

Molecular characterization of matrix metalloproteinase gene family across primates

Yinglian Pan¹, Yadan Fan², Yanda Lu¹, Siyuan Peng², Haixue Lin², Qingchun Deng²

¹Department of Medical Oncology, The First Affiliated Hospital of Hainan Medical University, Haikou 570102, Hainan, People's Republic of China

²Department of Gynecology, The Second Affiliated Hospital of Hainan Medical University, Haikou 570311, Hainan, People's Republic of China

Correspondence to: Qingchun Deng; **email:** qingchun0503@163.com, <https://orcid.org/0000-0001-7390-5271>

Keywords: matrix metalloproteinase, primate, gene structure, genomic organization, co-expression

Received: October 12, 2021

Accepted: April 12, 2022

Published: April 20, 2022

Copyright: © 2022 Pan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Deregulation of matrix metalloproteinases (MMPs) contributes considerably to cancers, psychiatric disorders, macular degeneration and bone diseases. The use of humans in the development of MMPs as prognostic biomarkers and therapeutic targets is complicated by many factors, while primate models can be useful alternatives for this purpose. Here, we performed genome-enabled identification of putative MMPs across primate species, and comprehensively investigated the genes. Phylogenetic topology of the MMP family showed each type formulates a distinct clade, and was further clustered to classes, largely agreeing with classification based on biochemical properties and domain organization. Across primates, the excess of candidate sites of positive selection was detected for MMP-19, in addition to 1-3 sites in MMP-8, MMP-10 and MMP-26. MMP-26 showed Ka/Ks value above 1 between human and chimpanzee copies. We observed two copies of MMP-19 in the old-world monkey genomes, suggesting gene duplication at the early stage of or prior to the emergence of the lineage. Furin-activatable MMPs demonstrate the most variable properties regarding Domain organization and gene structure. During human aging, MMP-11 showed gradually decreased expression in testis, so as MMP-2, MMP-14, MMP15 and MMP-28 in ovary, while MMP-7 and MMP-21 showed elevated expression, implying their distinct roles in different reproductive organs. Co-expression clusters were formed among human MMPs both within and across classes, and expression correlation was observed in MMP genes across primates. Our results illuminate the utilization of MMPs for the discovery of prognostic biomarkers and therapeutic targets for aging-related diseases and carry new messages on MMP classification.

INTRODUCTION

Matrix metalloproteinases (MMPs), also called matrixins, comprise a family of zinc-dependent endoproteinase. Collectively they are capable of degrading all extracellular matrix components (ECM) as well as non-matrix proteins [1]. The proteolytic activities of MMPs influence many fundamental physiological events involving tissue remodeling, embryogenesis, morphogenesis, angiogenesis, wound healing, and apoptosis [2–4]. Under pathological

conditions, deregulation of MMP activity causes a variety of pathological outcomes including matrix weakening, tissue destruction, and fibrosis [2, 5].

Over recent years, aging-related diseases involved with abnormal regulation of MMPs have attracted the attentions of physicians, medical scientists and pharmaceutical developers. In Brunch's membrane, amounts of MMPs such as MMP-2 and MMP-9 were observed to increase and the degree of covalent modification was enhanced for these MMPs with age

[6, 7]. The aging changes in the MMP system were shown to be aggressively advanced in age-related macular degeneration, resulting in a large accumulation of abnormal ECM material [8]. Over-expression of MMPs can cause Alzheimer's disease, a progressive neurologic disorder commonly affecting people over the age of 65 [9]. A recent study showed that the alteration of amyloid precursor proteolysis, a hallmark of Alzheimer's disease, was mediated by MMP-14 in human neuronal cells [10]. Parkinson's disease, which ordinarily begins in late life, has been reported to be associated with aberrant expression of MMP-3 and MMP-9 [11, 12]. MMPs like MMP-2, MMP-9 and MMP-13 have been shown to participate in bone formation and development, and abnormal regulation of these genes contributes to the development of bone diseases [13]. Cancers, occurring mostly in middle age and above, have been linked to MMPs with evidence as they are an important factor promoting tumor progression [14]. This is especially true for tumors like pancreatic adenocarcinoma, prostate adenocarcinoma, bladder urothelial carcinoma and melanoma, which are commonly at risk above the age of 50 [15–18]. Collectively they reveal that MMPs widely participate in physiological and pathological processes, and thus are promising targets for diagnosis and drug development of aging-related diseases.

Twenty-four MMP genes are present in the human genome; among them, MMP-23A is a duplicated copy of MMP-23B and likely a pseudogene [19]. Based on biochemical properties, MMPs can be classified into collagenases (including MMP-1, MMP-8 and MMP-13, whose key feature is the ability to cleave interstitial collagens), stromelysins (MMP-3, MMP-10, and MMP-11), gelatinases (MMP-2 and MMP-9, which can digest gelatins), matrilysins (MMP-7 and MMP-26, which can digest ECM components and activate proMMPs), membrane-type MMPs (MMP-14, MMP-15, MMP-16, MMP-17, MMP-24, and MMP-25, which can digest ECM molecules) and unclassified MMPs (MMP-12, MMP-19, MMP-20, MMP-21, MMP-23, MMP-27, and MMP-28) (Table 1) [20]. Among the unclassified MMPs, MMP-19 is used as a T-cell-derived auto-antigen from patients with rheumatoid arthritis, but its exact function is not fully characterized [21]. MMP-20 can digest amelogenin [22]. The limitation of this classification is obvious, as multiple substrates have been revealed for several MMPs. For example, MMP-1, a collagenase, can also cleave tenascin and aggrecan. MMP-14, biochemically classified as a membrane-type MMP, serves as a collagenase, too [23]. Another weakness is that several (7) MMPs are still not classified. These have limited the accuracy of this classification method. An alternative classification was

proposed based on domain organization, which categorizes MMPs into four groups, namely archetypal MMPs, matrilysins, gelatinases and furin-activatable MMPs (Table 1) [24]. This classification approach still holds weaknesses as domains defined in this context are not so common as conserved domains commonly referred to like Pfam domains. Therefore, careful investigation is required when allocating a new MMP gene into this classification [25].

In this article, we investigate primate MMPs with evolutionary analyses together with gene structure, genome organization and comparative gene expression analyses. We compare our results with the two stereotypical classification approaches, in the hope to reach consensus for classification approaches. We also identify MMPs of potential value for future studies.

RESULTS

Comprehensive identification of putative MMPs from 11 primate genomes

Primates arose 63–74 million years ago from small terrestrial mammals [26]. To make sure major lineages of primates are covered, 11 species were chosen for identification of putative MMPs in this manuscript (Figure 1). In addition to the human genome, those of bonobo (*Pan paniscus*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo abelii*) and gibbon (*Nomascus leucogenys*) were included to represent Hominoids. Olive baboon (*Papio anubis*), gelada (*Theropithecus gelada*) and golden snub-nosed monkey (*Rhinopithecus roxellana*) were sampled to represent old-world monkeys. Marmoset (*Callithrix jacchus*) and mouse lemur (*Microcebus murinus*) represented new-world monkey and prosimian, respectively. All the genome assemblies are chromosome-level, except that of marmoset which has only scaffold-level assembly available.

Through a combination of Basic Local Alignment Search Tool-Protein (BLASTp) and domain-search-based HMMER, a total of 261 putative MMPs were identified (Figure 1). Looking at gene counting in genomes, 19 to 27 MMPs were obtained for each species. Three Hominini species, human, chimpanzee and bonobo, harbor 23, 24, and 22 genes, respectively. Three other Hominoid species, gorilla, orangutan, and gibbon contain 25, 19 and 24 MMPs, respectively. Compared to other primate groups, MMP copy numbers in old-world monkey genomes are generally greater, with 27, 24 and 26 genes in olive baboon, gelada and snub-nosed monkey, respectively. In marmoset, a new-world monkey species, 23 MMPs were obtained. Mouse lemur, one of the smallest living primates, harbor 24 MMPs.

Table 1. Classification of matrix metalloproteinase (MMP) enzymes.

Class	Type	Name
Archetypal MMP	Collagenase	MMP-1, MMP-8, MMP-13
	Stromelysin	MMP-3, MMP-10
	Others	MMP-12, MMP-19, MMP-20, MMP-27
Gelatinase		MMP-2, MMP-9
Matrilysin		MMP-7, MMP-26
Furin-activatable MMP	Secreted	MMP-11, MMP-21, MMP-28
	Type I transmembrane MT-MMP	MMP-14, MMP-15, MMP-16, MMP-24
	GPI-anchored MMP	MMP-17, MMP-25
	Type II transmembrane MT-MMP	MMP-23A, MMP-23B

Classification of primate MMPs based on phylogenetic inference

By developing a phylogenetic tree, the peptides were classified into 23 distinct clades, with high bootstrap support at basal nodes (Figure 2). One human MMP was categorized into each clade, confirming the reliability of the tree topology. Each of the 23 MMP genes was identified in each of the most genomes, suggesting conserved evolution of the MMP family in primates (Supplementary Table 1). In this manuscript, very limited number of genes were missing in certain clades. For example, MMP-3 was failed to identify in mouse lemur, and MMP-7 was not retrieved from the gorilla genome. This result indicates a limited number of MMPs are missing from certain primate genomes, but the possibility of genome assembly errors or annotation artifacts could not be ruled out.

Interestingly, two copies of MMP-19 were present in each of the old-world monkey genome. To resolve relationships of these genes, phylogenetic trees of the MMP-19 peptides alone were reconstructed, which resolved the two copies into two distinct clades with high bootstrap support (Figure 3). In each clade, 1 copy of MMP-19 was represented for each of the three species, suggesting duplication of MMP-19 at the early stage of or before the emergence of old-world monkeys.

Based on biochemical properties, MMPs can be classified into four groups, namely archetypal MMPs, matrilysins, gelatinases, and furin-activatable MMPs (Table 1) [24]. Our phylogenetic topology shows that MMPs are clustered in four major groups (shown in brown circles in Figure 2). MMP-1s, MMP-8s and MMP-13s are proximate to each other, which biochemically serve as collagenases; MMP-3 and MMP-10, two clades as stromelysins, form an

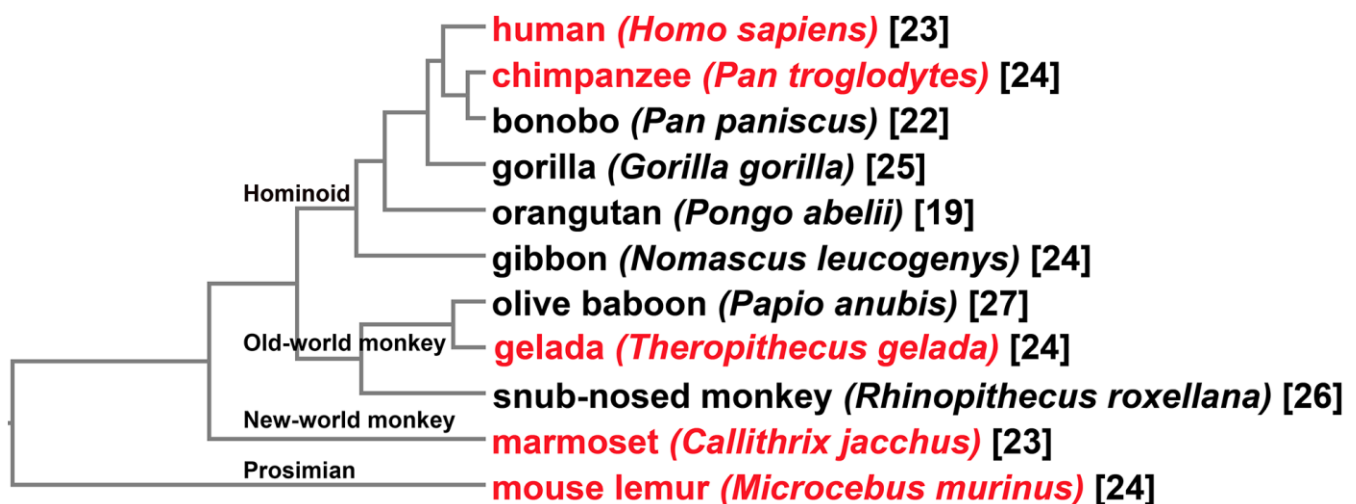


Figure 1. Taxonomy of the primate species for investigation of MMPs. At the tip of each clade shows the common name followed by the Latin name of the organism. Tips shown in red represent species with comparative gene expression analyses performed. Shown at nodes are four major primate groups. Numbers in brackets are counting of MMPs identified for each organism.

independent group. These peptides, together with MMP-12s and MMP-27s, are clustered together, which can be biochemically classified as archetypal MMPs. MMP-11s, MMP-21s and MMP-28s, three secreted MMPs, are proximate to each other. MMP-14s, MMP-15s, MMP-16s and MMP-24s, which are type-I transmembrane MT-MMPs, form a group. MMP-17 and MMP-25, GPI-anchored MMPs, clustered together. These peptides, together with MMP-19 and MMP-23, form a group and can be classified as furin-activatable MMPs, except MMP-19 which is referred to as an archetypal MMP-. MMP-7s

and MMP-26s serve as matrilysins, and MMP-2s and MMP-9s serve as gelatinases, and they form a group, respectively. In summary, the phylogenetic topology of primate MMPs largely agrees with biochemistry-based classification, which demonstrates the power of phylogenetic analysis as a guide in biochemical characterization of a certain MMP gene.

Evolution of primate MMPs

To investigate the possibility of positive selection among primate MMPs, a maximum-likelihood-based

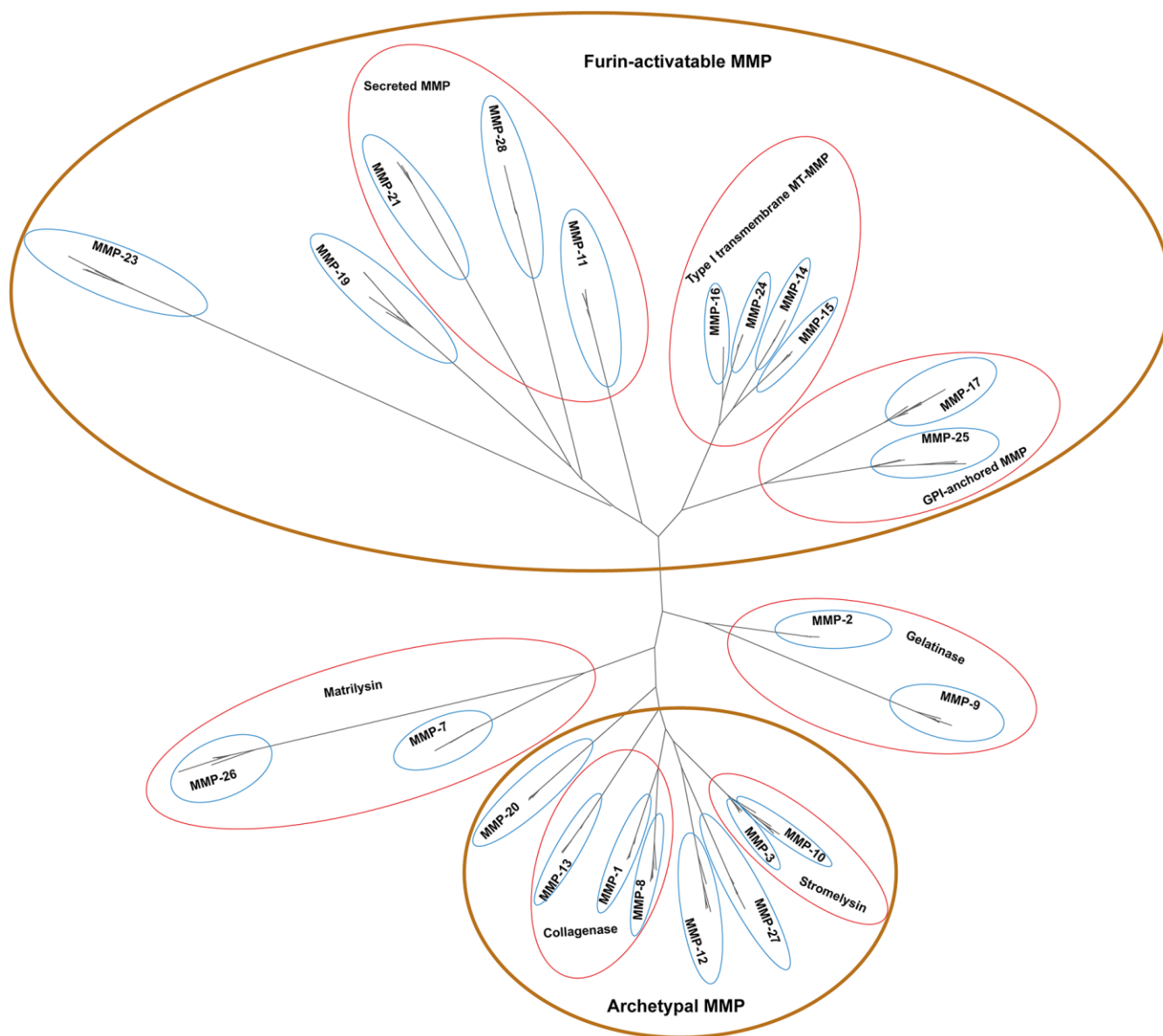


Figure 2. Unrooted phylogenetic tree of 261 primate MMP peptides. Looped in brown ovals are four major groups supported with biochemistry evidence, and enclosed in red circles are sub-groups based on biochemical properties under certain groups. Phylogenetic analyses were performed by using PhyML v3.3.3 with 1000 bootstraps.

estimation of codon substitution model was carried out, on the basis of codon-enabled sequence alignment. Hypothesis of positive selection was tested against purifying selection with a likelihood ratio test (LRT). Overall, primate genes of MMP-8, MMP-9, MMP-10, MMP-19 and MMP-26 showed positive selection, with the M8 model fitting the data significantly better than the M7 model (Table 2). Furthermore, predicted ω for each site was calculated as the weighted mean of ω categories, and candidate sites of positive selection were detected in MMP-8, MMP-10, MMP-19 and MMP-26. The excess of candidate sites of positive selection is striking for MMP-19, with 12 positive selected sites detected (Table 2).

To further investigate molecular evolution of human MMPs, evolutionary rates between human and chimpanzee orthologs were calculated. Results showed that K_a (nonsynonymous substitution), K_s (synonymous substitution) and K_a/K_s showed variation among orthologs of MMPs. Only MMP-26 showed K_a/K_s ratio above 1, implying selection of human MMP-26 gene (Table 3). MMP-17 showed the highest synonymous substitution rate ($K_s = 0.0398$), and MMP-3, MMP-11 and MMP-15 also showed K_s values above 0.025, suggesting relatively relaxed pressure which results in sequence divergence for

these genes. Notably, orthologs of MMP-24 showed 0 value of K_a , which indicates that function of human MMP-24 is strictly restricted.

Conservation and divergence of domain organization and gene structure in primate MMPs

Like those of many other families, MMP peptides contain a conserved domain structure, which has been revealed in human and many other organisms [19, 24, 27]. By querying the databases of Pfam and SignalP, we identified a cascade of domains in each MMP peptide (Figure 4; function information is described in Supplementary Table 2). All the MMPs contain a domain with Pfam ID PF00413, which is a catalytic metalloproteinase domain (Supplementary Table 2). Near the N-terminal, there is a domain resembling bacterium peptidoglycan binding domain (PF01471) for all the MMPs except MMP-11, MMP-26 and MMP-23B. At the C-terminal, hemopexin-like repeats were contained in all MMPs except MMP-23B and the two matrilysins. A signal peptide was identified in all MMPs except MMP-16 and MMP-23B. In many non-human primate species, each gene shares the same domain organization like that in human, except limited peptides (Supplementary Table 3). For example, for MMP-24, only the human and

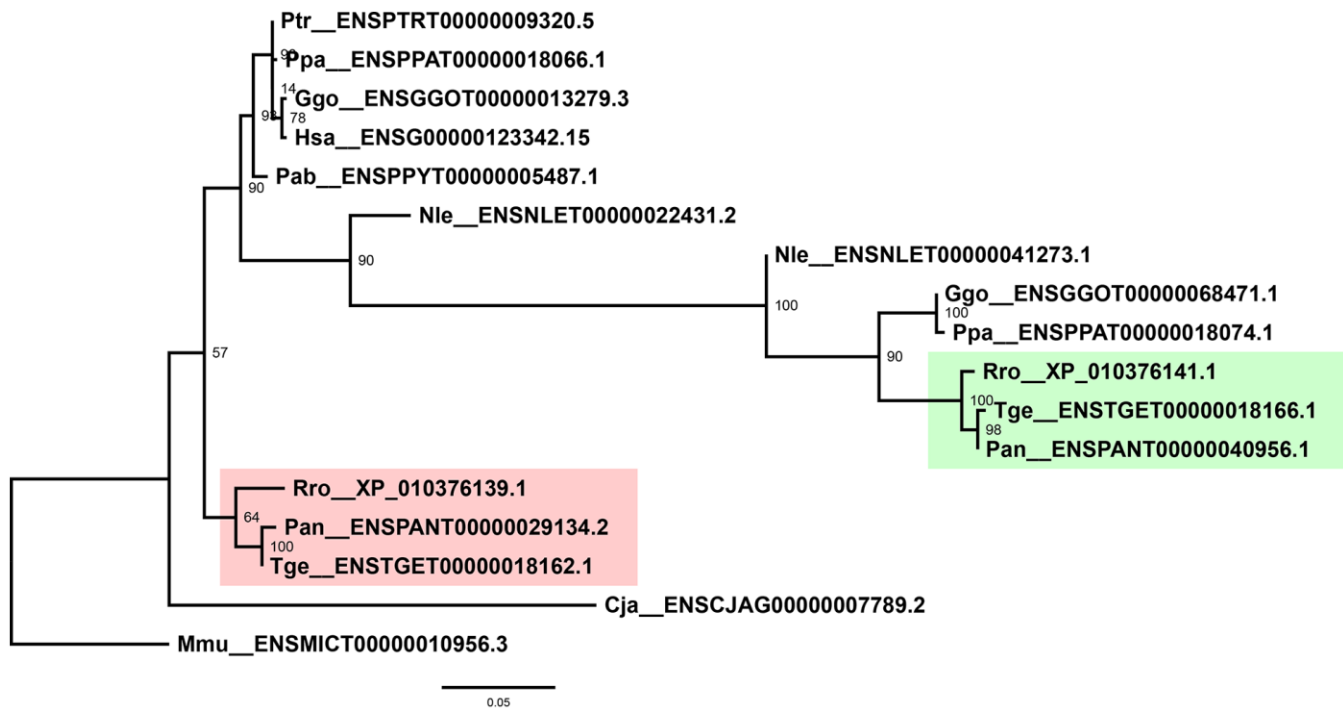


Figure 3. Phylogenetic tree of primate MMP-19s. Shaded are two independent clades as a result of duplication of old-world monkey MMP-19s. Labeled at nodes are bootstrap support values. Phylogenetic analyses were performed by using PhyML v3.3.3 with 1000 bootstraps.

Table 2. Test of positive selection among clades of primate MMPs.

	lnL(M7)	lnL(M8)	P-value	Positive selection sites
MMP-1	-3332.31	-3331.97	0.713	-
MMP-2	-3230.67	-3229.95	0.486	-
MMP-3	-3662.88	-3662.19	0.499	-
MMP-7	-2007.11	-2004.66	0.086	-
MMP-8	-3203.69	-3195.15	0.000	H35,G67,T124
MMP-9	-5414.77	-5408.82	0.003	-
MMP-10	-3099.71	-3095.91	0.022	G113
MMP-11	-2908.63	-2907.80	0.435	-
MMP-12	-3597.70	-3597.22	0.619	-
MMP-13	-3276.51	-3276.22	0.747	-
MMP-14	-3505.05	-3505.05	1.000	-
MMP-15	-4486.82	-4486.59	0.794	-
MMP-16	-3171.86	-3171.86	0.999	-
MMP-17	-3222.31	-3220.89	0.242	-
MMP-19	-1882.99	-1874.37	0.000	P93,A98,G122,L135,P164,L186, G187,L189,F190,S196,Y203,T206
MMP-20	-3309.21	-3307.24	0.140	-
MMP-21	-4067.37	-4067.37	1.000	-
MMP-23	-1816.71	-1816.71	1.000	-
MMP-24	-2961.92	-2960.96	0.385	-
MMP-25	-1513.27	-1513.14	0.874	-
MMP-26	-1531.90	-1527.78	0.016	M6
MMP-27	-3414.02	-3412.24	0.168	-
MMP-28	-3443.36	-3443.19	0.844	-

Table 3. Nonsynonymous and synonymous substitutions between human and chimpanzee MMP orthologs.

MMP	human	chimpanzee	Ka	Ks	Ka/Ks
MMP-1	ENSG00000196611.4	ENSPTRT00000007870.6	0.0031	0.0071	0.444
MMP-2	ENSG00000087245.12	ENSPTRT00000092414.1	0.0007	0.0247	0.027
MMP-3	ENSG00000149968.11	ENSPTRT00000007868.3	0.0063	0.0295	0.2139
MMP-7	ENSG00000137673.8	ENSPTRT00000007862.4	0.0075	0.0242	0.3092
MMP-8	ENSG00000118113.11	ENSPTRT00000007866.5	0.0071	0.0156	0.4591
MMP-9	ENSG00000100985.7	ENSPTRT00000025282.4	0.0024	0.0117	0.2077
MMP-10	ENSG00000166670.9	ENSPTRT00000007867.5	0.0049	0.0127	0.3881
MMP-11	ENSG00000099953.9	ENSPTRT00000026441.6	0.0029	0.0261	0.1099
MMP-13	ENSG00000137745.11	ENSPTRT00000088050.1	0.0065	0.0217	0.3017
MMP-14	ENSG00000157227.12	ENSPTRT00000011266.6	0.0008	0.0225	0.0356
MMP-15	ENSG00000102996.4	ENSPTRT00000015071.5	0.0019	0.0279	0.0679
MMP-16	ENSG00000156103.15	ENSPTRT00000099969.1	0.0015	0.0174	0.086
MMP-17	ENSG00000198598.6	ENSPTRT00000010368.6	0.0035	0.0398	0.0889
MMP-19	ENSG00000123342.15	ENSPTRT00000009320.5	0.0029	0.0109	0.2691
MMP-20	ENSG00000137674.3	ENSPTRT00000007863.3	0.0041	0.0089	0.4612
MMP-21	ENSG00000154485.4	ENSPTRT00000005789.2	0.0025	0.0122	0.2076
MMP-23	ENSG00000189409.13	ENSPTRT00000000071.5	0.001	0.0239	0.0404
MMP-24	ENSG00000125966.9	ENSPTRT00000024960.5	0	0.0119	0.001
MMP-25	ENSG00000008516.16	ENSPTRT00000014165.7	0.0038	0.0089	0.4229
MMP-26	ENSG00000167346.7	ENSPTRT00000006170.3	0.0126	0.0036	3.5059
MMP-27	ENSG00000137675.4	ENSPTRT00000007864.3	0.0061	0.0135	0.4547
MMP-28	ENSG00000271447.5	ENSPTRT00000045218.3	0.0026	0.014	0.1825

chimpanzee sequences have the signal peptide, and we failed to identify it in all the other primate species. Most conserved domains we failed to identify were located at N- or C-terminal, suggesting that they are results of genome annotation artefacts.

As the domain organization is different, we investigated gene structure among MMPs. Gelatinases (MMP-2 and MMP-9) have the largest molecular weights, which might be attributable to the extra domain (PF00040) they have. Matrilysin (MMP-7 and MMP-26) show the smallest molecular weights, as they have the simplest domain organization; they have the largest CDS (coding sequence)/gene ratios, which demonstrates that they have the most compact gene structure. Furin-activatable MMPs are the longest genes, and CDS/gene ratios are low, which indicates that they have the longest introns. Exon numbers vary both among the different MMP types and within the same type; exon numbers of gelatinases are overall the largest. Among the groups of MMPs, furin-activatable MMPs show the largest variation for

molecular weight, gene length and exon counting (Figure 5).

Genomic organization of primate MMPs

As gene sequence, domain organization and gene structure of the MMP genes show both conservation and divergence, we investigated the conservation of genomic organization for the primate genes. We first mapped the human MMP genes to chromosomes and discovered that the 23 human genes are located at 10 chromosomes. Interestingly, 9 MMP genes are located together as a single cluster in the 11th chromosome, of which most are archetypal MMPs (MMP-1, -3, -8, -10, -12, -13, -20, and -27), and one encodes matrilysin (MMP-7) (Figure 6).

We then investigated co-synteny of the primate genomes and the MMPs in the co-synteny clusters. We compared the human genome with the genomes of three other primates, chimpanzee, a Hominoid, olive baboon, an old-world monkey, and mouse lemur, a Prosimian species. The genome of marmoset, the only new-world

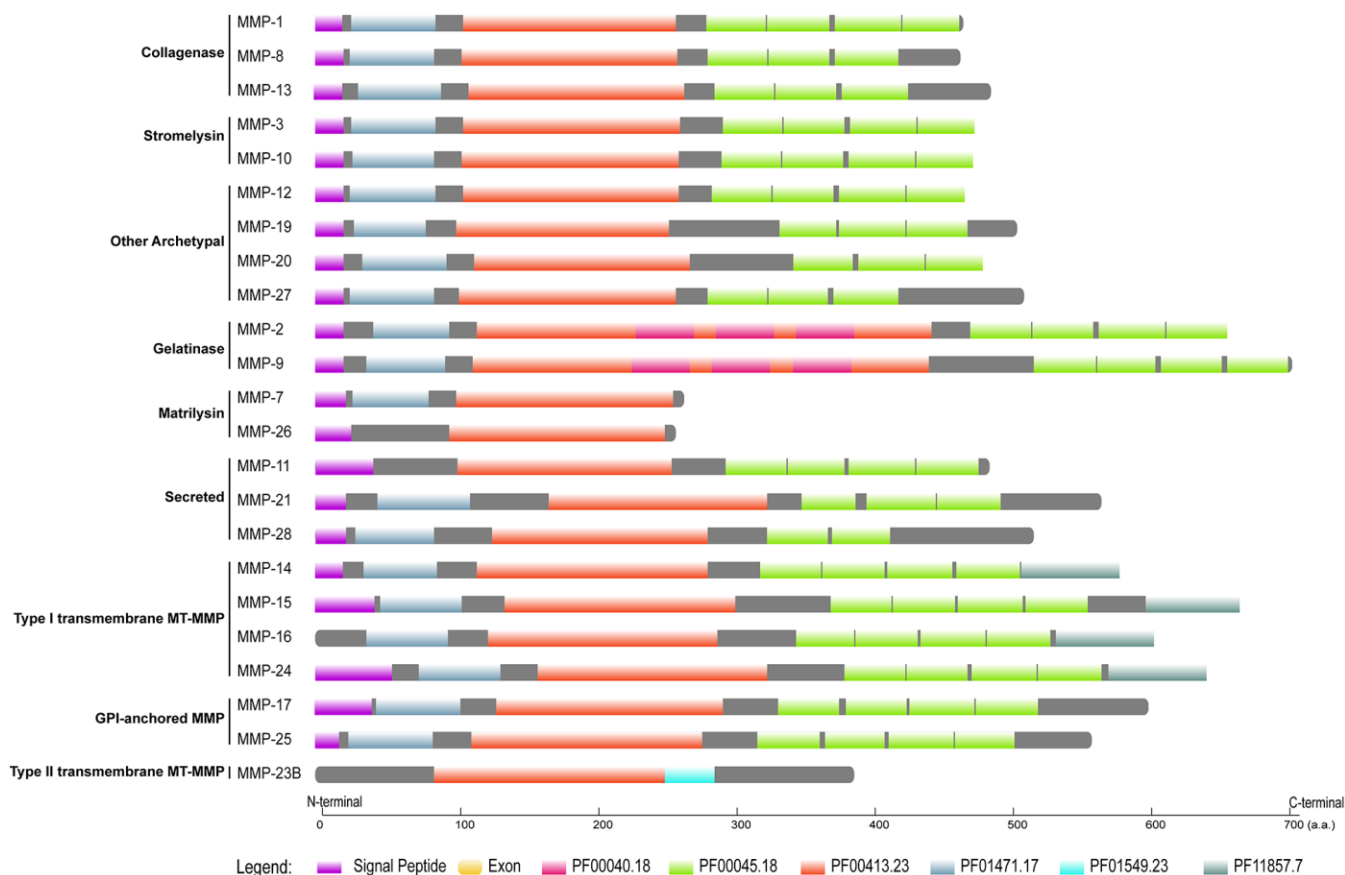


Figure 4. Domain organization of human MMPs. Pfam domains were identified by using HMMER 3.3 against Pfam database, and signal peptides were obtained by querying SignalP-5.0 server. Pfam domains and signal peptides are shown as colored boxes.

monkey representative, was not assembled to the chromosome level, and thus was excluded from the co-synteny analysis [28]. There was high co-synteny between the human and chimpanzee genomes, and between chimpanzee and olive baboon genomes. There was moderate co-synteny between the olive baboon and mouse lemur genomes. Twenty-one of 23 human genes are located in 12 co-synteny clusters. A cluster in the human 16th chromosome, where MMP-2 and -15 are located, is not shared between chimpanzee and olive baboon. Only 16 MMPs are within the 8 co-synteny clusters between olive baboon and mouse lemur. The co-synteny cluster in the human 11th chromosome, where the 9-gene cluster is located, is conserved across all the four primate species (Figure 7). Together, these results demonstrate that both conservation and divergence of MMPs are attributable to the conserved genome organization over primates and changes of co-synteny as the genetic relationship gets further.

Expression of MMP genes in human and across primates

As MMPs are clustered into different groups, little is known about the relatedness among the MMPs on the gene expression level. We extracted MMP gene expression data of 11 human tissues from Human BodyMap and performed co-expression analysis among the genes. Gene expression data show that, although genes like MMP-2 and MMP-14 show ubiquitous high expression, many genes have tissue-specific expression patterns, and MMPs are expressed in all tissues except skeletal muscles, in which only limited genes like MMP-2 and MMP-14 are expressed (Figure 8A).

MMPs play important roles in human aging [29]. As effects of aging are particularly evident on both male and female reproductive system [30, 31], we further investigated longitudinal expression of the MMP genes in male and female reproductive organs, testis and

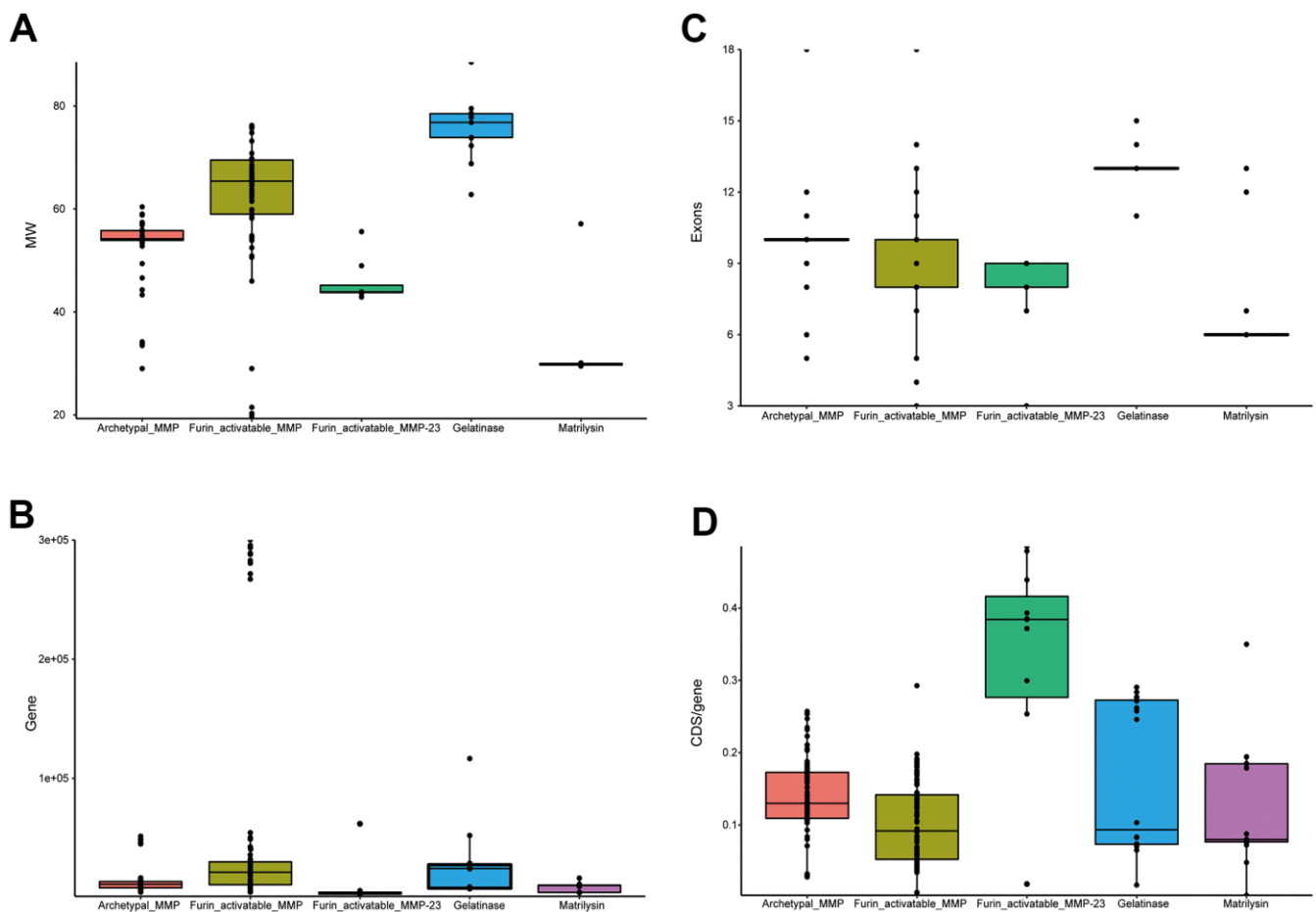


Figure 5. Statistics of primate MMP gene structure. (A) Box plots of molecular weights of MMP peptides. (B) Box plots of gene lengths. (C) Box plots of CDS/gene ratio. (D) Box plots of exon counting. While MMP-23s are classified in furin-activatable MMPs, they demonstrate very different characteristics than other MMPs, and thus were investigated separately.

ovary, respectively, by using GTEx data [32]. In testis, the male sexual organ, MMP-2, MMP-11, MMP-14, MMP-15, MMP-17 and MMP-28 are highly or moderately expressed across all ages from 20-79 (Figure 8B). MMP-11 showed gradually decreased expression during human aging, and all the other moderately or highly expressed genes showed lower expression at the ages of 20-29 than the older ages in testis (Figure 8C). In ovary, the female sexual organ, MMP-2, MMP-11, MMP-14, MMP-17 and MMP-28 showed universally moderate or high expression across all ages, the same as those in testis; in addition, MMP-19 and MMP-23B were highly expressed, too, which is different to those in testis (Figure 8C). Two lowly-expressed genes, MMP-7 and MMP-21, showed elevated

expression during aging, and MMP-2, MMP-14, MMP-15 and MMP-28 showed the reversed trend in ovary (Figure 8D).

Co-expression analysis shows that all the three genes encoding collagenases (MMP-1, MMP-8 and MMP-13) are co-expressed (Figure 9A). In addition, they formed a co-expression cluster with MMP-10 and MMP-19. Interestingly, the expression of some genes is correlated even they belong to different groups based on stereotypical classifications. For example, a gene cluster formed for five furin-activatable MMP genes, MMP-11, MMP-17, MMP-21, MMP-23A and MMP-23B, although they belong to different groups according to biochemistry-based classification. In addition, groups were formed for genes in different

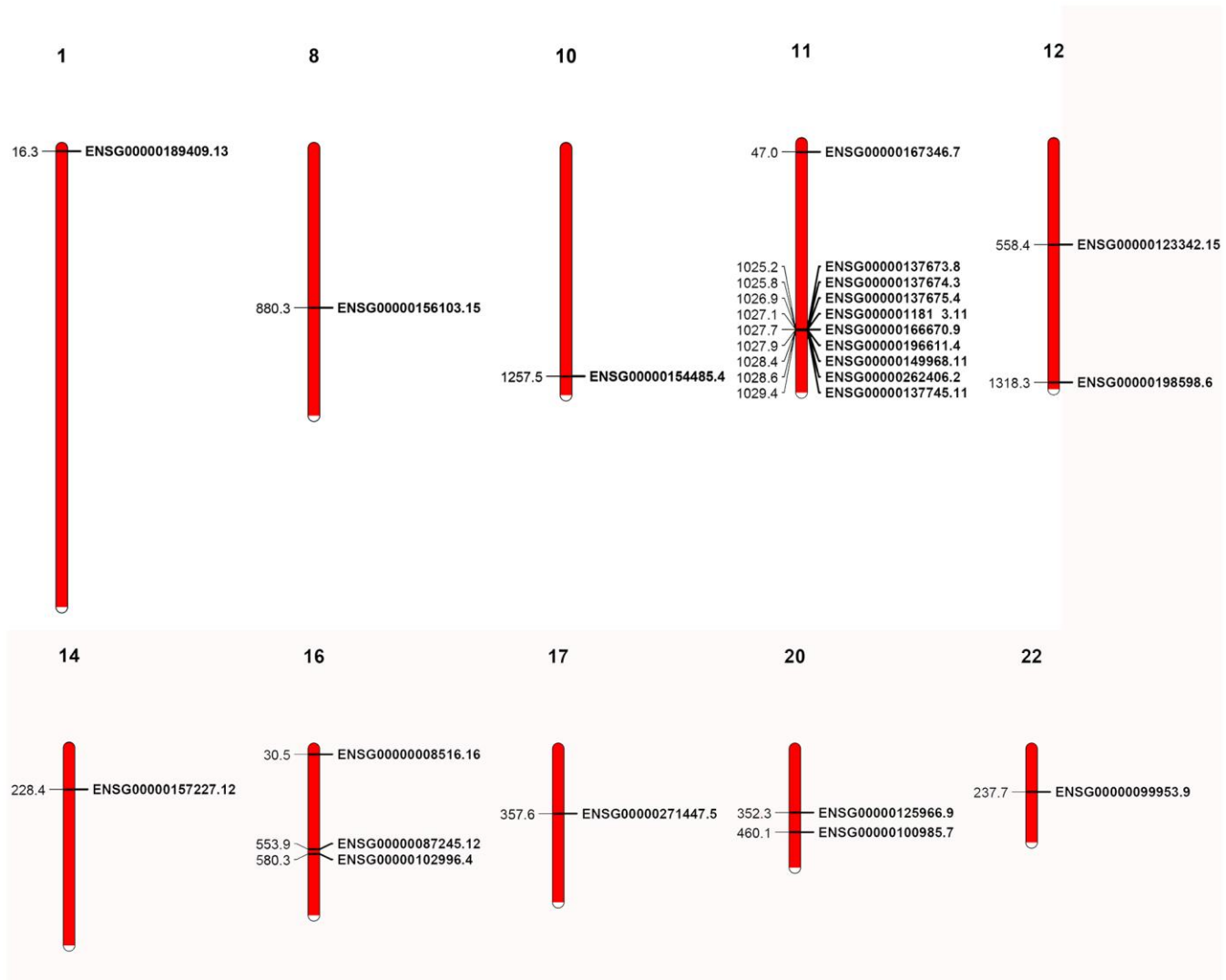


Figure 6. Localization of human MMP genes. A cluster is shown in the 11th chromosome. The localization of the genes was retrieved from GFF3 files.

groups by domain organization, such as pair of MMP-7 and MMP-24, and pair of MMP-9 and MMP-25 (Figure 9A).

To explore the relationship of gene expression patterns across primates, we compared gene expression of human MMP genes with those in chimpanzee, olive baboon, marmoset and mouse lemur. Between human and chimpanzee, 13 MMP genes (MMP-2, MMP-7, MMP-8, MMP-9, MMP-11, MMP-13, MMP-14, MMP-16, MMP-19, MMP-23A, MMP-23B, MMP-24, MMP-28) demonstrated co-expression. Fewer co-expression pairs were obtained between human and other species having longer genetic distance. Only 7 genes (MMP-2, MMP-8, MMP-9, MMP-16, MMP-17, MMP-19 and MMP-25) showed as co-expression pairs.

Expression of MMP-2 and MMP-8 showed the most conserved, as they showed co-expression between human and all the four non-human primate species (Figure 9B).

DISCUSSION

Non-human primates are essential models for modern medical and biochemical studies, because of their genetic, physiological and psychological similarity to humans. Even mouse lemur, which is placed at the basal node of primate taxonomy and roughly half the genetic distance between mice and humans, has been shown as a suitable model organism for drug development and genetics studies [33]. Investigation of gene evolution among primates can shed light on biochemical role, physiological and pathological

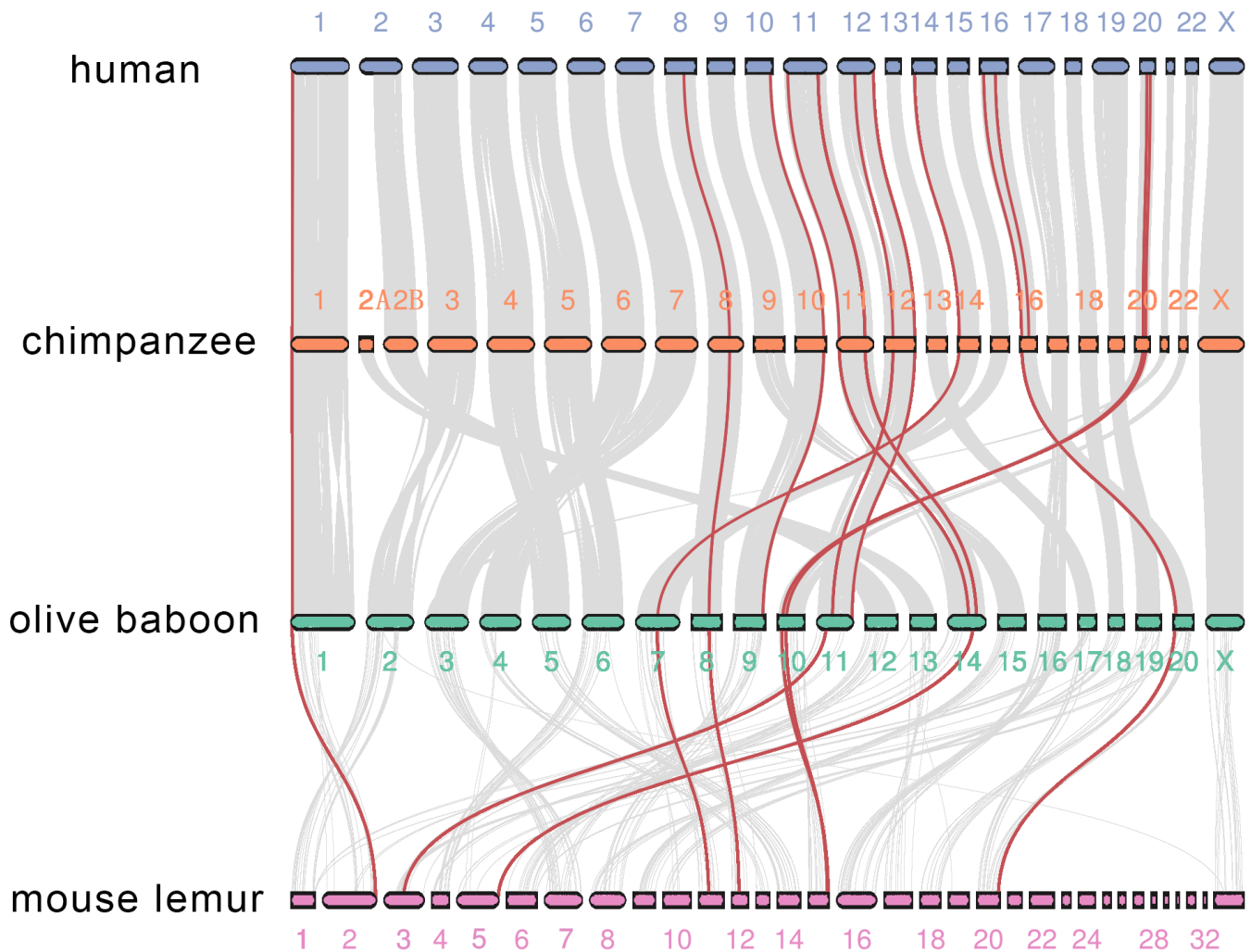


Figure 7. Co-synteny of the MMP genes across human, chimpanzee, olive baboon and mouse lemur. Co-synteny clusters are shown as grey curves, and MMPs in the co-synteny clusters are shown in red.

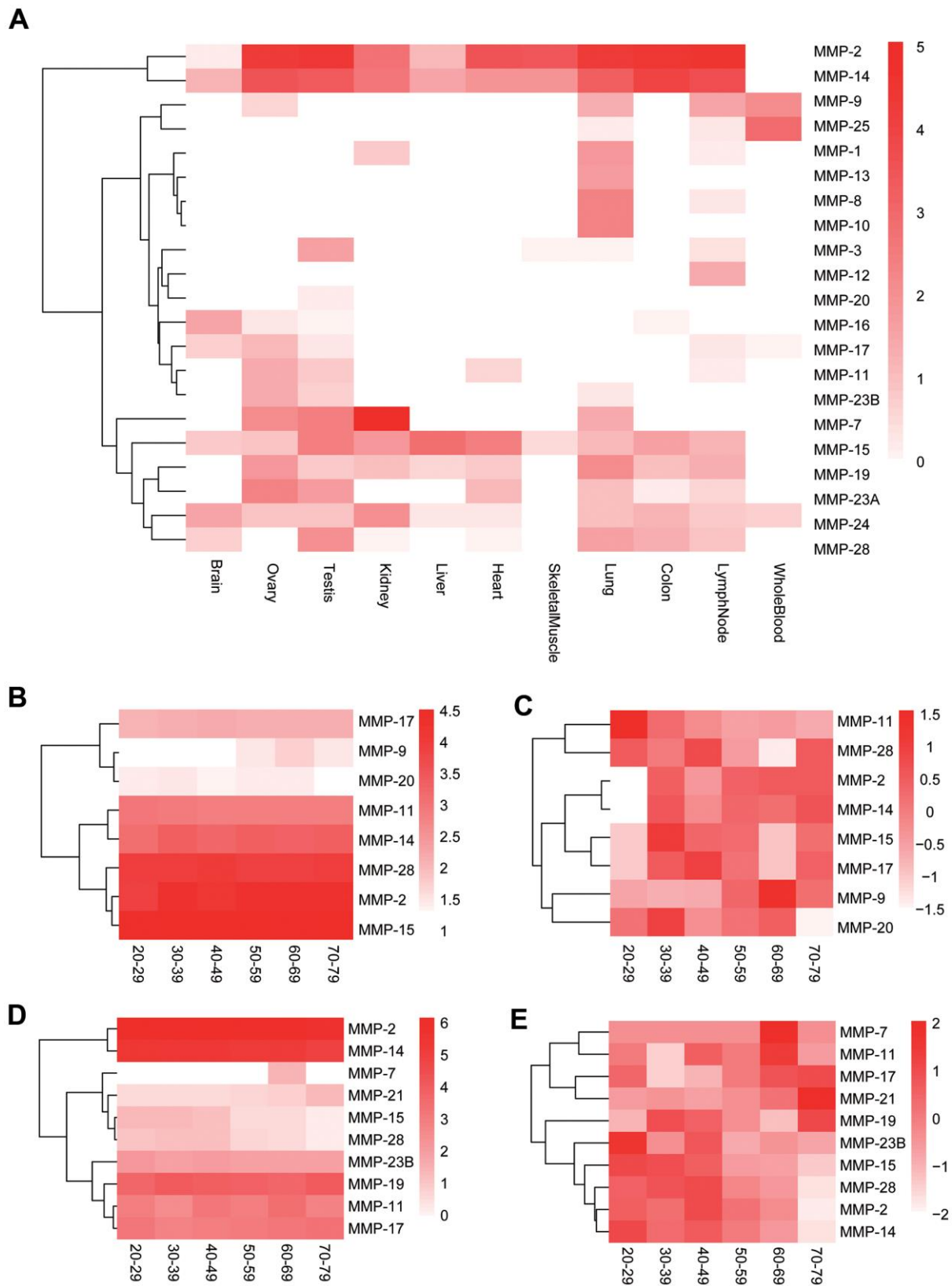


Figure 8. Expression of human MMP genes. (A) Heatmap of human MMP gene expression in 11 tissues. (B) Longitudinal expression of human MMPs in testis (not normalized). (C) Longitudinal expression of human MMPs in testis (normalized). (D) Longitudinal expression of human MMPs in ovary (not normalized). (E) Longitudinal expression of human MMPs in ovary (normalized). The MMP genes are clustered based on expression patterns.

functions of a certain gene, gene family, signaling network or metabolic pathway, and can offer insights in the discovery of prognostic biomarkers and therapeutic targets, which have been demonstrated in

numerous articles (for example, [34–39]). In this article, we provide an example of how this can be utilized in the analysis of the MMP family enabled by the high-quality genomes across primates.

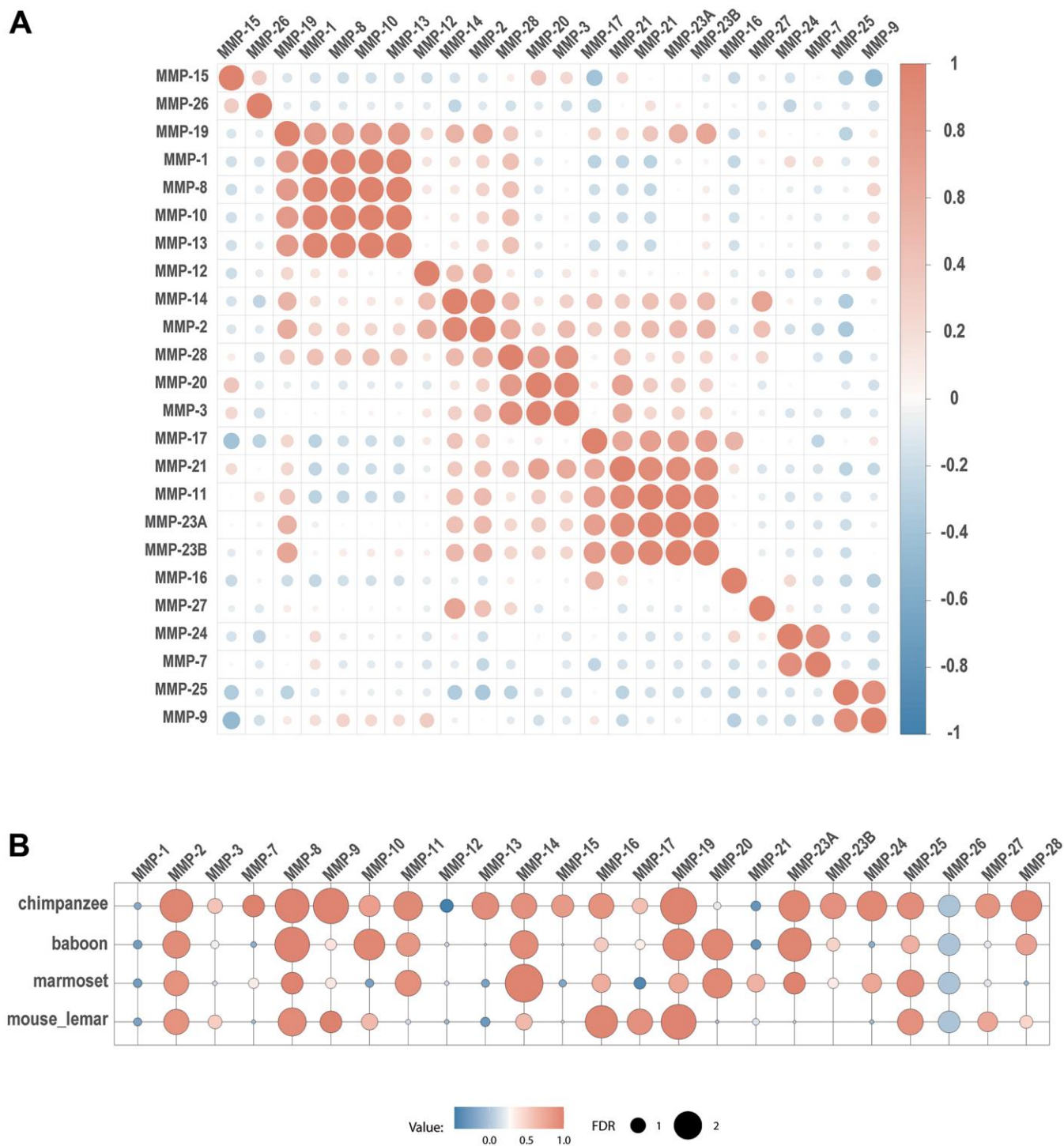


Figure 9. Co-expression of MMPs. (A) Expression correlation between human MMP genes. (B) Expression correlation of each MMP gene between human and four other primate species. Spearman's correlation coefficients are depicted as heatmap, with each circle denoting a pair.

MMPs comprise a family of zinc-dependent endoproteinase, responsible for degrading extracellular matrix components [1]. Deregulation of the MMPs has been implicated in many aging-related diseases, such as Alzheimer's disease, Parkinson's disease, age-related macular degeneration, bone diseases and cancers [6, 9, 11–13, 15–18]. Although biochemical properties of MMPs has been intensively studied, accumulated evidences have shown that our understanding of the MMPs is far from absolute, and some discoveries even have conflicted conclusions with previous findings [23, 40]. In this article, we comprehensively investigate the MMP family across the primates from the perspectives of phylogeny, sequence divergence, domain organization, gene structure, genome organization and gene expression. To our knowledge, this is the first comprehensive investigation of MMPs across primates.

We identified signal peptides at the N-termini and Pfam domains at the C-termini in most MMP peptides. Very limited peptides have conserved domains failing to identify, most of which are located at N- or C-termini (Supplementary Table 2). One possible reason is that they represent bona fide loss of the conserved domains during the evolution of certain primate species. But there is a possibility of incomplete genome assembly and genome annotation artifact which can result in incomplete sequence prediction of these MMPs. This is very common in eukaryotic genome assembly and annotation despite rapid advancement of genome sequencing technology [41, 42].

Our results showed that furin-activatable MMPs are the most variable among different classes of MMPs, including sequence divergence inferred by phylogeny, domain organization and gene structure. There are four types of MMPs within this class, and biochemical properties vary considerably among them [1, 19, 43]. Topology of MMP phylogeny shows that sequence divergence among furin-activatable MMPs is considerably large compared to others. Among the furin-activatable MMPs, MMP-19 showed positive selection in primates with the excess of candidate sites of positive selection. Biochemical and physiological roles of MMP-19 have not been characterized, although MMP-19 is used as a T-cell-derived autoantigen from patients with rheumatoid arthritis [21]. Our results indicate that MMP-19 may have important roles, highlighting its value for further studies. Furthermore, MMP-11, MMP-15 and MMP-17 showed relatively high Ks values than the others.

Our results showed that MMP gene sequences are conserved across primate organisms, and gene organization and genome localization are less conserved as the taxonomic relationships are getting more genetically distant. Additionally, we revealed that MMPs of different types can form a single co-expression cluster. For example, MMP-11 and MMP-21 are secreted MMPs, MMP-17 is a GPI-anchored MMP-, and MMP-23 is a type-II transmembrane MMP- [19] (Figure 9A). Expression patterns of these genes are correlated to each other, suggesting their functional relationships. Another notable co-expression cluster is formed with MMP-1, MMP-8, MMP-10, MMP-13 and MMP-19, all of which are archetypal MMPs. Particularly, all the collagenases (MMP-1, MMP-8 and MMP-13) are co-expressed, which have been shown to play roles of collagenases in wounds and ulcers [44, 45]. These results suggest functional conservation within archetypal MMPs.

Both gelatinases (MMP-2 and MMP-9) are secreted, zinc-dependent endopeptidases and are cell-surface transducers [46]. They are associated with several cancers and thus can be served as cancer biomarkers [46]. Our results showed that the genes are comprised by the longest peptide sequences. However, they showed different characteristics in expression. MMP-2 had high expression across all organs we investigated, while MMP-9 showed tissue-specific expression (Figure 8A). MMP-2 showed conserved expression across all five primate species (Figure 9B), while MMP-9 only showed expression between human and chimpanzee. These results indicate functional divergence between these two gelatinase genes, which deserve further investigations. In addition to MMP-2, several MMP genes show conserved expression across all five primate species (human, chimpanzee, baboon, marmoset and mouse lemur), including MMP-8, MMP-19 and MMP-25. In addition, MMP-11, MMP-14, MMP-23 showed conserved expression in at least three species when compared to human. Less divergence in expression of these genes imply their conserved function across primates. MMP-24, MMP-25 and MMP-27 showed expression in specific organs; MMP-1, MMP-3, MMP-12, MMP-26 showed very weak conservation among the primates; MMP-3 and MMP-26 also showed divergence in sequence. These results indicate their functional divergence among primates.

Life-history diversity are evident among primates. Human life histories have been described as “slow”, as humans have delayed maturity and lower adult mortality compared to chimpanzee [47, 48]. As MMPs

can mediate aging-associated remodeling of extracellular matrix (ECM) [29], we can expect divergence of evolutionary rates among MMPs. In this article, we report positive selection for MMP-26, and evident sequence divergence for MMP-3, MMP-11, MMP-15 and MMP-17. MMP-26 has been indicated to associate with a wide range of pathological process, such as breast cancer, glioma, wounds and ischemic stroke, but lack of evidence highlights necessity of further investigation of this gene [49–52]. We also showed decline of expression for MMP-11 in testis and the same trend for MMP-15 in ovary during human aging. However, expression of MMP-15 and MMP-17 was elevated during aging in testis, suggesting their positive regulatory roles in aging-associated ECM remodeling. These results indicate that certain MMPs such as MMP-11, MMP-15 and MMP-17 may play important roles in life-history variation among primates. Interestingly, several genes including MMP-2, MMP-14 and MMP-15 showed reversed trend expression across ages in testis and ovary; MMP-20 showed longitudinal trend of expression in testis, and MMP-28 only showed trend in ovary. These results suggest sexual difference of functions among these genes.

In addition to cancers, evidence has been reported on associations with psychiatric disorders for some MMPs. For example, MMP-14 can mediate alteration of amyloid precursor proteolysis, a hallmark of Alzheimer's disease, in human neuronal cells [10]. Parkinson's disease has been associated with aberrant expression of MMP-3 and MMP-9 [11, 12]. Our results showed that MMP genes such as MMP-15, MMP-16, MMP-17, MMP-24 and MMP-28 were expressed highly or moderately in brain, implying the values in investigation of more genes in association with psychiatric diseases.

In conclusion, we carried out genome-enabled identification and characterization of putative MMPs across 11 primate organisms from the perspectives of phylogeny, domain organization, gene structure, genome organization and gene expression, and we identified MMPs which need attentions in future studies (As summarized in Figure 10). Our results shed new light on the classification of MMP family, and our findings can illuminate the discovery of prognostic biomarkers and therapeutic targets for aging-related diseases. [29, 47, 48].

Both conserved and divergent		Co-expression clusters	
Gene copy number		MMP-1, MMP-8, MMP-13, MMP-19	
Sequence divergence		MMP-2, MMP-14	
Gene structure		MMP-3, MMP-20, MMP-28	
Domain organization		MMP-11, MMP-17, MMP-21	
Genomic organization		MMP-7, MMP-24	
Expression patterns across primates		MMP-9, MMP-25	

Genes of prospective attention	
MMP-19	Excessive sites of positive selection; two-copies in old-world monkeys
MMP-8	Positive selection
MMP-10	Positive selection
MMP-26	Positive selection
MMP-2, MMP-7, MMP-11, MMP-14, MMP-15, MMP-21, MMP-28	Longitudinally decreased or increased expression in testis or ovary
MMP-15, MMP-16, MMP-17, MMP-24, MMP-28	Possibly involved in brain-related physiology

Figure 10. Summary and prospect of the findings in this article.

MATERIALS AND METHODS

Sources of genome sequence data

Genome sequences, coding sequences, peptides and GFF files were obtained from Ensembl for human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gelada (*Theropithecus gelada*), orangutan (*Pongo abelii*), olive baboon (*Papio anubis*), bonobo (*Pan paniscus*), gibbon (*Nomascus leucogenys*), mouse lemur (*Microcebus murinus*), gorilla (*Gorilla gorilla*) and marmoset (*Callithrix jacchus*) [53]. Genome files of golden snub-nosed monkey (*Rhinopithecus roxellana*) were downloaded from NCBI Genome Browser [54]. Evolutionary relationships of the above species were retrieved from TimeTree [55].

Identification of putative MMPs

Sequences of the human MMPs were retrieved from UCSC genome browser [56]. MMPs from non-human primate species were identified by a combination of BLASTp and HMMER searches. By using BLASTp, searches were performed using an E-value cutoff of $1e-10$ with 24 human MMP peptide sequences as queries [57]. For HMMER, hmmsearch (HMMER v3.3 package) was used to search for Pfam domain PF00413.23 as a profile with default settings [25, 58]. The searching results were combined and redundant sequences were removed. The identified genes were validated by searching the peptide sequences for the Hidden Markov Model profile of Pfam domain PF00413.23, the signature conserved domain of MMPs.

Alignment and phylogenetic inference

Multiple sequence alignments of peptides were performed using MUSCLE v3.8.1551 [59]. Maximum likelihood phylogenetic trees were reconstructed with PhyML v3.3.3 package, with the confidence level of each branch estimated by using 1000 bootstrap replications and WAG model of amino acid evolution [60]. The phylogenetic trees were visualized with Figtree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>). The primate MMP genes identified in this article were classified and designated based on the phylogenetic topology and Fanjul-Fenandex et al. [24].

Evolution analysis of MMPs

Sequence divergence of each MMP clade was analyzed using codon substitution model implemented in CodeML under PAML 4.0 [61]. Two models, M7 and M8, of the distribution across amino acid sites were used, by which M7 assumes that ω is β

distributed along the sequence with a range limited by 0 and 1, while M8 assumes positive selection at some sites and the likelihood ratio test (LRT) of M8 against M7 provides a statistical test for positive selection. Sites of positive selection with P-value above 0.95 ($\omega > 1$) were retrieved. Nonsynonymous substitutions (Ka), synonymous substitutions (Ks) and Ka/Ks ratio between human and gorilla MMP orthologs were also calculated, which was implemented in codeML under PAML 4.0 [61].

Identification of domains and signal peptides

Conserved domains of the peptide sequences were identified by using hmmsearch under HMMER v3.3 package against Pfam-A database with default trusted cutoffs [25, 58]. Signal peptides for the peptides were identified by querying the SignalP-5.0 server [62]. The location of domain and signal peptides were visualized by querying Gene Structure Display Server 2.0 [63].

Gene structure characterization

Peptide molecular weights were calculated by using Bachem's Peptide Calculator (<https://www.bachem.com/knowledge-center/peptide-calculator/>). Gene structure information was retrieved from GFF3 files. Gene length was calculated from the start to the end point of each gene, and CDS length was calculated by deducting intron lengths from gene length. While MMP-23s are classified in furin-activatable MMPs, they demonstrate very different characteristics than other MMPs, and thus were investigated separately.

Genomic localization of MMP genes and co-syteny analysis

Localization of genes was retrieved from GFF3 files. The site of the start codon was considered as the gene localization of each gene. The length of each gene was calculated based on the genome sequences. The visualization of the genomics localization was performed by using MapChart 2.32 [64]. Co-syteny analysis was performed by using the MCScanX package [65].

Gene expression analysis

Data of Illumina's Human BodyMap 2.0 were downloaded from Expression Atlas at EBI (<https://www.ebi.ac.uk/gxa/>), and FPKM values of human MMPs were extracted. Longitudinal expression data for human testis and ovary were retrieved from GTEx Portal [32]. FPKM values of four non-human primates were extracted from non-human primate reference transcriptome resource (NHPRTR) [66].

Spearman's correlation coefficient was calculated for each human gene pair and each MMP gene between primate species, and heatmaps were drawn using R package corrplot [67].

AUTHOR CONTRIBUTIONS

The authors contributed in the following ways: Yinglian Pan: data collection, data analysis, Yadan Fan, Yanda Lu: study design, study supervision; Siyuan Peng: data collection; Haixue Lin: study design; Qingchun Deng: drafting, technical support and final approval of the manuscript. All the authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The data used for the human longitudinal expression analyses of testis and ovary described in this manuscript were obtained from the GTEx Portal on 03/30/2022.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This study was supported by: Dr. Yinglian Pan: Hainan Provincial Health Commission Project (20A200431). Dr. Qingchun Deng: Hainan Provincial Health Commission Project (20A200247) and Young Talents' Science and Technology Innovation Project of Hainan Association for Science and Technology (QCXM202018).

REFERENCES

1. Cui N, Hu M, Khalil RA. Biochemical and Biological Attributes of Matrix Metalloproteinases. *Prog Mol Biol Transl Sci.* 2017; 147:1–73. <https://doi.org/10.1016/bs.pmbts.2017.02.005> PMID:28413025
2. Tokuhara CK, Santesso MR, Oliveira GS, Ventura TM, Doyama JT, Zambuzzi WF, Oliveira RC. Updating the role of matrix metalloproteinases in mineralized tissue and related diseases. *J Appl Oral Sci.* 2019; 27:e20180596. <https://doi.org/10.1590/1678-7757-2018-0596> PMID:31508793
3. Amălinei C, Căruntu ID, Bălan RA. Biology of metalloproteinases. *Rom J Morphol Embryol.* 2007; 48:323–34. PMID:18060181

4. Page-McCaw A, Ewald AJ, Werb Z. Matrix metalloproteinases and the regulation of tissue remodelling. *Nat Rev Mol Cell Biol.* 2007; 8:221–33. <https://doi.org/10.1038/nrm2125> PMID:17318226
5. Li K, Tay FR, Yiu CK. The past, present and future perspectives of matrix metalloproteinase inhibitors. *Pharmacol Ther.* 2020; 207:107465. <https://doi.org/10.1016/j.pharmthera.2019.107465> PMID:31863819
6. Hussain AA, Lee Y, Marshall J. High molecular-weight gelatinase species of human Bruch's membrane: compositional analyses and age-related changes. *Invest Ophthalmol Vis Sci.* 2010; 51:2363–71. <https://doi.org/10.1167/iovs.09-4259> PMID:20007823
7. Kumar A, El-Osta A, Hussain AA, Marshall J. Increased sequestration of matrix metalloproteinases in ageing human Bruch's membrane: implications for ECM turnover. *Invest Ophthalmol Vis Sci.* 2010; 51:2664–70. <https://doi.org/10.1167/iovs.09-4195> PMID:20042661
8. Hussain AA, Lee Y, Marshall J. Understanding the complexity of the matrix metalloproteinase system and its relevance to age-related diseases: Age-related macular degeneration and Alzheimer's disease. *Prog Retin Eye Res.* 2020; 74:100775. <https://doi.org/10.1016/j.preteyeres.2019.100775> PMID:31473329
9. Peress N, Perillo E, Zucker S. Localization of tissue inhibitor of matrix metalloproteinases in Alzheimer's disease and normal brain. *J Neuropathol Exp Neurol.* 1995; 54:16–22. <https://doi.org/10.1097/00005072-199501000-00002> PMID:7815076
10. Llorente P, Martins S, Sastre I, Aldudo J, Recuero M, Adjaye J, Bullido MJ. Matrix Metalloproteinase 14 Mediates APP Proteolysis and Lysosomal Alterations Induced by Oxidative Stress in Human Neuronal Cells. *Oxid Med Cell Longev.* 2020; 2020:5917187. <https://doi.org/10.1155/2020/5917187> PMID:33282112
11. Lech AM, Wiera G, Mozrzymas JW. Matrix metalloproteinase-3 in brain physiology and neurodegeneration. *Adv Clin Exp Med.* 2019; 28:1717–22. <https://doi.org/10.17219/acem/110319> PMID:31851789
12. Lorenzl S, Albers DS, Relkin N, Ngyuen T, Hilgenberg SL, Chirichigno J, Cudkowicz ME, Beal MF. Increased plasma levels of matrix metalloproteinase-9 in patients with Alzheimer's disease. *Neurochem Int.* 2003; 43:191–6.

- [https://doi.org/10.1016/s0197-0186\(03\)00004-4](https://doi.org/10.1016/s0197-0186(03)00004-4)
PMID:[12689599](https://pubmed.ncbi.nlm.nih.gov/12689599/)
13. Paiva KB, Granjeiro JM. Matrix Metalloproteinases in Bone Resorption, Remodeling, and Repair. *Prog Mol Biol Transl Sci.* 2017; 148:203–303.
<https://doi.org/10.1016/bs.pmbts.2017.05.001>
PMID:[28662823](https://pubmed.ncbi.nlm.nih.gov/28662823/)
 14. Coussens LM, Fingleton B, Matrisian LM. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science.* 2002; 295:2387–92.
<https://doi.org/10.1126/science.1067100>
PMID:[11923519](https://pubmed.ncbi.nlm.nih.gov/11923519/)
 15. Liu N, Zhong L, Ni G, Lin J, Xie L, Li T, Dan H, Chen Q. High Matrix Metalloproteinase 28 Expression is Associated with Poor Prognosis in Pancreatic Adenocarcinoma. *Onco Targets Ther.* 2021; 14:4391–406.
<https://doi.org/10.2147/OTT.S309576>
PMID:[34408436](https://pubmed.ncbi.nlm.nih.gov/34408436/)
 16. Kim S, Kang M, Ko J. Small leucine zipper protein promotes the metastasis of castration-resistant prostate cancer through transcriptional regulation of matrix metalloproteinase-13. *Carcinogenesis.* 2021; 42:1089–99.
<https://doi.org/10.1093/carcin/bgab045>
PMID:[34050742](https://pubmed.ncbi.nlm.nih.gov/34050742/)
 17. Wu GJ, Bao JS, Yue ZJ, Zeng FC, Cen S, Tang ZY, Kang XL. Elevated expression of matrix metalloproteinase-9 is associated with bladder cancer pathogenesis. *J Cancer Res Ther.* 2018; 14:S54–9.
<https://doi.org/10.4103/0973-1482.163761>
PMID:[29578150](https://pubmed.ncbi.nlm.nih.gov/29578150/)
 18. Thakur V, Bedogni B. The membrane tethered matrix metalloproteinase MT1-MMP at the forefront of melanoma cell invasion and metastasis. *Pharmacol Res.* 2016; 111:17–22.
<https://doi.org/10.1016/j.phrs.2016.05.019>
PMID:[27221755](https://pubmed.ncbi.nlm.nih.gov/27221755/)
 19. Laronha H, Caldeira J. Structure and Function of Human Matrix Metalloproteinases. *Cells.* 2020; 9:1076.
<https://doi.org/10.3390/cells9051076> PMID:[32357580](https://pubmed.ncbi.nlm.nih.gov/32357580/)
 20. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res.* 2003; 92:827–39.
<https://doi.org/10.1161/01.RES.0000070112.80711.3D>
PMID:[12730128](https://pubmed.ncbi.nlm.nih.gov/12730128/)
 21. Kolb C, Mauch S, Peter HH, Krawinkel U, Sedlacek R. The matrix metalloproteinase RASI-1 is expressed in synovial blood vessels of a rheumatoid arthritis patient. *Immunol Lett.* 1997; 57:83–8.
[https://doi.org/10.1016/s0165-2478\(97\)00057-6](https://doi.org/10.1016/s0165-2478(97)00057-6)
PMID:[9232430](https://pubmed.ncbi.nlm.nih.gov/9232430/)
 22. Li W, Gibson CW, Abrams WR, Andrews DW, DenBesten PK. Reduced hydrolysis of amelogenin may result in X-linked amelogenesis imperfecta. *Matrix Biol.* 2001; 19:755–60.
[https://doi.org/10.1016/s0945-053x\(00\)00121-9](https://doi.org/10.1016/s0945-053x(00)00121-9)
PMID:[11223334](https://pubmed.ncbi.nlm.nih.gov/11223334/)
 23. Chakraborti S, Mandal M, Das S, Mandal A, Chakraborti T. Regulation of matrix metalloproteinases: an overview. *Mol Cell Biochem.* 2003; 253:269–85.
<https://doi.org/10.1023/a:1026028303196>
PMID:[14619979](https://pubmed.ncbi.nlm.nih.gov/14619979/)
 24. Fanjul-Fernández M, Folgueras AR, Cabrera S, López-Otín C. Matrix metalloproteinases: evolution, gene regulation and functional analysis in mouse models. *Biochim Biophys Acta.* 2010; 1803:3–19.
<https://doi.org/10.1016/j.bbamcr.2009.07.004>
PMID:[19631700](https://pubmed.ncbi.nlm.nih.gov/19631700/)
 25. Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M. Pfam: the protein families database. *Nucleic Acids Res.* 2014; 42:D222–30.
<https://doi.org/10.1093/nar/gkt1223>
PMID:[24288371](https://pubmed.ncbi.nlm.nih.gov/24288371/)
 26. Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL, Disotell TR. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Mol Phylogenet Evol.* 2014; 75:165–83.
<https://doi.org/10.1016/j.ympev.2014.02.023>
PMID:[24583291](https://pubmed.ncbi.nlm.nih.gov/24583291/)
 27. Nagase H, Visse R, Murphy G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res.* 2006; 69:562–73.
<https://doi.org/10.1016/j.cardiores.2005.12.002>
PMID:[16405877](https://pubmed.ncbi.nlm.nih.gov/16405877/)
 28. Marmoset Genome Sequencing and Analysis Consortium. The common marmoset genome provides insight into primate biology and evolution. *Nat Genet.* 2014; 46:850–7.
<https://doi.org/10.1038/ng.3042>
PMID:[25038751](https://pubmed.ncbi.nlm.nih.gov/25038751/)
 29. Freitas-Rodríguez S, Folgueras AR, López-Otín C. The role of matrix metalloproteinases in aging: Tissue remodeling and beyond. *Biochim Biophys Acta Mol Cell Res.* 2017; 1864:2015–25.
<https://doi.org/10.1016/j.bbamcr.2017.05.007>
PMID:[28499917](https://pubmed.ncbi.nlm.nih.gov/28499917/)
 30. Gunes S, Hekim GN, Arslan MA, Asci R. Effects of aging on the male reproductive system. *J Assist Reprod Genet.* 2016; 33:441–54.
<https://doi.org/10.1007/s10815-016-0663-y>
PMID:[26867640](https://pubmed.ncbi.nlm.nih.gov/26867640/)

31. Shirasuna K, Iwata H. Effect of aging on the female reproductive function. *Contracept Reprod Med*. 2017; 2:23.
<https://doi.org/10.1186/s40834-017-0050-9>
PMID:[29201428](https://pubmed.ncbi.nlm.nih.gov/29201428/)
32. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, Hasz R, Walters G, Garcia F, Young N, Foster B, Moser M, Karasik E, et al, and GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013; 45:580–5.
<https://doi.org/10.1038/ng.2653>
PMID:[23715323](https://pubmed.ncbi.nlm.nih.gov/23715323/)
33. Ezran C, Karanewsky CJ, Pendleton JL, Sholtz A, Krasnow MR, Willick J, Razafindrakoto A, Zohdy S, Albertelli MA, Krasnow MA. The Mouse Lemur, a Genetic Model Organism for Primate Biology, Behavior, and Health. *Genetics*. 2017; 206:651–64.
<https://doi.org/10.1534/genetics.116.199448>
PMID:[28592502](https://pubmed.ncbi.nlm.nih.gov/28592502/)
34. Dorigatti AO, Hussong SA, Hernandez SF, Sills AM, Salmon AB, Galvan V. Primary neuron and astrocyte cultures from postnatal *Callithrix jacchus*: a non-human primate *in vitro* model for research in neuroscience, nervous system aging, and neurological diseases of aging. *Geroscience*. 2021; 43:115–24.
<https://doi.org/10.1007/s11357-020-00284-z>
PMID:[33063253](https://pubmed.ncbi.nlm.nih.gov/33063253/)
35. Kelberman D, Fife M, Rockman MV, Brull DJ, Woo P, Humphries SE. Analysis of common IL-6 promoter SNP variants and the AnTn tract in humans and primates and effects on plasma IL-6 levels following coronary artery bypass graft surgery. *Biochim Biophys Acta*. 2004; 1688:160–7.
<https://doi.org/10.1016/j.bbadis.2003.11.010>
PMID:[14990346](https://pubmed.ncbi.nlm.nih.gov/14990346/)
36. Roy S, Vandenberghe LH, Kryazhimskiy S, Grant R, Calcedo R, Yuan X, Keough M, Sandhu A, Wang Q, Medina-Jaszek CA, Plotkin JB, Wilson JM. Isolation and characterization of adenoviruses persistently shed from the gastrointestinal tract of non-human primates. *PLoS Pathog*. 2009; 5:e1000503.
<https://doi.org/10.1371/journal.ppat.1000503>
PMID:[19578438](https://pubmed.ncbi.nlm.nih.gov/19578438/)
37. Van Dam D, De Deyn PP. Non human primate models for Alzheimer's disease-related research and drug discovery. *Expert Opin Drug Discov*. 2017; 12:187–200.
<https://doi.org/10.1080/17460441.2017.1271320>
PMID:[27960560](https://pubmed.ncbi.nlm.nih.gov/27960560/)
38. Willett CS, Wilson EM. Evolution of Melanoma Antigen-A11 (MAGEA11) During Primate Phylogeny. *J Mol Evol*. 2018; 86:240–53.
<https://doi.org/10.1007/s00239-018-9838-8>
PMID:[29574604](https://pubmed.ncbi.nlm.nih.gov/29574604/)
39. Yasukochi Y, Satta Y. Evolution of the CYP2D gene cluster in humans and four non-human primates. *Genes Genet Syst*. 2011; 86:109–16.
<https://doi.org/10.1266/ggs.86.109>
PMID:[21670550](https://pubmed.ncbi.nlm.nih.gov/21670550/)
40. Tam EM, Moore TR, Butler GS, Overall CM. Characterization of the distinct collagen binding, helicase and cleavage mechanisms of matrix metalloproteinase 2 and 14 (gelatinase A and MT1-MMP): the differential roles of the MMP hemopexin c domains and the MMP-2 fibronectin type II modules in collagen triple helicase activities. *J Biol Chem*. 2004; 279:43336–44.
<https://doi.org/10.1074/jbc.M407186200>
PMID:[15292230](https://pubmed.ncbi.nlm.nih.gov/15292230/)
41. Heydari M, Miclotte G, Van de Peer Y, Fostier J. Illumina error correction near highly repetitive DNA regions improves de novo genome assembly. *BMC Bioinformatics*. 2019; 20:298.
<https://doi.org/10.1186/s12859-019-2906-2>
PMID:[31159722](https://pubmed.ncbi.nlm.nih.gov/31159722/)
42. Howe K, Chow W, Collins J, Pelan S, Pointon DL, Sims Y, Torrance J, Tracey A, Wood J. Significantly improving the quality of genome assemblies through curation. *Gigascience*. 2021; 10:giaa153.
<https://doi.org/10.1093/gigascience/giaa153>
PMID:[33420778](https://pubmed.ncbi.nlm.nih.gov/33420778/)
43. Liu X, Liu D, Shen Y, Huang M, Gao L, Zhang Z, Liu B, Zhao B, Pang Q. Matrix Metalloproteinases in Invertebrates. *Protein Pept Lett*. 2020; 27:1068–81.
<https://doi.org/10.2174/0929866527666200429110945> PMID:[32348206](https://pubmed.ncbi.nlm.nih.gov/32348206/)
44. Donahue TR, Hiatt JR, Busuttill RW. Collagenase and surgical disease. *Hernia*. 2006; 10:478–85.
<https://doi.org/10.1007/s10029-006-0146-7>
PMID:[16977344](https://pubmed.ncbi.nlm.nih.gov/16977344/)
45. Patry J, Blanchette V. Enzymatic debridement with collagenase in wounds and ulcers: a systematic review and meta-analysis. *Int Wound J*. 2017; 14:1055–65.
<https://doi.org/10.1111/iwj.12760>
PMID:[28440050](https://pubmed.ncbi.nlm.nih.gov/28440050/)
46. Bauvois B. New facets of matrix metalloproteinases MMP-2 and MMP-9 as cell surface transducers: outside-in signaling and relationship to tumor progression. *Biochim Biophys Acta*. 2012; 1825:29–36.
<https://doi.org/10.1016/j.bbcan.2011.10.001>
PMID:[22020293](https://pubmed.ncbi.nlm.nih.gov/22020293/)
47. Davison RJ, Gurven MD. Human uniqueness? Life history diversity among small-scale societies and chimpanzees. *PLoS One*. 2021; 16:e0239170.
<https://doi.org/10.1371/journal.pone.0239170>
PMID:[33617556](https://pubmed.ncbi.nlm.nih.gov/33617556/)

48. Burger O, Baudisch A, Vaupel JW. Human mortality improvement in evolutionary context. *Proc Natl Acad Sci USA*. 2012; 109:18210–4.
<https://doi.org/10.1073/pnas.1215627109>
PMID:[23071331](https://pubmed.ncbi.nlm.nih.gov/23071331/)
49. Guo JG, Guo CC, He ZQ, Cai XY, Mou YG. High MMP-26 expression in glioma is correlated with poor clinical outcome of patients. *Oncol Lett*. 2018; 16:2237–42.
<https://doi.org/10.3892/ol.2018.8880>
PMID:[30008924](https://pubmed.ncbi.nlm.nih.gov/30008924/)
50. Hsieh FI, Chiou HY, Hu CJ, Jeng JS, Lin HJ, Lee JT, Lien LM. Combined Effects of MMP-7, MMP-8 and MMP-26 on the Risk of Ischemic Stroke. *J Clin Med*. 2019; 8:2011.
<https://doi.org/10.3390/jcm8112011>
PMID:[31752174](https://pubmed.ncbi.nlm.nih.gov/31752174/)
51. Pirilä E, Korpi JT, Korkiamäki T, Jahkola T, Gutierrez-Fernandez A, Lopez-Otin C, Saarialho-Kere U, Salo T, Sorsa T. Collagenase-2 (MMP-8) and matrilysin-2 (MMP-26) expression in human wounds of different etiologies. *Wound Repair Regen*. 2007; 15:47–57.
<https://doi.org/10.1111/j.1524-475X.2006.00184.x>
PMID:[17244319](https://pubmed.ncbi.nlm.nih.gov/17244319/)
52. Piskór BM, Przyłipiak A, Dąbrowska E, Sidorkiewicz I, Niczyporuk M, Szmitkowski M, Ławicki S. Plasma Concentrations of Matrilysins MMP-7 and MMP-26 as Diagnostic Biomarkers in Breast Cancer. *J Clin Med*. 2021; 10:1436.
<https://doi.org/10.3390/jcm10071436>
PMID:[33916127](https://pubmed.ncbi.nlm.nih.gov/33916127/)
53. Herrero J, Muffato M, Beal K, Fitzgerald S, Gordon L, Pignatelli M, Vilella AJ, Searle SM, Amode R, Brent S, Spooner W, Kulesha E, Yates A, Flicek P. Ensembl comparative genomics resources. *Database (Oxford)*. 2016; 2016:baw053.
<https://doi.org/10.1093/database/baw053>
PMID:[27141089](https://pubmed.ncbi.nlm.nih.gov/27141089/)
54. Kuznetsov A, Bollin CJ. NCBI Genome Workbench: Desktop Software for Comparative Genomics, Visualization, and GenBank Data Submission. *Methods Mol Biol*. 2021; 2231:261–95.
https://doi.org/10.1007/978-1-0716-1036-7_16
PMID:[33289898](https://pubmed.ncbi.nlm.nih.gov/33289898/)
55. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol*. 2017; 34:1812–9.
<https://doi.org/10.1093/molbev/msx116>
PMID:[28387841](https://pubmed.ncbi.nlm.nih.gov/28387841/)
56. Karolchik D, Hinrichs AS, Kent WJ. The UCSC Genome Browser. *Curr Protoc Bioinformatics*. 2012; 40:1.4.1–33.
<https://doi.org/10.1002/0471250953.bi0104s40>
PMID:[23255150](https://pubmed.ncbi.nlm.nih.gov/23255150/)
57. Johnson M, Zaretskaya I, Raytselis Y, Merezhuik Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res*. 2008; 36:W5–9.
<https://doi.org/10.1093/nar/gkn201>
PMID:[18440982](https://pubmed.ncbi.nlm.nih.gov/18440982/)
58. Prakash A, Jeffryes M, Bateman A, Finn RD. The HMMER Web Server for Protein Sequence Similarity Search. *Curr Protoc Bioinformatics*. 2017; 60:3.15.1–23.
<https://doi.org/10.1002/cpbi.40>
PMID:[29220076](https://pubmed.ncbi.nlm.nih.gov/29220076/)
59. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32:1792–7.
<https://doi.org/10.1093/nar/gkh340>
PMID:[15034147](https://pubmed.ncbi.nlm.nih.gov/15034147/)
60. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003; 52:696–704.
<https://doi.org/10.1080/10635150390235520>
PMID:[14530136](https://pubmed.ncbi.nlm.nih.gov/14530136/)
61. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007; 24:1586–91.
<https://doi.org/10.1093/molbev/msm088>
PMID:[17483113](https://pubmed.ncbi.nlm.nih.gov/17483113/)
62. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019; 37:420–3.
<https://doi.org/10.1038/s41587-019-0036-z>
PMID:[30778233](https://pubmed.ncbi.nlm.nih.gov/30778233/)
63. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 2015; 31:1296–7.
<https://doi.org/10.1093/bioinformatics/btu817>
PMID:[25504850](https://pubmed.ncbi.nlm.nih.gov/25504850/)
64. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002; 93:77–8.
<https://doi.org/10.1093/jhered/93.1.77>
PMID:[12011185](https://pubmed.ncbi.nlm.nih.gov/12011185/)
65. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012; 40:e49.
<https://doi.org/10.1093/nar/gkr1293>
PMID:[22217600](https://pubmed.ncbi.nlm.nih.gov/22217600/)
66. Peng X, Thierry-Mieg J, Thierry-Mieg D, Nishida A, Pipes L, Bozinoski M, Thomas MJ, Kelly S, Weiss JM, Raveendran M, Muzny D, Gibbs RA, Rogers J, et al. Tissue-specific transcriptome sequencing analysis

expands the non-human primate reference transcriptome resource (NHPRTR). *Nucleic Acids Res.* 2015; 43:D737–42.
<https://doi.org/10.1093/nar/gku1110> PMID:[25392405](https://pubmed.ncbi.nlm.nih.gov/25392405/)

67. Hahsler M, Hornik K, Buchta C. Getting things in order: An introduction to the R package seriation. *J Stat Softw.* 2008; 25:1–34.
<https://doi.org/10.18637/jss.v025.i03>

SUPPLEMENTARY MATERIALS

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–3.

Supplementary Table 1. List of MMP genes in 11 primate species.

Supplementary Table 2. Function information of Pfam domains in MMPs, which is extracted from Pfam database.

Supplementary Table 3. Signal peptides and Pfam domains in primate MMPs.