Chapter 7: Perspective
Chapter 7: Perspective

Discovery of the new vertebrate SPRN gene family has implications for better understanding of both normal and pathogenic roles of the human/mammalian PRNP gene. Evolutionary trajectories of the two genes differ. Whereas the evolution of SPRN is conservative, the PRNP gene has more relaxed evolutionary constraints. This evolutionary dialectic is the key to understanding these two genes better.

A better understanding of PRNP emerged from my cloning and analysis of PRNP in the Australian mammal tammar wallaby. The new insights that emerged about the evolution and function of mammalian PRNP best supported the signal transduction hypothesis. This hypothesis states that the PrP<sup>C</sup> is a signal transduction protein, which mediates the signalling pathways in both neuronal and non-neuronal cells.

7.1 Shadow of Prion Protein: Discovery of New Vertebrate Gene

I compiled evidence for the new vertebrate SPRN gene, based on the direct evidence, gene homology and ab initio gene prediction in mammals and fish. Such a manual analysis of the data from public databases is an example of targeted, conceptually driven, low-cost gene discovery, applicable also to discovery of other genes. The zebrafish sprn cDNA sequence sequenced by Professor Tatjana Simonic’s group was critical for the discovery.

The discovery of SPRN expanded the number of mammalian paralogues of the PRNP gene. There are only two functional, protein-coding human PRNP paralogues known. The first is PRND that may have diverged significantly from PRNP, and the second is SPRN.

Similarity of overall protein sequence features and the presence of the conserved middle hydrophobic region may suggest functional relationship between Sho and PrP. Of particular note here is the fact that Sho is the first (and at present the only) known
human/mammalian PrP paralogue that contains the middle hydrophobic region, critical for both normal function and pathogenic potential of the mammalian PrP. This conservation may define at least two functional and/or structural characteristics for Sho. Firstly, the role of the conserved region could be the same in two proteins. It is unknown at present how the hydrophobic region contributes to the PrP function and pathogenic characteristics, but the structural flexibility of this conserved region and its conformational heterogeneity is a hallmark of prion protein. Perhaps future comparative analyses of Sho and PrP may clarify structural and/or functional role of the hydrophobic region. Secondly, the presence of the conserved hydrophobic sequence in Sho could make it as conformationally flexible and as structurally promiscuous as is PrP. The Sho structure and its characteristics have to be determined to clarify this possibility. If this indeed is the case, than Sho may be a new human/mammalian protein that exhibits conformational plasticity, enabling better understanding of this structural phenomenon. Could Sho also adopt pathogenic conformation?

Mammalian Sho, a protein of 100-odd amino acids in its mature form, could be also very useful for investigating the most basic steps in protein folding.

SPRN expression seems to be confined to brain. However, RT-PCR experiments showed the faint bands also found in gut and lungs. Thus, SPRN expression could be more similar to that of PRNP than assumed initially (Premzl et al., 2003), and this requires careful re-examination of patterns of the SPRN mRNA and Sho expression.

The mammalian SPRN cDNAs and ESTs were isolated mostly from hippocampus, the brain region in which memory is formed. Perhaps this is the brain region in which the SPRN expression is highest, enabling isolation of the cDNAs and ESTs. Of particular note here is that PRNP in mammalian brain has highest expression exactly in the hippocampus. Therefore, it is likely that the two genes contribute to the hippocampal metabolism, and this brain region appears to be a logical choice for future comparative studies of the functions of two genes.
7.2. Evolution of *SPRN* and *PRNP*

By applying this approach for manual gene discovery, I defined the whole new family of *SPRN* vertebrate genes and expanded the number of known fish genes related to the *PRNP*. The total number of new ORFs that I discovered in public databases is 11.

I conducted comparative genomic analysis of the two vertebrate gene families, and determined different evolutionary trajectories for mammalian *SPRN* and *PRNP*. This lead to the new paradigm: the vertebrate *SPRN* gene family is more conserved than the vertebrate *PRNP* gene family. This evolutionary dialectic suggests that the gene that is redundant with *Prnp* could be more conserved *Sprn*, substituting for the loss of dispensable *Prnp* in the knock-out mice. The conserved protein features argue in favour of this hypothesis, but the gene and protein expression patterns will have to be similar as well.

The conservative evolution of *SPRN* compared with less constrained evolution of *PRNP* indicates more prominent, and perhaps more important, function for the *SPRN*, and the pathogenic potential of *PRNP* could be a consequence of the relaxed evolutionary constraints. This challenges a common viewpoint on “a singular protein that is as highly conserved among species as PrP, from turtles to frogs, fish, and humans” (Aguzzi and Polymenidou, 2004). I propose therefore this paradigmatic shift for understanding evolution, function and disfunction of the two genes.

To address the function of *SPRN*, the *Sprn* knock-out mice will have to be constructed. Knock-out of the conserved *Sprn* gene in mice is likely to produce phenotype that will provide insight to its function. However, due to the conservation of *SPRN*, the gene disruption could have detrimental effect on the development. If *Sprn* knock-out mice will be viable, they should be crossed with the *Prnp* knock-out mice. If an additional new phenotype appears in the double *Sprn/Prnp* knock-out mice, redundancy between the two genes will be demonstrated.
The conserved elements within the promoter and intron of SPRN identified by phylogenetic footprinting are potential regulatory elements. At this stage, these should be regarded as hypothetical, and the question whether they are really functional must be addressed experimentally.

One possibility is that the SPRN is involved in the ER stress-response, together with the other molecular chaperones and enzymes. I am tempted to speculate here that this may point to a link between the Sho and the pathogenic transformation of prion protein: Sho could be the protein X, molecular chaperone involved in the transformation of PrP. Protein X is expected to be expressed in the tissues where PrP transformation occurs, to be present in the same cellular compartment where PrP transformation takes place, and to interact with PrP* (Chapter 1.4.2). SPRN and PRNP are both expressed in brain where PrPSc accumulates, and Sho was predicted to be the extracellular, glycosylated and GPI-anchored protein, as is PrPc. Sho contains the critical conserved hydrophobic region that is involved in heterodimerization of the two PrP isoforms. Therefore, Sho could have potential to interact with PrP. If Sho indeed participates in the pathogenic transformation of PrP as the essential factor, the Sprn knock-out mice will have to be resistant to prion infection. This hypothesis could also be studied in the cell lines in which prions replicate: inhibiting expression of Sho should inhibit prion replication. Perhaps Sho could also assist the PrP transformation in vitro. Discovery of the protein X will prove the template assisted polymerisation model (Chapter 1.4.2) for prion propagation.

7.3 Utility of Tammar Wallaby Genomic Sequence in Comparative Genomic Analysis of PRNP

Marsupial DNA is interesting for comparative genomics due to the unique position of marsupial branch on the vertebrate evolutionary tree.

I showed the utility of the marsupial sequence for the cross-species analysis of the human disease-related PRNP gene. Using the tammar wallaby genomic sequence as the outgroup in comparison with the genomic sequences from species in which prion
diseases occur naturally (human, bovine, ovine) or experimentally (mouse), I gained new insights into the function and evolution of PRNP gene. This is one more argument in favour of the Kangaroo Genome Project.

The marsupial sequence enabled definition of the regions in PRNP that are conserved mammalian-wide, across 180 million years. These are potential regulatory elements determining PRNP gene expression and posttranscriptional activity of mRNA. The next step in proving their functional significance is experimental analysis.

By providing two new marsupial PrP sequences, those from tammar wallaby and Brazilian opossum, and by their comparisons with PrPs from other vertebrates, I analysed the patterns of the evolution of PrP regions. Whereas the variant repeats show lineage-specific conservation, the essential hydrophobic region shows highest conservation. Thus, whereas the conserved hydrophobic region is likely to have the same functional/structural role in all members of the vertebrate PrP protein family (and in the vertebrate Sho family as well), the repeats could have evolved the lineage-specific functions.

Marsupial and eutherian PrPs are conserved, suggesting that marsupial PrPs could have the same pathogenic potential as eutherian PrPs. One way to approach this possibility is construction of transgenic mice expressing marsupial PrP\(^C\). These mice could be inoculated with various prion strains. Alternatively, marsupials may be directly inoculated with prions.

The comparative genomic analysis of PRNP complemented current knowledge about its function, supporting best the signal transduction hypothesis. Understanding of the PRNP function is perhaps best within this framework. Therefore, PrP\(^C\) may be the signal transduction protein. The signalling pathways mediated by PrP\(^C\) will contribute to the cell-cell interactions, and will act anti-apoptotically. Yet, in the heterogenous set of cells expressing PRNP, this will contribute to various phenotypes, such as the synaptic plasticity and lymphocyte activation.
7.4. Biomedicine Tomorrow

The number of genomes sequenced increases almost on a weekly basis. Refinement of the current methods for DNA sequencing, and development of new sequencing methods, will further foster the momentum of genomic biology. This progress has already boosted development of the large-scale methods for analysis of genes and proteins, such as microarray and SAGE, and proteomic methods. Development of better computational methods for analysis and mining of the already overwhelming data is *conditio sine qua non*. Progress of biomedicine in the XXI century will be cultivated by integration of the experimental and computational reasoning. At this stage, available genome and transcriptome dataset allows elegant, rapid and cheap discovery of new genes and regulatory elements. The discovery of *SPRN in silico* and comparative genomic analysis of *PRNP* are such examples.