

# ResearchOnline@JCU

This file is part of the following reference:

**Go, Hiroki (2005) *Characterization of doublesex and nanos genes in Acropora millepora*. PhD thesis, James Cook University.**

Access to this file is available from:

<http://researchonline.jcu.edu.au/1204/>

*The author has certified to JCU that they have made a reasonable effort to gain permission and acknowledge the owner of any third party copyright material included in this document. If you believe that this is not the case, please contact [ResearchOnline@jcu.edu.au](mailto:ResearchOnline@jcu.edu.au) and quote <http://researchonline.jcu.edu.au/1204/>*

Characterization of *doublesex* and *nanos* genes  
in *Acropora millepora*

Thesis submitted by Hiroki Go (Chikuei Wu) (Doctor of  
Medicine) in January 2005

the Degree of Doctor of Philosophy in the School of Pharmacy and  
Molecular Sciences at James Cook University

# Statement of Access

I, the undersigned author of this thesis, understand that James Cook University will make it available for use within the University library and by microfilm or other photographic means, allow access to users in other approved libraries. all users consulting this thesis will have to sign the following statement:

*“In consulting this thesis, I agree not to copy or partially paraphrase it in whole or in part without the written consent of the author, and to make proper written acknowledgement of any assistance which I have obtained from it.”*

Beyond this, I do not wish to place any restriction on access to this thesis.

Hiroki Go (Chikuei Wu)

Date

-----

-----

# Abstract

Scleractinian coral, *Acropora millepora* is a hermaphroditic species which develops both male and female gonads in each polyp at the same time. Corals are diploblastic metazoa and have a simple body structure without well-defined organs. They lack typical reproductive system. Instead, eggs and sperms form independent clusters in the mesoglea of mesenteries.

Little has been known for the molecular mechanisms of sex-determination in coral. Recently, a homolog of *doublesex* (*AmDM1*) was cloned for *Acropora millepora*. *doublesex* was identified as the somatic sex-determinant of *Drosophila melanogaster* and has been shown to be well conserved throughout the metazoan evolution. The gene product, Dsx, is a transcription factor containing the DM domain, a highly conserved DNA-binding domain. Thus it was speculated that *AmDM1* took part in the sex-determination cascade by controlling the transcription of genes which regulated the female and/or male gonad formation.

In order to study the involvement of *AmDM1* in the gonad development, I tried to isolate the candidates for the downstream genes of *AmDM1* by applying the yeast one-hybrid screening system. This selection system enabled me to isolate the genomic DNA fragments including the potential binding elements for the DM domain. 5 clones were finally isolated by this screening. 4 of them contain the repeat sequences which are specific to genus *Acropora*. The interaction of these sequences and the DM domain of *AmDM1*

was demonstrated by the electrophoretic mobility shift assay (EMSA).

A genomic clone was isolated by using the fifth clone for the hybridization probe and identified as a coral homolog of T-type amino acid transporters 1 (TAT1) which was supposed to mediate the membrane transport of aromatic amino acids. Therefore this gene was named *AmTAT1*. The nucleotide sequence of the genomic clone indicated that the DNA fragment isolated by one-hybrid screening was located in the second intron of *AmTAT1*. According to the deduced amino acid sequence, AmTAT1 was predicted to have common features of membrane transporters with 12 transmembrane domains.

The responsive sequence to DM domain in the second intron of *AmTAT1* was identified by EMSA. Comparison of this sequence with that in the repeat sequence revealed that the recognition motif was AATGT(GT)/(TG)C. The mutant study showed the AATGT pentanucleotide stretch was essential to the interaction with the DM domain. Moreover alignment of DM recognition elements of various species indicates that this stretch is absolutely conserved throughout evolution.

The regulatory function of AmDM1 on the DM recognition motif was investigated by CAT assay with Hep-2 cell line. In this system, AmDM1 acted on the recognition motif in an inhibitory manner. The transcriptional inhibition by AmDM1 was insensitive to the orientation of the responsive element. Therefore I concluded that the DM responsive element in the second intron of *AmTAT1* was the *cis*-acting on the gene expression of this gene. This observation is consistent with the result of expression analysis on *AmTAT1*. In adult

branch tips of *Acropora millepora*, the expression of *AmDM1* was reported to be detected in October, the season of sexual maturation. On the contrary, the expression of *AmTAT1* is down-regulated in October but is detectable before and after October.

Although the biological significance of the inhibition of *AmTAT1* expression in the process of sex-determination is yet to be known, it may lead to the sexual maturation of coral polyps by blocking the influx of specific amino acids or derivatives.

*vitellogenin* gene expression is controlled by Doublesex homologs in *Caenorhabditis elegans* and *Bombyx mori*. I isolated cDNA and genomic DNA of coral *vitellogenin* (*AmVit*). The genomic DNA was partially sequenced. However, no potential *AmDM1* binding site was found in the upstream regulatory region and introns.

*nanos* plays crucial roles in abdominal formation and proper germline differentiation in *Drosophila*. Lower animals including sponges and cnidarians also have *nanos* homologs. To know the roles of *nanos* genes in the germline speciation in coral, I isolated cDNAs and genomic DNAs of *nanos* homologs of *Acropora millepora* and *Nematostella vectensis*. Cnidarians have two *nanos* homologs, *cnnos1* and *cnnos2*. All Nanos proteins have the highly conserved Zinc finger domains. The phylogenetic analysis demonstrated that two *cnnos* genes segregated before the cnidarian radiation.

The sequence of 5'-flanking regions of these *cnnos* genes were subjected to the prediction of potential transcription factor binding sites. Their features turned out to be quite different. This finding is in

agreement with their expression pattern: the expression of *Cnnos1(Am)* was only observed in the early embryonal stage and no expression was detected in the adult coral tips and *Cnnos2(Am)* expression was detected in all stages of development (eggs, prawn-chip and pre-settlement) and all samples for the adult colony (September, October and December).

Although the difference in functions between *cnnos1(Am)* and *cnnos2(Am)* is suggested by these data, the roles of *cnnos* genes in the germline speciation is still unknown.

# Acknowledgement

This work was supervised by Dr. David J. Miller.



# Contents

## Chapter 1. General Introduction 1

## Chapter 2. Characterization of *AmDM1*, a *doublesex* Homolog of Coral *Acropora millepora* 5

### Section 1. Introduction 6

*doublesex* is near the bottom of the sex determination hierarchy  
in *Drosophila* 6

Identification of DNA-binding DM domains 8

Structure and properties of DM domain 9

Downstream genes of *dsx* 10

Evolutionary conservation of DM domain genes 15

*dsx* and DM domain genes in insects 15

*mab-3* and DM domain genes in nematode 16

DM domain genes in vertebrate 18

*AmDM1*, a *doublesex* homolog in coral 21

Strategy for the research on *AmDM1* and sex-determination  
21

### Section 2. Materials and Methods 23

*E. coli* strains 23

Manipulation of bacteria 23

Manipulation of  $\lambda$ -bacteriophage 23

Manipulation of nucleic acids and protein 23

Enzymes for molecular biology 24

Oligonucleotide primers 24

Animal sampling	24
Screening <i>Acropora millepora</i> cDNA library	24
Screening <i>Nematostella vectensis</i> cDNA library	28
Screening <i>Acropora millepora</i> genomic library	29
Sequencing lambda phage DNA	30
Southern-blot hybridization	30
RNA extraction from coral eggs and embryos	31
RNA extraction from adult coral tips	32
RT-PCR for <i>AmTat1</i> and <i>AmVit</i>	33
Yeast one-hybrid screen	35
Yeast strain and culture	36
Construction of Activator plasmid containing DM domain	37
Construction of Reporter library	38
Yeast transformation with plasmid vectors	39
Selection of yeast clones	40
Recovery of plasmid from yeast transformants	41
Prokaryotic expression vector construction	42
Large-scale expression and purification of the GST fusion protein	42
Preparation of radioactive probes for EMSA by PCR	44
Preparation of radioactive probes for EMSA by Klenow reaction	45
<i>in vitro</i> binding reaction of DNA and fusion protein	46
Non-denaturing polyacrylamide gel electrophoresis	46
Cell line and culture condition	47
CAT reporter vector construction	47
<i>AmDM1</i> expression vector construction	48
Transfection of plasmids into the mammalian cell line	49

CAT reaction	50
Bradford assay	51
Liquid Scintillation Counting	51
Data processing and statistical analyses	51
Section 3. Results	53
cDNA isolation of <i>AmDM1</i> homolog for <i>Nematostella vectensis</i>	53
NvDM5 protein has both the DM and DMA domain	54
DM domain of NvDM5 is highly conserve through evolution	54
5 clones obtained by yeast one-hybrid screen	56
Isolation of the Reporter plasmid from the screened clones	57
Nucleotide sequences of the recovered genomic fragments	59
Isolation of $\lambda$ -phage clone containing the 5B region	60
5B region was located within an intron of a homolog of the transmembrane transporter gene	60
<i>AmTAT1</i> is a single copy gene in <i>Acropora millepora</i> genome	61
Isolation of the full-length cDNA for AmTAT1	62
AmTAT1 is a transmembrane protein	63
Structural feature of AmTAT1	64
The conserved structure in evolution	65
Phylogenetic analysis on AmTAT1 within MCT family	66
Intron-exon organization of <i>AmTAT1</i> gene	66
5' regulatory region of <i>AmTAT1</i> has multiple potential sites for the binding with transcription factors	67

Gene expression of <i>AmTAT1</i>	68
Computer analyses could not find the DM-binding motif in the 5B fragment	69
Finding the responsible subfragment of 5B for the association with DM domain by yeast one-hybrid assay	69
Finding the responsible element in 5B $\alpha$ subregion for the association with DM domain by electrophoretic mobility shift assay (EMSA)	70
Transcriptional regulation by AmDM1 on <i>AmTAT1</i> expression	73
Strategy for CAT assay	74
AmDM1 acts as a transcriptional repressor on DM enhancer	75
AmDM1 activates SV40 promoter in Hep-2 cells	76
cDNA cloning of <i>vitellogenin (Acropora millepora)</i>	77
Isolation of the full-length cDNA	78
Molecular properties of <i>Acropora millepora</i> Vitellogenin (AmVit)	80
Genomic DNA cloning of <i>Acropora millepora vitellogenin</i>	81
Complicated and conserved intron-exon organization of <i>Acropora millepora vitellogenin</i> gene	81
No DM binding motif in the upstream regulatory region and introns	82
Section 4. Discussion	83
<i>Acropora</i> -specific repetitive sequences contain DM-binding elements	83
Identification of the AmDM1 recognition motif	84

AmDM1 downregulates the expression of <i>AmTAT1</i>	84
<i>AmTAT1</i> encodes a membrane transporter protein	85
Molecular evolution of <i>AmTAT1</i>	87
Roles of AmTAT1 in sex-determination	88
AmDM1 contains the activation domain in the C-terminal region	89
No DM-binding motif has been found in <i>vitellogenin</i> gene of <i>Acropora millepora</i>	90

### Chapter 3 Characterization of Two *Cnnos* Homologs of *Acropora millepora* 91

Section 1. Introduction	92
<i>nanos</i> was identified as the posterior determinant	92
<i>nanos</i> is required for development and function of germline cells	93
The functional aspects of Nanos protein	95
Biochemical features of Nanos	97
Posterior localization of <i>nanos</i> mRNA	98
Regulatory mechanism of <i>nanos</i> translation	99
Evolutionary conservation of <i>nanos</i>	100
<i>nanos</i> homologs in <i>C. elegans</i> (Nematoda)	100
<i>nanos</i> homologs in leech, <i>Helobdella robusta</i> (Annelida)	102
<i>nanos</i> homologs in vertebrates (Chordata)	102
Two <i>nanos</i> homologs in Hydrozoa (Cnidaria)	104
Strategy for the research on <i>Cnnos</i> and germline speciation	105

### Section 2. Materials and Methods 106

<i>E. coli</i> strains	106
Manipulation of bacteria	106
Manipulation of $\lambda$ -bacteriophage	106
Manipulation of nucleic acids and protein	106
Enzymes for molecular biology	106
Primers	106
Animal sampling	106
Screening <i>Acropora millepora</i> genomic library	106
Screening <i>Nematostella vectensis</i> genomic library	108
Sequencing lambda phage DNA	109
Screening <i>Acropora millepora</i> cDNA library	110
Screening <i>Nematostella vectensis</i> cDNA library	110
RT-PCR	111
Inverted PCR	112
Section 3. Results	115
<i>Cnnos1</i> cDNA isolation for <i>Acropora millepora</i>	115
<i>Cnnos1</i> cDNA isolation for <i>Nematostella vectensis</i>	116
<i>Nematostella vectensis</i> also has the second Nanos homolog	118
High conservation in the Zinc finger domains throughout evolution	119
Genomic DNA cloning for <i>Cnnos1</i> ( <i>Am</i> ) and <i>Cnnos2</i> ( <i>Am</i> )	119
Genomic DNA cloning for <i>Cnnos1</i> ( <i>Nv</i> ) and <i>Cnnos2</i> ( <i>Nv</i> )	120
5'-regulatory regions of cnidarian <i>nanos</i> genes have multiple potential transcription factor binding sites	121
Gene expression patterns of <i>Cnnos1</i> and <i>Cnnos2</i> in <i>Acropora millepora</i>	121

Section 4. Discussion	122
<i>nanos</i> gene duplicated before cnidarian radiation	122
Comparison of genomic structures of cnidarian <i>nanos</i>	123
Different expression patterns of <i>Cnnos1</i> and <i>Cnnos2</i> in <i>Acropora millepora</i>	124
<b>Figures</b>	125
<b>Tables</b>	209
<b>References</b>	212
<b>Appendix</b>	236
Abbreviation	237

# List of Figures

Figure 1	Anatomy of a coral polyp	127
Figure 2	Development of genital organs in <i>Drosophila</i>	129
Figure 3	Principles of yeast one-hybrid selection	131
Figure 4	CAT reporter vectors for DM-enhancer activity	132
Figure 5	Nested deletion constructs for <i>AmDM1</i> expression	132
Figure 6	NvDM5 cDNA sequence and translation	134
Figure 7	Alignment of DM domains	135
Figure 8	Phylogeny tree of DM domains	136
Figure 9	Alignment of DMA domains	137
Figure 10	Phylogeny tree of DMA domains	138
Figure 11	Alignment of cnidarian DM proteins	139
Figure 12	Nucleotide sequences of the DM-binding fragments	140
Figure 13	Alignment of repeat sequences	141
Figure 14	Phylogeny tree of the repeat sequences	142
Figure 15	Detection of the homologous proteins for 5B1b	143
Figure 16	Sequencing strategy for $\lambda$ -5B1b clone	144
Figure 17	Southern-blot analysis on <i>AmTAT1</i>	145
Figure 18	cDNA sequence of <i>AmTAT1</i>	147
Figure 19	Conserved domains in AmTAT1 protein	149
Figure 20	Structural features of AmTAT1 protein	151
Figure 21	Alignment of AmTAT1, X-PEST and hTAT1	153
Figure 22	Phylogenetic tree for MCTs including AmTAT1	154
Figure 23	Intron-exon organization of <i>AmTAT1</i>	156
Figure 24	Potential transcription factor binding sites in the 5'-flanking region of <i>AmTAT1</i>	157



Figure 25	Expression analysis on <i>AmTAT1</i> by RT-PCR	158
Figure 26	Subfragments for the analysis on the association with DM domain	160
Figure 27	One-hybrid assay with subfragments of 5B	162
Figure 28	EMSA on the subfragments of 5B	164
Figure 29	EMSA on the subfragments of 5B $\alpha$ A	166
Figure 30	Competitive EMSA with specific and non-specific cold probes	168
Figure 31	EMSA with genomic fragments	170
Figure 32	Nest-deleted probes for EMSA and the repetitive sequence	172
Figure 33	EMSA on double-strand oligonucleotides for the A2 primer region and the repetitive sequence	174
Figure 34	EMSA with the mutant A2 probe	176
Figure 35	Alignment of DM recognition sequences	177
Figure 36	Strategy for CAT assay	179
Figure 37	Construction of CAT reporter vectors	180
Figure 38	CAT assay on the enhancer activity of DM recognition motif	182
Figure 39	CAT assay with nested deletion mutants of AmDM1	184
Figure 40	Amino acid sequence and conserved domains of AmVit	186
Figure 41	Comparison of genomic organization of <i>AmVit</i> and <i>vgII</i>	188
Figure 42	Nucleotide sequences and potential transcription factor binding sites in <i>AmVit</i> gene	190
Figure 43	Nucleotide sequence of <i>Cnnos1(Am)</i> cDNA and the deduced amino acid sequence	192
Figure 44	Nucleotide sequences of <i>Cnnos1(Nv)</i> and <i>Cnnos2(Nv)</i>	

	cDNA	194
Figure 45	Alignment of the Zinc finger domains of Nanos proteins	196
Figure 46	Phylogeny tree for the Zinc finger domains of Nanos proteins	197
Figure 47	Genomic DNA sequences and deduced amino acid sequences of <i>Cnnos1(Am)</i> and <i>Cnnos2(Am)</i>	199
Figure 48	Southern-blot analysis on <i>Cnnos1(Am)</i> and <i>Cnnos2(Am)</i>	203
Figure 49	Potential transcription factor binding sites in the 5'-flanking regions of cnidarian <i>nanos</i> genes	204
Figure 50	Expression analysis on <i>Cnnos1(Am)</i> and <i>Cnnos2(Am)</i> by RT-PCR	207

## List of Tables

Table 1	Radiolabelled probes for EMSA and the combination of oligonucleotide primers	209
Table 2	Potential phosphorylation sites for Protein kinase C and Casein kinase II in NvDM5 protein	209
Table 3	Summary of 5-FOA test	210
Table 4	Introns of <i>AmVit</i>	211

## Statement of Sources

I declare that the work presented in this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institute or tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Hiroki Go (Chikuei Wu)

Date

-----

-----

# Chapter 1

## General Introduction

Sexual reproduction is a common feature in the animal kingdom and its prototype is also seen in protozoan life cycle. In general, animals develop female and/or male sexual organs (gonads) that consist of somatic cells, while the germ line cells separate from the somatic development during the early embryonal stages and subsequently differentiate to mature germs, eggs and sperms.

Although the sexual reproduction is widely seen and conserved through evolution, the mechanisms of sex-determination are quite different from species to species. Taking a look at vertebrates, for example, we can find several kinds of sex-determination system. Mammalian sexes are genetically determined but the sex of some species of reptiles is temperature-dependent (Pieau and Dorizzi, 2004). A lot of fish species are reported to switch their sex in accordance with the maturity. As for the germ line specification, two distinct strategies are identified. In mice, the germ line specification and establishment have been shown to be controlled by the cellular interactions mediated by signal molecules, bone morphogenetic proteins BMP4 (Winnier et al., 1995; Lawson et al., 1999; Fujiwara et al., 2001) and BMP8b (Ying et al., 2001). The other strategy is dependent on the germ plasm in the eggs (reviewed in Houston and King, 2000).

In the phylogeny of metazoa, phylum Cnidaria is the basal animal to all the higher and triploblastic phyla. The body structure is very simple and at the same time primitive. However, they have functionally differentiated body parts which are also morphologically distinctive. Within this phylum, current views tell us that class Anthozoa (including corals and sea anemones) is the basal (Medina et

al., 2001).

Corals, our model system, are well-known for their characteristic strategy of sexual reproduction. Annual mass spawnings are observed in the spring in Great Barrier Reef region. The manners of sex-determination and reproduction are complex. Approximately 80% of corals are spawners which release eggs and sperms into water column. The rest are brooders which keep fertilized eggs in the polyps until the embryos develop to planulae. Among the spawners, roughly 80% are hermaphrodites, while the others are gonochorists. Moreover, hermaphrodites are classified into simultaneous and sequential hermaphrodites: the distributions are 90% and 10%, respectively (Harrison and Wallace, 1990). Our model species, *Acropora millepora*, is a simultaneously hermaphroditic spawner that is the most common type especially in Great Barrier Reef.

Cnidaria has a radial and diploblastic body structure without clearly defined organs. Although reproductive organs are absent, they develop temporary gonads where gametogenesis takes place. In corals, the development of gonads and the differentiation of germ cells occur in their mesentery (Figure 1).

As mentioned above, corals are basal animals to all the higher metazoa with complicated morphology and function. The knowledge about the mechanism of coral reproduction will provide us important clues to understand evolution and developmental process of the reproductive system. However, very little has been known so far about the molecular mechanisms of sex-determination and germ line differentiation of order Scleractinia.

Now I focus on two genes that are highly conserved throughout metazoan evolution, *nanos* and *doublesex (dsx)*. *nanos* probably plays some critical roles in germ line cells, while *doublesex* takes part in the cascade of sex-determination.



## Chapter 2

Involvement of *AmDM1*, the  
*doublesex* Homolog of Coral  
*Acropora millepora*

## Section 1. Introduction

### *doublesex* is near the bottom of the sex determination hierarchy in *Drosophila*

The molecular mechanism of sex-determination has long been studied for *Drosophila melanogaster* (Christiansen et al., 2002; Baker, 1989). *doublesex* was identified as a gene of which defect caused the intersex phenotype in both chromosomal male and female (Hildreth, 1965). And this gene is known to be near the bottom of the sex-determination hierarchy.

The somatic sexual dimorphism in *Drosophila* is cell-autonomous and dependent on the X chromosome / autosome ratio (X/A ratio). And the principle is that the sex-determinant genes function for feminization. In other words, maleness is the result of 'default' pathway. In wild type females, where the X/A ratio is 1, *sex lethal* (*sxl*) is activated by alternative splicing (Bell et al., 1988; Salz et al., 1989). The female-specific splicing is enhanced by Sxl protein (Sakamoto et al., 1992). The female type Sxl is only functional. While Sxl prevents the dosage compensation by inhibiting *msl-2* translation (Marín et al., 2000; Amrein, 2000), it binds to a splice site in *transformer* (*tra*) primary transcript leading to the female-specific alternative splicing (Inoue et al., 1990). The female type Tra protein is only functional as well.

Two pathways in the sex-determination cascade have been known to be in the downstream of Transformer. Tra protein coupled with the

constitutively expressed Tra2 protein brings about the female-specific alternative splicing to *fruitless* and *doublesex* transcripts. In *fruitless* pathway, only male type protein is functional and it is responsible for the male-specific function (the courtship behavior) of the central nervous system (Ryner et al., 1996). *doublesex* transcripts are translated into male or female type proteins which primarily take part in the somatic sexual dimorphism in each sex.

Female-specific ( $dsx^F$ ) and male-specific ( $dsx^M$ ) *doublesex* mRNAs are generated by alternative splicing (Baker, 1988). In *doublesex* transcripts, the 3' splice site of intron 3 is cryptic and skipped in the default state (male). Consequently, the 4th exon is removed and the mRNA is polyadenylated at exon 6. Instead, the exon 4 has *cis*-acting splicing enhancers (Nagoshi and Baker, 1990; Ryner and Baker, 1991). This enhancer is composed of 6 repeats of a 13 nucleotide sequence and one purine-rich element and provides the binding regions to the protein complex including Tra and Tra 2. Both Tra and Tra 2 contain SR (serine-arginine-rich) domains which are involved in protein-protein interaction. Tra and Tra 2 molecules form a higher order complex with SR family proteins and bind to the splice enhancer in the exon 4 (Lynch and Maniatis, 1996). Tra doesn't have RNA-binding motifs, whereas Tra 2 and SR proteins contain the RNP-type RNA-binding motifs. As a result, the 3' splice site is activated and female-specific *doublesex* mRNA ( $dsx^F$ ) has the 4th exon. Due to the polyadenylation in the exon 4, the C-terminal is shorter in  $Dsx^F$  in comparison with  $Dsx^M$ .

*doublesex* needs *intersex* (*ix*) and *hermaphrodite* (*her*) for its proper function. *ix* is required for female sexual development and the mutant

of this gene shows female-specific intersex phenotype. *ix* and *dsx<sup>F</sup>* are interdependent for the regulation of *yp1* (*yolk protein 1*) expression and several female phenotypes (Waterbury et al., 1999). And the physical interaction has been shown for IX and Dsx<sup>F</sup> proteins. It is suggested that IX provides the transactivating function to Dsx<sup>F</sup> which lacks the activation domain (Garrett-Engle et al., 2002). Recently, the mammalian homolog of Intersex was reported to be a subunit of Mediator complex which was a multi-protein co-activator of RNA polymerase II (Sato et al., 2003). *her* is expressed in all developmental stages in both sexes and participates in various aspects of sex-specific differentiation. *her* and *dsx* work dependently or independently (Li and Baker, 1998).

### **Identification of DNA-binding DM domains**

The *yolk protein* genes in *Drosophila* provide the classical model for the transcriptional control regulated by Dsx proteins. The regulatory region contains an enhancer element (*dsxA*) to which Dsx proteins bind. While Dsx<sup>F</sup> enhances the transcription of *yolk protein* genes, Dsx<sup>M</sup> represses it. Thus the protein is expressed in a sex-specific manner (Coschigano and Wensink, 1993; An and Wensink, 1995a; An and Wensink, 1995b). The N-terminal region of Dsx<sup>F</sup> and Dsx<sup>M</sup> is identical and the physical association of *dsxA* element with both Dsx proteins were demonstrated (Cho and Wensink, 1997). The *in vitro* study also indicated that two Dsx protein molecules formed a homodimer with high affinity to *dsxA* which contained the pseudo-palindromic sequence (Cho and Wensink, 1996). In fact, the N-terminal domain has both a DNA-binding domain and an oligomerization domain (Cho and Wensink, 1998).

The male sexual regulatory gene, *mab-3* (*male abnormal-3*) was isolated from nematode *Caenorhabditis elegans*. Surprisingly, Mab-3 turned out to be homologous to Dsx. A conserved domain was identified in their N-terminal region by aligning the amino acid sequences and it was named DM (Doublesex and Mab-3) domain . Moreover, a human homolog, *dmrt1*, was isolated. It was shown that the expression of *dmrt1* was restricted in testis and this gene was mapped in the distal arm of chromosome 9, a locus close to congenital 46, XY sex-reversal. Dsx<sup>F</sup> and Dsx<sup>M</sup> proteins have one DM domain in the N-terminal sex-non-specific region. On the other hand, Mab-3 has two slightly different DM domains (Mab-3a and Mab-3b). *mab-3* promotes the differentiation of sensory rays and represses the transcription of *vitellogenin* gene. The transgene of *Drosophila* Dsx<sup>M</sup> restored the ray formation in the *mab-3* mutant animals (Raymond et al., 1998).

### **Structure and properties of DM domain**

The DM domain of Dsx protein is approximately 70 amino acid residues in length and composed of two regions: the Zinc finger module (about 40 amino acids at the N-terminal) and the tail region (about 20 amino acids at the C-terminal). The former contains two atypical Zinc fingers. These fingers are CCHC-HCCC type and provide sites for two Zinc divalents. The tail region is a part of  $\alpha$ -helix and recognizes the minor groove of the consensus DNA element. There is a strong functional and structural constraint in the primary sequence of DM domain. And mutations at a variety of sites in both the Zinc finger module and the tail region disrupt the DNA-binding ability and

result in the intersex phenotype ( Zhu et al., 2000; Narendra et al., 2002).

The recognition consensus for the DM domain of Dsx is (A/G)NNAC(A/T)T (T/A)GTNN(C/T), forming a pseudo-palindrome. The homodimerization of two DM domains is induced by their association with the recognition element. In *C. elegans*, the DM recognition consensus is asymmetric. It has been suggested that Mab-3a domain and Mab-3b domain form an internal heterodimer on the recognition sequence.

Mab-3 functions as a transcription factor in *C. elegans* as well as Dsx in *Drosophila*. Aforementioned, *vitellogenin* gene is under the control of *mab-3* and the gene expression is repressed in male worms. In the promoter region of six *vitellogenin* genes in *C. elegans*, one or more Mab-3 responsive elements have been found (Yi and Zarkower, 1999). Thus the structure and function of DM containing genes are conserved in distant phyla, Arthropoda and Nematoda.

### **Downstream genes of *dsx***

Dsx and Mab-3 proteins were shown to be transcription factors directly regulating *yolk protein* and *vitellogenin* genes, respectively. These examples suggest that both DM domain containing genes are involved in the proper egg production. In addition to them, several genes related to somatic sex-determination have been found to be regulated directly or indirectly by *dsx*.

The sexual dimorphism is the most apparent in genital organs. The

development of *Drosophila* genital system has been intensively studied with respect to the regulatory function of *dsx* (reviewed in Sánchez and Guerrero, 2001; Vincent et al., 2001; Burtis, 2002; Christiansen et al., 2002). During metamorphosis in *Drosophila*, the adult structure forms by differentiation of imaginal discs. The genital organs of both sexes develop from the genital imaginal disc which is derived from the abdominal segment 8 (A8), 9 (A9) and 10 (A10). The analia develops from A10 in both sexes. In female, the genital develops mainly from A8, whereas A9 forms parovaria. On the contrary, the male genital develops from A9. The massive growth of A8 is repressed and the 8th tergite is formed.

The product of *branchless* (*bnl*) is the fibroblast growth factor (FGF). In mature genital disc, FGF and FGF receptor are expressed in male but not in female. It was shown that  $Dsx^M$  promoted the transcription of *bnl* gene and as a result, mesodermal cells were incorporated in the genital disc epithelia subsequently developing into paragonia and vas deferens (parts of male genital organ). In female,  $Dsx^F$  represses the *bnl* transcription (Ahmad and Baker, 2002). In the upstream regulatory region of *bnl* gene, multiple binding motifs for Dsx were found. This mechanism looks somewhat conserved in evolution because in mouse FGF9 is required for the testicular embryogenesis and defect of this gene leads to XY sex-reversal phenotype (Colvin et al., 2001).

*dsx* functions in A/P organizers by influencing the *hedgehog* pathway to regulate the sexually dimorphic pattern of development (Figure 2) (Keisman et al., 2001). Decapentaplegic (Dpp) and Wingless (Wg) are expressed in the A/P organizers in A8 and A9. Both are in the

downstream of the *hedgehog* cascade. In female, Dsx<sup>F</sup> inhibits the production of Dpp by blocking the signal in A9. On the other hand, Dsx<sup>F</sup> promotes the cellular response to Wg protein in A8. In contrast, Dsx<sup>M</sup> inhibits the response to Wg in A8 and promotes the production of Dpp in A9. These promotion and repression by *dsx* is supposed to be in collaboration with a homeotic gene, *abdominal-B (abd-B)* (Sánchez et al., 2001). *dachshund (dac)* was found to be another downstream gene of *dsx* in the genital development. *dac* expression is controlled by Wg and Dpp. Evidence shows that in cells expressing Wg, Dsx<sup>F</sup> promotes but Dsx<sup>M</sup> represses *dac* expression. In cells expressing Dpp, Dsx<sup>F</sup> represses but Dsx<sup>M</sup> promotes *dac* expression (Keisman and Baker, 2001).

Outside the genital organs, the pigmentation of abdomen is also different between two sexes of *Drosophila*. The black pigmentation on A5 and A6 is a representation of male sex. Bric-a-brac (*Bab*) is a transcription factor that represses the abdominal pigmentation. The homeotic gene *abd-B* blocks the expression of *bab* in male. However, Dsx<sup>F</sup> inhibits this blocking action in these segments of female abdomen (Kopp et al., 2000).

The regulatory effects of *dsx* genes expressed in *Drosophila* head have been described. Applying serial analysis of gene expression (SAGE) to find male- and female-specific genes in the head, 46 genes were found to be expressed in sex-biased manners (Fuji and Amrein, 2002). Four novel genes among them were characterized: *tsx* (turn on sex-specificity), *sxe1* (*sex-specific enzyme 1*), *sxe2* (*sex-specific enzyme 2*) and *fit* (*female-specific independent of transformer*). *tsx* is expressed in fat cells of males. And the encoded protein is the



odorant/pheromone-binding protein. Expression of *tsx* turned out to be dependent on *transformer* and *doublesex*. *sxe1* and *sxe2* are specific to male fat cells as well. Expression of these genes are *tra*- and *dsx*-dependent. *sxe1* encodes a cytochrome P450 protein and is implicated to take part in the ecdysone metabolism. *sxe2* encodes a phospholipase. Both *sxe1* and *sxe2* were shown to be dependent on *tra* and *dsx* pathway. Transgene study of *tsx* demonstrated that it did not affect the mating performance in male, whereas the escape response of female animals were increased resulting in a lower mating performance.

*takeout* was reported to be at downstream of *dsx* pathway (Dauwalder et al., 2002). It is male-specifically expressed in fat body around brains but not in central nervous system. The *in situ* hybridization study detected the *takeout* in the 3rd antennal segment in a non sex-specific manner. This gene encodes a secreted protein and its expression is circadian-responsive and starvation-inducible. The upstream regulatory region of *takeout* gene has potential *doublesex* recognition motifs revealing the direct transcriptional regulation by Dsx. Male flies homozygous for *takeout* mutant show a significantly low level for courtship behaviour, which is restorable by introduction of the wild type *takeout*. *takeout* expression is dependent not only on *dsx* but on *fruitless* (*fru*). Dsx<sup>M</sup> and Fruitless cooperatively activate *takeout* expression, whereas Dsx<sup>F</sup> represses independently of *fruitless*. The possibility has been suggested that *takeout* plays a redundant role with *fru*-dependent factors in male courtship behaviour.

The roles of Dsx in *Drosophila* neurobiology has been analyzed by mapping the expression in the central and peripheral nervous

systems. Dsx proteins are detected CNS in 3rd instar larvae, pupae and adults. Several specific neurons in CNS and PNS are detected positive for Dsx proteins in sexually differentiated patterns (Lee et al., 2002). According to the analysis, Dsx was shown to be expressed in regions responsible for the early steps in courtship (the dorso-posterior brain) and the male song sounds (the anterior ventral nerve cord region). It was also revealed that Dsx was involved in the sexually dimorphic patterns of mitosis in the abdominal ganglionic neuroblasts.

Mammalian PTB (polypyrimidine tract binding protein) is a heterogeneous ribonucleoprotein which is expressed ubiquitously. It has four RNA recognition motifs and regulates the splicing of multiple mRNAs. Unlike mammalian PTB, *Drosophila* PTB (*dmPTB*) is expressed only in male germline and has been shown to be required for the germline differentiation (Robida and Singh, 2003). The mutant male for PTB doesn't produce normal sperm. A close link between *dsx* and *dmPTB* has been demonstrated. In XY flies with *dsx* mutation, the expression of *dmPTB* is significantly decreased, whereas in XX flies with *dsx* mutation, *dmPTB* is not expressed. It is likely that *dmPTB* is at the downstream of *dsx* pathway. The evidence shows that Dsx<sup>M</sup> is necessary but not sufficient for *dmPTB* expression.

*dsx* has been identified as the somatic sex-determinant. However, as seen for *dmPTB* regulation, *dsx* is also involved in the pathway for the germline sex-determination *via* soma-germline communication. The participation of Dsx<sup>F</sup> in the germline sex-determination was reported as well (Waterbury et al., 2000).

## Evolutionary conservation of DM domain genes

Since the identification of DM domain, a conserved zinc finger domain with DNA-binding potential, a number of DM domain genes have been found or isolated for various metazoan phyla including Arthropoda, Nematoda, Chordata and Cnidaria (reviewed in Hodgkin, 2002). And the molecular evolution of genes in this category was studied (Volff et al., 2003).

### *dsx* and DM domain genes in insects

Although the sex-determination systems in insects are only poorly conserved even in dipterans (Schütt and Nöthiger, 2000; Graham et al., 2002), the function and the gene structure of the terminal regulator, *dsx* is well conserved.

4 DM domain genes including *dsx* have been found in *Drosophila melanogaster* and 3 in malaria mosquito, *Anopheles gambiae*. As for *dsx* in mosquito, the female-specific expression of hexamerin was reported and potential Dsx-binding elements were found in the 5' upstream region of this gene (Zakharkin et al., 2001).

In silkworm, *Bombyx mori*, female and male type *dsx* transcripts are generated by alternative splicing with a different pattern from that of *Drosophila* (Suzuki et al., 2001). Transgene study of female type *dsx* (*BmDsx<sup>f</sup>*) into male animals resulted in changes in expression level of some genes (Suzuki et al., 2003). The *vitellogenin* gene (evolutionarily unrelated to *yolk protein*) expression was increased and the *SP1* (*hexamerin*) gene was also increased. In the 5' flanking

region, a potential consensus for DM-binding site (ACATTGT) for BmDSX was found and the association of Dsx and this motif was shown in EMSA. PBP (pheromone binding protein) is normally expressed in male antennae and functions as a receptor for female pheromone. The expression of this gene was shown to be repressed by *BmDsx<sup>F</sup>* transgene. Albeit *BmDsx<sup>F</sup>* have substantial effects on sex-specific gene expression, it did not bring about any morphological change. It was speculated that co-factors might be required for BmDsx function as *dsx* needs *intersex* for its activity in *Drosophila*.

*Musca domestica* has a different sex-determination cascade from that of *Drosophila melanogaster* at the level of primary sex-determinant. While *sex lethal* is the primary determinant in *Drosophila*, F factor is the signal at the top of the hierarchy in *Musca*. However, the two fly species have *dsx* in common at the bottom of these systems. *Mddsx* (*musca domestica dsx*) undergoes the sex-specific alternative splicing generating female- and male-specific *Mddsx<sup>F</sup>* and *Mddsx<sup>M</sup>* (Hediger et al., 2004). Both variants have the DM domain. *MdDsx<sup>F</sup>* expression in *Drosophila* and *Musca males* led to the activation of *vitellogenin* gene, whereas *MdDsx<sup>M</sup>* in *Drosophila* female caused male-like pigmentation in 5th and 6th tergites. Interference study on *Mddsx* with RNAi showed the intersexual differentiation of gonads in both sexes. Taken together, *Musca* and *Drosophila* share the common gene and function at the bottom of the sex-determination cascades.

### ***mab-3* and DM domain genes in nematode**

In nematode, *Caenorhabditis elegans*, 11 DM domain genes have been found. Apart from *mab-3*, only one gene, *mab-23*, is known to be

related to sex-determination. As described before, *mab-3* represses *vitellogenin* gene expression in male intestine and promotes V ray differentiation of which defect in *mab-3* mutant males is restorable by *dsx<sup>M</sup>* transgene. And *mab-3* was directly regulated in the intestine by *tra-1A* (Yi et al., 2000). In the 5' upstream region of *mab-3*, a TRA-1A binding site was found.

For the male sensory ray differentiation, MAB-3 functions synergistically with a neurogenic bHLH (basic helix-loop-helix) transcription factor, LIN-32. It is also suggested that *mab-3* plays roles in various aspects in the formation of male sensory organs. Two genes, *srd-1* and *lov-1* have been shown to be at the downstream of *mab-3* pathway (Yi et al., 2000). *srd-1* encodes a putative chemosensory receptor protein (Troemel et al., 1995) and *lov-1* encodes a putative cell-surface protein required for the normal mating behavior of male worms (Barr and Sternberg, 1999).

Another DM domain gene, *mab-23* in *C. elegans* was reported to be necessary for sex-specific differentiation and mating behavior of male. The male defective mutants for *mab-23* are abnormal in tail morphology and gonadal differentiation, and defective in the ventral posterior muscle and copulatory behavior. The differentiation defects of male sensory neurons and sex muscle lead to deficiency of tail turning during copulation. Gonadal abnormality causes the obstruction in sperm-passage. As a result, the *mab-23* defective males are infertile. Contrary to these observation in male, the defect does not bring about any change in hermaphrodites. For ray neuron differentiation, a model in which *mab-23* functions through interaction with *egl-5* (a *Hox* gene) and DBL-1 (TGF $\beta$ -like pathway)

has been proposed (Lints and Emmons, 2002).

### **DM domain genes in vertebrate**

DM domain containing genes have been identified in all vertebrate groups: mammals, birds (Shan et al., 2000), reptiles (Smith et al., 1999); Kettlewell et al., 2000), amphibians (Shibata et al., 2002) and fish. *dmrt1* (*DM-related transcription factor 1*) homologue is Z-linked in bird resulting in a higher dosage in ZZ males. In reptiles in which the sex-determination depends on the temperature, the expression of *dmrt1* has been shown to be high in embryos exposed to male-determining temperature. In particular, a lot of study has been done for mammals and fish. 8 orthologous groups have been found in vertebrate. As for fish, DM domain genes have been identified for 5 of those groups (Volff et al., 2003).

*terra* is the first DM homologue found in fish (Meng et al., 1999). It was identified in zebra fish. However, unlike *dsx*, *terra* is involved in somitogenesis. It was shown that human and mouse had *terra* homologues and the gene expression was detected in mouse somites. *dmrt2*, the human homologue of *terra* was also shown to be involved in somitogenesis (Ottolenghi et al., 2000).

5 subfamilies of *dmrt* genes have been identified in several species of fish such as rice field eel (Huang et al., 2002), tilapia (Guan et al., 2000), puffer fish (Brunner et al., 2001) and medaka fish (Brunner et al., 2001; Matsuda et al., 2002; Nanda et al., 2002). Sexually dimorphic expression was reported for *tdmrt1* and *tdmo* (*DM-domain gene in ovary*) in tilapia (Guan et al., 2000). These are encoded in

distinct genes. While *tdmrt1* is specifically expressed in testis, *tdmo* is expressed in ovary. The comparative genomics about *dmrt1*, *dmrt2* and *dmrt3* between *Fugu rubripes* and human revealed conservation of synteny of human chromosome 9 in *Fugu*. Also, comparison of genomic sequences led to the identification of 3 regions conserved in the non-coding region, which were supposed to be involved in the regulation of gene expression (Brunner et al., 2001). However, gene expression study for *dmrt1*, 2 and 3 did not reveal any function in the primary step in male sex-determination of these fish.

Medaka fish is a male heterogametic species (male: XY, female: XX). No major difference has been found between X and Y chromosomes other than the sex-determinant. A DM family gene, *DMY* was identified and located in Y chromosome (Matsuda et al., 2002). The expression of *DMY* is detected exclusively in the somatic cells surrounding germ cells in XY embryos and Sertoli cells in adult testes (Nanda et al., 2002). Interestingly, XY females are found in wild population with mutation in *DMY* or transcriptional repression of this gene. In the genome of medaka fish, *dmrt1*, 2 and 3 are autosomal whereas *DMY* (*dmrt1Y*) is Y-linked. It has been shown that *DMY* is a duplicated copy of the autosomal *dmrt1* (Nanda et al., 2002).

Identification of the DM domain as a evolutionarily conserved domain led to the finding of a human homologue, *DMT1* (*dmrt1*) (Raymond et al., 1998) and DM family genes in mammals. 8 *dmrt* orthologous groups have been reported so far (reviewed in Volff et al., 2003).

The involvement of *dmrt1* in sexual development has been reported repeatedly. In mice, *dmrt1* expression is detected in the genital ridges

of early embryos for both XX and XY animals and in later stages, the expression is only observed XY-specifically in the seminiferous tubules (germ cells and Sertoli cells), but not in interstitial cells including Leydig cells (Raymond et al., 1999). The function of *dmrt1* is required for the normal differentiation of mammalian testes as evidenced by the knockout study (Raymond et al., 2000). The male knockout mice show severe defects in adult testes and infertility, whereas the female mice have normally differentiated ovaries and fertility. 7 *dmrt* genes (*dmrt1-7*) have been identified in mouse. Among them, *dmrt1*, 2, 3, 5 and 7 are expressed in sex-specific patterns (Kim et al., 2003) during development. The expression of these genes are also detected outside gonads.

Human *dmrt1* was found as a candidate for the responsible gene for 46, XY sex-reversal. *dmrt1* is located in chromosome 9p24.3 which has been known to be critical for the sex-reversal. However, according to a case study on 46, XY sex-reversed sisters with 700kb deletion at 9p24.3 region, *dmrt1* and *dmrt2* were intact (Calvari et al., 2000). In this case, the break point was less than 30kb telomeric to *dmrt1* gene. Thus it was speculated that haploinsufficiency of *dmrt1* caused the sex-reversed phenotype. Analyses on the 550kb contig on 9p24.3 revealed that this region contained several transcripts including *dmrt* genes, which were expressed in the urogenital system (Ottolenghi et al., 2000). In spite of the accumulating data on the involvement of *dmrt* genes in 46, XY sex-reversal, a majority of the cases remains unexplained (Ottolenghi and McElreavey, 2000). 8 DM genes (*dmrt1-8*) have been identified in human genome, most of which are revealed to be expressed in gonads. They are mapped in three well-defined autosomal regions (1p, 9p and 19p) and one gene



on 19p (*dmrt7*) has an additional homologue (*dmrt8*) on Xq (Ottolenghi et al., 2002).

### ***AmDM1*, a *doublesex* homolog in coral**

A DM-containing gene has already been isolated and characterized for *Acropora millepora* (*AmDM1*) (Miller et al., 2003). The cDNA of *AmDM1* encodes a protein of 470 amino acids in length. The homology search of the amino acid sequence outside the DM domain led way to the discovery of a novel conserved motif named DMA domain. Seven genes including two hypothetically deduced proteins have been found to contain DMA domains and some are sex-related genes. All of them also contain DM domains. Expression analysis of *AmDM1* detected four transcripts of different length. Combined with the information on the structure of genomic DNA, they are suggested to be generated by making use of the distinct transcription starts, the differential patterns of splicing and the alternative polyadenylation. As a result, 3.4k and 2.5k mRNA encode both DM and DMA domains, whereas 2.7k and 1.8k messengers only contain DMA domain. The level of gene expression showed all transcripts were highly expressed at the coral tips in October (the spawning season). But only 2.7k and 1.8k mRNAs (non-DM transcripts) were detected in the early stage embryos (pre-settlement). This finding suggests that non-DM protein acts inhibitory to sexual development in the embryo.

### **Strategy for the research on *AmDM1* and sex-determination**

Like other DM domain-containing genes, the coral homolog of *doublesex*, *AmDM1* is thought to encode a transcription factor.

According to the ‘reverse-order’ hypothesis (Wilkins, 1995), genes at the upstream in the sex-determination cascade are likely to evolve faster than genes at the downstream. Consequently, the genes at or near the bottom such as *doublesex* is the most conserved in structure and function.

The aim of my study is to know the function of *AmDM1* in the sex-determination hierarchy of *Acropora millepora*. The information about immediate downstream gene(s) of which transcription is *trans*-regulated by AmDM1 will give us the clues to investigate the functional aspects of this gene. Thus, I started my research by trying to find the downstream genes of *AmDM1*.

## Section 2. Materials and Methods

### ***E. coli* strains**

NM522, DH5 $\alpha$  and JM109 were used for the general plasmid propagation and cloning. BL21 was for protein expression. LE392, XL1-Blue MRF' and Y1090(r-) were for propagation of  $\lambda$ -phage strain,  $\lambda$ -GEM12,  $\lambda$ -UniZap XR and  $\lambda$ -ZipLox, respectively. SOLR was for the phagemid excision.

### **Manipulation of bacteria**

Bacteria were manipulated with the standard methods (Sambrook and Russell, 2001).

### **Manipulation of $\lambda$ -bacteriophage**

$\lambda$ -bacteriophages were manipulated with the standard methods (Sambrook and Russell, 2001) unless specified.

### **Manipulation of nucleic acids and protein**

Nucleic acids and proteins were manipulated with the standard molecular biology technique (Sambrook and Russell, 2001) or according to the manuals provided by the manufacturers of the materials. Plasmids were prepared with QIAGEN Plasmid Mini Kit (QIAGEN, Valencia, CA). DNA fragments were extracted from agarose gel by using QIAquick Gel Extraction Kit (QIAGEN, Valencia, CA). Other

materials and kits are specified in the following sections.

### **Enzymes for molecular biology**

All enzymes except those specified were purchased from Promega (Promega, Madison, WI).

### **Oligonucleotide primers**

Primers were synthesized by Sigma Genosys Australia (Sigma Genosys Australia, Castle Hill, NSW) or Proligo Australia (Proligo Australia, Lismore, NSW).

### **Animal sampling**

Eggs and larvae of *Acropora millepora* were sampled on spawning in November 2001 in Geoffrey Bay, Magnetic Island, Queensland. For RNA extraction, animals were collected in cryostat tubes and frozen in liquid nitrogen. A mature colony of *Acropora millepora* in Nelly Bay, Magnetic Island was tagged for repeated sampling (from November 2001 to December 2003). Branch tips of this colony were snapped by hand and kept frozen in liquid nitrogen until needed. All sampling was permitted by Great Barrier Reef Marine Park Authority.

### **Screening *Acropora millepora* cDNA library**

$\lambda$ -UniZap XR *Acropora millepora* cDNA library was constructed by Dr. David Hayward (Australian National University) with the messenger RNA from adult coral tips sampled in October. Approximately

500,000 clones were screened by the standard plaque hybridization in the primary screening. Hybond-N nylon membranes and Hybond-C nitrocellulose membranes were purchased from Amersham (Amersham, Piscataway, NJ) and used for the plaque lift and hybridization. Radiolabelled probes were prepared by PCR amplification.

For the cloning of *AmTAT1* cDNA, two probes were designed for screening. One fragment of 690bp which contained the entire exon3 of *AmTAT1* was amplified and used for the probe (exon3 probe). The genomic DNA of *Acropora millepora* was the template for this reaction. The primer pair for this amplification was:

tkzUp : 5'-TTGTCCTTGTTAGCCTGGGTT-3'

tkzDwn : 5'-GCAATTCGGCATGTATCTTCC-3'

The program of the PCR consisted of 30 times at 95°C for 45sec., 50°C for 30sec. and 72°C for 1min.

The other fragment of 185bp which contained the partial exon2 of *AmTAT1* was amplified and used for the probe (exon2 probe). The genomic DNA of *Acropora millepora* was the template for this reaction. The primer pair for this amplification was:

tkzf1Up : 5'-ATATTCCTCCCTGGAGATGC-3'

tkzf1Dwn : 5'-TCTCCCAAGCTGTGCTCATT-3'

The program of the PCR consisted of 30 times at 95°C for 45sec., 50°C for 30sec. and 72°C for 30sec. The labelled probe was purified with Ultra Clean PCR Clean-up / DNA Purification Kit (GeneWorks, Adelaide, SA).

The hybridization was performed at 65°C overnight in 5X SSC, 5X Denhardt's solution, 0.5% SDS. The hybridized membranes were

washed according to the manufacturer's instruction: in 500ml of 2X SSC / 0.5% SDS at room temperature for 5min., in 500ml of 2X SSC / 0.1% SDS at 65°C for 30min. and in 500ml of 1X SSC / 0.1% SDS at 65°C for 30min. The membranes were air-dried and the radioactive signals were exposed on Phosphoscreen and visualized with Phosphorimager (Molecular Dynamics, Sunnyvale, CA). The areas on the agar plates with promising signals on the phosphoimage were picked up and stored in SM buffer at 4°C. The viral suspension for the samples picked-up as positive on the primary screen was tested by PCR to confirm that the samples contained the targeted clone. Only positive samples for this PCR-based test were brought to the secondary screening. The plaque hybridization for the secondary screening was carried out in the same condition and with the same probe as those for the primary screening.

For the cloning of *vitellogenin* (*AmVit*) cDNA, the von Willebrand domain of the cDNA was amplified with a primer pair which corresponded to the highly conserved regions in the vitellogenin of *Favites chinensis* (Imagawa, 2004). The sequences of these primers were:

FcVit(F1) : 5'-TGCCCTTACGTCTTGGTTCAGGA-3'

FcVit(R1) : 5'-CCATTAGCGTCACCACAAAGCCC-3'

The thermal cycling protocol comprised 5 times at 95°C for 1min., 42°C for 30sec. and 72°C for 30sec., followed by 30 times at 95°C for 1min., 50°C for 30sec. and 72°C for 30sec., and an additional extension at 72°C for 2min. The product was purified and sequenced to confirm that the encoded peptide had homology to the von Willebrand domain of vitellogenin of *Favites chinensis*. The 1st strand cDNA synthesized with total RNA extracted from the adult coral tip of

*Acropora millepora* sampled in October 2003 was used for the template.

This fragment was further used as the template for preparation of radioactive probe. The probe was radiolabelled by PCR with the primer pair mentioned above. The program of the PCR consisted of 30 times at 94°C for 45sec., 50°C for 45sec. and 72°C for 1min., followed by a final extension at 72°C for 3min. The labelled probe was purified as well. The hybridization was performed at 65°C overnight in 5X SSC, 5X Denhardt's solution, 0.5% SDS. The hybridized membranes were washed with following protocol : in 500ml of 2X SSC / 0.1% SDS at room temperature for 10min., in 500ml of 1X SSC / 0.1% SDS at 65°C for 15min. The picked-up samples were selected by amplifying the probed region in the cDNA prior to the secondary screening. The plaque hybridization for the secondary screening was carried out in the same condition and with the same probe as those of the primary screening.

Phagemid vectors, pBluescript SK(-), which harbored the cloned cDNAs were recircularized by the excision protocol with ExAssist Interference-Resistant Helper Phage (Stratagene, La Jolla, CA) according to the manufacturer's instruction. The cDNAs were sequenced by primer walking in combination with subcloning. The sequencing reaction was performed with ET-Terminator (Amersham Biosciences, Piscataway, NJ) according to the provided manual. The sequence data were analyzed and processed with Sequencher version 3.0 and version 4.2(Gene Code Corporation) and subjected to the homology search. Sequence homology search for the possible translation was carried out using the blast-t program in the BLAST

server at the National Center for Biotechnology Information (<http://www.ncbi.nih.gov/index.html>).

### Screening *Nematostella vectensis* cDNA library

$\lambda$ -ZipLox cDNA library of *Nematostella vectensis* was provided by Dr. John Finnerty (Boston University). Approximately 300,000 clones were screened. With the genomic DNA of *Acropora millepora* as the template, a fragment corresponding to the entire length of DM domain was amplified and made radioactive by PCR. The primers used for this reaction are:

DM-FW : 5'-CGCGGCCGCTCTCCAAAATGCGCCCGAT-3'

DM-RV : 5'-GCTCGAGGGCCTCCGGCTTCTTCGCTTT-3'

The thermal cycling protocol consisted of 30 times at 94°C for 45sec., 60°C for 45sec. and 72°C for 1min., followed by one time at 94°C for 1min., 55°C for 30sec. and 72°C for 1min. Probe was purified and used for the plaque hybridization. After overnight hybridization at 60°C the nylon membranes were washed in 500ml of 2X SSC / 0.5% SDS at room temperature for 5min., in 500ml of 2X SSC / 0.1% SDS at 60°C for 30min., in 500ml of 1X SSC / 0.1% SDS at 60°C for 30min. and 0.1X SSC / 0.1% SDS at 60°C for 30min. The secondary screening was carried out in the same condition.

$\lambda$ -phage DNA was extracted from the screened viral clones with QIAGEN Lambda System (QIAGEN, Valencia, CA). And the DNA was sequenced by primer walking. The sequencing reaction was performed with BigDye Terminator version 3.1 (Applied Biosystems, San Francisco, CA). The sequence data were analyzed and processed with Sequencher version 3.0 (Gene Code Corporation).



## Screening *Acropora millepora* genomic library

$\lambda$ -GEM 12 *Acropora millepora* genomic DNA library was constructed by Dr. David Hayward (Australian National University). For screening for *AmTAT1* gene, approximately 500,000 clones were screened by the standard plaque hybridization in the primary screening. Hybond-N nylon membranes were used for plaque lift and hybridization. The insert, 5B fragment, in the rescued plasmid from yeast one-hybrid screen was excised by ECoRI digestion and purified. This fragment was radiolabelled with  $\alpha$ -<sup>32</sup>P-dATP by random priming method with Megaprime DNA labelling systems (Amersham Biosciences, Piscataway, NJ). The labelled probe was purified with Ultra Clean PCR Clean-up / DNA Purification Kit. The hybridization was performed at 65°C overnight in 5X SSC, 5X Denhardt's solution, 0.5% SDS. The hybridized membranes were washed according to the manufacturer's instruction: in 500ml of 2X SSC / 0.5% SDS at room temperature for 5min., in 500ml of 2X SSC / 0.1% SDS at 65°C for 30min., in 500ml of 1X SSC / 0.1% SDS at 65°C for 30min. and 0.1X SSC / 0.1% SDS at 65°C for 30min. The membranes were air-dried and the radioactive signals were exposed on Phosphoscreen and visualized with Phosphorimager (Molecular Dynamics, Sunnyvale, CA). The areas on the agar plates with signals were picked up and stored in SM buffer at 4°C. The secondary screening was carried out with probe radiolabelled by PCR. The primer pair shown below was designed to amplify a 460bp fragment corresponding to a region within 5B fragment. The rescued plasmid was used as the template. Reactions were thermal-cycled 30 times at 95°C for 45sec., 50°C for 30sec. and 72°C for 1min. followed by a 2min. extension at 72°C.

5BF1 : 5'-AAGAGGAAAGCGTGTAAGCC-3'

5BR1 : 5'-AGGTCTTCTTGACTGGTCAC-3'

For cloning *vitellogenin* of *Acropora millepora* (*AmVit*), approximately 200,000 clones were screened by the standard plaque hybridization in the primary screening. With the plasmid clone for *AmVit* as the template, the probe was amplified and radiolabelled by PCR. The primer pair was designed to amplified a 318bp region which was specific to the *AmVit* cDNA. The sequences of the primers were:

avDF3 : 5'-GATCTCCGGTCCGCTTGATCAGC-3'

avDR3 : 5'-TGATCCTTCGATTGCGTCAGC-3'

The protocol of the thermal cycling was: 30 times at 95°C for 45sec., 50°C for 30sec. and 72°C for 1min. followed by a 2min. extension at 72°C.

### **Sequencing lambda phage DNA**

The lambda phage DNA of the isolated clone was extracted with QIAGEN Lambda System (QIAGEN, Valencia, CA) according to the supplier's manual. The lambda phage DNA was sequenced by primer walking. The sequencing reaction was performed with BigDye Terminator version 3.1 (Applied Biosystems, San Francisco, CA) or ET-Terminator (Amersham Bioscience, Piscataway, NJ). The sequence data were analyzed and processed with Sequencher version 3.0 or 4.2 (Gene Code Corporation). Sequence homology search for the possible translation was carried out using the blast-t program in the BLAST server at the National Center for Biotechnology Information.

### **Southern-blot hybridization**

10µg of high molecular weight genomic DNA of *Acropora millepora* was digested by ECoRI, BamHI and KpnI. The samples were fractionated by electrophoresis in 0.7% agarose gel at 10V for 21 hours. Southern-transfer was carried out with the standard capillary transfer method and the DNA was immobilized on Hybond-N membrane by UV irradiation. The region corresponding to the exon3 of *AmTAT1* was chosen for the probe and radiolabelled by PCR with the primer pair:

tkzUp : 5'-TTGTCCTTGTTAGCCTGGGTT-3'

tkzDwn : 5'-GCAATTCGGCATGTATCTTCC-3'

Genomic DNA of *Acropora millepora* was used for the template. The reaction was performed with 30 cycles of denaturation at 94°C for 45sec., annealing at 50°C for 45sec. and extension at 72°C for 1min. followed by an additional extension at 72°C for 3min. The radioactive probe was purified with the spin column and hybridized to the immobilized DNA at 65°C overnight. The wash condition was: in 250ml of 2X SSC / 0.5% SDS briefly at room temperature, in 250ml of 2X SSC / 0.5% SDS at room temperature for 5min., in 500ml of 2X SSC / 0.1% SDS at 60°C for 30min. and in 500ml of 1X SSC / 0.1% SDS at 60°C for 30min. The membrane was exposed on Phosphoscreen for 4 hours and the signals were detected by Phosphoimager.

### **RNA extraction from coral eggs and embryos**

Total RNA was extracted by the standard guanidinium thiocyanate / ultracentrifuge method. The sample stocked in liquid nitrogen was thawed at room temperature and 5 volumes of 4M guanidinium thiocyanate / 0.1M Tris-HCl (pH7.5) / 0.5% N-lauroylsarcosine /

0.1M  $\beta$ -mercaptoethanol was added to the sample. The suspension was passed through a 18G needle 10 times and laid on 3ml of 5.7M cesium chloride / 10mM EDTA layer in a polyalomer ultracentrifuge tube. The ultracentrifuge was carried out in SW41Ti rotor (BECKMAN COULTER, Palo Alto, CA) at 32krpm for 20 hours at 20°C. The upper layer was carefully removed and the pellet was redissolved in 300 $\mu$ l of 10mM Tris-HCl (pH7.4) / 5mM EDTA / 1% SDS. This solution was extracted once with chloroform / 1-butanol. RNA was ethanol-precipitated and stored at -80°C. For reconstitution, it was recovered by centrifugation and redissolve in an appropriate volume of RNase-free water.

### **RNA extraction from adult coral tips**

From adult coral tips, RNeasy (Ambion, Austin, TX) was used for RNA extraction by the modified method from that in manufacturer's manual. Approximately 2.5g of the proximal portion of the coral tip was pulverized in liquid nitrogen using an iron pestle and a hammer. The powder was suspended in 25ml of RNeasy and mixed well by grinding in a mortar. The mixture was incubated at room temperature for 10min. for dissociation of nucleoprotein from nucleic acid. The debris of calcium exoskeleton was removed by centrifugation at 4krpm for 10min. in a swinging bucket rotor and the supernatant was recovered in a 50ml tube. The supernatant was mixed with 5ml of chloroform by shaking vigorously for 20sec. to form complete emulsion. After incubation at room temperature for 10min., the emulsion was centrifuged at 9250rpm for 10min. The aqueous phase was carefully recovered. Chloroform extraction was repeated again with 4ml of chloroform. An equal volume of RNase-free water was

added to the recovered aqueous phase (approximately 15ml) and mixed well by swirling. 15ml of isopropanol was added and incubated at room temperature for 10min. The precipitation was spinned down by centrifugation at 9250rpm for 10min., rinsed with 5ml of cold 75% ethanol, air-dried and redissolved in 100µl of RNase-free water. The RNA was stored at -80°C.

### RT-PCR for *AmTAT1* and *AmVit*

mRNA was reverse-transcribed for the 1st strand cDNA synthesis by using First-strand cDNA Synthesis Kit (Amersham Biosciences, Piscataway, NJ). 30µg of total RNA in 40µl of RNase-free water was incubated at 65°C for 10min. to dissociate the RNA-RNA annealing and stored on ice until needed. 25µl of Bulk reaction mix was mixed with 5µl of 200mM dithiothreitol and 5µl of 0.2µg/µl NotI-d(T)<sub>18</sub> primer. This mixture was combined with the heat-denatured total RNA and mixed well. Then the reaction was incubated at 37°C for 1hr for reverse-transcription. The enzyme was heat-inactivated by incubation at 90°C for 5min.

15µl of the 1st strand solution (equivalent to 5µg of total RNA at start) was used for the subsequent PCR amplification with 25pmoles of each primer.

For PCR amplification of *AmTAT1*, the primer pair was:

tkzf1.7-9R1 : 5'-CATGCTCCCTTAGAAGTTGG-3'

5BR6 : 5'-TCGTACGTACCTTCAATTGTC-3'

These sequences are located in exon 3 and exon 4, respectively, so that the product can be distinguishable from those amplified with the

genomic DNA contamination as the template. The amplification protocol consisted of 30 cycles of 95°C for 1min., 50°C for 30sec. and 72°C for 45sec. followed by an additional extension step at 72°C for 2min.

For AmVit, the primers were:

avDF3 : 5'-GATCTCCGGTCCGCTTGATCAGC-3'

avDR3 : 5'-TGATCCTTCGTAAGCGTCAGC-3'

These sequences are located in exon 2 and exon 4, respectively. The amplification protocol consisted of 35 cycles of 95°C for 1min., 50°C for 30sec. and 72°C for 1.5min. followed by an additional extension step at 72°C for 3min.

As the control for the fidelity and efficiency of the amplification, partial cDNA of coral *EF1 $\alpha$*  (elongation factor 1 $\alpha$ ) was amplified by PCR. The primer pairs were:

EF(FW) : 5'-TGCCTGGGTCTTGACAAACT-3'

EF(RV) : 5'-GGGATGGTATCCCACTTTCTG-3'

These primers were designed on basis of the sequence of *EF1 $\alpha$*  cDNA cloned previously in our laboratory (unpublished data). The amplification protocol consisted of 30 cycles of 95°C for 1min., 50°C for 30sec. and 72°C for 2min. followed by an additional extension step at 72°C for 4min.

To exclude the possibility of contamination, control PCR was also performed for each reaction without template.

20 $\mu$ l of the 50 $\mu$ l reaction was analyzed by agarose gel electrophoresis. And the products were visualized by UV transillumination and

photographed.

### Yeast one-hybrid screen

Large-scale yeast one-hybrid screening was carried out according to the method by Jingdong Liu (Liu et al., 1993) with minor modification.

The yeast one-hybrid screen is a detection system that was developed to identify DNA-binding sequences and DNA-binding domains. This system makes use of a yeast strain, YM4271, of which growth requires adenine, uracil, tryptophan, lysine and leucine. Two plasmids are introduced to this yeast step by step. First, the yeast is transformed with Activator plasmid, pBM2463, which contains two marker genes, *ADE5* and *URA3*, and encodes a fusion protein of the DNA-binding domain of interest, *lexA*, and *GAL4* activation domain. This transformant can grow on -Uracil plate and forms red colonies under the influence of *ADE5* marker. In the next step, Reporter plasmid, pBM2389, is introduced to the yeast. Reporter plasmid has exogenous genomic fragments inserted 5' to *GAL1* promoter and *HIS3* as a reporter gene ; the *GAL1* promoter lacks the Upstream Activating Sequence (UAS). Transcriptional activation of *HIS3* gene can occur only when the genomic sequence harbors the binding sequence for the DNA-binding domain which is derived from the Activator plasmid. Thus the yeast clone grows on -Histidine plate. The problem is that *HIS3* gene can be activated by endogenous transcription factors in yeast. However, these clones are excluded due to *ADE5*, sectoring marker. Clones with endogenous activation of *HIS3* gene do not need the Activation plasmid in the presence of uracil because

*HIS3* transcription is independent of the fusion protein. Consequently, the pseudo-positive clones start to grow without the Activator plasmid resulting loss of red color. The colonies appear to be sectored with red and white regions. Finally, true positivity of the double transformants which are negative for sectoring is confirmed by 5-fluoroorotic (5-FOA) acid test. This substance is toxic to cells carrying *URA3* gene. On the culture plate containig 5-FOA, the pseudo-positive clones survive because they can eliminate the Activator plasmid and still be viable. On the other hand, the true positive clones eventually die on the plate. Reporter plasmids from the clones which pass all these screens are isolated and used for further study.

In my case, one-hybrid system was made use of to isolate responsible elements for DM domain of AmDM1. A large-scale screening on the reporter library of *Acropora millepora* genomic DNA was carried out. Figure 3 explains the summary of the selection strategy.

### **Yeast strain and culture**

The genotype of *Saccharomyces cerevisiae*, YM4271 is *MATa ura3-52 his3-Δ200 ade2-101 lys2-801 trp1-901 gal4-Δ512 gal80-Δ538 ade5::hisG*. Culture media were prepared according to the standard protocol (Ausubel et al, 1995). YM4271 was cultured in YPD medium or selection medium shown below. For preparation of the culture agar plates, agar was added to make 2% gel.

-Uracil medium

0.17% yeast nitrogen base w/o ammonium sulfate



0.5% ammonium sulfate  
2% glucose  
10µg/ml L-histidine  
10µg/ml adenine  
20µg/ml L-tryptophan  
40µg/ml L-lysine  
60µg/ml L-leucine

**-Histidine medium**

0.17% yeast nitrogen base w/o ammonium sulfate  
0.5% ammonium sulfate  
2% glucose  
40µg/ml uracil  
10µg/ml adenine  
20µg/ml L-tryptophan  
40µg/ml L-lysine  
60µg/ml L-leucine

**-Tryptophan medium**

0.17% yeast nitrogen base w/o ammonium sulfate  
0.5% ammonium sulfate  
2% glucose  
10µg/ml L-histidine  
40µg/ml uracil  
10µg/ml adenine  
40µg/ml L-lysine  
60µg/ml L-leucine

**Construction of Activator plasmid containing DM domain**

pBM2463 was linearized by double digestion with NotI and XhoI followed by dephosphorylation with Shrimp alkaline phosphatase. The DM domain of AmDM1 was amplified by PCR with *Acropora millepora* genomic DNA as the template. The primer pair used for this amplification was:

DM-FW : 5'-CGCGGCCGCTCTCCAAAATGCGCCCGAT-3'

DM-RV : 5'-GCTCGAGGGCCTCCGGCTTCTTCGCTTT-3'

The underlined sequences were added to provide NotI and XhoI sites, respectively, for the subsequent ligation. After agarose gel electrophoresis, the PCR product was extracted from the gel and inserted into pGEM-T Easy vector (Promega, Madison, WI) . The insert was excised from pGEM-T Easy by cutting with NotI and XhoI. Ligation reaction was performed with the linearized pBM2463 and the excised fragment which contained cDNA sequence for the entire DM domain. The sequence of the insert and the in-frame ligation were checked by sequencing. The construct was named pDMact5 (Figure 3).

### **Construction of Reporter library**

The Reporter library was established by Dr. Nikki Hislop and described elsewhere (Hislop, 2003). Briefly, 5µg of *Acropora millepora* genomic DNA was partially digested with Sau3AI so that about a half of the fragments were less than 1kb in length. The digested DNA was fractionated by agarose gel electrophoresis and fragments of 0.5-1.5kb were extracted from the gel. Then, they were inserted into Reporter plasmid, pBM2389, at BamHI sites and introduced to *E. coli*, DH5α by electroporation. On the basis of

theoretical calculation, this library has up to 48% genome coverage. They were stored at -80°C as glycerol stock and plasmids were recovered by the standard culture and plasmid preparation methods.

### **Yeast transformation with plasmid vectors**

The standard lithium acetate/PEG method was used for yeast transformation. A single colony of freshly grown yeast on a YPD or an appropriate selection plate was inoculated in 5ml of YPD medium and cultured overnight at 30°C with vigorous agitation. The optical density at 600nm was measured with spectrophotometer and an aliquot of the culture was added to 100ml of YPD to make the final density 0.1 at 600nm. Culture was started again and continued until the density reached 0.8-1.0 (about 7 hours). Cells were harvested by centrifugation at 2krpm for 5min. And the pellet of yeast was resuspended in 1.1ml of TEL buffer (10mM Tris-HCl pH8.0, 1mM EDTA, 100mM Lithium acetate pH7.5). After incubation at 30°C for 20min. with gentle agitation, transformation mixture shown below was assembled in a 2ml tube.

Reporter library solution	6µl (less than 1µg DNA)
salmon sperm DNA (10mg/ml)	10µl
yeast suspension	200µl

The salmon sperm DNA was heat-danatured at 95°C for 20min. and snap-chilled on ice just before assembly. The mixture was mixed well and incubated at 30°C for 20min. with gentle agitation. Then 1.4ml of PLATE solution (40% PEG 4000, 0.1M Lithium acetate pH7.5, 10mM Tris-HCl ph8.0, 1mM EDTA) was added to the mixture and again

mixed well immediately. After incubation at 30°C for 20min., it was heat-shocked by incubation at 42°C for 5min. The yeast cells were recovered by spin down

at 13.2krpm for 5 seconds and resuspended in 300µl of sterilized double distilled H<sub>2</sub>O.

100µl of the suspension was spread on an appropriate selection agar plate. The plates were incubated at 30°C for 5-6 days until the colonies appeared.

### **Selection of yeast clones**

Yeast strain YM4271 was first transformed with the Activator plasmid pDMact5. This transformant, YM4271/pDMact5, forms red colonies on the -Uracil plate as described above. The clone YM4271/pDMact5 was used as a host for a large-scale one-hybrid screening in the next step.

Approximately, 99,000 clones were screened in total. 109 clones showed red colonies without sectoring. These colonies were spread again on -Histidine plates to exclude the pseudo-positive ones. This step was necessary because some of the colonies that seemed negative for sectoring were so small that it was impossible to recognize them 'non-sectoring' on the primary plates. After the secondary plating test, 53 out of 109 clones remained negative for sectoring. Thus, these clones were brought to the final selection with 5-Fluoroorotic acid.

5-FOA was purchased from Sigma (Sigma-Aldrich, St. Louis, MO). For this selection, culture plates were prepared by adding 5-FOA to the -

Histidine medium to make the concentration 1mg/ml. The 53 clones were screened by -Histidine + 5-FOA culture at 30°C for up to 9 days. Finally, 5 clones passed this screening.

### **Recovery of plasmid from yeast transformants**

The standard glass beads method was used to rescue the reporter plasmids. Prior to the rescue process, the screened clones were spread and cultured on -Tryptophan plate. This step makes it possible to eliminate Activator plasmid, pDMact5, from the yeast clones because they no longer need the activator for proliferation in presence of histidine while they still need Reporter plasmid for the source of tryptophan. As a result, the colonies grown on -Tryptophan plate were white.

A single colony of each clone was inoculated in 3ml of -Tryptophan medium and culture overnight at 30°C with vigorous agitation. 1.5ml of the culture were harvested in an Eppendorf tube by spinning down at 13.2krpm for 5sec. After removal of the supernatant, the pellet was resuspended in the residual medium by vortexing. 0.3g of acid-washed glass beads and 200µl of Lysis buffer (2% Triton X-100, 1% SDS, 0.1M NaCl, 10mM Tris-HCl pH8.0, 1mM EDTA) were added to the suspension and mixed. 200µl of phenol : chloroform (1:1) was added and mixed well by vortexing for 2min. After centrifugation at 13.2krpm for 5min., the aqueous phase was recovered carefully. Competent *E.coli* cells were transformed with approximately 5µl of the recovered solution with the standard transformation technique. The transformant had the Reporter plasmid rescued from the yeast clone. Then, Reporter plasmids were extracted and the genomic fragments in

the plasmids were sequenced.

### **Prokaryotic expression vector construction**

pGEX-6P-2 vector was purchased from Promega (Madison, WI). The partial cDNA sequence of *AmDM1* which corresponded to a 64 amino acid region including the entire DM domain was amplified by PCR with the primer pair shown below.

DMXFF : 5'-GGATCCCGATCTCCAAAATGCGCC-3'

DM-RV : 5'-GCTCGAGGGCCTCCGGCTTCTTCGCTTT-3'

The underlined nucleotides were arbitrarily attached to the cDNA sequences to provide the recognition sites for BamHI and XhoI, respectively. The PCR product was purified and ligated into pGEM-T Easy (Promega, Madison, WI). The insert was excised by digestion with BamHI and XhoI. And inserted into pGEX-6P-2 linearized with the identical combination of restriction enzymes. The expression construct (pGEX-DMZ+1.8) coded a fusion protein of glutathione-S-transferase and DM domain of *AmDM1* at its C-terminal. The nucleotide sequence of the insert and the cloning sites of the vector was checked out by sequencing.

### **Large-scale expression and purification of the GST fusion protein**

An *E. coli* strain BL21 was transformed with pGEX-DMZ+1.8 plasmid. Prior to the large-scale preparation of the fusion protein, the condition for the bacterial culture and incubation with IPTG was optimized and the fidelity of the expression construct was confirmed

by performing a small-scale experiment.

A single colony of BL21/pGEX-DMZ+1.8 clone was inoculated in 5ml of LB-broth (ampicillin 100µg/ml) and culture overnight. 2.5ml of the culture was added to 250ml of LB-broth (ampicillin 100µg/ml) and cultured until the optical density at 600nm reached 0.5-1.0. IPTG solution was added to make the final concentration 1mM. After 3 hours of incubation, the bacteria were harvested by centrifugation (4krpm for 15min. at 4°C). The pellet was suspended in 10ml of cold PBS(without calcium and magnesium) and lysed by sonication. 200µl of 25% TritonX-100 was added to the lysate and mixed gently for 30min. at room temperature. Cell debris was pelleted to be removed by 2 step centrifugation: at 4krpm. for 15min. at 4°C and at 10,000 x g for 5min. at 4°C. The supernatant was subjected to the purification with glutathione-agarose beads.

To the recovered supernatant, 0.5ml of equilibrated 50% Glutathione Sepharose 4B suspension (Pharmacia Biotech, Piscataway, NJ) was added. Purification process was carried out according to the supplier's manual. Briefly, the mixture was incubated at room temperature for 30min. with gently agitation, the beads were recovered by discarding the supernatant after centrifugation at 500 x g for 5min. and washed 4 times with 5ml of ice-cold PBS(without calcium and magnesium). Finally, the solution of the purified fusion protein was obtained by incubation with 10mM reduced glutathione in 50mM Tris-HCl (ph8.0) at room temperature for 30min. and subsequent recovery by centrifugation. Using Vector NTI Suite software (InforMax, Bethesda, MD), the value of the protein concentration which corresponded to 1.0 of light absorbance at

280nm was theoretically predicted. The concentration of the protein solution was determined by measuring the ultraviolet absorbance at 280nm. The purity of the fusion protein sample was checked by standard SDS polyacrylamide gel electrophoresis (SDS-PAGE).

### Preparation of radioactive probes for EMSA by PCR

To detect the responsible sequences for the binding with the DM domain of AmDM1, radioactive probes were synthesized by PCR. The amplified fragments corresponded to the separate or overlapping regions in the 5B fragment that was isolated by yeast one-hybrid screen. Names of the probes, the primer pairs used for the amplification of each probe are shown in Table 1. And the sequences of the primers are shown below. The genomic DNA of *Acropora millepora* or pDMBF-5B plasmid which was isolated in the one-hybrid screen and identified to harbor a DNA fragment containing a DM-binding region were used as the template. Reactions were thermal-cycled 30 times at 95°C for 45sec., 50°C for 30sec. and 72°C for 30sec.

5BF1 : 5'-AAGAGGAAAGCGTGTAAGCC-3'

5BF2 : 5'-TCCACAGACACTGCCATATG-3'

5BR1 : 5'-AGGTCTTCTTGACTGGTCAC-3'

5BR2 : 5'-CATATGGCAGTGTCTGTGGA-3'

5BR5 : 5'-CAAAGATGATGGCTACTTTGTA-3'

5BF6 : 5'-TACAAAGTAGCCATCATCTTTG-3'

5BF5 : 5'-GAAAAGCGTGGTTGAGAGAC-3'

5B $\alpha$ A0 : 5'-GTCAAGATGATGGCTACTTTGTATTTTTT-3'



5B $\alpha$ A1 : 5'-GTCATTTTTTTCATTACGTTGAAGC-3'  
5B $\alpha$ A2 : 5'-GTCAGTTGAAGCGACACATTTTGATACAA-3'  
5B $\alpha$ A3 : 5'-GTCAGATACAAATTTTCAATTTGAGTATC-3'

After the amplification by PCR, the radiolabelled products were purified by removing unincorporated radioactive or non-radioactive dNTP with Ultra Clean PCR Clean-up / DNA Purification Kit (MO BIO Laboratories, Solena Beach, CA).

### **Preparation of radioactive probes for EMSA by Klenow reaction**

Double-strand oligonucleotide probes were labelled by end-filling reaction catalyzed by Klenow fragment. Prior to radiolabelling, complementary oligonucleotides were annealed to provide a 5'-protruding stretch as the substrate for end-labelling. One of those oligonucleotides was designed to have an additional 4 nucleotide sequence (-GTCA-). The sequence of the pairs of complementary oligonucleotides were shown below. The additional protruding stretches were underlined.

5B $\alpha$ A2-u : 5'-GTCAGTTGAAGCGACACATTTTGATACAA-3'  
5B $\alpha$ A2-d : 5'-TTGTATCAAAATGTGTCGCTTCAAC-3'  
  
repetII-u : 5'-GTCAGTTTTTTTGCAAAGTACCGTGCA-3'  
repetII-d : 5'-TGCACGGTACTTTGCAAAAAAAC-3'  
  
mutA2-u : 5'-GATCGTTGAAGCGACGAGCCTTGATACAA

mutA2-d : 5'-TTGTATCAAGGCTCGTGCGTTCAAC

For annealing reaction, 10 $\mu$ l of both oligonucleotides (5 $\mu$ M) were mixed and incubated at 85°C for 5min. followed by slow cooling down to room temperature. The annealed double-strand oligonucleotides were radiolabelled with  $\alpha$ -<sup>32</sup>P-dATP catalyzed by Klenow fragment (Promega, Madison, WI) according to the manufacturer's instruction. Before the binding reaction with the fusion protein, the radioactive probes were purified by using QIA quick Nucleotide Removal kit (QIAGEN, Valencia, CA).

#### ***in vitro* binding reaction of DNA and fusion protein**

10fmole of the purified radioactive probe was mixed with 0 to 400mg of GST-DM domain fusion protein. The binding reaction was performed at room temperature for 1 hour in 16mM HEPES-NaOH(ph7.9), 64mM KCl, 4mM MgCl<sub>2</sub>, 6% glycerol, 0.8mM Dithiothreitol, 0.04% Nonidet P-40, 0.1mM ZnCl<sub>2</sub> and 50ng poly(dI)(dC). The volume of this reaction was adjusted to 20 $\mu$ l with double distilled H<sub>2</sub>O. Glutathione-S-transferase synthesized with the pGEX-6P-2 system was used for the negative control for binding.

#### **Non-denaturing polyacrylamide gel electrophoresis**

After incubation, 3 $\mu$ l of 6X xylene cyanol dye (2.5mg/ml xylene cyanol/400mg dextrose) was added to the binding reaction. Then samples were applied to 8% acrylamide mix (acrylamide:N,N'-methylene-bis-acrylamide=29:1)/0.5X TBE gel and fractionated by electrophoresis in 0.5X TBE (Tris-borate-EDTA) at 25mV for

appropriate length of time : 2.5-3.5 hours depending on the size of the probes. The electrophoresis apparatus was kept at low temperature by circulating ice-water. The gel was blotted on Whatmann 3MM blotting paper and dried in a vacuum gel drier. The radioactive signals were exposed on Phosphoscreen and visualized with Phosphorimager (Molecular Dynamics, Sunnyvale, CA).

### **Cell line and culture condition**

Human cancer cell line HEP-2 was purchased from American Type Culture Collection (ATCC, Rockville, MD). This cell line was maintained and cultured in MEM $\alpha$  (#12561 Gibco/Invitrogen, SanDiego, CA) supplemented with 10% Fetal Calf Serum (Gibco/Invitrogen, SanDiego, CA) and 100 $\mu$ g/ml Penicillin G and Streptomycin. Cells were incubated in a 5% CO<sub>2</sub>-incubater at 37°C with humidification.

### **CAT reporter vector construction**

pCAT3-Promoter vector was purchased from Promega (Promega, Madison, WI). This plasmid was exclusively designed for the evaluation of the enhancer activity in mammalian cells. A DNA fragment corresponding to a 259bp portion within the intron 2 of *AmTAT1* which included the AmDM1-binding motif was amplified by PCR. The primer pair was:

DME(Forward) : 5'-GTCGACATATGGCAGTGTCTGTGGAA-3'

DME(Reverse) : 5'-GTCGACTATCGAATGACCCCGC-3'

At the 5'-end of each oligonucleotide, a site for XhoI recognition was

created by adding an appropriate nucleotide stretch (underlined). The genomic DNA of *Acropora millepora* was used for the template. After purification, the fragment (=DM enhancer) was inserted into pGEM-T Easy (Promega, Madison, WI) by ligation and excised by cutting with XhoI. pCAT3-Promoter plasmid was linearized by XhoI digestion, dephosphorylated with Shrimp alkaline phosphatase and inserted with the DM enhancer. Two distinct clones were obtained with respect to the orientation of the DM enhancer: pCATDMEn6 (forward orientation) and pCATDMEn1 (Reverse orientation)(Figure 4).

### ***AmDM1* expression vector construction**

I made use of the eukaryotic expression vector pRc/CMV (Invitrogen, SanDiago, CA). pRc/CMV was double-digested with XbaI and ApaI. A partial cDNA fragment of *AmDM1*, which included the entire open reading frame was excised from pAmdsx8-1 plasmid (Miller et al., 2003) by cutting with XbaI and ApaI. The fragment was inserted into the linearized pRc/CMV vector (=pCMVdsx).

To analyze the functional structure of AmDM1 protein, expression vectors for the two species of nested deletion mutants were constructed.  $\Delta$ ACT fragment which lacked the C-terminal portion to the DMA domain of AmDM1 was amplified by PCR with the primer pair:

T3 (in pBluescript)

del.ACT : 5'-GGGCCCTTACTCGTGACTTGGAAGAAC-3'

$\Delta$ DMA fragment which only have the N-terminal and DM domain was amplified with :

T3 (in pBluescript)

del.DMA : 5'-GGGCCCCCACATGACTTTTCCATCGCC-3'

Both downstream primers have ApaI sites at their 5'-ends. The additional nucleotides to the cDNA sequences were underlines. The amplified fragments were ligated pGEM-T Easy vector and excised by XbaI+ApaI double digestion. After purification, they were inserted into the linearized pRc/CMV vector generating p $\Delta$  ACT and p $\Delta$  DMA, respectively (Figure 5).

### **Transfection of plasmids into the mammalian cell line**

The plasmid vectors were transfected into HEP-2 cell by using Lipofectamine 2000 (Invitrogen, San Diego, CA). HEP-2 cells were grown to a 90-95% confluent state in culture dishes. After trypsinization, cells were suspended in the growth medium (MEM $\alpha$ /10% FCS) without antibiotics. The cell density was counted with a hemocytometer and adjusted to approximately  $1 \times 10^6$  cells/ml and 2ml of the suspension was inoculated in each well of a 6-well format. After incubation for 1 day, cells were subjected to the transfection experiment.

For the analysis on the DM enhancer activity, 3 $\mu$ g of pCMVdsx (the expression vector for the full-length AmDM1) or pRc/CMV (empty control) was combined with 1 $\mu$ g of CAT reporter vectors, pCAT3-Promoter (empty control), pCATDMEn6 (forward orientation) or pCATDMEn1 (reverse orientation). For the analysis on the functional structure of AmDM1, 3 $\mu$ g of the expression vectors (pCMVdsx, p $\Delta$  ACT and p $\Delta$  DMA) were transfected.

Transfection was performed with Lipofectamine 2000 according to

the manufacturer's instruction. Briefly, plasmid vectors were mixed in 250µl of MEM $\alpha$  without FCS and antibiotics. In parallel, 10µl of Lipofectamine 2000 was mixed in the same volume of MEM $\alpha$  and incubated at room temperature for 5min. The plasmid solution was mixed in this suspension and incubated at room temperature for 20min. Finally, the mixture was added to the growth medium of the HEP-2 monolayer culture.

### CAT reaction

The transfected cells were cultured for 48-72 hours. The CAT reaction was carried out with CAT Enzyme Assay System With Reporter Lysis Buffer (Promega, Madison, WI). The monolayer cells were washed with PBS(without calcium and magnesium) twice and 200µl of 1X Reporter Lysis Buffer was added. After 15min. of incubation at room temperature with gentle rocking, the cells were harvested by scraping with a rubber policeman. The suspension was kept on ice for several minutes and vortexed for 15sec. Then it was incubated at 60°C for 10min. After centrifugation for 2min. at 13.2krpm, the supernatant (cell lysate) was transfer to a fresh tube. A part of this lysate was subjected to Bradford assay for the quantification of protein.

The reaction mixture for CAT assay was assembled as below.

cell lysate	100µl
[ <sup>14</sup> C] Chloramphenicol (25µCi/ml)	6µl
n-Butyryl CoA (5mg/ml)	5µl
ddH <sub>2</sub> O	14µl

[<sup>14</sup>C] Chloramphenicol was purchased from Amersham Biosciences (Amersham Biosciences, Piscataway, NJ). The reaction mixture was

incubated at 37°C for 3 hours and extracted with mixed xylene. The reaction and 300µl of o-,p-, m-mixed xylene were mixed well by vortexing for 30min followed by centrifugation for 3min. at the top speed. The upper xylene phase was carefully recovered and extracted with 100µl of 0.25M Tris-HCl (pH 8.0). Again, the upper xylene phase was carefully recovered for the liquid scintillation counting.

### **Bradford assay**

Protein concentration in the cell lysate was quantified by using Bio-Rad Protein Assay kit (Bio-Rad, Hercules, CA). 10µl of the 10-fold diluted cell lysate was mixed with 90µl of distilled water and 1ml of 1x protein assay dye prior to a 5min. incubation at room temperature. The absorbance at 595nm was measured with a spectrophotometer.

### **Liquid Scintillation Counting**

The recovered xylene samples were transferred to scintillation vials. 3ml of STARSCINT (Packard Biosciences, Netherland) was added and the cpm was measured with Wallac 1410 Liquid Scintillation Counter with the <sup>14</sup>C window for 3min.

### **Data processing and statistical analyses**

The scintillation counts were standardized by the measurement of Bradford assay and further normalized by the negative control, the values for the empty pCAT3-Promoter vector or the empty expression vector pRc/CMV. The studies were carried out in triplicate. The

Student-t test was applied for the statistical analyses.



## Section 3. Results

### cDNA isolation of *AmDM1* homolog for *Nematostella vectensis*

Sea anemone, *Nematostella vectensis* is one of Cnidarian model experimental animals, which belongs to class Anthozoa. For better understanding the molecular evolution of DM-containing proteins, especially the divergence within the class, the cDNA clone for the homolog of *AmDM1* was isolated from the  $\lambda$ Zip-lox *Nematostella vectensis* cDNA library. With the radioactive probe corresponding to the DM domain of *AmDM1*, approximately 300,000 clones were screened by standard plaque hybridization on the primary screening. 32 separate areas were detected positive and stored for the secondary screening. Finally, 6 single plaques were isolated as clones.  $\lambda$ -phage DNA was extracted from one of these clones, *NvDM5*, and the sequence was determined by primer walk. Both strands of *NvDM5* was sequenced.

The length of the cDNA is 1912 bp without the polyA tail (Figure 6). The 5' untranslated region (5'-UTR) is 134 bp and the 3'-UTR is 410 bp. In comparison with *AmDM1*, the length of 5'-UTR almost the same as that of *AmDM1*. On the other hand, *AmDM1* has a much longer 3'-UTR (1880 bp). The open reading frame of *NvDM5* is 1362 bases encoding a translation of 453 amino acids. An atypical polyadenylation signal (TATAAA) appears 32 base upstream from the polyA tail.

## **NvDM5 protein has both the DM and DMA domain**

The deduced amino acid sequence of NvDM5 was subjected to the further analyses. As shown in Figure 6, NvDM5 protein also has both DM and DMA domains. The sizes of these domains are identical to those of the *Acropora millepora*. Other regions outside DM and DMA domains show almost the same size except the N-terminal region; 47 amino acids for NvDM5 and 71 amino acids for AmDM1. The homology search by blast-p showed that the DM and DMA domain have remarkable homology to those of other organisms. However, for regions outside these domains the search did not find significantly homologous proteins in the data base. This observation is consistent with the other DM family proteins. As shown in Table 2, there are several potential sites for phosphorylation by Protein kinase C and Casein Kinase II. Note that the sites for Protein kinase C form a cluster in the Carboxy-terminal region.

## **DM domain of NvDM5 is highly conserve through evolution**

The amino acid sequences of the DM domains of various organisms were aligned to the DM sequences of AmDM1 and NvDM5. A cDNA clone encoding another DM-containing protein was identified in the *Nematostella vectensis* EST project (Technau, personal communication) and the sequence was aligned as well. The sequences were aligned with ClustalW program and graphically represented by Boxshade (Figure 7). As expected, AmDM1 shows the highest homology with NvDM5 (94% identity). The primary structure of DM domain was highly conserved throughout evolution. In particular, amino acid residues which participate in forming the Zinc finger

(CCHC-HCCC) are absolutely conserved. It is also true for NvDM5 and NvDM-like EST clone. The NvDM5 is more similar to AmDM1 than NvDM-like. Moreover, DM domains of other phyla have higher homology to AmDM1 than NvDM-like. According to the phylogenetic analysis (Figure 8), AmDM1 and NvDM5 are the closest sisters making a cluster together with DMRT1 (mouse). On the other hand, NvDM-like is located distant from the other Cnidarian proteins and close to DM proteins of *C. elegans*.

DMA domains were recently identified in DM family proteins. While all the DMA-containing proteins are members of DM family, not all the DM family contain DMA domains. In comparison with the sequence homology between the DM domains of AmDM1 and NvDM5, the similarity for their DMA domains is not high (85% identity). The alignment of DMA domains is shown in Figure 9. It is to be noted that the phenylalanine at the 9th position and the glutamate at the 33rd are conserved in all the DMA domains. NvDM-like also contain DMA domain as well as DM domain. The primary structure of DMA domain of NvDM-like shows less similarity to AmDM1 than that of NvDM5. This observation is consistent with the finding for the Cnidarian DM domains. The dendrogram for DMA domains supports this idea (Figure 10). NvDM-like was located distant from the clade of AmDM1 and NvDM5.

Three Cnidarian DM-containing proteins are aligned in Figure 11. It is obvious that the sequence conservation is particularly high within and around the regions for DM and DMA domains of AmDM1 and NvDM5, and the poor conservation is observed outside those regions. However, in the C-terminal portion, there are small 'islands' of

conserved elements.

## 5 clones obtained by yeast one-hybrid screen

To isolate the genomic fragment of *Acropora millepora*, which contained the consensus motif recognized by DM domain, large-scale screening was carried out with yeast one-hybrid selection system. Figure 3 shows the schematic representation for the strategy of one-hybrid screen. This system had primarily been developed to identify the protein-DNA interaction in eukaryotic cells and to isolate the downstream genes of eukaryotic transcription factors. The cDNA fragment corresponding to the entire DM domain of AmDM1, the Doublesex homolog of *Acropora millepora* was inserted properly in-frame in pBM2463 to make the Activator plasmid, pDMact5. This recombinant produces lexA-DM-GAL4 fusion domain which has the DNA-binding domain and the transcription activation domain. Yeast strain YM4271 was transformed with this construct to YM4271/pDMact5, the recipient of the Reporter library.

The primary screening was performed by transforming YM4271/pDMact5 with the library and selecting clones by growth on agar plates of the synthetic dropout medium lacking histidine (SD/-Histidine plates). All red colonies without sectoring were picked up and spread on large areas of SD/-Histidine plates to further confirm that the clones were negative for sectoring. This subsequent step of plating test was necessary because some of the colonies looked red but were too small to identify the sectoring. In total, 98,729 colonies were screened among which 109 clones seemed red without sectoring. The second plating test excluded 56 clones. Thus 53 out of

98,729 clones were suspected positive for the protein-DNA association.

Those clones were subjected to the secondary screening by testing the growth on medium with 5-fluoroorotic (5-FOA) acid. This substance is toxic to the organism with *URA3* gene which is on the Activator plasmid. When the Activator plasmid is dispensable for the growth of a yeast clone, the yeast can abandon the plasmid and grow on - Histidine plate with 5-fluoroorotic acid. It means *HIS3* gene on the Reporter plasmid is activated not by the *lexA-DM-GAL4* fusion protein but the endogenous transcription factors of the yeast. Table 3 shows the summary of 5-FOA test. After all, 5 clones, 4U, 5B, 6A, 6E and 6M remained suspected of being positive for the interaction between DM domain and the genomic DNA fragments.

### **Isolation of the Reporter plasmid from the screened clones**

For further investigation, I tried to recover the Reporter plasmids which were supposed to harbor the genomic DNA fragments containing potential DM domain-binding sequences.

To simplify the isolation protocol, pDMact5, the Activator plasmids should be eliminated from the yeast clones. As shown in Figure 3, the Activator plasmid has *URA3* while the Reporter plasmid has *TRP1*. On agar plates of the synthetic dropout medium lacking tryptophan (SD/-TRP), it is possible to select the yeasts only with the Reporter plasmids because in the presence of histidine and uracil, the *HIS3* gene on the Reporter plasmid doesn't need to be trans-activated for the expression by *lexA-DM-GAL4* fusion protein encoded by the pDMact5

and the *URA3* gene on the same plasmid is no longer necessary for growth. Thus the Activators are eventually lost during colony formation. The loss of pDMact5 could easily be recognized on the plate because the colonies were white instead of being red. By this selection, I obtained white colonies for clone 4U, 5B, 6A and 6E. The 6M clone grew on the selection plate but still formed red colonies. By applying the standard glass beads method, the recombinant pBM2463 were successfully rescued from the yeast clones except clone 6M.

Repeated attempts to rescue the plasmid from clone 6M were unsuccessful. I tried to amplify the inserted sequence by PCR with the external primers but no product was observed. This clone had quite different appearance from those of the others. The colonies were fairly pink rather than red from the beginning of the formation. Aforementioned, the color did not disappear after the selection by SD/-TRP culture. Moreover, its growth was by far the fastest. Taking these observation into consideration, I concluded that an aberrant genetic event happened on the transformation with the Reporter library and the clone was selected due to its color and auxotrophy. One of possible explanations of the mechanism is that the recipient yeast was made competent for genome integration during the transformation and both Activator and Reporter plasmids were integrated into the downstream of strong promoters in the genome. Anyway, I didn't further investigate what happened to clone 6M.

For confirmation of the interaction between DM domain and the DNA fragments rescued from clone 4U, 5B, 6A and 6E, one-hybrid assay was performed with the isolated Reporter constructs. The plasmids were introduced into both YM4271 and YM4271/pDMact5. All

transformants of YM4271 and YM4271/pDMact5 grew on SD/-TRP plates and SD/-TRP-URA plates, respectively. This observation showed that the Reporters actually provided the binding regions for DM domain of AmDM1.

### **Nucleotide sequences of the recovered genomic fragments**

The *Acropora millepora* genomic fragments inserted in the Reporter plasmid pBM2389 were sequenced by using the external primers and internal primer if necessary. Figure 12 shows the partial sequences of 4U, 6A and 6E and the complete sequence of 5B. For 4U, 6A and 6E, nucleotide sequences were readable only partially probably because they were ‘difficult sequences’. The sequences and the characterization with restriction enzymes of 4U and 6A fragments showed that these inserts were identical. This result indicated the fidelity of the one-hybrid system since the two clones were obtained from independent transformation.

The sequence data were subjected to homology search with the blast-n program in the BLAST server at the National Center for Biotechnology Information. The result for 4U (=6A) and 6E showed that they had regions with high homology to the repeat sequences reported from our laboratory (McMillan and Miller 1989 ; McMillan and Miller 1990). Outside the repeat sequences, no significant homology was detected. Nucleotide sequences of the repeats were reported for *Acropora formosa* and *Acropora latistella*. And homologous repeats were detected 12 other species of genus *Acropora* (McMillan and Miller 1990). The unit of repeat is 118bp in length and they cover 5% of *Acropora* genome. The repeat sequences

in the insert 4U (=6A) and 6E are 85-93% identical to those of *Acropora formosa* and 95-100% to *Acropora latistella*, respectively. In 4U (=6A) fragment, there are a tandem repeat of two repeat units (4U#1 and 4U#2). They are slightly different from each other. On the other hand, 6E contains only one repeat unit, of which 3'-half is connected to the 5' end of the 5'-half in the same orientation. Figure X shows the nucleotide sequences of 4U#1, 4U#2 and 6E aligned with those of *Acropora latistella* and *Acropora formosa*. They are highly conserved and form three well-defined clusters in accordance with taxa.

In contrast to the repeat sequences, the genomic DNA fragment in clone 5B did not show any homology. Therefore I attempted to isolate the gene which contained the 5B sequence from the genomic library of *Acropora millepora*.

### **Isolation of $\lambda$ -phage clone containing the 5B region**

For the primary screening, I performed plaque hybridization on 500,000 clones. The entire insert in the Reporter plasmid of clone 5B was radiolabelled by the random priming method and used as the probe. 3 signals were detected and the corresponding areas on the plaque-forming agar plates were sampled. The secondary screening was carried out for these samples and one clone, 5B1b was finally isolated as a single plaque. This clone was subjected to further analyses.

**5B region was located within an intron of a homolog of the transmembrane transporter gene**



$\lambda$ -phage DNA was extracted from the clone 5B1b and sequenced by primer walk. The primer walk started from within the 5B region with internal primers in both direction. When the nucleotide sequence of a substantial length of the  $\lambda$ -5B1b DNA (approximately 3kb) was determined, it was analyzed to identify possible translation frames homologous to known proteins in data base by applying blast-x program. As shown in Figure 15, two separate reading frames in this sequence were detected. They had significant homology to more than 160 proteins which belonged to transmembrane carriers. Among them, monocarboxylic acid transporters (solute carrier family 16, member 10, T-type amino acid transporters: SLC16A10) of mouse and rat showed the highest similarity. The region between the two potential reading frames in the 5B1b sequence did not show any homology with respect to the translation. The potential reading frames were located at the 5'- and 3'-termini of the analyzed 5B1b sequence and the 5B fragment isolated by one-hybrid screen was located within the central region. These findings were consistent with the idea that 5B fragment was a part of an intron of a transmembrane amino acid transporter gene of *Acropora millepora*. Thus this gene was named *Acropora millepora* T-type amino acid transporter gene (*AmTAT1*). Figure 16 shows the sequencing strategy for  $\lambda$ -5B1b. 7108 bases were sequenced in total and, by homology search, it was predicted that the  $\lambda$ -5B1b DNA covered the entire region of the *AmTAT1* gene. The identity and genomic organization of this gene will be described later.

***AmTAT1* is a single copy gene in *Acropora millepora* genome**

To study the genomic structure of *AmTAT1*, Southern-blot analysis was carried out. A DNA fragment which corresponded to one of the downstream translation frames (in exon 3) detected by blast-x was used as the probe. As shown in Figure 17, a single band was detected on each lane. It was concluded that *Acropora millepora* had a single copy of this gene in its genome.

### **Isolation of the full-length cDNA for AmTAT1**

For isolation of *AmTAT1* cDNA, 500,000 clones of *Acropora millepora* cDNA library constructed with the mRNA from tips of the adult coral branch (sampled in October) was screened by plaque hybridization. I chose this library because the sample was taken in the season of sexual maturity of the colony and the coral *doublesex* gene, *AmDM1* is known to be expressed at the highest level (Miller et al., 2003). For the primary screening, a region of genomic DNA which covered the entire stretch of the downstream reading frame (corresponding to exon 3) was used as the probe. 8 samples were detected positive for hybridization and picked up. In order to isolate the full-length cDNA, a different probe was prepared to detect the sequence corresponding to the upstream reading frame (exon 2) in the genomic DNA of *AmTAT1*. After the secondary screening, a clone was isolated and the cDNA was rescued by the phagemid excision protocol. And the inserted cDNA was sequenced by both primer walk and subcloning. The entire nucleotide sequence of the insert was determined by reading both strands.

Figure 18 shows the nucleotide sequence with the deduced amino acid sequence and the sequencing strategy. This cDNA has all the

translation frames predicted on the genomic DNA sequence. Therefore I concluded that it was the full length cDNA for the *AmTAT1*. The cDNA is 2405 bp in length including 31 bp of poly A tail. The longest and uninterrupted open reading frame is from nucleotide 125 to 1541, which encodes the AmTAT1 protein. The open reading frame is followed by a long 3'-untranslated region of 861 bp. The 5'-untranslated region is 124 bp. The polyadenylation signal is located 51 bp upstream from the poly A tail. The sequence is AATAAT instead of the typical AATAAA. In the 3'-UTR, the G+C content is 36.9%.

### **AmTAT1 is a transmembrane protein**

The amino acid sequence of the translation was subjected to the homology search with blast-p program. 5 conserved domains were identified (Figure 19) with some extent of overlaps: AraJ, Arabinose efflux permease; Nark, Nitrate / nitrite transporter; Permease of the major facilitator superfamily; UhpC, Sugar phosphate permease; CynX, Cyanate permease; a part of MelB, Na<sup>+</sup> / melibiose synporter. These domains share the function in membrane transportation of various substances. Thus, it is concluded that AmTAT1 had apparent features of membrane transporters.

562 homologues were detected by blast-p. All proteins with the high homology were transmembrane transporters which belonged to monocarboxylic acid transporters (MCTs), a large group of solute carrier family. Among those, human T-type amino acid transporter 1 (TAT1) (Kim et al. 2001) shows the highest homology score and 31% identity. A hypothetical protein of mosquito (*Anopheles gambiae*),

ENSANGP00000022234 has 33% identity. Two hypothetical proteins of *Drosophila melanogaster* (GH12726p & CG12286-PA) also show 33% and Mammalian X-PEST (X-linked PEST containing transporter), also named MCT8, has 32% identity. MCT8 and TAT1 are the closest sisters in the MCT family (Halestrap and Meredith, 2004). Thus, it is acceptable to conclude that AmTAT1 is a homologue to these proteins. By comparison, there is relatively low homology between AmTAT1 and other groups of MCTs. For example, the vertebrate MCT4s and MCT3s show 26- 29% identity.

### **Structural feature of AmTAT1**

AmTAT1 consists of 472 amino acids and its theoretically calculated molecular mass is 51.6 kDa. The transmembrane domains were predicted by the online service TMpred ([www.ch.embnet.org/software/TMPRED](http://www.ch.embnet.org/software/TMPRED)). There are a number of clusters of hydrophobic residues in the molecule which represent the putative transmembrane regions. Figure 20 (panel A) shows the amino acid sequence of AmTAT1 with the predicted transmembrane domains (underlined). The molecule of AmTAT1 protein contains 12 strongly supported transmembrane helices.

The most probable structure was predicted that both the amino- and carboxy-terminals were intracellularly located (Figure 20, panel B). These domains are divided into two by the longest intracellular loop between the 6th and 7th transmembrane domains. And the portion comprising through the 7th to 12th domains corresponds to the AraJ permease domain. The possible post-translational modification was predicted by Predict Protein server (<http://cubic.bioc.columbia.edu>).

On the primary structure, AmTAT1 has several potential amino acid residues for modification; 3 sites for the phosphorylation by Protein kinase C and 3 sites for Casein kinase II. However, except Ser433, Thr2 and Thr215 which were predicted to be in the intracellular domains (Figure 20), the modification is supposed to be implausible because they are located in the transmembrane domains or extracellular loops. Ser433 can be the substrate for Protein kinase C, while Thr2 and Thr215 provide potential sites for Casein kinase II dependent phosphorylation. Asn181 is located in the extracellular loop between the 5th and 6th transmembrane domains. The primary sequence, NYIL is predicted to be the target of N-linked glycosylation.

### **The conserved structure in evolution**

The structural features of AmTAT1 was compared with those of the close homologues, hTAT1 (human TAT1) and X-PEST (human MCT8). The amino acid sequences were aligned by ClustalW program and shown in Figure 21. The predicted transmembrane domains were also indicated. Three proteins shared several characters with respect to their structure. In the first place, the number and positions of their transmembrane helices were almost identical. As a result, the length of extra- and intracellular loops between the transmembrane domains are nearly conserved. Secondly, the similarity of amino acid sequences is higher in the transmembrane domains. It should be noted that they are more conserved at the boundary of the transmembrane domains and adjacent extra- and intracellular loops. The intracellular loop between 6th and 7th transmembrane domains are the longest of all loops in each protein. However, this feature is not specific to AmTAT1, hTAT1 and MCT8 but common to all MCT

family proteins. The N-terminal portion of MCT8 is particularly long and rich in P, E, S and T residues. And it does not show matching on alignment. This is the exclusive feature of MCT8 proteins from which its original name (X-linked PEST-containing transporter) derived (Lafreniere et al., 1993).

### **Phylogenetic analysis on AmTAT1 within MCT family**

The molecular evolution of AmTAT1 protein was analyzed by neighbour-joining method. The entire amino acid sequences of AmTAT1 and its homologues were subjected to this analysis. Although a plethora of data were available for the sequences of MCT family, the human ortholog for each MCT member was chosen because the cloning of genes of this family had been intensively carried out mainly for vertebrates, especially mammals. Therefore the human proteins are supposed to represent the whole orthologous groups. According to the NJ tree shown in Figure 22, AmTAT1 is classified in the orthologous group together with mammalian TAT1 and X-PEST (MCT8) proteins. Three Arthropoda proteins (Karmoisin and GH12716p of *Drosophila* and ENSANGP00000022243 of *Anopheles gambiae*) are the members of the same group and form the small clade within the group. It is demonstrated that AmTAT1 segregated before the separation of TAT1 and MCT8.

### **Intron-exon organization of *AmTAT1* gene**

By comparing the genomic DNA sequence with that of cDNA, the structure of *AmTAT1* gene was deduced. The entire region of this gene spans approximately 7kb and it is composed of 5 exons and 4

introns. The diagram (panel A) in Figure 23 shows the genomic organization of *AmTAT1*. The 1st exon is a non-coding exon. The start codon of translation (ATG) appears in the second exon and the termination codon (TAA) is located in the 5th exon with the entire 3' untranslated region. The intron 1 is almost 2kb in length and the longest of all. The table below the diagram (panel B) shows the summary of information on the introns. The intron 3 is the phase 0 intron. Others are phase 1. The boundary sequences of the introns are absolutely in agreement with 'GT-AG' rule. The nucleotide sequence of the genomic DNA and corresponding sequence of cDNA are almost identical. However, some minor differences are observed only in the 3' untranslated region. The cDNA library used for the cDNA cloning was constructed with the mRNA from a different adult colony of *Acropora millepora* from the colony from which the genomic DNA was prepared. It is likely that the minor sequence variation is due to the polymorphism within the species.

### **5' regulatory region of *AmTAT1* has multiple potential sites for the binding with transcription factors**

To investigate the function of the 5'-flanking region of *AmTAT1* gene, the genomic DNA was sequenced up to about 500 bases upstream from the start site of cDNA. And the nucleotide sequence was subjected to the online search (MOTIF sequence motif search: <http://motif.genome.jp/>) for the potential binding motifs for transcription factors. The search was done with the high stringency filter (score > 90) and TRANSFAC (Heinemeyer et al., 1998) was used as the data base. In this region, multiple potential *cis*-acting elements were found *in silico*. Figure 24 shows the sequence of the 5'-flanking

region and the cis-regulatory elements with the transcription factors. No putative TATA box was found. This region includes the site for inducible transcription factors such as HSF and NIT2 and tissue-specific/developmental gene regulator such as HFH-8. It is yet to be known whether these factors actually function as the transcriptional regulator on the regulatory region of *AmTAT1*.

### **Gene expression of *AmTAT1***

For understanding the expression pattern and the influence of *AmDM1* on *AmTAT1* gene, the gene expression was studied by RT-PCR. Total RNA of egg, two developmental stages (prawn-chip and pre-settlement) and three adult coral branch tips, sampled in September, October and December 2003, was extracted and used as the template. Tips were sampled repeatedly from a same colony right after the full moon. For the October sample, the branch tip was taken before spawning. *AmDM1* is a coral homolog of DM transcription factors which includes Doublesex (*Drosophila*), Mab-3 (*C. elegans*) and Dmrt (vertebrate). They are primarily involved in sex-determination pathway. And in mature colonies of hermaphroditic *Acropora millepora*, the gene expression of *AmDM1* is exclusively detected in October, the season of sexual maturation. Taken together, it had been speculated that *AmTAT1* was also up-regulated in adult coral in October. To my surprise, the expression of *AmTAT1* was not detectable on RT-PCR for the October sample (Figure 25). On the contrary, the expression was detected positive for September and December. The positive study was also seen for the eggs and both early and late developmental stages.



## **Computer analyses could not find the DM-binding motif in the 5B fragment**

Now that 5B fragment, the genomic DNA fragment isolated by the large-scale yeast one-hybrid screen was known to be a part of an intron of *AmTAT1* gene, a homolog of amino acid transporters, I attempted to show that a motif for binding with the DM domain of AmDM1 was actually present in the 5B region. It would further provide the evidence that AmTAT1 was one of downstream genes in the pathway regulated by AmDM1.

First, a computer based search for binding motifs was carried out via MOTIF sequence motif search (<http://motif.genome.jp/>). However, it did not detect any potential DNA elements for the association with known Doublesex homologs in the data base. The search showed the same result when the stringency of the filter was relaxed.

## **Finding the responsible subfragment of 5B for the association with DM domain by yeast one-hybrid assay**

The 5B fragment, approximately 487 bp in length, was divided into two subfragments, 5B $\alpha$  and 5B $\beta$ , which partially overlapped with each other (Figure 26). The sizes of these subfragments were 260 and 220 bp, respectively. These fragments were synthesized by PCR amplification with appropriate primer pairs shown in Figure 26 and subcloned into pBM2589, the vector used for the construction of the Reporter plasmids. The molecular interaction between the DNA subfragments and DM domain was examine by one-hybrid assay. The assay was performed by transforming the recipient yeast clone

YM4271/pDMact5 with the Reporter construct. The transformants were selected on SD/-HIS plates. Colonies were streaked again to be photographed. The pictures in Figure 27 shows the massive growth of the transformed yeast with 5B $\alpha$  subfragment after a 3-day incubation but no growth was observed for either 5B $\beta$  or non-inserted pBM2389 vector after a extended period of time (up to 9 days). These result suggested that the binding element was included in the 260 bases of 5B $\alpha$  fragment.

### **Finding the responsible element in 5B $\alpha$ subregion for the association with DM domain by electrophoretic mobility shift assay (EMSA)**

I tried to narrow the region of DM-binding element by applying electrophoretic mobility shift assay (EMSA). This technique is applicable for the DNA fragments with sizes up to 300 bp. The entire DM domain of AmDM1 was bacterially expressed as a fusion protein with Glutathione-S-transferase. The chimeric protein had a short stretch of 67 amino acids corresponding to the DM domain at the carboxy terminus of Glutathione-S-transferase. The fusion protein, GST-DM was successfully expressed and purified.

In the first step, the 5B $\alpha$  fragment was again divided into two subfragments. As shown in Figure Xa, 5B $\alpha$ A was a 130 bp fragment which covered the 3' half of the 5B $\alpha$  and 5B $\alpha$ B was a 150 bp fragment covering the 5' half. Note that they had a small overlapped region that corresponded to the middle of 5B $\alpha$ . These subfragments were radiolabelled by PCR amplification and used as the probes. The Reporter plasmid that had been isolated by the one-hybrid screen was

used as the template. The result of EMSA is shown in Figure 28. On the upper panel, retarded bands is visible only for the 5B $\alpha$ A probe. The DNA-protein complexes were detected when 100ng of the fusion protein GST-DM was incubated with the probe. On the contrary, no mobility shift was observed for the 5B $\alpha$ B probe. 5B $\beta$  probe was also incubated and electrophoresed as the negative control. As seen in the lower panel of Figure 28, GST itself did not show any affinity to the three probes. These result indicates that 5B $\alpha$ A has the binding activity to the DM domain. In addition, it was proved that the DM domain of AmDM1 was a DNA-binding domain.

To further narrow the suspected region for the interaction, nested primers were designed and used for preparation of the probes. As in Figure 29, three nest-deleted probes were generated (5B $\alpha$ A0, A1 and A2) by PCR. And they were incubated with 400ng of GST-DM protein. Hereafter up to 400ng of the protein was used for incubation in order to emphasize the retarded signals. All probes showed the retarded bands for DNA-protein complex.

The binding specificity of the sequence including in the 5B $\alpha$ A2 probe, competition study was performed with specific and non-specific cold probes. Larger amounts of cold probes (up to x100 excess) were incubated with the radioactive 5B $\alpha$ A2 probe. The specific competitor reduced the intensity of the retarded signal (Figure 30). On the other hand, non-specific 5B $\alpha$ B competitor did not affect the complex formation. It was concluded that the mobility shift was the result of specific binding of 5B $\alpha$ A2 probe with DM domain.

The diagrams in Figure 26 were drawn in accordance with the

nucleotide sequence of the 5B insert of the Reporter plasmid. By aligning the sequence with that of the corresponding region in *AmTAT1* genomic DNA, it was found that 17 bases at the 5'- and 70 bases at the 3'-terminus of 5B insert did not appear in the  $\lambda$ -5B1b DNA. The latter is included in the 5B $\alpha$ A subfragment. Although there is no idea about the origin of these sequences, it is likely that they are the artifacts during the library construction. To exclude the possibility that the 70bp stretch at the 3'-end of 5B $\alpha$ A includes the DM-binding motif, I assayed the interaction between DM-domain and the DNA probes which corresponded to the genome-specific sequence. Two probes were prepared as shown in Figure 31 (lower diagram) and tested for binding. In the phosphorimage, genomic A2 probe formed complex with GST-DM protein while genomic A3 probe did not. The genomic A3 probe was another nested fragment (Figure 31 and Figure 32). These findings demonstrated that the 18 bp stretch which was present in the A2 probe but lacking in the A3 was responsible for the association with DM domain. In other words, this stretch includes the DM-binding element. It was also suggested that either the plasmid- specific sequence or the genome-specific sequence (in the  $\lambda$ -5B1b DNA) in the probe did not affect the DNA-protein association.

As mentioned in the section for the large-scale one-hybrid screening, the *Acropora*-specific repetitive sequences were identified as potential binding regions to the DM domain. Next, I tried to demonstrate the *in vitro* association between DM domain and the repetitive sequence. The repetitive sequences and the suspected DM-responsible sequence in 5B $\alpha$ A2 probe have a pentanucleotide stretch of the identical nucleotide sequence (Figure 32). Two double stranded

oligonucleotide probes including the short stretch (ACATT /AATGT) were prepared in accordance with the sequences of the repetitive sequence and A2 primer region. EMSA (Figure 33) shows the mobility retardation for both A2 and repetitive oligonucleotide. This result indicates that the repetitive sequence as well as the A2 region has a motif for the interaction with DM domain. And it is likely that the pentanucleotide stretch is necessary for the binding.

To test this hypothesis, a mutant probe (mutA2) was designed substituting ACATT/AATGT to GAGCC/GGCTC. As expected, no complex formation was detected on EMSA (Figure 34). When a large excess of the cold mutant oligonucleotide was incubated with non-mutant 5B $\alpha$ A2 probe, it did not compete for binding with DM domain (Figure 34, right panel).

The nucleotide sequences of DM recognition element in 5B $\alpha$ A2 and the repetitive sequence were compared and aligned with known binding motifs for DM domains of Doublesex homologs of various organisms (Figure 35). As for the motifs of *Acropora millepora*, their consensus is AATGT(GT/TG)C. The pentanucleotide core (AATGT) is conserved throughout evolution. Two motifs of insects (Doublesex of *Drosophila melanogaster* and *Bombyx mori*) and one motif of worm (Mab-3) share the AATGT sequence. Taken together, the universal consensus AATGT is supposed to play a pivotal role in the interaction with DM domains.

### **Transcriptional regulation by AmDM1 on *AmTAT1* expression**

As mentioned in the previous section, *AmTAT1* gene has the

recognition motif for the molecular interaction with DM domain of AmDM1, the *Acropora* homolog of Doublesex. It was demonstrated in yeast by one-hybrid assay and *in vitro* by EMSA. There is no doubt that AmDM1 is the DNA-binding protein as well as the other DM domain-containing proteins and the DM-domain of AmDM1 functions as the DNA-binding domain. However, it is still not clear that the DM recognition element has the *cis*-regulatory activity for the gene expression of *AmTAT1* and at the same time, AmDM1 acts as a transcription factor on the element. Thus, I planned to investigate the *cis*-acting regulation by DM recognition element by deploying the CAT reporter system.

### Strategy for CAT assay

There were two points to note for performing the CAT reporter system in this case. Firstly, the DM binding motif of *Acropora millepora* was found in the 2nd intron of *AmTAT1* gene. It means that this *cis*-acting element can function as an enhancer. Therefore, the assay system to detect enhancer activity was designed. In the second place, CAT assay is an *in vivo* detection system which makes use of mammalian cultured cells. The efficiency of this assay depends primarily on the nature of the cell lines. It is unlikely that mammalian cells by themselves perform well to detect the regulatory activity of DM recognition motif not only because of the evolutionary distance between mammals and coral but also it is unfeasible to find cell lines expressing the transcription factors which act on the DM recognition motif such as the DM family transcription factor, Dmrt1. For these reasons, I tried to carry out the double transfection experiment in which the expression (activation) vector for AmDM1 is introduced to

the cells together with the reporter plasmid for CAT (Figure 35).

For construction of the expression vector, a partial cDNA including the entire open reading frame of AmDM1 was inserted into the multiple cloning site of pRc/CMV. A 259bp DNA fragment corresponding to the sequence of the intron 2 of *AmTAT1* was generated by PCR amplification. The DM recognition motif (DM enhancer) was present at the center of this fragment. And this fragment was inserted into the cloning site of pCAT3 promoter vector in both forward and reverse orientation (Figure 36). As mentioned before, pCAT3 promoter was developed to assay the enhancer activity.

#### **AmDM1 acts as a transcriptional repressor on DM enhancer**

The co-transfection into Hep-2 cell line was carried out with all possible combination of the expression vectors and reporter plasmids (Figure 37, lower panel). The values of CAT activity for the DM enhancer were standardized by the values for the empty pCAT3 promoter. And the standardized value of DM enhancer with the empty expression vector was compared with those with AmDM1 expression vector. For example, in the first standardization step, the values for <4>, <5> and <6> were divided by the values for <1>, <2> and <3>, respectively (Figure 37). In the next step, the standardized value for <5> was compared with the standardized value for <6>. Values for <4> provided the control.

The summary of triplicate assays is shown in the histogram in Figure 37. Marked decrease with AmDM1 transfection was observed in

comparison with empty vector transfection. It means that AmDM1 acted on DM enhancer as a repressor. The statistical analyses revealed 5% and 10% significance for forward and reverse orientation of DM enhancer, respectively. This observation is in agreement with that of gene expression in adult coral branch tips. RT-PCR showed that the expression of *AmTAT1* was not detected in tips for October when *AmDM1* was expressed. In contrast, *AmTAT1* was detected in tips for both September and December.

### **AmDM1 activates SV40 promoter in Hep-2 cells**

During the analysis on data of the CAT assay, I noticed that the ‘AmDM1 expression vector + empty control reporter’ assay resulted in a marked increase in CAT activity in comparison with the ‘empty expression vector + empty control reporter’ assay. It means that introduction of the AmDM1 expression vector into the cells up-regulated the SV40 promoter activity in the pCAT3promoter. This increase was also observed for the assays with the inserted reporter. Although SV40 promoter does not have potential sites for the association with DM domain, it is possible that AmDM1 activates the transcription of genes involved in the up-regulation of SV40 promoter. It is also speculated that AmDM1 associates with factors which have negative or positive influences on the promoter.

In any case, I thought this observation made it possible to study the functional structure of AmDM1 simply with the empty control reporter (pCAT3promoter). As mentioned before, AmDM1 is known to have two conserved domains, one DM domain and one DMA domain. While DM domains are Zinc-finger type DNA-binding



domains, the functional properties of DMA domains are yet to be known. And no conserved domains have been identified outside these domains.

To test the modular structure of AmDM1 in view of the function, I constructed the nest-deletion mutants of AmDM1 (Figure 39, panel A). Two deletion mutants were designed.  $\Delta$ C-terminal mutant lacks C-terminal region but retains DM and DMA domains.  $\Delta$ DMA mutant does not have both C-terminal and DMA domains. The cDNA fragments for these mutants were generated by PCR and inserted in the expression vector. And these vectors and the expression vector with the full-length AmDM1 were co-transfected with pCAT3 promoter into Hep-2 cells. The histogram in Figure 39 (panel B) shows the summary of the triplicate assays. The relative CAT activity was increased significantly by the introduction of full-length AmDM1. On the other hand, both the two deletion mutants,  $\Delta$ C-terminal and  $\Delta$ DMA did not change the expression level. The increase for the full-length AmDM1 compared with the level for  $\Delta$ C-terminal mutant shows the statistical significance ( $p < 0.05$ ). In conclusion, the C-terminal region of AmDM1 protein was shown to be essential for the transcriptional activation suggesting that this region contained at least a part of the activation domain. The participation of DMA domain in the trans-activation is not clear.

### **cDNA cloning of *vitellogenin* (*Acropora millepora*)**

In *Drosophila*, the expression of *yolk protein 1* gene was reported to be regulated by Doublesex transcription factor ( $Dsx^M$  and  $Dsx^F$ ) and the 5'-regulatory region of this gene has the recognition motif for DM domain of Doublesex. This finding provided a model for the function

of DM-containing proteins. As for *Caenorhabditis elegans*, *vitellogenin* genes were shown to be controlled by Mab-3, the homolog of Doublesex. The control region of *vitellogenin* genes contain the multiple DM recognition elements. It is possible that in *Acropora millepora* the homolog of *vitellogenin* is also expressed under the control of AmDM1 in a *trans*-acting manner. To study the role of AmDM1 in the transcriptional regulation of *vitellogenin* gene, I planned to identify the AmDM1 responsible element in the control region of *vitellogenin* genomic DNA.

Species *Favites chinensis* is a scleractinian coral which belongs to Suborder Faviina, Family Faviidae. Recently, the coral homolog of *vitellogenin* was identified for *F. chinensis* (Imagawa S. et al. 2004) and the nucleotide sequence of the cDNA has been available. This species of coral was expected to have genes similar to *Acropora*. Thus the nucleotide sequence of *F. chinensis vitellogenin* (*FcEP-1*) enabled me to design the hybridization probes to isolate the cDNA and genomic DNA of *Acropora* homolog of *vitellogenin*.

### **Isolation of the full-length cDNA**

Although the cDNA was not full-length, the deduced amino acid sequence of FcEP-1 showed that the protein had a highly conserved region in its Carboxy terminal portion. This region was identified the von Willebrand domain (VWD). Primer pairs were designed on the basis of *FcEP-1* cDNA sequence to amplify the DNA fragment corresponding to the VWD of Vitellogenin. The 1st strand cDNA was synthesized by reverse transcription with the total RNA extracted from adult coral tips of *Acropora millepora* (sampled in September

and October 2003). By RT-PCR with a low stringency condition, the DNA fragment for VWD was successfully amplified with both September and October samples. The fragment was purified and sequenced for confirmation.

For isolation of *Acropora millepora vitellogenin* cDNA, 500,000 clones of *Acropora millepora* cDNA library constructed with the mRNA from tips of the adult coral branch (sampled in October) was screened by plaque hybridization. For the primary screening, the DNA fragment corresponding to the VWD was amplified and radiolabelled by PCR. Surprisingly, more than 3000 signals were detected on hybridization. 50 areas with more or less strong signals were picked up and brought to further screening. Before the secondary screening by hybridization, these samples were subjected to a test by PCR amplification of VWD region. Two samples with the positive study were screened by plaque hybridization and 4 well-isolated plaques were obtained for each sample. From one of those clones, the cDNA was rescued by the phagemid excision protocol. And the inserted cDNA was sequenced by both primer walk and subcloning. The entire nucleotide sequence of the insert was determined by reading both strands.

The cDNA, *AmVitD* is 4630 bp in length without poly A tail and harbours an uninterrupted open reading frame for Vitellogenin which starts at 17th nucleotide followed by a 344 bp 3'-untranslated region. The polyadenylation signal is located 20 bp upstream from the poly A tail. The sequence is the typical AATAAA. The deduced translated protein consists of 1414 amino acids. The length is equivalent to Vitellogenins of other organisms. And, as described below, the homology search showed it had high homology to Vitellogenins with

two conserved domains, Lipoprotein N-terminal domain (LPD-N) and von Willbrand domain (VWD), which were characteristic to a number of Vitellogenin homologs. Taken together, I concluded that *AmVitD* was the full-length cDNA for *Acropora millepora* Vitellogenin.

### **Molecular properties of *Acropora millepora* Vitellogenin (AmVit)**

Figure 40 shows the deduce amino acid sequence of Vitellogenin. The protein consists of 1414 amino acids. As shown in the lower diagram in Figure 40, the Lipoprotein N-terminal domain (LPD-N) spans from 21nd to 606th residue in its N-terminal region. And von Willbrand domain (VWD) starts at 1239 and ends at 1396.

The homology search by blast-p for the LPD-N revealed it was highly conserved through metazoan evolution. Among them, Vitellogenin A2 of *Xenopus laevis* showed the highest homology (22% identity) followed by Vitellogenin of *Crassostrea gigas* (20% identity), *X. laevis* B1 (22%) and *C. elegans* Vit6 (21%). And the phylogenetic analysis (Figure X) indicates that the LPD-N domain of AmVit forms a clade with those of vertebrate homologs. Vitellogenin of *Favites chinensis* did not appear on the list of homologous proteins because the cDNA for FcEP-1 was not full-length including only 648 amino acids in the C-terminal region.

As for VWD, FcEP-1 shows the highest homology of 51% identity. Vit5 and Vit6 of *C. elegans* are 26% identical. However, apart from FcEP-1, VWDs of other phyla only shows moderate homology scores revealing the structural or functional constraint is weaker for this domain than

for LPD-N.

### **Genomic DNA cloning of *Acropora millepora vitellogenin***

To find the DM responsible elements in the control region of *Acropora millepora vitellogenin* gene, genomic DNA for AmVit was isolated from the  $\lambda$ -GEM12 *Acropora millepora* genomic library. For the primary screening approximately 200,000 clones were screened by plaque hybridization. A 310bp DNA fragment corresponding to a part of Lipoprotein N-terminal domain was amplified and radiolabelled by PCR for the hybridization probe. One signal was detected and the area including the signal was picked up. After the secondary hybridization, 5 well-isolated plaques were obtained.  $\lambda$ -phage DNA was extracted from one of the clone and partially sequenced by primer walk.

### **Complicated and conserved intron-exon organization of *Acropora millepora vitellogenin* gene**

About 3kb of the AmVit genomic clone was sequenced. The primer walk was started near the 5'-end of the putative transcribed region. As seen in Figure 41, the region corresponding to the N-terminal 235 amino acids was interrupted by multiple introns. Comparison with the chicken vitellogenin gene *vgtII* (Figure 41, below *AmVit*) demonstrated that the intron-exon organization of two genes was almost identical. Although there is difference between the sizes of corresponding introns to some extent, the intron positions are surprisingly conserved. The sizes, phases and sequences around the intron-exon boundaries are shown in Table 4. The splicing sites are

absolutely in accordance with the GT-AG rule.

### **No DM binding motif in the upstream regulatory region and introns**

As shown for the *AmTAT1* gene, DM domain of AmDM1 binds to the DNA consensus AATGT(GT/TG)C. In the sequence of the 5'-flanking region and intron 1-5, I could not find the potential DM-binding motif. However, multiple transcription factor binding sites were found in the upstream region including TATA box (Figure 42). And intron 1 and intron 2 were found to contain potential *cis*-acting elements as well. Thus it is concluded that the isolated genomic DNA represents the functional gene for *vitellogenin*.

## Section 4. Discussion

AmDM1, the *Acropora millepora* homolog for Dsx, contains the DNA-binding DM domain and is thought to be a transcription factor as well as other DM family proteins. Information about genes which are transcriptionally regulated by AmDM1 is essential to know the actual mechanisms for sexually differentiated formation of gonads because DM domain family genes function at or close to the bottom of sex-determination hierarchy. To find gene(s) at the downstream of *AmDM1* in this hierarchy, large-scale screening was carried out by applying the yeast one-hybrid system. This system was designed to pick out DNA elements with affinity to a protein domain of interest.

### ***Acropora*-specific repetitive sequences contain DM-binding elements**

By the one-hybrid screening, 5 clones were selected positive for the interaction between the inserted genomic DNA and DM domain of AmDM1. 4 of them were sequenced and identified. The inserts of 3 clones included the repetitive sequences which were specific to genus *Acropora* (McMillan and Miller, 1990; McMillan and Miller, 1989; McMillan et al., 1991). Subsequent EMSA (electrophoretic mobility shift assay) showed the physical interaction between DM domain of AmDM1 and DNA motifs within the repetitive sequences. According to the estimation (McMillan and Miller, 1989), there are 80 to 300 x 10<sup>3</sup> copies of the repetitive sequences accounting for approximately 5% of *Acropora* genome. In spite of the physical interaction, it seems impossible that all DM-binding motifs in the repetitive sequences are

accessible by AmDM1 in cells of living coral. Should it happen, a large proportion of AmDM1 molecules in the cells would be consumed by the repetitive sequences. I speculate that all or most of the repetitive sequences are silent for the association with transcription factors and non-functional as *cis*-acting elements probably due to the local structure of chromatin.

The fact that the various clones for the repetitive sequences were isolated by the one-hybrid selection in independent experiments shows the fidelity of this screening system.

### **Identification of the AmDM1 recognition motif**

I found that the 2nd intron of *AmTAT1* contained a binding site for the DM domain of AmDM1. By comparing its nucleotide sequence with that in the repetitive sequence, it was concluded that the recognition motif is AATGT(GT/TG)C. This motif is not pseudopalindromic like the Dsx-binding element on which two Dsx protein molecules form a homodimer (Cho and Wensink, 1996). Thus, it is highly likely that one AmDM1 molecule binds to the recognition motif as a monomer. The core sequence AATGT appears in all DM domain binding motifs known so far (Mab-3 in *C. elegans*, Dsx in *Drosophila* and BmDsx in *Bombyx mori*). As shown by EMSA, this pentanucleotide stretch is essential to the protein-DNA interaction. The high sequence conservation in these motifs reveals a strong structural constraint during metazoan evolution not only in DM domains but also in the DM recognition elements.

### **AmDM1 downregulates the expression of *AmTAT1***



The fact that a binding motif for AmDM1 is found in the 2nd intron of *AmTAT1* gene implies that this motif plays a *cis*-acting role as an enhancer for the gene expression. The enhancer activity was shown by CAT assay. In my CAT assay system with Hep-2 cells, AmDM1 acted as a repressor. The introduction of AmDM1 expression vector significantly downregulated the CAT activity. This result is consistent with the gene expression pattern of *AmTAT1*. In adult coral colonies, the gene expression of *AmDM1* was shown to be at the peak in October, the season of sexual maturity (Miller et al., 2003). On the contrary, I could not detect the expression of *AmTAT1* by RT-PCR for the October sample, whereas it was detectable for the September and December samples. However, it is possible that AmDM1 has both activating and inhibiting functions as a transcription factor and *AmTAT1* is one of the genes which are downregulated by AmDM1. In fact, the female type Dsx of *Bombyx mori* (BmDsx<sup>F</sup>), when introduced in male animals, activates the gene expression of *vitellogenin* and represses *PBP* (pheromone binding protein) at the same time (Suzuki et al., 2003).

### ***AmTAT1* encodes a membrane transporter protein**

By homology search, the protein encoded by *AmTAT1* turned out to be a member of monocarboxylate transporter (MCT) family. hTAT1 (human T-type amino acid transporter) (Kim et al., 2002) shows the highest homology to the translated protein that was named after TAT1.

The monocarboxylate transporter family --*SLC16* gene family--

comprises 14 members (MCT1-9, TAT1 and MCT11-14) (reviewed in Halestrap and Meredith, 2004) and they have been well studied especially for mammals. All members have a typical structural feature of membrane proteins. Predominant substrates have been demonstrated for several MCTs. For example, MCT1 co-transport lactate, pyruvate and ketone bodies with H<sup>+</sup> (Garcia et al., 1994; Kim et al, 1992); MCT2 also co-transport lactate, pyruvate and ketone bodies with different affinities from those of MCT1 (Garcia et al., 1995; Lin et al., 1998; Bröer et al., 1999).

MCT proteins have some biochemical and structural characteristics in common. 1) They contain 10 to 12  $\alpha$ -helical transmembrane domains with N- and C-termini located intracellularly. 2) The C-terminus is usually long. 3) And the intracellular loop between the 6th and 7th transmembrane domains is the largest. The length and the amino acid sequence of this portion is poorly conserved. 4) In the transmembrane domains, the amino acid sequences are well conserved. 5) Two highly conserved sequences are found in the transitional portion to the 1st and 5th transmembrane domains. The consensus sequences are **[D/E]G[G/S][W/F][G/A]W** and **YFXK[R/K][R/L]XLAX[G/A]XAXAG**, respectively (X stands for any amino acid residue and bold letters indicate absolutely conserved residues.) (Halestrap and Price, 1999). All of these characteristics are retained by AmTAT1. In addition, no glycosylation has been demonstrated for MCTs. There is a general tendency that higher conservation is observed in transmembrane domain 1-6 (N-terminal domains) than in 7-12 (C-terminal domains). The N-terminal domains has been supposed to take part in energy coupling, membrane insertion and maintenance of the structure. On the other hand, the C-

terminal domains function for determination of the substrate specificity (Price et al., 1998) and constitute a conserved Arabinose efflux permease domain revealing the ancient origin of this transporter family. As a matter of fact, MCT homologues are found in yeast, *S. cerevisiae* and archaebacterium *Sulfolobus solfataricus* (Price et al., 1998).

### **Molecular evolution of *AmTAT1***

Phylogenetic analysis indicates that TAT1 and MCT8 (XPCT) constitute an orthologous group (Halestrap and Meredith, 2004). In my analysis, AmTAT1 segregates before the separation of TAT1 and MCT8 (XPCT) and the three proteins form a subgroup (Figure 22).

A cDNA encoding TAT1 was first isolated from rat small intestine (Kim et al., 2001). Functional analyses for rat TAT1 (rTAT1) revealed properties of the classical amino acid transport system T. Unlike proton-linked MCTs, rTAT1 is not coupled with H<sup>+</sup>. The transport is carried out in an Na<sup>+</sup>-, Cl<sup>-</sup>- and pH-independent manner. rTAT1 transports aromatic amino acids (tryptophan, tyrosine and phenylalanine) and the derivatives (L-DOPA and 3-o-methyl-DOPA). The expression of rTAT1 was observed in intestine, placenta and liver. The basolateral membrane and cytoplasm of the small intestinal epithelium showed the highest expression implying the role in transmembrane transport from the epithelium to the capillary (Kim et al., 2001). The human TAT1 was isolated and characterized later (Kim et al., 2002). Both TAT1s share the structural, biochemical and functional properties. However, tissue distribution of the proteins seems different. The human TAT1 was detected in kidney (the most

abundant), skeletal muscle, heart and placenta.

MCT8 (XPCT) was found during the analysis on X-inactivation (Lafreniere et al., 1994) and mapped close to the X-inactivation center in both human and mouse (Debrand et al., 1998). The original name, X-linked PEST-containing transporter was derived from the 75 amino acid extension at the N-terminus which was rich in P (proline), E (glutamic acid), S (serine) and T (threonine). This motif is thought to be related to the rapid degradation of the protein. The expression of *MCT8* is detected in several organs including liver, kidney, brain and heart but not detectable in spleen, testis and skeletal muscle. Interestingly, MCT8 transports thyroid hormones (thyroxine and triiodothyronine) but does not transport leucine, phenylalanine, tryptophan and tyrosine (Friesema et al., 2003).

Although *AmTAT1* seems to have evolved from the common ancestor of *MCT8* (*Xpct*) and *Tat1* according to the phylogeny tree, it is probably that AmTAT1 functions as a t-type amino acid transporter because AmTAT1 lacks the PEST-motif in the N-terminal region. However, the name and identity of AmTAT1 may be revised by the investigation on the substrate specificity.

### **Roles of AmTAT1 in sex-determination**

*AmTAT1* was identified as an immediate downstream gene of which transcription was under the direct control of AmDM1. Albeit the involvement of this gene in the sex-determination is yet to be studied, it is possible to propose some hypothetical models for the roles of AmTAT1. 1) AmTAT1 uptakes particular amino acids or derivatives

from the environment, which inhibit sexual differentiation of the gonads in adult *Acropora millepora*. By downregulating the expression of *AmTAT1* in the reproduction season, AmDM1 stops this inhibitory signal in order to start the sexual differentiation of gonads or the germ cell maturation. 2) Downregulation of AmTAT1 leads to an imbalance in the amounts of intracellular amino acids, which favors the promotion of sexual maturation. 3) In F9 murine embryonal carcinoma cell line, *dmrt1* was shown to be one of the growth arrest genes induced by amino acid starvation (Fontanier-Razzaq et al., 2002). AmDM1 inhibits *AmTAT1* expression to keep the growth arrest status of particular cells to save energy for reproduction.

In any case, it can be concluded that the regulation of *AmTAT1* expression by AmDM1 is an important part of the sex-determination cascade in *Acropora millepora*. And the function of AmTAT1 probably is indispensable for the reproduction of coral.

### **AmDM1 contains the activation domain in the C-terminal region**

In the deletion mutant study, I showed that the C-terminal region of AmDM1 was indispensable for the transcriptional regulation. Although the boundaries of the indispensable portion has not yet been defined, it is suggested that the C-terminal region includes the activation domain of AmDM1. The alignment of C-terminal regions of AmDM1 and NvDM5 led to the finding of small 'islands' of conserved elements between these orthologs (Figure 11). It is likely that these islands represent parts of the activation domains.

In *Drosophila*, *doublesex* needs *intersex* (*ix*) for its proper function. It is suggested that IX provides the transactivating function to Dsx<sup>F</sup> which lacks the activation domain (Garrett-Engele et al., 2002). A recent report indicated that IX was a homolog of a subunit of Mediator complex, a co-activator of RNA polymerase II (Sato et al., 2003). It is also likely that the C-terminal portion of Cnidarian DM proteins (AmDM1 and NvDM5) interact with transcription regulators to control the expression of the target genes.

### **No DM-binding motif has been found in *vitellogenin* gene of *Acropora millepora***

The genomic DNA of *AmVit*, *Acropora millepora vitellogenin* gene was partially sequenced. I concluded the genomic clone represents the functional *vitellogenin* gene because the nucleotide sequences in the coding regions are identical to the cDNA for *AmVit*, the genomic clone has multiple exons and the 5'-regulatory region has potential binding sites for multiple transcription factors. However, I could not find the DM-binding motif in the 1st to 4th introns or in approximately 600bp of the upstream regulatory region. Although in *vitellogenin* genes of *C. elegans*, Mab-3 binding sequences located close to the transcriptional start sites (within the 350bp upstream region) (Yi and Zarkower, 1999), it is possible that DM-binding elements will be found at the distant upstream in *AmVit* gene. Or *Acropora millepora* may have another gene for *vitellogenin* which is transcriptionally controlled by AmDM1. In *Favites chinensis*, Vitellogenin was identified one of two major soluble egg proteins on SDS-PAGE (Imanaga et al., 2004). The other protein (FcEP-2) is yet to be characterized but its homolog in *Acropora millepora* may be the

candidate for a gene which is controlled by AmDM1.

The partial intron-exon organization is almost identical to that of vertebrates (Spieth et al., 1991). This gene will provide an example of high conservation in the genomic organization between the lower animal and vertebrate.

## Chapter 3

# Characterization of Two *Cnnos* Homologs of *Acropora millepora*



## Section 1. Introduction

### ***nanos* was identified as the posterior determinant**

Germ plasm is a morphologically distinct region in eggs of vertebrates and invertebrates. In *Drosophila*, it is found at the posterior pole of the egg cells and includes specialized structures, polar granules (Illmensee et al., 1976). The polar granules are electron-dense aggregates of proteins and RNA molecules. At least 8 genes have been known to be indispensable for the germ plasm formation. They are *cappuccino* (*cap*), *spire* (*spir*), *staufen* (*stau*), *oskar* (*osk*), *vasa* (*vas*), *tudor* (*tud*), *valois* (*vls*) and *mago nashi* (Schüpbach and Wieschaus, 1986; Boswell and Mahowald, 1985; Lehmann and Nüsslein-Volhard, 1986; Manseau and Schüpbach, 1989; Boswell et al., 1991). These genes are maternal effect genes of which RNAs are supplied during oogenesis. Mutations in the genes cause lack in pole plasm and abdominal defect in the embryo. *nanos* (*nos*) and *pumilio* (*pum*) were found to be necessary for abdominal morphogenesis and classified posterior class genes together with the 8 genes mentioned above (Lehmann, 1985). Mutants for *nanos* or *pumilio* only leads to abdominal defect and neither has been found necessary for the pole plasm formation.

By genetic and cytoplasmic transfer study, it was shown that *nanos* was the posterior determinant . Among the posterior group genes, *staufen*, *oskar*, *vasa*, *valois* and *tudor* function at the upstream of *nanos* (Lehmann and Nüsslein-Volhard, 1991), and *pumilio* is the downstream gene in this hierarchy. For example, *oskar* plays a role in

organizing the germ plasm (Ephrussi, et al., 1991).

The *nanos* gene was mapped on the right arm of the 3rd chromosome (Lehmann and Nüsslein-Volhard, 1991) and isolated. The gene is interrupted by two introns and encodes a protein of 400 amino acids. By whole-mount in situ hybridization, *nanos* transcripts were detected in the pole plasm of the eggs and in the pole cells during early embryogenesis (Wang and Lehmann, 1991). It was also found that localization of the *nanos* transcript at the posterior pole was crucial for its function and the 3'-untranslated region of *nanos* mRNA contained the signal for posterior localization which was mediated by *oskar* (Gavis and Lehmann, 1992).

### ***nanos* is required for development and function of germline cells**

As in the case of other posterior group genes, the posterior determinant *nanos* is necessary for the proper development of abdomen in *Drosophila* embryo. Although *nanos* RNA is one of the components of pole plasm in *Drosophila* eggs, the pole plasm is formed in the mutant without the normal function of this gene (Nüsslein-Volhard et al., 1987). However, the embryo from the *nanos* mutant mother grows infertile. During the development, pole plasm is incorporated into pole cells. Pole cells remain segregated from the somatic cells during gastrulation and are brought inward by the posterior midgut invagination. Subsequently they migrate to the mesoderm and split into two clusters. Each cluster of cells move laterally and reach the embryonal gonads. Nanos protein remains detectable in the pole cells through this process.

The requirement of *nanos* gene for germline development was first reported by Kobayashi (Kobayashi et al., 1996). They found that pole cells lacking Nanos failed to migrate into gonads. And premature expression of several germline-specific markers was observed. This observation suggested that the normal development of germline precursors were disturbed in these pole cells. In addition to the premature expression, ectopic transcription of *sex-lethal* (*Sxl* : sex-determination gene), *fushi tarazu* and *even-skipped* (segmentation genes) were also observed in *nanos* mutant germ cells in the syncytial blastoderm stage. (Deshpande et al., 1999). These cells are unable to remain quiescent for mitosis. Thus *nanos* plays a crucial role in the normal development of the germline cells. *sxl* is supposed to be the potential target of the downregulation by Nanos because removal of *sxl* from *nanos* mutant rescued the defect in germ cell migration.

In many metazoa, germline cells are maintained in adult organisms. Germline stem cells keep producing gametes by asymmetric division. This division is self-renewing mitosis and at the same time produces a differentiated daughter cell. In female *Drosophila*, a cystoblast, the differentiated daughter cell divides four times to give rise to a cluster of cystocytes. One of these cystocytes becomes the oocyte and the others become nurse cells.

The adult mutant females for *nanos* have abnormal features for egg production. They show the egg laying defect to various extents. And the number of cystocytes decreases progressively in an age-dependent manner. The rate of cystoblast production by stem cell division is slower than wild type (Bhat, 1999). The stem cell undergoes mitosis

only a few times and degenerates. Cell death of these germline is observed with degeneration of the plasma membrane and subsequent mitochondrial accumulation. As well as in female, *nanos* mutant male shows defect in gamete production. They show massive accumulation of spermatocytes. Their spermatogenesis is affected and the males eventually become sterile. Thus *nanos* is required for the maintenance of germline stem cells in both sexes.

In addition to the abdominal determination, *nanos* is a multifunctional gene with respect to the germ cell formation in *Drosophila* : it is indispensable for the germline development, germline cell migration in embryos and the maintenance of germline stem cells in adult flies.

### **The functional aspects of Nanos protein**

The function the molecular properties have been intensively studied about *Drosophila* Nanos protein. In 1989, it was reported that *nanos* functioned as a repressor on the activity of a gap gene, *hunchback* (Irish et al., 1989). Embryos lacking the normal activity of *nanos* show abdominal defect. The defect was found to be rescued by eliminating the maternal *hunchback* gene. *hunchback* is one of the gap segmentation gene which is necessary for the formation of segments in the head and anterior thorax. The maternal *hunchback* transcripts were produced during oogenesis and uniformly distributed in the eggs (Tautz et al., 1987). After fertilization, the Hunchback protein shows an asymmetric gradient ; the concentration is highest at the anterior pole and it decreases toward the posterior. The anterior factor, Bicoid promotes the zygotic transcription of

*hunchback* in the anterior portion, whereas Nanos blocks the translation of this gene (Tautz, 1988). The blockade on the translation of *hunchback* is mediated by the regulatory sequences in the 3' untranslated region of the mRNA (Wharton and Struhl, 1991). The sequences, NREs (*nos* response elements) is necessary and sufficient for the translational inhibition by Nanos protein. The *hunchback* mRNA has two NREs. The NRE is a 32 bases long sequence including the consensus GUUGU and AUUGUA spaced by 5 nucleotides. Interestingly, this consensus was also found in the 3' untranslated region of the *bicoid* mRNA. Actually the translation of *bicoid* is blocked in the presence of Nanos protein resulting the inverse gradient of Bicoid to Nanos.

*pumilio* is necessary for the proper function of *nanos* in abdominal formation (Barker, 1992; Lehmann and Nüsslein-Volhard, 1991). It was speculated that Nanos interacted with *hunchback* mRNA in collaboration with Pumilio . In fact, the downregulation on *hunchback* translation is dependent not only on Nanos but also on Pumilio protein (Wreden et al., 1997). It was also suggested that the translational repression was due to accelerated deadenylation in the poly A tail of *hunchback* mRNA. The molecular mechanism of Nanos action on the NRE has also been studied. Pumilio has the Pum RNA-binding domain which provides the binding site to NRE. And Nanos is recruited by Pumilio into an ternary complex (Sonoda and Wharton, 1999).

The collaboration of Nanos and Pumilio on the translational repression is required for the germline development as well as for the embryonic patterning (Forbes and Lehmann, 1998; Parisi and Lin,

2000). Albeit the phenotypes of germline cells in *nanos* mutant and *pumilio* mutant are slightly different from each other (Forbes and Lehmann, 1998), the two proteins have shown to function in parallel (Asaoka-Taniguchi et al., 1999). Recently it was reported that Nanos and Pumilio were involved in the dendrite morphogenesis of peripheral neurons in *Drosophila* (Ye, et al., 2004). However, this control is independent from *hunchback*. In addition to *hunchback* mRNA, *Cyclin B* mRNA in the pole cells is regulated by Nanos-Pumilio collaboration (Asaoka-Taniguchi et al., 1999; Deshpande et al., 1999).

Apart from Pumilio, two *Drosophila* proteins have been reported separately to interact with Nanos. One is Brain tumor (Brat) , a growth repressor in the larval brain belonging to NHL family proteins (Sonoda and Wharton, 2001). It was identified by 4-hybrid screen with Pumilio, Nanos and *hunchback* mRNA. Both Brat-Nanos and Brat-Pumilio interactions are essential to abdominal segmentation. The other protein is Cup. Nanos protein associates with Cup via subdomains in its N-terminal region. The co-localization of Nanos and Cup is detected in the cytoplasm of germline stem cells, cystoblasts and cysts. In this case, Cup plays an inhibitory role on Nanos function in germline cells (Verrotti and Wharton, 2000).

### **Biochemical features of Nanos**

Nanos protein comprises 401 amino acids. Analyses on *nanos* mutants indicated that a small C-terminal region (from C319 to C370) was required for Nanos functions. This region contains two motifs of cystein-cystein-histidine-cystein (CCHC) type Zinc finger (Curtis et al., 1997) highly conserved throughout evolution (described later). Each

finger binds to one Zinc ion and the two fingers are separated by a spacer region of 7 amino acids. It was also shown that this conserved domain was required for all the functions of *nanos*: abdominal formation (*via* the translational repression of *hunchback* and *bicoid*), oogenesis, germ cell migration and germline development. The tail region of Nanos is poorly conserved and required only for germ cell migration and translational repression (Arrizabalaga and Lehmann, 1999).

### **Posterior localization of *nanos* mRNA**

In *Drosophila*, localization of *nanos* mRNA at posterior pole of eggs is a crucial step for the normal oogenesis and the anterior-posterior organization. The mechanism of this localization is mediated by the 3' untranslated region of *nanos* mRNA (Gavis and Lehmann, 1992). Genetic analyses demonstrated that three posterior group genes participate in the localization process. Mutants for *vasa*, *oskar* and *tudor* do not show the posterior localization in eggs and early embryos (Wang et al., 1994; Gavis and Lehmann, 1994). The localization takes part in the gradient of Nanos protein in the embryo. Maternal *nanos* mRNA is synthesized by nurse cells. During the late oogenesis, *nanos* mRNA is infused by nurse cell dumping, actin-dependent contraction of nurse cells. The mRNA diffuses passively in the ooplasm and is entrapped at the posterior region (Forrest and Gavis, 2003). The signal element for *nanos* localization has already been identified in the 3' untranslated region that is also conserved in that of *Drosophila virilis* (Gavis et al, 1996a) and a 75kDa protein is supposed to be the factor which binds to a subregion in the localization signal (Bergsten et al., 2001).

## Regulatory mechanism of *nanos* translation

While the *nanos* mRNA is localized at the posterior pole of *Drosophila* embryo and, as a consequence, the activity of Nanos protein is at the highest level therein, translation of unlocalized RNA is repressed (Gavis and Lehmann, 1994). The correct spatial distribution of Nanos is mainly due to this regulatory mechanism on translation. The repression is known to be mediated by TCE (translational control element), a *cis*-acting element of 90 nucleotides in the 3' untranslated region of *nanos* mRNA (Dahanukar and Wharton, 1996). The downstream region of TCE participates in the mRNA localization (Bergsten and Gavis, 1999).

The TCE binds a 135kDa embryonic protein, Smaug (Smibert et al., 1996) which represses the translation of *nanos* by still unknown mechanism (Clark et al., 2000). The fine structure of TCE and its relationship with Smaug interaction has already been studied (Gavis et al., 1996b; Cruce et al., 2000; Aviv et al., 2003). The secondary structure of TCEs of *Drosophila melanogaster* and *Drosophila virilis* reveals that they have the three stem-loop structure in common. According to the detailed mutation analyses, the second stem-loop interacts with Smaug protein and contributes to the translational repression. The third stem-loop also takes part in the repression but it does not bind Smaug. Instead, it is suggested that the third stem-loop interacts with a novel RNA-binding protein. TCE is weakly involved in the mRNA localization.

The RNA-binding domain of Smaug protein, a 19.2kDa fragment, was



crystallized for structural analysis (Green et al., 2002). It has two characteristic features. First, the amino acid sequence doesn't have homology to any known RNA-binding domain. Second, it has an  $\alpha$ -helical structure that is unusual for RNA-binding domains. Further analyses demonstrated that Smaug RNA-binding domains was composed of 13 helices which formed two subdomains: helix 1 to helix 4 form SAM domain and helix 7 to helix 13 form PHAT domain (Green et al., 2003). The SAM domain is globular in shape and has the RNA-binding surface. This domain actually interacts with TCE in *nanos* mRNA. On the other hand, the PHAT domain does not bind TCE but supports the interaction between SAM and TCE. Recently, the SAM domain of *Saccharomyces cerevisiae* Vts1 was shown to bind TCE with 1 : 1 stoichiometry leading to RNA degradation *via* the activity of yeast cytoplasmic deadenylase CCR4 (Aviv et al., 2003). This observation provides analogy for the mechanism of *nanos* translational repression controlled by Smaug.

### **Evolutionary conservation of *nanos***

*nanos* is highly conserved throughout metazoan evolution. The homologs were found in all metazoan phyla studied so far. The only common feature in Nanos molecules is the Zinc finger domain. Otherwise the overall similarity of those amino acid sequences are pretty low : only 19% between dipteran Nanos (Curtis et al., 1995).

### ***nanos* homologs in *C. elegans* (Nematoda)**

In Nematoda, *Caenorhabditis elegans*, three *nanos* orthologues have been identified (Subramaniam and Seydoux, 1999; Kraemer et al.,

1999). Nanos-1 and Nanos-2 function somewhat redundantly. They have nothing to do with the body patterning during development. They are not required for germline fate, whereas both play a crucial role in maintaining germ cells and preventing primordial germ cells from dividing. In the double-mutant for *nanos-1* and *nanos-2*, up to 99% of germ cells die in apoptotic or non-apoptotic manners. *nanos-2* is necessary for the incorporation of primordial germ cells into gonads (Subramaniam and Seydoux, 1999). *Caenorhabditis elegans* has a Pumilio homolog, FBF. However, there has been no evidence of physical interaction between FBF and either Nanos-1 or Nanos-2.

In comparison with Nanos-1 and Nanos-2, Nanos-3 protein has a much longer N-terminal region. Nanos-3 interacts with FBF via a binding domain in this region (Kraemer et al., 1999). FBF belongs to a Puf family RNA-binding proteins, one of which is *Drosophila* Pumilio. Coupled with Nanos-3, FBF binds to the 3'-untranslated region of *fem-3* mRNA and represses the translation in hermaphrodite worms. Consequently, the spermatogenesis is switched to oogenesis (Zhang et al., 1997; Barton et al., 1987; Ahringer and Kimble, 1991). The collaboration of Nanos and Pumilio in *Drosophila* is also conserved in *Caenorhabditis*. However, while the coupling of Pumilio and Nanos is dependent on the mRNA, Nanos-3 and FBF are able to associate without mRNA (Kraemer et al., 1999). Recently, the important but complicated role of Nanos-3 in germline development was reported (Hansen et al., 2003). It was demonstrated that *nanos-3* promoted the accumulation of GLD-1 protein (an RNA-binding protein with KH domain) in the germ cell and an increased level of GLD-1 promoted the entry into meiosis. Surprisingly, FBF is necessary for the germ cell proliferation and it inhibits the accumulation of GLD-1 in the distal

end of germline by repressing the translation of *gld-1* mRNA (Crittenden et al., 2002). These observations mean that Nanos-3 and FBF function cooperatively and antagonistically for the spermatogenesis switch and germline development, respectively.

### ***nanos* homologs in leech, *Helobdella robusta* (Annelida)**

A *nanos* homolog, *Hro-nos* was isolated for *Helobdella robusta* (Pilon and Weisblat, 1997). The maternal transcripts of *Hro-nos* are detected uniformly in zygotes. In 4 cell embryo, they move toward the animal pole in D quadrant and after the next cleavage, they are more abundant in the DNOPQ cell than in the DM cell. The DNOPQ cell is the ectodermal precursor, whereas the DM cell is the precursor cell for mesoderm. The maternal *Hro-nos* mRNA decays rapidly as the development progresses and in the later stage, the zygotic transcripts increase in the M-lineage teloblasts and subsequently show a wide distribution in the germinal plate. In the next stage, one pair of *Hro-nos* positive spots appears in each of 11 mid-body segments. The cells expressing *Hro-nos* are derived from mesoderm and correspond to primordial germ cells. HRO-NOS proteins is supposed to function as a translational repressor because in the germinal plate in which the *Hro-nos* is expressed, LZF2 protein, the leech homolog of Hunchback is not detectable (Kang et al., 2002).

### ***nanos* homologs in vertebrates (Chordata)**

In zebra fish, a *nanos*-related gene, *nanos1* is essential for survival and proper migration of the primordial germ cells. While the *nanos1* transcripts are stabilized in primordial germ cells, they undergo rapid

degradation outside these cells (Köprunner et al. 2001).

The maternally supplied *Xcat-2* (a *nanos* homolog of *Xenopus laevis*) mRNA is detected in the vegetal cortex of *Xenopus* oocytes (Mosquera et al., 1993). *Xcat-2* protein physically interacts with *Xenopus Pumilio* which is expressed in the oocytes as well as *Xcat-2*. Moreover, *Pumilio* was shown to bind to an element in the 3' untranslated region of *cyclin B1* mRNA and CPEB (cytoplasmic polyadenylation element binding protein) which is a key regulator of translation in oocytes. In spite of these observations, the collaboration of *Xcat-2* and *Pumilio* for the translational regulation in *Xenopus* oocytes remains unclear (Nakahata et al., 2001).

Three genes have been identified as *nanos* homologs in mammalian (human and mouse). Human *nanos1* was cloned from the testis cDNA library (Jaruzelska et al., 2003). *Nanos1* protein forms a complex with human *Pumilio2* via the conserved Zinc finger domain and Puf domain in an RNA-independent manner. Both proteins are abundantly expressed in the germline stem cells (Moore et al., 2003). In addition to *nanos1*, two distinct *nanos* homologs have been found in the human genome sequence (Tsuda et al., 2003).

In mice, *nanos1* was detected not only in mature gonads of both sexes and oocytes but also in the central nervous system, predominantly in the hippocampal formation (Haraguchi et al., 2003). No expression was detected in the primordial germ cells. The knock out study didn't show any phenotypic aberration. The knock out mice were fertile and normal for learning ability. As for mouse *nanos2*, the expression is observed in developing male germ cells. And in knock out mice for

*nanos2*, spermatogonia were eliminated by apoptotic mechanism. *nanos3* is expressed in migrating germ cells in both sexes and male gonads. Germ cell loss was observed in the knock out mice for *nanos2* and *nanos3* (Tsuda et al., 2003).

### Two *nanos* homologs in Hydrozoa (Cnidaria)

The cloning of Cnidarian *nanos* genes was first reported for *Hydra magnipapillata* (Mochizuki et al., 2000). *Hydra* has two *nanos* orthologs, *Cnnos1(Hm)* and *Cnnos2(Hm)*. *Cnnos1* mRNA was specifically expressed in the germline cells and multipotent stem cell, but not in somatic cells. On the other hand, *Cnnos2* was detected in not only in the germline cells but in the endoderm of hypostome. By introducing *Cnnos1*, the defect in egg production of *Drosophila* was partially rescued in *nanos* mutant flies.

Recently, orthologs for *Cnnos1(Hm)* and *Cnnos2(Hm)* were identified in *Podocoryne carnea* (*Pcnos1* and *Pcnos2*, respectively) and reported by Torras (Torras et al., 2004). By comparing the amino acid sequences, they found a novel conserved domain in the N-terminal regions of Nanos proteins in those of sponge, Cnidaria and vertebrate. The domain is not found for protostomes. In this biphasic animal, both *nanos* homologs are expressed only in medusae and early embryos: from the early buds to mature medusae. In later stages, the expression is restricted in the gonads of both sexes. This expression pattern is more or less similar to that of *Helobdella robusta* (Kang et al., 2002). In early embryo, *Pcnos1* and *Pcnos2* are expressed at the posterior pole of blastula and gastrula and disappear later. No expression was detected in the invaginating cells in gastrula. They

discussed that this observation indicated the *nanos* involvement in formation of axial polarity in Cnidarian embryos and that the process of *nanos* expression was not related to the germline establishment.

### **Strategy for the research on *Cnnos* and germline speciation**

In *Drosophila*, a single *nanos* gene has been known to have two distinct functional features, the abdominal determination and germline differentiation and maintenance. Cnidaria has two *nanos* homologues. It is possible that the two genes evolved with distinct functional constraints. To study this possibility, I compared the primary structure, the structures of the upstream regulatory regions and the expression of *Cnnos1* and *2* of *Acropora millepora* and *Nematostella vectensis*.

## Section 2. Materials and Methods

### ***E. coli* strains**

See Chapter 1, Section 2.

### **Manipulation of bacteria**

See Chapter 1, Section 2.

### **Manipulation of $\lambda$ -bacteriophage**

See Chapter 1, Section 2.

### **Manipulation of nucleic acids and protein**

See Chapter 1, Section 2.

### **Enzymes for molecular biology**

See Chapter 1, Section 2.

### **Primers**

See Chapter 1, Section 2.

### **Animal sampling**

See Chapter 1, Section 2.

### **Screening *Acropora millepora* genomic library**

$\lambda$ -GEM 12 *Acropora millepora* genomic DNA library was constructed by Dr. David Hayward (Australian National University). For screening for *Cnnos1(Am)* and *Cnnos2(Am)* genes, approximately 560,000 and 280,000 clones were screened, respectively, by the standard plaque

hybridization in the primary screening. Hybond-N nylon membranes were purchased from Amersham Biosciences (Piscataway, NJ) and used for plaque lift and hybridization. The probes were radiolabelled with  $\alpha$ -<sup>32</sup>P-dATP by PCR. Primers for the PCR were designed to amplify DNA fragments of 162bp corresponding to the conserved Zinc finger domains of Nanos proteins.

For *Cnnos1(Am)* cloning, a DNA fragment for the Zinc finger was amplified by PCR from the viral suspension of *Acropora millepora* genomic library and purified. This fragment was used as template. The primer pairs for this amplification and radiolabelling was:

Cnnos1(FW) :5'-CGTGCGTGTTTTGTAGGAACAA-3'

Cnnos1(RV) :5'-CGGGACAATACTTAATGGTATG-3'

These sequences were based on the cDNA sequence of *Cnnos1(Ad)*, a nanos gene of *Acropora digitifera* (Dr. T. Fujisawa, National Institute of Genetics, Japan: personal communication). The thermal cycling protocol was 30 cycles of 94°C for 45sec., 60°C for 45sec. and 72°C for 1min. and 1cycle of 94°C for 1min., 55°C for 30sec, and 72°C for 1min.

For *Cnnos2(Am)* cloning, the 3' portion of *Cnnos2(Am)* cDNA was amplified by PCR from the viral suspension of *Acropora millepora* cDNA library (pre-settlement stage) with Cnnos1(FW) primer and T3 primer. The amplified product was purified and sequenced to confirm that it corresponded to *Cnnos2(Am)*. This fragment was ligated into pGEM T-Easy (Promega, Madison, WI) and used for the template. PCR-based radiolabelling was performed with the same primer pair and PCR protocol as those for *Cnnos1(Am)* cloning mentioned above.



The labelled probe was purified with Ultra Clean PCR Clean-up / DNA Purification Kit (GeneWorks, Adelaide, SA). The hybridization was performed at 65°C overnight in 5X SSC, 5X Denhardt's solution, 0.5% SDS. The hybridized membranes were washed according to the manufacturer's instruction: in 500ml of 2X SSC / 0.5% SDS at room temperature for 5min., in 500ml of 2X SSC / 0.1% SDS at 65°C for 30min., in 500ml of 1X SSC / 0.1% SDS at 65°C for 30min. and 0.1X SSC / 0.1% SDS at 65°C for 30min. The membranes were air-dried and the radioactive signals were exposed on Phosphoscreen and visualized with Phosphorimager (Molecular Dynamics, Sunnyvale, CA). The areas on the agar plates with promising signals on the phosphorimage were picked up and stored in SM buffer at 4°C.

The viral samples picked-up in the primary screening were further screened by PCR. With the viral suspension as the template, the Zinc finger region of cDNA was amplified. The same primers and protocol as in the primary screening were applied again. The positive samples for PCR products were brought to the secondary screening by plaque hybridization.

### **Screening *Nematostella vectensis* genomic library**

The *Nematostella vectensis* genomic library was provided by Dr. John Finnerty (Boston University, Boston, MA). The genomic DNA screening was carried out by plaque hybridization. Before the primary screening, cDNA fragments encoding the conserved zinc finger domains for *Cnnos1* and *Cnnos2* were amplified by PCR and cloned in pGEM T-Easy vector. The inserts were sequenced and clones for *Cnnos1* and *Cnnos2* were identified. The primer pair for this

amplification was Cnnos1(FW) and Cnnos1(RV) for the *Cnnos1*-specific probe which were used for the genomic cloning of *Acropora millepora* and those shown below for the *Cnnos2*-specific probe.

nos2F1 : 5'-TGCGTGTTCTGCAGGAACAA-3'

nos2R1 : 5'-GGGACAATACTTTATGGTGTG-3'

The thermal cycling protocol was: 30 cycles at 95°C for 1min., 50°C for 1min. and 72°C for 30sec. followed by 1 cycle a 72°C for 2min.

The *Cnnos1(Nv)*- and *Cnnos2(Nv)*-specific fragments were radiolabelled for the hybridization for *Cnnos1(Nv)* and *Cnnos2(Nv)*, respectively. Labelling reaction was performed with the protocol: 30 cycles at 94°C for 45sec., 60°C for 45sec. and 62°C for 1min. followed by 1 cycle at 94°C for 1min., 55°C for 30sec. and 72°C for 1min.

Approximately 500,000 clones were screened in the primary screening. The hybridization and membrane washing were same as described for *Acropora millepora* genomic screening. PCR-based screening and the secondary screening were also done in the same way.

### **Sequencing lambda phage DNA**

The lambda phage DNA of the isolated clone was extracted with QIAGEN Lambda System (QIAGEN, Valencia, CA) according to the supplier's manual. The lambda phage DNA was sequenced by primer walking. The sequencing reaction was performed with BigDye Terminator version 3.1 (Applied Biosystems, Piscataway, NJ). The sequence data were analyzed and processed with Sequencher version 3.0 (Gene Code Corporation). Sequence homology search for the possible translation was carried out using the blast-n program in the BLAST server at the National Center for Biotechnology Information.

### **Screening *Acropora millepora* cDNA library**

$\lambda$ -UniZap XR *Acropora millepora* cDNA library from adult coral tip (sampled in October) was constructed by Dr. David Hayward (Australian National University). Approximately 1200,000 clones were screened by the standard plaque hybridization in the primary screening for *Cnnos1*(*Am*) cDNA cloning. Plaque hybridization was carried out with the same method described in the genomic cloning.

Phagemid vectors, pBluescript SK(-), which harbored the cloned cDNAs were recircularized by the excision protocol with ExAssist Interference-Resistant Helper Phage (Stratagene, La Jolla, CA) according to the manufacturer's instruction. The cloned cDNA was sequenced by primer walking with BigDye Terminator ver.3.1 (Applied Biosystems, San Francisco, CA). Sequence data were processed and analyzed as described for the genomic cloning.

*Cnnos2*(*Am*) cDNA was cloned by Karin Nördstrom in our laboratory.

### **Screening *Nematostella vectensis* cDNA library**

$\lambda$ -ZipLox *Nematostella vectensis* cDNA library constructed from all developmental stages of the organism was a gift from Dr. John Finnerty (Boston University, Boston, MA). Approximately 500,000 clones were screened. The same probes as those for cloning genomic DNA of *Cnnos1*(*Nv*) and *Cnnos2*(*Nv*) were used for this screening.  $\lambda$ -DNA was extracted from the cloned phages and sequenced as described above.

## RT-PCR

5µg of total RNA (for extraction protocol, see Chapter 1, Section 2) was digested with Deoxyribonuclease I, Amplification Grade (Invitrogen, San Diego, CA). RNA was mixed with Reaction buffer provided by the manufacturer and 5 units (1 unit for 1µg of RNA) of enzyme. After 15min. of incubation, the reaction was stopped by adding EDTA with the final concentration at 2.5mM and the enzyme was heat-inactivated at 65°C for 10min. The RNA was recovered by ethanol precipitation and redissolved in an appropriate volume of RNase-free water.

mRNA was reverse-transcribed for the 1st strand cDNA synthesis by using First-strand cDNA Synthesis Kit (Amersham Biosciences, Piscataway, NJ). The reaction was primed by NotI-d(T)<sub>18</sub> primer. An amount of the reaction that was equivalent to 5µg of RNA in the DNaseI digestion was brought to the subsequent PCR amplification.

For the amplification, the primer pair was:

*Cnnos1(Am)*

nos1F1 : 5'-TGCGTGTTTTGTAGGAACAA-3'

1R3 : 5'-TGGACATGACTAACAAGCAC-3'

*Cnnos2(Am)*

2F3/bf : 5'-TAAGCTTAGTGATGGCTGAA-3'

Cnnos2Rev : 5'-GGCAGCATTAAACCTACTT-3'

The thermal cycling consisted of 40 cycles at 95°C for 1min., 50°C for 30sec. and 72°C for 45sec. and an additional extension at 72°C for

2min.

After the Deoxyribonuclease I digestion, a part of each RNA sample was subjected to PCR amplification without reverse transcription. The result provided the negative control for the genomic DNA contamination.

To check the overall process, a partial cDNA of coral EF1 $\alpha$  was amplified by PCR. The primer pairs were:

EF(FW) : 5'-TGCCTGGGTTCTTGACAAACT-3'

EF(RV) : 5'-GGGATGGTATCCCACTTTCTG-3'

The amplification protocol consisted of 30 cycles of 95°C for 1min., 50°C for 30sec. and 72°C for 2min. followed by an additional extension step at 72°C for 4min.

### **Inverted PCR**

500ng of  $\lambda$ -phage DNA for *Cnnos1(Nv)* genomic DNA was digested with various restriction enzymes. The fragments of this  $\lambda$ -phage clone were circularized by self-ligation. An appropriate volume of the reaction which was equivalent to 50ng of DNA was used as the template for inverted PCR. The protocol of thermal cycling consisted of 35 cycles at 95°C for 45sec., 50°C for 30sec. and 72°C for 4min. followed by 5min. extension at 72°C. The combinations of restriction enzymes and primer pairs are shown below:

SacI n1F1 : 5'-TTCAAACCGCGGTGTGGTT-3'

n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'

XhoI n1F1 : 5'-TTCAAACCGCGGTGTGGTT-3'

	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
BamHI	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
ECoRI	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
KpnI	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
PstI	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
Sall	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'
XbaI	n1F5 : 5'-CTGCTTGTAAGGCATGTGTC-3'
	n1R5 : 5'-TCACTCTTGCATCAAGTCCC-3'

For *Cnnos2(Nv)* genomic DNA, 50ng of the cloned  $\lambda$ -phage DNA was digested. The PCR protocol was 30 cycles at 95°C for 1min., 50°C for 30sec. and 72°C for 4min. followed by one cycle at 72°C for 5min. The combinations of restriction enzymes and primer pairs are shown below.

PstI	n2F1 : 5'-ATTCGGCGAGGGATCATTTTC-3'
	n2R5 : 5'-AGCCGTGATCAGAGGCGTAA-3'
XbaI	n2F2 : 5'-GAACGATCAAGTGATCGCAG-3'
	n2R5 : 5'-AGCCGTGATCAGAGGCGTAA-3'
BamHI	n2F3 : 5'-GGCAGTAATCCTTCGTATCAT-3'
	n2R5 : 5'-AGCCGTGATCAGAGGCGTAA-3'
ECoRI	n2F3 : 5'-GGCAGTAATCCTTCGTATCAT-3'
	n2R5 : 5'-AGCCGTGATCAGAGGCGTAA-3'

The products of PCR amplification were separated by agarose gel electrophoresis and extracted. These DNA fragments were sequenced with the same primer pairs as those for the amplification. The sequence data were analyzed and aligned as mentioned before.

## Section 3. Results

For understanding the evolutionary conservation of the function of cnidarian nanos proteins and the functional speciation of two nanos homologues in Cnidaria, I carried out cDNA isolation for *Cnnos1* of *Acropora millepora* and *Cnnos1* and *Cnnos2* of *Nematostella vectensis*. *Cnnos2*(*Am*) had been cloned and identified previously in our laboratory (unpublished data).

### ***Cnnos1* cDNA isolation for *Acropora millepora***

First, I attempted to isolate the cDNA for *Cnnos1*(*Am*) from the cDNA library which was constructed for a late embryonal stage, pre-settlement. The cloning was unsuccessful probably due to the poor expression of *Cnnos1*(*Am*) gene in this stage. I thought the sexually mature coral could express this gene because in hydra, *Cnnos1* was expressed in germline cells (Mochizuki et al. 2000). Therefore I tried to isolate *Cnnos1*(*Am*) from the cDNA library for the adult coral branch tips sampled in October, the season for sexual maturation. Approximately 1,200,000 clones of UniZap XR cDNA library were screened in the primary screening with the standard plaque hybridization technique. A DNA fragment corresponding to the highly conserved Zinc finger domain was amplified and radiolabelled by PCR for the probe preparation. 15 signals were detected and picked up for further screening. 12 out of the 15 samples showed the positive study for PCR-amplification of the DNA fragment used for the plaque hybridization. The secondary screening was carried out on these phage samples. Positive and well-isolated signals were obtained for 8



samples. To isolate the clones with the entire open reading frame, the length of 5'-extension of these clones was checked by PCR with an internal and an external (T7 primer) primer. Finally, a cDNA clone which showed the longest 5'-extension was rescued by excision protocol with the helper phage. The inserted cDNA was sequenced by primer walk and subcloning. The nucleotide sequence was determined by reading both strands.

Figure 43 shows the nucleotide sequence and deduced amino acid sequence of *Cnnos1(Am)*. The cDNA is 1685bp long in total without poly A tail. The 5'- and 3'-untranslated regions are 300bp and 705bp, respectively. A typical polyadenylation signal, AATAAA, appears at 21 bases upstream from the start site of poly A tail. The translated protein comprises 227 amino acids and contains the Zinc finger RNA binding domain in the C-terminal region. The homology search for the Zinc finger domain by blast-p program detected more than 40 homologs with high and moderate homology. Among them, cnidarian Nanos proteins have high homology ; *Cnnos1(Pc)*, *Cnnos1(Hm)*, *Cnnos2(Pc)* and *Cnnos2(Hm)* show 84%, 83%, 81% and 79% identity, respectively (*Pc* : *Podocoryne carnea*, *Hm* : *Hydra magnipapillata*). In addition, the sponge Nanos, *PoNOS* shows 75% identity. In contrast, outside the RNA-binding Zinc finger, *Cnnos1(Am)* are not conserved. This finding is the common feature of Nanos family proteins.

### ***Cnnos1* cDNA isolation for *Nematostella vectensis***

For the better insight into the divergence of Nanos proteins in Cnidaria, cDNA cloning of *nanos* homologs was also tried for *Nematostella vectensis*. Approximately 500,000 clones of  $\lambda$ -Ziplox

cDNA library were screened in the primary screening with the standard plaque hybridization technique. A DNA fragment corresponding to the highly conserved Zinc finger domain was amplified and radiolabelled by PCR. 22 signals were detected and 4 of them were chosen for further screening. 3 out of the 4 samples showed the positive study for PCR-amplification of the DNA fragment used for the plaque hybridization. One of them showed the longest PCR product on the 5'-extension. The secondary hybridization was carried out on this phage sample and well-isolated plaques with positive hybridization signals were picked up.  $\lambda$ -phage DNA was extracted from the phage clone and subjected to sequencing. The entire insert of this clone was sequenced by primer walk. The sequence was determined by reading both strands.

Figure 44 (panel A) shows the nucleotide sequence and deduced amino acid sequence of *Cnnos1(Nv)*. The cDNA is 1886bp in total without poly A tail. The 5'- and 3'-untranslated regions are 190bp and 1048bp, respectively. A typical polyadenylation signal, AATAAA, appears at 22 base upstream from the start site of poly A tail. The translated protein comprises 216 amino acids and contains the Zinc finger RNA binding domain in the C-terminal region. The homology search for the Zinc finger domain by blast-p program detected more than 40 homologs with high and moderate homology. Among them, cnidarian Nanos proteins have high homology ; *Pcnos1*, *Cnnos1(Hm)*, *Pcnos2* and *Cnnos2(Hm)* show 85%, 83%, 80% and 78% identity, respectively. In addition, the sponge Nanos, *PoNOS* shows 74% identity. As well as in *Acropora millepora*, other regions in the protein does not show significant homology.

## ***Nematostella vectensis* also has the second Nanos homolog**

The existence of the Cnnos2 homolog for sea anemone was identified when DNA fragments for the Zinc finger domains specific to Cnnos1 and Cnnos2 of *Nematostella vectensis* were obtained by PCR amplification with the cDNA library as the template. A cDNA clone of *Cnnos2(Nv)* was successfully isolated from the same cDNA library. The library screening was carried out in almost the same manner as that for the *Cnnos1(Nv)* cloning. The primary plaque hybridization showed 81 signals in total. 19 promising signals were picked up and 6 of them were detected positive by PCR. After the selection by 5'-extension, one clone was finally isolated by secondary plaque hybridization. The entire insert of this clone was sequenced by primer walk. As my routine, the sequence was determined by reading both strands.

Figure 44 (panel B) shows the nucleotide sequence and deduced amino acid sequence of *Cnnos2(Nv)*. The cDNA is 1852bp in total without poly A tail. The 5'- and 3'-untranslated regions are 158bp and 950bp, respectively. An atypical polyadenylation signal, ATTAAA, appears at 18 bases upstream from the start site of poly A tail. The translated protein comprises 247 amino acids and contains the Zinc finger RNA binding domain in the C-terminal region. The homology to cnidarian proteins were 85%, 85%, 75% and 72% to Cnnos2(Hm), Pcnos2, Pcnos1 and Cnnos1(Hm), respectively. The PoNOS shows 74% identity. Between *Acropora millepora* and *Nematostella vectensis*, the identity is 96% and 93% for Cnnos1 and Cnnos2, respectively.

## **High conservation in the Zinc finger domains throughout**

## evolution

The amino acid sequences of the Nanos family Zinc finger domains were aligned by ClustalW algorithm. The Boxshade representation is shown in Figure 45. The overall conservation is significantly high through metazoan evolution. It is to be noted that cysteine and histidine residues which take part in the CCHC-CCHC type double Zinc fingers are absolutely conserved except for a few homologs such as Nanos1 of *C. elegans*.

The Neighbour Joining tree (Figure 46) shows that cnidarian Nanos proteins form a clade and indicates that the two cnidarian nanos genes segregated before the class radiation in this phylum. *Cnnos1* and *Cnnos2* form separate subclusters and the branches in this group are congruent with each other.

### Genomic DNA cloning for *Cnnos1(Am)* and *Cnnos2(Am)*

To obtain the information about the genomic organization of cnidarian *nanos* genes and the regulatory mechanism of the gene expression, I isolated the genomic DNA for both *Cnnos1* and *Cnnos2* from *Acropora millepora* genomic DNA library.

Screening was carried out in almost the same method in those for the cDNA cloning. For cloning *Cnnos1(Am)*, 560,000 clones were screened by plaque hybridization *Cnnos1(Am)*-specific probe. and 78 signals were detected and picked up. 74 of them were shown positive for PCR amplification corresponding to the DNA region used for the probe. 10 samples were brought to the secondary hybridization and

finally 7 clones were obtained. For *Cnnos2(Am)*, 280,000 clones were screened and 6 signals were detected and picked up on the primary screening. All of them were shown positive for PCR and the secondary hybridization was performed on them. 2 plaques were picked up as clones.

$\lambda$ -phage DNA was extracted from these clones. The inserts were sequenced by primer walk. The 5'- and 3'-flanking regions were also sequenced. As shown in Figure 47, both genes are intronless. The nucleotide sequences of the transcribed regions are identical to those of the cDNAs. Southern-blot analyses on *Cnnos1(Am)* and *Cnnos2(Am)* detected single copies of these genes with the specific probes (Figure 48). Taken together, it is concluded that these genomic clones represent the functional genes.

### **Genomic DNA cloning for *Cnnos1(Nv)* and *Cnnos2(Nv)***

The genomic DNA for *nanos* genes was isolated from *Nematostella vectensis* genomic DNA library. 500,000 clones were screened for *Cnnos1(Nv)* and *Cnnos2(Nv)* by plaque hybridization with specific probes. After the PCR-based screening and the secondary hybridization, one clone was isolated for each gene.

$\lambda$ -phage DNA was extracted from both clone and sequencing was performed. However, with unknown reason, the sequencing reaction did not perform well. Therefore, I tried to amplify the flanking regions by inverted PCR. The amplification was successful and the products were sequenced.

## **5'-regulatory regions of cnidarian *nanos* genes have multiple potential transcription factor binding sites**

Potential transcription factor binding sites were searched in the 5'-flanking regions of *Cnnos1* and *Cnnos2* genes (*Acropora millepora* and *Nematostella vectensis*) with TRANSFAC data base (Heinemeyer et al., 1998). As shown in Figure 49, they have multiple binding elements. The transcription factors with high scores of probability (>95.0) are shown. It seems that there is no similar feature for the factor binding sites between each other. Therefore it is suggested the expression and regulatory mechanism of two *nanos* homologs is quite different in the same organism.

## **Gene expression patterns of *Cnnos1* and *Cnnos2* in *Acropora millepora***

The gene expression of *Cnnos1*(*Am*) and *Cnnos2*(*Am*) was investigated by RT-PCR (Figure 50). During the development, *Cnnos1*(*Am*) transcript was detected only for the early embryonal stage (prawn-chip). The amplification was not observed for the later embryo (pre-settlement) or for eggs. As for *Cnnos2*(*Am*), all the stages tested were positive for the amplification. For adult colony, *Cnnos1*(*Am*) was not detected either before or after the reproductive season. On the other hand, *Cnnos2*(*Am*) was detected in all three months and all the examined developmental stages. It is suggested that gene expression of the two *nanos* homologs in *Acropora millepora* is controlled by different mechanisms.

## Section 4. Discussion

### ***nanos* gene duplicated before cnidarian radiation**

Cnidaria has two *nanos* homologs, *Cnnos1* and *Cnnos2*. The cDNAs for both *Cnnos* genes have been cloned from Hydrozoa—*Hydra magnipapillata* (Mochizuki et al., 2000) and *Podocoryne carnea* (Torras et al., 2004) and Anthozoa—*Acropora millepora* and *Nematostella vectensis*. The amino acid sequences of Cnnos proteins showed the typical features of Nanos proteins: the highly conserved zinc finger domains in their C-terminal region and poor conservation outside these domains. Very recently, Torres et al. reported the finding of another conserved domain in the N-terminal regions of cnidarian Nanos proteins (Torres et al., 2004). The novel domain was also found in those of *Acropora millepora* and *Nematostella vectensis*. According to their observation, this domain is present in the Nanos proteins of sponge, cnidarians and vertebrates. Although the function and biological significance of this domain are unknown, it provides evidence for the idea that cnidarian and vertebrate genes have a lot in common.

My phylogenetic analysis on Nanos proteins of lower animals indicated that the *nanos* gene underwent duplication before the radiation of phylum Cnidaria. Cnnos1 and Cnnos2 proteins form independent clusters. The trees in both clusters are congruent with each other and absolutely parallel with the currently accepted cnidarian phylogeny (Rupper et al., 2004). Therefore it is speculated that there have been the same functional constraints for Cnnos1 and

Cnnos2.

A single copy of *nanos* homolog has been found for Porifera (*PoNOS*). Although the function of Cnnos and PoNOS proteins are to be investigated, it is likely that the common ancestor to cnidarians had to develop two *nanos* genes in order to acquire novel biological characteristics such as the diploblastic structure, which were not present in Porifera.

### **Comparison of genomic structures of cnidarian *nanos***

The genomic DNA of *Cnnos1*(*Am*) and *Cnnos2*(*Am*) were isolated and characterized. Both are intron-less genes. The nucleotide sequences for the coding regions are identical to those of cDNAs and Southern-blot analyses with the specific probe to each gene detected a single band. Therefore, it is concluded that the genomic clones represent the functional genes for both *nanos* homologs of *Acropora millepora*. The features of the upstream regulatory regions for *Cnnos1* and *Cnnos2* of *Acropora millepora* and *Nematostella vectensis* look different. However, all four regions lack TATA-box and have potential binding sites for CdxA (Caudal-type homeodomain protein) (Frumkin et al., 1991). And they have a number of potential binding sites for a variety of transcription factors including the factors regulating the constitutive transcription such as C/EBP (CCAAT enhancer binding protein) and those for the responsive transcription such as CREB (cAMP-responsive element binding protein). The observation that *Cnnos1* and *Cnnos2* have distinct features in their upstream regulatory regions further supports the speculation that each gene has the differentiated function and the specific constraint in



evolution.

### Different expression patterns of *Cnnos1* and *Cnnos2* in *Acropora millepora*

As expected from the different features of the 5'-regulatory regions of the two *nanos* homologs, the expression patterns of these genes do not look like each other. The expression of *Cnnos1*(*Am*) was only observed in the early embryonal stage and no expression was detected in the adult coral tips. In hydrozoan *Podocoryne carnea*, *Pcnos1* was detected in the early stages of development including eggs and in medusae (Torras et al., 2004). The pattern of *Cnnos1*(*Am*) expression is like that of *Podocoryne* which shows no expression in polyp. For *Cnnos2*(*Am*), the expression was detected in all stages of development (eggs, prawn-chip and pre-settlement) and all samples for the adult colony (September, October and December). This pattern looks quite different from those of *Podocoryne carnea* and *Hydra magnipapillata* (Mochizuki et al., 2000). Information of the spatial patterns of expression is necessary for understanding the function of *cnnos* genes in *Acropora millepora*.

# Figures

Fig.1

A schematic representation of a single coral polyp. The reproductive organs do not appear in coral. Instead, germ cells differentiate in the mesoglea of mesenteries forming temporary gonads.

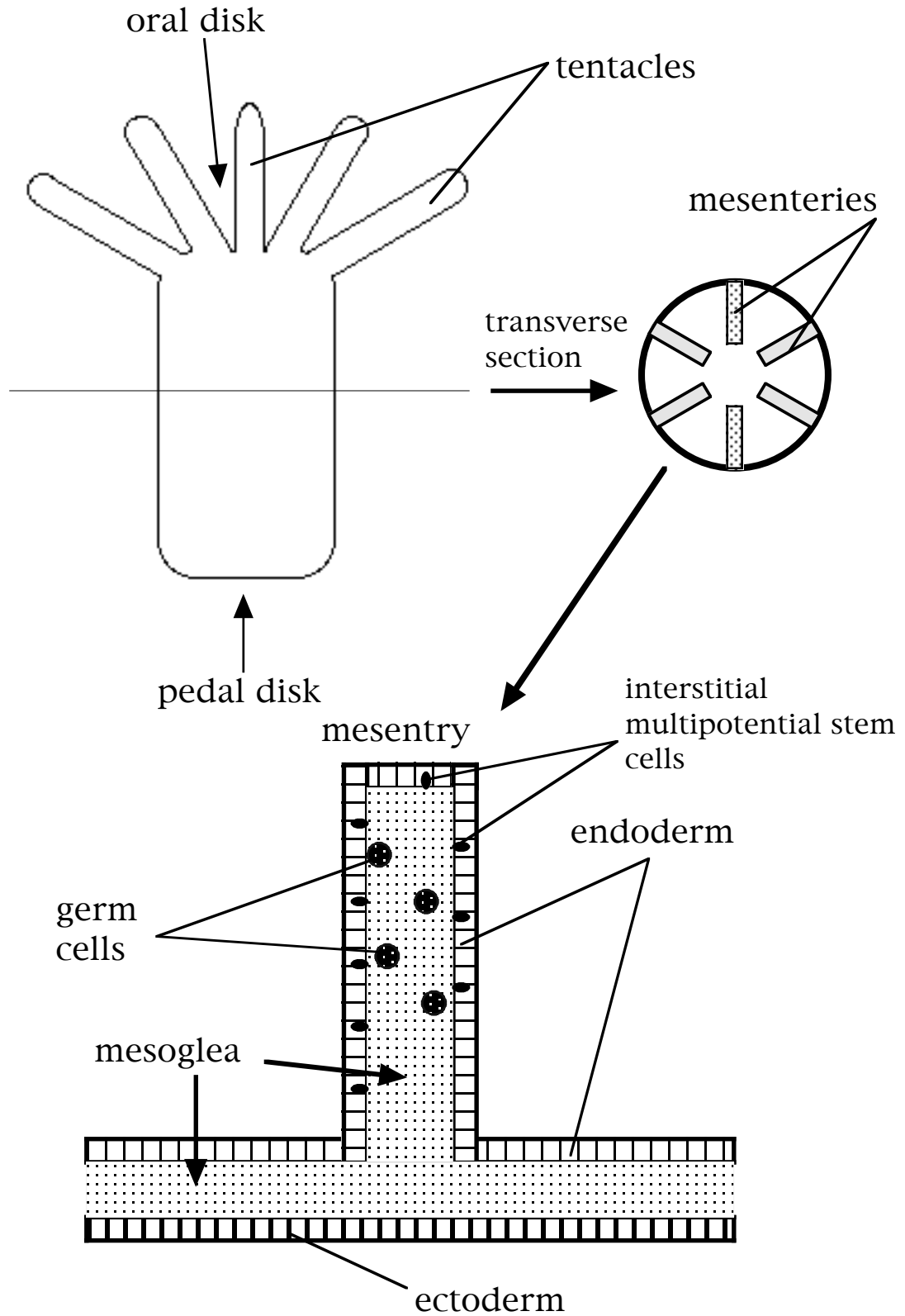


Fig.1 Anatomy of a coral polyp

Fig.2

In *Drosophila*, both male- and female-specific Dsx play roles in the sexually dimorphic development of the genital organs in collaboration with a Hox protein, Abd-B. Wingless pathway is activated in the female A8 resulting in the female gonadal primordia, whereas Decapentaplegic pathway is activated in the male A9 resulting in the gonadal primordia. Wg: Wingless, Dpp: Decapentaplegic, Abd-B: Abdominal-B, Hh: Hedgehog. Squares represent the Abdominal segments. Dotted areas show the A/P organizers. Arrows mean the activating pathways and lines with shot bars represent the inhibiting effects.

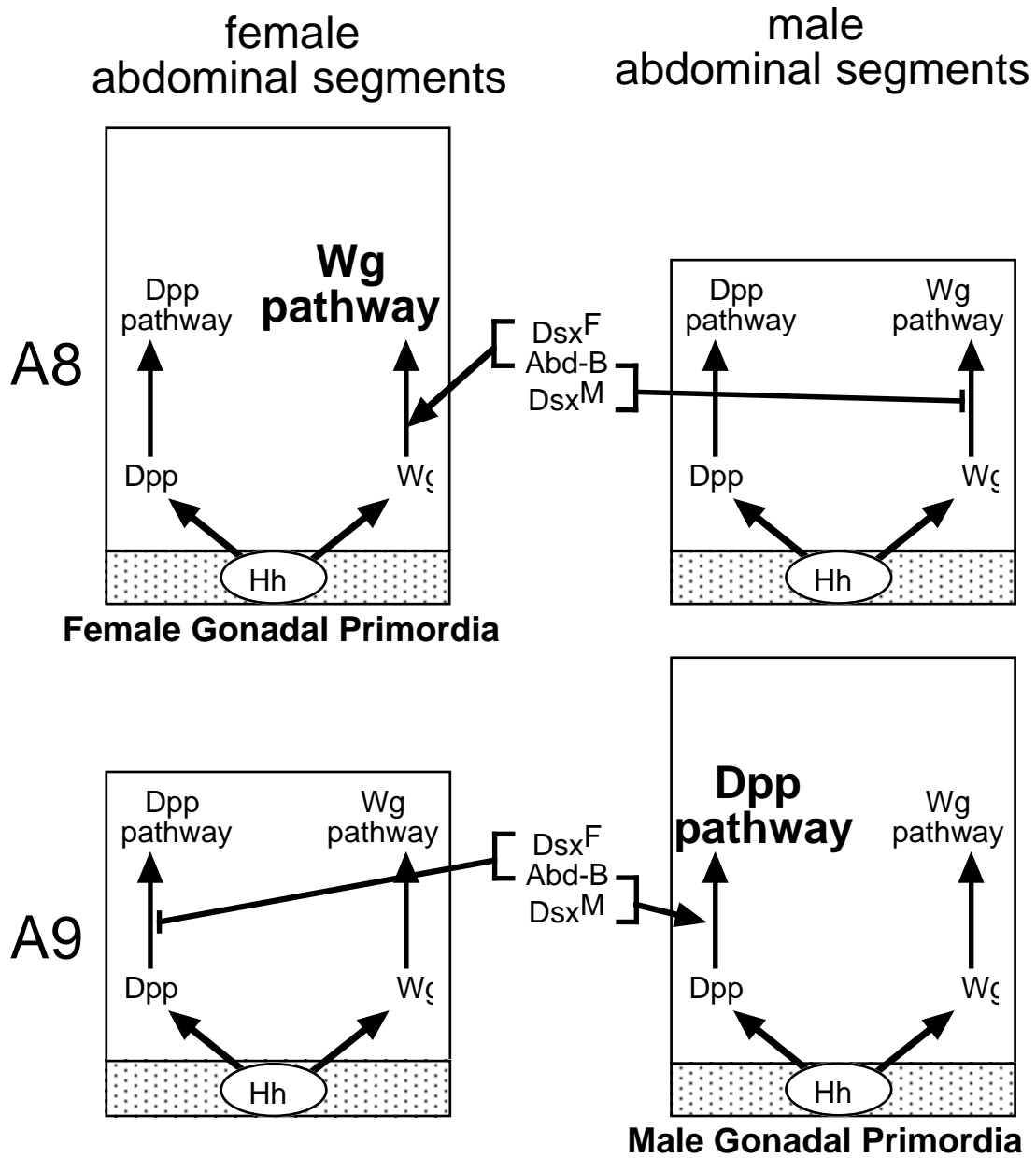


Fig.2 Development of genital organs in *Drosophila*

Fig.3

Strategy for yeast one-hybrid screening to isolate the genomic DNA fragment of *Acropora millepora* with the binding sequence to DM domain. The Activation plasmid produces the fusion protein containing the DNA-binding DM domain and the activation domain of GAL4, which binds to the DM-binding sequence in the DNA fragment inserted in the upstream of *HIS3* gene in the Reporter plasmid. The clone with the Reporter plasmid harboring the DM domain binding sequence is selected on the -Histidine plate. Under the influence of *ADE5* gene, the clone forms a red colony without sectoring.

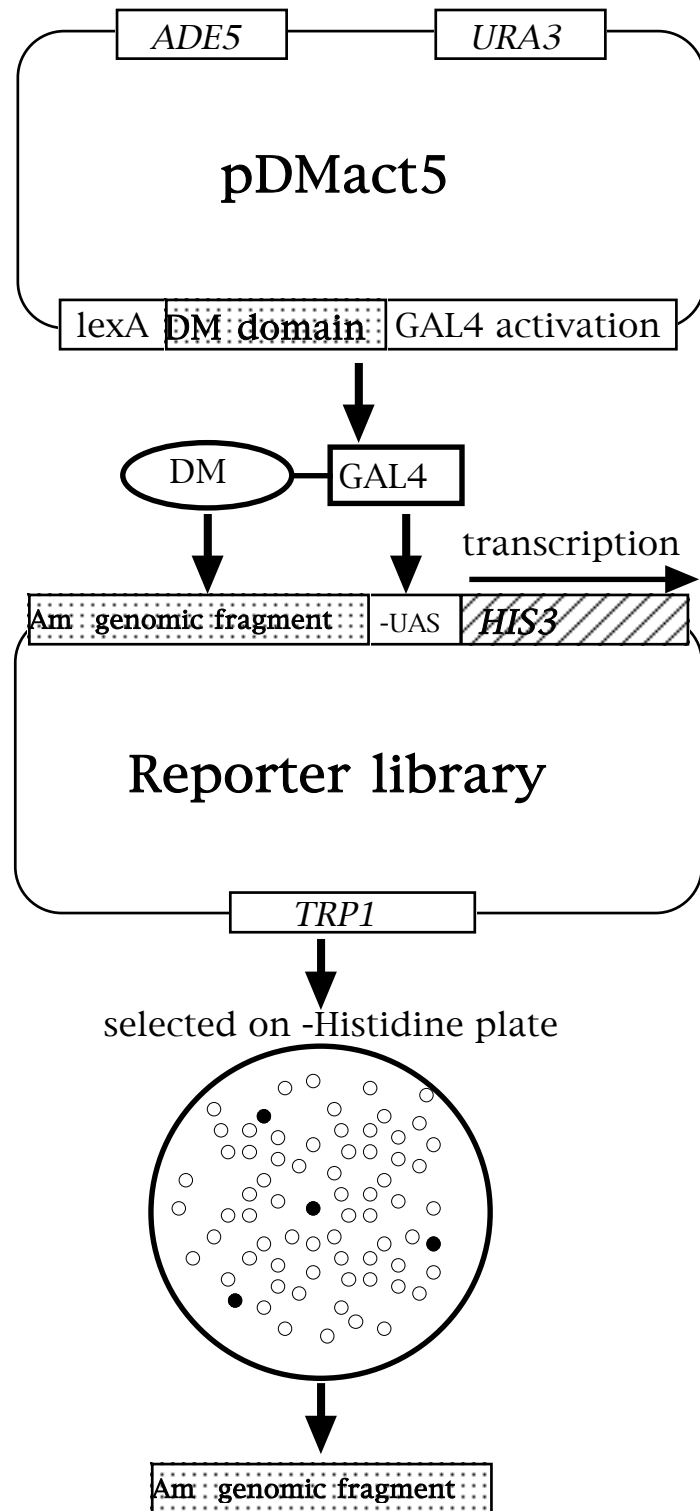


Fig.3 Principles of yeast one-hybrid selection



Fig.4

CAT reporter vectors for detection of the DM enhancer activity.

pCAT3 promoter vector was inserted with the DM enhancer fragment in both forward and reverse orientation.

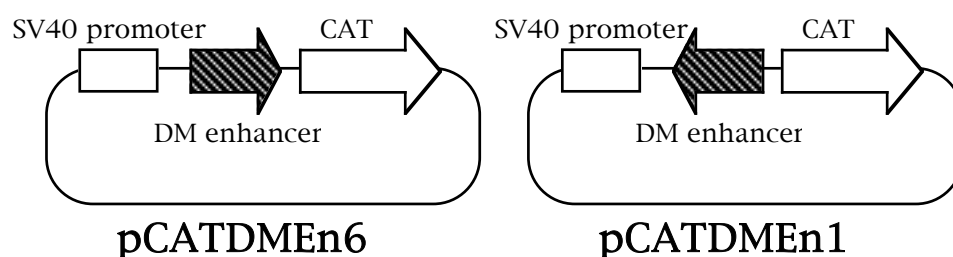


Fig.5

Nested deletion constructs for *AmDM1* expression. Partial cDNAs

corresponding to the deletion mutants were inserted in the eukaryotic expression vector pRc/CMV.

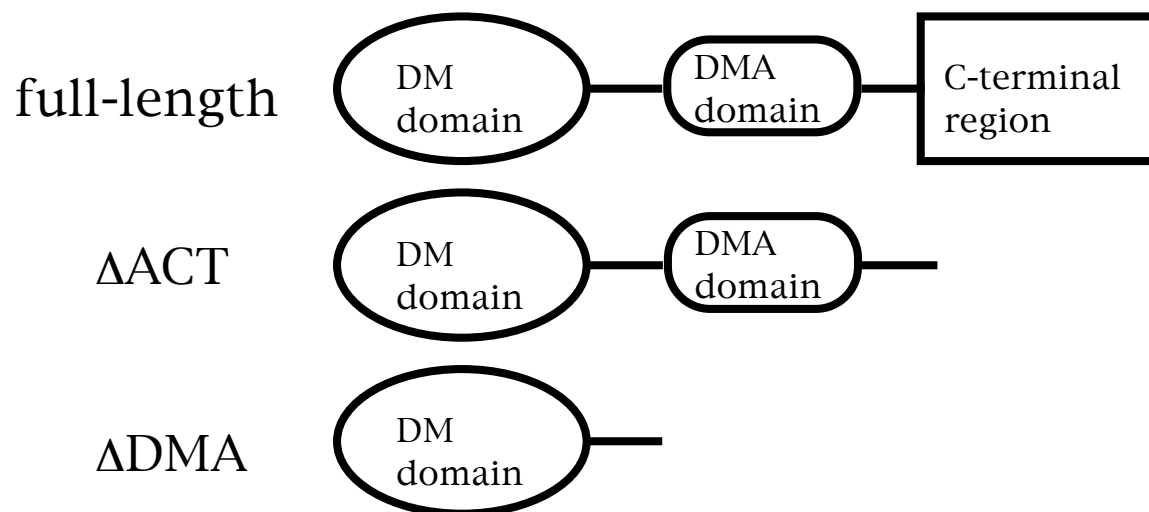


Fig.6

The nucleotide sequence and the deduced amino acid sequence of *NvDM5* cDNA, a *doublesex* homolog of sea anemone *Nematostella vectensis*. The DM domain is double-underlined and the DMA domain is single-underlined. A putative and atypical (TATAAA) polyadenylation signal is written in bold letters.

1 GTCTGGGAGGTAAGCTTGCTCTAGACGGCCTTCTCTGTTAAATGGGCACACATTAACAAGATGTTGCTGTATATAC'  
M F Q D I S P  
 81 GCATCGTAAGGAAATGGAGCCGATACTGATGGAGTGCTAACCCCTAGGGTTTACAATGTTTCAGGACATTTCTCCT.  
 L L M P S T S L S A S A Q T T T T T A Q Q Q A Q Q  
 161 GCTTCTAATGCCTTCGACAAGTCTCTCGGCCTCGGCGCAGACCACTACCACAACGGCCCAACAACAAGCCCAACA  
 A Q Q Q A A A A A A Q Q R S P K C A R C R N H G V  
 241 CGCAACAGCAGGCGGCGGCAGCAGCCGCTCAACAACGCTCCCCAAAATGCGCGAGGTGCCGAAATCACGGCGTCA'  
I L K G H K R F C K W K D C T C P D C N L I A E R  
 321 ATCTAAAAGGCCACAAGCGCTTCTGCAAATGGAAGGACTGCACATGCCCTGACTGCAACCTCATAGCGGAACGG  
V M A A Q V A L R R Q Q E S D A E H F T N F P Y N  
 401 CGTCATGGCTGCGCAGGTAGCGCTGAGAAGACAACAAGAAAGTGACGCAGAGCATTTACCAACTTTCTTTACAA'  
 V Q F Y Y N S S G Q R G Q D G G S S K G P M S P N  
 481 TGCAGTTCTACTACAACCTTCCGGACAGCGTGGTCAGGATGGAGGAAGTCCAAGGGTCCGATGTCTCCGAATG.  
 L A A V K T P V S P V S S N G T A F S E H G G S P  
 561 CTAGCAGCGGTGAAGACGCCAGTGTGCGCTGTTTCGAGTAATGGTACAGCGTTCAGCGAGCATGGGGGTTACCCG'  
 T P T V S S A E P R M S D M A L K R Q R S E E D D  
 641 CACCCCAACCGTCAGCTCAGCTGAACCAAGAATGTCTGATATGGCCCTCAAAGGCAACGTTGAGAAGAGGACGA  
 P P N S D R R S P D Q K R P R P E N I V L S N G F  
 721 CCCCAAACAGTGACCGAAGATCGCCCGATCAAAGCGACCACGTCGCGAAAACATTGTTCTATCAAATGGCTTTT  
 D V H G R M F S A E H A R A M N I L T R L F P E Q  
 801 GATGTTTCATGGACGTATGTTTTTCAGCGGAGCACGCACGAGCTATGAATATCTTGACGAGACTTTTCCCGGAGCAA  
N V L E L I L K G C S G D I V Q T I E C V L P S H  
 881 CAATGTGTTAGAGCTGATTCTCAAAGGGTGCAGTGGGGACATTGTACAGACTATAGAATGCGTTCGCAAGTCA  
 E A L A R A Q M F T T G P P R G M F P M P S P L R  
 961 AAGCCTTAGCAAGGGCTCAAATGTTCACTACTGGACCCCAAGGGGAATGTTCCCATGCCCTCCCCACTGAGGA  
 L S A F S P I T P G L R P S L P I M P L E L H T A  
 1041 TTGTCAGCCTTCTCGCCTATCACACCAGGCTTACGGCCATCCTTACCCATTATGCCTTTAGAGCTTCACACTGCA  
 C A S G K C P G C V Y Y P G M V P R S P R D A P K  
 1121 ATGTGCCAGTGGAAAGTGCCCTGGGTGTGTGACTACCCAGGGATGGTCCCTAGGAGTCCCAGAGATGCCCCCAA  
 P R E Y A S P P P L P S K S P R S P R T S E S R S  
 1201 CAAGGGAGTATGCCTCTCCCCACCCTGCCCTCTAAGTCACCTCGGTCACCAAGAACATCAGAAAAGTCGTTCCG  
 R E E M L D S Q R V H S A T A A L I S M S N S A Q  
 1281 AGGGAAGAAATGCTGGACAGTCAAAGAGTGCACCTCTGCTACTGCTGCACTCATTTCATGAGCAACAGTGCACAG  
 H N E R S S V I R N N M S M V G S P T S V K S H D  
 1361 CCACAATGAACGCTCATCTGTCTATTAGGAACAATATGAGCATGGTGGGATCCCCGACAAGTGTCAAGTCACATGA  
 V R G S P I R E L N \*  
 1441 TCAGGGGATCCCCTATTAGGGAACCTCAACTAGGGGAATCTACTCCTACCCTAGTAATTAATCTTTGACTACTGTA  
 1521 CATTCAACCCATTTTCACTTTTTTTCTTAGTTATTTAGTGCTAGTTAGACTTAAAAGAAATTAATTTGGTAAGGAAA  
 1601 GATTTGACAGTTTAGATCATTTAGCTAAGAAAGAAGCTATCTTTGGATATCAGTGCTTTAGTATTTTGCTGATGA  
 1681 CGGGATGTTGAAACACAATCTTTTATCTATTATTCATATTCATAGCAACAATAAATATCTAAAAATCTTATGACCAT  
 1761 TAAAATAAAAAAACTAAAAGTTTATAAGTTGAACAACCTTCCAGGCTGTTCTACACTTCAGTATGTCTACGAC  
 1841 TGTAGTTATGATAGTGATGATGTTTGTGGGTTAGGGGTTGTATAAACATTGCAATTGATAAGTTCTTAAAATAAA  
 1921 AAAAAAA

**Fig.6 NvDM5 cDNA sequence and translation**

Fig.7

Boxshade representation for the alignment of DM domain. Cnidarian DM domains (AmDM1, NvDM5 and NvDM-like) are aligned with those of various organisms including vertebrates, *C. elegans* and *Drosophila*. The two intertwined Zinc fingers are indicated above. All the cysteine and histidine residues (CCHC and HCCC) of these fingers are absolutely conserved throughout evolution.

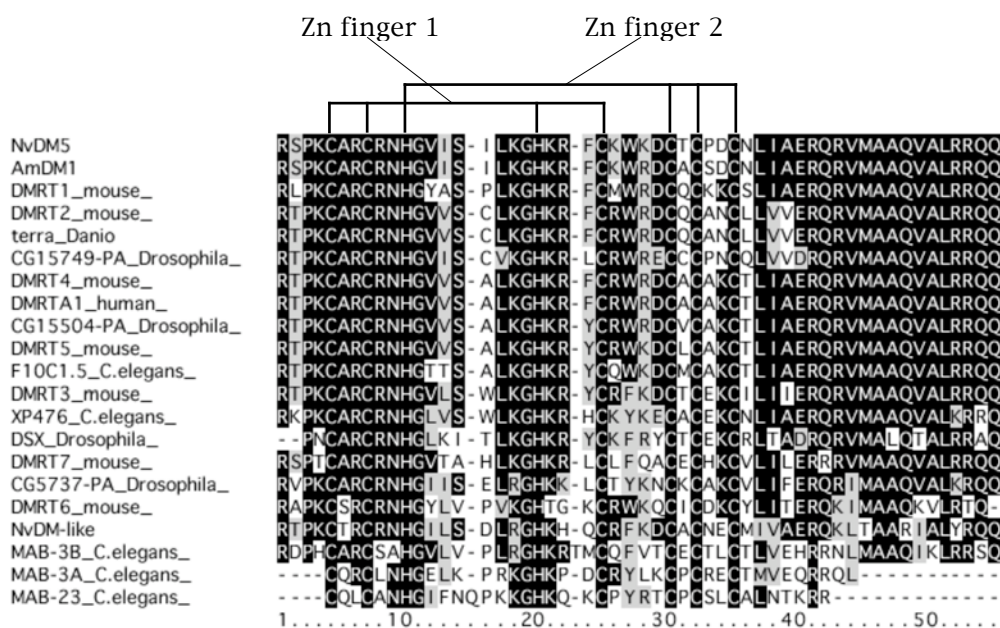


Fig.7 Alignment of DM domains

Fig.8

The neighbour-joining tree for the phylogeny of the DM domains of various organisms: vertebrate (DMRT1-7 and terra), fly (DSX, CG15749-PA, CG15504-PA and CG5737-PA), nematode (F10C1.5, MAB-3A and 3B, MAB-23 and X476), coral (AmDM1) and sea anemone (NvDM5 and NvDM-like).

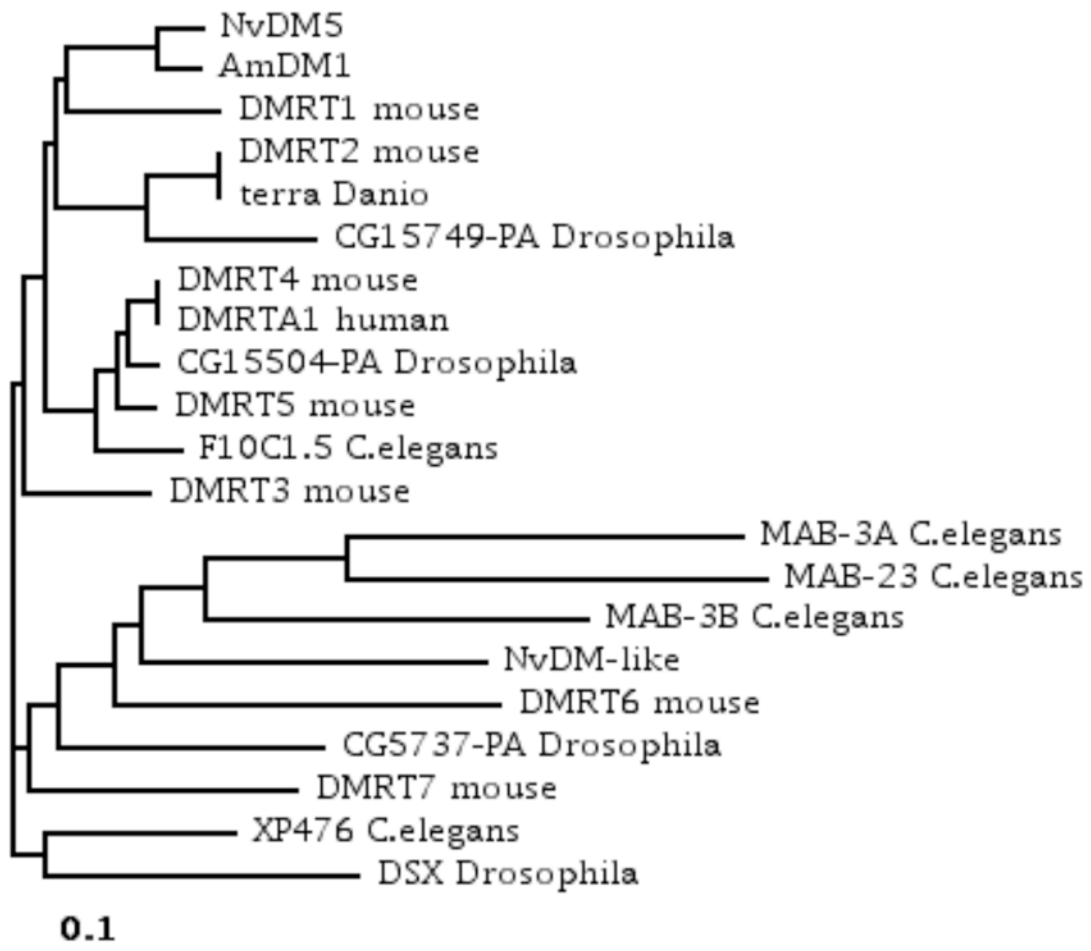


Fig.8 Phylogeny tree of the DM domain

Fig.9

Boxshade representation for the alignment of DMA domain. Cnidarian DMA domains (AmDM1, NvDM5 and NvDM-like) are aligned with those of various organisms including vertebrates, *C. elegans* and insects (*Anopheles* and *Drosophila*).

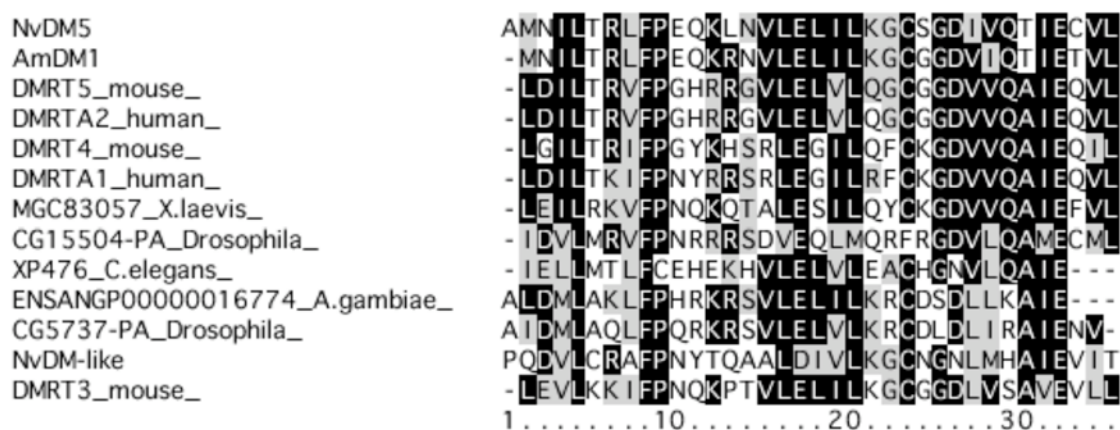


Fig.9 Alignment of DMA domains

Fig.10

The neighbour-joining tree for the phylogeny of the DMA domains of various organisms: vertebrate (DMRTs and MGC83057), fly (CG15504-PA and CG5737-PA), mosquito (ENSANGP00000016774), nematode (X476), coral (AmDM1) and sea anemone (NvDM5 and NvDM-like).

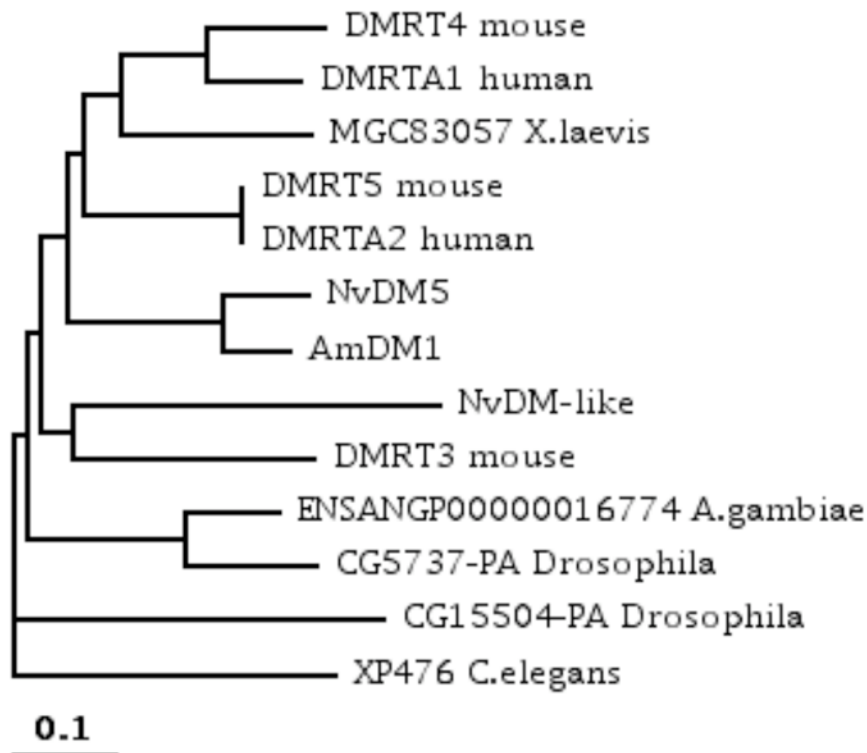


Fig.10 Phylogeny tree of the DMA domains

Fig.11

Alignment of the entire DM proteins of AmDM1, NvDM5 and NvDM-like. The regions corresponding to the DM domains are indicated by an arrow over the sequences. And the DMA domains are boxed. The sequence conservation is high within and around DM and DMA domains especially between AmDM1 and NvDM5. In addition, small ‘islands’ with high conservation are scattered in the C-terminal region to the DMA domains of AmDM1 and NvDM5.

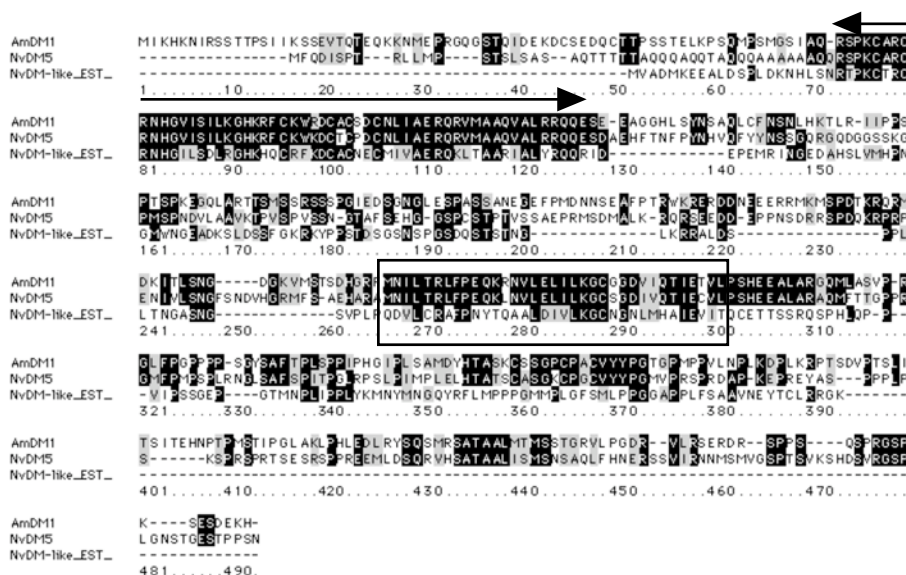


Fig.11 Alignment of Cnidarian DM proteins



Fig.12

Partial of complete nucleotide sequences of DM-binding DNA fragments isolated from clone 4U, 5B, 6A and 6E. 4U and 6A turned out to be identical fragments. The *Acropora* specific repeat sequences are underlined. 5B fragment does not contain the repeat.

#### 4U(=6A) partial sequence

GATCTCCTCGAGATAAGCAACTTTTTGTGCAAAGTACCGTGCCTAGGGTTTTGACCCAATTTTTAGAAAAAT  
GTTGCAAAAAGTAGGGCATTGTCTGAAAAACCACCAAACTTGTTCGATCTCCTCGAGATAAGCAACTTTTTCT  
GCAAAGTACCGTGCCTAGGGTTTTGACCCAATTTTTAGAAAAATGTTGCAAAAAGTAGGGCATTGTCTGAAA  
AACCACCAAACTTGTTCGATCTCCTAAAAAGTGACCTAGACAACACTTTCTCAACAATGTGACAAAAGTAGGC  
CATTGACGGACAGCTGCAGGTCATTTCAATTGCAAATCGGTGATAAATGTGGTCAAATCGGTCAATACAAAAGTA  
CGGCCAACTGATTTGCGCAAAATGCGCGCACTACCACAGTTGCGATAGCAAATACAGCCAGCGCCAGAGACA  
TGATGGTGCCCTGACAGATTTTTTTCTATTTTTTTTTAAGTCAAAATGG

#### 5B complete sequence

GGATCTTTCTTGTACAAAAGAGGAAAGCGTGTAAAGCCGTACGCGCTTGCCTGACTTACTTCATATTTCCCGTG  
ATACTCAAATGAAAAATTTGTATCAAAATGTGTGCGCTTCAACGTAATGAAAAATACAAAAGTAGCCATCATCT  
TTGATAAAAAACAAACAAAAACAAAAACAAAGGAAAGCAAAATACAGAGCAACTTTCAAATTTGCATGACA  
AAGTAGAAAAAAGTTGACAAAAATGCAAGCAGTTTTCCACAGACACTGCCATATGCATTATGCACAAGCT  
GGGGCCGATACCACAAGAAATGTGCACATCCCACGCCCATTTGGGCTGCTATGACCAGTTATTTCCATTCCTTTT  
AGGTTTCATAAGTATACTATAACCCATGTCAATGCAGATGTGTAGTCAGATACTGGTATTTAAACAGCAGCTCTA  
TACTACTAGTTGTAGAAGCTGTGACCAGTCAAGAAGACCTAAGAAGATCC

#### 6E partial sequence

TAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTTCAAGAATTTCCCGGATCCAAATAGCATTAAAA  
GAGCCGCCACTGTTAAAAACGTCCAATTTAAATGAACTTCCCTAAGCCAAATTTAAAGTCCAAGCAAATAAAATATA  
GCCTCATAACACCCGACGGGTGTTTTTCTATAGCCCGCTATTTAATCTAAGAACTACGCTCCGATTGGCTT  
ATTTTCGCGGGCACGTAATGAAGGTTTTGGTTTTTCGCAAGCTCATGTGCTGTTTTGAATAAAATACCAGTTGTGC  
GGAGGAAAGTTTTGTTTTTCCGACGCGGAAAGCTTTACAGTAAAGGAAAAATCAATTTCAAAAAATTTGCGAC  
GTTTGTGTAATGGTACCAACGAAAGCTCCACGTACCCTGCCACTCGAATAAAAGTTCTGCGAAGCTTGATACG  
CGGTACGCAGAACTAATAAAATCAAGTATGGAAATTTATTTCAAAACAGTATTTCTCGTTCTTAAAGCGTGACT  
CGCCAAATTAAGTAGTGATCAAAGGAGTCTGCCTTTTTTACCCTTGGTATGTTCTTAGACAGGAGACGTTGGTA  
ATCTTTGCCACATTACCGTTCCCGAGGAAAAACGGTCATCTTCTGAAACAACCTAAGGAAATTCAGAAAAAA  
AAAAATGAAACAGTCAATCTATGTTGTATCCATTTACATAAATGCTCTCTTAAGACTGTTTCATTTAGTTGTT  
TAGAGACTAAAAGGAAAAAATAATTAATTTGAAAATCTATCAAAATCGTGTTTAAAACCTATTTCTTTCAA  
TTGCTAATACCTGCCTTGGTATTTGATAGACCCAAATTTTTAGAAAAATGTTGCAAAAAGTAGGGCATTGGTG  
AAAAACCACCAAACTTGTTCGATCTCCTCGAGATAAGCAAGTTTTTTTTGCAAAGTACCGTGCCTAGGGTTTT  
GACCTTAGGGATGATTGGAAGGAAATGAACTAAGGTAACGATAAGAATGGAAAGTAATTTGCTGTGACGCTGC  
CTCATGTGTTGACCTACGTAAGAAACGTGTTACCATTTAAATGAAAGCTACTGCGCAGGTCTTACCTTTTCGAG  
ATTTGAGTATTAATGACTCAAAGAGTCTATGAGTCAATAACCATGAGGCAATGAGTCAAGTGCAGTTACCCAAA  
TCCATAATATGAAATCTAACC

Fig.13

Alignment of the repeat sequences specific to genus *Acropora*. The nucleotide sequences of the repeats in the DM-binding fragments (one in 6E and two in 4U) are compared to those of *Acropora latistella* (pJCAL1, 2 and 4) and *Acropora formosa* (pJCAF9 and 10).

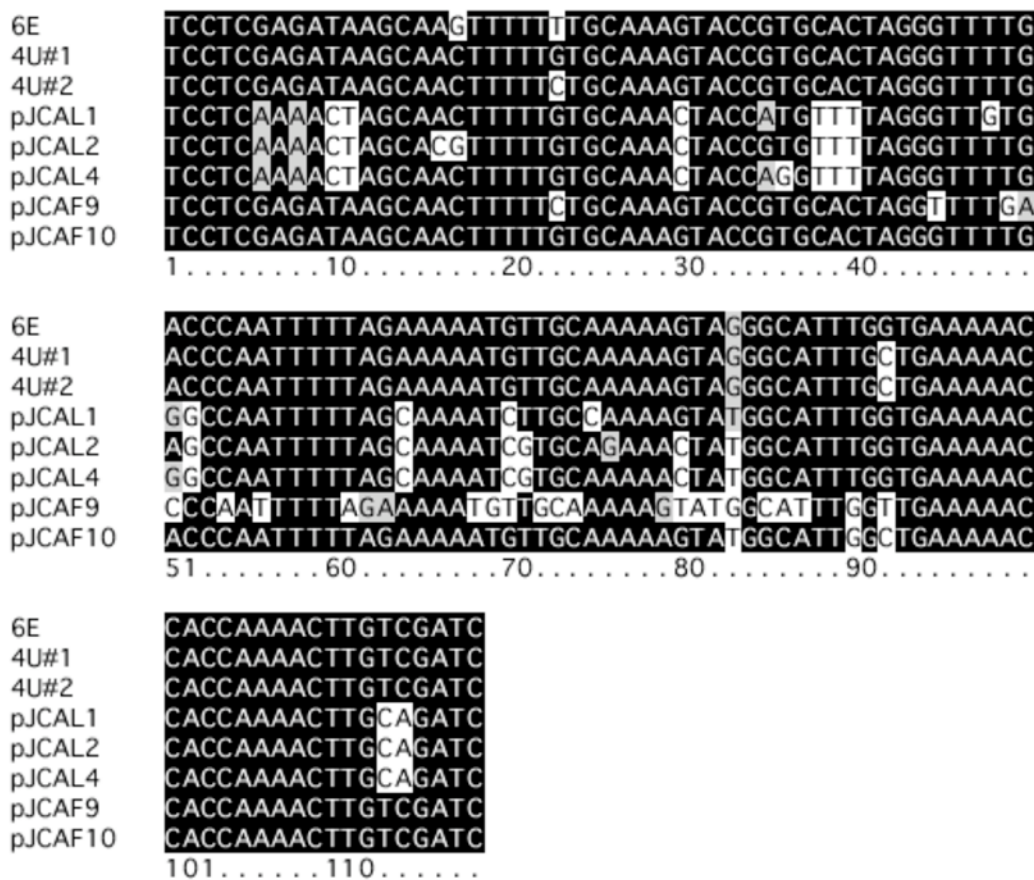


Fig.14

The neighbor-joining tree of *Acropora*-specific repeat sequences. pJCAL1, 2 and 4 are sequences of *Acropora latistella*. pJCAF9 and 10 are sequences of *Acropora formosa*. The sequences found in the DM-binding fragments were 6E, 4U(=6A)#1 and 4U(=6A)#2. These sequences form 3 separate groups according to the species.

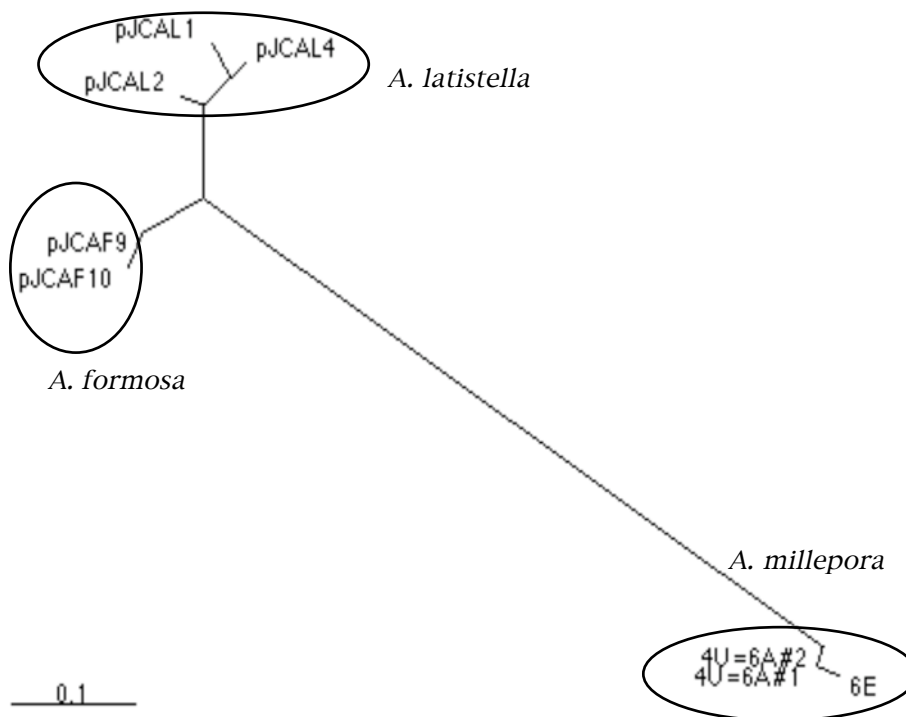


Fig.15

The nucleotide sequence of the partial 5B1b clone was subjected to blast-x to detect the homologous proteins for the putative translation frames encoded by 5B1b. A number of proteins, most of which were members of the membrane transporter family, were detected with an approximately 1.2kb interruption that was thought to represent an intron.

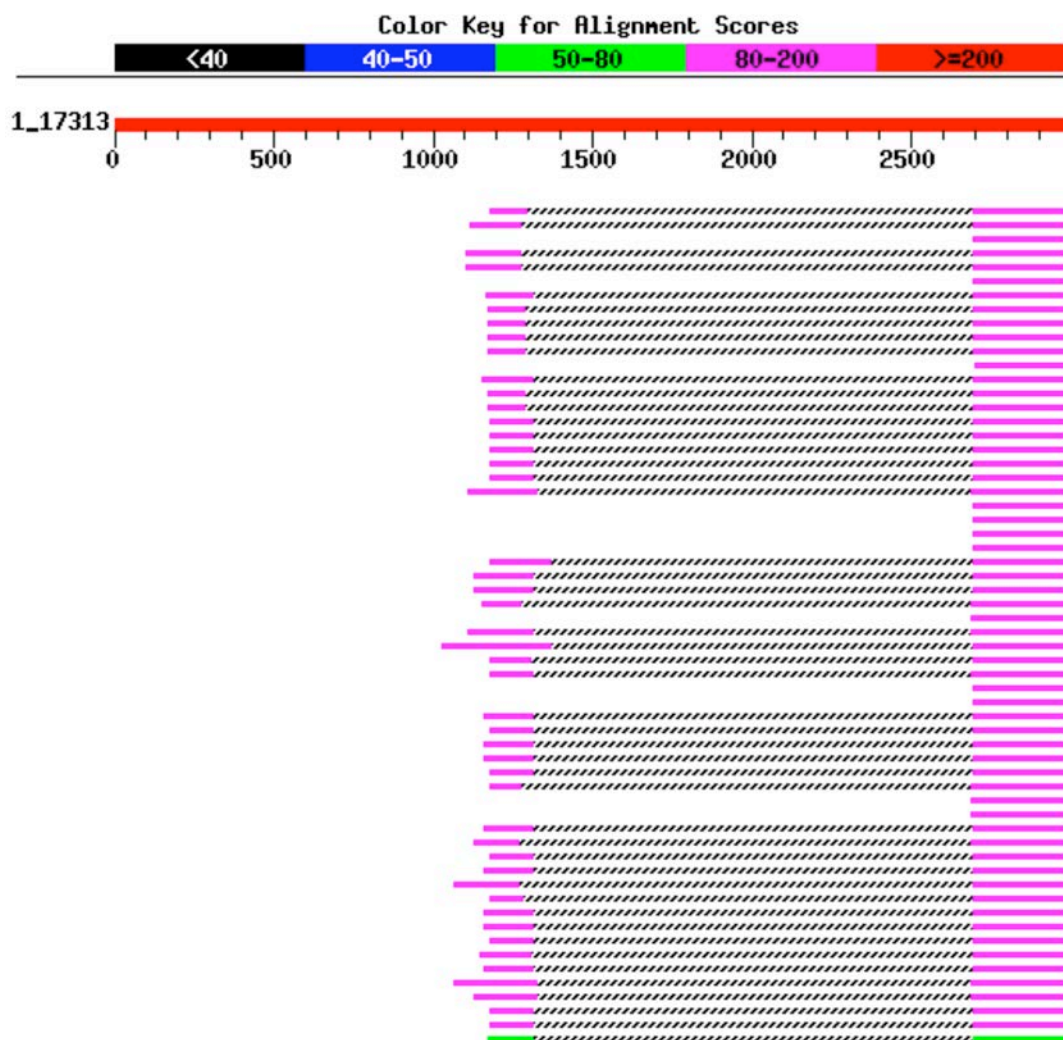


Fig.16

Sequencing strategy for  $\lambda$ -5B1b clone. The sequencing shown by arrows was performed by primer walk started within the region corresponding to the 5B fragment. The 5B fragment which was isolated by one-hybrid screen is shown by the solid bar.

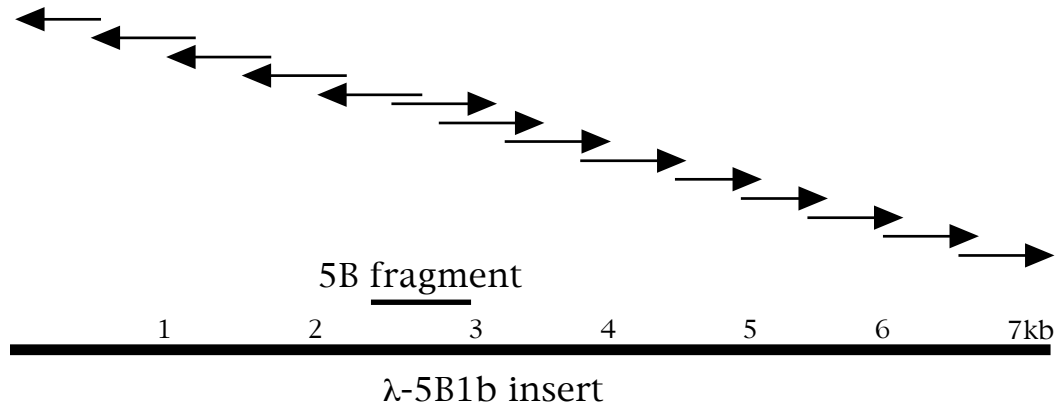


Fig.17

Southern-blot analysis on *AmTAT1* gene. The genomic DNA of *Acropora millepora* was digested with ECoRI, BamHI and KpnI and hybridized with the probe corresponding to the exon 3 of *AmTAT1*. A single band is seen on each lane.

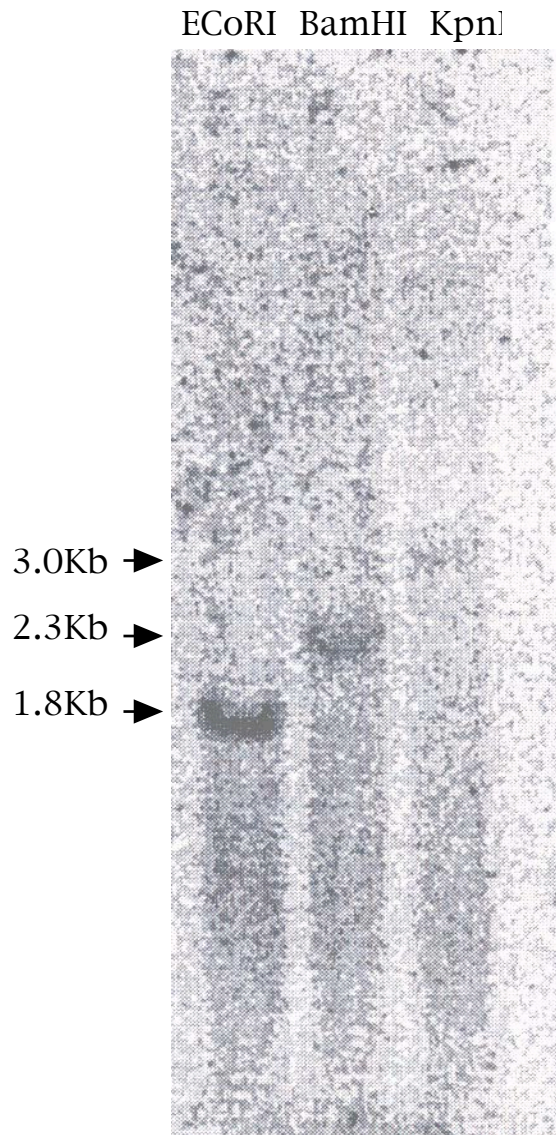


Fig.18

A: cDNA sequence and the translation of *AmTAT1*. The putative polyadenylation signal in the 3'-UTR is underlined.

B: The strategy for sequencing *AmTAT1* cDNA clone. The entire cDNA was sequenced by reading both strands.

A

1 CAACAAGGTGCCTGCTGTCTAGGATCCTGTTTCGCTGAGGTAATCACTGC  
51 TTCTTTGTCATCTGCTGAACTACATGTCAACAATCTTCCCAAATCAACTC  
M T N I D D K E  
101 TTCAACTGCCTGCAAATACTCAGTCATGACAAACATAGATGATAAAGAA  
I F L P G D A A S E H S P G P Y E  
151 TATTCCTCCCTGGAGATGCAGCCTCAGAACATTCTCCAGGCCCTTATGA  
C E K P D G G W G W I V C C G T I  
201 TGTGAAAAACCTGACGGGGGTTGGGGATGGATAGTTTGTGTGGTACAT  
L V N F I V F G I H N S F G V V  
251 TCTAGTCAATTTTCATTGTGTTTGGAAATTCATAACTCGTTCGGTGTGTG  
Y E Y L V N E H S L G E A E T A W  
301 ACGAATACCTGGTGAATGAGCACAGCTTGGGAGAGGCAGAGACAGCCTGC  
V G S M A A S L N F L C G P L S S  
351 GTTGGCTCAATGGCCGCTTCTTTGAACTTCCTGTGTGGACCACTCTCCA  
V L C D R F G C R R V A F F G A  
401 TGTTTTGTGCGACCGTTTGGCTGTCGTCGTGTAGCTTTCTTCGGAGCC  
I L S V T G L F L T S F I T R L H  
451 TCCTTTCAGTTACTGGACTTTTCCTCACGTCATTTATCACTCGTCTGCA  
Y M Y I T Y G L L W G V G S S F S  
501 TACATGTACATTACTTATGGACTTCTTTGGGGGGTGGGATCGAGCTTTTC  
Y V P S I V M L G N Y F D R R L  
551 GTACGTACCTTCAATTGTCATGTTGGGGAATTACTTTGACAGAAGACTTC  
A L A N G L G T S G S G V G S L V  
601 CGTTAGCCAATGGACTAGGCACTAGCGGAAGTGGAGTAGGTTTCATTAGT  
A S P V I N Y T L R A F G W K T S  
651 GCTTCGCCAGTGATCAACTACTCTGCGGGCCTTCGGTTGGAAGACTTC  
L R I I S G A A A F L A F A C I  
701 TCTGCGTATAATCAGTGGTGCGGCAGCGTTCCTTGCTTTTCGCGTGCATT  
V Y R P L R T K N E Y R R S G V K  
751 TTTATCGACCATTAAGAACGAAAAACGAGTATCGCAGAAGTGGCGTTAA  
R P R S K L C D S S I W K N R A S  
801 CGACCTCGAAGCAAATTGTGTGACAGTTCTATCTGGAAAAATCGAGCCT  
V T W V L T I A L F Q F G Y P V  
851 TGTAACATGGGTCCTTACAATAGCACTTTTTTCAGTTTGGATACCCAGTC  
P F V H M V K Y A N D L G V P T S  
901 CTTTTGTTACATGGTAAAGTACGCTAATGACCTTGGTGTACCAACTTC  
K G A W L V G F L S I M S T L G I  
951 AAGGGAGCATGGCTGGTTGGTTTCTTATCGATCATGTCTACATTGGGGCC  
V I F G K I C E H Q R V N K L H  
1001 TGTGATTTTTCGGAAAGATTTGCGAACATCAACGCGTCAACAAGTTGCAC  
V Y Q L S L F V I G M S T V I C P  
1051 TGTATCAGTTATCACTGTTTGTATTATTGGAATGAGTACCGTTATATGTCC  
F T E S Y A G F V V Y S L V F G I



1101 TTCACGGAAAGTTATGCGGGGTTTGTGGTGTACTCGCTTGTGTTTGGATF  
       F D G C F V G Q V A V I T G E I  
 1151 TTTTCGATGGCTGTTTTGTTGGTCAGGTCGCCGTTATCACTGGTGAAATA(  
       A G K G K L S Q A I G N M F G V V  
 1201 CCGGCAAGGGAAAACCTCTCTCAAGCCATAGGAAATATGTTTGGTGTAGT(  
       A V P M I L G A P V A G W L H E I  
 1251 GCAGTTCCCATGATTCTGGGTGCACCTGTTGCCGGCTGGCTTCACGAAG(  
       F G S Y R Q A F F V C G S F A I  
 1301 TTTTGGTTCGTATCGTCAAGCCTTCTTTGTTTGCGGTAGCTTTGCAATT(  
       A S S L L L F V V S C M V N K Q R  
 1351 CGTCATCTCTCCTTCTCTTCGTTCGTTTCTTGTATGGTTAACAAACAGAG/  
       R R V I L S T S S R K T K E L H I  
 1401 AGAAGGGTTATCTTATCCACAAGTTCAAGAAAGACAAAAGAGCTTCATG/  
       E S L V P G S E S C A L S G F S  
 1451 AGAAAGCCTTGTGCCCGGAAGCGAATCATGTGCTTTAAGTGGATTTTCA/  
       S H Q T L V I A S R E T V L \*  
 1501 CACACCAGACGCTTGTAATTGCCAGTAGGGGAGACTGTTCTCTAATGACT/  
 1551 GGATTCCACTGAATTAGATGCATGCCGAACTTAATAAGCGTCAGAAAGT(  
 1601 TCCACCCACCTTTCTCGTAAACTTCACGATCTAAGGACAATAAAAAGTCA(  
 1651 TGTCTTCCACTTCCCTATACTCCTTAAAGAGACACAAACAGTAACCTTA/  
 1701 TCGGCCTAATCTAAAAATATCCTGAACACAAACACTAACTCTGATGTTG(  
 1751 AAACAGGTGCTGTTGGCTTAGACATGAGGGGAGACTTCGTGAAGTTTAT(  
 1801 AAAATTGAGCATGCTTTTAATGAGGTGCATTTCTGGATATTAGTGCTAA/  
 1851 GACTGCATGGAAAAGTCGAGGTTCTTAACAAAGCTGCCCAGGATCCGGT/  
 1901 GTTCAAACCGCGGTGAACGGCAGGCCGGGATTCAAACCTAACAGCTGAG(  
 1951 TTTTGAATTTTAGAAAAACGCCCTACCAAGTAGACTTACGAAAGGTCTT/  
 2001 TTGAAACTGAAGGACCAAGCCCTCATGAAAATGTCTCTAAGTTAATGGT/  
 2051 TTAAAAGTTTCTGCTAATCCTAAATTAATTTAATCCGGGTTTGAACAAC/  
 2101 GGGCATAGAAGGACTCGTGAAGGGGAACTTTGACCAAATCTGAAACGT/  
 2151 TGCCATCGAGGTTTCTTCCATTTTCATAATCATGTCACTTTTGGAAATG/  
 2201 CAATTTTCAATTTTCAAGAATGCTAAACTCCATTGCTGTTGGCTTGTTTG/  
 2251 GGATGAATATGTGAGGTAATGAATTAATGCACTGTTTAATTTACTATGT/  
 2301 AATTTAATTTATTTATTTAAGACAATAATGTCTTTATTGGCAAATGTCT/  
 2351 TAACTAAGATTTTATTTATTTTGAAAAAAAAAAAAAAAAAAAAAAAAAA/  
 2401 AAAAA

B

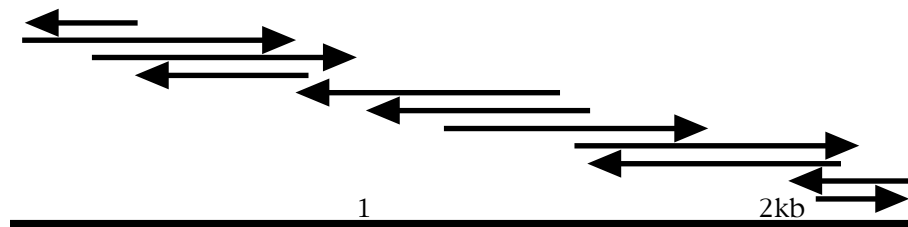


Fig.19

5 conserved regions were found in AmTAT1 protein.

AraJ, Arabinose efflux permease

NarK, Nitrate / nitrite transporter

Permease of the major facilitator superfamily

UhpC, Sugar phosphate permease

CynX, Cyanate permease

MelB, Na<sup>+</sup> / melibiose synporter (partial)

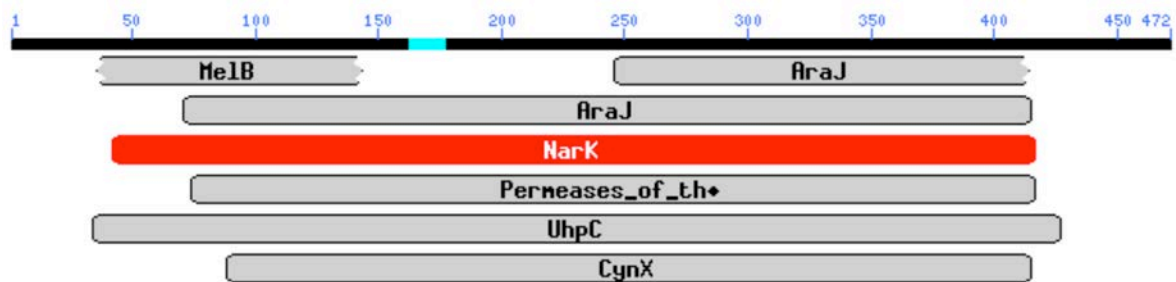


Fig.20

A: Amino acid sequence of AmTAT1. 12 predicted transmembrane domains are underlined. Potential sites for phosphorylation by Protein kinase C (#), phosphorylation by Casein kinase II (\*) and N-glycosylation (@) are marked over each residue.

B: A schematic representation of the predicted structure of the AmTAT1 molecule localized in the plasma membrane. Both the N- and C-termini are located in the cell. The intracellular loop between the transmembrane domain 6th and 7th is the largest.

A

\*  
MTNIDDKKEIFLPGDAASEHSPGPYECEKPDGGGWIVCCGTFLVNFIVFGIHNSFGVV  
TM1  
YEYLVNEHSLGEAETAWVGSMAASLNFLCGPLSSVLCDFGCRVVAFFGAILSVTGLF  
TM2 TM3  
LTSFITRLHYMYITYGLLWGVGSSFSYVPSIVMLGNYFDRRLALANGLGTSGSGVGSL  
@ TM4 \* TM5  
VASPVINYTLRAFGWKTTLRRIISGAAFLAFACIVYRPLRRTKNEYRRSGVKRPRSKLC  
TM6  
DSSIWKNRAYVTWVLTIALFQFGYPVPFVHMVKYANDLGVPTSKGAWLVGFLSIMSTL  
TM7 TM8  
GRVIFGKICEHQRVNKLHVYQLSLFVIGMSTVICPFTESYAGFVVYSLVFGFFDGCFFV  
TM9 TM10  
GOVAVITGEIAGK GKLSQAIGNMFGVVAVPMILGAPVAGWLHEAFGSYRQAFFVCGSF  
# TM11  
AIASSLLL FVSCMVNKQRRRVILSTSSRKTKEELHEESLVPGSESCALSGFSSHQTLV  
TM12  
IASRETVL

@ N-glycosylation  
\* Casein kinase II  
# Protein kinase II

B

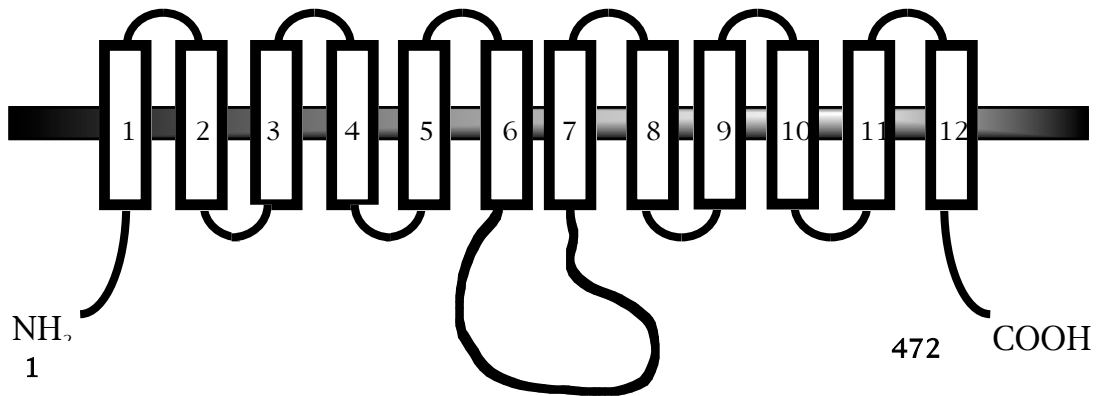


Fig.20 Structural features of AmTAT1

Fig.21

Alignment of AmTAT1 with hTAT1 and X-PCT. The predicted transmembrane domains are in bold letters. There is a general tendency that the conservation is high in the transmembrane domain. The amino acid sequences in the underlined portions are absolutely in accordance with the consensus conserved in all the monocarboxylic acid transporters.

```

AmTAT1 -----
hmCT8 (X-PEST) MGRGGGGLDVGGGEGSRDRLSRDGLASWGAEPGGGGSGSGSSSPSSSSC SSRNKYC
hTAT1 -----

AmTAT1 -----
hmCT8 (X-PEST) SGSSGSSSHSPPAAMALQSQASEEAKGPWQEADQEQQEPVGSPEPESEPEPEPEPEPV
hTAT1 -----MVLSEQEEDSARG-----TSEAQPLG-PAPTG-----

AmTAT1 -----MTNIDDKEIFLPGDAASEHSPGP-----YECEKPDGGGWIVC
hmCT8 (X-PEST) PPPEPQPEPQPLPDPAPLPELEFESERVHEPEPTPTVETRGTARGFQPPPEGGFGWVVV
hTAT1 PPPGPGPSDSPE-AAVEKVEVELAGPATAEPHEPP-----EPPEGGGWLVV
                * : . . * * : * : * : * : *

AmTAT1 TFLVNFIVFGIHNSFGVVYEYLVN-EHSLG----EAETAWVGSMAASLNFLCGPLSSV
hmCT8 (X-PEST) ATWCNGSIFGIHNSVGILYSMLLEEEKEKNR-QVEFQAAWVGALAMGMIFFCSPIVSI
hTAT1 AMWCNGSVFGIQNACGVLFVSMLETFGSKDDDKMVFKTAWVGSLSMGMIFFCCPIVSV
: * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

AmTAT1 DRFGCRRVAFFGAILSVTGFLTSFITRLHMYITYGLLWGVGSSFSYVPSIVMLGNY
hmCT8 (X-PEST) DRLGCRITATAGAAVAFIGLHTSSFTSLSLRYFTYGILFGCGCSFAFQPSLVILGHY
hTAT1 DLFGCRKTAVVGAAVFGVGLMSSSFVSSIEPLYLTYGIIFACGCSFAYQPSLVILGHY
* : * * . * * * : * * : : * : * : * : * : * : * : * : * : *

AmTAT1 RRLALANGLGTSGSGVGLVASPVINYTLRAFGWKTSLRRIISGAAFLAFACIVYRPI
hmCT8 (X-PEST) RRLGLANGVVSAGSSIFSMSFPFLIRMLGDKIKLAQTFQVLSTFMFVLMLLSLTYRPI
hTAT1 KRLGLVNGIVTAGSSVFTILLPLLLRVLIDSVGLFYTLRVLCIFMFVFLAGFTYRPI
: * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

AmTAT1 KNEYRRS---GV---KR---PRSKLCDSSIWKNRAYVTWVLTIALFQFGYVPPFVHM
hmCT8 (X-PEST) SSQDTPSK-RGVRTLHQRFLAQLRKYFNMRVRQRTYRIWAFGIAAALGYFVPYVH
hTAT1 STKDKESGGSGSSLFSRKKFSPPKIFNFAIFKVTAYAVWAVGIPLLALFGYFVPYVH
.. : * * : : * : : : * * . * : * * : * : * : * : * : *

AmTAT1 YANDLGVPTSKGAWLVGF-LSIMSTLGRVIFGKICEHQR-VNKLHVYQLSLFVIGMST
hmCT8 (X-PEST) YVEEE-FSEIKETWLLVCIGATSGLGRLVSGHISDSIPLKKIYLQVLSFLLLGLMS
hTAT1 HVNER-FQDEKNEVLMCIGTVSGVGRLLFGRIADYVPGVKKVYLQVLSFFFIGLMS
: : : . * : : . : * : * : * : * : * : * : * : * : * : * : * : *

AmTAT1 CPFTESYAGFVVYSLVFGFFDGCFVGQVAVITGEIAGKGKLSQAIGNMFGVVAVPMII
hmCT8 (X-PEST) IPLCRDFGGLIVVCLFLGLCDGFFITIMAPIAFELVGPMQASQAIGYLLGMMALPMIA
hTAT1 IPLCSIFGALIAVCLIMGLFDGCFISIMAPIAFELVGAQDVSQAIGFLLGFMSIPMTV
* : : . : : . * : * : * : * : * : * : * : * : * : * : * : *

AmTAT1 PVAGWLHEAFGSYRQAFFVCGSFAIASLLLFVSCMVNKQRRRVILSTSSRKTKELE
hmCT8 (X-PEST) PIAGLLRNCFGDYHVAFYFAGVPPIIGAVILFFVPLMHQRMFKQRDSSKDKMLAPI
hTAT1 PIAGLLRDKLGSYDVAFYLAGVPPIIGAVLCFIPWIHSKKQREISKTTGKEKMEKMI
* : * * : : * : * * : * : * : * : * : * : * : * : * : * : *

AmTAT1 SLVPGSESCALSGFSSHQTLVIASRETVL
hmCT8 (X-PEST) ---PNGE---LLPGS-----PNPEEPI-
hTAT1 ---QNSL---LSSSSG---MFKKESDSII
.. * * : :

```

Fig.21 Alignment of AmTAT1, X-PEST and hTAT1

Fig.22

Neighbor-joining tree for the monocarboxylic acid transporter family including AmTAT1.

hMCTs: human monocarboxylic acid transporter

h,rTAT: human or rat T-type amino acid transporter

Kar: Karmoisin

GH12716p: a hypothetical protein of *Drosophila melanogaster*

Ano. gambiae: ENSANGP00000022234 (a hypothetical protein of *Anopheles gambiae*)

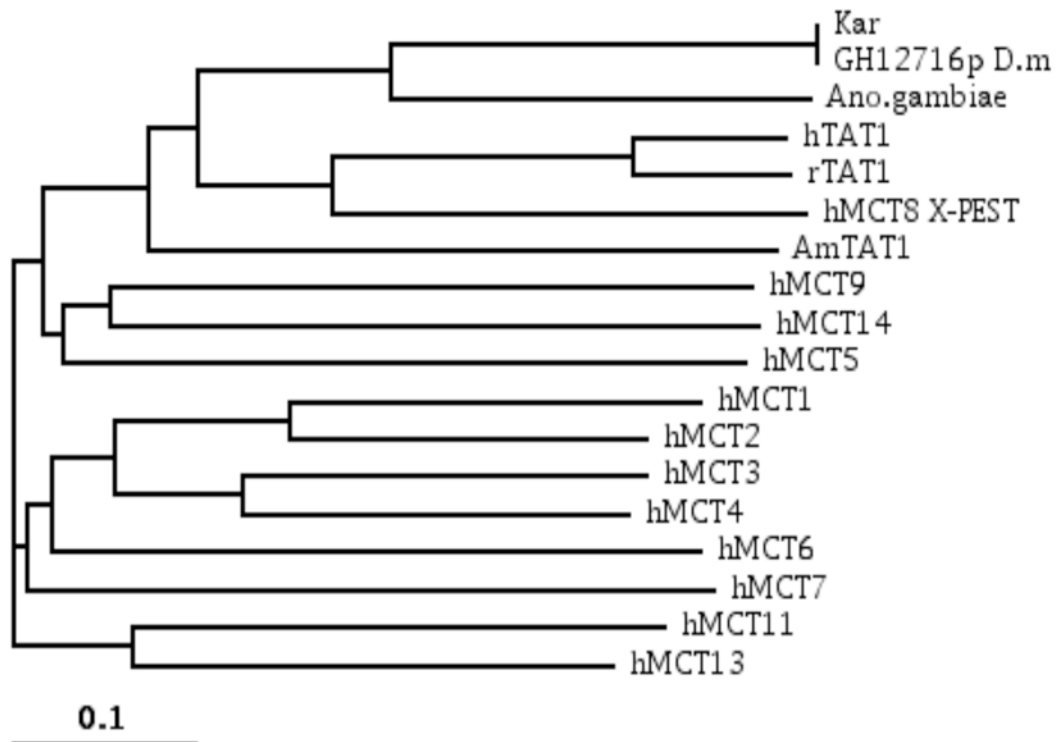


Fig.23

The genomic structure of *AmTAT1* gene.

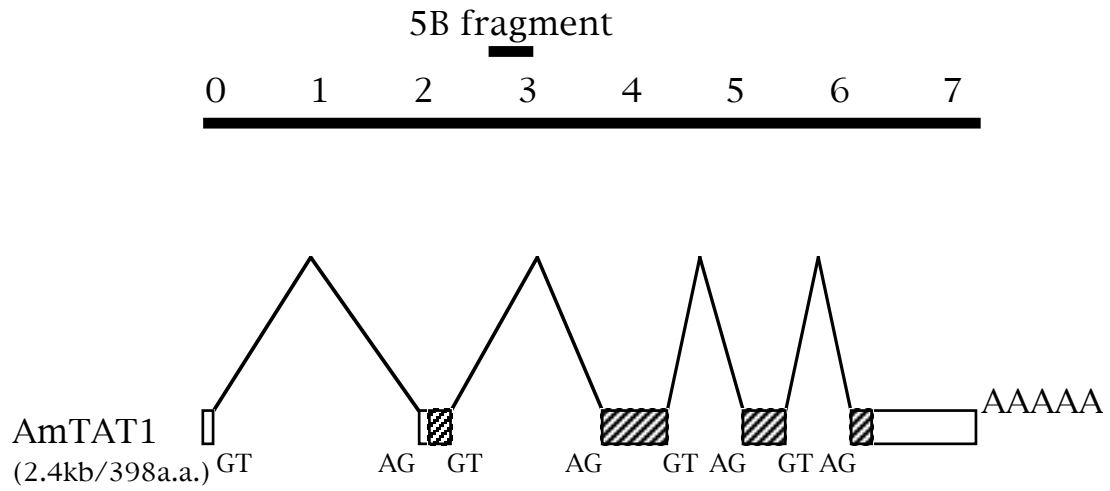
A: A diagram showing the intron-exon organization of *AmTAT1* in parallel with the scale (kb). The 5B fragment is also mapped on the scale. Open box: untranslated regions in the *AmTAT1* cDNA.

Hatched box: coding regions for AmTAT1 protein.

B: Table for the character of introns of *AmTAT1*. Sequences of all the exon-intron boundaries are subject to 'GT-AG' rule.



A



B

	length	phase	3' of exon	boundary sequence of intron	5' of exon
intron1	1958	X	GAG	<b>GTTGG---CTATTGAAAG</b>	GT
intron2	1395	1	CAG	<b>GTAAG---TCCTTGTTAG</b>	CC
intron3	724	0	ATG	<b>GTAAG---ACCCATACAG</b>	GT
intron4	602	1	CCG	<b>GTAAG---TATTTTCTAG</b>	GC

Fig.23 Intron-exon organization of *AmTAT1*

Fig.24

Nucleotide sequence of the upstream regulatory region of *AmTAT1* was subjected to the computer search for potential *cis*-acting elements and the transcription factors. The promoter lacks TATA box. and contains multiple sites for transcription factors.

CGATACTACG

-500  
CAACAAAACGTCTAATAAAGATATATTTTCCTGTCCGGAAGCTCGCGGAGAG  
NIT2  
CTTCAAAGTTGAAAATATCAGTATACCGGGTGCCGAGCTCACTAAAATTC  
-400  
ACATTACGGTGGAATAGTGAAAGCTTCAGCCATATAAAACAACGATTG  
SRY HFH-8  
TATTATTAACTGAAAATTAACTAAAAATTTGACTGCCGGTCACGGAGTAT  
-300  
TTTCAACAACCTGTCGGCTGCCGGACAACAAAATCCATTCTTTTCCATA  
  
AAATTTGTCTAACTGGTTAAAATTCTCTCCACTCCAGGCTTTTTTCTCAT  
-200  
CATCTCAGAATCGGAGGTGATCATCGTAGAAAATATTAGCGTTCCTTGGA  
PBF HSF  
TCGCTTTTTCATATTTCCACAATCTGTTGGTCCTAGACCAGCAAGTCCAAT  
-100  
TCAAATTTTACGAGTCGCTGTCATAATGTTGCATCATCTAATTCCATTCT  
  
TCTCACGGTATGTTTTTACTAGCCACGGTCACAAACAAAAATACGCGCAA  
HSF SRY

Fig.25

Expression analysis on *AmTAT1* by RT-PCR. The amplification was observed for all the developmental stages and adult branch tips in September and December. However, no expression was detected for samples in October. *EF1 $\alpha$*  was amplified for the positive control. Lane 1: egg, Lane 2: prawn-chip (early developmental stage), Lane 3: pre-settlement (late developmental stage), Lane 4: branch tips (September), Lane 5: branch tips (October), Lane 6: branch tips (December), Lane 7: water for the negative control.

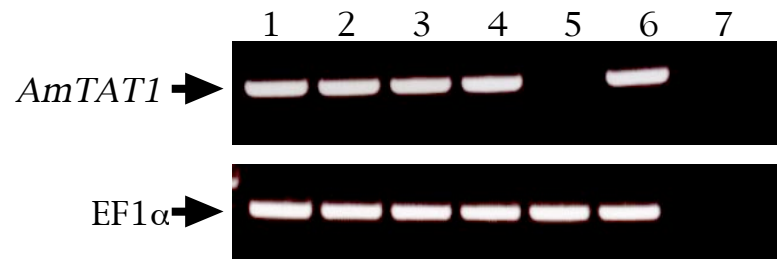


Fig.26

Schematic representation for the subfragments of DM-binding 5B clone designed for the analyses on DNA-DM domain association. The name and length of each fragment is shown. The primers for amplifying the DNA fragments are indicated by arrows. Hatched boxes correspond to the specific regions to DM-binding fragment 5B, which are not found in the genome.

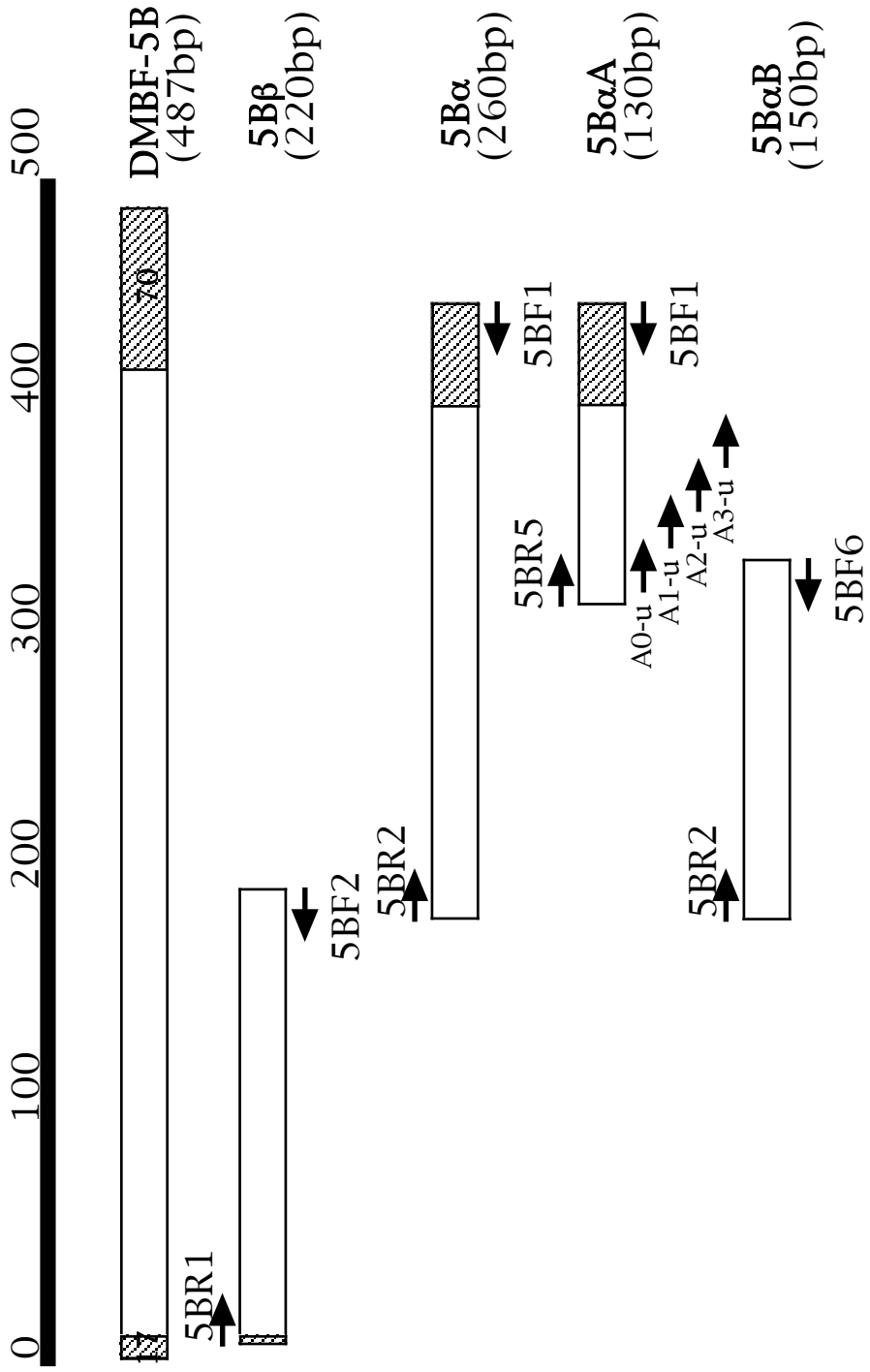


Fig.2.6 Subfranchments for the analyses on the association with DM domain

Fig.27

One-hybrid assay with the subfragment of 5B.

A: 5B fragment was divided into two subfragments, 5B $\alpha$  and 5B $\beta$  and inserted into the Reporter plasmid for one-hybrid assay in yeast.

B: Results of the one-hybrid assay. Assays with 5B and empty vector provided the positive and negative control, respectively. Massive growth was observed for 5B $\alpha$  and no growth was observed for 5B $\beta$ .

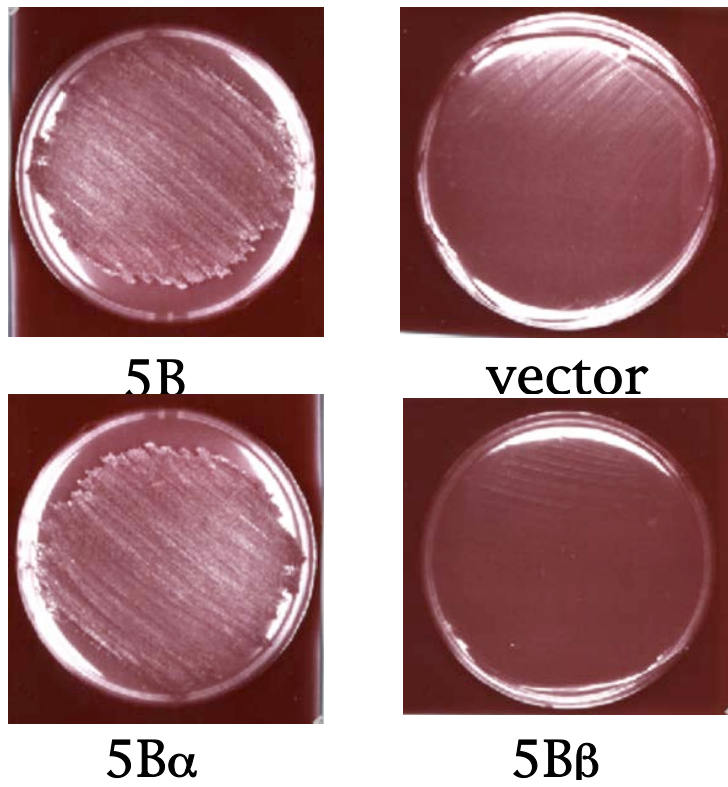
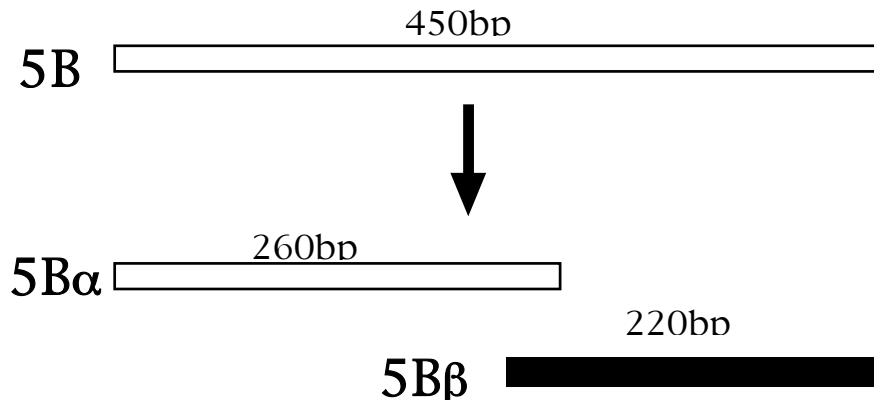


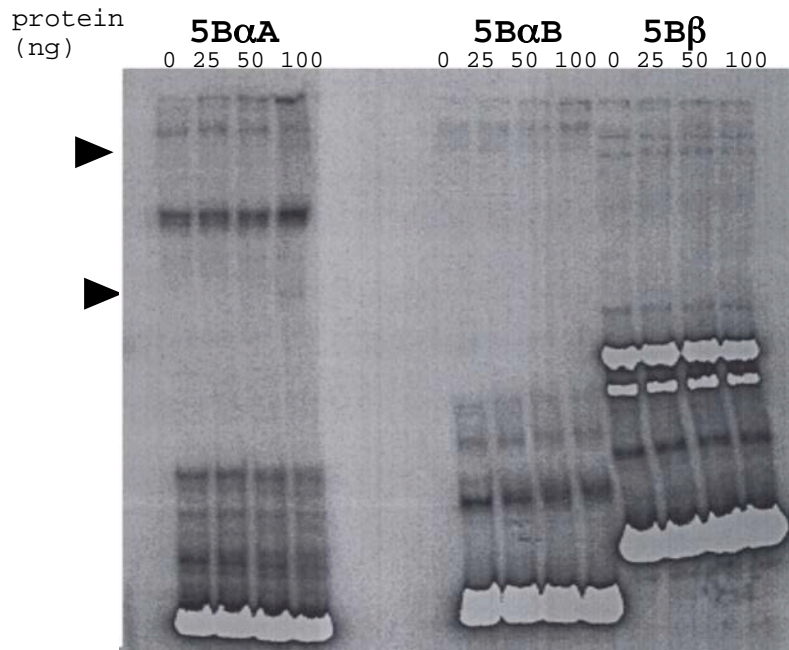
Fig.27 One-hybrid assay with subfragments of 5B

Fig.28

Electrophoretic mobility shift assay on the interaction of DM domain and the three subfragments of 5B. Only 5B $\alpha$ A was detected positive for the binding to GST-DM fusion domain. Arrow head indicate the retarded bands of the DNA-protein complex (lane for 100ng of protein). No retarded bands were observed for subfragments 5B $\alpha$ B and 5B $\beta$ . The lower panel shows the result with GST protein (not fused with DM domain) providing the negative control.



## GST-DM fusion protein



## GST control

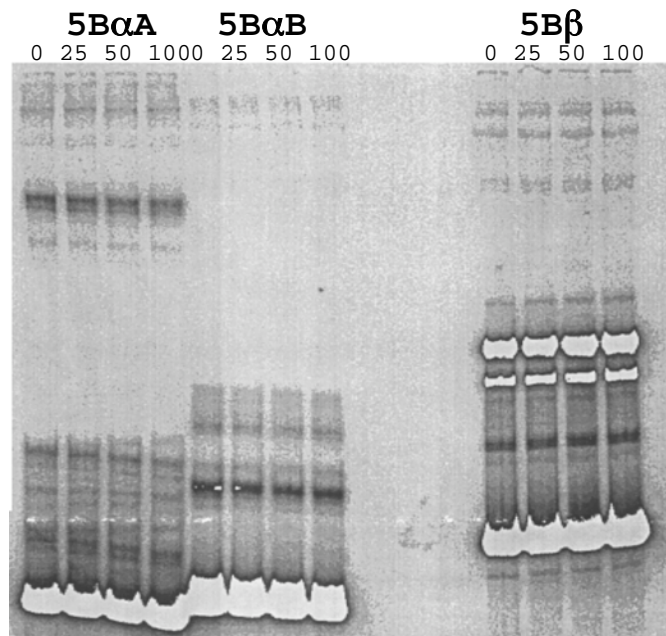


Fig.28 EMSA on the subfragments of 5B

Fig.29

The 5B $\alpha$ A fragment was further divided into nest-deleted subfragments 5B $\alpha$ A0, 5B $\alpha$ A1 and 5B $\alpha$ A2 with A0, A1 and A2 primer, respectively. The binding reaction was performed with 400ng of GST-DM fusion protein (+) or without protein (-) for the negative control. Retarded bands representing the complex formation were detected for all probes. The lower panel shows the partial nucleotide sequence of 5B $\alpha$ A fragment and the primers used for the amplification of the nest-deleted probes.

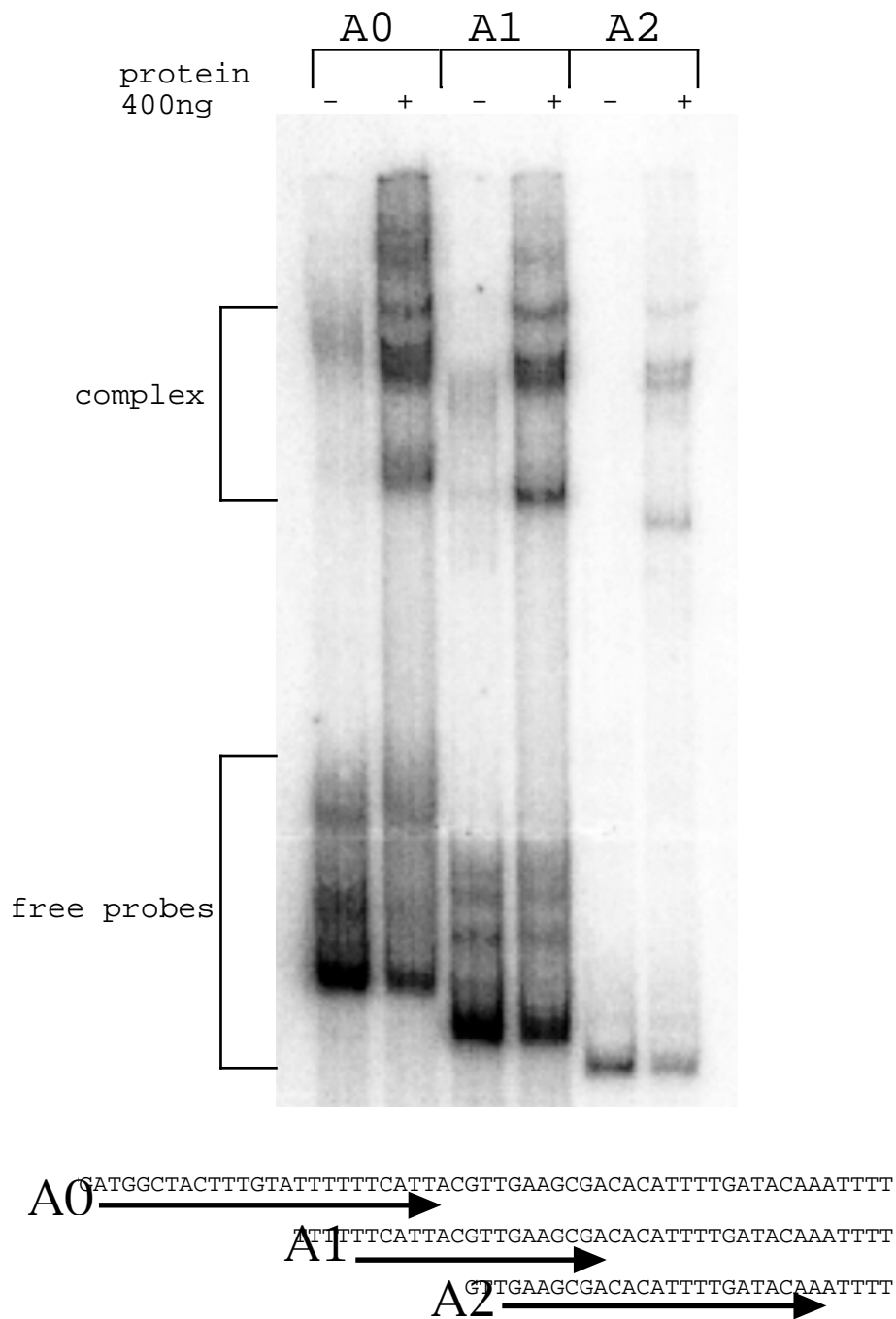


Fig.29 EMSA on the subfragments of 5B $\alpha$ A

Fig.30

400ng of GST-DM fusion protein was incubated with 5B $\alpha$ A2 probe in the presence of cold 5B $\alpha$ A2 or 5B $\alpha$ B DNA fragments as the competitors. The molar excess of the incubated competitors is indicated above each lane. Note that the intensity of signals for the retarded complex is decreased significantly by the addition of excess amount of the specific 5B $\alpha$ A2. Non-specific 5B $\alpha$ B does not inhibit the association of DM domain and 5B $\alpha$ A2 probe.

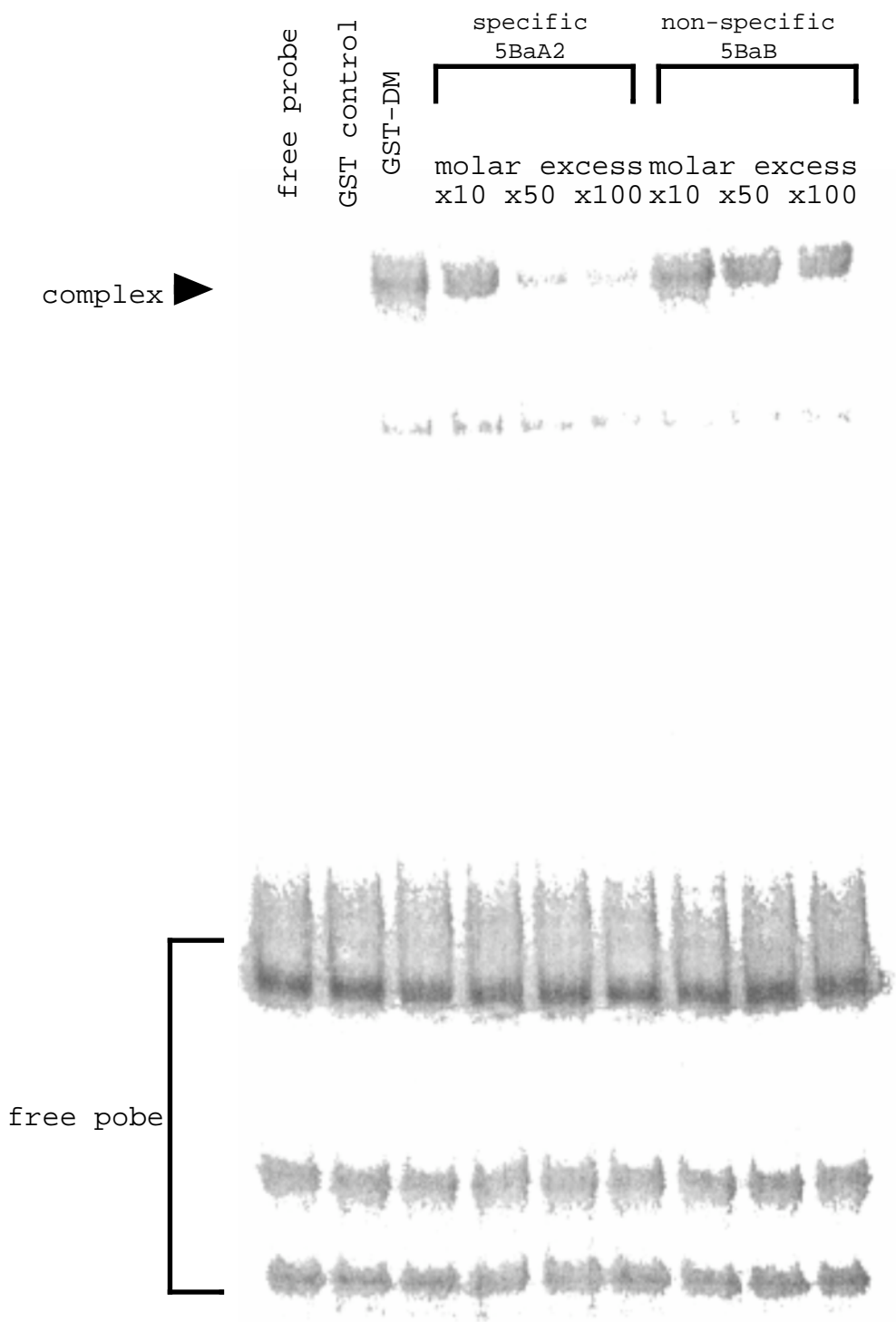
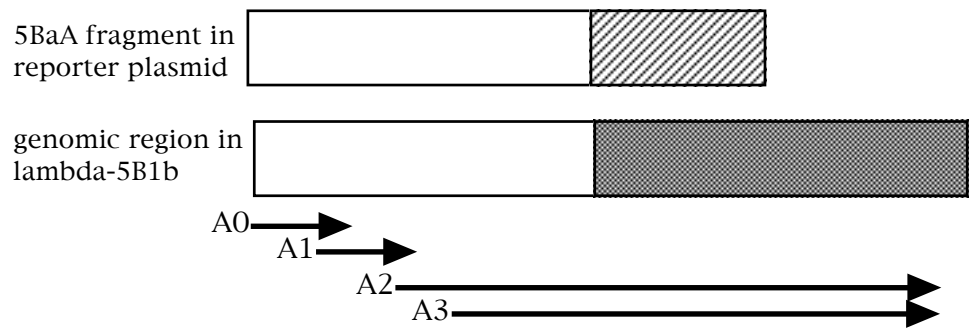
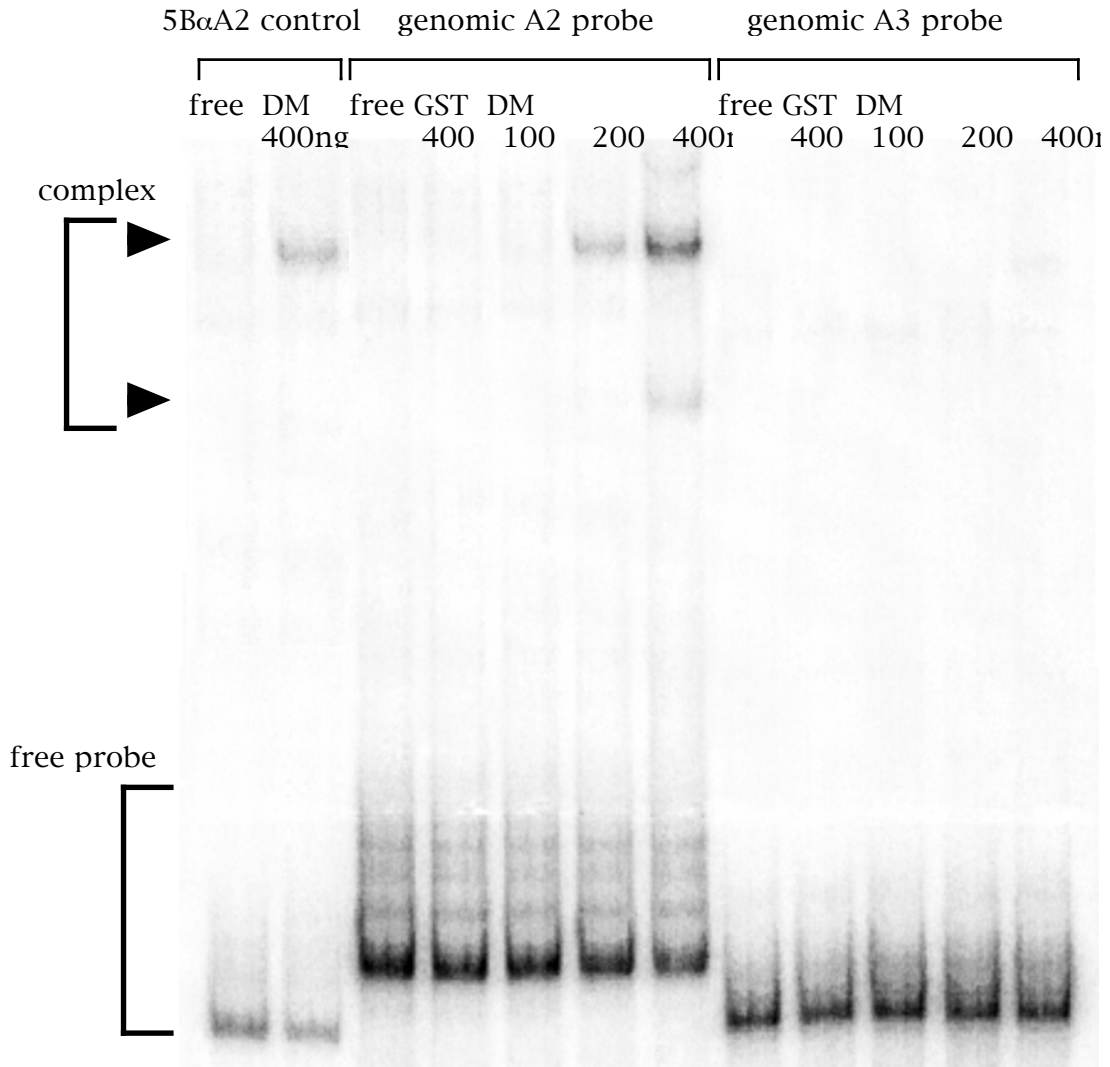


Fig.30 Competitive EMSA with specific and non-specific cold probes

Fig.31

Radiolabelled probes were generated with the genomic DNA template for EMSA. DM domain forms complex with the genomic DNA fragment containing the region for A2 primer but the mobility shift is not detected for the probe with further nest-deletion (A3 probe). The diagram below the phosphorimage explains the structural difference between the 5B $\alpha$ A2 probe which contains a specific region to the rescued Reporter plasmid (hatched box) and the genomic A3 probe which has a specific region to the genomic clone  $\lambda$ -5B1b (shaded box).



**Fig.31 EMSA with genomic fragments**

Fig.32

Nucleotide sequences of the nest-deleted probe. The primers used for amplification of the probes are indicated by arrows. The corresponding region to the nest-deletion is shown in the diagram above the sequences. EMSA studies demonstrated that the 18 nt. region in the A2 primer was responsible for the association with DM domain. Compare the nucleotide sequences of the *Acropora*-specific repetitive sequence and the 18nt. region. The pentanucleotide stretch (ACATT) common to both sequences are written in bold letters.



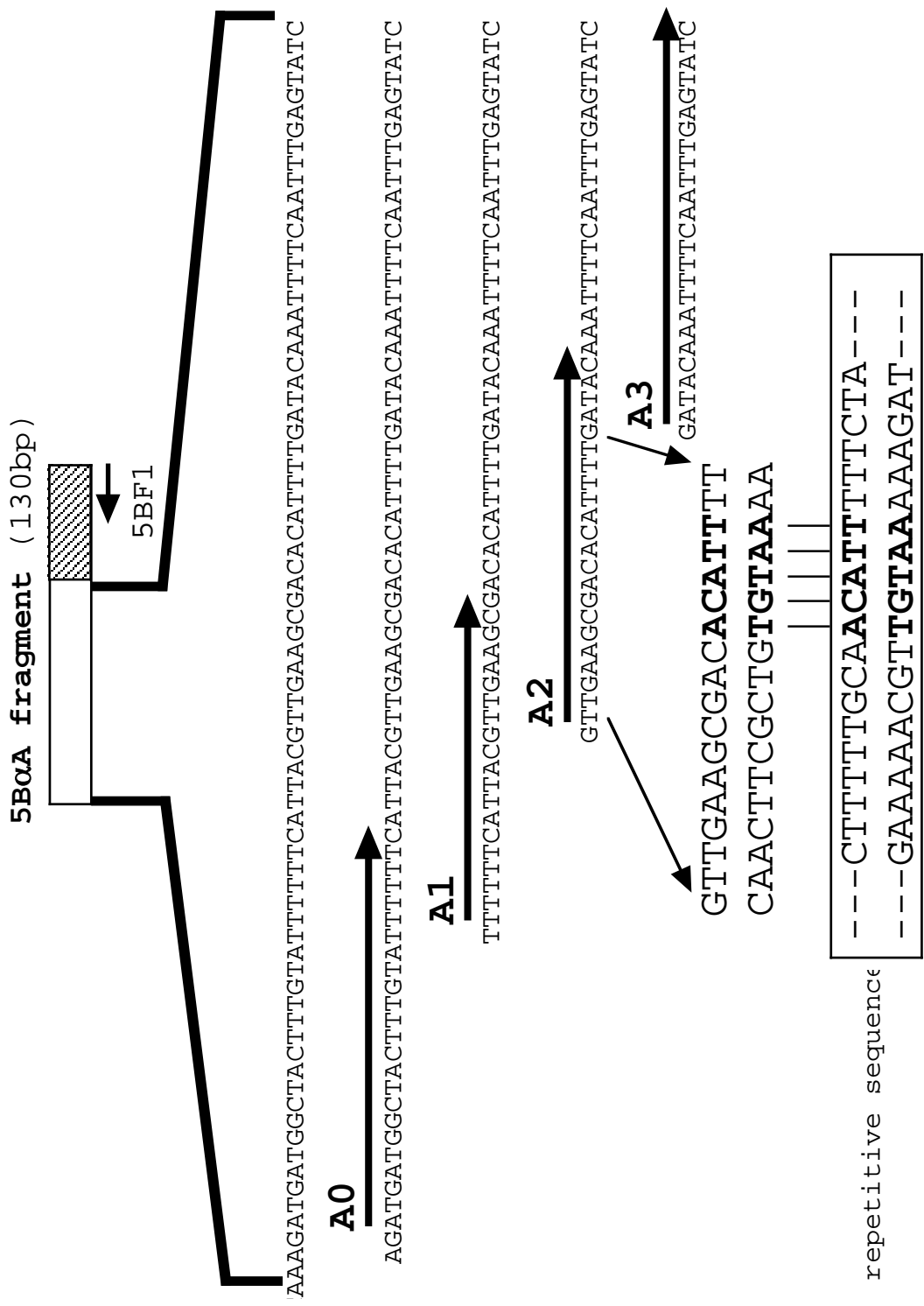
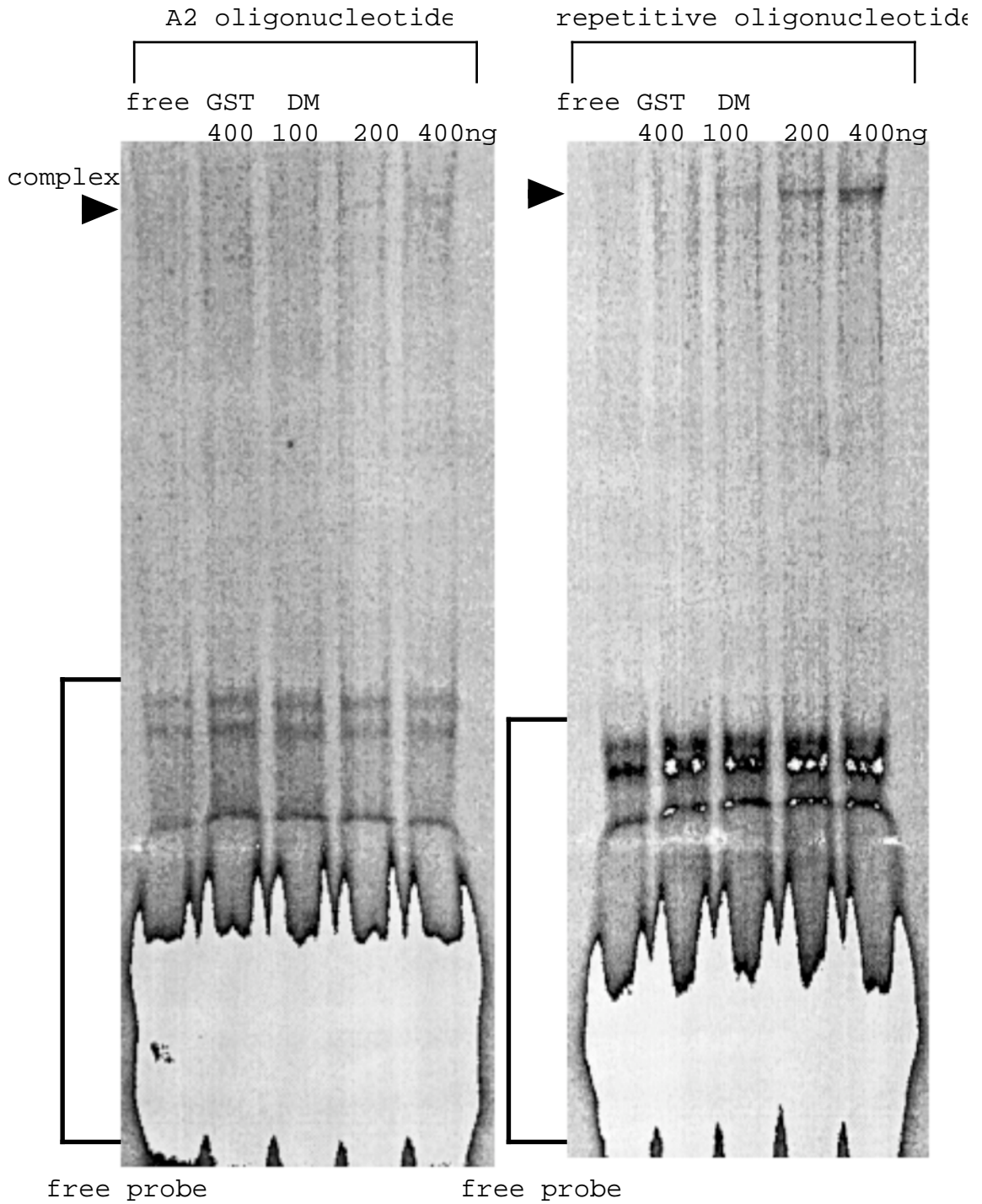


Fig.32 Nest-deleted probes for EMSA and the repetitive sequence

Fig.33

Double strand oligonucleotide probes corresponding to the A2 primer region and the repetitive sequence were incubated with GST-DM fusion protein and *in vitro* interaction was analyzed by EMSA. On the left panel, the DNA-protein complex (arrow head) is seen on the lanes for 200ng and 400ng of GST-DM fusion protein. For the repetitive sequence probe, the complex was observed even when the amount of incubated protein was decreased to 100ng (right panel).



**Fig.33 EMSA on double-strand oligonucleotides for the A2 primer region and the *Acropora*-specific repetitive sequence**

Fig.34

A mutant oligonucleotide probe was designed by substituting the 'core' element of 5 nucleotides (shown below the left panel). As seen in the left panel, mutant A2 probe did not form complex with GST-DM fusion protein. The cold mutant A2 oligonucleotide did not compete for binding with the non-mutant A2 probe (right panel) whereas non-mutant A2 efficiently decreased the strength of the signal for the DNA-protein complex (right panel, lane x10 to x400). The molar excess of the cold competitors is indicated over each lane. The retarded bands are indicated by arrow heads.

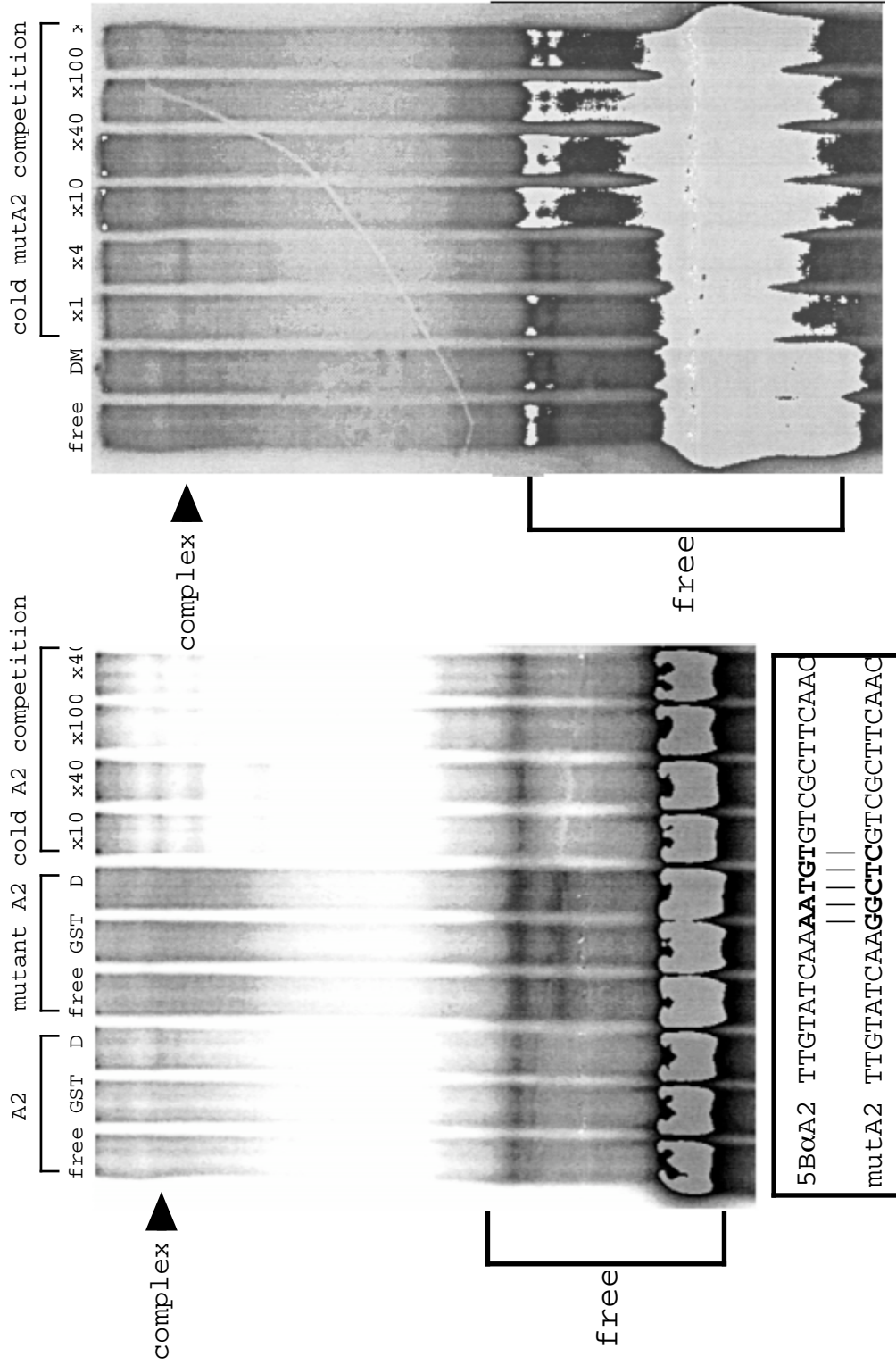


Fig.34 EMSA with the mutant A2 probe

Fig.35

Alignment of the recognition sequence of Dsx proteins of various organisms. The 'core' element is written in bold letters and the conserved nucleotides are indicated by bars. AmTAT1 and repeat: for AmDM1. Dsx: for Dsx of *Drosophila melanogaster*. BmDsx: for Dsx of *Bombyx mori*. Mab-3: for Mab-3 of *Caenorhabditis elegans*.

```

AmTAT1      TTGTATCAAAATGTGTCGCTTCAAC
              ||  |  |||||  |
repeat      TAGAAAAATGTTGCAAAA
              |  |||||  |
Dsx         ACTACAATGTAGC
            GAA          TT
            CT           A
              |  |||||  |
BmDsx       ACAAATGT
              |||||  |
Mab-3       AATGTTGCGATNT
                        A
    
```

Fig.36

Schematic representation for CAT assay to detect the enhancer activity of the DM-binding motif. CAT reporter vector (pCAT3 promoter inserted with the DM enhancer sequence) is co-transfected with AmDM1 expression vector into Hep-2 cells. Inside cells, the DM domain of AmDM1 protein is recruited to the DM enhancer sequence and regulates (promotes or inhibits) the transcription of CAT gene.

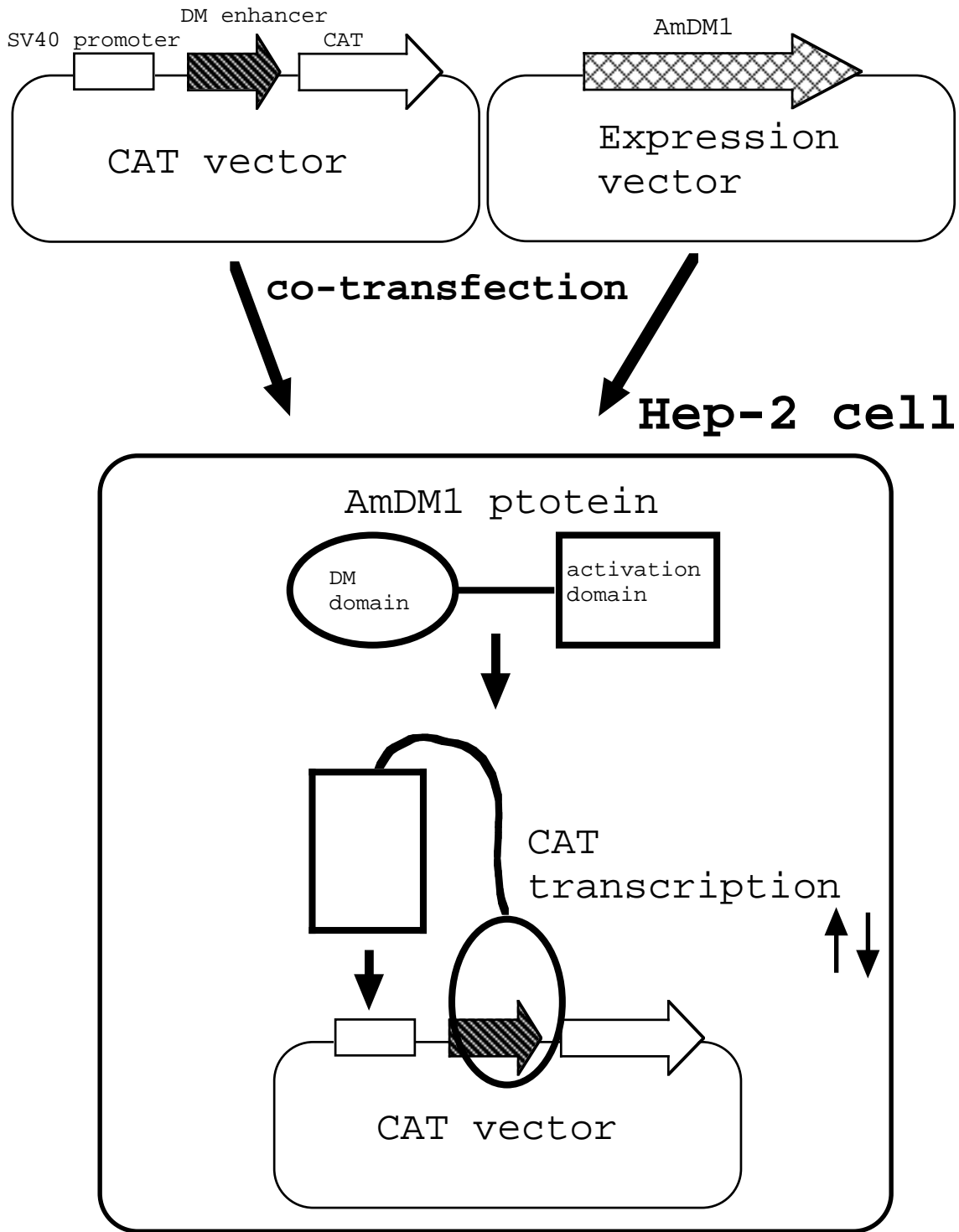


Fig.36 Strategy for CAT assay



Fig.37

A region within the 2nd intron of *AmTAT1* which included the DM recognition motif was amplified by PCR and inserted into pCAT3 promoter vector (immediately downstream of the SV40 promoter) with both forward and reverse orientation. The gene structure with the intron-exon organization is shown above the nucleotide sequence of the amplified region. The DM-binding motif (bold and large scripts) and the primers (arrows) use for the amplification are also indicated in the sequence.

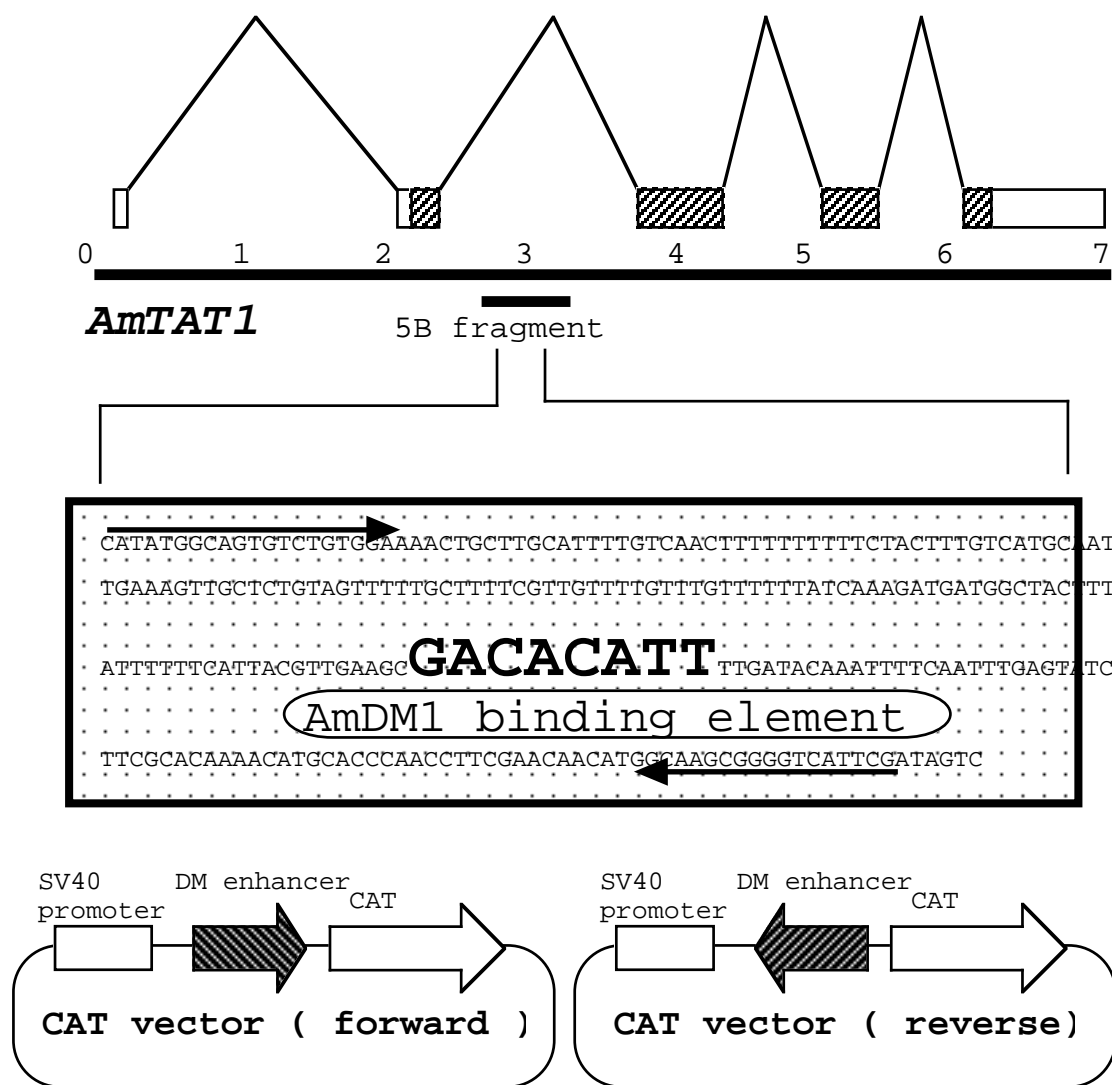
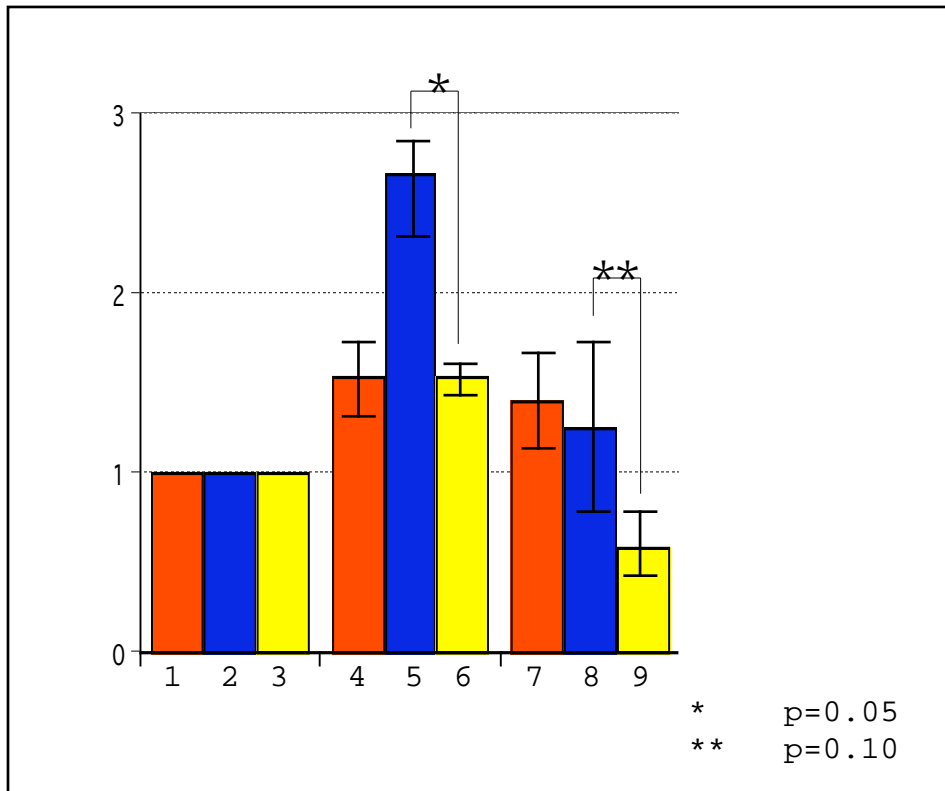


Fig.37 Construction of CAT reporter vectors

Fig.38

A: Relative CAT activities for DM enhancer with or without AmDM1 expression vector. The counts by liquid scintillation were standardized by dividing by values with empty pCAT3 reporter vector. The patterns of combination of CAT reporter vectors and AmDM1 expression vectors are shown in the matrix on panel B. \*: statistically significant at 5%; \*\*: statistically significant at 10%.

A



B

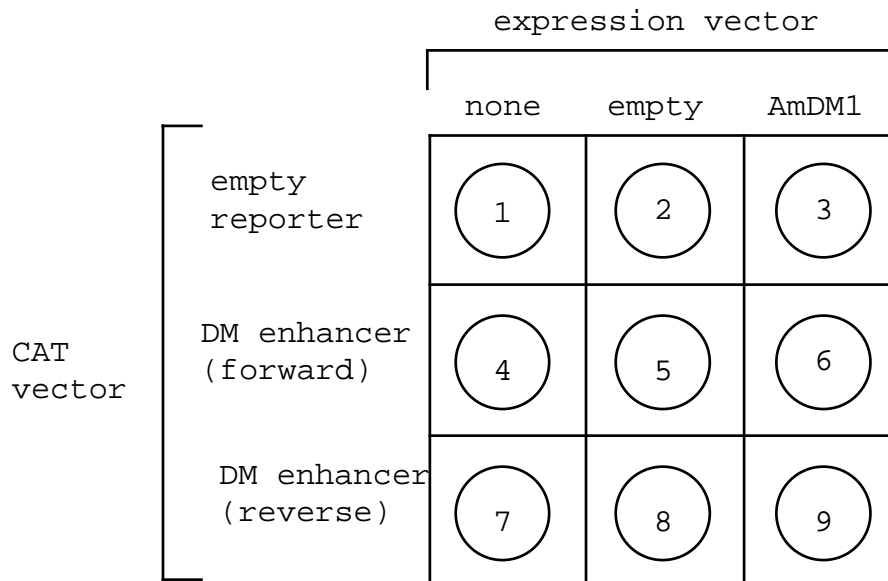


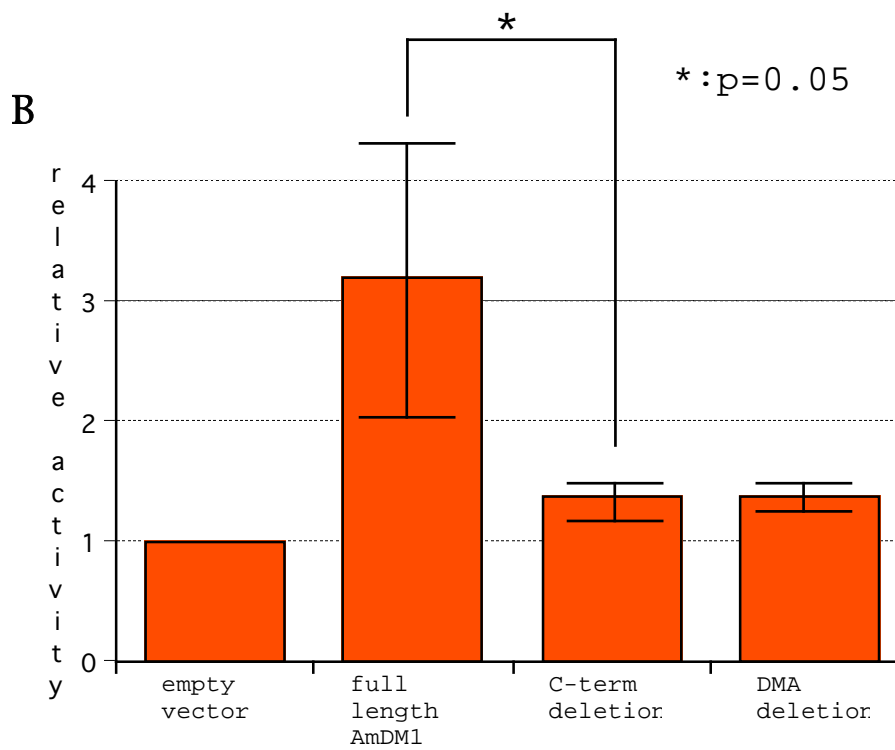
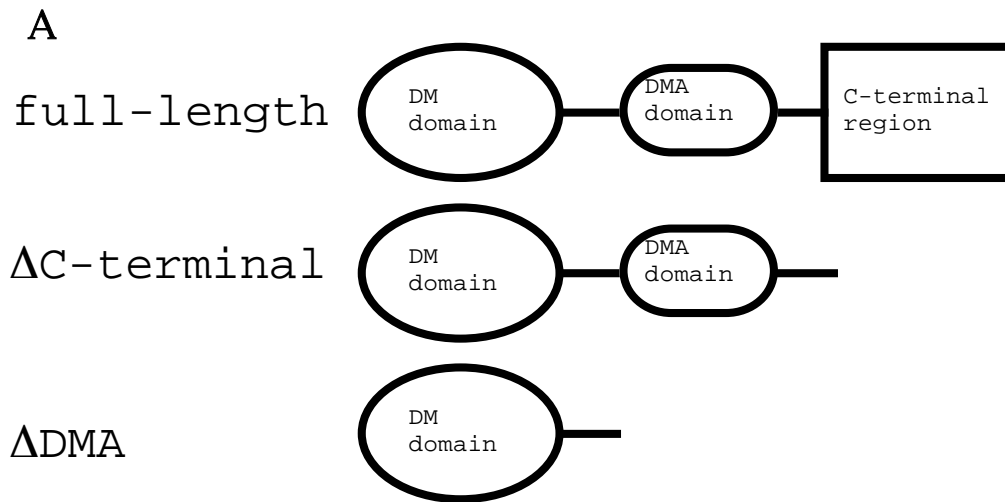
Fig.38 CAT assay on the enhancer activity of DM recognition motif

Fig.39

Deletion study to find the activation domain in AmDM1 protein.

A: Two deletion mutants were designed for AmDM1.  $\Delta$ C-terminal mutant lacks the region C-terminal to the DMA domain.  $\Delta$ DMA mutant lacks both the C-terminal region and the DMA domain.

B: Summary of the CAT assay with the deletion mutants. pCAT3 promoter vector was co-transfected with the expression vectors with or without deletion. The relative CAT activity was plotted on the X-axis. The activity dramatically decreased with both the deletion mutant. \*: statistically significant at 5%.



**Fig.39** CAT assay with nested deletion mutants of AmDM1

Fig.40

Deduced amino acid sequence of AmVit and a diagram for the conserved domains. Dotted box: Lipoprotein N-terminal domain. Hatched box: von Willebrand domain. Both conserved domains are written in bold letters in the amino acid sequence.

MKAFWVILFSLGAVSAMHLGYKVGQRYTHTYECGVTKAMPGSSTK  
 TVGLQLKCDAHVDVITKTEMQVTLANIQLNPVQGTRTNPTQQKDN  
 LKQNVLKAMQEQLSHPVKFFVKASDGSINKLRVHRKDTDWSLNVK  
 RGLVNLFQISGPLDQQYFIQPESTAVGECQVAYAIRRRTNAGVLE  
 LTKVINYNKCQHRAEFKQVNYLGEACQECQQSLGYSPYHRAVGQI  
 DHTLTGSKDKGLHHQPSWTLTKDHVITPF SRAAGQFTARARQTLK  
 VSKEHCSRRTIDRTNETVRMTFSAENPEAFYKYQKQQQQQIVKAR  
 ECIQEMFKAAKDIAGPNVGEKFAATVSGLRRCTKETLKTLDVDEM  
 KRKDQTERKLFMDAASFVGTCAFEVLKENAQLFKTRELSRLFLG  
 LAATPEPRACHIKALRDLCNHEKVTADQSCHRQCLLSGAVLNKI  
 SRRPSFQQNPKDVSEYGNLSKEMRSRLNKENAKSDEKRMYYIQALG  
 NAGSCQVQDDLQNILKNKQQPLFVRVECVWALRRITRQKRGKAKA  
 YPCLISIFADPEESPELMAIFVQILNTGPNFPTLQALASIVRRE  
 VSIGHRGPQSNQLASFVFSHLFALAYNNNILT KRRAMQARLVLRL  
 MPEMKYGLSYSKGVMAFNSENYQSGIEIEANKLDLPNSGLPRNL  
 NARLRFNILGYRLNALEFGARIENMDDIVDYIASKIHKRSKRSVM  
 GKICLILQSWTKWMTAMVIRQPWVAINPAKGLPFQSRVDKEIKQR  
 RMSAFVKFFGNEMTFLEVRQKDQDQLAQDLSDLLVGPCKKLEYPE  
 QGIEITRHGVKIEQTLQKAFLGYHARHMIPTLAGIMLDMEFRAQT  
 SAIARTTANFGVSSSSWLTFWQYNRMTGNLRITPRVNTHAHALVG  
 IHTPLVRVGIQMKLNIRSYFDKRVVVAEKGKQCSVKFNI PDRKW  
 DVWTVKVDTEGFTQECNPDTLKTTEKAI SMELNRDHITQLQRRCY  
 GKDMFGVQYCVQGTFPDLTALRLQQVQLLPVIGQFDGRFTGEPAE  
 DKPTDIEWLHKHPVISKLGTQEIHG EINVLAASKAVTRKIPYKIT  
 YKRSHQRELLIKFDDLQVKGYEKAMMTLKVDSNELKLNFGTKAGQ  
 KGPKYEITVSGQVENQGNVCEHG HAGTSFQSSGKTFSTNGSHRFG  
 MLFNSSLGLDES IKKRSNNEFQIQ LKTATTADWQLKVPNADV KRT  
 NVRLPLRVERFASSYDDIMNFVF **ARCDVQGGQNI R VFDHLQYVANI**  
**KAGCPYVLVQEHNTDRPSRIMFTVQMNGKGQKTLRILIRKSKERV**  
**EIKANDDHPIVLIDGQLCAKQDCQTKLQEVTVHTAQMPGGKNQVD**  
**LHTKFGLCARIKGGQISVYASPLLQGHVRGLCGDANGEQWNEFRD**  
 RKDKIRDRQTFVKSQKQC .

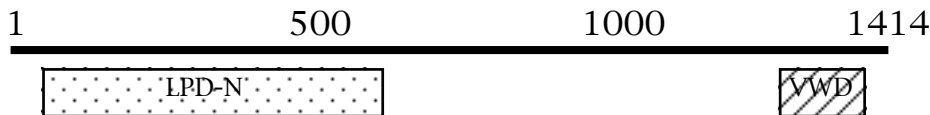
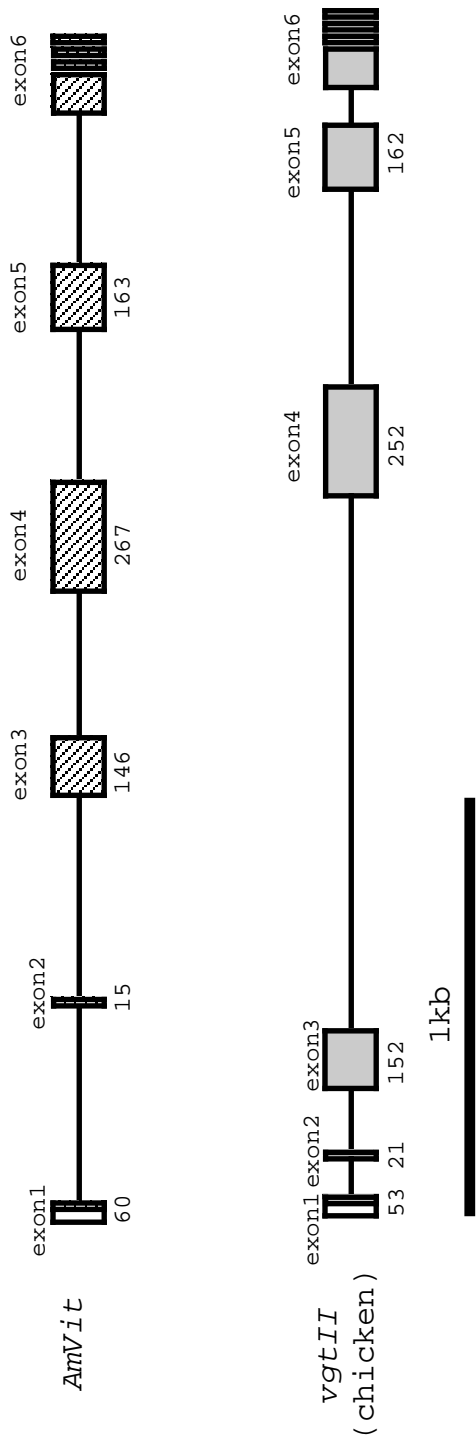


Fig.40 Amino acid sequence and conserved domains of AmVit

Fig.41

A diagram showing the comparison of the partial (corresponding to the N-terminal 235 amino acids of AmVit protein) intron-exon organization between *AmVit* (*Acropora millepora*) and *vgII* (chicken) genes. Note that the length of each exon is highly conserved. Hatched boxes and shaded boxes correspond to the translated region of *AmVit* and *vgII*, respectively. Properties of the introns of *AmVit* are shown in the table below.





	length	phase	3' of exon	boundary sequence of intron
intron1	443	1	TTT	GTAAGT-----TCATTACAG
intron2	494	1	TGG	GTAATG-----ATTTTTCAG
intron3	315	0	ACG	GTAAG-----CACATCATAG
intron4	454	0	GAG	GTTAGG-----GTTGTCTAG
intron5	364	1	CAG	GTAAGT-----TCCGTCACAG

Fig.41 Comparison of genomic organization of *AmVit* and *vgtII*

Fig.42

Nucleotide sequences of the 5'-flanking region, intron 1 and intron 2 of *AmVit* are shown with the potential binding factors. Binding motifs for AmDM1 were not found.

<5'-flanking region>

AAGCTTAAACAGGAAGTAGGTCTCTTAAATTGTTTTGTAAATAACTTAGC  
c-Ets-1  
TCTTTGTTGCAAACAGGGGCAAAAAAATATTTGTGAATTCAAAGGTCAAC  
GGCAATCCTAGGAACGCGTTCTAACAAGTTAACAAAACACTACACTTTTATG  
G TTCAGATTGTAAATAATTGCCATGTGTTTTATCTGCTCTCGAGAAGCA  
GATA-1  
CTTTTAGCTACATTTTGTACAACATGGGAGATTATGCTCCGAGTTGTTAA  
TTTTAAACTTATTTCTGGAAGACTATCCACAGATAGAAATCCCATTG  
STATx HSF  
ATTAAATTATGGCCGTTTTTGATATGCTTATTAAAACAGAAATTAATAAC  
CdxA NIT2 HSF  
GTGCTGTCTCGGAATTCGCGAGGGTCTAGTTTATGATAAATGTAATGAAC  
CCAAATTCTATAACTTCTTGCAGGACTTGTGTGGCTGTTTGTATCAGAC  
GTTTCTGTTCCCTAACTGCGCTCTACTCCGGTACAAAGATCGAGAATGTCT  
HSF  
GGCTTATTGATCAAATCAAATAAAAAGCCTTTATGTTGTGACTAAGAGGA  
CdxA  
GTTTCAACGCGCACTTTGAGACTTTGCGACAGAATATATAAGCCCTAGAA  
StuAp TATA-box  
TCATTTGCTTGGCTACTTGTGGGATCG

<intron1>

GTAAGTACTATAGCGTAGTCCTCTTGATTTGTCATGTTGCTCCCTTGGAT  
TCF11  
CGTTGACTCAATGGGTAATGGAAGGAAAGCATCGTCAGGATTTGAAAAAT  
ACTTGAACACTACAATAATATCTCCTTTTTCGCTTTAGTCTTTTCAGTGAA  
NIT2  
CdxA  
TGACGAATATTTTTGGTAAGAAAATTACCACAACACTGAATGCTGAATTTGA

HSF AML-1a  
ATTATTCTGAAGCATCAAAGAACTATTCTGAATTCAGTGAGATGACGATTAC  
HSF Dof 3  
TCF11  
GAACTTATGCCAGGCTTGGTGTCAATCACAGTTAGCCATTTAATTT  
TCAAGACAGCTCCCCAACTACTGATTTTCTCATAATGACGTTTGATCTA  
HSF  
Oct-1  
CTTGATCCTTTTGACAAATTTCTGTGTTAAATCACGAATTTTCATGCCCTCA  
TTACGCGGAGTTAACTTAGTTTTTAATCTTATTTTCATTTACAG  
Oct-1 SRY

**<intron2>**

GTAAGTACCATAAATGTACTTTGCTGCTCTAGCTAGCCCGCACTAAATTTCTC  
CdX HSF  
TTTTTGTAGTAAGTAACTGCAAAAAACAGCGCATTTTCTAGTAATAACTTTGT  
HSF  
CTCCACAATATCCTTCAACGTTTCAGCTCAAAAATTGTGGATAGTTAGTGC  
GATCCAAACCAAATTAATTATGCAATAGATATCAAGCTTCTAACCATAGC  
NIT2  
C/EBP- $\alpha$   
CATCTGCAGGAGGAAATCCGGAGCCTCAAGCAGGTACTGCGGCGAACATA  
GACCTAATTACCTACCAATTAGAGAGAACTTCGAGAAAATTTACTTCTAT  
HSF HSF  
TGTGATATGAACTATGCATGAGCCGCTGCTGCAAGTCTGTCTTTTATGC  
Sox-5  
TGTGATTGATGTTTCCAAGTTTCAACAGCAAACCTTGATAAGGAGGAAGGC  
GAGGATAAGGAGGATCAACCCCAGCAGTATTAACCTCACTTGCTAGATTCT  
ADR1  
TATTTGAAACACTGATTCAAAGTTATCATGCCAATTTTTTTCAG  
C/EBP

Fig.43

Nucleotide sequence of *Cnnos1*(*Am*) cDNA. The deduced amino acid sequence of the open reading frame is also shown. The amino acid residues corresponding to the Zinc finger domain are written in bold letters. The putative polyadenylation signal is indicated with bold letters and an underline.

```
1  GAAAAGCTCTCGGTCGGCAGCCGGTACGGTGGACGGATCGTTGTATAAATTGCCATTTTCAT
61  GCACACATTAAGTCCGAACACTCACGCTTGTAGGTGCTTCCAAAGCTTTTGTGGCTTT
121  AAATCCCCTGCCCTTTATTACAAAAGTTAATCATCACGATTCGAAAACGACTGGAAACTG
181  AATATATACTTTTCGCTCTAGCTGCCTCGGTCACGCGGAAACGTCGCTGAGGGAAATTTTCG
241  TTAATACAGTTGCTAGTCTGCACGCGATAAATCACTCATCGAACCAGGAGCCAGGAACAA
M  G  D  E  M  L  S  V  F  Y  S  G  D  I  F  P  S  H  T  K
301  TGGGGGACGAGATGCTGAGCGTGTCTACAGCGGTGACATTTTCCCAGCCATACTAAGA
M  I  S  A  R  E  V  A  S  L  K  V  T  F  D  E  K  S  E  F
361  TGATAAGTGCCCGGGAAGTAGCAAGCCTCAAAGTCACTTTTCGATGAAAAGAGCGAATTC
R  L  F  R  D  Y  F  G  L  I  R  L  V  Q  N  A  S  D  L  E
421  GACTCTTCAGGGATTACTTTGGCTTAATACGACTAGTTCAAAAACGCGTCGGATTTGGAAA
R  A  E  P  G  F  D  S  L  Y  D  F  T  Y  T  K  R  S  R  V
481  GAGCGGAACCCGGATTTGATAGTTTGTATGATTTCACTTACACCAAAAAGATCCCGGGTTG
D  S  M  G  S  D  H  S  L  T  S  S  S  S  G  G  N  S  E  I
541  ACAGCATGGGATCAGATCATTCTTAACAAGTTC AAGCTCTGGTGGAAACTCCGAAATTC
E  L  C  I  E  A  S  K  T  P  P  P  S  L  A  P  G  A  P  T
601  AACTTTGCATAGAGGCGTCTAAGACGCTCCGCCGTCGCTAGCGCCCGGTGCGCCAATC
R  V  S  K  R  N  I  E  N  S  S  D  V  A  N  R  N  R  K  N
661  GAGTTTCGAAGCGGAACATTGAAAACAGCAGCGATGTAGCCAATCGCAACCGGAATC
R  E  S  K  K  N  R  N  A  N  V  C  V  F  C  R  N  N  G  E
721  GGGAAAGCAAGAAAAATCGGAACGCAACGTGTGCGTGTTTTGTAGGAATAATGGCGAGA
S  K  K  V  Y  S  S  H  V  L  K  D  A  E  G  N  T  T  C  P
781  GTAAAAAAGTCTATTCAAGCCATGTTTTGAAAAGATGCTGAAGGCAACACAACATGTCCTA
I  L  R  A  Y  T  C  P  L  C  K  A  S  G  N  E  S  H  T  I
841  TTCTGCGCGCATATACCTGCCCTCTTTGCAAAGCGTCCGGGAATGAATCTCATACAATCA
K  Y  C  P  K  N  K  A  A  S  K  Q  Q  Q  Q  Q  N  A  T  A
901  AGTACTGTCCCAAGAACAAGGCCGCATCAAACAACAGCAACAACAAAACGCAACAGCAG
V  A  A  V  G  Q  V  *
961  TTGCTGCAGTTGGCCAAGTTTGACCGAACTATTAACACTGAAAGATCACAAATACACGTG
1021  GGTCCATTTTGAATCGCTGAGCTATGTGCTCAAGGTACATGCACCTGCAGGAGGAAGTTA
1081  ACTGAAACGGTAAAAATATTTCTCGCACAGTTTCTCAAAGGCTGATGCCCTGAAAAACA
1141  GGCCTCTAAACTATATGGATAAATAATAGCGGGATATTTAACCAGTATTCGGATACACCG
1201  TTTAAGTTTTTATAATTTGATGATACAAACCTTAATATTTTGTGCTTGTGTAGTCATGTCC
1261  ACAAATAAAAAATAAAATAACTGTAGATTTTTAGATATTTGTAAAAATGTCATGACTCTTTA
1321  TGGCACATAGCTCAAATGCGGCGGTAACTTTCATTTGGGCTCTGCAATATTTGCCTTATTT
1381  TGCAAGTTCGCTATTTTGTCTGATTTACCCAGTAAAGAAATTTTGGGCTCAACGCCCTCT
1441  TTTCAAGCTATGATAAAAATTTCAAAGAATTCGGAATATCAGTAATTTATGTAAACCGAA
1501  ATTTGCTCATACGATATTCGGTTTTTTCAGAAATTCATAACTTATACAAAATTAATTTGTGT
1561  ACAGGACTGCGGAGAAAAATATGTTAACTTTTGGTGTGAAATCATATAACCAGTGAATA
1621  AAAATTTGAAGCACAAAGTGAATTTCTAGTTTTTAAGTCAAAACAAATAAACTATCGACAG
1681  GTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Fig.44

Nucleotide sequences of *Cnnos1(Nv)* (A) and *Cnnos2(Nv)* (B) cDNA.

The deduced amino acid sequence of the open reading frame is also shown. The amino acid residues corresponding to the Zinc finger domain are written in bold letters. The putative polyadenylation signal is indicated with bold letters and an underline.

## A *Cnnos1(Nv)* cDNA

1 GAGAGCGGCACCCAACTTCCTCTCCGCACCCGGGAAAGGCTAATAAGAGGGAGGCTATAA  
61 AAACAGCGGGCCAGACTTGAGTGTGAGGGGACTTGATGCAAGAGTGATTATGAATACTT  
121 GAAGGAGGTGTAAGAGCTGCTAAAAGCGTGCAACAACCTGTCAACAACACATCGACTTTTG  
M G D E M L S V F Y H G D L Y G G  
181 GGCTCGAAGCATGGGAGATGAAATGCTAAGTGTTCCTACCACGGTGATCTTTATGGTGG  
S G G K L S M G G V E T F K D T L D D Q  
241 TAGTGGCGGTAAATGAGCATGGGAGGAGTGGAGACTTTC AAGGACACCTTGGATGATCA  
D E F R L F R D Y F G L I H L V Q N R G  
301 GGACGAGTTTAGACTCTTTTCGCGACTACTTTGGCTTAATTCACCTCGTTCAAACCCGCGG  
V V E D E D S V V D Q F F N F T Y K R T  
361 TGTGGTTGAAGACGAAGACAGCGTGGTGGACCAGTTTTTTAATTTACGTATAAACGCAC  
R M D S I G S D R D S V T S S G S S A G  
421 TCGCATGGATAGCATCGGCTCCGATCGTGATTCTGTCACGAGCTCCGGCTCGAGCGCGGG  
S A E F E F K M E P H F G S K A V G S S  
481 GAGTGCCGAGTTTCGAGTTTAAGATGGAGCCACACTTTGGCTCCAAGGCTGTTGGGTCGTC  
L L P G T R S L T K R T D I D T A A R N  
541 GCTTCTTCCAGGACAAGTCCCTAACAAAACGAACTGACATTGACACCGCGGCGCGCAA  
R K N R E N K K N R N A N V C V F C R N  
601 TCGCAAAAACCGCGAGAACAAGAAAAACCGCAACGCGAACGTGTGCGTGTTCGTAGAAA  
N G E S K K V Y S S H V L K D A E G N T  
661 TAACGGAGAGAGCAAGAAAGTTTACTCCAGCCATGTGCTCAAAGATGCGGAGGGAAATAC  
T C P I L R A Y T C P L C K A S G S Q S  
721 TACATGTCCCATTCTGCGGGCGTACACATGCCCTCTGTGCAAGGCCTCGGGAAGTCAATC  
H T I K Y C P K N K N G S K L Q A K V \*  
781 ACACACCATCAAGTACTGTCCGAAAAACAAAAATGGGAGCAAACCTTCAAGCAAAAGTCTG  
841 AAAGCGTCTCAGGAACTATGGAGAATTGAGGGTACACAAAAGTTAAGCCTTTAAATAATGT  
901 TTGAATTACCTTTTAGCAGAATTGGCTTTTTGTTTGCTAATGTGTTAGGCGAGTTTTGGG  
961 TTTTGCAAGATATATTAATTTGTAACACTCATTGAACGGACCTCCCAGGGAAGGGAAAA  
1021 TGGGAAACAGTCATTCGCATACTCTTCATACTAGTGACCAAGGTACAAGCCTCAACTATT  
1081 GAGGTAACCAACAACAAATACAAGAGTCAATTTAAAAGGAAAAGTTAATTCAAACTGTCGA  
1141 TGTGTTAGAAAGCACTAGACGATAATGTGATAAATAATTCCTTTAACGGCTGCAATATGTT  
1201 ATTGAAACTCCCCTTATTAGTGGGAATACAATTGATTTTTTCACAACACTTTTTATCGCGTT  
1261 AAGCTTTTTTCAAATTCATGTGATTTCATGGGTATCACAATCTCACTGCTAAACAGAAAA  
1321 GTTGCACCTCAATTTGCGGCGGGTTGCCACAAAATAACCGTCAAAAATTTCTCTATTAAT  
1381 TCTATATCAAATATTATACGAATCAAATGAATCCAACGTATGGAATATCAATATATTTCT  
1441 CAAACAAAAACCTTTACCCTATTGGTAAAGTGATTATTATTGTTGATTGTAAAAATATAT  
1501 TTAGATCAATAATGTGCGGTTAGAAAAGAAATGCAGAGAGATATTCATTGAAATTCCTTA  
1561 AATCCATTGACCTTTGGTATTACCAGTCAATGCAGTTGTACTCTATATACTCAAATCAAA  
1621 AAACGTGTTCTGTACGCGTTAGACTTAATGTAATAAGTTTATTTCTGAAAAATTAACGCCCT  
1681 ATTTGATATAGTCTCCAAGAGACCTAAAGAATGGTAGTTCTTATGTCTTGTAGCAAAGTTC  
1741 CTGCTTGTAAAGCATGTGTCTGAGAAAAGATCGATTTTGAATATTTGACCATGCTGGTGT  
1801 TGTATTATGGCTAAGGGCTTTGTGCATACAAATCTATTAATTCATTCTTCATCAATTTGTT  
1861 ATGG**AATAAA**TCGAAAACGCAAACTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## B *Cnnos2(Nv)* cDNA

1 GCTGTTTCACATGCGAGTAGTGCGTTGCATTGACAAAACAACCGCAACTGCCCTGTGGCGG  
61 CATTACACTGTAATTTTTTACGCCTCTGATCACGGCTGGAACCCAAGCAGAAAAGGCCACTC  
M S L S E V M  
121 TGCACTGTCAAACCTGGTTATCGCTAAGCTGCGTCCACGATGTCTTTGAGTGAGGTGATGT  
L K T L I P S N F L D S D D E E I K S L  
181 TAAAAACGCTGATTCCCAGCAACTTTTTGGACTCAGACGACGAAGAGATAAAGTCGTTAA  
T C T T E S D D L Q K N Q Y T N Y N M F  
241 CGTGCACTACGGAGAGCGATGATTTGCGAAAAATCAGTATACCAACTACAACATGTTCA  
N D Y F G L S N L V Q K V R L A N N D F  
301 ACGACTACTTCGGGCTCAGTAATCTGGTGCAAAAAGTGCGGCTCGCTAATAATGATTTCC  
P Y L G S D G Y G G K L H R R D R F D S  
361 CGTATCTCGGCAGCGATGGCTACGGCGCAAACTGCACCGTCGTGACCGATTTGACAGCA  
I D S E N S S S G S C D S F T D H G L T  
421 TTGATTCGGAGAACTCGTCTTCTGGAAGCTGTGATTCCTTCACGGATCACGGCTGACAT  
F G E G S F Q P F S T S S I L H S A G K  
481 TCGGCGAGGGATCATTTCAGCCATTTAGCACATCATCTATCCTGCACAGTGCAGGGAAAA  
R K E V A L P P S R R L G K P T A R S S  
541 GGAAAGAAGTTGCGTTGCCGCCAAGTCAAGGCTCGGAAAGCCGACCGCGCGAAGCTCAG  
A P G A N R Q V C V F C R N N G E S E E  
601 CTCCCGGTGCAAATCGTCAAGTGTGCGTGTCTGCGAGAAAACAACGGTGAGAGCGAGGAGG  
V Y A S H V L K S A D G K T T C P I L R  
661 TGTACGCTTCACATGTGTTAAAAAGTGCCGATGGGAAAACTACTTGCCCGATCCTGAGAG  
A Y T C P I C K A S G D D S H T I K Y C  
721 CCTACACTTGTCCGATTTGCAAGGCAAGCGCGCATGATTCGCATACGATAAAAATTTGCC  
P Q N Q Q T Q G N G Q L P P P P V K P P  
781 CGCAAAATCAGCAAACGCAAGGAAATGGCCAGCTTCTCCTCCACCAGTCAAACCACCAA  
T S T A Q P I A R S T R G I M P P K R R  
841 CCAGCACAGCACAGCCAATCGCTCGTCCACGCGGGGAATAATGCCGCCAAAACGCCGCT  
\*  
901 AGACTAAACTCCAACAACCTACCATATTTAAATAGTTATACCCTAACAGGGGTTTCGTGCACA  
961 GAGTTTGTAGGAAATTCATAGTTTATATTCACATTTTGTGTTGGGATTCGCTACGGCGACGG  
1021 TTGTTGTGTTATTTGTTTACCCGACATGGCGATGTGCTAATTTATTAGGAGTTTCTTAAAC  
1081 ACACGACTTCTACATGAAAAATGTAGAAAGGAATATTTATTTATCTTAATTCACGTTGAA  
1141 CGATCAAGTGATCGCAGTATTTTTCGTCAGATTATCGGATGCGAGGCTTTTCGCTTTTCGAT  
1201 TTTTGCAGCCGTGAATTTTCATGAATGAAATGGCGCGTTTTTTTTTCGTACGTACTATTATT  
1261 ATAAGGAATGATATTTTACTCACAATCCACTTACTACCCAGTCTCGCCAGACGTTGTGTA  
1321 AATTTCAAGAAATCCAGTCTTGCTCTAGAGGCTTAATTTATCGAATATATACCATTGTAAA  
1381 ATTTTTGAATTTATAGCAGAAATATTCTAAGAGATATTGAATTTTACCCTTTAAGTGTA  
1441 GTATTTTAATATCAAGTGTGATTGAGAAAATTTCACTGTAAGGATTCAGAGAAAATGATTT  
1501 AATAATGTTTTTCAAATTCCTGCTCTTAGTTTCAGGGACGTTATTTTATGCTGTTTGTAA  
1561 AAAAAAATTCAGTGCAAAAAATTTGACACTACTTGGAAAAGTAGTAGCAATGACTTTAGTA  
1621 GAATTAAGATTTTTCAATGAAATTTATTATAATCTTTAAACAGCATTTGCTGTGAGTTAC  
1681 AGATGAAAAGTTAAGTTTAATATTTCTAAAAAGTTTTTAATATTTATGGCAGTAATCCTTC  
1741 GTATCATAACTTTTAGAAGTTTGAAGCATGTAAAGCTTATTTATATAAAATTCCAAATAA  
1801 GAAACCAGAAAATTTGTACGGTTTGATCGTAAAGCATTAAAGCTTGTTTTGAGAAAAAAA  
1861 AA



Fig.45

Boxshade representation for the alignment of Zinc finger domains of Nanos proteins. The conserved cysteine and histidine residues are indicated with arrow heads. Am1: Cnnos1(Am), Am2: Cnnos2(Am), Nv1: Cnnos1(Nv), Nv2: Cnnos2(Nv), Pc1: Pcnos1, Pc2: Pcnos2, Hm1: Cnnos1(Hm), Hm2: Cnnos2(Hm), Sponge: PoNOS, Dm: Nanos (*Drosophila melanogaster*), mosquito: Nanos (*Anopheles gambiae*), Nasonia: Nanos (*Nasonia vitripennis*), Sa: Nanos (*Schistocerca americana*), human1-3: Nanos1-3 (human), rat1: Nanos1 (rat), mouse1-3: Nanos1-3 (mouse), Xl: Xcat-2 (*Xenopus laevis*), Xt: Xcat-2 (*Xenopus tropicalis*), fugu\_GAG01247,04666,06036: hypothetical proteins (fugu), zebra: Nanos (zebra fish), leech: Hro-Nos (*Helobdella robusta*), Ce1-3: Nanos1-3 (*Caenorhabditis elegans*).

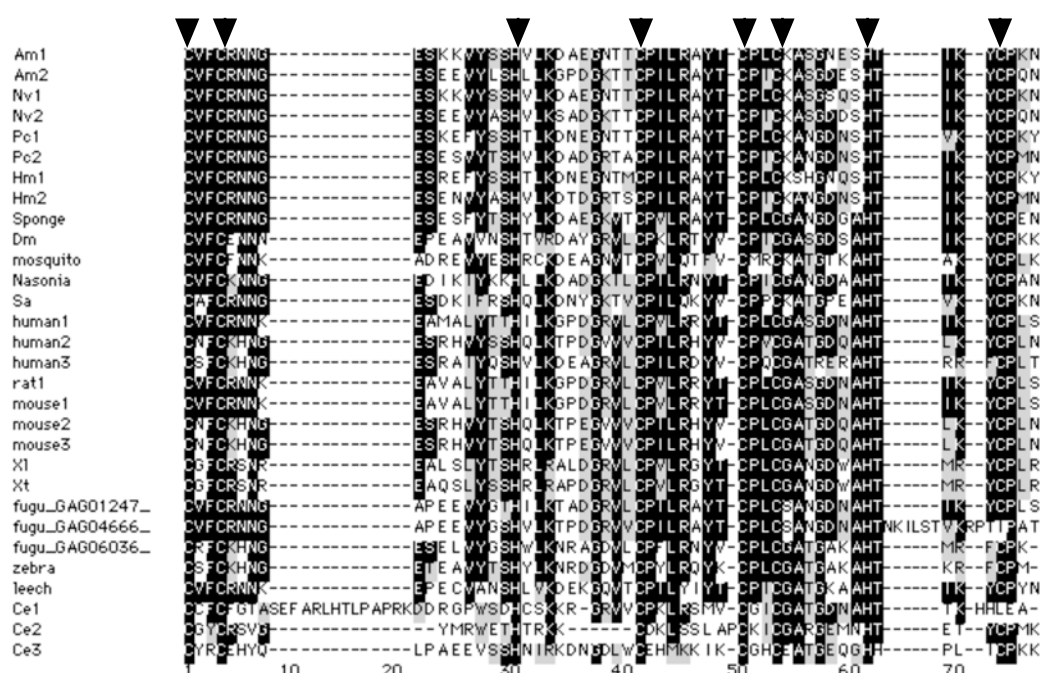


Fig.46

Neighbour joining tree for the Zinc finger domains of Nanos proteins. Proteins analyzed for this tree are described in Fig.45. Note that the cnidarian Nanos proteins form a well-defined clade and two Nanos homologs show congruent phylogeny with each other.

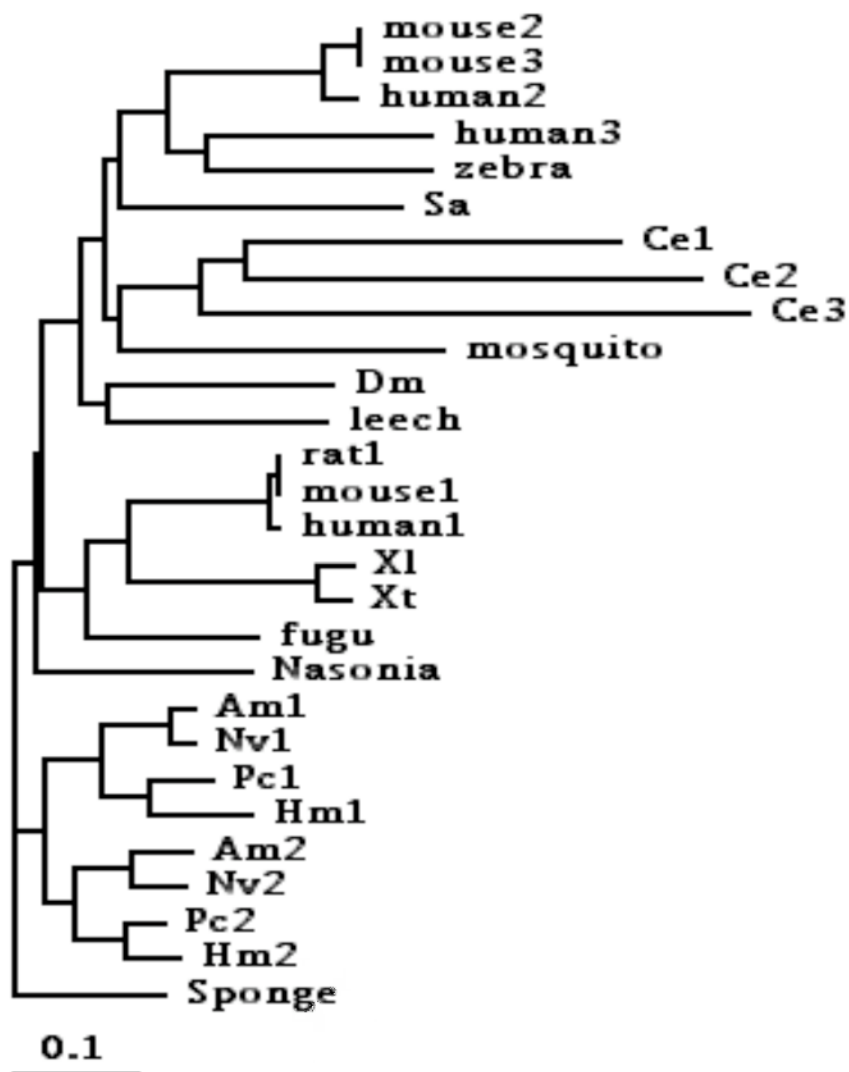


Fig.47

Genomic DNA sequences and deduced amino acid sequences of A: *Cnnos1*(*Am*) and B: *Cnnos2*(*Am*). Both genes are intronless. The sequences in the 5'- and 3'-flanking regions are also shown.

A

-613 TAAAGTGCGCATTAGTCTGACTGTTTACAGAAACCGCGTTCCTTTCAACGGCGTTTACAA  
GCCTTTTTTGGTAACGATGTGTTTCGTTAATGTGCGATAACCGAAGTTTATAAACGCCCA  
GCTGGACAAGTTGACAAAATTGTCAAAGAAATTAAGCGCTGCATTCACAAAACAAAACAT  
TTCATGGTATTTTGCGGAGAAAATTGATGACTTTTTTGTGTCAACTTTAAACGCTAATGAA  
GACACTGCTGATAATTTTTTAGACTATTTTTAACAGTCAGGCTTCCCTGTCTTCTCCAC  
GTGCGTTCGTTGTTTTGTCTTGAATAATAACACTTAAGCTCGTCGTTGGAAAACAATTACT  
GATGAAACCGGTCTGCCGTTTGACAGGACGAACTCTCTTTGACCAATCACACCAAAGTAA  
AATGACCCCATTTGCTTTTCAGAAATGCTGAGTTTTTAAAAAACTCCGCGATTCACTCAA  
AACGCATCTAGAGAAGTTTTGCGATTGGCTGATCTTAATGATCGTTTTCTATAGAAACAA  
TGAGACCACATGTGCGTTGCTTCTCTTAAAGCCATTTTCAATGCGTTTTTCAAATGCAT  
TCAAAGTAGCGAT

1 GAAAAGCTCTCGGTCCGGCAGCCGGTACGGTGGACGGATCGTTGTATAAATTGCCATTTTCAT  
61 GCACACATTAAGTCCGAACACTCACGCTTGTTAGGTGCTTCCAAAGCTTTTGTGGCTTT  
121 AAATCCCGCTGCCCTTTATTACAAAAGTTAATCATCACGATTCGAAACGACTGGAAACTG  
181 AATATATACTTTTCGCTCTAGCTGCCTCGGTACGCGGAAACGTCGCTGAGGGAAATTTTCG  
241 TTAATACAGTTGCTAGTCTGCACGCGATAAATCACTCATCGAACCAGGAGCCAGGAACAA  
M G D E M L S V F Y S G D I F P S H T K  
301 TGGGGGACGAGATGCTGAGCGTGTCTACAGCGGTGACATTTTCCCCAGCCATACTAAGA  
M I S A R E V A S L K V T F D E K S E F  
361 TGATAAGTGCCCGGAAAGTAGCAAGCCTCAAAGTCACTTTTCGATGAAAAGAGCGAATTTTC  
R L F R D Y F G L I R L V Q N A S D L E  
421 GACTCTTCAGGGATTACTTTGGCTTAATACGACTAGTTCAAACGCGTCGGATTTGGAAA  
R A E P G F D S L Y D F T Y T K R S R V  
481 GAGCGGAACCCGGATTTGATAGTTTGTATGATTTCACTTACACCAAAAAGATCCCGGGTTG  
D S M G S D H S L T S S S S G G N S E I  
541 ACAGCATGGGATCAGATCATTCTTAACAAGTTCAAGCTCTGGTGGAAACTCCGAAATTTG  
E L C I E A S K T P P P S L A P G A P T  
601 AACTTTGCTATAGAGCGTCTAAGACGCTCCGCGTCTAGCGCCCGTGCGCCAACCTC  
R V S K R N I E N S S D V A N R N R K N  
661 GAGTTTCGAAGCGGAACATTGAAAACAGCAGCGATGTAGCCAATCGCAACCGCAAGAATC  
R E S K K N R N A N V C V F C R N N G E  
721 GGGAAAGCAAGAAAAATCGGAACGCGAACGTGTGCGTGTTTTGTAGGAATAATGGCGAGA  
S K K V Y S S H V L K D A E G N T T C P  
781 GTAAAAAAGTCTATTCAAGCCATGTTTTGAAAGATGCTGAAGGCAACACAACATGTCCTA  
I L R A Y T C P L C K A S G N E S H T I  
841 TTCTGCGCGCATATACCTGCCCTCTTTGCAAAGCGTCCGGGAATGAATCTCATACAATCA  
K Y C P K N K A A S K Q Q Q Q Q N A T A  
901 AGTACTGTCCCAAGAACAAAGCCGCATCAAAAACAACAGCAACAACAAAACGCAACAGCAG  
V A A V G Q V \*  
961 TTGCTGCAGTTGGCCAAGTTTGACCGAACTATTAACACTGAAAGATCACAAATACACGTG  
1021 GGTCCATTTTGAATCGCTGAGCTATGTGCTCAAGGTACATGCACCTGCAGGAGGAAGTTA  
1081 ACTGAAACGGTAAAAATATTTCTCGCACAGTTTCTCAAAAAGGCTGATGCCCTGAAAAACA  
1141 GGCTCTAAACTATATGGATAAAATAATAGCGGGATATTTAACAGTATTCGGATACACCG  
1201 TTTAAGTTTTTATAATTTGATGATACAAACCTTAATAATTTTGTGCTTGTAGTCATGTCC  
1261 ACAAATAAAAATAAAAATAACTGTAGATTTTTAGATATTTGTAATAATGTCATGACTCTTTA  
1321 TGGCACATAGCTCAAATGCGGCGGTAACTTTCATTGGGCTCTGCAATATTTGCCTTATTT  
1381 TGCAAGTTCGCTATTTTGTCTGATTTACCCAGTAAAGAAATTTTTGGGCTCAAACGCCCTCT  
1441 TTTCAAGCTATGATAAAAATTTCAAAAAGAAATTCGGAATATCAGTAATTTATTTGTAACCGAA  
1501 ATTTGCTCATAACGATATTCGGTTTTTTCAGAAATTCATAACTTATACAAAATTAATTTGTGT  
1561 ACAGGACTGCGGAGAAAAATATGTTAACTTTTTGGTGTGAAATCATATAACAGTGAATA  
1621 AAAATTTGAAGCACAAAGTGAATTTCTAGTTTTTAAGTCAAAAACAATAAACTATCGACAG  
1681 GGTG  
AATTTTCTTGATGGAAACGAATTTCTTTTTGAACCTTAAAAGGTTAGCAACAACACTAGGGT  
AATATTTAATAGAAGTTCGAGAGGAAAACCTTACTGAAGTTGTAAATTCAGGATTAGATA  
AATTTTCAGAAATGTCATCGTGTAA

**B**

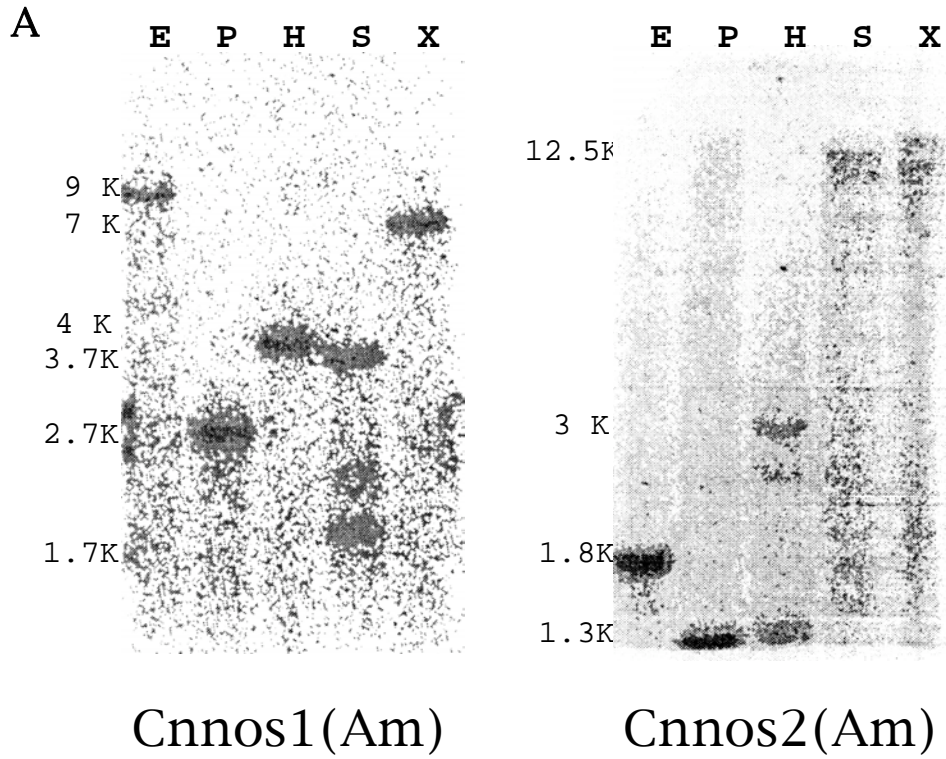
-539 CTGAAAACAAAGTGAAATGTCAGCGTGA CTCCAAAGGCGATTGTTTACACTTCATTTCCA  
CGACAATTAATATTTTAAAACTGCTTTATGATTGGGTCAAAAACGATCAAGTAATTCGA  
GTGGCTTGGCGCCGTCAGTCTTGCTGTTTTCAACATGGAAGGAACGCTGTCAAAAA  
TTATAGGTCACAAATCAATGCGCACTTGCTTTGAATCTTAGGGAGGGGAGACTCAGTCT  
TACCTATCCCCGCGCGCTTTAAGAAATGACCGCAAAGTTTTTTCTAATACTACCTGATGT  
TAGCGCCTTTTAAAGCTATCTTTGACGTTAGCTTTGTAAATTGAGAACATGGATGATAACTT  
TGGGCGATTTGAAAGGCGTAAAAATGCACTTCTTTACTAGCATGGTGTGTATCTCATT  
TCGGCTCATTTGACGTCATAAAGCGCAAGCGCAGTGTGTACTTTTTGACAAAGTATTGA  
GGGAAAGTCACGCCCTCAATTTTTTCGCTGAAGAAATACATGTCCCCTAATTTCCATG  
1 CGAGGTTTGGTGTAGCATTGACAAAACAGCCAGTAATCTGCCCTCTGGCGAACGAGAACGT  
61 AAATTTCTACACCGTGATTACAGTCCGTAAGAAAAAGCCTGACCTTTCAACGTGGGTTTCGT  
M S L S E V M L K T L I P  
121 GAAAGTTTTTGTCAAACGAAAAATGTCGCTGAGCGAAGTGATGCTAAAACTTTATTTCC  
S N F L D S D D E E I K S L T C P P E T  
181 CAGCAACTTCTTGATTGACGACGAGGAGATAAAGTCGTTAACCTGTCTCCCGAGAC  
G D L Q R S E F H M F N D Y F G L S N L  
241 GGGCGATCTGCAAAGAAGTGAATTCACATGTTTAATGACTATTTTCGGGTTAAGCAATTT  
V Q R V R Q A N N D F D Y F A S E N F G  
301 GGTGCAAAGAGTCAGGCAAGCGAACAAATGATTTTACTACTTCGCGAGTGAAAAATTTGG  
V S K L P S R R D R F D S I D S E N S S  
361 AGTCAGCAAATGCGCTCGAGACGTGACCGATTTGACAGCATTGATTTCGGAAAAATTCATC  
S G S C D S S T D H F G D G P F F S F N  
421 ATCTGGAAGCTGTGATTCCAGCACAGATCACTTTGGAGATGGGCCTTTCTTCTCATTCAA  
S S P L L Q G V A S K R K D A V V P P S  
481 CTCATCCCCCTTTTTGCAAGGTGTAGCTCCCAAGAGGAAAGATGCCGTTGTCCCTCCAAG  
R R I K P S A R A T A P V S N R G A V C  
541 TCGGCGGATTAAACCGTCTGCAAGAGCTACCGCTCCGGTCTCGAATCGCGGCGCAGTGTG  
V F C R N N G E S E E V Y L S H L L K G  
601 CGTGTCTGTCAGGAACAATGGTAAAAGCGAAGAAAGTCTATCTGTCTCATCTGTTGAAAGG  
P D G K T T C P I L R A Y T C P I C K A  
661 CCCAGATGGCAAGACGACATGCCCAATTTTGAAGAGCATATACCTGTCCGATTTGCAAGGC  
S G D E S H T I K Y C P Q N Q Q A Q A N  
721 AAGCGGTGATGAATCTCACACCATAAAGTATTGTCCCCAAAACCAGCAAGCACAAAGCCAA  
G G Q L P K V P R N N R V L P K R S R \*  
781 TGGCGGTCAAATGCCAAAAGGTGCCAGGAATAATCGCGTCCCTTCCAAAAGGAGCCGATG  
841 AACTGTGCTTGCAGATGCAGATGTAAGAGTTAGAGTGGAATCATTTTGTTCGGATTTTCA  
901 ATGGGTAAAATTTATAGTTAAGATGCCGATTTTTCAATTTATATTTCGGGTTGTTGTATT  
961 TTGTTTACGACTGGATCGCTCACTGTATTTTTCAAGACGGCGTGTACTGCGCAAAGATCGC  
1021 AACGAGTTAAATTTATTATTGTGTTTATGTAACAAAACCCACAAATGATGGCGCAGATTGC  
1081 AGCTCTTAGCTTTGTCAAGTCCCATTCATTTCATGGCGTGTGTTGTATCAAAAAGTTAAATTT  
1141 CATTCCCTTTGAGGCCATTTGAAATTTCTAAGCTTAGTGATGGCTGCCAAGTTCAGTTCT  
1201 GTTAGATCAGTTTTTCGCGCAACGATCTACATATTTAGATCGGAAATTTGAGGAAGATTAT  
1261 TGTTAAAATTTAGTATTTAGCTCCCCACAGAGGATGTTGTGTATATGTTACTGAAATTTAC  
1321 AATATTTTACCCCAAAGAGGTCCGATTTTTCCCTAGTTAGAGGTTAACTTATCAAATATGT  
1381 ATCTATGTAAAATTTTTGAATTTGTAGCACAAATGCAAAAATGAAATTTTACCCTCTAAGTG  
1441 TAAAGTATTTAATATGAAAGCGCTGATTAAGAAATTTTCGTGCGAGAAACGGAGAGATTTA  
1501 TGATTAATAATGTGTTTTATAAATTACAAGTTCTTCTCTAAGGGAAAATTTTCCAAATAATT  
1561 TGGCGATATTTACGATGTAGAAAATTTCACTACTTTGAAAAGTAGGTTTAAATGCTGCCTGG  
1621 TTGTAGAGGTGGATTTTTGAATGAAATTTTCTAATTTTTTACCAGCTGTGTTCTAAAT  
1681 ACTGGCTGCAAATGATGAAATGAAAAATCAATGGTTTATGAGGAAAGTTTTTGTAAATTTA  
1741 AAAACCTACATTTGTGCTATTCTAACATCAAAAAGGAGTTTGCATACGGTTGTATTTAAGC  
1801 TGAAGTTTATAACATGTAATCTTAATTTCTAATATCAGATGAGAAATTTTGTATGGTTTATT  
1861 GAAAGGATGAATCACATTAACCTTCGTTTTAC  
AATTTCTATGCCACTAGCTGCAGCTTATTGCGGTGCGGGAAAAATTCCTAGTCACTAAG  
CATATTTGTTTTGTTATTTGCGGCCATTTCTTGCTTTACTCGACCGAAGCAAATTAATTT

ACGCTTATGATGAAGCCAGTGG AACATTAAATTAACCAATATTTTGTGGTTTTACCACAAG  
TCCATCA

Fig.48

A: Southern-blot analysis on *Cnnos1(Am)* and *Cnnos2(Am)* in the genome of *Acropora millepora*. The letters above the lanes indicate the restriction enzymes used for the genomic DNA digestion. E: EcoRI, P: PstI, H: HindIII, S: SpeI, X: XhoI. The sizes of the bands are shown on the left.

B: Cut maps for *Cnnos1(Am)* and *Cnnos2(Am)*. The solid bars represent the probes for the hybridization.



**Fig.48** Southern-blot analysis on *Cnnos1(Am)* and *Cnnos2(Am)* in *Acropora millepora* genome



Fig.49

Potential transcription factor binding sites in the 5'-flanking region of  
A: *Cnnos1(Am)*, B: *Cnnos2(Am)*, C: *Cnnos1(Nv)* and D: *Cnnos2(Nv)*.

The transcription factors with high scores for binding (>95.0) are shown. The recognition sites and their orientation are indicated with dotted arrows.

**A**            *Cnnos1(Am)*

	1 TAAAGTGC GC ATTAGTCTGA CTGTTTACAG AAACCGCGTT CCTTTCAACG	entry	score
	----->	<a href="#">M00028</a> HSF	100.0
	----->	<a href="#">M00029</a> HSF	96.0
51 GCGTTTACAA GCCTTTTTTGT GTAACGATGT GTTTCGTTAA TGTGCGATAA			
101 CCGCAAGTTT ATAAACGCC AGTGGACAAG TTGACAAATT GTCAAAGAAA			
	----->	<a href="#">M00028</a> HSF	100.0
	--	<a href="#">M00101</a> CdxA	98.6
	----->	<a href="#">M00029</a> HSF	96.0
151 TTATAAGCGC TGCATTCACA AACAAAACAT TTCATGGTAT TTTGCGGAGA			
	----->	<a href="#">M00148</a> SRY	100.0
	---	<a href="#">M00028</a> HSF	100.0
	----->	<a href="#">M00101</a> CdxA	98.6
	---	<a href="#">M00029</a> HSF	96.0
201 AATTGATGAC TTTTTTGTGT CAACTTTAAA CGCTAATGAA GACACTGCTG			
	->	<a href="#">M00028</a> HSF	100.0
	->	<a href="#">M00029</a> HSF	96.0
251 ATAATTTTTT AGACTATTTT TAACAGTCAG GCTTCCCTGT CCTTCTCCAC			
	----	<a href="#">M00217</a> USF	100.0
	<----	<a href="#">M00028</a> HSF	95.3
301 GTGCGTTCGT TGTTTTGTCT TGAATAATAA CACTTAAGCT CGTCGTTGGA			
	---->	<a href="#">M00217</a> USF	100.0
	----->	<a href="#">M00101</a> CdxA	98.6
	<-----	<a href="#">M00240</a> Nkx-2.	97.7
	<-----	<a href="#">M00148</a> SRY	96.4
	--	<a href="#">M00019</a> Dfd	95.5
351 AACAACTACT GATGAAACCG GTCTGCCGTT TGACAGGACG AACTCTCTTT			
	----->	<a href="#">M00185</a> NF-Y	96.8
	----->	<a href="#">M00019</a> Dfd	95.5
401 GACCAATCAC ACCAAAGTAA AATGACCCCA TTGCTTTTCA GAAATGCTGA			
	<-----	<a href="#">M00048</a> ADR1	100.0
	----->	<a href="#">M00028</a> HSF	100.0
	----->	<a href="#">M00185</a> NF-Y	96.8
	----->	<a href="#">M00029</a> HSF	96.0
451 GTTTTTTAAAA AACTCCGCG ATTCACTCAA AACGCATCTA GAGAAGTTTT			
	----->	<a href="#">M00028</a> HSF	95.3
501 GCGATTGGCT GATCTTAATG ATCGTTTTCT ATAGAAACAA TGAGACCACA			
	<----	<a href="#">M00028</a> HSF	100.0
	----->	<a href="#">M00028</a> HSF	100.0
	<-----	<a href="#">M00211</a> Poly A	100.0
	<-----	<a href="#">M00271</a> AML-1a	100.0
	<----	<a href="#">M00029</a> HSF	96.0
	----->	<a href="#">M00029</a> HSF	96.0

```

----->
551 TGTGCGTTGC TTCTCCTTAA AGCCATTTTC AATGCGCTTT TCAAATGCAT
<-----
601 TCAAAGTAGC GAT

```

```

M00148 SRY 90.0
M00028 HSF 95.3

```

## B *Cnnos2(Am)*

	entry	score
1 CTGAAAACAA AGTGAAATGT CAGCGTGACT CCAAAGGCGA TTGTTTACAC ----->	<a href="#">M00148</a> SRY	100.0
51 TTCATTTCCA CGACAATTAA TATTTTAAAA ACTGCTTTAT GATTGGGTCA ----->	<a href="#">M00101</a> CdxA	100.0
101 AAAACGATCA AGTAATTCGA GTGGCTTGCC GCCGTCAAGT CTTGCTGTTT ----->	<a href="#">M00100</a> CdxA	96.2
151 TCAACATGGA AGGAACGCTG TCAAAAAAAA TTATAGGTCA CAAATTCAAT ----->	<a href="#">M00019</a> Dfd	97.0
201 GCGCACTTGC TTTGAATCTT AGGGAGGGGA GACTCAGTCT TACCTATCCC ----->	<a href="#">M00101</a> CdxA	98.6
251 CGCGCGCGTT TAAGAAATGA CCGCAAAGTT TTTCTAATAC TACCTGATGT ----->	<a href="#">M00253</a> cap	100.0
-	<a href="#">M00076</a> GATA-2	96.0
-	<a href="#">M00028</a> HSF	100.0
-	<a href="#">M00028</a> HSF	100.0
-	<a href="#">M00076</a> GATA-2	96.0
-	<a href="#">M00029</a> HSF	96.0
-	<a href="#">M00029</a> HSF	96.0
301 TAGCGCCTTT TAAGCTATCT TTGACGTTAG CTTTGTAATT GAGAACATGG ----->	<a href="#">M00041</a> CRE-BP	100.0
351 ATGATAACTT TGGGCGATTT GAAAGGCGTA AAAATGCACT TCTTTCACCTA ----->	<a href="#">M00029</a> HSF	100.0
401 GCATGGTGTG TATCTCATT TCGGCTCATT TGACGTCATA AAGCGCAAGC ----->	<a href="#">M00142</a> NIT2	95.0
451 GCAGTGTGT ACTTTTGGAC AAGTGATTGA GGGAAAGTCA CGCCTCCAAT ----->	<a href="#">M00028</a> HSF	95.3
501 TTTTCGCTG AAGAAATACA TGTCCCGCTA ATTTCCATG ----->	<a href="#">M00041</a> CRE-BP	100.0
	<a href="#">M00039</a> CREB	100.0
	<a href="#">M00039</a> CREB	100.0
	<a href="#">M00041</a> CRE-BP	100.0
	<a href="#">M00142</a> NIT2	98.8
	<a href="#">M00100</a> CdxA	96.2
	<a href="#">M00028</a> HSF	100.0
	<a href="#">M00029</a> HSF	96.0

## C *Cnnos1(Nv)*

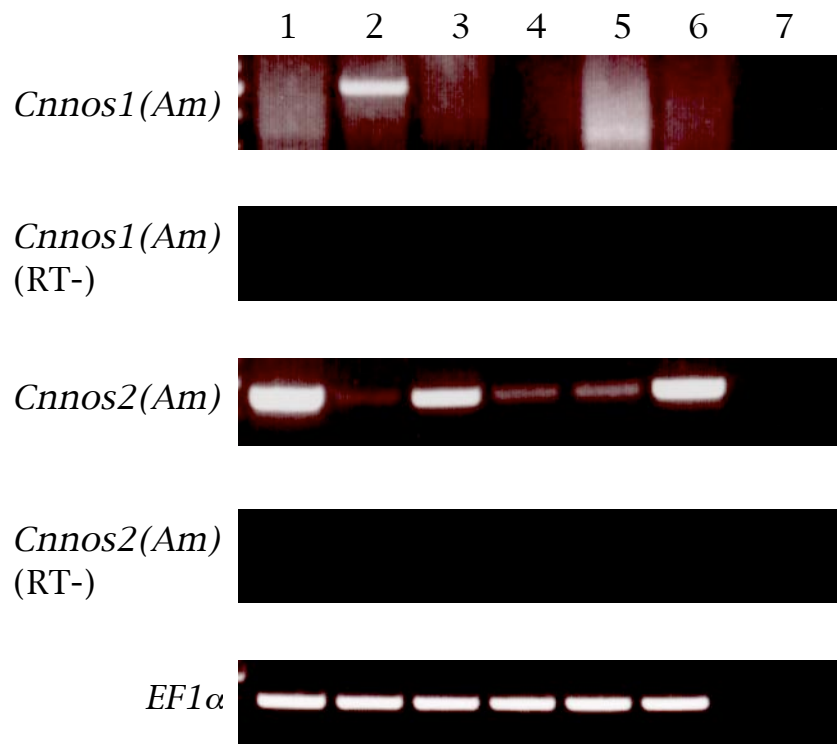
					entry	score
1	GAATTCGCTA	ATGTTAACGA	AAAATAAAAG	AATAAACTG	ACCGAGACGA	<- <a href="#">M00142</a> NIT2 98.8
						<- <a href="#">M00029</a> HSF 95.4
51	GATAAAGTGC	GAAATTTAAA	ATCTAAAATG	GAAAGAACGT	GACCTATATA	
						<a href="#">M00029</a> HSF 100.0
						<a href="#">M00142</a> NIT2 98.8
101	AGGTTACAT	TTTGAATAA	AAACAAGTTT	CACCTTTCTT	TTCGGTGAGG	
						<a href="#">M00028</a> HSF 100.0
						<a href="#">M00029</a> HSF 96.0
151	GATTCGAAAA	CATCAACAAT	TTATAAGCAA	TTCAAAAAGG	CAAACGCCTG	
						<a href="#">M00101</a> CdxA 100.0
						<a href="#">M00100</a> CdxA 96.2
201	GTTATAGTAT	GTTATTAGTC	AAGATTTGGC	TCCATAAAGA	TGTTTTAAAC	
						<a href="#">M00100</a> CdxA 96.2
251	TCCCCTAATC	AAGGGTGTCC	TTCTGAGTGC	TCAGTTTCGC	TATGACATCA	
						<a href="#">M00028</a> HSF 95.3
						<a href="#">M00253</a> cap 95.3
301	AGCCCAATCG	GCCAATAACA	GTGAGGCTCC	GGGATCATCG	GTTGTCTCAC	
351	CAGAATCGTT	TGCTTGCCTT	AACACGCTGC	GAATTGCCTT	GCGAATAAGT	
						<a href="#">M00029</a> HSF 95.4
401	GAGCACGCGC	GCTGATTGGT	GCGCGAAGTT	AATCCATTTT	ACTGTGAACA	
451	ATGAGGCCAC	ATGGTTGGTG	TTCTGGATAC	TGCTATGAGC	TATTTTCCGC	
						<a href="#">M00029</a> HSF 100.0
						<a href="#">M00226</a> P 95.0
501	GTCTTTTTTG	AAAAACATTC	AGCGGGCCCC	TGATTAGTGG	ACAATACAAT	
						<a href="#">M00154</a> STRE 98.1
551	TCGCAAAGAG	CCGTCGGCAG	CTGCGTGACC	TGGCGAGCGG	ATAAAACCGA	
601	ACGCGAGTGC	ACCTCTCGTA	ACTTAGCTCT	GACCATTGA		

## D *Cnnos2(Nv)*

					entry	score
1	AACAGAAGCA	ATTCAAATAA	TCGGTCTCTG	ATTGGTAGAG	AAACAGTGAC	<- <a href="#">M00028</a> HSF 100.0
						<- <a href="#">M00029</a> HSF 96.0
51	AAGAAAGGAC	ATTGGCTCCA	ATGAGTCTGC	AAAATGACAA	TGCAGTCATT	
						<a href="#">M00028</a> HSF 100.0
						<a href="#">M00029</a> HSF 96.0
101	CGCCCTAAGA	GCTGTGAGAT	TGCACTAGTC	TTAGATTTGT	CAATATTGTG	
						<a href="#">M00048</a> ADR1 98.5
151	GGGGAGGTAG	TGTTGTCACG	AAGTGTTCAT	ACCACGAAAT	TCTTCAGAAT	
						<a href="#">M00048</a> ADR1 98.5
						<a href="#">M00029</a> HSF 95.4
						<a href="#">M00029</a> HSF 95.4
201	ATCAAGCGAT	CATAGCAAAC	ATAGACTAGA	GCGTTGGATT	GAAACAACCT	
						<a href="#">M00148</a> SRY 96.4
251	ATCGTGTGGC	CTTGCTTAGC	GACTTGCAAG	AATAATTTTG	GAAAGGATTT	
						<a href="#">M00029</a> HSF 95.4
301	TGTAGCTTGT	TTACCATCCT	TGCACAACCC	CCTCCCTAGT	GGAATACTTT	
351	TGCTCTGTCT	GTTTGCAGCA	TTGCGCACGC	GCACTAGTAG	CACATTTGAC	
401	GTATTGAGCC	CTTATAAAGG	TCACGCCTCC	TCTCCTCTCA	AATCGAAAAA	
451	AAATCCC					

Fig.50

Expression analysis on *Cnnos1(Am)* and *Cnnos2(Am)* by RT-PCR. *EF1 $\alpha$*  was amplified for the positive control. RNA templates without reverse transcription (RT-) were subjected to the PCR amplification to exclude the possibility of the genomic DNA contamination. Lane 1: egg, Lane 2: prawn-chip (early developmental stage), Lane 3: pre-settlement (late developmental stage), Lane 4: branch tips (September), Lane 5: branch tips (October), Lane 6: branch tips (December), Lane 7: water for the negative control.



# Tables

PROBE	PRIMER PAIR
5B $\alpha$	5BF1 + 5BR2
5B $\alpha$ A	5BF1 + 5BR5
5B $\alpha$ B	5BF6 + 5BR2
5B $\beta$	5BF2 + 5BR1
A0	5B $\alpha$ A0 + 5BF1
A1	5B $\alpha$ A1 + 5BF1
A2	5B $\alpha$ A2 + 5BF1
A3	5B $\alpha$ A3 + 5BF1

Table 1.  
Radiolabelled probes for EMSA and the combination of oligonucleotide primers for each probe.

	Protein kinase C phosphorylation	Casein kinase II phosphorylation
DM domain	38 SPK	66 TCPD
Bridging region (between DM & DMA)	121 SSK	92 SDAE
	188 SDR	127 SPND
		163 SSAE
		179 SEED
C-terminal region	321 SGK	262 SHEE
	337 SPR	337 SPRD
	358 SPR	433 STGE
	361 SPR	
	378 SQR	
	418 SVK	
	424 SVR	

Table 2.  
Potential phosphorylation sites for Protein kinase C and Casein kinase II in NvDM5 protein.

clone	growth (6days)	clone	growth (6days)	clone	growth (6days)
2B	+	4J	+	5H	+
2I	+	4J'	+	5L	+
2J	+	4K	+	5P	+
2P	+	4K'	+	5R	+
2S	+	4L	+	5T	+
2W	+	4N	+	5U	+
2A'	+	4O	+	6A	-
2B'	+	4R	+	6E	-
2C'	+	4S	+	6F	+
2D'	+	4T	+	6I	+
2E'	+	4U	-	6L	+
2G'	+	5A	+	6M	-
2H'	+	5B	-	6N	+
2I'	+	5C	+	6P	+
2J'	+	5D	+	6Q	+
3E'	+	5E	+	6R	+
4B	+	5F	+	6S	+
4I	+	5G	+	control	+

Table 3. Summary of 5-FOA test

	length	phase	3' of exon	boundary sequence of intron	5' of exon
intron1	443	1	TTT	<b>GTAAGT-----TCATTTACAG</b>	CT
intron2	494	1	TGG	<b>GTAATG-----ATTTTTTCAG</b>	GA
intron3	315	0	ACG	<b>GTAAAG-----CACATCATAG</b>	CT
intron4	454	0	GAG	<b>GTTAGG-----GTTGTTCTAG</b>	TC
intron5	364	1	CAG	<b>GTAAGT-----TCCGTCACAG</b>	AG

Table 4. Introns of *AmVit*



# References

Asaoka-Taniguchi, M., Yamada, M., Nakamura, A., Hanyu, K., and Kobayashi, S. (1999)  
Maternal Pumilio acts together with Nanos in germline development in *Drosophila* embryos.  
*Nature Cell Biology* 1, 431-437

Ahmad, S. M., and Baker, B. S. (2002)  
Sex-specific deployment of FGF signaling in *Drosophila* recruits mesodermal cells into the male genital imaginal disc.  
*Cell* 109, 651-661

Ahringer, J., and Kimble, J. (1991)  
Control of the sperm-oocyte switch in *Caenorhabditis elegans* hermaphrodites by the *fem-3* 3' untranslated region.  
*Nature* 349, 346-348

Amrein, H. (2000)  
Multiple RNA-protein interactions in *Drosophila* dosage compensation.  
*Genome Biology* 1, 1030.1-1030.5

An, W., and Wensink, P. C. (1995a)  
Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer.  
*Genes & Development* 9, 256-266

An, W., and Wensink, P. C. (1995b)  
Three protein binding sites form an enhancer that regulates sex- and fat body-specific transcription of *Drosophila* yolk protein genes.  
*The EMBO Journal* 14, 1221-1230

Arrizabalaga, G., and Lehmann, R. (1999)  
A selective screen reveals discrete functional domains in *Drosophila* Nanos.  
*Genetics* 153, 1825-1838

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1995)  
*Current Protocols in Molecular Biology*, John Wiley & Sons Inc., USA

- Aviv, T., Lin, Z., Lau, S., Rendl, L. M., Sicheri, F., and Smibert, C. A. (2003)  
The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators.  
Nature Structural Biology 10, 614-621
- Baker, B. S., and Wolfner, M. F. (1988)  
A molecular analysis of *doublesex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*.  
Gene & Development 2, 477-489
- Baker, B. S. (1989)  
Sex in flies: the splice of life.  
Nature 340, 521-524
- Barker, D. D., Wang, C., Moore, J., Dickinson, L. K., and Lehmann, R. (1992)  
Pumilio is essential for function but not distribution of the *Drosophila* abdominal determinant Nanos.  
Genes and Development 6, 2312-2326
- Barr, M. M., and Sternberg, P. W. (1999)  
A polycystic kidney-disease gene homologue required for male mating behavior in *C. elegans*.  
Nature 401, 386-389
- Barton, M. K., Schedl, T. B., and Kimble, J. (1987)  
Gain-of-function mutations of *fem-3*, a sex-determination gene in *Caenorhabditis elegans*.  
Genetics 115, 107-119
- Bell, L. R., Maine, E. M., Schele, P., and Cline, T. W. (1988)  
Sex-lethal, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins.  
Cell 55, 1037-1046
- Bergsten, S. E., and Gavis, E. R. (1999)

Role for mRNA localization in translational activation but not spatial restriction of *nanos* RNA.

Development 126, 659-669

Bergsten, S. E., Huang, T., Chatterjee, S., and Gavis, E. R. (2001)  
Recognition and long-range interactions of a minimal *nanos* RNA localization signal element.

Development 128, 427-435

Bhat, K. M. (1999)

The posterior determinant gene *nanos* is required for the maintenance of the adult germline stem cells during *Drosophila* oogenesis.

Genetics 151, 1479-1492

Boswell, R. E., and Mahowald, A. P. (1985)

*tudor*, a gene required for assembly of the germ plasm in *Drosophila melanogaster*.

Cell 43, 97-104

Boswell, R. E., Prout, M. E., and Steichen, J. C. (1991)

Mutations in a newly identified *Drosophila melanogaster* gene, *mago nachi*, disrupt germ cell formation and result in the formation of mirror-image symmetrical double abdomen embryos.

Development 113, 373-384

Bröer, S., Bröer, A., Schneider, H., Stegen, C., Halestrap, A. P., and Deitmer, J. W. (1999)

Characterization of the high-affinity monocarboxylate transporter MCT2 in *Xenopus laevis* oocytes.

The Biochemical Journal 341, 529-535

Brunner, B., Hornung, U., Shan, Z., Nanda, I., Kondo, M., Zend-Ajus, E., Haaf, T., Ropers, H., Shima, A., Schmid, M., Kalscheuer, V. M., and Schartl, M. (2001)

Genomic organization and expression of the *doublesex*-related gene cluster in vertebrates and detection of putative regulatory regions for *DMRT1*.

Genomics 77, 8-17

Burtis, K. C.  
Doublesex in the middle  
*Science* 297, 1135-1136

Calvari, V., Bertini, V., De Grandi, A., Peverali, G., Zuffardi, O.,  
Ferguson-Smith, M., Knudtson, J., Camerio, G., Borsani, G., and Guioli,  
S. (2000)  
A new submicroscopic deletion that refines the 9p region for sex  
reversal.  
*Genomics* 65, 203-212

Cho, S., and Wensink, P. C. (1996)  
Purification and physical properties of the male and female double sex  
proteins of *Drosophila*.  
*Proceedings in the National Academy of Sciences USA* 93, 2043-2047

Cho, S., and Wensink, P. C. (1997)  
DNA binding by the male and female Doublesex proteins of  
*Drosophila melanogaster*.  
*The Journal of Biological Chemistry* 272, 3185-3189

Cho, S., and Wensink, P. C. (1998)  
Linkage between oligomerization and DNA binding in *Drosophila*  
Doublesex proteins.  
*Biochemistry* 37, 11301-11308

Christiansen, A. E., Keisman, E. L., Ahmad, S. M., and Baker, B. S.  
(2002)  
Sex comes in from the cold: the integration of sex and pattern.  
*Trends in Genetics* 18, 510-516

Clark, I. E., Wyckoff, D., and Gavis, E. R. (2000)  
Synthesis of the posterior determinant Nanos is spatially restricted by  
a novel cotranslational regulatory mechanism.  
*Current Biology* 10, 1311-1314

Colvin, J. S., Green, R. P., Schmahl, J., Capel, B., and Ornitz, D. M.  
(2001)

Male-to-female sex reversal in mice lacking fibroblast growth factor 9.

Cell 104, 875-889

Coschigano, K. T., and Wensink, C. (1993)

Sex-specific transcriptional regulation by the male and female doublesex proteins of *Drosophila*

Genes & Development 7, 42-54

Crittenden, S. L., Bernstein, D. S., Bachorik, J. L., Thompson, B. E., Gallegos, M., Petcherski, A. G., Moulder, G., Barstead, R., Wickens, M., and Kimble, J. (2002)

A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*.

Nature 417, 660-663

Crucs, S., Chatterjee, S., and Gavis, E. R. (2000)

Overlapping but distinct RNA elements control repression and activation of *nanos* translation.

Molecular Cell 5, 457-467

Curtis, D., Apfeld, J., and Lehmann, R. (1995)

*nanos* is an evolutionarily conserved organizer of anterior-posterior polarity.

Development 121, 1899-1910

Curtis, D., Treiber, D. K., Tao, F., Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997)

A CCHC metal-binding domain in Nanos is essential for translational regulation.

The EMBO Journal 16, 834-843

Dananukar, A., and Wharton, R. P. (1996)

The Nanos gradient in *Drosophila* embryos is generated by translational regulation.

Genes & Development 10, 2610-2620

Dauwalder, B., Tsujimoto, S., Moss, J., and Mattox, W. (2002)

The *Drosophila takeout* gene is regulated by the somatic sex-

determination pathway and affects male courtship behavior.  
*Genes & Development* 16, 2879-2892

Debrand, E., Heard, E., and Avner, P. (1998)  
Cloning and localization of the murine *Xpct* gene: evidence for complex rearrangements during the evolution of the region around the *Xist* gene.  
*Genomics* 48, 296-303

Deshpande, G., Gretchen, C., Yanowitz, J. L., and Schedl, P. D. (1999)  
Novel functions of nanos in downregulating mitosis and transcription during the development of the *Drosophila* germline.  
*Cell* 99, 271-281

Ephrussi, A., Dickinson, E., and Lehmann, R. (1991)  
oskar organizes the germ plasm and directs localization of the posterior determinant nanos  
*Cell* 66, 37-50

Fontanier-Razzaq, N., Haries, D. N., Hay, S. M., and Rees, W. D. (2002)  
Amino acid deficiency up-regulates specific mRNAs in murine embryonic cells.  
*The Journal of Nutrition* 132, 2137-2142

Forbes, A., and Lehmann, R. (1998)  
Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells.  
*Development* 125, 679-690

Forrest, K. M., and Gavis, E. R. (2003)  
Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*.  
*Current Biology* 13, 1159-1168

Friesema, E. C. H., Ganguly, S., Abdalla, A., Fox, J. E. M., Halestrap, A. P., and Visser, T. J. (2003)  
Identification of monocarboxylate transporter 8 as a specific thyroid hormone transporter.  
*The Journal of Biological Chemistry* 278, 40128-40135

Frumkin, A., Rangini, Z., Ben-Yehuda, A., Gruenbaum, Y., and Fainsod, A. (1991)

A chicken Caudal homologue C<sub>Hox</sub>-cad is expressed in the epiblast with posterior localization and in the early endodermal lineage. *Development* 112, 207-219

Fujii, S., and Amrein, H. (2002)

Genes expressed in the *Drosophila* head reveal a role for fat cells in sex-specific physiology. *The EMBO journal* 21, 5353-5363

Fujiwara, T., Dunn, N. R., and Hogan, B. L. M. (2001)

Bone morphogenetic protein 4 in the extraembryonic mesoderm is required for allantois development and the localization and survival of primordial germ cells in the mouse.

*Proceedings of the National Academy of Sciences of USA* 98, 13739-13744

Garcia, C. K., Brown, M. S., Pathak, R. K., and Goldstein, J. L. (1995) cDNA cloning of MCT2, a second monocarboxylate transporter expressed in different cells than MCT1.

*The Journal of Biological Chemistry* 270, 1843-1849

Garcia, C. K., Goldstein, J. L., Pathak, R. K., Anderson, R. G. W., and Brown, M. S. (1994)

Molecular characterization of a membrane transporter for lactate, pyruvate, and other monocarboxylates: Implications for the cori cycle.

*Cell* 76, 865-873

Garrett-Engle, C. M., Siegal, M. L., Manoli, D. S., Williams, B. C., Li, H., and Baker B. S. (2002)

*intersex*, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with *doublesex* to regulate terminal differentiation.

*Development* 129, 4661-4675

Gavis, E. R., Curtis, D., and Lehmann, R. (1996a)



- Identification of *cis*-acting sequences that control *nanos* RNA localization.  
Developmental Biology 176, 36-50
- Gavis, E. R., and Lehmann, R. (1992)  
Localization of *nanos* RNA controls embryonic polarity  
Cell 71, 301-313
- Gavis, E. R., and Lehmann, R. (1994)  
Translational regulation of *nanos* by RNA localization.  
Nature 369, 315-318
- Gavis, E. R., Lunsford, L., Bergsten, S. E., and Lehmann, R. (1996b)  
A conserved 90 nucleotide element mediates translational repression of *nanos* RNA.  
Development 122, 2791-2800
- Graham, P., Penn, J. K. M., and Schedl, P. (2002)  
Masters change, slaves remain.  
BioEssays 25, 1-4
- Green, J. B., Gardner, C. D., Wharton, R. P., and Aggarwal, A. K. (2003)  
RNA recognition via the SAM domain of Smaug  
Molecular Cell 11, 1537-1548
- Green, J. B., Edwards, T. A., Trincao, J., Escalante, C. R., Wharton, R. P., and Aggarwal, A. K. (2002)  
Crystallization and characterization of Smaug: a novel RNA-binding motif  
Biochemical and Biophysical Research Communication 297, 1085-1088
- Guan, G., Kobayashi, T., and Nagahama, Y. (2000)  
Sexually dimorphic expression of two types of DM (*Doublesex* / *Mab-3*)-domain genes in a teleost fish, the tilapia (*Oreochromis niloticus*).  
Biochemical and Biophysical Research Communications 272, 662-666
- Halestrap, A. P., and Meredith, D. (2004)  
The *SLC16* gene family--from monocarboxylate transporters (MCTs)

to aromatic amino acid transporters and beyond.  
European Journal of Physiology 447, 619-628

Halestrap, A. P., and Price, N. T. (1999)  
The proton linked monocarboxylate transporter (MCT) family:  
structure, function and regulation.  
The Biochemical Journal 343, 281-299

Hansen, D., Wilson-Berry, L., Dang, T., and Schedl, T. (2003)  
Control of the proliferation versus meiotic development decision in  
the *C. elegans* germline through regulation of GLD-1 protein  
accumulation.  
Development 131, 93-104

Haraguchi, S., Tsuda, M., Kitajima, S., Sasaoka, Y., Kitabayashi-  
Nomura, A., Kurokawa, K., and Saga, Y. (2003)  
*nanos1*: a mouse *nanos* gene expressed in the central nervous system  
is dispensable for normal development.  
Mechanisms of Development 120, 721-731

Harrison, P. L. and Wallace, C. C. (1990)  
Reproduction, Dispersal and Recruitment of Scleractinian Coral.  
Coral Reefs, Ecosystem of the World 25, 113-207

Hediger, M., Burghardt, G., Siegenthaler, C., Buser, N., Hilfiker-  
Kleiner, D., Dübendorfer, A., and Bopp D. (2004)  
Sex determination in *Drosophila melanogaster* and *Musca domestica*  
converges at the level of the terminal regulator *doublesex*.  
Development, Genes and Evolution 214, 29-42

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E.,  
Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A.,  
Kolpakov, F. A., Podkolodny, N. L., and Kolchanov, N. A. (1998)  
Databases on transcriptional regulation: TRANSFAC, TRRD and  
COMPEL.  
Nucleic Acids Research 26, 362-367

Hildreth, P. E. (1965)  
Doublesex, a recessive gene that transforms both males and females

of *Drosophila* into intersexes.  
*Genetics* 51, 659-678

Hislop, N. R. (2003)  
Thesis, James Cook University

Hodgkin, J. (2002)  
The remarkable ubiquity of DM domain factors as regulators of sexual phenotype: ancestry or aptitude?  
*Genes & Development* 16, 2322-2326

Houston, D. and King, M. (2000)  
Germ plasm and molecular determinants of germ cell fate.  
*Current Topics in Developmental Biology* 50, 155-181

Huang, X., Cheng, H., Guo, Y., Liu, L., Gui, J., and Zhou, R. (2002)  
A conserved family of *doublesex*-related genes from fishes.  
*Journal of Experimental Zoology* 294, 63-67

Illmensee, K., Mahowald, A. P., and Loomis, M. R. (1976).  
The ontogeny of the germ plasm during oogenesis in *Drosophila*.  
*Developmental Biology* 49, 40-65

Imanaga, S., Nakano, Y., and Watanabe, T. (2004)  
Molecular analysis of a major soluble egg protein in the scleractinian coral *Favites chinensis*.  
*Comparative Biochemistry and Physiology Part B* 137, 11-19

Inoue, K., Hoshijima, K., Sakamoto, H., and Shimura, Y. (1990)  
Binding of the *Drosophila* sex-lethal gene product to the alternative splice site of transformer primary transcript.  
*Nature* 29, 461-463

Irish, V., Lehmann, R., and Akam, M (1989)  
The *Drosophila* posterior-group gene *nanos* functions by repressing hunchback activity.  
*Nature* 338, 646-648

Jaruzelska, J., Kotecki, M., Kusz, K., Spik, A., Firpo, M., and Pera, R. A.

- R. (2003)  
Conservation of a Pumilio-Nanos complex from *Drosophila* germ plasm to human germ cells.  
*Development Genes & Evolution* 213, 120-126
- Kang, D., Pilon, M., and Weisblat, D. A. (2002)  
Maternal and zygotic expression of a *nanos*-class gene in the leech *Helobdella robusta*: primordial germ cells arise from segmental mesoderm.  
*Developmental Biology* 245, 28-41
- Keisman, E. L., and Baker, B. S. (2001)  
The *Drosophila* sex determination hierarchy modulates *wingless* and *decapentaplegic* signaling to deploy *dachshund* sex-specificity in the genital imaginal disc.  
*Development* 128, 1643-1656
- Keisman, E. L., Christiansen, A. E., and Baker, B. S. (2001)  
The sex determination gene *doublesex* regulates the A/P organizer to direct sex-specific patterns of growth in the *Drosophila* genital imaginal disc.  
*Developmental Cell* 1, 215-225
- Kettelwell, J. R., Raymond, C. S., and Zarkower, D. (2000)  
Temperature-dependent expression of turtle *Dmrt1* prior to sexual differentiation.  
*Genesis* 26, 174-178
- Kim, C. M., Golsstein, J. L., and Brown, M. S. (1992)  
cDNA cloning of *mev*, a mutant protein that facilitates cellular uptake of mevalonate, and identification of the point mutation responsible for its gain of function.  
*The Journal of Biological Chemistry* 267, 23113-23121
- Kim, D. K., Kanai, Y., Chairoungdua, A., Matsuo, H., Cha, S. H., and Endou, H. (2001)  
Expression cloning of a Na<sup>+</sup>-independent aromatic amino acid transporter with structural similarity to H<sup>+</sup>/monocarboxylate transporters.

The Journal of Biological Chemistry 276, 17221-17228

Kim, D. K., Kanai, Y., Matsuo, H., Kim, J. Y., Chairoungdua, A., Kobayashi, Y., Enomoto, A., Cha, S. H., Goya, T., and Endou, H. (2002)  
The human T-type amino acid transporter-1: characterization, gene organization, and chromosomal location.  
Genomics 79, 95-103

Kim, S., Kettlewell, J. R., Anderson, R. C., Bardwell, V. J., and Zarkower, D. (2003)  
Sexually dimorphic expression of multiple *doublesex*-related genes in the embryonic mouse gonad.  
Gene Expression Patterns 3, 77-82

Kobayashi, S., Yamada, M., Asaoka, M., and Kitamura, T. (1996)  
Essential role of the posterior morphogen nanos for germline development in *Drosophila*.  
Nature 380, 708-711

Kopp, A., Duncan, I., and Carroll, S. B. (2000)  
Genetic control and evolution of sexually dimorphic characters in *Drosophila*.  
Nature 408, 553-559

Köprunner, M., Thisse, C., Thisse, B., and Raz, E. (2001)  
A zebrafish *nanos*-related gene is essential for the development of primordial germ cells.  
Genes & Development 15, 2877-2885

Kraemer, B., Crittenden, S., Gallegos, M., Moulder, G., Barstead, R., Kimble, J., and Wickens, M. (1999)  
NANOS-3 and FBF proteins physically interact to control the spermatocyte switch in *Caenorhabditis elegans*.  
Current Biology 9, 1009-1008

Lafreniere, R. G., Carrel, L., and Willard, H. F. (1994)  
A novel transmembrane transporter encoded by the XPCT gene in Xq13.2.  
Human Molecular Genetics 3, 1133-1139

Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P., and Hogan, B. L. M. (1999)  
*Bmp4* is required for the generation of primordial germ cells in the mouse embryo.

Genes & Development 13, 424-436

Lee, G., Hall, J. C., and Park, J. H. (2002)  
*doublesex* gene expression in the central nervous system of *Drosophila melanogaster*.

Journal of Neurogenetics 16, 229-248

Lehmann, R., and Nüsslein-Volhard, C. (1991)  
The maternal gene *nanos* has a central role in posterior pattern formation of the *Drosophila* embryo.

Development 112, 679-691

Lehmann, R., and Nüsslein-Volhard, C. (1986)  
Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*.

Cell 47, 141-152

Lehmann, R. (1985)

Thesis, Eberhard-Karls-universität, Tübingen

Li, H., and Baker, B. S. (1998)

*hermaphrodite* and *doublesex* function both dependently and independently to control various aspects of sexual differentiation in *Drosophila*.

Development 125, 2641-2651

Liu, J., Wilson, T. E., Milbrandt, J., and Johnson, M. (1993)

Identifying DNA-binding sites and analyzing DNA-binding domains using a yeast selection system.

Methods: A Companion to Methods in Enzymology 5, 125-137

Lin, R. Y., Vera, C., Chaganti, R. S. K., and Golde, D. W. (1998)

Human monocarboxylate transporter 2 (MCT2) is a high affinity pyruvate transporter.

The Journal of Biological Chemistry 273, 28959-28965

Lints, R., and Emmons, S. W. (2002)  
Regulation of sex-specific differentiation and mating behavior in *C. elegans* by a new member of the DM domain transcription factor family.  
*Genes & Development* 16, 2390-2402

Lynch, K. W., and Maniatis, T. (1996)  
Assembly of specific SR protein complexed on distinct regulatory elements of the *Drosophila doublesex* splicing enhancer.  
*Genes and Development* 10, 2089-2101

Manseau, L. J., and Schüpbach, T. (1986)  
cappuccino and spire: two unique maternal -effect loci required for both the anteroposterior and dorsoventral patterns of the *Drosophila* embryo.  
*Genes and Development* 3, 1437-1452

Marín, I., Siegal, M. L., and Baker, B. S. (2000)  
The evolution of dosage-compensation mechanisms.  
*Bioessays* 22, 1106-1114

Matsuda, M., Nagahama, y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C. E., Shibata, N., Asakawa, S., Shimizu, N., Hori, H., Hamaguchi, S., and Sakaizumi, M. (2002)  
*DMY* is a Y-specific DM-domain gene required for male development in the medaka fish.  
*Nature* 417, 559-563

McMillan, J., Mahony, T., Veron, J. E. N., and Miller, D. J. (1991)  
Nucleotide sequencing of highly repetitive DNA from seven species in the coral genus *Acropora* (Cnidaria: Scleractinia) implies a division contrary to morphological criteria.  
*Marine Biology* 110, 323-327

McMillan, J., and Miller, D. J. (1989)  
Nucleotide sequences of highly repetitive DNA from scleractinian corals.

Gene 83, 185-186

McMillan, J., and Miller, D. J. (1990)

Highly repeated DNA sequences in the scleractinian coral genus *Acropora*: evaluation of cloned repeats as taxonomic probes.

Marine Biology 104, 483-487

Medina, M., Collins, A. G., Silberman, J. D. and Sogin, M. L. (2001)

Evaluation Hypothesis of Basal Animal Phylogeny Using Complete Sequence of Large and Small Subunit rRNA.

Proceedings of the National Academy of Sciences of USA 98, 9707-9712

Meng, A., Moore, B., Tang, H., Yuan, B., and Lin, S. (1999)

A *Drosophila doublesex*-related gene, *terra*, is involved in somitogenesis in vertebrates.

Development 126, 1259-1268

Miller, S. W., Hayward, D. C., Bunch, T. A., Miller, D. J., Ball, E. E.,

Bardwell, V. J., Zarkower, D., and Brower, D. L. (2003)

A DM domain protein from a coral, *Acropora millepora*, homologous to proteins important for sex determination.

Evolution & Development 5, 251-258

Mochizuki, K., Sano, H., Kobayashi, S., Fujisawa-Nishimiya, C., and

Fujisawa, T. (2000)

Expression and evolutionary conservation of *nanos*-related genes in *Hydra*.

Development Genes & Evolution

Moore, F. L., Jaruzelska, J., Fox, M. S., Urano, J., Firpo, M. T., Turek, P. J., Dorfman, D. M., and Pera, R. A. R. (2003)

Human *Pumilio-2* is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins.

Proceedings of the National Academy of Sciences of USA 100, 528-543

Mosquera, L., Forristall, C., Zhou, Y., and King, M. L. (1993)

A mRNA localized to the vegetal cortex of *Xenopus* oocytes encodes a



protein with a *nanos*-like zinc finger domain.  
Development 117, 377-386

Nagoshi, R. N., and Baker, B. S. (1990)  
Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns.  
Genes & Development 4, 89-97

Nakahata, S., Katsu, Y., Mita, Ko, Inoue, K., Nagahama, Y., and Yamashita, M. (2001)  
Biochemical identification of *Xenopus Pumilio* as a sequence-specific Cyclin B1 mRNA-binding protein that physically interacts with a Nanos homolog, Xcat-2, and a cytoplasmic polyadenylation element-binding protein  
The Journal of Biological Chemistry 276, 20945-20953

Narendra, U., Zhu, L., Li, B., Wilken, J., and Weiss, M. A. (2002)  
Sex-specific gene regulation : the *Doublesex* DM motif is a bipartite DNA-binding domain.  
The Journal of Biological Chemistry 277, 43463-43473

Nüsslein-Volhard, C., Frohnhofer, H. G., and Lehmann, R. (1987)  
Determination of anteroposterior polarity in *Drosophila*.  
Science 238, 1675-1681

Ottolenghi, C., Fellous, M., Barbieri, M., and McElreavey, K. (2002)  
Novel paralogy relations among human chromosomes support a link between the phylogeny of *doublesex*-related genes and the evolution of sex determination.  
Genomics 79, 333-343

Ottolenghi, C., and McElreavey, K. (2000)  
Deletions of 9p and the quest for a conserved mechanism of sex determination.  
Molecular Genetics and Metabolism 71, 397-404

Ottolenghi, C., Veitia, R., Barbieri, M., Fellous, M., and McElreavey, K. (2000)

The human *doublesex*-related gene, *DMRT2*, is homologous to a gene involved in somitogenesis and encodes a potential bicistronic transcript.

Genomics 64, 179-186

Ottolenghi, C., Veitia, R., Quintana-Murci, L., Torchard, D., Scapoli, L., Souleyreau-Therville, N., Beckmann, J., Fellous M., and McElreavey, K. (2000)

The region on 9p associated with 46, XY sex reversal contains several transcripts expressed in the urogenital system and a novel *doublesex*-related domain.

Genomics 64, 170-178

Parisi, M., and Lin, H. (2000)

A duet of Nanos and Pumilio

Current Biology 10, R81-R83

Pieau, C., and Dorizzi, M. (2004)

Oestrogens and temperature-dependent sex determination in reptiles: all is in the gonads.

Journal of Endocrinology 181, 367-377

Pilon, M., and Weisblat, A. (1997)

A *nanos* homolog in leech.

Development 124, 1771-1780

Price, N. T., Jackson, V. N., and Halestrap, A. P. (1998)

Cloning and sequencing of four new mammalian monocarboxylate transporter (MCT) homologues confirms the existence of a transporter family with an ancient past.

The Biochemical Journal 329, 321-328

Raymond, C. S., Kettlewell, J. R., Hirsch, B., Vivian, J. B., and Zarkower, D. (1999)

Expression of *dmrt1* in the general ridge of mouse and chicken embryos suggests a role in vertebrate sexual development.

Developmental Biology 215, 208-220

Raymond, C. S., Murphy, M. W., O'Sullivan, M. G., Bardwell, V. J., and

- Zarkower, D. (2000)  
*Dmrt1*, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation.  
Genes & Development 14, 2587-2595
- Raymond, C. S., Shamu, C. E., Shen, M. M., Seifert, K. J., Hirsch, B., Hodgkin, J., and Zarkower, D. (1998)  
Evidence for evolutionary conservation of sex-determining genes.  
Nature 391, 691-694
- Robida, M. D., and Singh, R. (2003)  
*Drosophila* polypyrimidine-tract binding protein (PTB) functions specifically in the male germline.  
The EMBO Journal 22, 2924-2933
- Ruppert, E. E., Fox, R. S., and Barnes, R. D. (2004)  
Invertebrate Zoology 7th Edition, 175-176  
Brooks/Cole--Thomson Learning, Belmont
- Ryner, L. C., Goodwin, S. F., Castrillon, D. H., Anand, A., Vilella, A., Baker, B.S., Hall, J. C., Taylor, B. J., and Wasserman, S. A. (1996)  
Control of male sexual behavior and sexual orientation in *Drosophila* by the fruitless gene.  
Cell 13, 1079-1089
- Ryner, L. C., and Baker, B. S. (1991)  
Regulation of *doublesex* pre-mRNA processing occurs by 3'-splice site activation.  
Genes & Development 5, 2071-2085
- Sakamoto, H., Inoue, K., Higuchi, I., Ono, Y., and Shimura, Y. (1992)  
Control of *Drosophila* Sex-lethal pre-mRNA splicing by its own female-specific product.  
Nucleic Acids Research 20, 5533-5540
- Sambrook, J., and Russell, D. W. (2001)  
Molecular Cloning (3rd Edition)  
Cold Spring Harbor Laboratory Press, USA

- Salz, H. K., Maine, E. M., Keyes, L. N., Samuels, M. E., Cline, T. W., and Schedl, P. (1989)  
The *Drosophila* female-specific sex-determination gene, *Sex-lethal*, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation.  
*Genes & Development* 3, 708-719
- Sánchez, L., and Guerrero, I. (2001)  
The development of the *Drosophila* genital disc.  
*BioEssays* 23, 698-707
- Sánchez, L., Gorfinkiel, N., and Guerrero, I. (2001)  
Sex determination genes control the development of the *Drosophila* genital disc, modulating the response to Hedgehog, Wingless and Decapentaplegic signals.  
*Development* 128, 1033-1043
- Sato, S., Tomomori-Sato, C., Banks, C. A. S., Parmely, T. J., Sorokina, I., Brower, C. S., Conaway, R. C., and Conaway, J. W. (2003)  
A mammalian homolog of *Drosophila melanogaster* transcriptional coactivator *Intersex* is a subunit of the mammalian Mediator complex.  
*The Journal of Biological Chemistry* 278, 49671-49674
- Schüpbach, T., and Wieschaus, E. F. (1986)  
Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo  
*Roux's Archives in Developmental Biology* 195, 302-317
- Schütt, C., and Nöthiger, R. (2000)  
Structure, function and evolution of sex-determining systems in Dipteran insects.  
*Development* 127, 667-677
- Shan, Z., Nanda, I., Wang, Y., Schmid, M., Vortkamp, A., and Haaf, T. (2000)  
Sex-specific expression of an evolutionarily conserved male regulatory gene, *DMRT1*, in birds.  
*Cytogenetics and Cell Genetics* 89, 252-257

- Shibata, K., Takase, M., and Nakamura, M. (2002)  
The *Dmrt1* expression in sex-reversed gonads of amphibians  
General and Comparative Endocrinology 127, 232-241
- Smibert, C. A., Wilson, J. E., Kerr, K., and Macdonald, P. M. (1996)  
smaug protein represses translation of unlocalized *nanos* mRNA in the  
*Drosophila* embryo.  
Genes & Development 10, 2600-2609
- Smith, C. A., McClive, P. J., Western, P. S., Reed, K. J., and Sinclair, A.  
H. (1999)  
Conservation of a sex-determining gene.  
Nature 402, 601-602
- Sonoda, J., and Wharton, R. P. (1999)  
Recruitment of Nanos to hunchback mRNA by Pumilio  
Genes & Development 13, 2704-2712
- Sonoda, J., and Wharton, R. P. (2001)  
*Drosophila* Brain Tumor is a translational repressor.  
Genes & Development 15, 762-773
- Spieth, J., Nettleton, M., Zucker-Aprison, E., Lea, K., and Blumenthal,  
T. (1991)  
Vitellogenin motifs conserved in nematodes and vertebrates.  
Journal of Molecular Evolution 32, 429-438
- Subramaniam, K., and Seydoux, G. (1999)  
*nos-1* and *nos-2*, two genes related to *Drosophila nanos*, regulate  
primordial germ cell development and survival in *Caenorhabditis*  
*elegans*.  
Development 126, 4861-4871
- Suzuki, M. G., Funaguma, S., Kanda, T., Tamura, T., and Shimada, T.  
(2003)  
Analysis of the biological functions of a *doublesex* homologue in  
*Bombyx mori*.  
Development, Genes and Evolution 213, 345-354

- Suzuki, M. G., Ohbayashi, F., Mita, K., and Shimada, T. (2001)  
The mechanism of sex-specific splicing at the *doublesex* gene is different between *Drosophila melanogaster* and *Bombyx mori*.  
*Insect Biochemistry and Molecular Biology* 31, 1201-1211
- Tauz, D. (1988)  
Regulation of the *Drosophila* segmentation gene *hunchback* by two maternal morphogenetic centres.  
*Nature* 332, 281-284
- Tauz, D., Lehman, R., Schnurch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., and Jäckle, H. (1987)  
Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes.  
*Nature* 327, 383-389
- Torras, R., Yanze, N., Schmid, V., and González-Crespo, S. (2004)  
*nanos* expression at the embryonic posterior pole and the medusa phase in the hydrozoan *Podocoryne carnea*.  
*Evolution & Development* 6, 362-371
- Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. a., and Bargmann, C. I. (1995)  
Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*.  
*Cell* 83, 207-218
- Tsuda, M., Sasaoka, Y., Kiso, M., Abe, K., Haraguchi, S., Kobayashi, S., and Saga, Y. (2003)  
Conserved Role of *nanos* proteins in germ cell development.  
*Science* 29, 1239-1241
- Verrotti, A. C., and Wharton, R. P. (2000)  
*Nanos* interacts with *Cup* in the female germline of *Drosophila*.  
*Development* 127, 5225-5232
- Vincent, S., Perkins, L. A., and Perrimon, N. (2001)  
*Doublesex* surprises.  
*Cell* 106, 399-402

- Wang, C., and Lehmann, R. (1991)  
Nanos is the localized posterior determinant in *Drosophila*.  
*Cell* 66, 637-647
- Waterbury, J. A., Horabin, J. I., Bopp, D., and Schedl, P. (2000)  
Sex determination in the *Drosophila* germline is dictated by the sexual identity of the surrounding soma.  
*Genetics* 155, 1741-1756
- Waterbury, J. A., Jackson, L. L., and Schedl, P. (1999)  
Analysis of the Doublesex female protein in *Drosophila melanogaster*: role in sexual differentiation and behavior and dependence on Intersex.  
*Genetics* 152, 1653-1667
- Wilkins, A. S. (1995)  
Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway.  
*BioEssays* 17, 71-77
- Winnier, G., Blessing, M., Labosky, P. A., and Hogan, B. L. M. (1995)  
Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse.  
*Genes & Development* 9, 2105-2116
- Yi, W., Ross, J. M., and Zarkower, D. (2000)  
*mab-3* is a direct *tra-1* target gene regulating diverse aspects of *C. elegans* male sexual development and behavior.  
*Development* 127, 4469-4480
- Yi, W., and Zarkower, D. (1999)  
Similarity of DNA binding and transcriptional regulation by *Caenorhabditis elegans* MAB-3 and *Drosophila melanogaster* DSX suggests conservation of sex determining mechanisms.  
*Development* 126, 873-881
- Ying, Y., Qi, X., and Zhao, G. Q. (2001)  
Induction of primordial germ cells from murine epiblasts by

synergistic action of BMP4 and BMP8B signaling pathways.  
Proceeding of National Academy of Science 98, 7858-7862

Zakharkin, S. O., Headley, V. V., Kumar, N. K., Buck, N. A., Wheeler, D. E., and Benes, H. (2001)  
Female-specific expression of a hexamerin gene in larvae of an autochthonous mosquito.  
European Journal of Biochemistry 268, 5713-5722

Zhai, Y., Tchieu, J., and Saier, M. H. Jr. (2002)  
A web-based Tree View (TV) program for the visualization of phylogenetic trees.  
Journal of Molecular Microbiology and Biotechnology 4, 69-70

Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Field, S., and Kimble, J. (1997)  
A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphroditic germ line.  
Nature 390, 477-484

Zhu, L., Wilken, J., Phillips, N. B., Narendra, U., Chan, G., Stratton, S. M., Kent, S. B., and Weiss, M. A. (2000)  
Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers.  
Genes & Development 14, 1750-1764



# Appendix

## Abbreviation

bp	base pair(s)
CAT	chloramphenicol acetyltransferase
cDNA	complementary DNA
CoA	co-enzyme A
DNA	deoxyribonucleic acid
DOPA	dihydroxyphenylalanine
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
IPTG	isopropylthio- $\beta$ -D-galactoside
kb	kilobase(s)
kDa	kilodalton
MCT	monocarboxylic acid transporter
min	minute(s)
nt	nucleotide(s)
ORF	open reading frame
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
rpm	rounds per minute
RT-PCR	reverse transcriptase-PCR
SDS	sodium lauryl sulfate
sec	second(s)
UTR	untranslated region