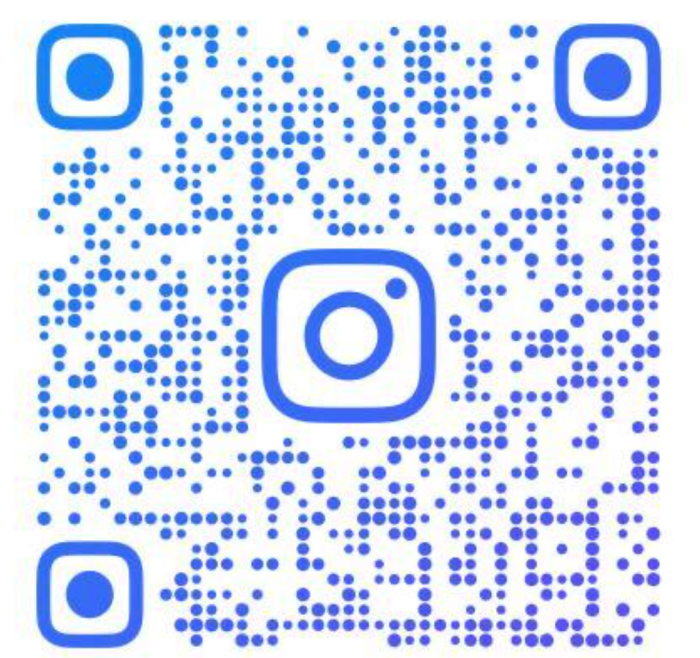# PHYLOBONE: A NEW RESOURCE FOR OSTEOPOROSIS RESEARCH

Margalida Fontcuberta-Rigo[1,2], Miho Nakamura[1,3,4] and Pere Puigbò[1,2,5]

[1]University of Turku; [2]University Rovira i Virgili; [3]Tokyo Medical and Dental University;
[4]Tohoku University; [5]Eurecat, Technology Center of Catalonia

@PHYLOBONE

INTRODUCTION

Osteoporosis is a major disease in our society, causing nine million fractures per year. It mostly affects middle age and elderly population worldwide.

Our goal is to create a resource for osteoporosis research based on the identification of bone extracellular matrix (ECM) proteins in human and model organisms.
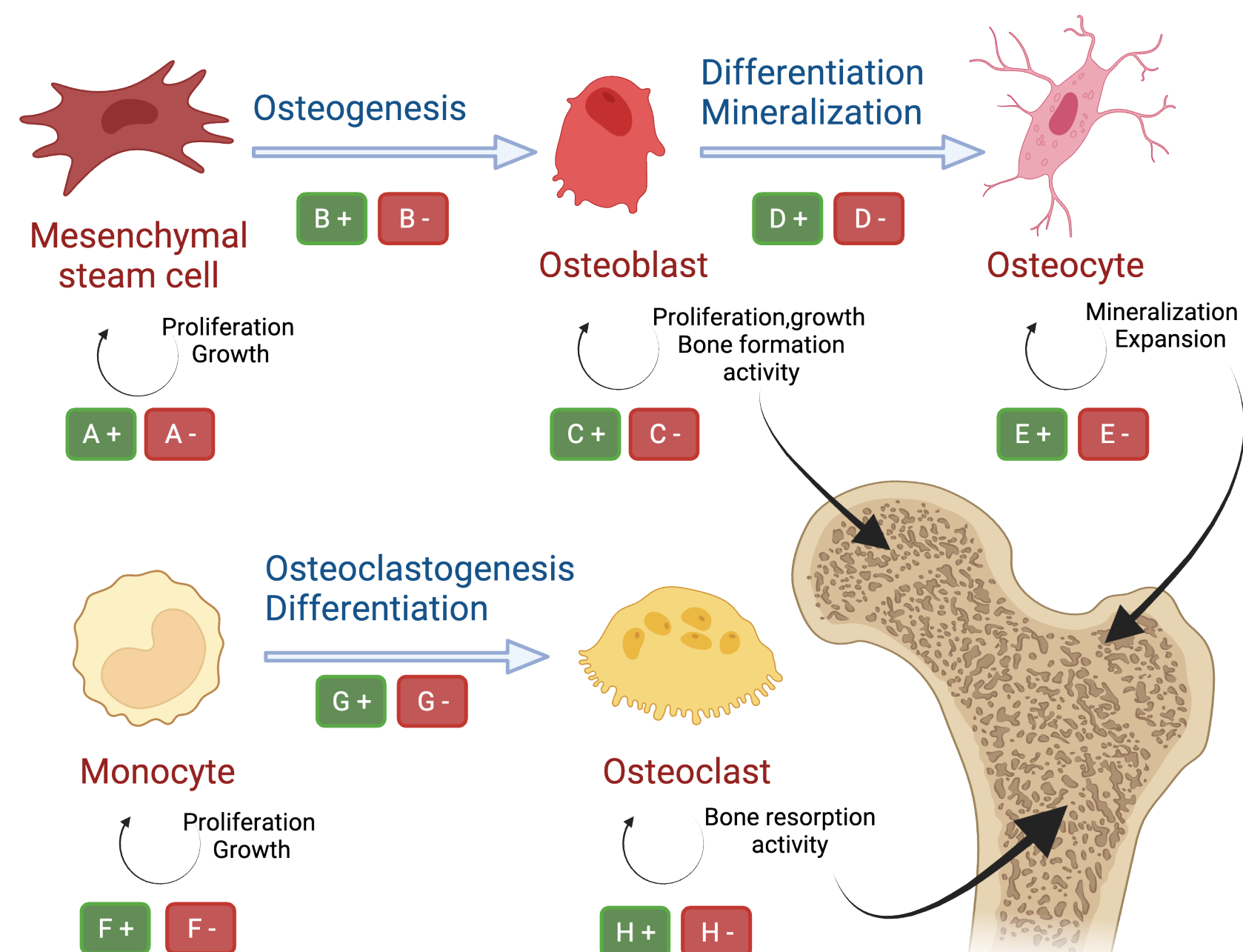

Figure 1. ECM proteins are involved in the proliferation and differentiation processes of bone cell types.

METHODS

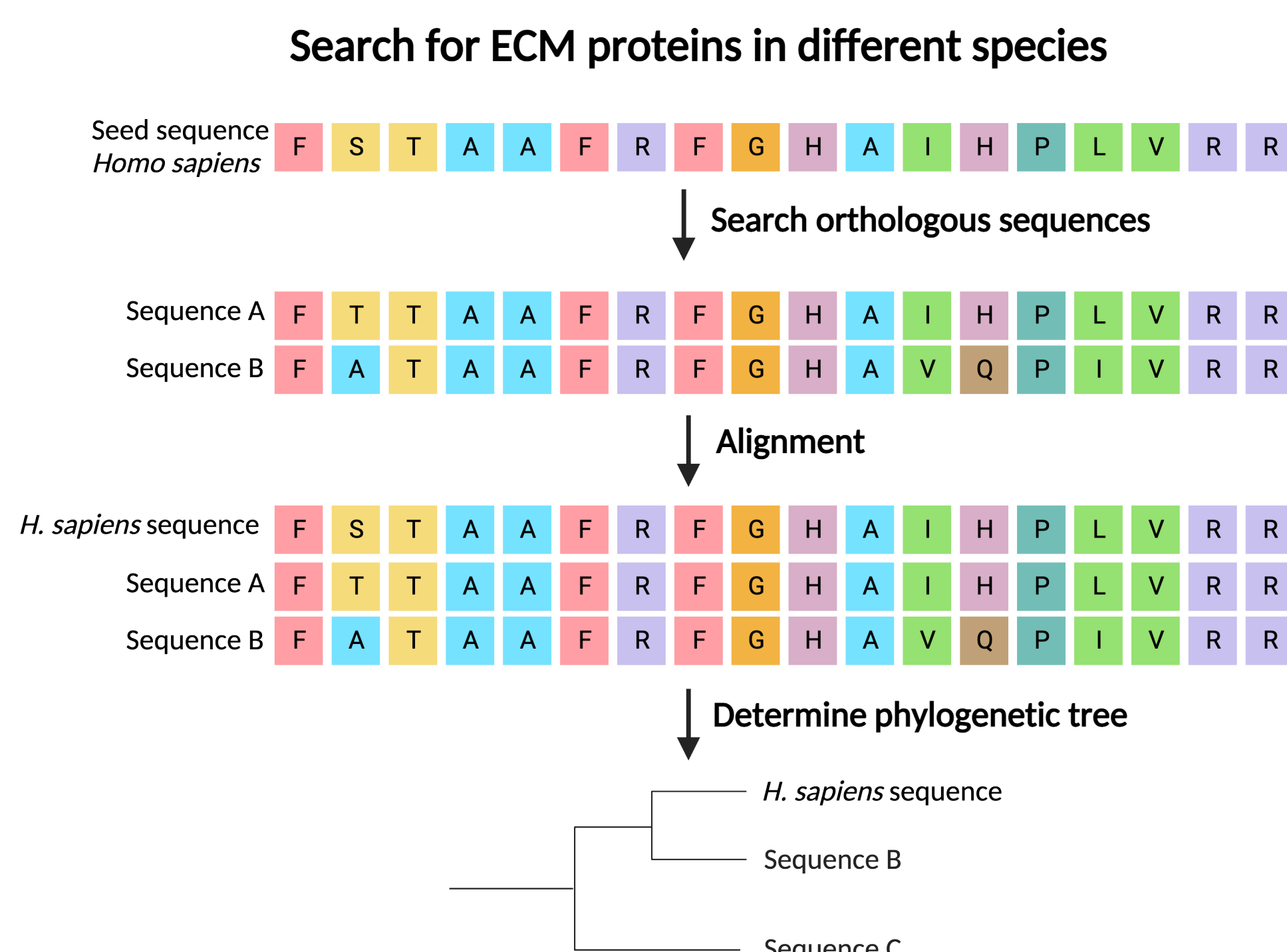### Search for ECM proteins in different species

Figure 2. We have used state-of-the-art bioinformatics analysis to identify 255 collagenous and non-collagenous proteins potentially expressed in the bone ECM of human, zebrafish, sika deer, and other 28 species of vertebrates. For each seed sequence in *Homo sapiens*, sequences from the same gene were searched in those species and analysed phylogenetically in order to find their evolutionary relationship.

CONCLUSION

The Phylobone database is the most comprehensive resource of bone ECM proteins in model organisms and can be relevant in osteoporosis research.

Our study, in agreement with other studies from the literature, indicates that several non-collagenous proteins are determinant to regulate bone formation and regeneration *via* signaling pathways.
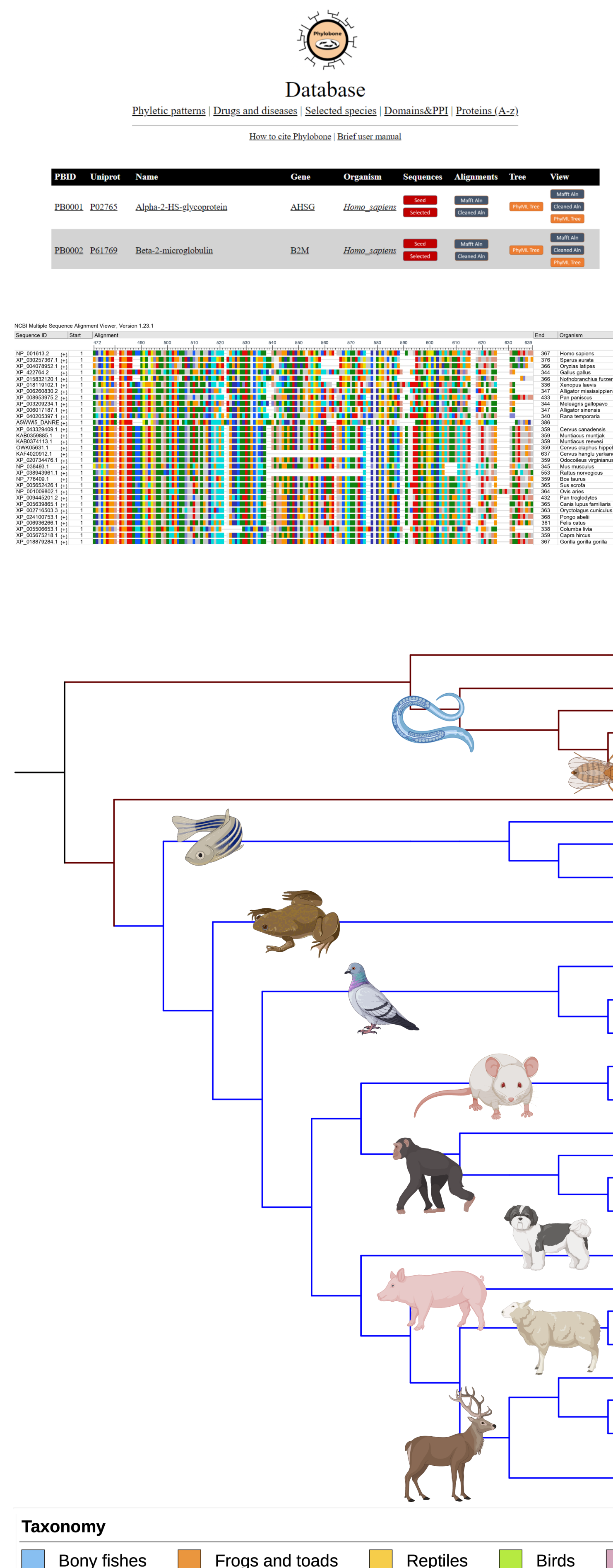
RESULTS

Figure 3. Screenshot of the main menu of Phylobone database. There is access to protein group's information, alignments, sequences and phylogenetic trees for the 8,615 putative bone ECM proteins from 39 species of vertebrates and invertebrates. Alignment and phylogenetic tree of Alpha-2-HS-glycoprotein (PB0001), a protein that influences bone mineralization, are shown. The Phylobone resource (1) is freely accessible at https://phylobone.com.
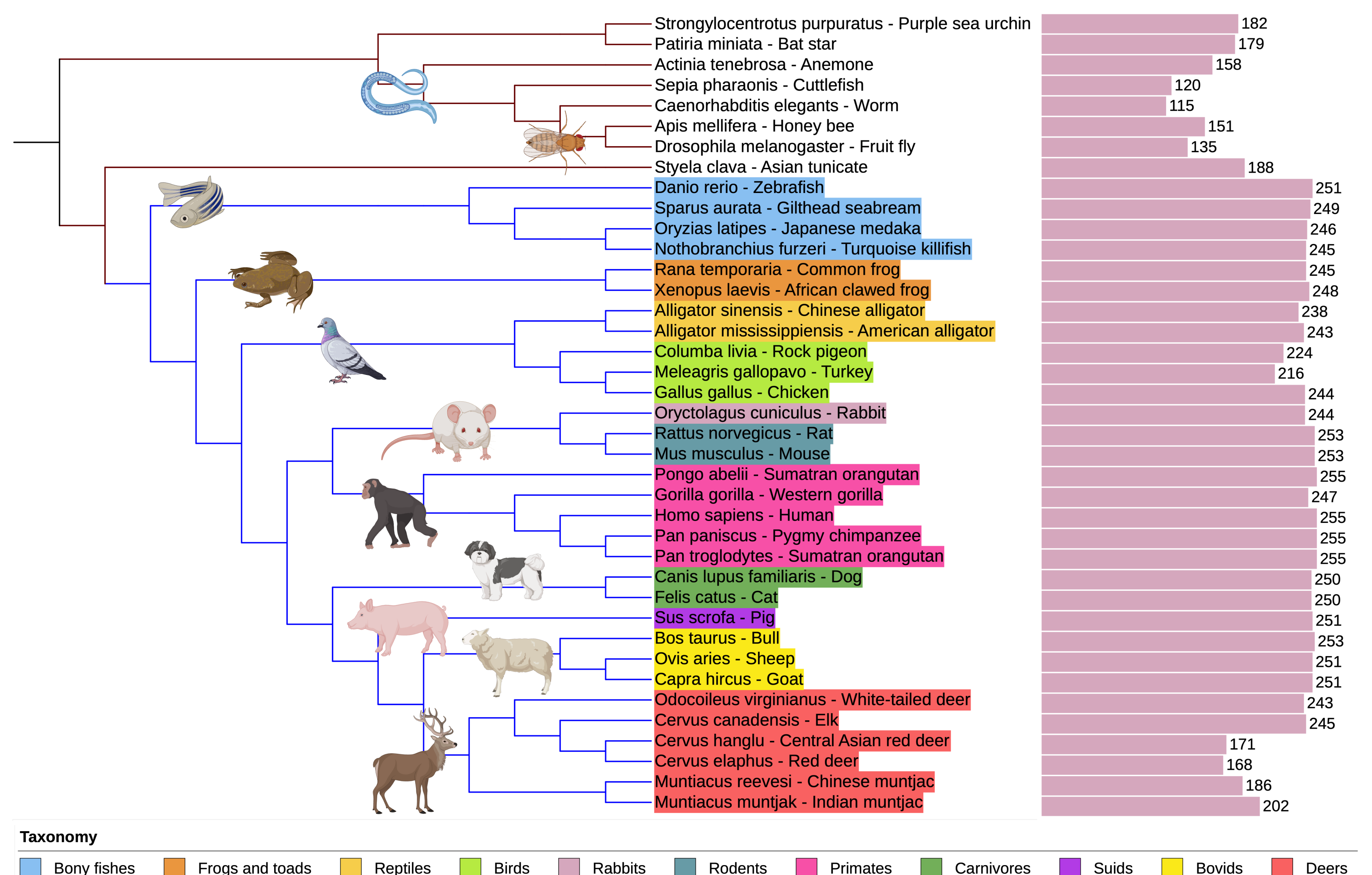

Figure 4. Tree with the 39 species in Phylobone database. The phylogenetic analysis of orthologous proteins across vertebrates allows the identification of potential model organisms to study molecular mechanisms in bone regeneration (e.g. some species have known orthologs for more than 90% of the proteins in the dataset, including human, rat, mouse, pig, rabbit, and members of Cervidae family).
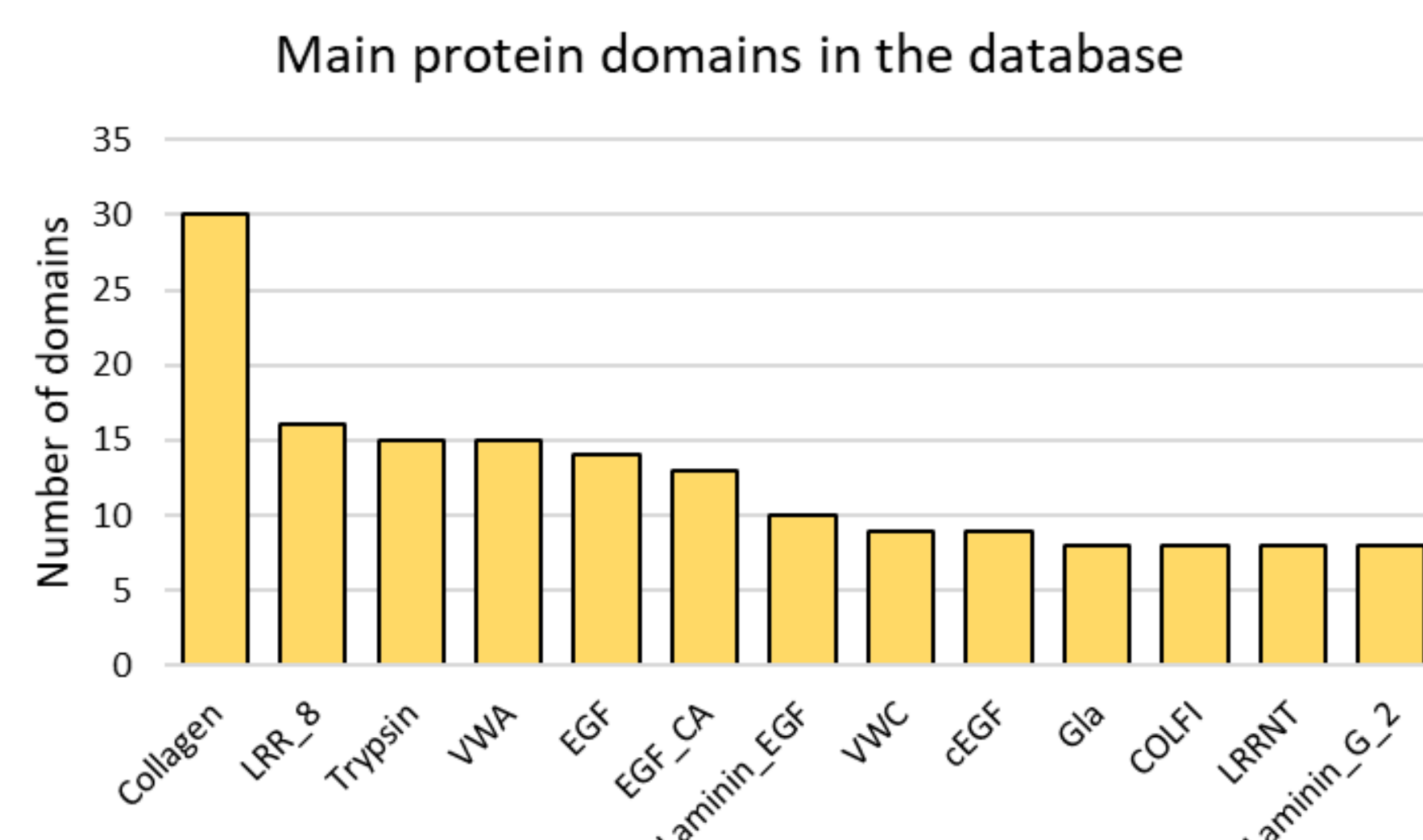

Figure 5. A total of 214 functional protein domains define functions of bone ECM proteins. As expected, collagen is the most common domain. In addition, domains such as leucine-rich repeats (LRR), laminin, epidermal growth factor (EGF), and von Willebrand factors (VWF: VWA, VWC) are highly frequent. LRR has a key role in bone formation and homeostasis, VWF is important for the inhibition of osteoclastogenesis, and EGF can stimulate bone resorption. Bone ECM proteins are also highly inter- and intra-connected. In the database the interactions and domains of each protein can be found.
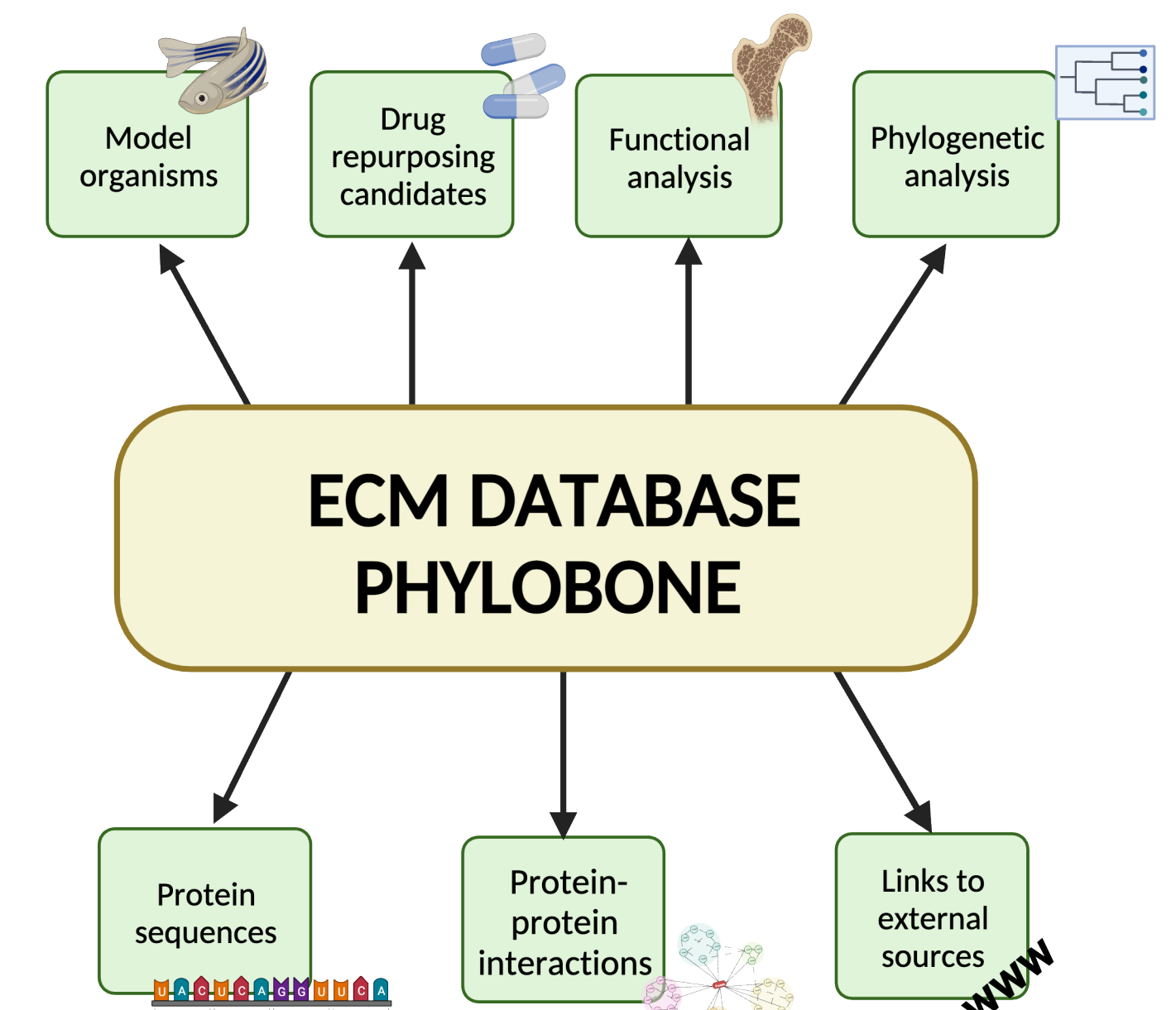

Figure 6. The Phylobone database contains a comprehensive list of proteins from model organism sequences. Moreover, it includes a list of drug repurposing candidates, that can be utilized for the identification of new treatments for bone regeneration and osteoporosis. It also contains functional and Gene Ontology information to assist in future research experiments.

1. Fontcuberta-Rigo M, Nakamura M, Puigbò P. Phylobone: A comprehensive database of bone extracellular matrix proteins in human and model organisms. BioRxiv. 2023 Mar 29;