Supplementary Note

Meunier et al : Birth and functional evolution of mammalian microRNA genes

RNA-seq read mapping and genome quality. The proportion of reads used for miRNA gene detection that mapped perfectly on their corresponding genomes ranged from 0.62 to 0.77 in each sample, and was not biased towards a species or a tissue (two-way anova test, $P > 0.5$, see sample description in Supplementary Table 1). Thus, we detected no significant bias that might have been caused by heterogeneous quality in the genome sequences we used or in our RNA-sequencing data.

Proportion of miRNA reads across tissue samples. The proportions of mapped reads with 15-23 nucleotides (nt) that corresponded to miRNA genes detected by our procedure was very high in somatic tissues (0.93 +/– 0.04) and varied very little among samples (two-way ANOVA, $P > 0.2$). However, they were significantly lower in the testis (0.57 +/– 0.27; Mann-Whitney *U* test, $P < 0.0001$), especially in platypus and chicken. In all the testis samples, the majority of the reads corresponded to piRNAs, small RNAs that are slightly longer (24-31 nt) than miRNAs in humans (Kim et al. 2009) (Supplementary Figure 1). The length of miRNAs and piRNAs seems to overlap in platypus and chicken (Supplementary Figure 1), which might explain the low proportions of miRNA reads in the testis of these species. Generally, we note that only reads corresponding to miRNAs and piRNAs were detected in significant amounts in our samples, contrary to other studies (Cole et al. 2009).

We therefore assessed the influence of variable testis read sampling on miRNA detection in each species. To do so, we reran our detection procedure using 44 million 15-23 nt mapped reads sampled from the somatic tissue and adding a variable amount of reads sampled from the testis. Using between 1.5 and 3 million reads sampled from the testis did not result in any notable variations (<5%) in the number of detected miRNAs (Supplementary Figure 2), which suggests that the number of available testis reads does not limit miRNA detection in any species and can therefore not explain the differences in the number of miRNA genes we detect for each species (Table 1).

References

Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JWS, Green PJ, Barton GJ, Hutvagner G. 2009. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* **15** 2147-2160.

Kim VN, Han J, Siomi MC. 2009. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**: 126-139.