As secondary filter feeders, pearl oysters are exposed to various kinds of stressors like bacteria, viruses, pesticides, industrial wastes, toxic metals and petroleum derivatives, making them vulnerable to diseases. Environmental changes and ambient stress also affect non-specific immunity, making the organisms susceptible to infections. One of the strategies to combat these problems is to get insight into the disease resistance genes, and use them for disease control and health management. Similarly, although it is known that formation of pearl in mollusces is mediated by specialized proteins which are in turn regulated by specific genes encoding them, there is a paucity of sufficient information on these genes. Therefore, work on these aspects were carried out in the Indian pearl oyster *P. fucata* in the present study viz, molecular characterization and expression analysis of the genes involved in disease resistance and pearl formation, and the outcome are discussed hereunder.

### 5.1 Molecular Identification and Characterization of Functional Genes

#### 5.1.1 Antioxidant Enzyme Genes

Oxygen is a vital factor for the existence of life and it is an important element for all flora and fauna of terrestrial and aquatic origin. Reactive oxygen species (ROS) are produced in abundance when the animals are exposed to pathogenic attack and other abiotic stressors. Excessive ROS can damage a number of cellular macromolecules (Hartog *et al.*, 2003). Consequently, organisms contain a complex
network of antioxidant metabolites and enzymes that work together to prevent oxidative damage to cellular components such as DNA, proteins and lipids. (Sies, 1997; Vertuani et al., 2004). Super oxide dismutase (SOD) is the key defence molecule which fight ROS followed by others such as glutathione peroxidise (GPX) and glutathione-S-transferase playing vital roles in the detoxification pathways. Molecular detection and characterization of antioxidant enzyme genes are helpful for detecting cellular stress and also for bio-monitoring of environment. Though, study of Cu/Zn SOD gene including cloning have been carried out in several animals including Pacific oyster, Crassostrea gigas (Boutet et al., 2004), clam, Ruditapes decussatus (Geret et al., 2004), prawn, Macrobrachium rosenbergii (Cheng et al., 2006), and the abalones Haliotis discus discus (Kim et al., 2007) and Haliotis diversicolor supertexta (Zhang et al., 2009), no such study in pearl oysters has been made so far. The present study, for the first time, report the molecular identification, complete characterization and expression analysis of the gene encoding the most prominent antioxidant enzyme viz., superoxide dismutase (SOD). However, the cloning and mRNA expression of other antioxidant enzyme genes viz. GPX and GST in pearl oyster *P. fucata* from China have been reported by Jiang et al (2014). In the present study we also have carried out the PCR amplification and partial characterization of the GPX and GST genes in Indian pearl oyster *P. fucata*.

5.1.1.1 Superoxide Dismutase (SOD)

SOD is the first line of defence against superoxide radicals which exist in the haemocytes. In the present study, the full length nucleotide sequence of the cytoplasmic SOD of the pearl oyster, *Pinctada fucata*, and the amino acid sequence encoded by it were elucidated and reported. This is the first report of the kind from
this species. Out of the 924 bp long cDNA of SOD, the ORF consisted of only 471 bp encoding 156 amino acids. The InterProscan analysis of the amino acid sequence have revealed that it is a Cu/Zn SOD with copper/zinc binding domains within the sequence. There was no signal peptide indicating that it should be a member of the cytoplasmic SOD family. Multiple alignment of the deduced amino acid sequence of SOD with other cytoplasmic SODs revealed that two cysteines (Cys 60 and Cys 149) were conserved in all SODs, which were predicted to form one disulphide bond. Ni et al. (2007) have reported that Cu/Zn SODs can be discriminated from other types of SODs by their copper and zinc binding amino acids which are highly conserved. In the present study, it was found that *P. fucata* SOD contains copper and zinc binding regions, suggesting that pearl oyster SOD is a Cu/Zn SOD. Several reports have indicated that copper and zinc ions have critical functions in the quaternary structure and kinetic properties of Cu/Zn SOD (Assfalg et al., 2003; Cioni et al., 2003; Lynch and Colon 2006). Matsuo et al. (1997) found that limiting the dietary copper and zinc decreased SOD activity in the onion fly, *Delia antiqua*. The two Cu/Zn SOD family signature sequences of *P. fucata* SOD gene found in this study (signature 1 47GFHVHQFDNT57 and signature 2 141GNAGGRLACGVI152) are similar to the family signature sequence reported from *Chlamys farreri* and *Crassostrea gigas* (Ni et al., 2007; Boutet et al., 2004). Multiple alignments and phylogenetic analysis of Cu/Zn SOD genes of *P. fucata* with other molluscs and vertebrates have shown that these family signatures are conserved across the species (*Crassostrea gigas* CAD42722, *Mytilus edulis* CAE46443, *Candida albicans* AAC12872, *Gallus gallus* NP_990395, *Bos taurus* NP_777040, *Mus musculus* NP_035564).
BLAST analysis carried out using the deduced amino acid sequence with that of other species revealed extremely high identity (99%) with *C. hongkongensis*, *C. gigas* and *M. chilensis*.

Close evolutionary relationship between *P. fucata* and other aquatic invertebrates was revealed from the phylogenetic tree constructed using SOD amino acid sequence of *P. fucata* and other invertebrates as well as vertebrates. Vertebrates were evolutionarily distinctly separated.

5.1.1.2 *Glutathione Peroxidase (GPX)*

GPX enzyme constitute the second line of defence against lipid peroxidation, preventing the breakdown of lipid hydroperoxide which would otherwise lead to the formation of new radicals (Lawrence and Burk, 1976; Ren *et al.*, 1997; Bao and Williamson, 2000; Cossio-Bayugar *et al.*, 2005). In the present study the nucleotide sequence of GPX gene and the amino acids encoded by it were elucidated and reported. The amino acid sequence deduced from the ORF region was found to have many charactacteristic features. Absence of a signal peptide suggests that the glutathione peroxidase of *P. fucata* should be classified as a cytoplasmic GPX. The present study has shown that the structurally important amino acid residue viz., selenocysteine (Sel), the GPX signature motif, and the active site motif are highly conserved in the deduced amino acid. Similar pattern have been reported in *Haliotis discus discus* and *Chlamys farreri* by De Zoysa *et al.* (2008) and Mu *et al.* (2010). Though, generally the active site of Se-GPX and other selenoproteins is based on a selenium moiety (generally a single covalently bound atom of selenium) present as the amino acid selenocysteine, Gamble *et al.* (1997) reported an extension of the genetic code to include the codon TGA as the 21st codon specifying the presence of
the selenocysteine in the polypeptide structure of selenoproteins (selenium dependent proteins). In the present study, it was found that *P. fucata* GPX contains a TGA codon corresponding to Sel residue, suggesting that pearl oyster GPX is a selenium dependent protein and could be considered as a member of the selenoprotein family. The deduced amino acid sequence displays two important motifs, namely the Se-GPx signature 2 at $^{75}$LGFCNQ$^{82}$ and the active site at $^{163}$WNFEKF$^{168}$, which contain Gln(Q) and Trp(W). Depege *et al.* (1998) had reported that these residues are important for the catalytic activity of GPX. Earlier workers have identified the amino acids important for different functional roles of Selenium dependent glutathione peroxidase (Se-GPX) in other species. It has been reported that two residues, Gln (Q85) and Trp (W153), are involved in the fixation of selenium (Ladenstein *et al.*, 1986). Also, two arginine residues (R101 and R199), contribute to the electrostatic architecture that directs the glutathione donor substrate toward the catalytic center (Aumann *et al.*, 1997). The present study have also identified Gln85, Trp153, Arg101, and Arg199 amino acid residues in the *P. fucata* Se-GPX, supporting the view that they are involved in the functional and biochemical activities. Multiple alignments analysis carried out in the present study revealed that the selenocysteine residue, the GPX signature motif and the active site motif are highly conserved in the species such as pacific oyster, *Crassostrea gigas* (EF692639), mouse, *Mus musculus* (NM_008160) and insect, *Ixodes scapularis* (AAY66814). On account of these characteristic features, the *P. fucata* GPX can be considered as belonging to the selenium dependent glutathione peroxidase family.
5.1.1.3 Glutathione-S-transferase

Glutathione-S-transferases (GST) are the major phase II detoxification enzymes that play central roles in the defense against various environment toxicants. In the present study, the nucleotide sequence of GST and the amino acid sequence encoded by it have been elucidated. The full ORF sequence was found to be consisting of 231 amino acids. There was no signal peptide indicating that it should be a member of the cytoplasmic GST family. Previous reports suggest that GST commonly possess two binding domains. Mannervik and Danieison, (1988); Ivarsson et al. (2003) and Sheehan et al. (2001) reported that the N-terminal domain contains a GSH-binding site (G-site) that is well conserved among different classes of GSTs, and the C-terminal domain contains a hydrophobic substrate binding site (H-site) that varies widely in different classes, resulting in different substrate specificities. Conserved domain search using the amino acid sequence of *P. fucata* GST deduced in the present study revealed that it has also a G-site in its N-terminal region and an H-site in the C-terminus. GST genes are ubiquitous in almost all the existing organisms from bacteria to humans. Earlier workers have identified the co-presence of GST isoforms in a number of species, but with similar enzymatic features and physiological functions. In the case of disk abalone, *Haliotis discus discus*, six GST genes including one mu class GST, three sigma class GSTs and two omega class GST have been cloned and characterized, exhibiting protective roles against several organic pollutants (Wan et al., 2008a,b; Wan et al., 2009). Multiple alignments and phylogenetic analysis of *P. fucata* GST genes with other molluscs and vertebrates showed that they share some conserved domains (Fig.4.7) and that the *P. fucata* GST gene identified belongs to the omega class of the GST gene family (Fig.4.8).
al. (2009) found that two omega GSTs in abalone have two differences in the G-site, Tyr30 and Thr67 in HdGST01 (Haliotis discus discus glutathione-S-transferase 01) and Phe30 and Val68 in HdGST02 (Haliotis discus discus glutathione-S-transferase 02) sequence. In the present study, similar G-site pattern of HdGST02 have been identified in GST sequence of *P. fucata*. The compositions of the cysteine-containing tetramers are as divergent as C–P–Y–C, C–P–F–A, C–P–Y–A, C–P–Y–S, C–P–W–A, etc (Wan et al., 2009). In this study it was found that *P. fucata* GST shared the same active tetramer of C–P–F–A with abalone HdGST02, human GST01 and pig omega GST.

5.1.2 Pattern Recognition Receptor Protein Genes

Microorganisms have highly conserved and widely distributed signature molecules in their cell walls such as LPSs. They are recognized by the molecules of the immune or defense systems that have been named as “pattern recognition molecules” (Medzhitov and Janeway, 1997). They are also referred to as “pattern recognition receptor proteins” (PRPs). To date, a number of LPS-recognition/binding protein molecules have been isolated and characterized from several animals, and these pattern recognition molecules have been shown to be involved in the innate immune system of both invertebrates and vertebrates (Medzhitov and Janeway, 1997; Carroll and Prodeus, 1998; Hoffmann et al., 1999). Three types of pattern recognition receptor proteins of *P. fucata* viz, F-type lectin, galectin and LGBP were taken for this study.

5.1.2.1 F-type lectin

F-type lectin is a PRP which participate in the innate immune defense system. They play important roles in the immune response of both invertebrates and
vertebrates either by recognizing the exposed glycans of potential pathogens or by their immunoregulatory roles through the binding to carbohydrates on the surfaces of immune competent cells (Kuhlman et al., 1987; Cooper et al., 1994; Tino and Wright, 1996; Goldstein et al., 1980; Rabinovich et al., 2007). In the present study, the full length nucleotide and the amino acid sequence encoded by the F-type lectin gene of the pearl oyster, Pinctada fucata, were elucidated. This study is the first report of full length F-type lectin in pearl oyster P. fucata. The ORF of F-type lectin consisted of 588 bp nucleotides encoding a putative peptide of 196 amino acids. A signal peptide was identified at the N-terminus of the deduced polypeptide. There are seven cysteines (Cys 12, Cys 92, Cys 124, Cys 125, Cys 147, Cys 164 and Cys 188) in the mature F-type lectin, and multiple alignment of deduced amino acid sequences with other closely related F-type lecin sequences showed that six of them (Cys 92, Cys 124, Cys 125, Cys 147, Cys 164 and Cys 188) are conserved. Similar pattern have been reported in P. martensii (Chen et al., 2011). It is believed that these are involved in an internal disulfide bond. Conserved domain search in the NCBI revealed that the F-type lectin domain extends from Lys55 to Val192. Several reports have indicated that the first carbohydrate binding residue in F-type lectin is a His residue in European eel (Anguilla anguilla), Japanese eel (Anguilla japonica) etc followed by 26 other residues and the RGD (Arginine, glycine, Aspartic acid) motif. (Bianchet et al., 2002; Cammarata et al., 2012; Pierschbacher et al., 1985; Salerno et al., 2009). However, the present study has indicated that in P. fucata, the first carbohydrate binding residue is a serine followed by 26 other residues and the RGD motif, which is similar to that reported in P. martensii. Pierschbacher et al. (1985) has reported that RGD motif is conserved in cell adhesion proteins. Many of the F-
type lectin contains multiple domains: the zebrafish FBPL-I contains two F-type lectin domain (NCRD and CCRD), the Pacific oyster bindin protein contains one to five F-type lectin domains, and the F-type lectin domain of *Xenopus* pentraxin is combined with other domains. The *P. fucata* F-type lectin possesses only one domain which is similar to fucolectin of *A. japonica* reported by Honda *et al.* (2000) and F-type lectin of *P. martensii* reported by Chen *et al.* (2011). Multiple alignment of *P. fucata* F-type lectin sequence of the present study also was in agreement. To evaluate the molecular evolutionary relationships of F-type lectin of *P. fucata*, a phylogenetic tree was constructed. While the F-type lectin of *P. fucata* of the present study and that from *P. martensii* formed one cluster, other molluscs formed different clusters indicating relationship of varying degrees with each other.

### 5.1.2.2 Galectin

Galectins have been reported to bind glycans on the cell-surface and extracellular matrix of potentially pathogenic microorganisms, thereby functioning as recognition molecule and effectors in innate immunity (Vasta, 2009). In the present study, the nucleotide sequence of galectin from the *P. fucata* and the amino acid sequence encoded by them were cloned and sequenced. The nucleotide sequence of galectin gene of *P. fucata* decipherd in the present study was 1750 bp long of which the ORF region consisted of 1727 bp nucleotides encoding a putative peptide of 551 amino acids. The amino acid sequence deduced from the ORF region has many characteristic features.

Sequence of these amino acids were highly homologous to bivalve galectins reported in the literature for *P. fucata* from China, *Crassostrea virginica* and *Argopectin irradians*. Bivalve galectins have been reported to possess four
carbohydrate recognition domains (CRDs) with conserved carbohydrate binding sites in each CRD, whereas, tandem-repeat galectins contain two CRDs covalently linked through a unique link peptide (Vasta et al., 2004b; Leffler et al., 2004). In the present study it was found that 7 amino acid residues are conserved in each of the four galectin specific domains, and each of these four domains are connected by unique 11 amino acid residues. Recently, galectin cDNAs were cloned and characterized from Eastern oyster (Tasumi and Vasta, 2007), Pacific oyster (Yamaura et al., 2008), Manila clam (Kim et al., 2008), abalone (EF392832), freshwater snail (Yoshino et al., 2008), and bay scallop (Song et al., 2010). Four CRDs have only been found in molluscan galectins, including Crassostrea virginica, Argopectin irradians, P. fucata from China and P. fucata of the present study. Hirabayashi and Kasai (1993) have classified the mammalian galectins, based on their CRD organization, into three subfamilies as proto-, chimera and tandem-repeat types. Since the tandem-repeat galectins could simultaneously bind to multivalent ligands and greatly enhance the binding activity (Arata et al., 1997), the joining of more number of different CRDs in the same galectin could open up the possibility to specifically crosslink more distinct types of ligands, as opposed to homo-functional crosslinking by multimeric prototype galectins (Cooper, 2002). The previous studies demonstrated that galectins might be secreted via a “nonclassical” secretory pathway, which involves direct translocation of proteins across the inner plasma membrane followed by membrane blabbing and secretion (Nickel, 2003). The present study indicated that P. fucata galectin did not contain typical signal sequence, supporting the “nonclassical” secretory pathway.
From multiple alignment of *P. fucata* galectin CRDs with other known galectin CRDs, Hirabayashi et al. (2002) showed that *P. fucata* galectin contained many evolutionally conserved structural features in each CRD, such as the typical motifs H-NPR and WG-ER. Multiple alignment of *P. fucata* galectin sequence of the present study also was in agreement. To evaluate the molecular evolutionary relationships of *P. fucata* galectin, a phylogenetic tree was constructed. Galectin of *P. fucata* of the present study and that from China formed one cluster, while the other molluscs formed different clusters indicating relationship of varying degrees with each other.

### 5.1.2.3 LGBP

LGBP is a glycosylated protein, which has the ability to bind with the glycosylated substrates like carbohydrate moieties. In the present study, the nucleotide sequence of LGBP gene from the *P. fucata* and the amino acid sequence encoded by them were cloned and sequenced. The nucleotide sequence of LGBP gene of *P. fucata* deciphered in the present study was of 1740 bp long. The amino acid sequence deduced from the ORF region has many characteristic features. The ORF of LGBP consisted of 1692 bp nucleotides encoding a putative peptide of 563 amino acids.

The amino acid sequence of LGBP gene of the pearl oyster in the present study contained a potential recognition motif for β1,3-linkage of polysaccharides. Similar pattern have been identified in the sequence of *P. fucata* from China. The amino acid sequence of this motif was slightly modified in the blue shrimp *Penaeus stylirostris* (Roux et al., 2002), white shrimp *L. vennamei* (Cheng et al., 2005) and crayfish *P. leniusculus* (Lee et al., 2000). A LPS-binding site also was observed in
the pearl oyster LGBP, which was similar to the zhikong scallop LGBP (Su et al., 2004). The carbohydrate recognition domains in the scallop might facilitate the interaction of immunocytes with the pathogen and subsequently induce cellular defenses (Su et al., 2004). It is generally believed that the invertebrate LGBP genes might have evolved from bacterial glucanase and lost the glucanase activity during evolution, but retained the glucan-binding properties and therefore, played a role in innate immune response (Lee et al., 2000). It was found that LGBP of P. fucata contains a threonine-rich region and a glycine-rich region. Similar pattern have been reported in the P. fucata of China by Zhang et al. (2010). Johansson (1999) suggested that the RGD (Arg–Gly–Asp) motif reported in other crustacean LGPBs (Du et al., 2007; Lin et al., 2008; Padhi and Verghese 2008) are putative cell adhesive sites. However, there is no RGD motif in zhikong scallop LGBP (Su et al., 2004) and disk abalone HdPRP (Nikapitiya et al., 2008), and the P. fucata (Zhang et al., 2010). But, a modified form viz. KGD (Lysine, Glycine, Aspartic acid) have been reported in the molluscs. It is generally believed that the invertebrate LGBP genes might have evolved from bacterial glucanase and lost the glucanase activity during evolution, but retained the glucan-binding properties and therefore, played a role in innate immune response (Lee et al., 2000). However, recently, Pauchet and coworkers identified a 40 kDa β-1,3-glucan-binding protein from midgut of lepidopteran, which was similar to previously characterized epidopteran βGRPs, and demonstrated that this βGRP was an active β-1,3-glucanase and was mainly expressed in midgut, and this type of PRP was named as catalytic PRP (Pauchet et al., 2009). At present, a similar glucanase domain with the conserved active sites has also been observed in the kuruma shrimp LGBP (Lin et al., 2008), earthworm LGBP
In the present study, it was revealed that the *P. fucata* LGBP also had a glucanase motif with the conserved active sites, similar to the lepidopteran βGRP (Pauchet et al., 2009). The *P. fucata* LGBP might also be an active glucanase. Multiple alignments of *P. fucata* LGBP genes with other species showed that they share some conserved motifs (Fig. 4.16). To evaluate the molecular evolutionary relationships of *P. fucata* LGBP, a phylogenetic tree was constructed. LGBP of *P. fucata* of the present study and that from China formed one cluster, while the other groups formed different clusters indicating relationship of varying degrees with each other.

### 5.1.3 Genes Involved in the Pearl Formation

Identification of pearl forming genes and their expression pattern during pearl formation are valuable, both scientifically and economically, as it can be used for improving the quality and productivity of pearls. The present study was focused on the identification and characterization of three genes involved in pearl formation, and study of their expression pattern during Mabe pearl formation.

#### 5.1.3.1 Nacrein

Pearl is produced by the process of bio-mineralization of calcium carbonate crystals. Formation of the nacreous layer on the pearl surfaces is closely related to the expression patterns of the gene involved in the matrix proteins (Takeuchi and Endo, 2006; Inoue *et al.*, 2010a, b). Nacrein constitutes a significant portion of the soluble matrix proteins of the nacreous layer of the pearl. It is widely believed that aspartic acid-rich calcium binding proteins are responsible for biomineralization in mollusk (Weiner and Hood, 1975; Weiner, 1979; Wheeler *et al.*, 1981; Wheeler and
Sikes, 1984; Cariolou and Morse, 1988). There are only a few studies on the gene expression patterns of the matrix proteins in the mantle epithelial cells of the pearl sac (Inoue et al. 2010a; Wang et al., 2009). In the present study attempt was made to characterize the nacrein gene. The partial ORF cDNA of nacrein was cloned from *P. fucata*. The amino acid sequence of *P. fucata* nacrein deduced in the present study was highly homologous to that from Japan. Boskey (1992) reported that the regularly spaced carboxyl groups of Asp or Glu side chains on the surface of such calcium binding protein participate directly in calcium binding, either alone or in concert with other Asp or Glu in the polypeptide, such as osteopontin. Nacrein has an acidic Gly-Xaa-Asn (Xaa = Asp, Asn, or Glu) domain, which suggests that nacrein is a calcium binding protein. Probably, nacrein is the major aspartic acid rich calcium binding protein in the nacreous layer. The Gly-Xaa-Yaa (Yaa = any amino acid) repeats like those in the collagenous domain are also found in a structural protein of the inner ear, which produces calcium carbonate crystals (Davis et al., 1995). Two different forms of calcium carbonate crystal exist in a shell: one is calcite in the prismatic layer and another is aragonite in the nacreous layer. The major difference between these forms is the position of the carbonate groups. Addadi, Weiner and colleagues (Addadi and Weiner, 1985; Addadi et al., 1989) have already demonstrated that acidic proteins from shells regulate biological calcite growth. Moreover, Falini *et al.* (1996) and Belcher *et al.* (1996) showed that acidic macromolecules extracted from the prismatic and nacreous shell layers induced calcite and aragonite formation respectively, *in vitro*. These results suggest that the macromolecules modulate the stereochemical position of carbonate groups in the process of biological calcium carbonate crystallization. The mantle of mollusc shellfish is known to contain
carbonic anhydrase (CA) activity (Freeman and Wilbur 1948). However, the exact nature of acidic proteins and CA activity in the formation of prismatic or nacreous layer is not well understood. Analysis of the nacrein amino acid sequence deduced from the DNA sequence in the present study showed that the presence of Gly-Xaa-Asn repeats (Xaa = Asp, Asn, or rarely Glu) may be the sites for interaction with calcium. A search of the Swissprot protein data base by Miyamoto et al. (1996) revealed homology between nacrein and CA. The CA domain of nacrein exhibited greatest identity to human CAII (Montgomery et al., 1987; Murakami et al., 1987). In particular, among the 36 residues of the active site (Tashian, 1989), 63% identity to human CAII was observed. Furthermore, nacrein retained the three histidines involving zinc-binding residues at positions 141, 152, and 164 (Murakami et al., 1987). In the present study the characteristic thirty six residues of the active site and three zinc binding histidine residues (H-141, H-152, H-164) were identified in the nacrein sequence which are similar to the previous report.

5.1.3.2 Prismanlin-14

Prismanlin-14 is the second macromolecule so far characterized from the prismatic layer of *P. fucata*. Prismanlin-14 is an acidic protein and is associated with calcification of the prismatic layer. In the present study, the full ORF cDNA of a prismanlin-14 was cloned from *P. fucata*. Prismanlin-14 have been reported to have a peculiar amino acid composition comprising only 11 kinds of amino acids with high proportions of glycine (27.6%) and tyrosine (20.0%) residues. According to a previous report (Wada, 1966), high contents of glycine (23.0%) and tyrosine (12.3%) were also found in the crude matrix proteins from the prismatic layer of *P. fucata*. In the present study, it was found that the deduced amino acid of prismanlin-14 also
comprises of 11 kinds of amino acids with high content of glycine and tyrosine residues. A signal peptide was identified at the N-terminus of the deduced polypeptide. Previous reports suggest that acidic amino acid residues are located near both the N- and C-termini. This acidic nature of prismalin-14 is similar to that of many other macromolecules isolated and characterized from various biominerals and is supposed to be essential for direct interactions with calcium ions and/or the surface of calcium carbonate crystals (Fujisawa et al., 1996; Gilbert et al., 2000). Prismalin-14 showed an ability to bind calcium ions and showed inhibitory activity on calcium carbonate crystallization by Suzuki et al. (2004). Tkatchenko et al. (2001) and Whitbread et al. (1991) reported that prismalin-14 has a Gly/Tyr-rich region. The amino acid sequence of this region is similar to that of keratin, produced by vertebrates to some extent, and prismalin-14 shares common functions with keratin as an extracellular matrix protein, although the functions of other parts of prismalin-14 are completely different. This region may also form the structural motif termed as the glycine loop. The glycine loop may contribute to the elasticity and flexibility of the molecular conformation (Steinert et al., 1984; Zhang et al., 2003).

5.1.3.3 N19

N19 is one of the predominant proteins found in the water-insoluble fraction of the nacreous layer. This 19 kDa protein was first isolated by Yano et al. (2007) from the nacreous layer of the pearl oyster Pinctada fucata, and was named N19. They suggested that N19 is localized in the nacre, and plays a negative regulatory role in calcification in the pearl oyster. N19 play important roles in governing both pearl production and nacreous shell growth, and the N19 gene should be in a deeper association with the pearl formation than any other known nacreous shell matrix
protein genes (Yano et al., 2007; Wang et al., 2009). In the present study, partial ORF region of N19 gene was identified and confirmed by similarity search in NCBI BLAST analysis. The primer was designed for the expression study and to investigate their role during pearl formation. The present study revealed that N19 gene was constitutively expressed in mantle tissue which indicates its role in pearl formation. The amplification of the target genes with the cDNA synthesized from the total RNA indicates the expression of the candidate genes in Indian Pearl oyster.

5.2 Gene Expression

5.2.1 Antioxidant Genes

While the reactive oxygen species (ROS) produced by the respiratory burst of haemocytes can kill the invading pathogen, excess of it may cause the degradation of the host’s biomolecules (Mates et al., 2000). As a safeguard against the accumulation of ROS, several enzymatic or nonenzymatic antioxidants do exist in the body which play important role in maintaining normal cellular physiology (Mates et al., 1999). In order to study the time dependent expression level of three antioxidant genes in the candidate species viz., *P. fucata*, LPS challenge followed by semiquantitative PCR analysis of the genes was carried out. LPS, the principal component of the cell wall of gram-negative bacteria, acts as a powerful stimulator of innate immunity (Ulevitch and Tobias, 1995; Lemaitre, 1996). Some marine invertebrates such as the Atlantic horseshoe crab and the giant African snail are highly responsive to LPS (Biswas and Mandal, 1999; Iwanaga, 2002). In the present study semi-quantitative PCR estimation revealed that LPS stimulation significantly increased mRNA expression of SOD, GPX and GST in the haemocytes in a time-dependent manner.
SOD is the first line of defense against superoxide radicals generated in haemocytes. Previous reports show that pathogens or environmental pollutants can induce expression of cytoplasmic Cu/Zn SODs in invertebrates. Various expression patterns of cytoplasmic Cu/Zn SOD have been reported which might be depending on the species and the challenging condition. In all cases significantly higher levels of expression were detected. For example, in giant freshwater prawn *Macrobrachium rosenbergii*, the cytoplasmic Cu/Zn SOD expression in haemocytes was upregulated after injection with bacterial suspension of *Lactococcus garvieae* (Cheng *et al.*, 2006). In bumble bee *Bombus ignites*, the cytoplasmic Cu/Zn SOD expression was upregulated after LPS injection (Choi *et al.*, 2006). In oyster *C. gigas*, mRNA expression of cytoplasmic Cu/Zn SOD increased after 7 days exposure to hydrocarbons (Boutet *et al.*, 2004). The present study revealed that SOD mRNA expression got up-regulated over the time, and showed a significant increase 4 h after exposure to LPS, compared to the control. The level further increased to reach the maximum at 8 h post treatment and then dropped to basal levels at 36 h.

GPX enzymes constitute the second line of defense against lipid peroxidation, preventing the breakdown of lipid hydroperoxide which would otherwise initiate new radicals (Lawrence and Burk, 1976; Ren *et al.*, 1997; Bao and Williamson, 2000; Cossio-Bayugar, *et al.*, 2005). In the present study, GPX mRNA expression had increased from 4 hour onwards, compared to the control, and reached maximal levels at 8 h, when it was significant. It is likely that these temporal changes in GPX expression correspond with increased ROS production, and the ensuring detoxification of the ROS might have resulted in the decrease of GPX-mRNA levels from 8 h to 12 h. Upregulation of GPX following exposure to inducing agents have
been reported in other invertebrates. Shan et al. (2011) showed that GPX mRNA expression in Japanese scallop, Mizuhopecten yessoensis after exposure to V. anguillarum, increased from 3 h onwards and reached maximal levels at 12 h before declining to the basal levels. Mu et al. (2010) also had reported significant up-regulation of GPX resulting in increased mRNA expression in scallop Chlamys farreri after challenge with L. anguillarum.

GSTs are a group of multifunctional enzymes that reduce organic hydroperoxides, or catalyse the conjugation of glutathione to a large variety of electrophilic alkylating compounds, including hydroxyl alkenals, thereby protecting the cell against their potential toxicity (Forman et al., 2009). The increased GST mRNA levels from 4 h post LPS challenge and reaching the maximum at 8 h in the present study is indicative of the role of GST in the immune response of pearl oysters. Similar observations were reported by Park et al. (2009) while searching for biomarkers. The authors showed increasing mRNA expression of the pi, rho, and sigma class GSTs in Antarctic bivalves, Laternula elliptica following exposure to PCB. Zhao et al. (2010) observed that GST mRNA expression in crab, Eriocheir sinensis was significantly upregulated at 6 h after Aeromonas hydrophila challenge and dropped to basal levels at 12 h, suggesting that GST is important in initiating the host response in acute infection.

The present study showed that messenger RNA levels of GPX and GST increased almost parallel to that of SOD reaching a maximum at 8 h, and thereafter decreasing slowly to control levels. At the maximum, the relative mRNA expression of SOD, GPX and GST increased to 48%, 34% and 34% over the control, respectively. The mRNA expression levels of all three antioxidant enzymes peaked
at 8 hours post challenge; SOD attained higher expression level earlier than GPX and GST, evidenced by the significance of SOD increase at 4 hours post challenge. The high level of GPX expression lasted longer (12 h) when the decline was slow. Assuming that mRNA expression reflects the timing of enzymatic action, it can be suggested that during respiratory burst, SOD is the first enzyme to react which is followed in short succession by GST and GPX. There is, however, a broad time window in which all three enzymes act in concert eliminating the different molecular forms of reactive oxygen species.

In conclusion, antioxidant genes are constitutively expressed but can also be induced, enabling them to play critical roles in the detoxification of ROS during the respiratory burst. As filter feeders, pearl oysters are exposed constantly to various pathogenic bacteria naturally present in the microflora of coastal environments, and the basal level of expression noticed in the unchallenged animals indicate their preparedness to face them. Under experimental conditions elevated levels of expressed antioxidant genes peaked at 8 hours after challenge. Under aquacultural conditions elevated levels of antioxidant enzymes can be of diagnostic value, as a measure of stress on the pearl oyster in the ponds.

5.2.2 Pattern Recognition Receptor Protein Genes (PRP)

A critical step in the immune response is the identification of an invading organism as foreign. This recognition step involves interactions between microbial structural motifs and host receptors. This immune recognition process is being carried out by the innate immune receptors called PRPs. In oysters, haemocytes are responsible for cell-mediated defense. A major defense exhibited by haemocytes involves the direct phagocytosis of antigens. During phagocytosis, the haemocyte
recognizes and binds to an antigen by the presence of specific lectins either in the haemolymph or in the membrane of the haemocyte (Ford and Tripp, 1996). Though these lectins cannot destroy foreign matter, they are involved in the recognition by different immune modulators leading to its destruction.

In order to study the time dependent expression level of three PRP genes, LPS challenge followed by semiquantitative PCR analysis of the candidate genes was carried out, and it was observed that the mRNA level of F-type lectin in the challenged animals increased, compared to the control, and reached maximal level at 4 h and then gradually decreased over the time. This may be due to the combined effect of progressive recognition of the bacterial LPS and the clearance of LPS by the activated cellular or humoral immune responses concurrently taking place. In the previous reports, various expression patterns of F-type lectin occurring in other invertebrate animals have been reported. In Japanese sea perch (*Lateolabrax japonicus*), upregulation of mRNA expression of F-type lectin at 4 h after LPS stimulation was reported, and from the 6 h the expression level decreased (Qiu *et al*., 2011). In *P. martensii*, the expression level of F-type lectin significantly increased at 3 h post challenge and then decreased gradually over the time (Chen *et al*., 2011). The basal level of expression in unchallenged animals and substantial increase of expression following challenge may indicate that it is a constitutive as well as inducible acute-phase protein, and maintains the animal in a constant state of readiness by providing immediate detection of an invading infectious threat.

Among the PRPs, galectin is important as it is used for recognizing β-galactoside ligand of the pathogens by their conserved carbohydrate recognition domains (CRD) (Barondes *et al*., 1994), and play crucial roles in the innate
immunity. Up-regulation of invertebrate galectin induced by bacteria, virus, fungi or parasites has been reported in oyster (Zhang et al., 2011a), amphioxus (Yu et al., 2007) and clam (Kim et al., 2008). Lectin–glycan interactions are ubiquitous and essential to biological systems. In bay scallop (Argopecten irradians), up-regulation of AiGal2 leading to increased the mRNA expression level after Vibrio anguillarum or Micrococcus luteus challenge was reported by Song et al. (2011). The present study has also revealed that the galectin mRNA expression increased following LPS challenge, and reached maximal levels at 8 h, and then dropped by 36 h. This is in agreement with the report of Zhang et al. (2011b) who found that the expression of galectin in P. fucata was significantly up-regulated between 8 h and 12 h after bacterial challenge. Fermino et al. (2011) reported that glycan-binding protein, galectin (Gal 3) secreted by activated macrophages and mast cells at inflammation sites play an important role in inflammatory diseases caused by bacteria and their products such as lipopolysaccharides (LPS). Infection by gram negative bacteria result in LPS-galectin interactions, and the galectin could serve as a sensor to detect small amounts of LPS and allow it to efficiently activate recruited neutrophils. These results suggest that galectin is involved in immune defense against a broad-spectrum of bacteria and their products.

LGBP is another important member of the PRPs in invertebrate which displays various biological functions. In scallop Chlamys farreri, LGBP gene expression was up-regulated initially after stimulation by V. anguilarum and subsequently reduced to the normal level (Su et al., 2004). Zhang et al. (2010) reported that LGBP gene expression was up-regulated at 8h and 12 h after bacterial and LPS stimulation in P. fucata. LPS could up-regulate the mRNA level of LGBPs.
in several marine invertebrates including kuruma shrimp (Lin et al., 2008), crayfish (Lee et al., 2000) and disk abalone (Nikapitiya et al., 2008). In the present study, the expression of LGBP gene was significantly increased at 8 h after LPS stimulation and then declined gradually.

In conclusion, the PRP genes are not only constitutively expressed genes but can also be induced by LPS, enabling them to play critical role in innate immune defense of pearl oyster, P. fucata. The overall data in the present study revealed that when a foreign object enters the body, the PRP gene shows an increasingly higher level of transcription. At the maximum the relative mRNA expression of F-type lectin, galectin and LGBP increased to 50%, 30% and 24% over the control, respectively. This is to be expected since it has been suggested that PRP functions by recognizing PAMP and may activate different immune genes to defend against these pathogens. Hence, information on this PRP in P. fucata may be useful in studies of PRPs in other marine invertebrates.

5.2.3 Expression of pearl forming genes following Mabe implantation

In the present study time-dependent expression analyses was carried out in three genes viz., nacrein, prismalin-14, and N19 which are reported to be involved in the bio-mineralization process of pearl formation.

Miyamoto et al. (1996) who purified a matrix protein from the nacreous layers of the pearl from the pearl oyster P. fucata named it as nacrein, and reported that it is involved in pearl formation. Nacrein has been identified in various kinds of molluscs (Norizuki and Samata, 2008) and appears to be conserved in the families Pteriidae and Turbinidae (Miyamoto et al., 2005).
In the present study the expression of nacrein gene (as indicated by the nacrein mRNA level) have shown an up-regulation soon after Mabe implantation. The increase was 53% more than the control at day 10. The increased level of nacrein mRNA in the experimental oysters was significantly higher than the control up to 30 days. The expression on the 30th day was only marginally lower than the 10th day level. Steep decrease in the nacrein level in the present study started only from the 40th day of Mabe implantation. Inoue et al. (2011) has reported an early expression of nacrein gene in P. fucata (15 days) following implantation.

Miyamoto et al. (1996) showed that nacrein contains a carbonic anhydrase (CA) domain and a number of characteristic Gly-Xaa-Asn repeats. They suggested that the CA domain plays a role in calcium carbonate crystal formation of the nacreous layer, and proposed a model for the nacreous layer formation controlled by nacrein. Medakovic (2000) described high CA activity preceding biomineralization in Mytilus edulis.

Although there is increasing agreement that the CA domain is involved in the biomineralization of shells, the underlying mechanism is not fully understood. It is thought that the carbonic anhydrase (CA) domain catalyzes bicarbonate-formation and supplies bicarbonate ions to form CaCO₃ crystals (Miyamoto et al., 1996, 2005). The enhanced expression of nacrein observed in the present study probably indicate the high CA activity during the biomineralization process taking place during the Mabe pearl formation following the implantation, taking the cue from Medakovic (2000) who observed increased CA activity in Mytilus edulis during the process of biomineralization.
Pearl has an inner “prismatic-layer and an outer “nacreous-layer” over the nucleus. Generally, in the early stages of pearl formation, the prismatic layer is first formed over the nucleus, followed by the nacreous layer formed on the prismatic layer (Wada, 1962; Hongyan et al. 2007). Inoue et al. (2011) reported that changes in gene expression are closely related to the layer types (nacreous layer or prismatic layer) on the nucleus. Nacrein presumably participates in the formation of both the nacreous and prismatic layers (Kono et al., 2000), and hence its early upregulation which is sustained for longer period.

Prismalin-14 is a matrix protein capable of interacting with the calcium carbonate and is associated with calcification of the prismatic layer. Suzuki et al. (2004) suggested that prismalin-14 is the main one involved in the prismatic layer. Suzuki and Nagasawa (2007) suggested that Prismalin-14 bind to chitin through its chitin-binding site and participate in calcium carbonate crystal formation through an acidic region on the protein surface. In the present study it can be seen that (Fig.4.28) endogenous mRNA expression increases significantly in the initial days. The expression at day 10 was higher compared to the day 30 and 40 expression. This suggests that the prismalin-14 participate in the early phase of pearl formation shortly after Mabe implantation when the prismatic layer is being formed. The present result is in agreement with the report of Inoue et al. (2011) who found that highest prismalin-14 gene expression of the round pearl implanted into the oyster mantle was at day 10. This might indicate that prismalin-14 is the key component of pearl biominerlization.
Gene expression pattern for nacrein and prismalin-14 in the early phase show a common trend with highest mRNA levels observed at day 10. Since nacrein participates in the formation of both the nacreous and prismatic layers (Kono et al., 2000), its enhanced expression is sustained during the nacrein layer formation also, unlike prismalin-14. Inoue et al. (2010) reported that the quality of pearl is determined by the ratio of the thickness of the lower prismatic layer to that of the upper nacreous layer, and suggested that inhibition of the prismatic layer in the early stages of pearl formation will improve pearl quality. It is well known that the nacreous pearl is of higher quality than the prismatic pearl.

Wang et al. (2009) observed that N19 gene is the only gene that exhibited a high expression level in the pearl sac, which was even higher than that in the pallial mantle, indicating that N19 plays an important role in governing both pearl production and nacreous shell growth. He has suggested that the N19 protein have a more profound association with pearl formation than any other known nacreous shell matrix protein (NSMP). It was reported that N19 protein plays an important role in CaCO$_3$ precipitation (Yano et al., 2007).

In the present study only a slight increase in the expression of N19 gene was observed at day 10 in the implanted animals compared to control, whereas, at day 30 gene expression of N19 was significantly higher than on day 10, indicating its late action. On the 40$^{th}$ day there was only a slight decrease in expression when compared to day 30. This indicates that the expression of the N19 gene for participation in pearl formation starts 10 days after implantation. The peak of mRNA expression, and therefore, the main activity occurs by 30$^{th}$ day, which is in agreement with a recent study of nacreous shell matrix proteins in $P. fucata$ by Liu et al. (2012) suggesting
that significant expression of N19 gene in the pearl sac occurs by day 30–35 after implantation.

The expression of nacreous layer forming genes in the present study support the view by Liu et al. (2012) that temporal expression of various matrix protein genes are involved in the nacreous layer deposition, and is crucial for pearl development.

In conclusion, the present study revealed the relative expression of three genes involved in pearl formation. The general trend was an up-regulation between day 10 and day 30 post Mabe implantation. Expression levels of nacrein, prismalin-14 and N19 showed an up-regulation with variations at different time points. Nacrein and prismalin-14 had an early start of expression, and reached highest level at day 10 after implantation, and then gradually decreased until day 40. On the other hand expression level of N19 was highest at day 30. Based on the knowledge on the time dependent expression of different genes involved in pearl formation, suitable genetic intervention strategies like promoting nacreous layer forming genes and silencing prismatic layer forming genes during pearl formation could be developed for stimulating the production of high quality nacreous pearls.