

## **Online Supplementary Materials**

**Journal:** Fisheries and Aquatic Sciences

**Title:** Diversification and domain evolution of molluscan metallothioneins – a mini review

**Author:** Yoon Kwon Nam (Corresponding author), Eun Jeong Kim

## Legends for Supplementary Materials

### Suppl. Table S1

Summarized information on MT genes and species used for analysis in this study.

### Suppl. Fig. S1.

Structural variations of MT proteins in vertebrates (mammals and teleosts; **A**) and molluscs (bivalves and gastropods; **B**) as addressed by polypeptide lengths, theoretical pI values and cysteine contents. For information on MT abbreviations and species, see **Suppl. Table S1**.

### Suppl. Fig. S2.

Common and non-canonical domain structures of gastropod MTs. Commonly most gastropod MTs represent  $\beta_2\beta_1$ -structure, while exceptionally certain MTs display tandem duplication of  $\beta_2$ -domains, giving rise to larger MTs with more than two metal-binding domains, exemplified by *L. littorea* (Caenogastropoda) MT (*Llit*-MT) and *B. glabrata* (Heterobranchia) MT (*Bgla*-novel-MT). Information on species and MT abbreviations can be referred to **Suppl. Table S1**.

### Suppl. Fig. S3.

Domain structures of Ostreidae (Bivalve) MT isoform families. **(A)** For MT-I isoform family, Ostreidae species represent commonly the prototypic  $\alpha\beta_1$ -structured MT. On the other hand, the domain structure of MT-II isoform family is species-specific. *C. virginica* represent tandem duplication(s) of solely the  $\alpha$ -domain without  $\beta$ -domain (here exemplified by *C. virginica* MT-IIH isoform). However, on the contrary, other Ostreidae species display the duplication of  $\beta_1$ -domain, giving rise to  $\alpha\beta_1\beta_1$ -structured MTs. **(B)** MT-III isoforms from *Crassostrea* oysters are featured by the  $\beta_2\beta_2$ -domain structure, while *Crassostrea* MT-IVs represent a modified  $\alpha\beta_1$ -structure showing additional acquisition of cysteine residues in both  $\alpha$  and  $\beta_1$ -domains. Information on species and MT abbreviations can be referred to **Suppl. Table S1**.

#### **Suppl. Fig. S4.**

The  $\alpha\beta_1$ -domain structure of Mytilidae (Bivalve) MTs. Mytilidae MTs can be grouped into intronless MT10B group, major MT10 group and MT20 group. And the major MT10 group can be further divided into three subgroups (Mytilid MT10, *P. viridis* MT10-I, and *P. viridis* MT10-II). In the alignment, some potentially MT group(s)-specific amino acid residues are indicated with section sign (§), dagger (†) and double dagger (‡) marks. Information on species and MT abbreviations can be referred to **Suppl. Table S1**.

#### **Suppl. Fig. S5.**

(A) The  $\alpha\beta_1$ -domain structure of MTs from bivalve species belonging to various taxonomic positions excluding the families Ostreidae and Mytilidae. Information on species and MT abbreviations can be referred to **Suppl. Table S1**. In addition to the MTs showing the canonical cysteine arrangement pattern (21 cysteines in  $\alpha\beta_1$ -structure), there are also various variant MT forms with altered cysteine motif patterns.

(B) Proposed domain structures of larger MTs from *A. irradians* (both MT1 and MT2 paralogs) and *P. coreanum*. Based on the cysteine arrangement pattern, *A. irradians* MTs are likely to have undergone the duplication of  $\beta_2$  domains followed by further divergent differentiations, giving rise to multiple  $\beta_2$ -domains. Within this context of this hypothesis, the domains structures of *A. irradians* MT1 and MT2 are designated to be  $\beta_2\beta_2\beta_2\beta_1$ -like and  $\beta_2\beta_2\beta_1$ -like shapes, respectively. On the other hand, *P. coreanum* MT is thought to have experienced the duplication of  $\beta_1$ -domain, giving rise to  $\alpha\beta_1\beta_1$ -like structure. For both species, the duplicated domains within a MT protein share little sequence homology. For each MT isoform, non-canonical cysteine residues are indicated by arrows.

#### **Suppl. Fig. S6.**

Analysis of duplicated  $\beta_2$ -domains (1<sup>st</sup> and 2<sup>nd</sup>  $\beta_2$ -domains; numbered from N-terminal) from the newly recognized *B. glabrata* MT (designated *Bgla*-novel MT) in comparison with corresponding  $\beta_2$ -domains from its own and other pulmonate Cu-MT, Cd-MT and Cd/Cu-MT isoforms. Information on species and MT abbreviations can be referred to **Suppl. Table S1**. From the multiple sequence alignment (*upper*), the duplicated  $\beta_2$ -domains of the *Bgla*-novel MT show more sequence homology to Cd-MTs than to Cu-MTs and Cd/Cu-MTs. Phylogenetic NJ-trees drawn with MEGA7 (*lower*) also suggest that  $\beta_2$ -domains of the *Bgla*-novel MT represent a closer relationship with those of Cd-MTs than with those of other MT types. Taken together, the duplication of multiple  $\beta_2$ -domains in the *Bgla*-novel MT has likely been a selective process towards needs to prepare multiple metal-binding domains for better detoxification functions of non-essential metals (*e.g.*, primarily Cd).

**Suppl. Fig. S7.**

Reconstructed neighbor-joining tree to show the phylogenetic relationship among  $\beta$ -domains of Ostreidae MTs. Information on species and MT abbreviations can be referred to **Suppl.**

**Table S1.** Ostreidae MT-I, MT-II (excluding *C. virginica* MT-IIIs) and MT-IV possess  $\beta_1$ -domain(s), giving rise to their overall shape of  $\alpha\beta_1$ - or  $\alpha\beta_1\beta_1$ -domain structure. On the other hand, Ostreidae MT-IIIs (yet only two *Crassostrea* species reported) show a  $\beta_2\beta_2$ -domain structure. Based on the bootstrap tests (1,000 replicates), clades supported by the confidence level higher than 50% are visualized with MEGA7 program. For multiple sequence alignment data, refer to **Suppl. Fig. S3.**

**Suppl. Fig. S8.**

Reconstructed neighbor-joining tree to show the phylogenetic relationship among  $\alpha$ -domains of bivalve MTs. Information on species and MT abbreviations can be referred to **Suppl.**

**Table S1.** Based on the bootstrap tests (1,000 replicates), clades supported by the confidence level higher than 50% are visualized with MEGA7 program.

**Suppl. Fig. S9.**

Reconstructed neighbor-joining tree to show the phylogenetic relationship among  $\beta$ -domains (all  $\beta_1$ -domains) of Mytilidae MTs. Information on species and MT abbreviations can be referred to **Suppl. Table S1.** Based on the bootstrap tests (1,000 replicates), clades are visualized with the condensed tree cutoff value at 50% using the MEGA7 program. *Mytilus* MT10 and *P. viridis* MT10 clades are separately placed, suggesting their earlier divergence. See also the **Suppl. Fig. S4** for multiple sequence alignments of Mytilidae MTs.

**Supplementary Table S1A** Summarized information on molluscan MT genes and species used for analysis in this study

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Biomphalaria glabrata</i>	<b>Gastropoda</b> ; Heterobranchia; Euthyneura; Panpulmonata; Hygrophila; Planorboidea; <b>Planorbidae</b> ;	CdMT	GQ205374	Bgla-CdMT	
		CuMT	GQ205373	Bgla-CuMT	
		Cd/Cu MT	GQ205375	Bgla-Cd/CuMT	
		MT	XP_013080485	Bgla-novel MT	Genomic scaffold (NW_013326649)
<i>Physa acuta</i>	<b>Gastropoda</b> ; Heterobranchia; Euthyneura; Panpulmonata; Hygrophila; Planorboidea; <b>Physidae</b> ;	MT	GU259686	Pacu-MT	
<i>Cepaea hortensis</i>	<b>Gastropoda</b> ; Heterobranchia; Euthyneura; Panpulmonata; Eupulmonata; Stylommatophora; Sigmurethra; Helicoidea; <b>Helicidae</b>	Intermediate MT	GQ996716	Chor-MT	
<i>Helix aspersa</i>		CuMT	EF178297	Hasp-CuMT	
<i>Helix pomatia</i>		CdMT	EF152281	Hasp-CdMT	Midgut gland Cd-MT
		Cd/Cu MT1	GU111728	Hpom-Cd/CuMT1	
		Cd/Cu MT2	GU111729	Hpom-Cd/CuMT2	
		CdMT	FJ755002	Hpom-CdMT	
	CuMT	AF399741	Hpom-CuMT		
<i>Nesiohelix samarangae</i>	<b>Gastropoda</b> ; Heterobranchia; Euthyneura; Panpulmonata; Eupulmonata; Stylommatophora; Sigmurethra; Helicoidea; <b>Bradybaenidae</b>	MT	EU437399	Nsam-MT	
<i>Haliotis discus hannai</i>	<b>Gastropoda</b> ; Vetigastropoda; Haliotoidea; <b>Haliotidae</b> ;	MT	KT895222	Hdis-MT	
<i>Haliotis diversicolor supertexta</i>		MT	EU071824	Hdiv-MT	
<i>Haliotis gigantea</i>		MT	Present study	Hgig-MT	
<i>Haliotis sieboldii</i>		MT	Present study	Hsie-MT	
<i>Megathura crenulata</i>	<b>Gastropoda</b> ; Vetigastropoda; Fissurelloidea; <b>Fissurellidae</b> ;	MT	AY102647	Mcre-MT	
<i>Littorina littorea</i>	<b>Gastropoda</b> ; Caenogastropoda; Hypsogastropoda; Littorinimorpha; Littorinoidea; <b>Littorinidae</b> ;	MT	AY034179	Llit-MT	
<i>Bathymodiolus azoricus</i>	<b>Bivalvia</b> ; Pteriomorpha; Mytiloidea; Mytiloidea; <b>Mytilidae</b>	MT10a	AJ630215	Bazo-MT10a	
		MT10b	AJ630216	Bazo-MT10b	Redundant AA with MT10c (AJ630217)
<i>Mytilus edulis</i>		MT-10	AJ577130	Medu-MT-10	
		MT-10A(1)	AJ577124	Medu-MT-10A(1)	
		MT-10A(2)	AJ577125	Medu-MT-10A(2)	
		MT-10B(1)	AJ577126	Medu-MT-10B(1)	Intronless
		MT-10B(2)	AJ577127	Medu-MT-10B(2)	Intronless
		MT-10Ia	AJ005451	Medu-MT-10Ia	
		MT-10Ib	AJ005452	Medu-MT-10Ib	
		MT-10II	AJ005453	Medu-MT-10II	
		MT-10III	AJ005454	Medu-MT-10III	
		MT-10IV	AJ007506	Medu-MT-10IV	
		MT-20	AJ577131	Medu-MT-20	
		MT-20II	AJ005456	Medu-MT-20II	

Supplementary Table S1A Continued

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Mytilus galloprovincialis</i>	<b>Bivalvia</b> ; Pteriomorphaia; Mytiloidea; Mytiloidea; <b>Mytilidae</b>	MT-I	AF199020	Mgal-MT-I	
		MT-10IIIa	AY566248	Mgal-MT-10IIIa	
		MT-10III	HQ681034	Mgal-MT-10III	
		MT-10B	DQ848984	Mgal-MT-10B	Intronless (HQ681035)
		MT-20IV	AY566247	Mgal-MT-20IV	
		MT-20	HQ681036	Mgal-MT-20	
<i>Mytilus sp. KL-2006</i>		MT-10	EF140765	Msp-MT-10	
		MT-20II	EF140766	Msp-MT-20II	
<i>Perna viridis</i>		MT	AF036904	Pvir-MT	
		MT-IA	JN596471	Pvir-MT-IA	
		MT-IB	JN596472	Pvir-MT-IB	
		MT-IC	JN596473	Pvir-MT-IC	
		MT-ID	JN596474	Pvir-MT-ID	
		MT-IE	JN596475	Pvir-MT-IE	
		MT-IF	JN596476	Pvir-MT-IF	
		MT-IIA	JN596477	Pvir-MT-IIA	
		MT-IIB	JN596478	Pvir-MT-IIB	
		MT-IIC	JN596479	Pvir-MT-IIC	
		MT-IID	JN596480	Pvir-MT-IID	
<i>Crassostrea angulata</i>	<b>Bivalvia</b> ; Pteriomorphaia; Ostreoida; Ostreoida; <b>Ostreidae</b>	MT	AF349907	Cang-MT	
<i>Crassostrea ariakensis</i>		MT	DQ342281	Cari-MT	
		Ca29 MT	JF919324	Cari-Ca29MT	
		Ca28 MT-IV	JF919323	Cari-Ca28MT-IV	
<i>Crassostrea gigas</i>		MT gene	AJ242657	Cgig-MT-I	MT-I
		MT	JF781299	Cgig-MT-III	MT-III
		MT-like	XM_011422542	Cgig-MT-II(1)	MT-II; Redundant AA with XP_011420844, EKC17900,
		MT2	AJ297818	Cgig-MT-II(2)	MT-II; Redundant AA with AJ243263.1
		mt3	AJ295157	Cgig-mt3	MT-I variant
		MT-IV	AM265551	Cgig-MT-IV(1)	MT-IV
		MT-II	EKC41855	Cgig-MT-II(3)	MT-II (genomic scaffold; JH819078)
		MT-IV-like	EKC32371	Cgig-MT-IV(2)	MT-IV like (genomic scaffold; JH816574)
MT-IV-like		EKC28510	Cgig-MT-IV(3)	MT-IV like(genomic scaffold; JH818394)	

Supplementary Table S1A Continued

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Crassostrea rivularis</i>	<b>Bivalvia</b> ; Pteriomorpha; Ostreoida; Ostreoida; <b>Ostreidae</b>	MT	JN225502	Criv-MT	
<i>Crassostrea virginica</i>		MTA	AF506977	Cvir-MTA	
		MT-IA	AY331695	Cvir-MT-IA	Redundant AA with X59862, AY331696, AY331697, AY331698
		MT-IB	AY331699	Cvir-MT-IB	
		MT-IIA	AY331700	Cvir-MT-IIA	
		MT-IIB	AY331701	Cvir-MT-IIB	
		MT-IIC	AY331702	Cvir-MT-IIC	
		MT-IID	AY331703	Cvir-MT-IID	
		MT-IIE	AY331704	Cvir-MT-IIE	
		MT-IIF	AY331705	Cvir-MT-IIF	
		MT-IIG	AY331706	Cvir-MT-IIG	
		MT-IIH	AY331707	Cvir-MT-IIH	
		MT-IIIA	AY331708	Cvir-MT-IIIA	
		MT-IIIB	AY331709	Cvir-MT-IIIB	
		MT-IIIC	AY331710	Cvir-MT-IIIC	
		MT-IVA	DQ117912	Cvir-MT-IVA	
		MT-IVB	DQ117913	Cvir-MT-IVB	
		MT-IVC	DQ117914	Cvir-MT-IVC	
<i>Crassostrea sp.</i> HZ-2013		MT	AB775344	Csp-MT	
<i>Ostrea edulis</i>		MTa	AJ306366	Oedu-MTa	
		MTb	AJ306365	Oedu-MTb	
<i>Alectryonella plicatula</i>		MT	KP875559	Apli-MT	
<i>Pinctada martensi</i>	<b>Bivalvia</b> ; Pteriomorpha; Pterioidea; Pterioidea; <b>Pteriidae</b>	MT	KC197172	Pmar-MT1	
		MT2	KC832833	Pmar-MT2	
<i>Pinctada maxima</i>		MT	FJ389580	Pmax-MT	
<i>Scapharca broughtonii</i>	<b>Bivalvia</b> ; Pteriomorpha; Arcoidea; Arcoidea; <b>Arcidae</b>	MT	FJ154101	Sbro-MT	
<i>Tegillarca granosa</i>		MT	AY568678	Tgra-MT	
<i>Argopecten irradians</i>	<b>Bivalvia</b> ; Pteriomorpha; Pectinoidea; Pectinoidea; <b>Pectinidae</b>	MT1	EF093795	Aira-MT1	
		MT2	EU734181	Aira-MT2	
<i>Laternula elliptica</i>	<b>Bivalvia</b> ; Anomalodesmata; Pholadomyoidea; Pandoroidea; Laternulidae	MT10a	DQ832722	Lell-MT10a	
		MT10b	DQ832723	Lell-MT10b	
<i>Cerastoderma edule</i>	<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Cardioidea; <b>Cardiidae</b> ;	MT	HF947027	Cedu-MT	Redundant AA with EF520705
<i>Cerastoderma glaucum</i>		MT	HF947022	Cgla-MT	

Supplementary Table S1A Continued

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Corbicula fluminea</i>	<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Corbiculoidea; <b>Corbiculidae</b>	MT	EF185126	Cflu-MT	
<i>Pisidium coreanum</i>	<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Corbiculoidea; <b>Sphaeriidae</b>	MT	GQ268325	Pcor-MT	
<i>Dreissena polymorpha</i>	<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Dreissenoidea; <b>Dreissenidae</b>	MT	U67347	Dpol-MT	
<i>Cyclina sinensis</i>	<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Veneroidea; <b>Veneridae</b>	MT	HM246244	Csin-MT	
<i>Mercenaria mercenaria</i>		MT	JQ691633	Mmerc-MT	
<i>Meretrix lusoria</i>		MT	AY525635	Mlus-MT	
<i>Meretrix meretrix</i>		MT	GU233466	Mmere-MT	
<i>Meretrix lamarckii</i>		MT	KT448843	Mlam-MT	
<i>Ruditapes philippinarum</i>		MT1	HQ419062	Rphi-MT1	
		MT2	KF214789	Rphi-MT2	
<i>Macra veneriformis</i>		<b>Bivalvia</b> ; Heteroconchia; Euheterodonta; Veneroidea; Mactroidea; <b>Mactridae</b>	MT	FJ611963.1	Mven-MT
<i>Hyriopsis cumingii</i>	<b>Bivalvia</b> ; Palaeoheterodonta; Unionoidea; Unionoidea; <b>Unionidae</b>	MT	FJ861993	Hcum-MT1	Intronless
		MT	GQ184290	Hcum-MT2	
<i>Hyriopsis schlegelii</i>		MT1	KJ019820	Hsch-MT1	
		MT2	KJ019821	Hsch-MT2	
<i>Unio tumidus</i>		MT	EF185127	Utum-MT	



**Supplementary Table S1B** Summarized information on the representative mammalian MTs used for the comparison of structural parameters with molluscan MTs

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Homo sapiens</i>	<b>Mammalia</b> ; Euarchontoglires; Primates; <b>Hominidae</b>	MT-IA	NP_005937	hMT-IA	Human
		MT-IB	NP_005938	hMT-IB	
		MT-IE	P04732	hMT-IE	
		MT-IF	NP_005940	hMT-IF	
		MT-IG	NP_005941	hMT-G	
		MT-IH	NP_005942	hMT-IH	
		MT-IJ	NP_783321	hMT-IJ	
		MT-IK(1)	P80296	hMT-IK(1)	
		MT-IK(2)	AAH28280	hMT-IK(2)	
		MT-IK(3)	NP_789846	hMT-IK(3)	
		MT-IL	P80297	hMT-IL	
		MT-IM	AAL83902	hMT-IM	
		MT-IQ	AAO49186	hMT-IQ	
		MT-IR	Q93083	hMT-IR	
		MT-IS	AAK26162	hMT-IS	
		MT-IX(1)	NP_005943	hMT-IX(1)	
		MT-IX(2)	AAH18190	hMT-IX(2)	
		MT-II(1)	P02795	hMT-II(1)	
		MT-II(2)	P80295	hMT-II(2)	
		MT-III	P25713	hMT-III	
MT-IV	P47944	hMT-IV			
<i>Cercopithecus aethiops</i>	<b>Mammalia</b> ; Euarchontoglires; Primates; <b>Cercopithecidae</b>	MT-I	P02797	CaeMT-I	Grivet
		MT-II	P02796	CaeMT-II	
<i>Oryctolagus cuniculus</i>	<b>Mammalia</b> ; Euarchontoglires; Lagomorpha; <b>Leporidae</b>	MT-I	AAA31147	OcuMT-I	Rabbit
<i>Mus musculus</i>	<b>Mammalia</b> ; Euarchontoglires; Rodentia; <b>Muridae</b>	MT-I	P02802	MmuMT-I	Mouse
		MT-II(1)	P02798	MmuMT-II(1)	
		MT-II(2)	SMMS2	MmuMT-II(2)	
		MT-III	P28184	MmuMT-III	
		MT-IV	P47945	MmuMT-IV	
<i>Rattus norvegicus</i>	<b>Mammalia</b> ; Euarchontoglires; Rodentia; <b>Muridae</b>	MT-I	P02803	RnoMT-I	Rat
		MT-II	P04355	RnoMT-II	
		MT-III	P37361	RnoMT-III	
<i>Cricetulus griseus</i>	<b>Mammalia</b> ; Euarchontoglires; Rodentia; <b>Cricetidae</b>	MT-I	P02804	CgrMT-I	Hamster
MT-II		P02799	CgrMT-II		
<i>Cricetulus longicaudatus</i>		MT-II	I48116	ClgMT-II	Long-tail hamster
<i>Mesocricetus auratus</i>		MT-II	P17808	MauMT-II	Golden hamster

Supplementary Table S1B Continued

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Bos taurus</i>	<b>Mammalia</b> ; Laurasiatheria; Cetartiodactyla; Ruminantia; <b>Bovidae</b>	MT-I(1)	P58280	BtaMT-I(1)	Cow
		MT-I(2)	P55942	BtaMT-I(2)	
		MT-II(1)	P09579	BtaMT-II(1)	
		MT-II(2)	P55943	BtaMT-II(2)	
		MT-II(3)	SMBO2	BtaMT-II(3)	
		MT-III	P37359	BtaMT-III	
<i>Ovis aries</i>		MT-IA	S00808	OarMT-IA	Sheep
		MT-IB	P09577	OarMT-IB	
		MT-IC	P09578	OarMT-IC	
		MT-II	S00811	OarMT-II	
		MT-III	AAM21134	OarMT-III	
<i>Sus scrofa</i>	<b>Mammalia</b> ; Laurasiatheria; Cetartiodactyla; Suina; <b>Suidae</b>	MT-IA	P49068	SscMT-IA	Pig
		MT-IC	P79376	SscMT-IC	
		MT-ID	P79377	SscMT-ID	
		MT-IE	P79431	SscMT-IE	
		MT-IF	P79378	SscMT-IF	
		MT-IIA	P79379	SscMT-IIA	
		MT-IIB	P79380	SscMT-IIB	
		MT-III	P55944	SscMT-III	
<i>Stenella coeruleoalba</i>	<b>Mammalia</b> ; Laurasiatheria; Cetartiodactyla; Cetacea; <b>Delphinidae</b>	MT-II	P14425	ScoMT-II	Dolphin
<i>Equus caballus</i>	<b>Mammalia</b> ; Laurasiatheria; Perissodactyla; <b>Equidae</b>	MT-IA(1)	P02800	EcaMT-IA(1)	Horse
		MT-IA(2)	SMHO1A	EcaMT-IA(2)	
		MT-IB(1)	P02801	EcaMT-IB(1)	
		MT-IB(2)	SMHOB	EcaMT-IB(2)	
		MT-III	P37360	EcaMT-III	
<i>Canis familiaris</i>	<b>Mammalia</b> ; Laurasiatheria; Carnivora; <b>Canidae</b>	MT-I	O19000	CfaMT-I	Dog
		MT-II	Q9XST5	CfaMT-II	
		MT-IV	Q9TUI5	CfaMT-IV	

**Supplementary Table S1C** Summarized information on the representative fish MTs used for the comparison of structural parameters with molluscan MTs

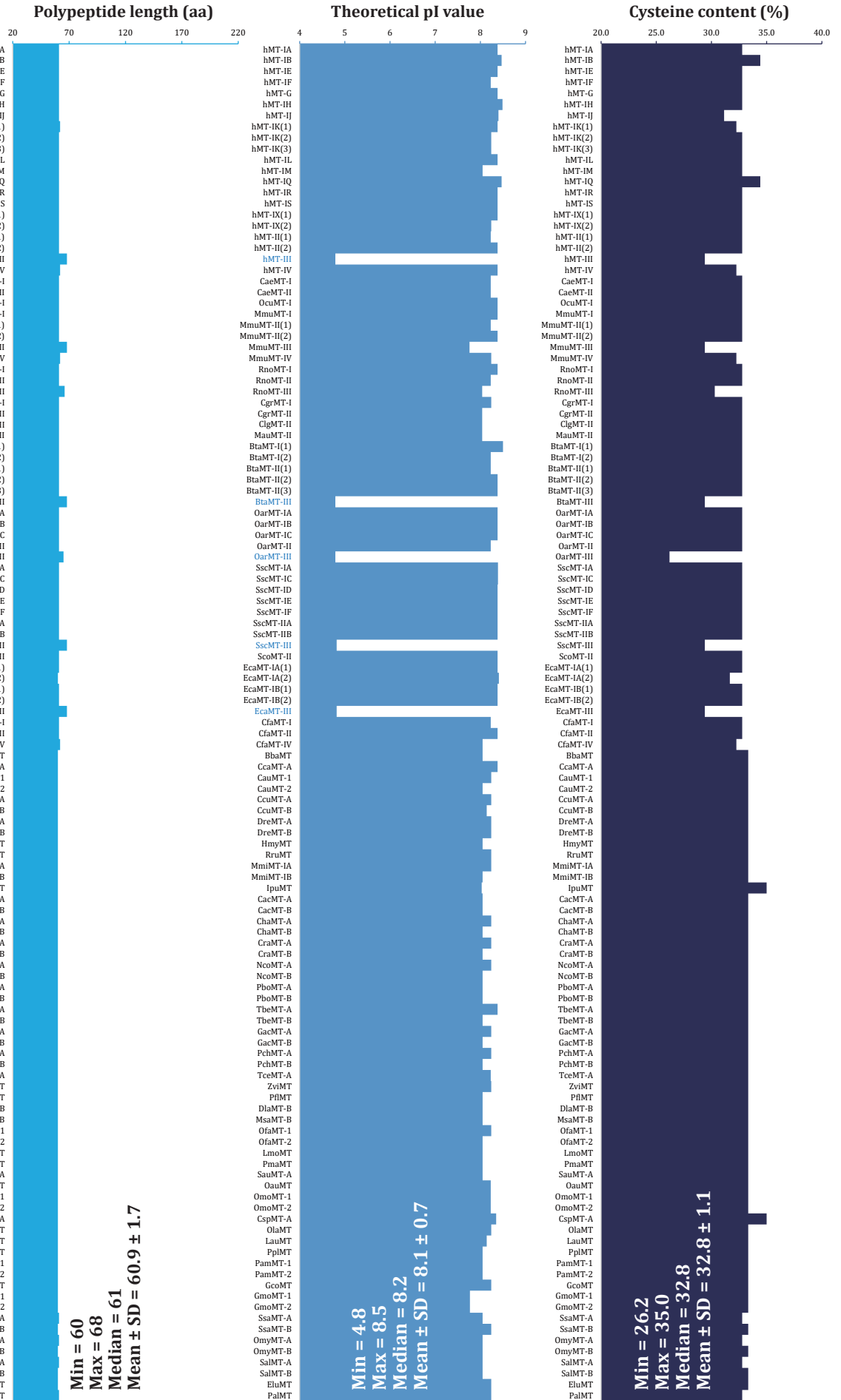
Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Barbatula barbatula</i>	<b>Teleostei</b> ; Otomorpha; Cypriniformes; <b>Nemacheilidae</b>	MT	P25128	BbaMT	
<i>Cyprinus carpio</i>	<b>Teleostei</b> ; Otomorpha; Cypriniformes; <b>Cyprinidae</b>	MT-A	O13269	CcaMT-A	
		MT-B	Q9I9I0	CcaMT-B	
<i>Carassius auratus</i>		MT-1	P52723	CauMT-1	
		MT-2	JC2419	CauMT-2	
<i>Carassius cuvieri</i>		MT-A	AAN85819	CcuMT-A	
		MT-B	AAN85820	CcuMT-B	
<i>Danio rerio</i>		MT-A	NP_571150	DreMT-A	
		MT-B	NP_919249	DreMT-B	
<i>Hemibarbus mylodon</i>		MT	EF689140	HmyMT	
<i>Rutilus rutilus</i>		MT	P80593	RruMT	
<i>Misgurnus mizolepis</i>	<b>Teleostei</b> ; Otomorpha; Cypriniformes; <b>Cobitidae</b>	MT-IA	FJ015060	MmiMT-IA	Tandemly (tail-to-head) duplicated
		MT-IB	FJ015059	MmiMT-IB	
<i>Ictalurus punctatus</i>	<b>Teleostei</b> ; Otomorpha; Siluriformes; <b>Ictaluridae</b>	MT	O93571	IpuMT	
<i>Chanocephalus aceratus</i>	<b>Teleostei</b> ; Euteleosteomorpha; Perciformes; <b>Channichthyidae</b>	MT-A	O93593	CacMT-A	
		MT-B	P52724	CacMT-B	
<i>Chionodraco hamatus</i>		MT-A	O13258	ChaMT-A	
		MT-B	P62711	ChaMT-B	
<i>Chionodraco rastrispinosus</i>		MT-A	CAA09714	CraMT-A	
		MT-B	P62679	CraMT-B	
<i>Notothenia coriiceps</i>	<b>Teleostei</b> ; Euteleosteomorpha; Perciformes; <b>Nototheniidae</b>	MT-A	P62339	NcoMT-A	
		MT-B	P62680	NcoMT-B	
<i>Pagothenia borchgrevinki</i>		MT-A	CAA07558	PboMT-A	
		MT-B	P62681	PboMT-B	
<i>Trematomus bernacchii</i>		MT-A	O93609	TbeMT-A	
		MT-B	P62678	TbeMT-B	
<i>Gymnodraco acuticeps</i>	<b>Teleostei</b> ; Euteleosteomorpha; Perciformes; <b>Bathydraconidae</b>	MT-A	CAA07555	GacMT-A	
		MT-B	P62713	GacMT-B	
<i>Parachaenichthys charcoti</i>		MT-A	O93450	PchMT-A	
		MT-B	P62682	PchMT-B	
<i>Thermarces cerberus</i>	<b>Teleostei</b> ; Euteleosteomorpha; Perciformes; <b>Zoarcidae</b>	MT-A	P52721	TceMT-A	
<i>Zoarces viviparus</i>		MT	P52728	ZviMT	
<i>Perca fluviatilis</i>	<b>Teleostei</b> ; Euteleosteomorpha; Perciformes; <b>Percidae</b>	MT	P52725	PflMT	
<i>Dicentrarchus labrax</i>	<b>Teleostei</b> ; Euteleosteomorpha; Eupercaria incertae sedis; <b>Moronidae</b>	MT-B	Q9PTG9	DlaMT-B	
<i>Morone saxatilis</i>		MT-B	P62712	MsaMT-B	

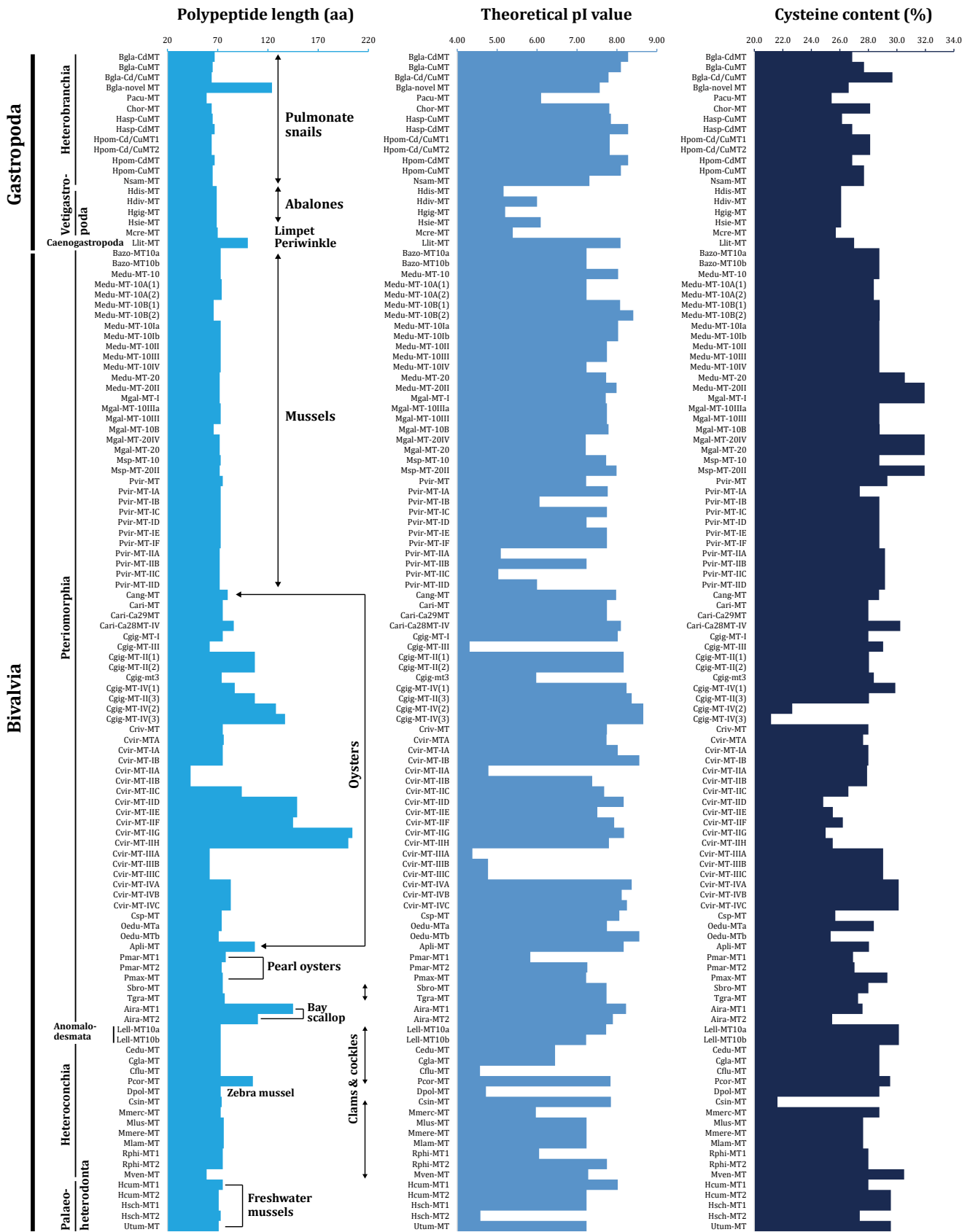
Supplementary Table S1C Continued

Species	Taxonomic position	MT proteins	Accession code	Abbreviated label used	Remark
<i>Oplegnathus fasciatus</i>	Teleostei; Euteleostomorpha; Centrarchiformes; <b>Oplegnathidae</b>	MT1	JF419527	OfaMT-1	
		MT2	JF419528	OfaMT-2	
<i>Lithognathus mormyrus</i>	Teleostei; Euteleostomorpha; Spariformes; <b>Sparidae</b>	MT	AAL37187	LmoMT	
<i>Pagrus major</i>		MT	Q9IB50	PmaMT	
<i>Sparus aurata</i>		MT-A	P52727	SauMT-A	
<i>Oreochromis aureus</i>	Teleostei; Euteleostomorpha; Cichliformes; <b>Cichlidae</b>	MT	AAP14677	OauMT	
<i>Oreochromis mossambicus</i>		MT-1	AAP14678	OmoMT-1	
		MT-2	P52726	OmoMT-2	
<i>Cyprinodon sp.</i>	Teleostei; Euteleostomorpha; Cyprinodontiformes; <b>Cyprinodontidae</b>	MT-A	Q92044	CspMT-A	
<i>Oryzias latipes</i>	Teleostei; Euteleostomorpha; Beloniformes; <b>Adrianichthyidae</b>	MT	AAR30249	OlaMT	
<i>Liza aurata</i>	Teleostei; Euteleostomorpha; Mugiliformes; <b>Mugilidae</b>	MT	O13257	LauMT	
<i>Pleuronectes platessa</i>	Teleostei; Euteleostomorpha; Pleuronectiformes; <b>Pleuronectidae</b>	MT	S30567	PplMT	
<i>Pseudopleuronectes americanus</i>		MT-1	P55945	PamMT-1	
		MT-2	CAA31930	PamMT-2	
<i>Gobiomorphus cotidianus</i>	Teleostei; Euteleostomorpha; Gobiiformes; <b>Eleotridae</b>	MT	AAO89258	GcoMT	
<i>Gadus morhua</i>	Teleostei; Euteleostomorpha; Gadiformes; <b>Gadidae</b>	MT-1	P51902	GmoMT-1	
		MT-2	CAA65924	GmoMT-2	
<i>Salmo salar</i>	Teleostei; Euteleostomorpha; Salmoniformes; <b>Salmonidae</b>	MT-A	CAA65929	SsaMT-A	
<i>Oncorhynchus mykiss</i>		MT-B	CAA65930	SsaMT-B	
		MT-A	P09861	OmyMT-A	
<i>Salvelinus alpinus</i>		MT-B	P09862	OmyMT-B	
		MT-A	AAB66342	SalMT-A	
		MT-B	AAB66343	SalMT-B	
<i>Esox lucius</i>	Teleostei; Euteleostomorpha; Esociformes; <b>Esocidae</b>	MT	P25127	EluMT	
<i>Plecoglossus altivelis</i>	Teleostei; Euteleostomorpha; Osmeriformes; <b>Plecoglossidae</b>	MT	AAP43669	PalMT	

Mammalia

Teleostei





# Gastropoda MT family

## Major $\beta_2\beta_1$ -form of gastropod MTs (18 cysteines)

### Vetigastropoda Heterobranchia

	$\beta_2$ -domain					$\beta_1$ -domain									
	CXXXC	→	CXC	→	CXC	→	CXC	→	C	→	CXC	→	CX	C	
<i>Pacu</i> -MT	MSGKG--PN	C	TEA	C	TGEQ	C	T	CGRQLQV	----	C	----	CK	-----	TCKCEDNACKCGEGCTGPSTCKCESSDCA	CK 59-aa
<i>Bgla</i> -CdMT	MSGKGKGEK	C	TSA	C	RSEP	C	Q	CGSKQCG-EG	CTCAACK	-----	T	C	NCTSDGCKCGKECTGPDSCKCGS-SCS	CK 67-aa	
<i>Hpom</i> -CdMT	MSGKGKGEK	C	TSA	C	RSEP	C	Q	CGSKQCG-EG	CTCAACK	-----	T	C	NCTSDGCKCGKECTGPDSCKCGS-SCS	CK 67-aa	
<i>Hasp</i> -CdMT	MSGKGKGEK	C	TAA	C	RNEP	C	Q	CGSKQCG-EG	CTCAACK	-----	T	C	NCTSDGCKCGKACTGPDSCTCGS-SCG	CK 67-aa	
<i>Nsam</i> -MT	MSGKG--EL	C	TSA	C	KNP	C	Q	CGDKCQCG-EG	CTSAACK	-----	S	H	CTNDGCKCGKECTGPTSCKCDT-SCS	CK 65-aa	
<i>Hpom</i> -Cd/CuMT1	MSGKG--SN	C	AGS	C	NSNP	C	S	CGDDCKCG-AG	CSVQCH	-----	S	Q	CNNDTCKCGNQCSASGSCKCG--SCG	CK 64-aa	
<i>Hpom</i> -Cd/CuMT2	MSGKG--SN	C	AGS	C	NSNP	C	S	CGDDCKCG-AG	CSAQCH	-----	S	Q	CNNDTCKCGNQCSASGSCKCG--SCG	CK 64-aa	
<i>Bgla</i> -Cd/CuMT	MSGKG--SV	C	ACS	C	NSNP	C	S	CGDDCKCG-AG	CSAQCY	-----	S	Q	CNNDTCKCGSQCTSTSGSCKCG--SCG	CK 64-aa	
<i>Chor</i> -MT	MSGKG--SA	C	AGS	C	NSNP	C	S	CGDDCKCG-AG	CSAQCY	-----	S	Q	CNNDTCKCGSQCTSTSGSCKCG--SCG	CK 64-aa	
<i>Bgla</i> -CuMT	MSGRGK--N	C	CGA	C	NSNP	C	S	CGNDCKCG-AG	CNDRCS	-----	S	H	CNDDCKCGSQCTGSGSCKCGS-ACG	CK 65-aa	
<i>Hpom</i> -CuMT	MSGRGK--N	C	CGA	C	NSNP	C	S	CGNDCKCG-AG	CNDRCS	-----	S	H	CNDDCKCGSQCTGSGSCKCGS-ACG	CK 65-aa	
<i>Hasp</i> -CuMT	MSGRGQ--N	C	CGA	C	NSNP	C	N	CGNDNCG-TG	CNDQCS	-----	A	R	HCSNDDCKCGSQCTRSQSCKCGN-ACG	CK 65-aa	
<i>Hdis</i> -MT	MSSP-QGAG	C	TGE	C	KTD	P	C	ACGTDCKCNPDD	CACDTC	---	V	K	TKCP-GSCECGKCTSGETCKCDD-SCT	CK 69-aa	
<i>Hgig</i> -MT	MSSP-QGAG	C	TGE	C	KTD	P	C	ACGTDCKCNPED	CACDTC	---	V	K	TKCP-GSCECGKCTSGETCKCDD-SCT	CK 69-aa	
<i>Hsie</i> -MT	MSSP-QGAG	C	TGE	C	KTD	P	C	ACGTDCKCNPDD	CACDTC	---	V	K	TKCP-GSCECGKCTSGETCKCDD-SCK	CK 69-aa	
<i>Hdiv</i> -MT	MSSP-QGPG	C	TAS	C	KSEP	C	ACGTDCKCNPSD	CPCTTC	---	D	K	TKCS-DG	CQCGKCTTGDI	CKCDD-SCS	CK 69-aa
<i>Mcre</i> -MT	MSGKG--EN	C	TAE	C	KSDP	C	ACGDS	CKCG-EG	CAC	TT	C	V	KTEAKTTCKCG-ESCK-EGCKEAGEACKCES-GCAS	CK 70-aa	

### Caenogastropoda

*Llit*-MT

1 <sup>st</sup> $\beta_2$ domain	MSSVFGAG	C	TDV	C	KQTP	CG	ATSG	C	N	CTDD	CK	Q	S	CKY																									
2 <sup>nd</sup> $\beta_2$ domain	GAG	C	TDT	C	KQTP	CG	-GSG	C	N	CKED	DR	Q	S	CK	100-aa																								
[NH <sub>3</sub> <sup>+</sup> $\beta_2\beta_2\beta_1$ COO <sup>-</sup> ]																																							
										C-terminal $B_1$ domain																													
										A	C	K	C	A	A	G	S	C	K	C	G	K	G	C	T	G	P	D	S	C	K	C	D	R	S	C	S	C	K

### Heterobranchia

*Bgla*-novel-MT

1 <sup>st</sup> $\beta_2$ domain	MSGKGPN	C	TEA	C	TGEQ	C	N	CGDS	CK	C	G	E	G	C	N	C	P	S	C	K	T	T											
2 <sup>nd</sup> $\beta_2$ domain	KGPN	C	TEA	C	TGK	Q	C	S	CGDS	Q	C	G	E	G	C	T	C	S	C	K	K	A											
3 <sup>rd</sup> $\beta_2$ domain		C	T	K	E	C	T	D	T	E	C	S	C	G	D	S	CK	C	G	E	G	CK	C	S	S	CK	A	G					
[NH <sub>3</sub> <sup>+</sup> $\beta_2\beta_2\beta_2\beta_P$ COO <sup>-</sup> ]															C-terminal partial domain																		
										K	C	T	K	S	D	E	G	C	K	T	E	G	H	C	A	K	G	C	C	K	S		124-aa

### Ostreidae MT-I family

[NH<sub>3</sub><sup>+</sup> αβ<sub>1</sub> COO<sup>-</sup>]

		α-domain (12 cysteines)						Linker	β <sub>1</sub> -domain (9 cysteines)																																
		CXC	→	CXC	→	C	→	CXC	→	CXC	→	CXC	→	C	→	CXC	→	CXC																							
<i>C. virginica</i> MT-Is	<i>Cvir</i> -MT-IA	--MSDP	CNC	IETGT	CACS	SDS	CPATG	-CK	CGPG	CK	CGDD	CK	CAG	CK	KVK	C	SCT	SEGG	CK	CG	-EK	CT	GPAT	CK	CG	SG	--	C	S	CKK	75-aa										
	<i>Cvir</i> -MT-IB	--MANP	CNC	NETGT	CVCS	NS	CPATG	-CK	CGPG	CK	CGDD	CK	CAG	CK	KVK	C	SCT	SEGG	CK	CG	-KK	CT	GPAT	CK	CG	SG	--	C	S	CKK	75-aa										
	<i>Cvir</i> -MTA	--MSDP	CNC	TETGT	CVCS	NS	CPATG	-CK	CGPG	CK	CGDE	CK	CAG	CK	KVK	C	NS	SEGG	CK	CG	DEK	CT	GPAT	CK	CG	ES	--	C	S	CKK	76-aa										
Other <i>Crassostrea</i> MT-Is	<i>Cgig</i> -MT-I	--MSDP	CNC	TESGT	CVCS	SDS	CPATG	-CK	CGPG	CK	CGDG	CK	CSG	CK	VVK	C	NS	SGT	--	CG	CG	-KG	CT	GPEN	CK	CA	ND	SG	CG	CKK	75-aa										
	<i>Cari</i> -MT	--MSDP	CNC	IEAGT	CVCS	SDS	CPPTG	-CK	CGPG	CK	CGDG	C	NSG	CK	VVK	C	NS	SGS	--	CG	CG	-KG	CT	GPEN	CK	CA	ND	SG	CG	CKK	75-aa										
	<i>Cari</i> -Ca29MT	--MFD	CNC	IETGT	CVCS	SDS	CPATG	-CK	CGPG	CK	CGDG	C	NSG	CK	VVK	C	NS	SGS	--	CG	CG	-KG	CT	GPEN	CK	CA	ND	SG	CG	CKK	75-aa										
	<i>Criv</i> -MT	--MSDP	CNC	IEAGT	CVCS	SDS	CPPTG	-CK	CGPG	CK	CGDG	C	NSG	CK	VVK	C	NS	SGS	--	CG	CG	-KG	CT	GPEN	CK	CA	ND	SG	CG	CKK	75-aa										
	<i>Cgig</i> -mt3	--MSDP	CNC	AETGS	CQ	SAQ	CTSEGG	CK	CG	PN	CK	CGDG	C	APG	CK	VV	C	NS	NGS	--	C	AG	-QD	CT	GQST	CK	CG	SD	--	C	S	CKK	74-aa								
	<i>Csp</i> -MT	MYSLAP	CNC	SETGK	C	NS	SDS	CPATG	-CK	CG	PR	CK	CGDG	CK	CSG	CK	VK	S	GS	SDS	--	C	V	D	-KD	CP	GPET	CK	K	ASE	--	C	P	CK-	74-aa						
	<i>O. edulis</i> MT-Is	<i>Oedu</i> -MTa	--MSDP	CNC	KTTGN	CL	SDT	CPSTG	-C	N	ASG	CK	CGAG	CK	CPG	Q	I	V	CK	NS	T	--	C	AG	-VG	CT	GPDT	CK	CA	AD	SG	CS	CK-	74-aa							
<i>Oedu</i> -MTb	--MSDP	CNC	QKR	-N	C	Q	SGI	CPSTG	-C	N	ASG	CK	CVAG	CK	CPG	CK	I	V	N	T	V	AS	F	--	C	AYG	-IE	C	I	R	P	V	T	CK	CA	AG	--	C	T	CK-	75-aa

### Ostreidae MT-II family

*C. virginica* MT-IIH  
[NH<sub>3</sub><sup>+</sup> α<sub>(n)</sub> COO<sup>-</sup>]

α<sub>1</sub> domain  
α<sub>2</sub> domain  
α<sub>3</sub> domain  
α<sub>4</sub> domain

		α-domain																	
<i>C. virginica</i> MT-IIH	α <sub>1</sub> domain	NH <sub>3</sub> <sup>+</sup>	MSDP	CNC	IETGT	CACS	SDS	CPATG	CK	CAPG	CK	CGDD	CK	CAG	CK	KVK	C	SGT	SENANIK
	α <sub>2</sub> domain		MADP	CNC	TETGT	CVCS	NS	CPATG	CK	CGPG	CK	CGDD	CK	CAG	CK	KVK	C	SGT	----NIK
	α <sub>3</sub> domain		MADP	CNC	TETGT	CVCS	NS	CPATG	CK	CGPG	CK	CGDD	CK	CAG	CK	KVK	C	SGT	----NIK
	α <sub>4</sub> domain		MADP	CNC	TETGT	CVCS	NS	CPATG	CK	CGPG	CK	CGDE	CK	CAG	CK	KV	COO <sup>-</sup>	200-aa	

Other Ostreidae MT-IIs  
[NH<sub>3</sub><sup>+</sup> αβ<sub>1</sub>β<sub>1</sub> COO<sup>-</sup>]

		α-domain												Linker	β <sub>1</sub> -domain														
Other Ostreidae MT-IIs	<i>Cang</i> -MT	MSDP	CNC	IETGT	CVCS	SES	CPATG	CK	CGLG	CK	CGDG	CK	CSG	CK	KVK	C	NS	SGS	CG	CG	KG	CT	GPEN	CK	CA	ND	SG	CVC	80-aa
	<i>Cgig</i> -MT-II(1)	MSDP	CNC	TATGT	CVCS	SDS	CPATG	CK	CGPG	CK	CGDG	CK	CSG	CK	KVK	C	NS	SGS	CG	CG	KG	CT	GPEN	CK	CA	ND	SE	CG	107-aa
	<i>Cgig</i> -MT-II(3)	MSDP	CNC	TKTGT	CVCS	SDS	CPATG	CK	CGPG	CK	CGDG	CK	CSG	CK	KVK	C	NS	SGS	CG	CG	KG	CT	GPEN	CK	CA	ND	SG	CG	107-aa
	<i>Cgig</i> -MT-II(2)	MSDP	CNC	TETGT	CVCS	SDS	CPATG	CK	CGPG	CK	CGDG	CK	CSG	CK	KVK	C	NS	SGS	CG	CG	KG	CT	GPEN	CK	CA	ND	SG	CG	107-aa
	<i>Apli</i> -MT	MSDP	CNC	SETGT	CVCS	SDS	CPATG	CK	CGPG	CK	CGDG	CK	CSG	CK	KVK	C	NS	SGS	CG	CG	KG	CT	GPEN	CK	CA	ND	SG	CVC	107-aa



**Ostreidae MT-III family**

[NH<sub>3</sub><sup>+</sup> β<sub>2</sub>β<sub>2</sub> COO<sup>-</sup>]

		β <sub>2</sub> -domain (9 cysteines)	Linker	β <sub>2</sub> -domain (9 cysteines)	
<i>C. virginica</i> MT-IIIs <i>C. gigas</i> MT-III	{ <i>Cvir</i> -MT-IIIA <i>Cvir</i> -MT-IIIB <i>Cvir</i> -MT-IIIC <i>Cgig</i> -MT-III	MPFETSCTCANGACECGENQCKTTDCAC	TTCNV	TCSETESECKCGADCNCSAECKCQTCKS	62-aa
		MPFETSCTCANGACECGENQCKTTDCAC	TTCNV	TCSESKSECKCGADCNCSAECKCQTCKS	62-aa
		MPFETNCTCANGACECGENQCKTTDCAC	TTCNV	TCSESKSECKCGADCNCSAECKCQTCKS	62-aa
		MPIETNCTCANGACNCGETCQCKTTDCAC	AI	CNNP	CGCTESECNCGAECQCPETCSCKTCKA

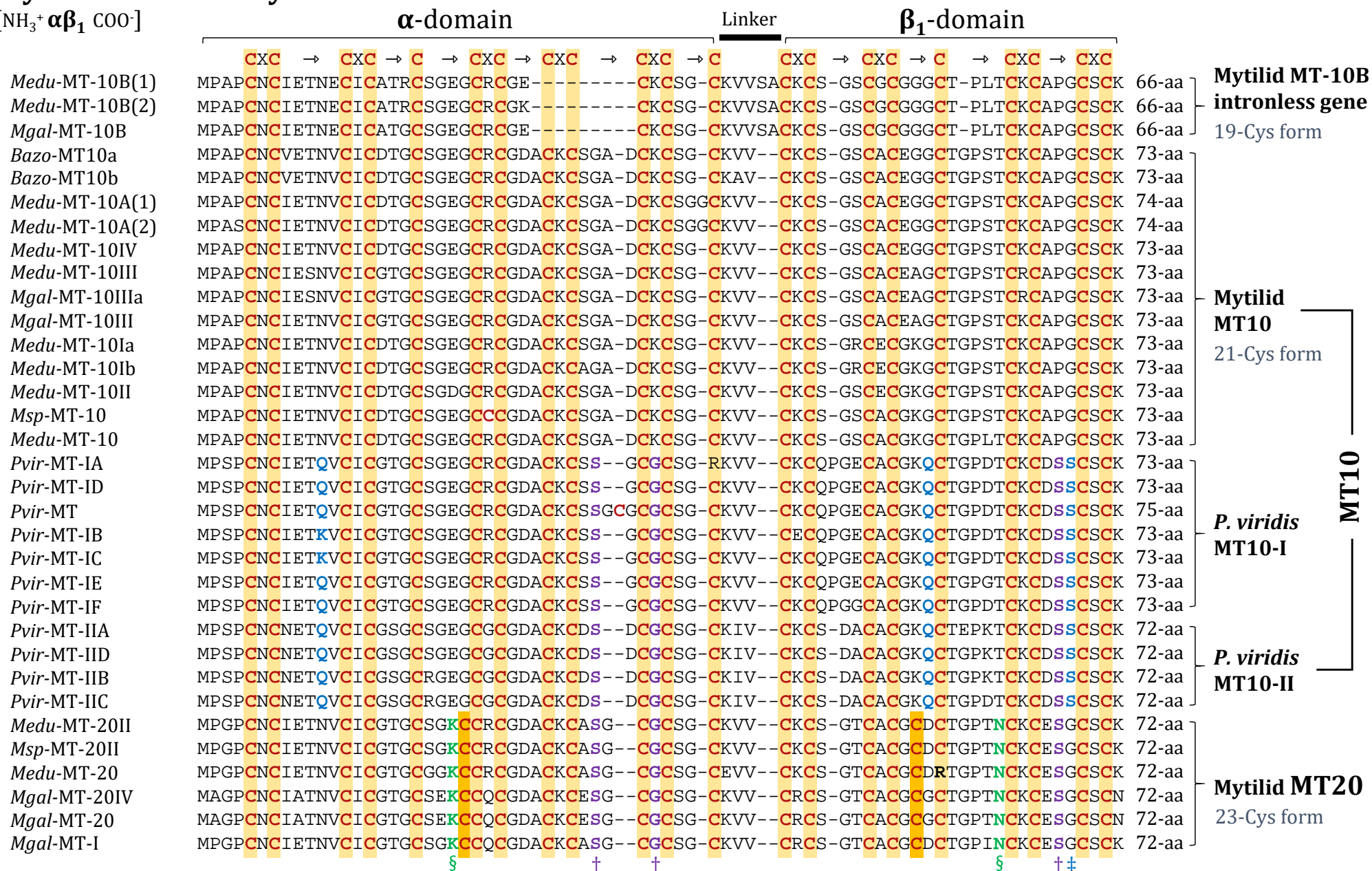
**Ostreidae MT-IV family**

[NH<sub>3</sub><sup>+</sup> α'β<sub>1</sub>' COO<sup>-</sup>]

		α'-domain	Linker	β <sub>1</sub> '-domain						
<i>C. virginica</i> MT-IVs	{ <i>Cvir</i> -MT-IVA <i>Cvir</i> -MT-IVB <i>Cvir</i> -MT-IVC	MSDTCACATTGKCVCCDTCGPDGASC	SGEAC	SCAKKTCNCKGCKVK--CCC	TKDKCCCGKGC	CHGPGTCKCDENCCC	KKQDKKQAT	83-aa		
		MSDTCACATTGKCVCCDTCGPDGASC	SGEAC	SCAEKTCNCKGCKVK--CCC	TKDKCCCGKGC	CHGPGTCKCDENCCC	KKQDKKQAT	83-aa		
		MSDTCACATTGKCVCCDTCGPDGASC	SGEAC	SCAKKTCNCKGCKVK--CCC	TKDKCCCGKGC	CDGPGTCKCDENCCC	KKQDKKQAT	83-aa		
<i>C. gigas</i> MT-IV <i>C. ariakensis</i> MT-IV	{ <i>Cgig</i> -MT-IV(1) <i>Cari</i> -Ca28MT-IV	MSDHCTCAQTGKCVCCDTCPPDGNCS	CGDKCK	CAKASCNCT-CKGKAKCCC	TKEKCCCGKGC	CHGPE	TKCPADCCC	KKEHDA	CSKAGH	87-aa
		MSDHCTCAETGKCVCCDTCPPDGNCS	CGDKCK	CAKASCNCKGCKVK--CCCT	KEKCCCGKGC	CHGPE	TKCPADCCC	KKEHDA	CSKAGH	86-aa

# Mytilidae MT family

[NH<sub>3</sub><sup>+</sup> αβ<sub>1</sub> COO<sup>-</sup>]



§: Amino acid residue specific to MT20

†: Amino acid residue shared by *Pvir*-MTs and MT20s

‡: Amino acid residue specific to *Pvir*-MTs

## Bivalve MTs excluding Ostreidae and Mytilidae

[NH<sub>3</sub><sup>+</sup> αβ<sub>1</sub> COO<sup>-</sup>]

	α-domain												Linker	β <sub>1</sub> -domain																							
<i>Pmax-MT</i>	MSDP	CR	CAKTKEP	-	CACTDA	--	CPASG	-	CCCGST	-	QCGDN	-	CKCPN	-	CKVK	----	CACADKGI	-	CGCGVG	--	CTTSLSD	CN	CTD	--	C	SCKN	75-aa	22 Cys; one extra Cys									
<i>Lell-MT10a</i>	MPSP	CN	CRETGK	--	CTCDGK	--	CSGDA	-	CCCGAN	-	CNCGEG	-	CKCPG	-	CKTVV	----	CKCSGD	--	CACGKG	--	CTGPDS	CK	CDAG	--	C	SCK	73-aa										
<i>Lell-MT10b</i>	MPSP	CN	CRETGK	--	CTCDGK	--	CSGDA	-	CCCGAN	-	CNCGED	-	CKCQG	-	CKTVV	----	CKCSGD	--	CACGKG	--	CTGPDS	CK	CDAG	--	C	SCK	73-aa										
<i>Hcum-MT1</i>	MSDP	CN	CIETGT	--	CVCS	-	SPATG	-	CKCGPG	-	CKCGDG	-	CKCSG	-	CKVK	----	CNCSGS	--	CGCGKG	--	CTGPEN	CK	CANDSG	-	C	GCKK	75-aa	21 Cys; all in canonical positions									
<i>Pmar-MT1</i>	MPSP	CN	CATTSD	--	CVCFN	-	NCPTSG	-	CNCAMS	-	CKCSDS	-	CPCPG	-	CKSAVNSTVN	CT	CPGS	--	CACGVG	--	CTGPSE	C	SCTS	--	C	SCKT	78-aa										
<i>Cflu-MT</i>	MSDP	CD	CATTGT	--	CKCED	--	CTAESG	CR	CGTG	-	CNCGDD	-	CKCPG	-	CKVV	----	CKCSGS	--	CDCGQG	--	CTGPST	CK	CESD	--	C	SCK	73-aa										
<i>Mlus-MT</i>	MSDP	CN	CIETGT	--	CKCED	--	CSSTTG	CR	CGAG	-	CKCGDA	-	CKCPG	-	CKVV	----	CKCAENNGE	CK	CGKG	--	CTGPDS	CK	CDQG	--	C	SCK	76-aa	21 Cys; all in canonical positions									
<i>Mmere-MT</i>	MSDP	CN	CIETGT	--	CKCED	--	CSSTTG	CR	CGAG	-	CKCGDA	-	CKCPG	-	CKVV	----	CKCAENNGE	CK	CGKG	--	CTGPDS	CK	CDQG	--	C	SCK	76-aa										
<i>Mlam-MT</i>	MSDP	CN	CIETGT	--	CKCED	--	CSSTTD	CR	CGAG	-	CKCGDA	-	CKCPG	-	CKVV	----	CKCADNNGE	CK	CGKG	--	CTGPSS	CK	CDQS	--	C	SCK	76-aa										
<i>Rphi-MT1</i>	MGDP	CN	CAETGT	--	CKCSDQ	--	CTSADG	CR	CGPN	-	CNCGSD	-	CSCPG	-	CKVAT	----	CKCSGS	--	CACGKG	--	CTREGT	CE	CGND	--	C	SCKK	75-aa	21 Cys; all in canonical positions									
<i>Rphi-MT2</i>	MGDP	CN	CAESGT	--	CKCSDQ	--	CTSADG	CR	CGPN	-	CKCGSD	-	CRCPG	-	CKVAT	----	CKCSGS	--	CACGKG	--	CTREGT	CE	CGND	--	C	SCKK	75-aa										
<i>Mmerc-MT</i>	MGDP	CN	CAETGS	--	CNCS	DQ	--	CTSDGG	CR	CGSN	-	CKCGAD	-	CKCPG	-	CKVV	----	CKCGDS	--	CACGKG	--	CTGPST	CQ	CDSG	--	C	SCK		73-aa								
<i>Sbro-MT</i>	MSDP	CK	CIEGGE	--	CKCED	--	CASDN	-	CRCDPAK	CK	CKAG	-	CKCSG	-	CGVK	----	CKCSGT	--	CDCGQN	--	CTGPAD	CK	CKPNS	-	CP	CNQ	75-aa	21 Cys; some non-canonical cysteine residues and motifs									
<i>Utum-MT</i>	MSDP	CN	CLETGE	--	CKCSGD	--	CTGD	--	CRCGDA	-	CKCGNA	-	CKCPG	-	CKVV	----	CKCSSS	--	CDCGKG	--	CTGPST	CK	CDSG	--	C	SCK	71-aa										
<i>Tgra-MT</i>	MSDP	CK	CVEGGE	--	CKCSDS	--	CAGDN	-	CRCDPEK	CK	CKDG	-	CACSN	-	CKVK	----	CKCSGT	--	CDPACGKN	CT	GP	CK	DK	CK	PPNS	-	CP		CNQ	77-aa							
<i>Hcum-MT2</i>	MSDA	CN	CLETGE	--	CKCCGE	--	STGD	-	CRCGKD	-	CKCGDA	-	CKCPG	-	CKVV	----	CKCSSS	--	CDCGKG	--	CTGPST	CK	CDSG	--	C	SCK	71-aa	21 Cys; some non-canonical cysteine residues and motifs									
<i>Hsch-MT1</i>	MSDA	CN	CLETGE	--	CKCCGE	--	STGD	-	CRCGKD	-	CKCGDA	-	CKCPG	-	CKVV	----	CKCSSS	--	CDCGKG	--	CTGPST	CK	CDSG	--	C	SCK	71-aa										
<i>Cedu-MT</i>	MGDP	CN	CAQTGGD	--	CKCAAGNC	SS	DTP	CR	CGSG	-	CGCGSE	-	CTC	--	HVK	----	CTCSGS	--	CACGNN	--	CTGPAN	CT	CGAG	--	C	SCK	73-aa										
<i>Cgla-MT</i>	MGDP	CN	CAQTGGD	--	CKCAAGNC	SS	GETP	CR	CGSG	-	CGCGSD	-	CTC	--	HVK	----	CS	SGS	--	CACGNN	--	CTGPAN	CT	CGAG	--	C	SCK	73-aa	20 Cys								
<i>Dpol-MT</i>	MSDP	CN	CVETG	-	D	-	CRCADGS	-	CSDCSN	CK	CGDS	-	CKCSKPN	CC	-	G	--	KNVT	----	CKCGEN	--	CQCGVG	--	CTGPDS	CT	CDSG	--	C		SCK	73-aa						
<i>Hsch-MT2</i>	MSES	CQ	CLEVADRE	CT	DV	--	CKTQ	--	CNCS	DT	-	CN	SED	-	CE	TGR	CK	IQ	----	CN	SEEE	--	CI	GHG	-	CK	GPRT	CK		CDPS	--	CE	CH	73-aa			
<i>Mven-MT</i>	MGDP	CD	CVKDGQ	--	CTCQTS	--																														59-aa	18 Cys

## Bivalve MTs excluding Ostreidae and Mytilidae

[Domain duplicated structure]

### *A. irradians* MT1 (145-aa)

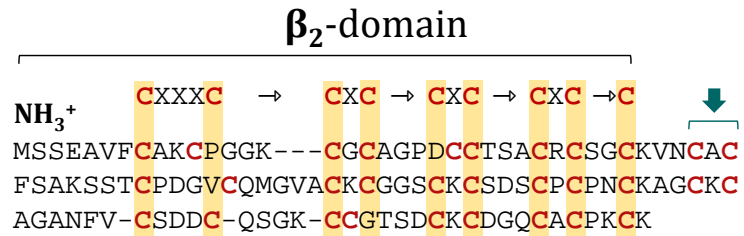
[ $\beta_2\beta_2\beta_2\beta_1$ -like]

N-terminal putative

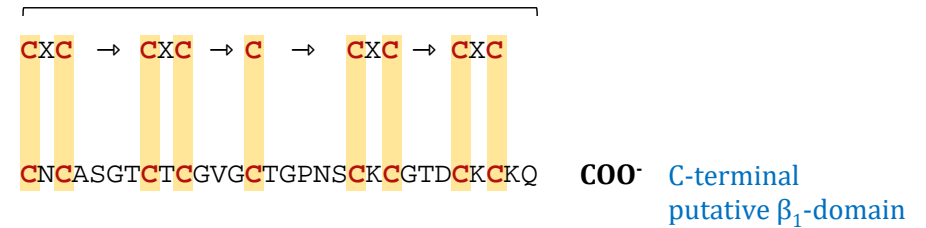
$\beta_2$ -domain (1)

$\beta_2$ -domain (2)

$\beta_2$ -domain (3)



### **$\beta_1$ -domain**



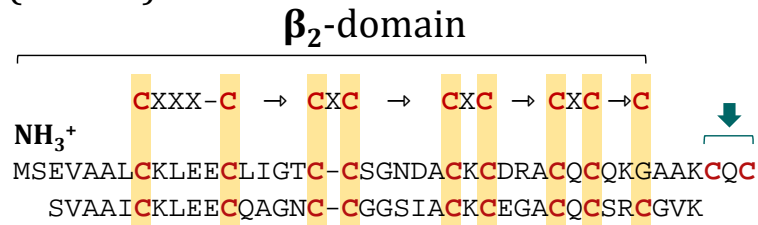
### *A. irradians* MT2 (110-aa)

[ $\beta_2\beta_2\beta_1$ -like]

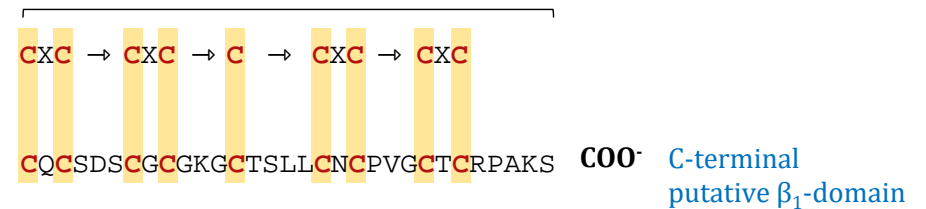
N-terminal putative

$\beta_2$ -domain (1)

$\beta_2$ -domain (2)



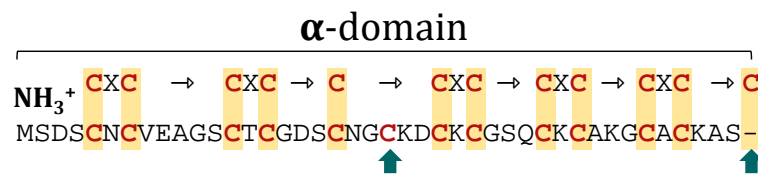
### **$\beta_1$ -domain**



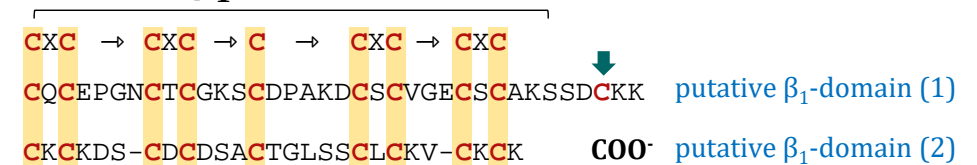
### *P. coreanum* MT (105-aa)

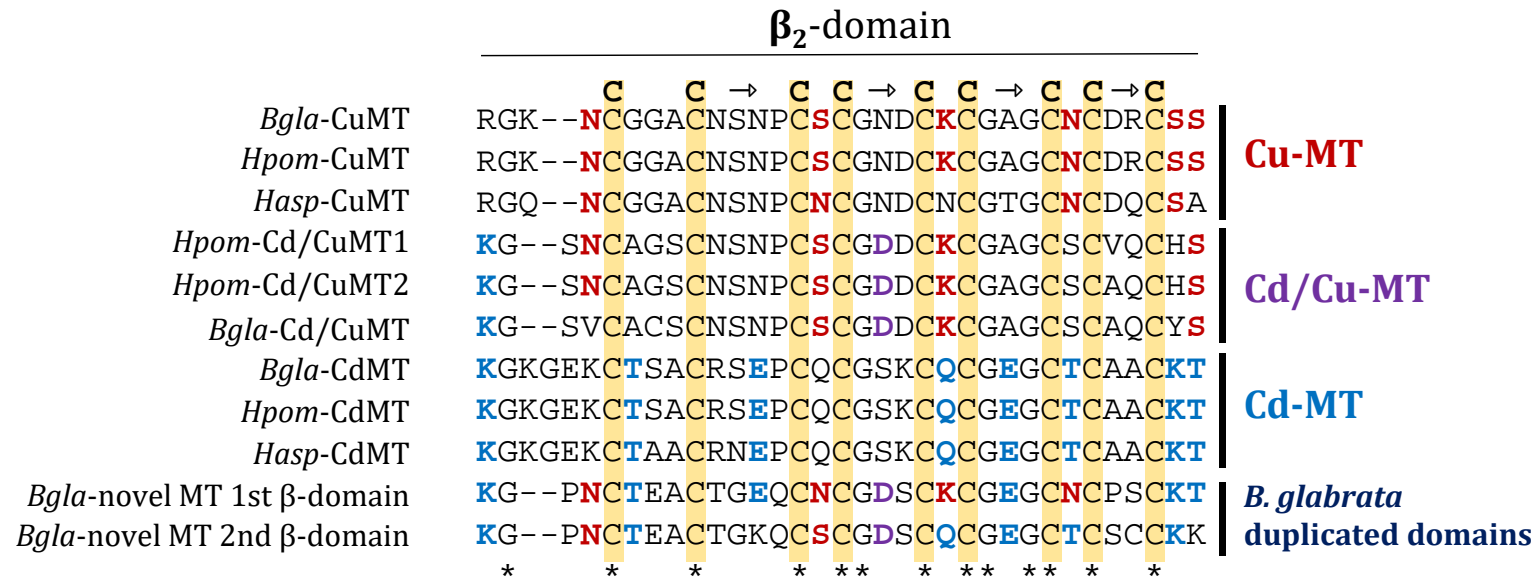
[ $\alpha\beta_1\beta_1$ -like]

N-terminal putative  $\alpha$ -domain

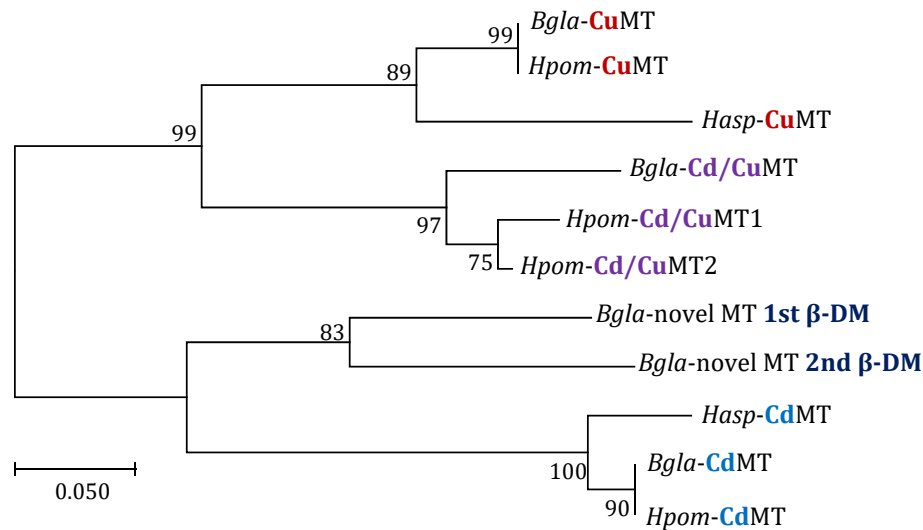


### **$\beta_1$ -domain**





Original tree



Bootstrap consensus tree

