

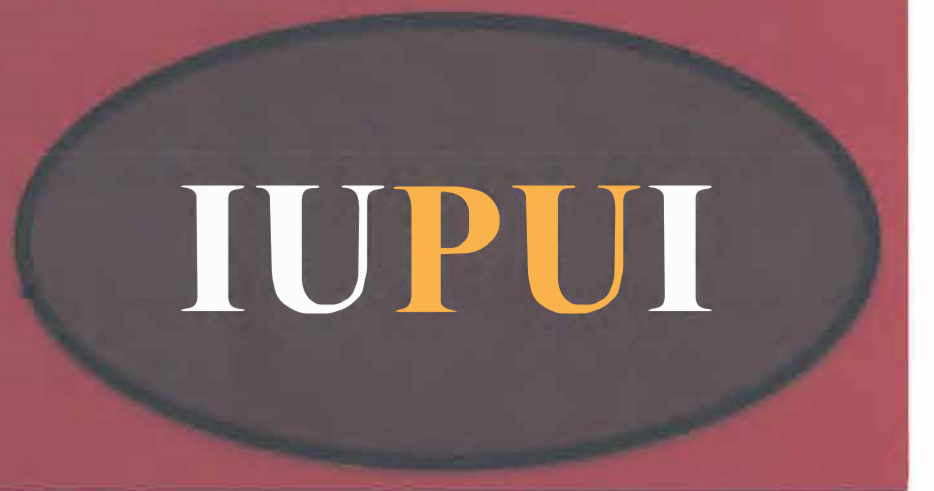
Evolutionary dynamics of RNA-binding proteins expression levels in mammals

Abhijit Badve¹, Vishal Sarsani¹, Yaseswini Neelamraju¹, Sarath Chandra Janga^{1, 2, 3,*}

¹ Department of Biohealth Informatics, School of Informatics and Computing, Indiana University Purdue University, Indianapolis

² Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, HITS, Indianapolis

³ Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana



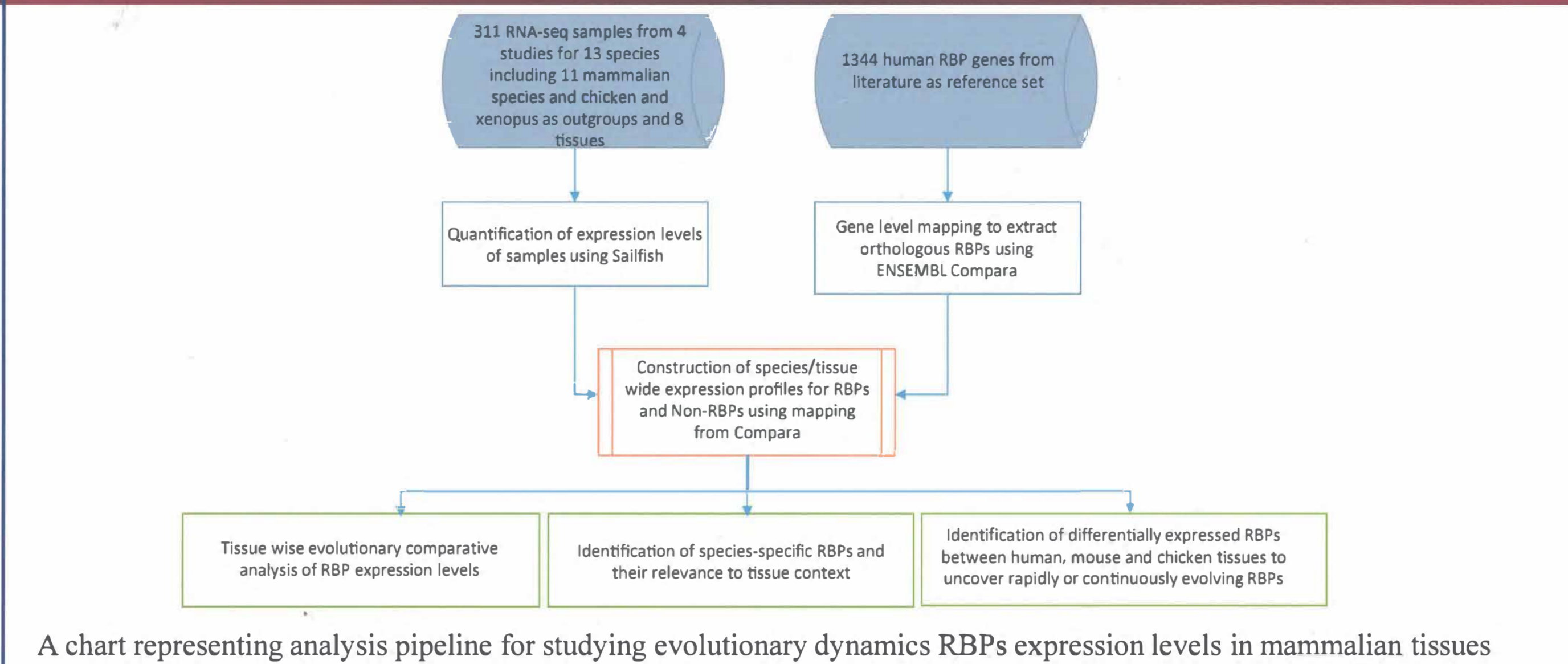
Abstract

RNA binding proteins (RBPs) play important roles in controlling the post-transcriptional fate of RNA molecules, yet their evolutionary dynamics remains largely unknown. As expression profiles of genes encoding for RBPs can yield insights about their evolutionary trajectories on the post-transcriptional regulatory networks across species, we performed a comparative analyses of RBP expression profiles across 8 tissues (brain, cerebellum, heart, lung, liver, lung, skeletal muscle, testis) in 11 mammals (human, chimpanzee, gorilla, orangutan, macaque, rat, mouse, platypus, opossum, cow) and chicken & frog (evolutionary outgroups). Noticeably, orthologous gene expression profiles suggest a significantly higher expression level for RBPs than their non-RBP gene counterparts - which include other protein-coding and non-coding genes, across all the mammalian tissues studied here. This trend is significant irrespective of the tissue and species being compared, though RBP gene expression distribution patterns were found to be generally diverse in nature. Our analysis also shows that RBPs are expressed at a significantly lower level in human and mouse tissues compared to their expression levels in equivalent tissues in other mammals chimpanzee, orangutan, rat, etc. which are all likely exposed to diverse natural habitats and ecological settings compared to more stable ecological environment humans and mice might have been exposed, thus reducing the need for complex and extensive post-transcriptional control. Further analysis of the similarity of orthologous RBP expression profiles between all pairs of tissue-mammal combinations clearly showed the grouping of RBP expression profiles across tissues in a given mammal, in contrast to the clustering of expression profiles for non-RBPs, which frequently grouped equivalent tissues across diverse mammalian species together, suggesting a significant evolution of RBPs expression after speciation events. Calculation of species specificity indices (SSIs) for RBPs across various tissues, to identify those that exhibited restricted expression to few mammals, revealed that about 30% of the RBPs are species-specific in at least one tissue studied here, with lung, liver, kidney & testis exhibiting a significantly higher proportion of specie-specifically expressed RBPs. We conducted a differential expression analysis of RBPs in human, mouse and chicken tissues to study the evolution of expression levels in recently evolved species i.e. humans and mice than evolutionarily distant specie i.e. chicken. We identified more than 50% of the orthologous RBPs to be differentially expressed in at-least one tissue compared between human and mouse but not so between human and the outgroup i.e. chicken in which RBP expression levels are relatively conserved. Among the studied tissues brain, liver and kidney showed a higher fraction of differentially expressed RBPs, which may suggest hyper regulatory activities by RBPs in these tissues with species evolution. Overall, this study forms a foundation for understanding the evolution of expression levels of RBPs in mammals, facilitating a snapshot of the wiring patterns of post-transcriptional regulatory networks in mammalian genomes.

Materials

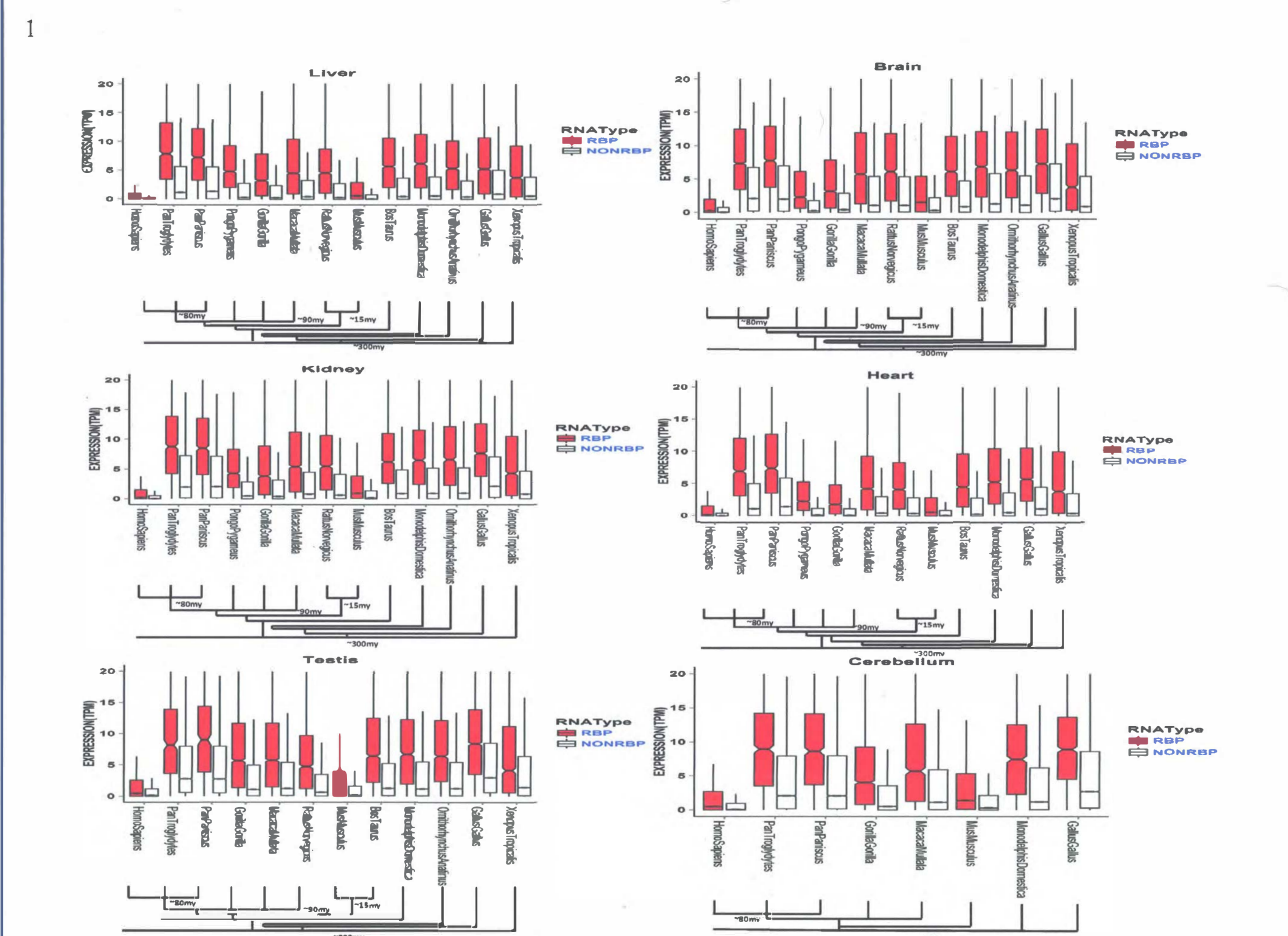
We used a reference set of 1344 human genes encoding for RBPs collected from current literature [2]. RNA-seq data-sets were collected from different studies [1,3]. In all, 311 RNA-seq samples from 4 studies for 13 species including 11 mammalian species and two evolutionary outgroup samples chicken and xenopus across 7 tissues were used in this study.

Methods

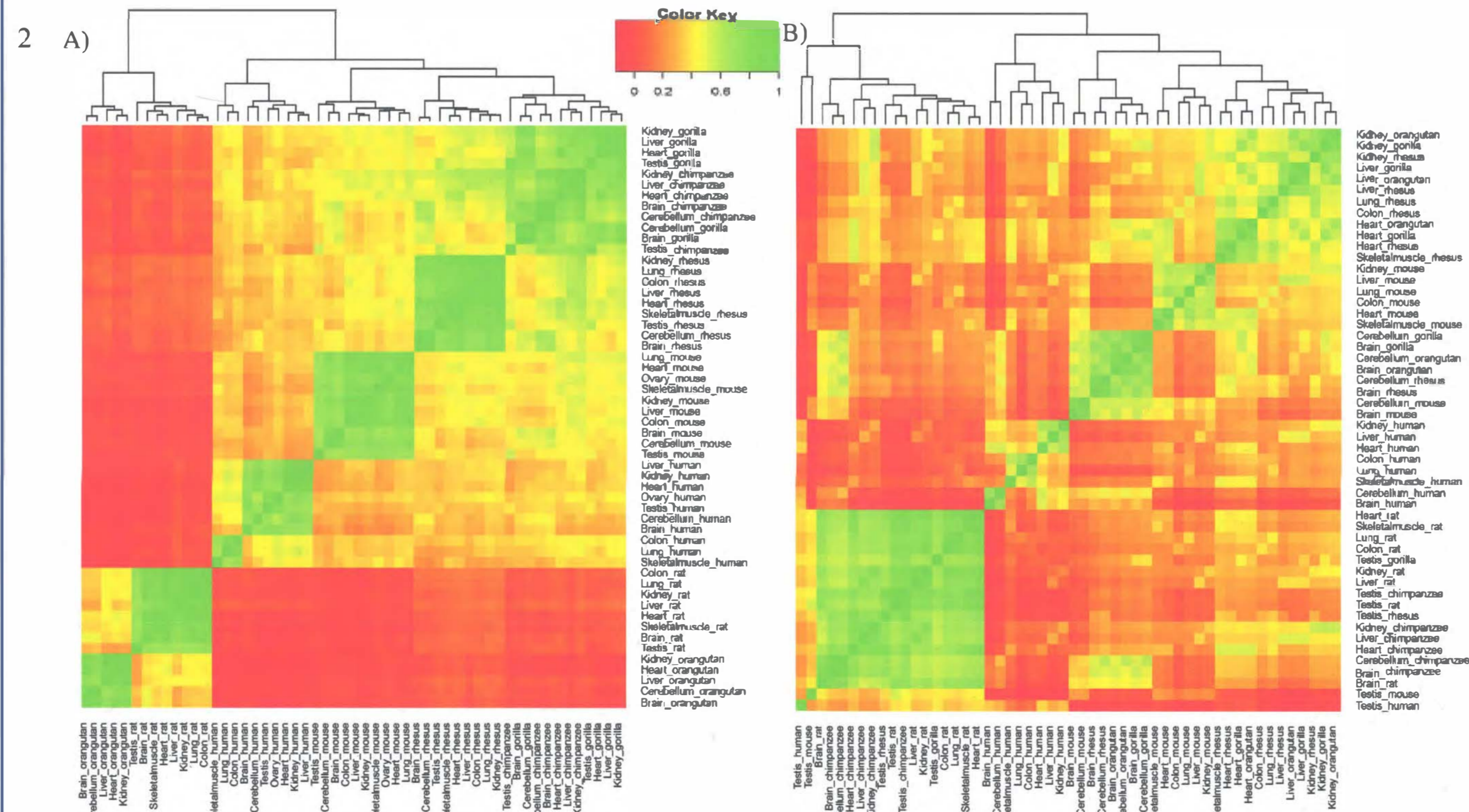


A chart representing analysis pipeline for studying evolutionary dynamics RBPs expression levels in mammalian tissues

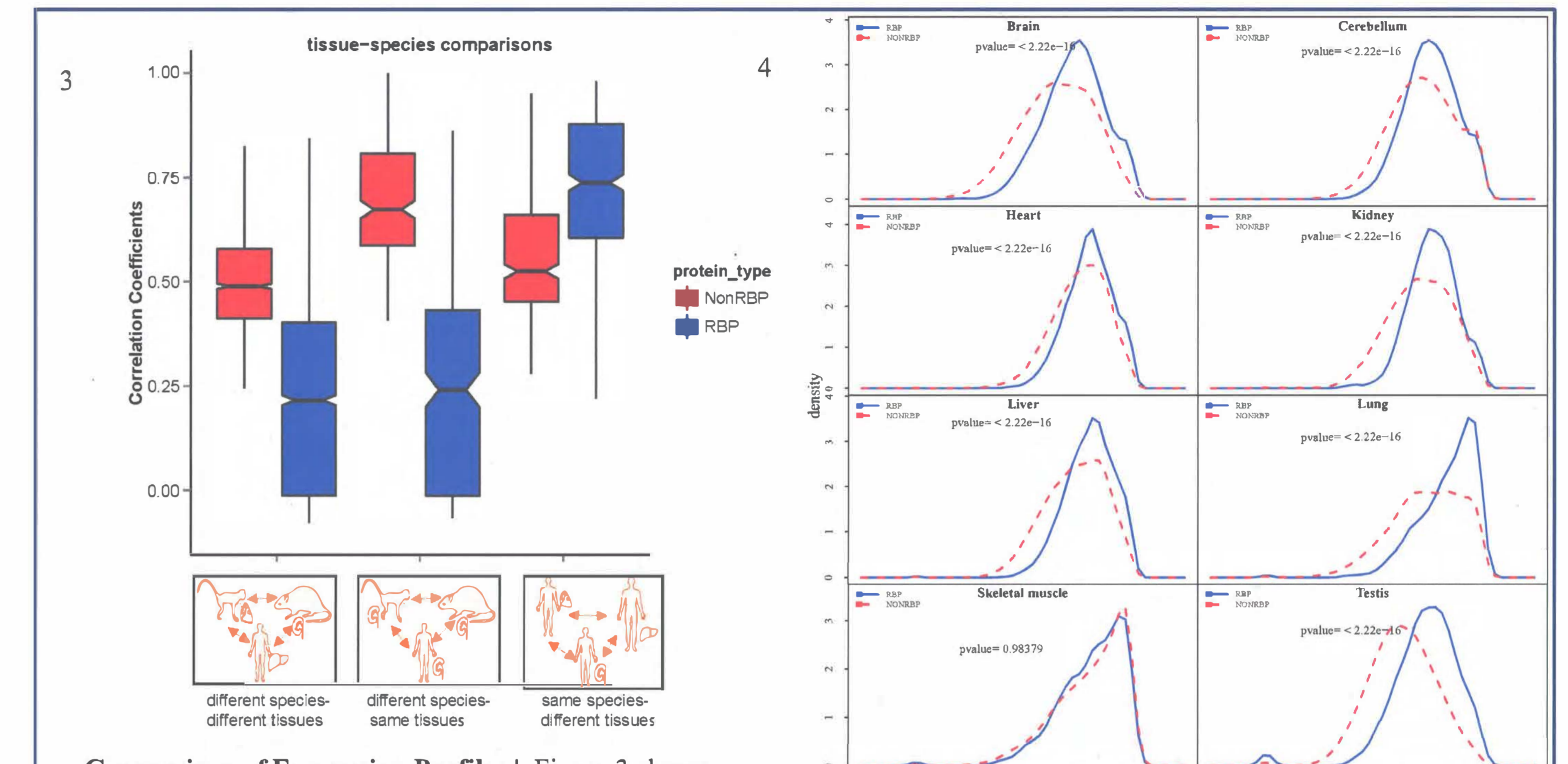
Results and Discussion



Expression Profiles | Figure 1: Multi-panel boxplots showing the expression level (TPM) comparisons between orthologous RBPs vs Non-RBPs across 6 tissues studied here (KS test p-values are significant at $2.2e^{-16}$ for each of the 6 tissues)

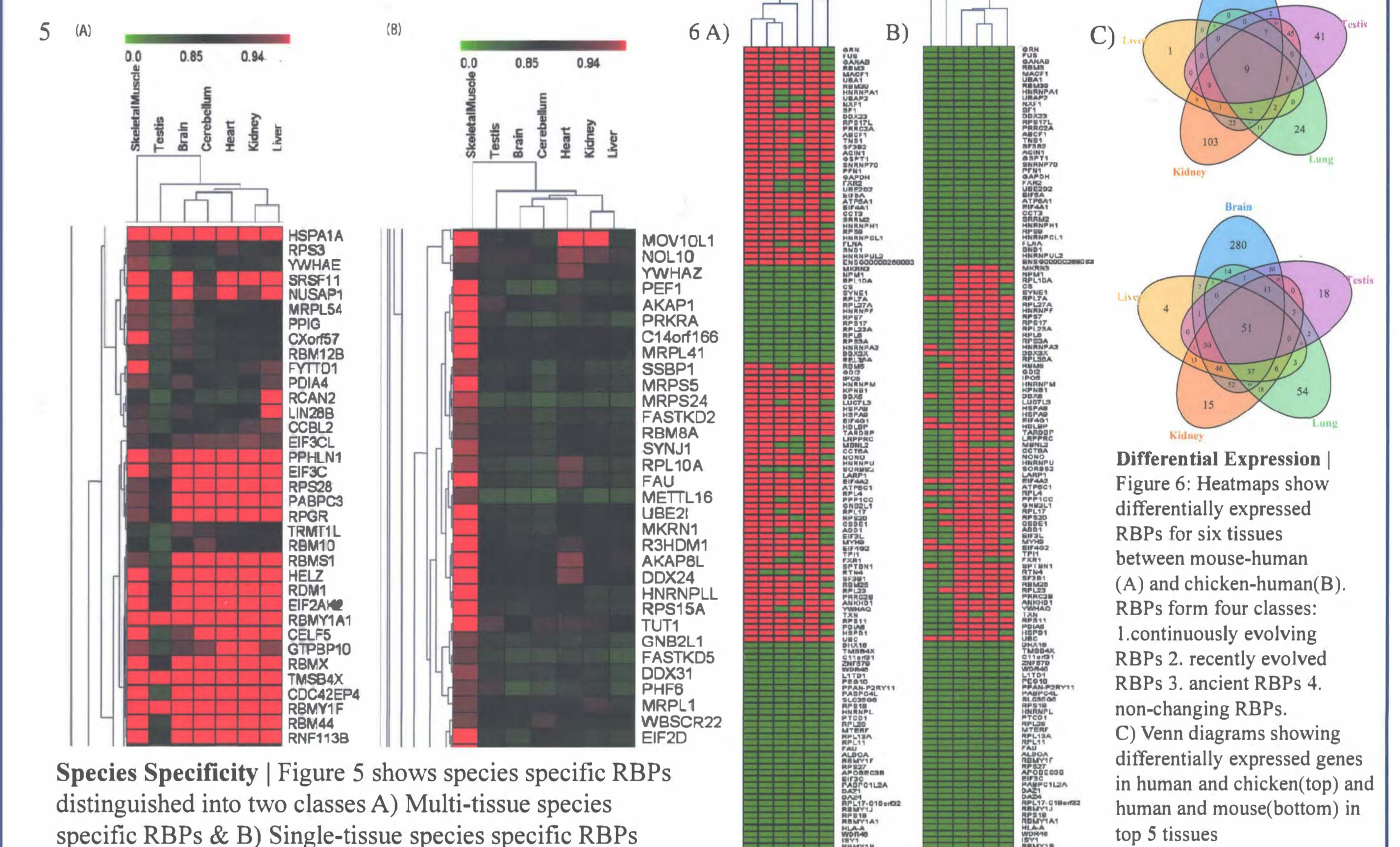


Clustering of expression correlations for RBPs and Non-RBPs | Figure 2: Heatmap shows clustering based on spearman correlation coefficients calculated from expression profiles of each tissue across species for RBPs (A) and Non-RBPs (B)



Comparison of Expression Profiles | Figure 3 shows comparisons of expression correlation coefficients for specie-tissue combinations classified into 3 mutually exclusive sets A) of different species and different tissues B) of different species and same tissues and C) of same species and different tissues for RBPs and non-RBPs .

Species Specificity | Figure 4 shows pairwise density distributions of species specificity indices (SSIs) for RBPs vs non-RBPs across various tissues under study



Species Specificity | Figure 5 shows species specific RBPs distinguished into two classes A) Multi-tissue species specific RBPs & B) Single-tissue species specific RBPs

Differential Expression | Figure 6: Heatmaps show differentially expressed RBPs for six tissues between mouse-human (A) and chicken-human (B). RBPs form four classes: 1. continuously evolving RBPs 2. recently evolved RBPs 3. ancient RBPs 4. non-changing RBPs. C) Venn diagrams showing differentially expressed genes in human and chicken (top) and human and mouse (bottom) in top 5 tissues

Conclusion

1. Orthologous gene expression profiles suggest a significantly higher expression level for RBPs than non-RBPs.
2. Expression profiles of orthologous RBPs across tissue-specie combinations reveal specie-wise clustering of RBPs while non RBPs cluster tissue-wise
3. About 30% of the RBP repertoire is species specific in at-least one tissue studied here, with several tissues exhibiting a significantly higher proportion of specie-specifically expressed RBPs .
4. Differential expression analysis of RBPs between human, mouse and outgroup i.e. chicken classify RBPs in distinct evolutionary groups as ancient (5%), recently evolving (12%), continuously evolving (8%) and non-changing (75%).

References

1. Necsulea, A., et al. (2014). "The evolution of lncRNA repertoires and expression patterns in tetrapods." Nature 505(7485): 635-640.
2. Kechavarzi, B. and S. C. Janga (2014). "Dissecting the expression landscape of RNA-binding proteins in human cancers." Genome Biol 15(1): R14.
3. Brawand, D., et al. (2011). "The evolution of gene expression levels in mammalian organs." Nature 478(7369): 343-348.