Chapter 7

Discussion

This thesis has described the analysis of 5' CGI methylation states in a large number of X-linked genes in three species, (human, mouse, and opossum) belonging to the two extant mammalian groups that use X chromosome inactivation for dosage compensation.

A dynamic landscape of X chromosome CGI methylation patterns has been revealed in the females of these three species. On one extreme, the vast majority of CGIs assayed in mouse were found to be methylated in mouse females following the RPMA analysis (Chapter 3). Single base pair resolution methylation maps generated by bisulphite sequencing revealed a heterogeneous pattern of methylation within individual islands, but most of the methylated island molecules tend to have more than half of the CpG dinucleotides methylated (Chapter 4). Only three CGIs were hypomethylated in females, where the lack of methylation was almost complete, indistinguishable from the methylation pattern seen in males. A different pattern was seen in the human CGIs. In the RPMA analysis, a much greater proportion of human islands were found to be hypomethylated

in females as well as in males (Chapter 3). The difference between human and mouse is also obvious in the levels of methylation in individual CGIs. As shown by bisulphite sequencing, the human islands were never methylated to the same extent as that of their mouse homologues (Chapter 4). In addition, intermediate levels of methylation were found in a good proportion of human CGIs, but never in mouse (Chapters 3 and 4). At the other end of the spectrum, the first large scale methylation analysis of marsupial CGIs to date revealed an overall hypomethylation of 5' CGIs on the opossum X chromosome.

Methylation of 5' CGIs is rarely seen on the autosomes, but very common on eutherian X chromosomes, presumably playing a role in maintaining the silencing state of genes on the inactive X (Kaslow and Migeon, 1987). One aim of this thesis was to test whether CGI methylation serves as a good indicator for a gene's XCI status, as previous research has established a correlation between CGI methylation and gene silencing from XCI (Tribioli *et al.*, 1992). The work presented in this thesis has greatly extended this correlation (Chapter 3). For the human genes, a comparison of their CGI methylation states with their XCI status determined using the somatic cell hybrid system (Carrel and Willard, 2005) demonstrated a very strong correlation between CGI methylation and gene silencing (49/53 in agreement). For the mouse genes with known XCI status, their CGI methylation states enabled perfect prediction of their XCI status.

The findings in this thesis have also provided novel XCI information for a great number of genes. For eleven human genes that could not be assayed in the previously published human XCI profile (Carrel and Willard, 2005), an XCI state prediction was made based on their 5' methylation profile in this study. In the case of mouse, XCI information uncovered for most genes in this thesis is novel.

The differences in CGI methylation between human and mouse correlate with the extent of escape from XCI seen in these species, providing some clue to the mechanism of escape. It is currently not clear if escapees occur from failure to establish or failure to maintain inactivation. The latter scenario is supported by experimental demonstration that a mouse escapee gene, *Smcx*, is inactivated in early embryonic development followed by re-activation (Lingenfelter *et al.*, 1998). The progressive loss of CGI methylation has been proposed as a possible reason for re-activation (Disteche *et al.*, 2002). The findings in this thesis are also in support of this hypothesis. In mouse, the vast majority of CGIs studied were hypermethylated, possibly contributing towards more complete XCI than in human. In human, a good proportion of CGIs studied showed low level or hypomethylation, in agreement with the extent of escape from XCI. The lower level of methylation in human than mouse is consistent with the suggestion that escape from XCI is associated with a failure to maintain the inactive state.

Very little is known about the inactivation status of X-linked genes in marsupials. Prior to this study, allozyme studies of a small number of genes showed that these genes are less stably inactivated in marsupials than in human and mouse (reviewed in Cooper *et al.*, 1993). The less stable XCI in marsupials has been linked to lack of CGI methylation (Kaslow and Migeon, 1987), which was confirmed for *G6PD* in several marsupial species (Kaslow and Migeon, 1987; Loebel and Johnston, 1996). Findings presented in this thesis have greatly expanded our knowledge about CGI methylation on the marsupial X chromosome. CGI methylation analysis was performed for 36 novel genes, comprising almost a third of all assayable genes on the X chromosome of a model marsupial, the grey short-tailed opossum (*M. domestica*). Four-fifths of them showed indistinguishable patterns

of hypomethylation between female and male. If escaping from XCI is a result of failure to maintain the inactive state, as speculated above, a high level of escape from XCI might be anticipated in marsupials. However, this may not be the case as there might be other maintenance mechanisms in place. For example, the CGI of *G6PD* is not methylated in the common wallaroo (Loebel and Johnston, 1996), but the paternal allele is completely repressed in this species, and re-activation was only seen in culture cell lines (Johnston *et al.*, 1978; Richardson *et al.*, 1971).

This thesis has also demonstrated for the first time existence of female-specific methylation on a marsupial X chromosome. Based on limited evidence, it is believed that methylation is not part of the molecular mechanism of marsupial XCI (Graves, 1996, 2006), but a fifth of all genes assayed in this study showed some degree of differential methylation between the two sexes (Chapter 5). Bisulphite sequencing analysis of seven CGIs in four different tissues confirmed the presence of female-specific methylation, and also revealed conservation of methylation patterns across different tissues. It is not clear whether such degree of methylation may bear any functional significance in the absence of expression data, but interestingly, the CGI methylation states of three genes uncovered in this study correlate with their expression from the inactive X in a related species, the Virginian opossum, in the same fashion that CGI methylation profiles of human and mouse genes correlate with their XCI status. It leaves open the possibility that DNA methylation may play a role in maintaining XCI for a subset of marsupial genes.

7.1 Future directions

Many of the features associated with XCI in the eutherian mammals are absent from the XCI system in the marsupial mammals, which is thought to represent a more ancestral mechanism. Studying the details of XCI in marsupials is therefore of key importance to help our understanding of the evolutionary origin of this unique form of dosage compensation in therian mammals. Contrary to the common view that CGI methylation is absent from the marsupial X chromosome, this thesis has described the first evidence of female-specific CGI methylation on a marsupial X (Chapter 5). It is of immediate interest to investigate whether this extent of methylation may have any functional significance in regulation of marsupial XCI. Expression studies could be carried out in parallel with methylation studies to see if level of methylation is correlated with expression from the inactive X. The imprinted XCI in marsupial enables easy identification of gene expression from the active or inactive X chromosome, which can only be achieved in human and mouse using experimental systems that don't necessarily represent situations in normal cells. This thesis has also shown a stable methylation pattern across different tissues in the same individual. This could be extended to comparison of methylation profiles of the same CGIs among different individuals. If such methylation bears functional significance, it is more likely to be preserved in multiple individuals. In addition to DNA methylation, it is worthy to search for other factors that may affect stability of XCI in marsupials, for example histone modifications.

In the eutherian domain, little is known about the XCI status of X-linked genes in most eutherian mammals, as the expression-based techniques used to

assess XCI in human and mouse are mostly difficult to extend to other species. In order to understand how the landscape of gene silencing has evolved, it is necessary to study a group of genes' XCI status in a number of species on different branches of the mammalian evolutionary tree. Moreover, multi-species comparison will give an opportunity to define more precisely the characteristics that influence the XCI status of a gene or region. This thesis has demonstrated an indirect method to predict rapidly the XCI status of a large number of genes based on their 5' CpG island methylation status (Chapter 3), and therefore provided a platform for constructing a comparative XCI profile in mammals. Recent advances in sequencing technology have greatly promoted the growth of genomic sequencing projects. During the course of this project, full genomic sequence has become available for an increasing number of mammals, covering a great total branch length of the evolutionary tree, and providing a good representation of the diversity of mammalian species (Figure 7.1). In this thesis, the S3 region, containing both inactivated and escapee genes, where different XCI patterns are known between different species, was studied in human and mouse (Chapters 3 and 4). This work can be extended to other mammalian species to make predictions of the ancestral XCI status for a number of genes. As demonstrated by Jegalian and Page's study of the XCI status of four genes in 18 species (1999), generation of XCI profiles for even a small number of genes in these species will greatly expand our understanding of the evolution of XCI.

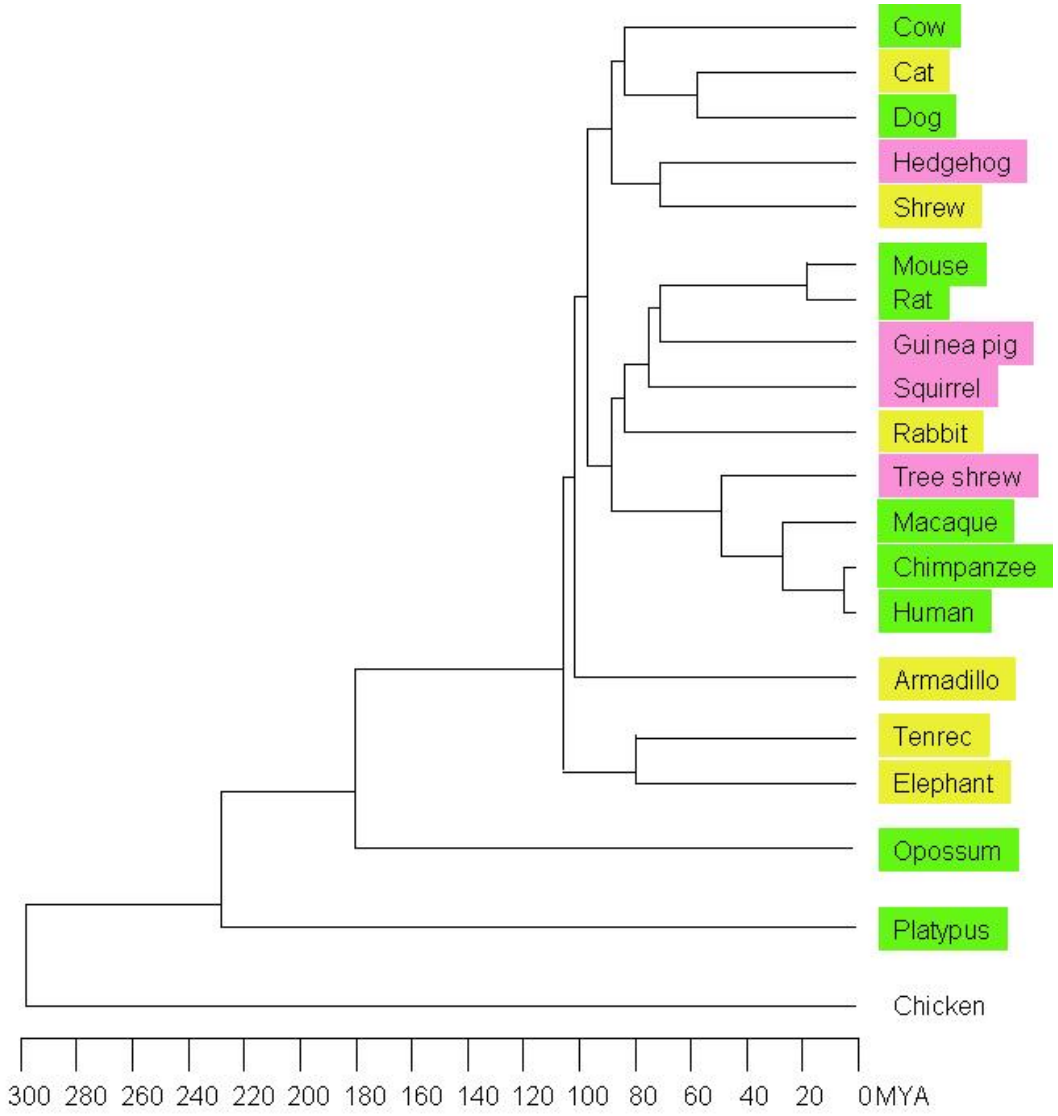

Figure 7.1: Phylogeny and time scale of mammalian species with available sequence information. Chicken is outgroup. Species highlighted in green have completed medium to deep coverage sequences; species in yellow have completed low coverage sequences; and sequencing are in process towards low-coverage for those in pink.

On the other dimension, more in-depth methylation analysis could be carried out on the basis of current findings. As many X-linked genes are not associated with a predicted CGI yet undergo stable XCI, it is of great interest to know whether DNA methylation still plays a role in maintaining their silencing. To this end, oligo-array-based technology could be applied to the entire human X chromosome. It may reveal CpG-rich domains at 5' region of genes with methylation patterns typically associated with inactivated or escapee genes as established in this study. Most recently, bisulphite sequencing has been combined with advanced sequencing technology to yield a single base resolution map of the Arabidopsis methylome (Lister *et al.*, 2008). With further development of technology, similar details may be revealed for the human and other X chromosomes and enable more thorough understanding of the dynamics of DNA methylation in X chromosome inactivation.

7.2 Conclusions

This thesis has shown the potential of DNA methylation analysis to further our understanding of X chromosome inactivation. Described here is the most extensive methylation analysis to date of 5' CGIs associated with genes subject to and escaping from XCI. The generation of fine methylation maps has revealed a great extent of variation in DNA methylation, on the levels of molecules, islands, and species. Comparative study of three species with different extent of escape has provided support to the nature of escape being a lack of maintenance of XCI. It has also provided the first evidence of female-specific methylation on any marsupial X chromosome. This study has formed a basis of more extensive multi-species studies to explore the evolutionary landscape of X chromosome inactivation.