Chapter 2

Review of Literature

“Before you discover you must explore.”

-Unknown
**REVIEW OF LITERATURE**

Genetic engineering provides a potential way of solving many of the problems regarding the stress, yield and increase in nutritional value related etc. by using a plethora of techniques to identify genes and processes that will lead to overall crop improvement. The use of mutant screens to identify genes and promoters associated with gamete formation has provided access to novel information regarding the processes of reproduction in higher plants (Xu *et al.* 2010; Chen *et al.* 2011a; Ariizumi and Toriyama 2011; Ma *et al.* 2012).

**2.1 Importance of study of plant reproduction**

There remains a large gap in elucidating the molecular basis of plant reproductive development. The capacity of plant populations to reproduce successfully is being challenged by factors such as, climate change, habitat fragmentation and the spread of invasive species, which will consequently affect their demography, evolution and long term persistence (Intergovernmental Panel on Climate Change 2007; Hedhly *et al.* 2009). The understanding of the biology of plant reproduction is of crucial importance to deal with these environmental challenges and for maintaining biodiversity, genetic resources and human well-being (Barrett 2000; Aguilar *et al.* 2006; Aizen and Vazquez 2006). The plant reproductive success determines the levels of resources that support both biodiversity and the food supply (Hedhly *et al.* 2009; Eckert *et al.* 2010). Immobility of flowering plants requires biological structures (pollen) to deliver the male gametes to the enclosed female gametophyte by means of pollination. There are several molecular mechanisms underlying in biological system which controls the overall developmental processes. Various kinds of plant research has been advanced since the establishment of principles of genetics, almost 150 years back by Gregor Mendel but still the researchers are not able to understand the overall biological meaning in plants (Suwabe *et al.* 2010). To better understand the biology of plant reproduction there is need to investigate several molecular, biochemical, physiological and environmental clues affecting the reproductive success. Although, the sole study of classical genetics is unable to reveal the comprehensive analysis of the entire gene networks, the emerging new technologies will continuously help the understanding of complex genetic phenomena (Suwabe *et al.* 2010). The study of haploid plant gametophytes provides an excellent opportunity to examine such genes essential to cell division or other fundamental cellular...
processes. Mutations that interfere with the production of male and female gametophytes are common and have been described in more than 100 species. Deficiency analysis and transmission studies in maize and *Arabidopsis* suggest that a large number of genes are required during the haploid gametophyte development (Patterson 1978; Buckner and Reeves 1994; Vizir et al. 1994; Vollbrecht and Hake 1995; Grossniklaus and Schneitz 1998).

### 2.2 Manipulation of reproduction

Modification of plants, animals, and microbes to produce desired traits has begun about 10,000 years ago. The construction of genotypes to meet specified requirements, along with development of reproductive system for ensuring the faithful reproduction of these genotypes may lead to sustainable maintenance of genotypes (Sybenga 1983). Advantageous outcomes of these genetic modifications include increased food production, reliability, and yields; enhanced taste and nutritional value; and decreased losses due to various biotic and abiotic stresses. Many of the hardeist, most productive crops are hybrids of two genetically disparate cultivars. But the beneficial combination of genes that makes the hybrids so robust disappears in the next generation because the genes are shuffled into new combinations during sexual reproduction (Barcaccia and Albertini 2013). Knowledge on the mode of reproduction and pollination is essential for making proper strategy for modification of traits desired and help in deciding the proper procedures to be used for the genetic improvement of plant species. In brief the plant reproduction is of two types i) sexual reproduction and ii) asexual reproduction (Fig. 2.1) (Bharathi and John 2013). Vegetative propagation can be exploited for maintenance of exact copies of a superior genotype however, this technique is usually not applicable to annual crops such as maize, rice and wheat, while in the species that exhibit an asexual type of seed production termed, ‘apomixis’, the fixation of a given genotype occurs naturally. Apomixis results in production of offsprings that are exact genetic replicas of the female parent because embryos are derived from the parthenogenic development of apomeiotic egg cells, is regarded as the consequence of sexual failure (Silvertown 2008; Bicknell and Koltunow 2004; Ozias-Akins 2006; Albertini et al. 2010; Pupilli and Barcaccia 2012; Koltunow et al. 2013). The important features of apomictic development are the absence of meiosis and fertilization-independent development of the egg cell (Singh et al. 2007). The potential benefits of harnessing apomixis are many and vary from full exploitation of heterosis (the phenomenon termed heterosis or hybrid vigor is achieved by crossing genetically distant breeding lines, Shull 1952; Spillane et al. 2004) by reseeding the best hybrids to clonal propagation of the superior genotypes in seed propagated
The clonally propagated crops could also be benefitted by Apomictic technology, as the pathogens effects their total production and the transfer of these through seed is limited and thus generating disease free material that can be more easily stored and transported. The development of apomixis technology needs a deeper knowledge of the mechanisms regulating reproductive development in plants, and necessitates identification of genes/promoters specifically or differentially expresses during the reproduction.

![Diagram of Sexual and Asexual Reproduction](image)

Fig. 2.1: Ways of reproduction for producing the offsprings by plants.

2.3 Role of gametophytes in manipulation of reproduction

In angiosperms, gametophytes play a major role in the reproductive process. For the manipulation of the reproductive system of sexually reproducing plant species the understanding of the proximate mechanisms governing pollination and mating and identification of the genes and promoters responsible for key reproductive traits is a prerequisite. The embryo and the endosperm are the edible parts of most of the seeds and is essential component for human and animal feeding and nutrition (Drews and Koltunow 2011). Therefore, an understanding of the processes underlying development of male and
female gametophytes that mediate fertilization is of great relevance for the genetic improvement of the plant species (Bolanos-Villegas et al. 2010). Hence, the knowledge of developmental process combined with identification of key genes and promoters at different developmental stages of male and female gametophytes is important.

2.3.1 Male gametophyte development

The process of male gamete formation in plants involves a series of events that lead to the production and release of mature pollen grains from the anther (Ma 2005). A typical anther contains four microsporangia, form sacs or pockets (locules) in the anther. Each microsporangium contains four separate maternal cell layers, the epidermis, endothecium, middle cell layer, tapetum, which surround the central microsporocytes, or pollen mother cells (PMCs). These PMCs undergo meiosis to form haploid spores. The spores may remain attached to each other in a tetrad or separate after meiosis. Each microspore then divides mitotically to form an immature microgametophyte called a pollen grain. The pollen is eventually released by the opening (dehiscence) of the anther. Anther development has been divided into 14 stages (Sanders et al. 1999). The development of microspore and its pollen wall formation needs the tapetum to undergo cellular degradation via programmed cell death (PCD) (Wu et al. 1997; Hsieh and Huang 2007; Parish and Li 2010) during subsequent developmental stages. The production and release of functional pollen from anthers is a complex developmental process that requires the coordinated participation of various cell and tissue types and their associated specific gene expression patterns. A number of genes defective at specific stages during male gametogenesis (Fig. 2.2) have previously been characterized and their mutants are being used as tools to investigate and manipulate anther and pollen development (Sanders et al. 1999; Scott et al. 2004; Ma 2005; Wilson and Zhang 2009; Feng and Dickinson 2010; Wilson et al. 2011). During the past decade, major advances in genetic and genomic technologies (Table 2.1) made significant progress in understanding of male gametophyte development at the molecular level (Twell et al. 2006; Sanders et al. 1999; Scott et al. 2004; Ma 2005; Wilson and Zhang 2009; Feng and Dickinson 2010). The anther specific genes are associated with anther cell division and differentiation (Nonomura et al. 2003; Xu et al. 2010; Wilson et al. 2011), tapetum development (Jung et al. 2005; Luo et al. 2006; Xu et al. 2006; Wu et al. 2008; Wilson and Zhang 2009; Niu et al. 2013), male meiosis (Yang et al. 2003; Kapoor and Takasuji 2006), pollen maturation (Park et al. 2005, 2006; Zhao et al. 2006; Gupta et al. 2007; de Azevedo Souza et al. 2009; Dobritsa et al. 2009; Li et al. 2010; Tang et al. 2009; Chen et al. 2011b; Chang et al. 2012), anther
dehiscence (Zhu et al. 2004), stamen filament development (Mariani et al. 1990), etc. Till date several mutants have been isolated that demonstrate the importance of genes and cellular processes in patterning male gametophyte development, PMRD 2012, (Cui et al. 2012).

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Resources</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Highly annotated <em>Arabidopsis thaliana</em> genome</td>
<td>The Arabidopsis Genome Initiative 2000</td>
</tr>
<tr>
<td>2</td>
<td>Public databases and readily available sequenced insertion site mutants</td>
<td>Alonso et al. 2003</td>
</tr>
<tr>
<td>3</td>
<td>Transcriptomic data sets from <em>Arabidopsis</em></td>
<td>Zimmermann et al. 2004</td>
</tr>
<tr>
<td></td>
<td>From the male gametophyte</td>
<td>Honys and Twell 2003; Pina et al. 2005</td>
</tr>
<tr>
<td></td>
<td>From isolated sperm cells</td>
<td>Borges et al. 2008; Engel et al. 2003</td>
</tr>
<tr>
<td></td>
<td>From generative cells</td>
<td>Okada et al. 2006),</td>
</tr>
<tr>
<td>4</td>
<td>Male gametophyte-specific techniques</td>
<td>Johnson-Brousseau and McCormick 2004</td>
</tr>
<tr>
<td>5</td>
<td>Identification of various gametophytic mutants through genetic screening</td>
<td>New mutants identified and characterised</td>
</tr>
<tr>
<td>6</td>
<td>Newly sequenced plant genomes</td>
<td>Nussbaumer et al. 2013</td>
</tr>
</tbody>
</table>

2.3.2 Female gametophyte development

The pattern of megasporogenesis in *Arabidopsis* is called monosporic pattern and female gametophyte development is of Polygonum type (Huang and Russell 1992). The female gametogenesis in *Arabidopsis* is divided into eight stages (Christensen et al. 1997). The development of female gametophyte includes megasporogenesis and megagametogenesis and is governed by several key genes and promoters at different developmental stages (Fig. 2.3). During megasporogenesis, a megaspore mother cell (MMC) undergoes meiosis to produce four haploid nuclei (Yang and Sundaresan 2000), three of which at the micropylar end degenerate by PCD.
Fig. 2.2: Depiction of different stages of microsporogenesis and key genes involved in regulating progression of gametogenesis development. Red coloured letters indicates key genes operating at particular stage while green letters indicates the promoters identified/characterized with maximal expression at particular stage.

The functional megaspore at the chalazal end undergoes megagametogenesis and gives rise to the female gametophyte. The genes and promoters identified to act at different developmental stages of male and female gametophytes have been targeted for improvement of plant species by using different approaches of genetic engineering.

Fig. 2.3: Stages of female gametophyte development and seed formation. A) Stages of microsporogenesis and microgametogenesis; B) double fertilization leading to seed formation. Red colour letters are genes identified by several authors to regulate particular stage of development. (Pictures of double fertilization adopted from http://orangegroupbiology.blogspot.in/2012/06/bow-double-fertilization.html)
2.3.3 Male gametophyte manipulation

The knowledge of key genes and their expression pattern is utilized for their manipulation and improvement of plant species for incorporating desired qualitative and quantitative traits. For example, i) generation of male sterility has been proved advantageous as these can be used for cost effective hybrid seed production, the male sterile transgenic plants or GM crops are safer to use to prevent the horizontal gene flow in the environment or to natural crops, to avoid sexual reproduction for more vegetative and floral growth in ornamental crop, to prevent the extinction of invasive but useful plant species, to overcome the reproductive isolation or cross-compatibility barriers (Gardner et al. 2009). Cytotoxic genes (e.g. barnase and RNase T1, diphtheria toxin A-chain (DTx-A) gene; Aprotinin) with the help of anther/pollen specific promoter, such as TA29, BcA9, PrMC2 (Mariani et al. 1990; Lee et al. 2003; Viswanathan et al. 2011; Zhang et al. 2012) can be used for engineering male sterility in plants. There are reports of successful development of male sterile lines by using this technique in cauliflower (Janssens et al. 1992), tomato (Zhang et al. 1998; Bai et al. 2002); Chinese cabbage (Yu et al. 2000); cabbage (Shen et al. 2001); watermelon (Chen and Zhang 2006); flowering Chinese cabbage (Cao et al. 2008) and eggplant (Cao et al. 2010); etc. Using antisense RNA or RNAi to silence relevant gene expression of pollen development, e.g. use of anti-gene CYP86MF encoding cytochrome P450 associated with the nuclear male sterility by Yu et al. (2004) and Huang et al. (2005) to develop male sterile lines into Chinese cabbage-pak-choi and broccoli, the antigene of actin gene and DAD1 gene (encoding phospholipase A1) has been used by Chen et al. (2009, 2010) for development of male sterile tomato and broccoli. And other approaches such as, transferring bacterial and other genes to disturb the pollen fertility (Schmulling et al. 1988), transferring mitochondrion genes relevant to cytoplasmic male sterility (He et al. 1996), fusing the specific promoter with a toxin gene of chemical-inducible expression by simulating chemical hybridizing to transformed plants, obtaining male sterile lines through double transgenic lines hybridization, and transposon mutation. But these approaches are uncontrolled and may affect the normal growth and development of plant species. ii) Improvement in floral quality, e.g., flavonoid 3’, 5’-hydroxylase (F3’5’H) is a key enzyme in the synthesis of delphinidin (Holton and Tanaka 1994) and over-expression of it in floral parts enhanced delphinidin production and produced blue flowers in roses (Holton and Tanaka 1994; Holton 1996; Katsumoto et al. 2007). Qi et al. (2013) expressed Phalaenopsis F3’5’H (PhF3’5’H) from the orchid under the control of the chalcone synthase promoter in
petunia flowers which resulted in production of deeper pink color of petals and they suggests that to produce a blue lily, many genes must be properly expressed and this may require promoters with different spatio-temporal expression property. iii) Alteration of flowering time for desired phenotype is of central importance to a plant reproductive success and quantity of flowers (and seed), delay or prevent flowering altogether, early flowering, longer duration of vegetative growth are crucial aspects of crop yield and quality (Rooney et al. 2007). Burks et al. (2013) reported that delayed flowering increased the size of stems of sweet sorghum (Sorghum bicolor subsp. bicolor) and its potential for sucrose accumulation, it is a key trait associated with high biomass yield and nitrogen use efficiency, in energy sorghum (Rooney et al. 2007; Olson et al. 2012, 2013). Mutasa-Gottgens et al. (2009) demonstrated that down-regulation of GA responses by transformation with the A. thaliana gai gene (which represses GA signalling and pollen poduction), under its own promoter (pgai::gai) or deactivation of GA by over-expression of the Phaseolus coccineus (bean) GA2ox1 gene, which inactivates GA, increased the required post vernalization thermal time. Similarly, Hong et al. (2011) filed a patent in which they claimed the molecular engineering of the Arabidopsis FT (flowering time) protein to produce a modified FT showing enhanced ability to trigger flowering, resulting in more flowers and an increase in seed yield. iv) Improvement of nutritional quality of food and fruits, e.g., expression of bHLH from (snapdragon Delila) and R2R3-MYB (snapdragon Rosea 1) under a fruit specific promoter in tomato yielded purple fruits accumulating anthocyanins at a concentration comparable to the level of anthocyanin in blackberries and blueberries and feeding of such tomatoes to cancer susceptible mice significantly increased their life span (Butelli et al. 2008). v) For treatment of allergic reactions, e.g., two proteins isolated from Japanese cedar pollen (Cryj1 and Cryj2) have been expressed at a high level in GM rice in Japan. Consumption of this rice to mice, successfully treated allergic reactions to Japanese cedar pollen by causing immune system tolerance to develop (Takagi et al. 2005). If this concept proves to work in people, treatments for a range of allergic reactions could be devised. vi) can be used in biopharming, since the first recombinant plant-derived pharmaceutical protein, human serum albumin production in transgenic tobacco and potato plants (Sijmons et al. 1990), the strategies has been formulated to improve the recombinant protein yield in plants which include i) the development of novel promoters, ii) the improvement of protein stability and accumulation through the use of signals that target the protein to intracellular compartments, and the improvement of downstream processing technologies (Menkhaus et al. 2004). The fundamental advantage of protein production in plants is the range and diversity of recombinant molecules that they can
potentially produce. As higher eukaryotes, plants are able to synthesize small peptides, polypeptides and complex multimeric proteins, many of which cannot be made in microbial systems (Ma et al. 2003). For bioenergy production, the growing demand of fuel and food necessitate the improvement of production of food and bioenergy crops (Brutnell and Frommer 2012). Although the risks of commercially-grown transgenic crops remains silent (Stewart 2004; Editorial 2012 nature biotechnology), commercial-scale production of certain combinations of transgenic traits and crops may confer some undesirable environmental and agricultural consequences (Stewart 2004; Wolfenbarger and Phifer 2000). The strategy to overcome these potential risks may be the confinement of transgenes, (Kausch et al. 2010; Moon et al. 2010; Zapiola et al. 2008; Watrud et al. 2004; Reichman et al. 2006). In this effort, Hague et al. (2012) used a promoter (Zm13Pro) from a maize pollen-specific gene (Zm13) for driving expression of the reporter gene GUS and the cytotoxic gene barnase in transgenic rice (Oryza sativa ssp. Japonica cv. Nipponbare) as a monocot proxy for bioenergy grasses and successfully demonstrated its usefulness for gametophytic transgene confinement and breeding strategies by pollen sterility in food and bioenergy crops.

2.3.4 Female gametophyte manipulation

The development of megaspore mother cell or a somatic nucellar cell is controlled by basic inheritance usually thought to depend on a single master regulatory gene or a few dominant key genes, which is controlled by “a delicate gene balance” (Muntzing 1940) of recessive genes and may allow the MMC or a somatic nucellar cell to change the mode of reproduction (Asker and Jerling 1992; Koltunow et al. 1995; Savidan 2000; Grossniklaus et al. 2001b). To understand this genetic balance, a deeper knowledge of the mechanisms and genes, that are specifically or differentially expressed during the formation of the embryo and embryo sac, regulating reproductive development in plants is prerequisite (Barcaccia and Albertini, 2013). Leblanc et al. (1997) were able to identify two candidate genes in Brachiaria specifically expressed in mature ovaries containing unreduced embryosacs, while Rodrigues et al. (2003) could identified 11 genes from the same species that were differentially expressed for regulating the switch in reproduction mode. The recent findings regarding the molecular mechanisms involved in controlling embryo sac development, fertilization, and endosperm development may be useful for determining genetic links between different forms of reproduction such as, apomeiosis, parthenogenesis, and autonomous or pseudogamous endosperm development. The basic structural and functional analyses of these candidate genes are crucial for engineering desired forms of reproduction.
The study on APOSTART suggests that this gene may be related to PCD that is involved in the non-functional megaspore and nucellar cell degeneration events and permit enlargement of maturing embryo sacs (Barcaccia and Albertini, 2013). In the embryo sac the development of MMC is largely controlled by sporocyteless/nozzle (spl) gene as its mutant is unable to develop a functional MMC and shows defects in nucellar cell identity (Schiefthaler et al. 1999; Sieber et al. 2004), while WUSCHEL (WUS) gene, is necessary for MMC specification (Gross-Hardt et al. 2002). Additionally, the multiple archesporial cells1 (MAC1) gene of maize (Sheridan et al. 1996, 1999) and the multiple sporocytes1 (MSPI) and tapetum determinant like1a (TDL1A) genes in rice (Nonomura et al. 2003; Zhao et al. 2008) are essential for the production of only one megasporocyte per single ovule. In Arabidopsis, argonaute (AGO) genes are involved in post-transcriptional gene silencing mediated by short RNAs (Baumberger and Baulcombe 2005; Olmedo-Monfil et al. 2010; Tucker et al. 2012), and mutations in the molecular pathways that generate sRNAs may dramatically affect fertility (Van Ex et al. 2011). AGO1, DCL1, HEN1, and HYL1 mutant showed that these are involved in disruption of reproductive development (Van Ex et al. 2011). These observations were further supported by AGO5 ortholog in rice, reported to be essential for the progression of pre-meiotic mitosis and meiosis (Nonomura et al. 2007), and the AGO9 ortholog in maize (Singh et al. 2011). Study of few Arabidopsis mutants have revealed that gametogenesis can be uncoupled from meiosis, e.g. loss of certain ARGONAUTE (i.e., AGO9) genes and other genes in the small RNA pathway, such as RNA dependent RNA polymerase6 (RDR6) and suppressor of gene silencing3 (SGS3), results in loss of restriction in gametic cell identity and fate (Slotkin et al. 2009; Olmedo-Monfil et al. 2010). In Arabidopsis, YUCCA (YUC) genes provides the evidence regarding the role of auxin in the cell fate specification of embryo sac development (Pagnussat et al. 2009). Li et al. (2008) reported that the sporocyteless/nozzle (SPL) represses the expression of YUCCA and it is likely that auxin plays a key role in the cell specification machinery that regulates differentiation of the MMC and/or maintains the undifferentiated state of nucellar cells once the MMC is formed. Several loss-of-function phenotypes related to megasporogenesis were recently discovered in Arabidopsis and monocots such as rice and maize, showing features of regulation of switch of genetic programmes in forms of reproductive development. DYAD/SWITCH1 (SWII) gene from Arabidopsis, is critical for proper megasporogenesis, which is responsible for sister chromatid cohesion and centromere organization at meiosis. A mutation in allelic variant of this gene, dyad, has been reported to be responsible for the production of few unreduced egg cells (Ravi et al. 2008). The combined effect of mutations in the two genes SPO11-1, which
prevents chromosome pairing and recombination, and REC8 (or SYN1), which modifies chromatid segregation along with OSD1 (MiMe-1 mutant; d’Erfurth et al. 2009) or CYCA1-2/TAM (MiMe-2 mutant; d’Erfurth et al. 2010) could mimic fully penetrant phenotypes of production of unreduced egg cells, similar to dyad, as these mutants did not undergo a second meiotic division. The maize elongate (ell) (Rhoades and Dempsey 1966) and Dominant non-reduction 4 (Dnr4) (Singh et al. 2011) mutant were able to form functional unreduced egg cells as occurs in the natural apomict Tripsacum (Grimanelli et al. 2003). Koltunow and Grossniklaus (2003) reported the identification of gene from Arabidopsis, termed FERTILIZATION-INDEPENDENT SEEDS (FIS), are known to initiate endosperm development without fertilization to varying extents. Matzk (1996) reported the ‘‘Salmon’’ system in wheat producing high numbers of haploid parthenogenic embryos, and the haploid inducer (hap) mutant of barley is associated with parthenogenesis (Hagberg and Hagberg 1980). Recently, Ravi and Chan (2010) shown that parthenogenic embryos can be generated at a relatively high frequency in transgenic lines of Arabidopsis expressing a modified centromere-specific histone CENH3 protein. The ectopic formation of embryo-like structures could be induced in Arabidopsis, by ubiquitous overexpression of several transcription factors, e.g. LEAFY COTYLEDON (LEC1 and LEC2) genes induce the expression of embryo-specific genes and trigger the development of embryo-like structures (Stone et al. 2001), while PICKLE (PKL) acts as repressor of embryogenesis, an upstream regulator of LEC genes (Henderson et al. 2004), while, BABY BOOM (BBM) leads to the development of embryos and cotyledons from vegetative tissues when overexpressed (Boutilier et al. 2002).

2.4 Ways of identification and manipulation of genes and promoters

Gene and promoters manipulation refers to the alteration of a specific DNA sequence in an endogenous gene at its original locus in the genome. It can be achieved by one of the methods depicted in the figure 2.4. Loss-of function i.e. T-DNA insertion, transposon insertion and mutation by virtue of chemical and physical mutagenesis agents leads to identification of responsible gene for a particular physiological, metabolic and other type of alterations in phenotype (Sundaresan et al. 1995; Martienssen 1998; AzpirozLeehan and Feldmann 1997). In the post genomic era, chemical and physical mutagenic agents are used in reverse genetics which were previously used primarily for forward genetics, with the development of new technologies such as TILLING (Targeting of Induced Local Leisons IN Genome) and Deleteagene (Slade and Knouf 2005; Jiang and Ramachandran 2010). Further,
reduced gene expression techniques *e.g.* RNAi and VIGS (*Virus Induced Gene Silencing*) and gene replacement methods (Shaked *et al.* 2005; Porteus and Caroll 2005) have also been used for targeting of genes and promoters for assigning function to an unknown gene (Ossowaski *et al.* 2008; Unver and Budak 2009).

![Manipulation of GENE/Promoter](image)

**Fig. 2.4:** Different methods and approaches applied for gene/promoter identification and manipulation

### 2.5 Importance of promoters

A gene is the entire nucleic acid sequence that is necessary for the controlled production of its final product (RNA or Protein). The Eukaryotic genome contains several thousands of genes which operate in an orchestrated mode to regulate development, growth, and survival of an organism. The expression of a gene is monitored by its regulatory elements commonly known as PROMOTER, usually located upstream (towards the 5' region) of a gene, providing proper activation or suppression of a gene that it controls (Arnone and Davidson 1997; Pedersen *et al.* 1999, Smale and Kadonaga 2003; Tirosh *et al.* 2009). Gene transcription is regulated by *trans*-acting sequence specific transcription factors that bind *cis* regulatory DNA sequences located in promoter or enhancer regions to regulate transcription by RNA polymerase (pol) *I, II* or *III*, which stands as a near universal mode of gene regulation (Remenyi *et al.* 2004). The organization and regulation of promoter elements in
plants has been comprehensively reviewed by Singh (1998) and Tyagi (2001). Promoters regulate the levels, sites and timing of gene expression (Pennacchio and Rubin 2001).

2.6 Types and characteristic features of promoters

2.6.1 Core promoter

The core promoter encompasses the transcription start site and typically extends -50 to +50 nt either upstream or downstream from the +1 start site and is known for its role in accurate transcription initiation and regulating basal transcription (Pedersen et al. 1999, Smale and Kadonaga 2003, Choi et al. 2004). A key function of the core promoter is to direct the initiation of transcription by the basal RNA polymerase II machinery. Some of the cis acting elements present in the core promoter (Fig. 2.5) regulating basal transcription are TATA Box, INR (Lo and Smale 1996), BRE (Lagrange et al. 1998, Deng and Roberts 2005) and CpG islands (Bird 2002; Rombauts et al. 2003) and the downstream core promoter element (DPE) (Butler and Kadonaga 2002).

2.6.2 Types of promoters

2.6.2.1 Based on transcriptional start site

On the basis of the patterns of transcription initiation promoters are divided in two groups, (a) single peak promoters (focused promoters), which have a relatively tightly defined TSS position and transcription starts at a single nucleotide or within a narrow region of several nucleotides and contains TATA box that makes gene regulation tissue specific; and (b) broad peak promoters (disperse promoters), which consists of multiple weak start sites over a broad region of about 50 to 100 nucleotides including CpG island and regulates constitutive gene expression (Carninci 2006; Juven-Gershon and Kadonaga 2010).

2.6.2.1.2 Based on nature of regulation

Based on the type and degree of expression, promoters are of three types viz. 1) constitutive promoters, 2) tissue specific promoters, and 3) inducible promoters (Mitra et al. 2009).

2.6.2.1.1 Constitutive promoters
Constitutive promoters express ubiquitously at about the same level in various plant tissues and organs, or throughout different developmental stages. These include the well known cauliflower mosaic virus 35S promoter (CaMV35S) (Benfey and Chua 1990), maize ubiquitin gene promoter (Christensen et al. 1992), rice actin 1 gene promoter (McElroy et al. 1990) and from Agrobacterium T-DNA nos (Shaw et al. 1984; Mitra and An 1989) and mas (Ni et al. 1995) genes promoter.

![Diagram of core promoter elements required for gene transcription](image)

**Fig. 2.5:** Diagrammatic representation of core promoter elements required for gene transcription. BREu= upstream TFIIB-recognition element; BREd= downstream TFIIB-recognition element; TFIIB= Transcription factor II B; TBP= the TATA-binding protein; TAF= TBP-associated factors; DCE= Downstream core element; MTE=Motif ten element; DPE= downstream promoter element; Inr= The initiator. Numbers indicate respective position of promoter element from Transcription start site and letters indicate the conserved domains for respective elements.

### 2.6.2.1.2 Tissue specific promoters

The expression of tissue specific promoters operates in particular tissues or organs and at certain developmental stages of a plant. Seed specific, fruit specific and root-preferential promoters are the most commonly known promoters of this type e.g. soybean (Glycine max) glycinin gene seed-specific promoter (Cho et al. 1989; Sims and Goldberg 1989), tomato fruit specific TFM promoter (Conner 1997), and maize root preferential ZRP promoter (Colbert et al. 1997). Distinct cis acting DNA sequences confer spatially-regulated activation of gene expression within different regions and in different cell types of the same region (Bustos et al. 1991).

### 2.6.2.1.3 Inducible promoters
These promoters are activated by one or more stimuli such as hormones (e.g. gibberellin, abscisic acid jasmonic acid, salicylic acid, auxin), chemicals, environmental conditions (water, salt, wounding) and biotic stress (microbes, insects, nematodes) (Singh et al. 2002; Gurr and Rushton 2005a, 2005b; Nakashima and Yamaguchi-Shinozaki 2006). This is particularly valuable when using transgenes whose constitutive expression is detrimental or even lethal to the host plants. For example, ethylene inducible promoter drives gene expression when induced by ethylene (Shinshi et al. 1995). Samalova et al. (2005) reported a promoter viz. pOp6/LhGR: a stringently regulated and highly responsive dexamethasone-inducible gene expression system for tobacco.

2.7 Ways of identification of functional promoters

How a gene is expressed differentially is a key to understanding of genetic regulation. One method to study this question is to map the functional sequence domains of a gene and determine what sequences are bound by proteins (presumably trans-acting factors) during expression in different tissues. Isolation and characterization of promoter leads to an understanding of what cis acting DNA sequences are responsible for the regulation of gene expression and how these sequences allow appropriate gene expression (Sarvestani et al. 2014). There are three main approaches to identify promoter sequences:-

a) Mapping of the cDNA in genomic DNA- (also classified as first generation promoter identification technique), and the second generation promoter identification techniques includes:- b) Use of promoter prediction programmes/computational method and c) Promoter tagging

2.7.1 Mapping of the cDNA in genomic DNA

cDNA methods of promoter isolation is one of the primary approach, in which differentially expressed or tissue specific or inducible cDNAs are identified by transcriptome analysis. Adjacent regulatory regions can then be isolated by screening the genomic DNA population or retrieved from the database (Koltunow et al. 1990). The desired mRNA is isolated and cDNA clone produced. Adjacent regulatory regions can then be isolated by screening genomic DNA (Maleck et al. 2000; Schenk et al. 2000; Seki et al. 2001; 2002a; 2002b; Rabbani et al. 2003; Yazaki et al. 2004). The most commonly used methods for isolating constitutive promoters have been cDNA based (Xiao et al. 2005). However, cryptic promoters cannot be easily identified using cDNA mapping strategy. cDNA microarray and
transcriptome analysis in rice and Arabidopsis could identify genes as cold, drought, high-salinity, abscisic acid (ABA) stress and gibberellin inducible genes (Maleck et al. 2000; Schenk et al. 2000; Seki et al. 2001; 2002a; 2002b; Rabbani et al. 2003; Yazaki et al. 2004).

2.7.2 Use of promoter prediction programmes/computational method

Since the DNA sequence of several higher eukaryotes is now available, new methods of promoter-finding have become popular. This provides the opportunity to identify and analyze the parts of a genome believed to be the promoter by using in silico Promoter Prediction Programmes (PPPs) generated on the basis of sequence properties such as GC content and other general physico-chemical properties (Fickett and Hatzegeorgiou 1997; Ohler and Niemann 2001; Rombauts et al. 2003; Bajic et al. 2004, 2006a; Sonnenburg et al. 2006; Huang et al. 2014). However, the identification of the core promoter region and the TSS remains a difficult problem (Bajic et al. 2006; Sonnenburg et al. 2006; Xie et al. 2006; Wang et al. 2007a) because of some limitations associated with these programmes that include training with high quality experimental data (Munch and Krogh 2006).

2.7.3 Promoter tagging

The knowledge of the expression patterns of the genes led to define the strategy for the identification and cloning of plant genes. Differential screening approaches like differential display or subtractive hybridization (Reuber and Ausubel 1995) has been used for identification and isolation of the genes that expresses spatially, temporally or in response to environmental signals. However, the development of microarray and gene chip technologies has allowed rapid detection of expression profiles of many genes in single experiments (Richmond and Sommerville 2000). These approaches has their own disadvantage in terms of the limited source of tissues used for preparation of RNA probe, and genes expressed transiently or at low levels are unlikely to be identified. Furthermore, these approaches might just provide clues about the functions of the gene in cell or organs in which they are expressed and fail to assign a biological function to a given gene with confidence. The mutants conferring a particular phenotype in comparison to WT provides the opportunity to isolate and characterize its upstream sequences as a promoter associated with the phenotypic events. The identification of the mutant site of insertion and its flanking sequence known as tagging has become more popular technique to tag a gene with its spatio-temporal expression pattern.
2.8 Traf like protein genes and their role in plant reproductive development

The family of proteins identified on the basis of their capabilities to interact with and regulating the different members of Tumor necrosis factor family receptors (TNFR) were named as TNF-receptor associated factors (TRAFs) (Chung et al. 2002; Zapata 2003; Bradley and Pober 2001; Bishop 2004). TNF- Receptors associated Factors (TRAFs) constitute a family of adapter proteins, regulates the recruitment of kinases and other effector proteins to the activated receptor and other signaling complexes (Abell and Johnson 2005). TRAFs also mediate the activation of downstream components of these pathways, control the subcellular relocalization of the receptor-ligand complexes, and modulate the extent of the response by controlling the degradation of key proteins in the pathway (Arch et al. 1998; Wallach et al. 1999; Deng et al. 2000). TNF receptor superfamily mainly confers a wide range of biological functions, such as adaptive and innate immunity, embryonic development, stress response and bone metabolism through the induction of cell activation, cell survival, and anti-apoptotic functions mostly mediated by the family of TRAFs (Park et al. 2000; Chung et al. 2002). Recent advances in TRAF research revealed that TRAFs function as E3 ubiquitin ligase (Deng et al. 2000; Liao et al. 2004; Brown et al. 2002; Hostager et al. 2003) and the study of Sakai et al. (2004) points the role of the ubiquitin system in reproductive development.

TRAF like proteins has been reported in plant system to be involved in several signalling processes directly or indirectly in many studies. Cosson et al. 2010 worked on RTM3-a member of TRAF like protein family and reported that it is able to form multiprotein complex and interact with Cullin3 (CUL3) proteins via their BTB domain to form functional E3 ligases targeting specific proteins for ubiquitination (Weber et al. 2005; Gingerich et al. 2007). The expression of RTM genes in floral organs points the probable role of TRAFs during floral development in Arabidopsis. Doelling et al. (2001) analysed an Arabidopsis ubiquitin protein (UBP14) mutant which is involved in multiubiquitin chains formation, cause an embryonic lethal phenotype and shows embryonic arrest at globular stage. Similarly, analysis of an Arabidopsis counterpart of human CAND1 (cullin-associated and neddylation-dissociated) by Feng et al. (2004), revealed the fact that it preferentially interact with unmodified CUL1. The null mutant displayed distinct phenotypes, including late flowering, floral organ defects and low fertility, due to compromised activity of CUL1-containing ubiquitin E3 ligases, indicating that CAND1 is required for their optimal activity and their targeted cellular and developmental pathways. In study of Pintard et al. (2003a), it
has been found CUL3-based ubiquitin ligase required for degradation of MEI-1 (microtubule serving protein) necessary for the meiosis-to-mitosis transition, in C. elegans, depletion of CUL3 causes defects in early embryogenesis (Kurz et al. 2002), while in Drosophila melanogaster CUL3 is required for the external sensory organ development, pattern formation, cell growth and survival (Mistry et al. 2004). Singer et al. (1999) and Winston et al. (1999) reported that the mammalian CUL3 ubiquitin ligase mutant mice die at embryonic day 6.5 with defects in both embryonic and extra embryonic compartments. A closely related CUL3 protein in Arabidopsis viz. AtCUL3 was found to be essential for embryo development (Risseeuw et al. 2003; Figueroa et al. 2005). The functional complementation for cdc53 mutation in yeast by AtCUL1 gene from Arabidopsis provides direct evidence about the role of AtCUL1 in cell cycle regulation, there are several reports in mammals and yeast where the SCF (for SKP1, Cullin/CDC53, F-box protein) ubiquitin ligase targets a number of cell cycle regulators, transcription factors, and other proteins for degradation. AtCUL1 has been reported to display an arrest in early embryogenesis, both the transcript and the protein of the AtCUL1 gene accumulate in embryos and the protein localized mainly in the nucleus with weak accumulation in the cytoplasm during interphase and colocalized with the mitotic spindle in metaphase (Shen et al. 2002). Some of the E3 ubiquitin ligases and its associated components displaying abnormality in normal reproductive development of plants have been enlisted in Table 2.2. The loss-of-function mutants of HISTONE MONOUBIQUITINATION 1 (HUB1) and HUB2 (the E3 ligases) involved in histone H2B ubiquitylation mediated by E2 and E3 ubiquitins (Xu et al. 2009; Liu et al. 2008). H2B ubiquitylation is essential step in the activation of floral repressor genes as well as other biological and molecular processes as mutation of these genes also showed the pleiotropic phenotypes of the mutants. The regulation of time of flowering is coordinated by a web of gene regulatory networks that integrates developmental and environmental cues in plants. The regulation of floral integrators, such as FLOWERING LOCUS T (FT) and SUPPRESSOR OF EXPRESSION OF CO 1 (SOC1) is coordinated by GIGANTEA (GI) and CONSTANS (CO) to genes. CO plays a key role in photoperiodic flowering, acts as central hub in integrating photoperiodic and cold stress signals into the flowering genetic pathways and its degradation under cold treatment is mediated via a ubiquitin/proteasome pathway that involves the E3 ubiquitin ligase (Jung et al. 2012).

Table 2.2: The ubiquitin genes from Arabidopsis reported to play role in reproductive development and cell cycle of plant
<table>
<thead>
<tr>
<th>Sr. no.</th>
<th>Gene</th>
<th>AGI ID</th>
<th>Protein type</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UFO</td>
<td>At1g30950</td>
<td>E3 (F-BOX)</td>
<td>Flower Development</td>
<td>Samach et al. 1999</td>
</tr>
<tr>
<td>2</td>
<td>CUL1</td>
<td>At4g02570</td>
<td>E3 (Cullin)</td>
<td>Embryogenesis</td>
<td>Shen et al. 2002</td>
</tr>
<tr>
<td>3</td>
<td>CER3</td>
<td>At5g57800</td>
<td>E3 (Ring HC)</td>
<td>Wax biosynthesis</td>
<td>Hannoufa et al. 1996</td>
</tr>
<tr>
<td>4</td>
<td>COII</td>
<td>At2g39940</td>
<td>E3 (F-box)</td>
<td>pollen fertility</td>
<td>Xu et al. 2002</td>
</tr>
<tr>
<td>5</td>
<td>FKF1/ ZTL/ LKP2</td>
<td>At1g68050</td>
<td>E3 (F-box)</td>
<td>transition to flowering</td>
<td>Nelson et al. 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>/At5g57360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>/At2g18915</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SKP2;1</td>
<td>At1g05080</td>
<td>E3 (F-box)</td>
<td>cell division and cell size</td>
<td>Del Pozo et al. 2002</td>
</tr>
<tr>
<td>7</td>
<td>ASK1</td>
<td>At1g10940</td>
<td>E3 (SKP)</td>
<td>male meiosis</td>
<td>Yang et al. 1999</td>
</tr>
<tr>
<td>8</td>
<td>HOBBI T</td>
<td>At2g20000</td>
<td>E3 (APC)</td>
<td>Cell division</td>
<td>Bliou et al. 2002</td>
</tr>
<tr>
<td>9</td>
<td>HOS1</td>
<td>At2g39810</td>
<td>E3 Ub-ligase</td>
<td>photoperiodic and cold stress signals into the flowering genetic pathways</td>
<td>Jung et al. 2012</td>
</tr>
<tr>
<td>10</td>
<td>PUB13</td>
<td>At3g46510</td>
<td>E3 Ub-ligase</td>
<td>SA-dependent defense signaling and flowering time regulation in Arabidopsis</td>
<td>Li et al. 2012</td>
</tr>
</tbody>
</table>

**2.9 Conclusion**

During the last few years, advancements in molecular biology have fuelled significant progress in understanding the development of male gametophyte. The mutant analysis is an important tool to investigate and isolate the promoters involved in male gametophyte development. The availability of detailed information about the structure, function and the regulatory networks of the anther- and pollen-specific genes will help to modify traits related to them which include a versatile and durable male sterility system, floral quality, nutritional quality, gene confinement, etc. Till date only a limited number of promoters as well as genes has been characterized and functionally validated. Besides the well-characterized pollen-specific promoters from different species, there is still the need for other characterized pollen-specific promoters that confer different expression strengths or different expression windows.