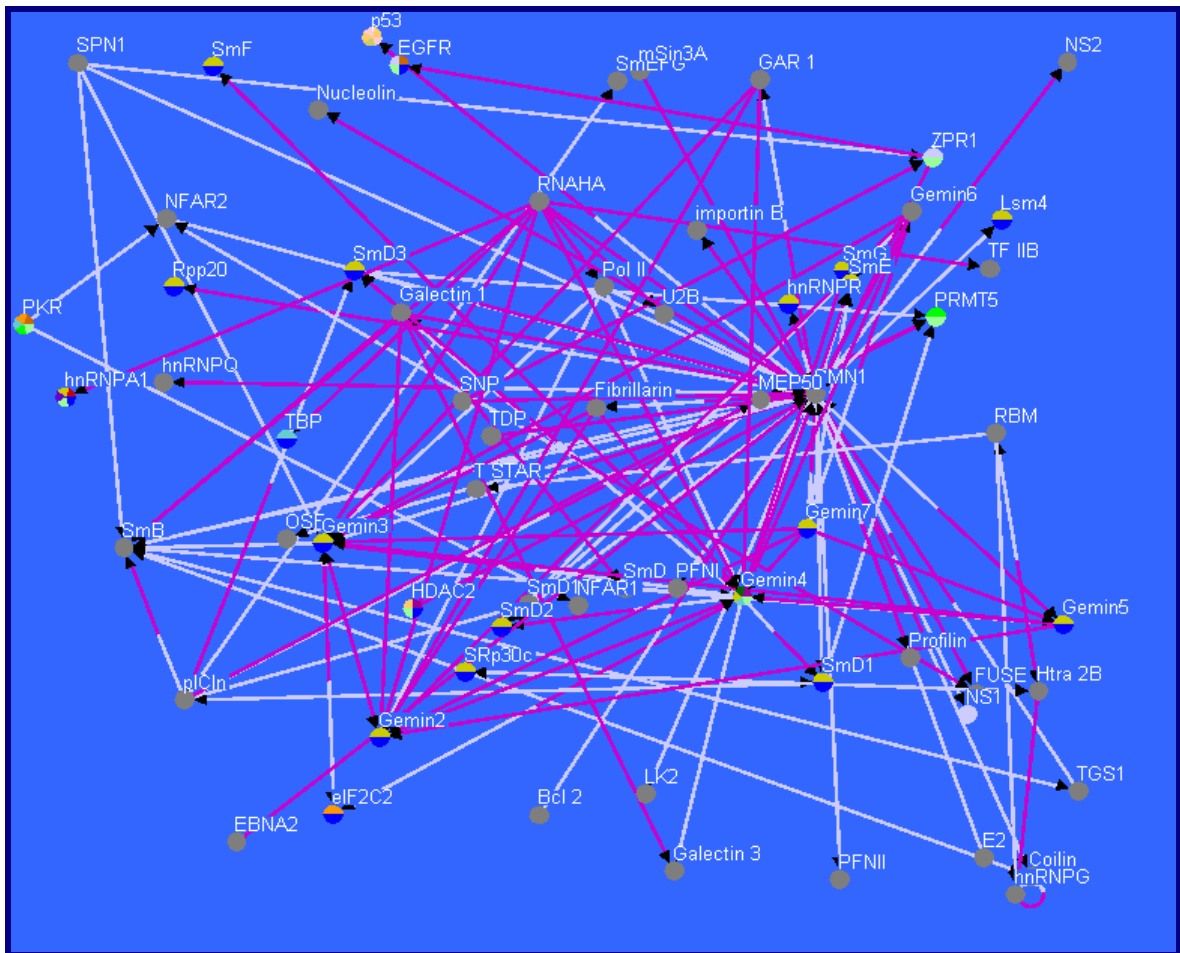


Figure 7: Unformatted superimposition of SMN:SMN protein interactions defined through in vitro (all types) and in vivo (all types) experiments.

2.2.1.4 Gene and Edge information

The left hand display area shows downloaded gene information when a node is highlighted. The ORF name is displayed, as well as the known names for the gene. In the network display area, it is also possible to change the gene screen name. The “add comment” box below the gene screen name allows the user to add comments for that particular gene. If available, information retrieved from the Gene Ontology Consortium (Ashburner, Ball et al. 2000) is also displayed, including GO component (where the gene/protein is found in the cell), GO Process, GO Function (‘abilities’ or jobs of the protein), and GO special. The node is colored according to the GO special data (Figure 9). Gene Ontology is a relatively new field which aims to classify proteins on several levels, enabling the clustering of proteins based on cell location and function.

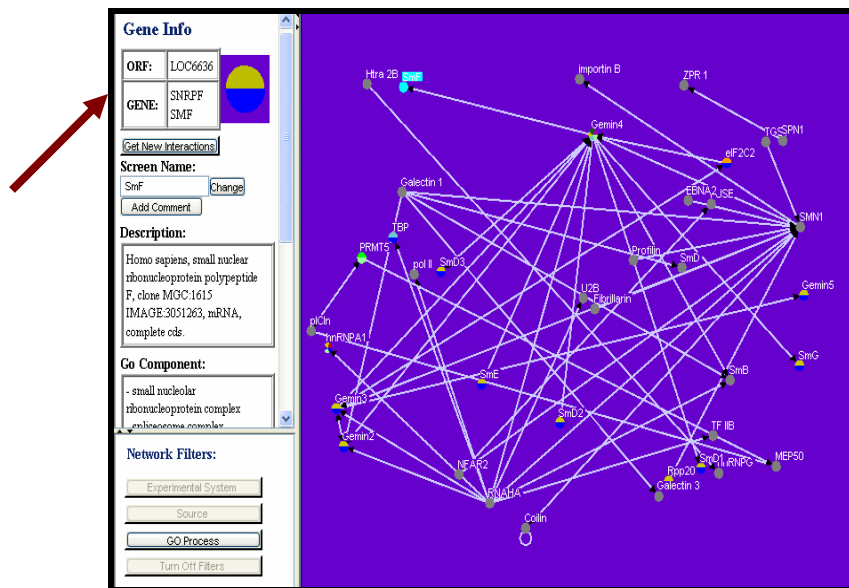


Figure 9: Gene information display.

When an edge is highlighted, limited information regarding both of the connected proteins is displayed. This information included the ORF names for each gene and known names of each gene. Below the gene names a table containing information on the experimental system used to prove the interaction, the source, and the PubMed link is displayed (Figure 10). Node information may also be viewed in table format from the View menu under the “Selected node report” option.

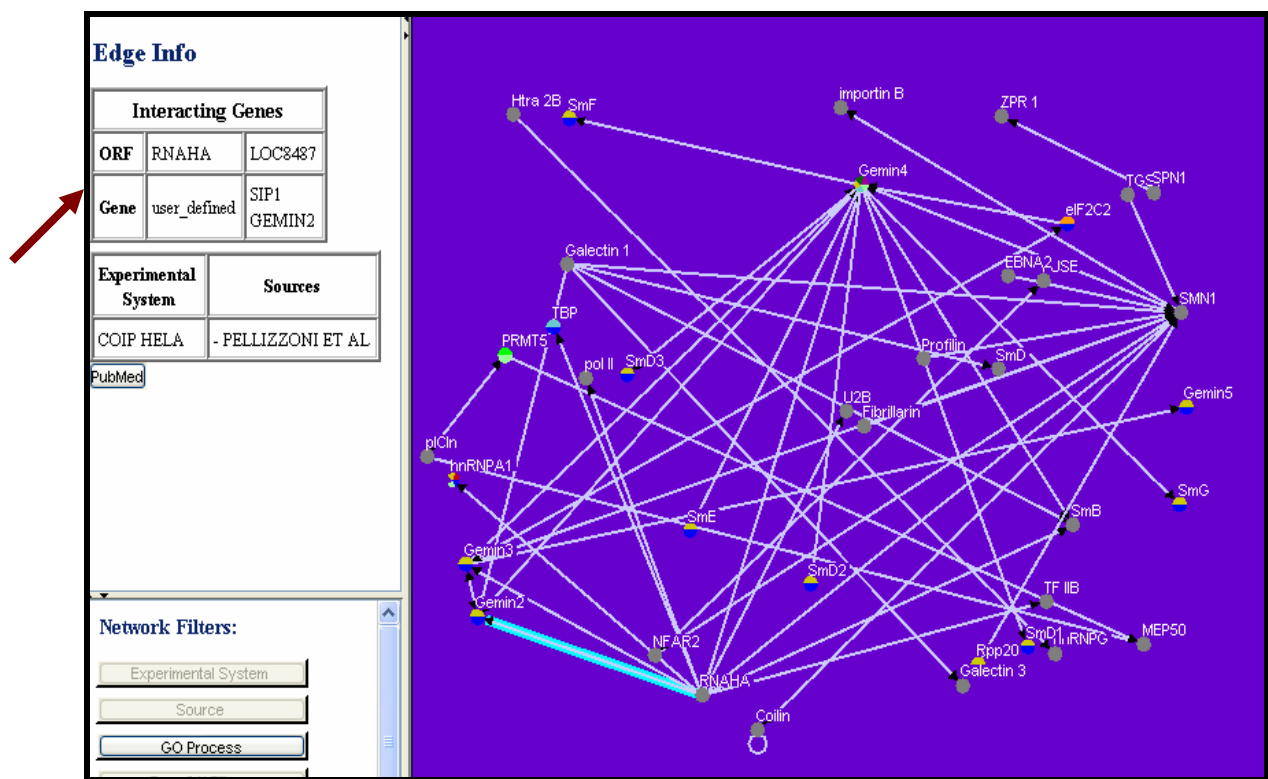


Figure 10: Edge information display.

2.2.1.4 Color indices

Osprey uses three major color indices when coloring a network. Nodes will automatically be colored according to their GO process, although the Format → advanced setting option allows user defined colors to be substituted. All GO processes the protein is involved in do not need to be shown on each node, the number of colors on the node may also be set in advanced settings (Figure 11).

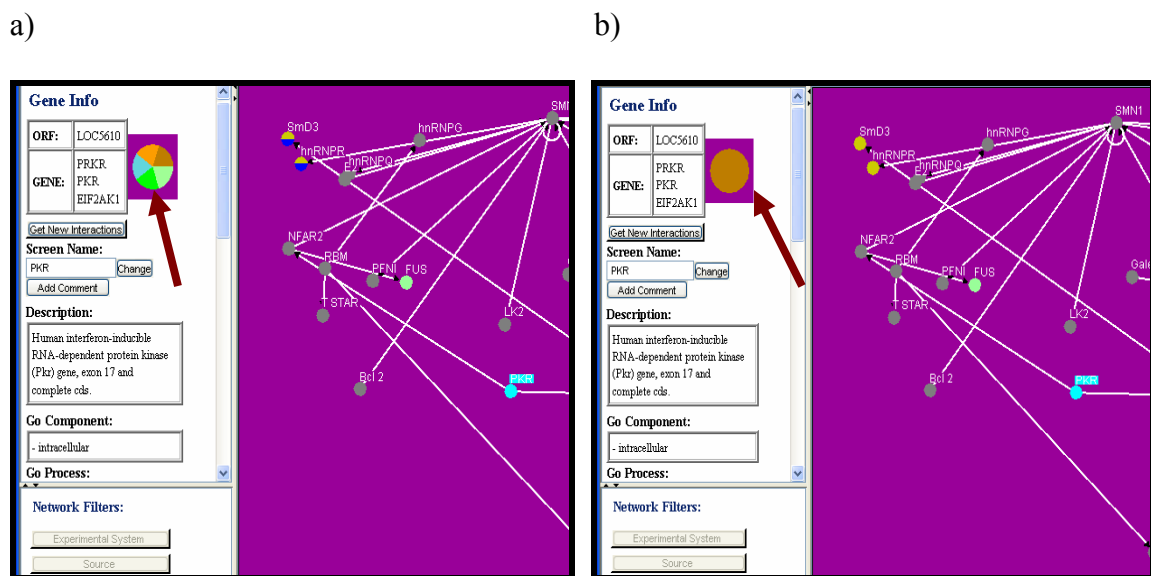
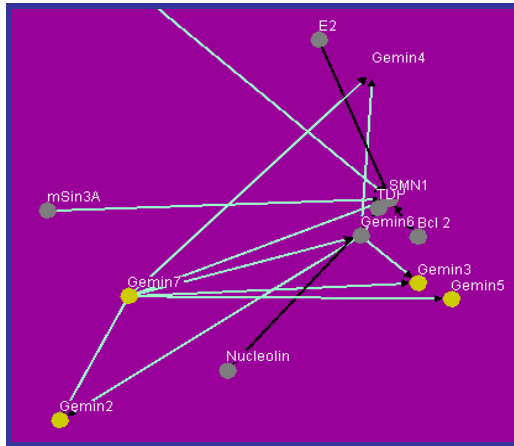


Figure 11: a) A node colored by only one of its GO processes b) the same node colored by all of its GO processes.

When the network display area is in “Color by experimental system” mode, also accessed through “Advanced settings,” the edges will be colored according to the experimental system that was used to define the interaction. When the display area is in “Color by source” mode, however, the edges will be colored according to the first authors listed in the original data file “source” column (Figure 12).

a)



b)

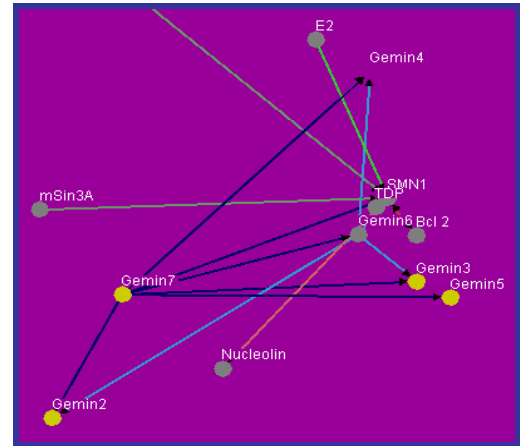


Figure 12: a) Edges colored by experimental system b) the same edges, colored by source.

2.2.1.5 Network and connectivity filters

Filters can be applied to the network display area so that only certain interactions are visible.

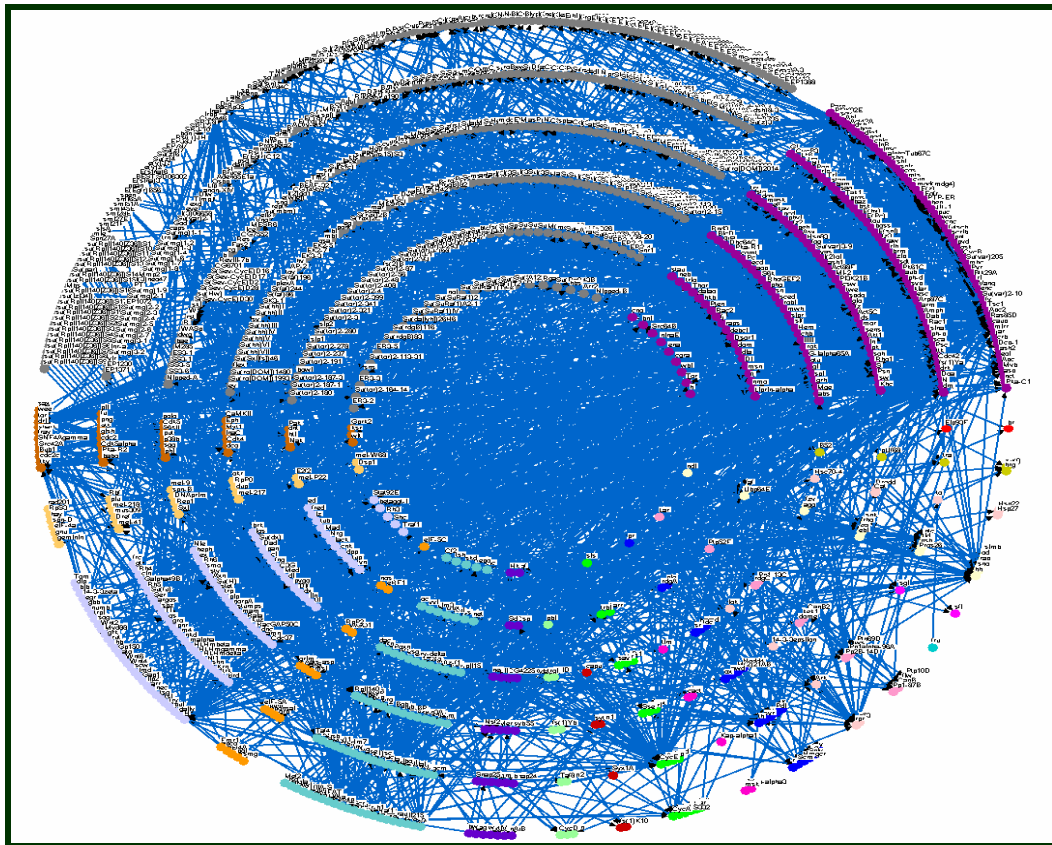
The network may be filtered by experimental system, source, or GO process. These options are accessed through the “Filters” pull down menu and only are applicable if the edges are colored by experimental system, the edges are colored by source, and the nodes are colored by GO process, respectively. Connectivity filters are also accessed through the Filters menu and allow the user to define a minimum number of connections a node must have in order to be displayed (“minimum,” “depth”). An iterative option is also available, where a minimum number of connections is displayed, nodes with less than that minimum are removed, the network is

reevaluated, and if on the second pass any nodes have less than the minimum they too are removed.

2.2.1.6 Network Layouts

One of the most powerful features of Osprey is its ability to rearrange input data in different graphical layouts. Java Sun graphing algorithms are used to relax nodes, arrange nodes in one circle or concentric circles, create dual rings with either highly connected nodes on the inside or outside, or created a spoke/ring layout (Figure 13).

A)



B)

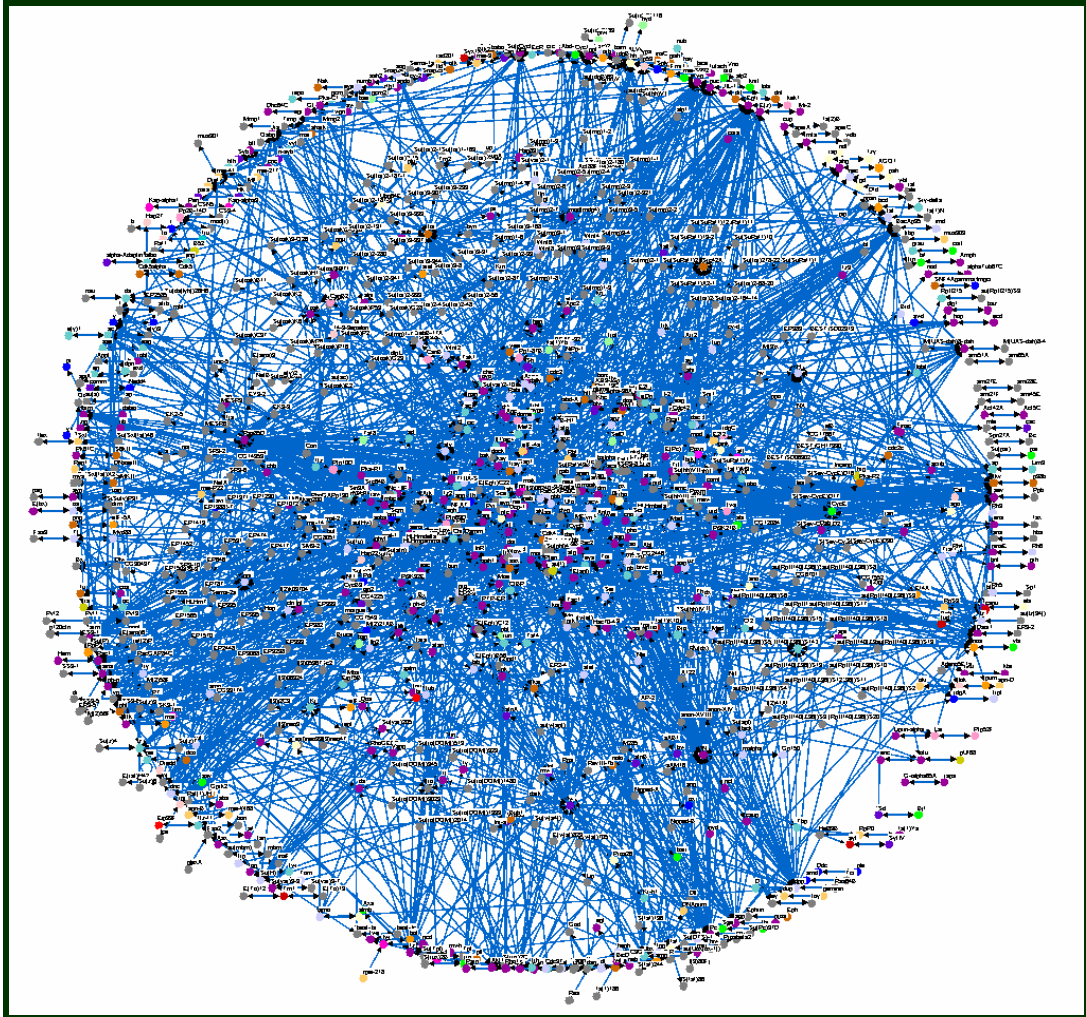


Figure 13: Examples of network layouts. A) Yeast two-hybrid data for *S. cerevisiae* in concentric circles B) Same data in spoked dual ring layout.

Using these layouts for the SMN data gave less visibly conclusive information, however, since many of the proteins with which SMN has been reported to interact have not been entered into the human GRID, so the dataset was limited to the user defined SMN files (Figure 14).

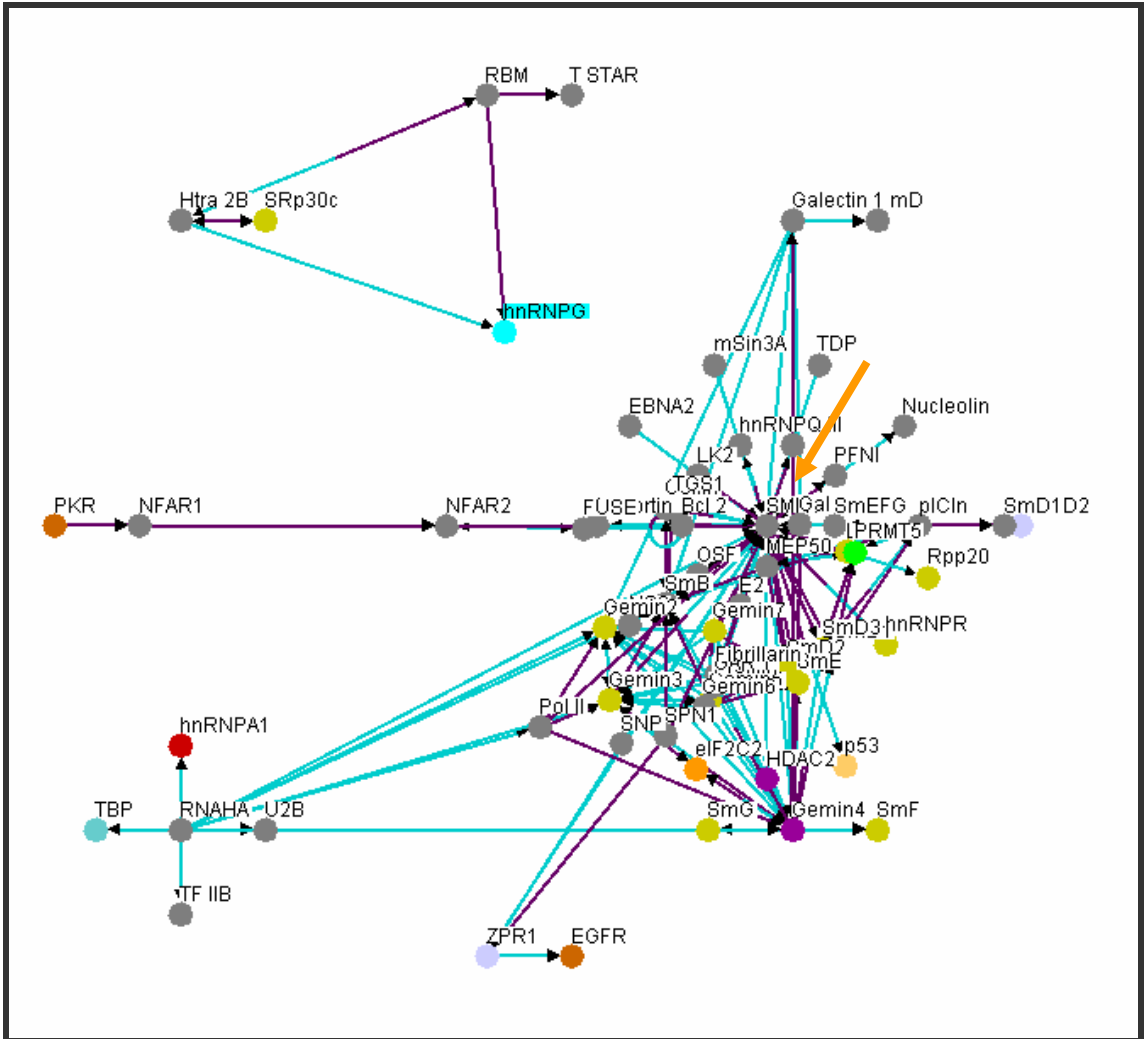


Figure 14: SMN interactions displayed in a spoked dual ring layout (orange arrow indicates position of SMN).

The display option that most characterized the extraordinary level of SMN interactions was the Dual ring – highly connected out option (Figure 15). Ideally, the interactions of SMN could be grouped according to location in the cell and GO process. Although a functional clustering option is available, many of the SMN-interacting proteins (including SMN itself) are not yet assigned gene ontology (GO) annotation, so a functional clustering method provides only limited

information (Figure 16) (Table 1). This is expected to change as more genes are annotated by the GO consortium.

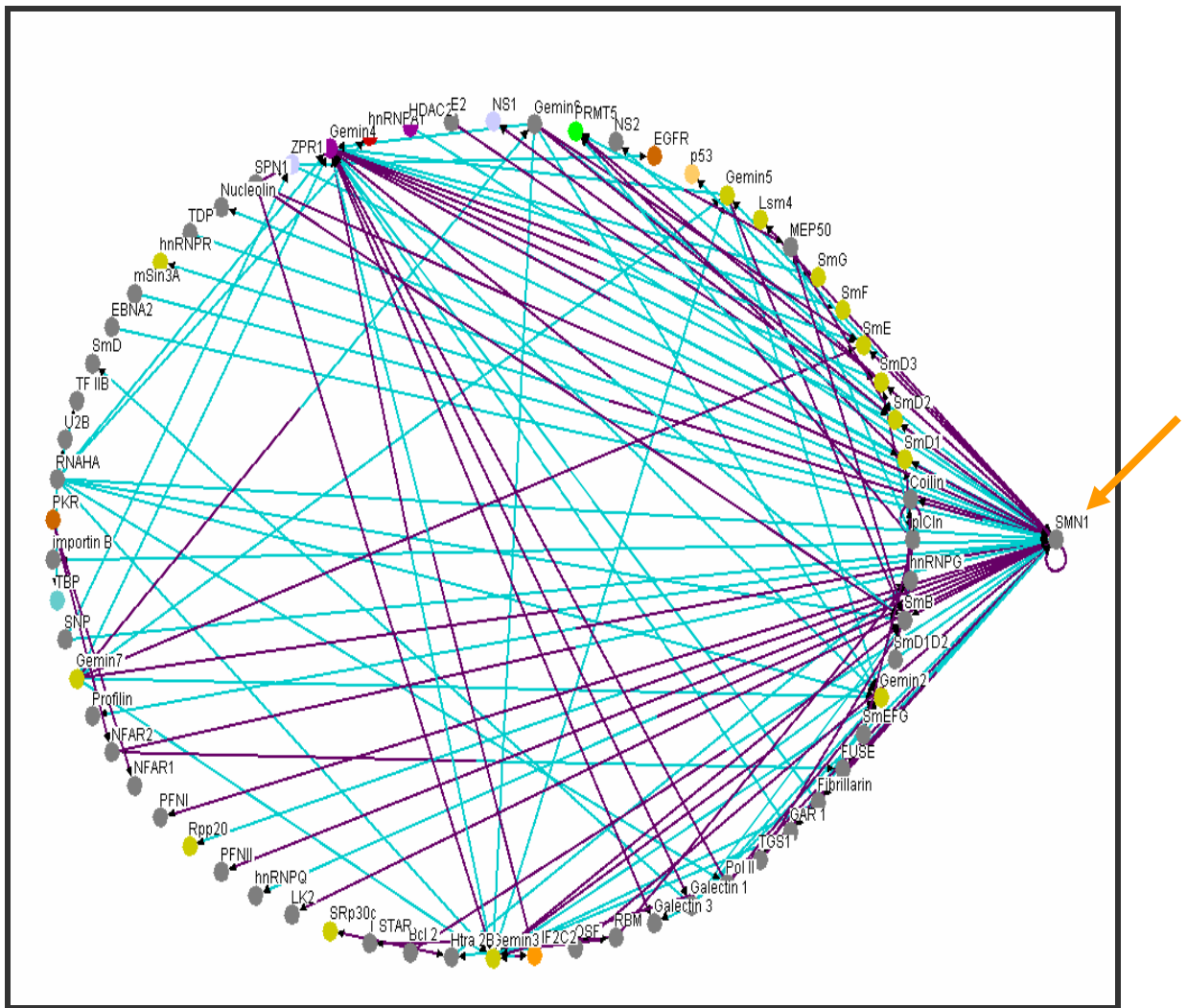


Figure 15: SMN interactions in a dual ring-highly connected out layout (orange arrow indicates position of SMN).

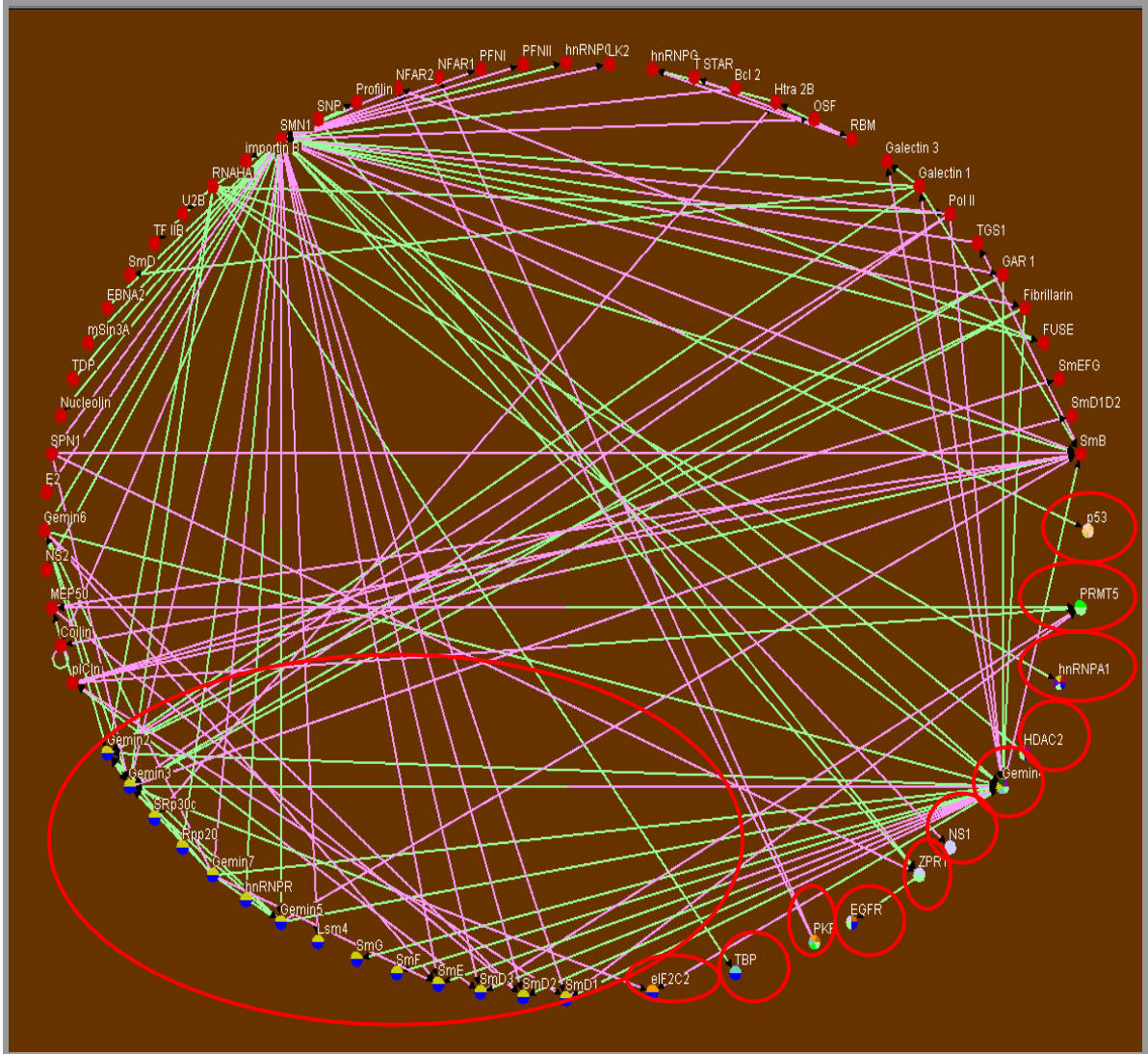


Figure 16: SMN interacting proteins clustered by GO process (Functional groups are circled in red).

Table 1: Known GO components of SMN interacting proteins shown in Figure 15

YORF	Gene	Description	GO Component	GO Function	GO Process	GO Special
LOC6908	TBP GTF2D SCA17 TFIID GTF2D1	Initiation of transcription by RNA polymerase II requires the activities of more than 70 polypeptides. The protein that coordinates these activities is transcription factor IID (TFIID), which binds to the core promoter to position the polymerase properly, serves as the scaffold for assembly of the remainder of the transcription complex, and acts as a channel for regulatory signals. TFIID is composed of the TATA-binding protein (TBP) and a group of evolutionarily conserved proteins known as TBP-associated factors or TAFs. TAFs may participate in basal transcription, serve as coactivators, function in promoter recognition or modify general transcription factors (GTFs) to facilitate complex assembly and transcription initiation. This gene encodes TBP, the TATA-binding protein. A distinctive feature of TBP is a long string of glutamines in the N-terminal. This region of the protein modulates the DNA binding activity of the C terminus, and modulation of DNA binding affects the rate of transcription complex formation and initiation of transcription. Mutations that expand the number of CAG repeats encoding this polyglutamine tract, and thus increase the length of the polyglutamine string, are associated with spinocerebellar ataxia 17, a neurodegenerative disorder classified as a polyglutamine disease.	- transcription factor TFIID complex - nucleus	- general RNA polymerase II transcription factor activity - DNA binding	- transcription initiation from Pol II promoter - transcription - regulation of transcription, DNA-dependent	- transcription - metabolism
LOC5610	PRKR PKR EIF2AK1	Human interferon-inducible RNA-dependent protein kinase (Pkr) gene, exon 17 and complete cds.	- intracellular	- eukaryotic translation initiation factor 2alpha kinase activity - double-stranded RNA binding - protein-tyrosine kinase activity - ATP binding - antiviral response protein activity - protein phosphatase type 2A, intrinsic regulator activity - transferase activity	- negative regulation of cell proliferation - cell cycle - protein amino acid phosphorylation - regulation of transcription, DNA-dependent - regulation of translational initiation - apoptosis - immune response	- cell growth and/or maintenance - cell cycle - protein amino acid phosphorylation - metabolism - transcription - protein biosynthesis
LOC1956	EGFR ERBB	Homo sapiens epidermal growth factor receptor (EGFR) gene, complete cds,	- endosome - cytoskeleton	- epidermal growth factor receptor	- EGF receptor signaling pathway	- signal transduction

	ERBB1	alternatively spliced; and 5S ribosomal RNA gene, complete sequence.	- integral to plasma membrane	activity - ATP binding - receptor activity - transferase activity	- protein amino acid phosphorylation - cell proliferation	- protein amino acid phosphorylation - metabolism - cell growth and/or maintenance
LOC8882	ZNF259 ZPR1	Homo sapiens, Similar to zinc finger protein 259, clone MGC:20397 IMAGE:4578136, mRNA, complete cds.	- nucleolus - cytoplasm	- protein binding - zinc ion binding	- cell proliferation - signal transduction	- cell growth and/or maintenance - signal transduction
LOC5781	PTPN11 CFC NS1 SHP2 BPTP3 PTP2C PTP-1D PRO1847 SH-PTP2 SH-PTP3 MGC14433	The protein encoded by this gene is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. This PTP contains two tandem Src homology-2 domains, which function as phospho-tyrosine binding domains and mediate the interaction of this PTP with its substrates. This PTP is widely expressed in most tissues and plays a regulatory role in various cell signaling events that are important for a diversity of cell functions, such as mitogenic activation, metabolic control, transcription regulation, and cell migration. Mutations in this gene are a cause of Noonan syndrome.	- NONE	- non-membrane spanning protein tyrosine phosphatase activity - protein tyrosine phosphatase activity	- signal transduction	- signal transduction
LOC3066	HDAC2 RPD3 YAF1	Histone deacetylase 2 (HDAC2), or transcriptional regulator homolog RPD3 L1, is highly homologous to the yeast transcription factor RPD3 (reduced potassium dependency 3) gene. As in yeast, human HDA2 is likely to be involved in regulating chromatin structure during transcription. It has been implicated to associate with YY1, a mammalian zinc-finger transcription factor, which negatively regulates transcription by tethering RPD3 to DNA as a cofactor. This process is highly conserved from yeast to human.	- nucleus	- histone deacetylase activity	- histone deacetylation	- metabolism - DNA metabolism - cell organization and biogenesis - cell growth and/or maintenance
LOC8487	SIP1 GEMIN2	Homo sapiens survival of motor neuron protein interacting protein 1 (SIP1) mRNA, complete cds.	- spliceosome complex - cytoplasm - nucleus	- pre-mRNA splicing factor activity	- mRNA processing - spliceosome assembly	- RNA processing - metabolism
LOC27161	EIF2C2 Q10 AGO2 MGC3183	This gene encodes a member of the Argonaute family of proteins which play a role in RNA interference. The encoded protein is highly basic, and contains a PAZ domain and a PIWI domain. It may interact with dicer1 and play a role in short-interfering-	- cellular_component unknown	- translation initiation factor activity	- protein biosynthesis	- protein biosynthesis - metabolism

		RNA-mediated gene silencing.				
LOC11218	DDX20 DP103 GEMIN3 DKFZP434H052 DKFZp434H052	DEAD box proteins, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD), are putative RNA helicases. They are implicated in a number of cellular processes involving alteration of RNA secondary structure such as translation initiation, nuclear and mitochondrial splicing, and ribosome and spliceosome assembly. Based on their distribution patterns, some members of this family are believed to be involved in embryogenesis, spermatogenesis, and cellular growth and division. This gene encodes a DEAD box protein, which has an ATPase activity and is a component of the survival of motor neurons (SMN) complex. This protein interacts directly with SMN, the spinal muscular atrophy gene product, and may play a catalytic role in the function of the SMN complex on RNPs.	- spliceosome complex - cytoskeleton	- ATP dependent RNA helicase activity - ATP binding - DNA binding	- assembly of spliceosomal tri-snRNP - mRNA processing	- RNA processing - metabolism
LOC8683	SFRS9 SRp30c	Cytochrome c oxidase (COX), the terminal enzyme of the mitochondrial respiratory chain, catalyzes the electron transfer from reduced cytochrome c to oxygen. It is a heteromeric complex consisting of 3 catalytic subunits encoded by mitochondrial genes and multiple structural subunits encoded by nuclear genes. The mitochondrially-encoded subunits function in the electron transfer and the nuclear-encoded subunits may function in the regulation and assembly of the complex. This nuclear gene encodes polypeptide 1 (liver isoform) of subunit VIa, and polypeptide 1 is found in all non-muscle tissues. Polypeptide 2 (heart/muscle isoform) of subunit VIa is encoded by a different gene, and is present only in striated muscles. These two polypeptides share 66% amino acid sequence identity. It has been reported that there may be several pseudogenes on chromosomes 1, 6, 7q21, 7q31-32 and 12. However, only one pseudogene (COX6A1P) on chromosome 1p31.1 has been documented.	- nucleus	- pre-mRNA splicing factor activity - RNA binding	- mRNA splice site selection - nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism
LOC10248	RPP20 POP7 RPP2	Homo sapiens, POP7 (processing of precursor, <i>S. cerevisiae</i>) homolog, clone MGC:1986 IMAGE:3138336, mRNA, complete cds.	- nucleolar ribonuclease P complex - nucleus	- ribonuclease P activity - hydrolase activity	- tRNA processing	- RNA processing - metabolism
LOC79760	GEMIN7 FLJ13956	Homo sapiens, hypothetical protein FLJ13956, clone MGC:14121 IMAGE:4053402, mRNA, complete cds.	- spliceosome complex	- NONE	- nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism

LOC10236	HNRPR HNRNPR hnRNP-R	This gene belongs to the subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs). The hnRNPs are RNA binding proteins and they complex with heterogeneous nuclear RNA (hnRNA). These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. While all of the hnRNPs are present in the nucleus, some seem to shuttle between the nucleus and the cytoplasm. The hnRNP proteins have distinct nucleic acid binding properties. The protein encoded by this gene has three repeats of quasi-RRM domains that bind to RNAs and also contains a nuclear localization motif.	- nucleus - ribonucleoprotein complex	- heterogeneous nuclear ribonucleoprotein - RNA binding	- mRNA processing	- RNA processing - metabolism
LOC50628	GEMIN4 HHRF-1 DKFZP434B131 DKFZP434D174	GEMIN4 is part of a high molecular mass multiprotein complex that also contains SMN (MIM 600354), GEMIN2 (SIP1; MIM 602595) and GEMIN3 (DDX20; MIM 606168). In the cytoplasm, this complex is associated with small nuclear ribonucleoproteins (snRNPs).[supplied by OMIM]	- nucleolus - spliceosome complex - cytoplasm - small nuclear ribonucleoprotein complex	- NONE	- rRNA processing - nuclear mRNA splicing, via spliceosome	- RNA processing - ribosome biogenesis - transcription - metabolism - cell organization and biogenesis - cell growth and/or maintenance
LOC25929	GEMIN5 DKFZP586M1824	GEMIN5 is part of a large macromolecular complex localized to both the cytoplasm and the nucleus that plays a role in the cytoplasmic assembly of small nuclear ribonucleoproteins (snRNPs). Other members of this complex include SMN (MIM 600354), GEMIN2 (SIP1; MIM 602595), GEMIN3 (DDX20; MIM 606168), and GEMIN4 (MIM 606969).[supplied by OMIM]	- spliceosome complex	- NONE	- nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism
LOC25804	LSM4 YER112W	Homo sapiens, U6 snRNA-associated Sm-like protein, clone MGC:29665 IMAGE:5015049, mRNA, complete cds.	- snRNP U6 - small nucleolar ribonucleoprotein complex	- pre-mRNA splicing factor activity - RNA binding	- mRNA splicing - nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism
LOC6637	SNRPG SMG	Homo sapiens, small nuclear ribonucleoprotein polypeptide G, clone MGC:24706 IMAGE:4276853, mRNA, complete cds.	- small nucleolar ribonucleoprotein complex - spliceosome complex - small nuclear ribonucleoprotein complex	- pre-mRNA splicing factor activity - RNA binding	- mRNA splicing - spliceosome assembly	- RNA processing - metabolism
LOC6636	SNRPF SMF	Homo sapiens, small nuclear ribonucleoprotein polypeptide F, clone MGC:1615 IMAGE:3051263, mRNA, complete cds.	- small nucleolar ribonucleoprotein complex - spliceosome complex	- pre-mRNA splicing factor activity - RNA binding	- mRNA splicing - nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism

			- small nuclear ribonucleoprotein complex			
LOC6635	SNRPE SME	Homo sapiens, small nuclear ribonucleoprotein polypeptide E, clone MGC:3652 IMAGE:3609230, mRNA, complete cds.	- small nucleolar ribonucleoprotein complex - spliceosome complex - small nuclear ribonucleoprotein complex	- pre-mRNA splicing factor activity - RNA binding	- spliceosome assembly	- RNA processing - metabolism
LOC6634	SNRPD3 SMD3	The protein encoded by this gene belongs to the small nuclear ribonucleoprotein core protein family. It is required for pre-mRNA splicing and small nuclear ribonucleoprotein biogenesis.	- small nucleolar ribonucleoprotein complex - spliceosome complex - small nuclear ribonucleoprotein complex	- pre-mRNA splicing factor activity	- mRNA splicing - nuclear mRNA splicing, via spliceosome	- RNA processing - metabolism
LOC6633	SNRPD2 SMD2	The protein encoded by this gene belongs to the small nuclear ribonucleoprotein core protein family. It is required for pre-mRNA splicing and small nuclear ribonucleoprotein biogenesis. Alternative splicing occurs at this locus and two transcript variants encoding the same protein have been identified.	- small nucleolar ribonucleoprotein complex - spliceosome complex - small nuclear ribonucleoprotein complex	- pre-mRNA splicing factor activity	- mRNA splicing - spliceosome assembly	- RNA processing - metabolism
LOC6632	SNRPD1 SMD1 SNRPD	This gene encodes a small nuclear ribonucleoprotein that belongs to the SNRNP core protein family. The protein may act as a charged protein scaffold to promote SNRNP assembly or strengthen SNRNP-SNRNP interactions through nonspecific electrostatic contacts with RNA.	- small nucleolar ribonucleoprotein complex - small nuclear ribonucleoprotein complex	- pre-mRNA splicing factor activity - RNA binding	- mRNA splicing - spliceosome assembly	- RNA processing - metabolism
LOC3178	HNRPA1 HNRNPA1	This gene belongs to the A/B subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs). The hnRNPs are RNA binding proteins and they complex with heterogeneous nuclear RNA (hnRNA). These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. While all of the hnRNPs are present in the nucleus, some seem to shuttle between the nucleus and the cytoplasm. The hnRNP proteins have distinct nucleic acid binding properties. The protein encoded by this gene has two repeats of quasi-RRM domains that bind to RNAs. It is one of the most abundant core proteins of hnRNP complexes and it is localized to the nucleoplasm. This protein, along with other hnRNP proteins, is exported	- nucleoplasm - cytoplasm - ribonucleoprotein complex	- heterogeneous nuclear ribonucleoprotein - RNA binding	- RNA-nucleus export - mRNA processing	- RNA localization - metabolism - transport - cell growth and/or maintenance - RNA processing

		from the nucleus, probably bound to mRNA, and is immediately re-imported. Its M9 domain acts as both a nuclear localization and nuclear export signal. The encoded protein is involved in the packaging of pre-mRNA into hnRNP particles, transport of poly A+ mRNA from the nucleus to the cytoplasm, and may modulate splice site selection. It is also thought have a primary role in the formation of specific myometrial protein species in parturition. Multiple alternatively spliced transcript variants have been found for this gene but only two transcripts are fully described. These variant have multiple alternative transcription initiation sites and multiple polyA sites.				
LOC10419	SKB1 JBP1 IBP72 PRMT5 SKB1Hs	Homo sapiens cDNA FLJ14831 fis, clone OVARC1001107, highly similar to Homo sapiens protein methyltransferase (JBP1) mRNA.	- NONE	- NONE	- regulation of mitosis - cell proliferation	- cell cycle - cell growth and/or maintenance
LOC7157	TP53 P53 p53 TRP53	Tumor protein p53, a nuclear protein, plays an essential role in the regulation of cell cycle, specifically in the transition from G0 to G1. It is found in very low levels in normal cells, however, in a variety of transformed cell lines, it is expressed in high amounts, and believed to contribute to transformation and malignancy. p53 is a DNA-binding protein containing DNA-binding, oligomerization and transcription activation domains. It is postulated to bind as a tetramer to a p53-binding site and activate expression of downstream genes that inhibit growth and/or invasion, and thus function as a tumor suppressor. Mutants of p53 that frequently occur in a number of different human cancers fail to bind the consensus DNA binding site, and hence cause the loss of tumor suppressor activity. Alterations of the TP53 gene occur not only as somatic mutations in human malignancies, but also as germline mutations in some cancer-prone families with Li-Fraumeni syndrome.	- nucleus	- nuclease activity - transcription factor activity	- induction of apoptosis by hormones - DNA damage response, signal transduction resulting in induction of apoptosis - cell cycle checkpoint - regulation of transcription, DNA-dependent - DNA repair - DNA recombination - apoptosis - negative regulation of cell cycle	- response to DNA damage stimulus - response to stress - cell cycle - cell growth and/or maintenance - transcription - metabolism - DNA repair - DNA metabolism - DNA recombination

For this project, SMN was manually clustered with respect to its proposed functions (sections

Functional clustering of SMN interactions represents a potent tool for understanding SMN function. Functional clustering visually groups reports of interactions that otherwise appear to be unrelated. The process of clustering leads to new hypotheses of assembly of SMN interaction complexes, and rapidly defines simple experiments to verify a proposed function. For example, the interaction of SMN with ZPR1, SPN1, Coilin, and Nucleolin have all been individually determined. Interactions between Coilin, SPN1, Nucleolin and/or ZPR1 would demonstrate that some or all of these proteins are involved in a new SMN complex. Further determination of interactions in this group would also help establish if SMN interactions to each component are primary, or through a secondary interaction with one of the other proteins.

3. **Contrasting methods of protein interaction visualization**

3.1 **“Cartoon” images**

The first detailed function ascribed to the SMN protein was that of snRNP biogenesis (section 3.1.3.1). Two detailed, but very different, pictures of how SMN operated in snRNP biogenesis were produced by different research groups. “Cartoon” images have been used to show the sequence and location of events in this complex process (Figures 18)(Paushkin, Gubitz et al. 2002). Cartoon images are a widely used approach to visualize and document protein interactions.

The first major critique of cartoons is that too little information is given. Interactions are viewable only in an isolated context and the picture is limited to what one particular research group deems relevant to their research. The second major critique is that cartoon images, particularly, are inherently biased and easily misleading. The issue in this case is that these figures provide *more* information than is known. For example, in Figure 18, the sequence of binding events is implied in the cartoon, but not thoroughly stated in the results. Although the figure may initiate as one research group’s theory, once published, it is often accepted as a result. Again, in Figure 18, pieces of information such as the number of SMN molecules that oligomerize in the SMN-Gemin complex, or the orientation of the Gemin and Sm proteins on the SMN-oligomer are implied but not proven. In addition, interactions recorded in 2001 (Friesen, Paushkin et al. 2001) by the same group that published an expanded version of Figure 18 are not shown. The most reasonable explanation is that the limits of “cartooning” have long since been reached, and even previously documented interactions can not be accurately described with this method.

3.2 Fluorescence and live-cell imaging

Fluorescence imaging and live-cell imaging provide another approach to protein interaction visualization (Pagliardini, Giavazzi et al. 2000; Zhang, Pan et al. 2003). Although indisputably providing a very real glimpse into the cell, these methods still present problems when attempts are made to accurately describe protein interactions. As with cartooning, an inherently biased picture is presented. The researcher determines which regions of the cells to view and even more specifically, which few, out of thousands, of proteins to image. Therefore while the cell carries out thousands of reactions involving protein interactions and proteins are shuttled around the cell on large cyto-skeletal highways, what we ultimately observe with our imaging systems is a tiny fraction of stained cellular proteins on an otherwise blank background.

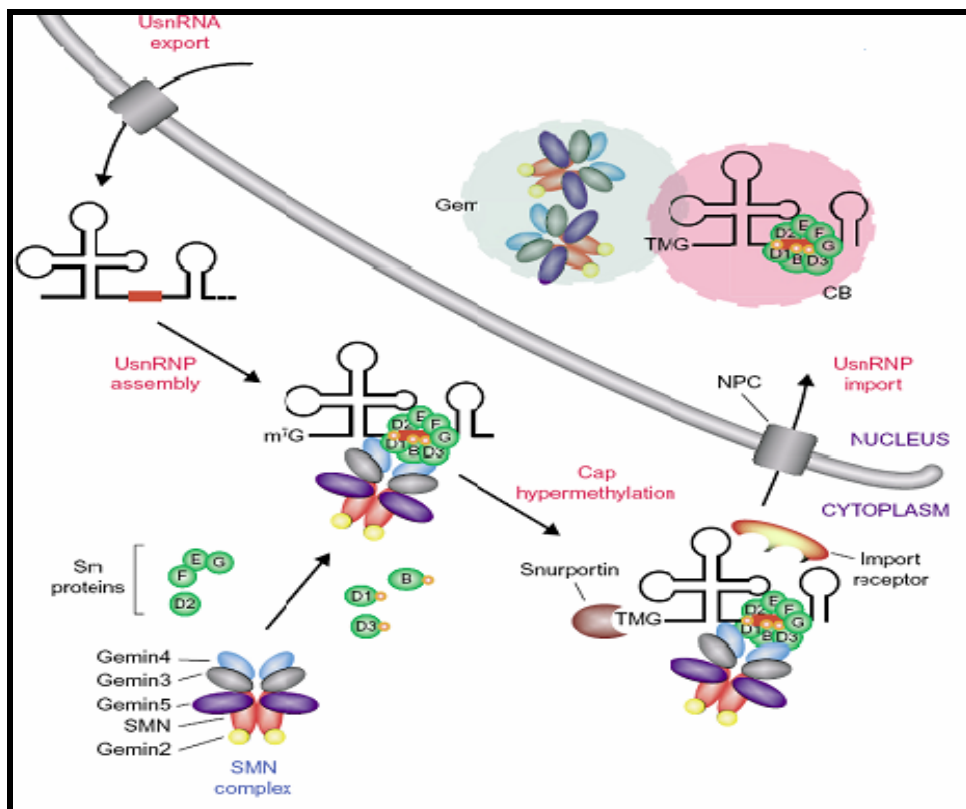


Figure 18: One representation of SMN function (adapted from Paushkin, Gubitzi et al. 2002)

Although immuno-fluorescence and live cell imaging studies have given clues to the location and function of full-length and truncated SMN proteins in the cell, there is clearly not enough information gained from these methods.

3.3 Osprey: advantages and disadvantages

In contrast to the cartoon and cell imaging approaches, using a network visualization system such as Osprey provides neither too little nor too much information. Protein-protein interactions are not shown with implications toward binding order or binding sites, although this information may be specifically included in the comments section if desired. The extremely flexible interface allows rapid manipulation of the network, which allows immediate and creative formation of hypotheses regarding the interactions. To illustrate this, Osprey was used to diagram two conceptual pictures of SMN. First, the neuron-specific interactions of SMN were compared to the non-neuron specific interactions. In SMA, it is specifically the α -motor neurons that are affected. Creating diagrams using Osprey and downloading additional interaction information from the GRID may provide clues to the specific dysfunction in motor neurons (Figure 19). Secondly, SMN is known to interact with several viral proteins (Figure 20).

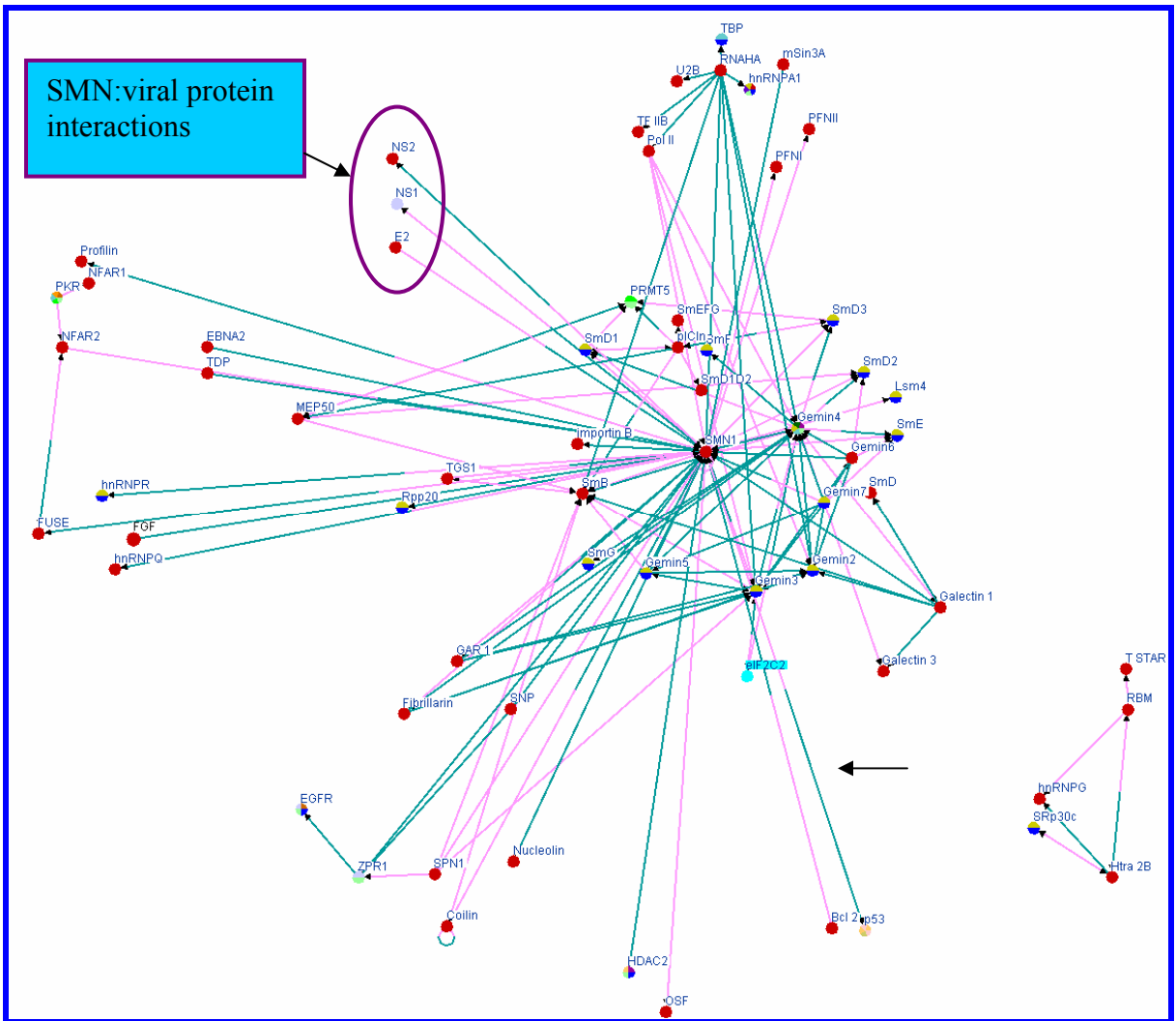


Figure 20: Interactions of SMN with viral proteins.

5. Discussion

For complete understanding on the proteome scale, tools such as Osprey will need to be employed, with the goal of creating highly accessible multi-tiered data that can be viewed in many different ways. For exceptionally interactive proteins such as SMN, the picture will improve as more datasets are publicly available. The data collected from this study gives a positive starting point for SMN interaction visualizations on a large scale, providing a valuable addition to the field of SMA research.

This project demonstrates that manual extraction of protein interactions from the primary literature is a necessary step in understanding the interaction network for a given protein. Unfortunately, manual extraction and input of interactions is highly time and labor intensive, and will continue to represent a large bottleneck to completion of protein-interaction networks. The situation will improve as authors increase use of automated deposition tools in a fashion now used for deposition of crystallographic and x-ray structures. Such deposition is now required with publication, and a similar requirement for authors of publications reporting protein-interactions could greatly facilitate completion of proteome-level interactions. However, multiple interaction databases, the lack of a common deposition standard, and the lack of standards for interaction quality and reproducibility will reduce the usefulness of the networks, at least in the near future. For example, multiple entries of the same protein under different names will make it difficult to uncover common interactions.

Finally, a second major limitation of existing binding network software is the inability to record temporal or post-translational modification dependence on binding interactions in the input files.

Appendix A: Bibliography for Osprey interaction datasets

1. Baccon, J., Pellizzoni, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J Biol Chem*, 277,31957-62.
2. Barth, S., Liss, M., Voss, M.D., Dobner, T., Fischer, U., Meister, G., and Grasser, F.A. (2003) Epstein-Barr virus nuclear antigen 2 binds via its methylated arginine-glycine repeat to the survival motor neuron protein. *J Virol*, 77,5008-13.
3. Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA*, 7,1531-42.
4. Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999) Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Hum Mol Genet*, 8,2351-7.
5. Campbell, L., Hunter, K.M., Mohaghegh, P., Tinsley, J.M., Brasch, M.A., and Davies, K.E. (2000) Direct interaction of Smn with dp103, a putative RNA helicase: a role for Smn in transcription regulation? *Hum Mol Genet*, 9,1093-100.
6. Charroux, B., Pellizzoni, L., Perkinson, R.A., Shevchenko, A., Mann, M., and Dreyfuss, G. (1999) Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J Cell Biol*, 147,1181-94.
7. Charroux, B., Pellizzoni, L., Perkinson, R.A., Yong, J., Shevchenko, A., Mann, M., and Dreyfuss, G. (2000) Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol*, 148,1177-86.
8. Claus, P., Doring, F., Gringel, S., Muller-Ostermeyer, F., Fuhlrott, J., Kraft, T., and Grothe, C. (2002) Differential intranuclear localization of fibroblast growth factor - 2 (FGF-2) isoforms and specific interaction with the survival of motoneuron protein (SMN). *J Biol Chem*.
9. Coovert, D.D., Le, T.T., Morris, G.E., Man, N.T., Kralewski, M., Sendtner, M., and Burghes, A.H. (2000) Does the survival motor neuron protein (SMN) interact with Bcl-2? *J Med Genet*, 37,536-9.
10. Friesen, W.J., Dreyfuss, G. (2000) Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN). *J Biol Chem*, 275,26370-5.
11. Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001) Smn, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol Cell*, 7,1111-7.
12. Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) The Methylosome, a 20S Complex Containing JBP1 and pICln, Produces Dimethylarginine-Modified Sm Proteins. *Mol Cell Biol*, 21,8289-300.
13. Friesen, W.J., Wyce, A., Paushkin, S., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) A novel WD repeat protein component of the methylosome binds Sm proteins. *J Biol Chem*, 277,8243-7.
14. Gangwani, L., Mikrut, M., Theroux, S., Sharma, M., and Davis, R.J. (2001) Spinal muscular atrophy disrupts the interaction of ZPR1 with the SMN protein. *Nat Cell Biol*, 3,376-83.

Appendix A: Bibliography for Osprey interaction datasets

15. Giesemann, T., Rathke-Hartlieb, S., Rothkegel, M., Bartsch, J.W., Buchmeier, S., Jockusch, B.M., and Jockusch, H. (1999) A role for polyproline motifs in the spinal muscular atrophy protein SMN. Profilins bind to and colocalize with smn in nuclear gems. *J Biol Chem*, 274,37908-14.
16. Gubitz, A.K., Mourelatos, Z., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J Biol Chem*, 277,5631-6.
17. Hebert, M.D., Matera, A.G. (2000) Self-association of coilin reveals a common theme in nuclear body localization. *Mol Biol Cell*, 11,4159-71.
18. Hebert, M.D., Shpargel, K.B., Ospina, J.K., Tucker, K.E., and Matera, A.G. (2002) Coilin methylation regulates nuclear body formation. *Dev Cell*, 3,329-37.
19. Hebert, M.D., Szymczyk, P.W., Shpargel, K.B., and Matera, A.G. (2001) Coilin forms the bridge between Cajal bodies and SMN, the Spinal Muscular Atrophy protein. *Genes Dev*, 15,2720-9.
20. Hofmann, Y., Wirth, B. (2002) hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. *Hum Mol Genet*, 11,2037-49.
21. Hua, Y., Zhou, J. (2004) Rpp20 interacts with SMN and is re-distributed into SMN granules in response to stress. *Biochem Biophys Res Commun*, 314,268-76.
22. Iwahashi, H., Eguchi, Y., Yasuhara, N., Hanafusa, T., Matsuzawa, Y., and Tsujimoto, Y. (1997) Synergistic anti-apoptotic activity between Bcl-2 and SMN implicated in spinal muscular atrophy. *Nature*, 390,413-7.
23. Jones, K.W., Gorzynski, K., Hales, C.M., Fischer, U., Badbanchi, F., Terns, R.M., and Terns, M.P. (2001) Direct interaction of the spinal muscular atrophy disease protein smn with the small nucleolar rna-associated protein fibrillarin. *J Biol Chem*, 276,38645-51.
24. Kurihara, N., Mena, C., Maeda, H., Haile, D.J., and Reddy, S.V. (2001) Osteoclast-stimulating Factor Interacts with the Spinal Muscular Atrophy Gene Product to Stimulate Osteoclast Formation. *J Biol Chem*, 276,41035-9.
25. Lefebvre, S., Burlet, P., Viollet, L., Bertrand, S., Huber, C., Belser, C., and Munnich, A. (2002) A novel association of the SMN protein with two major non-ribosomal nucleolar proteins and its implication in spinal muscular atrophy. *Hum Mol Genet*, 11,1017-27.
26. Lindholm, D., Mercer, E.A., Yu, L.Y., Chen, Y., Kukkonen, J., Korhonen, L., and Arumae, U. (2002) Neuronal apoptosis inhibitory protein: Structural requirements for hippocalcin binding and effects on survival of NGF-dependent sympathetic neurons. *Biochim Biophys Acta*, 1600,138-47.
27. Liu, Q., Dreyfuss, G. (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J*, 15,3555-65.
28. Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997) The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell*, 90,1013-21.
29. Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001) Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol*, 11,1990-4.

Appendix A: Bibliography for Osprey interaction datasets

30. Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A.G., Bertrand, E., Tazi, J., and Bordonne, R. (2003) Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep*, 4,616-22.
31. Mourelatos, Z., Abel, L., Yong, J., Kataoka, N., and Dreyfuss, G. (2001) SMN interacts with a novel family of hnRNP and spliceosomal proteins. *EMBO J*, 20,5443-52.
32. Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev*, 16,720-8.
33. Narayanan, U., Ospina, J.K., Frey, M.R., Hebert, M.D., and Matera, A.G. (2002) SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum Mol Genet*, 11,1785-95.
34. Park, J.W., Voss, P.G., Grabski, S., Wang, J.L., and Patterson, R.J. (2001) Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. *Nucleic Acids Res*, 29,3595-602.
35. Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001) The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Curr Biol*, 11,1079-88.
36. Pellizzoni, L., Baccon, J., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J Biol Chem*, 277,7540-5.
37. Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) A functional interaction between the survival motor neuron complex and RNA polymerase II. *J Cell Biol*, 152,75-85.
38. Rossoll, W., Kroning, A.K., Ohndorf, U.M., Steegborn, C., Jablonka, S., and Sendtner, M. (2002) Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Hum Mol Genet*, 11,93-105.
39. Saunders, L.R., Perkins, D.J., Balachandran, S., Michaels, R., Ford, R., Mayeda, A., and Barber, G.N. (2001) Characterization of two evolutionarily conserved, alternatively spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA processing and interact with the double-stranded RNA-dependent protein kinase, PKR. *J Biol Chem*, 276,32300-12.
40. Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., and Sattler, M. (2001) SMN tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol*, 8,27-31.
41. Strasswimmer, J., Lorson, C.L., Breiding, D.E., Chen, J.J., Le, T., Burghes, A.H., and Androphy, E.J. (1999) Identification of survival motor neuron as a transcriptional activator-binding protein. *Hum Mol Genet*, 8,1219-26.
42. Venables, J.P., Elliott, D.J., Makarova, O.V., Makarov, E.M., Cooke, H.J., and Eperon, I.C. (2000) Rbmy, a Probable Human Spermatogenesis Factor, and Other Hnrnp G Proteins Interact With Tra2 Beta and Affect Splicing. *Hum Mol Genet*, 9,685-694.
43. Voss, M.D., Hille, A., Barth, S., Spurk, A., Hennrich, F., Holzer, D., Mueller-Lantzsch, N., Kremmer, E., and Grasser, F.A. (2001) Functional Cooperation of Epstein-Barr Virus Nuclear Antigen 2 and the Survival Motor Neuron Protein in Transactivation of the Viral LMP1 Promoter. *J Virol*, 75,11781-90.

Appendix A: Bibliography for Osprey interaction datasets

44. Wang, I.F., Reddy, N.M., and Shen, C.K. (2002) Higher order arrangement of the eukaryotic nuclear bodies. *Proc Natl Acad Sci U S A*, 99,13583-8.
45. Whitehead, S.E., Jones, K.W., Zhang, X., Cheng, X., Terns, R.M., and Terns, M.P. (2002) Determinants of the interaction of the spinal muscular atrophy disease protein SMN with the dimethylarginine-modified box H/ACA small nucleolar ribonucleoprotein GAR1. *J Biol Chem*, 277,48087-93.
46. Williams, B.Y., Hamilton, S.L., and Sarkar, H.K. (2000) The survival motor neuron protein interacts with the transactivator FUSE binding protein from human fetal brain. *FEBS Lett*, 470,207-10.
47. Yan, X., Mouillet, J.F., Ou, Q., and Sadovskiy, Y. (2003) A Novel Domain within the DEAD-Box Protein DP103 Is Essential for Transcriptional Repression and Helicase Activity. *Mol Cell Biol*, 23,414-23.
48. Young, P.J., Day, P.M., Zhou, J., Androphy, E.J., Morris, G.E., and Lorson, C.L. (2002) A direct interaction between the survival motor neuron protein and p53 and its relationship to spinal muscular atrophy. *J Biol Chem*, 277,2852-9.
49. Young, P.J., DiDonato, C.J., Hu, D., Kothary, R., Androphy, E.J., and Lorson, C.L. (2002) SRp30c-dependent stimulation of survival motor neuron (SMN) exon 7 inclusion is facilitated by a direct interaction with hTra2 beta 1. *Hum Mol Genet*, 11,577-87.
50. Young, P.J., Jensen, K.T., Burger, L.R., Pintel, D.J., and Lorson, C.L. (2002) Minute virus of mice NS1 interacts with the SMN protein, and they colocalize in novel nuclear bodies induced by parvovirus infection. *J Virol*, 76 ,3892-904.
51. Young, P.J., Jensen, K.T., Burger, L.R., Pintel, D.J., and Lorson, C.L. (2002) Minute virus of mice small nonstructural protein NS2 interacts and colocalizes with the Smn protein. *J Virol*, 76,6364-9.
52. Young, P.J., Man, N.T., Lorson, C.L., Le, T.T., Androphy, E.J., Burghes, A.H., and Morris, G.E. (2000) The exon 2b region of the spinal muscular atrophy protein, SMN, is involved in self-association and SIP1 binding. *Hum Mol Genet*, 9,2869-77.
53. Zou, J., Barahmand-Pour, F., Blackburn, M.L., Matsui, Y., Chansky, H.A., and Yang, L. (2004) Survival of motor neuron protein SMN interacts with transcription corepressor mSin3A. *J Biol Chem*.

Appendix B: Interaction data files

BIA.txt

GeneA	GeneB	Experimental	System	Source	PubMed ID
SMN1 p53	BIA	Young et al			11704667
P53 SMN1	BIA	Young et al			11704667
SMN1 SIP1	BIA	Young et al			11092763
SIP1 SMN1	BIA	Young et al			11092763
Htra 2B	SRp30c	BIA	Young et al		11875052
SMN1 NS1	BIA	Young et al			11907229
NS2 SMN1	BIA	Young et al			12021369

GST pull down.txt

GeneA	GeneB	Experimental	System	Source	PubMed ID
pICln	PRMT5	GST pull down		Friesen et al	11756452
MEP50	PRMT5	GST pull down		Friesen et al	11756452
MEP50	SmD2	GST pull down		Friesen et al	11756452
MEP50	SmB	GST pull down		Friesen et al	11756452
Coilin	SMN1	GST pull down		Hebert et al	11641277
SMN1 Coilin		GST pull down		Hebert et al	11641277
SmB Coilin		GST pull down		Hebert et al	11641277
Coilin	SmB	GST pull down		Hebert et al	11641277
pICln	SmD1D2	GST pull down		Meister et al	11747828
pICln	SmEFG	GST pull down		Meister et al	11747828
SPN SMN1		GST pull down		Narayanan et al	12095920
SPN Gemin3		GST pull down		Narayanan et al	12095920
SPN SmB		GST pull down		Narayanan et al	12095920
SPN ZPR1		GST pull down		Narayanan et al	12095920
Fibrillarlin	SMN1	GST pull down		Pellizzoni et al	11509230
GAR1 SMN1		GST pull down		Pellizzoni et al	11509230
Gemin6	SmD2	GST pull down		Pellizzoni et al	11748230
Gemin6	SmE	GST pull down		Pellizzoni et al	11748230
SmB TGS1		GST pull down		Mouaikel et al	12776181
TGS1 SmB		GST pull down		Mouaikel et al	12776181
TGS1 SMN1		GST pull down		Mouaikel et al	12776181
SMN1 RHA		GST pull down		Pellizzoni et al	11149922
Pol II	SMN1	GST pull down		Pellizzoni et al	11149922
Pol II	Gemin2	GST pull down		Pellizzoni et al	11149922
Pol II	Gemin3	GST pull down		Pellizzoni et al	11149922
Pol II	Gemin4	GST pull down		Pellizzoni et al	11149922
Gemin4	Galectin 1	GST pull down		Park et al	11522829
Gemin4	Galectin 3	GST pull down		Park et al	11522829
SMN1 SIP1		GST pull down		Liu et al	9323129
SMN1 SmB		GST pull down		Liu et al	9323129
SMN1 SmD1		GST pull down		Liu et al	9323129

Appendix B: Interaction data files

GST pull down.txt

SMN1 SmD2	GST pull down	Liu et al	9323129	
SMN1 SmD3	GST pull down	Liu et al	9323129	
SMN1 SmE	GST pull down	Liu et al	9323129	
Fibrillarin	SMN1 GST pull down	Jones et al	11509571	
SMN1 Fibrillarin	GST pull down	Jones et al	11509571	
SMN1 Lsm4	GST pull down	Friesen et al	10851237	
Lsm4 SMN1	GST pull down	Friesen et al	10851237	
RBM htra 2B	GST pull down	Venables et al	10748875	
Gemin3	SMN1 GST pull down	Charroux et al	10601333	
Gemin3	SmB GST pull down	Charroux et al	10601333	
eIF2C2	Gemin3 GST pull down	Mourelatos et al	11914277	
eIF2C2	Gemin4 GST pull down	Mourelatos et al	11914277	
Gemin3	eIF2C2 GST pull down	Mourelatos et al	11914277	
Gemin4	eIF2C2 GST pull down	Mourelatos et al	11914277	
NS1 SMN1	GST pull down	Young et al	11907229	
SMN1 GAR1	GST pull down	Whitehead et al	12244096	
Gemin4	Gemin3 GST pull down	Charroux et al	10725331	
Gemin3	Gemin4 GST pull down	Charroux et al	10725331	
Gemin4	SmB GST pull down	Charroux et al	10725331	
Gemin4	SmD1 GST pull down	Charroux et al	10725331	
Gemin4	SmD2 GST pull down	Charroux et al	10725331	
Gemin4	SmD3 GST pull down	Charroux et al	10725331	
Gemin4	SmE GST pull down	Charroux et al	10725331	
E2 SMN1	GST pull down	Strasswimmer et al	10369867	
SMN1 E2	GST pull down	Strasswimmer et al	10369867	
SMN1 Gemin5	GST pull down	Gubitz et al	11714716	
Gemin5	SMN1 GST pull down	Gubitz et al	11714716	
Gemin5	SmB GST pull down	Gubitz et al	11714716	
SMN1 Rpp20	GST pull down	Hua et al	14715275	
SmD1 PRMT5	GST pull down	Friesen et al	11713266	
SmD1 pICln	GST pull down	Friesen et al	11713266	
SmD3 PRMT5	GST pull down	Friesen et al	11713266	
SmD3 pICln	GST pull down	Friesen et al	11713266	
pICln	SmB GST pull down	Pu et al	10330151	
pICln	SmD1 GST pull down	Pu et al	10330151	
pICln	SmD3 GST pull down	Pu et al	10330151	
SMN1 OSF	GST pull down	Kurihara et al	11551898	
Gemin7	Gemin6 GST pull down	Baccon et al	12065586	
Gemin7	SMN1 GST pull down	Baccon et al	12065586	
Gemin7	SmE GST pull down	Baccon et al	12065586	

Appendix B: Interaction data files

His pull down.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1	NS2	His pull down	Young et al	12021369
Gemin3	SF1	His pull down	Yan et al	12482992

MBP pull down.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1	FBP	MBP pull down	Williams et al	10734235
SMN1	Profilin	MBP pull down	Giesemann et al	10608857

CoIP from IVT.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
Coilin	Coilin	CoIPP from IVT	Hebert et al	11102515

Yeast two hybrid.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1	FBP	Yeast two hybrid	Williams et al	10734235
Galectin 1	Gemin4	Yeast two hybrid	Park et al	11522829
SMN1	SMN1	Yeast two hybrid	Liu et al	8670859
SMN1	Fibrillarlin	Yeast two hybrid	Liu et al	8670859
SMN1	SIP1	Yeast two hybrid	Liu et al	9323129
Bcl 2	SMN1	Yeast two hybrid	Iwahashi et al	9389483
Fibrillarlin	SMN1	Yeast two hybrid	Jones et al	11509571
RBM T STAR		Yeast two hybrid	Venables et al	10749975
RBM hnRNPGT		Yeast two hybrid	Venables et al	10749975
RBM htra2B		Yeast two hybrid	Venables et al	10749975
SRp30c	htra2B	Yeast two hybrid	Venables et al	10749975
SMN1	LK2	Yeast two hybrid	Campbell et al	10767334
SMN1	hnRNPR	Yeast two hybrid	Rossoll et al	11773003
SMN1	hnRNPQ	Yeast two hybrid	Rossoll et al	11773003
SMN1	PFNII	Yeast two hybrid	Giesemann et al	10608857
SMN1	PFNI	Yeast two hybrid	Giesemann et al	10608857
Coilin	Coilin	Yeast two hybrid	Hebert et al	11102515
E2	SMN1	Yeast two hybrid	Strasswimmer et al	10369867
SMN1	Rpp20	Yeast two hybrid	Hua et al	14715275
PKR	NFAR1	Yeast two hybrid	Saunders et al	11438536
PKR	NFAR2	Yeast two hybrid	Saunders et al	11438536
NFAR2	SMN1	Yeast two hybrid	Saunders et al	11438536
NFAR2	FUS	Yeast two hybrid	Saunders et al	11438536
pICln	SmD3	Yeast two hybrid	Pu et al	10330151
OSF	SMN1	Yeast two hybrid	Kurihara et al	11551898

Appendix B: Interaction data files

In vitro all.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
Coilin	Coilin	In vitro	Hebert et al	11102515
pICln	PRMT5	In vitro	Friesen et al	11756452
MEP50	PRMT5	In vitro	Friesen et al	11756452
MEP50	SmD2	In vitro	Friesen et al	11756452
MEP50	SmB	In vitro	Friesen et al	11756452
Coilin	SMN1	In vitro	Hebert et al	11641277
SMN1	Coilin	In vitro	Hebert et al	11641277
SmB	Coilin	In vitro	Hebert et al	11641277
Coilin	SmB	In vitro	Hebert et al	11641277
pICln	SmD1D2	In vitro	Meister et al	11747828
pICln	SmEFG	In vitro	Meister et al	11747828
SPN1	SMN1	In vitro	Narayanan et al	12095920
SPN1	Gemin3	In vitro	Narayanan et al	12095920
SPN1	SmB	In vitro	Narayanan et al	12095920
SPN1	ZPR1	In vitro	Narayanan et al	12095920
Fibrillarin	SMN1	In vitro	Pellizzoni et al	11509230
GAR 1	SMN1	In vitro	Pellizzoni et al	11509230
Gemin6	SmD2	In vitro	Pellizzoni et al	11748230
Gemin6	SmE	In vitro	Pellizzoni et al	11748230
SmB	TGS1	In vitro	Mouaikel et al	12776181
TGS1	SmB	In vitro	Mouaikel et al	12776181
TGS1	SMN1	In vitro	Mouaikel et al	12776181
SMN1	RHA	In vitro	Pellizzoni et al	11149922
Pol II	SMN1	In vitro	Pellizzoni et al	11149922
Pol II	Gemin2	In vitro	Pellizzoni et al	11149922
Pol II	Gemin3	In vitro	Pellizzoni et al	11149922
Pol II	Gemin4	In vitro	Pellizzoni et al	11149922
Gemin4	Galectin 1	In vitro	Park et al	11522829
Gemin4	Galectin 3	In vitro	Park et al	11522829
SMN1	SIP1	In vitro	Liu et al	9323129
SMN1	SmB	In vitro	Liu et al	9323129
SMN1	SmD1	In vitro	Liu et al	9323129
SMN1	SmD2	In vitro	Liu et al	9323129
SMN1	SmD3	In vitro	Liu et al	9323129
SMN1	SmE	In vitro	Liu et al	9323129
Fibrillarin	SMN1	In vitro	Jones et al	11509571
SMN1	Fibrillarin	In vitro	Jones et al	11509571
SMN1	Lsm4	In vitro	Friesen et al	10851237
Lsm4	SMN1	In vitro	Friesen et al	10851237
RBM	htra 2B	In vitro	Venables et al	10748875
Gemin3	SMN1	In vitro	Charroux et al	10601333
Gemin3	SmB	In vitro	Charroux et al	10601333
eIF2C2	Gemin3	In vitro	Mourelatos et al	11914277

Appendix B: Interaction data files

In vitro all.txt

eIF2C2	Gemin4	In vitro	Mourelatos et al	11914277
Gemin3	eIF2C2	In vitro	Mourelatos et al	11914277
Gemin4	eIF2C2	In vitro	Mourelatos et al	11914277
NS1	SMN1	In vitro	Young et al	11907229
SMN1	GAR 1	In vitro	Whitehead et al	12244096
Gemin4	Gemin3	In vitro	Charroux et al	10725331
Gemin3	Gemin4	In vitro	Charroux et al	10725331
Gemin4	SmB	In vitro	Charroux et al	10725331
Gemin4	SmD1	In vitro	Charroux et al	10725331
Gemin4	SmD2	In vitro	Charroux et al	10725331
Gemin4	SmD3	In vitro	Charroux et al	10725331
Gemin4	SmE	In vitro	Charroux et al	10725331
E2	SMN1	In vitro	Strasswimmer et al	10369867
SMN1	E2	In vitro	Strasswimmer et al	10369867
SMN1	Gemin5	In vitro	Gubitz et al	11714716
Gemin5	SMN1	In vitro	Gubitz et al	11714716
Gemin5	SmB	In vitro	Gubitz et al	11714716
SMN1	Rpp20	In vitro	Hua et al	14715275
SmD1	PRMT5	In vitro	Friesen et al	11713266
SmD1	pICln	In vitro	Friesen et al	11713266
SmD3	PRMT5	In vitro	Friesen et al	11713266
SmD3	pICln	In vitro	Friesen et al	11713266
pICln	SmB	In vitro	Pu et al	10330151
pICln	SmD1	In vitro	Pu et al	10330151
pICln	SmD3	In vitro	Pu et al	10330151
SMN1	OSF	In vitro	Kurihara et al	11551898
Gemin7	Gemin6	In vitro	Baccon et al	12065586
Gemin7	SMN1	In vitro	Baccon et al	12065586
Gemin7	SmE	In vitro	Baccon et al	12065586
SMN1	FUSE	In vitro	Williams et al	10734235
Galectin 1	Gemin4	In vitro	Park et al	11522829
SMN1	SMN1	In vitro	Liu et al	8670859
SMN1	Fibrillarin	In vitro	Liu et al	8670859
SMN1	SIP1	In vitro	Liu et al	9323129
Bcl 2	SMN1	In vitro	Iwahashi et al	9389483
Fibrillarin	SMN1	In vitro	Jones et al	11509571
RBM	T STAR	In vitro	Venables et al	10749975
RBM	hnRNPG	In vitro	Venables et al	10749975
RBM	htra2B	In vitro	Venables et al	10749975
SRp30c	htra2B	In vitro	Venables et al	10749975
SMN1	LK2	In vitro	Campbell et al	10767334
SMN1	hnRNPR	In vitro	Rossoll et al	11773003
SMN1	hnRNPG	In vitro	Rossoll et al	11773003
SMN1	PFNII	In vitro	Giesemann et al	10608857

Appendix B: Interaction data files

In vitro all.txt

SMN1	PFNI	In vitro	Gieseemann et al	10608857
Coilin	Coilin	In vitro	Hebert et al	11102515
E2	SMN1	In vitro	Strasswimmer et al	10369867
SMN1	Rpp20	In vitro	Hua et al	14715275
PKR	NFAR1	In vitro	Saunders et al	11438536
PKR	NFAR2	In vitro	Saunders et al	11438536
NFAR2	SMN1	In vitro	Saunders et al	11438536
NFAR2	FUSE	In vitro	Saunders et al	11438536
pICln	SmD3	In vitro	Pu et al	10330151
OSF	SMN1	In vitro	Kurihara et al	11551898
SMN1	FUSE	In vitro	Williams et al	10734235
SMN1	Profilin	In vitro	Gieseemann et al	10608857
SMN1	NS2	In vitro	Young et al	12021369
Gemin3	SF1	In vitro	Yan et al	12482992
SMN1	p53	In vitro	Young et al	11704667
P53	SMN1	In vitro	Young et al	11704667
SMN1	SIP1	In vitro	Young et al	11092763
SIP1	SMN1	In vitro	Young et al	11092763
Htra 2B	SRp30c	In vitro	Young et al	11875052
SMN1	NS1	In vitro	Young et al	11907229
NS2	SMN1	In vitro	Young et al	12021369

CoIP 293T.txt

GeneA	GeneB	Experimental	System	Source	PubMed ID
GAR1	Gemin2	CoIPP	293T	Pellizzoni et al	11509230
GAR1	Gemin3	CoIPP	293T	Pellizzoni et al	11509230
GAR1	Gemin4	CoIPP	293T	Pellizzoni et al	11509230
Gemin6	SMN1	CoIPP	293T	Pellizzoni et al	11748230
Gemin6	Gemin2	CoIPP	293T	Pellizzoni et al	11748230
Gemin6	Gemin3	CoIPP	293T	Pellizzoni et al	11748230
Gemin6	Gemin4	CoIPP	293T	Pellizzoni et al	11748230
mSin3A	SMN1	CoIPP	293T	Zou et al	14749338
HDAC2	SMN1	CoIPP	293T	Zou et al	14749338
TDP	SMN1	CoIPP	293T	Wang et al	12361981
Gemin7	SMN1	CoIPP	293T	Baccon et al	12065586
Gemin7	Gemin2	CoIPP	293T	Baccon et al	12065586
Gemin7	Gemin3	CoIPP	293T	Baccon et al	12065586
Gemin7	Gemin4	CoIPP	293T	Baccon et al	12065586
Gemin7	Gemin5	CoIPP	293T	Baccon et al	12065586
Gemin7	Gemin6	CoIPP	293T	Baccon et al	12065586

Appendix B: Interaction data files

CoIP A431.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
ZPR1	SMN1	CoIPP A431 cells	Gangwani et al	11283611
ZPR1	EGFR	CoIPP A431 cells	Gangwani et al	11283611
ZPR1	EF1a	CoIPP A431 cells	Gangwani et al	11283611

CoIP COS.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1	Bcl 2	CoIPP COS	Iwahashi et al	9389483
Bcl 2	SMN1	CoIPP COS	Iwahashi et al	9389483
Nucleolin	SMN1	CoIPP COS	Lefebvre et al	11978761
SMN1	Nucleolin	CoIPP COS	Lefebvre et al	11978761
E2	SMN1	CoIPP COS	Strasswimmer et al	10369867

CoIP HEK293.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1	FBP	CoIPP HEK 293	Williams et al	10734235
Htra 2B	hnRNPG	CoIPP HEK 293	Hofmann et al	12165565
Htra 2B	RBM	CoIPP HEK 293	Hofmann et al	12165565
hnRNPR	SMN1	CoIPP HEK 293	Rossoll et al	11773003
hnRNPO	SMN1	CoIPP HEK 293	Rossoll et al	11773003
Gemin5	SMN1	CoIPP HEK 293	Gubitz et al	11714716
Gemin5	Gemin2	CoIPP HEK 293	Gubitz et al	11714716
Gemin5	Gemin3	CoIPP HEK 293	Gubitz et al	11714716
Gemin5	Gemin4	CoIPP HEK 293	Gubitz et al	11714716

CoIP HEK293.txt

SMN1	Gemin5	CoIPP HEK 293	Gubitz et al	11714716
------	--------	---------------	--------------	----------

CoIP HeLa.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
PRMT5	pICln	CoIPP HeLa	Friesen et al	11756452
PRMT5	MEP50	CoIPP HeLa	Friesen et al	11756452
pICln	PRMT5	CoIPP HeLa	Friesen et al	11756452
pICln	MEP50	CoIPP HeLa	Friesen et al	11756452
SMN1	Coilin	CoIPP HeLa	Hebert et al	11641277
Coilin	SMN1	CoIPP HeLa	Hebert et al	11641277
SPN	SMN1	CoIPP HeLa	Narayanan et al	12095920
SMN1	importin B	CoIPP HeLa	Narayanan et al	12095920
SPN	ZPR1	CoIPP HeLa	Narayanan et al	12095920
TGS1	SMN1	CoIPP HeLa	Mouaikel et al	12776181
RHA	SMN1	CoIPP HeLa	Pellizzoni et al	11149922
RHA	Gemin2	CoIPP HeLa	Pellizzoni et al	11149922
RHA	Gemin3	CoIPP HeLa	Pellizzoni et al	11149922
RHA	Gemin4	CoIPP HeLa	Pellizzoni et al	11149922

Appendix B: Interaction data files

CoIP HeLa.txt

RHA	SmB	CoIPP	HeLa	Pellizzoni et al	11149922
RHA	U2B	CoIPP	HeLa	Pellizzoni et al	11149922
RHA	TF IIB	CoIPP	HeLa	Pellizzoni et al	11149922
RHA	TBP	CoIPP	HeLa	Pellizzoni et al	11149922
RHA	hnRNP A1	CoIPP	HeLa	Pellizzoni et al	11149922
RHA	pol II	CoIPP	HeLa	Pellizzoni et al	11149922
Galectin 1	SMN1	CoIPP	HeLa	Park et al	11522829
Galectin 1	Gemin2	CoIPP	HeLa	Park et al	11522829
Galectin 1	Galectin 3	CoIPP	HeLa	Park et al	11522829
Galectin 1	SmB	CoIPP	HeLa	Park et al	11522829
Galectin 1	SmD	CoIPP	HeLa	Park et al	11522829
Fibrillarlin	SMN1	CoIPP	HeLa	Jones et al	11509571
Htra 2B	hnRNP G	CoIPP	HeLa	Venables et al	10749975
Gemin3	SMN1	CoIPP	HeLa	Charroux et al	10601333
Gemin3	Gemin2	CoIPP	HeLa	Charroux et al	10601333
SMN1 Gemin3	CoIPP	HeLa	Charroux et al	10601333	
Gemin2	Gemin3	CoIPP	HeLa	Charroux et al	10601333
Gemin3	Gemin4	CoIPP	HeLa	Mourelatos et al	11914277
Gemin3	Gemin5	CoIPP	HeLa	Mourelatos et al	11914277
Gemin3	eIF2C2	CoIPP	HeLa	Mourelatos et al	11914277
eIF2C2	Gemin3	CoIPP	HeLa	Mourelatos et al	11914277
eIF2C2	Gemin4	CoIPP	HeLa	Mourelatos et al	11914277
EBNA2	SMN1	CoIPP	HeLa	Barth et al	12663808
Profilin	SMN1	CoIPP	HeLa	Gieseemann et al	10608857
Gemin4	SMN1	CoIPP	HeLa	Charroux et al	10725331
Gemin4	Gemin2	CoIPP	HeLa	Charroux et al	10725331
Gemin4	Gemin3	CoIPP	HeLa	Charroux et al	10725331
SMN1 Gemin4	CoIPP	HeLa	Charroux et al	10725331	
Gemin2	Gemin4	CoIPP	HeLa	Charroux et al	10725331
Gemin3	Gemin4	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmB	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmD1	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmD2	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmD3	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmE	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmF	CoIPP	HeLa	Charroux et al	10725331
Gemin4	SmG	CoIPP	HeLa	Charroux et al	10725331
Coilin	Coilin	CoIPP	HeLa	Hebert et al	11102515
Rpp20	SMN1	CoIPP	HeLa	Hua et al	14715275
NFAR2	SMN1	CoIPP	HeLa	Saunders et al	11438536
NFAR2	FUS	CoIPP	HeLa	Saunders et al	11438536
FUS	NFAR2	CoIPP	HeLa	Saunders et al	11438536

Appendix B: Interaction data files

CoIP MDCK.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
pICln	PRMT5	CoIPP MDCK cells	Pu et al	10330151
pICln	SmB	CoIPP MDCK cells	Pu et al	10330151
pICln	SmD3	CoIPP MDCK cells	Pu et al	10330151

CoIP mouse brain.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
Dp103	SMN1	CoIPP mouse brain	Campbell et al	10767334

CoIP mouse spinal cord.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1 Nucleoin		CoIPP mouse spinal cord	Lefebvre et al	11978761

CoIP murine A92L.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1 NS1		CoIPP murine A92L cells	Young et al	11907229
SMN1 NS2		CoIPP murine A92L cells	Young et al	12021369

CoIP Raji lymphocytes.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
Gemin3	SMN1	CoIPP Raji lymphocytes	Voss et al	11689659

CoIP Schwann.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
FGF2 SMN1		CoIPP Schwann cells	Claus et al	12397076
SMN1 FGF2		CoIPP Schwann cells	Claus et al	12397076

CoIP U2OS.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
SMN1 p53		CoIPP U2OS	Young et al	11704667

In vivo all cell types CoIP.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
PRMT5	pICln	CoIPP	Friesen et al	11756452
PRMT5	MEP50	CoIPP	Friesen et al	11756452
pICln	PRMT5	CoIPP	Friesen et al	11756452
pICln	MEP50	CoIPP	Friesen et al	11756452
SMN1 Coilin		CoIPP	Hebert et al	11641277
Coilin	SMN1	CoIPP	Hebert et al	11641277
SNP SMN1		CoIPP	Narayanan et al	12095920
SMN1 importinB		CoIPP	Narayanan et al	12095920
SNP ZPR1		CoIPP	Narayanan et al	12095920
TGS1 SMN1		CoIPP	Mouaikel et al	12776181

Appendix B: Interaction data files

In vivo all cell types CoIP.txt

RNAHA	SMN1	CoIPP	Pellizzoni et al	11149922
RNAHA	Gemin2	CoIPP	Pellizzoni et al	11149922
RNAHA	Gemin3	CoIPP	Pellizzoni et al	11149922
RNAHA	Gemin4	CoIPP	Pellizzoni et al	11149922
RNAHA	SmB	CoIPP	Pellizzoni et al	11149922
RNAHA	U2B	CoIPP	Pellizzoni et al	11149922
RNAHA	TFIIB	CoIPP	Pellizzoni et al	11149922
RNAHA	TBP	CoIPP	Pellizzoni et al	11149922
RNAHA	hnRNPA1	CoIPP	Pellizzoni et al	11149922
RNAHA	Pol II	CoIPP	Pellizzoni et al	11149922
Galectin1	SMN1	CoIPP	Park et al	11522829
Galectin1	Gemin2	CoIPP	Park et al	11522829
Galectin1	Galectin3	CoIPP	Park et al	11522829
Galectin1	SmB	CoIPP	Park et al	11522829
Galectin1	SmD	CoIPP	Park et al	11522829
Fibrillarin	SMN1	CoIPP	Jones et al	11509571
htra 2B	hnRNPA1	CoIPP	Venables et al	10749975
Gemin3	SMN1	CoIPP	Charroux et al	10601333
Gemin3	Gemin2	CoIPP	Charroux et al	10601333
SMN1	Gemin3	CoIPP	Charroux et al	10601333
Gemin2	Gemin3	CoIPP	Charroux et al	10601333
Gemin3	Gemin4	CoIPP	Mourelatos et al	11914277
Gemin3	Gemin5	CoIPP	Mourelatos et al	11914277
Gemin3	eIF2C2	CoIPP	Mourelatos et al	11914277
eIF2C2	Gemin3	CoIPP	Mourelatos et al	11914277
eIF2C2	Gemin4	CoIPP	Mourelatos et al	11914277
EBNA2	SMN1	CoIPP	Barth et al	12663808
Profilin	SMN1	CoIPP	Gieseemann et al	10608857
Gemin4	SMN1	CoIPP	Charroux et al	10725331
Gemin4	Gemin2	CoIPP	Charroux et al	10725331
Gemin4	Gemin3	CoIPP	Charroux et al	10725331
SMN1	Gemin4	CoIPP	Charroux et al	10725331
Gemin2	Gemin4	CoIPP	Charroux et al	10725331
Gemin3	Gemin4	CoIPP	Charroux et al	10725331
Gemin4	SmB	CoIPP	Charroux et al	10725331
Gemin4	SmD1	CoIPP	Charroux et al	10725331
Gemin4	SmD2	CoIPP	Charroux et al	10725331
Gemin4	SmD3	CoIPP	Charroux et al	10725331
Gemin4	SmE	CoIPP	Charroux et al	10725331
Gemin4	SmF	CoIPP	Charroux et al	10725331
Gemin4	SmG	CoIPP	Charroux et al	10725331
Coilin	Coilin	CoIPP	Hebert et al	11102515
Rpp20	SMN1	CoIPP	Hua et al	14715275
NFAR2	SMN1	CoIPP	Saunders et al	11438536

Appendix B: Interaction data files

In vivo all cell types CoIP.txt

NFAR2	FUSE	CoIPP	Saunders et al	11438536
FUSE	NFAR2	CoIPP	Saunders et al	11438536
SMN1 p53	CoIPP	Young et al	11704667	
SMN1 FUSE	CoIPP	Williams et al	10734235	
htra 2B	hnRNPG	CoIPP	Hofmann et al	12165565
htra 2B	RBM	CoIPP	Hofmann et al	12165565
hnRNPR	SMN1	CoIPP	Rossoll et al	11773003
hnRNPQ	SMN1	CoIPP	Rossoll et al	11773003
Gemin5	SMN1	CoIPP	Gubitiz et al	11714716
Gemin5	Gemin2	CoIPP	Gubitiz et al	11714716
Gemin5	Gemin3	CoIPP	Gubitiz et al	11714716
Gemin5	Gemin4	CoIPP	Gubitiz et al	11714716
SMN1 Gemin5	CoIPP	Gubitiz et al	11714716	
Fibrillarin	SMN1	CoIPP	Pellizzoni et al	11509230
Fibrillarin	Gemin2	CoIPP	Pellizzoni et al	11509230
Fibrillarin	Gemin3	CoIPP	Pellizzoni et al	11509230
Fibrillarin	Gemin4	CoIPP	Pellizzoni et al	11509230
GAR 1	SMN1	CoIPP	Pellizzoni et al	11509230
GAR 1	Gemin2	CoIPP	Pellizzoni et al	11509230
GAR 1	Gemin3	CoIPP	Pellizzoni et al	11509230
GAR 1	Gemin4	CoIPP	Pellizzoni et al	11509230
Gemin6	SMN1	CoIPP	Pellizzoni et al	11748230
Gemin6	Gemin2	CoIPP	Pellizzoni et al	11748230
Gemin6	Gemin3	CoIPP	Pellizzoni et al	11748230
Gemin6	Gemin4	CoIPP	Pellizzoni et al	11748230
mSin3A	SMN1	CoIPP	Zou et al	14749338
HDAC2	SMN1	CoIPP	Zou et al	14749338
TDP	SMN1	CoIPP	Wang et al	12361981
Gemin7	SMN1	CoIPP	Baccon et al	12065586
Gemin7	Gemin2	CoIPP	Baccon et al	12065586
Gemin7	Gemin3	CoIPP	Baccon et al	12065586
Gemin7	Gemin4	CoIPP	Baccon et al	12065586
Gemin7	Gemin5	CoIPP	Baccon et al	12065586
Gemin7	Gemin6	CoIPP	Baccon et al	12065586
SMN1 Bcl2	CoIPP	Iwahashi et al	9389483	
Bcl2	SMN1	CoIPP	Iwahashi et al	9389483
Nucleolin	SMN1	CoIPP	Lefebvre et al	11978761
SMN1 nucleolin	CoIPP	Lefebvre et al	11978761	
E2	SMN1	CoIPP	Strasswimmer et al	10369867
Dp103	SMN1	CoIPP	Campbell et al	10767334
SMN1 nucleolin	CoIPP	Lefebvre et al	11978761	
SMN1 NS1	CoIPP	Young et al	11907229	
SMN1 NS2	CoIPP	Young et al	12021369	
pICln	PRMT5	CoIPP	Pu et al	10330151

Appendix B: Interaction data files

In vivo all cell types CoIP.txt

pICln	SmB	CoIPP	Pu et al	10330151
pICln	SmD3	CoIPP	Pu et al	10330151
Gemin3	SMN1	CoIPP	Voss et al	11689659
FGF 2	SMN1	CoIPP	Claus et al	12397076
SMN1	FGF2	CoIPP	Claus et al	12397076
ZPR1	SMN1	CoIPP	Gangwani et al	11283611
ZPR1	EGFR	CoIPP	Gangwani et al	11283611
ZPR1	EF1A	CoIPP	Gangwani et al	11283611

Unknown.txt

GeneA	GeneB	Experimental System	Source	PubMed ID
ZPR1	EGFR	unknown	Galcheva-Gargova et al	8650580
ZPR1	eEF1a	unknown	Gangwani et al	9852145
ZPR1	SMN1	unknown	Gangwani et al	11283611
NAIP	Hippocalcin	unknown	Mercer et al	10899114
LMB	CRM1	unknown	Lain et al	10585254
Htra 2B	hnRNP G	unknown	Venables et al	10749975
Htra 2B	RBM	unknown	Venables et al	10749975
Nucleolin	B23	unknown	Li et al	8620867
Gemin3	EBNA2	unknown	Grundhoff et al	10383418
Gemin3	EBNA3	unknown	Grundhoff et al	10383418
PML	PML	unknown	Ishov et al	10525530
Sam68	Sam68	unknown	Chen et al	10473643
PPP4	Gemin3	unknown	Carnegie et al	12668731
PPP4	Gemin4	unknown	Carnegie et al	12668731

REFERENCES

- Ashburner, M., C. A. Ball, et al. (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet **25**(1): 25-9.
- Baccon, J., L. Pellizzoni, et al. (2002). "Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex." J Biol Chem **277**(35): 31957-62.
- Bader, G. D. and C. W. Hogue (2000). "BIND--a data specification for storing and describing biomolecular interactions, molecular complexes and pathways." Bioinformatics **16**(5): 465-77.
- Batagelj V, M. A. (1998). "Pajek - program for large network analysis." Connections **21**: 47-57.
- Battaglia, G., A. Princivale, et al. (1997). "Expression of the SMN gene, the spinal muscular atrophy determining gene, in the mammalian central nervous system." Hum Mol Genet **6**(11): 1961-71.
- Breitkreutz, B. J., C. Stark, et al. (2003). "The GRID: the General Repository for Interaction Datasets." Genome Biol **4**(3): R23.
- Breitkreutz, B. J., C. Stark, et al. (2003). "Osprey: a network visualization system." Genome Biol **4**(3): R22.
- Broccolini, A., W. K. Engel, et al. (1999). "Localization of survival motor neuron protein in human apoptotic-like and regenerating muscle fibers, and neuromuscular junctions." Neuroreport **10**(8): 1637-41.
- Campbell, L., K. M. Hunter, et al. (2000). "Direct interaction of Smn with dp103, a putative RNA helicase: a role for Smn in transcription regulation?" Hum Mol Genet **9**(7): 1093-100.
- Coovert, D. D., T. T. Le, et al. (1997). "The survival motor neuron protein in spinal muscular atrophy." Hum Mol Genet **6**(8): 1205-14.

Fischer, U., Q. Liu, et al. (1997). "The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis." Cell **90**(6): 1023-9.

Friesen, W. J., S. Paushkin, et al. (2001). "The Methylosome, a 20S Complex Containing JBP1 and pICln, Produces Dimethylarginine-Modified Sm Proteins." Mol Cell Biol **21**(24): 8289-300.

Gubitz, A. K., Z. Mourelatos, et al. (2002). "Gemin5, a Novel WD Repeat Protein Component of the SMN Complex That Binds Sm Proteins." J Biol Chem **277**(7): 5631-5636.

Jones, K. W., K. Gorzynski, et al. (2001). "Direct interaction of the spinal muscular atrophy disease protein smn with the small nucleolar rna-associated protein fibrillarin." J Biol Chem **276**(42): 38645-51.

Kurihara, N., C. Menea, et al. (2001). "Osteoclast stimulatory factor (OSF) interacts with the spinal muscular atrophy (SMA) gene product to stimulate osteoclast formation." J Biol Chem.

Lefebvre, S., L. Burglen, et al. (1995). "Identification and characterization of a spinal muscular atrophy-determining gene." Cell **80**(1): 155-65.

Lefebvre, S., P. Burlet, et al. (1997). "Correlation between severity and SMN protein level in spinal muscular atrophy." Nat Genet **16**(3): 265-9.

Liu, Q. and G. Dreyfuss (1996). "A novel nuclear structure containing the survival of motor neurons protein." EMBO J **15**(14): 3555-65.

Lorson, C. L., E. Hahnen, et al. (1999). "A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy." Proc Natl Acad Sci U S A **96**(11): 6307-11.

McAndrew, P. E., D. W. Parsons, et al. (1997). "Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMNT and SMNC gene copy number." Am J Hum Genet **60**(6): 1411-22.

Melki, J. (1997). "Spinal muscular atrophy." Curr Opin Neurol **10**(5): 381-5.

Narayanan, U., J. K. Ospina, et al. (2002). "SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta." Hum Mol Genet **11**(15): 1785-95.

Pagliardini, S., A. Giavazzi, et al. (2000). "Subcellular localization and axonal transport of the survival motor neuron (SMN) protein in the developing rat spinal cord." Hum Mol Genet **9**(1): 47-56.

Paquette-Murphy, M. R. (1996). "PRO-CITE as a bibliographic management tool." Natl Netw **21**(2): 12, 20.

Paushkin, S., B. Charroux, et al. (2000). "The survival motor neuron protein of *Schizosacharomyces pombe*. Conservation of survival motor neuron interaction domains in divergent organisms." J Biol Chem **275**(31): 23841-6.

Paushkin, S., A. K. Gubitz, et al. (2002). "The SMN complex, an assemblysome of ribonucleoproteins." Curr Opin Cell Biol **14**(3): 305-12.

Pearn, J. (1978). "Incidence, prevalence, and gene frequency studies of chronic childhood spinal muscular atrophy." J Med Genet **15**(6): 409-13.

Pearn, J. (1980). "Classification of spinal muscular atrophies." Lancet **1**(8174): 919-22.

Pellizzoni, L., J. Baccon, et al. (2001). "The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1." Curr Biol **11**(14): 1079-88.

Pellizzoni, L., B. Charroux, et al. (2001). "A functional interaction between the survival motor neuron complex and RNA polymerase II." J Cell Biol **152**(1): 75-85.

Sato, K., Y. Eguchi, et al. (2000). "Regions essential for the interaction between Bcl-2 and SMN, the spinal muscular atrophy disease gene product." Cell Death Differ **7**(4): 374-83.

Williams, B. Y., S. Vinnakota, et al. (1999). "Differential subcellular localization of the survival motor neuron protein in spinal cord and skeletal muscle." Biochem Biophys Res Commun **254**(1): 10-4.

Zhang, H. L., F. Pan, et al. (2003). "Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization." J Neurosci **23**(16): 6627-37.
s in motor neuron axons in SMA.

Zou, J., F. Barahmand-Pour, et al. (2004). "Survival of motor neuron protein SMN interacts with transcription corepressor mSin3A." J Biol Chem.